Project Summary

Face mask detection

I. Introduction

This study develops a deep-learning model to assess mask-wearing during the COVID-19 pandemic, focusing on proper mask usage as per World Health Organization guidelines. Our model analyzes over 50,000 images, classifying them based on correct mask-wearing and type (medical or non-medical). It aims to create a user-friendly tool for monitoring mask compliance in public areas, essential for public health during ongoing or future outbreaks.

II. Literature Review

The document "An Improved Target Detection Algorithm Based on EfficientNet".[1] by Tao Wu et al. is closely related to our project because it focuses on using EfficientNet to enhance target detection. Here's a simplified summary:

1. **EfficientNet Optimization**: The study improves target detection by modifying EfficientNet. They make the network more efficient, which helps in better feature extraction and faster processing.
2. **Feature Pyramid Networks (FPN)**: Just like your project, the study uses FPN for extracting features at multiple scales. This is especially useful for detecting small details, similar to your project's goal of identifying fine details in masks.
3. **Network Structure Improvements**: The researchers made changes to the YOLOv3 network to better detect small targets. This is relevant to your project's aim of recognizing high-resolution details.

Overall, both our project and this study emphasize the need for high-resolution detail recognition and efficient performance. The success of EfficientNet in the study for detecting small objects supports our decision to use it for differentiating mask types and fits. The use of FPN in both projects shows a common focus on extracting features at multiple scales for detailed recognition.

III. Models and Methods

Our project is dedicated to classifying various types of face masks, with a keen focus on discerning both prominent and nuanced features. This involves accurately identifying the type of mask as well as assessing the finer details of its fit. To achieve this, we have meticulously chosen our models and methodologies.

1. **Model Selection and Task Requirements**: In selecting a model for our task, it was imperative to find one adept at handling multiple categories and capable of recognizing high-resolution details. This need stems from our goal to differentiate between a wide range of mask types and the specifics of their fit on the wearer. After considering several options, we identified key criteria that our chosen model must meet: the ability to process complex image data and discern intricate details efficiently.

2. **Incorporating Feature Pyramid Networks (FPN)**: Our approach integrates Feature Pyramid Networks (FPN) to enhance our model's capability to capture details across various scales. This is particularly beneficial for our task, as it allows the model to identify not just the larger, more visible features of the masks (such as the type and design) but also the smaller, critical aspects (like the precise fit on the face). FPN' s multi-scale feature detection is instrumental in achieving the level of detail required for accurate classification.

3. **Choosing EfficientNet for Optimal Performance**: EfficientNet.[3] emerged as our model of choice. This decision was driven by EfficientNet's exemplary balance of accuracy and computational efficiency – a crucial factor in processing images that encompass both complex patterns and subtle features. EfficientNet excels at pinpointing detailed features, making it exceptionally suitable for our multi-class mask classification task. Notably, we opted not to use ImageNet's pre-training initially, given the rich diversity and high accuracy of our dataset. This choice ensured that our model was tailored specifically to the unique characteristics of our data.

IV. Experimental Setup

Dataset Overview: Our project utilizes a dataset comprising over 50,000 images, focusing on different types of face masks. The dataset includes categories like medical and standard masks, as well as classifications based on how well the mask is worn: correctly worn, mouth exposed, nose exposed, and both mouth and nose exposed.

Model Architecture.[2] and Implementation: We've developed a custom model using EfficientNet and Feature Pyramid Networks (FPN).[4] for our task. Our model, named `Efficient_Yolov3`, is structured to classify both the type of mask being worn and the correctness of its fit. It employs EfficientNet as the backbone for feature extraction and utilizes FPN for detailed multi-scale analysis. The model includes specif ic classifiers for mask-wearing, mask mutex (mutual exclusion), and mask type, each designed to handle different aspects of mask classification.

Hyperparameter Tuning and Loss Functions: We carefully adjusted our model's settings (hyperparameters) to best classify different face masks and how they're worn. We used two loss functions, CrossEntropyLoss and NLLLoss, because they work well for tasks that involve classifying into multiple categories, like our mask types and fit conditions.

Evaluation Strategy: To test our model, we used separate sets of data for training and testing. This helped us accurately judge the model's performance. The model's success was measured by how well it could identify different ways masks are worn and the types of masks.

V. Results

To better chase our model' s performance, we have made the test part of the model to record every single image test' s prediction and its actual label into CSV files to support the parameter adjustment. This result table compiles the most representative outcomes, illustrating how our team analyzes performance to refine model parameters.

Initial Parameter

parameter							accuracy	
epoch	w1	w2	batchsize	lr	loss function	optimizer	mask_wearing	mask_type
5	0.5	0.5	16	0.001	CrossEntropyLoss	Adam	34.00%	78.00%
5	0.5	0.5	16	0.0001	CrossEntropyLoss	Adam	98.59%	99.89%
5	0.7	0.3	16	0.0001	CrossEntropyLoss	Adam	98.54%	99.82%
5	0.7	0.3	32	0.0001	CrossEntropyLoss	Adam	98.63%	99.86%
5	0.7	0.3	20	0.0001	CrossEntropyLoss	Adam	98.63%	99.92%
5	0.7	0.3	16	0.0002	CrossEntropyLoss	Adam	98.20%	99.60%
5	0.7	0.3	16	0.0003	CrossEntropyLoss	Adam	97.87%	99.65%
5	0.7	0.3	16	0.00005	CrossEntropyLoss	Adam	98.38%	99.83%
5	0.7	0.3	16	0.0001	NLLLoss	Adam	31.33%	73.59%

Initially, our model had low accuracy with the first set of parameters (shown in green). Then, we reduced the learning rate by ten times. This change greatly improved our model's accuracy: it went over 98% for identifying how masks are worn and 99% for identifying the type of mask (shown in yellow).

