Proteksi Citra Foto KPM Mahasiswa Fasilkom UNSRI 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 Foto KPM Mahasiswa Fasilkom UNSRI 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 merupakan ancaman bagi kepentingan masyarakat umum (Borges et al., 2019; Qayyum et al., 2019). Berita palsu mengacu pada konten gaya berita fiktif yang dibuat untuk menipu publik (Aldwairi & Alwahedi, 2018; Jang & Kim, 2018). Informasi palsu menyebar dengan cepat melalui media sosial, yang dapat berdampak pada jutaan pengguna (Figueira & Oliveira, 2017). Saat ini, satu dari lima pengguna internet mendapatkan berita melalui *YouTube*, kedua setelah Facebook (Anderson, 2018). Peningkatan popularitas video ini menyoroti perlunya alat untuk mengonfirmasi keaslian konten media dan berita, karena teknologi baru memungkinkan manipulasi video yang meyakinkan (Anderson, 2018). Mengingat kemudahan dalam memperoleh dan menyebarkan informasi yang salah melalui platform media sosial, semakin sulit untuk mengetahui apa yang harus dipercaya, yang mengakibatkan konsekuensi berbahaya bagi pengambilan keputusan yang terinformasi (Borges et al., 2019; Britt et al., 2019). Memang, hari ini kita hidup

di apa yang oleh beberapa orang disebut era “pasca-kebenaran”, yang ditandai dengan disinformasi digital dan perang informasi yang dipimpin oleh aktor jahat yang menjalankan kampanye informasi palsu untuk memanipulasi opini publik (Anderson, 2018; Qayyum et al., 2019; Zannettou et al., 2019).

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 berupa video lucu, pornografi, atau politik seseorang yang mengatakan apa pun, tanpa persetujuan orang yang citra gambar dan suaranya terlibat (Day, 2019; Fletcher, 2018).

Situs resmi milik UNSRI versi lama[1] hingga saat ini dapat dengan mudah diakses oleh siapapun tanpa memerlukan verifikasi terlebih dahulu. Adanya laman ini juga memberikan informasi terkait mahasiswa, termasuk didalamnya foto diri mahasiswa. Setiap foto yang diunggah oleh pihak UNSRI di situs tersebut akan muncul dalam hasil pencarian gambar Google dan dapat dengan mudah diunduh. Informasi pribadi mahasiswa yang ada didalam situs ini sangat memungkinkan dapat dieksploitasi oleh seseorang untuk melakukan tindakan kejahatan, seperti mengatasnamakan identitas dan menggunakan wajah mahasiswa tersebut dengan menggunakan *deepfake*.

Adversarial watermark dapat digunakan untuk memerangi *deepfake model*, *adversarial watermark* dapat menghasilkan citra gambar yang terdistorsi (Ruiz et al., 2020). Namun metode ini masih kurang efisien karena memerlukan proses

[1] https://old.UNSRI.ac.id/?act=daftar_mahasiswa

pelatihan individu untuk setiap citra gambar wajah, untuk menghasilkan *adversarial attack model* terhadap *deepfake model* tertentu (Huang et al., 2021). 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 citra 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

1. Bagaimana cara memproteksi citra gambar foto KPM mahasiswa Fasilkom UNSRI dari *deepfake* dengan metode *CMUA-Watermark*?
2. Bagaimana tingkat keberhasilan metode *CMUA-Watermark* dalam memproteksi citra foto KPM mahasiswa Fasilkom UNSRI dari *deepfakes*?

1.4 Tujuan Masalah

Tujuan dari penelitian ini adalah:

1. Menghasilkan perangkat lunak yang dapat memproteksi citra gambar foto KPM mahasiswa Fasilkom UNSRI dari *deepfake* menggunakan metode *CMUA-Watermark.*
2. Mengetahui tingkat keberhasilan penggunaan metode *CMUA-Watermark* dalam memproteksi citra foto KPM mahasiswa Fasilkom 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 Fasilkom UNSRI dari *deepfake* 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 pelatihan merupakan dataset Celeb-a, dari penelitian *Deep Learning Face Attributes in the Wild* (2015).
2. Data uji yang digunakan merupakan dataset foto mahasiswa Fakultas Ilmu Komputer Universitas Sriwijaya Angkatan 2018.
3. Ekstensi citra yang didukung oleh perangkat lunak adalah .jpg.
4. Penelitian hanya fokus dalam proteksi citra dari *deepfake*.

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*, CMUA-*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 patokan pada penelitian , seperti latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan.

BAB II

KAJIAN LITERATUR

2.1 Pendahuluan

pada bab ini akan dijelaskan mengenai dasar-dasar teori digunakan pada penelitian ini. Serta penjelasan hasil dari penelitian-penelitian terkait mengenai citra, *Deepfakes*, *CMUA-Watermark* dan RUP. Pada bab ini pula dibahas mengenai penelitian terkait lainnya yang relevan.

