A publication of ISCA*:
International Society for Computers
and Their Applications

INTERNATIONAL JOURNAL OF COMPUTERS AND THEIR APPLICATIONS

TABLE OF CONTENTS

***"International Journal of Computers and Their Applications is Peer Reviewed".**

Volume 31, No. 1, March 2024 **ISSN 1076-5204**

International Journal of Computers and Their Applications
A publication of the International Society for Computers and Their Applications

ISCA Headquaters • 278 Mankato Ave, #220, Winona, Minnesota 55987, USA
Phone 507-458-4517 • Email isca@isca-hq.org • URL: http://www.isca-hq.org

Editorial

It is my distinct honor, pleasure, and privilege to serve as the Editor-in-Chief of the International Journal of Computers and Their Applications (IJCA) since 2022. I have a special passion for the International Society for Computers and their Applications. I have been a member of our society since 2014 and have served in various capacities. These have ranged from being on program committees of our conferences to being Program Chair of CATA since 2021 and currently serving as one of the Ex-Officio Board Members. I am very grateful to the ISCA Board of Directors for giving me this opportunity to serve society and the journal in this role.

I would also like to thank all the editorial board, editorial staff, and authors for their valuable contributions to the journal. Without everyone's help, the success of the journal would be impossible. I look forward to working with everyone in the coming years to maintain and further improve the journal's quality. I want to invite you to submit your quality work to the journal for consideration for publication. I also welcome proposals for special issues of the journal. If you have any suggestions to improve the journal, please feel free to contact me.

Dr. Ajay Bandi
School of Computer Science and Information Systems
Northwest Missouri State University
Maryville, MO 64468
Email: AJAY@nwmissouri.edu

In 2024, we are having four issues planned (March, June, September, and December). The next latest issue is taking shape with a collection of submitted papers.

I would also like to announce that I will begin searching for a few reviewers to add to our team. We want to strengthen our board in a few areas. If you would like to be considered, don't hesitate to get in touch with me via email with a cover letter and a copy of your CV.

Ajay Bandi, Editor-in-Chief
Email: AJAY@nwmissouri.edu

Guest Editorial
March 2024

This issue of the International Journal of Computers and their Applications (IJCA) has gone through the normal review process. The papers in this issue cover a broad range of research interests in the community of computers and their applications.

IJCA Contributed Papers: This issue comprises papers that were contributed to the International Journal of Computers and their Applications (IJCA). The topics and main contributions of the papers are briefly summarized below:

THITIVATR PATANASAKPINYO from Mahidol University, Thailand, and ADEL SULAIMAN from Najran University, Saudi Arabia present their work “Toward an Extension of Efficient Algorithm to Solve Derangement Problems by Dynamic Programming Approach”. The article discusses the derangement problem, denoted as !n, which calculates the total ways to rearrange n items without any item occupying its original position. Formally, the problem involves a finite collection C of pairs (x_i, y_j), where $|\mathscr{C}| = n$, and each pair must have distinct values for i and j. The article presents a dynamic programming algorithm to compute !n and explores extending the problem to cover scenarios with a finite set of variables $(a_1, a_2, ..., a_k)$ instead of the traditional two-variable case (x and y).

AMIRA BENDJEDDOU, and MOUNA HEMICI from the University of Annaba, Algeria present their work ” Energy Efficient Vice Low Adaptive Hierarchy Clustering Protocol: EEV-LEACH”. This paper introduces EEV-LEACH, an energy-efficient hierarchical routing protocol for Wireless Sensor Networks (WSNs), designed to enhance the performance of the LEACH protocol. EEV-LEACH aims to extend network lifetime by reducing energy consumption at sensor nodes and cluster heads (CHs). It achieves this by minimizing the periodic selection of CHs, reducing the number of association messages exchanged, and introducing vice-CHs to share the workload with CHs. MATLAB simulations demonstrate that EEV-LEACH outperforms LEACH, LEACH-S, and TL-LEACH protocols in terms of network lifetime extension and overall overhead reduction, thereby addressing key objectives of sensor network applications.

SIJO THOMAS, SAJIMON ABRAHAM, PRAVEEN KUMAR, FAIZAL BASHEER, and JITHINMARY RAPHEL from Mahatma Gandhi University, India present their work “Nodule Classification Using Custom Build 3D Convolution Neural Network Model”. This paper introduces an innovative method for classifying lung nodules using a 3D Convolutional Neural Network (3D CNN) model, aiming to improve early detection and prognosis assessment of lung-related diseases. The approach includes customized 3D augmentation techniques and a unique 3D CNN model with finely tuned parameters to optimize nodule classification while minimizing training requirements. The research focuses on determining the optimal CT patch size for classification and addresses data transfer bottlenecks by caching data chunks locally to enhance performance. This method contributes to more precise and efficient nodule classification, thus aiding in better medical diagnosis and treatment planning.

SANDIP SARKAR from Hijli College, India, DIPANKAR DAS from Jadavpur University, India, and DAVID PINTO from Benemerita Universidad Autonoma de Puebla, Mexico present their work "Enhancing Math Word Problem Solving Using Multi-Head-Attention Mechanism". This paper addresses the challenge of solving mathematical word problems (MWPs) by introducing two innovative models: a multi-head attention mechanism-based model and a GRU-LSTM-based Seq2seq model. These models aim to bridge the gap between machine-understandable logic and human-readable language in MWP solving, utilizing state-of-the-art metrics such as BLEU and ROUGE for evaluation. Through extensive experimentation on diverse datasets, the models demonstrate impressive results, surpassing previous benchmarks. The research contributes to advancing math word problem-solving using deep learning techniques and attention mechanisms, while also identifying areas for future improvement and expansion in this domain.

AHMED AL-NAKEEB, MOUNIR EL KHATIB from Hamdan Bin Mohammed Smart University, UAE, and RAED ABU ZITAR from Sorbonne University UAE present their work "From PMO to PMOCoE: How to Manage Project Knowledge Process Improves Quality of Organization Knowledge Management Assets Cases from UAE". The paper discusses the concept of the Project Management Office Center of Excellence (PMOCoE) as a strategy to enhance project management as a core competency within organizations. It involves establishing a multidisciplinary team to integrate project management disciplines into organizational practices. To drive desired outcomes, PMOCoE must promote organizational learning. The paper presents the fundamentals of PMOCoE, including requirements, expected outcomes, and pathways to improved organizational performance. Data collected from interviews with projected entities in the UAE, such as Dubai Police and the Road and Transportation Authority, demonstrates the successful implementation of PMOCoE, leading to improved overall performance and business growth.

AOULALAY AYOUB, ABDERRAHIM EL MHOUTI, ABDELLAH EL ZAAR, RACHIDA ASSAWAB, and MOHAMMED MASSAR from Abdelmalek Essaadi University, Morocco present their work "A Deep Learning Approach for Moroccan Date Types Recognition". This study explores the use of computer vision techniques to classify seven popular types of date fruits in Morocco, addressing the growing demand in the market. By constructing a dataset and employing convolutional neural networks (CNNs), two approaches are compared: standard feature extraction and fine-tuning of pre-trained models. Both methods achieve high performance, with a classification precision of 97%. This research highlights the effectiveness of artificial intelligence in recognizing and categorizing date fruits based on their visual characteristics, catering to the needs of the expanding market.

SHATHA A. BAKER, HESHAM HASHIM MOHAMMED, and OMAR I. ALSAIF from the Northern Technical University, Iraq present their work "Docker Container Security Analysis Based on Virtualization Technologies". This paper conducts a security analysis of Docker containers, focusing on internal security and their interaction with Linux kernel security features. Docker's isolation of resources like filesystems, networks, devices, and processes, as well as inter-process communication, is examined alongside its utilization of Linux Cgroup and Namespace. The paper also discusses Docker's interactions with SELinux, AppArmor, Seccomp, and Linux functions to enhance host system security. Recommendations for improving Docker's security include deactivating certain container functions to prevent breaches and enhancing interoperability with Linux kernel security mechanisms.

SOURAV BANERJEE, SUDIP BARIK, JACOB TAURO from Kalyani Government Engineering College, India DEBASHIS DAS from Meharry Medical College, USA, and NARAYAN C DEBNATH from Eastern International University, Vietnam present their work "Enhancing Acute Lymphoblastic Leukemia Image Segmentation: Unveiling the Impact of Color Spaces and Clustering Techniques". This paper investigates the challenges of accurately segmenting white blood cells (WBCs), particularly in cases of Acute Lymphoblastic Leukemia (ALL), which requires precise segmentation for computer-aided diagnosis. It explores contemporary segmentation techniques, focusing on clustering methods within ALL images and emphasizing the importance of suitable color spaces for segmenting the cytoplasmic region of WBCs. Additionally, the paper conducts a comparative evaluation of different color spaces and image clustering algorithms to identify the most effective combination for precise WBC segmentation in leukemia diagnosis.

As guest editors, we would like to express our deepest appreciation to the authors and the reviewers. We hope you will enjoy this issue of the IJCA. More information about ISCA society can be found at http://www.isca-hq.org.

Guest Editors:

Ajay Bandi, Northwest Missouri State University, USA

March 2024

Toward an Extension of Efficient Algorithm to Solve Derangement Problems by Dynamic Programming Approach

Thitivatr PatanasakPinyo*
Mahidol University, Salaya, Nakhon Pathom 73170, THAILAND
Adel Sulaiman†
Najran University, Najran, SAUDI ARABIA

Abstract

In statistics, probability theory, and computer science, a derangement, $!n$, is known to be a basic problem that computes total ways of rearranging $n \in \mathbb{N}$ items such that a result contains no item i that stands in the same position as it did in the input. Formally, the derangement problem has a problem instance of a finite collection $\mathscr{C} = \{(x_1,y_1),....,(x_n,y_n)\}, |\mathscr{C}| = n$. With this formulation, $!n$ is a total number of qualified collections $\mathscr{C}'$ where $\mathscr{C}'$ contains n members of (x_i,y_j) where $i \neq j$ and every x_i and y_j $(1 \leq i,j \leq n)$ must show up exactly once in $\mathscr{C}'$. In this article, we present a dynamic programming algorithm that computes $!n$ and its justification. We also provide a discussion about extending the limitation of the problem with an objective to cover a general case where we have a finite set of variables $a_1,a_2,....,a_k$ rather than the traditional scenario that has only two variables: x and y.

Key Words: derangement, dynamic programming, recursion, algorithm, complete derangement

1 Introduction

The total ways of an arrangement of $n \in \mathbb{N}$ distinct items is a fundamental statement that has been brought to introduce the concept of factorial ($n!$). A derangement problem ($!n$) has the same input instance as $n!$, which is a finite set that contains n distinct items, i.e., $\{x_1,x_2,....,x_n\}$. The derangement problem asks to compute how many ways of an arrangement of these n items such that for every item x_i, it is not located on location i [2, 3]. For the case of $n = 1$, it is impossible. For the case of $n = 2$, there is only one way to do, i.e., $!2 = 1$. Particularly, let us illustrate the case of $n = 2$. Let the input instance be $\{x_1,x_2\}$. There are two possible arrangements of this input. The first arrangement is (x_1,x_2), which is obviously not qualified for a derangement. The second arrangement is (x_2,x_1), which is a valid derangement. Thus, $!2 = 1$.

*Faculty of Information and Communication Technology. Email: thitivatr.pat@mahidol.edu.
†College of Computer Science and Information Systems. Email: aaalsulaiman@nu.edu.sa.

Without loss of generality, we create a new viewpoint to the derangement problem by modifying an instance of the problem. Particularly, the input instance contains a collection $\mathscr{C} = \{(x_1,y_1),....,(x_n,y_n)\}, |\mathscr{C}| = n$. The objective of the derangement problem will be treated as computing total ways that we can create a collection $\mathscr{C}' = \{(x_i,y_j)\}, 1 \leq i,j \leq n$ with the following constraints:

1. $|\mathscr{C}'| = n$.
2. $\forall i, 1 \leq i \leq n, x_i$ must show up in $\mathscr{C}'$ exactly once.
3. $\forall j, 1 \leq j \leq n, y_j$ must show up in $\mathscr{C}'$ exactly once.
4. $\forall (x_i,y_j) \in \mathscr{C}', i \neq j$

With the new input instance, we use an umbrella problem, which can be found as an exercise in several textbooks about probability theory and statistics, to illustrate how it represents the derangement problem. The umbrella problem states that there are $n \in \mathbb{N}$ customers, which are $x_1,x_2,....,x_n$. Each customer x_i has his own umbrella y_i. Every customer must drop his umbrella into the box in front of the restaurant. After every customer finishes his meal, he will take an umbrella from the box and leave. The umbrella problem asks us to compute how many ways that **no** customer x_i gets his umbrella y_i. This is obvious that it is equivalent to compute all possible collections $\mathscr{C}'$ that we have previously mentioned where $x_1,....,x_n$ represent n customers and $y_1,....,y_n$ represents n umbrellas where Umbrella y_k belongs to Customer x_k at the beginning. Table 1 shows $!n$ and its corresponded collections $\mathscr{C}'$ for $n = 1$ to 4. We also show results of $!n$ for small n in Table 2.

Even though it is impractical to list every valid $\mathscr{C}'$ when n is bigger, computing total number of every valid $\mathscr{C}'$ is still feasible. Several methodologies were suggested to compute $!n$, mostly were intricate mathematical formulas. In this paper, we have come up with an efficient algorithm that can solve $!n$ by implementing a dynamic programming approach, which is a technique that is widely used in the field of algorithm design. Many problems have been known to have an efficient dynamic programming algorithm, which is way better than a traditional recursive algorithm, e.g., the famous Fibonacci series where an

Table 1: Total Valid Collections $\mathscr{C}'$ for Different n.

n	$!n$	$\mathscr{C}'$
1	0	$\emptyset$
2	1	$\mathscr{C}'_1 = \{(x_1,y_2),(x_2,y_1)\}$
3	2	$\mathscr{C}'_1 = \{(x_1,y_2),(x_2,y_3),(x_3,y_1)\}$ $\mathscr{C}'_2 = \{(x_1,y_3),(x_2,y_1),(x_3,y_2)\}$
4	9	$\mathscr{C}'_1 = \{(x_1,y_2),(x_2,y_1),(x_3,y_4),(x_4,y_3)\}$ $\mathscr{C}'_2 = \{(x_1,y_2),(x_2,y_3),(x_3,y_4),(x_4,y_1)\}$ $\mathscr{C}'_3 = \{(x_1,y_2),(x_2,y_4),(x_3,y_1),(x_4,y_3)\}$ $\mathscr{C}'_4 = \{(x_1,y_3),(x_2,y_1),(x_3,y_4),(x_4,y_2)\}$ $\mathscr{C}'_5 = \{(x_1,y_3),(x_2,y_4),(x_3,y_1),(x_4,y_2)\}$ $\mathscr{C}'_6 = \{(x_1,y_3),(x_2,y_4),(x_3,y_2),(x_4,y_1)\}$ $\mathscr{C}'_7 = \{(x_1,y_4),(x_2,y_1),(x_3,y_2),(x_4,y_3)\}$ $\mathscr{C}'_8 = \{(x_1,y_4),(x_2,y_3),(x_3,y_1),(x_4,y_2)\}$ $\mathscr{C}'_9 = \{(x_1,y_4),(x_2,y_3),(x_3,y_2),(x_4,y_1)\}$

Table 2: Results of $!n$ for Small n.

n	$!n$
1	0
2	1
3	2
4	9
5	44
6	265
7	1854
8	14833
9	133496
10	1334961
11	14684570
12	176214841

objective is to compute a member of the series at any index $i \in \mathbb{N}$ has an efficient dynamic programming algorithm.

This paper follows the following structure. Section 2 reveals related work, which consists of descriptions and past studies about the derangement problem and the dynamic programming approach, with an objective to provide readers a story of the topic and its challenge in an aspect of algorithm study and theoretical computer science. Section 3 explains how we came up with a conceptualization to solve the derangement problem by formulating it as a recurrence problem that each layer has another subproblem inside. Section 4 illustrates the proposed algorithm along with efficiency analysis. Section 5 extends the work by initiating an idea to design a universal algorithm that can effectively solve a generalized version of the derangement problem. Section 6 summarizes all the work being presented in this paper and shed a light to possible future work.

2 Preliminaries

2.1 Derangement

2.1.1 Solving Problem Using Mathematical Approach

The derangement problem hes been interchangeably known as the **hat-check** problem where the instance of the problem contains $n \in \mathbb{N}$ people and each person has his own hat. The problem asks how many ways to assign the hat to each person such that no one gets his own hat. Multiple mathematical methodologies have been suggested. Here are some of them that we select to mention.

1. The first methodology treats the hat-check problem as two mutual-exclusive incidents. Particularly, let $X = \{x_1, x_2, ..., x_n\}$ be a set of all n people and $Y = \{y_1, y_2, ..., y_n\}$ be a set of all hats where x_i owns y_i. With this setup, all the derangement results can be partitioned into:
   (a) For a certain x_i, x_i is assigned y_j and x_j is assigned y_i.
   (b) For a certain x_i, x_i is assigned y_j but x_j is assigned y_k where $i \neq k$.

   For the first incident, it is equivalent to the exchange of hats between x_i and x_j. This causes the problem to be reduce to a smaller instance of the hat-check problem with a size $n-2$. The second incident also leads to a smaller instance where $X_{\text{new}} = X \setminus \{x_i\}$ and $Y_{\text{new}} = Y \setminus \{y_j\}$. X_{new} is a set of all people for this new instance. Y_{new} is a set of hats for the new instance. Combining both incidents together, a formula for $!n$ shall be constructed as:

$$!n = (n-1)(!(n-1) + !(n-2)), n \geq 2$$

   where $!1 = 0$ and $!0 = 1$ [17].
2. The second methodology is simply a formula that implementing a summation, particularly,

$$!n = n! \sum_{i=0}^{n} \frac{(-1)^i}{i!}, n \geq 0$$

3. Another formula that can also return a solution of $!n$ is defined using a floor function as:

$$!n = \lfloor \frac{n!}{e} + \frac{1}{2} \rfloor$$

   where e is a base of natural logarithm and $\lfloor x \rfloor$ is the maximum $z \in \mathbb{Z}$ subject to $z \leq x$ [18].

We are going to discuss in more detail about the hat-check problem later in this paper.

2.1.2 Formulating Problem Using Permutation Function

Let A and B be two finite sets where $A = \{x_1, x_2, ..., x_n\}$ and $B = \{y_1, y_2, ..., y_n\}$. We define a concatenation of A and B (denoted by $A \cdot B$ or AB) as

$$A \cdot B = \{(x_i, y_j) \in A \times B\}$$

subject to:

1. $\forall x_i \in A$, x_i must show up in exactly one member of $A \cdot B$
2. $\forall y_j \in A$, y_j must show up in exactly one member of $A \cdot B$
3. $\forall (x_i, y_j) \in A \cdot B, i \neq j$.

The concatenation $A \cdot B$ can be perceived as a bijection $A \to B$. The total ways that we can create $A \cdot B$ is equivalent to a solution of $!n$. To apply this methodology, we run an algorithm that generate all permutations on B. Next, we do a concatenate of A on each B returned from the algorithm. If the result of $A \cdot B$ satisfies all constraints of concatenation, we then keep it. Otherwise, we omit it. The total number concatenations qualified is nothing but $!n$. For instance, we use the case of $n = 3$. The instance of the problem are $A = \{x_1, x_2, x_3\}$ and $B = \{y_1, y_2, y_3\}$. All valid permutations of B are:

1. $B_1 = \{y_1, y_2, y_3\}$
2. $B_2 = \{y_1, y_3, y_2\}$
3. $B_3 = \{y_2, y_1, y_3\}$
4. $B_4 = \{y_2, y_3, y_1\}$
5. $B_5 = \{y_3, y_1, y_2\}$
6. $B_6 = \{y_3, y_2, y_1\}$

From the six permutations of B, only two of them can form a precise concatenations, which are $A \cdot B_4 = \{(x_1, y_2), (x_2, y_3), (x_3, y_1)\}$ and $A \cdot B_5 = \{(x_1, y_3), (x_2, y_1), (x_3, y_2)\}$. Therefore, $!3 = 2$. Although this methodology guarantees to return a correct answer but its performance in an aspect of running time is not quite efficient since its complexity is $n!$ and $n! \notin \Theta(n^k), k \in \mathbb{Z}$, i.e., it is not polynomial-bound.

2.1.3 Relationship with Bell's Number

Recall that a Bell's number is an infinite sequence of a natural number [15]. Let $\mathbb{B} = (b_1, b_2, \ldots)$ be a sequence of Bell's number. We define b_i, $i \in \mathbb{N}$, by:

$b_i =$ the total number of partitions of a finite set S where $|S| = i$

With this definition, $\mathbb{B}$ looks like:

$$\mathbb{B} = (1, 2, 5, 15, 52, 203, \ldots)$$

There exists a formula [18] to compute $!n$ by using $\mathbb{B}$, particularly, the formula is:

$$!n - \frac{n!}{e} - \sum_{k=1}^{m} (-1)^{n+k-1} \frac{b_k}{n^k} = O\left(\frac{1}{n^{m+1}}\right)$$

2.2 Dynamic Programming

A dynamic programming (DP) is a well known algorithm that implements a concept of recursion plus a usage of a memory to memorize a sub-solution previously computed by a recursive call for future use. The design of DP can be divided into two types, which are either top-down or bottom-up. Even though both always returns a correct solution, choosing between top-down or bottom up most is mostly based on the structure and the way that we formulate the problem. Several problems in both mathematics and computer science have an efficient algorithm that implementing DP as a backbone, e.g., Computing a Fibonacci number at arbitrary index i or solving a maximum-weight interval scheduling problem or solving a deadlock problem in an operating system [4, 7].

3 Understanding the Basis of Derangement

There is no better way to design a DP algorithm than unrolling the instance of the problem to the most inner layer and revealing the hidden pattern of the tentative solution. With this guidance, we start with the most basic case, particularly, the solution of $!n$ when $n = 2$. The instance of the case $n = 2$ contains $X = \{x_1, x_2\}$ and $y = \{y_1, y_2\}$. Clearly, there is only one valid candidate of $\mathscr{C}'$, which is

$$\mathscr{C}' = \{(x_1, y_2), (x_2, y_1)\}$$

This can be interpreted as Person x_1 gets Umbrella y_2 and Person x_2 gets Umbrella y_1. Hence, $!2 = 1$. To eliminate any ambiguity, let us define $C_{(i,j)}$ to be a total ways of assigning umbrellas such that $j \in \mathbb{N}$ out of $i \in \mathbb{N}$ people do not get their umbrellas. Equivalently, a solution of $!2$ is nothing but $C_{(2,2)}$. Hence, for any case of $n \in \mathbb{N}$, $!n = C_{(n,n)}$. We also define $\mathbb{S}_n$ for future use where $\mathbb{S}_n = n!$. The purpose of the symbol is just to simplify the proof in the next section. Therefore, with this basic case that $n = 2$, we can write:

$$!2 = C_{(2,2)} = 1 = \mathbb{S}_2 - 1$$

Although we have seen that there is only one candidate of $\mathscr{C}$ when $n = 2$ as described previously, it is worth here to mention all possible $X \cdot Y$ concatenations if there exists any clue that can lead us to the base case for DP algorithm. Since $n = 2$, there are totally $2! = 2$ permutations for Y, which are $\{y_1, y_2\}$ and $\{y_2, y_1\}$. Hence, all possible cross products of $X \times Y$ are:

$$\mathscr{C}'_1 = \{(x_1, y_1), (x_2, y_2)\}$$
$$\mathscr{C}'_2 = \{(x_1, y_2), (x_1, y_1)\}$$

We can see that there is only one concatenation, which is $\mathscr{C}'_2 = \{(x_1, y_2), (x_2, y_1)\}$. So the fact that $!2 = \mathbb{S}_2 - 1 = C_{(2,2)} = 1$ still has no change. The interesting question is, what else we can tell about $\mathscr{C}'_1 = \{(x_1, y_1), (x_2, y_2)\}$. It is quite straightforward that $\mathscr{C}'_1$ represents the case that at least one person gets his own umbrella. Since $n = 2$, the case that at least one person gets his umbrella is the same as the case that every person gets his umbrella, which can be represented by $\mathscr{C}_{(2,0)}$. We then can construct the following equation to represent this scenario:

$$\mathbb{S}_2 = C_{(2,2)} + C_{(2,0)}$$

We are going to observe one more example when $n = 3$ to witness if there exists any difference comparing to the previous case. For $n = 3$, the instance of the problem consists of $X = \{x_1, x_2, x_3\}$ and $Y = \{y_1, y_2, y_3\}$. All valid permutations are:

$$\begin{aligned}
\mathscr{C}'_1 &= \{(x_1,y_1),(x_2,y_2),(x_3,y_3)\}\\
\mathscr{C}'_2 &= \{(x_1,y_1),(x_2,y_3),(x_3,y_2)\}\\
\mathscr{C}'_3 &= \{(x_1,y_2),(x_2,y_1),(x_3,y_3)\}\\
\mathscr{C}'_4 &= \{(x_1,y_3),(x_2,y_2),(x_3,y_1)\}\\
\mathscr{C}'_5 &= \{(x_1,y_2),(x_2,y_3),(x_3,y_1)\}\\
\mathscr{C}'_6 &= \{(x_1,y_3),(x_2,y_1),(x_3,y_2)\}
\end{aligned}$$

From the list, out of $3! = 6$ permutations/assignments, only $\mathscr{C}'_5$ and $\mathscr{C}'_6$ are valid derangement. Hence, $!3 = 2$. However, all six permutations can be partitioned into a set A_i by the following manner:

$$\mathscr{C}'_k \in A_i \iff i \text{ people do not get their own umbrellas in } \mathscr{C}'$$

Following this criteria that we use to do A_i partition, it is clear that $|A_i| = C_{(3,i)}$. Particularly:

$$\begin{aligned}
A_3 &= \{\mathscr{C}'_5, \mathscr{C}'_6\}\\
A_2 &= \{\mathscr{C}'_2, \mathscr{C}'_3, \mathscr{C}'_4\}\\
A_1 &= \emptyset\\
A_0 &= \{\mathscr{C}'_1\}
\end{aligned}$$

With th definition of A_i, $|A_3|$ is equivalent to $!3$. Since it fits the concept of partition, we can construct the equation:

$$\begin{aligned}
\$_3 &= \sum_{i=0}^{3} |A_i| &\\
\$_3 &= |A_3| + |A_2| + |A_1| + |A_0| & \because \text{partition}\\
|A_3| &= \$_3 - (|A_2| + |A_1| + |A_0|) &\\
2 &= 6 - (3 + 0 + 1) & \because \$_3 = 3!\\
2 &= 2 & \blacksquare
\end{aligned}$$

At this point, there are a hidden pattern that reveals itself after we study both cases ($n = 2$ and $n = 3$). The official proof of the concept starts from here.

Lemma 1 For $n \in \mathbb{N}$, a collection $\mathscr{C}'_i$ of valid permutations of Y must belong to exactly one set $A_k, 0 \le k \le n$, such that $|A_k| = C_{(n,k)}$.

Proof We recall the complete proof of this lemma from PatanasakPinyo and Sulaiman [14]. Note that some variables have been changed according to the context of the paper. Suppose there exists a permutation $\mathscr{C}'_i$ such that it belongs to A_k and $A_m, m \neq k$, at the same time. Since $\mathscr{C}'_i \in A_k$, there are k umbrellas that are not assigned to their owners, i.e., there are k numbers of (x_a, y_b) pairs that $a \neq b$ in $\mathscr{C}'_i$. Since $\mathscr{C}'_i \in A_m$, there are m umbrellas that are not assigned to their owners, i.e., there are m numbers of (x_a, y_b) pairs that $a \neq b$ in $\mathscr{C}'_i$. To have these two statements satisfy true, $k = m$ must hold. Thus there exists a contradiction. Hence, proved $\blacksquare$

Theorem 2 For $n \in \mathbb{N}$, $!n = \$_n - \sum_{i=0}^{n-1} C_{(n,i)}$

Proof Again, we recall the formal proof of the theorem from PatanasakPinyo and Sulaiman [14]. This is a more formal version of the equation that we constructed previously when we discussed about the case of $n = 3$. For $n \in \mathbb{N}$:

$$\begin{aligned}
\$_n &= C_{(n,0)} + C_{(n,1)} + C_{(n,2)} + \cdots + C_{(n,n)} & \because \textbf{Lemma 1}\\
\$_n &= C_{(n,0)} + C_{(n,1)} + C_{(n,2)} + \cdots + C_{(n,n-1)} + C_{(n,n)} &\\
\$_n &= \sum_{i=0}^{n-1} C_{(n,i)} + C_{(n,n)} &\\
\$_n &= \sum_{i=0}^{n-1} C_{(n,i)} + !n &\\
!n &= \$_n - \sum_{i=0}^{n-1} C_{(n,i)} & \blacksquare
\end{aligned}$$

Lemma 3 For $n \in \mathbb{N}, C_{(n,1)} = 0$

Proof The proof of this lemma is very obvious since it is impossible, for any case of $n \in \mathbb{N}$, where there is only one person that is assigned a wrong umbrella. The total of people who are assigned the wrong umbrellas must be at least two. Hence, proved $\blacksquare$

We will use Theorem 2 and Lemma 3 to apply on the cases of $n = 2, 3, 4$, respectively.

1. Case $n = 2$

$$!2 = \$_2 - (C_{(2,0)} + C_{(2,1)}) \tag{1}$$
$$!2 = \$_2 - C_{(2,0)} \tag{2}$$
$$!2 = \$_2 - 1 \tag{3}$$

2. Case $n = 3$

$$!3 = \$_3 - (C_{(3,0)} + C_{(3,1)} + C_{(3,2)}) \tag{4}$$
$$!3 = \$_3 - (C_{(3,0)} + C_{(3,2)}) \tag{5}$$
$$!3 = \$_3 - (1 + C_{(3,2)}) \tag{6}$$

3. Case $n = 4$

$$!4 = \$_4 - (C_{(4,0)} + C_{(4,1)} + C_{(4,2)} + C_{(4,3)}) \tag{7}$$
$$!4 = \$_4 - (1 + C_{(4,2)} + C_{(4,3)}) \tag{8}$$

The case of $n = 2$ might cause zero problem since it can be directly computed (Line 3). However, the cases of $n = 3$ and $n = 4$ might do since there exist terms that need further computations, which are $C_{(3,2)}$ (Case $n = 3$, Line 6) and $C_{(4,2)} + C_{(4,3)}$ (Case $n = 4$, Line 8). These terms are leading us to a recursive property of the derangement problem. We begin unrolling $C_{(3,2)}$ from Line 6 first. By the definition, $C_{(3,2)}$ is equivalent to the scenario that there are two out of three people that are not assigned their umbrellas. The total ways of $C_{(3,2)}$ can be computed via two steps. We first compute $\binom{3}{1}$ to pick one person who correctly gets his umbrella. The second is to compute how many ways that the remaining two people are assigned wrong umbrellas, which is exactly computing !2. Now we can see that the recursive call has popped up, i.e.,

$$!3 = \mathbb{S}_3 - (1 + C_{(3,2)}) \tag{9}$$

$$!3 = \mathbb{S}_3 - (1 + \binom{3}{1}!2) \tag{10}$$

We continue with the case of $n = 4$ in which we have to compute $C_{(4,2)} + C_{(4,3)}$. First, to achieve $C_{(4,2)}$, we compute $\binom{4}{2}$ to get total ways of picking 2 people who are correctly assigned the umbrellas. Next, we compute total ways of the remaining two-people derangement by recalling !2. Similarly for $C_{(4,3)}$, we compute $\binom{4}{1}$ for one person who correctly gets his umbrella and !3 for the remaining three-people derangement. Hence:

$$!4 = \mathbb{S}_4 - (1 + C_{(4,2)} + C_{(4,3)}) \tag{11}$$

$$!4 = \mathbb{S}_4 - (1 + \binom{4}{2}!2 + \binom{4}{1}!3) \tag{12}$$

To compute !4 (Line 12), the first term that we can easily compute is $\mathbb{S}_4$, which is nothing but 4!. The second term is $\binom{4}{2}!2$. This one is also computably easy since !2 is a basis case of derangement. The last term that we have to get it done is $\binom{4}{1}!3$. We can compute !3 by Line 10, which we have to do !2. This is the place that the DP gets involved since we have already computed !2 in the previous term. We just recall the value of !2 that we previously computed and stored it somewhere in the memory rather than re-computing the same expression. Searching for the value stored in an array of size $n \in \mathbb{N}$ has $O(n)$ running time, which is considered efficient.

We finally decide to demonstrate the complete process of this DP algorithm for derangement problem. We select the case of $n = 5$. The instance for this case includes $X = \{x_1, x_2, x_3, x_4, x_5\}$ and $Y = \{y_1, y_2, y_3, y_4, y_5\}$. Let $\mathbb{O}$ represent an array of size n where $\mathbb{O}[i] = !i, 1 \leq i \leq n$. Tables 3 - 9 illustrate all the steps happen during the run-time of the algorithm.

Table 3: Initial Step of Computing !5.

Process	$!5 = \mathbb{S}_5 - (C_{(5,0)} + C_{(5,1)} + C_{(5,2)} + C_{(5,3)} + C_{(5,4)})$
$\mathbb{O}$	[0]

At the initial step (Table 3), We recall Theorem 2 to construct an equation for !5. $\mathbb{O}$ is initiated with only one member, $\mathbb{O}[1] = 0$, which represents !1.

Table 4: Step 2 of Computing !5 (Compute $\mathbb{S}_5$).

Process	$!5 = 5! - (C_{(5,0)} + C_{(5,1)} + C_{(5,2)} + C_{(5,3)} + C_{(5,4)})$
$\mathbb{O}$	[0]

At Step 2 (Table 4), we solve $\mathbb{S}_5$ by $\mathbb{S}_5 = 5!$. There is no change for $\mathbb{O}$ at this step.

At Step 3 (Table 5), we replace $C_{(5,0)}$ by 1 since there is only one way that everyone correctly gets his umbrella.

Table 5: Step 3 of Computing !5 (Compute $C_{(5,0)}$).

Process	$!5 = 5! - (1 + C_{(5,1)} + C_{(5,2)} + C_{(5,3)} + C_{(5,4)})$
$\mathbb{O}$	[0]

Table 6: Step 4 of Computing !5 (Compute $C_{(5,1)}$).

Process	$!5 = 5! - (1 + 0 + C_{(5,2)} + C_{(5,3)} + C_{(5,4)})$
$\mathbb{O}$	[0]

Table 7: Step 5 of Computing !5 (Compute $C_{(5,2)}$).

Process	$!5 = 5! - (1 + 0 + \binom{5}{3}!2 + C_{(5,3)} + C_{(5,4)})$ $!5 = 5! - (1 + 0 + \binom{5}{3}(1) + C_{(5,3)} + C_{(5,4)})$
$\mathbb{O}$	[0, 1]

At Step 4 (Table 6), we recall Lemma 1 for $C_{(5,1)}$. Hence, $\mathbb{O}$ still has no update.

At Step 5 (Table 7), we would like to compute $C_{(5,2)}$. So, we replace it with $\binom{5}{3}!2$. Since $!2 = 1$ is the base case of derangement, we replace !2 with 1 in the equation and add 1 to $\mathbb{O}[2]$.

Table 8: Step 6 of Computing !5 (Compute $C_{(5,3)}$).

Process	$!5 = 5! - (1 + 0 + \binom{5}{3}(1) + \binom{5}{2}!3 + C_{(5,4)})$ $!5 = 5! - (1 + 0 + \binom{5}{3}(1) + \binom{5}{2}(2) + C_{(5,4)})$
Recursive Process	$!3 = \mathbb{S}_3 - (C_{(3,0)} + C_{(3,1)} + C_{(3,2)})$ $!3 = 3! - (1 + 0 + \binom{3}{1}!2)$ $!3 = 2$
$\mathbb{O}$	[0, 1, 2]

At Step 6 (Table 8), we would like to compute $C_{(5,3)}$. So, we replace it with $\binom{5}{2}!3$. Unfortunately, we have not had the value of !3 stored in $\mathbb{O}$. We have to recursively call Theorem 2 to handle !3. After we constructed an equation for !3 (Row 2 of Table 8), we, rather than re-computing, search on $\mathbb{O}$ the value of !2 that we previously computed, which is required to compute !3. We end up with $!3 = 2$. So the term $\binom{5}{2}!3$ is done. We also add $!3 = 2$ to $\mathbb{O}[3]$.

Table 9: Step 7 of Computing !5 (Compute $C_{(5,4)}$).

Process	$!5 = 5! - (1 + 0 + \binom{5}{3}(1) + \binom{5}{2}!3 + \binom{5}{1}!4)$ $!5 = 5! - (1 + 0 + \binom{5}{3}(1) + \binom{5}{2}(2) + \binom{5}{1}(9))$ ■
Recursive Process	$!4 = \mathbb{S}_4 - (C_{(4,0)} + C_{(4,1)} + C_{(4,2)} + C_{(4,3)})$ $!4 = 4! - (1 + 0 + \binom{4}{2}!2 + \binom{4}{1}!3)$ $!4 = 9$
$\mathbb{O}$	[0, 1, 2, 9]

At Step 7 (Table 9), we would like to compute $C_{(5,4)}$. So, we replace it with $\binom{5}{1}!4$. Unfortunately, we have not had the value of !4 stored in $\mathbb{O}$. We have to recursively call Theorem 2 to handle !4. After we constructed an equation for !4 (Row 2 of Table 9), we, rather than re-computing, search on $\mathbb{O}$ both the values of !2 and !3 that we previously computed, which are required to compute !4. We end up with $!4 = 9$. So the term $\binom{5}{2}!3$ is done. We also add $!3 = 2$ to $\mathbb{O}[3]$. This is also the final step since we unlock every term required to compute !5. The algorithm halts and return the solution of !5. The complete pseudocode and the analysis part to verify the performance of the purposed DP algorithm will be discussed in the next section.

4 Dynamic Programming as Efficient Algorithm for Derangement

4.1 Design Phase

Algorithm 1 Dynamic Programming Algorithm for Derangement Problem.

```
1: function TOTAL-DERANGEMENT(n)
2:     return REC-DERANGE(n, n)
3: end function
```

Algorithm 2 Algorithm for Subproblem of Derangement.

```
1: function REC-DERANGE(n, i)
2:     𝕆 ← [0]
3:     if i = 0 then
4:         return 1
5:     end if
6:     if i = 1 then                                  ▷ Lemma 3
7:         return 0
8:     end if
9:     if 1 < i ≤ n − 1 then
10:        d_i ← 0
11:        if 𝕆[i] ≠ null then
12:            d_i ← 𝕆[i]
13:        else
14:            d_i ← REC-DERANGE(i, i)
15:            𝕆[i] ← d_i
16:        end if
17:        return (n choose n−i) d_i
18:    end if
19:    if i = n then
20:        sum ← 0
21:        for j = 0 : n − 1 do
22:            sum ← sum + REC-DERANGE(n, j)
23:        end for
24:        return n! − sum                 ▷ !n = 𝕊_n − Σ_{i=0}^{n−1} C_(n,i)
25:    end if
26: end function
```

We officially present two algorithms [14]. Algorithm 1 is the main procedure that retrieves an input $n \in \mathbb{N}$ and output $!n$. Algorithm 2 is a sub-procedure, which implementing DP approach previously described in the last section, called by Algorithm 1. The procedure that Algorithm 2 executes is totally based on Lemma 1, Theorem 2, and Lemma 3 as we mark in-line of the pseudocode. Furthermore for Algorithm 2, we added a modification to the previous version of PatansakPinyo and Sulaiman [14] on Lines 2, 10, 11, 12, 13, 14, 15, and 16 to fulfill the usage of Array $\mathbb{O}$, which makes it fits the concept of the DP algorithm.

4.2 Analysis Phase

The former analysis done by PatanasakPinyo and Sulaiman [14] is still valid even though we add more instructions to the pseudocode in Algorithm 2. Particularly, Algorithm 2 consists of two parts that affect the running time. The first part os located on Lines 14 and 22. This part involves with a recursive call. The upper bound of the recursive call is at most $n-1$ calls. Another part that causes a significant running time is an iteration on Line 21. Again, the upper bound of the iteration is $O(n)$. Combining these two parts together, we come up with the running time of $O(n^2)$, which is a polynomial running time. For Algorithm 1, the pseudocode has no change at all, so there is no effect on the running time. Hence, the running time of the proposed DP algorithm is dominated by Algorithm 2. We conclude that the running time is $O(n)$, which is technically considered efficient as it has a polynomial bound [5].

