Department of Electrical and Computer Engineering
North South University

Senior Design Project

Image Segmentation and Classification Using Radon Cumulative Distribution Transform and Nearest Subspace

Abeed Pasha	ID# 2131353042
Sirajus Salekin	ID# 2132275642
Mohiuddin Sarker	ID# 1821404042

Faculty Advisor:

Dr. Mohammad Shifat-E-Rabbi

Assistant Professor

ECE Department

Spring, 2025

LETTER OF TRANSMITTAL

April, 2025

To

Dr. Mohammad Abdul Matin

Chairman,

Department of Electrical and Computer Engineering

North South University, Dhaka

Subject: **Submission of Capstone Project Report on "Image Segmentation and Classification Using Radon Cumulative Distribution Transform and Nearest Subspace "**

Dear Sir,

With due respect, we would like to submit our **Capstone Project Report** on **"Image Segmentation and Classification Using Radon Cumulative Distribution Transform and Nearest Subspace"** as a part of our BSc program. This report elaborates on a novel image classification model system that provides numerous benefits compared to classic machine learning techniques in the same field. The project as a whole nurtured our knowledge on both a practical and theoretical level, and we hope that our growth as academics is showcased through the work we put into this endeavor.

We will be highly obliged if you kindly receive this report and provide your valuable judgment. It would be our immense pleasure if you found this report useful and informative and had an apparent perspective on the issue.

Sincerely Yours,

....

Abeed Pasha

ECE Department

North South University, Bangladesh

....

Sirajus Salekin

ECE Department

North South University, Bangladesh

....

Mohiuddin Sarker

ECE Department

North South University, Bangladesh

APPROVAL

Sirajus Salekin (ID # 2132275642), Abeed Pasha (ID # 2131353042) and Mohiuddin Sarker (ID # 1821404042) from Electrical and Computer Engineering Department of North South University, have worked on the Senior Design Project titled “Image Segmentation and Classification Using Radon Cumulative Distribution Transform and Nearest Subspace” under the supervision of Dr. Mohammad Shifat-E-Rabbi partial fulfillment of the requirement for the degree of Bachelors of Science in Engineering and has been accepted as satisfactory.

Supervisor’s Signature

………………………………………

Dr. Mohammad Shifat-E-Rabbi

Assistant Professor

Department of Electrical and Computer Engineering

North South University

Dhaka, Bangladesh.

Chairman’s Signature

………………………………………

Dr. Mohammad Abdul Matin

Professor

Department of Electrical and Computer Engineering

North South University

Dhaka, Bangladesh.

DECLARATION

This is to declare that this project is our original work. No part of this work has been submitted elsewhere partially or fully for the award of any other degree or diploma. All project related information will remain confidential and shall not be disclosed without the formal consent of the project supervisor. Relevant previous works presented in this report have been properly acknowledged and cited. The plagiarism policy, as stated by the supervisor, has been maintained.

Students' names & Signatures

1. Abeed Pasha

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

2. Sirajus Salekin

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

3. Mohiuddin Sarker

_ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _ _

ACKNOWLEDGEMENTS

The authors would like to express their heartfelt gratitude towards their project and research supervisor, Dr. Mohammad Shifat-E-Rabbi, Assistant Professor, Department of Electrical and Computer Engineering, North South University, Bangladesh, for his invaluable support, precise guidance, and advice pertaining to the experiments, research, and theoretical studies carried out during the course of the current project and also in the preparation of the current report.

Furthermore, the authors would like to thank the Department of Electrical and Computer Engineering, North South University, Bangladesh, for facilitating the research. The authors would also like to thank their loved ones for their countless sacrifices and continual support.

ABSTRACT

Image Segmentation and Classification Using Radon Cumulative Distribution Transform and Nearest Subspace

Traditional neural network-based image classification methods often require significant computational resources and time. In this work, we propose a novel classification approach that leverages the Radon Transform (RT), Cumulative Distribution Transform (CDT), and Nearest Subspace Classifier (NSC) to efficiently classify images while maintaining high accuracy. The Radon Transform enables us to extract more discriminative features by capturing directional information from images. The extracted features are then transformed using CDT, which enhances class separability. Finally, the NSC is employed to classify images based on their projections onto learned subspaces. For images containing multiple objects, we incorporate image segmentation techniques to classify each object individually. Our segmentation pipeline includes Binary Thresholding (Otsu's Method), Contour Detection, and Bounding Box Extraction, ensuring precise isolation of objects before classification. This approach significantly reduces computational complexity while achieving competitive classification performance compared to deep learning-based methods, for which we showcase an ablation study . We demonstrated the effectiveness of the proposed model on several datasets—such as MNIST digits, EMNIST, EMNIST letters, Bangla digits, and several others for which we got satisfactory results- 95.34%, 95.41%, 81.85%, and 88.84% accuracy respectively. Experimental results demonstrate that our method provides a robust and efficient alternative for image classification, particularly in resource-constrained environments.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1 Introduction

1.1 Background and Motivation

Image segmentation and classification are essential tasks in computer vision, widely used in fields like medical imaging, autonomous systems, and remote sensing. However, traditional segmentation methods often demand significant computational resources and time. This project explores the integration of Radon Cumulative Distribution Transform (R-CDT) and Nearest Subspace Classifier (NSC) to address these challenges. R-CDT simplifies image processing by transforming non-segmented images into a more manageable form, while NSC enhances classification efficiency by projecting data into lower-dimensional subspaces. Together, these methods offer a more resource-efficient alternative to traditional convolutional neural networks (CNNs), which are computationally expensive. The motivation for this project is to develop a fast and accurate solution for image classification, especially useful in resource-constrained environments. The integration of R-CDT and NSC can improve efficiency in applications like character recognition, medical diagnostics, autonomous systems, and remote sensing, where rapid image analysis is crucial.

1.2 Purpose and Goal of the Project

The purpose of this project is to explore the application of Radon Cumulative Distribution Transform (R-CDT) and Nearest Subspace Classifier (NSC) for image classification on non-segmented data. Traditional image segmentation methods often require significant computational resources and time, which can be a limitation in real-time or resource-constrained applications. This project aims to develop a solution that can classify images efficiently without the need for prior segmentation.

The goal is to improve the performance of image classification tasks by applying R-CDT to non-segmented images, transforming them into a format suitable for efficient classification. By integrating NSC, the project seeks to enhance classification speed and accuracy while minimizing computational overhead, making it ideal for applications where real-time processing and limited resources are key constraints.

