

Unmasked

Removing Face Masks with Machine Learning

Tian Yu Fan, Grant Perkins, Jean Claude Zarate, Mingjie
Zeng

Introduction

Facial Recognition

- ❖ Analyzing security footage after incidents
- ❖ Unlocking phones, other security devices.
- ❖ Partial Face Recognition
 - Weng et. al proposed using a scale-invariant feature transform keypoint detector + keypoint matching scheme. Requires unmasked image to identify against.
 - Others proposed focusing on key features (eyes and forehead)

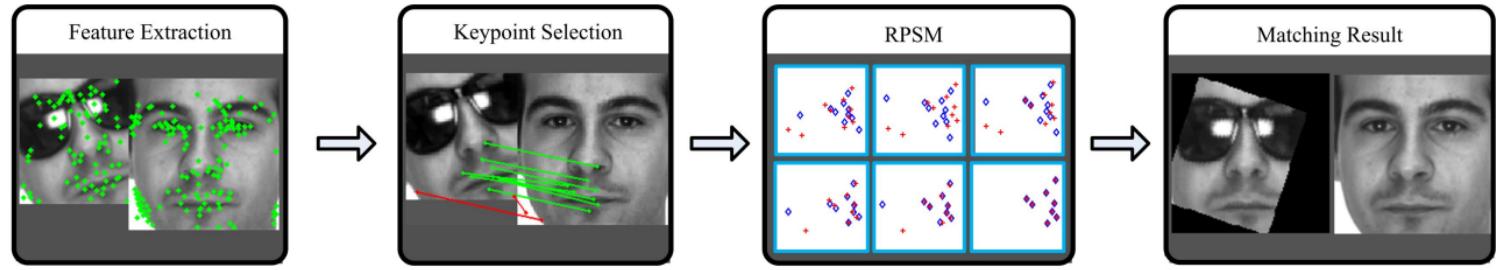


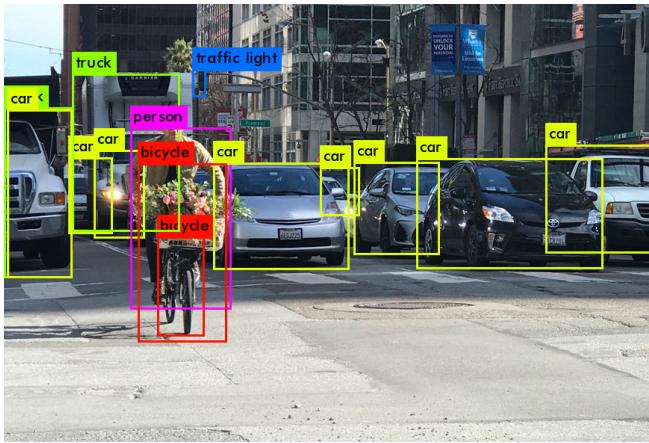
Image Inpainting

- ❖ Reconstruct missing regions of an image
- ❖ Achieving a realistic result
- ❖ Masks are commonplace now, defeating facial recognition systems.
- ❖ Can use image inpainting to regenerate the masked parts of faces

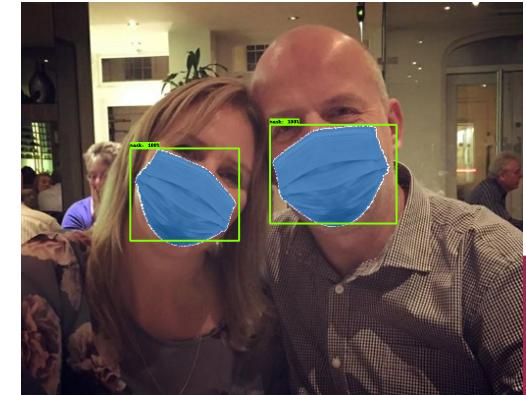


Object Detection

- ❖ YOLO (you only look once) is a famous detection system
- ❖ Creates detection candidates, runs classifier on candidates
- ❖ A network can be further trained to detect specific features



