1. PENDAHULUAN

1.1. Latar Belakang

Dengan kemajuan teknologi dalam bidang kecerdasan buatan, manipulasi media digital telah mencapai tingkat keahlian yang mengkhawatirkan. Banyak kejahatan digital yang sering dilakukan dengan memanfaatkan *Artificial Intelligent* (AI). Bisa kita temui di aplikasi ponsel ataupun sebuah iklan di *platform* media sosial, AI tampil bahwa mereka dapat membantu dalam berbagai pekerjaan. AI juga dapat membuat sebuah ilustrasi gambar yang hampir sama dengan seseorang asli di dunia nyata. AI juga dapat mengubah wajah seseorang dengan wajah seseorang yang lainnya. Dari sini, muncul suatu bentuk kejahatan digital dalam memanipulasi citra. Salah satu bentuk manipulasi yang semakin memprihatinkan adalah *deepfake* [1], di mana teknik *Generative Adversarial Networks* (GANs) merupakan salah satu pendekatan *deep learning* yang dapat digunakan untuk membuat video palsu yang sangat realistis [2] dan seringkali sulit untuk dibedakan dari rekaman asli. *Deepfake* memiliki potensi serius untuk digunakan dalam kegiatan kejahatan seperti penipuan, pemalsuan bukti, dan penggelapan identitas [3], sehingga meningkatkan risiko keamanan bagi individu dan organisasi.

Kini mulai banyak *tools* untuk membuat gambar baik 2D seperti Midjourney, Stable Diffusion, Fooocus, Roop, dan lain lain. *Tools* tersebut memanfaatkan AI dalam membuat sebuah gambar yang realistis. Namun sayangnya teknologi ini kerapkali digunakan untuk pemalsuan identitas pribadi dikarenakan AI yang dapat membuat sebuah gambar menyerupai wajah orang secara spesifik. Selain dapat membuat gambar yang menyerupai seseorang, teknologi *tools* dengan bantuan AI ini juga dapat melakukan *faceswap* yang mana AI hanya perlu merubah wajah seseorang yang satu dengan wajah orang yang lainnya tanpa perlu merubah keseluruhan gambar. Dari sini timbullah kejahatan *deepfake* [3] yang memanipulasi identitas seseorang yang satu seorang yang lain baik dalam bentuk foto, video, ataupun suara.

Sistem CCTV (*Closed-Circuit Television*) adalah alat penting dalam pengawasan dan keamanan modern [5]. CCTV memberikan bukti visual yang kritis dalam berbagai situasi, dari investigasi kecelakaan lalu lintas hingga penegakan hukum. CCTV kini sudah banyak digunakan di khalayak umum, seperti misalnya toko/kios, rumah, jalan raya, kantor, dan lain-lain. Namun, dengan maraknya *deepfake*, integritas dan keandalan bukti visual dari sistem CCTV [6] menjadi terancam. Dengan pemalsuan citra pada CCTV mengakibatkan tindakan kriminal yang terjadi dapat dimanipulasi.

Dikutip dari media berita CNN.com [16] bahwa pada tahun 2021 *hacker* Malaysia telah

meretas 5000 CCTV Israel. Mereka mengunggah hasil tangkapan gambar CCTV di media sosial yang menampilkan pemukiman penduduk hingga lembaga pemerintah Israel. Dikutip dari biometricupdate.com [13] menjelaskan bahwa kamera dapat dilakukan peretasan dengan serangan *video injection* yang nantinya akan digunakan untuk melakukan kejahatan menipu identitas *biometric*. Pelaku penipuan menggunakan *charged-coupled device* untuk melakukan *bypass* pada kamera agar dapat melakukan injeksi *pre-recorded content* [13]. Berdasarkan dua kasus tersebut, diketahui bahwa CCTV merupakan bagian dari kamera yang berfungsi untuk merekam aktivitas sebagai perangkat pembantu keamanan namun CCTV juga masih rentan dari aktivitas *hacking* yang dapat merugikan identitas seseorang tertentu.

