

ABSTRAK

FER (*Facial Expression Recognition*) merupakan bidang dalam pengenalan pola dan pengolahan citra yang berfokus untuk mengenali dan menganalisis ekspresi wajah manusia. Dengan kemajuan kecerdasan buatan dan pembelajaran mesin, banyak algoritma FER telah digunakan untuk memproses informasi dalam gambar dan video, memungkinkannya untuk mengenali berbagai ekspresi wajah dengan tingkat akurasi yang jauh lebih tinggi dibandingkan dengan pendekatan tradisional. Beberapa penelitian sebelumnya telah menggunakan algoritma *Viola-Jones* untuk mendeteksi wajah dan CNN dengan arsitektur VGG-16 untuk mengklasifikasi ekspresi wajahnya. Namun, masih terdapat kekurangan dimana penggunaan *Viola-Jones* hanya dapat mendeteksi satu wajah dalam satu frame dengan kondisi wajah harus frontal.

Penelitian ini bertujuan untuk mengimplementasikan YOLOv8 untuk mendeteksi wajah dan CNN dengan arsitektur VGG-16 untuk mengklasifikasi ekspresi wajah dasar. Selain itu, penelitian ini juga menganalisis performa YOLOv8 dalam mendeteksi wajah menggunakan dataset WIDER-FACE serta melakukan modifikasi pada arsitektur VGG-16 dengan melakukan penambahan normalisasi *batch* dan *dropout*, serta pengurangan jumlah *dense unit* pada *fully connected layer* guna meningkatkan akurasi klasifikasi ekspresi wajah dengan dataset FER-2013.

Hasil penelitian menunjukkan bahwa model terbaik YOLOv8 yang dilatih berhasil memperoleh nilai mAP50 sebesar 64,8%, nilai *Precision* sebesar 80,9%, dan nilai *Recall* sebesar 57,3% pada pengujian menggunakan data testing. Selain itu, model terbaik VGG-16 yang dilatih menggunakan arsitektur yang telah dimodifikasi berhasil memperoleh nilai *Accuracy* sebesar 70,58%, nilai *Precision* sebesar 71,04%, dan nilai *Recall* sebesar 67,76% pada pengujian menggunakan data testing. Secara keseluruhan, model YOLOv8 terbukti dapat mendeteksi beberapa wajah dalam satu frame dengan kondisi non-frontal sekalipun, serta modifikasi arsitektur VGG-16 terbukti lebih unggul bila dibandingkan dengan basis arsitektur VGG-16.

Kata kunci: YOLOv8, CNN, VGG-16, ekspresi wajah, deteksi wajah, klasifikasi

ABSTRACT

FER (Facial Expression Recognition) is a field within pattern recognition and image processing that focuses on recognizing and analyzing human facial expressions. With advancements in artificial intelligence and machine learning, many FER algorithms have been used to process information in images and videos, enabling them to recognize various facial expressions with a much higher degree of accuracy compared to traditional approaches. Some previous studies have used the Viola-Jones algorithm to detect faces and CNN with VGG-16 architecture to classify facial expressions. However, there are still limitations where the use of Viola-Jones can only detect one face in a single frame, and the face must be in a frontal position.

This study aims to implement YOLOv8 for face detection and CNN with VGG-16 architecture for basic facial expression classification. Additionally, this study analyzes the performance of YOLOv8 in detecting faces using the WIDER-FACE dataset and modifies the VGG-16 architecture by adding batch normalization and dropout, as well as reducing the number of dense units in the fully connected layer to improve facial expression classification accuracy using the FER-2013 dataset.

The results show that the best trained YOLOv8 model successfully achieved an mAP50 value of 64.8%, a Precision value of 80.9%, and a Recall value of 57.3% when tested on the testing data. Moreover, the best trained VGG-16 model using the modified architecture successfully achieved an Accuracy of 70.58%, a Precision of 71.04%, and a Recall of 67.76% when tested on the testing data. Overall, the YOLOv8 model proved capable of detecting multiple faces in a single frame, even in non-frontal conditions, while the modified VGG-16 architecture proved superior compared to the base VGG-16 architecture.

Keywords: YOLOv8, CNN, VGG-16, facial expression, face detection, classification