# The Librarian from Alexandria

Font Classification for Ancient Manuscripts


Mohammad Khair Hndauoi (794491)

Julia Milliet Sciorra (E05228)

Gabriele Goglia (806711)

# Project Overview

## CNN Model

CNN model to recognize 11 ancient manuscript fonts

## Dataset

Started with 1,200 images, expanded to 8,000

## Challenges

Data challenges: class imbalance, inconsistent formats, varying quality

## Goal

Goal: accurate classification despite limited training data

# Data Preparation

### Preprocessing

RGB conversion, resizing to 224×224, normalization

### Augmentation

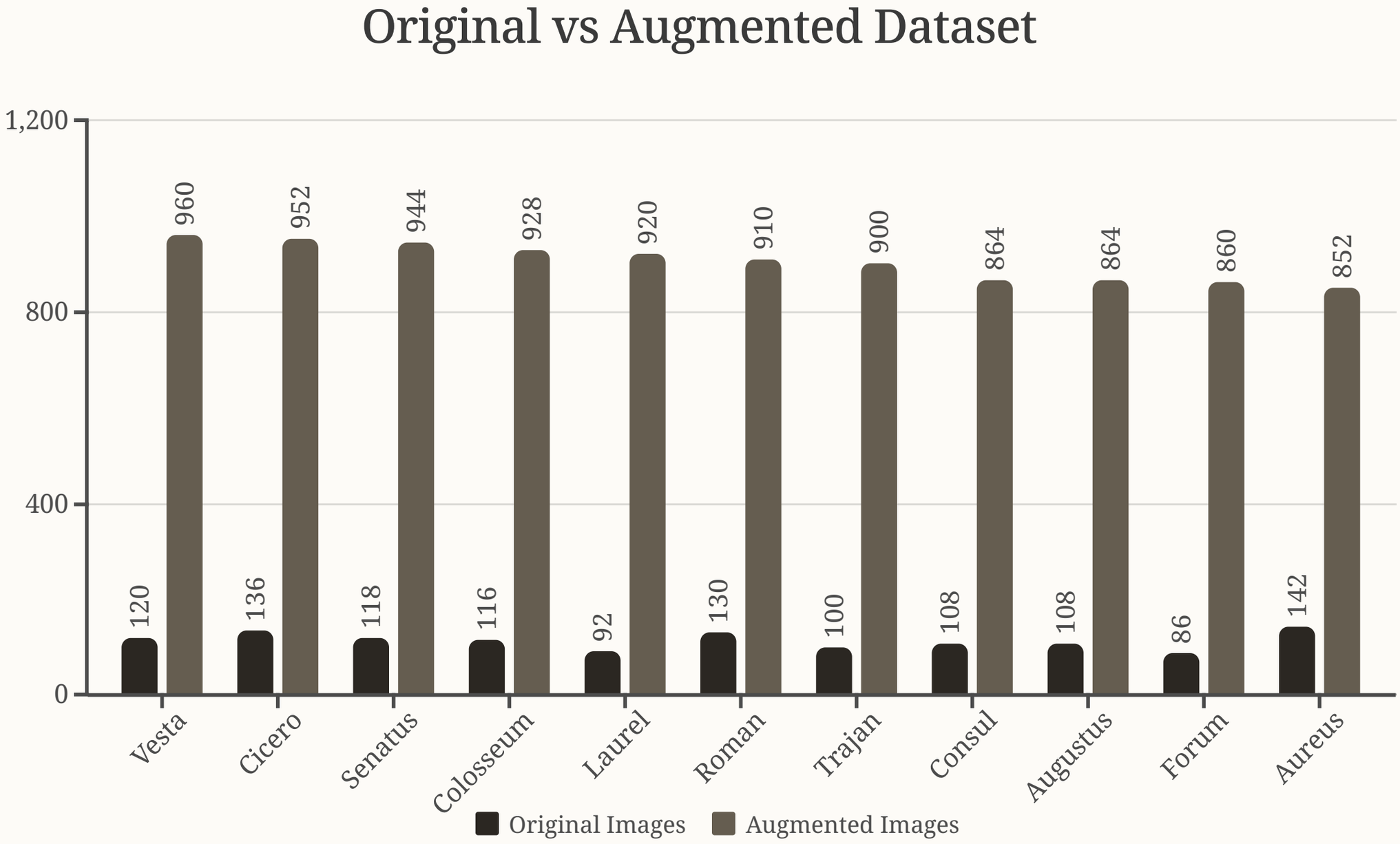
Rotation, brightness/contrast adjustments, blur, sharpening