2.2 Landasan Teori

2.2.1 Citra

Citra didefinisikan sebagai fungsi dari dua *variable* misalnya $a(x,y)$ dimana *a* sendiri sebagai amplitudo (misalnya kecerahan) citra pada koordinat (x,y). Citra digital $a[m,n]$ merupakan citra dalam ruang diskrit 2D yang berasal dari citral Analog $a(x,y)$ di ruang kontinu 2D melalui proses sampling yaitu yang biasa disebut sebagai digitalisasi (Young et al., 2006).

Menurut McAdrew citra digital adalah citra $f(x,y)$ yang telah didiskritkan pada koordinat spasial dan kecerahan. Citra digital direpresentasikan oleh *array* dua dimensi dimana setiap *array* merepresentasikan satu kanal warna. Nilai warna kecerahan yang didigitalkan ini dinamakan nilai tingkat keabuan . Setiap elemen array tersebut dinamakan *pixel* atau pel yang diambil dari istilah '*picture element*'. Dimensi pada citra ditulis dengan format panjang x tinggi. Namun pada citra digital didefinisikan dengan ukuran tinggi *M* dan panjang *N*.

$$f(x,y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0,N-1) \\ f(1,0) & f(1,0) & \cdots & f(1,N-1) \\ \vdots & \vdots & & \vdots \\ f(M,-1,0) & f(M,-1,1) & \dots & f(M,-1,N-1) \end{bmatrix}$$

Koordinat citra dimulai dari pojok kiri atas, secara Sistematis di dimulai dari (0,0) dan berakhir di (M-1,N-1)(McAndrew, 2014).

2.2.2 *DeepFake*

Deepfake adalah Kombinasi dari "pembelajaran mendalam" dan "palsu", *deepfake* adalah video hiper-realistis yang dimanipulasi secara digital untuk menggambarkan orang-orang yang mengatakan dan melakukan hal-hal yang tidak pernah benar-benar terjadi (Metz, 2019; Metz & O'Sullivan, 2019). *Deepfake* mengandalkan jaringan saraf yang menganalisis kumpulan besar sampel data untuk belajar meniru ekspresi wajah, tingkah laku, suara, dan infleksi seseorang (Dickson, 2018). Prosesnya melibatkan memasukkan rekaman dua orang ke dalam algoritme pembelajaran mendalam untuk melatihnya bertukar wajah(Rubenking & Eddy, 2019). Dengan kata lain, *deepfake* menggunakan teknologi pemetaan wajah dan AI yang menukar wajah seseorang di video menjadi wajah orang lain (Horowitz, 2019; Wallace, 2019). *Deepfake* muncul ke publisitas pada tahun 2017 ketika pengguna *Reddit* mem-*posting* video yang menunjukkan selebriti dalam situasi seksual yang membahayakan (Brown, 2019; Leetaru, 2019; Marr, 2019). *Deepfake* sulit dideteksi, karena mereka menggunakan rekaman nyata, dapat memiliki audio yang terdengar otentik, dan dioptimalkan untuk menyebar di media sosial dengan cepat (FRB05; WP01).

Dengan demikian, banyak pemirsa menganggap bahwa video yang mereka lihat adalah asli (CNET01; CNN10).

2.2.2.1 *Photo Deepfake*

2.2.2.1.1 *Face and Body Swapping*

Dalam hal ini, perubahan dilakukan pada wajah dan tubuh dengan mengganti atau memadukan tubuh dan wajah dengan wajah atau tubuh orang lain. Hasilnya adalah orang yang sama sekali berbeda dalam citra gambar aslinya. Contoh pendekatan ini dapat dilihat di banyak aplikasi menggunakan *Aging filter*. Ini dapat berguna bagi pelanggan untuk mencoba pakaian, kosmetik, atau gaya rambut secara virtual.

2.2.2.2 *Deepfake Creation*

Video *Deepfake* sangat sempurna sehingga dapat membodohi siapa pun. Berbagai alat dan aplikasi digunakan untuk mengembangkan video *deepfake* ini. Aplikasi ini sebagian besar menggunakan teknik pembelajaran mendalam untuk mengembangkan video ini. Video *deepfake* pertama dibuat menggunakan *FakeApp* yang dikembangkan oleh pengguna *Reddit.* Untuk memahaminya lebih jelas, mari kita ambil contoh gambar diam ini dari film "*Man of Steel*" di mana wajah aktris Amy Adams diganti dengan aktor lain Nicolas Cage seperti yang ditunjukkan pada Gambar II-1.

Gambar II-1, *Frame* dari klip *deepfake* film "*Man of Steel*".

