

**PENGEMBANGAN *DATASET* UNTUK DETEKSI  
*DEEPPFAKE* PADA VIDEO KUALITAS RENDAH DI  
INDONESIA**

**Proposal Tugas Akhir**

Oleh

**Harry Truman Suhalim  
18222081**



**PROGRAM STUDI SISTEM DAN TEKNOLOGI INFORMASI  
SEKOLAH TEKNIK ELEKTRO DAN INFORMATIKA  
INSTITUT TEKNOLOGI BANDUNG  
Desember 2025**

# **LEMBAR PENGESAHAN**

## **PENGEMBANGAN *DATASET* UNTUK DETEKSI *DEEPFAKE* PADA VIDEO KUALITAS RENDAH DI INDONESIA**

### **Proposal Tugas Akhir**

Oleh

**Harry Truman Suhalim**  
**1822081**

Program Studi Sistem dan Teknologi Informasi  
Sekolah Teknik Elektro dan Informatika  
Institut Teknologi Bandung

Proposal Tugas Akhir ini telah disetujui dan disahkan  
di Bandung, pada tanggal 04 Desember 2025

Pembimbing 1

Pembimbing 2

Dr. Fetty Fitriyanti Lubis, S.T., M.T.

NIP. 118110071

Dion Tanjung, S.Kom., M.Sc., Ph.D.

NIP. 124110057

## DAFTAR ISI

<b>DAFTAR GAMBAR . . . . .</b>	<b>iv</b>
<b>DAFTAR TABEL . . . . .</b>	<b>v</b>
<b>I PENDAHULUAN . . . . .</b>	<b>1</b>
I.1 Latar Belakang . . . . .	1
I.2 Rumusan Masalah . . . . .	3
I.3 Tujuan . . . . .	4
I.4 Batasan Masalah . . . . .	4
I.5 Metodologi . . . . .	5
<b>II STUDI LITERATUR . . . . .</b>	<b>8</b>
II.1 <i>Data-Centric Artificial Intelligence</i> . . . . .	8
II.2 Pendekatan-Pendekatan dalam Pengembangan <i>Dataset</i> . . . . .	9
II.2.1 Pendekatan <i>Crowdsourcing</i> . . . . .	10
II.2.2 <i>Human-Centric Approach</i> . . . . .	11
II.3 Metodologi Pengembangan <i>Dataset</i> . . . . .	12
II.4 Kelompok Suku Bangsa Indonesia . . . . .	14
II.5 Klasifikasi Deepfake . . . . .	15
II.5.1 <i>Visual Deepfake</i> . . . . .	16
II.5.2 <i>Audio Deepfake</i> . . . . .	17
II.5.3 <i>Multimodal Deepfake</i> . . . . .	18
II.6 <i>Dataset</i> Terdahulu . . . . .	18
II.7 <i>Deepfake</i> Pada Dunia Nyata . . . . .	20
II.8 <i>Benchmarking</i> FaceForensics++ . . . . .	21
II.9 Teknik Kompresi Video . . . . .	22
II.10 Kajian Pengembangan <i>Dataset</i> dan Benchmarking untuk Algoritma De- teksi . . . . .	23
II.10.1 Kategorisasi <i>Dataset</i> Berdasarkan Teknik Manipulasi . . . . .	23
II.10.2 Metodologi Pengembangan <i>Dataset Benchmark</i> . . . . .	25
II.11 Keterbatasan <i>Dataset Benchmark</i> dalam Skenario Dunia Nyata . . . . .	26
<b>III ANALISIS MASALAH . . . . .</b>	<b>28</b>
III.1 Analisis Kondisi Saat Ini . . . . .	28
III.2 Analisis Kebutuhan . . . . .	29
III.2.1 Identifikasi Masalah Pengguna . . . . .	29

III.2.2	Kebutuhan Fungsional . . . . .	30
III.2.3	Kebutuhan Nonfungsional . . . . .	30
III.3	Analisis Pemilihan Solusi . . . . .	31
III.4	Analisis Pemilihan Solusi . . . . .	31
III.4.1	Penjabaran Alternatif Solusi . . . . .	31
III.4.1.1	Alternatif Solusi 1: Kompresi <i>Dataset</i> Publik yang Ada . . . . .	31
III.4.1.2	Alternatif Solusi 2: Pengumpulan Pasif <i>Deepfake</i> Beredar . . . . .	32
III.4.1.3	Alternatif Solusi 3: Pengembangan <i>Dataset End-to-End</i> Konteks Lokal . . . . .	33
III.4.2	Analisis Penentuan Solusi . . . . .	34
<b>IV</b>	<b>DESAIN KONSEP SOLUSI . . . . .</b>	<b>35</b>
IV.1	Desain Konsep Solusi . . . . .	35
<b>V</b>	<b>RENCANA SELANJUTNYA . . . . .</b>	<b>38</b>
V.1	Rencana Implementasi . . . . .	38
V.1.1	Fase Pengumpulan <i>Collection</i> . . . . .	38
V.1.2	Fase Generasi dan Pemrosesan ( <i>Construction</i> ) . . . . .	41
V.1.3	Fase Pelabelan( <i>Labelling</i> ) . . . . .	41
V.1.4	Fase Validasi ( <i>Transition</i> ) . . . . .	42
V.2	Analisis Risiko dan Mitigasi . . . . .	43

## DAFTAR GAMBAR

II.1	Pilar Utama <i>Data-Centric AI</i> (Zha dkk. 2023) . . . . .	8
II.2	Jenis-Jenis <i>Deepfake</i> (Croitoru dkk. 2024) . . . . .	16
II.3	Perbedaan Kualitas Video Rendah dan Kualitas Video Tinggi (BarTech TV 2022) . . . . .	23
III.1	Model Konseptual Alur Deteksi <i>Deepfake</i> Saat Ini . . . . .	28
IV.1	Model Konseptual Solusi yang Diusulkan . . . . .	35
V.1	Fase Pengumpulan <i>Collection</i> . . . . .	38
V.2	Fase Generasi dan Pemrosesan ( <i>Construction</i> ) . . . . .	41
V.3	Fase Pelabelan ( <i>Labelling</i> ) . . . . .	42
V.4	Fase Validasi ( <i>Transition</i> ) . . . . .	43

## DAFTAR TABEL

II.1	Komposisi Penduduk Indonesia Berdasarkan Suku Bangsa (Badan Pusat Statistik 2012) . . . . .	14
II.2	Pengelompokan 2 suku Terbesar Pada Setiap Pulau di Indonesia Bersama Kelompok Representatif Berdasarkan Sensus Penduduk 2010 . . . . .	15
II.3	Informasi Fundamental <i>Dataset</i> Deteksi Video <i>Deepfake</i> Terdahulu (Sohan, Solaiman, dan Hasan 2023) . . . . .	19
III.1	Kebutuhan Fungsional <i>Dataset</i> . . . . .	30
III.2	Kebutuhan Non Fungsional <i>Dataset</i> . . . . .	31
III.3	Analisis Penentuan Solusi <i>Dataset</i> . . . . .	34
IV.1	Perbandingan Solusi dengan Kondisi Saat Ini . . . . .	36
V.1	Analisis Risiko dan Mitigasi . . . . .	44

