INTERNATIONAL UNIVERSITY
VIETNAM NATIONAL UNIVERSITY – HO CHI MINH CITY
School of Computer Science and Engineering
-----***-----

PROJECT REPORT

MedVisionAI Website

Deep Learning (IT157IU)
Semester 1 - Academic year 2024-2025
Course by Dr. Mai Hoang Bao An

CONTRIBUTION TABLE

No.	Full Name	Student's ID	Task	Contribution
1	Trần Thanh Nguyên(Leader)	ITCSIU21093	Collect Data,Implement Models,Evaluate Models,Deploy Models	100%

Contents

ABSTRACT

The advancement of computer vision has opened new possibilities in medical image processing, enabling precise analysis and interpretation of complex medical data. This project focuses on leveraging cutting-edge computer vision models to address critical challenges in medical imaging, including tumor detection, disease classification, and image segmentation. By integrating advanced deep learning architectures such as convolutional neural networks (CNNs) and transformer-based models, this work aims to enhance diagnostic accuracy, streamline clinical workflows, and improve patient outcomes. The proposed system is trained on diverse medical datasets to ensure robustness across varying imaging modalities such as MRI, CT, and X-rays. Preliminary results demonstrate promising performance in identifying pathological features, segmenting anatomical structures, and classifying diseases with high accuracy. This research contributes to the growing field of AI-driven healthcare, highlighting the transformative potential of computer vision in medical diagnostics and treatment planning.

I INTRODUCTION

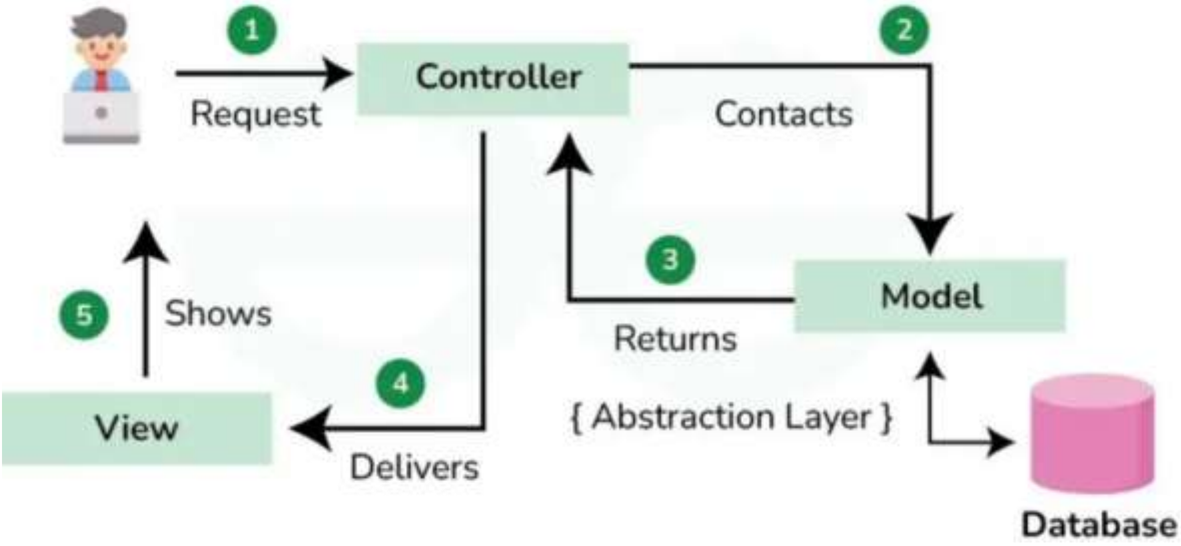
1. Objectives

The primary objectives of this project are:

1. Automated Tumor Detection:
   Develop a computer vision model capable of identifying tumors with high accuracy, reducing diagnostic time and assisting clinicians in early detection.
2. Disease Classification:
   Create a robust classification system to categorize medical images into specific diseases or conditions, aiding in faster and more reliable diagnoses.
3. Image Segmentation:
   Implement advanced segmentation techniques to delineate anatomical structures and pathological regions, enabling detailed analysis for treatment planning and monitoring.
4. Enhancement of Diagnostic Accuracy:
   Improve the reliability of diagnostic procedures by minimizing false positives and false negatives using state-of-the-art deep learning models.
5. Scalability and Adaptability:
   Ensure the system's adaptability across diverse imaging modalities (e.g., MRI, CT, and X-rays) and its scalability to handle large datasets from different medical domains.
6. Integration into Clinical Workflow:
   Facilitate the seamless integration of the developed system into clinical environments to assist radiologists and healthcare professionals.

2. ENVIRONMENT

Since this is a web-based product, the project is conducted using some Web Design and Programming Language under the model of MVC



The implementation of this project relies on a robust combination of programming languages, frameworks, and tools to ensure functionality, scalability, and ease of use. Key tools include:

1. Python Language: A versatile programming language used for data analysis, machine learning model development, and integration of AI-driven applications.
2. SQLAlchemy: A powerful database toolkit and Object Relational Mapper (ORM) for efficient data storage, retrieval, and management.
3. Flask Framework: A lightweight web framework that supports the development of web-based applications and APIs to enable user interaction and data accessibility.
4. HTML: Markup language used to structure and display content on web applications, facilitating a user-friendly interface.
5. CSS: Stylesheet language for designing visually appealing and responsive web pages, enhancing the user experience.

3. DEVELOPMENT PROCESS

Our team has chosen to adapt the agile method and the cycle of machine learning project into our project to ensure flexibility, iterative progress, and continuous improvement. By combining these approaches, we can remain responsive to changing requirements and integrate feedback from stakeholders at every stage.

The agile methodology will allow us to break down the project into manageable tasks, enabling faster delivery of results and easier adjustment of priorities.

Additionally, the machine learning project cycle will guide us through crucial phases such as data collection, preprocessing, model training, and evaluation. This structured approach will help ensure that our models are optimized for performance and can evolve as we gather more data and insights. By iterating through these cycles, we aim to continuously enhance the system’s accuracy and effectiveness, ultimately delivering a more robust solution.

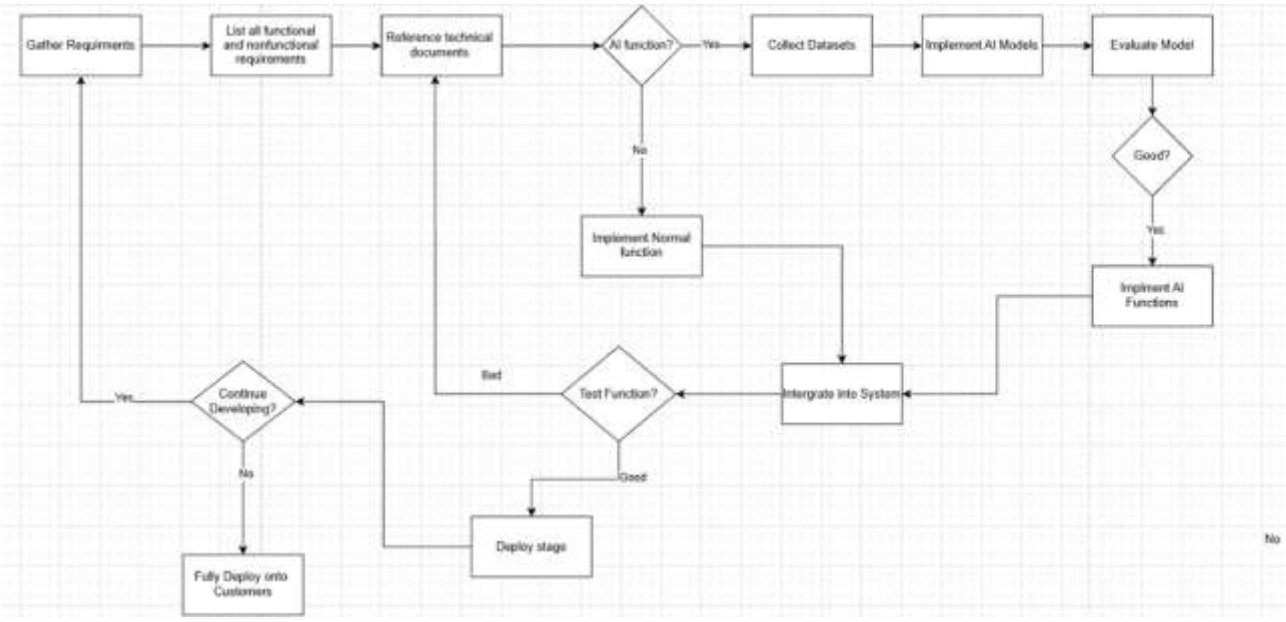


Figure 1.1 System Development Workflow

Project Development Process

Gather Requirements: Identify functional and non-functional requirements for the system.