Gambar II-1 menunjukkan gambar asli dari film *Man of Steel* dengan wajah aktris Amy Adams di sebelah kiri, dan di sebelah kanan adalah bingkai *deepfake* yang menggantikan wajah dengan Nicolas Cage. Contoh ini menunjukkan bagaimana wajah perempuan diganti dengan wajah laki-laki. Beginilah cara melakukannya:

1. Wilayah gambar yang menunjukkan wajah Amy Adams diambil dari video aslinya.
2. Gambar yang diekstrak ini digunakan sebagai *input* untuk di proses *deep learning*, teknik AI ini digunakan untuk secara otomatis menghasilkan gambar yang cocok, Nicolas Cage.
3. Gambar yang dihasilkan sekarang ditukar dengan wajah asli di dalam video asli dan menghasilkan video deepfake.

2.2.3 *CMUA-Watermark*

CMUA-Watermark adalah sebuah teknik *watermarking* yang digunakan untuk memberikan tanda air (watermark) pada data multi-media secara rahasia dan tangguh. Metode ini juga memiliki kemampuan untuk mengatasi masalah

kehilangan atau perubahan data yang disebabkan oleh proses kompresi, *cropping*, rotasi, dll (Li et al., 2019).

2.2.3.1 Metode *CMUA-Watermark*

Metode *CMUA-Watermark* menggabungkan teknik *deep learning* dan *adversarial learning* untuk memberikan watermark pada data multimedia. Pada tahap pelatihan, sebuah jaringan saraf berbasis CNN (*Convolutional Neural Network*) dilatih menggunakan pasangan data multimedia yang memiliki watermark dan tanpa watermark. Setelah pelatihan selesai, jaringan saraf akan menghasilkan watermark yang dapat diterapkan pada berbagai jenis data multimedia.

Selama proses *embedding*, data multimedia diubah menjadi representasi vektor dalam ruang fitur oleh jaringan saraf yang telah dilatih sebelumnya. Kemudian, watermark yang dihasilkan oleh jaringan saraf dimasukkan ke dalam representasi vektor tersebut. Setelah itu, representasi vektor yang telah dimodifikasi digunakan untuk menghasilkan data multimedia yang telah ditandai dengan watermark.

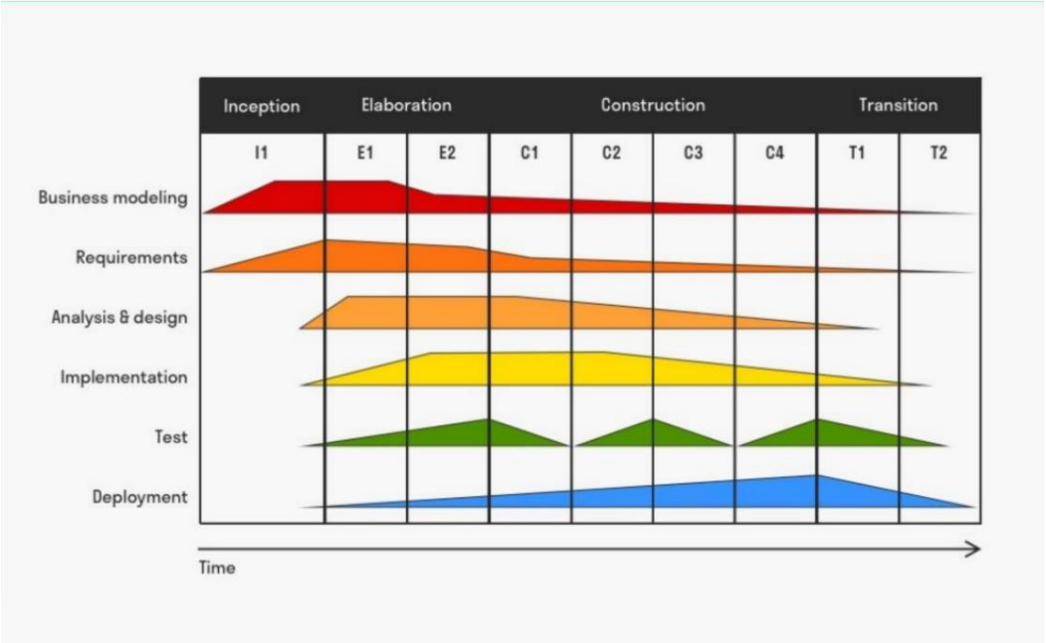
Proses pencarian watermark dilakukan dengan menggunakan sebuah jaringan saraf yang dibuat khusus untuk memisahkan watermark dari data multimedia. Jaringan saraf ini juga dilatih menggunakan pasangan data multimedia yang memiliki watermark dan tanpa watermark.

Metode *CMUA-Watermark* memiliki beberapa keunggulan dibandingkan dengan metode *watermarking* lainnya. Salah satu keunggulannya adalah

kemampuan untuk menghasilkan *watermark* yang dapat diterapkan pada berbagai jenis data multimedia dengan keandalan yang tinggi. Selain itu, metode ini juga dapat mengatasi masalah kehilangan atau perubahan data yang disebabkan oleh proses kompresi, cropping, rotasi, dll dengan baik (Li et al., 2019).