# BAB I

## PENDAHULUAN

### I.1 Latar Belakang

Perkembangan teknologi *Artificial Intelligence* menimbulkan cara baru untuk menciptakan berita hoaks melalui *deepfake*. *Deepfake* adalah teknik untuk menempatkan gambar wajah orang “sebenarnya” dalam suatu video menjadi wajah target sehingga seolah-olah target tersebut melakukan atau mengatakan hal-hal yang dilakukan orang “sebenarnya”. Menurut laporan dari Sensity AI, terjadi lonjakan kasus *deepfake* sebesar 550% sejak tahun 2019 (Kementerian Komunikasi dan Digital 2025), didapatkan bahwa 90% di antaranya digunakan untuk menyebarkan disinformasi, melakukan penipuan, pencemaran nama baik, hingga pelecehan yang menargetkan kelompok rentan seperti perempuan dan anak-anak. Di Indonesia sendiri, *deepfake* sering kali digunakan untuk memprovokasi sentimen negatif dari publik. Fenomena ini menunjukkan urgensi dari pengembangan teknologi untuk dapat mendeteksi *deepfake* secara efektif.

Salah satu tantangan dalam pendeteksian *deepfake* adalah ketika suatu video *deepfake* muncul, video tersebut akan diunggah dan disebarluaskan melalui berbagai *platform* media sosial. Proses kompresi yang dilakukan pada media sosial seperti YouTube, TikTok, WhatsApp, dan lain sebagainya akan menyebabkan kualitas video tersebut menurun. Kualitas video *deepfake* yang rendah akan semakin sulit untuk dideteksi, baik oleh manusia maupun oleh teknologi AI. Akibatnya, banyak model deteksi yang menunjukkan akurasi tinggi di lingkungan laboratorium (mencapai hampir 100%) mengalami penurunan kinerja drastis hingga hanya sekitar 65% saat dihadapkan pada skenario dunia nyata dengan video terkompresi (Alrashoud 2025). Masalah ini menjadi semakin relevan di Indonesia, terdapat disparitas kecepatan internet seluler yang signifikan antar regional. Studi menunjukkan bahwa Jawa memiliki kecepatan internet tertinggi sementara Papua menunjukkan kecepatan

an terendah, yang dipengaruhi oleh infrastruktur telekomunikasi yang tidak merata dan kondisi geografis (Anugrah dkk. 2025). Kualitas konektivitas yang rendah ini seringkali memaksa kompresi video yang lebih agresif agar konten dapat diakses dan dibagikan, yang secara langsung memperparah kesulitan dalam mendeteksi artefak deepfake pada video berkualitas rendah.

Berbagai pendekatan deteksi canggih sebenarnya telah dikembangkan untuk mengatasi manipulasi ini. Solusi tersebut meliputi analisis artefak spasial pada wajah menggunakan *Convolutional Neural Networks* (CNNs) (Li dan Lyu 2019), analisis inkonsistensi temporal antar-frame seperti kedipan mata yang tidak natural menggunakan *Recurrent Neural Networks* (RNNs) (Fernandes dan Fatma 2025; Alrashoud 2025), hingga metode deteksi sinyal biologis seperti detak jantung (rPPG) yang sulit ditiru generator *deepfake* (Ciftci, Demir, dan Yin 2020). Namun, efektivitas seluruh metode ini sangat bergantung pada kualitas data yang digunakan untuk melatihnya. Metode-metode tersebut sering kali gagal ketika diuji pada kondisi *out-of-distribution*, yaitu ketika data uji (video media sosial terkompresi) memiliki karakteristik yang jauh berbeda dari data latih (video berkualitas tinggi). Tanpa adanya data yang merepresentasikan degradasi kualitas ini, algoritma secanggih apa pun tidak akan optimal dalam skenario dunia nyata.

Keterbatasan ini menyoroti pergeseran paradigma penting dalam pengembangan AI saat ini, yaitu dari pendekatan *model-centric* (fokus pada perbaikan arsitektur) menuju pendekatan *data-centric* (fokus pada kualitas data). Menurut (Zha dkk. 2023), peningkatan kualitas dan representasi data sering kali memberikan dampak kinerja yang lebih signifikan dibanding sekadar memodifikasi model. Dalam konteks ini, *dataset engineering* bukan lagi sekadar pengumpulan data pasif, melainkan proses desain sistematis untuk memastikan cakupan fenomena sains terpenuhi (Vital Brazil dkk. 2023). Oleh karena itu, *bottleneck* deteksi deepfake di Indonesia saat ini bukanlah ketiadaan model deteksi, melainkan ketiadaan *dataset* yang representatif—baik dari segi demografi wajah Indonesia maupun karakteristik kompresi jaringan lokal.

Kesenjangan ketersediaan data ini terlihat jelas pada *dataset benchmark* yang ada. *Dataset* populer seperti yang dirangkum oleh (Sohan, Solaiman, dan Hasan 2023), mayoritas berisi video resolusi tinggi dengan subjek wajah non-Indonesia, sehingga validitasnya dipertanyakan untuk mendeteksi konten lokal (Banerjee dkk. 2024) grupner2025. Penelitian ini diusulkan untuk mengisi kekosongan tersebut dengan mengembangkan *dataset* baru yang secara spesifik dirancang untuk mencakup dua tantangan utama yaitu konteks demografis Indonesia dan simulasi artefak kompresi



video kualitas rendah yang realistis.

Penelitian ini didasarkan pada analisis mendalam terhadap metodologi yang digunakan oleh *benchmark* global, keberhasilan mereka terletak pada kontrol penuh atas *pipeline* data. Misalnya, FaceForensics++ yang merupakan salah satu dataset paling masif yang dibangun secara sistematis melalui tiga langkah terkontrol yaitu Pengumpulan (dengan kriteria video resolusi tinggi), Manipulasi (menggunakan empat teknik *swapping* dan *reenactment* terkemuka), dan Kompresi (*Post-processing*) yang menghasilkan tiga tingkat kualitas video (*Raw*, *High-Quality*, *Low-Quality*) untuk evaluasi yang terukur. Pendekatan ini adalah contoh sempurna dari metodologi *data-centric* yang menjamin *ground truth* dan memungkinkan *slicing* data berdasarkan artefak. Namun, seperti banyak *dataset* terdahulu lainnya, FaceForensics++ dan Celeb-DFv2 menggunakan subjek wajah non-Indonesia dan tidak secara akurat mereplikasi kompresi agresif dari media sosial yang relevan dengan konteks lokal. Oleh karena itu, penelitian ini mengadopsi struktur metodologis *end-to-end* yang sama yaitu mengontrol Pengumpulan, Generasi, dan Simulasi Kompresi. Namun, menerapkan hal tersebut pada konten lokal Indonesia dan teknik simulasi kompresi yang lebih realistis, untuk mengatasi kesenjangan *Out-of-Distribution* yang disorot oleh (Montibeller dkk. 2025). Pendekatan ini memastikan *dataset* yang dihasilkan valid secara ilmiah dan relevan dengan ancaman disinformasi di Indonesia.

