Electronic Theses and Dissertations

Theses, Dissertations, and Major Papers

5-28-2024

An Ensemble Deep Learning Approach for Enhanced Classification: A Case Study on Pituitary Tumors

Sumaiya Deen Muhammad
University of Windsor

An Ensemble Deep Learning Approach for Enhanced Classification: A Case Study on Pituitary Tumors

By

Sumaiya Deen Muhammad

A Thesis
Submitted to the Faculty of Graduate Studies
through the School of Computer Science
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2024

An Ensemble Deep Learning Approach for Enhanced Classification: A Case Study
on Pituitary Tumors

by

Sumaiya Deen Muhammad

APPROVED BY:

E. Abdel-Raheem
Department of Electrical and Computer Engineering

B. Boufama
School of Computer Science

Z. Kobti, Advisor
School of Computer Science

May 9, 2024

DECLARATION OF CO-AUTHORSHIP AND PREVIOUS PUBLICATION

I. Co-Authorship

I hereby declare that this thesis integrates findings derived from research carried out under the guidance of Dr. Ziad Kobti. Throughout the process, the author undertook the pivotal tasks, including formulating key concepts, making primary contributions, designing experiments, conducting data analysis, and interpreting the results. Although co-authors have played a crucial role in proofreading the published manuscripts and offering valuable suggestions, the core of the research is solely the outcome of my diligent endeavors.

I am aware of the University of Windsor Senate Policy on Authorship and I certify that I have properly acknowledged the contribution of other researchers to my thesis, and have obtained written permission from each of the co-author(s) to include the above material(s) in my thesis.

I certify that, with the above qualification, this thesis, and the research to which it refers, is the product of my own work.

II. Previous Publication

This thesis includes an original paper that have been previously published in a peer-reviewed conference, as follows:

Section	Full Citation	Publication Status
Full Paper	S. D. Muhammad and Z. Kobti, "An Ensemble Deep Learning Approach for Enhanced Classification of Pituitary Tumors," 2023 IEEE Symposium Series on Computational Intelligence (SSCI), Mexico City, Mexico, 2023, pp. 427-432, doi: 10.1109/SSCI52147.2023.10371824.	Published

I certify that I have obtained a written permission from the copyright owner(s)

to include the above published material(s) in my thesis. I certify that the above material describes work completed during my registration as a graduate student at the University of Windsor

III. General

I declare that, to the best of my knowledge, my thesis does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my thesis, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my thesis.

I declare that this is a true copy of my thesis, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this thesis has not been submitted for a higher degree to any other University or Institution.

ABSTRACT

The Segment Anything Model (SAM) by Meta AI Research, trained on an extensive collection of over 1 billion masks, has gained significant attention for its exceptional ability to segment "anything" in "any scene". SAM integrates a sophisticated image encoder, prompt encoder, and lightweight mask decoder, enabling flexible prompting and rapid mask generation in segmentation tasks. This segmentation model excels in granular, component-level segmentation, enriching our understanding of pixel semantics, critical for local feature learning. On a different note, the challenge of classifying small-scale objects persists, especially in sectors like medical imaging and remote sensing where objects of interest typically represent a small fraction of the entire image. In this study, we investigate the potential applications of SAM in the classification of small objects despite its primary design as a segmentation model. We introduce an ensemble deep learning methodology that leverages SAM within our custom dataset, specifically targeting the classification of tiny objects. Through comparative analysis between segmented data (processed by SAM) and non-segmented data (original data), our findings indicate a performance improvement in favor of the segmented data, underscoring the efficacy of our proposed approach.

DEDICATION

This thesis is dedicated to the unconditional support and love of my family, whose encouragement has been the foundation of my academic journey.

To my father, whose wisdom and guidance have been a beacon of inspiration throughout my life. Your enduring belief in my abilities has fueled my pursuit of knowledge.

To my mother, whose boundless love and sacrifices have been a source of strength. Your persistent support has been my foundation, enabling me to reach new heights.

To my elder sisters and younger brother, your constant encouragement and belief in my potential have been instrumental in shaping my academic endeavors. Your presence has been a comforting reminder that I am never alone in my pursuit of knowledge.

A special mention goes to my niece, Naila Nawar, whose youthful enthusiasm and unyielding belief in me have been a constant source of motivation. Your encouragement has been a driving force behind my academic achievements.

Last but not least, to my husband, Mohammed Kamruzzaman, whose consistent support, understanding, and encouragement have been my pillars of strength. Your belief in my capabilities has been a source of inspiration, making this academic journey more meaningful.

This thesis is a testament to the collective love, support, and encouragement of my family, without whom this achievement would not have been possible.

ACKNOWLEDGEMENTS

I extend my sincere gratitude to Dr. Ziad Kobti, my esteemed supervisor, whose indispensable guidance, expertise, and mentorship played a pivotal role in the successful completion of this thesis. His consistent support served as a driving force, inspiring me to persevere and achieve new milestones in my research journey. The constructive criticism and insightful feedback he provided not only facilitated my academic growth but also honed my research skills, fostering a profound understanding of the subject matter. I am truly honored to have had the privilege of learning from him, and I express my deep gratitude for his immeasurable contributions to my intellectual development.

I also extend my heartfelt thanks to my committee members, Dr. Boubakeur Boufama and Dr. Esam Abdel-Raheem, for their invaluable advice and feedback. Their expertise and rigorous evaluation significantly elevated the quality of my research, and I profoundly appreciate the time and effort they dedicated to this endeavor.

Lastly, I would like to humbly acknowledge the School of Computer Science and all individuals involved in assisting me throughout this process.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

SAM   Segment Anything Model

ViT   Vision Transformer

ViTDet   Vision-Transformer-Detector

MAE   Masked Autoencoder

MRI   Magnetic Resonance Imaging

CT Scan   Computed Tomography Scan

FCN   Fully Convolutional Network

DL   Deep Learning

ML   Machine Learning

CNN   Convolutional Neural Network

GAP   Global Average Pooling

MLP   Multi Layer Perceptron

GELU   Gaussian Error Linear Units

IoU   Intersection over Union

CHAPTER 1

Introduction

1.1   Background

The challenge of identifying small lesions such as tiny tumors in medical imaging is significantly exacerbated by their minimal pixel coverage within images, which often leads to indistinguishable features. These diminutive lesions lack the pronounced morphological and dynamic criteria usually relied upon to detect more conspicuous tumors, complicating both manual and automated segmentation. The difficulty is compounded in dynamic imaging scenarios like MRI (Magnetic Resonance Imaging), where the contrast between a small lesion and its environment may be insufficient for clear depiction. Additionally, the variability in the lesions' appearance—shape, size, and density—presents further complexities, as they often blend into surrounding tissue or are obscured by complex anatomical backgrounds, challenging traditional classification methods that depend on clear and consistent visual markers of pathology [9].

Pituitary tumors, categorized as the second most common type of primary brain tumor, embody a substantial challenge in neurodiagnostics due to their critical location and typically minute size [10]. Originating from the pituitary gland (Figure:1.1.1)—a crucial endocrine gland positioned at the base of the brain—these tumors constitute about 10-15% of all adult brain tumors [11]. According to diameter size, pituitary tumors have been categorized into three types: microadenomas (<10mm), macroadenomas (≥ 10mm), giant adenomas (≥ 40mm) [12]. The major-

ity of pituitary tumors are typically microadenomas, with a diameter ranging from 3 to 9 millimeters [13].
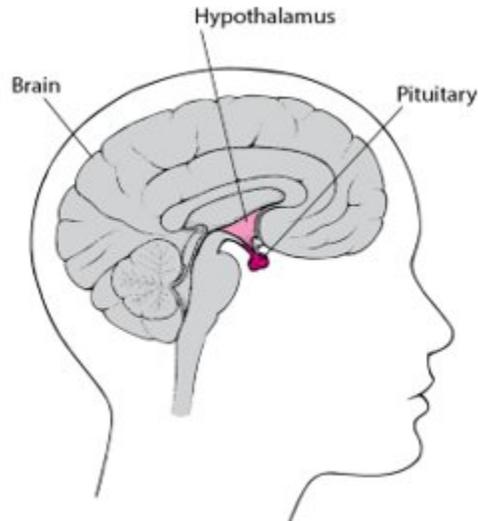
Fig. 1.1.1: Pituitary Gland [1]

Although most of the pituitary tumors are benign, known as pituitary adenomas, they can disrupt hormone production, leading to significant health issues such as diabetes, cardiovascular diseases, and fertility complications [11]. The detection and accurate classification of these tumors are hindered by their small size and the gland's position deep within the cranial cavity, with traditional diagnostic methods like MRI struggling to distinguish these tumors from surrounding brain tissue due to factors like tumor size variability, image resolution, and evaluator fatigue [14] [15].
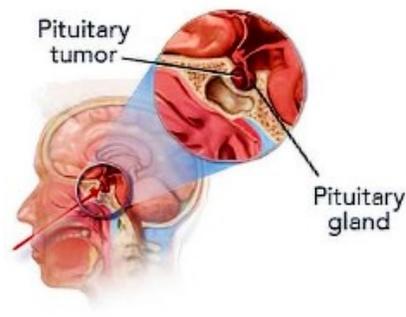
Fig. 1.1.2: Pituitary Tumor [2]

To address these diagnostic challenges, the advent of advanced deep learning

models, particularly SAM introduced by Meta AI in April 2023, marks a significant advancement in medical imaging [5]. SAM, developed on the largest segmentation dataset to date featuring over 1 billion masks, enables the creation of precise masks for a diverse range of objects across various scenes without prior training on new data. This thesis explores SAM's potential in enhancing the classification of pituitary tumors from MRI scans, aiming to significantly improve the accuracy and reliability of diagnostics by focusing on relevant image features and standardizing the appearance of objects of interest, thus potentially transforming the approach to detecting and managing this common yet complex type of brain tumor. The goal is to validate SAM's application in a real-world medical setting, assessing its efficiency against traditional diagnostic methods and its ability to contribute to more effective and timely medical interventions.

In the next section, we will discuss about image classification, segmentation, and SAM.

1.2   Image Classification

Image classification is a process in computer vision where an image is categorized into one of several predefined classes [16]. It involves categorizing and labeling groups of pixels or vectors within an image based on specific rules. The primary goal of image classification is to identify and portray, as a unique value, the features of an image for further analysis [17].

The evolution of image classification can be traced back to the late 20th century, where it started with relatively simple algorithms. With the advent of machine learning and neural networks, particularly Convolutional Neural Networks (CNNs), the field has witnessed a significant transformation. LeCun et al. [18] introduced CNNs, which revolutionized the way computers perceive and interpret images, leading to more accurate and efficient image classification.

Today, image classification leverages deep learning techniques to achieve unprecedented accuracies. Krizhevsky et al. [19] demonstrated the power of deep neural

networks with their groundbreaking work in the ImageNet Large Scale Visual Recognition Challenge, setting a new standard for image classification tasks.

These developments have found applications in various domains, from medical diagnosis [20] to autonomous vehicles, showcasing the versatility and importance of image classification in modern technology.

