Kathmandu University
Department of Computer Science and Engineering
Dhulikhel, Kavre

A Project Proposal On

"Netra.ai"

[Code No: COMP206]
(For partial fulfillment of Year II/Semester I in Computer Science)

Submitted by:
Dixit Bhandari(07)
Aarya Pathak (38)
Arogya Poudel (41)
Umanga Rayamajhi (43)
Rupak Raj Karki (64)

Submitted to:

Mr. Suman Shrestha
Department of Computer Science and Engineering

Submission Date: 11/8/2024

Abstract

Eye diseases, such as diabetic retinopathy, glaucoma, and cataracts, are significant causes of vision loss worldwide, often going undetected until advanced stages due to limited access to specialized eye care. In Nepal, for instance, cataracts are the principal cause of blindness, followed by glaucoma and diabetic retinopathy, with a significant portion of the blind population residing in rural areas where access to eye health services is limited (Ruit et al, 2014)

Early detection is crucial in preventing or reducing severe vision impairments. Netra.ai addresses this need by developing an automated diagnostic system that uses machine learning to analyze eye images and detect signs of common diseases. Utilizing a convolutional neural network (CNN) trained on a dataset of annotated images, the system aims for high accuracy in identifying abnormalities. Deployed as a web application using Flask, it allows users to scan images for real-time analysis, which is especially beneficial in underserved regions like rural Nepal, where awareness of ocular conditions is limited. The anticipated outcome is an accessible, low-cost solution that supports early diagnosis, timely intervention, and reduces the burden of eye disease. This aligns with the work of healthcare providers like Dr. Sanduk Ruit, who has restored vision for over 100,000 people, underscoring the transformative impact of accessible eye care on communities.

Keywords: (Machine Learning, Convolutional Neural Network(CNN), Flask, OpenCV, Eye Disease Detection)

Table of Content

List of Figures

Acronyms/Abbreviations

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
ReLU	Rectified Linear Unit
ODIR	Ocular Disease Intelligent Recognition
GPU	Graphics Processing Unit
OpenCV	Open Source Computer Vision Library
Flask	A Python web framework

Chapter 1 Introduction

Eye health is a critical component of overall well-being, impacting millions of people worldwide. Conditions like diabetic retinopathy, glaucoma, and cataracts are among the primary causes of vision impairment and blindness, particularly in areas with limited access to regular eye care. Untreated eye diseases can greatly diminish a person's quality of life, affecting their ability to work, learn, and fully engage in society. To address this gap, our project, Netra.ai, aims to provide an accessible solution that not only detects eye diseases but also connects users with the best specialists for their specific condition.

Netra.ai is designed to develop a reliable, automated system for analyzing eye images, identifying signs of common diseases, and recommending expert medical care based on the diagnosis. The system leverages a convolutional neural network (CNN) model trained on a diverse set of annotated images to ensure high accuracy. Additionally, by integrating this capability into web application, users can conveniently upload eye images for analysis and receive guidance on finding a specialist suited to their diagnosis, making expert medical advice more accessible.

With the growing role of machine learning in healthcare, Netra.ai stands at the forefront of AI-driven medical diagnostics, blending advanced image processing with accessible technology. This project not only seeks to improve early detection but also promotes awareness of eye health, especially in regions with limited access to specialized care. Our vision is to empower users with information and resources, bridging the gap between detection and treatment through personalized medical recommendations.

1.1 Background

Recent advancements in artificial intelligence (AI) have revolutionized the field of medical diagnostics, with a particular focus on improving accuracy and efficiency in disease identification. In ophthalmology, AI algorithms are being developed to analyze retinal images, offering new possibilities in diagnosing eye diseases such as diabetic retinopathy, glaucoma, and macular degeneration. These developments are especially significant in underserved regions, where access to specialists is limited, and delays in diagnosis can lead to preventable vision loss. Current trends emphasize improving AI model precision and accessibility, making these tools invaluable in both urban hospitals and rural clinics, where trained personnel may be scarce.