1. Reference Technical Documents: Consult relevant technical documents to ensure alignment with standards and practices.
2. AI Function? Determine if the system requires AI functionality.
3. If Yes (AI Function):
    - Collect Datasets: Gather data for AI training and testing.
    - Implement AI Models: Develop and train AI models.
    - Evaluate Model: Assess the AI model’s performance.
4. If No (No AI Function): Develop the system's core functionality without AI components.
5. Test Function: Thoroughly test the system's functionality.
6. Integrate into System: Integrate tested components (AI or non-AI) into the overall system.
7. Deploy Stage: Deploy the system in a controlled environment for validation.
8. Fully Deploy to Customers: If satisfactory, fully deploy the system to customers.

Additional Notes:

- The process is iterative, with a decision point to continue development.
- If testing fails, the system may need rework.
- If the AI model is deemed suitable, it will be integrated.

4. CUSOMER REQUIREMENTS

Functional:

- Tumor Detection:
  The system should accurately detect tumors in medical images and highlight them for review.
- Disease Classification:
  Enable classification of medical images into predefined categories of diseases or conditions.
- Segmentation of Medical Images:
  Provide precise segmentation of anatomical structures and pathological regions for detailed analysis.
- Multi-Modality Support:
  Process images from various medical imaging modalities, including MRI, CT, X-ray, and ultrasound.
- User-Friendly Interface:
  Design an intuitive interface for clinicians to upload images, view results, and interact with segmentation outputs.
- Reporting and Export:
  Generate detailed reports summarizing detection, classification, and segmentation outcomes, with export options in standard formats (e.g., PDF, DICOM).

Non-Functional

- 2.1 Accuracy and Reliability:
  Achieve high precision and recall rates, minimizing false positives and negatives in tumor detection and disease classification.
- 2.2 Performance and Scalability:
  Ensure fast processing of large datasets without compromising accuracy and support scaling to accommodate growing data volumes.
- 2.3 Security and Privacy:
  Comply with medical data security standards such as HIPAA and GDPR, ensuring patient data confidentiality and secure data handling.
- 2.4 Interoperability:
  Integrate seamlessly with existing clinical systems and Picture Archiving and Communication Systems (PACS).
- 2.5 Maintainability:
  Design the system for easy updates, model retraining, and long-term maintenance.
- 2.6 Robustness and Fault Tolerance:
  Ensure consistent performance even in the presence of noisy, incomplete, or low-quality images.
- 2.7 Accessibility:
  Support multiple platforms, including web-based and desktop applications, with cross-platform compatibility.

II. Model Analysis and Development

1. Overview

This project employs a multi-model architecture tailored to different medical image processing tasks. Specifically:

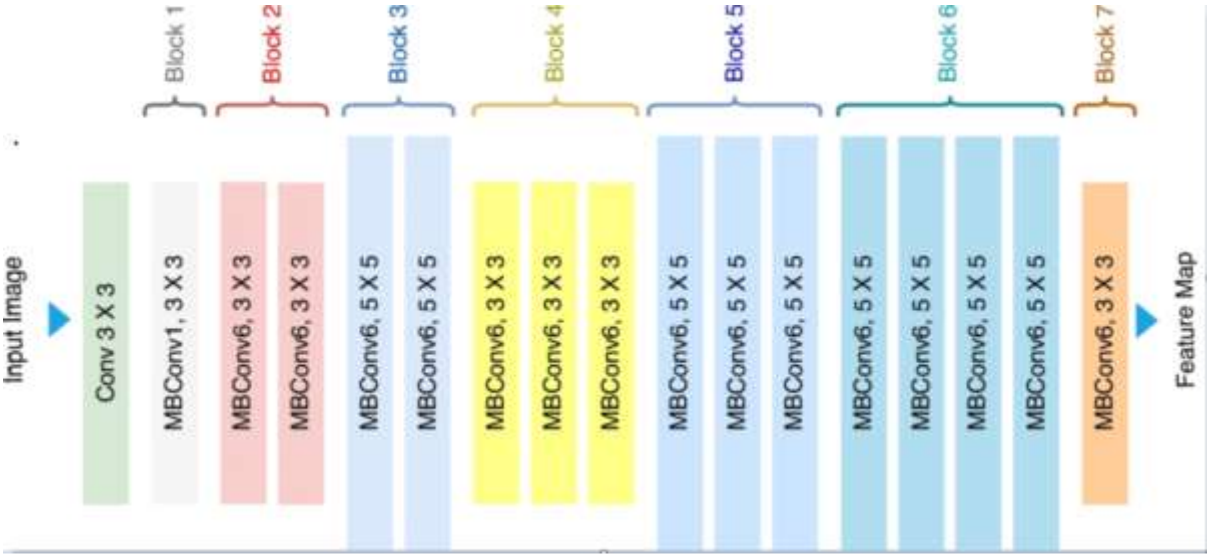
- EfficientNet-B0 is used for disease classification.
- Pyramid Vision Transformer (PVT) serves as the backbone for object detection (e.g., tumor identification).
- U-Net is implemented for pixel-wise segmentation of medical images.

2. Disease Classification with EfficientNet-B0

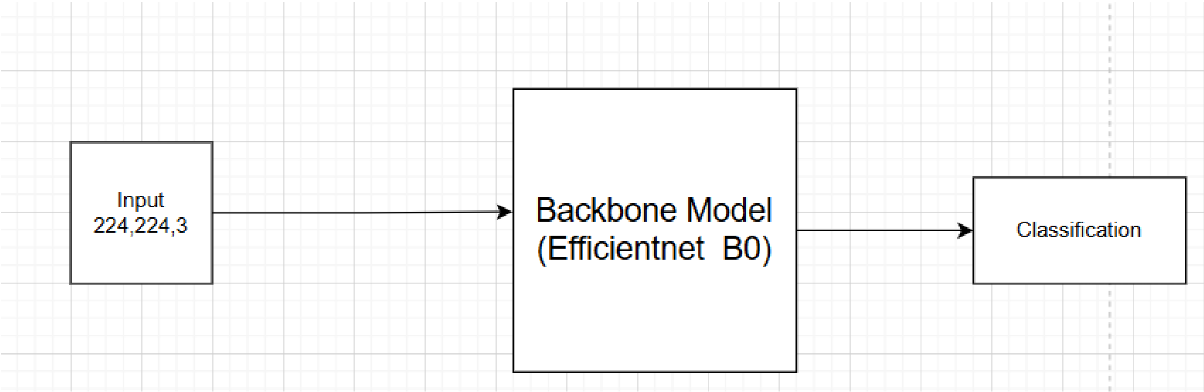
EfficientNet-B0, a highly efficient convolutional neural network (CNN), was chosen for its ability to balance accuracy and computational cost. It is pre-trained on ImageNet and fine-tuned on medical datasets to classify diseases from imaging modalities such as X-rays, CT, or MRI.

Model Architecture:

Backbone Architecture



Fully Architecture



- Input Layer: Accepts resized images (224x224).
- Feature Extraction: Utilizes EfficientNet-B0's pre-trained convolutional layers.
- Classification Head: Fully connected layer with Softmax (or Sigmoid) activation tailored to the number of disease classes.

Training Details:

- Loss Function: Cross-entropy loss for multi-class classification.
- Optimizer: Adam optimizer with learning rate scheduling.
- Metrics: Accuracy, precision, recall, and F1-score were used to evaluate performance.

Results:

Brain Tumor Classification

Dataset Overview

- Total Images: 253
  - Positive Cases: 155 images
  - Negative Cases: 98 images
- Original Image Dimensions: (630 x 630 pixels)
- Resized Image Dimensions: (224 x 224 pixels)

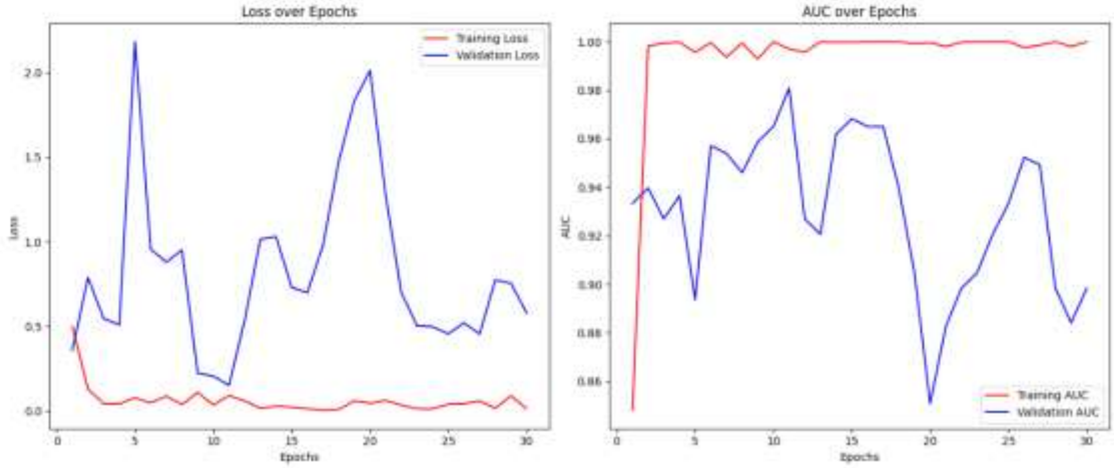
Data Preprocessing

- Normalization Parameters:
  - Mean: (0.485, 0.456, 0.406)
  - Standard Deviation (STD): (0.229, 0.224, 0.225)
- Train/Test Split: 80% training, 20% testing

Model Training Configuration

- Number of Epochs: 30
- Learning Rate: 0.01
- Number of classes:2

Diagram Results



AUC Explanation:

The **Area Under the Curve (AUC)**, often specifically referring to the **AUC-ROC** (Area Under the Receiver Operating Characteristic Curve), is a performance metric that evaluates the ability of a classification model to distinguish between classes. It measures the area under the curve of a plot that maps the True Positive Rate (TPR) against the False Positive Rate (FPR) at various classification thresholds.

- **AUC Score Range**:
  - An **AUC of 0.5** means the model performs no better than random guessing.
  - An **AUC of 1.0** means the model perfectly distinguishes between the classes.
  - An **AUC closer to 1.0** generally indicates better model performance.

AUC on Test Set: **0.95**
This indicates that the model has **excellent** discriminative power, achieving an AUC of 0.95. This score means that there is a **95% probability** that the model will correctly rank a randomly chosen positive instance higher than a randomly chosen negative instance.

Cataract Classification

Dataset Overview

- Total Images: 602
    - Normal cases: 300 images
    - Cataract cases: 100 images
    - Glaucoma cases: 100 images
    - Retina cases: 101 images
- Original Image Dimensions: (300 x 200 pixels)
- Resized Image Dimensions: (224 x 224 pixels)

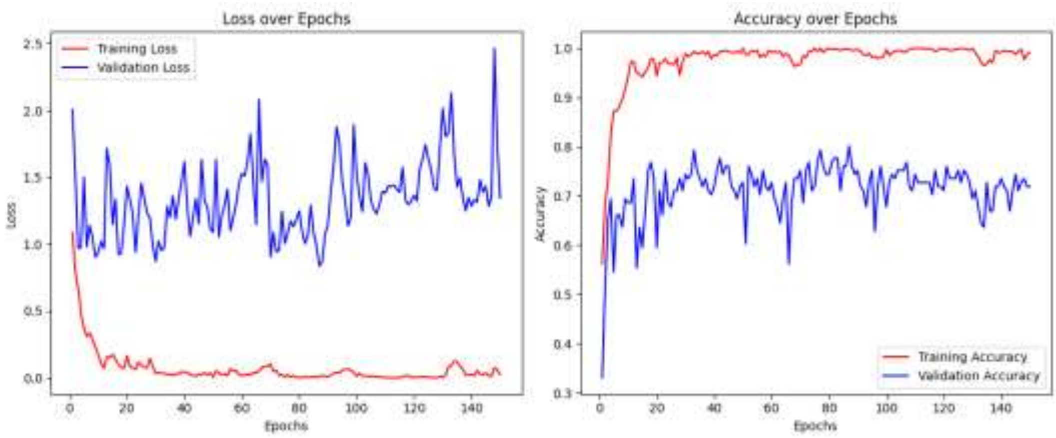
Data Preprocessing

- Normalization Parameters:
    - Mean: (0.485, 0.456, 0.406)
    - Standard Deviation (STD): (0.229, 0.224, 0.225)
- Train/Test Split: 80% training, 20% testing

Model Training Configuration

- Number of Epochs: 100
- Learning Rate: 0.01
- Number of classes:4

Diagram result



Accuracy Explanation:

Accuracy is a simple but widely-used metric in classification tasks. It is defined as the proportion of correct predictions (both true positives and true negatives) out of the total number of predictions made.

Mathematically:

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

Accuracy on Test Set: **0.80**
This indicates that **80%** of the model's predictions on the test set were correct. In other words, the model successfully identified the correct label for 80% of the test samples.

Oral Classification

Dataset Overview

- Total Images: 602
- Calculus cases: 1296 images
- Gingvitis cases: 2340 images
- Hypondita cases: 342 images
- Original Image Dimensions: (300 x 200 pixels)
- Resized Image Dimensions: (224 x 224 pixels)

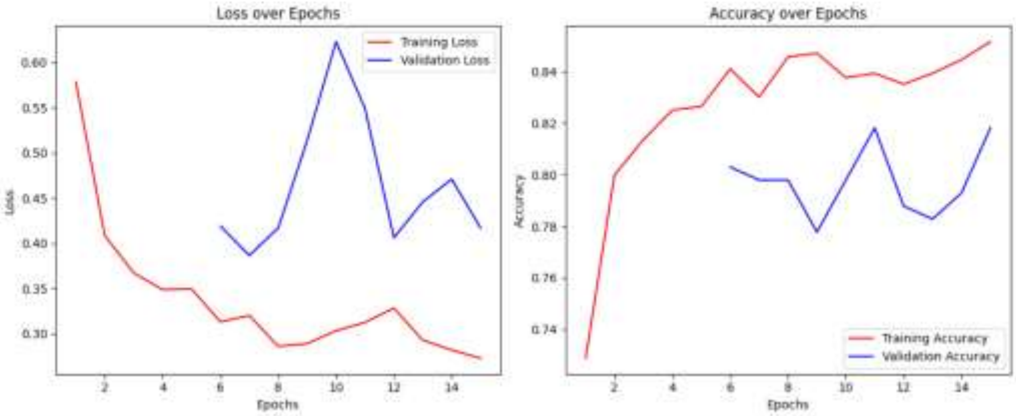
Data Preprocessing

- Normalization Parameters:
- Mean: (0.485, 0.456, 0.406)
- Standard Deviation (STD): (0.229, 0.224, 0.225)
- Train/Test Split: 80% training, 20% testing

Model Training Configuration

- Number of Epochs: 15
- Learning Rate: 0.01
- Number of classes:3

Diagram



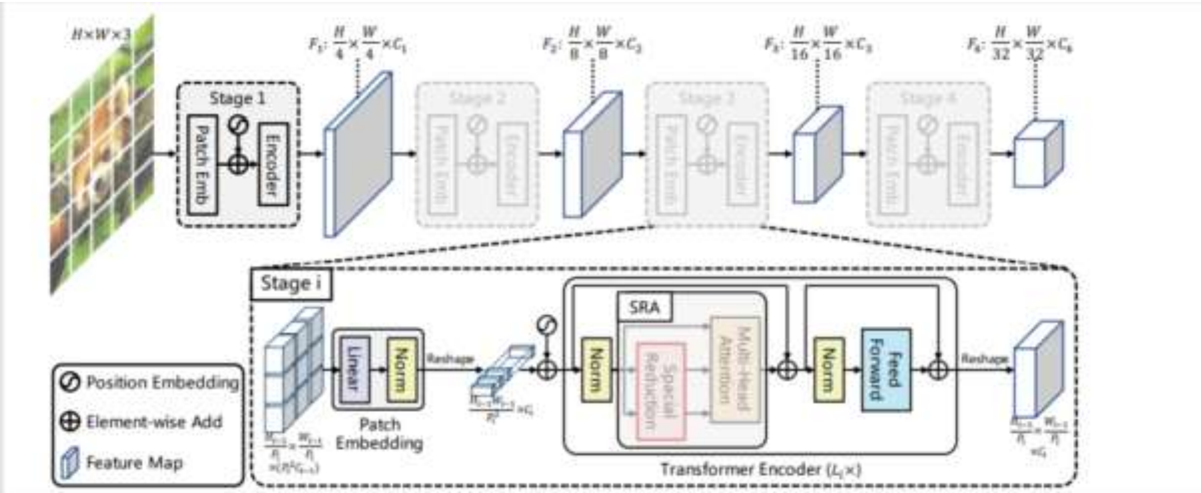
Accuracy on Test Set: **0.8550**
This indicates that **85.5%** of the model's predictions on the test set were correct. This is a strong result, suggesting that the model is performing well in identifying the correct labels for the majority of test samples.

3. Object Detection with PVT (Pyramid Vision Transformer)

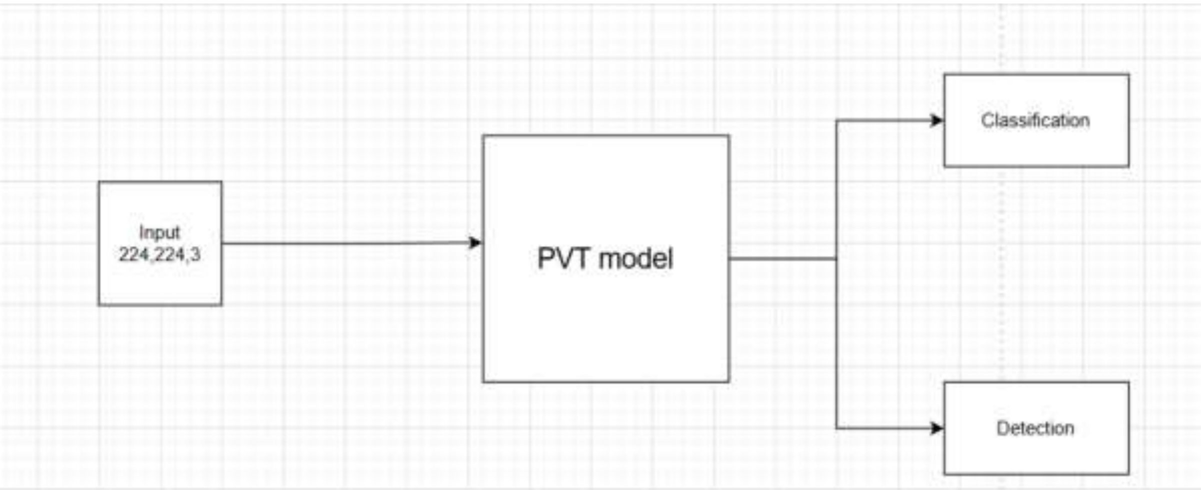
The Pyramid Vision Transformer was selected as the backbone for object detection tasks, such as tumor localization. PVT is well-suited for this application due to its ability to capture global context using self-attention mechanisms while maintaining computational efficiency.

Model Architecture:

PVT Architecture:



Fully Architecture



- Backbone: PVT layers for hierarchical feature extraction.
- Detection Head: Two sequential linear layers for detection and classification.
- Outputs: Bounding box coordinates and class predictions.

Training Details:

- Loss Function: A combination of classification loss (Cross entropy) and bounding box regression loss (Smooth L1 loss).

- Data Augmentation: Techniques such as random crops, scaling, and rotations were applied to increase dataset diversity.
- Optimizer: SGD with momentum for stable convergence.

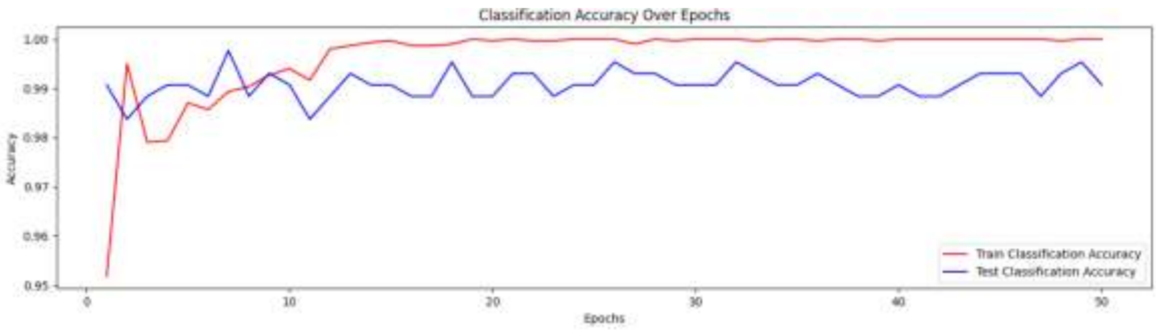
Evaluation Metrics:

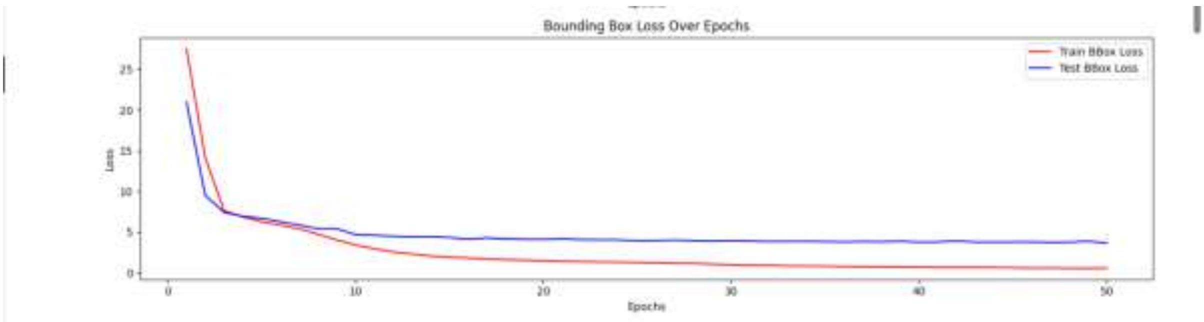
- Accuracy metric for classification bounding box.
- IOU metric for bounding box detection.

Dataset Overview

- Total Images: 1719
    - Train: 1503 images
    - Test: 216 images
    - Total of classes:3
    - Number of object in each images:2
- Original Image Dimensions: (640 x 640 pixels)
- Resized Image Dimensions: (224 x 224 pixels)

Results:





IoU

The **Intersection over Union (IoU)** is a metric commonly used in image segmentation tasks, especially when you want to evaluate how well the predicted segmentation matches the ground truth. It is defined as the ratio of the intersection of the predicted mask and the ground truth mask to the union of both masks.

Mathematically:

$$\text{IoU}=\frac{A \cap B}{A \cup B}$$

Where:

- A is the predicted mask (or region of interest),
- B is the ground truth mask (or true region of interest),
- $|A \cap B$ is the intersection of the predicted and ground truth masks (the overlapping area),
- $|A \cup B|$ is the union of the predicted and ground truth masks (the combined area).

Model Evaluation: Intersection over Union (IoU) Metric

For evaluating the performance of my proposal model the **Intersection over Union (IoU)** was used.

IoU on the test set: **0.67**
This indicates that the model correctly identified 67% of the area of interest (Area of Tumor in brain) in the test images, relative to the union of both the predicted and ground truth areas.

- **Interpretation**:
  An IoU of 0.67 suggests a moderate level of agreement between the model's predictions and the actual ground truth, indicating the model performs reasonably well, but there is still room for improvement. A higher IoU score (closer to 1.0) would indicate better model performance, suggesting that the predictions closely match the ground truth.

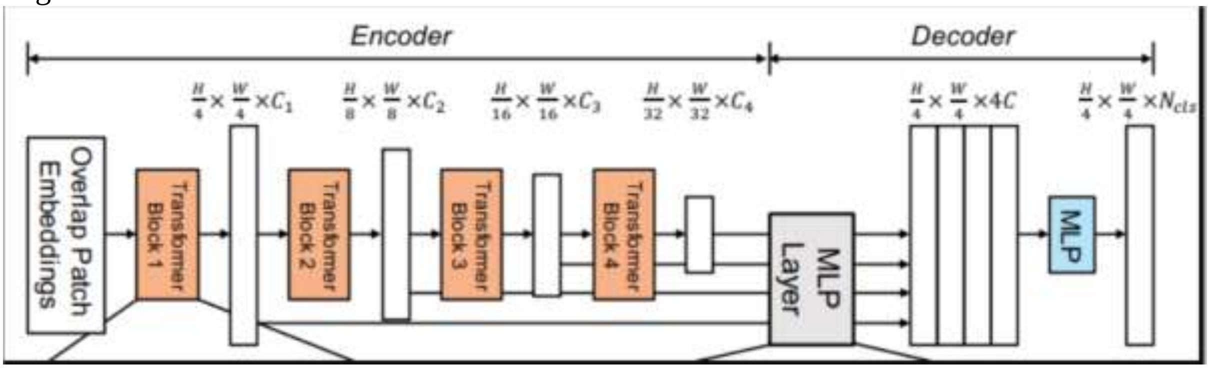
4. Segmentation with SegFormer-B0 (Modified with Custom Architecture)

Overview

SegFormer-B0, a cutting-edge transformer-based architecture, has been utilized for pixel-wise segmentation of anatomical structures and pathological regions. Pretrained on the ADE-20K dataset, the SegFormer-B0 model is fine-tuned to leverage multi-scale feature extraction and transformer-based attention mechanisms. However, the custom model architecture presented here introduces crucial modifications to further enhance performance for specific segmentation tasks, enabling more precise localization and improving output quality.

Fully Architecture

Segformer architecture



Enhanced Model Architecture
In this modified version of SegFormer-B0, several key adjustments were made to tailor the model for optimal performance in binary segmentation tasks, building on the power of the pretrained nvidia/segformer-b0-finetuned-ade-512-512 model. Key features include:

1. **SegFormer Backbone**:
   The backbone of the model remains the highly effective SegFormer-B0, pretrained by NVIDIA and fine-tuned on the ADE-20K dataset. The SegFormer architecture excels at extracting hierarchical features through a transformer-based encoder, providing superior semantic segmentation capabilities. The fine-tuned model from NVIDIA serves as an excellent foundation, bringing in advanced, pretrained knowledge for enhanced accuracy and context-awareness.
2. **Custom Upsampling Mechanism**:
   To improve resolution and feature preservation, the custom model introduces an **upsampling block** consisting of two ConvTranspose2d layers. These layers incrementally increase the resolution of the feature maps, ensuring that fine-grained details from the SegFormer encoder are preserved. This carefully designed upsampling step enables the model to refine segmentation outputs, making it more suitable for precise segmentation tasks like medical image analysis.
   - **First Layer**: ConvTranspose2d(150, 128, kernel_size=4, stride=2, padding=1)
   - **Second Layer**: ConvTranspose2d(128, 64, kernel_size=4, stride=2, padding=1)
     The introduction of the **ReLU activation** after each upsampling layer ensures that the model captures non-linear features, adding depth to the segmentation process.
3. **Refined Header for Final Predictions**:
   After the upsampling block, the model incorporates a **header** with a sequence of fully connected layers designed to generate the final segmentation output. This

modification serves to process the upsampled features and produce high-quality segmentation predictions:

- **Layer 1**: A linear layer that reduces the feature dimensions from 64 to 256, allowing the model to learn more abstract features.
- **Layer 2**: The final linear layer produces a single-channel output for binary segmentation, ensuring that the model is optimized for tasks such as anatomical structure delineation and pathological region identification.

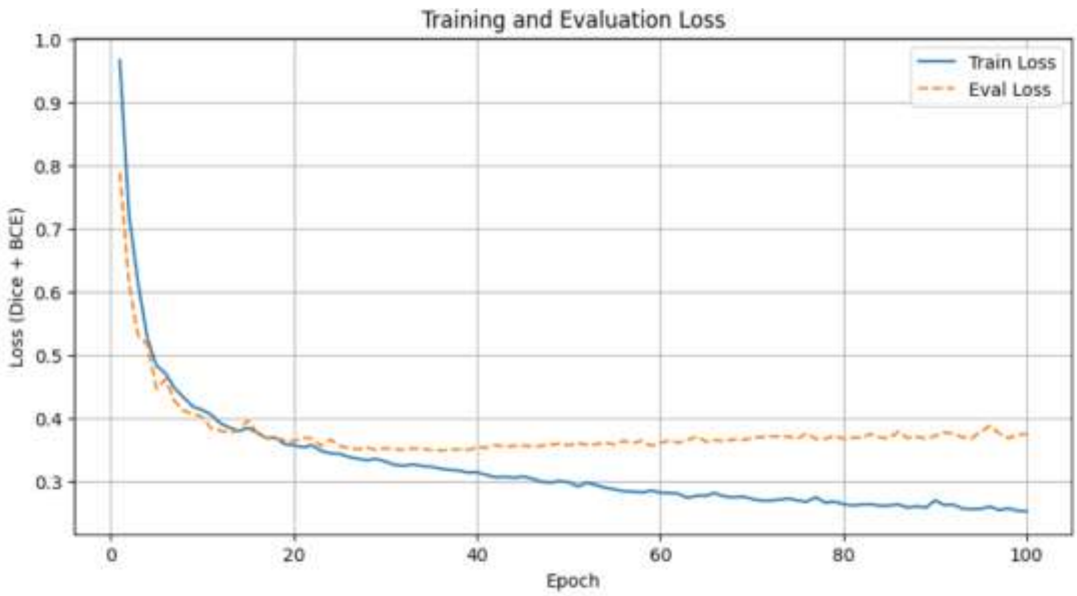
Training Details

For training, a combination of Dice Loss and Binary Cross-Entropy (BCE) Loss is used. This hybrid loss function balances pixel-wise accuracy (BCE) with region overlap (Dice), which helps in addressing class imbalance and ensuring robust segmentation. Key training parameters are:

- **Loss Function**: Dice Loss + BCE Loss
- **Optimizer**: Adam optimizer with a learning rate of 0.001
- **Data Augmentation**: Common techniques such as rotation, flipping, and intensity normalization were used for enhancing generalization.
- **Batch Size and Epochs**: 4 and 100

Evaluation Metrics

The model was evaluated using the Dice Coefficient, a metric particularly suited for binary masks as it measures the overlap between the predicted and ground truth masks. This ensures that both the segmentation accuracy and shape consistency of the predicted masks are taken into account.



Dice Loss and Dice Score Explanation:

The **Dice coefficient** (or **Dice Similarity Index**, DS or DSC) is a metric commonly used to evaluate the similarity between two samples, typically for segmentation tasks. It ranges from 0 (no similarity) to 1 (perfect similarity).

The **Dice Loss** is simply the complement of the Dice coefficient, meaning the Dice loss is often used as the objective function to minimize in many segmentation tasks. The relation between **Dice Score** and **Dice Loss** is as follows:

- **Dice Score**:

$$\text{Dice Score}=\frac{2\times A\cap B}{|A|+|B|}$$

  Where A is the predicted segmentation and B is the ground truth segmentation.

- **Dice Loss**:

$$1-Dice\ Score$$

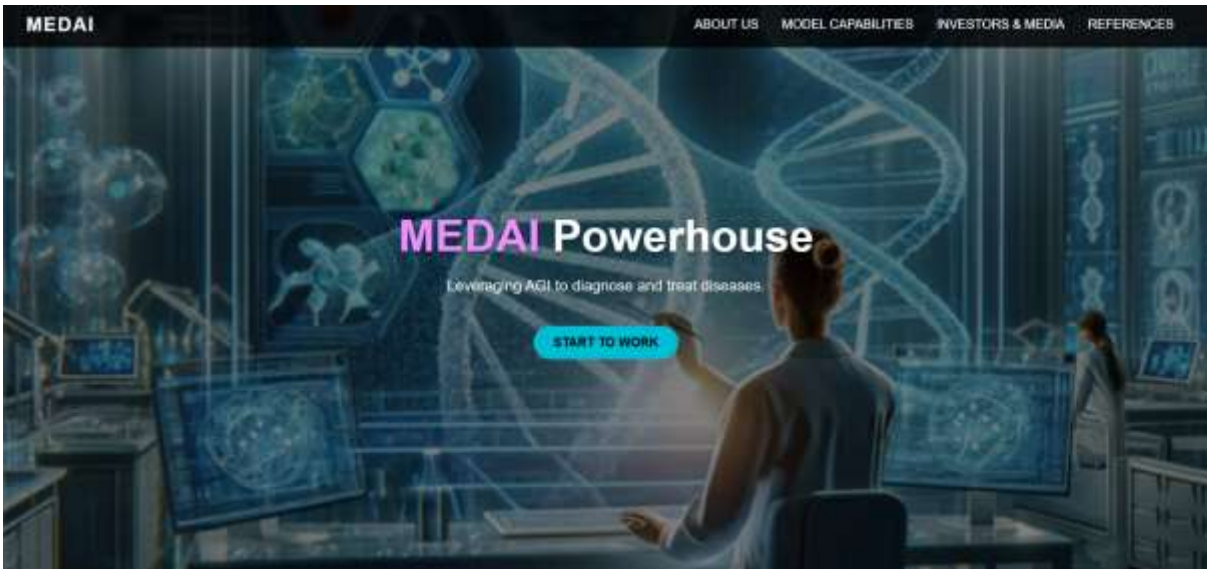
The **Dice Loss** was employed to evaluate the segmentation accuracy of the model on the oral disease classification task. The Dice Loss measures the dissimilarity between the predicted segmentation and the ground truth segmentation. It is particularly useful in segmentation tasks as it balances both precision and recall.

- **Dice Loss on Test Set**: **0.3483**
  This result indicates that the model has a **Dice Score** of approximately **0.6517**, meaning the predicted segmentation overlaps 65.17% with the ground truth segmentation, on average.
- **Interpretation**:
  A Dice Loss of **0.3483** suggests that the model's predictions are fairly close to the ground truth, but there is still significant room for improvement. A lower Dice Loss would indicate better performance, with a Dice Loss closer to 0 representing near-perfect segmentation accuracy.

III Implementation

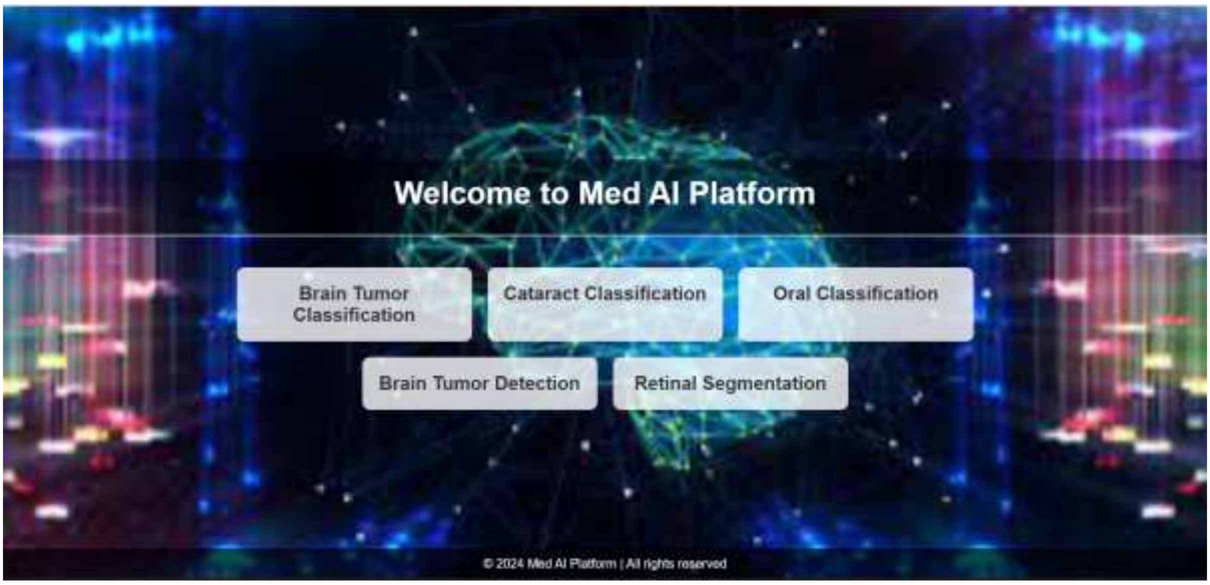
1 User at web page

- **Action**: User accesses the web page.



2 Go to MainPage

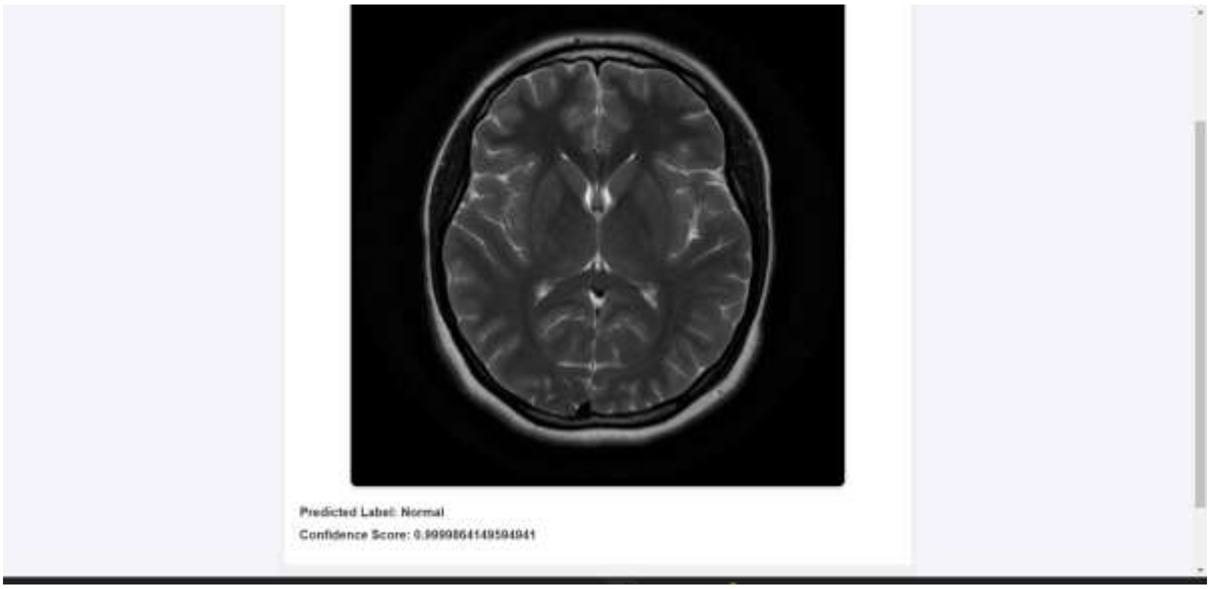
- **Action**: User clicks the "Start to Work" button to navigate to the main page.



3 Using Classification features

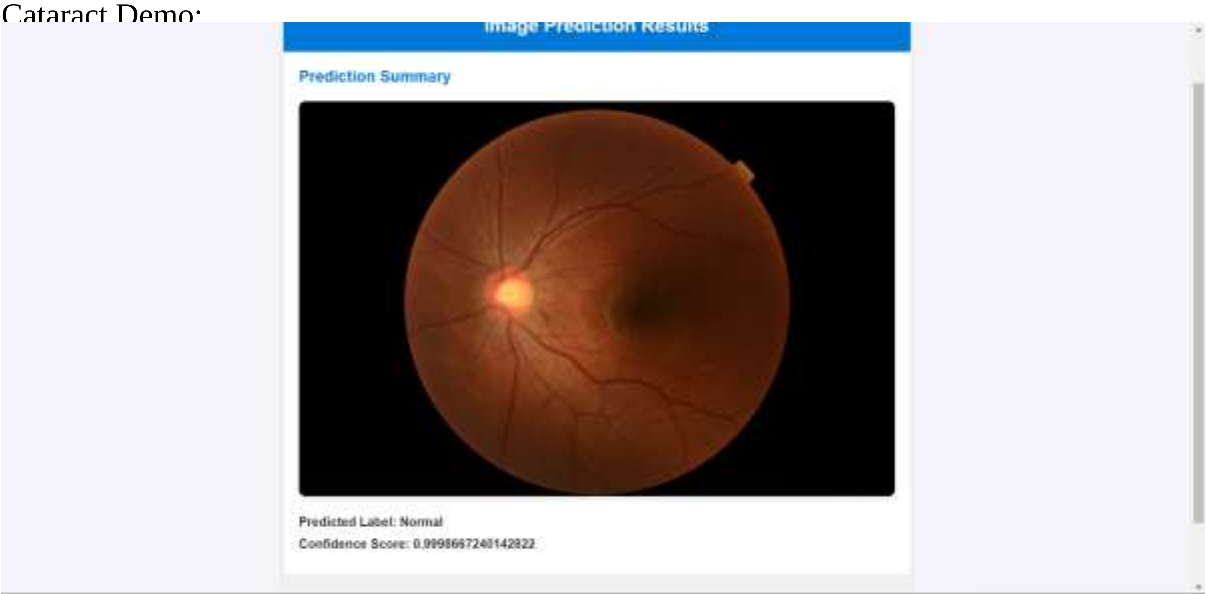
- **Step 1**: User clicks the "Brain Tumor Classification" button.
- **Step 2**: User uploads an image for classification.

- **Step 3**: User clicks the "Upload and Describe" button.
- **Step 4**: System processes the image and displays the result.



Similary, for two other classification features :Oral classification and Cataract classification

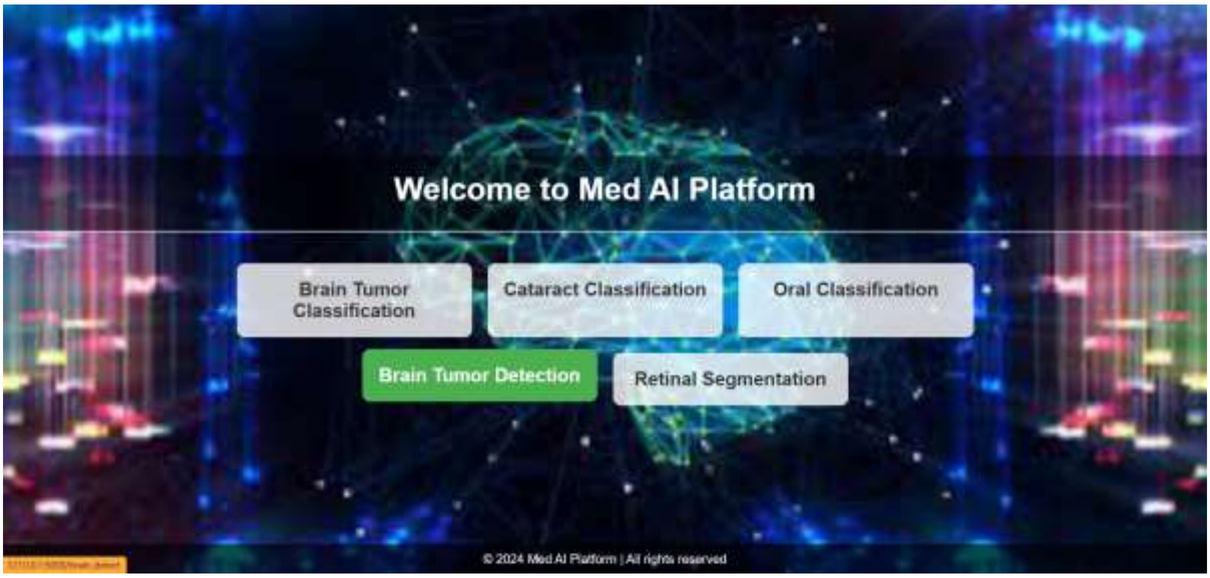
Cataract Demo:



Oral Demo:

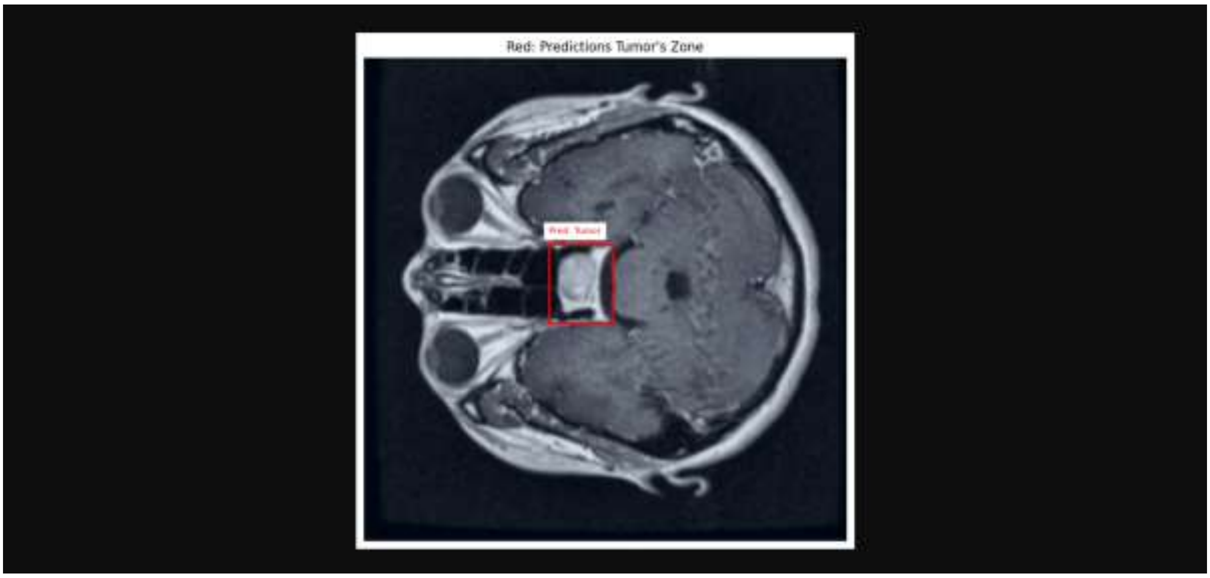
4 Using Brain Tumor Detection Features

Step1 Users click Brain Tumor Detection button



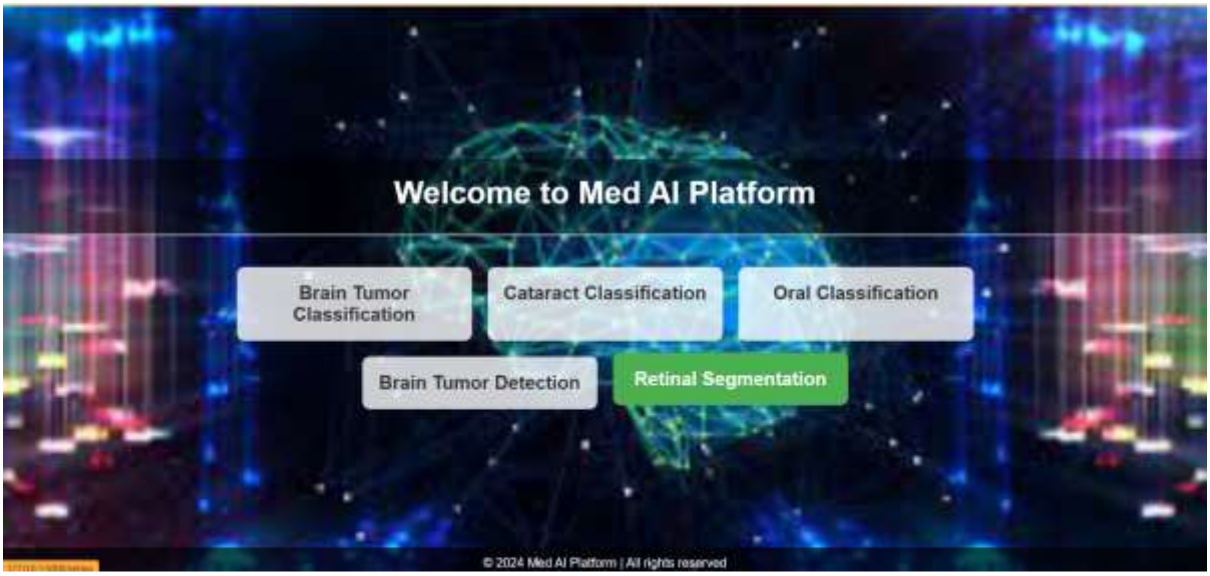
Step2 Upload an image

Step3 Obser the result



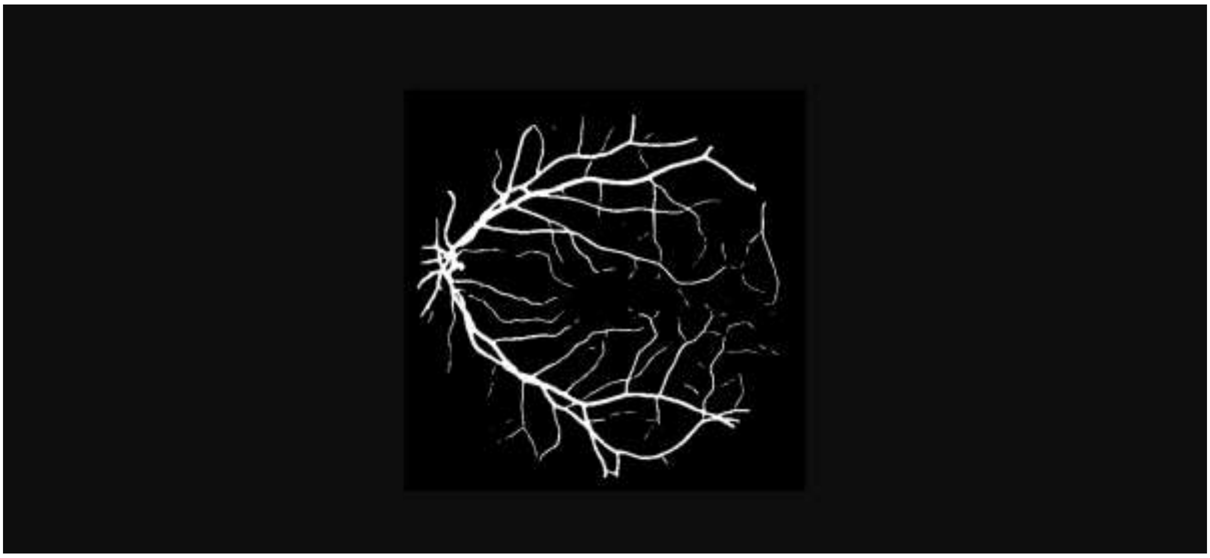
5 Using Segmantation features

Step1 User click Retinal Segmantation button



Step2 Upload an image

Step3 Observe the result



IV. Discussion

1 Implement AI Website Application

The AI website application was created to provide users with an easy way to access and use AI models for medical tasks, such as brain tumor classification, oral disease classification, cataract detection, and retinal segmentation. The application is designed to be user-friendly, so even people without technical knowledge can upload images and receive results quickly. This project helps show how AI can be used in medicine to assist with diagnosis and decision-making. The website can be easily updated as new AI models or datasets become available, making it a flexible tool for future improvements.

