AI-Driven Medical Imaging Platform: Advancements in Image Analysis and Healthcare Diagnosis

Waleed Salah Eldin[a], Ahmed Kaboudan[b]

[a] Research engineer, DigiBrain4, USA.

[b] CRO, DigiBrain4, USA, visiting professor Shorouk Academy, EG.

Abstract

In the realm of healthcare, the integration of artificial intelligence (AI) has revolutionized medical imaging analysis [1, 2, 3]. This research paper presents a comprehensive medical imaging platform powered by artificial intelligence (AI) techniques for automated image analysis. The platform incorporates specialized AI models for image classification using ResNet50 [4], object detection using YOLOv5l [7] and segmentation using Res50-U-Net [10]. The automated pipeline seamlessly categorizes incoming medical images, directing them to appropriate analysis modules - MRI and X-ray images trigger object detection to identify abnormalities; MRI brain images undergo additional tumor segmentation. Through three pivotal phases focused on classification, detection and segmentation, this research demonstrates an effective framework to harness AI for enhanced efficiency and precision in medical image evaluation, paving the pathway towards improved clinical diagnostics and patient care. The platform's automated workflow for progressive image analysis sets it apart from existing systems. Outcomes indicate AI's immense potential to transform medical imaging, assisting clinicians through actionable insights while mitigating subjectivity and variability in manual evaluation.

1. INTRODUCTION

The infusion of artificial intelligence (AI) techniques into medical imaging analysis has emerged as a transformative innovation in healthcare [1, 3, 11]. However, manual evaluation of scans is time-intensive and reliant on subjective human interpretation. This research paper embarks on developing an AI-based platform to automate and enhance efficiency in medical image analysis - focusing specifically on classification, abnormality detection and segmentation [2, 12, 13, 14].

The motivation for this undertaking is rooted in the relentless pursuit of precision and efficiency in healthcare diagnostics [3, 11]. Manual examination of medical images, while invaluable, often proves to be time-intensive and susceptible to variability. In response to these challenges, our project, delves into the core AI-driven components poised to revolutionize medical image analysis [1, 3, 11].

Our journey unfolds through three essential phases, each fortified by cutting-edge AI techniques and models [2, 3, 15,]. The classification phase harnesses the power of AI to categorize medical images with unparalleled accuracy

[12, 16, 17]. Object detection ushers in the ability to efficiently identify anomalies in X-ray images and detect tumors in MRI brain images [7, 18, 19]. The segmentation phase presents a meticulous endeavor, enabling the precise delineation of tumors from MRI brain images [20, 21, 22].

Our pipeline offers three pivotal contributions:

1- Automated workflow to direct incoming scans to appropriate AI models, ensuring specialized analysis. MRI and X-ray trigger object detection; Brain MRI activates additional tumor segmentation.

2- Ensemble of AI architectures including ResNet50 [4], YOLOv5l [7] and Res50-U-Net [10] to enable robust classification, anomaly detection and precise tumor delineation.

3- Seamless integration of multiple phases - classification, detection and segmentation - into one extremely high-precision diagnostic platform powered by deep learning.

The key motivations underlying this research are two-fold:

a) to mitigate variability and subjectivity inherent in manual evaluation of scans

b) to enable clinicians to leverage AI insights and devote greater attention towards understanding anomalies and determining optimal treatments.

By harnessing automated pipelines and diverse AI models, our platform aims to transform efficiency, consistency and accuracy in medical image diagnostics - transitioning towards data-driven precision care and ultimately driving better clinical decisions and patient outcomes. The methods and results sections elucidate the models and architectures employed to realize these goals.

2. *LITERATURE REVIEW*

The integration of artificial intelligence (AI) into the field of medical imaging has ignited a wave of transformative innovation and research [1, 2, 11]. This section conducts a comprehensive review of existing literature, illuminating the pivotal role of AI in revolutionizing medical image analysis and its wide-ranging applications.

2.1 Related work

Our review focuses on existing research at the intersection of AI and medical imaging spanning image classification, object detection and segmentation. While substantial literature explores individual methods in isolation, our platform's novelty lies in progressively automating the workflow - allowing seamless analysis via multiple AI models.

2.2 AI in Medical Imaging: A Paradigm Shift

The infusion of AI techniques into medical imaging represents a significant paradigm shift in healthcare [1, 3, 11]. Conventional diagnostic methods, though reliable, are often time-consuming and reliant on subjective human interpretation. AI offers a promising avenue to enhance diagnostic accuracy, reduce human variability, and expedite decision-making [3, 15, 23].

2.3 Image Classification and Diagnosis

AI-driven image classification has emerged as a cornerstone in medical image analysis [12, 16, 17]. Deep learning architectures have demonstrated exceptional capabilities in categorizing medical images accurately [13, 25, 26]. These models excel in distinguishing a wide range of medical conditions, from detecting abnormalities in X-ray images to categorizing MRI brain scans into distinct categories [12, 16, 17].