However, despite these promising advancements, existing diagnostic methods still face critical challenges. Manual analysis of retinal images is time-consuming, often leading to delays, especially in remote areas lacking specialist access. Many AI systems, while proficient at detecting diseases, lack the functionality to recommend the best available specialist for specific conditions, limiting the potential for comprehensive patient care. Addressing these gaps, our Netra.ai project seeks to enhance diagnostic accuracy by not only detecting eye diseases from retinal images but also guiding patients to the most suitable specialist. By combining automated disease detection with tailored specialist referrals, Netra.ai aims to improve healthcare accessibility, reduce diagnostic delays, and set a new standard in AI-driven medical support.

1.2 Objectives

- **Automate Eye Disease Detection**: Develop an AI-driven system that utilizes convolutional neural networks (CNNs) to detect and classify common eye diseases, such as cataracts, glaucoma, and diabetic retinopathy, from real-time webcam images.
- **Increase Accessibility to Eye Care**: Provide a low-cost, accessible web application, particularly benefiting underserved and rural communities in Nepal, where access to specialized eye care is limited.

- **Develop a User-Friendly Interface**: Create an intuitive, easy-to-navigate web application using Flask that allows users to upload eye images or use a real-time webcam feature, providing instant diagnostic feedback in a comprehensible format.

1.3 Motivation and Significance

In Nepal, particularly in rural areas, the healthcare system faces significant challenges in providing adequate eye care services. Many individuals suffer from various eye diseases, yet access to advanced diagnostic tools and specialized care is limited. This project is motivated by the urgent need for a more reliable, accurate, and efficient approach to diagnosing eye diseases. The objective is to streamline the diagnostic process for both healthcare providers and patients, ensuring faster and more precise results.

Eye conditions such as cataracts, diabetic retinopathy, and glaucoma often go undiagnosed or misdiagnosed due to a lack of accessible screening methods and trained professionals. By leveraging machine learning technologies, specifically utilizing deep learning models like CNN for image analysis, this project aims to automate the detection of these diseases from eye images. This automation will enhance diagnostic consistency and reliability, reducing human error and improving overall patient outcomes (Amna et al., 2023; Muchuchuti et al., 2023).

Compared to existing practices that rely heavily on manual observation and subjective assessments, this project focuses on creating a robust system capable of analyzing eye images to identify potential diseases. The implementation of such technology not only aids healthcare professionals in confirming their diagnoses but also empowers them to provide timely referrals for specialized treatment when necessary.

Globally, there has been a surge in the development of advanced diagnostic tools for eye care; however, in Nepal, these technologies remain underutilized. By introducing an automated eye disease detection system, we aim to modernize the country's eye healthcare services, aligning them with international standards while addressing pressing local needs (Nazir et al., 2020; Leng et al., 2023).

Ultimately, this project promotes preventive care through early detection of eye diseases, which is crucial in reducing the incidence of vision impairment and blindness. It strengthens local healthcare services by equipping non-specialist health workers with essential tools to deliver quality care in areas lacking specialists. Overall, this initiative contributes significantly to enhancing Nepal's eye healthcare landscape, supporting better health outcomes nationwide while addressing the critical shortage of specialized services in rural regions.

1.4 Expected Outcomes

The goal of this project is to develop a machine-learning model capable of automatically analyzing eye images, using real-world data annotated by qualified eye care professionals. The model will be deployed in a simple web application built with Flask, allowing healthcare providers to upload eye images and instantly receive a detailed analysis of potential eye conditions.

The system aims to enhance diagnostic accuracy by cross-checking manual observations made by eye care professionals, providing a faster, more reliable method for diagnosing patients.

Key Outcomes:

- **Accurate Eye Disease Detection**: A CNN-based model capable of accurately classifying common eye diseases, such as cataracts, glaucoma, and diabetic retinopathy, from real-time webcam images.

- **Accessible and User-Friendly Web Application**: A Flask-based web interface that allows users to access the diagnostic tool easily, facilitating eye health monitoring, especially for underserved communities with limited access to specialized healthcare.
- **Scalable Solution for Early Detection**: A foundation for a scalable diagnostic tool that could potentially integrate additional eye conditions or be deployed in broader healthcare applications, addressing early-stage detection needs.
- **Real-Time Analysis**: A system that delivers instant diagnostic results, enhancing user experience by enabling immediate feedback without the need for manual image uploads or delays in processing.