1.2.1 Formal Definition of Image Classification

Goal: Find a function $h: X{\rightarrow}Y$ that maps an input image $x \in X$ to a category label $y \in Y$, where X is the space of possible images and $Y=\{1,\ 2,\ ...,K\}$ is the set of category labels.
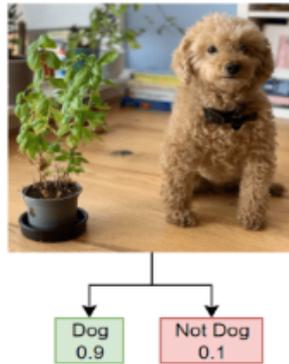
Fig. 1.2.1: Image Classification Example [3]

1.3 Image Segmentation

Image segmentation is a process that involves dividing an image into various segments or sections. This technique is commonly utilized to identify objects and outline boundaries within images [21]. It is a crucial process in computer vision that cleaves an image into segments or regions with similar attributes, aiming to simplify its representation. It is fundamental for tasks like object recognition, scene understanding, medical imaging, etc. Early methods relied on techniques such as thresholding, clustering, and edge detection. However, recent advancements leverage deep learning, notably Convolutional Neural Networks (CNNs), for more accurate and nuanced seg-

mentation. Deep learning-based approaches, such as Fully Convolutional Networks (FCN) and U-Net, have set new benchmarks in segmentation tasks by learning from vast amounts of data, offering precise and detailed segmentation outcomes [22] [23].

1.3.1 Formal Definition of Image Segmentation

Given: A digital image I represented as a two-dimensional array of pixels $I(x, y)$ where *(x, y)* are the coordinates of the pixels, and each pixel has a value corresponding to its intensity or colour.

Objective: To find a segmentation S of the image I, such that S is a set of regions $\{R_1, R_2, ..., R_n\}$ where each region R_i is a connected set of pixels with similar attributes (such as intensity, colour, texture), and n is the number of such regions.

Fig. 1.3.1: Image Segmentation Example [4]

1.4 Impact of Image Segmentation on Image Classification

The relationship between image segmentation and classification performance is a well-studied topic in the field of computer vision and medical imaging. In the realm of medical imaging, segmentation and classification can be closely related, and improvements in segmentation can lead to more accurate classification.

Segmentation allows for a hierarchical and multi-scale analysis of images, where classification can be performed at different levels of granularity. This is especially

important in medical imaging where abnormalities could be present at various scales.

- **Improved Interpretability and Reliability:** Segmentation provides a way to visualize and interpret the components of the image, which can be crucial for medical diagnosis. This interpretability ensures that the classification results are reliable and can be trusted by medical professionals.

- **Focus on Relevant Features**: Image segmentation allows for the isolation of important parts of an image, helping the classification algorithm to focus on the most relevant features. This is crucial in fields like medical imaging, where the area of interest may be a small part of the entire image.

- **Better Handling of Variability:** In cases where the objects of interest exhibit a lot of variability in size, shape, or appearance, segmentation can standardize these objects, making it easier for the classification algorithm to identify them.

1.5   Segment Anything Model (SAM)

SAM is a state-of-the-art approach in image segmentation that relies on a vast dataset and innovative architecture for prompt-based zero-shot transfer [24] across diverse image tasks. Supported by the SA-1B dataset, which includes over 1 billion masks from 11 million images, SAM showcases scalable and real-time interactive segmentation capabilities. This is achieved through a blend of an image encoder, prompt encoder, and mask decoder, enabling dynamic, high-quality segmentation tasks [5].
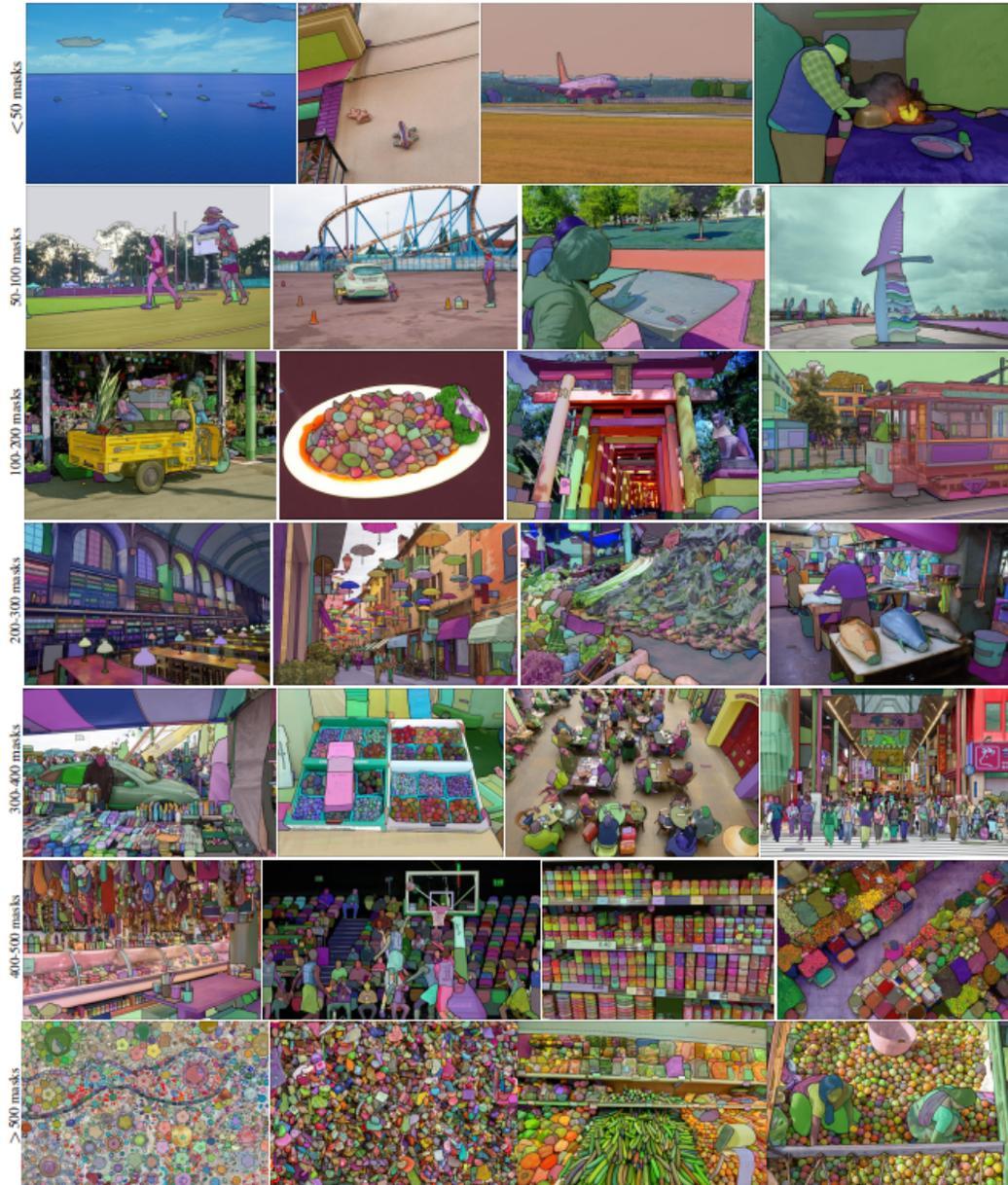
Fig. 1.5.1: Example images with overlaid masks from **SA-1B** dataset[5]

1.5.1   SAM Architecture Overview

The architecture of SAM is comprehensive and involves several key components designed for efficient and effective image segmentation. Here is a detailed breakdown of its architecture:
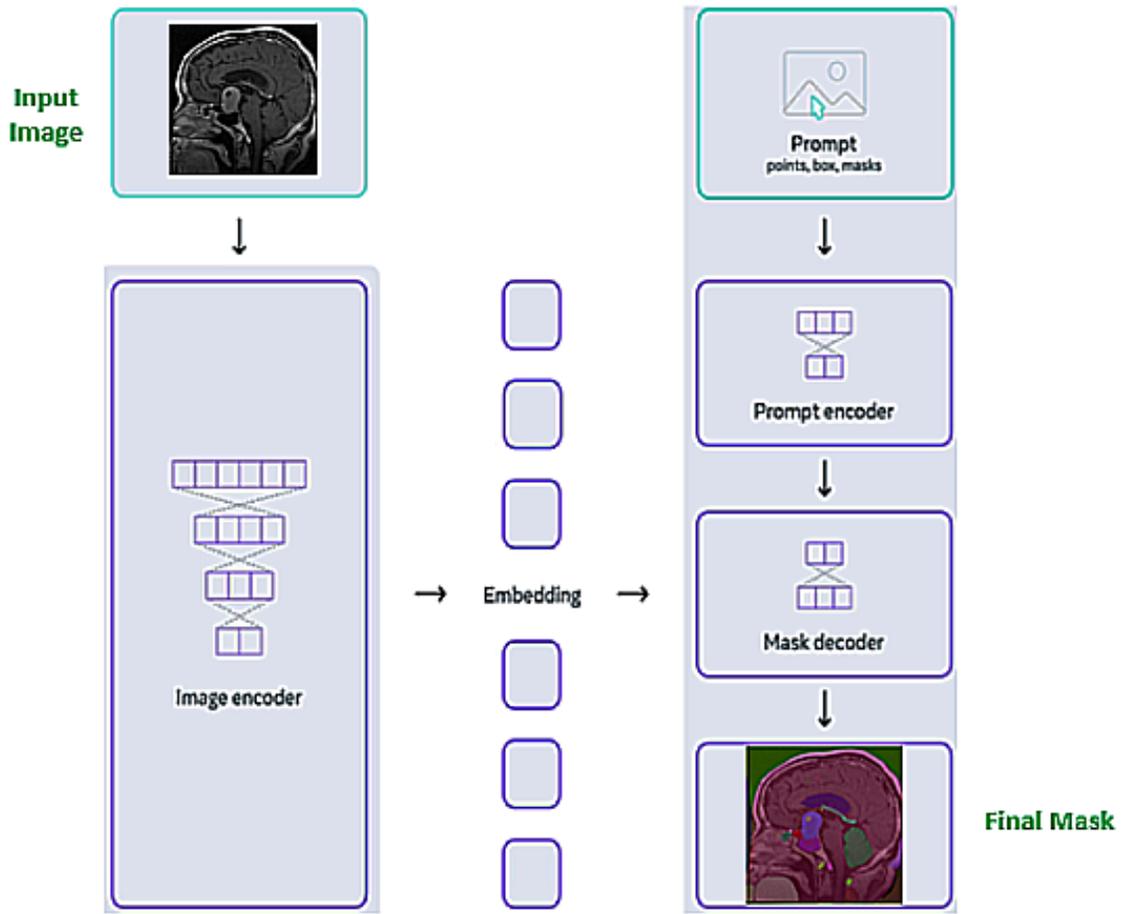
Fig. 1.5.2: SAM Architecture [6]

1.5.1.1 Image Encoder

The image encoder can be any network capable of outputting a C×H×W image embedding. SAM utilizes a Masked Autoencoder (MAE) pre-trained Vision Transformer (ViT) [25][26] with adaptations for high-resolution input processing. Specifically, a ViT-H/16 configuration is used with 14×14 windowed attention and four global attention blocks. The encoder outputs a 16× downscaled embedding of the input image. For input, images are rescaled to 1024×1024, and the embedding is processed to reduce the channel dimension to 256 through convolutions, followed by layer normalization [5].

1.5.1.2 Prompt Encoder

Sparse and Dense Prompts: The model maps sparse prompts (points or boxes) and dense prompts (masks) to vectorial embeddings. Points are represented by a combination of a positional encoding and learned embeddings indicating foreground or background. Boxes use embeddings for corners, and free-form text is processed using the CLIP text encoder [5].

Mask Processing: Masks are input at a lower resolution and further downscaled, with the processing including convolutional layers, GELU activations, and layer normalization to align with the 256-dimensional embedding space [5].