5 Generalization of Derangement Problems

5.1 Hat-check Problem: A Popular Special Case of Derangement Problems and the Way to Generalize It

The classic derangement is a special case that there exists the efficient algorithm (we also proposed one in this paper). However, there is no guarantee that the algorithm still works precisely when we generalize the problem. In this paper, we use the umbrella problem to depict the derangement problem. In mathematics and probability theory, this same approach is known by the name of the hat-check problem. An instance of the hat-check problem is very similar to the umbrella problem where there are two sets: X and Y. $X = \{x_1, x_2, ..., x_n\}$ is a set of $n \in \mathbb{N}$ customers. $Y = \{y_1, y_2, ..., y_n\}$ is a set of n hats where Hat y_i belongs to Customer x_i for all $1 \leq i \leq n$. We say that both the hat-check problem or the umbrella problem are a special case of derangement problem where each customer has possessed only one item (either a hat or an umbrella). Hence, we generalize the derangement problem that allows each customer to possess $k \in \mathbb{N}$ items. Readers can think of the case there are $n \in \mathbb{N}$ customers where each customer has $k \in \mathbb{N}$ items. The instance of the generalized version of the hat-check problem would include the following $(k+1)$ sets:

1. $X = \{x_1, x_2, ..., x_n\}$ where $x_i, 1 \leq i \leq n$, represents Customer i.
2. $Y_1 = \{y_{[1,1]}, y_{[1,2]}, ..., y_{[1,n]}\}$ where $y_{[1,i]}$ represents the first item of Customer i. We can assume that the first item is a

hat. Hence, Y_1 is a set of n hats.

3. $Y_2 = \{y_{[2,1]}, y_{[2,2]}, \ldots, y_{[2,n]}\}$ where $y_{[2,i]}$ represents the second item of Customer i. We can assume that the second item is an umbrella. Hence, Y_2 is a set of n umbrellas.

....

...

..

k. $Y_{k-1} = \{y_{[k-1,1]}, y_{[k-1,2]}, \ldots, y_{[k-1,n]}\}$ where $y_{[k-1,i]}$ represents the $(k-1)^{\text{th}}$ item of Customer i. We can assume that the $(k-1)^{\text{th}}$ item is a pair of sunglasses. Hence, Y_{k-1} is a set of n pairs of sunglasses.

$(k+1)$. $Y_k = \{y_{[k,1]}, y_{[k,2]}, \ldots, y_{[k,n]}\}$ where $y_{[k,i]}$ represents the k^{th} item of Customer i. We can assume that the k^{th} item is a mobile phone. Hence, Y_k is a set of n mobile phones

With this instance of generalized version, it is obvious that all the existing methodologies for classical derangement problem might not function correctly as we expect. Moreover, the previous definition of derangement does not fit with this version of the problem since the value of k is not necessary to be $k = 1$. We then categorize a derangement into two types: **complete derangement** and **partial derangement**.

Definition 4 Let an instance of the hat-check problem be $n \in \mathbb{N}$ people and $k \in \mathbb{N}$ items each. An arrangement $\mathscr{C}'$ is called **complete derangement** if for every ordered tuple $(x_i, y_{[1,j_1]}, \ldots, y_{[k,j_k]}) \in \mathscr{C}', \bigwedge_{\lambda=1}^{k} i \neq j_\lambda$.

Definition 5 Let an instance of the hat-check problem be $n \in \mathbb{N}$ people and $k \in \mathbb{N}$ items each. An arrangement $\mathscr{C}'$ is called **partial derangement** if for every ordered tuple $(x_i, y_{[1,j_1]}, \ldots, y_{[k,j_k]}) \in \mathscr{C}', \bigvee_{\lambda=1}^{k} i \neq j_\lambda$.

Definition 6 Let an instance of the hat-check problem be $n \in \mathbb{N}$ people and $k \in \mathbb{N}$ items each. An arrangement $\mathscr{C}'$ is called **non-derangement** if there exists an ordered tuple $(x_i, y_{[1,j_1]}, \ldots, y_{[k,j_k]}) \in \mathscr{C}', \bigwedge_{\lambda=1}^{k} i = j_\lambda$.

Definition 7 Let an instance of the hat-check problem be $n \in \mathbb{N}$ people and $k \in \mathbb{N}$ items each. An arrangement $\mathscr{C}'$ is called **antiderangement** if for every ordered tuple $(x_i, y_{[1,j_1]}, \ldots, y_{[k,j_k]}) \in \mathscr{C}', \bigwedge_{\lambda=1}^{k} i = j_\lambda$.

We use the following scenario to explain the difference between those two types of derangement. Let the instance be $n = 2$ customers and $k = 2$ items for each customer. To reduce ambiguity, we use Y rather than Y_1 and Z rather than Y_2. Particularly:

$$X = \{x_1, x_2\}$$
$$Y = \{y_1, y_2\}$$
$$Z = \{z_1, z_2\}$$

Initially, every customer comes with his own items, so the input collection would be $\mathscr{C}_{\text{in}} = \{(x_1, y_1, z_1), (x_2, y_2, z_2)\}$. The total arrangements, which is in a form of a collection of ordered triples (x_i, y_i, z_i), that can happen are:

$$\mathscr{C}'_1 = \{(x_1, y_1, z_1), (x_2, y_2, z_2)\}$$
$$\mathscr{C}'_2 = \{(x_1, y_1, z_2), (x_2, y_2, z_1)\}$$
$$\mathscr{C}'_3 = \{(x_1, y_2, z_1), (x_2, y_1, z_2)\}$$
$$\mathscr{C}'_4 = \{(x_1, y_2, z_2), (x_2, y_1, z_1)\}$$

From the list of outcomes above, $\mathscr{C}'_1$ is both non-derangement and antiderangement since every customer completely gets his own items. $\mathscr{C}'_2$ and $\mathscr{C}'_3$ are partial derangement since every customer gets at least one item that belongs to the other customers. Lastly, $\mathscr{C}'_4$ is the only complete derangement where every customer does not get his own items. The total number of arrangements is not computably difficult. Given n people and k items, the total number of arrangements, denoted by $\mathbb{S}(n,k)$, can be computed by:

$$\mathbb{S}(n,k) = (n!)^k$$

So for the example case ($n = k = 2$), we then get $\mathbb{S}(2,2) = (2!)^2 = 4$.

The next example that we would like to demonstrate is the case that $n = 2$ and $k = 3$. Again, we use Y rather than Y_1 and Z rather than Y_2 and W rather than Y_3. Hence, the total number of arrangements is $\mathbb{S}(2,3) = (2!)^3 = 8$. The list of those eight arrangements is:

$$\mathscr{C}'_1 = \{(x_1, y_1, z_1, w_1), (x_2, y_2, z_2, w_2)\}$$
$$\mathscr{C}'_2 = \{(x_1, y_1, z_1, w_2), (x_2, y_2, z_2, w_1)\}$$
$$\mathscr{C}'_3 = \{(x_1, y_1, z_2, w_1), (x_2, y_2, z_1, w_2)\}$$
$$\mathscr{C}'_4 = \{(x_1, y_1, z_2, w_2), (x_2, y_2, z_1, w_1)\}$$
$$\mathscr{C}'_5 = \{(x_1, y_2, z_1, w_1), (x_2, y_1, z_2, w_2)\}$$
$$\mathscr{C}'_6 = \{(x_1, y_2, z_1, w_2), (x_2, y_1, z_2, w_1)\}$$
$$\mathscr{C}'_7 = \{(x_1, y_2, z_2, w_1), (x_2, y_1, z_1, w_2)\}$$
$$\mathscr{C}'_8 = \{(x_1, y_2, z_2, w_2), (x_2, y_1, z_1, w_1)\}$$

Clearly from the list, $\mathscr{C}'_1$ is antiderangement. $\mathscr{C}'_8$ is complete derangement. $\mathscr{C}'_2$ to $\mathscr{C}'_7$ are partial derangement. Observe that for $n = 2$, the arrangement that is non-derangement but not antiderangement does not exist.

To show readers every type that we have a definition for, we demonstrate the last example with the case that $n = 3$ and $k = 2$. Again, we use Y rather than Y_1 and Z rather than Y_2. Hence, the total number of arrangements is $\mathbb{S}(3,2) = (3!)^2 = 36$. The list of those eight arrangements is:

$\mathscr{C}'_1 = \{(x_1,y_1,z_1),(x_2,y_2,z_2),(x_3,y_3,z_3)\}$ ♣
$\mathscr{C}'_2 = \{(x_1,y_1,z_1),(x_2,y_2,z_3),(x_3,y_3,z_2)\}$ ♦
$\mathscr{C}'_3 = \{(x_1,y_1,z_1),(x_2,y_3,z_2),(x_3,y_2,z_3)\}$ ♦
$\mathscr{C}'_4 = \{(x_1,y_1,z_1),(x_2,y_3,z_3),(x_3,y_2,z_2)\}$ ♦

$\mathscr{C}'_5 = \{(x_1,y_1,z_2),(x_2,y_2,z_1),(x_3,y_3,z_3)\}$ ♦
$\mathscr{C}'_6 = \{(x_1,y_1,z_2),(x_2,y_2,z_3),(x_3,y_3,z_1)\}$ ♡
$\mathscr{C}'_7 = \{(x_1,y_1,z_2),(x_2,y_3,z_1),(x_3,y_2,z_3)\}$ ♡
$\mathscr{C}'_8 = \{(x_1,y_1,z_2),(x_2,y_3,z_3),(x_3,y_2,z_1)\}$ ♡

$\mathscr{C}'_9 = \{(x_1,y_1,z_3),(x_2,y_2,z_1),(x_3,y_3,z_2)\}$ ♡
$\mathscr{C}'_{10} = \{(x_1,y_1,z_3),(x_2,y_2,z_2),(x_3,y_3,z_1)\}$ ♦
$\mathscr{C}'_{11} = \{(x_1,y_1,z_3),(x_2,y_3,z_1),(x_3,y_2,z_2)\}$ ♡
$\mathscr{C}'_{12} = \{(x_1,y_1,z_3),(x_2,y_3,z_2),(x_3,y_2,z_1)\}$ ♡

$\mathscr{C}'_{13} = \{(x_1,y_2,z_1),(x_2,y_1,z_2),(x_3,y_3,z_3)\}$ ♦
$\mathscr{C}'_{14} = \{(x_1,y_2,z_1),(x_2,y_1,z_3),(x_3,y_3,z_2)\}$ ♡
$\mathscr{C}'_{15} = \{(x_1,y_2,z_1),(x_2,y_3,z_2),(x_3,y_1,z_3)\}$ ♡
$\mathscr{C}'_{16} = \{(x_1,y_2,z_1),(x_2,y_3,z_3),(x_3,y_1,z_2)\}$ ♡

$\mathscr{C}'_{17} = \{(x_1,y_2,z_2),(x_2,y_1,z_1),(x_3,y_3,z_3)\}$ ♦
$\mathscr{C}'_{18} = \{(x_1,y_2,z_2),(x_2,y_1,z_3),(x_3,y_3,z_1)\}$ ♡
$\mathscr{C}'_{19} = \{(x_1,y_2,z_2),(x_2,y_3,z_1),(x_3,y_1,z_3)\}$ ♡
$\mathscr{C}'_{20} = \{(x_1,y_2,z_2),(x_2,y_3,z_3),(x_3,y_1,z_1)\}$ ♠

$\mathscr{C}'_{21} = \{(x_1,y_2,z_3),(x_2,y_1,z_1),(x_3,y_3,z_2)\}$ ♡
$\mathscr{C}'_{22} = \{(x_1,y_2,z_3),(x_2,y_1,z_2),(x_3,y_3,z_1)\}$ ♡
$\mathscr{C}'_{23} = \{(x_1,y_2,z_3),(x_2,y_3,z_1),(x_3,y_1,z_2)\}$ ♠
$\mathscr{C}'_{24} = \{(x_1,y_2,z_3),(x_2,y_3,z_2),(x_3,y_1,z_1)\}$ ♡

$\mathscr{C}'_{25} = \{(x_1,y_3,z_1),(x_2,y_1,z_2),(x_3,y_2,z_3)\}$ ♡
$\mathscr{C}'_{26} = \{(x_1,y_3,z_1),(x_2,y_1,z_3),(x_3,y_2,z_2)\}$ ♡
$\mathscr{C}'_{27} = \{(x_1,y_3,z_1),(x_2,y_2,z_2),(x_3,y_1,z_3)\}$ ♦
$\mathscr{C}'_{28} = \{(x_1,y_3,z_1),(x_2,y_2,z_3),(x_3,y_1,z_2)\}$ ♡

$\mathscr{C}'_{29} = \{(x_1,y_3,z_2),(x_2,y_1,z_1),(x_3,y_2,z_3)\}$ ♡
$\mathscr{C}'_{30} = \{(x_1,y_3,z_2),(x_2,y_1,z_3),(x_3,y_2,z_1)\}$ ♠
$\mathscr{C}'_{31} = \{(x_1,y_3,z_2),(x_2,y_2,z_1),(x_3,y_1,z_3)\}$ ♡
$\mathscr{C}'_{32} = \{(x_1,y_3,z_2),(x_2,y_2,z_3),(x_3,y_1,z_1)\}$ ♡

$\mathscr{C}'_{33} = \{(x_1,y_3,z_3),(x_2,y_1,z_1),(x_3,y_2,z_2)\}$ ♠
$\mathscr{C}'_{34} = \{(x_1,y_3,z_3),(x_2,y_1,z_2),(x_3,y_2,z_1)\}$ ♡
$\mathscr{C}'_{35} = \{(x_1,y_3,z_3),(x_2,y_2,z_1),(x_3,y_1,z_2)\}$ ♡
$\mathscr{C}'_{36} = \{(x_1,y_3,z_3),(x_2,y_2,z_2),(x_3,y_1,z_1)\}$ ♦

From the provided list, an arrangement marked with ♣ is antiderangement. An arrangement marked with ♦ is non-derangement. An arrangement marked with ♡ is partial derangement. Lastly, an arrangement marked with ♠ is complete derangement. For the case of $n = 3$ and $k = 2$, only one arrangement ($\mathscr{C}'_1$) is antiderangement. There are nine arrangements that are non-derangement but not antiderangement ($\mathscr{C}'_2,\mathscr{C}'_3,\mathscr{C}'_4,\mathscr{C}'_5,\mathscr{C}'_{10},\mathscr{C}'_{13},\mathscr{C}'_{17},\mathscr{C}'_{27}$ and $\mathscr{C}'_{36}$). There are four arrangements that are complete derangement ($\mathscr{C}'_{20},\mathscr{C}'_{23},\mathscr{C}'_{30}$ and $\mathscr{C}'_{33}$). The remaining twenty-two arrangements are partial derangement but not complete derangement.

Currently, as of our best knowledge, there exists no DP algorithm that is capable of tackling the generalized version of the hat-check problem where an instance of the problem contains $n \in \mathbb{N}$ customers and $k \in \mathbb{N}$ items.

6 Conclusions and Possible Future Direction

We report an algorithm that computes the total number of arrangements that are derangement with an input $n \in \mathbb{N}$ by using a concept of DP, which known to be efficient for several applications. An instance of the derangement problem comes with inputs $X = \{x_1,...,x_n\}$ and $Y = \{y_1,...,y_n\}$. The problem inquires how many $\mathscr{C}' = \{(x_i,y_j) \in X \times Y\}$ such that $i \neq j$ and every $x_i \in X$ shows up exactly once as well as every $y_j \in Y$. We invoke a methodology to compute from PatanasakPinyo and Sulaiman [14], which has an objective to extract a recursive call to solve a subproblem with smaller instance. The output of the proposed algorithm is correct and the algorithm itself has a polynomial running time, $O(n^2)$, that proves its efficiency.

We initiate an idea to generalize the hat-check problem by allowing the instance of the problem to be $n \in \mathbb{N}$ customers and $k \in \mathbb{N}$ items rather than $k = 1$, which is a default setup of the classic instance of the problem. A door is opened for theoretical computer scientists to come up with an efficient algorithm that overcomes this generalized version.

Furthermore, the derangement problem is potentially a key to several applications that need to rearrange a finite set such that the input arrangement should not to be returned. The example is designing a sequence of addresses for the address verification task [1, 6, 8, 9]. Another good example is designing a matching problems for the paper folding test, which is a test of individual spatial visualization (VZ) [10, 11, 12, 13, 16].

Acknowledgments

This study is partially supported by the Faculty of Information and Communication Technology, Mahidol University, Thailand.

References

[1] Georgi Batinov, Michelle Rusch, Tianyu Meng, Kofi Whitney, Thitivatr Patanasakpinyo, Les Miller, and Sarah Nusser. Understanding map operations in location-based surveys. In *Eighth International Conference on Advances in Computer-Human Interactions (ACHI 2015)*, pages 144–149, Lisbon, Portugal, 2015. International Academy, Research, and Industry Association (IARIA).

[2] Mehdi Hassani. Derangements and applications. *Journal of Integer Sequences*, 6(1):03–1, 2003.

[3] Mehdi Hassani. Derangements and alternating sum of permutations by integration. *Journal of Integer Sequences*, 23(2):3, 2020.

[4] Pallavi Joshi, Chang-Seo Park, Koushik Sen, and Mayur Naik. A randomized dynamic program analysis technique for detecting real deadlocks. *ACM Sigplan Notices*, 44(6):110–120, 2009.

[5] Jon Kleinberg and Eva Tardos. *Algorithm Design*. Pearson Education India, 2006.

[6] Thitivatr PatanasakPinyo. *Flattening methods for adaptive location-based software to user abilities*. Graduate Theses and Dissertations, Iowa State University, 2017.

[7] Thitivatr Patanasakpinyo. Model checking approach for deadlock detection in an operating system process-resource graph using dynamic model generating and computation tree logic specification. In Gordon Lee and Ying Jin, editors, *Proceedings of 34th International Conference on Computers and Their Applications*, volume 58 of *EPiC Series in Computing*, pages 55–64. EasyChair, 2019.

[8] Thitivatr Patanasakpinyo. Ameliorating accuracy of a map navigation when dealing with different altitude traffics that share exact geolocation. In Alex Redei, Rui Wu, and Frederick Harris, editors, *SEDE 2020. 29th International Conference on Software Engineering and Data Engineering*, volume 76 of *EPiC Series in Computing*, pages 95–104. EasyChair, 2021.

[9] Thitivatr PatanasakPinyo. Exploiting a real-time non-geolocation data to classify a road type with different altitudes for strengthening accuracy in navigation. *International Journal of Computers and Their Applications*, 28(1):55–64, 2021.

[10] Thitivatr PatanasakPinyo, Georgi Batinov, Kofi Whitney, and Les Miller. Methods that flatten the user space for individual differences in location-based surveys on portable devices. In *31st International Conference on Computers and Their Applications (CATA 2016)*, pages 65–70, Las Vegas, Nevada, 2016. International Society for Computers and their Applications (ISCA).

[11] Thitivatr Patanasakpinyo, Georgi Batinov, Kofi Whitney, Adel Sulaiman, and Les Miller. Object-indexing: A solution to grant accessibility to a traditional raster map in location-based application to accomplish a location-based task. *International Journal of Computing, Communication and Instrumentation Engineering (IJCCIE)*, 5(1), 2018.

[12] Thitivatr Patanasakpinyo, Georgi Batinov, Kofi Whitney, Adel Sulaiman, and Les Miller. Enhanced prediction models for predicting spatial visualization (vz) in address verification task. In Gordon Lee and Ying Jin, editors, *Proceedings of 34th International Conference on Computers and Their Applications*, volume 58 of *EPiC Series in Computing*, pages 247–256. EasyChair, 2019.

[13] Thitivatr Patanasakpinyo and Les Miller. Ui error reduction for high spatial visualization users when using adaptive software to verify addresses. In Gordon Lee and Ying Jin, editors, *Proceedings of 35th International Conference on Computers and Their Applications*, volume 69 of *EPiC Series in Computing*, pages 22–31. EasyChair, 2020.

[14] Thitivatr Patanasakpinyo and Adel Sulaiman. Alternative approach to achieve a solution of derangement problems by dynamic programming. In Ajay Bandi, Mohammad Hossain, and Ying Jin, editors, *Proceedings of 38th International Conference on Computers and Their Applications*, volume 91 of *EPiC Series in Computing*, pages 98–107. EasyChair, 2023.

[15] Stefano Pironio, Antonio Acín, Serge Massar, A Boyer de La Giroday, Dzmitry N Matsukevich, Peter Maunz, Steven Olmschenk, David Hayes, Le Luo, T Andrew Manning, et al. Random numbers certified by bell's theorem. *Nature*, 464(7291):1021–1024, 2010.

[16] Timothy A Salthouse, Renee L Babcock, Debora RD Mitchell, Roni Palmon, and Eric Skovronek. Sources of individual differences in spatial visualization ability. *Intelligence*, 14(2):187–230, 1990.

[17] Richard P Stanley. What is enumerative combinatorics? In *Enumerative combinatorics*, pages 1–63. Springer, 1986.

[18] Eric W Weisstein. Derangement. *https://mathworld. wolfram. com/*, 2002.

Thitivatr PatanasakPinyo is a Teaching Professor at the Faculty of Information and Communication Technology, Mahidol University. He received his PhD in Computer Science from Iowa State University of Science and Technology, Iowa, USA under a supervision of Prof. Dr. Les Miller and Prof. Dr. Pak Tavanapong. His dissertation is “Flattening methods for adaptive location-based software to user abilities”. His research area consists of HCI theory, algorithm, and theoretical computer science.

Adel Sulaiman is an Assistant Professor at the College of Computer Science and Information Systems, Najran University. He received his PhD in Computer Science from Iowa State University of Science and Technology, Iowa, USA under a supervision of Prof. Dr. Les Miller and Assoc. Prof. Dr. Stephen Gilbert. His research area consists of HCI and usability.

Energy Efficient Vice Low Adaptive Hierarchy Clustering Protocol: EEV-LEACH

Amira Bendjeddou*, Mouna Hemici†
University of Annaba, Algeria.

Abstract

The primary objectives for many sensor network applications are to reduce energy consumption and extend network lifetime. These goals have pushed the scientific community to develop new solutions to reduce total energy consumption while maintaining network performance, such as hierarchical routing protocols. This paper proposes an energy-efficient hierarchical routing protocol for Wireless Sensor Networks (WSNs) called EEV-LEACH (Energy Efficient Vice Low Adaptive Clustering Hierarchy) to improve the LEACH protocol. The EEV-LEACH protocol aims to extend the network lifetime by reducing energy consumption at each sensor node and cluster-head (CH). Minimizing the wasted energy by each sensor node is accomplished by minimizing the periodic selection of cluster heads (CHs) in each round; thus, the number of association messages exchanged between the CH and the nodes is reduced, reducing consumed energy and overhead. EEV-LEACH also reduces the energy consumed by CHs by employing vice-CHs, which share the workload with the CHs in an alternative way. The MATLAB simulator simulation demonstrates that the EEV-LEACH protocol extends network lifetime and reduces overall overhead compared to the LEACH, LEACH-S, andTL-LEACH protocols.

Key words: Wireless Sensor Network, hierarchical Routing Protocol, Network Lifetime, LEACH, overhead, vice-CH.

1 Introduction

Wireless Sensor Networks (WSNs) represent a relatively new technology that has emerged due to tremendous technological advances in developing smart sensors, powerful processors, and wireless communication protocols. This type of network comprises a few/thousand sensors that collect, process, and broadcast environmental data to the outside world [1]. WSNs are useful for various field applications [2][3]. However, energy conservation is the primary requirement of these networks because it directly affects the lifetime of network nodes. In fact, because of their small size, low cost, and deployment in hostile or difficult-to-reach areas, sensor nodes have several flaws [4], including limited network lifetime because sensor batteries are not typically rechargeable, low bandwidth, and limited acquisition and communication capabilities. Several research questions, primarily related to optimizing energy consumption to improve network lifetime, have emerged in recent years to address these limitations [5][6].

LEACH is a well-known energy-efficient hierarchical clustering routing protocol that operates in rounds of two phases. In the first phase (construction phase), CHs are chosen randomly from a pool of nodes, and clusters are formed. The data is forwarded to the base station (BS) using TDMA schedules in the second communication phase. The CH is changed in each round of the LEACH protocol. The periodic election of CHs results in excessive energy consumption. The exchange of association messages between CHs and nodes causes each CH election. It also causes overhead, particularly in dense networks. A hierarchical routing protocol for WSNs is proposed in this paper. This latter is a variant of the LEACH (Low Energy Adaptive Clustering Hierarchy) protocol based on the clustering approach. LEACH is one of the most popular hierarchical routing algorithms for WSNs.

The proposed energy-efficient protocol aims to increase the network's lifetime while decreasing overall overhead. Minimizing the selection of CHs in each round reduces the energy consumption of sensor nodes. When the periodic selection of CHs is limited, the number of association messages exchanged between the CH and the nodes is also reduced. The result is a reduction in energy consumption and overheads. Energy consumption by CHs can be reduced by using vice-CHs, who will share the workload with CHs in an alternative way. This can also help to reduce the amount of overhead generated in CH. The rest of this paper is structured as follows: Section 2 describes the most well-known LEACH variants, and Section 3 describes the proposed protocol EEV-LEACH. Section 4 evaluates the proposed protocol, and Section 5 concludes the paper.

* †LRS Laboratory, Department of Computer Science
*Amira.bendjeddou@univ-annaba.org

2 LEACH protocol and its variants

LEACH is a popular hierarchical clustering protocol in WSNs. LEACH has many advantages, but there are also some disadvantages. Many variants of the LEACH protocol have been introduced to address these shortcomings; these are discussed further below. [7] proposes LEACH, one of the first energy-efficient routing protocols and still the leading protocol in WSNs. The basic idea behind LEACH is to select CH from a number of nodes by rotation to extend the communication's energy dissipation to all nodes in the network. The operation is divided into two phases: set-up and steady-state. During the set-up phase, each node decides whether or not to become a CH for the current node round, which is determined by the percentage of CH proposed and the number of times a node has been CH. A random number between 0 and 1 is chosen; if the number is less than the threshold described in Eq. (1), the node is designated as a CH:

$$Tn = \begin{cases} \frac{P}{1-P(r \bmod \frac{1}{P})}, & if n \in G \\ 0 & else \end{cases} \qquad (1)$$

where, P is the percentage of CHs, r represents the current round, and G is the member nodes that were not selected as CHs in the last $1/P$ rounds.

The elected CH will advertise the other nodes, and the nodes will decide which cluster to join and send a membership message based on the signal strengths. The role of CH is rotated to maximize energy efficiency. The steady-state nodes receive and transmit data to their CH, aggregating and transmitting data to the Base Station (BS). The TDMA/CDMA MAC mechanism is used to avoid collisions. The authors of [8] proposed LEACH-C (Centralized Low Energy Adaptive Clustering Hierarchy). The primary distinction between LEACH and LEACH-C is in the set-up phase. The BS chooses the CH. Each node transmits its current location and energy level to the BS, which uses this information to select a CH. The BS chooses the node with the highest energy level to be the cluster's leader and broadcasts this information to all nodes in the network. The deterministic method of selecting the number of CH nodes in each round, which is predetermined at deployment, gives this protocol an advantage over LEACH. LEACH-C improves the network's distribution of CH nodes. LEACH-C, on the other hand, requires current position information from all nodes via GPS, which is unreliable.

[9] introduced TL-LEACH (Two Levels Low Energy Adaptive Clustering Hierarchy). In contrast to the LEACH protocol, the CH sends data to the BS in a single hop. The TL-LEACH protocol is organized in a two-level hierarchy. Instead of being sent directly to the BS, the CH collected the aggregated data from each CH located between the CH and the BS. This protocol has the advantage of reducing data transmission energy. The CH node dies faster than other nodes farther from the BS. TL-LEACH improves energy efficiency by using a Head cluster as a relay node between Head clusters.

The authors of [10] proposed Cell- LEACH (Cell level Low Energy Adaptive Clustering Hierarchy). In this protocol, the area is divided into seven cells, each containing nodes; one is chosen as the cell's head. There is no reassembling once formed. Using a TDMA schedule, each cell node sends data to the cell head in a time slot. The cell heads perform the data aggregation function, and the processed data is sent to the CHs. Data is transferred to the BS by CHs and cell heads. Following the first round, the cell and CHs are randomly chosen.

V-LEACH (Vice Low Energy Adaptive Clustering Hierarchy) was described by [11]. The CH uses more energy than the normal nodes in the LEACH protocol to send aggregated data to the BS (located a long distance away). As a result, the CH dies prematurely, rendering the entire cluster inoperable and resulting in data loss. V- LEACH mitigates this disadvantage by having a vice-CH in each cluster who takes over as CH when the latter dies. This protocol reduces the overhead generated by selecting a new CH each time a CH dies, increasing network lifetime because data is always delivered to the BS. M- LEACH (Mobile Low Energy Adaptive Clustering Hierarchy) is presented by the authors of [12]. Contrary to LEACH, this protocol ensures node mobility in the second phase. The nodes are homogeneous, and each node's position is calculated using GPS. The CH nodes are chosen based on their low mobility and attenuation, and their role is broadcast to all nodes within their transmission range.

In [13], the authors discuss EE-LEACH (Energy-Efficient Low Energy Adaptive Clustering Hierarchy). This protocol optimizes CH selection and data transmission by considering each node's energy and position. The node with the greatest residual energy will be chosen as the CH. Before sending the data to the BS, the CH aggregates it. It also sends the status of the nodes with the most residual energy. During the data transmission phase, some neighboring CHs are chosen as relay nodes to balance the communication's energy consumption. Compared to LEACH, this protocol is more stable and has a longer network lifetime. The authors propose LEACH-B (Balanced Low Energy Adaptive Clustering Hierarchy) in [14]. LEACH-B employs a decentralized approach to cluster formation in which each sensor node is aware of its position and the position of its final destination, regardless of the positions of other nodes in the network. LEACH-B is divided into CH selection, cluster formation, and multiple access data transmission. Each node chooses its CH based on the energy consumed between the node and the final receiver. LEACH-B uses less energy than the LEACH protocol.

LEACH-S (Low Energy Adaptive Clustering Hierarchy for Sensor Networks) was introduced by [15]. The LEACH protocol causes nodes to consume more energy, resulting in significant network overhead.

LEACH-S mitigates these drawbacks by reducing distributed energy and lowering overhead by reducing periodic CH selection.

The primary distinction between LEACH and LEACH-S is in the set-up phase. The CH checks its remaining energy in the second round to see if it is greater than a calculated threshold. In this case, it continues its activity as a CH; otherwise, another CH is chosen. The calculated threshold is shown in Eq. (2):

$$E_{moy} = (\sum_{i=1}^{i=n} E_i)/n \quad (2)$$

where, n is the number of nodes and E_i is the residual energy of the node in the cluster.

In [16], the authors discuss MOD-LEACH (Modified Low Energy Adaptive Clustering Hierarchy). The following are the primary distinctions between MOD-LEACH and LEACH: changing the CH in MOD-LEACH is unnecessary as long as it does not contain more energy than the required threshold. MOD-LEACH also does not equalize the amplitude of all signals. In LEACH, the CH is changed after each round to prevent the CH from dying prematurely, whereas, in MOD-LEACH, the current CH is replaced by a new one only if the energy of the current CH is greater than the required threshold. This reduces the energy required for cluster formation and the transmission of routing packets to a new CH. If the current CH's residual energy exceeds the minimum threshold value in each round, the current CH will remain the CH for the next round. MOD-LEACH categorizes communication into three types: intra-cluster communication, inter-cluster communication, and data transmission from CH to BS. The energy required for intra-cluster communication differs from that required for inter-cluster communication and communication from a CH to a BS. As a result, different types of amplification are required depending on the packet type. Previously, regardless of the communication type, all LEACH packets were amplified similarly.

The authors of [17] proposed MH-LEACH (Multi-Hop Low Energy Adaptive Clustering Hierarchy). LEACH has no effect on the distance between the BS and the CH. Data is transmitted from the CH to the BS in a single hop. The distance between the CH and the BS grows as the network diameter grows. As a result, energy consumption rises with distance. This protocol broadcasts data from the CH to the BS using Multi-Hop communication. The data is passed from one CH to the next, and so on, until it reaches the BS. The data is transmitted to the BS by the CH located near the BS. When CHs receive data, they aggregate it to reduce the total amount of data transmitted by the network. The authors introduce IB-LEACH (Low Energy Adaptive Clustering Hierarchy) in [18]. IB- LEACH is intended to reduce intra-group communication costs and CH load by dividing the task among the CH and its group members. IB-LEACH operates in rounds, each divided into three phases: set-up, pre-stabilized, and stabilized. The LEACH phase is similar to the set-up phase (the build phase). During the pre- stable phase, cluster nodes are classified into CH, sensing nodes, and aggregators. Sensor nodes sense their surroundings and transmit the information they gather to aggregators. The received data is aggregated and sent to the BS by the aggregators. This process reduces the CHs' energy consumption. The cluster activities are maintained and managed by the CHs. They develop and distribute the TDMA schedule to all cluster members. The CHs also choose the aggregator nodes in a given frame and broadcast the list to all cluster members. There are frames in the steady-state process. Each cluster member sends data in each frame according to its time slot. This data is aggregated by the aggregator and sent to the BS.

H-LEACH (Hybrid Low Energy Adaptive Clustering Hierarchy) is proposed by the authors of [19]. H-Leach divides the network's nodes into clusters, and the LEACH protocol is used to select a CH in each partition using a probabilistic method. This can lead to better CH distribution and a longer network lifetime. The authors of [20] proposed LEACH-E (Energy Low Energy Adaptive Clustering Hierarchy). In the LEACH-E protocol, all nodes start with the same amount of energy and the same chance of becoming a CH. Each node's energy level changes after the first round. The residual energy of each node is then used to select the CHs. The nodes with the highest residual energy are prioritized over others. LEACH-E extends the network's lifetime by balancing the energy load across all nodes. [21] proposes LEACH-F (Fixed number of cluster Low Energy Adaptive Clustering Hierarchy). LEACH-F, like LEACH-C, employs a centralized approach to cluster formation. After the cluster formation process is completed, there is no re-clustering phase in the following round. Only the rotation of CH nodes within its clusters is possible because clusters are fixed. LEACH-F reduces the overhead of the LEACH construction phase by forming a fixed number of clusters maintained throughout the network lifetime. However, once a cluster is formed, nodes cannot be added or removed, and nodes cannot adjust their behavior despite the death of nodes.

I- LEACH (Improved Low Energy Adaptive Clustering Hierarchy) was designed by the authors of [22]. I-LEACH is designed with two significant changes. First, instead of the probability used in LEACH, residual energy is used to select the CH, allowing it to be used for sensor nodes with varying initial energy. Second, nodes collaborate to form clusters, with one CH located near each sensor node. The LEACH protocol does not specify this criterion. LEACH- K (K-medoids low energy adaptive clustering hierarchy) was proposed in [23]. LEACH-K is proposed to eliminate the disadvantages of LEACH, specifically the random selection of CHs, which can result in very poor cluster formation. The K-LEACH protocol improves the cluster formation and head cluster selection procedures in the first phase because it uses the K-medoids

Protocols	*Data transmission*	*Mobility of BS*	*Homogenous*	*Needs for location information*	*Scalability*	*Distributed*	*Auto- organisation*
LEACH	Single-Hop	Fixed	Yes	Yes	Limited	Yes	Yes
LEACH-C	Single-Hop	Fixed	Yes	Yes	Good	No	Yes
TL-LEACH	Multi-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
Cell-LEACH	Multi-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
V-LEACH	Single-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
M-LEACH	Single-Hop	Mobile	Yes	Yes	Very good	Yes	Yes
EE-LEACH	Single-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
LEACH-B	Single-Hop	Fixed	Yes	Yes	Good	Yes	Yes
LEACH-S	Single-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
MOD-LEACH	Single-Hop	Fixed	Yes	Yes	Good	Yes	Yes
MH-LEACH	Multi-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
IB-LEACH	Single-Hop	Fixed	Yes	Yes	Good	Yes	Yes
H-LEACH	Single-Hop	Fixed	Yes	Yes	Good	Yes	Yes
I-LEACH	Single-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
F-LEACH	Single-Hop	Fixed	Yes	Yes	Limited	No	No
LEACH-E	Single-Hop	Fixed	Yes	Yes	Very good	Yes	Yes
LEACH-K	Single-Hop	Fixed	Yes	Yes	Good	Yes	Yes
EEM-LEACH	Multi-Hop	Fixed	Yes	Yes	Good	Yes	Yes

algorithm for efficient cluster formation at the start of each round and then selects the head of clusters using the Euclidean distance to the nearest center of each cluster. The maximum residual energy (MRE) is used until each cluster has a unique CH. LEACH-K's steady state is the same as LEACH's. [24] proposes EEM-LEACH (Energy Efficient Multi-Hop Low Energy Adaptive Clustering Hierarchy). It addresses the issue of direct communication between CHs and BSs and poor CH selection in the LEACH protocol. The EEM-LEACH protocol is distinguished by:

1) The selection of CH is based on residual energy and node average energy consumption.
2) Multi-hop inter-cluster communication chooses the shortest multi-hop path from each node to the BS.
3) Direct communication by nodes near the BS (if communication costs are low). Utilize relay nodes with higher residual energy to reduce communication costs per packet. This has the potential to increase network lifetime. Table 1 summarizes the protocols cited in the related work.

3 The Proposed protocol

3.1 Motivation

LEACH Protocol is a well-known hierarchical routing protocol. It is well-known in homogeneous wireless sensor networks for its energy efficiency. The LEACH protocol is divided into rounds of two phases: set-up and steady. The cluster head is elected randomly every round, forming a new cluster. Because of the association messages exchanged between nodes, this results in excessive energy consumption.

Furthermore, this adds unnecessary overhead to CHs and BS. If a CH did not expend much energy in the previous round, there is no need to elect a new CH. As a result, an energy-efficient CH replacement based on the current CH's residual energy is required. By taking into account the residual energy of CHs, the reflection of CHs is limited; as a result, the number of association messages is reduced; as a result, the consumed energy and overhead are reduced. Many LEACH protocol variants have been proposed to reduce the energy consumed by the periodic selection of CH. Some of them proposed the vice-CH position. When CH died, it took his place. However, this vice cluster only operates when the CH's energy runs out, and there is no reflection of the cluster or vice cluster when both die. Other protocols seek to replace the CH only when its energy level exceeds a certain threshold. This can increase network lifetime but generate an overhead at CH, especially in a homogeneous network, so residual energy is quickly depleted. This study proposes EEV-LEACH (Energy Efficient Vice Low Energy Adaptive Clustering Hierarchy), an energy-efficient clustering protocol to address the aforementioned issues.

3.2 EEV-LEACH protocol description

This work presents a novel hierarchical routing protocol, which is a new variant of LEACH. The proposed protocol is carried out in rounds, which include the set-up and steady phases. This contribution is primarily focused on the initial set-up. Two CHs are chosen in the first round: the main CH and the vice-CH. They both work in tandem. Based on certain criteria, CH and vice-CH are replaced in the second round. The residual energy of CH or vice-CH, who is responsible for this round, is compared to a calculated threshold for each subsequent round. If this latter value is less than the required threshold, a new CH/vice-CH is selected. The following section contains the specifics of the proposed protocol.

3.2.1 Protocol functioning

This paper only describes the set-up phase, as the steady phase aligns with the LEACH protocol.

In the first round

The CH is chosen in the same manner as in the LEACH protocol. After the cluster formation process is completed, the CH selects an assistant CH (vice-CH) from among its members. As an assistant, it selects the member closest to the BS. The chosen assistant serves as CH in the second round.

From the second round

The CH and vice-CH work in tandem; if the CH is active in the current round, the vice cluster will be active in the following round. The residual energy of the CH or vice- CH is compared to a calculated threshold (ENTR) in the second round. A new CH/vice-CH (the CH/vice-CH with the highest residual energy) is chosen if it is less than the required threshold. The required threshold is represented by Eq. (3):

$$ENTR = ((Et + Er) * n)/nbrcluster \qquad (3)$$

where, *Et* is the necessary energy for transmitting a packet, *Er* is the necessary energy for receiving a packet, *N* is the number of nodes in the network, and *Nbr cluster* represents the number of formed clusters.

The following pseudo-code summarizes the steps of the proposed protocol.

3.2.2.1 Pseudo-code of the proposed protocol (EEV-LEACH)

1st round: (r=0) Begin

While the nodes are alive, **do**

Begin

Random selection of the CHs

The elected CH sends a message to the other nodes.

Nodes decide which cluster to join

CH chooses the nearest node to the BS as an assistant (vice-CH)

Nodes collect and forward data to their CH CH aggregates and

sends data directly to BS

2nd round (r=1)

Begin

If (r mod 2! = 0)

Begin

If (Evch >= ((ET+ Er) * n) /nb cluster)

Stay as VCH

Else

Choose a new VCH

End if

Nodes collect and transmit data to its V-CH

The VCH aggregates and sends the data directly to BS

End else

Else

Begin

if (Ech >= ((ET+ Er) * n) /nb cluster)

Stay as CH

Else

Choose a new VCH

End if

Nodes collect and transmit data to their current CH

The CH aggregates and sends the data directly to BS

End else

End

End while

END

4 Results and discussion

This section describes the proposed protocol's performance evaluation. The evaluations are carried out with the help of a MATLAB simulator. A number of nodes are randomly distributed over a 100m*100m area. Every node is static and homogeneous. They are powered by a 0.5 joule battery. The BS is located in the area's center. Table 2 summarizes the simulation parameters. The proposed protocol is evaluated based on several metrics.