Left image: Redmon et al: "YOLOv3: An Incremental Improvement"

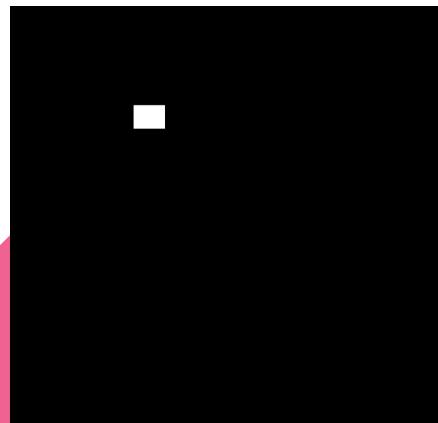
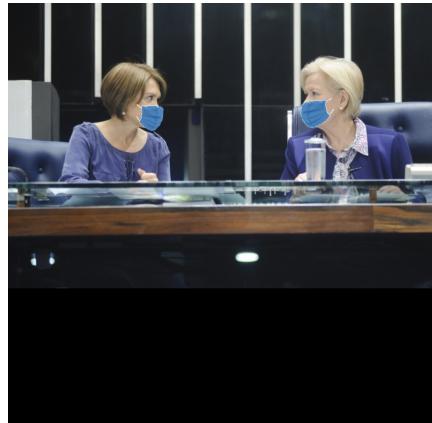


Right image:: picture from our dataset

Methodology

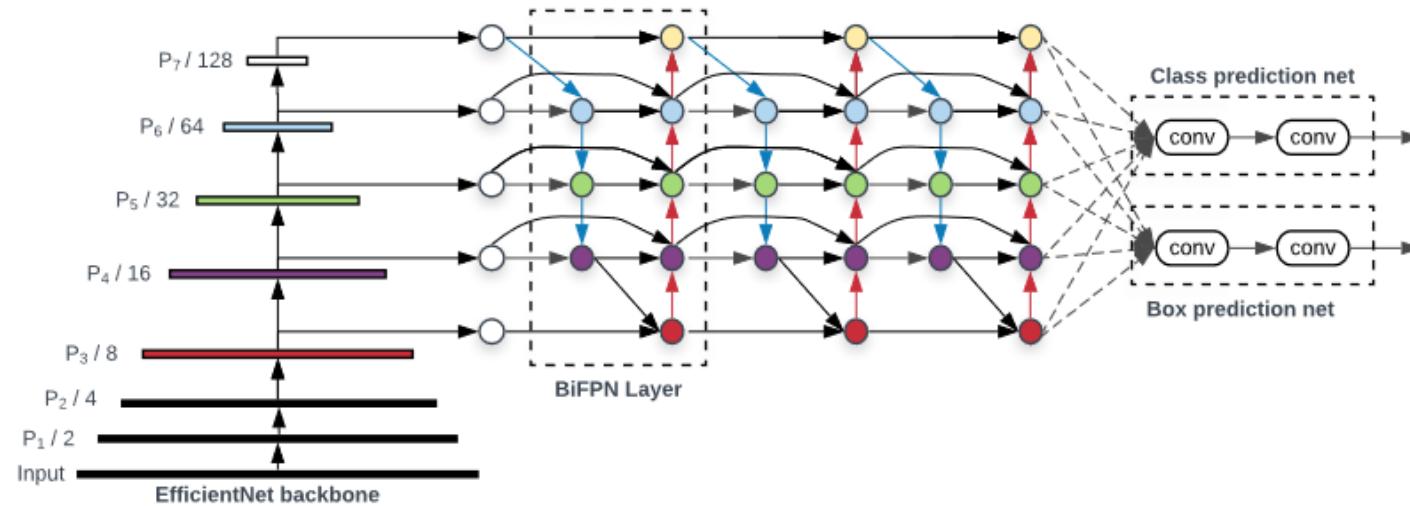
Data Formatting & Gathering

- ❖ Unmasked dataset - 2,932 Images from Open Images V6
- ❖ Masked dataset - MaskTheFace Facial Recognition Tool
 - CSV - image id, image dimensions, face bounding box locations, mask landmarks
- ❖ Resized dataset (all images resized to 800x800)
 - CSV - image id, mask bounding box locations
- ❖ Mask detection - TFRecord (80/20 for training/testing)



Mask Detector

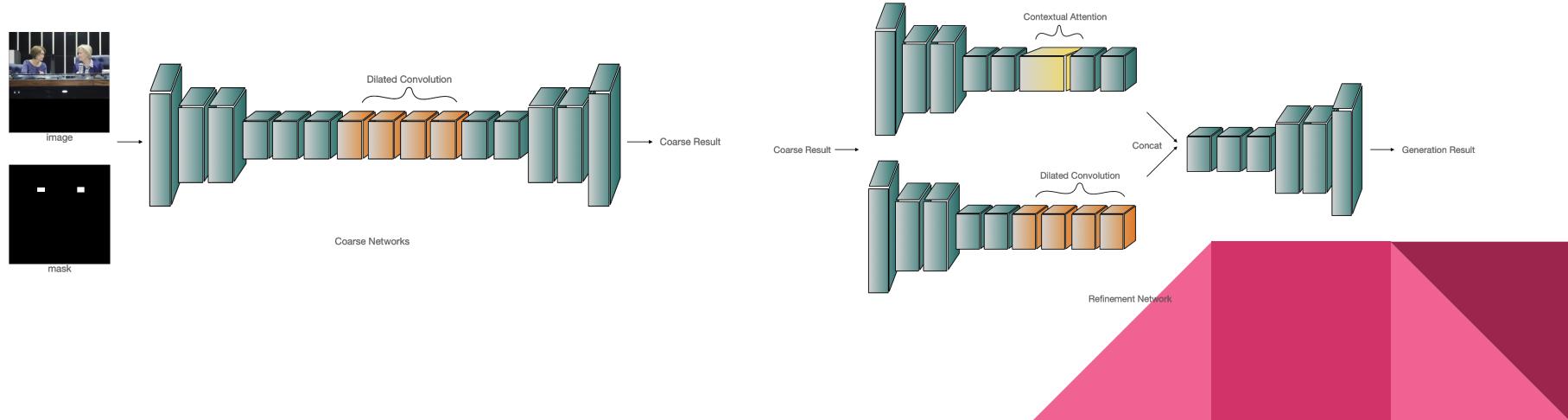
- ❖ EfficientDet-D0
 - Convolutional Neural Net (CNN) backbone
 - Bidirectional Feature Pyramid Network (BiFPN)
- ❖ Transfer learning from COCO weights
- ❖ Comparable mAP to YoloV3 (33.6). another object detection net



GAN

❖ Generator

- Learns from input data and generates reconstructed faces based on acquired knowledge
- Modeled for image inpainting
- Gated convolution



GAN

❖ Discriminator

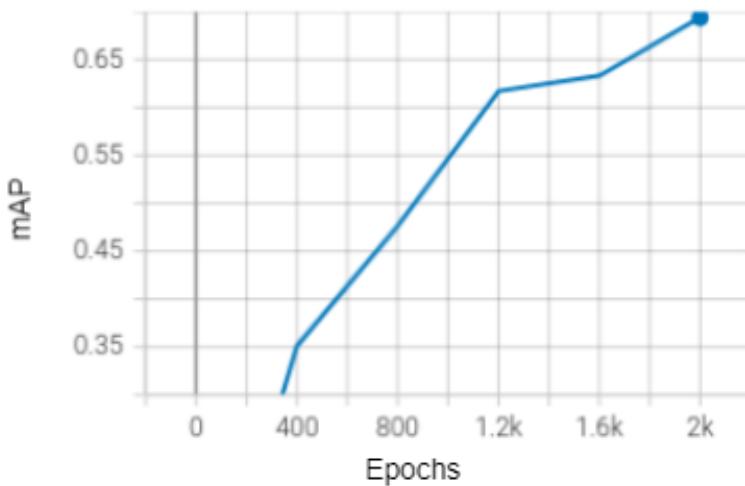
- Attempts to identify whether an image is real or generated
- Based on SN-PatchGAN
 - Markovian patches
 - Extract features patch by patch
 - Work as global discriminator
 - Spectral normalization
 - Stabilize the training of discriminator

