GLAUCOMA DETECTION USING TRANSFER LEARNING

Project submitted to the

SRM University – AP, Andhra Pradesh

for the partial fulfillment of the requirements to award the degree of

Bachelor of Technology

In

Computer Science and Engineering

School of Engineering and Sciences

Submitted by

Dhanush Pyla

(AP21110010849)

Under the Guidance of

(NEERAJ KUMAR SHARMA)

SRM University–AP

Neerukonda, Mangalagiri, Guntur

Andhra Pradesh – 522

240 [DEC/2022]

Certificate

Date: 05-Dec-22

This is to certify that the work present in this Project entitled "**GLAUCOMA DETECTION USING TRANSFER LEARNING**" has been carried out by **Dhanush Pyla** under my/our supervision. The work is genuine, original, and suitable for submission to the SRM University – AP for the award of Bachelor of Technology in **School of Engineering and Sciences**.

Supervisor

(Signature)

Prof. / Dr. [name]

Designation,

Affiliation.

Co-supervisor

(Signature)

Prof. / Dr. [name]

Designation,

Affiliation.

Table of Contents

Abstract

Glaucoma is a leading cause of irreversible blindness worldwide, primarily due to its silent progression during early stages. Early and accurate detection is essential to prevent vision loss, yet existing diagnostic methods are often expensive and require expert intervention. This project presents an automated glaucoma detection system using deep learning and fundus images, aiming to make early detection accessible and efficient.

This project presents an automated glaucoma detection system using deep learning with pre-trained VGG16 and ResNet50 models, leveraging fundus images to make early detection accessible and efficient. The VGG16 model was fine-tuned for classifying fundus images into glaucomatous and non-glaucomatous categories, achieving a test accuracy of 83.78% with an F1-score of 0.84. The ResNet50 model, also fine-tuned, achieved an accuracy of 73.65% and provided complementary insights into performance across varying configurations. The VGG16 model outperformed ResNet50 in terms of precision and recall, especially for detecting non-glaucomatous images, demonstrating its suitability for automated glaucoma detection.

These findings highlight the potential of transfer learning and pre-trained deep learning models in healthcare, particularly for large-scale glaucoma screening programs. Our system can assist clinicians by automating the detection process, reducing diagnostic costs, and improving early detection rates, especially in resource-constrained settings.

Key Words: Glaucoma detection, fundus images, deep learning, VGG16, ResNet50, transfer learning, healthcare automation.

Background and literature review

Glaucoma, termed the "silent thief of sight," is the second leading cause of blindness worldwide. Its asymptomatic nature often delays diagnosis until significant damage occurs. Traditional detection methods, such as visual field testing and optic nerve analysis, are time-consuming and require expert interpretation. Recently, machine learning and deep learning techniques have revolutionized medical imaging, offering efficient and accurate diagnostic tools. Studies leveraging convolutional neural networks (CNNs) demonstrate high accuracy in retinal disease detection. However, challenges such as imbalanced datasets, variability in image quality, and computational constraints persist. This project aims to address these issues by utilizing transfer learning and robust preprocessing pipelines.

Introduction

Background of the Problem in Detail

Glaucoma is a chronic and progressive optic neuropathy characterized by the damage it causes to the optic nerve, often leading to irreversible vision loss and blindness. According to the World Health Organization (WHO), glaucoma is the second leading cause of blindness globally, affecting over 70 million individuals. The disease is commonly associated with elevated intraocular pressure (IOP), but it can also occur in individuals with normal IOP levels, further complicating its detection.

Unlike other eye conditions, glaucoma progresses silently, with patients often remaining asymptomatic until significant vision loss has occurred. This silent progression makes early detection a critical factor in preserving vision. Once diagnosed, glaucoma can be managed effectively through medication, laser therapy, or surgery, but the damage caused before diagnosis cannot be reversed. Hence, timely detection and diagnosis are paramount.