## **I.2 Rumusan Masalah**

Masalah utama yang ingin diselesaikan dalam tugas akhir ini adalah belum adanya model deteksi *deepfake* yang andal dan teruji untuk menangani video berkualitas rendah dalam konteks Indonesia. Mayoritas solusi yang ada saat ini dilatih menggunakan *dataset* video resolusi tinggi dengan subjek non-Indonesia, sehingga performanya menurun drastis saat dihadapkan pada video terkompresi dan gagal mengenali karakteristik demografis lokal. Jika tidak diatasi, masyarakat akan semakin rentan terhadap ancaman disinformasi dan penipuan berbasis *deepfake*. Dalam menjawab permasalahan tersebut, penelitian ini akan berfokus pada pembuatan dataset yang representatif dan pengembangan model deteksi yang robust. Berikut adalah beberapa rumusan masalah spesifik yang akan diselesaikan

1. Bagaimana metodologi yang efektif untuk mengumpulkan dan menyeleksi video asli berkualitas tinggi yang beragam dari sumber lokal Indonesia sebagai bahan dasar pembuatan *dataset*, dengan memperhatikan keragaman fenotipe (visual) dan logat (audio) subjek?
2. Apa saja jenis-jenis manipulasi (*deepfake*) (visual, audio, dan multimodal)

yang perlu diterapkan pada video asli agar hasilnya relevan dengan ancaman disinformasi di Indonesia, dan bagaimana proses penerapannya dapat dikontrol untuk menjamin *ground truth*?

3. Bagaimana cara mensimulasikan efek kompresi media sosial (*lossy compression*) secara sistematis, terukur, dan realistis untuk menghasilkan video berkualitas rendah yang mengandung artefak forensik khas, agar model deteksi dapat bergeneralisasi di skenario dunia nyata?
4. Bagaimana cara melakukan validasi internal dan eksternal terhadap *dataset* yang akan dibuat untuk menjamin validitas label 100% dan mencapai keseimbangan yang wajar antara kelas video (asli/palsu) dan kualitas video (tinggi/rendah)?

### I.3 Tujuan

Penelitian ini bertujuan untuk mengembangkan *dataset* dalam konteks Indonesia guna mendukung deteksi video *deepfake*. *Dataset* disusun melalui pembuatan sampel video *deepfake* yang mereplikasi karakteristik dan kualitas visual video yang beredar di media sosial.

### I.4 Batasan Masalah

Berikut merupakan beberapa batasan yang ditetapkan untuk memfokuskan ruang lingkup dan memastikan hasil dari tugas akhir ini relevan dengan tujuan yang ditetapkan.

1. Tugas akhir ini dikerjakan secara berkelompok yang terdiri dari 3 orang mahasiswa, yaitu Alvin Fadhillah Akmal (NIM 18222079), Harry Truman Suhalim (NIM 18222081), dan Steven Adrian Corne (NIM 18222101). Adapun pembagian fokus dari ketiga mahasiswa tersebut ialah Harry fokus pada *dataset deepfake*, Steven fokus pada model untuk mendeteksi *deepfake*, dan Alvin fokus pada *browser extension*.
2. Penelitian akan menggunakan *dataset* video yang dirancang untuk merepresentasikan konteks Indonesia. Ini mencakup pengumpulan video asli yang menampilkan wajah orang Indonesia dan pembuatan video *deepfake* dari bahan tersebut.
3. *Dataset* ini dikembangkan hanya untuk kepentingan pendidikan, penelitian, dan penulisan karya ilmiah. Video asli dikumpulkan berdasarkan prinsip *Fair Use* dan Pasal 43 (a) UU RI No. 28 Tahun 2014, sehingga penggunaannya non-komersial dan tidak merugikan kepentingan Pencipta secara wajar.

4. Proses generasi *deepfake* dilakukan secara internal dan terkontrol hanya untuk tujuan penelitian. *Dataset* tidak akan memasukkan *deepfake* dengan konten sensitif atau yang menargetkan individu secara berbahaya (seperti pelecehan atau penipuan) untuk menghindari potensi penyalahgunaan.
5. *Dataset* yang dikembangkan dalam penelitian ini secara khusus berfokus pada konteks lokal Indonesia. Populasi subjek yang digunakan dalam *dataset* ini dibatasi pada perwakilan dari 13 kelompok suku bangsa utama di Indonesia untuk menjamin representasi visual dan audio yang relevan dengan keragaman lokal. *Dataset* ini ditujukan untuk skenario media sosial dan video berkualitas rendah, sehingga batasan kualitas video adalah pada resolusi dan simulasi kompresi yang meniru *platform in-the-wild*.
6. Pengembangan *dataset* dilakukan dengan mematuhi standar etika dan regulasi yang berlaku (UU Perlindungan Data Pribadi), di mana sumber data dibatasi pada video domain publik (*YouTube*) sesuai prinsip *Fair Use*, proses manipulasi dilakukan secara terkontrol hanya untuk tujuan penelitian defensif, serta distribusi *dataset* yang dibatasi untuk kalangan terbatas guna mencegah penyalahgunaan.

## I.5 Metodologi

Tahapan yang akan dilalui selama pelaksanaan tugas akhir ini terdiri dari tiga bagian, yaitu:

1. Perumusan masalah awal

Pada tahapan ini, dilakukan identifikasi masalah awal dengan cara *brainstorming* bersama dengan seluruh anggota kelompok tugas akhir untuk menentukan masalah yang akan diangkat menjadi topik tugas akhir. Bermula dari pengamatan salah seorang anggota kelompok, terkait sulit membedakan berita mana yang hoax dan berita mana yang asli. Seluruh anggota kelompok mengalami hal yang sama, dikarenakan berita hoax terlalu luas maka kami menyepakati untuk mendeteksi hoax berupa *deepfake*. Untuk mempelajari tentang *deepfake* terutama dalam konteks hoax kami melakukan riset lebih lanjut. Akhirnya ditemukan data dan fakta bahwa *deepfake* saat ini sering kali digunakan untuk menyebarkan hoax dan ditemukan juga beberapa jurnal yang membahas *deepfake*. Data dan informasi yang didapat dari jurnal tersebut akhirnya memperkuat masalah yang ingin kami selesaikan dikarenakan dalam beberapa jurnal disebutkan bahwa model deteksi *deepfake* yang mereka miliki belum dapat mendeteksi video dengan kualitas rendah dan juga model kurang relevan digunakan apabila konteks *deepfake* di Indonesia.

