ASTEEC Conference Proceeding: Computer Science
3rd International Conference on Information Science and Technology Innovation (ICoSTEC)
July 27, 2024, Yogyakarta, Indonesia

Worker Face Detection Using a Digital Image-Based YOLO Model

Jazmi Hadi Matondang[1], Rika Rosnelly[2], Roslina Roslina[3]

Magister of Computer Science, Potensi Utama University
JL. KL. Yos Sudarso Km. 6.5 No. 3-A, Medan, Indonesia
[1]jazmimatondang@gmail.com
[2]rikarosnelly@gmail.com , rika@potensi-utama.ac.id
[3]roslinanich@gmail.com

Abstract— **The large number of workers requires a long time for administrative officers to manually check worker attendance so that work efficiency cannot be achieved. It takes quite a long time to manually check worker attendance. Attendance reports are sent to office social media, including the names of each worker and so on. This is considered inefficient and sufficient as a basis for providing a solution by detecting worker faces using the YOLO (You Only Look Once) method, so that the process of checking worker attendance can be more efficient. The test results with low pixels and training of 100 epochs, namely 224x192 pixels, obtained an accuracy of 86%, while the best test results were using dimensions of 1088x640 pixels on worker photos as test data with original photo dimensions of 1080x1920 pixels in the YOLO model successfully detecting faces with 100% accuracy. So, it can be concluded that the higher the pixel value, the better the accuracy tends to be. However, in this case it also has a pixel limit that is recognized by the model.**

Keywords— **Face Detection, YOLO, Deep Learning**

I. INTRODUCTION

Facial recognition is a term used to identify a human face by the characteristics or prominent features of the face. Facial recognition technology is the fastest growing today. The goal of the facial recognition system with the traditional method is to develop an automatic system that can record attendance using technology that can identify the person present. The main goal is to make the attendance management and marking system fully automatic, simple and easy to use [1]. The impractical method of monitoring employee attendance manually is a problem for most of the institutional authorities. The process of manual signing on paper is tedious and unsafe. There should be an effective system in the workplace to monitor employee attendance [2]. The concept of presence applies to various activities and functions of daily life, often involving processes required for authentication. However, this concept is time-consuming and requires a lot of effort to automate this process with various technologies, although there are few attempts to use biometrics in attendance, such as facial recognition [3]. Research on human face recognition is one of the developments in computer vision. One application of a human face recognition system is an attendance system, which detects a person's face, identifies it as an identity, and stores it in a face database. Matching the face photo data taken by the camera with the face photo stored in the face database produces identification [4]. Object detection is one of the key and challenging problems in computer vision. Over the past decade, with the rapid evolution of deep learning, researchers have extensively experimented and contributed to improving the performance of object detection and related tasks such as object classification, localization, and segmentation using underlying deep models [5]. Previously, image processing for facial recognition using the YOLO algorithm has been carried out, including the Attendance Management System with Facial Recognition and GPS Features Using YOLO on the Android Platform with an average accuracy of 93.26%, then research on the implementation of YOLO for detecting victims of natural disasters obtained an accuracy of 92.9% in image detection [6]. Based on the explanation, it is necessary to develop a method that can be used in facial recognition with a fairly good level of accuracy, namely the YOLO (You Only Look Once) method. Human facial recognition can be utilized in the attendance recording system for each employee or worker in a company. We conclude that automatic detection of worker attendance through the digital image processing stage can be done so that work efficiency can be achieved. In achieving efficiency in worker facial recognition time, workers take pictures and the machine will carry out the facial recognition process according to the image data that has been labeled and trained previously.

II. RESEARCH METHODS

The development of the model used is using the YOLO (You Only Look Once) model. In the YOLO (You Only Look Once) model, a unified model is used to detect objects. In this model, a single convolutional network predicts several bounding boxes, or bounding boxes, while predicting the possible classes contained in those boxes. First, the YOLO system divides the input image into an S × S grid. One of the grid cells is responsible for detecting an object if its center falls within it. The bounding boxes and confidence scores of each bounding box are predicted by each grid cell [7]. Testing by performing various stages of the model to be tested using photo or image resolutions in the range of 224 pixels to 1088 pixels with the aim of producing the best accuracy.

In this case, the training data is 30 photos or images, the validation data is 15 photos and the test data is 1 image. The actual photos or test data have an average dimension of 1080x1920 pixels using the same cellphone camera. Training

is carried out with 100 Epochs to identify the "ilham" face class as the worker object to be detected. Figure. 1 shows the flowchart of Object Detection Using the YOLO Model.