2.4 Object Detection and Localization

AI-powered object detection has introduced efficient means to identify anomalies and specific features within medical images [2, 15, 26]. These models have proven highly effective in detecting abnormalities in X-ray images and localizing relevant structures in MRI brain images [7, 18, 19]. This capability expedites the diagnostic process and empowers healthcare professionals with precise insights.

2.5 Image Segmentation for Enhanced Precision

AI-based image segmentation techniques have redefined the precision of medical image analysis [2, 3, 15]. These techniques excel in delineating structures and anomalies within medical images [12, 16, 17]. In the context of MRI brain images, these methods have shown promise in segmenting tumors, offering invaluable support to medical practitioners [20, 21, 22].

2.6 Challenges and Future Directions

While the potential of AI in medical imaging is undeniable, challenges remain. Ensuring the robustness and interpretability of AI models is a priority [27, 28, 29], along with addressing ethical considerations and data privacy concerns [30, 31, 32]. Additionally, seamless integration into clinical workflows and collaborative research efforts present opportunities for further exploration [27, 28, 33].

However, current systems function as standalone solutions targeted towards specific tasks like classification or segmentation. By contrast, our pipeline orchestrates an end-to-end workflow - directly ingesting raw scans and channeling them through specialized AI architectures to generate insights tailored to the image type. This automated direction of scans eliminates manual sorting of images and piecing together of multiple models.

While aspects such as coordinated pipelines and ensemble modeling have gained traction in domains like autonomous vehicles, their application in medicine remains relatively nascent. Our platform aims to address this gap through its unique workflow.

In summary, although AI techniques have permeated medical imaging to assist aspects like classification and detection individually, few systems attempt to integrate multiple modalities into a streamlined workflow. Our platform generates a robust framework that leverages the collective strengths of classification, localization and segmentation models within one automated architecture. Findings validate enhanced efficiency alongside accuracy.

3. *METHODOLOGY*

The methodology section outlines the comprehensive approach undertaken in the research, focusing exclusively on the AI-driven components of the medical imaging platform. It details the AI models and architectures employed, data preprocessing strategies, dataset sources, the training process with performance evaluation metrics, and our meticulously designed automatic pipeline for medical image analysis [34, 35, 36].

3.1 AI Models and Architectures

This research leveraged a repertoire of state-of-the-art AI models to address distinct facets of medical image analysis. The models included ResNet50 [4], DenseNet121 [5], and VGG16 [6], for image classification. In object detection, YOLOv5l [7], was employed to efficiently identify anomalies in X-ray images and detect tumors in MRI brain images. For image segmentation, specialized architectures, namely U-Net [8], Attention U-Net [9], and Res50-U-Net [10], were developed to meticulously delineate structures within MRI brain images.

3.2 Data Preprocessing and Augmentation

Data preprocessing played a pivotal role in enhancing the quality of the training data. This involved standardizing pixel values, resizing images to a consistent format, and normalizing intensity levels. For object detection, bounding boxes were annotated to denote regions of interest.

Dataset sources encompassed a diverse range of medical institutions and publicly available repositories, ensuring the inclusion of varied clinical scenarios and imaging equipment, All the datasets were from Stanford AIMI Shared Datasets which are BrainMetShare, EchoNet-Dynamic, LERA, MRNET Knee MRI’s MURA and OL3I And only 1 dataset from Kaggle which is: Brain MRI segmentation, the only data augmentation we did was to make images in grayscale to enhance model robustness.

3.3 Training Process and Performance Evaluation

The training process commenced with the division of the dataset into training, validation, and test sets to assess model generalization. Models were trained on high-performance computing clusters, utilizing GPUs to expedite training.

Performance evaluation metrics varied according to the specific task. In image classification, metric was accuracy. For object detection, metrics involved mean average precision (mAP), intersection over union (IoU), and Confusion Matrix. In image segmentation, metrics encompassed dice coefficient, and intersection over union (IoU).

3.4 Automatic pipeline of Image Categorization

Upon uploading, medical images traverse an automated journey. The initial checkpoint involves the "General Classifier," which promptly identifies the primary image category among CT, MRI, Sonar, and X-ray. This swift categorization is the foundation of subsequent analysis.

3.5 Refined Subcategory Identification

Following the general classification, the pipeline seamlessly funnels the images into specialized subcategory classifiers. For MRI images, the "MRI Classifier" further classifies them into either "Brain" or "Knee" images. For MRI Knee images, an additional "MRI Knee Classifier" discerns the orientation, categorizing them as "Axial," "Sagittal," or "Coronal." Similarly, for X-ray images, the "X-ray Classifier" fine-tunes the classification, identifying specific regions of interest such as "Elbow," "Finger," "Forearm," "Hand," "Humerus," "Shoulder," "Wrist," and "Joint."

3.6 Object Detection and Segmentation

Elevating the capabilities of the pipeline, specific categories trigger additional layers of analysis. When the primary category is "X-ray" or "MRI Brain," the pipeline activates the "Object Detector" module. This component diligently identifies objects of interest, such as anomalies or structures. In the case of "MRI Brain" images, an additional "Segmentation Model" takes center stage, meticulously segmenting tumors, thereby aiding medical practitioners in precise diagnosis and treatment planning.

This methodology encapsulates the rigorous approach undertaken in the AI-driven components of the medical imaging platform. It sets the foundation for subsequent sections, providing insights into the strategies employed to develop and evaluate AI models for enhanced medical image analysis and our automated pipeline, characterized by its hierarchical classification and specialized analysis modules, sets our platform apart, ensuring the efficient processing of diverse medical images. It exemplifies our commitment to harnessing AI for the advancement of medical image analysis, ultimately contributing to improved healthcare diagnostics and patient care.

4. Classification Phase

In this phase we focused on developing a classification system that could accurately identify different distinct types of medical images we implemented the following classifiers: General, X-ray, MRI, and MRI Knee classifier.
The goal of this phase was to create a reliable tool that could help medical professionals quickly and accurately classify and identify different types of medical images.
To achieve this, we employed advanced deep learning architectures to develop a robust classification model that could accurately identify each type of medical image based on unique features and characteristics.
We took an experimental path in our project as we mentioned before and we implemented 3 different architectures and we ended up choosing only 1 of them.
This phase of the project was critical for establishing a strong foundation upon which we could build additional functionality and features in future phases.

By successfully developing a reliable classification system, we have taken an important first step towards building a comprehensive medical imaging platform.

4.1 Models description:

4.1.1 ResNet50:

short for Residual Network with 50 layers, is a deep convolutional neural network architecture that introduced the concept of residual learning [4]. It was designed to address the challenges of training very deep neural networks [4].

Components:

Basic Building Block (Residual Block): The key innovation in ResNet is the residual block, which contains skip connections or shortcuts [4]. These shortcuts enable the network to learn residual functions, making it easier to train deeper networks [4].

Bottleneck Architecture: ResNet50 uses a bottleneck architecture in its residual blocks [4]. This involves using 1x1, 3x3, and 1x1 convolutions to reduce the dimensionality before and after the 3x3 convolution, which helps in improving computational efficiency [4].

Global Average Pooling (GAP): Instead of using fully connected layers at the end of the network, ResNet50 employs Global Average Pooling to reduce spatial dimensions and generate a fixed-size vector regardless of the input size [4].

Fully Connected Layer (Dense Layer): The final layer of the network that produces the output predictions [4].

Layers:

Input Layer: Accepts the input image data [4].

Convolutional Layers: Several convolutional layers make up the initial part of the network, responsible for feature extraction [4].

Residual Blocks: These blocks contain the residual connections and bottleneck architecture [4]. ResNet50 consists of 16 residual blocks grouped into four stages [4].

Global Average Pooling Layer: Reduces the spatial dimensions of the feature maps to a fixed size [4].

Fully Connected Layer: Produces the final output predictions [4].

Total Parameters: Approximately 25.6 million parameters [4].

4.1.2 DenseNet121:

It’s a convolutional neural network architecture known for its densely connected blocks, where each layer receives input from all preceding layers [5]. This connectivity pattern encourages feature reuse and enhances model compactness [5].

Components:

Dense Block: The fundamental building block of DenseNet. In a dense block, each layer receives input from all preceding layers, fostering dense connectivity [5].

Transition Layer: Situated between two dense blocks, the transition layer reduces the spatial dimensions (width and height) of the feature maps by employing a combination of convolution and pooling operations [5]. It also reduces the number of feature maps [5].
Global Average Pooling (GAP): Similar to ResNet, DenseNet typically uses global average pooling to reduce spatial dimensions before the final fully connected layer [5].
Fully Connected Layer (Dense Layer): The last layer that produces the final output predictions [5].

Layers:
Input Layer: Accepts the input image data [5].
Initial Convolutional Layer: The first layer that performs convolutional operations to extract initial features [5].
Dense Blocks: Successive dense blocks form the core of DenseNet121 [5]. Each dense block consists of multiple densely connected layers [5].
Transition Layers: These layers reduce the spatial dimensions and the number of feature maps between dense blocks [5].
Global Average Pooling Layer: Reduces the spatial dimensions of the feature maps to a fixed size [5].
Fully Connected Layer: Produces the final output predictions [5].

Total Parameters: Approximately 8 million parameters [5].

4.1.3 VGG16:

Visual Geometry Group 16, is a deep convolutional neural network architecture known for its simplicity and uniform architecture [6]. It was proposed by the Visual Geometry Group at the University of Oxford [6].

Components:

Convolutional Layers: VGG16 consists of 13 convolutional layers, which are stacked one after another [6]. These layers have small 3x3 convolutional filters and are responsible for feature extraction [6].
Max Pooling Layers: After each group of convolutional layers, there is a max-pooling layer with a 2x2 filter and a stride of 2 [6]. Max pooling helps reduce the spatial dimensions of the feature maps [6].
Fully Connected Layers: The last three layers of VGG16 are fully connected layers responsible for classification [6]. The first two fully connected layers have 4,096 neurons each, and the final fully connected layer corresponds to the number of classes in the classification task [6].
Softmax Activation: The final layer uses the Softmax activation function to convert the network's output into class probabilities [6].

Layers:
Input Layer: Accepts the input image data [6].
Convolutional Layers (Conv Blocks): The 13 convolutional layers are organized into five groups or blocks [6]. Each block typically consists of multiple consecutive convolutional layers followed by a max-pooling layer [6].
Fully Connected Layers: The last three layers are fully connected layers [6]. The first two have 4,096 neurons each,

and the final layer has neurons equal to the number of classes for the specific task [6].
Softmax Layer: The final layer with Softmax activation for classification [6].

Total Parameters: Approximately 138 million parameters [6].

4.2 Dataset analysis:

We analyzed the datasets for our first phase and we found out the following results:
X-ray got 8 sub categories and they are 41276 photos in total and they are: Elbow: 5396, Finger: 5553, Forearm: 2124, Hand: 5989, Humerus: 1560, Shoulder: 8942, Wrist: 10415, Foot: 348, Joint {Knee: 534, Ankle: 321, Hip: 94}.

Due that Foot, Knee, Ankle and Hip had low amount of images we merged them into 1 category which is Joint

CT got 1 sub category and it is L3: 8139, MRI got 2 sub categories and they are in total 6595 and they are: Head: 2845, Knee: 3750, Sonar got 1 sub category and it is Heart: 10030.

We decided to split the datasets as follow:

General dataset which will be used for the General classifier is 12000 images and was split as follows:

Train	valid	Test
9600 images	1600 images	800 images
2400 images for each category: XR, MRI, Sonar and CT.	400 images for each category: XR, MRI, Sonar and CT.	200 images for each category: XR, MRI, Sonar and CT.

XR dataset which will be used for the XR classifier is 10400 images and was split as follows:

Train	valid	Test
8000 images	1600 images	800 images
1000 images for each category: Elbow, Finger, Forearm, Hand, Humerus, Shoulder, Wrist and Joint.	200 images for each category: Elbow, Finger, Forearm, Hand, Humerus, Shoulder, Wrist and Joint.	100 images for each category: Elbow, Finger, Forearm, Hand, Humerus, Shoulder, Wrist and Joint.

MRI dataset which will be used for the MRI classifier is 3000 images and was split as follows:

Train	valid	Test
2400 images	400 images	200 images
1200 images for each category: Brain and Knee	200 images for each category: Brain and Knee	100 images for each category: Brain and Knee

MRI_Knee dataset which will be used for the MRI_Knee classifier is 1497 images and was split as follows:

Train	valid	Test
1200 images	197 images	100 images
400 images for each category: Axial, coronal and sagittal	Axial: 68 images Coronal: 64 images and sagittal: 65 images	Axial: 33 images Coronal: 33 images and sagittal: 34 images

4.3 Classification phase results:

As we mentioned before we took an experimental path and we implemented 3 architectures ResNet50 [4], DenseNet101 [5], and VGG16 [6], now we going to show the results we got and which one did we choose.

Architectures: Datasets:	ResNet50	DenseNet121	VGG16
General	Train accuracy: 100% Valid accuracy: 99.8125% Test accuracy: 100%	Train accuracy: 99% Valid accuracy: 98% Test accuracy: 99%	Train accuracy: 99% Valid accuracy: 97% Test accuracy: 99%
XR	Train accuracy: 100% Valid accuracy: 94.1875% Test accuracy: 98%	Train accuracy: 97% Valid accuracy: 90% Test accuracy: 89%	Train accuracy: 91% Valid accuracy: 83% Test accuracy: 85%
MRI	Train accuracy: 100% Valid accuracy: 100% Test accuracy: 100%	Train accuracy: 100% Valid accuracy: 100% Test accuracy: 100%	Train accuracy: 100% Valid accuracy: 100% Test accuracy: 100%
MRI_Knee	Train accuracy: 100% Valid accuracy: 100% Test accuracy: 100%	Train accuracy: 100% Valid accuracy: 100% Test accuracy: 100%	Train accuracy: 99.9% Valid accuracy: 100% Test accuracy: 100%

We ended up choosing ResNet50 [4], because it had the best results among the 3 architectures on all datasets.

4.4 Impact on Accurate Classification

The impact of accurate medical image classification cannot be overstated. By curating a diverse dataset and standardizing image properties, we minimized bias and maximized the models' ability to handle real-world variations in medical images. This inclusivity was especially crucial in medical imaging, where variations in imaging devices and patient populations are common.

The dataset's richness also allowed our AI models to excel in distinguishing between a multitude of medical conditions within each category. Whether it was identifying abnormalities in X-ray images or classifying MRI scans into "Brain" or "Knee" categories, the dataset's comprehensiveness contributed to the models' exceptional accuracy.

5. OBJECT DETECTION PHASE

In this phase we extended our system to include object detection using state-of-the-art deep learning models [18, 19, 37].
This involved detecting and localizing specific objects within an image, rather than simply classifying the entire image as a certain category.
The main objective was to develop a system that could accurately detect and locate positive X-ray images and tumors in MRI brain images.
This is a critical task in medical imaging analysis, as early and accurate detection of abnormalities can have a significant impact on patient outcomes.
To achieve this objective, we used YOLOv5 L [7], which is one of the most popular and widely used deep learning models for object detection.
Our main task in this phase was to train the YOLOv5 [7], model on a dataset of X-ray and MRI images.
The dataset consisted of positive X-ray images and MRI brain images with tumors, and the model was trained to accurately detect the presence of tumors in the brain images and identify positive X-ray images.
Overall, this phase of this project represents a significant advancement in the capabilities of our medical imaging analysis system.
By incorporating object detection into our approach, we have enabled more nuanced and accurate analysis of medical images, which has the potential to improve patient outcomes and save lives.

5.1 Model description:

Components:

Backbone Architecture: YOLOv5L typically uses CSPDarknet53 as its backbone architecture [7]. CSPDarknet53 is a variant of Darknet, and CSP stands for "Cross-Stage Partial Networks," a feature designed to improve information flow [7].
Neck Architecture: YOLOv5L includes a PANet (Path Aggregation Network) as its neck architecture [7]. PANet helps aggregate multi-scale features for better object detection [7].

Detection Head: The detection head consists of multiple convolutional layers responsible for predicting bounding boxes, objects scores, and class probabilities [7].

Layers:

Backbone Layers: These include the layers in the CSPDarknet53 backbone architecture [7].

Neck Layers: The PANet layers, which help aggregate features from different scales [7].

Detection Head Layers: Multiple convolutional layers responsible for object detection predictions [7].

Total Parameters: Approximately 46.5 million parameters [7].

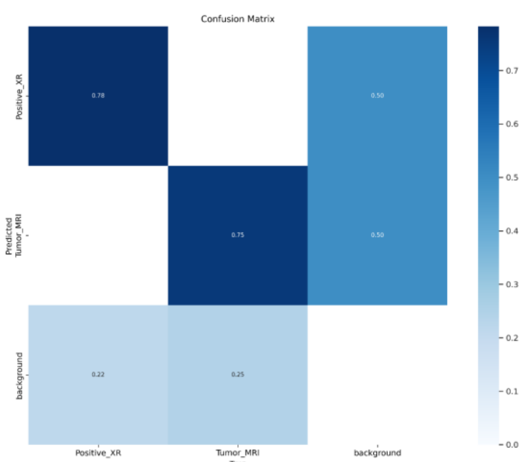
5.2 Dataset analysis:

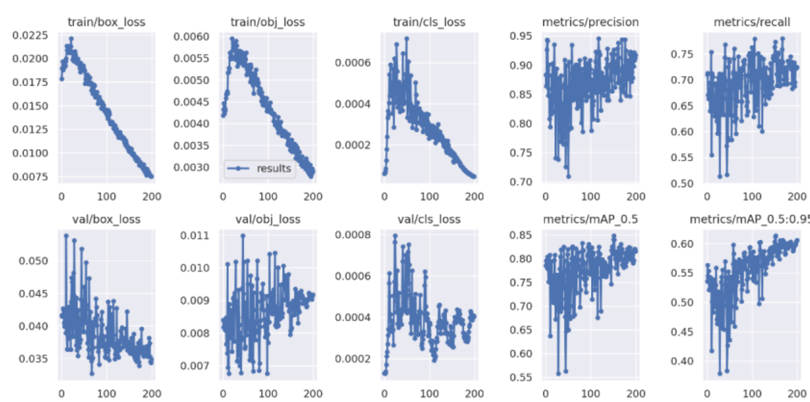
the dataset was as follow:

Train	Valid	Test
829 total images	205 total images	85 total images
437 MRI images 392 X-ray images	93 MRI images 112 X-ray images	30 MRI images 55 X-ray

