

BAB I

PENDAHULUAN

A. Latar Belakang

Perkembangan teknologi *Artificial Intelligence* atau kecerdasan buatan telah membawa transformasi besar dalam berbagai bidang, termasuk pengolahan citra digital. Salah satu kemajuan signifikan dalam kecerdasan buatan adalah munculnya *Generative Adversarial Networks* (GAN), yang mampu menghasilkan citra sintetis dengan kualitas tinggi dan menyerupai citra nyata. GAN bekerja melalui dua jaringan saraf yang saling bersaing *generator* dan *discriminator* untuk menghasilkan data yang semakin realistis [2]. Teknologi ini awalnya dikembangkan untuk keperluan positif seperti peningkatan resolusi gambar dan rekonstruksi citra medis. Namun, seiring waktu, GAN mulai disalahgunakan untuk merusak reputasi seseorang atau memanipulasi opini publik termasuk *deepfake*, yang memanipulasi citra wajah manusia secara realistis [1, 2].

Deepfake merupakan fenomena yang semakin mengkhawatirkan dalam era digital, di mana citra atau video seseorang dimanipulasi menggunakan kecerdasan sehingga menampilkan sesuatu yang sebenarnya tidak pernah terjadi. Dengan kualitas visual yang sangat meyakinkan, *deepfake* telah digunakan dalam berbagai bentuk penyalahgunaan, mulai dari pencemaran nama baik, penyebaran hoaks, manipulasi politik, hingga konten pornografi non-konsensual [3]. Beberapa laporan internasional dan nasional menunjukkan lonjakan kasus yang melibatkan *deepfake*, yang tidak hanya menyerang figur publik, tetapi juga individu biasa [4]. Penyebaran konten semacam ini melalui media sosial menyebabkan dampak sosial yang luas, karena publik semakin sulit membedakan antara informasi yang benar dan yang dimanipulasi.

Dalam upaya untuk mengatasi permasalahan ini, pendekatan berbasis *Convolutional Neural Network* (CNN) telah banyak digunakan dalam bidang

pengolahan citra digital. CNN merupakan jenis jaringan saraf tiruan yang secara khusus dirancang untuk mengenali pola visual melalui proses *convolution*, *pooling*, dan *activation* [5]. Dengan kemampuannya dalam mengekstraksi fitur spasial dari data visual, CNN telah terbukti efektif dalam berbagai tugas klasifikasi citra, seperti pengenalan wajah, deteksi objek, dan analisis medis. Keunggulan utama CNN terletak pada kemampuannya dalam mempelajari representasi visual secara otomatis dari data mentah, tanpa perlu melakukan *feature engineering* manual yang kompleks.

Terdapat berbagai arsitektur CNN yang telah dikembangkan untuk meningkatkan akurasi dan efisiensi pemrosesan citra, seperti *LeNet*, *AlexNet*, *VGGNet*, *ResNet*, dan *MovbileNet*. Setiap arsitektur memiliki karakteristik tersendiri dalam hal kedalaman jaringan, jumlah parameter, dan strategi pemrosesan fitur. Misalnya, *ResNet* menggunakan konsep *residual learning* untuk mengatasi masalah *vanishing gradient* pada jaringan yang sangat dalam, sementara *Inception* memanfaatkan berbagai ukuran filter dalam satu lapisan untuk menangkap informasi pada berbagai skala.

Penggunaan CNN sebagai solusi deteksi *deepfake* menjadi sangat relevan karena kemampuannya dalam mengidentifikasi anomali visual yang sering kali tidak dapat dikenali oleh mata manusia. Detail kecil seperti pencahayaan yang tidak alami, ketidakkonsistenan tekstur kulit, atau artefak pada tepi wajah dapat dikenali oleh model CNN yang telah dilatih dengan dataset yang sesuai. Dengan demikian, integrasi CNN dalam sistem deteksi *deepfake* memungkinkan proses identifikasi menjadi lebih cepat, akurat, dan andal, terutama di tengah meningkatnya kompleksitas konten digital yang dihasilkan oleh GAN. Hal ini menunjukkan bahwa CNN memiliki potensi besar sebagai alat bantu dalam mitigasi penyalahgunaan teknologi *deepfake*.

