

ABSTRAK

Video palsu yang dihasilkan oleh teknologi *deepfake* tersebar secara *online*. Sebagian besar berisi konten berbahaya dan mengakibatkan kerusakan, kecemasan sosial, ancaman politik serta masalah sosial lainnya. Beberapa teknik telah diajukan untuk mencegah penyalahgunaan teknologi *deepfake* dan mekanisme *attention* memungkinkan *Convolutional Neural Network* (CNN) untuk fokus pada bagian kritis pada gambar seperti mulut atau mata. Metode *Dual Attention Forgery Detection Network* (DAFDN) menggunakan dua mekanisme *attention*, yaitu *Spatial Reduction Attention Block* (SRAB) dan *Forgery Feature Attention Module* (FFAM). Metode DAFDN menghasilkan skor *Area Under the Curve* (AUC) sebesar 91,1% pada *dataset Deepfake Detection Challenge* (DFDC) dan 94,5% pada *dataset FaceForensics++* (FF++). Pada penelitian ini, penulis menggunakan metode DAFDN yang disematkan pada *EfficientNet-B4* untuk mendeteksi pemalsuan wajah pada *dataset* dari *Celeb-DF* dan menghasilkan skor AUC sebesar 90%.

Kata kunci: *deepfake*, DAFDN, SRAB, FFAM, *EfficientNet-B4*.

ABSTRACT

Fake videos generated by deepfake technology are spreading online. Most of them contain dangerous content and cause riots, social anxiety, political threats and other social problems. Several techniques have been proposed to prevent misuse of deepfake technology and attention mechanisms allow Convolutional Neural Networks (CNN) to focus on critical parts of an image such as the mouth or eyes. The Dual Attention Forgery Detection Network (DAFDN) method uses two attention mechanisms, namely Spatial Reduction Attention Block (SRAB) and Forgery Feature Attention Module (FFAM). The DAFDN method produces an Area Under the Curve (AUC) score of 91.1% on the Deepfake Detection Challenge (DFDC) dataset and 94.5% on the FaceForensics++ (FF++) dataset. In this research, the authors used the DAFDN method embedded in EfficientNet-B4 to detect facial forgeries in the Celeb-DF dataset and produced an AUC score of 90%.

Keywords: deepfake, DAFDN, SRAB, FFAM, EfficientNet-B4.

