Proteksi Citra dari *DeepFake* dengan CMUA-Watermark

Diajukan untuk Menyusun Skripsi
di jurusan Teknik Informatika Fakultas Ilmu Komputer UNSRI

Oleh :

Renaldi Budi Setiawan
NIM : 09021281823066

Jurusan Teknik Informatika
FAKULTAS ILMU KOMPUTER UNIVERSITAS SRIWIJAYA
2022

LEMBAR PENGESAHAN PROPOSAL SKRIPSI

Proteksi Citra dari *DeepFake* dengan CMUA-*Watermark*

Oleh :

Renaldi Budi Setiawan
NIM : 09021281823066

Indralaya, Oktober 2022

Pembimbing I	Pembimbing II,
Syamsuryadi, S.Si., M.Kom., Ph.D.	Muhammad Qurhanul Rizqie, S.KOM,. M.T., Ph.D.
NIP 197102041997021003	NIP 1671060312870008

Mengetahui,
ketua Jurusan

Alvi Syahrini Utami, M.Kom.
NIP. 19781222200642003

DAFTAR ISI

Halaman

BAB I

PENDAHULUAN

1.1 Pendahuluan

Pada bab ini akan dibahas berkenaan dengan garis besar pokok-pokok pikiran dalam penelitian ini. Pokok pikiran yang akan dibahas antara lain latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, dan manfaat penelitian. Pokok-pokok pikiran yang diuraikan akan dijadikan acuan dalam kajian penelitian ini.

1.2 Latar Belakang Masalah

Berita palsu telah menjadi isu yang mengancam bagi persepsi publik, masyarakat, dan demokrasi (Borges et al., 2019; Qayyum et al., 2019). Berita palsu mengacu pada konten berita fiktif yang dibuat untuk menipu publik (Aldwairi & Alwahedi, 2018; Jang &Kim, 2018), salah satu contohnya adalah *Deepfakes*.

Deepfake sendiri baru dipopulerkan di tahun 2017, berawal dari pengguna Reddit mengunggah video porno hasil editan. Pengguna Reddit ini mengembangkan GAN menggunakan TensorFlow. Teknologi *Deepfake* dapat menghasilkan berupa video lucu, pornografi, atau politik seseorang yang mengatakan apa pun, tanpa persetujuan orang yang gambar dan suaranya terlibat (Hari, 2018; Fletcher, 2018). Foto KPM mahasiswa unsri sangat mudah diakses pada situs resmi laman Unsri versi lama (https://old.unsri.ac.id/?act=daftar_mahasiswa). Hal ini sangat memungkinkan

terjadi penyalahgunaan foto tersebut oleh oknum yang dengan sengaja melakukan tindakkan tidak bertanggung jawab seperti pembuatan *deepfake*.

Untuk mencegah hal tersebut *adversarial watermark* dapat digunakan untuk memerangi *deepfake model*, *adversarial watermark* dapat menghasilkan gambar yang terdistorsi. Namun metode ini masih kurang efisien karena memerlukan proses pelatihan individu untuk setiap gambar wajah, untuk menghasilkan *adversarial attack model* terhadap *deepfake model* tertentu. Untuk mengatasi masalah ini, penelitian ini menggunakan metode *universal adversarial attack model* pada *deepfake model*, untuk menghasilkan *Cross-Model Universal Adversarial Watermark* (CMUA-Watermark) yang dapat melindungi ribuan gambar wajah dari beberapa model *deepfake* (Huang et al., 2021).

1.3 Rumusan Masalah

Berdasarkan permasalahan pada latar belakang yang telah diuraikan maka rumusan masalah dari penelitian ini adalah sebagai berikut :

1. Bagaimana cara memproteksi citra gambar foto KPM mahasiswa Unsri dengan metode CMUA-Watermark?
2. Bagaimana tingkat akurasi metode CMUA-*Watermark* dalam memproteksi citra foto KPM mahasiswa Unsri dari *deepfakes*?

1.4 Tujuan Masalah

Tujuan dari penelitian ini adalah :

1. Membangun perangkat lunak yang dapat memproteksi citra gambar foto KPM mahasiswa Unsri menggunakan metode CMUA-*Watermark*.

2. Mengetahui tingkat akurasi penggunaan metode CMUA-*Watermark* dalam memproteksi citra foto KPM mahasiswa Unsri dari *deepfakes*.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian ini adalah:

1. Sistem yang dibuat dapat memproteksi citra gambar foto KPM mahasiswa Unsri menggunakan metode CMUA-*Watermark*.
2. Hasil penelitian dapat dijadikan sebagai rujukan untuk penelitian terkait di masa mendatang.

1.6 Batasan Masalah

Batasan masalah dalam penelitian ini adalah sebagai berikut:

1. Dataset yang digunakan merupakan dataset Celeb-a, dari penelitian *Deep Learning Face Attributes in the Wild* (2015).
2. Data uji yang digunakan merupakan dataset foto mahasiswa jurusan Teknik Informatika Universitas Sriwijaya Angkatan 2018.
3. Ekstensi citra yang didukung oleh perangkat lunak adalah .jpg.

1.7 Sistematika Penulisan

Sistematika penulisan tugas akhir mengikuti standar penulisan tugas akhir Fakultas Ilmu Komputer Universitas Sriwijaya yaitu sebagai berikut:

BAB I. PENDAHULUAN

Pada bab ini akan membahas landasan dari penelitian, seperti latar belakang, rumusan masalah, tujuan dan manfaat penelitian, batasan masalah serta sistematika penulisan.

BAB II. KAJIAN LITERATUR

Pada bab ini membahas literatur pada penelitian, seperti pengertian Citra, *Deepfake*, CMUAI-*Watermark* dan penelitian yang relevan.

BAB III. METODOLOGI PENELITIAN

Pada Bab ini menjelaskan pelaksanaan alur penelitian, yakni pengumpulan data dan perancangan pembangunan perangkat lunak. Serta tahapan dijelaskan secara detail berdasarkan kerangka yang dibuat.

1.8 Kesimpulan

Pada bab ini telah menjelaskan dasar dan tolak ukur pada penelitian, seperti latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan.