Detection of Mask User Objects Using Deep Learning Algorithms with the YOLOv5 Method

COMPUTER VISION

Jerry Cahyo Setiawan

PROBLEM **STATEMENT**

In the era of the COVID-19 pandemic, using masks has become an important step in minimizing the spread of the virus. Until the pandemic ends, the use of masks will remain among the people because they feel comfortable wearing masks. Implementing manual controls to ensure mask use across a broad population is highly inefficient and requires a lot of human resources. Therefore, automatic detection of mask use through computer vision and deep learning is an effective and efficient solution.

TECHNICAL CHALLENGES

Point of View

The use of masks can vary depending on the user's facial position and point of view.

Lighting

Different light conditions can affect image quality and make mask detection difficult.

Scale

Mask use may vary depending on the size and scale of the user's face.

Dataset

To train a good mask detection model, it requires a dataset that covers enough mask usage variations.

Computation

Performing object detection using deep learning requires intensive computing and quite large memory.

RELATED WORKS

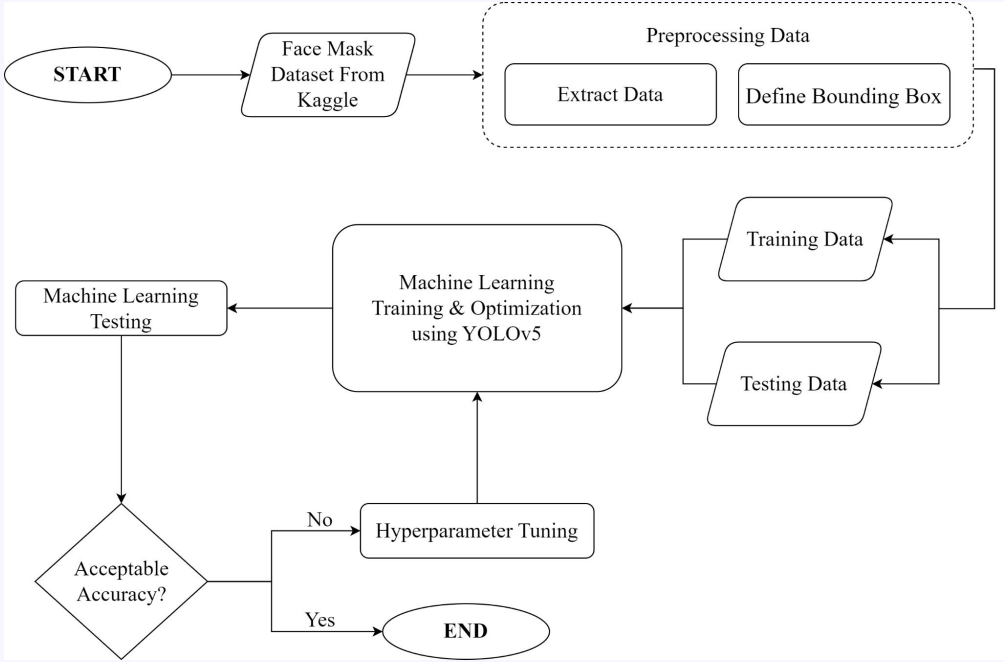
Related research as one of the backgrounds for conducting this research includes:

Face mask detection in COVID-19: a strategic review	This research discusses what strategies are quite effective in handling the detection of mask users during the COVID-19 pandemic. This research provides a solution in the form of suggestions using deep learning methods to overcome this problem.
SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2	This research uses a deep learning method with the MobileNetV2 algorithm to create a mask user detection model. The model evaluation results from this research has F1-Score reached 0.93.
Face Mask Recognition System with YOLOV5 Based on Image Recognition	This research aims to detect mask users at the world center in China for entrance automation. The dataset used comes from (https://github.com/AIZOOTech/FaceMaskDetection). The accuracy value obtained 0.97.