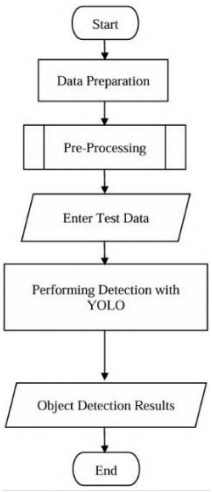

Fig. 1 The flowchart for Object Detection Using the YOLO Model

III. RESULTS AND DISCUSSIONS

A. *Data Preparation*

The research began by collecting image data or photos of workers as many as 30 photos for training data, 15 photos for validation data and 15 photos for test data. Where in the photo has 1 class of object name, namely 0 = "inspiration" which is stored in the training and validation folders.

TABLE I
DETECTION OBJECT CLASS

Class	Photo	Name of Worker/Label
0		ilham

B. *Pre-processing*

Pre-processing is done on the image data or photos of workers which are then labeled at the coordinate points of the object to the image, this is done so that the machine can study in depth the object to be predicted. The steps taken are as follows:

1. Create Object Labels
   Performing object labeling or annotation on the body part of the face for all worker photos in the training and validation folders. The results of the annotation or object labeling are stored in the train and validate data folders in the form of files containing the coordinate points of the facial objects in each photo. Figure 2 shows the process of creating object labels on the face.

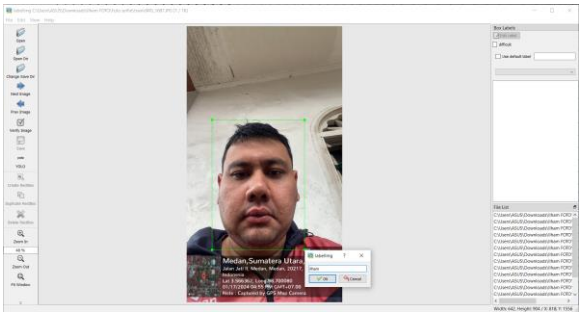

Fig. 2 Creating Object Labels on Faces

2. Create a YAML Format Data File
   Before conducting training, a YAML (Ain't Markup Language) format data file is required, which is the data serialization format required by the YOLO (You Only Look Once) model in conducting training.
3. Conduct Training
   The training was carried out by training 100 epochs with the YOLO model using ultralytics and the python programming language on Google Collab against train and valid data.
4. Training Results
   The training results of the YOLO (You Only Look Once) model can be seen from the precision-recall curve to explain the precision plot against the different threshold recall of the classifier confidence score. The precision-recall curve is a useful metric for evaluating classifiers, especially when the classes are imbalanced. The curve can be seen in Figure 3.

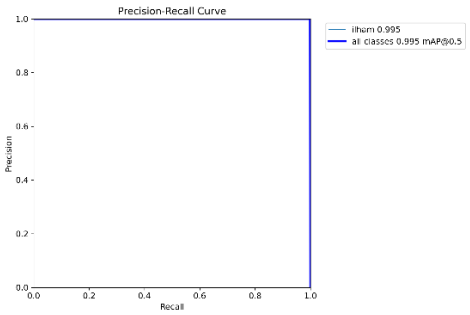

Fig. 3 Precision Recall Curve

In Figure 3. The curve is created by varying the confidence threshold, calculating the precision and recall for each threshold. The curve shows the mAP (mean Average Precision) score or average precision at the IoU (Intersection over Union) threshold of 0.5 which is a common evaluation metric for object detection. While the training results get a value of 0.995. So, the training results are considered very good because they

have passed the threshold value far to almost get a score of 1.0 or 100%.

5. Validation Results
   The results of the training process on the train data are validated using YOLO in the valid folder so that it produces a fairly good prediction in recognizing facial objects according to class. The dominant validation results get an accuracy of 1.0 or 100% which can be seen in Figure 4.

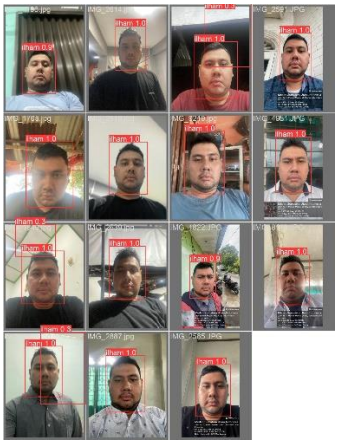

Fig. 4 Validation Results

C. Input Test Data

The test data is in the form of portrait mode photos with dimensions of 1080x1920 pixels or in other words, 1080 pixels wide and 1920 pixels high, where in the photo there are workers who have previously been labeled according to the coordinates of the object. The sample test data photos from the 15 photos can be seen in Figure 5.

Fig. 5 Sample Photo Image or Test Data

D. Prediction Using YOLO

Prediction testing uses the YOLO (You Only Look Once) model by adjusting the input pixel dimension size of the test photo data or portrait mode images with the previous dimension size, namely 1080x1920 pixels, where the photo contains the face of a worker according to the class that has been previously trained.

