**Abstract**

A research paper in the domain of face detection which has already been published, is replicated and the results have been reproduced, which might prove the appropriateness of the original work by authors, by achieving similar results. Multi-task and extra-supervised learning are utilized in performing a process of face localization with respect to each pixel in the image. This process is a part of a powerful single stage face detector which is termed as RetinaFace. The test dataset has been constructed in which each image from the Wider Face dataset has been annotated manually by recognizing the five facial positions. Pixel-wise shape prediction of these images in the dataset is done using a self-supervised algorithm/code. Further, the difference between the shape of pixels that are manually annotated and the existing data of images with actual pixel shape is observed.

1. **Source Paper Description**

The research paper that we are trying to replicate is going to be simple and will involve the stepwise reproducibility. The code and data of this research paper is handled in a manageable way which will include the management of data, explanation of implementation in each step. All the resources that are required for replication and reproduction of results will be mentioned clearly.

The research paper which is replicated is a face localization: RetinaFace research which deals with accurate face detection and efficient localization of a face in an image. Usually it has been seen that a single model cannot accommodate the training of multiple datasets and multiple models are required, but in case of RetinaFace, a single model can be used to train on different datasets containing images, which is an advantage as it is time-saving and localization of facial landmarks can be done in a flexible and simple manner.

* 1. **Evaluation Framework**

In the original work, the dataset contains 32,203 images and 393,703 bounding boxes which are created by annotating the five facial landmarks (eyes centre, nose tip and mouth corners). All these images are quite different from each other in scale, position, illumination, etc. The dataset is divided into training, validation, and test subsets, (40%,10%,50%) respectively. To get bigger sizes of training faces, the images are randomly cropped in a square shape and then resized to 640x640. SGD optimizer was used to train the RetinaFace and four GPUs were required. Box voting which uses a certain IoU threshold, flip, multi scale strategies were used to test the Wider Face dataset and perform evaluation.

The evaluation is done using the metrics such as average precision (AP) and mean AP (mAP) with different value of IoU to achieve accurate face detectors. The RetinaFace was compared with different algorithms which are used in face detection such as, multi-stage cascade CNN, two-stage CNN, CMS-RCNN, HR, MSCNN, Pyramid Box, ISRN and many more. RetinaFace outperforms other algorithms in case of average precision (AP) evaluation metric and gets the applauding AP in every subset of validation and test sets.

Other evaluation metrics such as normalized mean errors, cumulative error distribution can be used to evaluate the accuracy of five facial landmark. ArcFace verification accuracy is used to evaluate face recognition accuracy. Precision recall curves on the validation and test subsets of WiderFace are generated to interpret and compare the RetinaFace with different algorithms/models in which RetinaFace outstands others.

RetinaFace produces qualitative results and can detect numerous faces (upto 900 out of 1151) in a single image with accurate bounding boxes and efficient prediction of five facial landmarks in each face. The resolution, position, scale, illumination, pose of each face can be detected precisely. Techniques like dense regression can be used to overcome the failure in localization of dense face.

* 1. **Justification**

The quality of the research paper is justified by looking at the number of citations it has, the core rankings, google scholar discipline ranking. Our research paper, the RetinaFace was published in the IEEE Conference on Computer Vision and Pattern Recognition (May 2019) which has an A1 rating given by Qualis **[1]** and CORE rating of A\* **[2]**. According to Google Scholar, this research paper has 107 citations **[3].** A Github repository which contains all the codes, datasets, and other resources which were used in building this research paper is available **[4].**

[1] <http://www.conferenceranks.com/#data>

[2] <http://portal.core.edu.au/conf-ranks/604/>

[3] <https://scholar.google.com/scholar?hl=en&as_sdt=0%2C5&q=RetinaFace%3A+Single-stage+Dense+Face+Localisation+in+the+Wild+citations&btnG=>

[4] <https://github.com/deepinsight/insightface/tree/master/RetinaFace>