

## ABSTRAK

Kemajuan teknologi kecerdasan buatan telah memungkinkan penciptaan *deepfake*, sebuah teknik manipulasi video yang dapat menggantikan wajah seseorang dengan wajah orang lain secara meyakinkan. Meskipun memiliki potensi, teknologi ini juga menimbulkan kekhawatiran terkait penyalahgunaan dan penyebaran informasi palsu. Untuk mengatasi masalah tersebut, diperlukan adanya teknologi yang dapat mendeteksi video *deepfake* secara akurat. Pada penelitian ini mengatasi masalah tersebut dengan mengembangkan metode deteksi *deepfake* dengan menggabungkan arsitektur Xception untuk ekstraksi fitur spasial dan *Long Short-Term Memory* (LSTM) untuk analisis pola temporal dalam video.

Penelitian ini menggunakan dataset yang terdiri dari 600 video dari keseluruhan dataset Celeb-DF. Metode yang diusulkan mengintegrasikan Xception untuk mengekstrak fitur spasial dari *frame* wajah video, sementara LSTM digunakan untuk menganalisis pola informasi urutan antar *frame*. Multi-task Cascaded Convolutional Networks (MTCNN) diimplementasikan pada tahap *preprocessing* untuk meningkatkan deteksi dan *alignment* wajah, memungkinkan model untuk fokus pada area yang paling relevan. Pengujian dilakukan dengan menggunakan akurasi, *precision*, *recall*, dan *F1-score* untuk melihat efektivitas dan akurasi sistem deteksi *deepfake*.

Hasil pengujian menunjukkan bahwa model gabungan ini mencapai akurasi 91.67%, dengan *precision* 89.06%, *recall* 95%, dan *F1-score* 91.94%. Konfigurasi optimal diperoleh dengan panjang *sequence* 20 dan *learning rate*  $1e-4$ . Hal ini menunjukkan bahwa dengan menggabungkan Xception dan LSTM, sistem berhasil mengatasi kelemahan dari metode sebelumnya. Penelitian ini membuktikan efektivitas LSTM dalam menganalisis pola video *deepfake*, yang sangat penting dalam mendeteksi anomali seperti transisi wajah yang tidak wajar atau inkonsistensi visual lainnya.

**Kata kunci:** deteksi *deepfake*, Xception, *Long Short-Term Memory*, MTCNN, kecerdasan buatan.

## **ABSTRACT**

*The rapid advancement of artificial intelligence has facilitated the creation of deepfakes, a technique for manipulating videos that convincingly replaces one person's face with another. While this technology offers promising applications, it also raises significant concerns about misuse and the spread of misinformation. To combat these risks, it is essential to develop robust deepfake detection systems. This study addresses the challenge by proposing a method that combines Xception for spatial feature extraction and Long Short-Term Memory (LSTM) networks for temporal pattern analysis.*

*The research uses a dataset comprising 600 videos from the Celeb-DF dataset. The proposed approach leverages Xception to extract spatial features from facial frames, while LSTM models are employed to analyze sequential frame patterns. Additionally, Multi-task Cascaded Convolutional Networks (MTCNN) are implemented during preprocessing to enhance face detection and alignment, ensuring the model focuses on the most relevant regions of interest. The model's performance is evaluated using accuracy, precision, recall, and F1-score metrics.*

*Experimental results demonstrate that the combined model achieves an accuracy of 91.67%, a precision of 89.06%, a recall of 95%, and an F1-score of 91.94%. The optimal configuration was obtained with a sequence length of 20 frames and a learning rate of  $1e-4$ . These findings underscore the effectiveness of combining Xception and LSTM for deepfake detection, addressing the limitations of prior methods. This research highlights LSTM's capability to identify deepfake anomalies, such as unnatural facial transitions and visual inconsistencies, enhancing detection accuracy.*

**Keywords:** *deepfake detection, Xception, Long Short-Term Memory, MTCNN, artificial intelligence.*