2.2.4 Rational Unified Process

Rational Unified Process (RUP) adalah metode rekayasa pengembangan perangkat lunak yang digunakan untuk kedisiplinan dalam penetapan tugas dan tanggung jawab. Tujuan RUP adalah memastikan bahwa produk perangkat lunak yang dihasilkan akan berkualitas dan sesuai kebutuhan pengguna akhir (end-users) (Anwar, 2014). RUP yang baik akan tercipta lewat hasil kerja sama antara pengembang perangkat lunak, mitra dan pengguna. Salah satu perspektif dalam RUP merupakan *Dynamic Perspective* & *Lifecycle Phases* yang penggunaannya digambarkan dalam bidang dua dimensi. Bidang horizontal menyatakan lamanya waktu pengembangan dan aspek dinamis lainnya, sedangkan bidang vertikal menyatakan aspek statis dalam rekayasa pengembangan perangkat lunak. Perspektif RUP model ini dinyatakan seperti dalam gambar II-2.



Gambar II-2,Arsitektur Rasional Unified Process

Dalam bidang horizontal, terdapat fase atau tahap dalam proses rekayasa perangkat lunak yang memaparkan peran dari tiap unit. Fase dalam bidang ini terbagi ke dalam fase insepsi, elaborasi, konstruksi dan transisi.

1. Fase insepsi merupakan fase yang berfokus pada pendefinisian ruang lingkup atau batasan dalam proyek pengembangan dengan cara melakukan analisis desain berorientasi objek (*Object Oriented Analysis Design*). Tujuan dari fase ini adalah untuk mendapatkan seluruh pemahaman dari pihak yang berkaitan agar sistem yang diajukan sesuai dengan keinginan dan kebutuhan.
2. Fase elaborasi merupakan fase yang akan membuat arsitektur dasar sistem lewat hasil analisis sebelumnya. Fase ini juga akan menentukan perencanaan proyek serta spesifikasi dari fitur yang akan dimuat dalam

sistem. Hasil dari fase ini merupakan dokumen arsitektur yang berguna untuk fase selanjutnya.

3. Fase Konstruksi merupakan fase menerjemahkan spesifikasi fitur dari dokumen rancangan sebelumnya ke dalam bentuk program/sistem sesuai dengan arsitekturnya. Fase ini berfokus pada peningkatan fungsi serta implementasi yang lebih mendalam terhadap spesifikasi sistem.

Fase Transisi merupakan fase pengujian sistem ke pengguna akhir dimana sistem yang dibuat harus memenuhi kebutuhan perangkat lunak dan kebutuhan penggunanya. Kendali dalam fase ini mulai dipindah kepada tim pemeliharaan perangkat lunak.

2.3 Penelitian Lain yang Relevan

Penelitian yang telah dilakukan mengenai *Landmark Breaker: Obstructing DeepFake By Disturbing Landmark Extraction* (Sun et al., 2020). Tulisan ini menjelaskan metode baru, yaitu *Landmark Breaker*, untuk menghalangi generasi *DeepFake* dengan melanggar langkah prasyarat ekstraksi *landmark* wajah. Dengan menciptakan *adversarial perturbations* untuk mengganggu ekstraksi *landmark* wajah, sehingga wajah input ke model *DeepFake* tidak dapat disejajarkan dengan baik. *Landmark Breaker* divalidasi pada himpunan data Celeb-DF, yang menunjukkan kemanjuran *Landmark Breaker* pada ekstraksi *landmark* wajah yang mengganggu.

2.4 Kesimpulan

Pada bab ini telah dibahas teori yang akan digunakan sebagai dasar penelitian ini. Pada bab ini juga telah dibahas mengenai penelitian terkait yang mendukung literatur penelitian ini. Mekanisme pelaksanaan penelitian selengkapnya akan dibahas dalam bab selanjutnya.

BAB III

METODE PENELITIAN

3.1 Pendahuluan

Pada bab ini akan dijelaskan mengenai tahapan penelitian, metode penelitian serta manajemen proyek penelitian. Tahapan penelitian dijadikan sebagai acuan pada setiap fase pengembangan perangkat lunak agar dapat memberikan solusi untuk rumusan masalah dan tercapainya tujuan penelitian.

3.2 Pengumpulan Data

Pada bagian ini akan dijelaskan tahapan pengumpulan data meliputi jenis dan sumber data dan metode pengumpulan data yang digunakan dalam penelitian.

3.2.1 Jenis dan Sumber Data

Jenis data yang digunakan sebagai objek penelitian ini ada dua, data primer dan sekunder. Data primer berupa kumpulan data dari foto KPM mahasiswa Fasilkom UNSRI Angkatan 2018 yang didapatkan dari proses *Scraping* pada situs resmi laman UNSRI lama[2]. Sedangkan data sekunder berupa dataset Celeb-A[3] dari peneletian *Deep Learning Face Attributes in the Wild* (Liu et al., 2015).