2 Improve Existing AI model to work with Medical dataset

The AI models used in this project can be improved to work better with medical datasets, which often have specific challenges, such as different image qualities or rare conditions. Some ways to improve the models include:

1. **Data Augmentation**: This technique involves changing the images slightly (e.g., rotating or zooming in) to help the model learn to recognize different variations of the data.
2. **Transfer Learning**: Using pre-trained models from large datasets and fine-tuning them with medical images can improve the model's performance, even with limited medical data.
3. **Model Fine-Tuning**: Adjusting the model's settings and structure to make it more accurate for medical image analysis.

3 Achievements of the AI Medical Website Application

The development of the AI-powered medical website application has led to several key achievements:

1. **User-Friendly Interface**:
   The application features a clean, easy-to-navigate interface, designed to cater to both medical professionals and non-expert users. This allows healthcare providers to upload patient images directly for analysis, receiving results in real time. The simplicity of the design ensures quick adoption by users without technical backgrounds.
2. **Versatility Across Medical Specialties**:
   The website supports multiple medical domains, including **brain tumor classification**, **oral disease classification**, **cataract detection**, and **retinal segmentation**. This broad scope demonstrates the flexibility and potential of AI in assisting various diagnostic tasks, potentially improving the speed and accuracy of diagnoses across multiple specialties.
3. **Real-Time Predictions**:
   The application provides instant results after users upload medical images, offering a quick and reliable AI-powered second opinion. This capability is particularly valuable

in medical environments where timely decision-making is crucial, such as emergency care or remote consultations.

4. **Integration of State-of-the-Art Models**:
   The website integrates cutting-edge AI models that have been trained on large, diverse datasets. These models, designed to handle medical images, have been fine-tuned to ensure high accuracy and performance in real-world scenarios.
5. **Scalability and Flexibility**:
   The application is designed to be easily updated as new AI models or datasets become available. This scalability ensures that the website can evolve with advancements in AI and medicine, making it a long-term tool for healthcare professionals.
6. **Improved Accessibility**:
   With an easy-to-use web interface, medical professionals from all over the world, including those in remote locations, can access the application without requiring any specialized hardware or software. This promotes equitable access to advanced diagnostic tools.

4 Limitations of the AI Medical Website Application

Despite the significant achievements, there are several limitations and challenges that need to be addressed:

1. **Dataset Quality and Size**:
   Medical datasets are often smaller and less diverse compared to those used in other AI applications. This can result in overfitting, where models perform well on training data but fail to generalize to new, unseen cases. Continuous efforts to collect diverse, high-quality datasets are needed to improve the model's robustness.
2. **Handling Rare Conditions**:
   Medical datasets frequently include rare or atypical conditions that the AI models may struggle to recognize. To address this limitation, **transfer learning** and **data augmentation** techniques can be employed, but further research and development are needed to improve model performance in these cases.
3. **Image Quality Variability**:
   Medical images can vary significantly in quality due to factors such as equipment, patient conditions, or imaging techniques. The AI models may struggle with low-resolution or poorly captured images, leading to less accurate predictions. Implementing preprocessing steps like image enhancement could help, but some quality issues may remain challenging.
4. **Ethical and Regulatory Concerns**:
   The use of AI in healthcare raises important ethical and regulatory concerns, such as the need for accurate labeling, ensuring patient data privacy, and securing regulatory approval from medical authorities. The application must adhere to **HIPAA** or **GDPR** standards for patient confidentiality and safety, and ongoing compliance with local healthcare regulations will be necessary.
5. **Model Interpretability**:
   Many AI models, particularly deep learning models, are often considered "black boxes," meaning their decision-making processes are not easily interpretable. In medical applications, where explanations of AI predictions are important for clinician trust and decision-making, efforts to improve model transparency and explainability are critical.

6. **Dependence on High-Quality Data**:
   The performance of AI models is heavily dependent on the quality and quantity of the data used to train them. If the models are trained on biased or incomplete data, the resulting predictions may be inaccurate or biased, especially for underrepresented demographic groups. This issue underscores the need for diverse and representative datasets.
7. **Computational Resources**:
   Running AI models for medical image analysis, particularly deep learning models, can be computationally expensive. To ensure fast processing for real-time results, sufficient infrastructure and computational power are required. This may limit accessibility for healthcare providers with limited resources.
8. **Interpretation and Clinical Validation**:
   While the AI models provide predictions based on image analysis, they are not a substitute for professional medical judgment. The clinical validation of these predictions in real-world healthcare settings remains a challenge. AI should be used as a supplementary tool to support decision-making, not as a sole decision-maker.

5 Future Directions

To address these limitations and further enhance the AI website application, the following future improvements and strategies are recommended:

1. **Expand Dataset and Model Diversity**:
   Collecting more medical images, including rare conditions, and improving dataset diversity will be crucial for improving the robustness of AI models. Collaborations with hospitals, medical institutions, and research organizations can help broaden the dataset scope.
2. **Improve Image Quality Handling**:
   Techniques such as **image preprocessing**, **super-resolution algorithms**, and **denoising methods** can be explored to improve the model's ability to handle low-quality medical images.
3. **Enhance Model Explainability**:
   Developing methods for AI model interpretability will allow healthcare providers to better understand and trust AI predictions. This can be achieved through tools that explain model decisions, such as **Grad-CAM** (Gradient-weighted Class Activation Mapping).
4. **Integrate Multi-modal Data**:
   Combining medical image analysis with patient demographic and clinical data (such as lab results and medical history) could further enhance model predictions and lead to more comprehensive diagnostic support.
5. **Collaboration with Medical Professionals**:
   Ongoing collaboration with clinicians will be essential for improving the models, validating their predictions, and ensuring that the system aligns with real-world clinical needs and workflows.
6. **AI for Personalized Medicine**:
   With continuous advancements, AI models could be extended to provide personalized treatment recommendations based on medical image analysis, patient history, and genetic data.

V. Conclusion

1 Learning Outcomes

Through the development and implementation of this **Medical AI Web Application**, we have gained valuable experience in the following areas:

- **Building a Medical Web Application**:
  We successfully built a functional website using the **Flask framework** and **SQLAlchemy** for database management. This allowed us to create a platform where healthcare professionals can easily access AI-powered diagnostic tools.
- **AI Integration for Medical Diagnosis**:
  We integrated **AI models** for tasks like **brain tumor classification**, **oral disease classification**, **cataract detection**, and **retinal segmentation**. This enables the application to provide intelligent results, helping healthcare providers make faster and more accurate diagnoses.
- **Deployment for Real-World Use**:
  The application was deployed on **Ngrok** for demonstrations, showing how it performs in real-world conditions. This allowed us to gather feedback from users and improve the application.

2 Limitations

Although we have achieved a lot, there are areas that still need improvement:

1. **Feature Expansion**:
   - **User Authentication**: We plan to add better security features, like **OAuth 2.0**, allowing users to log in with accounts such as **Google** or **Facebook**.
   - **Automating Features**: Some tasks are currently done manually. We want to use **AI technologies**, like **Computer Vision** and **NLP**, to automate tasks and improve the user experience.
2. **AI Model Optimization**:
   - **Improving Performance**: We aim to make the AI models faster and more accurate, so they can handle more medical data and provide better predictions.
   - **User Satisfaction**: We want to make sure that AI features are reliable and easy to use, helping healthcare professionals trust the results.

3 Overall Assessment

This project has given us hands-on experience in building a medical website, integrating AI models, and deploying the application. While the current version works well, addressing the limitations—such as improving user authentication, automating tasks, and optimizing the AI models—will make the application even more useful.

Moving forward, we plan to improve the platform to meet the needs of healthcare professionals and patients. By making these improvements, the application will be able to provide faster, more accurate, and accessible AI-driven medical diagnosis.