1.3 Organization of the Report

This report is structured to provide a clear and detailed explanation of the project, its processes, findings, and the broader implications. The introduction, found in Chapter 1, provides the necessary context by outlining the background and motivation for the project, the specific goals and objectives, and an overview of the report’s structure. Chapter 2, the literature review, synthesizes existing research related to image segmentation and classification, focusing particularly on Radon Cumulative Distribution Transform (R-CDT) and Nearest Subspace Classifier (NSC), while highlighting the limitations and gaps that the project aims to address.

Chapter 3 presents the methodology, describing the system design, the hardware and software components used, and the implementation steps followed to achieve the project’s objectives. The core of the project, including the experimental setup and results, is detailed in Chapter 4, where the investigation is discussed alongside a comprehensive analysis of the findings. The chapter emphasizes the performance of the proposed approach and compares it to existing methods in terms of accuracy, computation time, and efficiency.

In Chapter 5, the impact of the project is analyzed from multiple perspectives, including societal, health, safety, legal, cultural, environmental, and sustainability considerations. This chapter aims to explore the broader implications and contributions of the work. Chapter 6 covers the project’s planning and budgeting process, offering insight into the timeline, resource allocation, and cost management aspects of the project.

The challenges faced during the project, categorized as Complex Engineering Problems (CEP) and Complex Engineering Activities (CEA), are discussed in Chapter 7, providing a deeper understanding of the engineering difficulties encountered and how they were addressed. Finally, Chapter 8 concludes the report by summarizing the key findings, discussing the limitations of the project, and proposing future improvements. This structured approach ensures that each section builds upon the previous one, providing a comprehensive view of the project’s lifecycle from conception to completion.

Chapter 2 Research Literature Review

2.1 Existing Research and Limitations

Convolutional Neural Networks (CNNs) come with significant drawbacks that limit their practicality, especially in specialized domains like biological or handwritten image classification. As highlighted by Jiaohua Qin et al. in their work on biological image classification, standard CNN architectures often require fixed-size input images, which can reduce their flexibility and effectiveness when dealing with real-world, variably sized data [2]. Furthermore, traditional CNNs tend to be computationally expensive and parameter-heavy, making them inefficient for deployment in resource-constrained environments such as mobile devices or embedded systems. These limitations not only slow down training and inference but also increase the risk of overfitting, especially when working with smaller or more complex datasets. Their analysis showcased the relatively low accuracies obtained even after several epochs of training the datasets, as low as 70.90% by InceptionV3 [3] and 73.40% with DenseNet201 [4].

In another study by Bhoomi Shah et al., the computational demands of Convolutional Neural Networks (CNNs) in image classification tasks were investigated. The study evaluates eight distinct CNN models, each varying in filter sizes, number of convolutional layers, filters, fully connected layers, and kernel sizes. Key findings indicate that factors such as optimizer choice, batch size, filter count, and neuron quantity significantly influence the time required for model training and inference [5]. Notably, convolutional layers, max pooling layers, and fully connected layers directly impact overall model performance. The results showcased the high complexity of classical CNN models in both computation requirements and time for training and testing.

A study by Samiha et al. introduces a novel approach for classifying breast tumors. The authors employ the Radon Cumulative Distribution Transform (RCDT) combined with a Nearest Subspace (NS) classifier to enhance classification accuracy. They evaluate their method using three datasets: two histopathology and one ultrasound. The RCDT-NS classifier achieved accuracies of 93.5%, 90.9%, and 90%, respectively [6]. When compared with deep CNNs like AlexNet and GoogLeNet on the same datasets, RCDT-NS demonstrated similar accuracy while being far more efficient. Notably, the method is nearly 10,000 times faster, requires no GPU, and performs well even with limited labeled data, making it highly suitable for resource-constrained environments.

Conventional CNNs are becoming increasingly obsolete, as demonstrated by the issues that have been arising in recent times. In response to this, RCDT-NS displays incredible potential to be a powerful alternative that is both lightweight and effective. This approach was initially explored by a study by M. Shifat-E-Rabbi et al., which showcased the effectiveness of the proposed RCDT-NS model in solving many mathematical issues previously faced by other image classification models. The method, which applies a nearest-subspace algorithm in the R-CDT domain, is straightforward to implement, does not rely on iterative processes, requires no hyperparameter tuning, and stands out for its computational and label efficiency. Despite its simplicity, it delivers classification accuracy comparable to that of cutting-edge neural network models across various tasks. The proposed method obtains accuracy figures similar to that of the CNN-based methods with 50 to 10,000 times reduction in the computational complexity while maintaining higher accuracies up to 2 to 4 times more for low sample datasets [1].

However, the method is in its infancy and has much room for development. Namely, the model cannot handle non-segmented data, which greatly reduces its versatility. Developing a comprehensive system that segments images before feeding it into the model is essential, and proper segmentation approaches need to be utilized. Otsu's thresholding comes to note, being extremely effective in most general purpose use cases. The study by Myat Thet Nyo et al. demonstrates the effectiveness of Otsu's thresholding technique in segmenting brain tumors from MRI images, achieving promising results in terms of accuracy and reliability. By evaluating different threshold classes, the authors found that Class 4 yielded the best performance, with a true positive rate of 68.80% and an overall accuracy of 95.56% [7]. These results highlight Otsu's method as a simple yet powerful tool for medical image segmentation. Its ability to achieve high accuracy without the need for complex computations or extensive training data makes it particularly suitable for resource-limited settings and real-time medical applications.

