

Facial Keypoints Detection – Project Summary

Author: Niher Ranjan Halder

Date: June 24, 2025

Model Type: Custom Convolutional Neural Network (CNN)

Tools Used: TensorFlow, Keras, Pandas, NumPy, Matplotlib, Seaborn

Final Model RMSE: 0.0230

Final Model MAE: 0.0163

Project Overview

This project focuses on detecting 30 facial keypoints (like eyes, eyebrows, nose, mouth corners) from 96×96 grayscale face images. It applies deep learning techniques to solve a computer vision problem with real-world applications in:

- ☐ Face tracking
-  Biometrics
- ☐ Medical diagnostics
-  Augmented reality (AR)

Dataset Summary

-  Training samples: 2140 images with labeled keypoints
- ☐ Test samples: 1783 unlabeled images
-  Image size: 96×96 pixels, grayscale
- ☐ Lookup Table: Specifies which keypoints must be predicted per test image

Model Architecture

A custom-built CNN (Convolutional Neural Network) with:

Multiple convolutional + max-pooling layers

ReLU activations

Dropout regularization

MSE loss function

RMSE + MAE evaluation

EarlyStopping and data augmentation applied

Model Performance

✓ Final Validation RMSE: 0.0230

✓ Final Validation MAE: 0.0163

🚀 Shows strong generalization and convergence

Visualizations

✖ predicted_vs_actual_keypoints.png – Red vs green overlay of prediction vs actual

 training_validation_rmse_curve.png – RMSE & MAE curves

 cnn_architecture_diagram.png – CNN model structure

 face_grid_keypoints_comparison.png – Grid of predicted test faces

 model_performance_metrics.png – Metric card

 submission_file_preview.png – Final test predictions snapshot

Deliverables

-  Jupyter notebooks (6 stages: EDA → Submission)
-  Trained model file (cnn_model.h5)
-  Submission CSV
-  Visual gallery of evaluation & learning curves
-  This PDF report

Conclusion

-  Successfully developed a deep learning pipeline for facial keypoint detection.
-  Achieved strong model performance with CNN (RMSE = 0.0230).
-  Model applicable to AR, biometric systems, facial analytics, and healthcare.
-  Demonstrated deep learning proficiency in a real-world, image-based problem.
-  Deliverables include clean code, trained model, submission file, and visual results.

“AI doesn't just recognize faces - it understands the points that make us human.”