

BAB I

PENDAHULUAN

1.1 Latar Belakang

Deepfake adalah teknologi yang memanfaatkan kecerdasan buatan dan algoritma *deep learning* untuk membuat sebuah konten digital, seperti gambar, video, atau audio. Dengan menggunakan teknologi *deepfake*, penggunaanya dapat mengganti wajah, suara, bahkan gerakan tubuh seseorang dalam sebuah media, menciptakan ilusi bahwa orang yang ada di dalamnya melakukan atau mengatakan hal yang sebenarnya tidak terjadi. Penyalahgunaan terhadap teknologi ini dapat menimbulkan masalah serius, karena kemampuannya untuk menyebarkan informasi palsu dan merusak reputasi.

Banyak hal dapat dibuat dengan teknologi ini, sekarang sering ditemui dalam berbagai kasus yang terjadi. Misalnya, di Indonesia, kasus video pornografi hasil rekayasa yang mirip seorang selebritas ternama (Hardiansyah dan Nistantato, 2024). Selain itu, kasus yang sama terjadi pada luar negeri terhadap selebritas ternama diganti ke wajah aktor pornografi, yang menyebabkan reputasinya memburuk (Badshah, 2024). Bahkan, dalam kampanye politik, beberapa pihak memanfaatkan teknologi ini untuk menarik dukungan orang lain dengan menyebarkan gambar hasil manipulasi (Robins-Early, 2024). Penyalahgunaan teknologi ini menciptakan tantangan besar terkait keaslian informasi yang beredar, terutama di media sosial.

Dalam hal ini, kebutuhan untuk mendeteksi dan mengidentifikasi konten *deepfake* semakin mendesak. Tanpa adanya teknologi yang mampu mendeteksi *deepfake*, orang akan sulit membedakan antara konten yang asli dan palsu. Masyarakat akan dengan mudah menyerap mentah-mentah terkait informasi yang didapatkannya tanpa melakukan verifikasi terlebih dahulu. Hal ini menimbulkan risiko yang besar, terutama ketika informasi palsu tersebut dapat menyebar dengan cepat di media sosial dan platform digital lainnya. Risiko ini yang akan menyesatkan banyak orang.

Untuk mendeteksi konten *deepfake*, teknologi berbasis *deep learning* (DL), seperti *Convolutional Neural Network* (CNN), telah berkembang menjadi alat utama dalam mendeteksi manipulasi visual.

CNN merupakan salah satu jenis algoritma *deep learning* yang dirancang khusus untuk mengolah data visual, baik dalam gambar maupun video. CNN bekerja dengan mendeteksi pola-pola tertentu dalam citra atau video. Algoritma ini bekerja dengan mengenali fitur-fitur spesifik pada gambar atau video, seperti tekstur, warna, perbedaan bayangan, dan perubahan gerakan. Dalam beberapa penelitian seperti Karandikar dkk. (2020), CNN berhasil dikembangkan dengan data *training* yang telah tersedia di situs internet seperti Kaggle untuk dilakukan pengenalan karakteristik spesifik *deepfake* pada video melalui analisis fitur visual. Meskipun demikian, akurasi dari hasil penelitian tersebut masih sangat rendah yakni sekitar 70%, yang menunjukkan adanya tantangan besar dalam mengembangkan model CNN yang lebih akurat untuk mendeteksi *deepfake*.

Berdasarkan permasalahan di atas, solusi yang akan dibuat dalam penelitian ini adalah dengan membangun sebuah aplikasi Android untuk mendeteksi *deepfake* pada video menggunakan algoritma CNN, melalui analisis fitur visual yang dimilikinya. meningkatkan akurasi prediksi yang lebih akurat. Aplikasi ini dirancang untuk memudahkan pengguna dalam mengidentifikasi konten *deepfake* melalui perangkat Android. Sehingga memberikan hasil akhir dengan akurasi yang tinggi terhadap klasifikasi konten video yang diberikan.

1.2 Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan maka perumusan masalahnya adalah:

- 1) Teknologi *deepfake* sering disalahgunakan untuk pembuatan video palsu yang bertujuan negatif, seperti penipuan, manipulasi informasi, serta merusak reputasi individu atau kelompok.
- 2) Masyarakat awam sulit dalam mengidentifikasi keaslian informasi atau konten digital, terutama terkait dengan penyebaran berita palsu dan manipulasi visual menggunakan teknologi *deepfake*.

- 3) Masyarakat awam sulit dalam mendeteksi *deepfake* yang dapat menyebabkan kerugian pada reputasi individu atau penyebaran informasi yang menyesatkan.
- 4) Penelitian sebelumnya yang dilakukan memiliki akurasi 70% yang masih memiliki tantangan besar dalam pengembangan model CNN yang lebih akurat untuk mendeteksi *deepfake*.

1.3 Batasan Masalah

Adapun batasan masalah dalam pembuatan proyek akhir ini adalah:

- 1) Deteksi *deepfake* hanya dilakukan pada konten visual video, tanpa melibatkan manipulasi audio.
- 2) Dataset yang digunakan berupa video yang diambil dari sumber Kaggle
(<https://www.kaggle.com/datasets/reubensuju/celeb-df-v2>).
- 3) Aplikasi ini dibangun dengan berbasis Android.

1.4 Tujuan dan Manfaat

1.4.1 Tujuan

Adapun tujuan dari penelitian ini adalah:

- 1) Merancang dan mengimplementasikan model CNN dalam aplikasi Android untuk mendeteksi video *deepfake*.
- 2) Mengevaluasi faktor-faktor yang mempengaruhi akurasi pada aplikasi video *deepfake* pemodelan CNN.

1.4.2 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah:

- 1) Menghasilkan model CNN yang memiliki akurasi deteksi lebih tinggi dari penelitian yang telah dilakukan sebelumnya.
- 2) Menghasilkan sebuah aplikasi berbasis Android yang mudah diakses oleh masyarakat umum supaya dapat

memudahkan dalam membedakan konten video asli dan *deepfake*.

1.5 Metodologi Penelitian

Metode penelitian yang dipakai dalam pembuatan proyek akhir ini adalah:

- 1) Studi Literatur
Pada tahapan ini akan dilakukan dengan cara membaca dan mempelajari referensi seperti jurnal atau penelitian sebelumnya yang berhubungan dengan proyek akhir.
- 2) Pengumpulan Data
Pada tahapan ini akan dilakukan dengan mencari *dataset* video asli dan hasil *deepfake* dari situs internet seperti Kaggle.
- 3) Perancangan
Pada tahapan ini akan dilakukan perancangan untuk membangun sistem aplikasi deteksi video *deepfake* berbasis Android.
- 4) Implementasi
Pada tahapan ini akan dilakukan implementasi untuk pengembangan sistem aplikasi deteksi video *deepfake* terhadap perangkat Android.
- 5) Pengujian dan Analisis
Pada tahapan ini akan dilakukan testing terhadap aplikasi yang telah dibuat. Dengan mencari keakuratan dan kekurangan yang ada pada aplikasi ini sehingga dapat ditingkatkan untuk pengembangan selanjutnya.

1.6 Sistematika Penulisan

Sistematika penulisan laporan proyek akhir ini secara keseluruhan terdiri dari lima bab, masing-masing terdiri dari beberapa sub bab. Adapun pokok pembahasan dari masing-masing bab tersebut secara garis besar sebagai berikut:

BAB I PENDAHULUAN

Pada bab ini menjelaskan tentang latar belakang masalah terkait penyalahgunaan teknologi *deepfake* oleh pengguna yang tidak

bertanggung jawab. Dengan perumusan masalah dari teknologi *deepfake* ini untuk penipuan, manipulasi informasi, serta merusak reputasi individu atau kelompok. Penulisan ini juga mencakup ruang lingkup masalah, tujuan dan manfaat penelitian, serta metodologi penelitian. Tujuan utama dari penelitian ini adalah merancang sebuah model CNN untuk mendeteksi video *deepfake* dan mengimplementasikannya ke dalam aplikasi Android. Selanjutnya, penulisan ini akan menjelaskan sistematika yang digunakan dalam penyusunan proyek.

BAB II TINJAUAN PUSTAKA

Pada bab ini menjelaskan beberapa hasil penelitian terdahulu mengenai pengembangan model untuk mendeteksi *deepfake*. Dengan mengambil 5 penelitian terdahulu untuk melakukan pertimbangan dan referensi dalam pembuatan model dan landasan teori yang diperlukan untuk mendukung pembuatan proyek ini seperti penjelasan dari *deepfake*, *deep learning*, *Convolutional Neural Network* (CNN), confusion matrix, dan Android.

BAB III PERANCANGAN

Bab ini menjelaskan tentang desain penelitian yang terdiri dari *model development* hingga perancangan aplikasi.

BAB IV JADWAL DAN PERKIRAAN BIAYA

Bab ini berisi informasi mengenai jadwal pengerjaan proyek akhir dan perkiraan biaya yang dibutuhkan untuk pengerjaan proyek akhir.