B. Rumusan Masalah

Berdasarkan masalah yang dipaparkan sebelumnya di latar belakang, dapat ditarik beberapa rumusan masalah untuk penelitian ini:

1. Bagaimana mengoptimalkan model *Convolutional Neural Network* (CNN) dalam mengklasifikasikan citra wajah asli dan citra wajah hasil rekayasa *deepfake* melalui penyesuaian *hyperparameter*?
2. Faktor-faktor apa saja, baik dalam pengaturan *hyperparameter* (seperti *epoch* dan *learning rate*) maupun fitur visual pada citra wajah, yang memengaruhi kinerja dan keputusan prediksi model CNN dalam membedakan citra asli dan *deepfake*?

C. Tujuan Penelitian

Atas permasalahan yang ada, maka penelitian ini memiliki beberapa tujuan yang hendak dicapai. Tujuan yang dimaksud antara lain adalah:

1. Mengoptimalkan model klasifikasi citra wajah *deepfake* menggunakan arsitektur *Convolutional Neural Network* melalui penyesuaian *hyperparameter* seperti *epoch* dan *learning rate*.
2. Mengevaluasi kinerja model CNN berdasarkan perubahan nilai *hyperparameter* terhadap akurasi klasifikasi citra wajah asli dan hasil rekayasa *deepfake*.
3. Mengidentifikasi fitur-fitur visual wajah yang berkontribusi terhadap keputusan prediksi model CNN dalam membedakan citra wajah asli dan *deepfake*.

D. Batasan Masalah

Untuk penelitian ini, peneliti merumuskan beberapa batasan agar penelitian dapat lebih terarah dan maksimal. Beberapa batasan ini antara lain adalah:

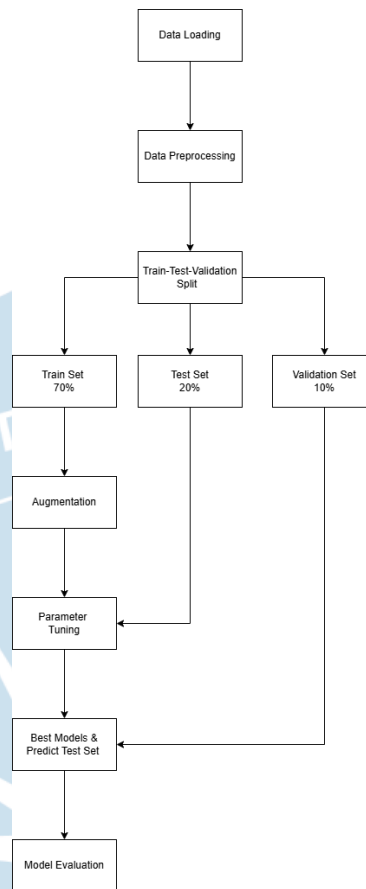
1. Fokus pada citra wajah: Penelitian ini hanya mencakup deteksi citra wajah hasil rekayasa Kecerdasan Buatan dan tidak mencakup media lain seperti video *deepfake* atau audio.
2. Algoritma Deep Learning: Penelitian akan menggunakan arsitektur *deep learning* berbasis *Convolutional Neural Networks* (CNN) dan tidak

akan mengeksplorasi metode lain di luar deep learning, seperti SVM atau decision tree.