Chapter 3 Methodology

3.1 System Design

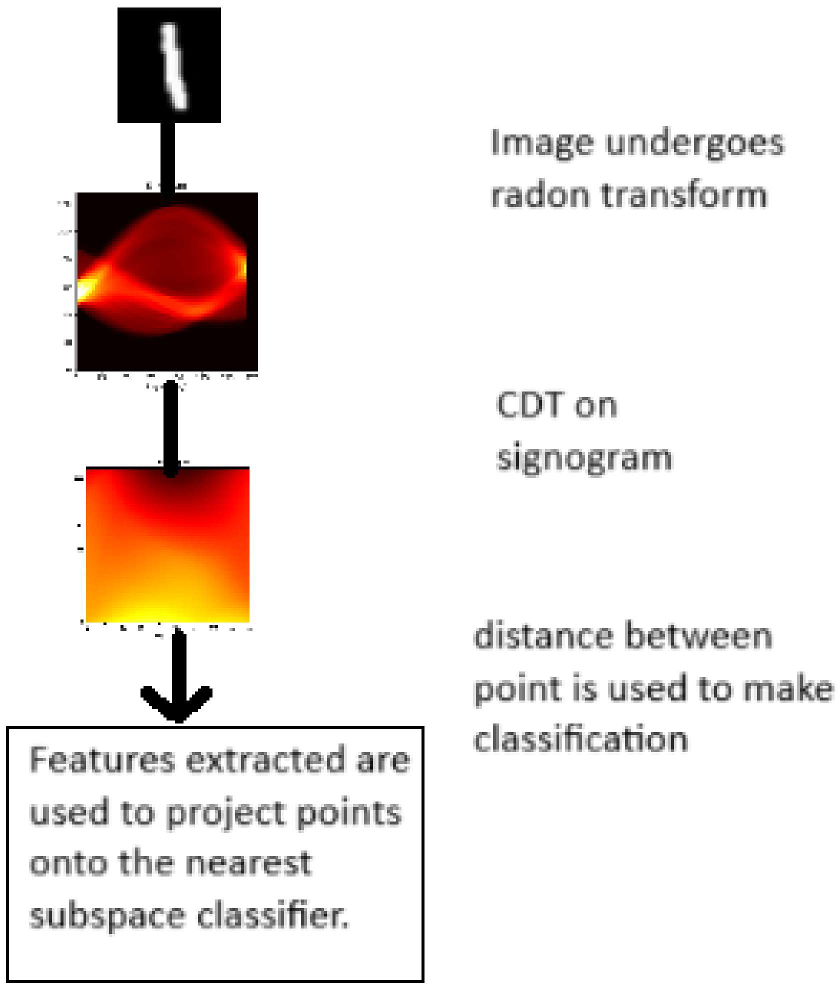

Fig 1: Pipeline of image classification using Rcdt

Fig 1 shows that the image processing begins with the Radon transform. The output of the Radon transform is a sinogram on which the Cumulative Distribution Function is applied. This allows for feature extraction that aids in the training and testing of the nearest subspace classifier

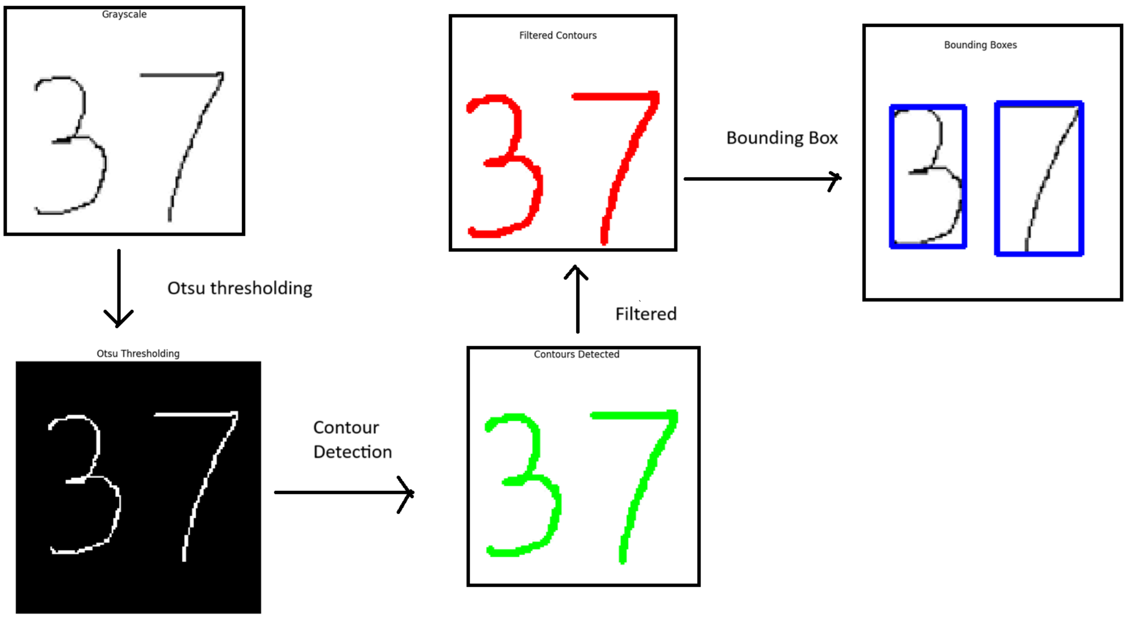

Fig 2: Steps included in the segmentation of image

Fig 2 shows the different stages involved in image segmentation. The image is converted to grayscale, after which Otsu thresholding is used on the image. On the binary imageContours are detected and later further refined. In the final stage Bounding box forms separating the two digits

In the two figures presented, the image processing and segmentation techniques are detailed, illustrating the steps taken to extract features and classify objects. Figure 1 demonstrates the process starting with the Radon Transform, which generates a sinogram from the input image. This sinogram is then processed using the Cumulative Distribution Function (CDF), facilitating feature extraction that is crucial for the training and testing of the Nearest Subspace Classifier. The application of these transformations allows the model to identify important patterns for classification.

Figure 2 outlines the image segmentation process, where the image is first converted to grayscale to simplify the processing. Subsequently, Otsu thresholding is applied to segment the image into a binary format. Contours are detected on the binary image, helping to highlight the edges of the objects. In the final step, a Bounding Box is formed, effectively separating the two digits, allowing for precise object classification in the subsequent analysis stages. Together, these processes enable efficient and accurate image classification.

The images are further preprocessed to fit the trained models in accordance with the respective dataset used for training. For example, with the MNIST dataset-trained model, we had to resize the sample images to work with the model. MNIST images are also white digits on a black background, meaning black digits on white background test images had to be color inverted to match the training images if the images were already segmented.

The images inside the bounding box are isolated and treated as separate images. They are fed to RCDT-NS for classification. This allows us to overcome the limitation RCDT had. It had failed to classify correctly when images had more than one object. The methodology we proposed allows RCDT to make the right prediction with any number of objects in the input image.