What differentiates this research from similar research is that the dataset used comes from Kaggle and has three classes. That are a dataset of people using masks, using masks unsuitable, and not using masks.

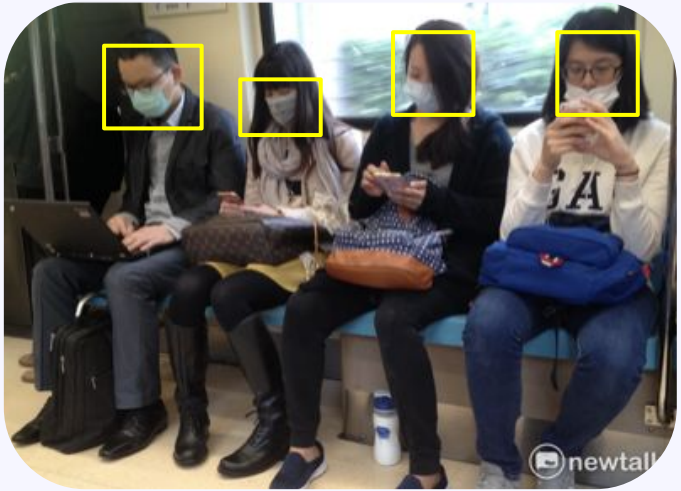
APPROACH AND RESULTS

Workflow



1. The workflow carried out in this research starts from collecting data from Kaggle.
2. Preprocessing data by extracting and determining bounding boxes.
3. Split data into training data and test data.
4. Data training using YOLOv5.
5. Model evaluation.
6. Verification model.

Define Bounding Box



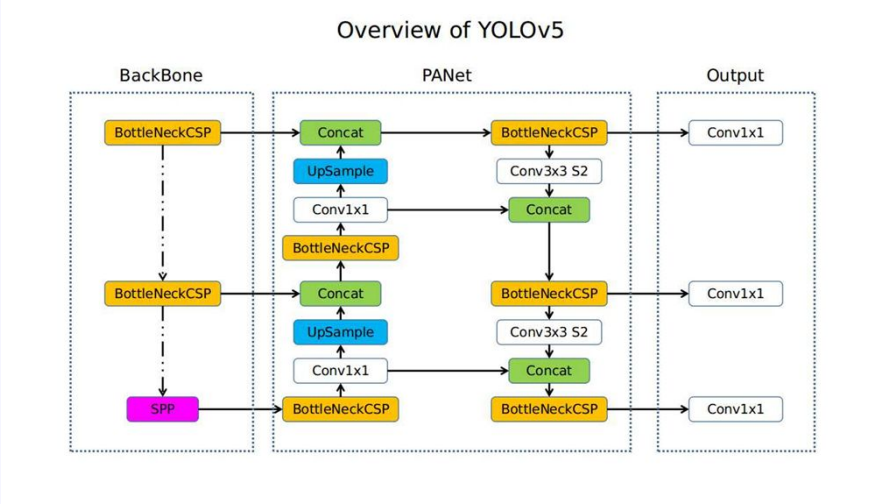
Picture

There are two classes in this image, namely with mask labeled number 2 and incorrect mask labeled number 1.

Bounding Box Result

```
2 0.21625 0.19137931034482758 0.0925 0.09310344827586207
2 0.415 0.20862068965517241 0.06 0.09310344827586207
2 0.625 0.17758620689655172 0.06 0.1206896551724138
1 0.87375 0.16551724137931034 0.0825 0.11724137931034483
```

Modeling YOLOv5



The YOLOv5 architecture is an object detection model developed gleaned from the "You Only Look Once" (YOLO) approach. This model uses a single-shot detection approach, which means that objects in the image can be detecting in one observation process.

In this research, the parameters used are Batch Size of 16, image size as input of 640, and epoch of 50.

Testing Result

Based on the test data and test results in the image above, the model can correctly classify objects that use masks and those that do not use masks.