[2] http://old.unsri.ac.id/?act=daftar_mahasiswa
[3] https://www.kaggle.com/datasets/nikhilbartwal001/celeba

Gambar III-1. Contoh data yang dari *dataset* celebA

3.2.2 Metode pengumpulan Data

Ada dua metode pengumpulan data yang digunakan dalam penelitian ini:

1. mengunduh dataset Celeb-A dari laman ini[4], yang terdapat dalam github[5], hasil ekstrasi data dari halaman[6] atau dapat di unduh dari halaman Kaggle[7].
2. *Crawling* data foto mahasiswa Fasilkom Unsri Angkatan 2018 dari laman situs UNSRI lama[8].

[4] https://www.dropbox.com/s/payjdk08292csra/celeba.zip?dl=0
[5] https://github.com/vipermu/ganimation
[6] http://mmlab.ie.cuhk.edu.hk/projects/CelebA.html
[7] https://www.kaggle.com/datasets/nikhilbartwal001/celeba
[8] https://old.unsri.ac.id/?act=daftar_mahasiswa

3.3 Tahapan Penelitian

Tahapan penelitian adalah rincian proses yang akan dilakukan pada penelitian. Tahapan penelitian yang akan dilakukan pada penelitian ini adalah sebagai berikut.

Gambar III-2. Diagram tahapan penelitian,

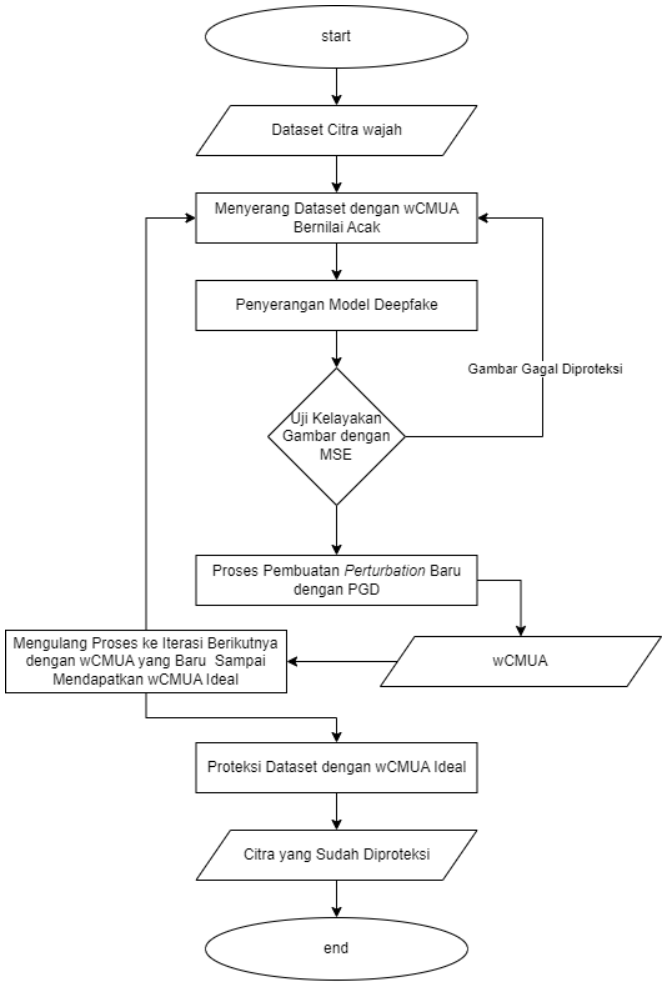
3.3.1 Mengumpulkan Data

pada tahapan ini akan dilakukan proses pengempulan dua data sekunder pertama citra gambar wajah seleb dari dataset Celeb-A, kedua citra gambar wajah mahasiswa Fasilkom Unsri Angkatan 2018. Dataset Celeb-A diambil dari link yang tertera pada repository di Github, proses

pengumpulan data dilakukan dengan menguduh secara langsung dataset tersebut berupa format .zip, sedangkan untuk citra gambar wajah mahasiswa Fasilkom Unsri Angkatan 2018 didapatkan melalui proses *crawing* secara langsung dari laman situs resmi laman kampus. Dataset ini memiliki nantinya harus berformat .jpg.

3.3.2 Menentukan Kerangka Kerja Penelitian

Kerangka kerja pada penelitian ini adalah sebagai berikut:



Gambar III-3. Diagram Alir Sistem Proteksi Citra dari deepfake dengan metode CMUA

Berdasarkan kerangka kerja penelitian pada Gambar III-3, sistem proteksi dari *deepfake* dengan metode CMUA memiliki alur sistem sebagai berikut:

1. **Input Dataset Citra Wajah**

   Proses penginputan dataset yang digunakan dalam sistem ini, yaitu berupa citra gambar wajah dari dataset Celeb-A yang digunakan sebagai data latih.

2. **Menyerang Dataset dengan wCMUA Bernilai Acak**

   Dalam tahapan ini, gambar akan di sisipkan dengan model proteksi awal,untuk menjadi awal pengujian model proteksi

3. **Penyerangan Model *Deepfake***

   Dalam tahapan ini, data gambar yang telah di sisipkan dengan model proteksi akan di serang dengan model deepfake tertentu.

4. **Uji Kelayakan Gambar dengan MSE**

   Pada tahapan ini gambar yang sebelumnya sudah diserang oleh model deepfake akan di uji dengan metode MSE, untuk mengetahui apakah model proteksi berhasil memproteksi atau tidak dari model deepfake yang digunakan. Jika gambar wajah yang di hasilkan cacat atau tidak dapat dikenali oleh Model MSE. Maka model proteksi berhasil begitu pula sebaliknya. Dan gambar yang berhasil akan di teruskan ke proses berikutnya sedangkan gagal yang gagal akan dicatat dan melanjutkan iterasi berikutnya

5. **Proses pengecekan PGD**

   Pada tahapan ini, gambar yang berhasil memproteksi akan dibandingkan dengan gambar asli sehingga memperoleh modification mask yang akan dijadikan sebagai model proteksi untuk iterasi berikutnya.

6. **Proses perulangan model sampai mendapatkan model proteksi yang ideal**

   Dalam tahapan ini akan terjadi perulang dimana model proteksi sebelumnya akan digunakan sebagai model proteksi awal untuk terasi berikutnya, sampai epochs yang ditentukan

7. **Didapatkan model proteksi ideal**

   Pada proses ini proteksi yang ideal akan langsung di implant pada semua data pada dataset.

8. **Didapatkan gambar yang telah di proteksi**

3.3.3 Menetukan Kriteria Pengujian

Untuk mengatasi masalah ini, kami memperkenalkan matriks *Mask*, yang lebih berkonsentrasi pada area yang dimodifikasi,

$$Mask_{(i,j)} = \begin{cases} 1, & if ||G(I)_{(i,j)} - I_{(i,j)}|| > 0.5, \\ 0, & else, \end{cases} \tag{10}$$

di mana (*i*, *j*) adalah koordinat piksel dalam gambar. Dengan cara ini, ketika menghitung L^2_{mask}, hanya piksel dengan perubahan besar yang akan dihitung dan area lainnya akan ditinggalkan,

$$L^2_{mask} = \frac{\sum_i \sum_j Mask_{(i,j)} \cdot ||G(I)_{(i,j)} - G(I + W_{CMUA})_{(i,j)}||}{\sum_i \sum_j Mask_{(i,j)}}. \tag{11}$$

Dalam eksperimen kami, jika $L^2_{mask} > 0{,}05$, kami menentukan bahwa gambar berhasil dilindungi, dan menggunakan SR_{mask} untuk merepresentasikan tingkat keberhasilan melindungi gambar wajah.

3.3.4 Menentukan Format Data Pengujian

Format data pengujian yang digunakan berupa table

- L^2_{mask} adalah perbandingan antara *original image* dengan *distorted* image
- SR_{mask} untuk merepresentasikan tingkat keberhasilan melindungi gambar wajah

dataset	$L^2_{mask}\uparrow$	$SR_{mask}\uparrow$
CelebA Training		
CelebA verifikasi		
Foto KPM mahasiswa Fasilkom Unsri 2018		

3.3.5 Menentukan Alat Bantu Penelitian

Alat bantu penelitian Proteksi Citra Foto KPM Mahasiswa Fasilkom UNSRI dari *DeepFake* dengan *CMUA-Watermark* yang akan digunakan dalam penelitian ini adalah sebagai berikut.

1. Perangkat Keras

   Proccessor : intel® core™ i7 8th *generation*

   RAM : 16GB RAM

   HDD : 1TB HDD strorage.

   VRAM : Nvidia Geforce GTX 1050 4GB RAM

2. Perangkat Lunak

   Sistem Operasi : Windows 10 64-bit

   Teks Editor : Visual Studio Code

Selain perangkat lunak diatas diatas.terdapat tambahan perangkat lunak ketiga

3.3.6 Membuat Kesimpulan

Setelah melakukan Analisa hasil pengujian, maka tahapan selanjutnya adalah pembuatan kesimpulan yang akan dimuat pada Bab V.

3.3.7 Melakukan Analisis dan Kesimpulan Hasil Pengujian Penetian

Analisi hasil pengujian dilakukan dengan memperhatikan nilai $L^2_{mask\ untuk}$ menentukan bahwa gambar berhasil dilindungi dan SR_{mask} untuk merepresentasikan tingkat keberhasilan melindungi gambar wajah. Serta perbandingan tingkat keberhasilan dari masing masing dataset yang diujikan.

3.4 Metode Pengembangan Perangkat Lunak

Metode pengembangan yang digunakan dalam penelitian ini adalah metode Rational Unified Process (RUP). Pengembangan sistem deteksi kemiripan kode sumber dibagi ke dalam empat tahap, yaitu fase insepsi, fase elaborasi, fase konstruksi dan fase transisi. Berikut merupakan tahapan pengembangan perangkat lunak yang akan dilakukan dalam tiap fasenya.

3.4.1 Face Insepsi

Tahapan yang akan dilakukan dalam fase ini adalah sebagai berikut.

1. Pemodelan Sistem : Menentukan ruang lingkup dan batasan masalah.
2. Kebutuhan : Mendefinisikan spesifikasi perangkat lunak.
3. Analisis dan Perancangan : Melakukan analisis terhadap kebutuhan perangkat lunak termasuk di dalamnya kebutuhan fungsional dan non fungsional dari spesifikasi perangkat lunak.
4. Implementasi : Membuat seluruh rancangan sistem ke dalam bentuk diagram use-case.

3.4.2 Fase Elaborasi

Tahapan yang akan dilakukan dalam fase ini adalah sebagai berikut.

1. Pemodelan Sistem: Membuat rancangan antarmuka (interface) sistem.
2. Kebutuhan: Menentukan spesifikasi dari sistem.
3. Analisis dan Perancangan: Membangun model *activity diagram* dan *sequence diagram* dari rancangan sistem.

4. Impelementasi: Membuat program berdasarkan diagram yang ditentukan sebelumnya.

3.4.3 Fase Konstruksi

Tahapan yang akan dilakukan dalam fase ini adalah sebagai berikut.

1. Pemodelan Bisnis : Menentukan bahasa pemrograman yang akan membangun sistem.
2. Kebutuhan : Menentukan kebutuhan sistem sesuai dengan fungsi yang telah ditentukan.
3. Analisis dan Perancangan : Membangun tampilan antar-muka sistem.
4. Implementasi: Membangun sistem dengan membuat program menggunakan bahasa pemrograman yang telah ditentukan.

3.4.4 Fase Transisi

Tahapan yang akan dilakukan dalam fase ini adalah sebagai berikut.

1. Pemodelan Sistem : Menentukan pengujian terhadap sistem.
2. Kebutuhan : Menentukan alat bantu pengujian terhadap sistem.
3. Analisis dan Perancangan : Merancang kasus penggunaan selama pengujian sistem.
4. Implementasi : Melaksanakan pengujian terhadap sistem menggunakan kasus penggunaan yang telah ditentukan.

3.5 Manajemen Proyek Perangkat Lunak

Rencana manajemen proyek penelitian merupakan perencanaan aktivitas penelitian dari tahap awal hingga selesai. Perencanaan aktivitas pada penelitian ini akan menggunakan *Gantt Chart* seperti pada Tabel III-4.

No.	Uraian Kegiatana	Tahun 2023 bulan ke-					
		1	2	3	4	5	6
1	**Melakukan Pengumpulan Data**						
a	Mengumpulkan data						
b	Melakukan pra-pengolahan data						
c	Melakukan modifikasi data sesuai skenario						
d	Tersedia dokumen hasil tahapan penelitian						
2	**Rekayasa Perangkat Lunak**						
2.1	**Insepsi**						
a	Menentukan pemodelan bisnis						
b	Menentukan kebutuhan pengguna						
c	Menentukan kebutuhan sistem						
2.2	**Elaborasi**						
a	Menentukan spesifikasi sistem						
b	Membangun model *acitvity diagram* dan *sequence diagram*						
c	Membangun rancangan tampilan antar-muka						
2.3	**Konstruksi**						
a	Membangun Model *Class Diagram*						
b	Membangun Sistem (impelementasi kode)						
c	Perbaikan Sistem						
2.4	**Transisi**						
a	Melakukan pengujian awal terhadap sistem						
b	Tersedia dokumen hasil tahapan penelitian						
3	**Melakukan Pengujian Penelitian Terhadap Sistem**						

a	Membuat rancangan hasil pengujian dalam penelitian						
b	Melakukan pengujian final terhadap sistem						
c	Tersedia dokumen hasil penelitian						
4	**Melakukan Analisis dan Kesimpulan dari Hasil Pengujian**						
a	Melakukan analisis terhadap hasil pengujian penelitian						
b	Membuat kesimpulan dan saran terhadap hasil pengujian penelitian						
c	Tersedia dokumen hasil penelitian						

Tabel Rencana Manajemen Proyek Penelitian

3.6 Kesimpulan

Pada bab ini telah dibahas tentang proses pengumpulan data yang digunakan sebagai bahan uji perangkat lunak, tahapan penelitian, metode pengembangan perangkat lunak yang akan digunakan serta kriteria pengujian penelitian yang akan dilakukan terhadap sistem.