1.5.1.3 Lightweight Mask Decoder

SAM utilizes a modified Transformer decoder. The decoder integrates prompt embeddings with the image embedding through self-attention, cross-attention, and pointwise MLPs. The decoder updates both image and prompt tokens via cross-attention, upscales the image embedding, and dynamically predicts masks using the updated tokens. It efficiently handles ambiguous prompts by predicting multiple masks and uses a small head to estimate the Intersection over Union (IoU) for ranking predicted masks [5].
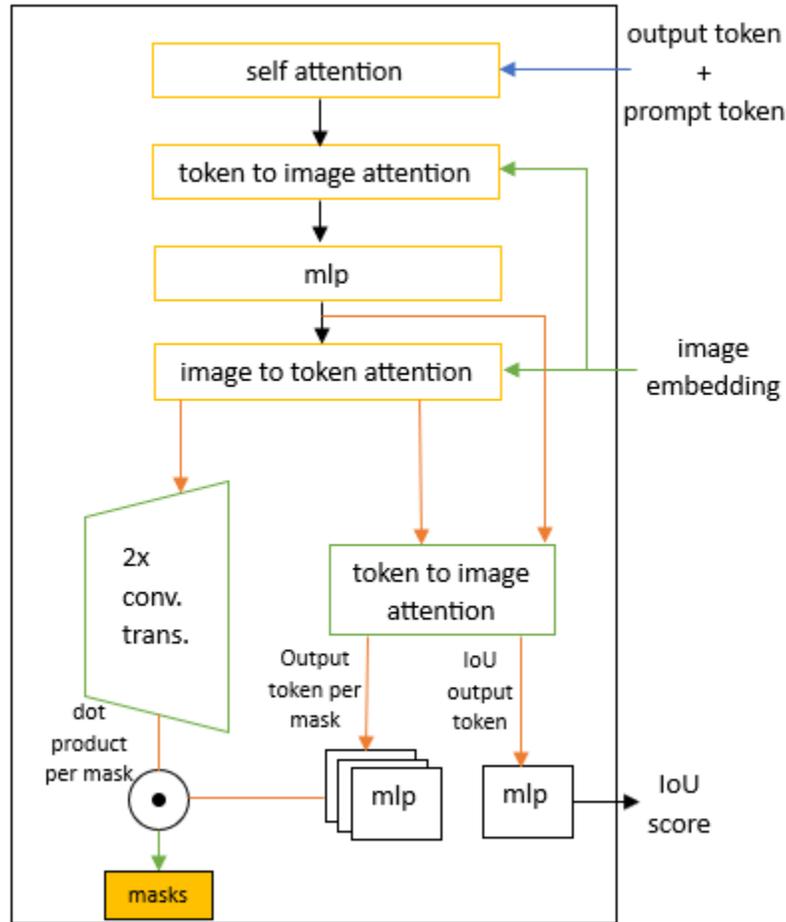
Fig. 1.5.3: Mask Decoder

1.5.2 Comparison Between SAM and Other Models

In the comparative analysis of segmentation models, SAM demonstrates notable advantages over established models such as ViTDet and U-Net [27][23] . SAM leverages a vast dataset to offer robust segmentation capabilities across diverse scenarios without the need for prior training on specific data. This flexibility is contrasted with ViTDet, a high-performance vision transformer-based object detection model, which still relies on conventional training datasets.
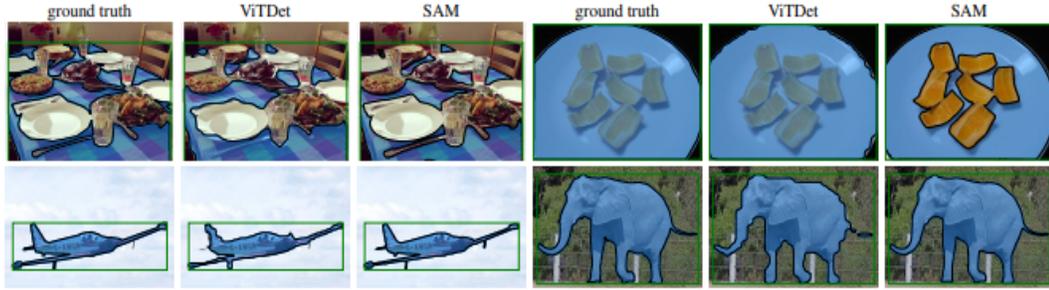
Fig. 1.5.4: Comparison between SAM and ViTDet [5]

Similarly, while U-Net has been a preferred choice for medical image segmentation, SAM's ability to generalize across different types of images and its zero-shot learning capabilities mark a significant improvement, particularly in handling varied and complex segmentation tasks more efficiently. [7] assesses the effectiveness of SAM and U-Net, in detecting cracks in civil infrastructure. SAM shows superior performance in detecting longitudinal cracks due to its ability to divide the image into various parts that help identify the location of the crack. This segmentation capability allows SAM to handle the variability in crack shapes effectively, which is essential for precise and comprehensive crack detection. On the other hand, while U-Net excels in identifying spalling cracks through positive label pixel detection, it may struggle with longitudinal cracks where the crack path is not as straightforward.
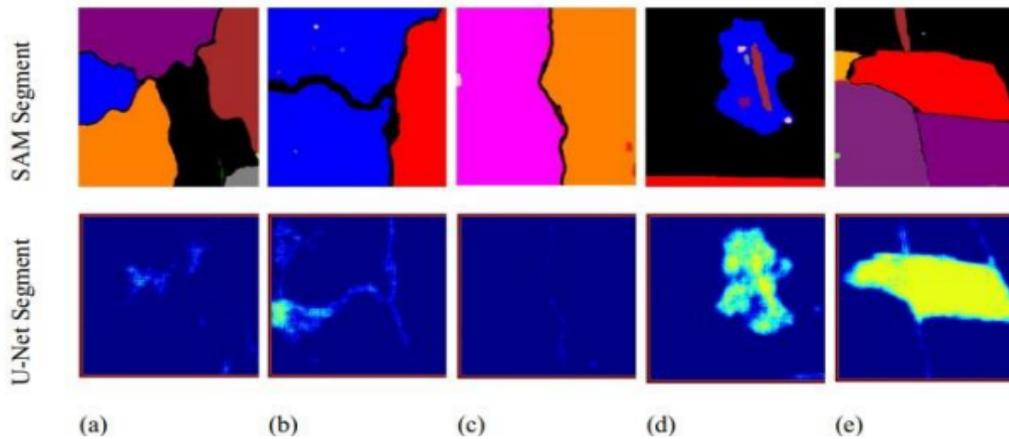
Fig. 1.5.5: Comparison between SAM and U-Net [7]

[8] shows SAM demonstrates superior performance over traditional semantic segmentation methods primarily through its advanced handling of local features and its

ability to generalize across diverse and complex visual scenes. SAM enhances local feature learning by incorporating detailed, fine-grained semantic information, which allows for more precise modeling of relationships between pixels. This capability is particularly useful in scenarios where semantic segmentation models falter due to their coarse-grained and category-specific nature. SAM, being trained on a massive dataset of 11 million images, offers robust zero-shot generalization capabilities that are not confined to predefined categories, enabling it to effectively handle a wide array of objects and scenes. This versatility and depth of understanding make SAM particularly adept at identifying and classifying features in images where traditional models may struggle due to limited scope and granularity.
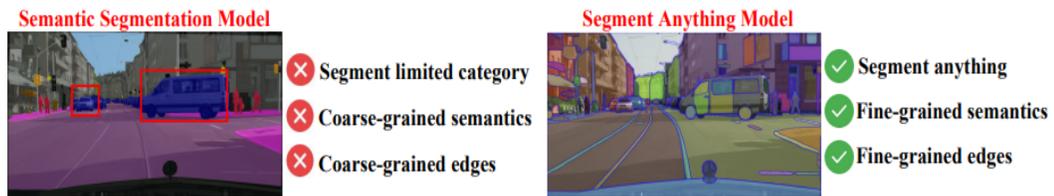
Fig. 1.5.6: Comparison between SAM and Semantic Segmentation [8]

These comparisons underline SAM's potential to enhance the accuracy and efficiency of medical image analysis, including the classification and segmentation of pituitary tumors, by providing more precise segmentation outputs than traditional models.

1.6 Problem Overview

SAM represents a novel approach in the realm of image segmentation, holding significant promise for the enhancement of small object classification. This innovative model introduces capabilities that could potentially revolutionize the accuracy with which small and often elusive objects are identified within various imaging contexts. Despite its advanced capabilities, the application of SAM in the specific context of classifying small objects has yet to be thoroughly investigated.

Research Hypothesis

This research is predicated on the hypothesis that leveraging SAM's inherent strengths for small object classification, when combined with strategic enhancements inspired by ensemble modeling techniques, can lead to substantial improvements in classification accuracy. By integrating SAM's robust segmentation capabilities with the collective intelligence of an ensemble model, this study develops a more precise and effective classification system. This could not only validate SAM's utility in a new application domain but also extend its operational framework to achieve superior performance outcomes in small object detection.

1.7 Motivation

The motivation for this thesis centers on the significant challenges in accurately classifying small objects, such as pituitary tumors, within images. Traditional classification methods often fall short when dealing with such small objects, leading to misclassification and potential diagnostic errors. This research is driven by the question of how well existing image segmentation and classification models perform on small objects, aiming to leverage SAM and an ensemble deep learning approach to enhance classification accuracy for small object detection in medical imaging, particularly pituitary tumors. This approach is expected to mitigate the limitations of current models, offering a more reliable and precise tool for medical diagnosis and treatment planning.

1.8 Thesis Statement

This thesis presents a comprehensive analysis of SAM in the context of small object classification, with a particular focus on addressing the prevalent challenges of precision and reliability in the domain of applied fields such as medical imaging. The core of this research lies in not only assessing the inherent capabilities of SAM but also in enhancing its classification performance through an innovative ensemble deep

learning technique tailored for small-sized entities. The culmination of this research is showcased through a case study on pituitary tumor classification from brain MRI scans—a critical endeavor, given the significance of accurate small object detection in healthcare applications.

1.9   Thesis Contributions

- **Evaluation of SAM's Baseline Performance:**

    – Conducted a thorough examination of SAM's ability to classify small objects, providing a nuanced understanding of its performance dynamics.

- **Development of an Enhanced Classification Technique:**

    – Presented an advanced ensemble methodology that synergizes multiple deep learning models with SAM, aimed at enhancing the performance of small object classification.

- **Comparative Analysis with Existing Methods:**

    – Performed a critical comparison of the enhanced SAM against established classification models, utilizing a range of performance metrics to ensure a comprehensive evaluation.

- **Real-world Application to Pituitary Tumor Classification:**

    – Applied the improved classification model to the domain of pituitary tumor identification in brain MRI scans, addressing a vital need within the medical field.

In summary, the thesis delineates a significant advancement in the field of small object detection, particularly in medical imaging, through the development and validation of a refined SAM-based model.

1.10 Thesis Outline

We have included six chapters for our thesis, which are organized in the following way:

- Chapter 2: In this chapter, we review existing literature on different deep learning based methods that have been implemented in order to classify different kind of brain tumors. We also explore the research works that have been conducted on SAM in different areas. And lastly, we study different ensemble deep learning approaches employed for different tasks.

- Chapter 3: This chapter provides a detailed explanation of our proposed methodology, including our custom dataset preparation.

- Chapter 4: This chapter focuses on the experimental setup, tuning hyperparameters for our models, and the evaluation metrics we consider for our experiment analysis.

- Chapter 5: In this chapter, we present an analysis and discussion of the results obtained from our experiments. We also conduct a comaparative analysis using segmented data and original non-segmented data.

- Chapter 6: In the final chapter, we provide a comprehensive conclusion summarizing our thesis work. We also suggest potential directions for future work and improvements.