3.2 Hardware and/or Software Components

TABLE 1. A SAMPLE SOFTWARE/HARDWARE TOOLS TABLE

Tool	Functions	Why selected this tool
Colab	**Cloud Computing**	**More computational resources**
Pycharm	**IDE**	**Lightweight**

Table 1 shows a list of software tools used to develop and implement the project. Each tool is selected based on its specific functionality and advantages that contribute to the project's success. The table highlights the primary functions of each tool and the reasons for their selection, such as the need for additional computational resources or a lightweight development environment.

3.3 Hardware and/or Software Implementation

The software system was implemented in Python, using a combination of Google Colab for GPU-accelerated training and PyCharm for local development and testing. The pipeline comprises two primary stages: image segmentation and image classification.

For the segmentation stage, we utilized Otsu's thresholding method from the OpenCV library (cv2) to perform automatic binary segmentation on grayscale images. This was followed by contour detection and bounding box extraction, ensuring that individual objects—such as handwritten characters—were isolated before classification. These preprocessing steps were crucial for datasets containing multiple elements within a single image.

In the classification stage, we implemented a model based on the Radon Cumulative Distribution Transform-Nearest Subspace (RCDT-NS) framework. The Radon Transform (RT) was used to extract directionally informative features from the segmented images, and the Cumulative Distribution Transform (CDT)—performed using the pytranskit library—was applied to enhance class separability. These transformed representations were then passed into a Nearest Subspace Classifier (NSC), which assigned class labels based on the minimum projection distance to learned subspaces. The model is notable for being non-iterative, computationally efficient, and hyperparameter-free.

To benchmark the performance of our approach, we conducted an ablation study comparing RCDT-NS with three convolutional neural network architectures: ResNet-18, DenseNet-121, and a custom Shallow-CNN. These networks were implemented using PyTorch and trained under similar conditions to ensure a fair comparison in terms of accuracy, training time, and computational requirements.

Supporting libraries such as NumPy, scikit-image, and matplotlib were used for data handling, visualization, and additional preprocessing. This hybrid implementation structure enabled us to balance performance and efficiency, validating the robustness of our method across multiple datasets, including MNIST, EMNIST, EMNIST Letters, and Bangla Digits.

Chapter 4 Investigation/Experiment, Result, Analysis and Discussion

The purpose of this research project is to showcase the superior performance of the proposed approach as compared to conventional neural networks. We carry out various analyses to demonstrate how the RCDT-NS model exhibits comparable or even better classification results while using less computational power, time, memory, and resources. For a more generalized analysis, we decided to test the model's performance on 7 different datasets, followed by a direct comparison between the proposed model and 3 state-of-the-art CNN models- resnet-18, shallow-CNN and DenseNet-121, each being vastly different from each other to show how each performs against the others.

Table 2 : Datasets used in testing

Dataset Name	**Image Size**	**No. of Classes**	**No. of Training Images**	**No. of Test Images**
MNIST	28x28	10	60,000	10,000
EMNIST	28x28	47	697,932	116,323
Roman Number	28x28	10	7,000	1,000
Chest X-Ray	1200x824	2 (Normal, Pneumonia)	5216	624
Fingerprint	160x160	203	5,800	1200
EMNIST Letters	28x28	26	88,800	14,800
Bangla Digit	180x180	10	12,492	3,128

Table 2 describes the datasets tested with the RCDT-NS model. The datasets cover a wide array of classification problems that can potentially be handled by the model to demonstrate its strengths and weaknesses. The radoncdt setting was set to (0,176,45) which starts at the positive x-axis and measures 45 angles up to 176 degrees (close to the negative x-axis).

MNIST is a well-established benchmark dataset consisting of 28x28 grayscale images of handwritten digits ranging from 0 to 9. It contains 60,000 training images and 10,000 test images. Due to its simplicity, MNIST is widely used in the early stages of image classification research and algorithm development, serving as a baseline for evaluating new models and techniques. EMNIST (Extended MNIST) expands upon the original MNIST dataset by including both digits and a comprehensive set of handwritten uppercase and lowercase letters. The version used in this work includes 47 balanced classes and a total of 697,932 training images with 116,323 test images. EMNIST provides a more challenging and realistic benchmark for handwritten character recognition compared to MNIST. Roman Number is a custom dataset of handwritten Roman numerals ranging from I to X, formatted as 28x28 grayscale images. It includes 7,000 training and 1,000 test images across 10 classes. This dataset is used for the recognition of Roman numeral representations and serves as a unique case for testing classification models on non-standard numerical systems. Chest X-Ray is a medical imaging dataset that consists of 1200x824 resolution grayscale X-ray images categorized into two classes: Normal and Pneumonia. It includes 5,216 training images and 624 test images. This binary classification task is clinically significant, as it focuses on identifying signs of pneumonia from chest radiographs, making it useful for evaluating the model's performance in real-world diagnostic scenarios. Fingerprint is a biometric dataset comprising 160x160 grayscale fingerprint images from 203 unique individuals. It contains 5,800 training images and 1,200 test images, with one class per individual. This dataset presents a multi-class classification challenge with a large number of categories, making it suitable for evaluating fine-grained identity recognition systems. EMNIST Letters is a subset of the EMNIST dataset, containing only handwritten English letters (A–Z), with 26 classes. The dataset includes 88,800 training images and 14,800 test images, each resized to 28x28 pixels. This dataset is particularly useful for evaluating classification models on alphabet recognition tasks, with increased difficulty due to the similarity between certain letter shapes. Bangla Digit is a dataset of handwritten Bengali digits (০–৯), represented in 180x180 grayscale images. It includes 12,492 training images and 3,128 test images across 10 classes. This dataset is important for regional language processing and digit recognition tasks in Bengali, our native language.

Table 3 : Accuracy of each tested dataset with the proposed RCDT-NS model

Dataset	**Accuracy(%)**
MNIST Digit	95.34
EMNIST	95.41
Roman Number	5.8
Chest X-Ray	8.0
Fingerprint	30.12
EMNIST Letters	81.85
Bangla Digits	88.84

The performance of our proposed Radon-CDT-NSC model was evaluated across a diverse set of datasets, as shown in table 3, demonstrating its adaptability to different image classification tasks.