## 2. Pencarian dan Penulisan Studi Literatur

Penentuan teori-teori yang dirasa perlu dalam penulisan tugas akhir ini dilakukan dengan mengidentifikasi hal-hal seputar fokus yang dibahas oleh setiap anggota kelompok tugas akhir. Tugas akhir ini fokus pada dataset sehingga teori-teori yang dibutuhkan adalah teori seputar metodologi pembuatan *dataset deepfake* publik, teknik generasi *deepfake*, serta dampak dan simulasi dari kompresi video. Metodologi pembuatan *dataset* sendiri dipilih dengan tujuan untuk menyamakan *dataset* dengan konten lokal yang ada di Indonesia sehingga lebih relevan. Selain itu, dapat juga disesuaikan kualitas dari video untuk menyesuaikan dengan kondisi video disebar luaskan di sosial media. Adapun penelusuran pustaka lebih dalam terkait setiap teori yang diperlukan dilakukan melalui internet. Untuk setiap teori, dilakukan pencarian terhadap jurnal-jurnal yang membahas spesifik terkait teori tersebut. Kata kunci pencarian yang digunakan antara lain "*deepfake dataset creation*", "*face forensics*", "*video compression simulation*", dan "*generative adversarial networks for face synthesis*". Untuk setiap area, proses dimulai dengan mencari survey paper guna mendapatkan gambaran umum, lalu dilanjutkan dengan penelusuran mendalam terhadap artikel-artikel kunci yang relevan untuk dipelajari metodologinya secara detail.

## 3. Pengembangan *Dataset*

Pengembangan *dataset* dibuat secara manual dengan membuat *deepfake* sendiri maupun mengambil *deepfake* yang ada di sosial media. Metodologi dalam pengembangan dataset dibagi menjadi empat fase yaitu:

### a. Fase Pengumpulan (*Collection*)

Pada fase ini, dilakukan pengumpulan bahan mentah untuk *dataset*. Fokus utama adalah mendapatkan data video asli yang merepresentasikan konten lokal Indonesia. Kebutuhan awal, seperti jumlah video, durasi, dan variasi subjek, ditetapkan untuk memastikan *dataset* yang dihasilkan cukup beragam. Video terkait tokoh publik, *vlogger*, dan individu lainnya pada sosial media diunduh dengan kualitas tinggi, video *deepfake* yang tersebar pada sosial media juga dikumpulkan.

### b. Fase Generasi dan Pemrosesan (*Construction*)

Setelah bahan mentah terkumpul, dilakukan proses pembuatan data *deepfake*. Fase ini diawali dengan penggunaan video asli untuk menghasilkan video manipulasi menggunakan beberapa teknik, seperti *face swapping* dan *lip-sync*, dengan perangkat lunak open-source untuk menciptakan variasi artefak. Setelah itu, dilakukan simulasi kompresi menggunakan *tools* seper-

ti FFmpeg untuk menurunkan kualitas video secara terkontrol, sehingga menghasilkan versi berkualitas rendah yang realistis.

c. Fase Pelabelan (*Labeling*)

Pada fase ini, setiap data yang telah diproses akan diberi label informatif untuk memberikan konteks bagi model *machine learning*. Setiap video akan diklasifikasikan dengan label utama (asli atau palsu). Selanjutnya, sebuah file metadata akan dibuat untuk mencatat informasi yang lebih detail untuk setiap data, seperti ID video, label utama, jenis manipulasi yang digunakan (jika palsu), dan tingkat kompresi yang diterapkan.

d. Fase Validasi (*Transition*)

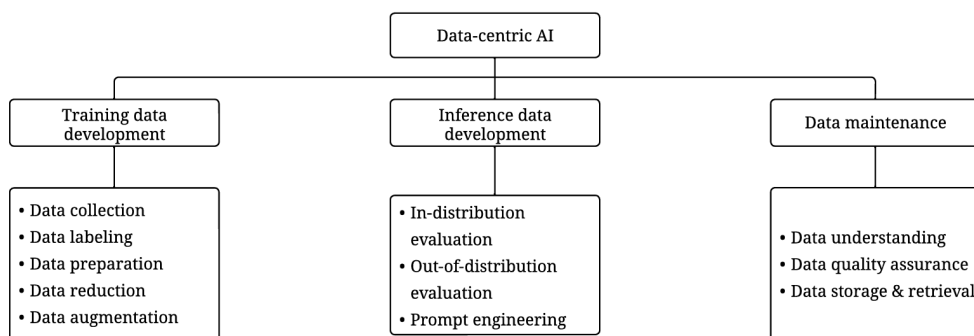
Setelah pelabelan selesai, *dataset* divalidasi untuk memastikan kualitas dan integritasnya. Dilakukan pemeriksaan silang secara acak terhadap sampel data untuk memastikan pelabelan sudah akurat. Selain itu, diperiksa keseimbangan antara kelas asli dan palsu serta antara data berkualitas tinggi dan rendah. Tujuannya adalah digunakan untuk melakukan perbaikan minor pada *dataset* sebelum digunakan untuk pelatihan model. Validasi teknis dilakukan melalui pengujian performa model, di mana akurasi model yang dilatih menggunakan *dataset* ini dibandingkan secara komparatif dengan model yang dilatih pada *dataset benchmark* untuk mengukur efektivitas dan reliabilitas *dataset* yang dikembangkan.

## BAB II

### STUDI LITERATUR

#### II.1 *Data-Centric Artificial Intelligence*

*Data-centric artificial intelligence* adalah disiplin yang diterapkan untuk merencanakan data yang digunakan dalam membangun sistem AI secara sistematis. Ini merupakan perubahan dari yang sebelumnya yang berupa *model-centric*, peneliti fokus mengubah model sambil membiarkan data tetap. Keberhasilan dalam AI tentu bergantung pada kualitas data yang dimiliki, dalam konteks *deepfake* terutama di Indonesia, sumber data yang relevan tentu dibutuhkan pula. Kualitas data sangat memengaruhi kinerja model dalam mendeteksi *deepfake* pula, model tentu sangat penting dalam mendeteksi *deepfake* tetapi *dataset* yang baik juga tidak kalah penting. Oleh karena itu, dibutuhkan *dataset* yang baik untuk mendukung pembentukan model yang baik pula.