The photo or image is resized or resizing the image on the photo, this is done to avoid differences in scale between horizontal and vertical dimensions in the input data (input) of the image or what is commonly called anisotropy so that the input image can approach a 1:1 scale precisely to the image dimensions with the aim of producing the best accuracy and understood by the YOLO model. The prediction test sample is carried out with several pixel dimensions as follows:

1. Detection by testing Dimension 224x192 Pixels
   This test produced a predicted score of 0.86 which can be seen in Figure 6.
   .

Fig. 6 Test Results with Dimensions 224x192 Pixels

2. Detection by testing Dimension 448x256 Pixels
   This test produces a predicted score of 0.97 or 97% which can be seen in Figure 7.

Fig. 7 Test Results with Dimensions 448x256 Pixels

3. Detection by testing Dimension 640x384 Pixels
   This test produces a predicted score of 0.98 or 98% which can be seen in Figure 8.

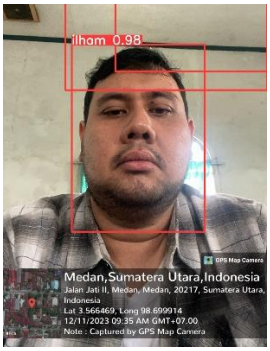

Fig. 8 Test Results with Dimensions 640x384 Pixels

4. Detection by testing Dimension 1088x640 Pixels
   This test produces a predicted score of 1.00 or 100% which can be seen in Figure 9.

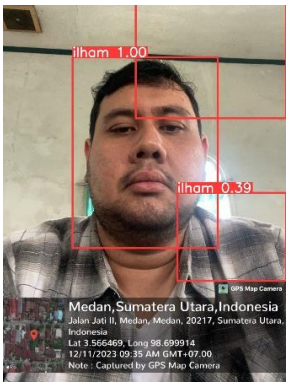

Fig. 9 Test Results with Dimensions 1088x640 Pixels

E. Object Detection Results

After conducting four tests using the YOLO (You Only Look Once) model on several different pixel dimensions, the facial prediction results can be seen in Table II.

TABLE II
FACIAL PREDICTION RESULTS

Class	224x192 Pixel	448x256 Pixel	640x384 Pixel	1088x640 Pixel
Ilham	0,86	0,97	0,98	1,00

IV. CONCLUSIONS

Based on the results of training and testing conducted in the research on digital image-based worker attendance detection, the following conclusions can be drawn:

1. The training results using 30 images or photos with the name class whose face will be detected obtained an accuracy of 99.5%.
2. The best test results using dimensions of 1088x640 pixels with dimensions of 1080x1920 pixels on the YOLO model obtained the highest accuracy of 100%.
3. The validation results for the images or photos whose faces will be detected get an accuracy of the "inspiration" class in the range of 90% to 100%.
4. The higher the pixels used during testing, the better the accuracy results tend to be.
5. The maximum pixel limit that can be used during testing that can be recognized by the YOLO model is up to 1088x640 pixels with photo or image test data having dimensions of 1080x1920 pixels.

REFERENCES

[1] Jenif D.S., W.S., Jothi, S., Chandrasekar, A., "Automated Attendance Marking and Management System by Facial Recognition Using Histogram," *An International Conference on Advanced Computing & Communication Systems (ICACCS)*, 2019.

[2] Koppikar, U., Hiremath, S., Shiralkar, A., "IoT based Smart Attendance Monitoring System using RFID," *An International Conference on Advances in Information Technology*, 2019.

[3] Helmi, R.A.A., Salsabil, S., Jamal, A., "Face Recognition Automatic Class Attendance System (FRACAS)," *An IEEE International Conference on Automatic Control and Intelligent Systems (12CACIS), Selangor:Malaysia*, 2019.

[4] Winarno, E., Amin, I.H.A., Februariyanti, H., Adi, P.W., Hadikurniawati, W., Anwar, M.T., "Attendance System Based on Face Recognition System Using CNN-PCA Method and Real-time Camera," An International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2019.

[5] Diwan, T., Anirudh, G., Themburne, J.V., "Object Detection Using YOLO: challenges, architectural, successors, datasets and applications," *An International Seminar on Research og Information Technology and Intelligent Systems (ISRITI)*, 2019.

[6] Salamah, I., Said, R.A., Soim, S., "Perancangan Alat Identifikasi Wajah Dengan Algoritma You Only Look Once (YOLO) untuk Presensi Mahasiswa," *Journal of Media Informatika Budidarma, Palembang: Indonesia*, 2022.

[7] Khairunnas, Yuniarno, E.M., Zaini, A., "Pembuatan Modul Deteksi Objek Manusia Menggunakan Metode YOLO untuk Mobile Robot," *Journal of Teknik Departemen Teknik Elektro Institut Teknologi Sepuluh November (ITS), Jawa Timur: Indonesia*, 2021.