CHAPTER 2

Related Works

2.1 Deep Learning Based Approaches For Different Brain Tumors Classification

The process of detecting a brain tumor is quite time-consuming and depends largely on the proficiency and expertise of the radiologist. With the growing patient population, the volume of data needing analysis has skyrocketed, rendering traditional methods inefficient and expensive. Numerous researchers have sought to devise algorithms that can both identify and classify brain tumors quickly and accurately. Recently, Deep Learning (DL) techniques have gained traction in creating automated systems adept at diagnosing or segmenting brain tumors in a reduced timeframe. DL leverages the power of pre-trained Convolutional Neural Network (CNN) models, specifically designed for categorizing medical images related to brain cancer.

Recent advancements in brain tumor detection and classification have demonstrated the potent application of deep learning and optimization techniques across various studies. ZainEldin et al. [28] introduced the Brain Tumor Classification Model (BCM-CNN) utilizing an Adaptive Dynamic Sine-Cosine Fitness Grey Wolf Optimizer (ADSCFGWO), achieving an exceptional accuracy of 99.98% on the BRaTS 2021 Task 1 dataset. This achievement parallels the work of Anita and Kumaran [29], who developed a convolutional neural network (CNN)-based method for meningioma tumor detection and segmentation, recording impressive sensitivity (99.1%), specificity

(99.5%), and tumor segmentation accuracy (99.4%). Meanwhile, Muezzinoglu et al. [30] proposed PatchResNet, a model leveraging patch-based deep feature engineering, which accomplished a 98.10% classification accuracy. Adding to the innovation, Kishanrao and Jondhale [31] presented a hybrid DCNN-DH framework for grade-based classification, showing superior performance in various metrics. JGate-AttResUNet, introduced by Ruba, Tamilselvi, and Beham [32], incorporated the J-Gate attention mechanism and achieved mean dice values of 0.896 and 0.913 on the BRATS datasets. The work of Farajzadeh, Sadeghzadeh, and Hashemzadeh [33] in deep hybrid representation learning method yielded 98.81% pixel-level accuracy and 98.93% classification accuracy on the MICCAI BraTS'20 dataset. Sarala et al. [34] introduced a Dual CNN method with a Histogram-Density Segmentation Algorithm (HDSA) achieving 98.9% sensitivity, 99.04% specificity, and 98.85% accuracy for both High-Grade and Low-Grade Glioma images. Sharma et al. [35]'s model using a Histogram of Gradient (HOG) technique for tumor detection noted an 88% accuracy. Additionally, Zahoor et al. [36]'s two-phase framework for MRI-based brain tumor analysis, with its detection scheme achieving a 99.56% accuracy and classification scheme outperforming existing methods with a 99.20% accuracy, underscores the field's progress. Mutual ensemble learning for segmentation by Hu, Gu, and Gu[37], utilizing Consensus Dice loss, consistently improved baseline network performance. Özkaraca et al. [38] merged DenseNet, VGG16, and basic CNN architectures to optimize MRI image classification, albeit with increased processing time. Evaluation of seven CNN models by Gómez-Guzmán et al. [39]on the Msoud dataset highlighted InceptionV3's superior accuracy of 97.12%. Shafi et al. [40]'s ensemble learning method for classifying brain tumors and autoimmune disease lesions marked a significant advancement with weighted rates in sensitivity (97.5%), specificity (98.838%), precision (98.011%), and accuracy (98.719%). These contributions collectively highlight the evolving landscape of medical imaging, showcasing both the efficacy of current deep learning applications and the continuous quest for improved diagnostic tools.

2.2 Segment Anything Model

The exploration of SAM in medical imaging has sparked a considerable amount of interest and research, leading to diverse applications and evaluations across various medical imaging modalities. de Oliveira et al. [41]investigated SAM's zero-shot capabilities across four imaging modalities, finding notable success in diverse medical imaging tasks and highlighting SAM's potential as a versatile tool in medical image segmentation. Mazurowski et al. [42] extended the evaluation of SAM across 19 medical imaging datasets, observing variability in SAM's performance which excelled in segmenting well-circumscribed objects and suggesting the utility of box prompts over point prompts for improved accuracy. Gong et al. [43] focused on adapting SAM from 2D to 3D medical imaging, particularly for tumor segmentation, proposing enhancements that surpass state-of-the-art models in some aspects, albeit noting the superior performance of CNN-based methods in specific segmentation tasks.

Further comparative studies such as those by Zhang et al. [44] and Ahmadi et al. [45], contrast SAM with established methods like FSL's BET and U-Net, respectively, showcasing SAM's advanced segmentation capabilities in challenging scenarios and its limitations in handling complex tumor shapes and text within images. Hu, Li, and Yang [46] applied SAM to breast tumor segmentation in ultrasound images, evaluating its effectiveness across different pre-trained model variants. Meanwhile, Höst et al. [47]introduced CellViT, demonstrating state-of-the-art performance in cell nuclei segmentation in tissue images, reflecting on the potential of integrating SAM with Vision Transformer architectures for enhanced medical image analysis.

Shin, Kim MD, and Baek [48] explored the synergy of SAM with condition embedding for ultrasound image segmentation, achieving significant improvements over traditional models. He et al. [49] benchmarked SAM across 12 medical image datasets, highlighting its underperformance compared to specialized medical image segmentation deep-learning algorithms, underscoring the necessity for further adaptation of SAM to the medical domain. This theme of customization and adaptation is further explored by Zhang and Liu [50] with SAMed, and Gao et al. [51] with DeSAM,

both aiming to bridge the gap between general image segmentation and the specific challenges of medical image analysis.

Wu et al. [52] presented the Medical SAM Adapter, enhancing SAM's application in medical image segmentation and achieving state-of-the-art results in several tasks, pointing towards the importance of domain-specific knowledge integration. Cheng et al. [53]'s comprehensive study on SAM's application across various medical datasets reveals the model's variable performance, emphasizing the critical role of prompt choice in segmentation tasks. Comparative studies like those by Mohapatra, Gosai, and Schlaug [54], and Bui et al. [55] with SAM3D, delve into the specifics of brain extraction techniques and 3D medical image segmentation, highlighting SAM's adaptability and areas for improvement.

Emerging studies such as Ma et al. [56] 's introduction of MedSAM, and Deng et al. [57]'s evaluation of SAM in digital pathology, further expand on SAM's applicability and limitations in medical image segmentation, suggesting avenues for future research and development. Huang et al. [58]'s examination of SAM in the context of a consolidated medical dataset underscores the model's potential and the challenges inherent in medical image segmentation. Putz et al. [59] evaluated SAM in the context of brain tumor segmentation for radiotherapy planning, noting the model's favorable accuracy but highlighting the influence of prompt number and view combination on segmentation performance.

In summary, the exploration of SAM and its adaptations in medical imaging reveals a landscape marked by innovative applications, challenges, and the ongoing quest to harness SAM's potential for enhancing medical diagnostics and treatment planning. These studies collectively underscore the dynamic interplay between model design, dataset specificity, and task complexity in leveraging SAM for medical image analysis [60] [61].

2.3   Ensemble Deep Learning Approaches

Various ensemble deep learning approaches have been explored to enhance the accuracy of medical diagnostics across different modalities. A notable integration of ResNet-152 and DenseNet-121 was introduced to detect COVID-19 from chest X-ray images [62]. Similarly, Xue et al. [63] proposed an ensemble framework that combines ResNet152, VGG16, ResNet50, and DenseNet121 to diagnose COVID-19 and pneumonia using CT scans and X-ray images. The field of cancer detection has also seen significant advancements through ensemble models. Barsha et al. [64] developed two models for detecting and grading Invasive Ductal Carcinoma (IDC), a common type of breast cancer. Their initial model uses DenseNet-121 and DenseNet-169 for IDC detection, while a subsequent model for IDC grading incorporates DenseNet-121, DenseNet-201, ResNet-101v2, and ResNet-50.

Further applications in cancer detection include an ensemble system employing DenseNet201, ResNet-101, ResNet-50, and AlexNet, coupled with a binary Support Vector Machine to evaluate malignancy in breast cancer [65]. Research into dermatological applications has seen implementations of deep learning methods like Resnet-50, VGG-16, Densenet, Mobilenet, Inceptionv3, and Xception for identifying skin cancer types, where multiple stacked models like inceptionv3-inceptionv3 and Resnet50-Vgg16 were evaluated [66]. However, these stacked models generally underperformed compared to existing methods.

Beyond medical imaging, ensemble methods have been applied in agriculture, where Vallabhajosyula et al. [67] utilized a weighted ensemble of deep neural networks to detect plant diseases, showing superior performance over other pre-trained models. In neurology, Anand et al. [68] implemented a weighted average ensemble model to detect brain tumors in MRI images, optimizing model weights via grid search. Additionally, Shaga Devan et al. [69] proposed an ensemble-based semantic segmentation process that evaluated various base learner models for optimizing segmentation tasks.

These studies illustrate the broad applicability and potential of ensemble deep

learning models in improving diagnostic accuracy across a spectrum of medical and environmental challenges.

CHAPTER 3

Methodology

In this chapter, we discuss our custom datasets and introduce our proposed methodology.

3.1   Dataset Preparation

Our research utilizes MRI images, focusing on pituitary tumors and healthy brain scans, sourced from four publicly available datasets. These datasets include Kaggle [70][70], Figshare [71], and BR35H [72], which collectively provide a diverse range of imaging data. We compiled these resources into a custom dataset categorized into two distinct classes: pituitary tumor images and normal or healthy brain images. The dataset comprises a total of 3,287 MRI scans, with 1,857 images representing pituitary tumors and 1,430 images depicting normal brain conditions.

This section presents a visual comparison of the dataset used in our study. Figure 3.1.1 in the first row displays the original MRI scans of pituitary tumors. These images serve as the baseline for assessing the effectiveness of SAM. The second row, shown in Figure 3.1.2, illustrates the results post-application of SAM. These segmented images highlight the detailed isolation and enhanced visibility of the pituitary tumors, demonstrating SAM's capability to refine and clarify the features critical for accurate medical diagnosis.

(a) (b) (c) (d)
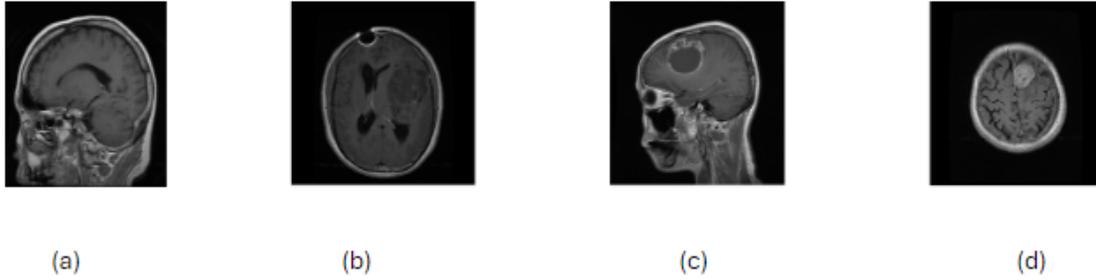
Fig. 3.1.1: Original Pituitary Tumor MRI Scans

(a) (b) (c) (d)
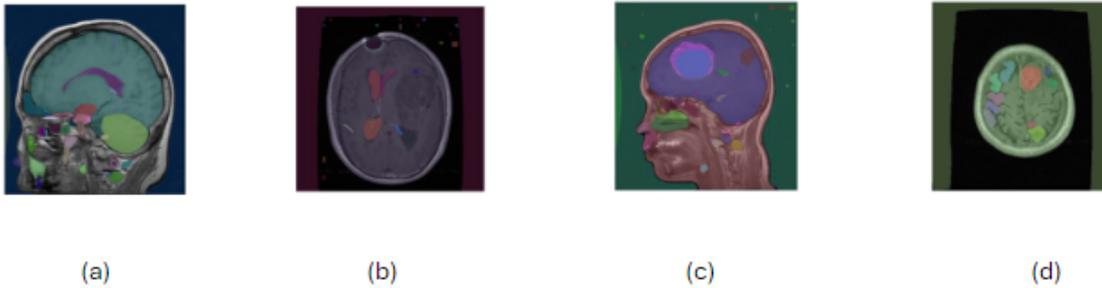
Fig. 3.1.2: Corresponding Segmented Images After Applying SAM