Gambar II.1 Pilar Utama *Data-Centric AI* (Zha dkk. 2023)

Berdasarkan Gambar II.1, terdapat tiga pilar utama dari *data-centric AI*. Pada penelitian ini, terdapat dua pilar utama dari *data-centric AI* yang difokuskan yaitu Pengembangan Data Pelatihan (*Training Data Development*) dan Pengembangan Data Inferensi (*Inference Data Development*). Tahapan dalam Pengembangan Data Pelatihan (*Training Data Development*) terdiri dari :

1. Pengumpulan Data (*Data Collection*)

Pengumpulan data adalah proses mengumpulkan dan memperoleh data dari berbagai sumber. Proses ini secara fundamental menentukan kualitas dan kuantitas data. Proses ini sangat bergantung pada *domain knowledge* untuk memastikan data yang dikumpulkan relevan, representatif, dan selaras dengan tujuan pemangku kepentingan.

2. Pelabelan Data (*Data Labeling*)

Pelabelan data adalah proses memberikan satu atau lebih *tag* atau label deskriptif ke *dataset*. Proses ini memungkinkan algoritma untuk belajar dari data dan membuat prediksi. Pelabelan memainkan peran krusial dalam memastikan bahwa model yang dilatih pada data tersebut secara akurat mencerminkan intensi atau harapan manusia.

3. Augmentasi Data (*Data Augmentation*)

Augmentasi data adalah teknik untuk meningkatkan ukuran (*size*) dan keragaman (*diversity*) data dengan cara membuat variasi buatan dari data yang ada. Teknik ini seringkali dapat meningkatkan kinerja model. Augmentasi data sangat penting karena algoritma pembelajaran mesin modern, terutama *deep learning*, seringkali membutuhkan data dalam jumlah besar untuk belajar secara efektif.

Sedangkan, tahapan dari Pengembangan Data Inferensi (*Inference Data Development*) terdiri dari :

1. Evaluasi *Out-of-Distribution*

Evaluasi *Out-of-Distribution* mengacu pada penggunaan sekumpulan sampel data yang mengikuti distribusi yang berbeda dari yang diamati dalam data pelatihan.

2. Pemilahan Data (*Data Slicing*)

*Data slicing* adalah bagian dari evaluasi *in-distribution*.

Ini didefinisikan sebagai proses mempartisi membagi *dataset* menjadi sub-populasi yang relevan dan mengevaluasi kinerja model pada setiap sub-populasi tersebut secara terpisah (Zha dkk. 2023).

## II.2 Pendekatan-Pendekatan dalam Pengembangan *Dataset*

Dalam pengembangan *dataset*, terdapat berbagai pendekatan metodologis yang dapat diterapkan sesuai dengan karakteristik masalah dan ketersediaan sumber daya. Pemilihan pendekatan yang tepat sangat krusial karena akan menentukan kualitas, skala, dan efisiensi biaya dari *dataset* yang dihasilkan. Berikut ini adalah uraian

mengenai beberapa pendekatan utama tersebut.

### II.2.1 Pendekatan *Crowdsourcing*

Pendekatan *Crowdsourcing* merupakan solusi efektif ketika menghadapi skenario ketiadaan data pelatihan yang masif atau tugas pelabelan yang terlalu kompleks bagi satu individu. Pendekatan ini berfokus pada dekomposisi tugas besar menjadi tugas-tugas mikro yang sederhana agar dapat dikerjakan oleh pekerja non-ahli dalam jumlah banyak. Contoh penerapan metode ini terlihat pada proyek 'In Codice Ratio' (Firmani dkk. 2020) yang melibatkan siswa sekolah untuk mentranskripsi manuskrip kuno Vatikan.

Secara metodologis, pendekatan ini dimulai dengan identifikasi masalah untuk memetakan batasan pekerja non-ahli. Selanjutnya, dilakukan pemrosesan awal otomatis oleh sistem untuk memecah data menjadi segmen kecil. Kunci utamanya adalah dekomposisi tugas, di mana pekerja tidak diminta melakukan analisis kompleks, melainkan hanya pencocokan pola sederhana. Untuk menjamin kualitas data, metode ini menerapkan sistem redundansi, di mana satu label baru dianggap valid jika disetujui oleh mayoritas pekerja.

Kelebihan utama dari pendekatan *crowdsourcing* adalah kemampuannya untuk menyelesaikan tugas pembuatan *dataset* berskala besar yang sebelumnya dianggap mustahil karena kendala biaya atau waktu, serta menghasilkan data yang sangat granular. Namun, pendekatan ini memiliki kelemahan pada ketergantungannya yang tinggi terhadap kualitas pra-pemrosesan otomatis. Jika segmentasi awal gagal, pekerja tidak dapat memberikan label yang benar. Selain itu, pendekatan ini seringkali menuntut pengembangan infrastruktur atau *platform* khusus untuk mengelola volume tugas mikro yang sangat besar.

Metode inti dari pendekatan *crowdsourcing* meliputi:

1. Identifikasi Masalah

Peneliti mengakui bahwa pekerja non-ahli tidak mampu mentranskripsi seluruh kata atau kalimat dari manuskrip kuno karena tulisannya sulit dibaca dan penuh singkatan.

2. Pemrosesan Awal Otomatis

Sistem secara otomatis memproses gambar manuskrip untuk mengidentifikasi baris dan kata. Setiap gambar kata kemudian *di-over-segmentasi* menjadi segmen-segmen yang lebih kecil dari satu karakter.

3. Dekomposisi Tugas



Alih-alih meminta pekerja untuk ”mentranskripsi”, sistem memberi mereka tugas *micro-task* yang jauh lebih sederhana yang menyerupai *pattern matching*.

#### 4. *Micro-Task*

Pekerja disajikan gambar kata yang sudah terpotong-potong (setiap segmen diberi warna berbeda) dan sebuah simbol target (misalnya, ‘a’). Tugasnya adalah “Tandai segmen-segmen yang membentuk simbol ‘a’”.

#### 5. Kontrol Kualitas

Untuk mengatasi kesalahan dari pekerja non-ahli, pendekatan ini sangat bergantung pada redundansi. Tugas yang sama diberikan kepada beberapa pekerja yang berbeda. Sebuah label baru diterima hanya jika telah mencapai ambang batas suara positif (misalnya, 3 suara).

### II.2.2 *Human-Centric Approach*

Berbeda dengan *crowdsourcing* yang menyederhanakan tugas, pendekatan *Human-Centric* adalah metodologi konstruksi *dataset* yang dirancang khusus untuk menangani tugas anotasi yang kompleks secara semantik, seperti penentuan keandalan konten berita. Sebagaimana diusulkan oleh (Bonet-Jover dkk. 2023), pendekatan ini tidak bergantung pada satu teknik saja, melainkan menggabungkan beberapa strategi dalam sebuah *pipeline* bertahap yang bertujuan memaksimalkan efisiensi anotasi. Metodologi ini membagi proses pengembangan menjadi tiga fase yang secara progresif meningkatkan keterlibatan otomatisasi mesin untuk membantu anotator manusia.