Consumed energy:

We are interested in the energy consumed during reception and transmission. The following formulas are used to calculate the consumed energy (Kirankumar & Katageri, 2014):

- *Consumed energy for transmission:*

$ETxs, d = ETxelecs + ETxamps, d$ (4)

$ETxs, d = ETxelec*s + (Eamp*s*d2)$ (5)

- *Consumed energy at reception:*

$ERxs = ERx\ elec*\ s$ (6)

$ERxs = ERxelec*s$ (7)

where, *Eelec* and *Eamp* represent the electronic transmission energy and amplification energy, respectively.

Parameters	***Values***
E0 (the initial energy of nodes)	0.5 J/node
E elec	50 nJ/bit
E fs	100 pJ/bit/m²
E mp	0.0013 pJ/bit/m4
E ag (energy of aggregation)	5 nJ/bit /signal
P (the desired percentage of CH)	0.1
Packet size in the Set-up phase « node-CH »	400 bits
Packet size in the Steady phase « CH-node »	1000 bits
Packet size in the Steady phase «CH-BS »	4000 bits
D0	0.000000005 Joule
Network size	100m *100m
Range	50
BS coordinates	X=50 ; Y= 50

Network overhead:

We are interested in the overhead generated at CHs and the overhead generated at the BS. It is represented by the number of packets sent to the CH and BS.

Network lifetime:.

It presents the amount of time until all nodes in the network die.

4 The consumed energy

4.1.1 The consumed energy per round

Figs. 1-3 show the results of 120 nodes' energy consumption for LEACH, LEACH-S, TL-LEACH, and EEV-LEACH. The simulations are performed during 100, 200, 500, 700 and 1000 rounds.

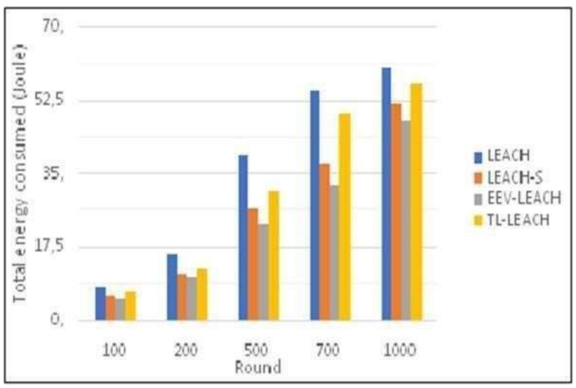


Fig 1. Total energy consumed/round

Fig. 1 represents the total energy consumed in the 4 protocols. Compared to LEACH, LEACH-S, and TL-LEACH during different rounds, the proposed EEV-LEACH protocol consumes less total energy in the network.

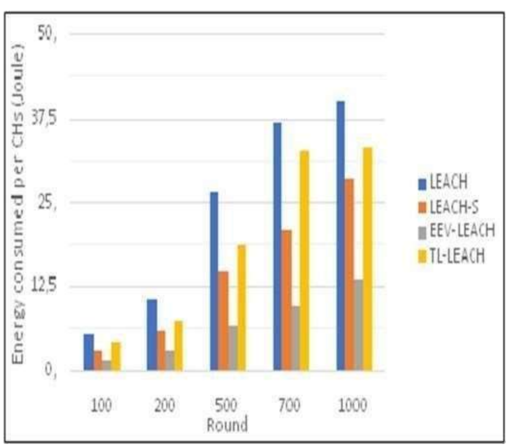


Fig. 2 Energy consumed per CH/round

Fig. 2 shows the energy consumed by the CHs in the 4 protocols. Because the CHs' role is shared with the assistants in an alternative way, the proposed EEV- LEACH protocol consumes less energy than LEACH, LEACH-S, and TL-LEACH during the different rounds.

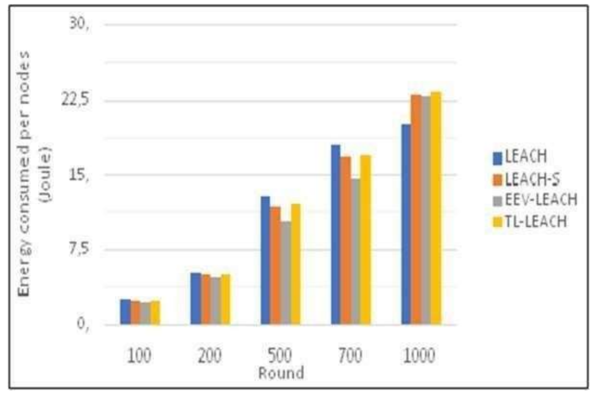


Fig. 3. Energy consumed per node/round

Fig. 3 represents the energy consumed by the nodes in the 4 protocols. The proposed EEV-LEACH protocol reduces node energy consumption. We have minimized the periodic selection of CHs in each round by comparing LEACH, LEACH-S, and TL-LEACH during the different rounds, decreasing the association messages between CHs and nodes and minimizing the energy consumed by nodes.

4.1.2 Energy consumed/number of nodes

Figs. 4-6 show the energy consumption results for 100 rounds of the four protocols with densities of 50, 100, 150, and 200 nodes.

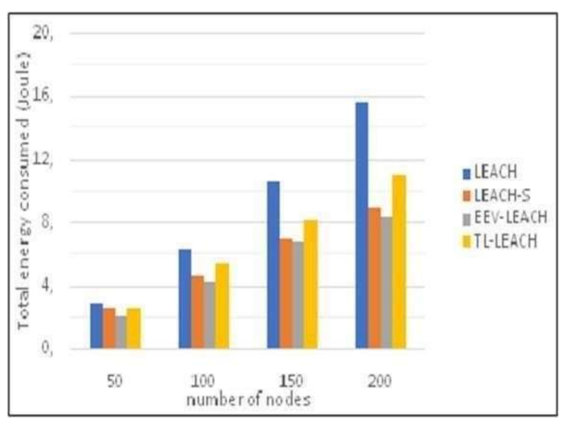


Fig. 4. Total energy consumed/number of nodes

Fig. 4 shows that the proposed EEV-LEACH protocol consumes less total energy than the three LEACH, LEACH-S, and TL-LEACH protocols with varying densities.

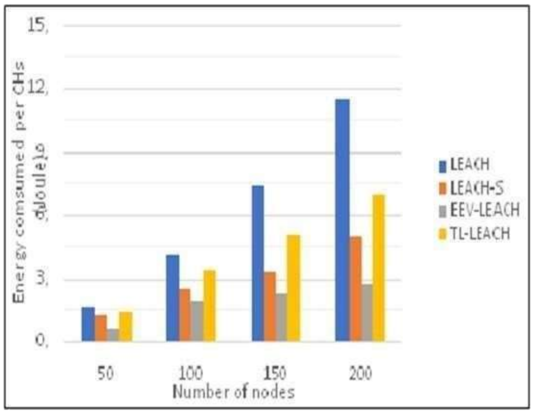


Fig. 5. Energy consumed per CH/number of nodes

Because the work of the CHs is shared with the assistants in an alternative way, the energy consumed by the CHs of the proposed EEV-LEACH protocol is lower than that of the LEACH, LEACH-S, and TL-LEACH protocols in different densities, as shown in Fig. 5.

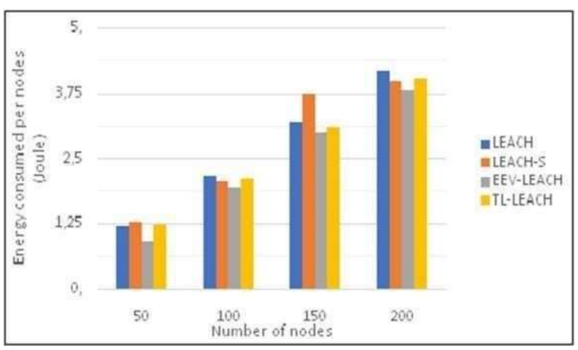


Fig. 6. Energy consumed per node/number of nodes

Fig. 6 shows that the proposed EEV-LEACH protocol consumes less energy than the three protocols LEACH, LEACH-S, and TL-LEACH in different densities because we minimized the selection of the CH in each round, which decreases the association messages between the CHs and the nodes and thus the energy consumed by the nodes. As a result, the proposed protocol reduces the energy consumed by nodes and CHs with varying network densities.

4.2 Network overhead

4.2.1. Network overhead/round

Fig. 7 and Fig. 8 show the number of packets sent to BS and CHs. This simulation employs 120 nodes and runs for 100, 200, 500, 700, and 1000 rounds, respectively.

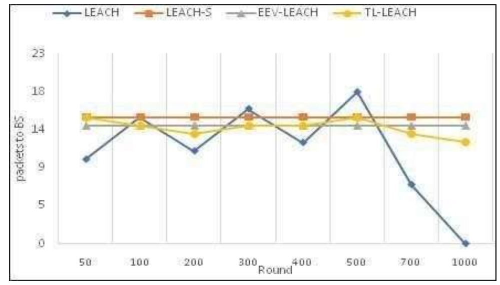


Fig. 7. Packets to BS/round

Fig. 7 depicts the packets sent to the four protocols' BS. The proposed EEV-LEACH protocol reduces the number of packets sent to BS compared to LEACH, LEACH-S, and TL-LEACH during different rounds, thereby minimizing network overhead.

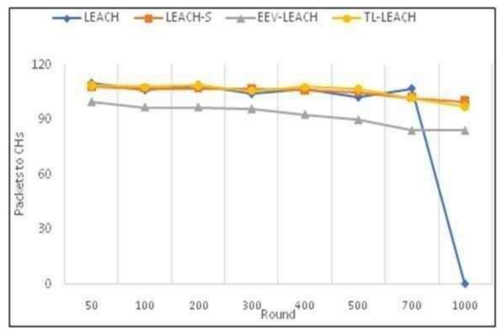


Fig. 8. Packets to CHs/round

Fig. 8 depicts the packets sent from nodes to CHs in the four protocols. Compared to LEACH, LEACH-S, and TL- LEACH during different rounds, the proposed EEV- LEACH protocol sends fewer packets to CHs. Reducing the number of CHs chosen can reduce the number of association messages between nodes and CHs. As a result, the overhead is reduced.

4.2.2 Network overhead/number of nodes

Fig. 9 and Fig. 10 show the number of packets sent to BS and CHs for 100 rounds of the four protocols at different densities of 50, 100, 150, and 200 nodes.

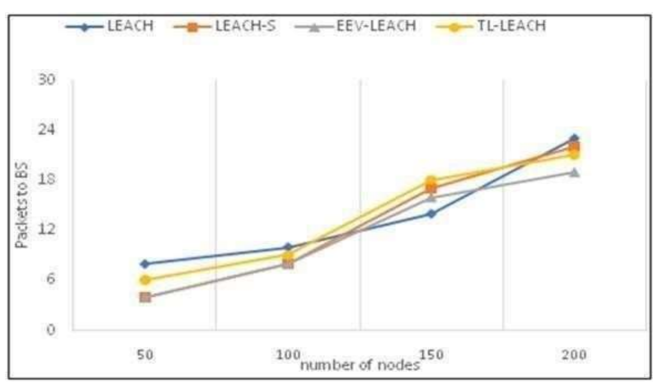


Fig. 9 depicts the packets to the BS in the four protocols at various densities.

Compared to LEACH, LEACH-S, and TL-LEACH during different rounds, the proposed EEV-LEACH protocol reduces the number of packets sent to BS; thus, the proposed protocol minimizes network overhead. Fig. 10 depicts the packets to the BS in the four protocols at various densities. The proposed EEV-LEACH protocol reduces the number of packets sent to BS compared to LEACH, LEACH-S, and TL-LEACH during different rounds, thereby minimizing network overhead.

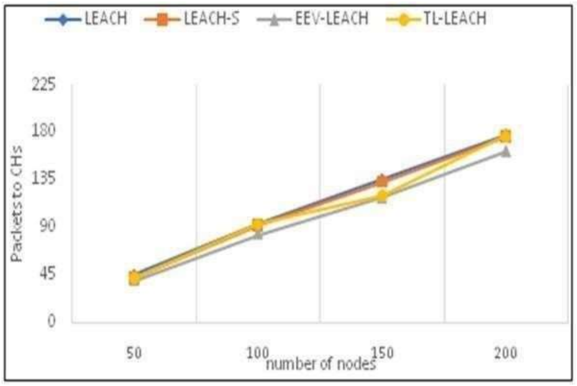


Fig.10. Packets to CHs/number of nodes

Fig. 10 depicts the packets sent from nodes to CHs in the four protocols at various densities. Compared to LEACH, LEACH-S, and TL-LEACH during different rounds, the proposed EEV-LEACH protocol sends fewer packets to CHs, reducing network overhead. This is due to reducing the number of association messages between nodes and CHs by minimizing the periodic selection of CHs.

4.3. Network lifetime

The network lifetime represents the time it takes for all nodes to die. The simulation is run with 120 nodes to evaluate the network lifetime.

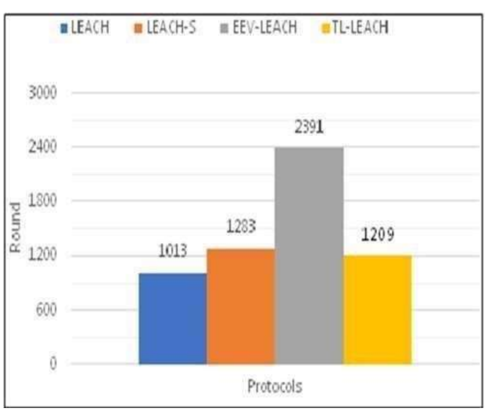


Fig. 11. Round of death of all nodes in the network for LEACH, LEACH-S, TL-LEACH and EEV-LEACH

Fig. 11 depicts the round in which all the nodes are dead. All nodes are dead in round r = 1013 of the LEACH protocol. All nodes are dead in round r = 1209 of the TL-LEACH protocol. All nodes are dead in round r = 1283 of the LEACH-S protocol. All nodes are dead in round r = 2391 of the proposed EEV-LEACH protocol. According to the results, the EEV-LEACH protocol outperforms LEACH, LEACH-S, and TL-LEACH regarding network lifetime.

5 Conclusion and future work

This paper introduced EEV-LEACH, a new hierarchical energy-efficient protocol. It is a new version of the LEACH protocol. The proposed protocol addressed the LEACH protocol's limitations. The objective is to minimize the overall energy consumption, thus extending the network's lifespan, by reducing the number of CHs selected on a regular basis in each round.

Additionally, the overall overhead can be reduced by reducing the number of association messages. Another goal of the proposed protocol is to reduce CH overhead by using an assistant CH in addition to CHs. Regarding consumed energy, overhead, and network lifetime, the simulation results show that EEV-LEACH outperforms LEACH, LEACH-S, and TL-LEACH. However, this work does not consider the topology changes and network coverage. These two drawbacks will be considered in our future works. In addition, more simulations will be performed to compare EEV-LEACH with other variants of LEACH protocols to better evaluate the performance of the proposed protocol.

References

[1] Mario, Di, N., & Haoxuan, Yu. (2021). Intelligent Ventilation Systems in Mining Engineering: Is ZigBee WSN Technology the Best Choice? . *Journal of applied system innovation*, (3) 4, 42.

[2] Mamoona, M., Shaista, H., Abdul Rehman, J., Muhammad, R., Gautam, S., Thippa, R, G., Jerry, C, W, L. (2022). Applications of Wireless Sensor Networks and Internet of Things Frameworks in the Industry Revolution 4.0: A Systematic Literature Review. *Sensors*, (22) 6, 2087.

[3] Mohanta, T, K., Das, D, K.(2021). Class Topper Optimization Based Improved Localization Algorithm in Wireless Sensor Network. *Wireless Pers Communication*, (119) 4, 3319-3338.

[4] Subramani, N., Mohan,P., Alotaibi, Y., Alghamdi, S., Khalaf, O, I. (2022). An Efficient Metaheuristic-Based Clustering with Routing Protocol for Underwater Wireless Sensor Networks. *Sensors*, (22) 2, 415.

[5] Priyanka, B, N., Jayaparvathy, R., Divya ,D, B. (2022). Efficient and Dynamic Cluster Head Selection for Improving Network Lifetime in WSN using Whale Optimization Algorithm. *Wireless Pers Commun*, (123), 2, 1467-1481.

[6] Mohiddin, M, K., Kohli, R., Dutt V, B, S, I., Dixit, P., Michal, G. (2021). Energy-Efficient Enhancement for the Prediction-Based Scheduling Algorithm for the Improvement of Network Lifetime in WSNs. Wireless Communications and Mobile Computing, (2021).

[7] Jan, B. , Farman, H., Javed,H., Montrucchio, B., Khan, M., Ali, S.(2017) . Energy Efficient.Hierarchical Clustering Approaches in Wireless Sensor Networks: A Survey. *Wireless Communications and Mobile Computing*, (2017), 1-14.

[8] Nayak, P.(2014). Comparison of Routing Protocols in WSN using NetSim Simulator: LEACH Vs LEACH-C. *International Journal of Computer Applications*, (106), 6.

[9] Loscri, V., Morabito,G., Marano, S. (2005) . A two-levels hierarchy for low-energy adaptive clustering hierarchy (TL-LEACH). *VTC-2005-Fall. 2005 IEEE 62nd Vehicular Technology Conference*, Dallas, TX, USA, *IEEE*

[10] Arezoo, Y., Fatemeh, H, N., Adel, S. (2012). An Improvement on LEACH Protocol (Cell-LEACH). *14th International Conference on Advanced Communication Technology (ICACT)*. PyeongChang, Korea (South), IEEE

[11] Singh, A., Rathkanthiwar, S., Kakde, S. (2016). Energy efficient routing of WSN using particle swarm optimization and V-LEACH protocol. *International Conference on Communication and Signal Processing (ICCSP)*, Melmaruvathur, India, IEEE.

[12] Tan, L., Wang, M., Wang,H., Liu,L. (2019). Improved LEACH-M Protocol for Aerial Sensor Networks. in *2019 International Conference on Computer, Information and Telecommunication Systems (CITS)*, Beijing, China, IEEE

[13] Arumugam, G,S., Ponnuchamy, T. (2015). EE-LEACH: development of energy-efficient LEACH Protocol for data gathering in WSN. *Journal of Wireless Communication Network*,(2015)1 p. 76.

[14] Tong, M., M. Tang, M. (2010). LEACH-B: An Improved LEACH Protocol for Wireless Sensor Network. *International. Conference on Computational Intelligence and Software Engineering*, Chengdu, China, IEEE

[15] Bendjeddou, A., Laoufi, H., Boudjit, S. (2018). LEACH-S: Low Energy Adaptive Clustering Hierarchy for Sensor Network . in *2018 International Symposium on Networks, Computers and Communications (ISNCC)*, Rome, Italy, IEEE

[16] Mahmood, D., Javaid, N., Mahmood, S., Qureshi, S., Memon, A., M., Zaman, T. (2013). MODLEACH: A Variant of LEACH for WSNs. *Eighth International Conference on Broadband and Wireless Computing, Communication and Applications*, Compiegne, France. IEEE

[17] Jose, H, B, N., Antoniel, S, R. (2014). MH-LEACH: A Distributed Algorithm for Multi-Hop Communication in Wireless Sensor Networks. *ICN 2014: The Thirteenth International Conference on Networks,* Nice, France, IARIA

[18] Salim, A., Osamy, W., Khedr, A., M. (2014). IBLEACH: intra-balanced LEACH protocol for wireless sensor networks. *Wireless Networks*, (20) 6, 1515-1525.

[19] Aaditya, Kumar, R., Rajpoot, P.(2018). Optimized H- LEACH algorithm for clustering to improve lifetime of WSN. *3rd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT)*, Bangalore, India, IEEE

[20] Xiangning, F., Yulin, S. (2007). Improvement on LEACH Protocol of Wireless Sensor Network. *2007 International Conference on Sensor Technologies and Applications (SENSORCOMM 2007).* Valencia, Spain. IEEE

[21] Singh, K. (2015). WSN LEACH based protocols: A structural analysis. *International Conference and Workshop on Computing and Communication (IEMCON)*, Vancouver, BC, Canada.

[22] Kumar, N., Kaur, J. (2011). Improved LEACH Protocol for Wireless Sensor Networks. 7th International Conference on Wireless Communications, Networking and Mobile Computing, Wuhan, China.

[23] Bakaraniya , P., Mehta, S. (2013). K-LEACH: An improved LEACH Protocol for Lifetime Improvement in WSN. *International Journal of Engineering Trends and Technology*, (4) 5, 1521- 1526.

[24] Antoo, A., Mohammed, A, R.(2014). EEM-LEACH: Energy efficient multi-hop LEACH routing protocol for clustered WSNs. *International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, Kanyakumari, India, IEEE.

[25] Kirankumar, B., Katageri, A,C. (2014). Multipath-LEACH an Energy Efficient Routing Algorithm for Wireless Sensor Network. Journal of Advances in Computer Networks, (2)3, 229- 232.

Biographies

1. Amira Bendjeddou

Amira Bendjeddou is an Assistant Professor at the University of Badji Mokhtar, Annaba, Algeria. She is a member of the LRS laboratory. Bendjeddou obtained her Ph.D. in Computer Science in April 2015 from the Computer Science Department of the University Badji Mokhtar, Annaba, Algeria. Her research interests focus on energy preservation in wireless sensor networks, coverage optimization and network security.

2. Mouna Hemici

Mouna Hemici received a master's degree in networks and security in 2022 from the Computer Science Department, University Badji Mokhtar, Annaba, Algeria. Her research interests focus on energy wireless sensor networks and security.

Nodule Classification Using Custom Build 3D Convolution Neural Network Model

Sijo Thomas,* Dr. Sajimon Abraham,† Praveen Kumar,‡ Faizal Basheer,§ Jithinmary Raphel¶
Mahatma Gandhi University, Kerala, 686560, India

Abstract

The classification of lung nodules plays a pivotal role in the early detection and prognosis assessment of lung-related diseases. This paper introduces an innovative approach to nodule classification by leveraging the capabilities of a 3D Convolutional Neural Network (3D CNN) model. The primary objective of this method is to enhance the precision and efficiency of nodule classification, ultimately contributing to improved medical diagnosis and treatment planning.

Our research introduces customized 3D augmentation techniques and a distinct 3D CNN model with finely-tuned parameters. The main focus of the model is to determine the optimal CT patch size for nodule classification while minimizing the training requirements on the proposed custom model. Our research begins by creating patch slice counts ranging from 64 to 24 and 2D spaces of 96x96, ensuring that the nodule center is positioned at the center of each patch. To enhance performance and reduce data transfer bottlenecks between the GPU and HDD, each data chunk is cached in the local environment.

Key Words: nodule classification; augmentation; Lung Cancer;Computed Tomography; Convolutional Neural Network,

1 Introduction

Deep learning models require substantial amounts of data and a significant computational capacity. In the earlier stages, such computational resources were prohibitively expensive and not accessible to a wide range of researchers. However, the recent availability of high-performance computing environments, such as GPUs, has had a profound impact on the processing of models that harness the capabilities of deep learning [1]. This development has been particularly influential in fields like medical image processing and Computer-Aided Diagnoses, where deep learning models have seen extensive use [2].

Deep learning-based systems now play a crucial role in identifying specific regions of interest in medical images, such as CT scans and X-rays, which require in-depth examination by medical professionals. Additionally, they can generate automated reports that aid doctors in correlating their observations [3].

Lung cancer ranks as the most prevalent form of cancer among humans [4] [5]. Each year, lung cancer claims more lives than breast, prostate, and colon cancers combined [5]. Achieving effective treatment necessitates the early detection of cancer. During the initial stages of cancer development, cancer cells manifest as tiny nodules. Given the extensive volume of computed tomography data, medical professionals must dedicate substantial effort to extract and analyze various features, such as size, morphology, contours, interval growth between CT examinations, multiplicity, and location [6].

Among the various deep learning models available, Convolutional Neural Networks (CNNs) have achieved a remarkable breakthrough by reducing computational parameters and enhancing accuracy [7]. Two-dimensional CNNs have shown superior performance when applied to 2D data. However, the application of two-dimensional CNNs is limited to single-slice CT (Computed Tomography) data, potentially resulting in the loss of inter-slice information [8][9].

Considering the size and spatial distribution of malignant nodules, they often extend across multiple slices of CT scans. This necessitates the use of 3D CNN models, but it comes with the added requirement of substantial GPU resources to handle the increased computational load.

In this paper, we present a method for addressing classification challenges in 3D data without causing excessive strain on GPU resources. Our research involves conducting experiments with segments of CT scans containing multiple dimensions, using a custom-built CNN model developed from the ground up. The model experimentation begins with a chunk size of 96x96x64 (XxYxZ), and we explore different combinations of dimensions to determine the most effective chunk dimension for accurate nodule classification. Additionally, we introduce custom augmentation techniques, such as rotation, flip, and offset, to enhance the dataset without putting undue stress on the hardware.

*Research Scholar, School of Computer Sciences. Email: sijothakadiyel@gmail.com
†Research Supervisor, School of Computer Sciences. Email: sajimonabraham@mgu.ac.in
‡Research Scholar, School of Computer Sciences, Email: praveenplavila@yahoo.co.in
§Research Scholar, School of Computer Science, School of Computer Science. Email: faizalbs777@gmail.com
¶Senior Quality Analyst, McFadyen Digital. Email: jraphel@mcfadyen.com

The experiments are assessed within a local environment using a single GPU with 8GB of memory (RTX 3060 Ti) and an Intel i5 processor with 16 GB of RAM.

Our primary contributions to this study can be summarized as follows:

1. We have estimated an optimized chunk of CT scan data to maximize the capabilities of the 3D CNN model.
2. We have introduced a custom-built 3D CNN model tailored for efficient classification tasks.
3. We have introduced effective data augmentation techniques specifically designed for CT scan data, all while ensuring GPU resources are not overwhelmed.
4. We have successfully conducted precise nodule classification on the LUNA / LIDC-IDRI dataset, achieving benchmark results.

The remainder of this paper is structured as follows: Section 2 provides an overview of related research, Section 3 elaborates on the proposed method, Section 4 presents experimental results and analysis, and finally, Section 5 offers the conclusion of this paper.

2 Related Work

This study reviewed multiple categories of articles published on lung nodule classification. The major categories of articles are the deep learning approach based on CNN and the deep learning approach based on non-CNN models. In the traditional non-deep learning approach, researchers used strategies available in computer vision with derivatives of machine learning models. One of the leading developments was the use of Gabr Local Binary Pattern [10] - LBP with support Vector Machine [11] algorithm to predict the probability of nodule.

Hussein [13] worked on creating a 3D CNN model for the classification of nodules. Due to the 3D volumetric nature of CT scans, and the unavailability of the pre-trained model at that time, authors trained their models on sports datasets with more than 1 million videos. In a later stage, they fine-tuned the results on the LIDC dataset.

Author Nibali [14] used a transfer learning strategy for their Resnet18 model. The model was trained on the CIFAR 10 data-set for additional generalization of the pre-trained network. Based on the views of the nodule, for example, axial, coronal, and sagittal, the authors have introduced a three-layer network approach for each view. In the final stage, outputs from three networks are fed into a fully connected layer for final classification. The additional overhead of three network models's memory and training time was a major drawback observed there.

Author Liu[15] proposed a multi-view approach for the classification of nodules. By introducing three scales and four views, the authors are able to introduce a new data set containing 12 unique images of one slice of a CT scan. The authors trained the model on each of the images separately and combined the model for final training. This method also introduced additional overhead of memory and training time.

Author Wentao Zhu[9] DeepLung is a comprehensive system that comprises two core components: one for detecting nodules and the other for classifying these nodules as either benign or malignant. Given the three-dimensional nature of lung CT data and the efficiency of dual path networks (DPN), two distinct deep 3D DPNs have been developed to handle nodule detection and classification tasks. To be specific, a 3D Faster Regions with Convolutional Neural Net (R-CNN) has been tailored for nodule detection, utilizing 3D dual path blocks and adopting a U- net-like encoder-decoder structure to effectively capture nodule features. In the realm of nodule classification, the approach utilizes gradient boosting machine (GBM) in conjunction with 3D dual path network features.

Author Al-Shabi[16] proposal involves the utilization of Residual Blocks employing a 3 × 3 kernel size to extract local features and the incorporation of Non-Local Blocks for the extraction of global features. The Non-Local Block is particularly advantageous in extracting global features without the need for an excessively large number of parameters. The fundamental concept underpinning the Non-Local Block is the application of matrix multiplications among features within the same feature maps.

Table 1: Performance of leading classification methods and baseline models

Method	Accuracy	AUC
DeepLung	90.44	97.13
Local-Global	88.46	95.62
Resnet50	77.62	86.82
Resnet18	78.21	86.41
Densenet121	84.57	92.50
Multi-Crop	89.27	94.32

3 Proposed

3.1 Methods

Our custom CNN model was implemented utilizing the PyTorch deep-learning library[23]. The model execution took place on a local GPU environment, specifically the NVIDIA RTX 3060 Ti with an 8 GB capacity. The operating system used was Ubuntu OS 22.04, and the Python version employed was 3.9 via Anaconda.

In order to effectively handle the dataset, it was partitioned into a 70% training set and a 30% validation set. Given the considerable and fluctuating slice count within the CT scans, it was not feasible to accommodate full-size CT scans within our limited GPU environment. To address this constraint, we structured our experiment to work with different slice counts, including 64, 48, 32, and 24 slices.

For each of these slice counts, we explored various 2D patch sizes, specifically 96x96, 64x64, 48x48, 32x32, and 24x24. Due to the constraints imposed by our hardware infrastructure, we maintained a fixed batch size of 16 across all patch sizes. These patches were derived using a mathematical approximation of the annotated nodule center, ensuring that it was positioned at the center of the patch in both 2D and 3D spaces.

Learning rate, hyper tuning parameter is tuned from $1x10^{-2}$ to $1x10^{-4}$[24]. The first 30 epochs are run with $1x10^{-2}$ and series of epochs 31 to 60 with $1x10^{-3}$ and the final epochs 61 to 75 with $1x10^{-4}$. The approach used for the learning rate selection was to introduce faster changes in the initial training phases and fine-tune in the last stage of training[25].

We followed an approach of Kaiming initialization of neural network weight to explicitly initialize the weight of kernels[28]. Since the CNN model relies on ReLU for activation, during the Kaiming initialization process, opted ReLU for non-linearity. Adam optimizer are used with default values for β_1 and β_2[26]. Checkpoints are saved on each model performance on each patch size with model state and neural weights.

3.2 Dataset Description

The dataset utilized in this study is the LIDC-IDRI [17] dataset, which has been made available by the National Cancer Institute, NIH. This dataset is renowned for being the most expansive and comprehensive public repository of lung nodules. It encompasses a compilation of 1,018 CT scans, derived from a total of 1,010 patients. Due to its substantial scale and public accessibility, the LIDC-IDRI dataset serves as a fitting resource for the development, comparison, and validation of diverse deep-learning methodologies. Moreover, it holds a significant presence in the existing literature, being frequently referenced [13] [15] in studies.

The image collection process encompassed contributions from four different institutions, each employing separate CT scanners. This approach introduced significant diversity in various image attributes. One notable area of diversity lies in image resolution, as evidenced by variations in pixel spacing across different CT scanners. Additionally, the thickness of CT scan slices ranged from 0.45 to 5.0 mm.

Leveraging a heterogeneous dataset like LIDC-IDRI for algorithm design presents the benefit of ensuring robustness to previously unencountered or broadly applicable data. This robustness arises because the algorithms are trained on a wide spectrum of cases, offering versatility in handling diverse scenarios.

Within the LIDC-IDRI dataset, the malignancy likelihood of each nodule underwent evaluation by a panel consisting of four highly experienced radiologists. Initially, all nodules present in each CT scan were annotated, and their boundaries were meticulously documented in individual XML files. It's worth noting that these annotations were restricted to nodules with diameters spanning from 3 to 30 mm.

Interestingly, despite the thorough evaluation by the four experienced radiologists, variations in the nodule annotations were observed, as noted in prior research. Specifically, only a subset of nodules received annotations by the majority of radiologists, which means they were annotated by at least three out of the four experts.

The radiologists also provided a malignancy rating for each nodule, employing a scale ranging from 1 to 5, where higher values indicated higher levels of malignancy (1 represented a benign nodule, and 5 indicated a highly malignant one). To consolidate the ratings provided by the radiologists, we computed the median malignancy score. Nodules receiving ratings below three were categorized as benign, while those with ratings exceeding 3 were classified as malignant. Following a common practice in similar studies, nodules with a rating of exactly 3 were omitted from the analysis due to their indeterminate malignancy status.

3.3 Preprocessing

The LIDC-IDRI dataset initially came in .mhd format along with corresponding raw data files. Our initial step involved converting these .mhd files into a tensor data type, while carefully considering variations in coordinate systems, vertical spacing, and differing slice counts.

In this research, we applied a clamping operation to the Hounsfield unit (HU) values of the scans, specifically restricting values below -1000 or above 400. This practice is well-established in the literature and is commonly used to exclude regions corresponding to air and bone tissues [25][22].

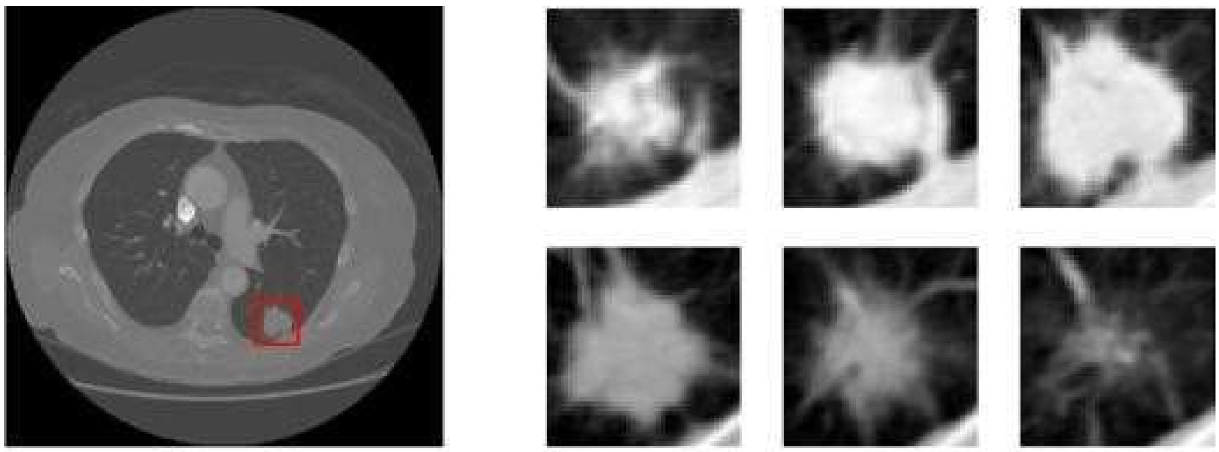
As part of our initial data preprocessing, we performed batch-level normalization of pixel intensities to ensure consistent mean and variance across the dataset. To precisely identify nodules, which are the areas of interest in CT scans, we utilized annotations to pinpoint the center of each nodule. We developed a versatile conversion function to prepare data chunks, ranging in size from 96x96x64 to 24x24x24, for input into our neural network.

To optimize performance and reduce data transfer bottlenecks between the GPU and HDD, each chunk of data was cached in the local environment.

3.4 Data Augmentation

The LIDC-IDRI dataset is imbalanced in terms of the nodule and non-nodule counts: 40,772 out of 62,492 patches are lung nodules, and 21,720 out of 62,492 patches are non-nodules [18]. Such scenarios make actual nodules have less impact on the training of networks. The resolution was to introduce and implement augmentation on chunks of CT scans. The following augmentation has been applied on the patches of CT.

Rotation: We carried out rotations at random angles of 0°, 90°, 180°, and 270°. The rotation logic was executed within a loop, repeating three times. During each iteration, an additional 90° angle was added based on a randomly generated number.



(a) A 2D lung CT slice. (b) Demonstration of a nodule at different Z-axis positions.

Figure 1: Represents sample CT image in 2D and 3D space

Augmentation of CT scans through flipping involves applying various transformations to the original scans to generate additional training data. Flipping typically includes horizontal or vertical flips. Horizontal Flip: This augmentation technique involves flipping the CT scan horizontally along its vertical axis. Vertical Flip: In contrast, a vertical flip flips the CT scan vertically along its horizontal axis[19].

Augmentation of CT scans through offset involves applying spatial translations or shifts to the original scans. This technique can create variations in the position of structures or objects within the CT images, which can help improve the robustness and generalization of machine learning models trained on these images[20]. The CT scan is shifted horizontally and/or vertically by a certain number of pixels. This simulates a change in the position of the patient or the scanner during image acquisition. Translations can be applied in both positive and negative directions along the x and y axes[19]

Augmentation techniques, which have a significant impact on the shape of nodules, are avoided [21].

3.5 Network Architecture

Figure 3 represents the overall architecture and methodology of research. The approach is modularized into 3 stages. In Stage 1, a 3D neural network model is developed using the PyTorch library 1.13. The model includes four blocks of 3D CNN where each set block has two 3D neural layer with ReLU activation function as represented in figure 2. Scale down has been applied using MaxPool 3D with kernel size and stride with 2 on the end of each block. Each layer of the 3D neural layer will have kernel size 3x3x3, stride 1, and padding 1.

The initial block comprises one input channel and yields eight output channels. Within this block, the first 3D CNN layer takes one input channel and produces eight output channels.

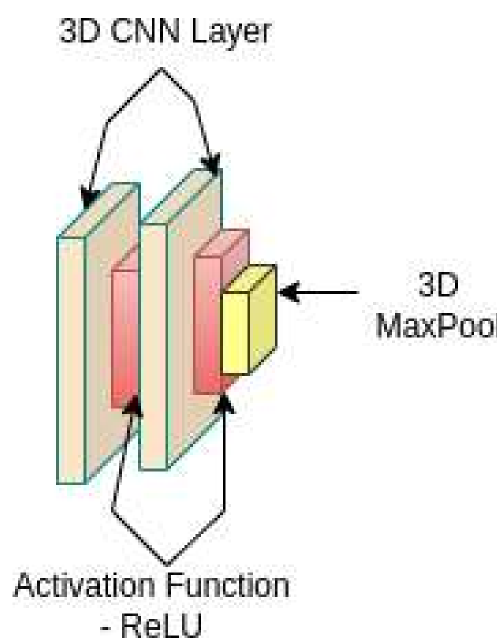


Figure 2: Represents unit block of our 3D CNN model

The output is then subjected to a ReLU activation function before being fed into the second 3D CNN layer. This second layer possesses eight input channels and generates eight output channels, once again going through a ReLU activation function. The outcome of this activation function is subsequently passed through a 3D Max pooling layer, where the pooling operation reduces the dimensionality by dividing the three-dimensional

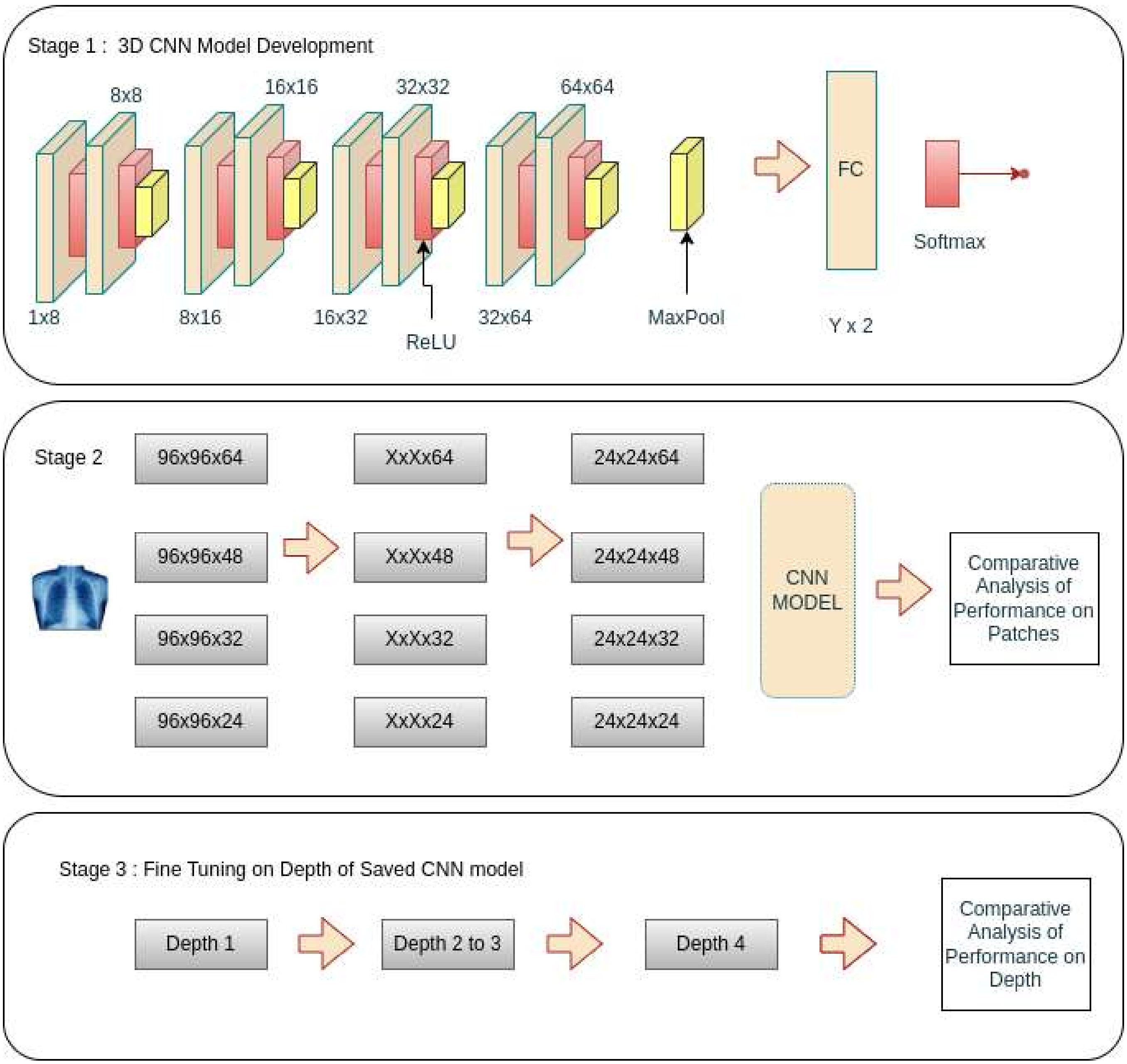


Figure 3: Represents the overall architecture of the study. In stage 1, the initial design of the 3D CNN model is built. Models are initialized with random parameters. In stage 2, the built model is evaluated on the different patch sizes and evaluate the model performance. The best model is saved for stage 3. In stage 3, the model is fine-tuned over different depths of CNN model

input into cuboidal regions and computing the maximum value in each region [source: network:maxpool].