Chapter 2 Related Works/ Existing Works

2.1 Computer Assisted Detector of Leukocoria (CRADLE)

The Cradle app scans photos on smartphones to identify unusual glares in the eyes that may indicate various eye diseases, including Coats disease and retinoblastoma. Developed by researchers at Baylor University, the app has been effective in detecting symptoms that are often overlooked during routine examinations. It aims to empower parents by enabling them to spot potential vision problems early, facilitating quicker medical consultations when necessary

Fig: 2.1
CRADLE

2.2 MIRAI

Although in a different field, MIT has developed several AI-based tools to enhance breast cancer detection, significantly improving early diagnosis. Their MIRAI system, predicts a patient's risk of developing breast cancer within five years by analyzing mammograms alongside patient history. This model has demonstrated high accuracy. MIRAI allows for screening schedules, reducing unnecessary surgeries and ensuring more precise diagnoses. Such AI-driven advancements are expected to revolutionize breast cancer screening and patient care globally.

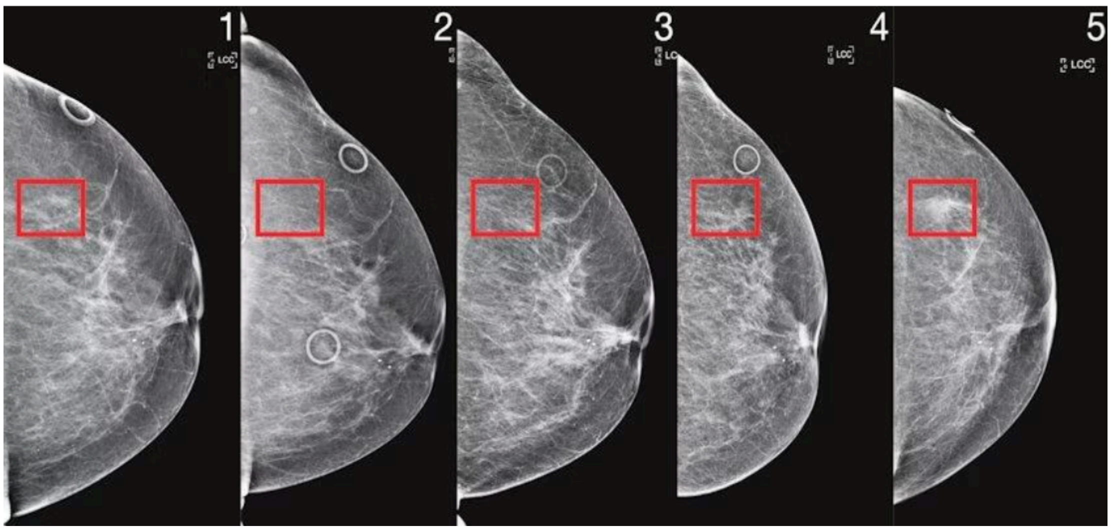

Fig: 2.2

Chapter 3 Existing Problem and Challenges

3.1 Challenges in Eye Health Among Nepal's Population

- **High Prevalence of Retinal Diseases:** The incidence of retinal diseases, such as diabetic retinopathy and age-related macular degeneration, is increasing in Nepal. These conditions are now the second leading cause of blindness in the country, exacerbated by a lack of effective screening methods and limited awareness among the population (Shrestha et al., 2020).
- **Limited Access to Eye Care Services:** Many rural areas in Nepal lack adequate eye care facilities and trained professionals. This scarcity leads to untreated eye conditions and prevents timely interventions, contributing to the growing burden of visual impairment (Tilganga Institute of Ophthalmology, 2020).
- **Inadequate Screening Programs:** Current screening initiatives, such as those conducted through community eye centers, often face challenges related to resource limitations and insufficient outreach. Many individuals at risk for retinal diseases remain undiagnosed due to ineffective screening strategies (Miller et al., 2022).
- **Lack of Comprehensive Patient Records:** Many clinics do not maintain detailed patient records, complicating efforts to track disease prevalence and manage public health initiatives effectively. This gap hinders the ability to monitor trends in eye health within communities (Tilganga Institute of Ophthalmology, 2020).