Penting untuk mengembangkan metode dan alat yang efektif untuk mendeteksi *deepfake* dalam rekaman CCTV. Dengan demikian, tugas akhir ini bertujuan untuk mengatasi tantangan ini dengan memanfaatkan pendekatan *deep learning* GANs untuk mendeteksi manipulasi visual yang terkait dengan *deepfake* dalam rekaman CCTV.

Melalui penelitian ini, dihasilkan solusi yang dapat memberikan perlindungan tambahan terhadap manipulasi video yang tidak sah [6] yang dilakukan oleh pelaku kriminalitas, meningkatkan keamanan dan integritas informasi visual pada sistem CCTV, dan memberikan alat penting dalam mendukung investigasi keamanan. Dengan memahami dan mengatasi ancaman *deepfake*, tugas akhir ini bertujuan untuk mendeteksi *image* asli dan palsu hasil *deepfake* yang dapat mendukung investigasi keamanan dan memastikan keaslian bukti rekaman. Pendekatan yang digunakan pada penelitian ini adalah metode *deep learning* yaitu GANs. Melihat bahwa GANs dapat menciptakan sebuah visual palsu seperti *deepfake*, dari penelitian ini juga akan dilihat apakah GANs juga dapat digunakan untuk mendeteksi visual tersebut merupakan gambar atau video asli atau merupakan sebuah *deepfake*.

1.2. Perumusan Masalah

Berdasarkan latar belakang yang telah diuraikan dapat dibentuk beberapa rumusan masalah untuk penelitian ini. Pada penelitian ini dibangun bagaimana mendeteksi *deepfake* menggunakan pendekatan GANs. Diperlukan juga untuk mengetahui sejauh mana ketepatan dan keandalan deteksi *deepfake* menggunakan model GANs ketika diuji dengan berbagai situasi pengawasan CCTV dalam scenario keamanan nyata.

1.3. Batasan Masalah

Terdapat beberapa batasan masalah pada penelitian ini agar penelitian ini dapat berfokus kepada topik yang diteliti dan berfokus pada satu penelitian. Batasan masalah yang dapat diuraikan diantaranya:

1. Dataset yang digunakan merupakan dataset *hybird* yang dikumpulkan sendiri oleh peneliti dan menggunakan *human detection dataset* yang diperoleh dari kaggle.com
2. Dataset yang digunakan adalah data *capture image* asli dari CCTV dan juga data *image* palsu yang diciptakan dengan bantuan AI.
3. Program yang dibangun menggunakan bahasa pemograman Python pada google colab dan jupyter
4. *Capture* yang direkam oleh CCTV harus direkam terlebih dahulu lalu dilakukan *screenshoot*, bukan video *real-time*

1.4. Tujuan

Tujuan dari penelitian ini adalah untuk mendeteksi *deepfake* pada CCTV menggunakan metode GANs. Pada penelitian ini dilakukan klasifikasi menggunakan GANs dan juga metode pembanding yang mendeteksi *capture* pada CCTV yang mana hasil *capture* tersebut apakah berupa asli atau berupa *deepfake*, serta dilakukan perbandingan antara GANs dan metode pembanding untuk melihat performansi dari metode-metode tersebut.

1.5. Organisasi Penulisan

Pada laporan ini, sistematika penulisan dapat dijabarkan sebagai berikut. Pada bab 1 kami menjelaskan secara umum menegnai teknologi AI Faceswap, CCTV, dan persiapan penelitian. Pada bab 2 kami menjelaskan mengenai studi terkait, *deepfake*, metode GANs, dan metode pembanding. Pada bab 3 kami menjelaskan mengenai alur yang terdiri dari proses pengumpulan dataset secara manual oleh peneliti, normalisasi dataset, *preprocessing* data, pembangunan model dan program. Pada bab 4 kami menjelaskan analisa hasil pengujian dan faktor yang mempengaruhi hasil, dan pada bab 5 kami menjelaskan kesimpulan pada penelitian ini.