The second block involves eight input channels and results in sixteen output channels. In this block, the first 3D CNN layer takes eight input channels and produces sixteen output channels, followed by a ReLU activation. The output is then fed into the second 3D CNN layer, which operates with sixteen input channels to yield sixteen output channels. Again, a ReLU activation is applied before the result is passed through a 3D Max pooling layer for downsampling.

Moving on to the third block, it encompasses sixteen input channels and delivers thirty-two output channels. The initial 3D CNN layer in this block takes sixteen input channels and generates thirty-two output channels, followed by ReLU activation. The subsequent 3D CNN layer within this block operates with thirty-two input channels to produce thirty-two output channels, once again undergoing ReLU activation. The result from this activation is then channeled through a 3D Max pooling layer for downsampling.

The fourth block has thirty-two input channels and yields sixty-four output channels. Here, the first 3D CNN layer in this block takes thirty-two input channels and generates sixty-four output channels, followed by ReLU activation. Subsequently, the second 3D CNN layer within this block operates with sixty-four input channels to yield sixty-four output channels, going through ReLU activation once more. The output from this activation function is passed through a 3D Max pooling layer for downsampling purposes.

The model's head consists of a single layer of fully connected linear neurons with two output units, which is subsequently followed by the application of a softmax activation function. In contrast, the model's tail includes batch-level normalization, where mean and standard deviation values are computed separately for each dimension across mini-batches.

During stage 2, the model undergoes evaluation on diverse patches extracted from chest scans, each with different dimensions. Promising patches that exhibit favorable performance will be chosen for further refinement during stage 3. In stage 3, the fine-tuning process takes place on the same dataset, with adjustments being made to various depths of the model.

Table 2: Model's detailed block architecture

Block	Layer	Input / Output	Kernel	Stride / Padding
Block 1	CNN	1 / 8	(3,3,3)	(1,1,1) (1,1,1)
Block 1	ReLU	-	-	-
Block 1	CNN	8 / 8	(3,3,3)	(1,1,1) (1,1,1)
Block 1	ReLU	-	-	-
Block 1	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)
Block 2	CNN	8 / 16	(3,3,3)	(1,1,1) (1,1,1)
Block 2	ReLU	-	-	-
Block 2	CNN	16 / 16	(3,3,3)	(1,1,1) (1,1,1)
Block 2	ReLU	-	-	-
Block 2	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)
Block 3	CNN	16 / 32	(3,3,3)	(1,1,1) (1,1,1)
Block 3	ReLU	-	-	-
Block 3	CNN	32 / 32	(3,3,3)	(1,1,1) (1,1,1)
Block 3	ReLU	-	-	-
Block 3	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)
Block 4	CNN	32 / 64	(3,3,3)	(1,1,1) (1,1,1)
Block 4	ReLU	-	-	-
Block 4	CNN	64 / 64	(3,3,3)	(1,1,1) (1,1,1)
Block 4	ReLU	-	-	-
Block 4	MaxPool3D	-	(2,2,2)	(2,2,2) (0,0,0)

4 Experimental Results and Analysis

Table 4 represents the performance of the classification model on a patch size of 96x9x96. Two-dimension size 96x96 with slice length of 64 to 24. Considering the model capacity and large patch size compared to nodule size, classification performance was not showing significant improvement compared to baseline models and other leading methods. Table 5 represents the performance of the classification model on a patch size of 96x96x64. Two-dimension size 64x64 with slice length of 64 to 24. Among the various patch sizes, 64x64x48 and 64x64x32 lead the classification performance.

Table 6 represents the performance of the classification model on a patch size of 48x48. Two-dimension size 48x48 with slice length of 64 to 24. Among the various patch sizes, 48x48x48 and 48x48x32 lead the classification performance.

Table 7 represents the performance of the classification model on a patch size of 32x32xZ Two-dimension size 32x32 with slice length of 64 to 24. Among the various patch sizes, 32x32x48 and 32x32x64 lead the classification performance, and patch size 32x32x48 has the highest accuracy across all the patches.

Table 3: model architecture for head and tail (32x32x48)

Block	Layer	Input / Output	Kernel	Stride / Padding
Tail	BatchNorm	-	-	- -
Head	Linear	1152 / 2	-	- -
Head	SoftMax	2/1	-	- -

Table 4: Performance of patch size 96x96 on CNN model

Patch Size	Accuracy	AUC
96x96x64	78.00	79.01
96x96x48	79.00	80.10
96x96x32	74.00	73.76
96x96x24	73.00	72.06

Table 8 represents the performance of the classification model on a patch size of 24x24xZ Two-dimension size 24x2 with slice length of 64 to 24. The performance of the classification model was below the other patches except the 96x96 patch.

The experimental results on the validation dataset show that patch sizes 32x32x48 and 48x48x32 have better performance yield compared to other patch sizes. These two patches are further taken for stage 3, the fine-tuning stage where depth-level fine-tuning of the CNN model is performed. Models are fine-tuned with depths 1 to 3 on the epoch of 50s. The model will be re-trained with all other layers of the model kept intact except layers up to specified depth. The table 9 and 10 shows the model performance of best classification patches on various depths

Table 9 represents the performance of the classification model on a patch size of 48x48x32 with fine-tuning at various depths. Among the various depth, depth 2 lead the classification performance.

Table 10 represents the performance of the classification model on a patch size of 32x32x48 with fine-tuning at various depths. Among the various depth, depth 3 lead the classification performance.

5 Discussion

The main objective of this work was to utilise the LIDC dataset for training and evaluating the acquired results. Additional validation can be attained through the implementation of extensive experiments that utilise diverse datasets. This approach serves to mitigate any biases and bolster the overall strength and reliability of the results. Further enhancements can be investigated by effectively implementing augmentation techniques, activation functions, and incorporating them in conjunction with loss functions, among other considerations.

Table 5: Performance of patch size 64x64 on CNN model

Patch Size	Accuracy	AUC
64x64x64	81.04	81.00
64x64x48	84.09	84.00
64x64x32	84.69	83.34
64x64x24	70.03	73.10

Table 6: Comparative performance of the model on best slice-count

Patch Size	Accuracy	AUC
48x48x64	83.90	84.00
48x48x48	90.89	91.09
48x48x32	91.29	91.98
48x48x24	81.01	81.24

6 Conclusions and Future work

In our research paper, we introduced a specialized CNN model designed for the classification of nodules as either benign or malignant. The rationale behind developing this custom CNN model was to gain a deeper understanding of the hardware needs and achieve enhanced accuracy in performance. Given the substantial size of CT scans, there could be challenges in processing them in resource-limited GPU environments. To address this, we devised a solution: generating CT scan patches centered around the annotated center of the nodules. Our study revealed that the 32x32x48 patch size yielded superior performance results when used with our custom classification model.

The augmentation strategy was created with the aim of minimizing the machine's workload and expediting network training through faster networking. Additionally, the augmentation approach has been devised to facilitate parallel execution, relieving the CPU from excessive burdens.

Furthermore, to enhance the utilization of GPU bandwidth, we pre-cache CT patches prior to conducting experiments. This approach helps minimize data loading latency on the GPU.

Our observations indicate that both higher-dimensional and lower-dimensional patch sizes yield classification accuracy that falls below that of middle-sized patches. The reduced accuracy observed in lower-dimensional patches may be attributed to the limited information contained within them. On the other hand, the diminished accuracy in higher-dimensional patches

Table 7: Performance of patch size 32x32 on CNN model

Patch Size	Accuracy	AUC
32x32x64	88.12	87.27
32x32x48	92.10	93.23
32x32x32	87.09	88.14
32x32x24	88.98	87.12

Table 8: Performance of patch size 24x24 on CNN model

Patch Size	Accuracy	AUC
24x24x64	86.12	85.28
24x24x48	85.95	85.00
24x24x32	82.63	80.12
24x24x24	80.18	79.65

could be due to the model's smaller receptive field and/or the need for more training epochs to effectively process the richer information in the data. Conducting further research in this domain would necessitate additional computational resources and in-depth analysis.

An alternative approach to consider is combining the best-performing patches with a common fully connected linear layer within the CNN model, which may lead to improved results. This avenue presents an opportunity for future research exploration in this field.

Additionally, in the context of multi-modal feature extraction, it is possible to customise the model architecture to accommodate different inputs, such as MRI and X-rays. This customization allows for the inclusion of additional characteristics related to nodules, hence enhancing the accuracy of classification. One significant obstacle that we anticipate encountering is the need for a one-to-one correspondence between computed tomography (CT) scans and their corresponding magnetic resonance imaging (MRI) or X-ray images in order to validate findings.

Table 9: Performance of fine-tuning on patch size 48x48x32

Patch Size	Accuracy	AUC
Depth 1	90.98	91.34
Depth 2	92.45	92.19
Depth 3	92.61	93.24
Depth 4	91.56	90.42

Table 10: Performance of fine-tuning on patch size 32x32x48

Patch Size	Accuracy	AUC
Depth 1	92.98	91.34
Depth 2	93.45	94.19
Depth 3	95.61	95.24
Depth 4	90.56	89.42

Acknowledgments

We extend our heartfelt gratitude to Professor Dr. Sajimon Abraham, my esteemed research supervisor, whose patient guidance, unwavering encouragement, and invaluable critiques have been the cornerstone of this research endeavor. His dedication to our growth and development as researchers has been instrumental in shaping this work.

Furthermore, we wish to convey our sincere appreciation to the National Cancer Institute and the Foundation for the National Institutes of Health for their pivotal role in creating and maintaining the LIDC-IDRI database, which has served as an indispensable resource for our study. Their commitment to facilitating access to this invaluable database has significantly enriched the scope and depth of our research.

References

[1] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in neural information processing systems, 2012, pp. 1097–1105.

[2] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: Fine tuning or full training?, IEEE Transactions on Medical Imaging 35 (5) (2016) 1299–1312.

[3] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, J. Dean, A guide to deep learning in healthcare, Nature medicine 25 (1) (2019) 24–29.

[4] W. Stewart, B. and Wild, C.P. (eds.), International Agency for Research on Cancer, "World Cancer Report 2014," 2014.

[5] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," CA. Cancer J. Clin., vol. 68, no. 1, pp. 7–30, Jan. 2018.

[6] A. McWilliams et al., "Probability of Cancer in Pulmonary Nodules Detected on First Screening CT," N. Engl. J. Med., vol. 369, no. 10, pp. 910–919, Sep. 2013.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Adv. Neural Inf. Process. Syst., pp. 1–9, 2012.

[8] W. Shen, M. Zhou, F. Yang, D. Yu, D. Dong, C. Yang, Y. Zang, J. Tian, Multi-crop convolutional neural networks for lung nodule malignancy suspiciousness classification, Pattern Recognition 61 (2017) 663–673.

[9] W. Zhu, C. Liu, W. Fan, X. Xie, Deeplung: Deep 3d dual path nets for automated pulmonary nodule detection and classification, in: 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), IEEE, 2018, pp. 673–681.

[10] Z. Fan, S. Yang, C. Weidong, L. Min-Zhao, Z. Yun, H. Heng, S. Shimin, M. J. Fulham, D. D. Feng, Lung nodule classification with multilevel patch-based context analysis, IEEE Trans Biomed Eng 61 (4) (2014) 1155–1166

[11] S. M. B. Netto, A. C. Silva, R. A. Nunes, M. Gattass, Automatic segmentation of lung nodules with growing neural gas and support vector machine, Computers in Biology and Medicine 42 (11) (2012) 1110–1121

[12] W. Sun, B. Zheng, and W. Qian, "Automatic feature learning using multichannel ROI based on deep structured algorithms for computerized lung cancer diagnosis," Comput. Biol. Med., vol. 89, no. January, pp. 530–539, 2017.

[13] S. Hussein, K. Cao, Q. Song, and U. Bagci, "Risk stratification of lung nodules using 3D CNN-based multi-task learning," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 10265 LNCS, pp. 249–260, 2017.

[14] A. Nibali, Z. He, and D. Wollersheim, "Pulmonary nodule classification with deep residual networks," Int. J. Comput. Assist. Radiol. Surg., vol. 12, no. 10, pp. 1799–1808, 2017.

[15] X. Liu, F. Hou, H. Qin, and A. Hao, "Multi-view multi-scale CNNs for lung nodule type classification from CT images," Pattern Recognit., vol. 77, pp. 262–275, 2018.

[16] Al-Shabi M, Lan BL, Chan WY, Ng KH, Tan M. Lung nodule classification using deep Local-Global networks. Int J Comput Assist Radiol Surg. 2019 Oct;14(10):1815-1819. doi: 10.1007/s11548-019-01981-7. Epub 2019 Apr 24. PMID: 31020576.

[17] S. G. Armato et al., "The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans," Med. Phys., vol. 38, no. 2, pp. 915–931, 2011.

[18] Li W., Cao P., Zhao D., Wang J. Pulmonary nodule classification with deep convolutional neural networks on computed tomography images. Computational and Mathematical Methods in Medicine. 2016;2016:7.

[19] Soni M, Singh AK, Babu KS, Kumar S, Kumar A, Singh S. Convolutional neural network based CT scan classification method for COVID-19 test validation. Smart Health (Amst). 2022 Sep;25:100296. doi: 10.1016/j.smhl.2022.100296. Epub 2022 Jun 11. PMID: 35722028; PMCID: PMC9188200.

[20] Khalifa, Nour Eldeen, Mohamed Loey, and Seyedali Mirjalili. "A comprehensive survey of recent trends in deep learning for digital images augmentation." Artificial Intelligence Review (2022): 1-27.

[21] Tran GS, Nghiem TP, Nguyen VT, Luong CM, Burie JC. Improving Accuracy of Lung Nodule Classification Using Deep Learning with Focal Loss. J Healthc Eng. 2019 Feb 4;2019:5156416. doi: 10.1155/2019/5156416. PMID: 30863524; PMCID: PMC6378763.

[22] X. Xu, C. Wang, J. Guo, Y. Gan, J. Wang, H. Bai, L. Zhang, W. Li, Z. Yi, MSCS-DeepLN: Evaluating lung nodule malignancy using multi-scale cost-sensitive neural networks, Med. Image Anal. 65 (2020) 101772. https://doi.org/10.1016/j.media.2020.101772.

[23] A. Paszke et al., "Automatic differentiation in PyTorch.".

[24] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770–778.

[25] M. Al-Shabi, H. K. Lee and M. Tan, "Gated-Dilated Networks for Lung Nodule Classification in CT Scans," in IEEE Access, vol. 7, pp. 178827-178838, 2019, doi: 10.1109/ACCESS.2019.2958663.

[26] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014

[27] Zhao, W., Jiang, W. and Qiu, X. Deep learning for COVID-19 detection based on CT images. Sci Rep 11, 14353 (2021). https://doi.org/10.1038/s41598-021-93832-2

[28] K. He, X. Zhang, S. Ren and J. Sun, "Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, pp. 1026-1034, doi: 10.1109/ICCV.2015.123.

Sijo Thomas serves as a research scholar affiliated with the School of Computer Sciences at Mahatma Gandhi University in Kerala, India. In addition to his academic pursuits, he assumes the role of a consultant software architect.His software solutions are utilised by prominent state universities, private universities, and autonomous colleges in India. Additionally he has academic experience of more than 5 years as faculty of science in colleges. Mr.Sijo Thomas holds a Master's degree in computer applications from Mahatma Gandhi University in Kerala, and he has also achieved a Master of Philosophy degree from Bharathidasan University in Tamil Nadu, India.

Dr.Sajimon Abraham earned his Ph.D. in Computer Science with a specialization in Spatio-Temporal Data Mining from Mahatma Gandhi University in Kerala, India. Starting in 2016, he has been serving as a research supervisor in the research department of the School of Computer Science. Additionally, Dr. Abraham holds the prestigious role of Honorary Director at the University Centre for International Co-operation at Mahatma Gandhi University, Kerala. Furthermore, he serves as a professor of computer and IT at the School of Management and Business Studies at Mahatma Gandhi University, Kerala. His distinguished academic career boasts an impressive total of 92 publications in the field of computer science and engineering.

Praveen Kumar holds the position of a research scholar at the School of Computer Sciences, Mahatma Gandhi University, located in Kerala, India. Additionally, he serves as an associate professor at a Government Aided College, boasting an impressive 24 years of teaching experience in undergraduate programs and 11 years in postgraduate programs. His academic interests primarily revolve around Spatio-Temporal Data Mining and the application of Artificial Intelligence in the realm of Human Rights. Sijo has made notable contributions to his field with three publications in international journals and one paper presented at an internal conference organized by the Indian Institute of Management (IIM).

Faizal Basheer is currently a research scholar affiliated with the School of Computer Sciences at Mahatma Gandhi University in Kerala, India. In addition to his academic pursuits, he actively serves as a consultant specializing in machine learning and artificial intelligence. Faizal brings a valuable perspective to his work with five years of experience as a faculty member in the field of computer science. His educational background includes a Master's degree in computer science and engineering, which he earned from Mahatma Gandhi University in India.

Jithinmary Raphel possessing a decade's worth of expertise in Information Technology (IT) as a quality engineer, is a committed professional deeply enthusiastic about the realms of Deep Learning and Artificial Intelligence (AI). Jithinmary's academic foundation was established at LBS College of Engineering in Kerala, India, where she earned a Bachelor of Technology (B.Tech) degree in Computer Science and Engineering.

Enhancing Math Word Problem Solving Using Multi-Head-Attention Mechanism

Sandip Sarkar*, Dipankar Das†, David Pinto‡
*Hijli College, India
†Jadavpur University, India
‡Computer Science of the Benem erita Universidad Aut onoma de Puebla

Abstract

Solving arithmetic word problems requires models that can effectively understand and generate mathematical equations based on textual information. However, bridging the gap between machine- understandable logic and human-readable language remains a significant challenge. While some researchers have made promising strides with limited datasets, there is a need for more comprehensive approaches. In this paper, we present two novel models for solving mathematical word problems (MWPs) involving multiple unknown variables: a multi-head attention mechanism-based model and a GRU-LSTM-based Seq2seq model. Our models are evaluated using state-of-the-art metrics, such as BLEU and ROUGE, considering factors such as the number of unknown variables and sentence structure. Through extensive experimentation on diverse datasets, we achieved impressive results and surpassed previous benchmarks. This research contributes to advancing math word problem-solving by harnessing the power of deep learning techniques and attention mechanisms. Our findings also highlight areas for future improvements and expansion in this domain.

Key Words: Math word problem, Deep learning, Seq2Seq learning, Text simplification

1 Introduction

Different factors contribute to the challenges students face when dealing with mathematical word problems. These obstacles encompass factors such as math-related anxiety, limitations in memory capacity, inadequate counting skills, language barriers, and a deficiency in problem-solving strategies. [1, 2]. Math anxiety manifests itself when dealing with numerical problems [3, 4]. Students experiencing anxiety necessitate additional time to solve math word problems, while those lacking counting and number sense skills also require extended durations for solving such problems [5, 6].

On the contrary, English assumes the status of the predominant language of instruction in a vast number of nations. Consequently, students hailing from non-English-speaking countries encounter certain challenges while grasping their academic subjects [7]. Specifically, if a student needs additional time to comprehend a math problem presented in English, a corresponding increase in time will also be essential for solving said problem. It is worth noting that math word problems often incorporate intricate vocabulary, posing difficulties in comprehension.

Creating an automated system to solve mathematical word problems poses a challenging endeavor within the realm of Natural Language Processing [8, 9]. Since the 1960s, researchers have put forward multiple methodologies aimed at creating an automatic mathematical problem-solving system [10, 11, 12, 13, 14]. The primary objective of math problem solving is to develop algorithms capable of solving math problems expressed in natural language. An automatic math solver aids students in comprehending mathematical word problems and necessitates the creation of a tutorial system. Several researchers argue that utilizing small datasets with limited variation can yield encouraging outcomes [15, 16, 17].

Mathematical word problems can be categorized into two distinct types: i) arithmetic word problems and ii) algebraic word problems [18]. Arithmetic word problems predominantly involve fundamental mathematical operations, while algebraic word problems encompass more intricate operators like square roots, exponentials, and logarithms, often involving multiple unknown variables. For the purposes of this paper, our focus will solely be on arithmetic word problems.

Table 1 shows examples of the math word problem. A word problem is a topic in mathematics that elementary school students typically resolve. Questions are presented in plain text using the symbols W_1, W_2, etc. The quantities are written as q_1, q_2, q_3 with X standing in for the unknown variable. The output of a math word problem only presents the four types of basic operators $O = +, -, \times, \div$.

Our primary objective is to assess the impact of varying the unknown variable on the generation of equations. Certain researchers focus on analyzing a single unknown variable, while others examine scenarios involving multiple unknown variables. Nevertheless, there is currently a lack of comprehensive research that compares and evaluates the system's performance for both types of unknown variables. Table 4 shows the data variation of the math word problems.

This paper uses two types of deep learning approaches 1) GRU-LSTM based Seq2seq Model, 2) Multi-Head Attention Model. In addition, we evaluated our models using BLEU and ROUGE matric scores. Math word problems not only depend on the number of unknown variables but also on the structure of the sentence (i.e., simple or complex forms). In

* Corresponding author - Sandipsarkar.ju@gmail.com

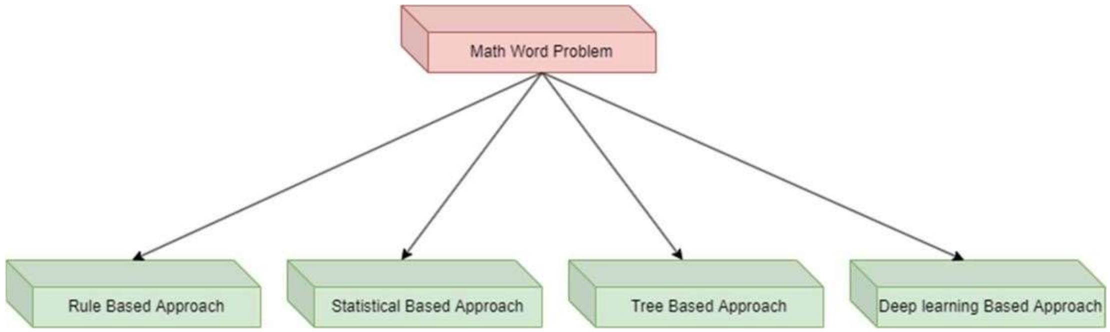


Figure 1: Categories of Math Word Problem

Problem 1	**Statement:**	Two numbers have a sum of 45 and a difference of 29. Can you find the values of these two numbers?
	Equation:	x+y=45.0 x-y=29.0
	Solution:	37, 8
Problem 2	**Statement:**	The student-teacher ratio for Washington High was reported to be 12 to 1. If there are 15 teachers, then how many students are there?
	Equation:	12*15=x*1
	Solution:	180.0

Table 1: Problem Definition

observation, we also give an overview of the reasons behind the poor performance of our system on some of the datasets.

The structure of our paper is outlined in the following manner. In Section 2, we conduct a thorough literature survey to examine existing research in the field. This provides a solid foundation for our study. In Section 3, we provide a detailed description of the dataset used, outlining its composition and characteristics. Moving forward, Section 4 presents our approach, discussing the methodologies and techniques employed to address the problem. Section 5 focuses on the training process, providing insights into the specific procedures and considerations taken. In Section 6, we present the experimental results of our approach. Additionally, Section 7 delves into our observations, discussing any limitations or drawbacks identified in our system. Finally, in Section 8, we conclude our work, summarizing the key findings and contributions made throughout the study, thereby bringing our research to a conclusive end.

2 Literature Survey

The advancement of systems designed to address arithmetic word problems has garnered significant attention in recent years. These methods can be categorized into four main groups: (i) Rule-based Approach, (ii) Statistical Approach, (iii) Tree-based Approach, and (iv) Neural Network-Based Approach. Each of these approaches presents distinct methodologies and techniques, contributing to the diverse landscape of arithmetic word problem-solving methodologies.

2.1 Rule based Approach

There are certain limitations associated with rule-based approaches to solve math word problems. One major drawback lies in its restricted coverage and the complexity associated with its implementation, as indicated by previous research [19, 20]. The rule-based approach may struggle to handle a wide range of math word problems effectively, and its intricate implementation can pose challenges in terms of development and maintenance. Moreover, scalability becomes a concern when dealing with large datasets. Notably, the WORDPRO system, proposed in 1985, introduced a rule-based approach for solving math word problems, incorporating four distinct schemes: change-in, change-out, combine, and compare [21]. Another noteworthy system, called ROBUST, was developed by Bakman specifically to address free-format multi-step arithmetic word problems [22].

2.2 Statistical Approach

In order to solve math word problems, statistical learning-based methods have been proposed since 2014 [23, 24]. By leveraging statistical techniques, such as machine learning algorithms, pattern recognition, and data analysis, this approach aims to extract relevant information, identify problem structures, and make informed predictions or decisions to solve math word problems efficiently. It offers a promising avenue for enhancing problem-solving capabilities and addressing the complexities often encountered in mathematical word problems. By Hosseini and others the open-domain nature of algebraic

Dataset	Problem Statement	Equations	Solutions
number_word_std	Find two numbers whose sum is 33 and one of the numbers is two times as large as the other?	x + y = 33, x = 2*y	x=22 , y =11
DRAW-1K	kevin 's sister mary is 5 years younger than he is . the total of their ages is 33 years . how old is kevin ?	x+y=33, y=x-5	x=19, y=14
Math23K	**Original Text:** 一种电视，电视现在售价450 元，比原来便宜10% ，比原来便宜多少元？ **Translated Text:** A kind of TV. The price of TV is 450 yuan now, which is 10% cheaper than before. How much is it cheaper than before?	x=450/(1-10%)-450	x = 50

Table 2: Description of Dataset

word problems is addressed by learning verb categorization from training data [25]. The operator classification method is used by the author of [24] to resolve math word problems. The operator classifier divides the math word problem by operator $op \in (+, -, *, /)$.

2.3 Tree based Approach

The utilization of tree structures is central to the tree-based approach for analyzing and solving mathematical problems presented in text format. This approach represents the problem as a tree, with nodes representing various problem components (such as numbers, operators, and variables), and edges representing the connections between them. By employing this method, a systematic understanding of the problem and the generation of solutions are made possible. Binary tree structures can be used to represent arithmetic expressions [26, 27]. In a binary tree structure, operators with higher priority are positioned at lower levels, and the operators with the lowest priority are found at the tree's root[28, 29].

2.4 Neural Network Based Approach

Since 2012, deep learning has been applied to a variety of natural language processing tasks, such as Question Answering [30], text simplification [31], sentiment analysis [32], machine translation [33] as well as math word problem [34, 35, 36, 37]. A system that can produce equation templates from a mathematical statement was developed by Wang in 2017 [38]. Next, a model that adds an attention-based mechanism to the prior model was presented by Huang [39]. Deep learning's main benefit is retrieval its ability to select features efficiently without human assistance [40]. On the other hand, it has one drawback: its learning process necessitates a lot of data [41].

3 Dataset

Math word problem datasets commonly suffer from limitations in terms of size and the variety of mathematical scenarios they cover. In our research, we sought to overcome this by meticulously selecting three datasets: Dolphin T2 Final, DRAW-1, and Math23K. Among these, Math23K and Dolphin T2 Final exhibited the most promising characteristics for math word problems. In the subsequent sections, we will provide comprehensive insights into these datasets, emphasizing their composition and significance in our research. The example of each dataset is given in Table 2. Similarly, Table 3 shows the statistics of each dataset.

3.1 Dolphin T2 Final

Dolphin T2 Final is the part of Dolphin18K dataset. [1] has a greater variety of problem types. Their main objective is to use simple mathematics to build a large dataset [42]. The mathematics category of the Yahoo! Answers website provided the dataset. The raw problem text and one or more solutions make up a math problem.

The Dolphin T2 Final dataset focuses on math word problems and is composed of 831 problem instances that have been submitted by users on the popular community driven platform Yahoo! Answers [43].

3.2 DRAW-1K

The Diverse Algebra Word Problem Set (DRAW) dataset [2] encompasses a collection of 1000 algebra word problems. Notably, DRAW surpasses the Alg-514 dataset with its tenfold increase in equation templates and double the number of problem instances. This significant expansion ensures that DRAW exhibits a more realistic representation of real-world problem scenarios. Moreover, DRAW holds the distinction of being the first dataset to provide alignments between equation coefficients and corresponding numbers within the problem text, along with annotated templates. These unique features make

[1]https://msropendata.com/datasets/f0e63bb3-717a-4a53-aa79-da339b0d7992

[2]https://www.microsoft.com/en-us/download/details.aspx?id=52628

Dataset	Problems	Sentences	Percentage of (> 1)Unknown Variables	Percentage of (1) Unknown Variable	Question Length	Equation Length
Dolphin T2 Final	831	831	26.23%	73.76%	75.64	12.93
DRAW-1K	1000	2330	74.5%	25.5%	103.90	12.27
Math23K	23162	70125	0%	100%	70.76	13.944

Table 3: Statistics of Dataset

Dataset Variation	Problem Statement	Equations
Single Variable	The sum of two consecutive whole numbers is 27. Find the numbers.	x+(x+1)=27
	Currently, a person is 35 years old, and his son is 6 years old. After how many years will the son be half the age of her father?	0.5*(35+x)=(6+x)
	To create a solution with a 35% acid concentration, how many liters of a 60% acid solution need to be combined with 11 liters of a 15% acid solution?	0.01*60*x+0.01*15*11=0.01*35*(x+11)
Multiple Variable	A wire with a length of 160 cm needs to be divided into two sections, with one part being 25 cm longer than the other. What is the length of each part?	x+y=160, x=y+25
	Two trains commence their journeys from towns that are 192 miles apart, moving towards each other on parallel tracks. After 1.6 hours, they cross paths. If one train travels at a speed 10 mph faster than the other, determine the speed of each train	x-y=10, (x+y)*1.6=192
	Rajesh is conducting an experiment in the laboratory where he is combining two alcohol-containing solutions. He uses a quantity of Solution A that is 500 milliliters less than Solution B. Solution A has an alcohol concentration of 16%, while Solution B has a concentration of 10%. If the resulting mixture contains 76 milliliters of pure alcohol, how many milliliters of Solution A does Rajesh use?	0.01*10*x+0.01*16*y=76, (x-500)=y

Table 4: Data Variation in Math Word Problem

DRAW an invaluable resource for advancing research in the field [44, 45].

3.3 Math23K

The Math23K dataset is predominantly derived from various online educational websites, making it a comprehensive resource for math word problems [38, 46]. Given the requirement of a large dataset for training deep learning models, Math23K's impressive collection of 23,161 problem instances makes it an ideal choice. This dataset is particularly well-suited for elementary school students, as it aligns with their educational level and curriculum. The compilation of Math23K involved the use of a rule-based extraction mechanism, enabling the retrieval of both problem statements and their corresponding solutions. Notably, the math word problems in Math23K can be effectively solved using linear algebra expressions that involve a single unknown variable.

4 System Architecture

The dataset has been thoroughly described in the preceding section. Our proposed models performed on three different datasets. All datasets, with the exception of Math23k, contain a number of unknown variables, as was already mentioned while the Dolphin T2 Final and DRAW 1K datasets contains multiple unknown variables. Table 3 shows the percentage of math word problems with one unknown variable and math word problems with multiple unknown variables in the dataset. On the other hand, Math23K only contains math word problems that contain one unknown variable.

To deal with we present two innovative models that offer

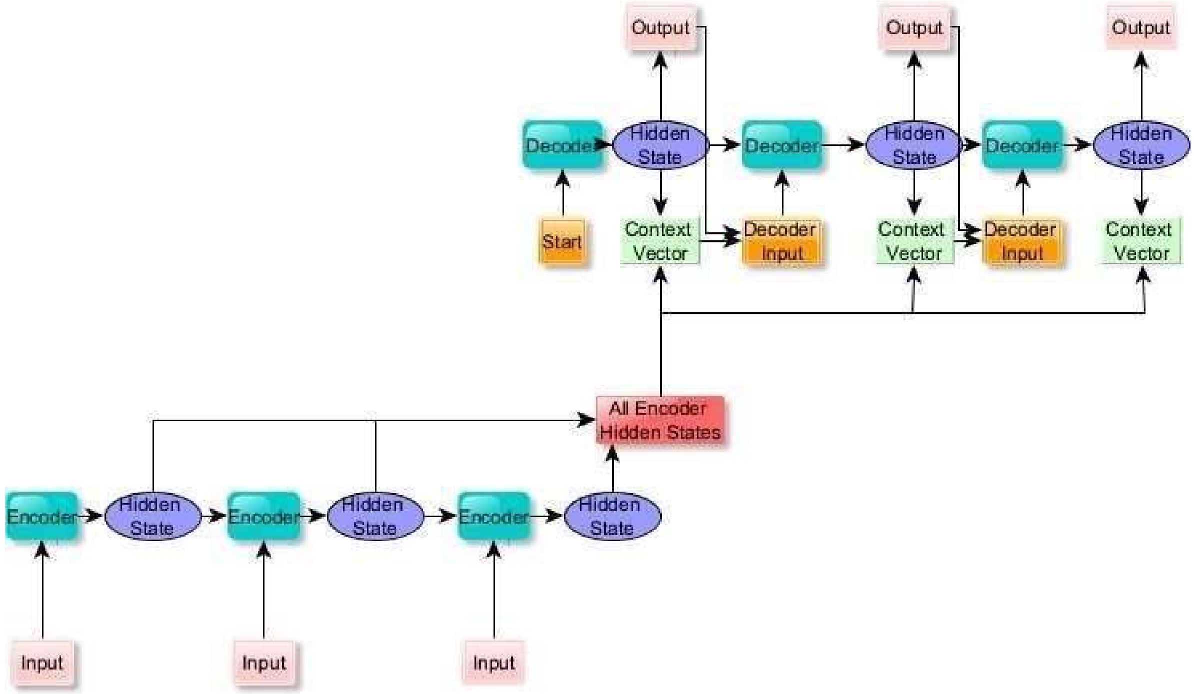


Figure 2: GRU-LSTM based Seq2seq Model

novel solutions to the challenges posed by mathematical word problems (MWPs) featuring multiple unknown variables. Our research introduces a cutting-edge approach utilizing a multi-head attention mechanism -based model, as well as a state-of-the-art Seq2seq model that combines GRU and LSTM. These models represent significant advancements in the field and provide promising avenues for effectively addressing complex MWPs. Table 4 illustrates various types of math word problems.

4.1 GRU-LSTM based Seq2seq Model

Our baseline model predominantly revolves around sequence -to-sequence modeling, which serves as its fundamental framework. Notably, it employs bidirectional GRU cells for the encoding phase and LSTM cells for the decoding phase. To preprocess the equation string, we employ a specific procedure. Initially, each character within the string is separated by a space. Subsequently, all characters in the math statement are converted to lowercase, while punctuation marks are accompanied by spaces inserted between them. As the final step, any remaining extra spaces and digits are removed from the processed equation. For the encoder component, we utilize a Bi-LSTM (Bidirectional Long Short-Term Memory) model, featuring 256 hidden units, 32 embedding dimensions, and a batch size of 32. Additionally, the LSTM cell within the model is configured with the same set of parameters.

Furthermore, the training process incorporates the Adam optimizer with a learning rate of 0.001, which aids in optimizing the model's performance. The training specifications include 10 epochs, a batch size of 100, and a dropout rate of 0.5. These settings ensure effective model training and help mitigate overfitting, leading to improved generalization capabilities. The architecture of the GRU-LSTM based Seq2seq Model is depicted in Figure 2.

4.2 Multi-Head Attention Mechanism-based Model

The Multi-head Attention module is an advanced component that operates in a complex manner through attention mechanisms. It iterates simultaneously and repeatedly within the attention mechanism. This process carefully combines the independent attention outputs using linear combinations, resulting in a meticulously crafted dimension. By using multiple attention heads, a system can focus on different elements in a sequence, making it easier to prioritize longer-term or shorter-term dependencies as needed. The Transformer architecture follows a hierarchical structure, using stacked self-attention and fully connected layers in both the encoder and decoder. This design allows for thorough modeling of relationships and information flow at various levels, resulting in improved understanding and generation of sequences. Figure 3 visually shows the architectural design, with the left and right portions representing different aspects. This graphical representation helps us understand how the components are arranged and interact with each other within the system.

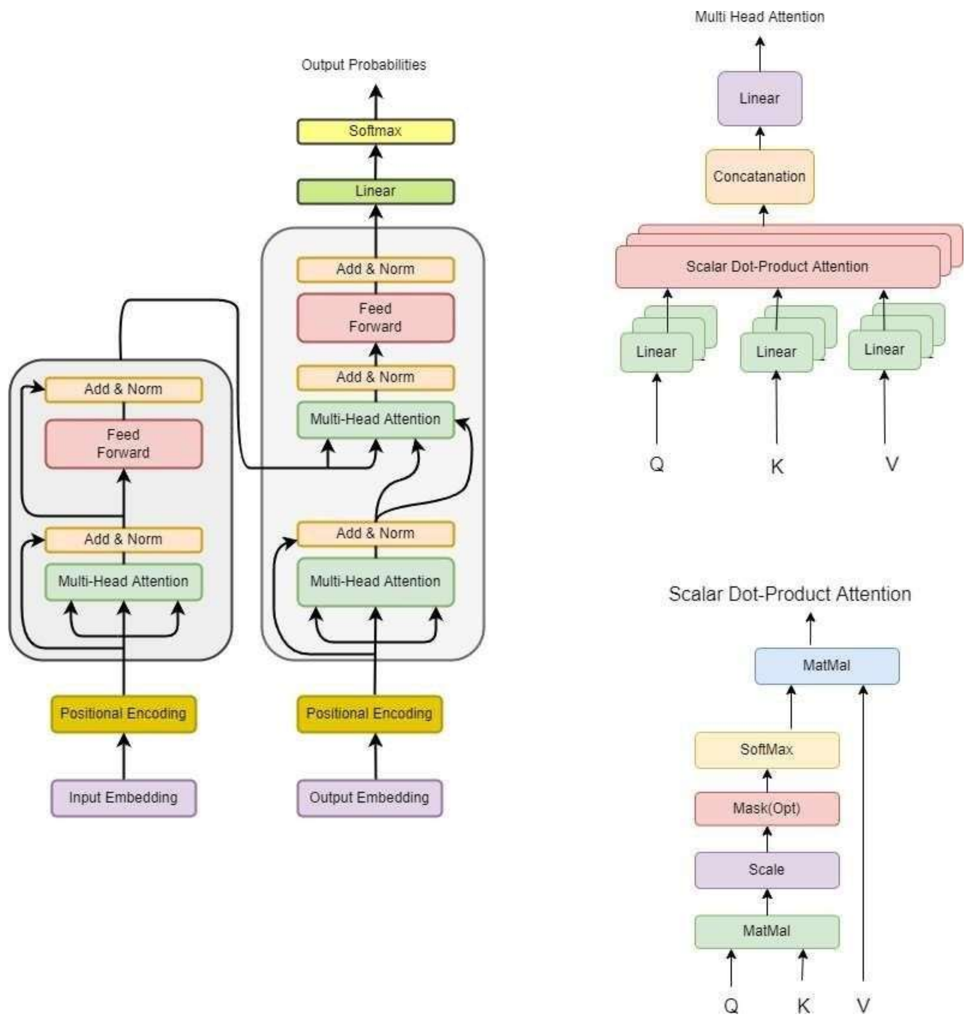


Figure 3: Multi-Head Attention Model

4.2.1 Encoder and Decoder Stacks

The Encoder and Decoder stacks have significant importance in the Multi-head Attention mechanism, which serves as a fundamental building block in the Transformer architecture widely utilized across various natural language processing applications. The Encoder and decoder of our multi-head attention mechanism-based model consist of 4 identical layers. Each layer of encoder and decoder is divided into two sub-layers. There are two components involved. The initial component is a self-attention mechanism with multiple heads, while the subsequent component is a positioned fully connected feed-forward network. Apart from the two sub-layers already present, the decoder introduces an additional sub-layer to each layer of the encoder.

