

ABSTRAK

Deteksi emosi berbasis citra wajah telah menjadi area penelitian yang berkembang dengan berbagai aplikasi, termasuk psikologi, neurosains, dan layanan masyarakat, karena dianggap dapat membantu memahami keadaan emosional individu dalam waktu singkat. Berbagai penelitian sebelumnya telah menggunakan metode seperti *Convolutional Neural Network* (CNN), *Support Vector Machine* (SVM), hingga pendekatan *hybrid* seperti *Convolutional Vision Transformer* (CVT) untuk mendeteksi emosi pada citra wajah dengan tingkat akurasi yang cukup tinggi. Namun, beberapa metode yang telah dilakukan memiliki keterbatasan, sebagian besar penelitian masih menggunakan tuning secara manual, tuning manual membutuhkan waktu dan sumber daya yang besar, serta masih adanya resiko *overfitting*, tidak scalable, serta terdapat kemungkinan kecil untuk menemukan solusi yang optimal. Studi ini bertujuan untuk mengimplementasikan model *Vision Transformer* (ViT) yang dioptimalkan menggunakan metode *Hyperparameter Tuning Grid Search* untuk mendeteksi emosi dari citra wajah. Penelitian ini menggunakan dataset AffectNet yang terdiri dari 450.000 citra wajah, mencakup delapan kelas emosi utama. Evaluasi model dilakukan dengan metrik akurasi, presisi, recall, dan F1-score untuk menilai performa model. Penelitian ini diharapkan memberikan kontribusi signifikan dalam mengembangkan metode deteksi emosi berbasis citra wajah dengan tingkat akurasi yang tinggi melalui pendekatan yang lebih efisien dan optimal.

Kata Kunci: deteksi emosi, *Vision Transformer*, citra wajah, hyperparameter tuning, AffectNet.