

ABSTRAK

Liveness detection merupakan komponen krusial untuk mengamankan sistem pengenalan wajah dari serangan presentasi (presentation attack). Meskipun model berbasis *Convolutional Neural Network* (CNN) dan Transformer menunjukkan potensi besar, keduanya masih menghadapi tantangan dalam hal generalisasi serangan dalam mendeteksi berbagai jenis serangan. Penelitian ini mengusulkan dan mengevaluasi sebuah model ensemble stacking yang mengombinasikan Swin Transformer dan EfficientNetV2 untuk meningkatkan kinerja deteksi serangan multi-kelas (*real*, *print attack*, *replay attack*). Metode ini dievaluasi secara kuantitatif menggunakan metrik akurasi, BPCER, dan APCER, serta diimplementasikan pada sebuah aplikasi web untuk validasi praktis. Hasil pengujian menunjukkan bahwa model *ensemble* yang diusulkan mencapai akurasi superior sebesar 98,17%, secara signifikan mengungguli performa model dasar individualnya. Model ini menghasilkan *Bona Fide Presentation Classification Error Rate* (BPCER) hingga 0,0049 dan nilai *Attack Presentation Classification Error Rate* (APCER) mencapai 0,0250. Meskipun implementasi pada aplikasi web *real-time* mengonfirmasi fungsionalitas model, pengujian tersebut juga mengungkap adanya beberapa kasus kegagalan pada skenario serangan dengan skala dan jarak yang ekstrem, yang menyoroti batasan dari model saat ini.

Kata Kunci: *Liveness Detection, Swin Transformer, EfficientNetV2, Ensemble Stacking.*