The sub-layer that has been newly introduced conducts multi-head attention operations on the output of the encoder stack, thereby enhancing the model's ability to capture intricate relationships and dependencies within the encoded information. Within both the encoder and decoder components, once a residual connection has been applied around each of the two sub-layers, the subsequent step involves the application of layer normalization. To provide a concise explanation, we can express each sub-output layer as LayerNorm(a + Sublayer(a)), where the term Sublayer(a) denotes the function performed by the sub-layer itself.

Moreover, within the decoder stack, we introduce modifications to the self-attention sub-layer in order to enforce a constraint that prevents positions from attending to future positions.

4.2.2 Scaled Dot-Product Attention

Scaled dot-product attention consists of a number of key-value pairs that map a query to an output. To determine the weights on the values, we compute the dot products of the query with each key and divide each result by d_k. The obtained weights are normalized using the Softmax function.

For the Query, Key, and Value, there are three different linear layers. The weights of each linear layer are unique. These linear layers process the input to create the Q, K, and V matrices. The equation is presented as follows:

$$\text{Attention}(Q, K, V) = softmax \frac{Qk^T}{d_k} \sqrt{} V \quad (1)$$

4.2.3 Multi-Head Attention

The Attention module of the Transformer carries out its computations in a repetitive and parallel manner, which are known as Attention Heads. The Attention module divides the Query, Key, and Value parameters into N parts and processes each split separately through individual Heads. After conducting all the relevant Attention calculations, a final Attention score is generated by aggregating the results of these computations. By employing a method called "Multi-head attention," the Transformer enhances its ability to capture diverse relationships and intricate details associated with each word during the encoding process.

$$head_i = attention(QW_i^Q, KW_i^K, VW_i^Q)_i \quad (2)$$

$$Multihead(Q, K, V) = concat(head_i, \ldots, head_m)W^o \quad (3)$$

Where the projections are parameter matrices $W_i^Q R^{d_{model} \times d_k}$, $W_i^K R^{d_{model} \times d_k}$, $W^V{}_i R^{d_{model} \times d_v}$ and $W_O R^{hd_v \times dmodel}$.

4.2.4 Self Attention

The input sequence is broken down into three parts during self-attention: queries, keys, and values. These components undergo linear transformations to generate representations. Attention scores are then computed to determine the significance of each word or token to others, based on the relationship between queries and keys. By dividing the values by the attention scores, a weighted sum is obtained, representing the final representation of each word or token.

The utilization of multiple heads in self-attention enables the model to capture a variety of relationships among words or tokens. Each head focuses on a distinct area of the input, allowing the model to process long-range dependencies and gather both local and global data with efficiency.

5 Results

Deep learning has found its way into various facets of our everyday lives, with certain applications demanding significant computational power and extensive datasets. Parallel computing has emerged as a solution to handle such demanding scenarios. One prominent example of parallel processing is the utilization of graphics processing units (GPUs). In our specific task of generating mathematical equations, we employed GPUs to leverage their immense processing capabilities.

It is already mentioned that Multi-head mechanism-based model has four Encoder-Decoder layers and a 128-embedding dimension for input/output. We have four self-attention heads with a batch size of 64. For the Math23k, number word std, DRAW-1K, and Dolphin 18K datasets, models are trained in 30, 50, 65, and 55 epochs respectively. We employ the Adam optimizer to fine-tune the models, utilizing a learning rate of 0.001.

Cloud-based solutions are popular because it do not require system maintenance or configuration. There are many companies that provide cloud-based services such as Amazon, Google, Azure, and Intel. We employed Google Colaboratory (also referred to as Colab) for our experiment. Colab is a cloud-centric platform built upon Jupyter Notebooks [47]. Jupyter is a tool that is open-source and accessible through a web browser. It is very easy to use and can be used both locally and in the cloud. Google Colaboratory has some limitations. After 12 hours we need to configure the Colab.

Our proposed Multi-Head attention mechanism produced promising results when compared to other rule-based and statistical methods. Our model, on the other hand, performs well on the large complex dataset. It can generate a wide range of equations. Both models perform well on the Math23K dataset but poorly when the math word problem contains two unknown variables.

As mentioned earlier, in order to address math word problems featuring multiple unknown variables, we partitioned the dataset (consisting of Dolphin T2 Final and DRAW 1K) into two subsets: one containing math problems with a single unknown variable, and the other containing math problems with multiple unknown variables. The resulting output generated by our system can be observed in Table 6, providing a comprehensive overview of the outcomes obtained through our approach.

On the other hand, stop words are generally considered to have minimal significance in sentences. Consequently, our model exhibits better performance when we employ customized stop words. The generated output produced by our system is presented in Table 6, showcasing the results achieved through our approach.

For evaluation, we have used two evaluation matrices. A detailed description of those evaluation matrices is given below.

BLEU: In comparison to using human judgments, automatic machine translation evaluation metrics offer a quicker and less expensive method of assessing translation quality. The benchmark evaluation metric (BLEU) for machine translation (MT) has been shown to have a fair amount of agreement with human judges despite being straightforward and independent of language[48]. It is frequently employed as a loss function for discriminative training in addition to being used in the evaluation. BLEU was created to evaluate MT output across a wide range of references and sizeable documents.

ROUGE: It stands for Recall-Oriented Understudy for Gisting Evaluation, is commonly employed as a reference to the evaluation metric used for assessing the quality of automatic summaries. It has several automated evaluation techniques that gauge how similar the summaries are to one another [49]. The measurements count the number of units that are provided by the computer-generated equations to be evaluated and the ideal equations, such as n-grams, word sequences, and word pairs.

In our experimental setup, we divided the dataset into two distinct subsets, allocating 80% of the data for training purposes and reserving the remaining 20% for testing. For evaluation, we used the BELU score. Table 5 provides a detailed summary of the outcomes for those datasets. On the Math23K dataset, our model performs exceptionally well.

6 Observation

Our proposed Multi-Head attention mechanism produced promising results when compared to other rule-based and statistical methods. In contrast, our model demonstrates strong performance when dealing with extensive and intricate datasets. It can generate a wide range of equations. Both models perform better on the Math23K dataset but poorly when the math word problem contains two unknown variables.

We analyze the observations based on the dataset, considering that our models yield distinct outputs for each dataset. This discussion allows us to gain insights into the performance and variations exhibited by our models when confronted with different datasets.

Furthermore, it is widely acknowledged that stop words have limited significance in sentences. Consequently, our model demonstrates enhanced performance when customized stop words are employed. The output generated by our system is presented in Table 6, showcasing the results achieved through this utilization. The subsequent sections discuss various observations based on the dataset, providing a comprehensive analysis of the different insights obtained from our study.

6.1 DRAW-1K

The DRAW-1K dataset contains 1,000 problems and 2,330 sentences. A significant majority, about 74.5%, of the problems in this dataset involve more than one unknown variable, while only 25.5% have a single unknown variable. The average question length is relatively longer at 103.90, suggesting that the questions in this dataset are more detailed or complex. Similar to the Dolphin T2 Final dataset, the equation length is relatively short, averaging at 12.27 characters.

Model 3 demonstrates superior performance in the DRAW-1K dataset when handling multiple variables compared to single variables alone. Its ability to effectively incorporate and analyze multiple variables leads to more accurate and robust outcomes. Furthermore, when considering the entire dataset, Model 3 achieves its best results by leveraging both single and multiple variables. By considering the full spectrum of available information, the model gains a comprehensive understanding of the data, resulting in optimal performance.

6.2 Dolphin T2 Final

Dolphin T2 Final dataset consists of 831 problems and 831 associated sentences. Approximately 26.23% of the problems in this dataset involve more than one unknown variable, while 73.76% have only one unknown variable. The average question length is 75.64, indicating that the questions in this dataset tend to be relatively concise. The equation length is relatively shorter with an average of 12.93 characters.

Table 5 shows the effectiveness of Model 3 in addressing both single-variable and multiple-variable problems within the Dolphin T2 Final dataset. Model 2 also showcases the notable performance, while Model 1 needs some improvement. Table 3 also shows that Model 3 performs better for the single variable variation in the Dolphin T2 Final dataset compared to the multiple variables and combined datasets. This observation suggests that Model 3 is more effective in generating accurate and relevant responses when there is only a single variable involved in the input data.

Dataset Name	Dataset Variation	Model	BLEU	ROUGE-1			ROUGE-2		
				r	p	f	r	p	f
DRAW 1K	Single Variable	Model 1	0.07	0.3002	0.3021	0.3142	0.1253	0.1224	0.1178
		Model 2	0.17	0.5012	0.5212	0.5112	0.2122	0.2431	0.2152
		Model 3	**0.22**	**0.550**	**0.5535**	**0.5346**	**0.2485**	**0.2890**	**0.2588**
	Multiple Variable	Model 1	0.21	0.4236	0.2113	0.2242	0.2315	0.2414	0.2486
		Model 2	0.38	0.5524	0.5621	0.5412	0.4215	0.4125	0.4211
		Model 3	**0.40**	**0.5700**	**0.5881**	**0.5760**	**0.4558**	**0.4358**	**0.4358**
	Combined Dataset	Model 1	0.20	0.4256	0.4235	0.4351	0.2453	0.2581	0.5413
		Model 2	0.34	0.5723	0.5653	0.5692	0.4122	0.4089	0.4156
		Model 3	**0.41**	**0.6198**	**0.6114**	**0.6029**	**0.4227**	**0.4187**	**0.4187**
Dolphin T2 Final	Single Variable	Model 1	0.35	0.5927	0.5837	0.5892	0.4567	0.4628	0.4582
		Model 2	0.68	0.7512	0.7215	0.7304	0.6033	0.5728	0.5641
		Model 3	**0.71**	**0.7738**	**0.7434**	**0.7562**	**0.6231**	**0.5734**	**0.5924**
	Multiple Variable	Model 1	0.31	0.5584	0.5428	0.5872	0.4235	0.4158	0.4251
		Model 2	0.47	0.7311	0.7135	0.7012	0.5145	0.5042	0.5095
		Model 3	**0.51**	**0.7823**	**0.7224**	**0.7499**	**0.5430**	**0.5107**	**0.5244**
	Combined Dataset	Model 1	0.52	0.7952	0.7825	0.7836	0.5825	0.5721	0.5821
		Model 2	0.64	0.7152	0.7145	0.7288	0.5672	0.5245	0.5289
		Model 3	**0.67**	**0.7577**	**0.7327**	**0.7425**	**0.6022**	**0.5701**	**0.5811**
Math 23K	Single Variable	Model 1	0.67	0.7125	0.7254	0.7173	0.5897	0.5862	0.5729
		Model 2	0.91	0.8521	0.8511	0.8580	0.6127	0.6158	0.6156
		Model 3	**0.92**	**0.8769**	**0.8712**	**0.8721**	**0.5990**	**0.5920**	**0.5939**

Table 5: Result of Our Models (i.e. Model 1: GRU-LSTM based Seq2seq Model, Model 2: Multi-Head Attention Mechanism Based Model, Model 3: Multi-Head Attention Mechanism Based Model without stop word+ Comparison with other model)

6.3 Math23k

The Math23K dataset is the largest among the three, consisting of 23,162 problems and a substantial 70,125 sentences. All problems in this dataset involve only a single unknown variable. The average question length is 70.76, indicating moderately sized problem statements. The equation length is slightly longer compared to the other datasets, averaging 13.944 characters.

Model 3 consistently gives better results than the other models in terms of BLEU, ROUGE-1, and ROUGE-2 scores, indicating its superior performance in generating accurate solutions for Math23K. Model 2 also performs exceptionally well, achieving high scores across all evaluation metrics.

The complexity of sentence structures within math word problems poses a challenge for our system's recognition capabilities, leading to less satisfactory results on that specific dataset.

On the other hand, the following complex sentence is also presented in the math word problem.

Solving math word problems becomes challenging when they involve complex forms. In Table 7, we provide additional examples of math word problems that exhibit complex structures. Our system's performance is influenced by the length of the problem statement. Furthermore, when confronted with extensive equations, our system may struggle to generate the correct equation.

Addressing these challenges is an important area for future work. Developing techniques to handle math word problems with complex forms will be crucial in improving the system's performance. Additionally, devising strategies to handle longer problem statements and large equations will contribute to more accurate equation generation.

By focusing on these areas of improvement, we can enhance the system's ability to tackle complex math word problems, irrespective of their length or equation size.

7 Conclusion

We proposed a seq2seq model that uses a Multi-Head attention mechanism to generate equations from math word problems. After conducting experiments on four renowned math word problem datasets, the obtained results provide compelling evidence that the proposed model surpasses the performance of the state-of-the-art statistical model. Notably, our model exhibits a greater level of complexity within the realm of math word problems, primarily due to its ability to effectively handle a substantial quantity of unknown variables. We can still improve the system's accuracy in some areas. Moving forward, our aspirations lie in expanding upon this research endeavor to delve into the realm of generating nonlinear equations. Furthermore, we intend to apply these techniques in a variety of domains related to word problems, such as physics, chemistry, and others.

DRAW 1K	**Single Variable**	**Problem**	A number added to 6 is equal to 30 less than four times the number. what is the number.
		Actual Output	6 + x = 4 * x - 30
		Predicted Output	8 + x = 6 x + 24
	Multiple Variable	**Problem**	the larger of two numbers is 7 more than three times the smaller number. When the smaller number is subtracted from the larger number, the result is 13. find the two numbers.
		Actual Output	x = 3 * y + 7 , x - y = 13
		Predicted Output	x = 5 * y + 7 , x - y = 45
	Combined Dataset	**Problem**	Jim can fill a pool carrying buckets of water in 30 minutes. Sue can do the same job in 45 minutes. Tony can do the same job in 90 minutes. how quickly can all three fill the pool together?
		Actual Output	x * (1 / 30 + 1 / 45 + 1 / 90) = 1
		Predicted Output	x * (1 / 25 - 1 / 52 + 1 / 2) = 1
Dolphin T2 Final	**Single Variable**	**Problem**	The sum of two consecutive integers is 237 . Find the two integers.
		Actual Output	x + (x + 1.0) = 237.0
		Predicted Output	x + (x + 1.0) = 125.0
	Multiple Variable	**Problem**	The sum of two numbers is 52. The larger number is three times the smaller number . Find the larger number
		Actual Output	x + y = 52.0 , y = 3.0 * x
		Predicted Output	x + y = 45.0 , y = 3.0 * x
	Combined Dataset	**Problem**	there are 3 consecutive integers. the sum of the first two integers is 16 more than the third. Find integers
		Actual Output	(x - 1.0) + x = (x + 1.0) + 16.0
		Predicted Output	(x - 1.0) + x = (x + 1.0) + 14.0
Math 23k	**Single Variable**	**Problem**	A certain kind of book is priced at 72 yuan per set of 6 . The bookstore sold 18 sets of such books yesterday. How much did they sell for?
		Actual Output	x = 72 * 18
		Predicted Output	x = 72 * 18

Table 6: Output Produced by Multi-Head Attention Mechanism

This paper addresses the challenges posed by math word problems containing more than two unknown variables. These complex problems can be particularly demanding to solve effectively. As is customary, math word problems often involve the four fundamental operations (+, -, /, *). To gain further insights, we plan to partition the dataset based on these fundamental operations. Subsequently, we conduct an in-depth examination of the performance of each partitioned dataset.

On the other hand, we also looked at how the system would perform in the attention model if we split the joint sentence and divided the multiple sentences into separate sentences.

Conflict of interest statement

The authors state that they do not have any conflicts of interest to disclose. All co-authors have reviewed and approved the manuscript's content, and no financial interests are to be reported. The authors affirm that the submission represents original work and is not being considered for publication elsewhere.

References

[1] M. Hickendorff, The demands of simple and complex arithmetic word problems on language and cognitive resources, Frontiers in Psychology 12 (2021). doi:10.3389/fpsyg.2021.727761.

	Problem Statement	Actual Output	Predicted Output
Case 1	A salesperson at a machine company is offered an incentive plan. They earn a commission of $40 for each machine they sell. Additionally, the commission increases by $0.04 for every machine sold beyond $600. The salesperson wants to know how many machines they need to sell to reach a total commission of $30,800.	30800=40*x+0.04*(x-600)	30800=40*x+0.04*(x-600)
Case 2	The total of a two-digit number is 9. When the digits are swapped, the resulting number is six times the sum of the original digits. Determine the original number.	smaller+larger=9, 10 * smaller + larger = 6 *(smaller + larger)	x + y = 9, 10 * x + y = 4 x + 5y

Table 7: Challenging Math Word Problem

URL https://www.frontiersin.org/articles/10.3389/fpsyg.2021.727761

[2] Y. Hong, Q. Li, D. Ciao, S. Huang, S. Zhu, Learning by fixing: Solving math word problems with weak supervision, CoRR abs/2012.10582 (2020). arXiv:2012.10582.
URL https://arxiv.org/abs/2012.10582

[3] S. A. Crossley, S. Karumbaiah, J. Ocumpaugh, M. J. Labrum, R. S. Baker, Predicting math identity through language and click-stream patterns in a blended learning mathematics program for elementary students, Journal of Learning Analytics 7 (1) (2020) 19–37. doi:10.18608/jla.2020.71.3.
URL https://learning-analytics.info/index.php/JLA/article/view/6585

[4] G. Daroczy, M. Wolska, W. D. Meurers, H.-C. Nuerk, Word problems: a review of linguistic and numerical factors contributing to their difficulty, Frontiers in Psychology 6 (2015). doi:10.3389/fpsyg.2015.00348.
URL https://www.frontiersin.org/articles/10.3389/fpsyg.2015.00348

[5] Z. Liang, X. Zhang, Solving math word problems with teacher supervision, in: Z.-H. Zhou (Ed.), Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21, International Joint Conferences on Artificial Intelligence Organization, 2021, pp. 3522–3528, main Track.

[6] A. Pathak, P. Pakray, S. Sarkar, D. Das, A. F. Gelbukh, Mathirs: Retrieval system for scientific documents, Computacio´n y Sistemas 21 (2) (2017).
URL http://www.cys.cic.ipn.mx/ojs/index.php/CyS/article/view/2743

[7] S. Sarkar, D. Das, P. Pakray, D. Pinto, Generating equations from math word problem using deep learning approach, in: S. Mukhopadhyay, S. Sarkar, P. Dutta, J. K. Mandal, S. Roy (Eds.), Computational Intelligence in Communications and Business Analytics, Springer International Publishing, Cham, 2022, pp. 252–259.

[8] A. Mukherjee, U. Garain, A review of methods for automatic understanding of natural language mathematical problems, Artif. Intell. Rev. 29 (2008) 93–122. doi:10.1007/s10462-009-9110-0.

[9] S. Powell, Solving word problems using schemas: A review of the literature, Learning disabilities research practice : a publication of the Division for Learning Disabilities, Council for Exceptional Children 26 (2011) 94–108. doi:10.1111/j.1540-5826.2011.00329.x.

[10] S. S. Sundaram, D. Khemani, Natural language processing for solving simple word problems, in: Proceedings of the 12th International Conference on Natural Language Processing, NLP Association of India, Trivandrum, India, 2015, pp. 394–402.
URL https://aclanthology.org/W15-5955

[11] B. Siyam, A. A. Saa, O. Alqaryouti, K. Shaalan, Arabic arithmetic word problems solver, Procedia Computer Science 117 (2017) 153–160, arabic Computational Linguistics. doi:https://doi.org/10.1016/j.procs.2017.10.104.
URL https://www.sciencedirect.com/science/article/pii/S1877050917321610

[12] K. Griffith, J. Kalita, Solving arithmetic word problems automatically using transformer and unambiguous representations, CoRR abs/1912.00871 (2019).

arXiv:1912.00871.
URL http://arxiv.org/abs/1912.00871

[13] K. Yokoi, A. Aizawa, An approach to similarity search for mathematical expressions using mathml, 2009.

[14] Q. Wu, Q. Zhang, X. Huang, Automatic math word problem generation with topic-expression co-attention mechanism and reinforcement learning, IEEE/ACM Transactions on Audio, Speech, and Language Processing 30 (2022) 1061–1072. doi:10.1109/TASLP.2022.3155284.

[15] B.-S. Ilany, Language and mathematics: Bridging between natural language and mathematical language in solving problems in mathematics, Creative Education 01 (2010) 138–148. doi:10.4236/ce.2010.13022.

[16] S. Mandal, S. K. Naskar, Classifying and solving arithmetic math word problems—an intelligent math solver, IEEE Transactions on Learning Technologies 14 (1) (2021) 28–41. doi:10.1109/TLT.2021.3057805.

[17] L. Verschaffel, S. Schukajlow, J. Star, W. Dooren, Word problems in mathematics education: a survey, ZDM 52 (2020) 1–16. doi:10.1007/s11858-020-01130-4.

[18] D. Zhang, L. Wang, N. Xu, B. T. Dai, H. T. Shen, The gap of semantic parsing: A survey on automatic math word problem solvers, CoRR abs/1808.07290 (2018). arXiv:1808.07290.
URL http://arxiv.org/abs/1808.07290

[19] S. Acharya, R. Basak, S. Mandal, Solving arithmetic word problems using natural language processing and rule-based classification, International Journal of Intelligent Systems and Applications in Engineering 10 (2022) 87–97. doi:10.18201/ijisae.2022.271.

[20] A. Amini, S. Gabriel, S. Lin, R. Koncel-Kedziorski, Y. Choi, H. Hajishirzi, Mathqa: Towards interpretable math word problem solving with operation-based formalisms, CoRR abs/1905.13319 (2019). arXiv:1905.13319.
URL http://arxiv.org/abs/1905.13319

[21] C. Fletcher, Understanding and solving arithmetic word problems: A computer simulation, Behavior Research Methods 17 (5) (1985) 565–571. doi:10.3758/BF03207654.

[22] Y. Bakman, Robust understanding of word problems with extraneous information (2007). doi:10.48550/ARXIV.MATH/0701393.
URL https://arxiv.org/abs/math/0701393

[23] N. Kushman, Y. Artzi, L. Zettlemoyer, R. Barzilay, Learning to automatically solve algebra word problems, in: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Baltimore, Maryland, 2014, pp. 271–281. doi:10.3115/v1/P14-1026.
URL https://aclanthology.org/P14-1026

[24] K. Wang, Z. Su, Dimensionally guided synthesis of mathematical word problems, in: Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence, IJCAI'16, AAAI Press, 2016, p. 2661–2668.

[25] M. J. Hosseini, H. Hajishirzi, O. Etzioni, N. Kushman, Learning to solve arithmetic word problems with verb categorization, in: Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Association for Computational Linguistics, Doha, Qatar, 2014, pp. 523–533. doi:10.3115/v1/D14-1058.
URL https://aclanthology.org/D14-1058

[26] T. Liu, Q. Fang, W. Ding, Z. Wu, Z. Liu, Mathematical word problem generation from commonsense knowledge graph and equations, CoRR abs/2010.06196 (2020). arXiv:2010.06196.
URL https://arxiv.org/abs/2010.06196

[27] D. Alvarez-Melis, T. S. Jaakkola, Tree-structured decoding with doubly-recurrent neural networks, in: International Conference on Learning Representations, 2017.
URL https://openreview.net/forum?id=HkYhZDqxg

[28] Q. Liu, W. Guan, S. Li, D. Kawahara, Tree-structured decoding for solving math word problems, in: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Association for Computational Linguistics, Hong Kong, China, 2019, pp. 2370–2379. doi:10.18653/v1/D19-1241.
URL https://aclanthology.org/D19-1241

[29] L. Wang, Y. Wang, D. Cai, D. Zhang, X. Liu, Translating a math word problem to a expression tree, in: Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Brussels, Belgium, 2018, pp. 1064–1069. doi:10.18653/v1/D18-1132.
URL https://aclanthology.org/D18-1132

[30] D. Singh, K. Suraksha, S. Nirmala, Question answering chatbot using deep learning with nlp, in: 2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2021, pp. 1–6. doi:10.1109/CONECCT52877.2021.9622709.

[31] S. Sarkar, D. Das, P. Pakray, D. Pinto, A hybrid sequential model for text simplification, in: N. Priyadarshi,

S. Padmanaban, R. K. Ghadai, A. R. Panda, R. Patel (Eds.), Advances in Power Systems and Energy Management, Springer Singapore, Singapore, 2021, pp. 33–42.

[32] P. C. Shilpa, R. Shereen, S. Jacob, P. Vinod, Sentiment analysis using deep learning, in: 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 930–937. doi:10.1109/ICICV50876.2021.9388382.

[33] S. P. Singh, A. Kumar, H. Darbari, L. Singh, A. Rastogi, S. Jain, Machine translation using deep learning: An overview, in: 2017 International Conference on Computer, Communications and Electronics (Comptelix), 2017, pp. 162–167. doi:10.1109/COMPTELIX.2017.8003957.

[34] B. Robaidek, R. Koncel-Kedziorski, H. Hajishirzi, Data-driven methods for solving algebra word problems, CoRR abs/1804.10718 (2018). arXiv:1804.10718.
URL http://arxiv.org/abs/1804.10718

[35] A. Patel, S. Bhattamishra, N. Goyal, Are NLP models really able to solve simple math word problems?, CoRR abs/2103.07191 (2021). arXiv:2103.07191.
URL https://arxiv.org/abs/2103.07191

[36] Z. Wang, A. S. Lan, R. G. Baraniuk, Math word problem generation with mathematical consistency and problem context constraints, CoRR abs/2109.04546 (2021). arXiv:2109.04546.
URL https://arxiv.org/abs/2109.04546

[37] W. Ling, D. Yogatama, C. Dyer, P. Blunsom, Program induction by rationale generation: Learning to solve and explain algebraic word problems, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Vancouver, Canada, 2017, pp. 158–167. doi:10.18653/v1/P17-1015.
URL https://aclanthology.org/P17-1015

[38] L. Wang, D. Zhang, J. Zhang, X. Xu, L. Gao, B. T. Dai, H. T. Shen, Template-based math word problem solvers with recursive neural networks, Proceedings of the AAAI Conference on Artificial Intelligence 33 (01) (2019) 7144–7151. doi:10.1609/aaai.v33i01.33017144.
URL https://ojs.aaai.org/index.php/AAAI/article/view/4697

[39] Q. Wu, Q. Zhang, Z. Wei, X. Huang, Math word problem solving with explicit numerical values, in: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 1: Long Papers), Association for Computational Linguistics, Online, 2021, pp. 5859–5869. doi:10.18653/v1/2021.acl-long.455.
URL https://aclanthology.org/2021.acl-long.455

[40] Z. Liang, J. Zhang, J. Shao, X. Zhang, MWP-BERT: A strong baseline for math word problems, CoRR abs/2107.13435 (2021). arXiv:2107.13435.
URL https://arxiv.org/abs/2107.13435

[41] P. Mehta, P. Mishra, V. Athavale, M. Shrivastava, D. Sharma, Deep neural network based system for solving arithmetic word problems, in: Proceedings of the IJCNLP 2017, System Demonstrations, Association for Computational Linguistics, Tapei, Taiwan, 2017, pp. 65–68.
URL https://aclanthology.org/I17-3017

[42] D. Huang, S. Shi, C.-Y. Lin, J. Yin, W.-Y. Ma, How well do computers solve math word problems? large-scale dataset construction and evaluation, in: Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Association for Computational Linguistics, Berlin, Germany, 2016, pp. 887–896. doi:10.18653/v1/P16-1084.
URL https://aclanthology.org/P16-1084

[43] Q. Zhou, D. Huang, Towards generating math word problems from equations and topics, in: Proceedings of the 12th International Conference on Natural Language Generation, Association for Computational Linguistics, Tokyo, Japan, 2019, pp. 494–503. doi:10.18653/v1/W19-8661.
URL https://aclanthology.org/W19-8661

[44] S. Upadhyay, M. Chang, Annotating derivations: A new evaluation strategy and dataset for algebra word problems, CoRR abs/1609.07197 (2016). arXiv:1609.07197.
URL http://arxiv.org/abs/1609.07197

[45] S. Upadhyay, M.-W. Chang, Draw: A challenging and diverse algebra word problem set, 2015.

[46] Y. Wang, X. Liu, S. Shi, Deep neural solver for math word problems, in: Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, Copenhagen, Denmark, 2017, pp. 845–854. doi:10.18653/v1/D17-1088.
URL https://aclanthology.org/D17-1088

[47] T. Carneiro, R. V. Medeiros Da No´Brega, T. Nepomuceno, G.-B. Bian, V. H. C. De Albuquerque, P. P. R. Filho, Performance analysis of google colaboratory as a tool for accelerating deep learning applications, IEEE Access 6 (2018) 61677–61685. doi:10.1109/ACCESS.2018.2874767.

[48] K. Papineni, S. Roukos, T. Ward, W.-J. Zhu, Bleu: a method for automatic evaluation of machine translation, in: Proceedings of the 40th Annual

Meeting of the Association for Computational Linguistics, Association for Computational Linguistics, Philadelphia, Pennsylvania, USA, 2002, pp. 311–318. doi:10.3115/1073083.1073135.
URL https://aclanthology.org/P02-1040

[49] C.-Y. Lin, ROUGE: A package for automatic evaluation of summaries, in: Text Summarization Branches Out, Association for Computational Linguistics, Barcelona, Spain, 2004, pp. 74–81.
URL https://aclanthology.org/W04-1013

Sandip Sarkar is working as an Assistant Professor in the Department of Computer Science and Application, Hijli College. He received his B.Tech in Computer Science and Engineering from Jalpaiguri Govt. Engineering College in 2011, followed by the completion of his M.E in Computer Science and Engineering from Jadavpur University in 2014. Currently, he is actively pursuing his Ph.D. in Computer Science and Engineering at the same university. His present research focus encompasses Question Answering, Information Retrieval, and Deep Learning.

Dipankar Das is an Assistant Professor in the Department of Computer Science and Engineering, Jadavpur University, India. He received Ph.D. and Master's degrees from the Department of Computer Science and Engineering, Jadavpur University in 2013 and 2009 respectively. He received Bachelor's degree in Computer Science and Engineering from West Bengal University of Technology in 2005. His research interests are in the area of Natural Language Processing, Emotion and Sentiment Analysis, Affect Computing, Information Extraction and Language Generation. He has more than 100 publications in top conferences and journals and has served as an author over 15 Book Chapters. He is a member of the IEEE, ACL, HUMAINE groups.

David Pinto obtained his PhD in artificial intelligence and pattern recognition at the Polytechnic University of Valencia, Spain in 2008. At present he is a full time professor at the Faculty of Computer Science of the Beneme´rita Universidad Auto´noma de Puebla (BUAP) leading the laboratory and PhD Program of Language & Knowledge Engineering. His areas of interest include clustering, information retrieval, crosslingual NLP tasks, computational linguistics, robotics, augmented reality, virtual reality, mobile devices and complexity theory. He has written more than 100 papers and has developed research projects, whose products have been registered in the Mexican Institute of Industrial Property. He has created various laboratories, being the last one, the Language & Knowledge Engineering Lab, visited by several national and international researchers for establishing collaboration in different topics of artificial intelligence.

From PMO to PMOCoE: How Manage Project Knowledge Process Improves Quality of Organization Knowledge Management Assets Cases from UAE

Ahmed Al-Nakeeb[*], Mounir El Khatib[†], Raed Abu Zitar[‡],

[*†]Hamdan Bin Mohammed Smart University, Dubai, U.A.E.
[‡]Sorbonne University, Abu Dhabi, U.A.E.

Abstract

Project Management Office Center of Excellence (PMOCoE) is one of the widely used strategies to promote project management as a core competency all over the organization. PMOCoE involves establishing a multidisciplinary team that unites various stakeholder representatives to facilitate the integration of project management disciplines into organizational practices, activities, and processes. For PMOCoE to drive the desired outcomes, it must promote organizational learning to become a "learning organization". This paper presents the fundamentals of PMOCoE, including general requirements, expected outcomes, and the pathways to improved organizational performance. In this research paper, the data was collected through conducting interviews with projectized entities based in UAE like Dubai Police and the Road and the Transportation Authority in Dubai City. This qualitative data helps to reach the best understanding of the literature review and analysis of some local cases in the UAE. The cases investigated show that the PMOCoE has been used successfully in the UAE. RTA and Dubai police are some of the organizations that have used this concept to improve their overall performance and achieve business growth.

Key words: PMO, PMOcoE, Project management office, Project management office centre of Excellence, manage project knowledge, knowledge management, knowledge management assets.

1 Introduction

Competition is intensifying daily in the global business environment. Various factors are changing, making the external operating environment very dynamic. Nevertheless, firms all over the world are struggling to gain strong competitiveness, which is the ultimate goal of every business. Project Management Office Center of Excellence (PMOCoE) is one of the widely used strategies that can help organizations to improve their competitiveness. Globally, companies have realized the need to promote project management as a core competency all over the organization. PMOCoE involves establishing a multidisciplinary team that unites various stakeholder representatives to facilitate the integration of project management disciplines into organizational practices, activities, and processes. This paper presents the fundamentals of PMOCoE, including general requirements, expected outcomes, and the pathways to improved organizational performance. It also focuses on case studies of companies that have used this concept effectively. For PMOCoE to drive the desired outcomes, it must promote organizational learning, which involves learning from the past, as well as creating and sharing knowledge within the organization and with external stakeholders. The creation and sharing of knowledge demand high levels of cooperation between departments or various functions of the organization. PMOCoE also utilizes the resulting knowledge to improve various elements of organizational performance, such as efficiency, framework, processes, and activities. Creating, sharing, and utilizing knowledge in an organization requires proper and effective management of the input, process, and output side of the PMOCoE. The management of the input side mainly involves evaluating the necessary resources and requirements, including planning, allocating, and controlling resources, as well as developing the frameworks. It is imperative to align the PMOCoE with the overall organizational goals, mission, vision, and values. The process aspect involves converting data and information into useful knowledge that is used to improve various elements of organizational performance. For that reason, the necessary tasks within the process elements include managing assets, processing data, and facilitating the creation, sharing, and utilization of knowledge. The proper management of tools, techniques, and systems

‡raed.zitar@sorbonne.ae, *a.alnakeeb@hbmsu.ac.ae, †m.elkhatib@hbmsu.ac.ae

contributes to attaining effective and efficient PMOCoE processes. Finally, the output phase concerns the application of the resulting knowledge. The common output elements include communication plans, budgets, and performance improvement. Data collection and analysis are critical aspects of PMOCoE. Consequently, the achievement of set objectives requires establishing proper and evidence-based practices for collecting and analysing data to generate useful knowledge. Most importantly, organizations should have clear and accurate documentation to complete their projects successfully. Documentation is necessary for storing data and information on experiences, which encourages continuous improvement. PMOCoE enables organizations to improve performance in various ways. The three main pathways include the reduction of business risks, resource optimization, and contribution to business growth. PMOCoE manages risks by enabling managers and business leaders to plan, accomplish, and monitor the core activities of the project or organizational activities. Resource optimization is also essential for improving overall performance. PMOCoE contributes to the growth of the business in terms of product and service portfolio, market share, productivity, revenues, and profits. PMOCoE has been used successfully in the UAE. RTA and Dubai police are two of many organizations that have used this concept to improve their overall performance and achieve business growth.

2 Research Methodology

To investigate the deployment and effectiveness of PMOCoE, a qualitative research approach was adopted, conducting in-depth interviews with key personnel at prominent projectized entities in the UAE, namely, the Roads and Transport Authority (RTA) in Dubai, and Dubai Police. These interviews were meticulously designed to elicit comprehensive insights into how these organizations integrate project management disciplines into their operational practices, activities, and processes. The qualitative nature of this study was critical for several reasons. It allowed for the gathering of detailed, context-rich data that quantitative methods might not capture, particularly relevant in understanding the nuanced concept of organizational learning within PMOCoEs. Furthermore, it facilitated a thorough exploration of the PMOCoE's implementation in distinct organizational cultures and structures, offering a granular perspective on the processes and outcomes associated with becoming a learning organization. Each interview was structured to not only corroborate the literature on PMOCoE but also to understand its practical application in leading UAE organizations like RTA and Dubai Police. The narratives collected from these discussions shed light on the successful application of PMOCoE principles and its tangible benefits in terms of business growth and performance enhancements. By focusing on entities that have successfully utilized PMOCoE, this paper contributes to the limited but growing body of knowledge on the subject, particularly within the Middle Eastern context. The qualitative insights derived from this research not only confirm the theoretical underpinnings of PMOCoE but also provide practical examples of its application, thereby offering valuable lessons for other organizations aspiring to achieve similar outcomes.

3 Literature Review

3.1 Introduction to PMO and PMOCoE:

Projects are composed of diverse set of people, processes, and resources, which are instrumental for project completion and success. Among these are the Project Management Office (PMO) and Center of Excellence (CoE) which hold important roles and functions for every project. The 21st century business environment offers challenges and important implications to businesses handling projects. According to [1], One of the key strengths of the Project Management Office (PMO) is its ability to provide clear visibility into project plans, ongoing progress, and project tracking. The use of standardized project reporting systems ensures that precise project details are readily accessible. This facilitates improved decision making and enhances coordination among different projects. Project Management Center of Excellence champions competences and best project management practices' use within the organization [2]. Center of Excellence PMOs provide the organization with methodologies, standards, and tools to enable project managers to deliver projects successfully. In addition, this type of PMO ensures organizational success through good practices, tools, and processes. As the central point of contact for project management in the organization, 35% of PMO practitioners reported having this type of PMO function [3][4]. It is essential for organizations of different types to carefully consider the different roles of stakeholders involved in projects and the different strategies and practices that could bring project success. According to PMI [5], a project is a temporary endeavour undertaken to create a unique product, service, or result. A project is temporary in that it has a defined beginning and end in time, and therefore defined scope and resources. A project is unique in that it is not a routine operation, but a specific set of operations designed to accomplish a singular goal. Project management is usually contained with five main processes: (1) initiating; (2) planning; (3) executing; (4) monitoring and controlling; and (5) closing. Meanwhile, project management knowledge draws on 10 areas, which are integration, scope, time, cost, quality, procurement, human resources, communications, risk management, and stakeholder management [6]. With these kinds of processes and areas to manage in projects, consulting Project Management Office (PMO) and Center of Excellence (CoE) can offer benefits that could help organizations combat the challenges of operating and managing projects in contemporary society. This paper will explore the roles and functions of the Project Management Office (PMO) and Center of Excellence (CoE) to best understand their essence in successful projects. How both PMOCoE works on adopting and integrating knowledge as assets for project management efficiency, will also be explored. Furthermore, the framework of project management and the best practices, alongside the case studies in the UAE setting, will also be discussed.

3.2 What is the PMO Center of Excellence (PMOCoE)

A Project Management Office Center of Excellence (PMO CoE) operates as a specialized division within the PMO, dedicated to the development and promotion of superior practices, standards, and methodologies pertinent to project management throughout the organization. This specialized sector is typically staffed by seasoned project managers and experts in various subject matters. These professionals offer essential guidance, support, and training to project teams and stakeholders. They play a pivotal role in fostering effective communication and cooperation among project managers, addressing, and resolving common project-related challenges, and ensuring that the project management methodologies are in sync with the organization's broader strategic goals and objectives. The primary objective of a PMO CoE is to instil a culture of excellence in project management within the organization. This initiative is aimed at enhancing the overall outcomes of projects, boosting efficiency, and minimizing risks. The PMO CoE achieves this by utilizing its wealth of expertise and resources to standardize project management procedures, elevate transparency in project progress and performance, and ultimately, contribute greater value to stakeholders [7].