Evaluation F1-Score

$$precision = \frac{TP}{TP + FP}$$

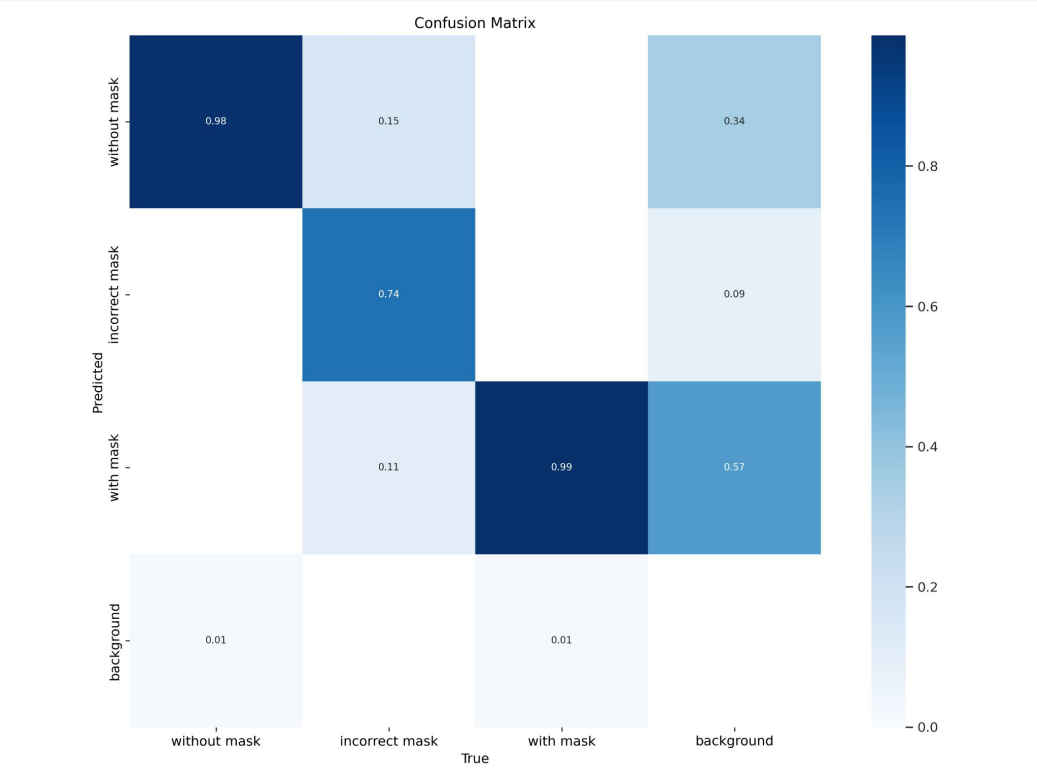
$$recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2\ x\ recall\ x\ precision}{recall + precision}$$

F1-Score is one of the evaluation metrics commonly used in classification tasks. This metric combines precision and recall to provide a more holistic picture of the model's performance in predicting target classes.

Based on the evaluation of the F1-Score calculation. Where the precision value obtained was 0.938 with the recall was 0.926. So the F1-Score of this model reaches **0.932**.

Evaluation Confusion Matrix



The picture show that the results of the Confusion Matrix visualization. The model can detected faces using masks very accurately, reaching 0.99 and faces without masks at 0.98. However, the model is still not good enough at predicting people who wear masks correctly.

BROADER IMPACT

Impact of object detection using masks using YOLOv5:

- Increase public awareness about the use of masks for those who are sick
- Effective in enforcing mask rules
- Reduces the risk of spreading viruses
- Obtain relevant data for compliance analysis
- Protection of individual privacy needs to be taken into account
- Not a single solution, other factors remain important
- Clear regulations and policies are needed
- Transparency and public participation are important
- Positive contribution in handling the COVID-19 pandemic and similar

Thank You!