These combined challenges highlight the urgent need for innovative solutions like Netra.ai to improve eye health diagnostics and access to care in Nepal. By addressing these issues through technology-driven approaches, it is possible to enhance early detection, streamline referrals, and ultimately reduce the incidence of preventable blindness in the population.

3.2 Regulatory or Ethical Challenges

- **Regulatory approvals:** AI-based systems used in healthcare, including ophthalmology, must go through rigorous approval processes by regulatory bodies like the Health Ministry in Nepal. This can slow down the deployment of AI in real-world practice.
- **Accountability and liability:** Determining accountability and liability becomes particularly complex when an AI system provides an incorrect diagnosis or treatment recommendation. In traditional medical practice, the a doctor is responsible for the decisions made regarding patient care. However,with AI involved, the question arises: who should be held liable if something goes wrong, is it the doctor who followed the AI's suggestion, the developer who created the AI system, or the institution that implemented the technology in their practice. This dilemma has led to questionable concerns with regards to the use of AI in healthcare practice.

3.3 Resistance to the Adoption of AI

- **Human factor**: This plays a significant role in the adoption of AI in ophthalmology. Ophthalmologists and eye care staff may exhibit resistance to embracing AI technology for several reasons. One key factor is a general lack of understanding about how AI works and its potential benefits. Many professionals may not be familiar with AI systems and may feel uncomfortable relying on technology for tasks traditionally performed by humans. This uncertainty can lead to skepticism about AI's accuracy and reliability, especially when it comes to making critical diagnostic and treatment decisions.
- **Training and education**: Integrating AI effectively requires ophthalmology professionals to receive training on how to use these technologies, which can take time and resources.

3.4 Ethical Concerns in the Adoption of AI

- **Bias in AI systems**: AI systems may inherit biases present in training data. If models are trained on eye images from specific demographic groups, they may perform poorly on patients from other backgrounds, potentially leading to inaccurate diagnoses or treatment recommendations. For example, an AI system trained primarily on images of one ethnicity might struggle to detect certain eye conditions in others, or models trained on data from affluent individuals might not be as effective for patients from underserved communities.
- **Doctor-patient relationship**: Over-reliance on AI could weaken the human aspect of ophthalmology, where trust, empathy, and personal connection are essential. Patients may feel uncomfortable if their care relies mainly on AI, lacking the reassurance and individualized understanding an experienced ophthalmologist provides. AI cannot fully capture patients' fears, medical histories, or personal preferences as a human doctor can.

Chapter 4 Procedures and Methods

This project follows a structured approach with distinct stages, each contributing to the development of an AI-based eye disease prediction system. The journey unfolds through the following sections:

4.1 Planning

We will begin by planning and distributing tasks for an efficient workflow. Initial research on real-time eye disease detection systems will guide best practices for webcam-based image analysis. Our training dataset, sourced from Kaggle, will include annotated eye images representing conditions like cataracts and glaucoma, ensuring it aligns with project objectives.

A project roadmap with clear milestones for data preparation, model training, validation, and deployment will be created. Each team member will have defined roles, and regular check-ins will keep the development process on track.

4.2 Data Collection and Preparation

The data collection phase for Netra.ai involves gathering and preparing a comprehensive dataset to train our model for eye disease detection. We utilized the ODIR-5K dataset from Kaggle, which includes 5,000 pairs of left and right eye fundus images from patients diagnosed with various ocular diseases, such as cataracts, glaucoma, and diabetic retinopathy. This dataset provides the diversity and annotation quality required to train our convolutional neural network (CNN) model effectively.

To prepare the dataset for machine learning, the following pre-processing steps are applied:

- **Scaling**: All images are resized to a uniform resolution to ensure consistency in model input.
- **Normalization**: Pixel intensities are standardized to improve the model's ability to learn from subtle features across different images.
- **Data Augmentation**: Techniques like rotation, flipping, and brightness adjustments are applied to increase dataset variability and enhance model robustness.
- **Noise Reduction**: Filters are applied to minimize noise and improve image clarity, which is essential for accurate disease detection.