Results

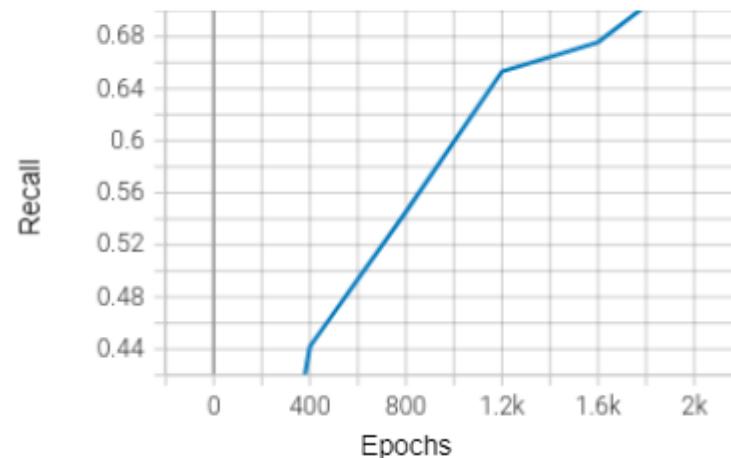
Mask Detection

- ❖ Quantitative analysis: mean Average Precision (mAP) and Recall
- ❖ 2000 epochs