3. **Dataset:** Penelitian akan menggunakan dataset publik seperti Kaggle yang terdiri dari citra wajah asli dan hasil rekayasa Kecerdasan Buatan. Data akan dibatasi pada citra wajah beresolusi standar yang telah banyak digunakan dalam penelitian sejenis.
4. **Metode Evaluasi:** Penelitian akan fokus pada klasifikasi biner, yaitu membedakan citra wajah asli dengan citra wajah hasil rekayasa Kecerdasan Buatan. Aspek lain seperti deteksi tingkat manipulasi (seberapa besar manipulasi yang terjadi pada citra wajah) tidak akan dibahas.
5. **Hyperparameter:** Penelitian ini akan memusatkan perhatian hanya terhadap 2 *hyperparameter* yaitu *epoch* dan *learning rate*.

E. Metode Penelitian

Penelitian ini dilakukan dengan pendekatan eksperimen kuantitatif yang bertujuan untuk mengembangkan model *deep learning* dalam mendeteksi dan mengidentifikasi citra wajah hasil rekayasa kecerdasan buatan. Metode penelitian akan dilakukan melalui beberapa tahapan berikut:



Gambar 1. 1 Diagram Metode Penelitian

1. Data Loading

Tahap awal dimulai dengan proses *data loading*, yakni pemuatan dataset citra wajah yang terdiri atas citra wajah asli dan citra hasil rekayasa berbasis *deepfake*. Dataset ini menjadi basis utama untuk pelatihan dan evaluasi model CNN.

2. Data Preprocessing

Setelah data dimuat, dilakukan proses *data preprocessing* yang mencakup penyesuaian ukuran gambar, normalisasi nilai piksel, dan konversi format yang sesuai untuk pemrosesan oleh jaringan saraf tiruan. Proses ini bertujuan untuk memastikan konsistensi dan kesiapan data sebelum dimasukkan ke dalam model.

3. Train-Test-Validation Split

Dataset kemudian dibagi menjadi tiga subset: *train set* sebesar 70%, *test set* sebesar 20%, dan *validation set* sebesar 10%. Pembagian ini bertujuan untuk memisahkan data pelatihan, pengujian, dan validasi guna menghindari *overfitting* dan memastikan generalisasi model.

4. Train Set dan Augmentation

Subset *train set* selanjutnya diperluas melalui proses *augmentation*, yaitu teknik manipulasi citra seperti rotasi, flipping, zooming, dan translasi. Tujuannya adalah untuk meningkatkan keragaman data latih dan memperkuat kemampuan generalisasi model terhadap variasi citra wajah.

5. Parameter Tuning

Setelah dilakukan proses *augmentation*, tahap selanjutnya adalah *parameter tuning* secara manual untuk mengevaluasi berbagai konfigurasi *hyperparameter* serta pilihan arsitektur model *Convolutional Neural Network* (CNN). Tahapan ini dilakukan dengan mengubah nilai-nilai *hyperparameter* secara sistematis, seperti jumlah *layer*, ukuran *filter*, fungsi aktivasi, dan *learning rate*, guna mengidentifikasi kombinasi yang memberikan performa optimal pada data pelatihan dan validasi.

6. Best Models & Predict Test Set

Model terbaik yang diperoleh dari proses *grid search* dan *cross-validation* kemudian digunakan untuk melakukan prediksi pada *test set*. Ini menjadi langkah penting dalam mengukur kinerja model terhadap data yang sebelumnya tidak dilihat.

7. Model Evaluation

Tahap akhir adalah *model evaluation*, yaitu proses evaluasi performa model dengan menggunakan metrik seperti *accuracy*, *precision*, *recall*, dan *F1-score*. Evaluasi ini dilakukan tidak hanya pada *test set*, tetapi juga dapat

divalidasi ulang menggunakan *validation set* untuk memastikan konsistensi kinerja.