DAFTAR PUSTAKA

Aldwairi, M., & Alwahedi, A. (2018). Detecting fake news in social media networks. *Procedia Computer Science*, *141*, 215–222. https://doi.org/10.1016/j.procs.2018.10.171

Anderson, K. E. (2018). Getting acquainted with social networks and apps: combating fake news on social media. *Library Hi Tech News*, *35*(3), 1–6. https://doi.org/10.1108/LHTN-02-2018-0010

Anwar, A. (2014). A Review of RUP (Rational Unified Process). *International Journal of Software Engineering*, *5*(2), 8–24. http://www.cscjournals.org/library/manuscriptinfo.php?mc=IJSE-142

Borges, L., Martins, B., & Calado, P. (2019). Combining similarity features and deep representation learning for stance detection in the context of checking fake news. *Journal of Data and Information Quality*, *11*(3). https://doi.org/10.1145/3287763

Britt, M. A., Rouet, J. F., Blaum, D., & Millis, K. (2019). A Reasoned Approach to Dealing With Fake News. *Policy Insights from the Behavioral and Brain Sciences*, *6*(1), 94–101. https://doi.org/10.1177/2372732218814855

Day, C. (2019). The Future of Misinformation. *Computing in Science and Engineering*, *21*(1), 108. https://doi.org/10.1109/MCSE.2018.2874117

Figueira, Á., & Oliveira, L. (2017). The current state of fake news: Challenges and opportunities. *Procedia Computer Science*, *121*, 817–825. https://doi.org/10.1016/j.procs.2017.11.106

Fletcher, J. (2018). Deepfakes, artificial intelligence, and some kind of dystopia: The new faces of online post-fact performance. *Theatre Journal*, *70*(4), 455–471. https://doi.org/10.1353/tj.2018.0097

Huang, H., Wang, Y., Chen, Z., Li, Y., Tang, Z., Chu, W., Chen, J., Lin, W., & Ma, K.-K. (2021). *CMUA-Watermark: A Cross-Model Universal Adversarial Watermark for Combating Deepfakes*. http://arxiv.org/abs/2105.10872

Jang, S. M., & Kim, J. K. (2018). Third person effects of fake news: Fake news regulation and media literacy interventions. *Computers in Human Behavior*, *80*, 295–302. https://doi.org/10.1016/j.chb.2017.11.034

Liu, Z., Luo, P., Wang, X., & Tang, X. (2015). Deep learning face attributes in the wild. *Proceedings of the IEEE International Conference on Computer Vision*, *2015 Inter*, 3730–3738. https://doi.org/10.1109/ICCV.2015.425

McAndrew, A. (2014). An Introduction to Digital Image Processing with Matlab, Notes for SCM2511 Image Processing 1. *Jurnal Ilmiah Elite Elektro*, *2*(2), 83–87.

Metz, R. (2019, June 12). *The fight to stay ahead of deepfake videos before the 2020 US election*. CNN BUSINESS. https://edition.cnn.com/2019/06/12/tech/deepfake-2020-detection/index.html

Metz, R., & O'Sullivan, D. (2019). A deepfake video of Mark Zuckerberg presents a new challenge for Facebook. *CNN BUSINESS*. https://edition.cnn.com/2019/06/11/tech/zuckerberg-deepfake/index.html

Qayyum, A., Qadir, J., Janjua, M. U., & Sher, F. (2019). Using Blockchain to Rein in the New Post-Truth World and Check the Spread of Fake News. *IT Professional*, *21*(4), 16–24. https://doi.org/10.1109/MITP.2019.2910503

Ruiz, N., Bargal, S. A., & Sclaroff, S. (2020). Disrupting Deepfakes: Adversarial Attacks Against Conditional Image Translation Networks and Facial Manipulation Systems. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12538 LNCS*, 236–251. https://doi.org/10.1007/978-3-030-66823-5_14

Sun, P., Li, Y., Qi, H., & Lyu, S. (2020). Landmark Breaker: Obstructing DeepFake by Disturbing Landmark Extraction. *2020 IEEE International Workshop on Information Forensics and Security, WIFS 2020*, 6–11. https://doi.org/10.1109/WIFS49906.2020.9360910

Young, I. T., Gerbrands, J. J., Vliet, L. J. van, Theodore, I., Jacob, J., Vliet, V., & Jozef, L. (2006). Fundamentals of image-processing. *Seimitsu Kogaku Kaishi/Journal of the Japan Society for Precision Engineering*, *72*(5), 583–586. https://doi.org/10.2493/jjspe.72.583

Zannettou, S., Sirivianos, M., Blackburn, J., & Kourtellis, N. (2019). The web of false information: Rumors, fake news, hoaxes, clickbait, and various other shenanigans. *Journal of Data and Information Quality*, *11*(3). https://doi.org/10.1145/3309699