### Dataset Expansion

Expanded dataset from 1,200 to 8,000 images

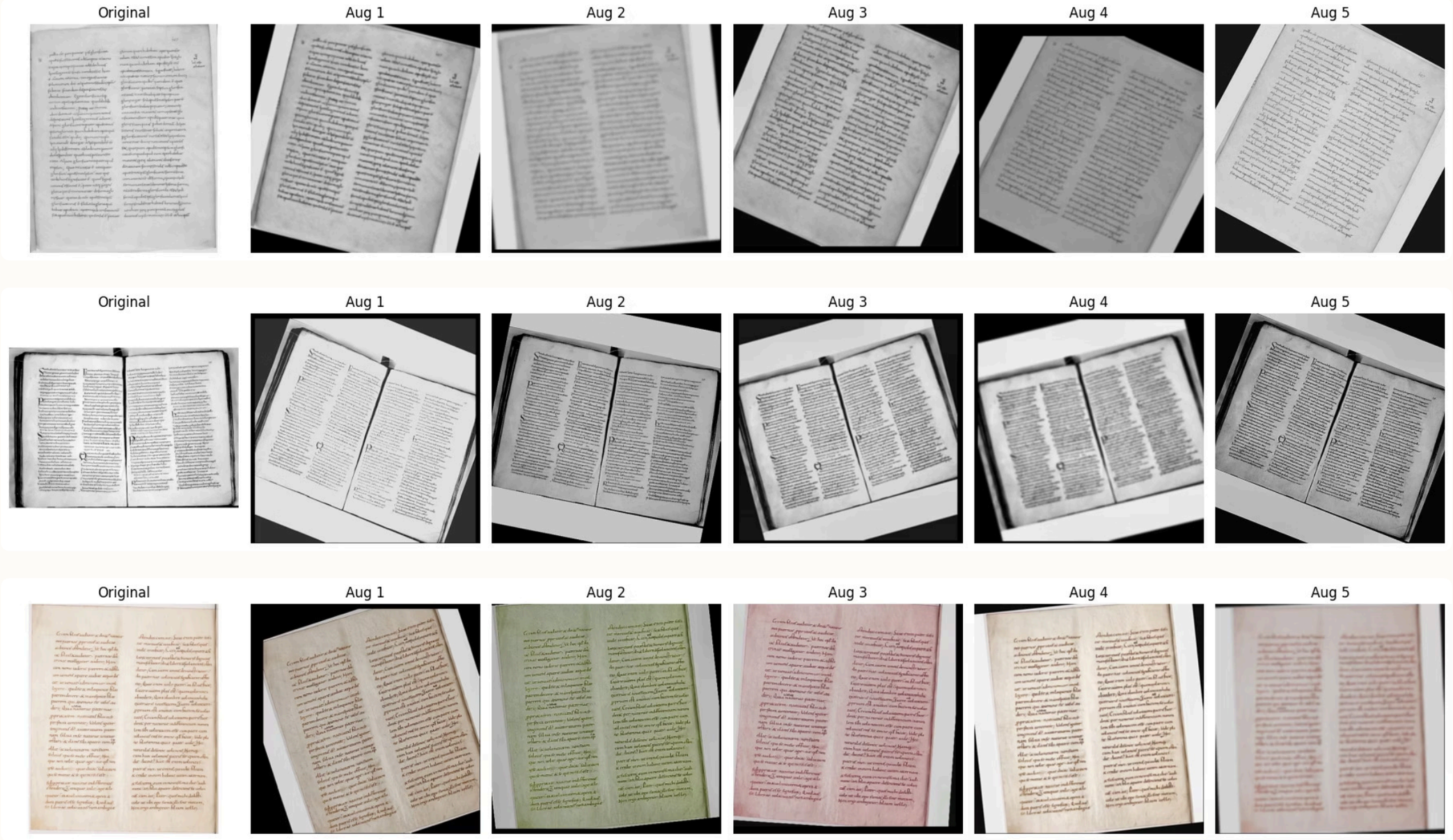
### Class Balancing

Balanced classes at ~850 samples each