3.2 Methodology Overview

In this section we include the overview of proposed methodology with details expla-
nation of each steps. The methodology illustrated in the figure 3.2.1 is a structured
approach to classifying MRI images for the presence of pituitary tumors. It encom-
passes several key stages, as follows:

a. **Original MRI:** The starting point is the collection of MRI scans. These scans
   are complex and high-dimensional data that require preprocessing to highlight
   the features of interest, namely, the anatomical structures within the brain.
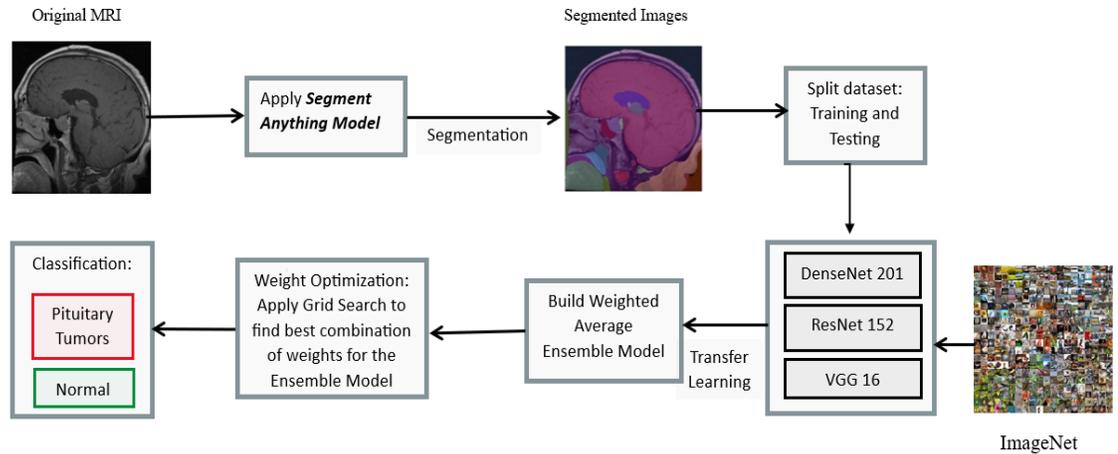
Fig. 3.2.1: Methodology Overview

b. **Apply Segment Anything Model:** SAM is applied to each MRI scan. It leads to the process of partitioning an image into multiple segments to simplify or change the representation of an image into something that is more meaningful and easier to analyze.

c. **Segmentation:** As a result of segmentation, the regions of interest within the MRI scans, such as potential tumors or specific brain structures, are isolated. This process is crucial for medical imaging as it aids in distinguishing pathological areas from normal tissue.

d. **Segmented Images:** The output from the segmentation process is a set of images where the regions of interest have been clearly marked or highlighted. A segmented image is the result of dividing a digital image into multiple segments (sets of pixels), with the aim of simplifying or changing the representation of the image into something that is more meaningful and easier to analyze. These images are now ready for further processing and analysis.

e. **Split dataset:** The segmented images are divided into two groups: a training set and a testing set. The training set is used by the machine learning algorithms to learn the patterns associated with normal scans and those with pituitary tumors. The testing set, which the algorithms have not seen during training, is used to evaluate the performance and generalizability of the trained models.

f. **Transfer Learning using ImageNet:** The methodology employs transfer learning [73], which involves taking a pre-trained model (trained on a different but related task) and fine-tuning it for a specific task. In this case, the pre-trained models (DenseNet 201, ResNet 152, VGG 16) have been trained on ImageNet, a large dataset containing millions of labeled images of thousands of categories. By using these pre-trained models, the system can leverage the rich feature representations these networks have developed for general image recognition tasks, which can be beneficial for medical image analysis despite the difference in domain [74]. We include brief description of each base models in the following:

• **DenseNet 201:** Huang et al. [75] introduce the concept of a dense convolutional network, known as DenseNet, which redefines connectivity in deep learning architectures. Diverging from the norm where a conventional convolutional network with 'n' layers has 'n' connections, DenseNet innovatively forges a total of *n(n+1)/2* direct connections. This architectural enhancement facilitates each layer receiving input from all preceding layer feature maps, with its outputs serving as inputs to every successive layer.

The operation of DenseNet can be visualized when an initial image x_0 progresses through a network comprised of 'n' layers, each responsible for a unique non-linear transformation denoted as $H_n(.)$. Here, 'n' symbolizes the ordered sequence of layers. In this setting, the nth layer obtains an aggregated set of feature maps from all preceding layers $[x_0, x_1, \ldots, x_{n-1}]$, producing the output:

$$x_n = H_n([x_0, x_1, \ldots, x_{n-1}]) \tag{1}$$

In this equation, $[x_0, x_1, \ldots, x_{n-1}]$ signifies the concatenated feature maps from layers *0* to *n-1*. DenseNet-201, a particular adaptation within the DenseNet family, is distinguished by its depth, with '201' indicating the total count of

layers it contains. This depth allows DenseNet-201 to learn highly complex patterns and nuances from the data, making it exceptionally suitable for tasks requiring detailed feature extraction, such as image classification.

- **ResNet 152:** In the domain of convolutional neural network architectures, Residual Networks (ResNets) represent a significant advancement, primarily introduced to address the vanishing gradient issue that hampers deep network training [76]. The cornerstone of ResNet's efficacy is the 'skip connection' strategy, which facilitates the bypassing of one or several layers. This technique involves the forwarding of the activation from an earlier layer directly to a subsequent layer, effectively enabling deeper networks to learn without the hindrance of diminishing gradients [77]. The ResNet-152 model is a variant within the ResNet family, characterized by its depth, with the number '152' signifying the total layers it comprises. This depth endows ResNet-152 with the capacity to perform intricate feature extractions, rendering it particularly potent for a wide array of complex tasks in computer vision.

- **VGG 16:** The VGG-16 architecture, conceived by the Visual Geometry Group at the University of Oxford [78], is recognized for its depth and utilization of multiple convolutional layers. Structurally, it consists of 16 layers that include 13 convolutional layers complemented by three fully connected layers. The model adopts compact 3x3 convolutional filters, systematically arranged in sequential blocks punctuated by max pooling to reduce dimensions progressively. Initial blocks are engineered to capture fundamental image attributes, whereas subsequent blocks are adept at recognizing more sophisticated patterns. The VGG-16 model has been seminal in demonstrating the profound capabilities of deep neural networks in the realm of image recognition. Its design underscores the significance of network depth in enhancing feature extraction capabilities and showcases the effectiveness of using small convolutional filters to discern detailed patterns within images.

g. **Build Weighted Average Ensemble Model:** An ensemble approach com-

bines multiple models to improve the overall predictive performance [79]. A weighted average ensemble assigns a weight to the outputs of individual models before averaging, giving more influence to certain models in the decision-making process.

h. **Weight Optimization:** We apply Grid Search to find the best combination of weights for the Ensemble Model: Grid Search is an exhaustive search technique that tries out all possible combinations of weights within a specified range for the models in the ensemble [80]. The goal is to find the set of weights that yield the best performance, usually measured by accuracy, precision, or recall. This process involves training the ensemble multiple times with different weight combinations and evaluating the results.

i. **Classification:** After training, the ensemble model is used to classify new MRI scans into either having pituitary tumors or being normal. The final classification is made based on the learned features and patterns from the training process, taking into account the optimized weights from the grid search. "Pituitary Tumor" and "Normal" are the final categories into which the MRI scans are classified. "Pituitary Tumors" would be assigned to scans where the model identifies patterns consistent with known tumors, whereas "Normal" would be assigned to scans that do not show these patterns.

In this methodology, the initial segmentation step is crucial as it enhances the models' focus on relevant features, potentially improving classification performance. The use of transfer learning can significantly reduce the training time and data requirements, which is especially important in the medical domain where labeled data can be scarce and expensive to obtain. The ensemble approach aims to create a robust system that is less likely to be affected by the weaknesses of individual models. Grid search ensures that the ensemble's classification ability is optimized for the highest possible accuracy. Each step in this process would typically be accompanied by rigorous testing to ensure the models are generalizable and not overfitting to the training data.

3.3 Algorithm for the Proposed System

Input: Original MRI Image

Output: Classification Result: "Pituitary Tumors" or "Normal"

Steps:

1. **Segmentation of MRI Image**

   1.1 Load the original MRI image into the system.
   1.2 Apply the "Segment Anything Model" to segment the MRI image.
   1.3 Store the segmented image for further processing.

2. **Dataset Preparation**

   2.1 Collect a dataset of segmented MRI images.
   2.2 Split this dataset into two parts: (i)Training Dataset, (ii) Testing Dataset

3. **Transfer Learning**

   3.1 Initialize the following pre-trained models using ImageNet weights: DenseNet 201, ResNet 152, VGG 16.
   3.2 Fine-tune each of these models on the Training Dataset of segmented MRI images.
   3.3 Store the weights and architecture of the fine-tuned models for ensemble.

4. **Build Weighted Average Ensemble Model**

   4.1 Create an ensemble of the fine-tuned models (DenseNet 201, ResNet 152, VGG 16) using a weighted average mechanism.
   4.2 Initialize the weights for each model in the ensemble.

5. **Weight Optimization**

   5.1 Apply a Grid Search algorithm on the Training Dataset to find the best combination of weights that optimize the ensemble model's performance.
   5.2 Store the optimized weights for classification.

6. **Classification**

   6.1 Use the Weighted Average Ensemble Model with optimized weights to classify the segmented MRI images from the Testing Dataset.
   6.2 Compare the predicted classification with the actual labels to evaluate the model's performance.
   6.3 Return the classification result: "Pituitary Tumors" or "Normal".

CHAPTER 4

Experiments

This section explains the experiments and settings utilized for this study, encompassing the assortment of tools and libraries employed in the construction of our proposed model, the system configuration details, the specific hyperparameters selected during the training phase, comprehensive information about the dataset, and an in-depth explanation of the evaluation metrics applied to assess the performance of our model.

4.1   Environmental Setup

The technical specifications for the environment in which the experiments were performed are as follows:

A system with Windows 11 64-bit, powered by an Intel(R) Core (TM) i7-8750H processor with a clock speed of 2.2 GHz, complemented by 16 GB of memory and a 8 GB NVIDIA GeForce GTX 1070 GPU laptop. Python was the primary programming language employed during the experimental procedures.

4.2   Tools and Libraries:

• NumPy (numpy): This is a library for numerical computing in Python. It provides support for large, multi-dimensional arrays and matrices, along with a collection of mathematical functions to operate on these arrays.

• TensorFlow (tensorflow): TensorFlow is an end-to-end open-source platform for

29

machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that lets researchers and developers build and deploy machine learning powered applications.

• Keras API (tensorflow.keras): Keras is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library. Within the Keras module, several classes are used:

-applications (ResNet152, DenseNet201, VGG16): Pre-trained models on the ImageNet dataset, used for transfer learning.

-models.Model, models.load_model: For constructing the neural network model and loading the saved model, respectively.

-layers: Different types of neural network layers, such as Input, Average, GlobalAveragePooling2D, Dense, and Dropout, for constructing custom model architectures.

-optimizers.Adam: Optimization algorithm used for training the neural network.

-callbacks.EarlyStopping: To stop training when a monitored metric has stopped improving.

• scikit-learn (sklearn.metrics): This library is used for performing various machine learning tasks. It provides tools for model fitting, data preprocessing, model selection, and evaluation metrics.

