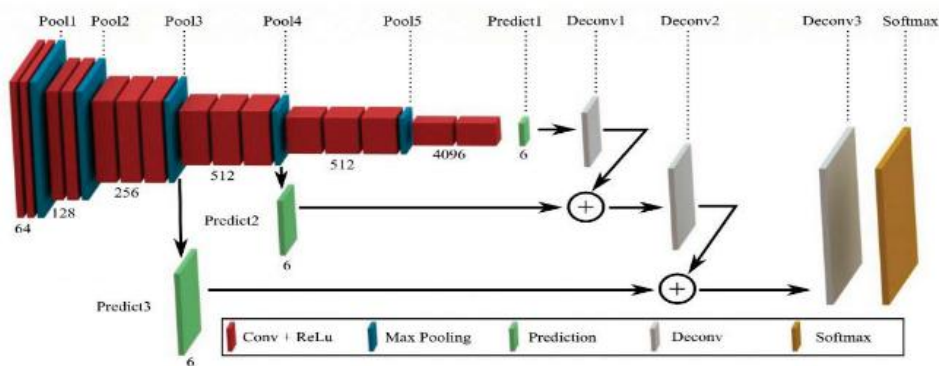


Summary of Facial Mask Detection using Semantic Segmentation:

In this paper they try to creating a binary face classifier which can detect faces in any orientation irrespective of alignment and train it in an appropriate neural network to get accurate result. Their model's basic function is feature extraction and class prediction.

In their proposed method they like to use Gradient descent algorithm for getting their output which is a feature vector. Their model detect multi faces in one frame. The input image of any arbitrary size is resized to $224 \times 224 \times 3$ and fed to the FCN network for feature extraction and prediction. Initially the pixel values of the face and background are subjected to global threshold. After that it's passed through median filter to remove the high frequency noise and then subjected to closing operation to fill the gaps in the segmented area. Here they using pre-defined training weights of VGG 16 architecture for the feature extraction and prediction. They use VGG 16 basic architecture algorithm which figure is given below -



Their contribution is their model consists of a total of 17 convolutional layers and 5 Max pooling layers. Convolutional layer convolutes the input image with another window while the max pooling operation ensures that the size of the feature vector being produced in every layer is halved so as to reduce the number of parameters. The initial layers extract the lower level features while as the subsequent layers extract the mid-level and higher level features. After the final max pooling layer the image size will be reduced. Post processing on the predicted mask obtained is performed so that the irregularities in the region can be filled and to remove the unwanted errors. They perform by first passing the mask through Median filter and then performing the Closing Operation. This ensures that the gaps in the segmented region are filled and most of the unwanted false erroneous prediction removed. They designed the model such that all those erroneous predictions are not considered while showing the final detected faces. They find the mean pixel level accuracy for facial masks for their model is 93.884%.

They don't describe the algorithms properly which they use here and how they use. They told here they perform a post processing on the predicted mask for removing unwanted data but they don't describe here how they processing this data's for the result.

In our study we can get help from this paper. They use face mask detection for multiple faces so in our study we also need to do this. They done post processing on the predicted mask obtained for remove unwanted errors it will also help us in our study.

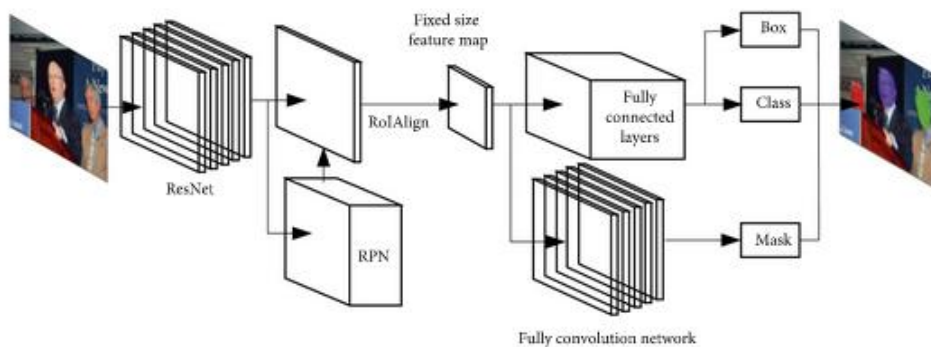
Link:

https://www.researchgate.net/publication/336952877_Facial_Mask_Detection_using_Semantic_Segmentation

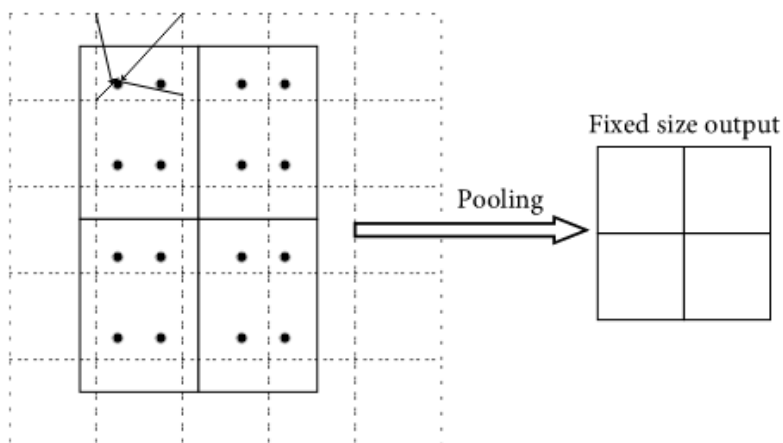
Summary of Face Detection and Segmentation Based on Improved Mask R-CNN:

They think most of the existing detection methods only localize each face using a bounding box, which cannot segment each face from the background image simultaneously. To overcome this drawback, this paper presents a face detection and segmentation method based on improved Mask R-CNN, named G-Mask, which incorporates face detection and segmentation into one framework aiming to obtain more fine-grained information of face.

In their proposed method they used ResNet-101 for feature extraction, they used RPN to generate RoIs, and RoIAlign faithfully preserves the exact spatial locations to generate binary mask through Fully Convolution Network (FCN). They also used Generalized Intersection over Union (GIoU) as the bounding box loss function to improve the detection accuracy. They compared with Faster R-CNN, Mask R-CNN, and Multitask Cascade CNN, the proposed G-Mask method has achieved promising results on FDDB, AFW, and WIDER FACE benchmarks. Their used G-mask framework is given below –



Their used RoIAlign layer is look like –



A new dataset was created which annotated 5115 images randomly selected from the FDDB and Choke Point datasets. A face detection and segmentation method based on improved Mask R-CNN was proposed, which can detect faces correctly while also precisely segmenting each face in an image. Furthermore, the proposed method improves the detection performance by introducing GIoU as a bounding box loss function.

The experimental results verify that our proposed G-Mask method achieves promising performance on several mainstream benchmarks, including the FDDB, AFW and WIDER FACE benchmarks.

Here they used many methods for their task. So this will take a long time for working. They don't use anything to reduce time.

Their detection methods only cannot localize each face using a bounding box, which can segment each face from the background image simultaneously. This part in this paper will be helpful for our project.

Link:

<https://www.hindawi.com/journals/ddns/2020/9242917/#abstract>