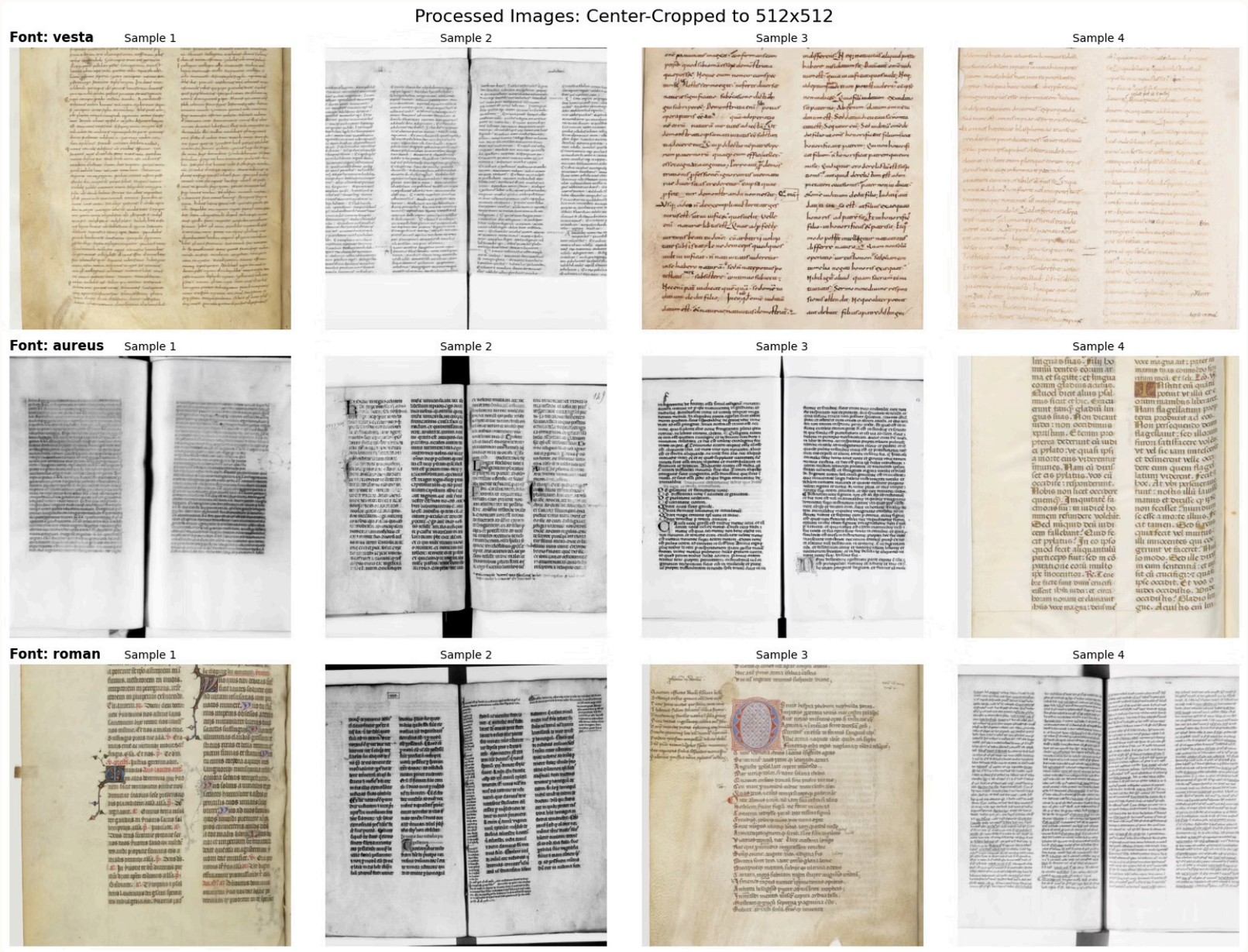



The chart shows a 7.93x multiplication factor. Initially underrepresented classes received more augmentation, creating balanced training data across all font types.