On standard benchmarks like MNIST Digits and EMNIST, the model achieved high accuracies of 95.34% and 95.41%, respectively, showcasing its strength in recognizing handwritten characters and digits. The model also performed well on more complex scripts such as Bangla Digits, achieving 88.84%, and on EMNIST Letters, with 81.85%, despite the high inter-class similarity in letter shapes. On the Fingerprint dataset, which contains 203 identity classes, the model reached 30.12% accuracy—reasonable given the high number of classes and the subtle inter-class differences. For the Chest X-Ray dataset, which involves binary classification between pneumonia and normal cases, the model achieved 8.0% accuracy, indicating limitations in processing high-resolution natural medical images without task-specific preprocessing or feature extraction. Similarly, performance on the Roman Number dataset was 5.8%, possibly due to limited training data and high intra-class variation in handwritten Roman numerals. These results suggest that while the R-CDT-NSC model is highly effective on structured and clean datasets like MNIST, its performance degrades on complex, high-resolution, or data-scarce domains, highlighting potential areas for future improvement such as integrating adaptive preprocessing, attention mechanisms, or hybrid learning strategies.

Moving on to the next phase of our experimentation, using the MNIST dataset as the standard, we provide an ablation study comparing RCDT-NS to our 3 chosen CNN models. Each of the models was trained using 10% of the dataset for 10 epochs and a 0.0001 learning rate. The radoncdt setting was set to (0, 176, 45), which starts at the positive x-axis and measures 45 angles up to 176 degrees (close to the negative x-axis).

ResNet-18 is a moderately deep convolutional neural network with 18 layers and approximately 11.7 million parameters. It utilizes residual connections that allow gradients to bypass certain layers, effectively addressing the vanishing gradient problem and enabling deeper networks to be trained more efficiently. With a computational requirement of roughly 1.8 billion FLOPs for a 224×224 image, ResNet-18 offers a balanced trade-off between performance and efficiency. Its architecture is well-suited for moderately complex tasks and datasets, often delivering strong classification performance. However, due to its depth and parameter count, it is still more computationally intensive than simpler architectures and may be unsuitable for low-resource environments or small datasets, where overfitting can become a concern.

Shallow CNN refers to a convolutional neural network with a small number of layers—typically between three and five—resulting in significantly fewer parameters, usually less than one million. Its computational footprint is minimal, making it highly efficient in terms of both memory and processing power, often requiring fewer than 200 million FLOPs. This simplicity allows for rapid training and inference, making it an excellent choice for applications on limited hardware or for datasets with low variability, such as MNIST. However, shallow CNNs lack the hierarchical feature extraction capabilities of deeper networks, leading to poorer performance on complex or high-resolution datasets. Their limited capacity can result in underfitting when faced with tasks requiring detailed pattern recognition.

DenseNet-121 is a deep convolutional neural network that incorporates dense connections between layers, ensuring each layer receives input from all previous layers. This results in more efficient feature reuse and improved gradient flow, helping the network generalize better with fewer parameters—about 7.98 million in total. Despite having fewer parameters than ResNet-18, DenseNet-121 is computationally more expensive, requiring around 2.9 billion FLOPs per 224×224 image, and is more memory-intensive due to the concatenation of feature maps across layers. DenseNet-121 excels in complex classification tasks and often achieves superior accuracy with fewer resources compared to other deep models. However, its dense connectivity can lead to increased memory usage during training and slower inference, making it less ideal for deployment on memory-constrained devices.

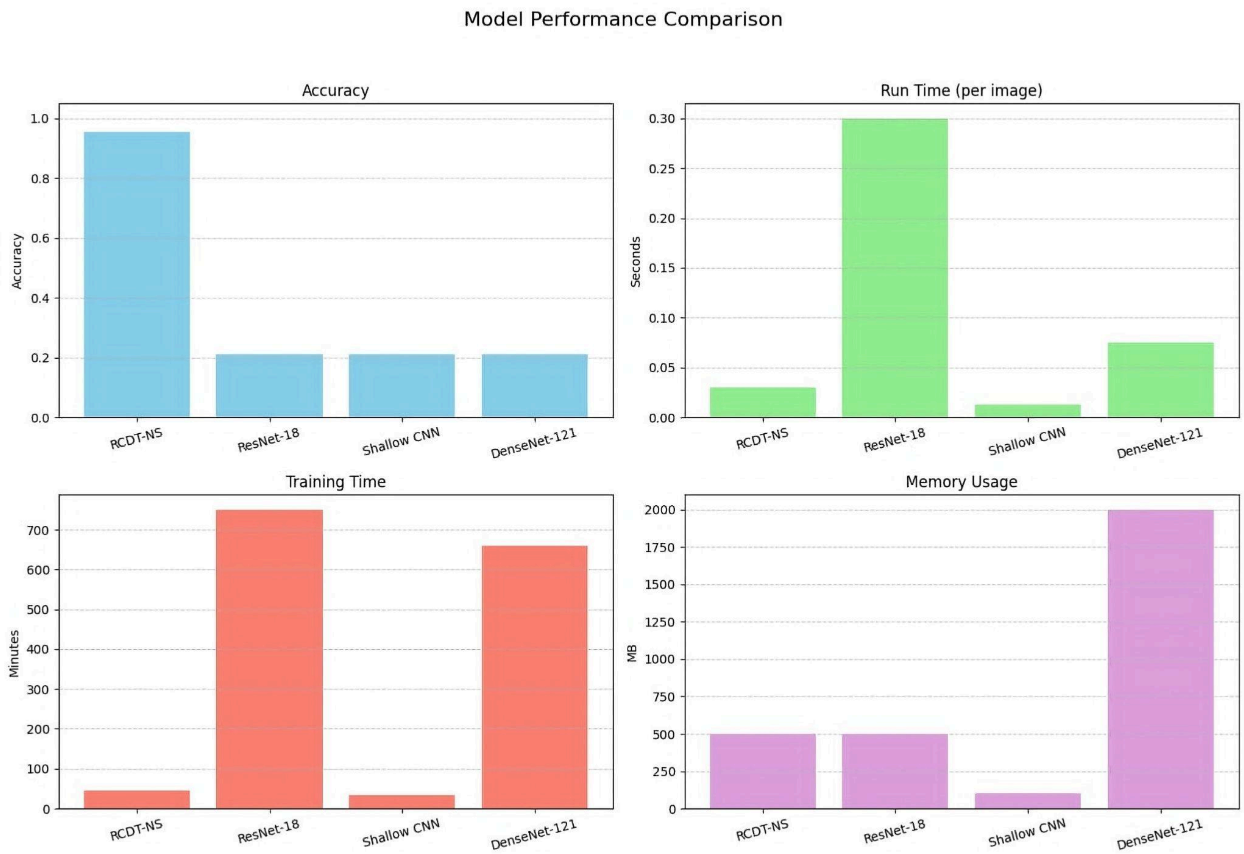

Fig 3: Performance Metrics of Compared Models for MNIST dataset

The figure 3 shows a comparative analysis of four models—RCDT-MS, ResNet-18, Shallow CNN, and DenseNet-121—evaluated on the MNIST dataset using four performance metrics: Accuracy, Run Time (per image), Training Time, and Memory Usage.