Moreover, there are processes to follow when initiating PMOCoE, as initiating a project, which starts from planning to reporting and ensuring the enhancement. Some of the risks possible to accrue when deciding to build PMOCoE are considered as 1) Not being able to decide where to allocate the PMOCoE within the organization, 2) Unidentified values and the purpose of PMOCoE establishment [8]. Moreover, the Project Management Office most likely falls within the operational level of the organization, specifically within Information Technology since it's the favourable department that brings growth within the industry and adopts the newest innovation and technology [9]. In addition, the Project Management Office (PMO) and Center of Excellence (CoE) are both teams and partners working toward the same objective of completing the project successfully and achieving the desired results. The Project Management Office is more likely to drive the project structure and governance, whereas the Center of Excellence (CoE) is members within the organization who support the project. Furthermore, four main processes should be considered when an organization decides to build a Center of Excellence (CoE) such as, 1) Organizing the team and allocating members within PMOCoE, 2) Preparing a plan for action and jointly tracking the performance and feedback, 3) Communicate with the stakeholder and make alignments, 4) Understand the business value for PMOCoE and map the capabilities within the project life cycle [10]. Also, the main objectives of PMOCoE existence within the organization are A) to Provide guidance and ability to adopt organization strategies of the project management, B) to Ensure Project Managers (PM) have enough capabilities and knowledge, C) to present as core of implementing the best practices for the project management, D) to ensure the organization continuity and operation is maintained, E) to support PM with required resources to accomplish a project and agility [11].

3.3 PMOCoE and Project Knowledge Adoption

Managing knowledge will enable PMOCoE to capture the knowledge, gain understanding, and share the knowledge among teams in an effective and efficient manner. And the more knowledge is visible and accessible, the more it becomes useful to staff [12]. By provisioning types of knowledge within the departments of the organization, the Project Management Office as Center of Excellence is able to share the current information and merge this for building new knowledge. The evidence shows Project Management Office is the most likable place for combining all the knowledge than the other places since PMO is able to access knowledge frameworks, whether internally or externally. Some departments are involved in creating the knowledge; it acknowledges those departments as the creators of the knowledge and represents them as contributors. As the result, Project Management Office is responsible to choose a specific and suitable method to support the knowledge development and thus to be shared among entities. Also, certain reasons need to consider about knowledge which is not only one item to be shared among the departments, but it is the way those entities react to different organization techniques. Thus, the responsibility of Project Management Office is control, encourage and maintain those departments to contribute to developing the knowledge thus to be shared effectively. Moreover, to develop the knowledge, it demands members of the organization to corporate to result discovering new methods to be adopted. Also, knowledge is categorized into local and global with different attributes such as, local knowledge is developed within one department to be adequate within its own market state specifically individual and team, and global knowledge is developed by variety of collaborating organizations consequently to serve different markets specifically organizations. Though there is globally and locally developed and shared knowledge in the Project Management Office, the lessons learned are not always documented among the members of the group, resulting in a gap in certain important topics that need to be covered, such as A) Which organization technique is suitable followed with which conditions? B) What knowledge processes are to be considered? and the responsibility of the Project Management Office turns into updating the template and to be leveraged within the level of the project. Moreover, developing the knowledge itself is done locally within the Center of Excellence then transformed from one team to another, and it's a success whereas the shared knowledge turns useful item to team members [13].

3.4 How PMOCoE Can Integrate Lesson Learned and Organization Knowledge Assets to Improve Project Management Efficiency

The integration of the Project Management Office Center of Excellence, or the PMOCoE, to improve management efficiency within an organizational framework, requires constructive evaluation, especially in terms of the organizational processes undertaken within a company or institution. PMOCoE has been established in firms that have focused on attaining successful projects within their framework. However, literature analysts provide diverse views on the use of the related concepts. In providing a critical evaluation of the literature review, an in-depth analysis of the PMOCoE is needed on aspects of project management, including the input, the output, the processing aspects, and the tools and techniques used in achieving the targets of the projects. The Project Management Office Center of Excellence has been available to companies for a long time. However, the purposes, functions, and descriptions of these concepts have transformed over time. Initially, PMOCoE could integrate lessons learned through the planning of organizational knowledge assets. PMOCoE has been concerned with one project as a whole; with time, the aspects have engaged in multi-faceted parts of a project and even multi-project processes [14]. The change is necessary due to the demanding needs of project management. The transformation of the PMOCoE has seen variations in purpose and functions; thus, no ideal structure of the project management tool exists that could be followed for every project [15]. Through knowledge asset management, various concepts should be considered, including direct project management, standardization, and facilitation of resource-sharing and management. The management of input is very important. Management of the input aspects majorly requires proper evaluation of the resources needed for starting the project. The main input issues considered include project planning, allocation and management of resources, creation of the project framework, and determination of human resource roles. Project planning is a key stage that the PMOCoE has to achieve. In this phase, the project management team should consider the strategies used for the allocation of resources [16]. The Project Management Office Center of Excellence helps in determining the process of resource-allocation and management of assets during this time [17]. The projects should be aligned based on issues such as prioritization of their phases and processes. With these guidelines set in place, the PMOCoE has developed a framework that helps minimize waste in the project while finding the most appropriate business strategies that would assist in achieving the completion of the project [16]. Still, regarding the input and planning level, the Project Management Office Center of Excellence is designed to map the blueprint of the whole idea as far as goals and assets are concerned. Full executive support is meted towards achieving the link between the project and organizational assets and goals [17]. The PMOCoE also establishes the specific ways or strategies developed for the decision-making processes, especially in critical stages of its development. Project management does not just focus on how decisions are made but also discusses the relationship between decision-making and the long term impacts involved. They must adhere to the goals of asset management within the organization. The management of processes must also be taken into consideration. When conducting an assessment on project management, it is evident that knowledge management should be taken into consideration. The processing stages therein require the establishment of knowledge management through the PMOCoE [18]. According to [19], projects lose considerable knowledge through rotation or resignation of project managers, particularly during the processing phase. A new management structure inherits contractual issues, hence failing to accomplish the targets set in the beginning. In some cases, replacing the old structure would require more resources and energy that would have improved the performance of the project, since the previous framework had a unique way of managing it. In this case, knowledge management of assets is a critical aspect; hence, the PMOCoE has to be involved in creating a management structure that effectively reduces losses [20]. Another critical issue within the PMOCoE framework in the processing phase is the management of data on administrative functions. The project guidelines provide concepts of project management through assessing how information would be shared during the processing stages [17]. Such details include the status of the project in every stage, identification of the emerging risks occurring in the developmental phases, as well as addressing the potential risks or challenges that might take place due to emergence of certain concerns. Such responsibility is given to the information manager, who advises on the tracking and reporting updates. The management of tools and techniques is also a critical issue that the PMOCoE should achieve when addressing the longevity and success of a project. [21] discuss the relevance of a project model that puts into interest the management of data – inclusive of techniques – and tools. The PMOCoE establishes a cohesive structure that integrates tools necessary for management, project design, as well as reporting [17]. These concepts require proper evaluation to ensure that the whole project adequately utilizes all the major techniques to achieve the set objectives [22]. In addition, the involved coaching is important to ensure various techniques are passed to all employees. Project coaching involves offering training and mentoring as well as other project management assistance to the workers, particularly project managers. The rapid advancements in technology alongside globalization prompted organizations to invest largely on their IT capabilities. This is also their strategy to overcome the pressures brought by competition as well as to effectively manage their projects. Project management information system (PMIS) is acquired and utilized by the organizations to help managers in making key decisions required in planning, organizing, and controlling projects. As mentioned by Raymond & Bergeron (2007), 75% of huge IT projects that are facilitated through PMIS will succeed. Although it is not a standalone factor to project success, organizations recognize PMIS as a necessity. Apart from providing decision-making support to project managers, PMIS is also a prerequisite in the effective and efficient management of projects. In a complex business environment today, business organizations ensure that achieving efficiency and success in project management is guaranteed. According to [23] the use of PMIS does just that, as it is a "comprehensive system that supports the entire lifecycle of projects, project programs, and project portfolios". It is further noted that there are several essential factors that influence project managers to utilize PMIS. These include (1) the quality of information generated by the system; (2) a sufficient level of detail that the PMIS provides the managers particularly relevant to their needs; (3) information generated by the system is understandable and easy to interpret and share to the members of the project team; and (4) PMIS supports continuous progress tracking and evaluation [23]. As efficiency and success in project management are emphasized through the use of PMIS, it is denoted that this may be accomplished because the system supports three primary functions, including communication, collaboration and community. Communication is reflected in how information and knowledge are promptly shared between team members via external or internal networks. Collaboration and community, on the other hand, are reflected in the cooperative management fostered between and among project team members and in the process of accumulating and sharing data and information [24]. Additionally, Karim noted that the main role of PMIS is described as acquiescent to the implementation of project strategies and the accomplishment of project goals, thereby reiterating the fact that PMIS is a key component in project management and project success. The management of output is just as important as the management of the input. The Project Management Centre of Excellence is important in addressing the main issues, especially on the output processing. Specifically, the PMOCoE concepts can be utilized in implementing the project through the management of a communication plan, budget baseline, assessment of procurement aspects and individual project goals. The latter aspect mainly addresses the completion and implementation of the project. In addition, it focuses on major effects of the project in achieving the organizational goals. The PMOCoE guidelines provide interest in the quality of the project and the analysis of whether it achieved the objectives set by the organization. The continuity of the project should be assessed, weighing on the challenges of the project versus the outcomes. In addition, this aspect should address how the project may affect the succeeding projects that would be developed by the organization.

3.5 Framework of Project Management

The major framework of the Project Management Centre of Excellence helps in understanding the current needs of project management practices. When addressing project management, various concepts should be taken into account. The discussion above has mainly cited concerns about management of assets, proper planning and delivery of services, establishment of project and organizational goals, assessment of risks, and other related factors such as change leadership [17]. Based on the framework below (Figure 1), five aspects should be assessed in achieving a successful project management process. These facets of the PMOCoE are interconnected and, therefore, depend on each other. When they are successfully implemented, a proper project management system will be established.

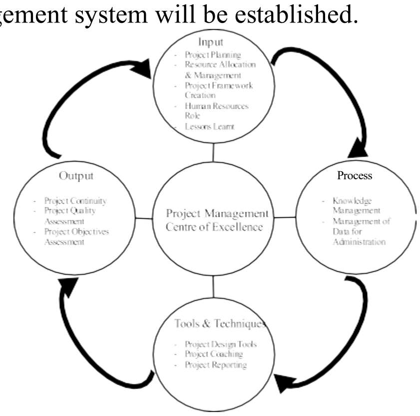


Figure 1: Project Management Centre of Excellence Framework

The first concept is the development of 'input' aspects, especially project planning, Resource Allocation and management, Project framework creation, and assessment of human resource roles. The planning, along with a clear vision, enhances the development of the project based on various inputs, such as resources and management of practices, which would ensure the success of the project. The 'process' aspects that should be considered by PMOCoE include knowledge management and management of data for administrative purposes. These issues require the concept of competency, which aims at fulfilling the process of prioritizing various aspects. In addition, 'tools and techniques' that should be integrated for PMOCoE include project design tools, reporting, and coaching. Management of workers is a major concern; therefore, a need arises to identify activities associated with training, coaching, and development, which reduce the challenges of incompetency. 'Output' concerns include Project Continuity, Project Quality Assessment, and Project Objectives Assessment. This project phase mainly addresses the supportive structures that would be acceptable for the project and its success.

3.6 Best Practices of PMO

A Project Management Office (PMO) is established in order to make sure that the management and support of projects in the organization are efficiently and successfully achieved. As [26] noted, organizations that have PMOs reportedly have more projects completed on time, within budget, and meeting the goals and objectives. With that said, the best practices of the PMO are deemed to be correlated with its functions which include portfolio management, talent management, organizational change management, knowledge management and strategic planning, among others. It is consequently identified that successful PMOs observe certain practices that relate to sufficient discipline and rigor in the processes they undertake in order to ensure the achievement of required reporting accuracy. Moreover, they provide support to the implementation of innovation and development, which in turn, helps the organization accomplish its strategic goals. Another best practice of PMOs is the support they give to good project governance, which includes the need to guarantee the usefulness, relevance, accuracy, and completeness of information in the reports and to provide interpretative and predictive evaluations to the senior management to support the decision-making process involved in portfolio management [27]. Furthermore, it has been identified that scope management is one of the most important best practices that PMOs perform, considering that poor scope management or the lack of it would lead to project failure. To achieve project success, it is purported that PMOs make sure that they have clear and accurate documentation of the project scope, which is necessary in setting the project goals and objectives as well as in establishing project management measures, budgets, and plans. If scope management is not taken into careful regard, adverse consequences can be encountered, such as cost overruns, project delays, poor project quality, and unmet customer needs [28].

According to [29], there are three other best practices that may be gained from effective PMOs. These include the following:

- *Reduce Business Risk*: A flexible, end-to-end project management process is formed by PMOs in order to reduce business risk as it helps balance rigor with the overhead. At the same time, it involves proper planning, management, monitoring, and reporting of all the key activities in the project.
- *Optimize Resources:* The PMOs optimize resources "by expanding PMO oversight to include business and IT projects, and projects sources externally." Integrating the project management discipline into the culture is equally essential as it paves the way for freeing up resources necessary to focus on program management.
- *Contribute to Business Growth:* PMOs expand the PMO influence by focusing on the benefits organizations yield from establishing one, therefore contributing to the growth of the business. It also centers on the strategic issues that the senior management should tackle in order to improve governance in the organization.

All these aforementioned practices are the most significant aspects that PMOs have to consider in order to improve their performance. Moreover, they are also necessary in meeting the goals and objectives and in accomplishing project success in general.

4. Data Collection

As mentioned in the Research Methodology section, to investigate the deployment and effectiveness of PMOCoE, a qualitative research approach was adopted, conducting in-depth interviews with key personnel at prominent projectized entities in the UAE, namely, the Roads and Transport Authority (RTA) in Dubai, and Dubai Police.

4.1 Road and Transportation Authority (RTA) -Project Management Office

RTA Background – The Roads and Transport Authority, commonly known as RTA, is the major independent government roads & transportation authority in Dubai, United Arab Emirates. It was founded in 2005 and is responsible for planning and executing transport and traffic projects, along with legislation and strategic plans of transportation in the city. It is a department of the Government of Dubai [30]. The interview with the Project Management Office (PMO) manager in RTA [31] has fostered several learnings and realizations on the functions, roles, and essence of PMO in the organization. The primary functions and roles of PMO include monitoring the project's annual plans, progress, and updates. The deliverables and milestones are also validated. Indeed, one can see how essential PMO managers are in organizations, especially in initiating projects and ensuring that a project pushes through accomplishing its goals and objectives. It was also learned that for PMO managers to perform their functions, the organization should invest in advanced and innovative tools and systems, as exemplified by the centralized system established and implemented called OPMS. Other than these, certain tools and systems are also required to be used to support project knowledge and organization knowledge capability, other than the Project Management Information System. The efforts made towards project knowledge support also underline the importance of knowledge among the members of the organization and the project managers. By knowledge sharing, the organization can gain insights and experiences that could improve the value of the business. Therefore, knowledge sharing and innovation should be developed or enhanced and employed in organizations like RTA in order to add or gain value and benefits. However, PMO managers and the rest of the members of the project team should be committed to sharing knowledge while also performing their respective jobs. The PMO has a critical role in engaging the rest of the team in knowledge sharing and usage of advanced tools.

4.2 Road and Transportation Authority (RTA) – Public Transport Agency

An interview with Mr. Bahrozyan, the CEO of the Public Transport Agency, has been conducted [32]. In general, the interview consisted of nine question that were collectively aimed at defining the role of PMO within the daily proceedings of the Public Transport Agency. As a matter of fact, Mr. Bahrozyan mentioned that PMO is generally used as the tool for implementing the Project Management Maturity incentives while simultaneously playing a pivotal role in the proceedings of the Program/Project Management Center of Excellence in RTA. It was stated that EPMO is the one practiced within PTA. The CEO went on explaining that whenever a lesson is learned or received, it is EPMO that is responsible for its approval and logging it within the system. Nonetheless, the number of benefits experienced by the organization due to the successful usage of PMO is not limited to the aforementioned ones. Mr. Bahrozyan has found a pleasure in stating that RTA

EPMO ensure complete maintenance and update of OPMS – the major electronic system in RTA that is used for the portfolio and project management along with the processing of educational documentation of the organization. Also, the organization's PMO is used as the tool for reviewing and approving the newly-added PM processes and educational activities. While OPMS is applied as the pivotal project management platform, the knowledge received by the organization from using it is being successfully applied to merging people, processes, and tools into a unified self-sustaining system. Finally, the usage of PMO within the organization can be classified as one of the top drivers of having a constantly updated Lessons Learned template. Thus, the entire life cycle of any project within the company is being updated constantly for future teams to stay away from the mistakes of the past. Mr. Bahrozyan's ending statements were the ones that hailed the integration between PM and Lesson Learned as of the pivotal benefit of organizational management.

4.3 Dubai Police – Quality Department

Dubai Police Background – The Dubai Police Force, commonly referred to as Dubai Police, is the police force of the Emirate of Dubai, United Arab Emirates. Founded on 1 June 1956, the Dubai Police Force now has over 17,500 sworn members and a force of more than 15,000 highly trained and multi-skilled employees who are responsible for policing an area of 4,114 square kilometres and a population of over 3 million people. They come under the jurisdiction of the Ruler of Dubai [33]. The interview conducted with the police in Dubai [34] improved the knowledge and awareness of the major functions of the Project Management Office (PMO). The exercise revealed that the primary role of the PMO is to check and evaluate all the operations of a project. It also identified the levels of the PMO and knowledge management, where both played a coordinative role in enhancing employees' skills and boosting their performance in their assigned duties. The integration of the employees and the management helped the police department achieve its goals and objectives. The PMO's role is to report and document all activities involved in a project. The record allows the people implementing a project to benefit from the experience as errors in the previous undertakings are avoided in the ongoing work. The previous schemes act as a lesson where the mistakes made in one project are recorded to help employee prevent them in their future undertakings. The Dubai police department requires every employee to submit a lesson-learned report at the end of a project. Moreover, the interview [33] enhanced understanding of the importance of the PMO in the improvement of the organization's knowledge and culture through the explanation of the various aspects of the project. The Dubai police shares knowledge and experiences with all other sectors for future advantages. It also uses two functions from the International Organization for Standardization, namely the ISO 21500:2012 and ISO 10006:2018. The former promotes guidance on the project management, while the latter provides guidelines for quality management systems in projects. The PMO sector documents all procedures in every stage of a project and preserves them for future reference. It adopts proper strategies to enhance project knowledge management and the implementation of methodology updates. Consequently, best practices can then be realized, which improves controls through risk management. From the Dubai police, there was no response on the source of knowledge for PM professionals, but the lessons learned proved to have a high-level integration between the PM and the studies. Finally, the PMO was recognized to be in the first stages but was considered to have significant impact on the organization knowledge management.

5. Analysis

The interview with Dubai Police [34] was interactive and engaging. Throughout the interview, the respondent provides comprehensive and consistent answers about the effectiveness of the PMO department in the organization. The interviewer asked questions that left room for discussion and gave the respondent an opportunity to uncover issues that provided more indepth insights about the PMO in Dubai Police. Additionally, the interviewee related solutions to daily life issues. For example, in the third question, the interviewee explained how the experiences of Project "A" benefited Project "B." The interview shows that Dubai Police uses the framework of project management to understand the current needs of its projects and develop strategies to achieve them. RTA interview with Mr. Bahrozyan [32] was of special meaning to the evaluation of PMO perspectives in the domain in the near future. It has proven itself, in accordance with Mr. Bahrozyan, to be an effective managerial tool that helps the organization implement more efficient educational methods and technologies while simultaneously staying away from repeating the mistakes of the past. Further integration of the LL sheet with PMO is the avenue that the organization is most likely to follow in the future. The two interviews have similarities and differences. For example, they were well structured because the interviewer started with simple questions to prepare the respondents and give them time to get comfortable with answering the challenging questions. The first four questions were closed questions, while the rest were mostly open and aimed to seek details about the two organizations. In the two interviews, the depth of their answers is different. For example, the interviewees in Dubai Police and RTA expanded and elaborated their solutions to help the interviewer know more about the PMO in their organizations. The implementation and the effectiveness of the project management office are advantageous to RTA, and Dubai police. The PMO in both companies increased the effectiveness of project management. Additionally, the office provided sufficient data to match project managers with the appropriate projects. According to the interviews, RTA and Dubai police use PMOCoE to document lessons learned from previous projects and integrate them to increase the efficiency of project management. The senior management of Dubai police ensures efficient project management by conducting a thorough assessment of processes, inputs, tools, and techniques. The project leaders of the two successful companies estimate the resources required and do a proper resource allocation for the project. Process assessment demands knowledge management through PMOCoE. PMO ensures proficient management of resources, tools, and techniques of a project. Efficient project management depends on the interaction between the knowledge and the technology departments in a firm. According to the interviewees, Dubai police and RTA communicated well with all their departments regarding PMO. Evidently, the two organizations have a stable project management information system that enables the project managers to share information on a new methodology with the rest of the employees for future use. Additionally, the project managers can easily make critical decisions related to the project. Additionally, PMOCoE guidelines help companies to evaluate the effectiveness of the project in determining organizational goals. It is evident that RTA and Dubai Police used this strategy to weigh on the expectations of the project versus the outcomes. Both companies, (RTA and Dubai Police), have opportunities for improvement. For example, they should build a Center of Excellence (CoE) and combine it with a Project Management Office because the goal of both offices is to complete projects successfully. While the former drives the employees responsible for supporting the project, the latter motivates its structure and governance. The adoption of PMOCoE will ensure the achievement of the desired results in a project. Additionally, organizations need to access and manage knowledge from different departments through the project management office. PMO, as the CoE, supports knowledge development since it combines information from various departments and uses it to create one significant idea.

6. Recommendation

Existing literature provides the development of the Project Management Centre of Excellence and dwells on how it provides guidelines for the successful management of projects through strategic asset management. From the analysis, the PMOCoE framework advances the concepts of competency, quality, vision, strategies, culture, and processes in assessing the success of a project. It is imperative to note that related management tools and techniques can be customized to ensure the reduction of resource wastage while enhancing the maximum output from the available resources. Organizations should, therefore, align their projects to meet the demands of the PMOCoE while addressing the business strategies aiming at the proper management of assets and resources. In order to avoid the risks considered when establishing a PMO-CoE, it’s important for an organization to identify the exact values and the purpose of the establishment. Moreover, it’s important for each individual to contribute to documenting the lessons learned from the previous project, and because some members are drawn by their busy schedules, we recommend upgrading the system with a notification alarm that appears as an urgent task to be completed. Moreover, it is important to benchmark with third parties, whether local or global parties, since it supports the current organization to enhance its own infrastructure and performance and thus to be able to compete with other organizations. It is also important to provide training to internal staff on how PMO-CoE functions and the main activities to be undertaken; therefore, it is recommended that organizations allocate a certain amount of their budget or partnership with certified organizations to educate their staff on a proper matter and demonstrate the expectations clearly.

7. Conclusion

To implement a successful and fully optimized PMOCoE will eventually require the top management attention to make the business call. Therefore, the project managers must be highly encouraged to gather all the important data and justifications to debate and demonstrate to the seniors why the role of PMO-CoE is important and how it can bring benefits to the organization. The tangible and intangible values brought by the PMO-CoE are much appreciated by organizations. Moreover, it's highly important for the senior management to reward the members of PMOCoE on a regular basis to encourage more hard work and dedication to the organization. Also, knowledge sharing within organizations needs to ensure that every project manager is authorized to access the knowledge and to re-use the knowledge to build the new on top, thus delivering the project on time, with the intended quality, cost, and with less error. Moreover, using effective tools is highly encouraged to allow the project managers to update the status of the project and document the lessons learned, and the role and responsibilities of the PMO-CoE to follow up with the provided updates and ensure its correct and thus to share it among the other departments effectively and efficiently.

8. Research Limitations

Sample Size and Generalizability: Due to time constraints, the research is limited to in-depth interviews with personnel from two selected organizations in the UAE (RTA, and Dubai Police). This small, localized sample may not be representative of all organizations using PMOCoE, which limits the generalizability of the findings to other contexts or regions. Subjectivity and Bias: Given that the data collection relies on interviews, there is an inherent subjectivity in the responses. Participants' views may be influenced by their personal experiences, biases, or the desire to portray the organization in a positive light. Lack of Comparative Analysis: The research focuses on entities that have successfully implemented PMOCoE, without comparing these cases to ones where PMOCoE may not have been as successful or not used at all. This could lead to an incomplete understanding of the factors that contribute to the success or failure of PMOCoE implementations.

9. Future Research Recommendations

Researchers may want to consider expanding the scope of their study in future research efforts to encompass a more diverse range of geographical locations and cultural contexts for a more comprehensive understanding of the topic. It is also recommended that future research should capture a wider range of industries. It is also recommended to conduct studies over a longer period to understand the sustainability and long-term impact of PMOCoE on organizational performance. A final recommendation for future research is to compare organizations with and without PMOCoE implementations to identify the specific impact of PMOCoE on project management and organizational learning.

References

[1] Van der Linde, J. and Steyn, H., 2016. The effect of a project management office on project and organisational performance: A case study. South African Journal of Industrial Engineering, 27(1), pp.151-161.

[2] Wyrozębski, P., PMO in Focus of Management Sciences.

[3] Salameh, H., 2014. A framework to establish a project management office. European Journal of Business and Management, 6(9), pp.19-26

[4] Görög, M. and PMI-RMP, P.S., 2013, August. Strategic-Oriented Implementation of Projects. Project Management Institute.

[5] The University of Chicago, 2023 https://pmo.uchicago.edu/project-definition-and-initiation/ .

[6] Peter Landau, Apr 27, 2023 available on https://www.projectmanager.com/blog/10-project- management-knowledge-areas

[7] Bakkah, 2023, available on: https://bakkah.com/knowledge-center/center-excellence-coe-pmo

[8] Dennis Bolles (2002), Library of Congress Cataloging – in – Publications Data Building Project Management Centers of Excellence, United Sates of America, New York, Available at:https://books.google.ae/books?hl=ar&lr=&id=gyqUwbfl6QC&oi=fnd&pg=PR11&dq=project+AND+management+AND+Officer+AND+center+of+excellence+establishment&ots=zXVGl1P0G&sig=yZN_115x0zVN47XymfZa7M0YXJM&redir_esc=y#v=onepage&q&f=false Accessed at: 9-Jul-2019

[9] Bolles, D. (2004). Project Management Centres of Excellence: a guide to getting started. Paper presented at PMI Global Congress 2004-EMEA, Prague, Czech Republic. Newtown Square, PA: Project Management Institute. Available at: https://www.pmi.org/learning/library/pm-centersexcellence-getting-started-guide-8410 Accessed at: 7-Jul-2019

[10] Carl A. Singer (2007), Building Your Project Management Center of Excellence, Available at: http://www.processmakesperfect.net/BuildingPM.pdf Accessed at: 7-Jul-2019

[11] Luara Barnard (2018), What is a project management center of excellence and does it belong in the PMO?, PMO Strategies, Available at: https://pmostrategies.com/project-management- centerexcellence-belong-pmo/

[12] Zamir, Z. (2019) 'The Impact of Knowledge Capture and Knowledge Sharing on Learning, Adapting, Job Satisfaction and Staying Interaction: A Study of the Banking Industry in Bangladesh', International Journal of Entreprenurial Knowledge, 7(1), pp. 46-64.

[13] Maria Adenfelt & Katarina Lagerstrom (August 2006), Knowledge Development and Sharing in Multinational corporations: The case of a centre

of excellence and a transactional team, V.15, Issue 4, Pages 381-400.

[14] Hobbs, B, & Aubry, M 2007, 'A multi-phase research program investigating project management offices (PMOs): The results of phase 1'. Project Management Journal. 38(1), 74–86.
[15] Müller, R 2009, Project Governance. Aldershot, UK: Gower Publishing.
[16] Crawford, JK 2010, The Strategic Project Office: A Guide to Improving Organizational Performance, Second Edition. Boca Raton, FL: CRC Press.
[17] Garfein, SJ 2005, 'Strategic portfolio management: A smart, realistic and relatively fast way to gain sustainable competitive advantage', Proceedings of the PMI Global Congress (North America), Toronto, Canada.
[18] Maake, H, & Potgieter, A 2016, 'September. Knowledge sharing as means of enhancing intellectual capital within the project environment'. In European Conference on Knowledge Management (p. 543). Academic Conferences International Limited.
[19] Eskom (2014). Transfer of responsibilities between project managers. Johannesburg: Eskom.
[20] Eskom (2014). Transfer of responsibilities between project managers. Johannesburg: Eskom.
[21] Monteiroa, A, Santosa, V, & Varajãoc, J 2016, 'Project management office models – A review'. Procedia Computer Science, 100, pp.1085-1094
[22] Project Management Institute 2013, Pulse of the profession TM: PMO frameworks. Newtown Square, PA: Project Management Institute
[23] Caniëls, M.C.J. & Bakens, R.J.J.M. (2012). The effects of Project Management Information Systems on decision making in a multi project environment. International Journal of Project Management 30, pp. 162-175.
[24] Kahura, M.N. (2013). The Role of Project Management Information Systems towards the Success of a Project: The Case of Construction Projects in Nairobi Kenya. International Journal of Academic Research in Business and Social Sciences 3(9), pp. 104- 116.
[25] Karim, A.J. (2011). Project Management Information Systems (PMIS) Factors: an Empirical Study of Their Impact on Project Management Decision Making (PMDM) Performance. Research Journal of Economics, Business and ICT 2, pp. 22-27.
[26] Salameh, H. (2014). A Framework to Establish a Project Management Office. European Journal of Business and Management 6(9), pp. 19-26.
[27] Too, E. & Weaver, P. (2014). The management of project management: A conceptual framework for project governance. International Journal of Project Management 32(8), pp. 1382-1394.
[28] Eby, K. (2017). From the Experts: All About Project Management Offices and PMO Best Practices. Available at: <https://www.smartsheet.com/experts-all- about-project-management-offices-and-pmobest-practices>, Accessed at: 9-Jul-2019
[29] Oracle (2009). Project Management Office Best Practices: A step-by-step plan to build and improve your PMO. Available at:<http://www.oracle.com/oms/eppm/project-magament- pmo-wp302976.pdf>, Accessed at: 10-Jul-2010.
[30] Roads and Transport Authority (Dubai). (2023, October 9). In Wikipedia. https://en.wikipedia.org/wiki/Roads_and_Transport_Au thority_(Dubai)
[31] RTA (2021). Interview with Road and Transportation Authority (RTA) – Project Management Office. Dubai, UAE.
[32] Bahrozyan (2021). Interview with Mr. Bahrozyan from RTA, The Public Transport Agency. Dubai, UAE.
[33] Dubai Police Force. (2023, November 21). In Wikipedia.https://en.wikipedia.org/wiki/Dubai_Police_Force
[34] Dubai Police Force (2021) Interview with Dubai Police – Quality Office, Dubai, UAE.

A deep learning approach for Moroccan dates types recognition.

AOULALAY Ayoub*, abderrahim El MHOUTI†, Abdellah El zaar‡, Rachida Assawab and Mohammed MASSAR§

Abdelmalek Essaadi University, Tetouan, Morocco

Abstract

With the growth of the date fruit market, it becomes necessary to use artificial intelligence techniques to recognize dates categories. In this work, we use computer vision approaches to classify date fruits produced in Morocco according to their type. To do this, a dataset is constructed. This dataset contains images of seven types of date fruit that are the most popular in Morocco. Two computer vision approaches are compared and evaluated in terms of performance. Both approaches are based on transfer learning of a convolutional neural network CNN. These two approaches are standard feature extraction where deep features are used to train a machine learning classifier, and fine-tuning where we fit the pre-trained model to our dataset. The approaches used in this work achieve very good performance on our dataset, with a classification precision of 97 %.

Key Words: dates fruit classification, transfer learning, fine tuning , features extraction, deep convolutional neural network .

1 Introduction

In recent years, deep learning has emerged as a breakthrough in artificial intelligence techniques that has overwhelmed the field of pattern recognition and computer vision research by delivering cutting-edge results. In the agricultural field, this technology is widely used to improve productivity, including the classification of agricultural products from images[1], [2]. Recently, Convolutional Neural Network (CNN) has emerged as a powerful tool for image processing tasks, achieving remarkable results, making it the leading technique for vision applications. In the present study, we will use recent methods based on CNN applied to the classification and recognition of Moroccan dates according to their type. Indeed, Morocco is one of the largest producers of date fruit. According to statistics from the Ministry of Agriculture, Morocco is the seventh-largest area of date palms. The regions that produce the most dates in Morocco are mainly located in the eastern south, in the regions of Erachidia Ouarzazate, Tata Zagora, and Figuig. The date palm chain covers about 59,600 hectares (1 % of the arable land nationwide). Daraa-Tafilalt region is responsible for producing an average of 100 000 Tons of dates yearly.

Morocco produces a variety of dates fruit types with 455 types, but across Morocco, only a few are well known. These types range from the best quality to less quality. Al Majhoul is the date fruit that is considered to have the best quality. And this is what makes it the most expensive date fruit. Erachidia and Ouarzazate are responsible for 90 % of the production of this type. However, the production of this type of dates fruit in Morocco is low (0.3 %) compared to other countries. The most produced dates fruit in Morocco are Bofgous and Jihle with 12.2 % and 11.9 %. The most consumed type of dates fruit is the Tathmout, due to it's high-quality while being very cheap. Bouzkri is widely used in pastries when it is dry and unlike the other types, Bouzkri reaches its best quality when it ages a little. Aziza is rare in Morocco, and it represents a percentage of 0.2 % of the date production in Morocco, it is located in the region of Fuigig mainly. Although it is imported and not cultivated in Morocco, Sokari is very popular in Morocco and has earned its place on Moroccan tables alongside other types of dates. In this work, we aim to automatically recognize the seven types of dates that are produced in morocco.

Different types of dates are produced in Morocco with high quality types like almajhoul which is well known internationally and other local types only known in their local places. Each type of dates has a price that defines its quality and rarity, Almajhoul is the most expensive because of its high quality. The other types each have their own price. The type of dates is one of the factors that define their marketable price. And thus, it is important to know the type of dates before buying it.

The main problem in distinguishing the type of dates is that dates can be very similar, and are often confused with each other. Distinguishing one type from another requires a certain level of expertise. Figure 1 shows the seven types discussed in this work. some dates fruit share the same visual properties. For example, the texture of Majhoul and Jihl is similar. Bouzkri and Bofgous also have a similar texture. And for the case of shape, Tathmout and Almajhoul are similar. And in terms of color, Almajhoul, Bofgous, and Sokari are similar. another factor that makes dates fruit classification very challenging is that dates

*Email: aoulalayayoub@gmail.com

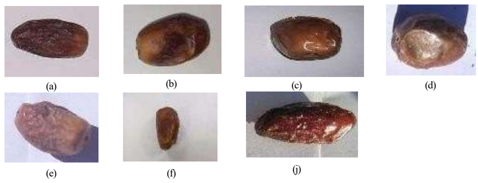


Figure 1: **images of the seven Moroccan cultivars of dates used in this work (a): al majhoul ,(b): aziza, (c): bofgous, (d): bouzkri , (e): jihl, (f): sokari , (j): tathmout.**

Figure 2: **difference between two dates belonging to bofgous cultivar.**

fruit from the same class can be very different as in the example in figure 2. Here we can see two images of the type Bofgous that are visually very different. This makes classifying dates fruit into types a very difficult process.

In this work we use computer vision to automatically classify these seven types of dates fruit. Our approach is based on transfer learning of convolutional neural network CNN.

In recent years CNN has shown a great performance in a variety of visual recognition tasks, like face recognition [3], handwritten character recognition [4], medical image classification [5], Vehicles identification [6] and deepfake detection[7] . CNN architectures like AlexNet[8] VGG [9], ResNet [10] have reached a precision that is close to perfection in the ILSVRC challenge.

It has been shown that increasing the depth of the CNN network leads to increased performance [8] [9].And as shown in many works The CNN relies on training data to build its image description. Each CNN filter is adjusted during training to extract a discriminative feature type for the current task. This makes the CNN a data-intensive model, since the essential part of the CNN, as well as its depth, is the data. To train a deep CNN, a large amount of data is needed, as adding more layers means having more trainable parameters. with limited data, a state-of-the-art architecture will suffer from overfitting [11]. The need for a large amount of data is the main drawback of CNN. In many tasks, data collection can be expensive and is not available to everyone. In our case, we have 700 images split between training and testing. This is too little compared to ImageNet (14 million images), and it is not enough to train a state-of-the-art network. In this work, to overcome this problem, we use transfer learning [12] [13]. It is the process of using knowledge from one model to improve the other. In tasks with limited data, it is useful to use a model that has performed well on a task closer to the task at hand. In object recognition tasks the models used to transfer knowledge are trained on ImageNet. The weights of these models are used to extract the saliency representations from the limited data of the task at hand. This can be done in two ways. The first is to extract off the shelf features that will be used to train a new classifier. The second is to adapt the source network to our task, replacing the last layers with new ones and fine-tuning them to our task.

In this work, we compare the two transfer learning strategies for the classification of Moroccan dates fruit. With transfer learning, we were able to achieve a very good performance with a classification precision of 97 % using both approaches.

This paper is organized as follows: firstly, the document sets out the related work. Then the paper presents the materials and methods. In this section, we describe the data, and the theory of the approaches used for the automatic recognition of Moroc-can dates fruits. The paper then presents and discusses the results obtained. Finally, the article ends by discussing future work.

2 Related works

The categorization and classification of dates fruit have attracted the attention of many researchers in recent years. Especially in countries that produce dates. The categorization of dates is very important because it is a factor in the marketing of dates. thus, many works have proposed the use of chemical or organic characteristics of dates to categorize them.

[14] suggested the use of chemical characteristics as well as mor-phological and textural characteristics to classify date fruit. They studied the most important characteristics to classify 20 types of dates grown in the United Arab Emirates (UAE). Their results show that textural characteristics are not essential for date fruit classification.

Although classifying date fruit is a fine-grained classification challenge, distinguishing between them is mostly a visual task. It only takes a glance for experts to know which type of date fruit it is. For this reason, much work has focused on date fruit classification using visual information alone.

[15] used machine learning to automatically classify dates fruit into categories. Their study was conducted on dates produced in the Middle East. To do so, they used three image description methods, these methods are based on extracting color, shape, and morphology features. The extracted features are then fed into logistic regression and artificial neural network models for classification. Their study shows that the best results were obtained by combining logistic regression and neural network, with an accuracy of 92.8 %.

[16] proposed a new approach for the detection and classification of date fruit. Their approach is based on the extraction of color, shape, and size features. These features will be used to train a neural network classifier. They obtained an accuracy of 97.2 % on the classification of three types of dates that grow in Pakistan.

[17] used shape, color, size, and texture feature extraction

for the classification of seven date fruit produced in the UAE. They compared three classification techniques: KNN, LDA, and neural network. They found that the best results were obtained using the neural network with 98.6 % accuracy.

[18] proposed a pipeline for the classification of four types of date fruit grown in the Arabian Gulf countries. Their system is based on combining texture, size, and shape features to describe date fruits images. In their work, texture features are extracted from three color components. for this they compared several texture descriptors and achieved the best performance using a combination of the three feature extraction methods, with 98.1 % accuracy.

Transfer learning was utilized by [19], where they used a model based on MobileNetV2 to classify dates from Saudi Arabia. To train their model they used a database comprising images captured under stable lighting conditions, obtained using a 48 cm, 55 W 5500 K Dimmable LED Ring Light outdoor mobile tripod. This device was also employed to remove shadows. The camera-to-date fruit distance was fixed for all images, and the images were captured in daylight to avoid texture changes. Consequently, the database contains high-quality images, and their model achieves an accuracy of 99 %. The model is an excellent tool for classifying date fruits in industrial applications where such conditions and devices are available. However, such conditions are not always feasible. In our study, we use images with various lighting conditions, varying image qualities, and the presence or absence of shadows, as shown in the figure. We aim to collect images in different scenarios and qualities to present a challenge to the classification model. We believe this will lead to a model that generalizes better in real-world scenarios where stable conditions are not always available.

The works described above have focused on the classification of dates from the East and the Middle East. They achieved good results, sometimes close to optimal classification accuracy, using only handmade descriptors. This indicates that in this domain, perfect hand-designed features have been discovered. Moroccan dates have some similarities with the middle east date fruit types and some differences. Many types that are grown in Morocco can be very similar to each other. Besides, a type of dates fruit may have different subtypes like Bofgous, and Tathmout with black and red Tathmout. This makes the classification of Moroccan date types a very difficult task. In our work, we investigate the power of deep learned features for Moroccan dates fruit categorization. For this purpose, we use transfer learning of a deep convolutional neural network.