These steps ensure that the model can accurately interpret and classify eye conditions based on the input images. This approach aligns with best practices in medical image analysis, as discussed in the research by Li et al. (2021), which emphasizes the importance of pre-processing in enhancing model performance for diagnostic applications.

4.3 Model Design and Training

The model design phase focuses on selecting and training an appropriate machine learning model for accurate eye disease detection. We chose a **Convolutional Neural Network (CNN)** model due to its proven effectiveness in image classification tasks. The dataset from Kaggle (ODIR-5K) will be split into three parts:

- **Training set (70%)**
- **Validation set (15%)**
- **Test set (15%)**

The training process involves leveraging transfer learning with pre-trained CNN models to accelerate learning and enhance accuracy. These pre-trained models will be fine-tuned to recognize specific features in fundus images related to eye diseases. Key hyperparameters like learning rate, batch size, and regularization parameters will be optimized to improve model performance and prevent overfitting.

4.4 Debugging and Model Evaluation

The debugging phase is aimed to refine the model by reducing errors and inconsistencies. Extensive testing will be conducted using the validation set, followed by performance evaluation through several metrics:

- Accuracy: To assess the model's overall performance,
- Precision, Recall, and F1 Score: To measure the model's ability to correctly detect decay, missing teeth, and fillings,
- Confusion Matrix: To analyze true positive, false positive, and false negatives, offering insight into the model's effectiveness in identifying dental conditions.

Cross-validation techniques will be applied to ensure the model's robustness across different data subsets.

4.5 Model Deployment

The trained model will be deployed as a web-based application, accessible via an intuitive, user-friendly interface. Flask will be used to build the backend infrastructure, enabling real-time integration with a webcam. Users can access the application directly from their browsers, where the webcam captures live eye images for instant analysis. The model processes these images on the backend and returns diagnostic results in real time.

4.6 Prototyping and Testing

As the project progresses, the prototyping phase will involve refining Netra.ai based on real-world usage and feedback. Initial testing will focus on functionality and accuracy of the model in detecting eye diseases. Users and healthcare professionals will provide feedback on the system's usability and diagnostic reliability. Controlled testing and feedback loops will lead to iterative improvements, optimizing the model for practical performance and user experience. This continuous refinement process will ensure that the application is reliable, accurate, and effective for end users in real-time scenarios.

4.7 Workflow Diagram

The workflow diagram encapsulates the project's journey. It starts with planning, proceeds through data collection and preparation, and moves into model design and training. Debugging and evaluation phases ensure a high level of accuracy before the system is deployed and tested in real-world environments. Continuous feedback and prototyping refine the system, leading to the final deployment and validation of the model in clinical settings.