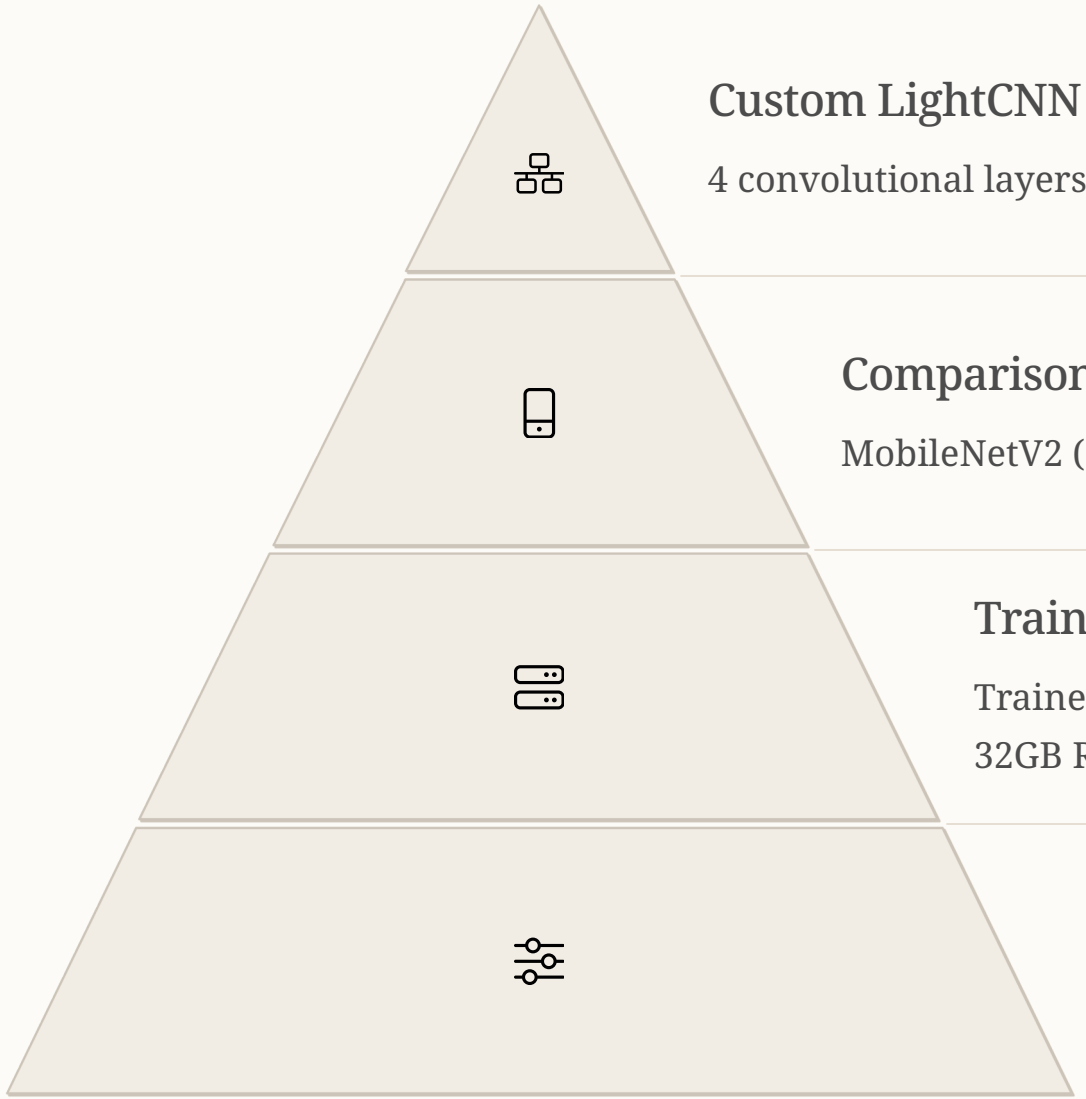
# Data Augmentation Samples



# Preprocessed Image Samples - White Space From Canvas Removed



# Model Architecture



## Custom LightCNN

4 convolutional layers, dropout for preventing overfitting

## Comparison Baseline

MobileNetV2 (other models crashed the GPU) with transfer learning

## Training Environment

Trained with PyTorch with torch_directml and on AMD GPU with 32GB RAM

## Optimization

Optimized parameters with Optuna, implemented early stopping

# Experimental Design

## Dataset Splitting

Stratified dataset splitting for balanced evaluation

## Evaluation Metrics

Metrics: accuracy, precision, recall, F1-score, confusion matrix

## Goals

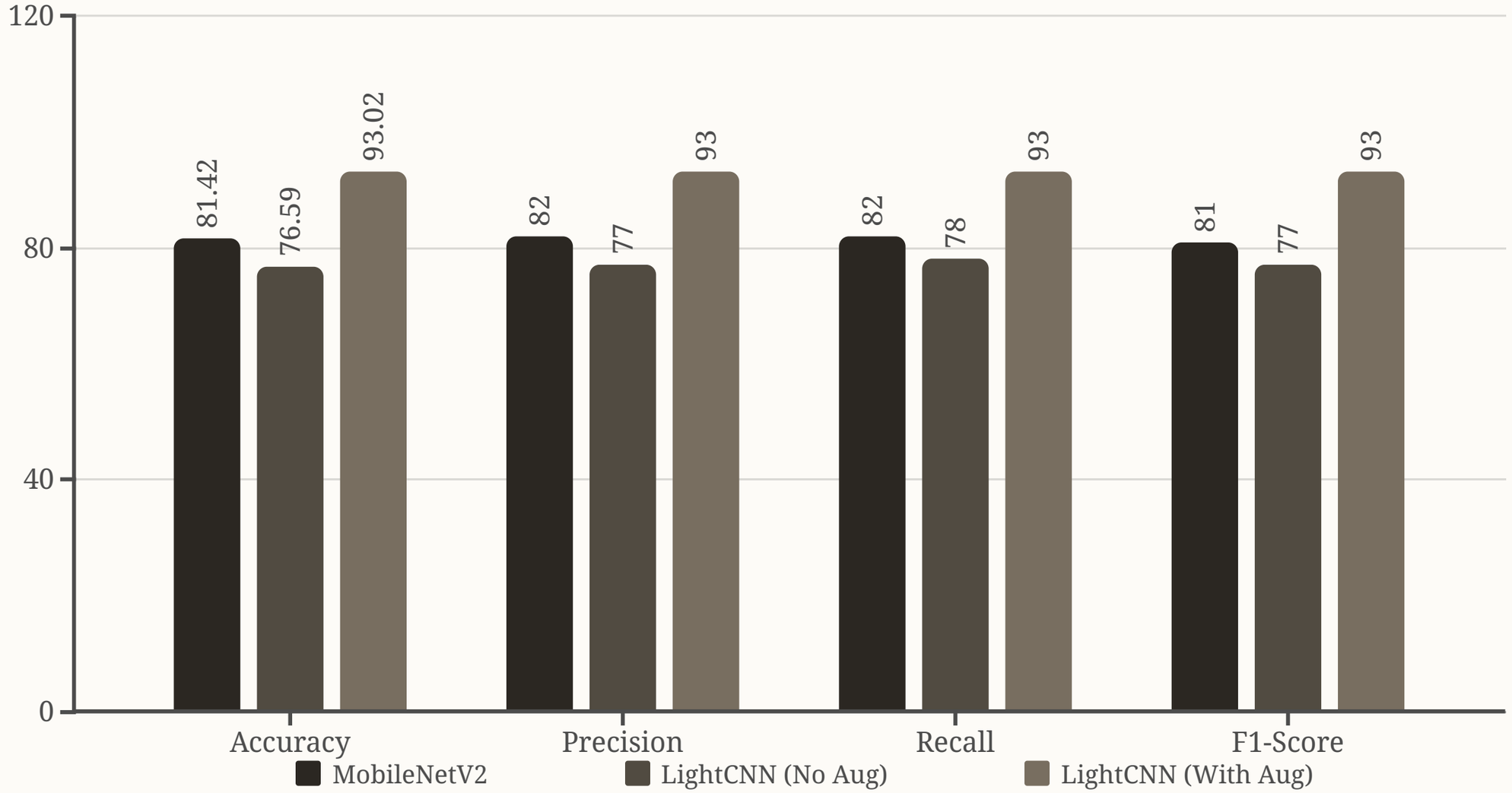
Goal: evaluate CNN effectiveness, impact of augmentation, architecture comparison

## Performance Focus

Focus on realistic performance to avoid overfitting

# Results: Model Performance Comparison

The chart shows performance metrics across three models. LightCNN with augmentations dramatically outperformed both alternatives.




Data augmentation provided +16.4% accuracy improvement over the baseline LightCNN and +11.6% over MobileNetV2.

# Key Takeaways

## Better Architecture

Task-specific CNN outperformed MobileNetV2

## Augmentation Impact

Data augmentation crucial for performance

## Effective Preprocessing

Effective preprocessing enabled strong results despite limited data

# Limitations & Future Work

## Dataset Limitations

Limited original dataset size

## Hardware Constraints

Hardware constraints affected training efficiency

## Advanced Augmentation

Potential for more advanced augmentation techniques

## Hardware Optimization

Room for optimization with better hardware

# Thank You!