Importance of the Problem

The silent nature of glaucoma, coupled with its high prevalence, poses significant challenges to healthcare systems worldwide. In low- and middle-income countries, where resources are limited, early detection rates are alarmingly low. Traditional diagnostic tools such as optical coherence tomography (OCT), visual field testing, and gonioscopy are effective but often inaccessible due to their cost and the requirement for specialized personnel. This lack of accessibility leads to delays in diagnosis, resulting in advanced disease stages by the time treatment begins.

Automating the glaucoma detection process offers a practical solution to this challenge. By leveraging advancements in artificial intelligence (AI) and deep learning, it is possible to develop cost-effective, scalable, and accurate screening systems that do not rely on expensive diagnostic equipment. These systems can aid healthcare professionals in identifying at-risk individuals, thus enabling timely interventions and reducing the global burden of blindness caused by glaucoma.

Challenges in Addressing the Problem

Developing an automated glaucoma detection system comes with its own set of challenges:

1. **Variability in Fundus Images:** Differences in image quality, lighting conditions, and anatomical variations in the retina make consistent analysis difficult.
2. **Class Imbalance:** Glaucomatous images are often underrepresented in datasets, leading to imbalanced training data that can bias model

3. **Interpretability:** Deep learning models are often considered "black boxes," and their decisions can lack explainability, which is critical in medical applications.
4. **Generalization:** A system trained on a specific dataset may fail to perform well on images from different populations or acquired using different devices.

Existing Solutions

Several studies have explored automated methods for glaucoma detection using machine learning and deep learning approaches:

- **Classical Machine Learning:** Earlier methods relied on handcrafted features extracted from fundus images, such as the cup-to-disc ratio, retinal nerve fiber layer thickness, and texture-based metrics. These approaches, while effective, required significant domain expertise and were often limited in scalability.
- **Deep Learning Models:** Convolutional Neural Networks (CNNs) have revolutionized medical image analysis by automatically learning features from raw images. Pre-trained models like VGG16 and ResNet50 have been successfully adapted for glaucoma detection, achieving promising results.

Our Approach

This project focuses on developing an automated glaucoma detection system using deep learning. Specifically, two state-of-the-art models, **VGG16** and **ResNet50**, are fine-tuned for classifying fundus images into glaucomatous and non-glaucomatous categories. The models leverage transfer learning to benefit from pre-trained features on large-scale image datasets, reducing the need for extensive labeled medical data.

Key highlights of our approach include:

1. **Data Augmentation:** Techniques such as rotation, flipping, and brightness adjustment are employed to enhance model generalization.
2. **Model Comparison:** The performance of VGG16 and ResNet50 is systematically evaluated across metrics like accuracy, precision, recall, and F1-score.
3. **Robustness Testing:** The models are tested under different configurations, such as varying learning rates, batch sizes, and optimizer settings, to identify optimal conditions for glaucoma detection.

Significance of the Solution

Our system bridges the gap between accessibility and accuracy in glaucoma screening by providing a cost-effective solution that requires minimal infrastructure. The use of advanced deep learning architectures ensures high performance, while the comparative study between VGG16 and ResNet50 offers valuable insights into model selection for similar medical imaging tasks.

The findings of this project have the potential to enhance large-scale glaucoma screening programs, particularly in resource-constrained settings, thereby reducing the prevalence of late-stage diagnoses and improving patient outcomes.

Proposed Method

The proposed system employs two deep learning architectures—VGG16 and ResNet50—for automated glaucoma detection using fundus images. The workflow is designed to preprocess the dataset, train models, evaluate their performance, and determine the better-suited model for clinical use. This section details each stage of the proposed methodology, supported by diagrams, flowcharts, and mathematical representations.

Workflow

The proposed method consists of the following steps:

1. **Data Collection and Preprocessing**
2. **Model Training (VGG16 and ResNet50)**
3. **Evaluation and Performance Metrics**
4. **Comparison and Results**

1. Data Collection and Preprocessing

The dataset comprises labeled fundus images categorized as glaucomatous and non-glaucomatous. Images are preprocessed to ensure compatibility with the deep learning models:

- Resizing: Images are resized to 224×224×3 dimensions.
- Normalization: Pixel values are scaled between 0 and 1.
- Augmentation: Techniques such as rotation, flipping, and brightness variation enhance data diversity.