From the results in figure 3, it is clear that our proposed model performs the best among the 4 systems. The results fall in line with expectations. The Accuracy chart clearly shows that RCDT-NS significantly outperforms the other models, achieving close to 0.95 (95.0%) accuracy, whereas ResNet-18, Shallow CNN, and DenseNet-121 all hover around the 0.2 (20%) mark.

This stark contrast highlights the effectiveness of the Radon-CDT-NSC pipeline on the evaluated dataset, particularly for structured or well-preprocessed image data.

In terms of inference speed, the Shallow CNN is the fastest, followed closely by RCDT-NS, both taking under 0.05 seconds per image. In contrast, ResNet-18 exhibits the slowest runtime, taking around 0.3 seconds per image, while DenseNet-121 is moderately slower than RCDT-NS, landing just above 0.1 seconds per image. This indicates that RCDT-NS, despite its high accuracy, maintains efficient inference performance, making it suitable for real-time or low-latency applications.

The Training Time chart reveals a major advantage of the RCDT-NS model—it requires the least training time, well under 100 minutes. In comparison, ResNet-18 takes over 700 minutes, followed closely by DenseNet-121 at around 650 minutes, indicating a significant training cost. Shallow CNN is also fast to train but still slower than RCDT-NS. This result highlights RCDT-NS's efficiency, which stems from the fact that it does not require gradient-based backpropagation training typical of deep learning models.

For memory consumption, DenseNet-121 is the most demanding, requiring nearly 2000 MB, largely due to its dense connectivity and large intermediate feature maps. ResNet-18 and RCDT-NS have comparable memory footprints, each using about 500 MB, while Shallow CNN is the most lightweight in this aspect, using less than 200 MB. These findings suggest that while RCDT-NS is not the absolute lowest in memory, it achieves a good balance between memory usage and accuracy.

Overall, these comparisons show that RCDT-NS outperforms traditional deep learning models in terms of accuracy, training time, and run-time efficiency, while also maintaining a reasonable memory footprint. In contrast, while ResNet-18 and DenseNet-121 are powerful models for general-purpose image classification, they incur significant training costs and slower inference with limited performance on the evaluated dataset. Shallow CNN is efficient in terms of training and memory but falls short in accuracy. These results underscore the robustness and practicality of the RCDT-NS model, especially in resource-constrained or real-time environments.

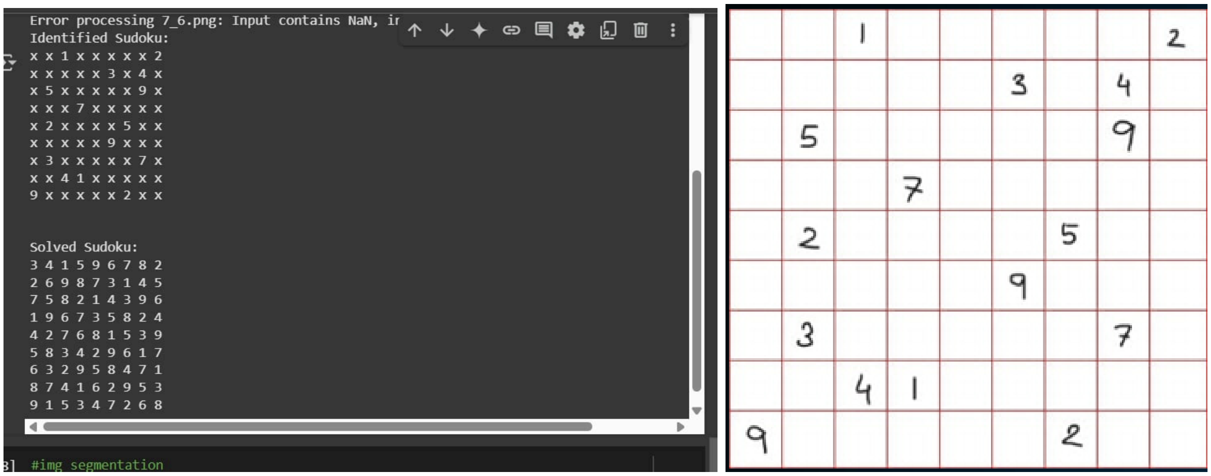

Fig 4: Sudoku solver on scanned handwritten image

Fig 4 shows our demonstration of a practical application of the RCDT-NS model, developed to highlight its capabilities through a system that takes in scanned images of a sudoku board with handwritten numerical digits written and outputs a solved sudoku board.

A sudoku grid will have physically handwritten digits on a white page that will be scanned and input into the system. The system segments the grid cells and separates them into individual images that will be preprocessed and passed into the prediction model. The model detects the digits and places them on a virtual sudoku grid, solves it, and displays the completed board. The above image shows how our system accurately recognizes the digits and uses them in a practical scenario.

