

# COMPARISON OF CNN AND CNN-LSTM PERFORMANCE IN FACIAL EXPRESSION CLASSIFICATION BASED ON FER2013 DATASET

Written By

**Putu Ananda Adi Savitri, NIM 2215101057**

**Information Engineering Department**

## ABSTRACT

Facial expressions are a crucial form of nonverbal communication in human interaction, but their automatic recognition still faces challenges due to lighting variations and individual differences. This study aims to compare the performance of Deep Learning models between pure Convolutional Neural Network (CNN) and hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) models in classifying seven basic facial expressions (anger, disgust, fear, happiness, sadness, surprise, and neutral) using the FER2013 dataset. The research method used is quantitative experimental. The designed CNN architecture consists of four hierarchical convolution layers for spatial feature extraction, while the CNN-LSTM architecture uses a TimeDistributed approach to process image pixel rows as sequential data. Evaluation was carried out using accuracy, precision, recall, F1-score, and Area Under Curve (AUC) metrics. The test results showed that the pure CNN model provided superior performance with a better accuracy, where in contrast, the CNN-LSTM model produced a much lower accuracy. The low performance of CNN-LSTM is due to the static nature of the FER2013 dataset, which prevents the LSTM layer from obtaining the temporal information necessary for optimal sequential modeling. Implementation in real-time using a camera also confirms that CNN is more stable in detecting expressions with contrasting visual features such as 'Happy' and 'Surprised' compared to CNN-LSTM, which tends to be biased towards the 'Neutral' category. This study concludes that for static single-image-based facial expression classification, a pure CNN architecture is more effective than a hybrid CNN-LSTM approach.

Keywords: *Convolutional Neural Network, Long Short Term Memory, Facial Expression Recognition, FER2013, Deep Learning.*

# PERBANDINGAN KINERJA CNN DAN CNN-LSTM DALAM KLASIFIKASI EKSPRESI WAJAH BERDASARKAN DATASET FER2013

Oleh

**Putu Ananda Adi Savitri, NIM 2215101057**

**Jurusan Teknik Informatika**

## ABSTRAK

Ekspresi wajah merupakan bentuk komunikasi nonverbal yang penting dalam interaksi manusia, tetapi pengenalan otomatisnya masih menghadapi tantangan karena variasi pencahayaan dan perbedaan individu. Studi ini bertujuan untuk membandingkan kinerja model Deep Learning antara Convolutional Neural Network (CNN) murni dan model hybrid Convolutional Neural Network-Long Short-Term Memory (CNN-LSTM) dalam mengklasifikasikan tujuh ekspresi wajah dasar (marah, jijik, takut, bahagia, sedih, terkejut, dan netral) menggunakan dataset FER2013. Metode penelitian yang digunakan adalah eksperimen kuantitatif. Arsitektur CNN yang dirancang terdiri dari empat lapisan konvolusi hierarkis untuk ekstraksi fitur spasial, sedangkan arsitektur CNN-LSTM menggunakan pendekatan TimeDistributed untuk memproses baris piksel gambar sebagai data sekuensial. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, F1-score, dan Area Under Curve (AUC). Hasil pengujian menunjukkan bahwa model CNN murni memberikan kinerja yang lebih unggul dengan akurasi yang lebih baik, sedangkan sebaliknya, model CNN-LSTM menghasilkan akurasi yang jauh lebih rendah. Performa rendah CNN-LSTM disebabkan oleh sifat statis dataset FER2013, yang mencegah lapisan LSTM memperoleh informasi temporal yang diperlukan untuk pemodelan sekuensial optimal. Implementasi secara real-time menggunakan kamera juga menegaskan bahwa CNN lebih stabil dalam mendeteksi ekspresi dengan fitur visual yang kontras seperti 'Senang' dan 'Terkejut' dibandingkan dengan CNN-LSTM, yang cenderung bias terhadap kategori 'Netral'. Studi ini menyimpulkan bahwa untuk klasifikasi ekspresi wajah berbasis gambar tunggal statis, arsitektur CNN murni lebih efektif daripada pendekatan hibrida CNN-LSTM.

Kata Kunci: *Convolutional Neural Network, Long Short Term Memory, Pengenalan Ekspresi Wajah, FER2013, Deep Learning.*