Batch Size and Weights: We tried different weights and batch sizes in our experiments. The changes made only a small difference, about 0.1%, in accuracy. However, by carefully adjusting both these factors, we did manage to slightly improve the model's accuracy.

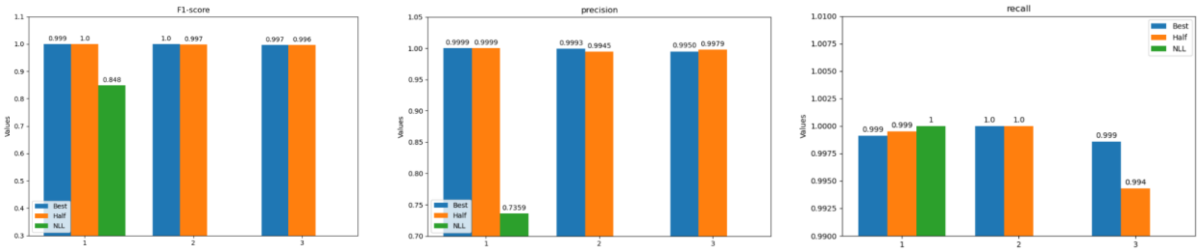
Learning Rate: We also experimented with various learning rates. We found that 0.0001 is the best learning rate for our model. Any increase or decrease from this rate resulted in lower accuracy.

Loss Function: We tested different loss functions and found that NLL (Negative Log Likelihood) Loss worked with our data, but it could only recognize one class at a time, which was not our goal. This limitation made NLL less accurate than CrossEntropyLoss in our tests.

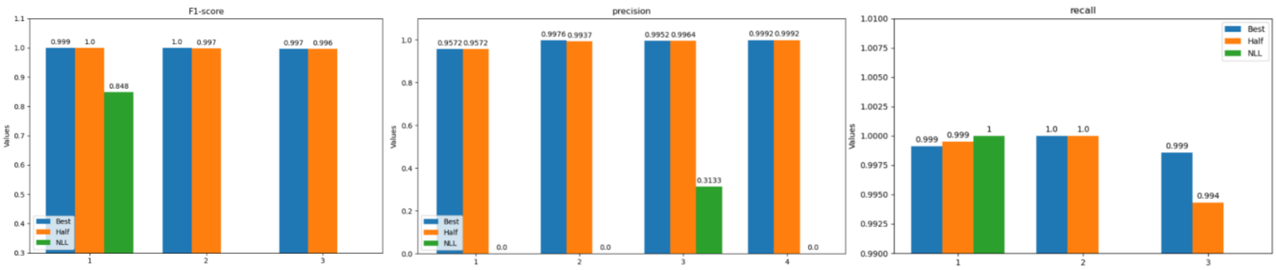
Accuracy

Compared with the public paper model' s 90% accuracy we have referred to, our adjusted parameter and model achieves a much higher accuracy that up to 98% with wear type and 99% with mask type. below are the comparison between our team' s balanced weight, best parameter setting and NLL loss function.

Mask Wear



Wear Type



VI. Conclusions

Here' s an analysis along with future recommendations:

Strengths

High Accuracy: The model uses EfficientNet and FPN for precise identification in complex scenes, accurately classifying different mask types and fits.
Adaptability: It adapts well to various races and mask types, thanks to a diverse dataset, which is important for real-world use.
Advanced Feature Recognition: This can detect subtle differences in mask-wearing, important for detailed classification.

Weaknesses

High Hardware Requirements: Needs high-end GPUs, limiting its use in environments with less computational power.
Complexity and Overfitting Risk: The model's complexity could lead to overfitting if the training data isn' t diverse or large enough.

Key Limitations

Computational Efficiency: Processing speed and resource use are concerns, especially for real-time or low-power hardware applications.

Generalizability: It’s unclear if the model can adapt to very different datasets or real-world scenarios without substantial retraining.

Future Work

These improvements could make the model more versatile, efficient, and useful in various situations:

Parameter Optimization: Adjusting model parameters to maintain accuracy while reducing complexity.

Enhancement of ImageNet Training: Improving this aspect could speed up learning and efficiency.

Data Augmentation & Diversity: Expanding the dataset to enhance robustness and generalizability.

Hardware Optimization: Making the model work on less advanced hardware for wider use.

Exploring Lightweight Models: investigating simpler models or compression techniques for a balance between accuracy and efficiency.

Reference

1. Tao Wu1, Hongjin Zhu2, Honghui Fan2 and Hongyan Zhou1. “An improved target detection algorithm based on EfficientNet.” Journal of Physics: Conference Series, Volume 1983, The Fourth International Conference on Mechanical, Electric and Industrial Engineering (MEIE2021), 22-24 May 2021, DOI 10.1088/1742-6596/1983/1/012017.
2. A. Chavda, J. Dsouza, S. Badgujar and A. Damani, "Multi-Stage CNN Architecture for Face Mask Detection," 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2021, pp. 1-8, doi: 10.1109/I2CT51068.2021.9418207.
3. J. Ieamsaard, S. N. Charoensook and S. Yammen, "Deep Learning-based Face Mask Detection Using YoloV5," 2021 9th International Electrical Engineering Congress (iEECON), Pattaya, Thailand, 2021, pp. 428-431, doi: 10.1109/iEECON51072.2021.9440346.
4. Jignesh Chowdary, G., Punn, N.S., Sonbhadra, S.K., Agarwal, S. (2020). Face Mask Detection Using Transfer Learning of InceptionV3. In: Bellatreche, L., Goyal, V., Fujita, H., Mondal, A., Reddy, P.K. (eds) Big Data Analytics. BDA 2020. Lecture Notes in Computer Science(), vol 12581. Springer, Cham.