3 Data

Our dataset was collected manually by taking photos of seven types of Moroccan dates. The photos of Al majhoul, Jihl, bofgous, Tathmout, Sokari Aziza, and Bouzkri are taken using cameras and smartphones. Each image in the dataset represents a single dates fruit on a white background. Our dataset contains seven classes with 100 images per class. Each class has a

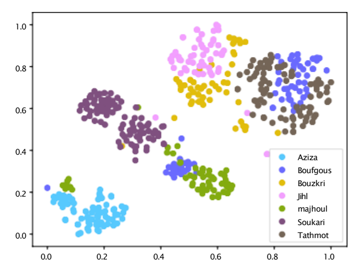


Figure 3: **T-distributed stochastic neighbor embedding(t-sne) visualization of the dataset using ResNet50 deep features. different colors represent the different cultivars. Samples that are similar are close to each other in the graph**.

large intra-class variance because each class contains dates with different harvest dates. Each date cultivar appears different as it ages. And the appearance as a function of harvest date differs from one cultivar to another. For example, Bozkri and jihl harden over time, and other cultivars may look the same with just a small difference. Dates are purchased in local markets, so we are not sure of the dates and places of harvest, nor of their post-harvest history. The dataset we collected contains images of date cultivars with different picking dates. The images are taken under different conditions (different lighting, different image quality, different cameras) in the training and evaluation sets. The reason we use images with different conditions is to impose some challenge on our model. This will make the model robust to changes in the input and image conditions, as we want it to be usable in real-world applications, such as smartphones. To ensure that the model makes the decision based only on the discriminating features of the date presented in the image, and is not sensitive to the state of the input image with different lightning and conditions. We use the model visualization tool, gradcam. Grad-cam is a neural network visualization technique introduced by [20]. As the name implies, Grad-cam uses the gradient of the classification prediction against the feature maps of the last convolutional layers to show the parts of the input image that lead to the decision made by the CNN, and that are most important for distinguishing one class from another.

We divide the dataset into 80 % for training and 20 % for testing. To analyze the distribution of the data, we use t-distributed stochastic neighbor embedding T-SNE visualization. We use as input to T-SNE, the features extracted by the ResNet50 model pre-trained on ImageNet. The results are presented in figure4. We can see that the most similar date fruit types are clustered next to each other.

4 Convolutional Neural Network and Transfer Learning

In recent years, CNN has become the state of the art in various image classification and pattern recognition tasks. CNN is designed to represent the image content by extracting a hierarchical representation that starts at a low level by extracting low-level features such as edges and color blobs, up to high-level features that are class-specific. Since 2012, CNN has outperformed other machine learning techniques such as random forest [21] and SVM [22] in a variety of image classification challenges like the ImageNet Large Scale Visual Recognition Challenge ILSVRC. The CNN has benefited from advances in computing resources and technology, with advanced systems such as GPUs for deep learning applications. These systems are able to com-pute operations in tensors, which are the building blocks of a neural network, more quickly. The main difference between CNN and other machine learning models is that the image representation part is learned from with convolutional layers in CNN, unlike other methods where the image representation has to be extracted using hand-made descriptors. It is believed that a deep network should be able to discover by hidden layer the effective representation of a given task [11]. As shown in many works the parameter that led to better accuracy is the depth of the networks. A deeper CNN network will succeed in extracting features that are highly related to the task and that are discriminative enough to separate classes. The only problem is that a deeper network has a large number of parameters, and they need a large training set to be trained. In our work, we only have 100 images per class, which is not sufficient to train a deep CNN.

One solution for this problem is transfer learning. Transfer learning [12], [13], [23], is the process of using the knowledge of a source network to improve the performance of another one. This is based on the hypothesis that a network that has been trained on a dataset to perform a task has built some knowledge that allows him to perform well on another task close to the one it's trained on. Transfer learning is defined as an operation of using a model to improv another.

a domain D is composed of a features space χ and a marginal distribution P(X), D=$\{\chi$, P(X)$\}$where X is the set of instances X= $\{$x1, x2,. . . . . . .x$\}$n , and a learning task is composed of a label space γ, and a prediction function f. T=$\{\gamma$,f$\}$. note that the prediction function is learned from the data. Transfer learning involves two elements the first one is the source model that has a source domain Ds and a source task Tt. The second element is the target model that has the target domain Dt and the target task Tt. Transfer learning can be defined as the process of using the knowledge gained from training the source model trained on the source domain Ds and the source task Ts, to improve the prediction function of the target model Ft on the target domain Dt. The condition of transfer learning is that both source and target domains have some similarities. For example, a professional swimmer will need less training to be a Water polo player, than a regular person. In transfer learning, the learning model from one task is reused as a starting point for a second task [11] In our case, Moroccan dates fruit classification, the source model is a state of art network, trained on the ImageNet. The source and the target domain have some similarities as they are all composed of images. The tasks also have similarities as they focus on the classification of images. ImageNet is also very close to our dataset, as it contains a variety of fine-grained classes like different breeds of dogs for example. It was shown that ImageNet is a powerful source for transfer learning tasks. And many schools of thought have been raised on the reason why ImageNet performs well on transfer learning tasks [24]. The overall structure of Image Net is what makes it suitable for the transfer learning application, as it has a large number of images divided into 1000 classes. These classes vary from coarse-grained to fine-grained classes. The diversity of classes as well as the number of images and classes makes ImageNet a perfect data source for transfer learning.

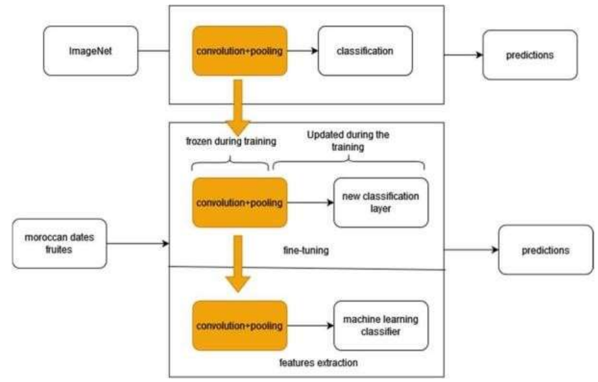


Figure 4: **Representation of transfer learning, the convolutional part of a model trained on ImageNet is transferred to our model which will be trained on our dataset. For fine-tuning, some of the first layers are frozen and the others are updated during training. For feature extraction, the features extracted using the transferred part are used to train a machine learning classifier.**

In this work, we investigated two transfer learning strategies. The first one is by using off-the-shelf features of the source model to extract features from our dataset. For this work, we utilized the convolutional parts of several models (VGG19, VGG16, ResNet50, and Inception V2) as feature extractors. Essentially, convolutional filters with ImageNet pretrained weights were employed as a filter bank to extract abstract features. These features are then used as an input of a machine learning classifier. Image representation extracted with a pre-trained network has shown a great performance in fine-grained image classification challenges [23].

The second strategy is fine-tuning. Fine-tuning consists of using the weights of the pre-trained model as a starting point for the target model. CNN is a hierarchical features extractor, as it extracts different levels of features. We can categorize these features into two types, general and specific [25]. general features are the ones that can be used in all image classification tasks. Specific features represent the characteristics that are related to the task at hand. General features as shown in many works are extracted in the first and the mid-layers of the network while the specific features are extracted in deeper layers. The basic idea of fine tuning is using the layers of the source network that it's believed extract features related to both tasks (ILSVRC and Moroccan dates fruits classification) meaning that these features are general. And we replace the task specific layers with new ones. For this we use a state of art network trained on ImageNet as a source network. The first layers that extract general features are frozen, (i.e. they are not updated during the training) and we replace the fully connected layers with a new randomly initialized one, and we train the whole network on our dataset.

For this approach, we compared four models, similar to the first strategy. The fully connected layer of the base model has been deleted and replaced with a new one. The new fully connected layer consists of four layers: three fully connected layers activated with the ReLU activation function, with the number of filters being (128, 64, and 32). The fourth layer is a SoftMax layer that makes the final decision. The network is trained for 50 epochs with the RMSprop optimizer.

5 Results and Discussion

In this section, we present the results and analysis of the two transfer learning strategies used in this work on Moroccan date fruit categories datasets. To evaluate the classification performance, we divided our database into 80 % training images and 20 % test images. We then extracted the precision and the accuracy of the five classes along with other evaluation measures.We used the [8:2] split due to the small size of the data set. We can confirm that human-selected hyperparameter bias did not contribute to our results because in our experiment we use a standard set of hyperparameters. This set was not changed at any point in the experiment.Before feeding the images to the model we preprocess them by subtracting the mean RGB pixel intensity from the ImageNet dataset.

To evaluate the performance, we use the classification report tool of scikit-learn library in python, it provides us with three metrics.

- the precision is the ratio of correct forecasts to supposedly positive ones.

$$\text{Precision} = \frac{TP}{TP + FP} \tag{1}$$

- the recall determines the detection capability of all images that represent the target class in a dataset.

$$\text{Recall} = \frac{TP}{TP + FN} \tag{2}$$

- the F1-score allows to establish the balance between precision and recall.

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{3}$$

5.1 Features extraction results

Feature extraction involves using the pre-trained model to extract features from the dataset. These features will be used to train a machine learning classifier. We have tested several pre-trained source models, which are trained on ImageNet. And we compare the performance of a variety of machine learning classification algorithms. The models tested are VGG16, VGG19, ResNet50, and Inception V2. We compared four classifiers: KNN, SVM, logistic regression, and Random Forest. The results are presented in Table 1.

Table 1: Results of features extraction approach using different source models and different classifiers.

	SVM	KNN	LR	Random Forest
VGG19	97%	83%	96%	88%
VGG16	95%	94%	95%	91%
ResNet50	95%	92%	95%	91%
Inception V2	87%	82%	88%	82%

The best precision is obtained with VGG19 as feature extractor and SVM as a classifier. on the other hand The lowest precision was obtained with Inception V2 as features extractor and KNN and random forest as a classifier. Random forest and KNN have the lowest prediction precision compared to SVM and logistic regression which perform better. And for feature extraction, Inception V2 performs worse than other source models.

5.2 Fine-tuning results

Fine-tuning involves replacing the task-specific layers of a pre-trained CNN with new layers and training the entire network for the target task. As we did in the feature extraction approach, we compare four architectures, VGG16, VGG19, Inception V3, and ResNet. We replace the set of fully connected layers with a new one whose weights are chosen stochastically. The new classifier is composed of four layers whose number of filters is respectively (128, 64, 32, and 7). The last layer is a Soft-Max layer, while the others are activated with Relu[8]. Dropout [26] is used to reduce over-fitting and the whole network is trained for 50 epochs with stochastic gradient descent as an optimizer. The learning rate is reduced during training.

the results are reported in fig5 the best precision is obtained with VGG16 as the source model with 97 %, and the lowest when using InceptionV3 with 93 %. The fine-tuning approach shows the same performance as the feature extraction approach, achieving the same accuracy of 97 %. The only difference is in which model has the best accuracy, in the fine-tuning approach, VGG16 is the best performing model, while it is better to use VGG19 in the features extraction approach. The accuracy and loss curves in figure7 show that the model does not suffer from overfitting despite having limited data. In the first few epochs, the validation accuracy is better than the training accuracy, but after a few epochs, they start to be equal.

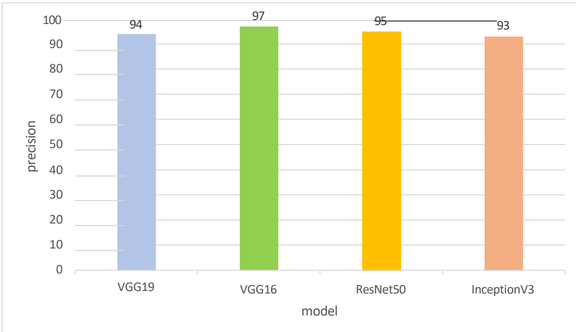


Figure 5: Precision of fine-tuning different Convolutional neural network models in %. VGG16 gives the best performance and InceptionV3 is the worst performing.

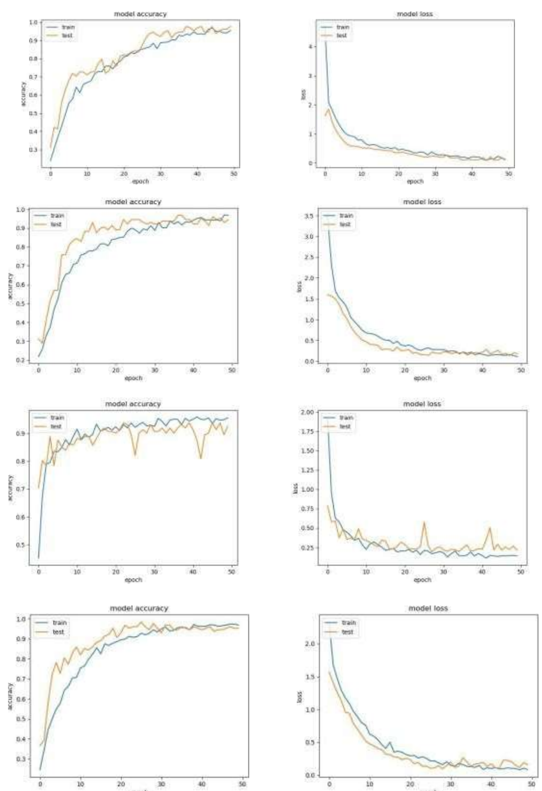


Figure 6: accuracy and loss curves of fine-tuning different models the first column belongs to VGG16 the second for VGG19 the third for INCEPTION and the last one for RESNET50

Transferring the image representation of a network trained on ImageNet helps in date fruit classification. Keep in mind that date fruit classification is a fine-grained challenge, and one of the key elements that make ImageNet ideal for transfer learning tasks is that it has a large number of fine-grained classes. A network trained on ImageNet always learns to distinguish subclasses, and filters learned from ImageNet has specialized in this way. Fine-tuning this filter towards dates fruit is more efficient and faster than learning this representation from scratch.

5.3 Class precision

Table 2: the precision, the recall and f1 score off each class using features extraction of VGG19 and SVM classifier in %. Bouzkri , soukari , jihl and aziza have the perfect precision and recall.

	Al majhoul	tathmout	bouzkri	boufgous	soukari	jihl	aziza
Precision	100	91	100	86	100	100	100
Recall	85	100	100	90	100	100	100
F1-score	92	95	100	88	100	100	100

Table 3: the precision, the recall and f1 score off each class of fine tuning VGG16 in % . Bouzkri , soukari , jihl and aziza have the perfect precision and recall.

	Al majhoul	tathmout	bouzkri	boufgous	soukari	jihl	aziza
Precision	100	95	100	86	100	100	100
Recall	90	95	100	95	100	100	100
F1-score	95	95	100	90	100	100	100

We extract classification precision off each of the Moroccan date fruit types used in this work. for this purpose, we use the best-performing model in the features extraction approach and fine-tuning approach. For features extraction approach, we use VGG19 and SVM as it gives the best results. And for fine-tuning we use VGG16. We extract precision, recall, and f1-score and show them in the table2 and table3. table2 shows the class-performance off the features extraction approach .it shows that only two classes don't have perfect precision. These classes are Tathmout and Bouzkri which obtain respectively 91 % and 86 %. The other classes are classified perfectly. For the recall, only two classes are not detected perfectly. These are al Majhoul and Bofgous with 85 % and 90 % respectively. Table3 represent the fine-tuning approach. the precision of the classes is perfect except for the precision off Tathmout and Boufgous that reaches 95 % and 86 % respectively. Three classes have a recall lower than 100 %. these are al majhoul with 90 %, tathmout 95 % and boufgous 95 %. Bofgous and Tathmout represent a challenge for our approach, as these are the classes that our model has difficulty recognizing using both methods. Bofgous have high intra-class variance because there are a variety of Bofgous types and thus there will be subclasses within the same class. For Tathmouts, we can explain the model's performance by the fact that Tathmouts vary visually with their level of maturity as well as their age. aging only few days can make a huge difference in the appearance of the Tathmout.

5.4 The level of features

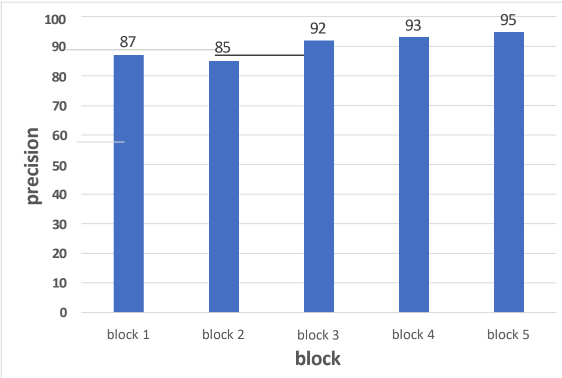


Figure 7: the precision of features extracted from different blocks of VGG16 network in %. Extracting features from deeper layers (block 5) gives the best accuracy, compared to earlier layer (block1 and block2).

It is believed that the general image representation suitable for both the source and target task resides in the intermediate layers[11]. We tested this assumption by testing the power of the extracted features at different levels of CNN. For this, we use different blocks of the VGG16 network pre-trained on ImageNet. The convolutional part of VGG16 is divided into five blocks as shown in the figure. Each block ends with a maximum pooling layer. we extract features from different levels of CNN by using the output off the max polling layer at the end of each block. these features will be fed into an SVM model for classification. The model's blocks indicate the level of the extracted features. From low-level features extracted in the first block to high-level task-specific features extracted in block5. figure8 shows the classification accuracy corresponding to each block. The best results are obtained by using the features of the last convolutional block with 95 %. At the same time, the accuracy increases with increasing feature levels. This indicates that the optimal representation shared by ImageNet classification and Moroccan date fruit classification is in deep layers close to the classification part. We can say that the similarity between the two tasks is huge. these also shows the power of ImageNet deep features because even using the low-level representation in block1, we still get a good precision with 87 %.

5.5 model visualization

We use gradCam to visualize the concentration of the last convoluted layer. For this purpose, we used a ResNet50 model trained on our dataset. The results of applying gradCam to images of the seven date cultivars used in this work are shown in Figure8. The model focuses on the surface of the date. This means that date texture and color are the main features that contribute to the final decision.

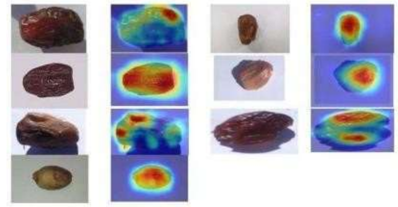


Figure 8: Visualization of the gradient concentration of the last convolution layer of a ResNet50 model using Grad-Cam. The images represent seven date cultivar samples used in this work, the corresponding Grad-cam map is to the right of each sample, and shows that the algorithm focuses on the date surface.

5.6 deep features Vs hand-crafted features

Many works have used hand-created features prior to classification and have performed well on different datasets. The problem with this approach is that we have to use improved image quality in the training and test dataset, either by using a system to obtain images with the same lighting and conditions as in [15], which achieves 92.8 % accuracy. This is a drawback because the ordinary everyday user would not have access to the same conditions. The same is true for [16], which achieved 97.2 % accuracy, but used the same conditions for data collection. They used the same camera and manual preprocessing before feeding images to the classification algorithm. [17] achieved an accuracy of 98.6 %. But a thresholding was performed to segment the date from its background in order to optimally apply the feature extraction. In our work, the only preprocessing we use is image normalization by subtracting the average RGB pixel intensity from the ImageNet dataset. We use images of different conditions and qualities to impose some challenge on the model, and to exploit real-world scenarios. For this, we use deep learning features that we believe are robust to changes in input conditions, and we obtained an accuracy of 97 % in a challenging dataset.

6 Conclusion

Dates are an important product in the countries that produce them, this type of fruit has become in recent years an important aspect of the diet. Indeed, the market for dates has developed rapidly throughout the world. One of the important factors that define the price and quality of dates is their type. In this work, we introduce an approach for automatic recognition of Moroccan dates fruit types based on transfer learning of a convolutional neural network. The first step is the collection of data on which we built a dataset of seven well-known date types in Morocco. Our model is used to classify the images of this dataset. We compare two approaches: feature extraction and fine-tuning. Our results show that both approaches achieve a very good performance with a classification precision of 97 %. We find that in the feature extraction approach, VGG19 combined with SVM performs best, while in the fine-tuning approach, VGG16 is the best performing model.Although the obtained results demonstrated very good performance on data captured in real-world scenarios with varying lighting and viewpoints, the main drawback is the use of a white background. In the real world, date fruits are often captured against various backgrounds. This motivates us to further investigate this challenge in the future by improving the generalization of the recognition model for date fruits captured in more realistic scenarios.

7 Statements and Declarations

We the undersigned declare that this manuscript is original, has not been published before and is not currently being considered for publication elsewhere. We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We understand that the Corresponding Author is the sole contact for the Editorial process. He/she is responsible for communicating with the other authors about progress, submissions of revisions and final approval of proofs Signed by all authors as follows

References

[1] P. T. Q. Anh, D. Q. Thuyet, and Y. Kobayashi, "Image classification of root-trimmed garlic using multi-label and multi-class classification with deep convolutional neural network," *Postharvest Biology and Technology*, vol. 190, p. 111956, 2022.

[2] A. Septiarini, A. Sunyoto, H. Hamdani, A. A. Kasim, F. Utaminingrum, and H. R. Hatta, "Machine vision for the maturity classification of oil palm fresh fruit bunches based on color and texture features," *Scientia Horticulturae*, vol. 286, p. 110245, 2021.

[3] G. Hu, Y. Yang, D. Yi, J. Kittler, W. Christmas, S. Z. Li, and T. Hospedales, "When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition," in *Proceedings of the IEEE international conference on computer vision workshops*, 2015, pp. 142–150.

[4] A. Elzaar, R. Assawab, A. Aoulalay, L. O. Ali, N. Benaya, A. El Mhouti, M. Massar, and A. El Allati, "Amazigh handwriting recognition system—multiple dcnn strategies applied to new wide and challenging database," in *International Conference on Advanced Technologies for Humanity*. Springer, 2021, pp. 241–252.

[5] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–18, 2019.

[6] Z. Dorrani, H. Farsi, and S. Mohammadzadeh, "Edge detection and identification using deep learning to identify vehicles," *Journal of Information Systems and Telecommunication (JIST)*, vol. 3, no. 39, p. 201, 2022.

[7] A. Biswas, D. Bhattacharya, and K. A. Kumar, "Deepfake detection using 3d-xception net with discrete fourier transformation," *Journal of Information Systems and Telecommunication (JIST)*, vol. 3, no. 35, p. 161, 2021.

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, 2012.

[9] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[11] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," in *International conference on machine learning*. PMLR, 2014, pp. 647–655.

[12] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, no. 1, pp. 1–40, 2016.

[13] F. Zhuang, Z. Qi, K. Duan, D. Xi, Y. Zhu, H. Zhu, H. Xiong, and Q. He, "A comprehensive survey on transfer learning," *Proceedings of the IEEE*, vol. 109, no. 1, pp. 43–76, 2020.

[14] A. Kamal-Eldin and S. Ghnimi, "Classification of date fruit (phoenix dactylifera, l.) based on chemometric analysis with multivariate approach," *Journal of Food Measurement and Characterization*, vol. 12, no. 2, pp. 1020–1027, 2018.

[15] M. Koklu, R. Kursun, Y. S. Taspinar, and I. Cinar, "Classification of date fruits into genetic varieties using image analysis," *Mathematical Problems in Engineering*, vol. 2021, 2021.

[16] A. Magsi, J. A. Mahar, S. H. Danwar *et al.*, "Date fruit recognition using feature extraction techniques and deep convolutional neural network," *Indian Journal of Science and Technology*, vol. 12, no. 32, pp. 1–12, 2019.

[17] A. Haidar, H. Dong, and N. Mavridis, "Image-based date fruit classification," in *2012 IV International Congress on Ultra Modern Telecommunications and Control Systems*. IEEE, 2012, pp. 357–363.

[18] G. Muhammad, "Automatic date fruit classification by using local texture descriptors and shape-size features," in *2014 European Modelling Symposium*. IEEE, 2014, pp. 174–179.

[19] K. Albarrak, Y. Gulzar, Y. Hamid, A. Mehmood, and A. B. Soomro, "A deep learning-based model for date fruit classification," *Sustainability*, vol. 14, no. 10, p. 6339, 2022.

[20] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," pp. 618–626, 2017.

[21] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[22] G. Chandrarathne, K. Thanikasalam, and A. Pinidiyaarachchi, "A comprehensive study on deep image classification with small datasets," in *Advances in Electronics Engineering*. Springer, 2020, pp. 93–106.

[23] A. Sharif Razavian, H. Azizpour, J. Sullivan, and S. Carlsson, "Cnn features off-the-shelf: an astounding baseline for recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2014, pp. 806–813.

[24] M. Huh, P. Agrawal, and A. A. Efros, "What makes imagenet good for transfer learning?" *arXiv preprint arXiv:1608.08614*, 2016.

[25] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" *Advances in neural information processing systems*, vol. 27, 2014.

[26] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, 2014.

Docker Container Security Analysis Based On Virtualization Technologies

Shatha A. Baker*, Hesham Hashim Mohammed‡, Omar I. Alsaif §
* ‡ § Northern Technical University, Iraq.

Abstract

The utilization of virtualization technology, particularly Docker containers, has increased significantly in recent years. As Docker provides a lightweight and efficient virtualization environment for software packages, ensuring its security becomes crucial. This paper performs a security analysis of Docker with two perspectives: the security within Docker and its relationship to the security features of the Linux kernel. Resources that are isolated, controlled, and limited are all examined in the Docker internal security. Linux Cgroup is used by Docker to manage computer resources, while Linux Namespace is used to securely isolate running environments. In this paper discussed how to separate resources such as filesystems, networks, devices, and processes, as well as how to isolate inter-process communication. Docker's interactions with Linux kernel security characteristics such as SELinux, AppArmor, Seccomp, and Linux functions were also discussed. These capabilities boost host system security by deploying Docker containers. AppArmor maintains security policies, SELinux offers further permission checks, and the Linux function limits container rights. Network-based assaults are defended against with the aid of Seccomp and the network framework. Additionally, the study makes recommendations for possible enhancements to improve Docker's security. This involves configuring Docker to deactivate specific functions within containers to thwart possible breaches and enhancing interoperability with Linux kernel security mechanisms.
Keywords: Docker Container, Virtual Machine, Security Analysis, Linux Kernel, SELinux, AppArmor, Seccomp.

1 Introduction

In today's world, virtualization has become essential. It provides a way to share resources among several users and is an affordable way to cut down on resource underuse, especially as cloud services become more and more common. Virtual machines (VM) and containers are two widely recognized technologies that facilitate virtualization [1].

Docker is a significant platform utilized by many developers. It is an open-source platform for packaging apps and running them in containers. The engine's objective is to deliver a quick, light environment for running the applications of developer and making deployment easy and efficient. However, ensuring the security of Docker containers is crucial to protect applications and sensitive data from potential threats [2].

Several approaches have been proposed to address Docker container security. Bui [3] offered a comprehensive examination of Docker technology, concentrating on both its aspects of internal securityand the modules of external security that may be integrated with Docker. Bélair et al in [4] introduced a novel taxonomy of container security that focuses on the infrastructure level, distinguishing it from previous works. They propose a classification of defense frameworks based on this taxonomy, specifically highlighting the methods by which the host OS can enhance container security. The authors in [5], extensively described the various attack within Docker technology and proposed effective strategies to mitigate these security risks. Another study [6] introduced a method of dynamic analysis that evaluates the Docker images security by analyzing their activities. The dynamic analysis method has been proven to be a useful addition to static analyses, which are frequently used in security evaluations of Docker images. In [7], The authors investigated denial of service attacks against the architecture of Docker and advocated the usage of a memory limit technique to mitigate against resource exhaustion induced by malicious containers.

In this paper, we are going to highlight the Docker architecture and its key components and analyze the Docker container security. The analysis looks at several topics: Docker's internal security depends on the isolation level that Docker provides for its virtual environments, the way Docker handles Linux kernel's security features like AppArmor and SELinux to

* shathaab@ntu.edu.iq
* ‡ § Computer Systems Department

strengthen the host system, and the challenges which face Docker container.

The paper is set up like follows: The second section provides a brief explanation to the Docker platform. The third section explains the Architecture of Docker. Docker security analysis is introduced in the fourth section. The fifth section focuses on discussing Docker security analysis and what can be done to improve it. The conclusion forms the final section.

2 Docker Overview

The "Docker" was initially established in March 2013 following the introduction of an innovative approach termed containerization, which prompted OS-level virtualization. Docker provides automated deployment capabilities for applications through containers. It adds an additional layer to the container environment, allowing virtualization and execution of applications. Docker is intended to build a lightweight and fast environment for efficient code execution. Additionally, it provides a convenient workflow for testing code before deploying it to production [8].

The container model is often misrepresented. While container technology appears secure due to its ability to contain all dependencies in one package, it does not guarantee overall security. Container platforms, like other cloud platforms, are also vulnerable to various threats from both internal and external sources. Operating secure services within a multi-tenant cloud system poses several challenges in virtual environments. It is widely acknowledged that VMs created through hypervisor-based virtualization methods offer higher levels of security compared to containers, table 1 shows the differences between the two technologies [9].

Table 1: The distinction between Docker containers and virtual machines [9]

Docker container	Virtual Machine
Within a few seconds, boots	It needs a few minutes to boot.
The Docker engine is utilized by Docker.	VMs are operated by Hypervisor.
The utilization mechanism for Docker is complex.	The use of and interaction with VM tools is simple.

The VMs increase the level of isolation between the host and the apps. VM-based applications are restricted to communicating solely with the VM kernel and are unable to access the host kernel directly. Therefore, for an attacker to target the host kernel, they would need to bypass both the hypervisor and the kernel of VM. Conversely, the container model enables direct application access to and communication with the kernel of host, as illustrated in figure 1.

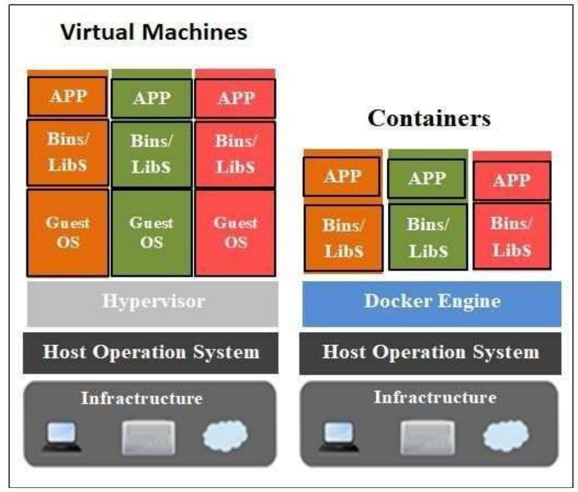


Figure 1: Docker and VM [10]

This situation enables an attacker to directly attack the host kernel, making container technology more susceptible to security concerns when compared to VM platforms [10].

3 Architecture of Docker

Docker architecture consists primarily of major components, as shown in Figure 2: Docker host and client, Docker image, Docker registries, and Docker container[11]. In the sections that follow, these elements will be covered in more detail:

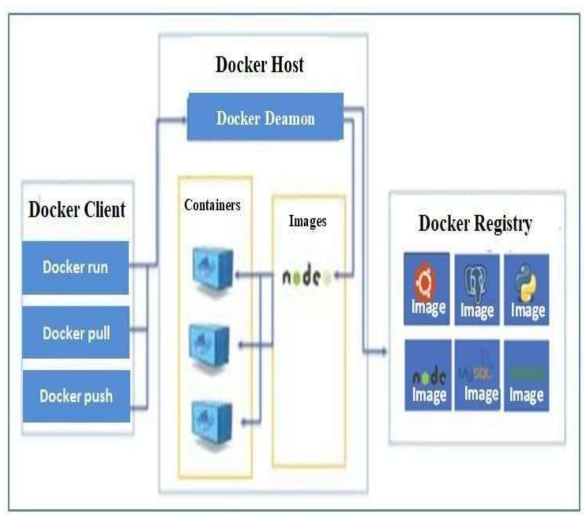


Figure 2: Docker Architecture [11]

3.1 Docker Host and Client

Docker client and Docker host or daemon can operate on the same machine, or a local Docker client can connect to a host or daemon operating on a different machine.

Docker clients allow users to interact with Docker. When a Docker command is executed, the client sends it to the Docker daemon, that executes it and is in responsible of all container-related activities and collects commands. The Docker client can communicate with multiple daemons. A complete environment for executing and running applications is provided by the Docker host. This includes the Docker daemon, image, container, storage, and network [12].

3.2 Docker Images

It is significant to build docker images based on two methods. The first one includes building an picture from a read-only pattern. A base image serves as the foundation for all images. These operating system images allow the creation of a container with a fully functional operating system. Alternatively, a center image can be produced from scrape and modified by adding the required applications. The second step involves creating a Dockerfile containing a list of instructions. When the "Docker build" command is executed, it uses these instructions to construct the image. With the use of this technology, creating images can be automated [13].

3.3 Docker Registries

Registries, which function similarly to source code repositories, are where Docker images are kept. These registries serve as centralized locations for uploading and downloading images. Private and public registrations are the two different categories of registers. Anyone may push and pull their images from public registries like Docker Hub without having to start from scratch. It is commonly known that Docker Hub is a public registry that makes it simple to access and share Docker images. Docker has the possibility to construct private registries in addition to public ones. These private registries use Docker's hub feature to limit access and let businesses or individuals share images to particular public or private spaces. Finally, Docker registries are essential for organizing and distributing Docker images because they provide both private and public choices to meet varying security needs and use cases. [14].

3.4 Docker Containers

A Docker image is used to create a Docker container, which contains all the parts needed to run an application in an isolated manner. Docker containers have a number of advantages. By doing so, developers may ensure consistent behavior across development, testing, and production environments and package apps and their dependencies into a single unit. Additionally, containers offer isolation, which eliminates issues while running many applications on the same host. They provide easy orchestration, scalability, and rapid deployment, which simplify the management and scaling of applications in a distributed, cloud-native environment. [15].

4 Docker Container Security

When it comes to running services in virtual environments like Docker, which is built on container-based virtualization technology security is a major concern. A number of factors are examined in order to evaluate the security that Docker offers for applications that are executing. This analysis explores the internal security measures implemented by Docker containers, along with the enhanced security features provided by the underlying kernel. It also addresses the challenges that arise when working with Docker containers from a security standpoint.

4.1 Docker Internal Security

Docker containers differ from traditional virtualization methods as they do not rely on virtualizing hardware or employing a separate operating system. Instead, Docker utilizes the Linux Namespace mechanism to achieve safe isolation of operating environment. Furthermore, it leverages the Linux Cgroup mechanism effectively to manage computer resources. The employing process of these mechanisms, Docker ensures the secure separation of virtual objects without the need for hardware virtualization or an independent operating system [16 - 17].

4.1.1 Isolation of a resource

Docker ensures secure isolation by leveraging the Linux Namespace mechanism. By using this mechanism, isolated containers are certain not to be able to access resources associated with other containers, thereby providing transparency in computer resource allocation. The details of resource isolation achieved through namespaces are presented in table 2 [18].

Table 2: Details of resource isolation [18]

Contents Isolated	Call Arguments of System	Namespace
Isolation of Process	CLONE _ NEWPID	PID
Isolation of Inter-Process Communication	CLONE _ NEWIPC	IPC
Isolation of a device	CLONE _ NEWUTS	UTS
Isolation of a network	CLONE _ NEWNET	Network
Isolation of a Filesystem	CLONE _ NEWNS	Mount

Isolation of Process PID: Process isolation in Docker is achieved through the utilization of PID (Process Identifier) namespaces. The main objective of this process, prevent compromised containers from interfering with other containers through management process interfaces. Docker implement this procedure by encapsulating processes within containers and restricting their permissions and access to other containers that deals with host device [18].

Isolation of Inter-Process Communication (IPC): shared memory, semaphores and message queues blocks are good examples of IPC objects used for exchanging data among processes. When processes operate in containers, it is important to restrict their communication to a specific set of IPC resources and prevent interference with processes running on the host machine or other containers. Docker employs IPC namespaces to achieve IPC isolation. By setting the parameter CLONE_NEWIPC during the clone operation, Docker creates separate IPC namespaces. Each IPC namespace consists of a collection of IPC object identifiers. Processes within a namespace cannot access or modify IPC resources in another namespace[18].

Isolation of a device: In Unix, Device drivers can be accessed by device nodes, which are special files, allowing the kernel and programs to access hardware. However, if a container has unrestricted access to critical device nodes, it can potentially cause significant damage to the host system. As a result, it is critical to limit a container's access to device nodes [3].

Isolation of a network: Network isolation is crucial to prevent network-based attacks like ManintheMiddle and ARP spoofing. Docker creates separate network stacks for each container, allowing them to interact through their respective network. By default, connectivity is provided via the Virtual Ethernet Bridge, which forwards packets between its network interfaces. However, this model is vulnerable to ARP spoofing and Mac flood attacks as it forwards all incoming packets without filtering [19].

Isolation of a Filesystem: Docker uses mount namespaces to isolate the filesystem hierarchy associated with different containers, providing processes with different views of the file system structure. However, some kernel file systems do not have a namespace, causing containers to inherit the host's view and access them directly. Docker limits threats through two filesystem protection mechanisms: revokes container permissions to write to these filesystems and prohibits processes from remounting a file system inside the container. Additionally, Docker uses a copy on write file system, allowing each container to write content to its specific filesystem[20].

4.1.2. Control of Resources

Docker containers use Linux's Control Group, or Cgroup to ensure efficient utilization of system resources like memory, CPU, block I/O and bandwidth. As a result, each container instance is able to compete for resources equally, and it is nearly impossible for any container instance to run out of the host computer's system resources. When a system resource is depleted, the Linux kernel will trigger out of Memory, that will terminate all active containers or processes. As a result of the Cgroup mechanism, Denial-of-Service (DoS) attacks are able to successfully controlled [21].

4.1.3. Limiting of Resources

A common attack on multi-tenant systems is a DoS attack, where a process consumes all system's resources, disrupting the normal operation of remaining processes. Limiting the resources allotted to each container ought to be achievable in order to stop this kind of attack. The primary tool used by Docker to address this problem is Cgroups. Resource management in Docker involves overseeing the allocation of key resources like CPU disk I/O and memory for each container. This ensures equitable distribution of resources among containers and prevents any single container from monopolizing them[3].

4.2. Docker and Kernel Security System

Two kernel security mechanisms, the Linux function and the Linux Security Module (LSM), that can improve the Docker security and the underlying kernel. The Linux function limits each process's permissions, while the Linux kernel manage many security models using the infrastructure provided by the LSM. LSM integrates security features like AppArmor, SELinux and other implemented measures into the official Linux kernel. This kernel security system contains the following items:

4.2.1. Linux function

Users are classified into two categories on traditional UNIX systems: root users and non root users. The first type possess high privileges and can bypass the kernel's permission checks, while the second type must adhere strictly to execution authorizations to perform tasks. Linux introduced the capability mechanism in version 2.2, which enables non-root users to execute tasks previously restricted to root users. This mechanism operates on processes or files, facilitating access control.

Furthermore, in the event that an intrusion manages to gain root access within the containers, Docker employs measures to mitigate its impact on the host system. This is achieved by restricting a specific set of Linux functions. Table 3 displays some of the disabled functions within a Docker container [19] .

4.2.2. SELinux

A security improvement for the Linux system is SELinux. Linux includes the common Discretionary Access Controls (DAC) method to manage access to objects, including owner/group and permission flags. Following the normal DAC, SELinux offers a second level of permission checking known as Mandatory Access Control (MAC). MAC can offer greater security than DAC. As users are unable to alter their security level or the object security features, MAC relies on matching of user and data security levels to make specific and objective decisions. The subjective factors of user are secured, resulting in an overall improvement in system security. MAC is typically used in conjunction with DAC in the Linux system, and specialized security modules are developed within the LSM framework. Everything is managed via labels in SELinux. Each process, system object, and file/directory has a label. These labels are used by the system administrator to create rules that regulate access among system objects and processes. They are referred to as policies [22].

Table 3: list of disabled Docker container functions[19]

process name of Docker	Disabled functions
CAP _ SETPCAP	Modifying process functions
CAP _ SYS _ MODULE Insert	delete kernel module
CAP _ SYS _ RAWIO	Change the memory of kernel
CAP _ SYS _ PACCT	Configuration process record
CAP _ SYS _ RESOURCE	Covering resource constraints
CAP _ SYS _ NICE	Change the process priority
CAP _ SYS _ TIME	Change the clock of system
CAP _ AUDIT _ WRITE	Writing audit logs
CAP _ MAC _ OV ERRIDE	Disregard the MAC policy in the kernel.
CAP _ SYSLOG	Changing the behavior of kernel printk
CAP _ SYS _ TTY _ CONFIG	Configuring TTY devices
CAP _AUDIT _ CONTROL	Configuring audit subsystem
CAP _ MAC _ ADMIN	Set up the MAC configuration
CAP _ SYS _ ADMIN	Select all
CAP _ NET _ ADMIN	Configure the network

4.2.3 AppArmor

AppArmor is a model for improving Linux security that relies on required access control, but it limits its scope to individual programs. This restricts the functionality of the program by downloading files of security configuration by administrators in each file. When a new container is started by Docker on a system that supports AppArmor, an interface is provided for loading a AppArmor profile. The profile is installed in application mode, ensuring that processes within the container adhere to the profile's restrictions. If no profile is specified during container startup, a default profile is automatically loaded into the container by the Docker daemon. This default profile prevents access to the host's crucial filesystems [23].

