

INTISARI

KLASIFIKASI EKSPRESI WAJAH MENGGUNAKAN *CONVOLUTIONAL NEURAL NETWORK* DALAM KEADAAN *WILD SETTING* PADA *VIRTUAL MEETING*

Oleh
Isnan Firmansyah
19/445567/PA/19391

Ekspresi wajah merepresentasikan perasaan dalam diri manusia serta dapat menjadi sebuah mediator dalam dunia sosial. Dalam rapat virtual, pemahaman terhadap suasana hati dan emosi peserta sangat dibutuhkan untuk menciptakan interaksi dan kerjasama yang baik. Emosi alami akan muncul ketika peserta berbicara secara spontan dengan keadaan lingkungan alami tanpa dikondisikan sebelumnya. Oleh karena itu, dibutuhkan sebuah sistem yang mampu mengetahui emosi manusia secara otomatis dalam keadaan spontan (*wild setting*) pada *virtual meeting* sehingga ekspresi lebih alami. Teknik *machine learning* digunakan untuk melakukan klasifikasi ekspresi wajah pada *virtual meeting*.

Penelitian ini membandingkan performa arsitektur LightCNN dan EfficientNet. Pelatihan dilakukan pada dataset gabungan antara FER-2013, *Extended and Augmented Google FER* dan CK+. Dataset berjumlah 67.362 citra terbagi menjadi *data training* 60.184 buah, *data validation* 3.589 buah, dan *data testing* 3.589 buah. Input model arsitektur EfficientNet divariasikan menjadi 48x48 dan 224x224 *pixels*. Optimasi *learning rate* dilakukan untuk menemukan performa tertinggi dari arsitektur terbaik.

Hasil penelitian menunjukkan bahwa arsitektur terbaik adalah EfficientNet dengan input 48x48 *pixel*. Nilai parameter learning rate paling optimal yaitu 0,0005. performa model dalam klasifikasi ekspresi wajah mencapai akurasi 90,50%, presisi 89,50%, recall 90,69% dan F1-score 90,06%. Implementasi EfficientNet pada video *virtual meeting* untuk mengklasifikasikan ekspresi wajah mendapatkan performa akurasi sebesar 96,18%.

Kata kunci – Deteksi Ekspresi Wajah, Pengenalan Ekspresi Wajah (FER), *Virtual Meeting*, LightCNN, EfficientNet

ABSTRACT

FACIAL EXPRESSION RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK IN A WILD SETTING FOR VIRTUAL MEETING

by

Isnan Firmansyah
19/445567/PA/19391

Facial expressions represent feelings in humans and can be a mediator in the social world. In virtual meetings, understanding the moods and emotions of participants is necessary to create good interaction and cooperation. Natural emotions will arise when participants speak spontaneously in a natural environment without being conditioned beforehand. Therefore, a system is needed that is able to automatically determine human emotions in a spontaneous state (wild setting) at virtual meetings so that expressions are more natural. Machine learning techniques are used to classify facial expressions in virtual meetings.

This research compares the performance of LightCNN and EfficientNet architectures. Training was conducted on a combined dataset between FER-2013, Extended and Augmented Google FER and CK+. The dataset amounted to 67,362 images divided into 60,184 training data, 3,589 validation data, and 3,589 testing data. The input of EfficientNet architecture model is varied into 48x48 and 224x224 pixels. Learning rate optimization is performed to find the highest performance of the best architecture.

The results showed that the best architecture is EfficientNet with 48x48 pixel input. The most optimal learning rate parameter value is 0.0005. model performance in facial expression classification achieved accuracy of 90.50%, precision 89.50%, recall 90.69% and F1-score 90.06%. EfficientNet implementation on *virtual meeting* videos to classify facial expressions gets an accuracy performance of 96.18%.

Keywords - *Facial Emotion Detection, Facial Emotion Recognition, Facial Expression Recognition (FER), Virtual Meeting, LightCNN, EfficientNet*