RetinaFace- Single-stage Dense Face Localisation in the Wild

1

Proposal

This paper presents a robust single-stage face detector, named RetinaFace, which performs pixel-wise face localisation on various scales of faces by taking advantages of joint extra-supervised and self-supervised multi-task learning

Robust · single-stage detector

RetinaFace

pixelwise face localisation

1

(1) We manually annotate five facial landmarks on the WIDER FACE dataset and observe significant improvement in hard face detection with the assistance of this extra supervision signal. (2) We further add a selfsupervised mesh decoder branch for predicting a pixel-wise 3D shape face information in parallel with the existing supervised branches. (3) On the WIDER FACE hard test set, RetinaFace outperforms the state of the art average precision (AP) by 1.1% (achieving AP equal to 91.4%). (4) On the IJB-C test set, RetinaFace enables state of the art methods (ArcFace) to improve their results in face verification (TAR=89.59% for FAR=1e-6).
(5) By employing light-weight backbone networks, RetinaFace can run real-time on a single CPU core for a VGA-resolution image. Extra annotations and code have been made available

Five aspects

1

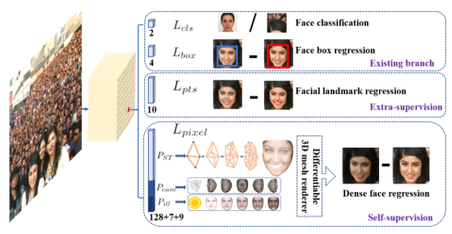

Figure 1. The proposed single-stage pixel-wise face localisation method employs extra-supervised and self-supervised multi-task learning in parallel with the existing box classification and regression branches. Each positive anchor outputs (1) a face score, (2) a face box, (3) five facial landmarks, and (4) dense 3D face vertices projected on the image plane.

1

face detection features smaller ratio variations (from 1:1 to 1:1.5) but much larger scale variations (from several pixels to thousand pixels).

Anchor Box => 1:1 or 1:1.5

SIFT - How can we get?

2

Inspired by [6], MTCNN [66] and STN [5] simultaneously detected faces and five facial landmarks