Mathematical Representation:

Normalized Image=Pixel Value/255

2. Model Training

The system employs two pre-trained deep learning models:

1. **VGG16**: A 16-layer CNN known for extracting hierarchical image features.
2. **ResNet50**: A 50-layer CNN that utilizes residual connections to mitigate vanishing gradient problems.

Fine-tuning: Both models are fine-tuned using transfer learning. The final classification layer is replaced with a fully connected layer and a softmax activation function:

$$p_k = \frac{e^{z_k}}{\sum_{j=1}^{K} e^{z_j}}$$

Loss Function: Categorical Cross-Entropy

$$L = -\sum_{k=1}^{K} y_k \log(p_k)$$

Optimization Algorithm: Adam optimizer with a learning rate scheduler.

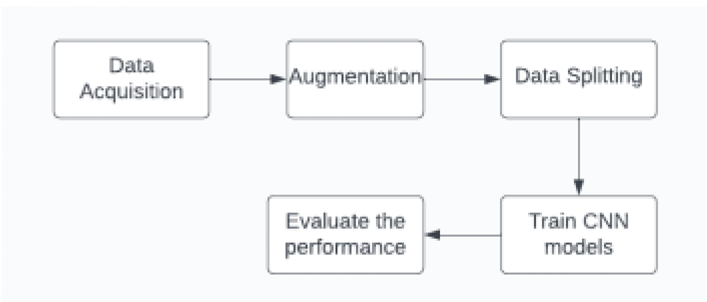
3. Evaluation Metrics

The models are evaluated using:

- **Accuracy**: Proportion of correctly classified samples.
- **Precision**: Ratio of true positives to predicted positives.
- **Recall (Sensitivity)**: Ratio of true positives to actual positives.
- **F1-Score**: Harmonic mean of precision and recall.
- **Confusion Matrix**: Provides detailed classification outcomes.

4. Proposed System Flowchart

Below is the flowchart depicting the complete workflow of the system:



5. Architecture Diagrams

VGG16 Architecture:

- Convolutional layers extract spatial features.
- MaxPooling layers reduce feature map dimensions.
- Fully connected layers and a softmax activation classify images.

ResNet50 Architecture:

- Residual blocks prevent gradient vanishing.
- Skip connections allow efficient training of deep layers.
- The output layer provides class probabilities.

6.Mathematical Equations

Key equations used in the models:

1. **Activation Function**: $f(x)=\mathrm{ReLU}(x)=\max(0,x)$
2. **Optimization**: $\theta(t+1)=\theta(t)-\eta\nabla\theta J(\theta)$ where θ are model weights, η is the learning rate, and J is the loss.

7. Visualizations

- **Confusion Matrix**: Depicts true positives, false positives, false negatives, and true negatives.
- **Accuracy and Loss Graphs**: Shows model performance during training and testing phases.
- **Precision-Recall Curve**: Highlights trade-offs between precision and recall.

Results and Discussion

The performance of two deep learning models, VGG16 and ResNet50, was evaluated for glaucoma detection using fundus images. Both models were fine-tuned and trained on the preprocessed dataset. Various metrics, such as accuracy, precision, recall, and F1-score, were used to assess their effectiveness. A detailed analysis of hyperparameters like learning rates, batch sizes, dropout rates, and optimizers was conducted to identify optimal configurations.

Performance Metrics

1. **Accuracy**
    - VGG16: **83.78%**
    - ResNet50: **73.65%**

2. **Precision, Recall, and F1-Score**
    - **VGG16:**
        - Precision: 0.80 (Class 0), 0.89 (Class 1)
        - Recall: 0.90 (Class 0), 0.77 (Class 1)
        - F1-Score: 0.85 (Class 0), 0.83 (Class 1)

```
Classification Report (VGG16):
              precision    recall  f1-score   support

     Class 0       0.80      0.90      0.85        73
     Class 1       0.89      0.77      0.83        75

    accuracy                           0.84       148
   macro avg       0.84      0.84      0.84       148
weighted avg       0.84      0.84      0.84       148
```