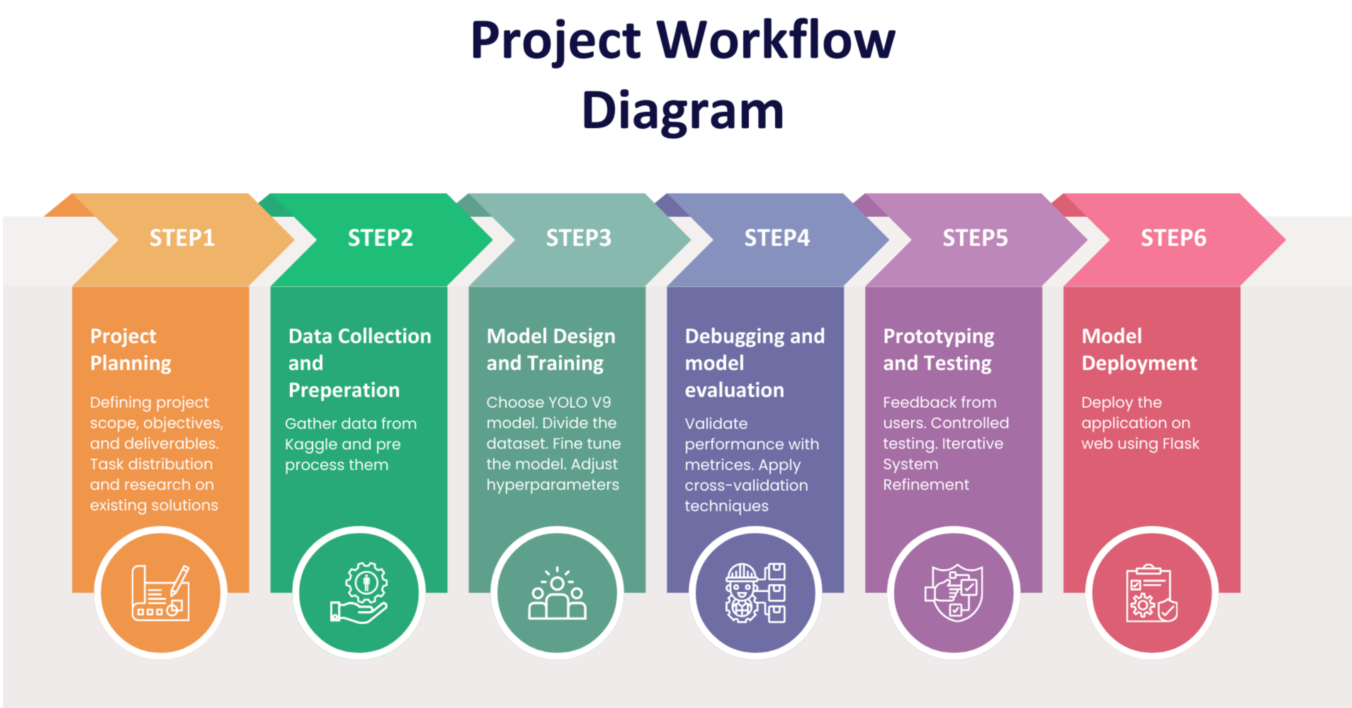

Fig: 4.7 Project Workflow Diagram

4.8 Model Architecture

A Convolutional Neural Network (CNN) is made up of several key layers, each designed to handle specific tasks in feature extraction and classification.

1. **Convolutional Layers**:
   - These are the primary layers in a CNN, tasked with extracting features from input images.
   - Each layer uses filters (also known as kernels) that slide over the image, calculating the dot product between the filter and small sections of the image.
   - This process creates feature maps that highlight important visual elements like edges, corners, and textures.

2. **Pooling Layers**:
   - After the convolutional layers, pooling layers reduce the size of the feature maps, lowering the number of parameters and computational load.
   - Pooling helps prevent overfitting and improves efficiency. Common pooling methods include max pooling and average pooling.
3. **Activation Layers**:
   - Activation functions are applied to the output of convolutional and pooling layers to introduce non-linear behavior, allowing the network to learn more complex patterns.
   - Popular activation functions are ReLU (Rectified Linear Unit), Sigmoid, and Tanh.
4. **Fully Connected Layers**:
   - These layers appear near the end of the CNN architecture and are responsible for making final predictions.
   - The output from the previous layers is flattened into a one-dimensional vector and passed through one or more fully connected layers.
   - The final layer, or output layer, provides the classification or regression results.

Example CNN Architecture:

- **Input Layer**: Receives the raw image data.
- **Convolutional Layer**: Applies filters to extract basic features.
- **Activation Layer**: Adds non-linearity (e.g., using ReLU).
- **Pooling Layer**: Reduces the spatial size of the feature maps.
- **Repeat**: Several rounds of convolution, activation, and pooling layers for deeper feature extraction.
- **Flattening Layer**: Converts the final feature maps into a single, flat vector.
- **Fully Connected Layers**: Process the flattened vector to make predictions.

● **Output Layer**: Produces the final classification or regression output.

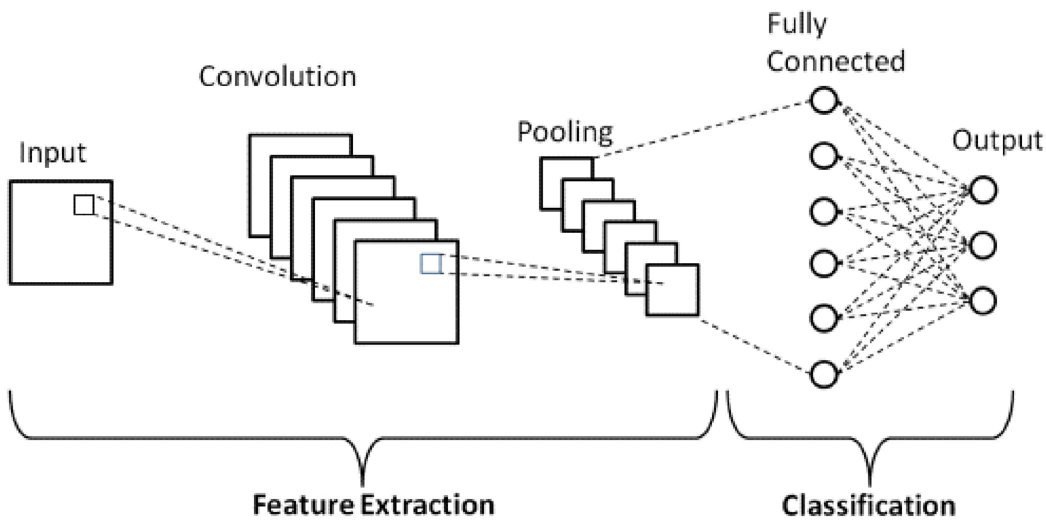

Fig: 4.81

Outline of CNN (Source: Medium.com)

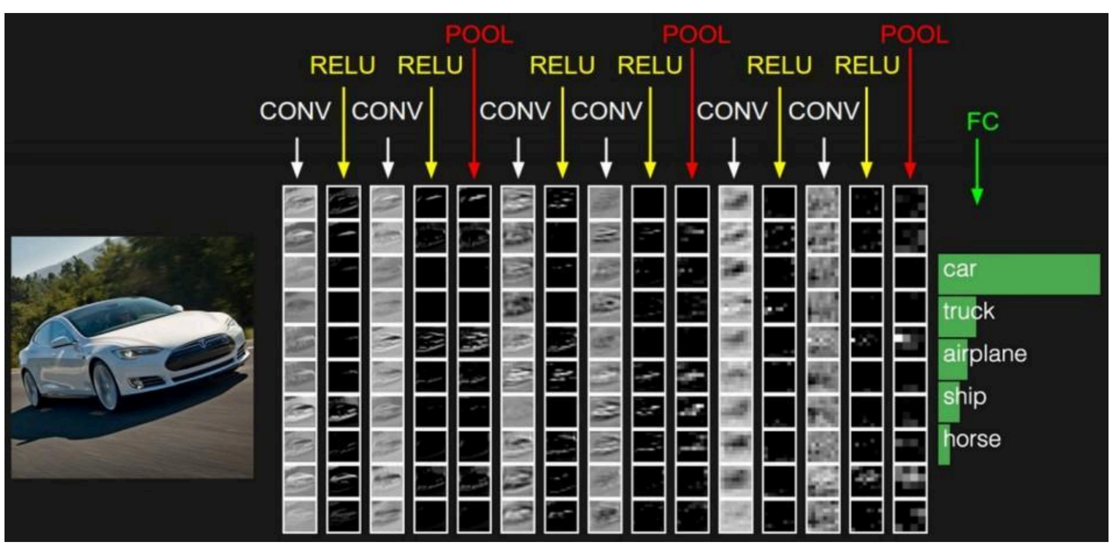

Fig: 4.82

CNN layers processing for image classification

(Source: Deep Learning for Computer Vision)

This structure enables CNNs to effectively learn and classify features in image data.

4.9 System Design

The process begins with an annotated dataset of eye images from the ODIR-5K dataset, split into 70% for training, 15% for validation, and 15% for testing. The model training uses a Convolutional Neural Network (CNN) architecture, optimized for image classification to detect various eye diseases. During training, transfer learning is applied by fine-tuning a pre-trained CNN model to enhance performance for specific eye conditions. Hyperparameters are adjusted, and the model's accuracy is validated using metrics such as accuracy, precision, recall, F1 score, and a confusion matrix, along with cross-validation techniques to ensure robustness. Once trained, the model is ready for application.