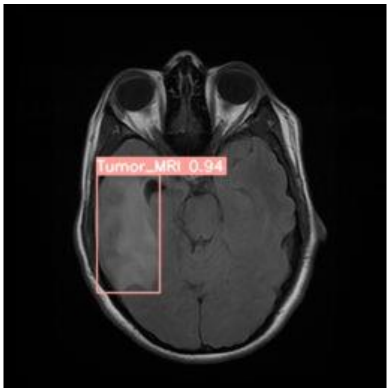
5.3 Object Detection phase results:

The Positive XR: 78%, Tumor MRI:75%





Here are some examples of the medical images results:



6. *SEGMENTATION PHASE*

In the third phase of our project, we focused on developing a segmentation model to accurately segment tumors from MRI brain images [20, 21, 22].

Segmentation is an important technique in medical image analysis as it can aid in the diagnosis and treatment of various diseases.

Our goal was to create a model that could accurately segment brain tumors with high precision and recall rates.

Accurate tumor segmentation in medical imaging is crucial for accurate diagnosis, treatment planning, and monitoring of cancer patients.
In medical imaging, segmentation refers to the process of separating the region of interest, such as a tumor, from the surrounding tissue.
This enables clinicians to precisely locate the tumor, determine its size and shape, and monitor changes over time.
Accurate tumor segmentation can have a significant impact on the field of healthcare as it can improve the accuracy of cancer diagnosis and staging, leading to more appropriate treatment planning and better patient outcomes.
It can also help clinicians monitor the response to treatment and detect any recurrence or metastasis.
Moreover, accurate segmentation can help reduce the subjectivity of visual inspection, thereby improving the consistency and reproducibility of diagnosis and treatment planning.
We also approached this phase in an experimental way, we implemented 3 architectures U-Net [9], Attention U-Net [9], and Res50 U-Net [10].
In medical imaging, segmentation plays a crucial role in the accurate diagnosis of brain tumors, making this phase a critical aspect of our overall project.

In summary, accurate tumor segmentation is a critical component of medical imaging and has the potential to transform cancer diagnosis, treatment, and research.

6.1 Models description:

6.1.1 U-Net:

U-Net is a convolutional neural network architecture designed for semantic image segmentation [8]. It consists of a contracting path, a bottleneck, and an expansive path [8].

Components:

Contracting Path (Encoder): The top part of the U-Net architecture involves a series of convolutional and pooling layers that gradually reduce the spatial dimensions while increasing the number of channels [8]. These layers capture context and features from the input image [8].
Bottleneck: The bottleneck is a central part of the network where the spatial information is compressed into a feature representation [8]. It typically involves multiple convolutional layers [8].
Expansive Path (Decoder): The bottom part of the U-Net architecture consists of up-sampling and concatenation operations to restore the spatial dimensions [8]. This path helps generate a high-resolution segmentation map [8].

Layers:

Input Layer: Accepts the input image data [8].
Contracting Path (Encoder) Layers: Typically consists of a series of convolutional layers followed by a pooling operation (commonly max pooling or stride-2 convolution) [8].
Bottleneck Layers: Multiple convolutional layers to compress spatial information [8].
Expansive Path (Decoder) Layers: Consists of up-sampling operations (commonly transposed convolutions or bilinear interpolation) and concatenation with corresponding feature maps from the contracting path [8]. This path gradually restores the spatial dimensions [8].
Output Layer: Produces the final segmentation map [8].

6.1.2 Attention U-Net:

The Attention U-Net is an extension of the U-Net architecture that incorporates attention mechanisms to improve its performance in semantic segmentation tasks [9].

Components:

Encoder (Contracting Path):
Convolutional layers with 3x3 kernels and rectified linear unit (ReLU) activation [9].
Down-sampling through max-pooling (2x2) [9]. Attention modules added to capture important contextual information [9].
Bottleneck:
Convolutional layers with 3x3 kernels and ReLU activation [9].
Attention module to enhance the relevant features [9].
Decoder (Expansive Path):
Up-sampling through transposed convolutions [9].
Concatenation with corresponding feature maps from the contracting path [9].
Convolutional layers with 3x3 kernels and ReLU activation [9].
Attention modules for refined feature selection [9].

Output Layer: Convolutional layer with a single channel for binary segmentation or multiple channels for multiclass segmentation [9].
Attention Modules: Attention mechanisms such as Channel Attention and Spatial Attention are integrated to help the model focus on important regions and features [9].

Layers:

Input Layer: Accepts the input image data [9].
Encoder (Contracting Path) Layers: Convolutional layers with down-sampling through max-pooling [9]. Attention modules are typically inserted in this path to capture contextual information [9].
Bottleneck Layers: Convolutional layers in the central part of the network with an attention module for enhanced feature selection [9].
Decoder (Expansive Path) Layers: Convolutional layers with up-sampling through transposed convolutions [9]. Concatenation with feature maps from the contracting path, and attention modules for refined feature integration [9].
Output Layer: Produces the final segmentation map [9].