• Matplotlib (matplotlib.pyplot): A plotting library for creating static, interactive, and animated visualizations in Python.

• Pillow (PIL): The Python Imaging Library adds image processing capabilities to your Python interpreter. It allows for opening, manipulating, and saving many different image file formats.

• OpenCV (cv2): Open Source Computer Vision Library, which is aimed at real-time computer vision.

• OS (os): This module provides a way of using operating system-dependent functionality like reading or writing to a filesystem.

• Jupyter is used as IDE for development.

4.3 Hyperparameter tuning

Hyperparameter tuning is a critical step in the construction of machine learning models, which involves finding the most effective combination of parameters that defines the model architecture and training process. Unlike model parameters, hyperparameters are not directly learned from the training process and must be set prior to the training [81].

In the development of our ensemble CNN models for image classification, we employed various hyperparameter tuning techniques to optimize the performance. These techniques are integral for improving model accuracy and preventing overfitting.

Here is the summary of hyperparameter tuning:

Learning rate = 0.0001

Train-test split ratio = 80:20

Batch Size = 32

Number of Epochs = 50

Image size = (224, 224)

4.4 Weight Optimization in Ensemble Models Using Grid Search

To enhance the predictive performance of our ensemble model, which integrates outputs from ResNet152, DenseNet201, and VGG16 models, we implemented weight optimization using the grid search technique. This approach plays a crucial role in determining the optimal weights assigned to each constituent model within the ensemble. Grid search is a methodical approach used to tune hyperparameters where a model is evaluated for each combination of hyperparameter settings specified in a predefined grid. In the context of our ensemble model, the grid comprised various possible weight combinations for the predictions from the individual models.

The specific implementation involved defining a set of potential weights for each

model, and then exhaustively testing every possible combination to identify which set of weights maximized the ensemble's performance on the validation dataset. This process ensures that the selected weights are those that best leverage the strengths of each individual model. For example, the predefined grid for our ensemble might look like this:

$$weights\_grid = [[0.5,\ 0.1,\ 0.4],\ [0.4,\ 0.3,\ 0.3],\ [0.3,\ 0.4,\ 0.3],\ ...]$$

Each combination of weights is applied to the predictions of the individual models, and the ensemble's performance with each weight set is evaluated. The optimal weights are those that achieve the highest accuracy.

This method of hyperparameter tuning via grid search not only fine-tunes the ensemble for better accuracy but also provides insights into the relative effectiveness of each individual model within the ensemble.

4.5   Learning Rate Optimization

The learning rate for the Adam optimizer was set to a small value of 0.0001. We experiment with other values, like 0.01, 0.001, etc, but this learning rate gives us highest accuracy. The learning rate is a hyperparameter that controls how much to update the model in response to the estimated error each time the model weights are updated. Choosing a smaller learning rate can lead to precise convergence but may require more training epochs:

$$opt = Adam(learning\_rate = 0.0001)$$

4.6   Evaluation Metrics

In our experiments, we employed four evaluation metrics: accuracy, precision, recall, and F1-score.

(i) Precision: Precision measures the proportion of correctly predicted positive observations to the total predicted positives. It evaluates a model's capacity to minimize false positives [82]. Mathematically, it is expressed as:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

In our thesis, achieving high precision indicates that our classification model effectively classifies the tumor images with minimal incorrection.

(ii) Recall: Recall, also known as sensitivity, is the ratio of correctly predicted positive observations to all observations in the actual class. It assesses the model's ability to capture all relevant instances [82]. Mathematically, it is defined as:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

A high recall value in our experiments suggests that our model proficiently classifies the tumor images.

(iii) F1-Score: The F1-score is a metric that quantifies a model's accuracy on a dataset by combining the precision and recall into a single measure. This is important as balancing these two metrics is a common challenge [82]. It is calculated as the harmonic mean of precision and recall:

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (3)$$

A high F1-score indicates that both precision and recall are high, suggesting a more accurate and consistent performance from our classification model.

(iv) Accuracy: Accuracy is a widely-used evaluation metric in the field of machine learning and statistics, particularly in classification tasks. It is defined as the ratio of correctly predicted instances to the total number of instances evaluated. In essence,

it measures the proportion of true results, both true positives and true negatives, in the data [82]. The formula for accuracy is given by:

$$Accuracy = \frac{True\ Positives\ +\ True\ Negatives}{True\ Positives + True\ Negatives + False\ Positives + False\ Negatives} \quad (4)$$

CHAPTER 5

Results and Discussions

In this chapter, we embark on a detailed analysis of the experimental outcomes previously delineated. We will delve into a comparative assessment, examining the performance metrics and results in a comprehensive discussion that synthesizes our findings within the broader context of our research objectives.

5.1 Analysis of DL Models for Classification

In our research, we explored a variety of DL architectures to evaluate their performances on the task at hand. These architectures included DenseNet-201, InceptionV3, ResNet50, ResNet-101, ResNet-152, MobileNetV2, EfficientNetB0, Xception, VGG-19, VGG-16, and NasNet-Large. Upon careful examination of their accuracy, three models stood out: ResNet-152, DenseNet-201, and VGG-16. These models achieved accuracy scores of 0.9312, 0.9103, and 0.9475, respectively, when applied to data segmented by SAM.

Building on the individual strengths of these top-performing models, we constructed an ensemble model. The ensemble approach was based on a weighted average methodology, which considers the predictive confidence of each model by assigning it a weight. To optimize these weights (*w1* for ResNet-152, *w2* for DenseNet-201, and *w3* for VGG-16), we employed the grid search technique, which exhaustively searched for all possible weight combinations to enhance the ensemble's predictive accuracy.

Table 5.1.1: Weight Optimization Using Grid Search

w1	w2	w3	Train Loss	Val Loss	Val Acc
0.1	0.1	0.8	0.0089	0.1019	0.9627
0.1	0.2	0.7	0.0057	0.1335	0.9539
0.1	0.3	0.6	0.0026	0.0954	0.9676
0.1	0.4	0.5	0.0054	0.0695	0.9753
0.1	0.5	0.4	0.0086	0.0881	0.9741
0.1	0.6	0.3	0.0037	0.0768	0.9753
0.1	0.7	0.2	0.0040	0.0769	0.9724
0.1	0.8	0.1	0.0033	0.0704	0.9803
0.2	0.1	0.7	0.0092	0.1116	0.9638
0.2	0.2	0.6	0.0110	0.0825	0.9655
0.2	0.3	0.5	0.0057	0.0748	0.9741
0.2	0.4	0.4	0.0077	0.0933	0.9688
0.2	0.5	0.3	0.0096	0.0887	0.9692
0.2	0.6	0.2	0.0030	0.0742	0.9532
0.2	0.7	0.1	0.0122	0.0680	0.9786
0.3	0.1	0.6	0.0206	0.1341	0.9572
0.3	0.2	0.5	0.0186	0.0663	0.9803
0.3	0.3	0.4	0.0095	0.0723	0.9757
0.3	0.4	0.3	0.0101	0.0878	0.9770
0.3	0.5	0.2	0.0127	0.0912	0.9819
0.3	0.6	0.1	0.0184	0.0547	0.9786
0.4	0.1	0.5	0.0157	0.0792	0.9741
0.4	0.2	0.4	0.0152	0.0758	0.9753
0.4	0.3	0.3	0.0133	0.0521	0.9235
0.4	0.4	0.2	0.0365	0.1036	0.9720
0.4	0.5	0.1	0.0157	0.0607	0.9852
0.5	**0.1**	**0.4**	**0.0154**	**0.0880**	**0.9885**
0.5	0.2	0.3	0.0217	0.0726	0.9836
0.5	0.3	0.2	0.0334	0.1552	0.9836
0.5	0.4	0.1	0.0298	0.0692	0.9852
0.6	0.1	0.3	0.0728	0.4178	0.6891
0.6	0.2	0.2	0.0316	0.1066	0.9424
0.6	0.3	0.1	0.0278	0.0797	0.9737
0.7	0.1	0.2	0.0393	0.2109	0.8717
0.7	0.2	0.1	0.0562	0.0402	0.9622
0.8	0.1	0.1	0.0683	0.1243	0.9465

Our grid search revealed that the ensemble model attained the highest accuracy, at 0.9885, with the weights set to w1 = 0.5, w2 = 0.1, and w3 = 0.4. This outcome demonstrates the efficacy of leveraging ensemble techniques and hyperparameter tuning to improve model performance significantly beyond what the individual models achieved on their own.

5.2 Comparative Analysis

In our comparative analysis, we carefully executed the implementation of the three foundational models—namely, ResNet-152, DenseNet-201, and VGG-16—as well as their integrative ensemble model using the original MRI datasets as inputs. This step was critical to establish a performance baseline for models utilizing unprocessed imaging data.

Subsequently, we compared the performance of the same models when applied to data that had been pre-processed through segmentation using SAM. Our findings revealed a noticeable enhancement in model performance metrics when leveraging the segmented data. The segmentation effectively isolates and emphasizes the regions of interest within the MRI scans, thus providing the models with refined inputs that are more conducive to accurate classification.

This improvement in performance on segmented data underscores the efficacy of the SAM in pre-processing complex medical images for subsequent analysis. Such segmentation aids in reducing the background noise and focusing the model's attention on salient features, which is paramount in medical imaging tasks where precision is of the utmost importance.

Further, these findings align with the hypothesis that models trained on data that are closely aligned with the task-specific features can outperform those trained on raw, unprocessed data. This comparative analysis not only serves as a testament to the importance of data pre-processing in deep learning workflows but also sets the stage for further discussion on the integration of segmentation models into clinical diagnostic processes.

Table 5.2.1: Experiment results using original MRI (data without segmentation)

Evaluation Metrics	ResNet152	DenseNet201	VGG16	Ensemble Model
Recall	0.9251	0.8818	0.9244	0.9681
Precision	0.9146	0.8997	0.9211	0.9697
F1 Score	0.9198	0.8907	0.9227	0.9689
Accuracy	0.9173	0.8953	0.9271	0.9708
Training Time(second)	2167.9955	1269.9225	1117.3872	4021.1835
Inference Time(second)	11.0752	9.7861	5.9373	19.5717

Table 5.2.2: Experiment results using segmented data (segmented by SAM)

Evaluation Metrics	ResNet152	DenseNet201	VGG16	Proposed Ensemble Model
Recall	0.9399	0.9018	0.9389	**0.9791**
Precision	0.9286	0.9233	0.9586	**0.9879**
F1 Score	0.9342	0.9124	0.9486	**0.9835**
Accuracy	0.9312	0.9103	0.9475	**0.9885**
Training Time (second)	1996.7869	1159.5039	1040.4228	4016.5057
Inference Time(second)	**9.7003**	**9.4863**	**5.6003**	**11.1178**

In the process of evaluating performance of our model, a critical factor considered was the computational efficiency, particularly the time overhead for mask generation. The segmentation task involves the creation of masks that outline the target regions within each image. Our experimental results indicate that the average time required to generate these segmentation masks across all tested images was approximately 20.0259 seconds per image. This measurement was taken from the initiation of the segmentation process to the completion of the mask generation for each image, reflecting the computational overhead associated with SAM. This duration is a significant metric, as it impacts the practical deployment of the model in real-time applications, where speed is often as critical as accuracy.

5.3 Findings

In this research, our objective extended to a comprehensive comparative analysis of the ensemble model's performance using both segmented and non-segmented MRI datasets. The segmented dataset, processed by SAM, and the original, non-segmented MRI data formed the basis for this evaluation. The ensemble model, intricately weighted with w1 = 0.5 for ResNet-152, w2 = 0.1 for DenseNet-201, and w3 = 0.4 for VGG-16, served as the analytical tool for this comparison.