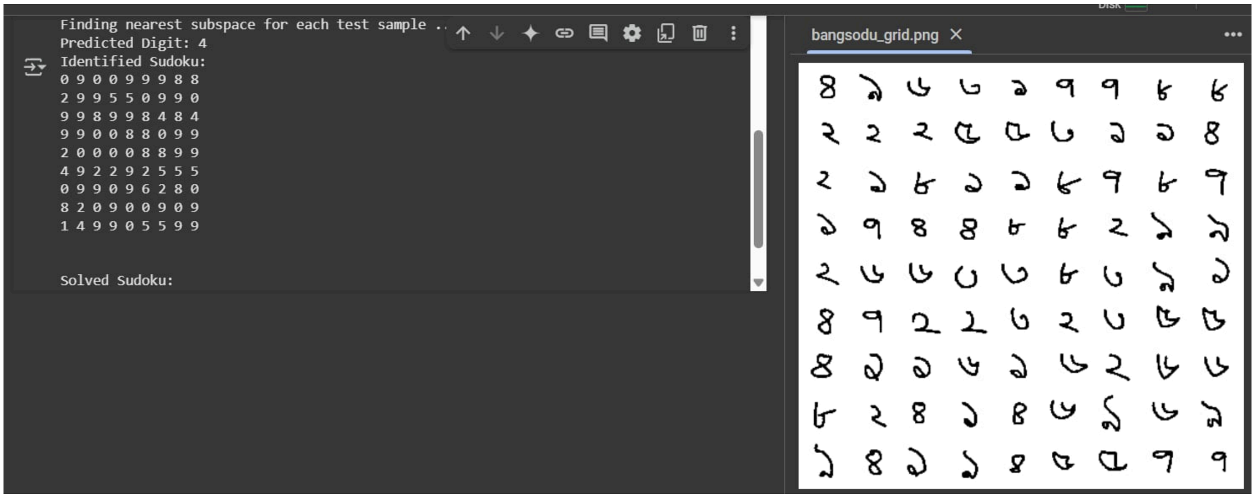

Fig 5: Sudoku solver on scanned handwritten image in bangla

Figure 5 shows results of RCDT using the Bangla letter dataset in the Sudoku system after inputting the scanned image of a Sudoku grid with handwritten Bangla digits.

We attempted to replicate the same system for a Bangla digit Sudoku using the same pipeline as can be seen in fig 5. However, the dataset used to train the model was not suited for our implementation. The dataset had too much noise, washed backgrounds, inconsistent images, and too few variations of each digit to account for the various kinds of handwriting in reality, despite having a respectable accuracy performance in the training and testing within the dataset. In particular, 1, 4, 3, and 7 showed the most fluctuations in predictions.

Chapter 5 Impacts of the Project

5.1 Impact of this project on societal, health, safety, legal and cultural issues

This project has significant potential to positively impact societal, health, safety, legal, and cultural aspects. By enabling some disease detection with minimal resources, it enhances healthcare accessibility in low-resource and remote areas, improving early detection and reducing the burden on overburdened healthcare systems. This could lead to better health outcomes, especially in regions with limited access to medical professionals or technology. The simplicity and transparency of the system minimize legal and ethical concerns, ensuring more trust and understanding in its use. Additionally, the project addresses cultural and societal issues by providing a solution that can be adapted to different regions, supporting cultural diversity and reducing stigma surrounding diseases. Ultimately, the project promotes health equity, safety, and social well-being, offering a valuable tool for improving public health, particularly in underserved communities.

5.2 Impact of this project on environment and sustainability

This project has notable implications for the environment and sustainability, especially in areas with limited resources. By enabling disease detection with minimal resources, it reduces the reliance on complex and resource-heavy healthcare systems, thereby minimizing energy consumption and electronic waste. The use of lightweight, low-energy solutions ensures that the project is both environmentally friendly and sustainable, as it avoids the environmental burden associated with large-scale medical equipment and deep learning systems. Additionally, by improving early disease detection and reducing the need for frequent and costly medical interventions, the project can contribute to the long-term sustainability of healthcare systems, ensuring they can better manage resources and provide services to a wider population. This focus on efficiency and minimal resource usage aligns with sustainability goals, reducing the ecological footprint of healthcare delivery in remote areas while improving health outcomes.

Chapter 6 Project Planning and Budget

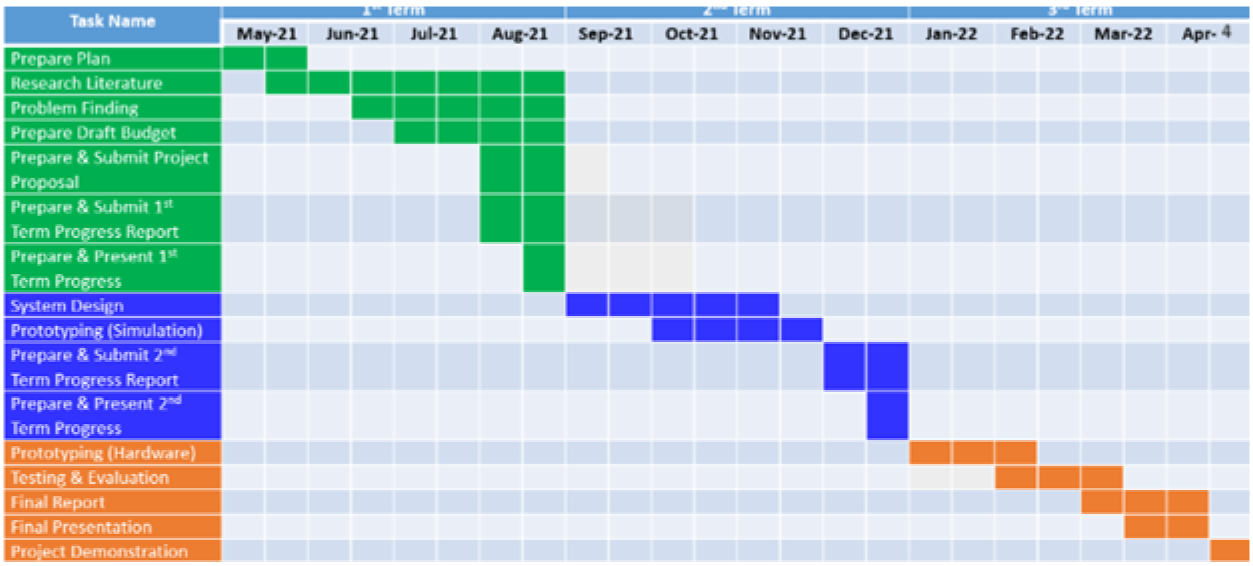

Fig 6. Gantt chart.

Figure 6 shows a sample Gantt chart that illustrates the project timeline and key milestones. The chart provides a visual representation of the project's schedule, highlighting the start and end dates for each major task and phase.