DetectionBoxes_Precision/mAP
tag: DetectionBoxes_Precision/mAP



DetectionBoxes_Recall/AR@10
tag: DetectionBoxes_Recall/AR@10



GAN

- ❖ Loss: Total loss of GAN and L1 loss and Hinge loss of GAN vs. Epochs

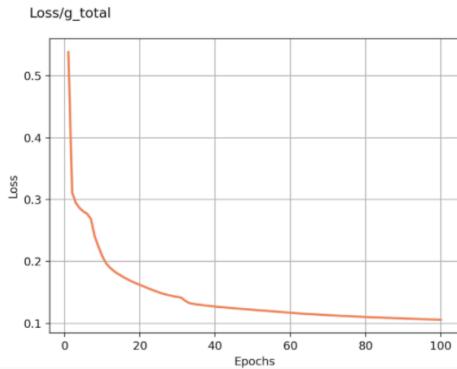


Figure 13. Total loss of GAN vs. Epochs

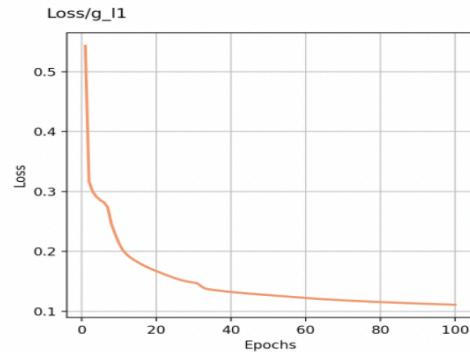


Figure 14. L1 loss of GAN vs. Epochs

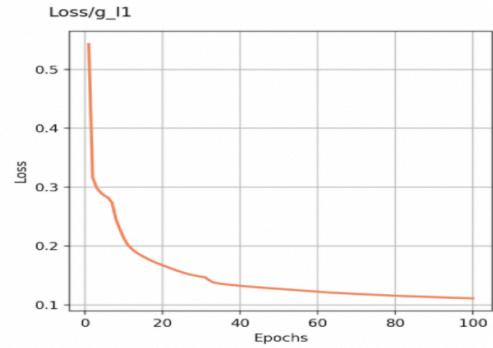
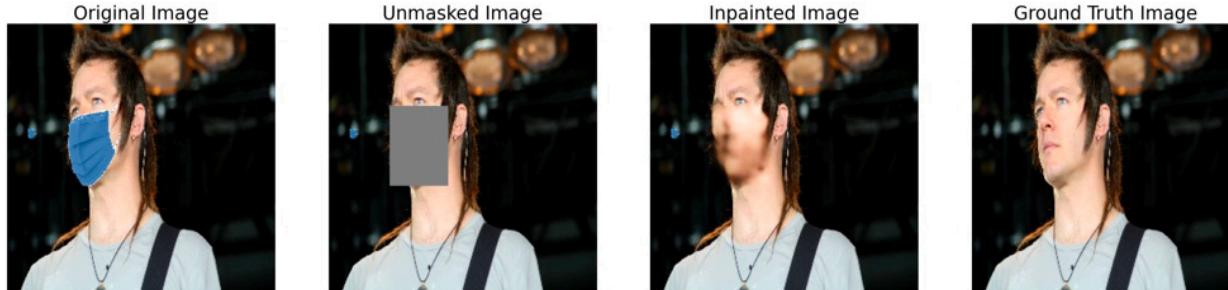


Figure 15. Hinge loss of GAN vs. Epochs

GAN

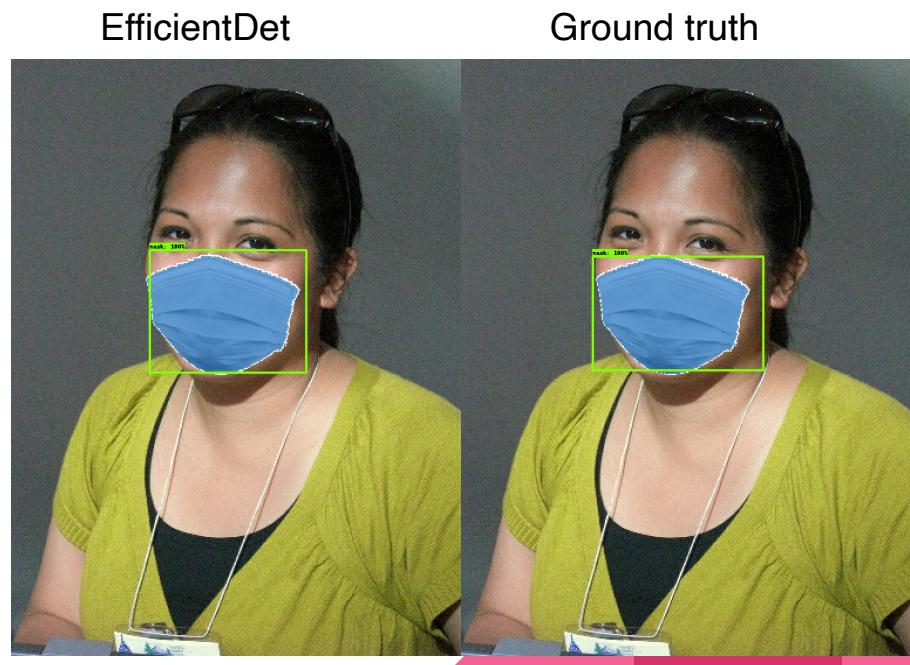
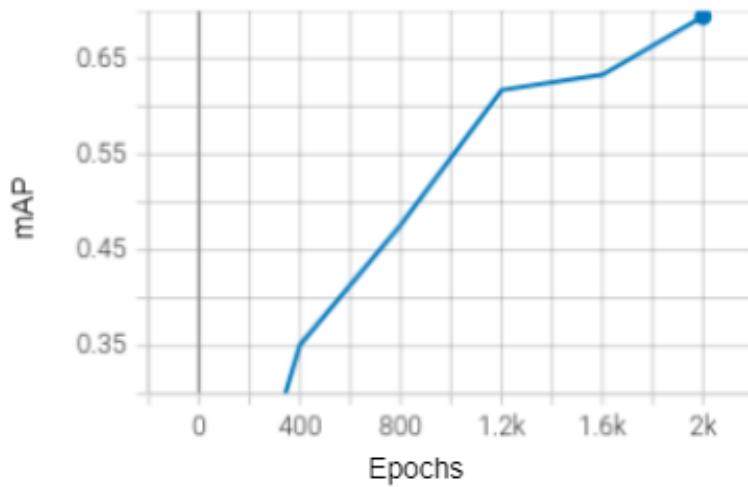
- ❖ Input for testing: masked images + location of masks
 - 402 masked images
 - Coordinates of every mask in a CSV file
- ❖ Trained for 300 epoches:



Conclusion

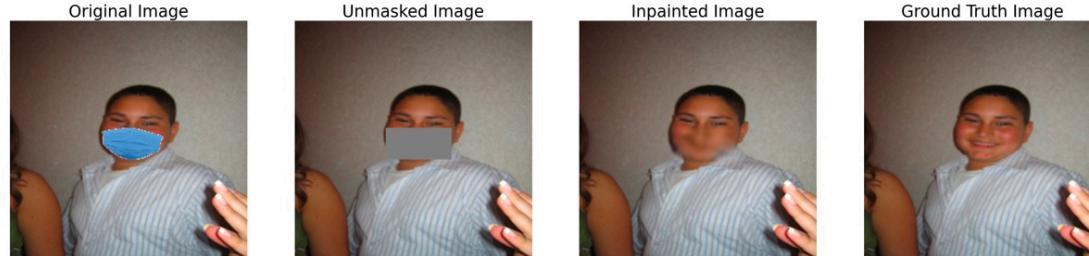
Mask Detection

DetectionBoxes_Precision/mAP
tag: DetectionBoxes_Precision/mAP



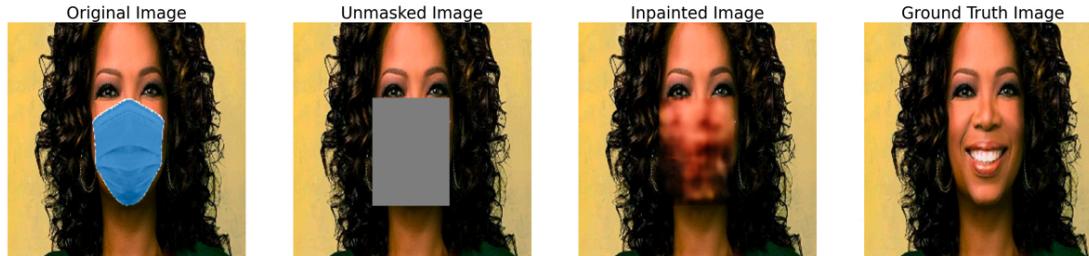
Evaluation of Results

- ❖ Calculated L1 and L2 loss between original unmasked images and reconstructed faces after 300 epochs
 - Average L1 loss: 6169 & Average L2 loss: 1345
- ❖ Positive example



- ❖ Negative example

Machine may think she has a beard!!!



Future Work and Sources of Error

- ❖ Time and hardware limitations
 - 300 training epochs in 36 hours
- ❖ Train with larger (10k+) and more diverse dataset
 - Considering race, gender, age
- ❖ Improve efficiency
 - 3-day training
- ❖ Live video demonstration application
- ❖ Dataset errors
 - Mask channel fails to cover entire surgical mask
 - Surgical mask cover upper facial features

Questions?