6.1.3 Res50 U-Net

The Res50 U-Net architecture marries the distinctive strengths of the U-Net, renowned for its effective feature mapping and spatial localization, with the robust ResNet-50, celebrated for its deep residual learning capabilities [10]. This amalgamation is engineered to address the challenges posed by intricate segmentation tasks, particularly in the context of medical imaging or fine-grained object delineation [10].

Components:

ResNet-50 Backbone:

ResNet-50 is a deep residual network with 50 layers [10]. It consists of residual blocks, each containing multiple convolutional layers [10].

Encoder (Contracting Path):
Utilizes the ResNet-50 backbone to extract hierarchical features [10].
May involves down-sampling through strides or pooling operations [10].

Bottleneck: May include additional convolutional layers to further compress features [10].

Decoder (Expansive Path): Up-sampling through transposed convolutions or other up-sampling methods [10]. Concatenation with corresponding feature maps from the contracting path [10].

Output Layer: Convolutional layer with a single channel for binary segmentation or multiple channels for multiclass segmentation [10].

6.2 Dataset analysis:

The dataset was as follow:

Train	Valid	Test
1073 total images	200 total images	100 total images

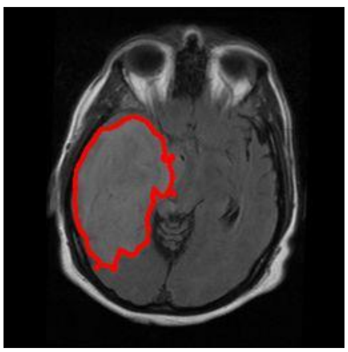
Segmentation phase results:

The results were as following:

U-Net	Attention U-Net	Res50 U-Net
0.78 IOU	0.85 IOU	0.93 IOU

We ended up choosing Res50 U-Net [10], because it had the best results among the 3 architectures.

Here are some examples of the medical images results:



7. RESULTS, DISCUSSION AND CONCLUSIONS

In this project, we have employed various techniques and methodologies to develop an AI-based medical image classification, Object detection and segmentation system. As with any project, the results, discussion, and conclusions section are a crucial aspect that summarizes the findings and analysis [25, 27, 33].

7.1 Summary of achieved results

During the course of this project, we have achieved significant results in the field of medical imaging. We started by developing a classification system using ResNet50 [4], DenseNet121 [5], and VGG16 [6], architectures to classify different types of medical images. We ended up choosing ResNet50 [4], and Our model achieved high accuracy rates and proved to be effective in differentiating between different classes of medical images.

In the second phase of the project, we worked on object detection using YOLOv5 [7], to detect positive XR images and tumors in MRI brain images. We successfully implemented the model and achieved high precision and recall rates.

In the third phase of the project, we focused on segmentation of tumors from MRI brain images using different U-Net architectures, including U-Net [8], Attention U-Net [9], and Res50 U-Net [10]. Our results showed that the Res50 U-Net [10], architecture performed better than the other architectures in terms of accuracy and speed.

Overall, the achieved results demonstrate the potential of using AI in medical imaging to assist doctors in accurate diagnosis and treatment. Our models have the potential to be integrated into healthcare systems to improve patient outcomes.

7.2 Possible extensions and further development

Possible extensions and further development for the project can be considered to enhance its functionality and capabilities.
Some potential areas for improvement include:
Integration with Electronic Health Records (EHR) - Integrating the system with EHR can allow for a more comprehensive patient history and facilitate more accurate diagnosis.

Improved Object Detection Models - Continuously improving the Object Detection models can increase the accuracy of tumor detection and positive XR and reduce false positives.

Improved Segmentation Models - Continuously improving the segmentation models can increase the accuracy of tumor segmentation and reduce false positives.

Expansion of organs and human parts - Expanding the system to support more organs and human parts such as XR on lungs for example, can increase its usefulness and applicability.

Enhanced Reporting - Developing enhanced reporting capabilities can provide more detailed and actionable insights to medical professionals.

Real-time Image Processing - Implementing real-time image processing capabilities can allow for faster diagnosis and treatment.

Integration with AI-assisted Diagnosis - Integrating the system with AI-assisted diagnosis can provide additional support to medical professionals and increase the accuracy of diagnosis.

Overall, these extensions and further developments can enhance the system's functionality and usefulness, ultimately improving patient outcomes and advancing the field of medical imaging.

References

[1] Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. Annual review of biomedical engineering, 19, 221-248.

[2] Greenspan, H., van Ginneken, B., & Summers, R. M. (2016). Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. IEEE Transactions on Medical Imaging, 35(5), 1153-1159.

[3] Suzuki, K. (2017). Overview of deep learning in medical imaging. Radiological physics and technology, 10(3), 257-273.

[4] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[5] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[6] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[7] Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767.

[8] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.

[9] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M., Heinrich, M., Misawa, K., ... & Rueckert, D. (2018). Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999.

[10] Alom, M. Z., Hasan, M., Yakopcic, C., Taha, T. M., & Asari, V. K. (2018). Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. arXiv preprint arXiv:1802.06955.

[11] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. Medical image analysis, 42, 60-88.

[12] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.

[13] Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., ... & Kim, R. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22), 2402-2410.

[14] Chartrand, G., Cheng, P. M., Vorontsov, E., Drozdzal, M., Turcotte, S., Pal, C. J., ... & Tang, A. (2017). Deep learning: a primer for radiologists. Radiographics, 37(7), 2113-2131.

[15] Han, X., & Leung, T. (2018). Applications of deep learning for fundus image analysis: A review. Computational and mathematical methods in medicine, 2018.

[16] Anthimopoulos, M., Christodoulidis, S., Ebner, L., Christe, A., & Mougiakakou, S. (2016). Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. IEEE transactions on medical imaging, 35(5), 1207-1216.

[17] Bychkov, D., Linder, N., Turkki, R., Nordling, S., Kovanen, P. E., Verrill, C., ... & Lundin, J. (2018). Deep learning based tissue analysis predicts outcome in colorectal cancer. Scientific reports, 8(1), 1-11.

[18] Li, Z., Wang, C., Han, M., Xue, Y., Wei, W., Li, L. J., & Fei, B. (2018). Thoracic disease identification and localization with limited supervision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 8290-8299).

[19] Yan, K., Wang, X., Lu, L., Zhang, L., Harrison, A. P., Bagheri, M., & Summers, R. M. (2018). Deep lesion graphs in the wild: relationship learning and organization of significant radiology image findings in a diverse large-scale lesion database. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 9261-9270).

[20] Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., ... & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. Medical image analysis, 35, 18-31.

[21] Pereira, S., Pinto, A., Alves, V., & Silva, C. A. (2016). Brain tumor segmentation using convolutional neural networks in MRI images. IEEE transactions on medical imaging, 35(5), 1240-1251.

[22] Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., ... & Glocker, B. (2017). Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical image analysis, 36, 61-78.

[23] Lee, J. G., Jun, S., Cho, Y. W., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep learning in medical imaging: general overview. Korean journal of radiology, 18(4), 570-584.

[24] Bi, L., Kim, J., Ahn, E., & Kumar, A. (2019). Automatic prostate zonal segmentation using fully convolutional network with feature pyramid attention. Medical physics, 46(10), 4646-4655.

[25] Liu, X., Faes, L., Kale, A. U., Wagner, S. K., Fu, D. J., Bruynseels, A., ... & Shamdas, M. (2019). A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. The Lancet Digital Health, 1(6), e271-e297.

[26] Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., ... & Yang, J. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell, 172(5), 1122-1131.

[27] Topol, E. J. (2019). High-performance medicine: the convergence of human and artificial intelligence. Nature medicine, 25(1), 44-56.

[28] Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., ... & Wang, Y. (2017). Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology, 2(4), 230-243.

[29] Benjamens, S., Dhunnoo, P., & Meskó, B. (2020). The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database. NPJ digital medicine, 3(1), 1-8.

[30] Char, D. S., Shah, N. H., & Magnus, D. (2018). Implementing machine learning in health care—addressing ethical challenges. The New England journal of medicine, 378(11), 981.

[31] Wiens, J., Saria, S., Sendak, M., Ghassemi, M., Liu, V. X., Doshi-Velez, F., ... & Najarian, K. (2019). Do no harm: a roadmap for responsible machine learning for health care. Nature medicine, 25(9), 1337-1340.

[32] Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. BMC medicine, 17(1), 1-9.

[33] Sendak, M., Gao, M., Nichols, M., Lin, A., & Balu, S. (2020). Presenting machine learning model information to clinical end users with model facts labels. NPJ digital medicine, 3(1), 1-7.

[34] Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 1-48.

[35] Wang, G., Li, W., Ourselin, S., & Vercauteren, T. (2019). Automatic brain tumor segmentation using convolutional neural networks with test-time augmentation. In International MICCAI Brainlesion Workshop (pp. 61-72). Springer, Cham.

[36] Wang, G., Li, W., Aertsen, M., Deprest, J., Ourselin, S., & Vercauteren, T. (2019). Aleatoric uncertainty estimation with test-time augmentation for medical image segmentation with convolutional neural networks. Neurocomputing, 338, 34-45.

[37] Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology, 284(2), 574-582.