Proses dimulai dengan fase Anotasi Manual Penuh (Inisialisasi), di mana sejumlah kecil data dikumpulkan dan dianotasi sepenuhnya dari awal oleh manusia guna menciptakan seed dataset berkualitas tinggi untuk melatih model awal. Setelah model terbentuk, proses berlanjut ke fase Kompilasi Otomatis, di mana model menganalisis kumpulan data besar dan secara cerdas memilih dokumen yang paling informatif untuk dianotasi, sehingga menghemat waktu kurasi data. Tahap yang paling efisien adalah fase ketiga, yaitu Kompilasi dan Anotasi Semi-Otomatis (Pra-pelabelan). Pada tahap ini, sistem memberikan pra-anotasi otomatis pada data baru, sehingga peran manusia bergeser drastis dari melakukan anotasi manual menjadi sekadar memverifikasi, mengoreksi, dan melengkapi hasil kerja mesin.

Penerapan pendekatan ini terbukti memberikan dampak signifikan terhadap efisiensi dan kualitas data. Studi kasus menunjukkan adanya pengurangan waktu anotasi hingga hampir 64% dibandingkan metode manual penuh, serta peningkatan konsis-

tensi antar-anotator karena adanya basis pra-anotasi yang seragam. Meskipun demikian, pendekatan ini memiliki tantangan tersendiri, seperti kebutuhan infrastruktur teknis yang kompleks untuk melatih model secara iteratif dan risiko munculnya bias model. Risiko bias terjadi ketika model hanya menyarankan data yang sesuai dengan pola yang sudah dipelajarinya, yang berpotensi mengurangi keragaman dataset jika tidak diawasi dengan cermat.

Berikut ini adalah tiga fase utama dari pendekatan *Human-Centric*:

1. Anotasi Manual Penuh (Inisialisasi)

Sejumlah kecil data awal dikumpulkan dan dianotasi sepenuhnya secara manual dari awal. Ini adalah proses yang lambat dan melelahkan, tetapi penting untuk menciptakan *seed dataset* berkualitas tinggi untuk melatih model awal.

2. Kompilasi Otomatis

Model ini dilatih kemudian menganalisis kumpulan data besar yang belum berlabel dan secara otomatis memilih dokumen yang paling informatif untuk dianotasi oleh manusia. Ini menghemat waktu kompilasi, tetapi anotasi masih dilakukan secara manual.

3. Kompilasi dan Anotasi Semi-Otomatis (Pra-pelabelan)

Ini adalah fase yang paling efisien. Sistem tidak hanya memilih dokumen baru, tetapi juga secara otomatis memberikan pra-anotasi pada dokumen tersebut menggunakan model yang dilatih dari semua data sebelumnya. Tugas manusia kemudian beralih dari "menganotasi dari nol" menjadi "memeriksa, mengoreksi, dan melengkapi" hasil pra-anotasi dari mesin.

## II.3 Metodologi Pengembangan *Dataset*

Pengembangan *dataset* dalam *machine learning* modern semakin diakui sebagai proses yang krusial dan kompleks, yang membutuhkan metodologi terstruktur untuk memastikan transparansi, akuntabilitas, dan tanggung jawab etis. Mengambil analogi dari lembar data di industri elektronik yang merinci karakteristik komponen, (Gebru dkk. 2021) mengusulkan agar setiap dataset disertai dengan lembar data serupa. Tujuan utamanya adalah untuk meningkatkan transparansi dan akuntabilitas dalam komunitas *machine learning*. Metodologi ini dirancang untuk melayani dua pemangku kepentingan utama: pembuat dataset dan konsumen *dataset*. Bagi pembuat, proses ini mendorong refleksi yang cermat terhadap asumsi, potensi risiko, dan implikasi penggunaan *dataset*; proses refleksi ini sengaja dirancang untuk tidak otomatis. Bagi konsumen, lembar data menyediakan informasi yang diperlukan untuk membuat keputusan yang tepat tentang penggunaan *dataset* dan menghindari

penyalahgunaan (Gebru dkk. 2021).

Melengkapi kerangka kerja dokumentasi tersebut, (Orr dan Crawford 2024) mengeksplorasi praktik dan proses penciptaan *dataset* yang bertanggung jawab dari perspektif pembuatnya. Melalui wawancara mendalam dengan 18 pembuat *dataset* terkemuka, mereka mengidentifikasi bahwa pekerjaan *dataset* seringkali terfragmentasi, kurang dihargai, dan para pembuatnya sering belajar dari kesalahan secara terisolasi (Orr dan Crawford 2024). Studi ini menyajikan tujuh rekomendasi metodologis utama yang berasal dari pengalaman praktis para pembuat *dataset*, yang memberikan panduan tentang bagaimana melaksanakan siklus hidup yang diidentifikasi oleh (Gebru dkk. 2021).

Studi Orr dan Crawford menyajikan tujuh rekomendasi metodologis utama yang berasal dari pengalaman praktis para pembuat *dataset*. Rekomendasi ini memberikan panduan praktis tentang bagaimana melaksanakan tahapan siklus hidup yang diidentifikasi oleh Gebru et al.

1. Diversifikasi dan Audit

Metodologi pengembangan harus secara aktif mendiversifikasi *dataset* tidak hanya secara demografis tetapi juga dalam atribut data untuk menghindari sinyal palsu. Ini harus disertai dengan audit menyeluruh untuk melaporkan distribusi data secara jujur.

2. Upayakan Kualitas Tinggi

Kualitas adalah pilar metodologis, yang dicapai melalui validasi data dan inspeksi manual, serta proses kurasi dan pembersihan yang disiplin.

3. Mulai Lebih Awal dan Iterasi

Proses pengembangan *dataset* bersifat iteratif. Para pembuat merekomendasikan untuk menerima kesalahan sebagai hal yang tak terhindarkan dan menggunakan *rapid iteration* untuk perbaikan berkelanjutan, terutama saat bekerja dengan *crowd worker*.

4. Dokumentasi Terbuka dan Komunikasi Keterbatasan

Menegaskan kembali pentingnya "Datasheets", Orr dan Crawford menyoroti perlunya mendokumentasikan secara terbuka untuk mengatasi "documentation debt" dan memastikan reproduktifitas. Ini juga mencakup komunikasi berkelanjutan tentang keterbatasan bahkan setelah *dataset* dirilis.

5. Desain Berpusat Pengguna dan Pembatasan Penggunaan

Metodologi yang bertanggung jawab menuntut pembuat untuk mendefinisikan kasus penggunaan yang dimaksud dan secara aktif mengantisipasi potensi bahaya atau penggunaan yang tidak diinginkan. Jika risiko penyalahgunaan

terlalu tinggi, pembuat harus berani untuk tidak merilis *dataset*.

#### 6. Atasi Privasi dan Persetujuan

Pembuat *dataset* didesak untuk mempertimbangkan privasi melampaui kewajiban hukum. Ini termasuk praktik seperti *scrapping* yang ”sopan” (*polite*) dan mengevaluasi kembali asumsi bahwa data yang tersedia untuk umum menyiratkan persetujuan untuk digunakan dalam pelatihan model.