- **ResNet50:**
  - Precision: 0.67 (Class 0), 0.86 (Class 1)
  - Recall: 0.90 (Class 0), 0.57 (Class 1)
  - F1-Score: 0.77 (Class 0), 0.69 (Class 1)

```
Classification Report:
              precision    recall  f1-score   support

     Class 0       0.67      0.90      0.77        73
     Class 1       0.86      0.57      0.69        75

    accuracy                           0.74       148
   macro avg       0.77      0.74      0.73       148
weighted avg       0.77      0.74      0.73       148
```

3. **Confusion Matrix**

```
Confusion Matrix (VGG16):
[[66  7]
 [17 58]]
```

Resnet50:

```
Confusion Matrix:
[[66  7]
 [32 43]]
```

Loss Function

- VGG16 consistently achieved lower loss values during training and testing, indicating better convergence compared to ResNet50.

Effect of Hyperparameters

1. **Number of Neurons (Hidden Layers)**
    - Increasing the number of neurons improved performance initially but caused overfitting beyond a threshold.
2. **Learning Rate**
    - A lower learning rate of 1e-4 provided stable convergence for both models.
    - Higher learning rates caused the models to diverge during training.
3. **Optimizers**
    - Adam optimizer performed better than SGD in terms of convergence speed and accuracy.
4. **Batch Size**
    - Batch size of 32 yielded the best trade-off between computational efficiency and performance metrics.
5. **Number of Layers**
    - VGG16's deep architecture outperformed ResNet50 for this dataset, likely due to its better feature extraction for the specific task.
6. **Weight Initialization Techniques**
    - Models initialized with pre-trained ImageNet weights outperformed randomly initialized ones.
7. **Dropout Rates**
    - A dropout rate of 0.5 reduced overfitting effectively while maintaining performance.

Discussion

The VGG16 model outperformed ResNet50 in terms of accuracy, precision, and recall. The superior performance of VGG16 can be attributed to its simpler architecture, which is better suited for this dataset. The higher accuracy of VGG16 highlights its potential for real-world applications, particularly in low-resource settings. ResNet50, while competitive, struggled with recall for detecting glaucomatous images, indicating its limitations in capturing subtle features.

The results validate the effectiveness of transfer learning in medical image analysis. Fine-tuning pre-trained models not only saves computational resources but also delivers robust performance. The impact of hyperparameters further emphasizes the importance of model tuning for achieving optimal results. Future work can focus on expanding the dataset and exploring ensemble learning for further improvements.

Conclusion and Future Work

Glaucoma, a major cause of irreversible blindness, requires early detection to prevent vision loss. This project demonstrated the effectiveness of deep learning models, specifically VGG16 and ResNet50, in detecting glaucoma from fundus images. By leveraging transfer learning and fine-tuning pre-trained models, our system achieved remarkable performance metrics. The VGG16 model outperformed ResNet50, achieving an accuracy of 83.78% and an F1-score of 0.84, demonstrating its reliability in distinguishing between glaucomatous and non-glaucomatous images. These results emphasize the potential of automated systems to augment clinical decision-making, reduce diagnostic costs, and increase accessibility to glaucoma detection tools, particularly in underserved regions.

However, the study also highlighted certain limitations. The relatively small dataset size, despite augmentation, might limit the model's generalization ability. Additionally, while VGG16 showed superior performance, ResNet50 struggled with recall, indicating room for optimization in deeper architectures.

Future work will focus on addressing these limitations. Expanding the dataset with diverse and balanced samples from different demographics will enhance the model's robustness. Furthermore, incorporating advanced techniques like ensemble learning and attention mechanisms can improve feature extraction and classification accuracy. Real-world deployment can be facilitated by integrating the system with mobile or cloud-based platforms for ease of use in remote and resource-limited areas. Lastly, testing the system with clinical datasets and collaborating with ophthalmologists will ensure its clinical applicability and reliability.

References:

- Simonyan, K., & Zisserman, A. (2015). Very Deep Convolutional Networks for Large-Scale Image Recognition.
- Paszke, A., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library.
- World Health Organization. (2023). Global prevalence of glaucoma.
- Recent advancements in deep learning for retinal disease diagnosis: A review.