MediaPipe Face Detection

Overview

MediaPipe Face Detection is an ultrafast face detection solution that comes with 6 landmarks and multi-face support. It is based on BlazeFace, a lightweight and well-performing face detector tailored for mobile GPU inference. The detector's super-realtime performance enables it to be applied to any live viewfinder experience that requires an accurate facial region of interest as an input for other task-specific models, such as 3D facial keypoint estimation (e.g., MediaPipe Face Mesh), facial features or expression classification, and face region segmentation. BlazeFace uses a lightweight feature extraction network inspired by, but distinct from MobileNetV1/V2, a GPU-friendly anchor scheme modified from Single Shot MultiBox Detector (SSD), and an improved tie resolution strategy alternative to non-maximum suppression.

Supported configuration options:

- model_selection
- min_detection_confidence

MODEL_SELECTION

An integer index `0` or `1`. Use `0` to select a short-range model that works best for faces within 2 meters from the camera, and `1` for a full-range model best for faces within 5 meters. For the full-range option, a sparse model is used for its improved inference speed. Default to `0` if not specified.

MIN_DETECTION_CONFIDENCE

Minimum confidence value (`[0.0, 1.0]`) from the face detection model for the detection to be considered successful. Default to `0.5`.

DETECTIONS

Collection of detected faces, where each face is represented as a detection proto message that contains a bounding box and 6 key points (right eye, left eye, nose tip, mouth center, right ear tragion, and left ear tragion). The bounding box is composed of `xmin` and `width` (both normalized to `[0.0, 1.0]` by the image width) and `ymin` and `height` (both normalized to `[0.0, 1.0]` by the image height). Each key point is composed of `x` and `y`, which are normalized to `[0.0, 1.0]` by the image width and height respectively.