#### 7. Buat *Dataset* yang Dibutuhkan

Terakhir, para praktisi merekomendasikan untuk tidak hanya bergantung pada data yang ditemukan. Metodologi yang kuat harus mencakup upaya untuk membuat *dataset* baru yang sesuai *fit-for-purpose*, meskipun membutuhkan lebih banyak sumber daya.

## II.4 Kelompok Suku Bangsa Indonesia

Indonesia merupakan negara kepulauan dengan keragaman etnis yang luar biasa. Berdasarkan data Sensus Penduduk tahun 2010 yang dirilis oleh Badan Pusat Statistik, komposisi penduduk Indonesia terdiri dari ribuan suku bangsa yang tersebar di berbagai pulau besar. Untuk memberikan gambaran kuantitatif mengenai persebaran ini, data demografi suku bangsa utama dapat dilihat pada Tabel berikut ini.

Tabel II.1 Komposisi Penduduk Indonesia Berdasarkan Suku Bangsa (Badan Pusat Statistik 2012)

Kelompok Suku Bangsa	Jumlah Penduduk	Persentase dari Total Penduduk (%)	Ranking
Suku asal Aceh	4.091.451	1,73	14
Batak	8.466.969	3,58	3
Nias	1.041.925	0,44	30
Melayu	5.365.399	2,27	10
Minangkabau	6.462.713	2,73	7
Suku asal Jambi	1.415.547	0,60	25
Suku asal Sumatera Selatan	5.119.581	2,16	10
Suku asal Lampung	1.381.660	0,58	26
Suku asal Sumatera Lainnya	2.204.472	0,93	21
Betawi	6.807.968	2,88	6
Suku asal Banten	4.657.784	1,97	11
Sunda	36.701.670	15,50	2
Jawa	95.217.022	40,22	1
Cirebon	1.877.514	0,79	24
Madura	7.179.356	3,03	5
Bali	3.946.416	1,67	15
Sasak	3.173.127	1,34	16
Suku Nusa Tenggara Barat lainnya	1.280.094	0,54	27
Suku asal Nusa Tenggara Timur	4.184.923	1,77	12
Dayak	3.009.494	1,27	17
Banjar	4.127.124	1,74	13
Suku asal Kalimantan lainnya	1.968.620	0,83	22
Makassar	2.672.590	1,13	20
Bugis	6.359.700	2,69	8
Minahasa	1.237.177	0,52	29
Gorontalo	1.251.494	0,53	28
Suku asal Sulawesi lainnya	7.634.262	3,22	4
Suku asal Maluku	2.203.415	0,93	22
Suku asal Papua	2.693.630	1,14	19
Cina	2.832.510	1,20	18
Asing/Luar Negeri	162.772	0,07	31

Berdasarkan data dari Sensus Penduduk 2010 pada Tabel II.1, didapat bahwa Indonesia memiliki kelompok suku bangsa yang sangat beragam (Badan Pusat Statistik 2012). Oleh karena itu untuk merepresentasikan Indonesia diputuskan untuk mengambil 2 suku terbanyak di setiap pulau di Indonesia beserta kelompok suku representatif lainnya.

Berikut ini adalah tabel pengelompokan 2 suku terbesar pada setiap pulau di Indonesia beserta kelompok representatif yang akan digunakan.

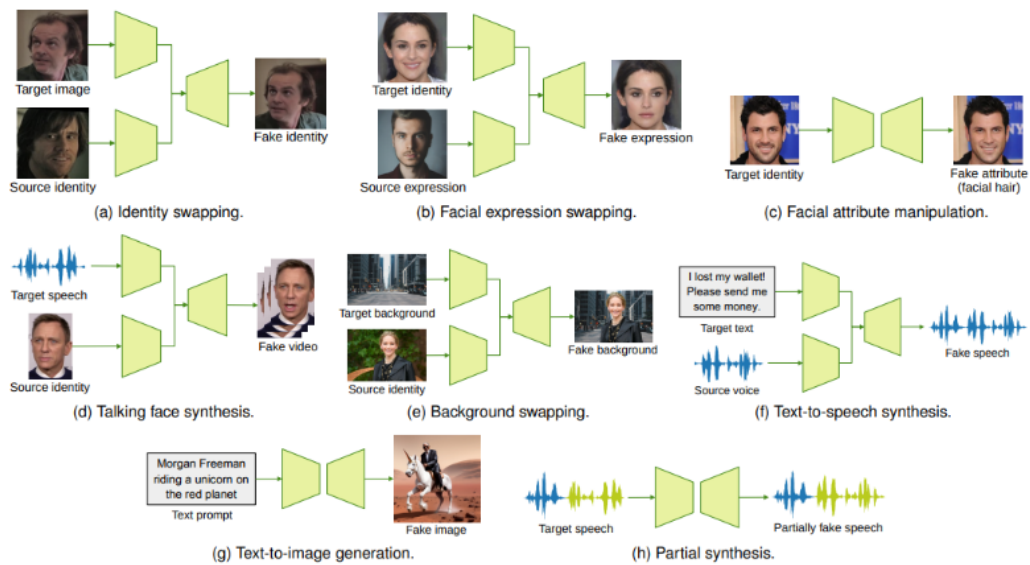
Tabel II.2 Pengelompokan 2 suku Terbesar Pada Setiap Pulau di Indonesia Beserta Kelompok Representatif Berdasarkan Sensus Penduduk 2010

Wilayah	Suku	Persentase
Jawa	Jawa	40,22%
	Sunda	15,5%
Sumatra	Batak	3,58%
	Minangkabau	2,73%
Kalimantan	Banjar	1,74%
	Dayak	1,27%
Sulawesi	Bugis	2,69%
	Makassar	1,13%
Bali & Nusa Tenggara	Bali	1,67%
	Sasak	1,34%
Kelompok Representatif	Cina	1,2%

Berdasarkan Tabel II.2, dengan strategi pengambilan sampel berstrata berdasarkan wilayah ini, kesebelas kelompok suku bangsa tersebut dipilih untuk memastikan dataset yang dibangun dapat mewakili keragaman suku utama di Indonesia secara lebih berimbang, tidak hanya terkonsentrasi pada populasi mayoritas.

## II.5 Klasifikasi Deepfake

Secara umum, teknologi *deepfake* dapat dikategorikan berdasarkan modalitas media yang dimanipulasi, yaitu visual, audio, dan multimodal. Masing-masing kategori memiliki metode pembangkitan dan karakteristik artefak yang berbeda. Klasifikasi mendalam mengenai jenis-jenis manipulasi deepfake ini divisualisasikan dalam diagram taksonomi berikut.



