End-to-End Strabismus Detection Using Optimized CNN Architectures

By

Jayant Khandelwal(102215158)

Ashwin Rustagi(102215204)

Shubh Garg(102215216)

Rishit Garg(102215281)

Problem Statement

Strabismus, a misalignment of the eyes, affects both vision and appearance and often requires early detection to prevent long-term complications such as amblyopia or vision loss. Traditional diagnosis relies heavily on expert ophthalmologists and manual assessment, which can be time-consuming, subjective, and inaccessible in remote or under-resourced settings. With the growing capabilities of deep learning in medical imaging, there is a need to develop and evaluate robust, automated systems for accurate strabismus detection.

This study aims to design and compare multiple convolutional neural network (CNN) architectures—for the task of classifying strabismus into five categories: Exotropia, Esotropia, Hypertropia, Hypotropia, and Normal. The goal is to identify the most efficient and accurate model for deployment in real-time, portable diagnostic systems, potentially integrated into mobile or embedded platforms. The final system should support clinicians by providing fast, objective, and consistent screening results, especially in low-resource environments.

Methodology

Data Collection: Clinical eye images were collected from reliable sources, ensuring a balanced distribution across five classes: *Exotropia, Esotropia, Hypertropia, Hypotropia,* and *Normal*.

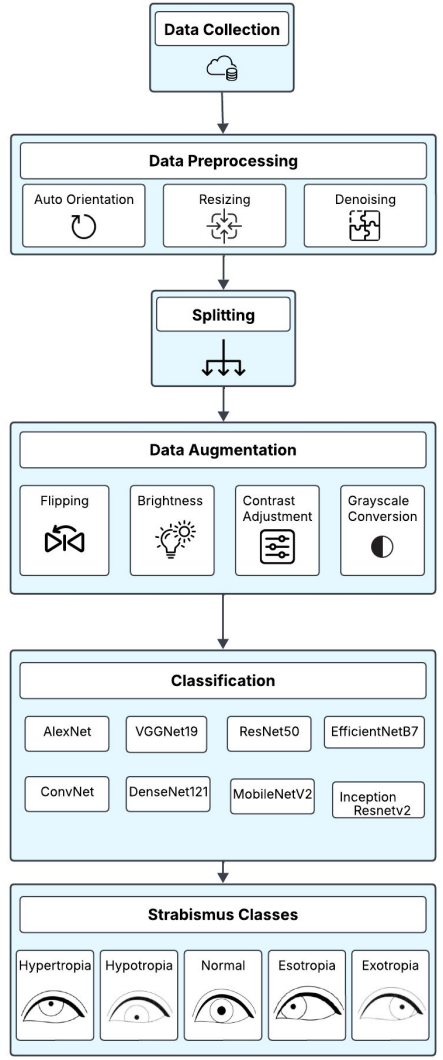
Data Preprocessing: Images were preprocessed with:Auto-orientation to correct rotated images,Resizing to a fixed input size (224×224 pixels),and Denoising to enhance image clarity and remove noise artifacts.

Dataset Splitting: The dataset was split into training, validation, and testing sets of 70:15:15 to enable robust model evaluation and prevent overfitting.

Data Augmentation: Applied only to the training set for generalization, including Horizontal flipping,Brightness adjustment,Contrast enhancement and Grayscale conversion .

Classification Architectures:Eight different CNN architectures were trained and evaluated AlexNet, VGGNet19, ResNet50, Inception-ResNetV2, EfficientNetB7,DenseNet121, MobileNetV2, and a custom ConvNet

Output Classes: Each model predicts one of the five **strabismus conditions**:*Hypertropia, Hypotropia, Normal, Esotropia, Exotropia*



Results

	Accuracy	Precision	Recall	F1-score
Efficient Net B7	84.00%	85.00%	83.00%	84.00%
Mobilenet V2	77.00%	70.00%	69.00%	69.00%
ResNet50	74.00%	79.00%	73.00%	74.00%
ConvNet	69.00%	70.00%	69.00%	69.00%
Inception Resnet V2	67.00%	68.00%	67.00	66.00%
VGGNet 19	66.67%	67.45%	66.67%	67.00%
Alexnet	59.26%	58.88%	59.26%	58.71%
Densenet 121	53.00%	52.00%	53.00%	52.00%

Individual Contributions

Jayant Khandelwal(102215158):

1.Trained and tested Mobilenetv2
2.Supported data augmentation experiments

Ashwin Rustagi(102215204):

1.Trained and evaluated DenseNet 121
2.Helped in curating ML pipeline

Shubh Garg(102215216):

1. Defined problem statement and project scope
2.Dataset curation, preprocessing
3.Trained and evaluated Inception ResNet V2

Rishit Garg (102215281):

1.Trained and Evaluated Custom ConvNet
2.Assisted in hyperparameter tuning