Upon implementation, the ensemble model, utilizing the optimally tuned weight configuration, demonstrated a remarkable accuracy of 98.85% on the segmented data. In contrast, the same model achieved a lower accuracy of 97.08% when applied to the non-segmented original dataset. This differential outcome underscores a notable accuracy improvement of 1.77% as a result of employing SAM for data pre-processing.

Moreover, an assessment of the individual base models revealed a consistent trend, with an average performance enhancement of 1.6430% observed upon the application of SAM to the datasets. This enhancement was not limited to accuracy metrics alone; there was a discernible improvement in computational efficiency as well, evidenced by reductions in both training and inference times for the base models.

These observations unequivocally suggest that the integration of segmentation

model, SAM can play a pivotal role in boosting the performance of deep learning models in medical image analysis. The segmentation process aids in distilling critical features from the MRI scans, thereby enhancing the model's ability to learn and generalize from the data. This enhancement in accuracy, coupled with the efficiencies gained in computational performance, holds significant promise for the application of such methodologies in clinical settings, potentially improving diagnostic processes and patient outcomes.

5.4   Significance of Our Study

In the realm of medical imaging, radiologists are frequently confronted with the intricate task of specifying diminutive anatomical structures, such as minuscule brain tumors or skin lesions. The challenge is compounded by the often-indistinct boundaries of such objects, which can blend imperceptibly into surrounding tissues.

Current automated and semi-automated segmentation strategies are not without their drawbacks. They necessitate extensive customization to suit specific datasets and are hindered by a dearth of dependable methods for result validation, as highlighted in reference [9]. These hurdles can lead to inefficiencies in the segmentation process and may introduce uncertainties in subsequent diagnostic analyses.

To address these impediments, our research introduces an ensemble deep learning framework, incorporating SAM. This innovative approach presents a robust solution to the quandaries posed by small object segmentation. By leveraging SAM within our ensemble method, we seek to pioneer advancements in the classification of minute medical entities.

Our contributions have the potential to markedly reduce the health risks associated with minuscule, yet clinically significant abnormalities. These include small lung nodules, incipient pituitary tumors, and diminutive polyps detectable in colonoscopy imagery. Such conditions, if undiagnosed, may escalate undetected, precipitating grave health complications. The detection of these anomalies in their nascent stages is vital, as it can enable timely medical interventions. By facilitating earlier detec-

tion, our research endeavors to enhance therapeutic outcomes and may obviate the necessity for more aggressive, invasive procedures down the line.

The implications of this work are profound. It not only stands to refine diagnostic precision but also to usher in an era of improved patient prognoses through the early interception of conditions that, if left unchecked, could prove detrimental to patient health.

CHAPTER 6

Conclusion and Future Work

6.1 Conclusion

In conclusion, this thesis has successfully demonstrated the considerable potential of an ensemble deep learning approach in enhancing the accuracy of medical image classification. Through extensive experimentation, it became evident that the application of SAM to preprocess the data significantly improved the performance of the models under investigation. The precise segmentation of elusive medical structures like small tumors and lesions, achieved by the SAM, has proven pivotal, leading to an increase in classification accuracy by 1.77% when compared to original non-segmented data.

The findings confirm the hypothesis that the integration of sophisticated segmentation techniques with advanced deep learning models can lead to more precise and reliable medical imaging analysis. Specifically, the optimized ensemble model, harnessing the collective strengths of ResNet-152, DenseNet-201, and VGG-16 with finely-tuned weights, emerged as a powerful tool, outperforming the individual base models with an impressive accuracy of 98.85% on segmented data.

This work contributes a significant step forward in the application of artificial intelligence in medical diagnostics, particularly in the detection and classification of small-scale pathologies, which are often challenging yet critical for early diagnosis and treatment. It sets the stage for future advancements where such AI-driven tools could be routinely used in clinical practice, aiding radiologists in making faster, more

accurate diagnoses and thus improving patient outcomes.

Moreover, this thesis lays the groundwork for future research to expand upon. Subsequent studies could explore a broader range of medical imaging modalities, apply the ensemble model to larger and more diverse datasets, and investigate the implementation of such AI systems in real-world clinical settings. The promising results obtained here advocate for the continued intersection of artificial intelligence and healthcare, with the goal of developing non-invasive, efficient, and highly accurate diagnostic techniques for the betterment of global health.

6.2 Future Work

In future investigations, it would be beneficial to refine our segmentation approach by specifically targeting the segments containing the pituitary gland, rather than processing entire images. This focused segmentation could potentially reduce computational overhead and enhance the accuracy of our model. By isolating the region of interest—namely, the pituitary gland—the model can dedicate more computational resources to analyzing features critical for accurate classification of pituitary tumors.

Moreover, the application of these methodologies to other types of medical imaging data, such as CT scans or advanced MRI modalities, could broaden the impact of our research. By adapting our models to handle different imaging technologies and conditions, we can contribute to the broader field of medical image analysis, potentially aiding in the detection and classification of other diseases.

Through these focused efforts, our future work aims to not only advance the state of technology in medical image processing but also significantly enhance the tools available to radiologists and medical professionals, thereby improving patient outcomes and diagnostic efficiency.

REFERENCES

[1] "Enlargement of the pituitary gland," https://www.msdmanuals.com/home/quick-facts-hormonal-and-metabolic-disorders/pituitary-gland-disorders/enlargement-of-the-pituitary-gland, [Accessed 24-04-2024].

[2] "Skull base tumors," https://www.cancer.org/content/dam/CRC/PDF/Public/8790.00.pdf, [Accessed 22-11-2023].

[3] MathWorks, "Multilabel Image Classification Using Deep Learning - MATLAB & Simulink — mathworks.com," https://www.mathworks.com/help/deeplearning/ug/multilabel-image-classification-using-deep-learning.html, [Accessed 24-04-2024].

[4] IndiaAi, "Image Segmentation: The Deep Learning Approach — indiaai.gov.in," https://indiaai.gov.in/article/image-segmentation-the-deep-learning-approach, August 2022, [Accessed 24-04-2024].

[5] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo *et al.*, "Segment anything," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2023, pp. 4015–4026.

[6] Meta, "Introducing Segment Anything: Working toward the first foundation model for image segmentation — ai.meta.com," https://ai.meta.com/blog/segment-anything-foundation-model-image-segmentation/, April 2023, [Accessed 24-11-2023].

[7] M. Ahmadi, A. G. Lonbar, A. Sharifi, A. T. Beris, M. Nouri, and A. S. Javidi, "Application of segment anything model for civil infrastructure defect assessment," *arXiv preprint arXiv:2304.12600*, 2023.

[8] J. Wu, R. Xu, Z. Wood-Doughty, and C. Wang, "Segment anything model is a good teacher for local feature learning," *arXiv preprint arXiv:2309.16992*, 2023.

[9] M. B. Nagarajan, M. B. Huber, T. Schlossbauer, G. Leinsinger, A. Krol, and A. Wismüller, "Classification of small lesions in dynamic breast mri: eliminating the need for precise lesion segmentation through spatio-temporal analysis of contrast enhancement," *Machine vision and applications*, vol. 24, no. 7, pp. 1371–1381, 2013.

[10] E. Laws, D. Penn, and C. Repetti, "Advances and controversies in the classification and grading of pituitary tumors," *Journal of Endocrinological Investigation*, vol. 42, pp. 129–135, 2019.

[11] L. E. Castellanos, C. Gutierrez, T. Smith, E. R. Laws, and J. B. Iorgulescu, "Epidemiology of common and uncommon adult pituitary tumors in the us according to the 2017 world health organization classification," *Pituitary*, pp. 1–9, 2022.

[12] A. D. Bhimani, A. J. Schupper, G. D. Arnone, D. Chada, A. N. Chaker, N. Mohammadi, C. G. Hadjipanayis, and A. I. Mehta, "Size matters: rethinking of the sizing classification of pituitary adenomas based on the rates of surgery: a multi-institutional retrospective study of 29,651 patients," *Journal of Neurological Surgery Part B: Skull Base*, pp. 066–075, 2020.

[13] UCSF Health, "Pituitary tumors," https://www.ucsfhealth.org/conditions/pituitary-tumors, 2023, accessed: 2023-04-24.

[14] T. Tsukamoto and Y. Miki, "Imaging of pituitary tumors: an update with the 5th who classifications—part 1. pituitary neuroendocrine tumor (pitnet)/pituitary adenoma," *Japanese Journal of Radiology*, vol. 41, no. 8, pp. 789–806, 2023.

[15] "Pituitary tumors early detection, diagnosis, and staging," https://www.cancer.org/content/dam/CRC/PDF/Public/8790.00.pdf, [Accessed 22-11-2023].

[16] K. Huang, "Image classification using the method of convolutional neural networks," in *2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS)*. IEEE, 2022, pp. 827–832.

[17] Y. Zhang, "Understanding image fusion," *Photogramm. Eng. Remote Sens*, vol. 70, no. 6, pp. 657–661, 2004.

[18] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[20] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *nature*, vol. 542, no. 7639, pp. 115–118, 2017.

[21] Y. Tan, *GPU-based parallel implementation of swarm intelligence algorithms*. Morgan Kaufmann, 2016.

[22] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 3431–3440.

[23] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*. Springer, 2015, pp. 234–241.

[24] Z. Wang, J. Liang, Z. Wang, and T. Tan, "Exploiting semantic attributes for transductive zero-shot learning," in *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2023, pp. 1–5.

[25] K. He, X. Chen, S. Xie, Y. Li, P. Dollár, and R. Girshick, "Masked autoencoders are scalable vision learners," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 16 000–16 009.

[26] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.

[27] Y. Li, H. Mao, R. Girshick, and K. He, "Exploring plain vision transformer backbones for object detection," in *European Conference on Computer Vision*. Springer, 2022, pp. 280–296.

[28] H. ZainEldin, S. A. Gamel, E.-S. M. El-Kenawy, A. H. Alharbi, D. S. Khafaga, A. Ibrahim, and F. M. Talaat, "Brain tumor detection and classification using deep learning and sine-cosine fitness grey wolf optimization," *Bioengineering*, vol. 10, no. 1, p. 18, 2022.

[29] J. N. Anita and S. Kumaran, "A deep learning architecture for meningioma brain tumor detection and segmentation," *Journal of Cancer Prevention*, vol. 27, no. 3, p. 192, 2022.

[30] T. Muezzinoglu, N. Baygin, I. Tuncer, P. D. Barua, M. Baygin, S. Dogan, T. Tuncer, E. E. Palmer, K. H. Cheong, and U. R. Acharya, "Patchresnet: multiple patch division–based deep feature fusion framework for brain tumor classification using mri images," *Journal of Digital Imaging*, vol. 36, no. 3, pp. 973–987, 2023.

[31] S. A. Kishanrao and K. C. Jondhale, "An improved grade based mri brain tumor classification using hybrid dcnn-dh framework," *Biomedical Signal Processing and Control*, vol. 85, p. 104973, 2023.

[32] T. Ruba, R. Tamilselvi, and M. P. Beham, "Brain tumor segmentation using jgate-attresunet–a novel deep learning approach," *Biomedical Signal Processing and Control*, vol. 84, p. 104926, 2023.

[33] N. Farajzadeh, N. Sadeghzadeh, and M. Hashemzadeh, "Brain tumor segmentation and classification on mri via deep hybrid representation learning," *Expert Systems with Applications*, vol. 224, p. 119963, 2023.

[34] B. Sarala, G. Sumathy, A. Kalpana, and J. J. Hephzipah, "Glioma brain tumor detection using dual convolutional neural networks and histogram density segmentation algorithm," *Biomedical Signal Processing and Control*, vol. 85, p. 104859, 2023.