4.2.4. Seccomp

Seccomp is a mechanism that enables the restriction of system calls made by user processes and allows the filtering of system call parameters. System calls play a crucial role in connecting kernel states and user. Processes can function within a secure and controlled range through restricting system calls.

In Docker, a whitelist approach is adopted for Seccomp, where a configuration file is utilized to specify the allowed system calls, while over 50 specific system calls are blacklisted. This approach ensures that processes running within Docker containers are restricted to a safe subset of system calls and mitigates potential security risks[24].

5 Enhancing Docker Containers Security

Docker Containers have gained popularity in the software development community because they allow developers to bypass time-consuming library and dependency setting. However, along with the benefits of containerization, there are also security challenges that may make data vulnerable to attackers so must be addressed to ensure the integrity and protection of containerized applications[26]. Let's explore some of the key security challenges facing Docker containers and the solutions to address them shown in table 4.

Docker Containers supports the usage of SELinux and AppArmor frameworks to improve the Docker containers security. These frameworks enables to specify rules of the activities and access rights for containers, such as Docker-sec and Lic-sec.

Docker-sec offers an extra layer of security on top of Docker's default security by automatically constructing per-container AppArmor profiles.Docker-sec creates profiles of container based on behavior of application and configuration instructions, see figure 3. This procedure combines dynamic monitoring and static analysis to generate and enhance profiles of container operations during a predetermined test period.

Table 4: Security challenges for container networks

Security Challenge	Description	Solution
Container Breakouts	Compromise within a container can lead to the entire host being compromised.	Implement strong isolation between containers, apply container runtime security measures, and regularly update container runtimes.
Image Vulnerabilities	Vulnerabilities within container images can be exploited to compromise containerized applications.	Use trusted base images, regularly update container images, perform vulnerability scanning, and enforce secure image sourcing practices.
Inadequate Isolation	Weak isolation between containers can result in unauthorized access and data leakage.	Utilize container network segmentation, apply proper access controls, and enforce least privilege principles for containerized applications
Network Security	Ensuring secure communication, traffic control, and protection against network-based attacks.	Implement network security policies, utilize secure overlay networks or service meshes, and employ network monitoring and intrusion detection systems.
Dynamic IP Changes	Dynamic IP addresses make it challenging to implement IP-based access control and modify firewall rules.	Explore container network security solutions that offer dynamic IP-based access controls and leverage container orchestrators' network policies.
Orchestration Platform Security	Misconfigurations or vulnerabilities in the container orchestration platform can lead to unauthorized access and data exposure.	Follow security best practices for container orchestration platforms, regularly update them, enforce strong access controls, and implement security auditing.
Insider Threats	Malicious insiders with legitimate access can misuse privileges and compromise containers or data.	Implement strong access controls, enforce the principle of least privilege, monitor and audit user activities, and provide security awareness training.

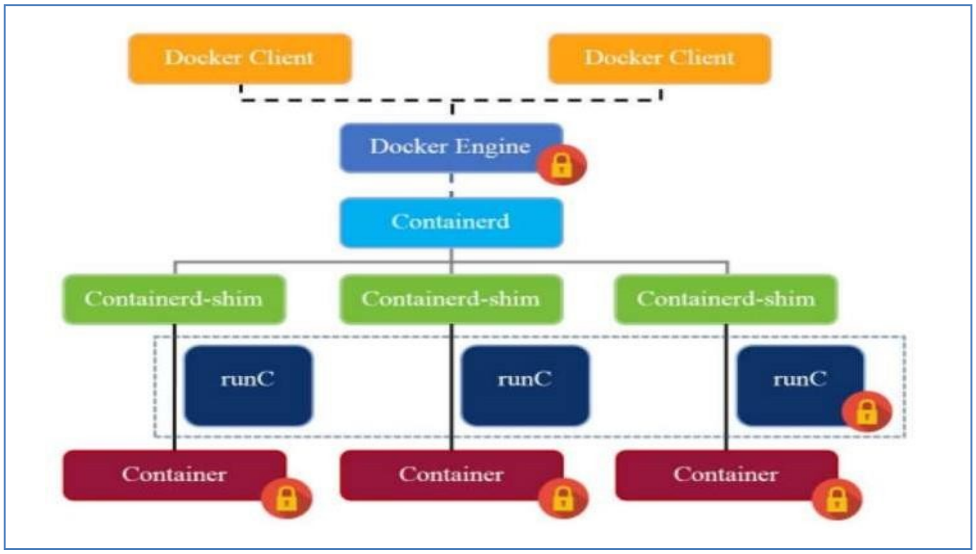


Figure 3. Components of the counter that the Docker-sec can protect

Dynamic Monitoring allows users to schedule container training time, allowing Docker-sec to capture behavior data, identify necessary permissions, examine audit log, adjust execution profile, potentially decreasing capabilities, and repeat until needed functionality is documented. Static Analysis gathers container and operation data from command-line arguments supplied by the user or data produced by Docker. The data is employed to create rules basic security and presentation new containers profiles. Docker-sec collects critical details from arguments of command line and utilities of Docker, like container volumes, privileges of user, and the SHA256 checksum of the container. Docker-sec may briefly enforce an AppArmor profile during the container setup process before switching to the one used during the container's runtime. Despite the capabilities provided by Docker-sec, its protection against user space program targeted exploits is limited[27].

Lic-Sec is an AppArmor profile generator which merges Docker-sec and LiCShield's best features in order to offer a more complete and effective security solution for Docker containers, figure 4 provides an overview of the Lic-Sec. It concentrates on creating AppArmor profiles for all Docker components automatically, assuring greater confinement of rights inside the container, and facilitating expansion of functionality, although it still has limits when it comes to web-server attacks. The LiCShield framework secures Docker containers and workloads by automatically generating AppArmor rules for both the host and container. It traces kernel operations using SystemTap5, converts traces to AppArmor rules, and constructs two profiles: one for operations inside the container and one for operations on the host[28].

6 Discussion of the Docker Container Security

Docker Containers provide an effective method and lightweight of packaging an application and all of its dependencies. However, several security concerns are preventing their wider deployment [29]. According to the analysis, Docker allows a high level isolation of process, inter-process communication, device, network, and filesystem. It also has the ability to control resources and limiting resources for its containers.

LiCShield and Docker-sec are popular ways to improve Docker container security based on MAC and allow container protection without manual configurations. The design principle of LiCShield is comparable to Docker-sec. It does not, however, build rules for capabilities and network accesses, but it does generate other critical rules that Docker-sec cannot generate, such as file access rules, mount rules, and so on. Table 5 shows the differences between the two methods [28]. Lic-Sec an AppArmor profile generator combines the strengths of both methods and provides stronger protection. Lic-Sec provides a mechanism for creating AppArmor rules that are specific to each container. It will be possible to create a generator that provides a greater level of safety with more rules to protect various portions.

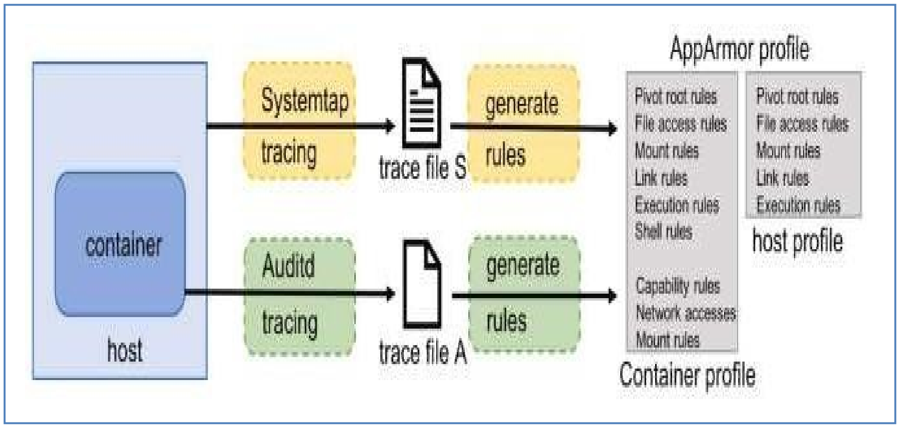


Figure 4. Overview of the Lic-Sec framwork

Table 5 : Comparison between and LiCShield and Docker-sec.

LSM	MAC	Year	Tracing tool	Generated profiles	Generated rules	Pre-defined profiles	Effective protective range
LiCShield	AppArmor	2015	SystemTap	Container runtime Docker daemon	Pivot root rule access rule mount rule link rule execution rule	-----	Docker daemon
Docker-sec	AppArmor	2018	Auditd	Container runtime	Capability rule network rule	RunC Docker daemon	Container RunC Docker daemon

7 Conclusions and Future Work

The migration of critical applications to containers is anticipated to accelerate as container applications continue to gain popularity. While assuring convenient deployment and optimum resource usage, improving container security will become an increasingly important problem.

This paper illustrates the fundamental concepts of Docker and explains how Docker achieves secure isolation through the use of Linux Namespace and Cgroup mechanisms, which ensure resource isolation and control within containers, as well as the importance of considering Linux kernel security features such as AppArmor and SELinux to strengthen the host system when using Docker. A variety of additional measures have been proposed to improve the Docker containers security, and some of the ways to enhance container security are reviewed.

In the future work, we want to improve Docker's scheduling feature and create a safer container variant that will loosen security restrictions.

References

[1] MP, A. R., Kumar, A., Pai, S. J., Gopal, A. “Enhancing security of docker using linux hardening techniques”. 2nd International Conference on Applied and Theoretical Computing and Communication Technology, pp.94–99, 2016.

[2] Lee, Haneul, Soonhong Kwon, Jong-Hyouk Lee. “Experimental Analysis of Security Attacks for Docker Container Communications”. Electronics 12.4 , 2023.

[3] Bui, Thanh. “Analysis of docker security”. arXiv preprint arXiv:1501.02967 , 2015.

[4] Bélair, M., Laniepce, S., & Menaud, J.-M. Leveraging kernel security mechanisms to improve container security: A survey. Proceedings of the 14th international conference on availability, reliability and security, 1–6, 2019.

[5] Souppaya, Murugiah, John Morello, Karen Scarfone. “Application container security guide” (2nd draft). No. NIST Special Publication 800-190. National Institute of Standards and Technology, 2017.

[6] Brady, K.; Moon, S.; Nguyen, T.; Coffman, J. "Docker Container Security in Cloud Computing". In Proceedings of the 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 6–8 January; pp. 975–980, 2020.

[7] Chelladhurai, Jeeva, Pethuru Raj Chelliah, Sathish Alampalayam Kumar. "Securing docker containers from denial of service (dos) attacks", International Conference on Services Computing (SCC). IEEE, 2016.

[8] T. Combe, A. Martin, R. Di Pietro, "To Docker or Not to Docker: A Security Perspective" in IEEE Cloud Computing, vol. 3, no. 5, pp. 54-62, Sept.-Oct. 2016.

[9] Jain, V., Singh, B., Khenwar, M., Sharma, M. "Static vulnerability analysis of docker images", IOP Conference Series: Materials Science and Engineering. Vol. 1131. No. 1. IOP Publishing, 2021.

[10] Yasrab, Robail. "Mitigating docker security issues". arXiv preprint arXiv:1804.05039 , 2018.

[11] Garg, Somya, Satvik Garg. "Automated cloud infrastructure, continuous integration and continuous delivery using docker with robust container security." IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, 2019.

[12] Loukidis-Andreou, F., Giannakopoulos, I., Doka, K., Koziris, N. "Docker-sec: A fully automated container security enhancement mechanism." IEEE 38th International Conference on Distributed Computing Systems (ICDCS). IEEE, 2018.

[13] Rad, Babak Bashari, Harrison John Bhatti, Mohammad Ahmadi. "An introduction to docker and analysis of its performance." International Journal of Computer Science and Network Security (IJCSNS) 17.3 ,228, 2017.

[14] Brogi, Antonio, Davide Neri, Jacopo Soldani. "DockerFinder: Multi-attribute search of docker images." 2017 IEEE international conference on cloud engineering (ic2e). IEEE, 2017.

[15] S. Winkel, "Security Assurance of Docker Containers: Part 1," ISSA Journal, April 2017.

[16] Alyas, Tahir, et al. "Container Performance and Vulnerability Management for Container Security Using Docker Engine." Security and Communication Networks 2022.

[17] E. Reshetova, J. Karhunen, T. Nyman, N. Asokan. "Security of OS-level virtualization technologies". In Proceedings of the 2014 NordSec Conference, pages 77–93, Norway, 2014.

[18] D. Huang, H. Cui, S. Wen, C. Huang, "Security Analysis and Threats Detection Techniques on Docker Container", IEEE 5th International Conference on Computer and Communications (ICCC), Chengdu, China, 2019, pp. 1214-1220

[19] J. Wenhao, L. Zheng, "Vulnerability Analysis and Security Research of Docker Container", IEEE 3rd International Conference on Information Systems and Computer Aided Education (ICISCAE), Dalian, China, 2020, pp. 354-357,

[20] Sever D , Kisasondi T . "Efficiency and security of Docker based honeypot systems", International Convention on Information &Communication Technology, Electronics & Microelectronics. 2018.

[21] S. Wu, K. Wang, H. Jin, "Research Situation and Prospects of Operating System Virtualization", Journal of Computer Research and Development, vol. 56, no. 1, pp. 58-68, 2019.

[22] Bhalerao, Amol. "Docker and analysis of its security." Journal homepage: www. ijrpr. com ISSN 2582: 7421.

[23] Zhu, Hui, Christian Gehrmann. "Lic-Sec: an enhanced AppArmor Docker security profile generator." Journal of Information Security and Applications 61: 102924, 2021.

[24] Tolaram, Nanik. "Docker Security." Software Development with Go: Cloud-Native Programming using Golang with Linux and Docker". Berkeley, CA: Apress, 89-107, 2022.

[25] Xiang, Jie, Long Chen. "A method of docker container forensics based on api". Proceedings of the 2nd International Conference on Cryptography, Security and Privacy. 2018.

[26] Baker, S. A., Nori, A. S. "Internet of things security: a survey". Advances in Cyber Security: Second International Conference, ACeS 2020, Penang, Malaysia, December 8-9, 2020, Revised Selected Papers 2. Springer Singapore, 2021.

[27]Loukidis-Andreou, F., Giannakopoulos, I., Doka, K., & Koziris, N. "Dockersec: A fully automated container security enhancement mechanism". 2018 IEEE 38thInternational Conference on Distributed Computing Systems (ICDCS), 1561–1564, 2018.

[28] Zhu, Hui, and Christian Gehrmann. "Lic-Sec: an enhanced AppArmor Docker security profile generator." Journal of Information Security and Applications 61 (2021): 102924.

[29] Baker, S. A., Nori, A. S. "A secure proof of work to enhance scalability and transaction speed in blockchain technology for IoT". In AIP Conference Proceedings (Vol. 2830, No. 1). AIP Publishing , 2023.

BIOGRAPHY / BIOGRAPHIES

Dr. Shatha A. Bakr has a Bachelor's degree in Computer Science from the University of Mosul, which she obtained in 1997. In 2013, she earned a master's degree in Computer Science from the same university. Later, in 2022, she completed her Ph.D. from the University of Mosul. Dr. Bakr worked as a Lecturer at the Northern Technical University in Mosul, Iraq. Her research interests encompass mobile phone programming, information security, multimedia communications, and artificial intelligence.

Dr. Hisham Hashem earned his bachelor's, master's, and doctoral degrees from the University of Mosul, Faculty of Computer Science and Mathematics, Department of Computer Science, in 2010, 2013, and 2022, respectively. Currently, he works as a Lecturer at the Northern Technical University in Mosul, Iraq. His research areas encompass big data, artificial intelligence applications, image processing, and bioinformatics

Omar I. Alsaif is currently a lecturer in the Mosul Technical Institute/Northern Technical University in Mosul, Iraq. He received his B.Sc. in electrical engineering from the University of Mosul in 1992. In 2005 and 2018, he obtained his M.Sc. and Ph.D. degrees in Electronics and Microelectronic Engineering from Mosul University, respectively. His research interests encompass microelectronic and solid-state systems, renewable energy, and nanotechnology devices.

Enhancing Acute Lymphoblastic Leukemia Image Segmentation: Unveiling the Impact of Color Spaces and Clustering Techniques

Sourav Banerjee*, Sudip Barik*, Jacob Tauro*, Debashis Das§, and Narayan C Debnath†
Kalyani Government Engineering College, India*
Meharry Medical College, USA §
Eastern International University, Vietnam†

Abstract

Leukemia, a cancer characterized by an excess of white blood cells, particularly Acute Lymphoblastic Leukemia (ALL), necessitates precise white blood cell (WBC) segmentation for computer-aided diagnosis. Despite the complex nature of WBC morphology and background, achieving accurate segmentation remains challenging. This paper explores contemporary WBC segmentation techniques, emphasizing clustering methods within ALL images, and highlighting the significance of suitable color spaces for cytoplasmic region segmentation. Additionally, it conducts a comparative evaluation of various color spaces and image clustering algorithms, with the aim of identifying the most effective combination for precise WBC segmentation.

Key words: Acute Lymphoblastic Leukemia (ALL), Clustering, Pathology Image, Image Segmentation, Color Spaces · White Blood Cell.

1 Introduction

1.1 The Role of Digital Pathology Images in Disease Diagnosis

Digital pathology images revolutionize disease diagnosis, with computer-aided diagnosis (CAD) providing efficient quantitative analysis. CAD mitigates observer discrepancies, emphasizing the importance of precise segmentation de- spite challenges such as noise and variations in cell characteristics [6, 9].

1.2 Acute Lymphocytic Leukemia and the Need for Accurate WBC Segmentation

Acute Lymphocytic Leukemia, also known as Acute Lymphoblastic Leukemia, is a severe hematological cancer that can quickly become fatal if not promptly treated. This disease predominantly affects children, with the highest incidence occurring between the ages of 2 and 5 years. Detecting this condition in its early stages is paramount, especially for children's well-being. The immature white blood cells (WBCs) associated with this condition are called lymphoblasts or blasts. These blast cells are categorized into three types based on their morphology: L1, L2, and L3, as per the FAB (French American-British) classification system [10]. Each category displays distinct shapes and patterns. L1 cells are characterized by their small size and uniform shape, accompanied by a minimal cytoplasmic portion. The nucleus is well-defined and circular. In contrast, L2 cells exhibit larger size and varying shapes. They have an asymmetrical nucleus and variable cytoplasmic area. L3 cells possess a circular or oval nucleus and are similar in size and shape. The nucleus contains vacuoles, and there's a noticeable cytoplasmic area. Generally, L3 cells are larger than L1 cells. Refer to Figure 1 for visual representations of L1, L2, and L3 cells.

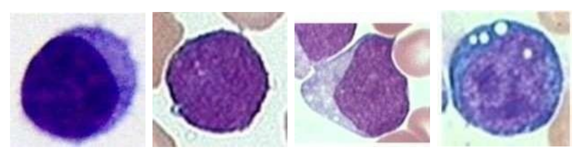


Fig. 1: Variability in morphology observed in blast cells as per the FAB classification. (a) Healthy lymphocytes cell; (b) L1- lymphoblasts Cell; (c) L2- lymphoblasts Cell; (d) L3- lymphoblasts Cell.

Accurate classification of these cell types offers vital insights to medical professionals, aiding them in formulating appropriate treatment plans. Hematologists perform differential white blood cell counts during blood tests to diagnose the disease [16]. However, this process is time-consuming, laborious, and expensive when done manually by experts. Furthermore, manual results might lack precision and reproducibility. Recent advancements have introduced automatic white blood cell classification technology with lower costs and enhanced accuracy, sparking interest in the field of hematological disease diagnosis.

The automated classification of white blood cells typically involves three main stages: segmentation, feature extraction, classification, and counting. It's evident that precise segmentation significantly influences the accuracy of

* Computer science and Engineering
§ Computer science and Engineering.
¶ School of Computing and Information Technology

classification and counting [16]. Hence, achieving accurate segmentation of white blood cells is paramount for the detection of Acute Lymphoblastic Leukemia. Numerous automated image segmentation techniques have been developed for early-stage Leukemia detection, including automatic thresholding [11], clustering [5], watershed [15], active contour [8], and deep learning [14]. Clustering, especially in various color spaces [18] (e.g., RGB, HSI, HSV, YCbCr, CIELAB), has shown promise for WBC segmentation. This study focuses on clustering- based WBC segmentation and aims to identify optimal color space and clustering technique combinations for precise ALL image segmentation. The following sections outline the structure of this paper: In Section 2, we provide detailed descriptions of the dataset used for White Blood Cell (WBC) segmentation. Section 3 elaborates on the clustering method employed, while Section 4 delves into the discussion of experimental results. Lastly, Section 5 offers concluding remarks, along with summaries, and outlines directions for future research.

2 Dataset Description

Two prominent publicly available datasets have been widely utilized in the academic literature for White Blood Cell (WBC) segmentation: the ALL-IDB Dataset [10]and the LISC Dataset [13]. These datasets are described in detail as follows:

ALL-IDB Dataset: The images contained within the ALL-IDB dataset were captured using a PowerShot G5 camera and were saved in JPG format, featuring a 24-bit color depth and a native resolution of 2592 x 1944 pixels. These images encompass a range of microscope magnifications, spanning from 300 to 500. The ALL-IDB database is categorized into two versions, specifically, ALL-IDB1 and ALL-IDB2, both designed to facilitate tasks related to segmentation and classification.

ALL-IDB1 constitutes a dataset comprising 108 images, which serves as a benchmark for the evaluation of segmentation and classification algorithms. Expanding upon this, ALL-IDB2 extends the capabilities of ALL-IDB1 and comprises 260 images, encompassing both normal and blast cells, with half of them representing lymphoblasts. It's worth noting that the grey-level characteristics of the images in ALL-IDB2 align with those found in ALL-IDB1.

LISC Dataset: The LISC Dataset comprises 400 color images of stained peripheral blood samples from 8 healthy subjects. These images were taken using a light microscope with a 100x magnification and saved in BMP format with a resolution of 720x576 pixels. The dataset comes from research centers in Tehran, Iran, and includes various normal leukocyte types. Additionally, an expert has outlined nucleus and cytoplasm regions in some of the images, but only 250 images have manual annotations.

3 Clustering Methods

Clustering embodies a systematic process wherein data entities are amalgamated into distinct clusters, guided by the principles of fostering high intra- cluster similarities and minimal inter-cluster similarities. The pursuit of optimizing this arrangement is encapsulated in defined objective functions, as elucidated by reference [5]. Leveraging these functions, the quest for the optimal partition of a given dataset is undertaken through the lens of minimizing or maximizing relevant metrics. These objective functions often encapsulate intricate statistical-mathematical relationships linking individual data elements to the prospective representatives of each cluster, colloquially known as cluster centers [1]. An optimal partition, delineating entities into 'c' classes, necessitates adherence to a specific set of properties that uphold the structural integrity of the resulting clusters:

1. At least one data vector should be assigned to each cluster, i.e.,
$$c_i \neq \emptyset, \forall i \in \{1, 2, \ldots, c\} \tag{1}$$
2. There should be no data vector in common between two clusters, i.e.,
$$c_i \cap c_j = \emptyset, \forall i \neq j \text{ and } i, j \in \{1, 2, \ldots, c\} \tag{2}$$
3. Each data vector must be associated with a cluster, i.e.,
$$\bigcup_{i=1}^{c} = D_v \tag{3}$$

Where, D_v is stands for total number of objects.

4 Experimental Results

The assessment of the proposed methodologies transpired across a comprehensive collection of a hundred images sourced from all IDB datasets. Within this evaluative framework, three widely acknowledged clustering algorithms, namely K-Means (KM), Fuzzy C-Means (FCM), and Self Organizing Map (SOM) were harnessed for the crucial task of segmentation. The vital parameters underpinning these distinct clustering techniques were meticulously selected as follows:

The selection of cluster prototype values varies according to the specific clustering algorithm: it is set to 5 for Fuzzy C-Means (FCM), 4 for K-Means (KM), and 3 for Self-Organizing Maps (SOM). In the case of FCM, we employ a fuzzification parameter of 2. The algorithms are designed to terminate when the maximum difference between two successive partition matrices U falls below predefined minimal error threshold η, which is mathematically represented as Max $U^t - U^{t+1} < \eta$, with η being set to 10^{-5}. For K-Means, the procedure halts when the change in centroid values becomes smaller than η.

This approach ensures that the clustering process terminates when the desired level of convergence or accuracy is achieved for each respective algorithm.

Figure 2 depicts the original images of several Acute Lymphoblastic Leukemia blast cells. The assessment of segmentation accuracy for the proposed models involves the use of seven performance evaluation parameters based on ground truth data. These parameters include Accuracy (AC) [12], Precision Index (PI) [7], Specificity Index (SI) [2], Recall Index (RI) [3], Dice Index (DI) (also referred to as F-score or F-measure or Dice similarity coefficient) [19], Jaccard Index (JI) [17], and Matthews's correlation coefficient (MCC) [4].

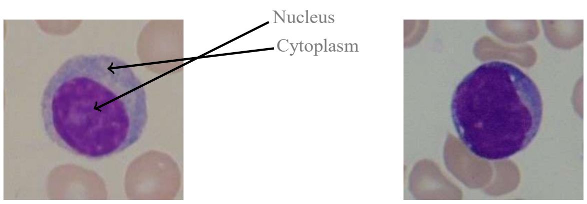
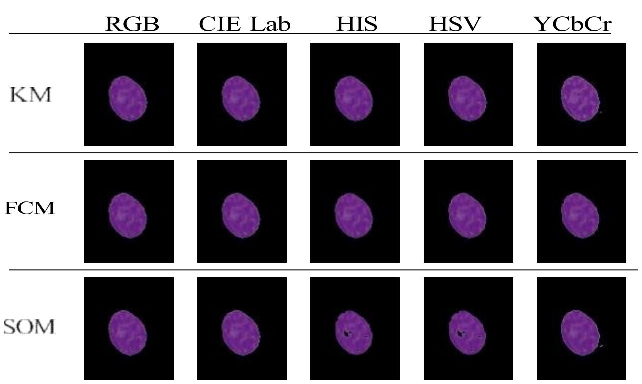


Fig. 2: Original Acute Lymphoblastic Leukemia images

4.1 Results over Nucleus Segmentation

This section evaluates nucleus segmentation using FCM, KM, and SOM across five color spaces. Tables 1 and 2 illustrate results for two images, while Tables 3,4,5 present average quality parameters of over 100 images. FCM and KM excel in CIE Lab, while SOM performs best in HSV and HSI, with SOM-HSV being the optimal pair.

Table 1: Segmenting Nuclei with Cluster Analysis in Various Color Spaces for Figure 2(a)



4.2 Results over Nucleus Segmentation

This section evaluates WBC (nucleus with cytoplasm) segmentation using FCM, KM, and SOM across five color spaces, presenting visual results in Table-5, 6, 7 and numerical data in Tables 8, 9, 10. FCM and KM excel in the CIE Lab color space, while SOM performs well with HSI and HSV. The most effective clustering-color space pairs are FCM-CIELab, KM-CIELab, SOM-HSI, and SOM-HSV, emphasizing the importance of proper technique and color space selection for WBC segmentation.

Table 2: Segmenting Nuclei with Cluster Analysis in Various Color Spaces for Figure 2(b)

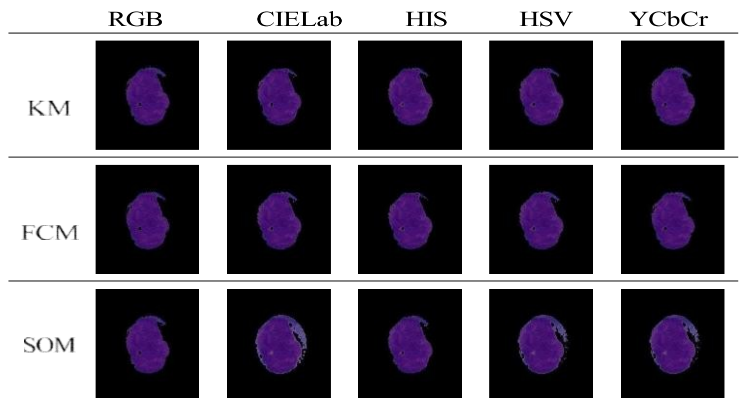


Table 6: Segmentation of Nucleus with Cytoplasm using Clustering Techniques in Various Color Spaces for Fig. 2(a)

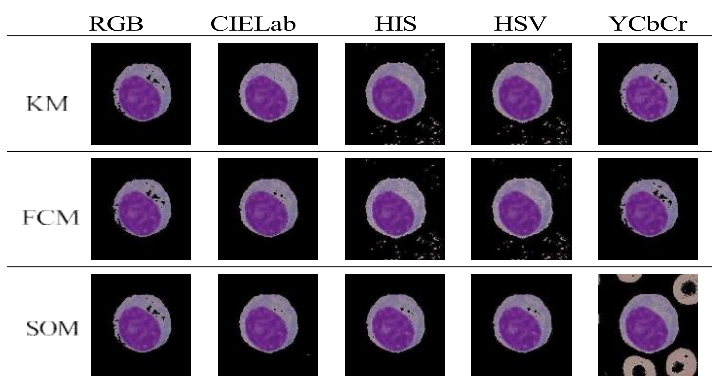


Table 3: Quality Parameters for FCM over different Color Spaces for nucleus detection.

Method	Accuracy	Recall	Precision	MCC	Dice	Jaccard	Specitivity
RGB	0.9839	0.9636	0.9602	0.9513	0.9610	0.9255	0.9870
CIELab	**0.9862**	0.9547	**0.9851**	**0.9560**	**0.9643**	**0.9314**	0.9927
HSI	0.9741	0.9026	0.9849	0.9265	0.9396	0.8894	**0.9963**
HSV	0.9801	0.9326	0.9742	0.9404	0.9520	0.9100	0.9930
YCbCr	0.9796	**0.9682**	0.9440	0.9421	0.9535	0.9127	0.9803

Table 4: Quality Parameters for KM over different Color Spaces for nucleus detection

Method	Accuracy	Recall	Precision	MCC	Dice	Jaccard	Specifivity
RGB	0.9831	0.9617	0.9596	0.9496	0.9601	0.9241	0.9872
CIELab	0.9858	0.9526	0.9818	0.9565	0.9644	0.9332	0.9944
HSI	**0.9844**	0.9478	0.9812	**0.9524**	**0.9616**	0.9268	**0.9952**
HSV	0.9825	0.9436	**0.9721**	0.9464	0.9567	**0.9183**	0.9922
YCbCr	0.9812	**0.9520**	0.9640	0.9456	0.9569	0.9190	0.9900

Table 5: Quality Parameters for SOM over different Color Spaces for nucleus detection.

Method	Accuracy	Recall	Precision	MCC	Dice	Jaccard	Specitivity
RGB	0.9801	0.9419	0.9818	0.9482	0.9604	0.9246	**0.9932**
CIELab	0.9471	0.9360	0.9179	0.8893	0.9195	0.8770	0.9505
HSI	**0.9844**	**0.9663**	0.9737	**0.9593**	**0.9694**	**0.9408**	**0.9903**
HSV	**0.9844**	0.9662	**0.9739**	**0.9593**	**0.9694**	**0.9408**	**0.9903**
YCbCr	0.8318	0.8759	0.7110	0.6688	0.7670	0.6706	0.8149

Table 8: Quality Metrics for Fuzzy C-Means in Various Color Spaces for Nucleus and Cytoplasm Detection

Method	Accuracy	Recall	Precision	MCC	Dice	Jaccard	Specitivity
RGB	0.9646	0.9414	0.9513	0.9215	0.9408	0.8939	0.9722
CIELab	**0.9763**	**0.9815**	**0.9588**	**0.9407**	**0.9555**	**0.9156**	**0.9842**
HSI	**0.9763**	0.9812	0.9327	0.9405	**0.9555**	**0.9156**	0.9736
HSV	0.9755	**0.9881**	0.9217	0.9379	0.9534	0.9120	0.9705
YCbCr	0.9599	0.9344	0.9407	0.9096	0.9314	0.8810	0.9700

Table 9: Quality Metrics for KM in Various Color Spaces for Nucleus and Cytoplasm Detection.

Method	Accuracy	Recall	Precision	MCC	Dice	Jaccard	Specitivity
RGB	0.9812	0.9477	0.9810	0.9515	0.9633	0.9296	**0.9929**
CIELab	**0.9850**	0.9649	0.9772	**0.9608**	**0.9705**	**0.9430**	0.9915
HSI	0.9754	**0.9803**	0.9311	0.9386	0.9539	0.9127	0.9726
HSV	0.9800	0.9788	0.9487	0.9495	0.9630	0.9289	0.9789
YCbCr	0.9799	0.9424	**0.9812**	0.9480	0.9604	0.9245	**0.9929**

Table 10: Quality Metrics for SOM in Various Color Spaces for Nucleus and Cytoplasm Detection

Method	Accuracy	Recall	Precision	MCC	Dice	Jaccard	Specitivity
RGB	0.9801	0.9419	0.9818	0.9482	0.9604	0.9246	**0.9932**
CIELab	0.9471	0.9360	0.9179	0.8893	0.9195	0.8770	0.9505
HSI	**0.9844**	**0.9663**	0.9737	**0.9593**	**0.9694**	**0.9408**	**0.9903**
HSV	**0.9844**	0.9662	**0.9739**	**0.9593**	**0.9694**	**0.9408**	**0.9903**
YCbCr	0.8318	0.8759	0.7110	0.6688	0.7670	0.6706	0.8149

5 Conclusion and Future works

This paper explores clustering-based techniques for segmenting white blood cells (WBCs) in Acute Lymphoblastic Leukemia (ALL) images, using various color spaces to enhance accuracy. Fuzzy C-Means (FCM) and K-Means (KM) with CIELab showed optimal results for nucleus and whole WBC segmentation. Self-Organizing Maps (SOM) with HSI excelled in nucleus segmentation, while HSI and HSV with SOM performed well for whole WBCs.

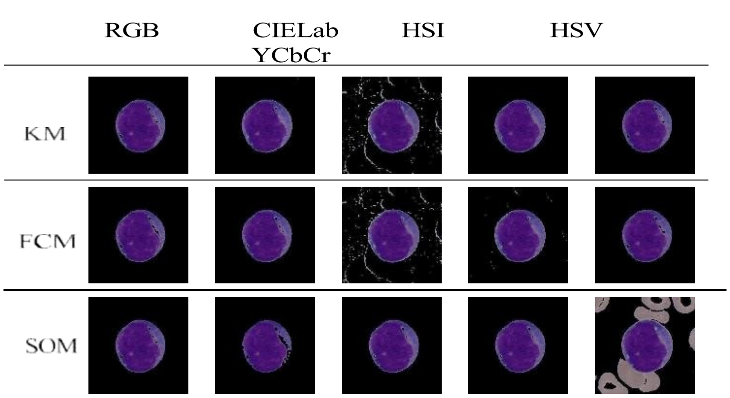
These findings emphasize that the choice of color space significantly impacts segmentation efficacy in different clustering approaches, highlighting the need for a tailored pairing of clustering method and color space.

The study also outlines potential future research directions, including:

- Explore advanced clustering techniques, including Rough-Fuzzy clustering, various Fuzzy clustering variants, and nature-inspired optimization algorithms (NIOA), to improve WBC segmentation.
- Investigate the use of diverse NIOAs, expanding beyond distance-based objective functions, to broaden the scope of research in WBC segmentation.
- Focus on optimizing color spaces for WBC segmentation, including under- utilized ones, and explore hybrid color spaces and color component-based segmentation to address illumination challenges.

In conclusion, this study underscores the significance of combining ap- propriate clustering techniques with suitable color spaces for effective WBC segmentation. It suggests avenues for future research to advance the field, encompassing the exploration of advanced clustering methods, diverse objective functions for NIOAs, and innovative approaches to leveraging color spaces for segmentation improvement.

Table 7: Segmentation of Nucleus with Cytoplasm using Clustering Techniques in Various Color Spaces for Fig. 2(b)



6 References

[1] Ajala Funmilola, A., Oke, O., Adedeji, T., Alade, O., Adewusi, E.: Fuzzy kc-means clustering algorithm for medical image segmentation. Journal of information Engineering and Applications, ISSN 22245782, 2225–0506 (2012).

[2] Alagarsamy, S., Govindaraj, V., et al.: Automated brain tumor segmentation for mr brain images using artificial bee colony combined with interval type-ii fuzzy technique. IEEE Transactions on Industrial Informatics (2023)

[3] Chen, J.: Survey of evaluation methods and metrics for face clustering. In: 2023 IEEE 13th International Conference on Electronics Information and Emergency Communication (ICEIEC), pp. 181–186. IEEE (2023)

[4] Chicco, D., Jurman, G.: The matthews correlation coefficient (mcc) should replace the roc auc as the standard metric for assessing binary classification. Biodata Mining **16**(1), 1–23 (2023)

[5] Dhal, K.G., Ray, S., Barik, S., Das, A.: Illumination-free clustering using improved slime mould algorithm for acute lymphoblastic leukemia image segmentation. Journal of Bionic Engineering pp. 1–19 (2023)

[6] Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M., Yener, B.: Histopathological image analysis: A review. IEEE reviews in biomedical engineering **2**, 147–171 (2009)

[7] He, Y., Yang, G., Ge, R., Chen, Y., Coatrieux, J.L., Wang, B., Li, S.: Geometric visual similarity learning in 3d medical image self-supervised pre-training. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 9538– 9547 (2023)

[8] Hegde, R.B., Prasad, K., Hebbar, H., Singh, B.M.K.: Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybernetics and Biomedical Engineering **39**(2), 382–392 (2019)

[9] Irshad, H., Veillard, A., Roux, L., Racoceanu, D.: Methods for nuclei detection, segmentation, and classification in digital histopathology: a review—current status and future potential. IEEE reviews in biomedical engineering **7**, 97–114 (2013)

[10] Labati, R.D., Piuri, V., Scotti, F.: All-idb: The acute lymphoblastic leukemia image database for image processing. In: 2011 18th IEEE international conference on image processing, pp. 2045–2048. IEEE (2011)

[11] Rawat, J., Singh, A., Bhadauria, H., Virmani, J., Devgun, J.S.: Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia. Biocybernetics and Biomedical Engineering **37**(4), 637–654 (2017)

[12] Ren, Z., Kong, X., Zhang, Y., Wang, S.: Ukssl: Underlying knowledge based semi- supervised learning for medical image classification. IEEE Open Journal of Engineering in Medicine and Biology (2023)

[13] Rezatofighi, S.H., Soltanian-Zadeh, H.: Automatic recognition of five types of white blood cells in peripheral blood. Computerized Medical Imaging and Graphics **35**(4), 333–343 (2011)

[14] Shahin, A.I., Guo, Y., Amin, K.M., Sharawi, A.A.: White blood cells identification system based on convolutional deep neural learning networks. Computer methods and programs in biomedicine **168**, 69–80 (2019)

[15] Suryani, E., Palgunadi, S., Pradana, T.N., et al.: Classification of acute myelogenous leukemia (aml m2 and aml m3) using momentum back propagation from watershed distance transform segmented images. In: Journal of Physics: Conference Series, vol. 801, p. 012044. IOP Publishing (2017)

[16] Swanson, K., Wu, E., Zhang, A., Alizadeh, A.A., Zou, J.: From patterns to patients: Advances in clinical machine learning for cancer diagnosis, prognosis, and treatment. Cell (2023)

[17] Tang, T., Zhang, R., Lin, K., Li, F., Xia, X.: Sm-rnet: A scale-aware-based multi- attention guided reverse network for pulmonary nodules segmentation. IEEE Transac- tions on Instrumentation and Measurement (2023)

[18] Zhang, N., Jiang, Z., Li, J., Zhang, D.: Multiple color representation and fusion for diabetes mellitus diagnosis based on back tongue images. Computers in Biology and Medicine **155**, 106652 (2023)

[19] Zhang, W., Cao, Y., Hu, X., Mi, J., Zhang, P., Mukhopadhyay, S.C., Li, Y., Liu, Z.: Iat: A full-scale ivus analysis toolbox based on deep learning algorithms for clinical diagnosis. IEEE Transactions on Instrumentation and Measurement **72**, 1–15 (2023)