In the application phase, real-time images from the user's webcam are processed by the model. The system analyzes each frame, detecting and classifying potential eye conditions, such as cataracts, glaucoma, or diabetic retinopathy. The output provides users with instant diagnostic feedback, highlighting detected conditions to aid in early intervention and awareness. The system is divided into two primary stages: Model Training, where the classification model is built and refined, and Model Application, where the model is deployed in a Flask-based web app for real-time analysis using webcam input.

Chapter 5: System Requirement Specifications

5.1 Software Requirements:

To build and train the CNN model, a deep learning framework like TensorFlow is recommended due to its popularity and robust support, while OpenCV 4.5 or later is essential for handling webcam input and image processing tasks with the latest computer vision updates.

Operating System: A 64-bit operating system is required.

- Recommended options: Windows 10 or 11, or a Linux distribution like Ubuntu or MacOS (M2 chip or higher)

Python Version: Python 3.8 or newer, as it is optimized for most deep learning libraries.

Additional Libraries: Other libraries may be required for data preprocessing, model evaluation, and visualization.

- Recommended libraries: NumPy, Pandas, Scikit-learn, and Matplotlib for general data science and visualization tasks.

5.2 Hardware Requirements

Processor: Intel Core i5 or equivalent (4+ cores)

• Ram: 8 GB or more

• Graphics Card (optional): A CUDA enabled GPU with at least 6 GB VRAM (for GPU-based training)

• Operating System: Windows, Linux, or macOS (64-bit)

Chapter 6 Project Planning and Scheduling

This project spans seven weeks, starting with Week 1, where the scope, objectives, and roles are defined. Weeks 1-3 focus on Data Collection and Preparation, gathering and organizing relevant eye disease data. Weeks 3-4 involve AI Model Development, creating and refining machine learning algorithms, while Weeks 4-5 are dedicated to Prototype Development, where a user interface is integrated with the AI model for real-world testing.

The project moves into Weeks 5-6 for Testing and Training, rigorously evaluating the system and making necessary improvements. Finally, Week 7 is for Project Reporting, where a comprehensive report is compiled to document development, findings, and future recommendations. Each phase ensures the project progresses efficiently within the seven-week period.

TASK	WEEK 1	WEEK 1	WEEK 3	WEEK 4	WEEK 5	WEEK 6	WEEK 7
PROJECT INITIALIZATION	✓						
DATA COLLECTION AND PREPARATION		✓					
AI MODEL DEVELOPMENT			✓				
PROTOTYPE DEVELOPMENT				✓			
TESTING AND TRAINING					✓	✓	
PROJECT REPORTING							✓

Fig: 6
Gantt Chart

References

Ruit, S., Tabin, G., Paudyal, G., Gurung, R., & Paudyal, B. R. (2014). The epidemiology of blindness in Nepal: 2012. *BMC Ophthalmology, 14*(2). Retrieved from https://bmcophthalmol.biomedcentral.com/articles/10.1186/1471-2415-14-2

Baylor University. (2019, October 23). *Prototype smartphone app can help parents detect early signs of eye disorders in children*. Baylor University News. Retrieved from https://news.web.baylor.edu/news/story/2019/prototype-smartphone-app-can-help-parents-detect-early-signs-eye-disorders-children

Chan, C. (2021, January 28). *Robust artificial intelligence tools predict future cancer*. MIT News. Retrieved from https://news.mit.edu/2021/robust-artificial-intelligence-tools-predict-future-cancer-0128

Li, Z., Li, K., Hu, Y., Yu, J., Zhao, J., & Ma, C. (2021). A robust artificial intelligence model in healthcare. *arXiv preprint arXiv:2102.07978*. Retrieved from https://arxiv.org/pdf/2102.07978

Ahmed, H. M., & Tariq, M. (2020). The application of AI in medical imaging: Future outlook and challenges. *Journal of Applied Health Technology, 10*(3), 237-250. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC7382583/

TrueConf. (2020). *TrueConf Server in Tilganga Institute of Ophthalmology: Success story*. TrueConf. Retrieved from https://trueconf.com/blog/wp-content/uploads/2020/05/success_story_trueconf_server_in_tilganga_institute_of_ophthalmology_en.pdf