F. Sistematika Penulisan

Penelitian ini disusun secara sistematis untuk memberikan pemahaman yang jelas mengenai pengembangan sistem klasifikasi otomatis produk gaharu berbasis *supervised learning*. Untuk mempermudah dalam membaca dan memahami keseluruhan pembahasan dalam penelitian ini, diperlukan sistematika penulisan yang terstruktur. Sistematika penulisan ini mencakup beberapa bab utama, yaitu:

BAB I PENDAHULUAN

Bab ini membahas latar belakang penelitian yang menjelaskan pentingnya pengembangan sistem deteksi *deepfake* di era digital saat ini. Di dalamnya dijabarkan pula rumusan masalah yang menjadi dasar penelitian, serta tujuan yang ingin dicapai melalui penerapan teknik *deep learning* menggunakan *Convolutional Neural Network*.

Selain itu, bab ini juga memuat manfaat penelitian bagi masyarakat, akademisi, dan pengembang teknologi, ruang lingkup penelitian untuk membatasi fokus kajian, serta sistematika penulisan yang menggambarkan alur isi dari keseluruhan laporan tugas akhir.

BAB II TINJAUAN PUSTAKA

Bab ini menguraikan berbagai penelitian terdahulu yang memiliki keterkaitan dengan topik deteksi *deepfake*. Kajian pustaka ini bertujuan untuk mengetahui perkembangan teknologi di bidang deteksi manipulasi citra wajah dan mengidentifikasi metode yang telah digunakan dalam penelitian sebelumnya.

Melalui tinjauan pustaka, penulis dapat memahami posisi penelitian ini dibandingkan dengan penelitian lain, serta memperkuat landasan dalam memilih pendekatan *deep learning* sebagai metode utama yang digunakan dalam sistem yang dikembangkan.

BAB III LANDASAN TEORI

Bab ini menjelaskan konsep dan teori dasar yang mendasari pengembangan sistem, seperti kecerdasan buatan, *machine learning*, *deep learning*, serta arsitektur *Convolutional Neural Network*. Penjelasan mengenai cara kerja *CNN* dalam mengenali pola pada citra wajah juga dibahas secara rinci.

Selain itu, bab ini membahas teori mengenai manipulasi citra digital dan karakteristik *deepfake*, termasuk tantangan dalam proses pendeteksian serta pendekatan teknis yang relevan untuk menangani isu tersebut.

BAB IV DATASET DAN PENGEMBANGAN MODEL

Bab ini memaparkan proses pemilihan dan pengumpulan dataset yang digunakan dalam pelatihan model deteksi *deepfake*. Penjelasan meliputi jenis dataset, format data, distribusi kelas, serta langkah-langkah pra-pemrosesan data seperti augmentasi, dan konversi citra.

Kemudian dijelaskan tahapan pengembangan model, mulai dari pemilihan arsitektur *CNN*, konfigurasi parameter pelatihan, hingga proses training dan validasi model. Tujuannya adalah membangun model deteksi yang mampu membedakan citra asli dan citra hasil rekayasa dengan akurasi tinggi.

BAB V IMPLEMENTASI MODEL DAN PENGUJIAN SISTEM

Bab ini membahas proses implementasi model CNN ke dalam sistem deteksi yang telah dirancang. Disajikan pula antarmuka pengguna (jika ada), alur kerja sistem, serta cara penggunaan sistem dalam mendeteksi citra *deepfake*.

Selanjutnya, dilakukan pengujian terhadap sistem untuk mengevaluasi performa model menggunakan metrik evaluasi seperti akurasi, presisi, recall, dan F1-score. Hasil pengujian dianalisis untuk mengetahui kekuatan dan kelemahan sistem, serta validitas model dalam kondisi nyata.

BAB VI PENUTUP

Bab terakhir ini menyajikan kesimpulan dari keseluruhan proses penelitian, mencakup hasil yang dicapai dan efektivitas model deteksi yang dibangun. Kesimpulan disusun berdasarkan analisis data dan pencapaian tujuan penelitian.

Selain itu, penulis juga menyampaikan saran-saran untuk penelitian selanjutnya, seperti pengembangan model dengan dataset yang lebih besar, peningkatan arsitektur jaringan, atau penerapan sistem pada platform nyata sebagai alat bantu pendeteksi *deepfake* secara langsung.