Table 4. Budget table

Component	Unit Price (USD)	Quantity	Total Cost (USD)	Dimensions	Weight
High-performance Laptop	$1,200	1	$1,200	14" x 9" x 1.5"	5 lbs
GPU for Training (e.g., NVIDIA RTX 3060)	$500	1	$500	10.5" x 4.5" x 2"	1.5 lbs
External Storage (1TB SSD)	$80	1	$80	3" x 4" x 0.5"	0.2 lbs
Pre-trained Models for R-CDT	Free	1	Free	N/A	N/A
Cloud Services for Model Training	$20	1	$20	N/A	N/A
Miscellaneous Printing/Documentation	$20	1	$20	N/A	N/A
TOTAL			$1,820		6.7lbs

Table 4 shows a detailed breakdown of the project's budget, outlining the costs associated with each phase and component of the project. It includes expenses for hardware, software, tools, materials, and any other resources required for the successful completion of the project.

Chapter 7 Complex Engineering Problems and Activities

7.1 Complex Engineering Problems (CEP)

TABLE 5. COMPLEX ENGINEERING PROBLEM ATTRIBUTES TABLE

Attributes		Addressing the complex engineering problems (P) in the project
P1	Depth of knowledge required (K3-K8)	The project requires knowledge of image processing (K3), Radon/Cumulative Distribution, Statistics, Linear Algebra (K4), Machine Learning (K5), Python Tools (K6), Scientific Research Papers (K8).
P2	Range of conflicting requirements	In the model choice, accuracy vs speed, size vs performance, efficiency vs complexity
P3	Depth of analysis required	Among the many possible image processing machine learning models, we chose the proposed RCDT–NS model over the other available CNN models.
P4	Familiarity of issues	The proposed method is novel and relatively not mainstream compared to CNN models.
P5	Extent of applicable codes	The proposed solution was not based on conventionally available code.
P6	Extent of stakeholder involvement	Stakeholder involvement is minimal unless applied practically in a particular discipline/industry.
P7	Interdependence	The project involves a number of interdependent sections, such as preprocessing, transformation, classification, and visualization, that are all interconnected.

Table 5 shows the various attributes associated with the complex engineering problems (CEPs) encountered during the project. It highlights the different factors that contribute to the complexity, such as the technical challenges, resource requirements, and time constraints.

7.2 Complex Engineering Activities (CEA)

TABLE 6. COMPLEX ENGINEERING PROBLEM ACTIVITIES TABLE

Attributes		Addressing the complex engineering activities (A) in the project
A1	Range of resources	This project involves Python math, graphical analysis libraries, machine learning models, image datasets, and a camera for scanning.
A2	Level of interactions	Involves interactions between different stakeholders, including group members, to design, develop, test, and debug the model and in model evaluation.
A3	Innovation	RCDT-NS is a novel approach that is lightweight, efficient, effective, and underused.
A4	Consequences to society / Environment	Low-power, fast models can be developed for use in implementations where larger models are not suitable, reducing energy use and increasing access.
A5	Familiarity	Learned new concepts regarding radon transform and CDT. Supports SDG #9: Industry, Innovation and Infrastructure

Table 6 shows the various activities related to the complex engineering problems (CEAs) encountered throughout the project. It outlines the specific engineering tasks performed to address these problems, detailing the steps involved in finding solutions.

Chapter 8 Conclusions

8.1 Summary

This project focuses on non-segmented image classification using Radon Cumulative Distribution Transform (R-CDT) and Nearest Subspace Classifier (NSC). The goal is to improve image classification by reducing computational complexity and time compared to traditional deep learning methods, such as Convolutional Neural Networks (CNNs). The project uses R-CDT for feature extraction and NSC for classification. Additionally, image segmentation techniques like Otsu's method are applied to handle multiple objects within an image. The project compares accuracy and performance (runtime, memory, etc.) between traditional CNN models and the proposed R-CDT-based approach.

8.2 Limitations

The R-CDT and NSC approach may face limitations in accuracy when dealing with highly complex or varied images, as compared to deep learning models like CNNs. Its performance is heavily reliant on the accuracy of the image segmentation step, and incorrect segmentation can lead to misclassification or reduced accuracy. Additionally, the method may struggle with scalability when applied to very large image datasets, facing challenges in both computation time and memory usage. Furthermore, the R-CDT approach primarily focuses on geometric features derived from the Radon transform, potentially missing out on capturing texture or color information as effectively as deep learning models attributed to a limited feature set.

8.3 Future Improvements

Hybrid Models: Future improvements could involve hybridizing the R-CDT-based approach with deep learning techniques, leveraging both traditional and neural network-based features to improve classification accuracy.

Real-time Segmentation: Integrating faster and more efficient segmentation algorithms, such as those based on deep learning (e.g., U-Net), could lead to better performance on a wider range of images.

Extension to More Complex Datasets: Exploring the use of R-CDT and NSC on more diverse datasets, including those with complex backgrounds, dynamic scenes, and higher resolutions, would test the robustness and versatility of the approach.

Automating Feature Selection: Implementing automated feature selection techniques could help improve the performance of the NSC by selecting the most relevant features from R-CDT, rather than using all extracted features.

References

1. M. Shifat-E-Rabbi *et al.*, "Radon Cumulative Distribution Transform Subspace Modeling for Image Classification," *Journal of mathematical imaging and vision*, vol. 63, no. 9, pp. 1185–1203, Aug. 2021, doi: https://doi.org/10.1007/s10851-021-01052-0.

2. "A biological image classification method based on improved CNN," *Ecological Informatics*, vol. 58, p. 101093, Jul. 2020, doi: https://doi.org/10.1016/j.ecoinf.2020.101093.

3. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *www.cv-foundation.org*, 2016. https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html

4. G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *openaccess.thecvf.com*, 2017. https://openaccess.thecvf.com/content_cvpr_2017/html/Huang_Densely_Connected_Convolutional_CVPR_2017_paper.html

5. B. Shah and H. Bhavsar, "Time Complexity in Deep Learning Models," *Procedia Computer Science*, vol. 215, pp. 202–210, 2022, doi: https://doi.org/10.1016/j.procs.2022.12.023.

6. S. Ahmed, M. Jahan, N. S. Pathan, and Q. D. Hossain, "Classification of Breast Tumor Using Radon Cumulative Distribution Transform Nearest Subspace Classifier," pp. 01–06, Feb. 2023, doi: https://doi.org/10.1109/ecce57851.2023.10101642.

7. M. T. Nyo, F. Mebarek-Oudina, S. S. Hlaing, and N. A. Khan, "Otsu's thresholding technique for MRI image brain tumor segmentation," *Multimedia Tools and Applications*, May 2022, doi: https://doi.org/10.1007/s11042-022-13215-1.