

Face Landmarks

MODEL DETAILS

This model is designed to perform multi-task facial analysis from a single grayscale face image. It estimates 23 facial landmarks, validates the presence of a face, assesses whether the face is in a frontal pose, and predicts face translation. These outputs support downstream tasks such as facial alignment, tracking, and pose-aware recognition in embedded vision systems.

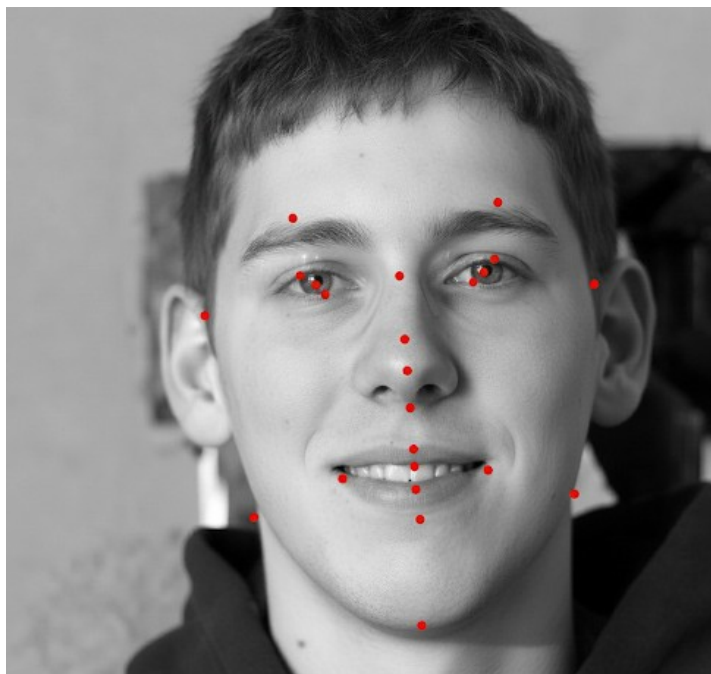


Figure 1: Facial Landmarks Model Output.

MODEL SPECIFICATIONS

Inputs

- 96×96×1 grayscale image of a face, typically extracted as a region of interest (ROI) from a face detector or tracked from a previous frame.

Outputs

- Face validation confidence
- Frontal pose indicator (frontal if yaw < 60 degrees)
- Face translation vector
- 23 facial landmark coordinates

Architecture

- MobileNetV2-like lightweight convolutional neural network

Parameters

- 657,202 (643,170 trainable, 14,032 non-trainable)

AUTHORS Lattice Semiconductor

VERSION fl-fpga-8.1.1

RELEASE 2025-12-18

SOURCE CODE [Training Source Code](#)

TOOLCHAIN [LATTE](#)

[LSCQuant](#)

PERFORMANCE EVALUATION

Live Evaluation

Evaluations were performed on Lattice CLNX-33 FPGA and IMX219 camera, under a controlled, standardized environment to ensure consistency and reproducibility of results. While these conditions shaped the reported metrics, the model architecture is designed for flexibility, supporting potential deployment across a wide range of platforms, including non-FPGA environments.

The model was tested on single-face inputs where the face was not required to be centered within the region of interest (ROI). Some facial landmarks were permitted to fall outside the ROI boundaries. While the test setup focused on constrained ROI positioning, the model is capable of handling more flexible face placements.

Offline Evaluation

The offline evaluations were performed under the following conditions:

- Face datasets: Multi-PIE, Wider Facial Landmarks in the Wild (WFLW), IBug, Face Ext, totaling 50,000 face images.
- Negative samples (images without faces) were sourced from Lattice internal datasets, also totaling 50,000 images.
- Input Constraints: Facial landmark bounding boxes were constrained to sizes between 38 and 84 pixels. The face center was required to be within 24 pixels of the input image center to ensure consistent performance.
- Note: Because the datasets are anonymized, individual identities cannot be tracked. Consequently, it is possible that the same individuals appear in both the training and test splits.

The KPIs for the evaluation data are reported in Table 1. Categorical metadata for quantitative analysis were generated using [Gender Classifier](#), [Age Classifier](#), and [Deep Face](#). Figure 2 illustrates the false negative rate by age, ethnicity and gender categories.

Table 1: Model performance across datasets.

Metric	Value
Landmarks MAE	1.22 pixels
Face Translation MAE	0.57 pixels
Face Validation Accuracy	99.95%
Frontal Pose Accuracy	99.21%

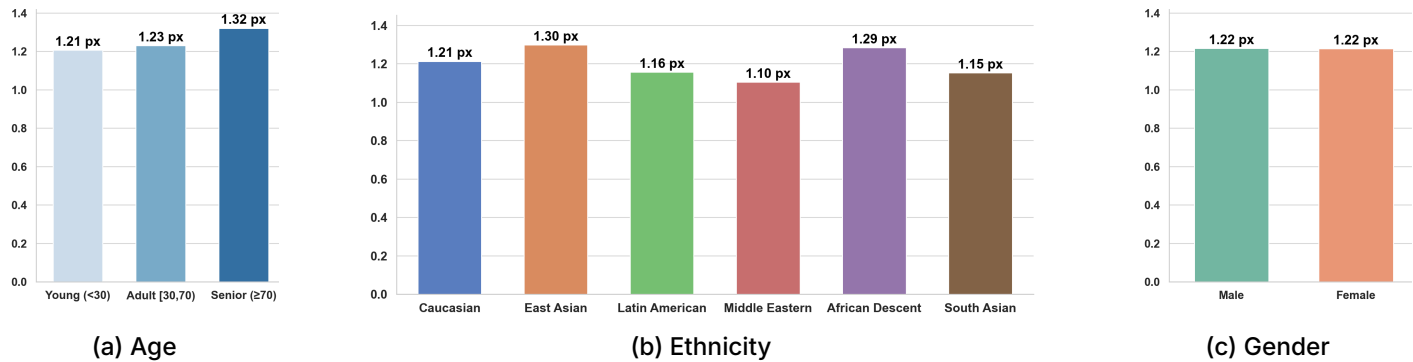


Figure 2: Face landmarks MAE by age, ethnicity, and gender categories.