[35] A. K. Sharma, A. Nandal, A. Dhaka, K. Polat, R. Alwadie, F. Alenezi, and A. Alhudhaif, "Hog transformation based feature extraction framework in modified resnet50 model for brain tumor detection," *Biomedical Signal Processing and Control*, vol. 84, p. 104737, 2023.

[36] M. M. Zahoor, S. A. Qureshi, S. Bibi, S. H. Khan, A. Khan, U. Ghafoor, and M. R. Bhutta, "A new deep hybrid boosted and ensemble learning-based brain tumor analysis using mri," *Sensors*, vol. 22, no. 7, p. 2726, 2022.

[37] J. Hu, X. Gu, and X. Gu, "Mutual ensemble learning for brain tumor segmentation," *Neurocomputing*, vol. 504, pp. 68–81, 2022.

[38] O. Özkaraca, O. İ. Bağrıaçık, H. Gürüler, F. Khan, J. Hussain, J. Khan, and U. e. Laila, "Multiple brain tumor classification with dense cnn architecture using brain mri images," *Life*, vol. 13, no. 2, p. 349, 2023.

[39] M. A. Gómez-Guzmán, L. Jiménez-Beristaín, E. E. García-Guerrero, O. R. López-Bonilla, U. J. Tamayo-Perez, J. J. Esqueda-Elizondo, K. Palomino-Vizcaino, and E. Inzunza-González, "Classifying brain tumors on magnetic resonance imaging by using convolutional neural networks," *Electronics*, vol. 12, no. 4, p. 955, 2023.

[40] A. Shafi, M. B. Rahman, T. Anwar, R. S. Halder, and H. E. Kays, "Classification of brain tumors and auto-immune disease using ensemble learning," *Informatics in Medicine Unlocked*, vol. 24, p. 100608, 2021.

[41] C. M. de Oliveira, L. V. de Moura, R. C. Ravazio, L. S. Kupssinskü, O. Parraga, M. M. Delucis, and R. C. Barros, "Zero-shot performance of the segment anything model (sam) in 2d medical imaging: A comprehensive evaluation and practical guidelines," *CoRR*, 2023.

[42] M. A. Mazurowski, H. Dong, H. Gu, J. Yang, N. Konz, and Y. Zhang, "Segment anything model for medical image analysis: an experimental study," *Medical Image Analysis*, vol. 89, p. 102918, 2023.

[43] S. Gong, Y. Zhong, W. Ma, J. Li, Z. Wang, J. Zhang, P.-A. Heng, and Q. Dou, "3dsam-adapter: Holistic adaptation of sam from 2d to 3d for promptable medical image segmentation," *arXiv preprint arXiv:2306.13465*, 2023.

[44] C. Zhang, L. Liu, Y. Cui, G. Huang, W. Lin, Y. Yang, and Y. Hu, "A comprehensive survey on segment anything model for vision and beyond," *arXiv preprint arXiv:2305.08196*, 2023.

[45] M. Ahmadi, M. F. Nia, S. Asgarian, K. Danesh, E. Irankhah, A. G. Lonbar, and A. Sharifi, "Comparative analysis of segment anything model and u-net for breast tumor detection in ultrasound and mammography images," *arXiv preprint arXiv:2306.12510*, 2023.

[46] M. Hu, Y. Li, and X. Yang, "Breastsam: A study of segment anything model for breast tumor detection in ultrasound images," *arXiv preprint arXiv:2305.12447*, 2023.

[47] F. Hörst, M. Rempe, L. Heine, C. Seibold, J. Keyl, G. Baldini, S. Ugurel, J. Siveke, B. Grünwald, J. Egger *et al.*, "Cellvit: Vision transformers for precise cell segmentation and classification," *Medical Image Analysis*, p. 103143, 2024.

[48] D. Shin, B. Kim, MD, and S. Baek, "Cemb-sam : Segment anything model with condition embedding for joint learning from heterogeneous datasets," in *International Conference on Medical Image Computing and Computer-Assisted Intervention.* Springer, 2023, pp. 275–284.

[49] S. He, R. Bao, J. Li, P. E. Grant, and Y. Ou, "Computer-vision benchmark segment-anything model (sam) in medical images: Accuracy in 12 datasets," *arXiv preprint arXiv:2304.09324*, 2023.

[50] K. Zhang and D. Liu, "Customized segment anything model for medical image segmentation," *arXiv preprint arXiv:2304.13785*, 2023.

[51] Y. Gao, W. Xia, D. Hu, and X. Gao, "Desam: Decoupling segment anything model for generalizable medical image segmentation," *arXiv preprint arXiv:2306.00499*, 2023.

[52] J. Wu, R. Fu, H. Fang, Y. Liu, Z. Wang, Y. Xu, Y. Jin, and T. Arbel, "Medical sam adapter: Adapting segment anything model for medical image segmentation," *arXiv preprint arXiv:2304.12620*, 2023.

[53] D. Cheng, Z. Qin, Z. Jiang, S. Zhang, Q. Lao, and K. Li, "Sam on medical images: A comprehensive study on three prompt modes," *arXiv preprint arXiv:2305.00035*, 2023.

[54] S. Mohapatra, A. Gosai, and G. Schlaug, "Brain extraction comparing segment anything model (sam) and fsl brain extraction tool," *arXiv preprint arXiv:2304.04738*, 2023.

[55] N.-T. Bui, D.-H. Hoang, M.-T. Tran, and N. Le, "Sam3d: Segment anything model in volumetric medical images," *arXiv preprint arXiv:2309.03493*, 2023.

[56] J. Ma, Y. He, F. Li, L. Han, C. You, and B. Wang, "Segment anything in medical images," *Nature Communications*, vol. 15, no. 1, p. 654, 2024.

[57] R. Deng, C. Cui, Q. Liu, T. Yao, L. W. Remedios, S. Bao, B. A. Landman, L. E. Wheless, L. A. Coburn, K. T. Wilson *et al.*, "Segment anything model (sam) for digital pathology: Assess zero-shot segmentation on whole slide imaging," *arXiv preprint arXiv:2304.04155*, 2023.

[58] Y. Huang, X. Yang, L. Liu, H. Zhou, A. Chang, X. Zhou, R. Chen, J. Yu, J. Chen, C. Chen *et al.*, "Segment anything model for medical images?" *Medical Image Analysis*, vol. 92, p. 103061, 2024.

[59] F. Putz, J. Grigo, T. Weissmann, P. Schubert, D. Hoefler, A. Gomaa, H. B. Tkhayat, A. Hagag, S. Lettmaier, B. Frey *et al.*, "The segment anything foundation model achieves favorable brain tumor autosegmentation accuracy on mri to support radiotherapy treatment planning," *arXiv preprint arXiv:2304.07875*, 2023.

[60] Y. Zhang and R. Jiao, "Towards segment anything model (sam) for medical image segmentation: a survey," *arXiv [Preprint]*, 2023.

[61] C. Hu and X. Li, "When sam meets medical images: An investigation of segment anything model (sam) on multi-phase liver tumor segmentation," *arXiv preprint arXiv:2304.08506*, 2023.

[62] T. H. Rafi, "An ensemble deep transfer-learning approach to identify covid-19 cases from chest x-ray images," in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. Via del Mar, Chile: IEEE, Oct 2020, p. 1–5. [Online]. Available: https://ieeexplore.ieee.org/document/9277695/

[63] X. Xue *et al.*, "Design and analysis of a deep learning ensemble framework model for the detection of covid-19 and pneumonia using large-scale ct scan and x-ray image datasets," *Bioengineering*, vol. 10, no. 3, p. 363, Mar 2023.

[64] N. A. Barsha, A. Rahman, and M. Mahdy, "Automated detection and grading of invasive ductal carcinoma breast cancer using ensemble of deep learning models," *Computers in Biology and Medicine*, vol. 139, p. 104931, Dec 2021.

[65] A. M. Hafiz, R. A. Bhat, and M. Hassaballah, "Image classification using convolutional neural network tree ensembles," *Multimedia Tools and Applications*, vol. 82, no. 5, p. 6867–6884, Feb 2023.

[66] M. S. Akter *et al.*, "Multi-class skin cancer classification architecture based on deep convolutional neural network," in *2022 IEEE Intl Conference on Big Data*

(Big Data). Osaka, Japan: IEEE, Dec 2022, p. 5404–5413. [Online]. Available: https://ieeexplore.ieee.org/document/10020302/

[67] S. Vallabhajosyula *et al.*, "Transfer learning-based deep ensemble neural network for plant leaf disease detection," *Journal of Plant Diseases and Protection*, vol. 129, no. 3, pp. 545–558, 2022.

[68] V. Anand *et al.*, "Weighted average ensemble deep learning model for stratification of brain tumor in mri images," *Diagnostics*, vol. 13, no. 7, p. 1320, Apr 2023.

[69] K. Shaga Devan *et al.*, "Weighted average ensemble-based semantic segmentation in biological electron microscopy images," *Histochemistry and Cell Biology*, pp. 1–16, 2022.

[70] Sartaj, "Brain Tumor Classification (MRI) — kaggle.com," https://www.kaggle.com/datasets/sartajbhuvaji/brain-tumor-classification-mri, May 2020.

[71] HASHIRA, "Br35h :: Brain tumor detection 2020 dataset," https://universe.roboflow.com/hashira-fhxpj/br35h-::-brain-tumor-detection-2020, oct 2022, visited on 2024-04-24. [Online]. Available: https://universe.roboflow.com/hashira-fhxpj/br35h-::-brain-tumor-detection-2020

[72] N. Chakrabarty, "Brain mri images for brain tumor detection," Apr 2019. [Online]. Available: https://www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection

[73] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of transfer learning," *Journal of Big data*, vol. 3, pp. 1–40, 2016.

[74] J. A. L. Marques, F. N. B. Gois, J. P. do Vale Madeiro, T. Li, and S. J. Fong, "Artificial neural network-based approaches for computer-aided disease diagnosis and treatment," in *Cognitive and Soft Computing Techniques for the Analysis of Healthcare Data*, ser. Intelligent Data-Centric Systems, A. K. Bhoi, V. H. C. de Albuquerque, P. N. Srinivasu, and G. Marques, Eds. Academic Press, 2022, pp. 79–99. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780323857512000086

[75] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[76] 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.

[77] A. Abedalla, M. Abdullah, M. Al-Ayyoub, and E. Benkhelifa, "Chest x-ray pneumothorax segmentation using u-net with efficientnet and resnet architectures," *PeerJ Computer Science*, vol. 7, p. e607, 2021.

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

[79] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," *Engineering Applications of Artificial Intelligence*, vol. 115, p. 105151, 2022.

[80] P. Liashchynskyi and P. Liashchynskyi, "Grid search, random search, genetic algorithm: a big comparison for nas," *arXiv preprint arXiv:1912.06059*, 2019.

[81] "What is Hyperparameter Tuning? - Hyperparameter Tuning Methods Explained - AWS — aws.amazon.com," https://aws.amazon.com/what-is/hyperparameter-tuning, [Accessed 24-04-2024].

[82] M. Hossin and M. N. Sulaiman, "A review on evaluation metrics for data classification evaluations," *International journal of data mining & knowledge management process*, vol. 5, no. 2, p. 1, 2015.

VITA AUCTORIS

NAME: Sumaiya Deen Muhammad

PLACE OF BIRTH: Chattogram, Bangladesh

YEAR OF BIRTH: 1992

EDUCATION: International Islamic University Chittagong, B.Sc in Computer Science and Engineering, Chattogram, Bangladesh, 2014

University of Windsor, M.Sc in Computer Science (Artificial Intelligence Specialization), Windsor, Ontario, 2024