Gambar II.2 Jenis-Jenis *Deepfake* (Croitoru dkk. 2024)

Berdasarkan survei komprehensif oleh (Croitoru dkk. 2024), media *deepfake* didefinisikan sebagai file gambar, video, atau audio yang diubah secara digital atau dibuat dari awal menggunakan alat *AI*. Tinjauan literatur dalam penelitian tersebut secara khusus mencakup semua jenis media *deepfake*, yang dikategorikan ke dalam empat domain utama: gambar, video, audio, dan konten multimodal (*audio-visual*). Subbab berikut akan menguraikan klasifikasi ini secara rinci, dimulai dari *deepfake* visual (yang mencakup gambar dan video), diikuti oleh *deepfake* audio, dan *deepfake* multimodal. Prosedur-prosedur ini, yang diilustrasikan pada Gambar II.2, dapat bersifat *domain-agnostic* (berlaku di berbagai jenis media) atau *domain-specific* (hanya berlaku untuk media tertentu saja).

### II.5.1 Visual Deepfake

*Visual deepfake* merujuk pada media gambar atau video yang telah diubah atau dibuat dari awal menggunakan alat *AI*. Tujuannya adalah untuk memanipulasi konten visual, paling sering menargetkan wajah manusia. Berikut ini adalah jenis-jenis utama visual *deepfake*:

#### 1. Identity Swapping (Face Swapping)

Ini adalah jenis *deepfake* yang paling umum, identitas (wajah) seseorang dalam gambar atau video target diganti dengan identitas orang lain (sumber). Teknik ini berusaha mempertahankan atribut non-identitas dari target, seperti ekspresi wajah.

#### 2. Facial Expression Swapping (Face Reenactment)

Berbeda dengan *face swapping*, teknik ini tidak mengubah identitas seseorang, melainkan memodifikasi ekspresi atau emosi wajah. Dalam domain video, ini sering disebut *face reenactment*, gerakan wajah target diubah, seringkali meniru gerakan dari video sumber.

### 3. *Facial Attribute Manipulation*

Teknik ini berfokus pada pengubahan atribut semantik tertentu pada wajah sambil mempertahankan identitas aslinya. Atribut yang umum diubah meliputi usia, jenis kelamin, warna kulit, atau gaya rambut.

### 4. *Text-to-Image/Video Generation*

Dengan kemajuan model difusi, *deepfake* visual kini dapat dibuat hanya dengan menggunakan perintah teks (*prompt*). Pengguna dapat menentukan detail seperti nama orang, atribut wajah, dan tindakan yang harus dilakukan.

### 5. *Background Swapping*

Jenis manipulasi ini mengubah pemandangan latar belakang pada konten visual, biasanya dengan mensegmentasi orang di latar depan dan menempatkannya di latar belakang yang baru (Croitoru dkk. 2024).

## II.5.2 *Audio Deepfake*

*Audio deepfake* adalah file audio yang dimanipulasi atau dihasilkan oleh AI. Tujuannya sering kali untuk meniru suara seseorang atau mengubah konten ucapan. Jenis-jenis utama *audio deepfake* yang dibahas meliputi:

### 1. *Identity Swapping (Voice Conversion)*

Dalam konteks audio, ini dikenal sebagai *voice conversion* atau *voice swapping*. Tujuannya adalah untuk mengubah warna suara dan irama dari seorang pembicara agar terdengar seperti pembicara lain, sambil tetap mempertahankan isi pidato aslinya.

### 2. *Emotion Swapping*

Teknik ini mengubah emosi dalam sebuah rekaman ucapan (misalnya, dari tenang menjadi marah) tanpa mengubah isi konten atau identitas pembicara.

### 3. *Text-to-Speech (TTS) Synthesis*

Ini adalah proses pembuatan *audio deepfake* menggunakan model *machine learning* untuk sintesis ucapan dari input teks. Model TTS modern dapat menghasilkan ucapan yang terdengar alami dan mampu meniru suara identitas sumber tertentu.

### 4. *Partial Synthesis*

Jenis manipulasi ini hanya mengubah sebagian dari file media. Dalam domain audio, ini berarti mengubah sebagian kata dalam sebuah ucapan, sambil

memastikan identitas target tetap terjaga di seluruh klip audio (Croitoru dkk. 2024).

### II.5.3 *Multimodal Deepfake*

*Multimodal deepfake* adalah manipulasi yang melibatkan lebih dari satu jenis media secara bersamaan. Berikut ini adalah contoh dari *multimodal deepfake*:

#### 1. *Talking Face Synthesis*

Ini adalah contoh utama dari *deepfake* multimodal, yang digambarkan sebagai prosedur kompleks untuk menghasilkan file *audio-video* dari wajah yang berbicara. Proses ini dapat dikondisikan oleh berbagai input seperti teks, audio, atau video. Tujuannya adalah untuk menghasilkan video gerakan bibir, ekspresi wajah, gerakan kepala, dan ucapan yang dihasilkan semuanya konsisten dan tersinkronisasi (Croitoru dkk. 2024).

## II.6 *Dataset Terdahulu*

Pengembangan algoritma deteksi *deepfake* yang handal sangat bergantung pada ketersediaan dan kualitas data pelatihan. Seiring dengan kemajuan teknik manipulasi wajah, berbagai *dataset* publik telah dirilis oleh komunitas peneliti sebagai standar evaluasi. Evolusi *dataset* ini terlihat jelas dari peningkatan skala data, mulai dari ribuan hingga jutaan video, serta variasi metode pembuatan yang semakin kompleks. Sumber data yang digunakan dalam *dataset-dataset* ini bervariasi, mencakup video yang dikumpulkan dari *platform* publik seperti YouTube untuk menangkap kondisi nyata, hingga video yang direkam dalam lingkungan terkontrol melibatkan aktor berbayar untuk menjamin kualitas visual. Pemahaman terhadap karakteristik *dataset* terdahulu ini menjadi landasan penting untuk mengidentifikasi celah kekurangan yang ada saat ini, khususnya terkait representasi demografis dan kualitas video rendah. Informasi fundamental mengenai spesifikasi *dataset-dataset* utama tersebut dirangkum dalam tabel berikut.



Tabel II.3 Informasi Fundamental *Dataset* Deteksi Video *Deepfake* Terdahulu (Sohan, Solaiman, dan Hasan 2023)

Dataset	Rilis	Sumber	Real		Fake	
			Video	Frame	Video	Frame
FaceForensics (Rössler dkk. 2018)	Mar 2018	YouTube	1.004	519k	1.004	521k
FaceForensics++ (Rossler dkk. 2019)	Jan 2019	YouTube	1.000	509k	4.000	18M
DeepFakeDetection (Dufour dan Gully 2019)	Sep 2019	Partisipan berbayar	363	315k	3.068	2.242k
UADFV (M. Li dkk. 2021)	Nov 2018	YouTube	49	-	49	-
DeepfakeTIMIT (Korshunov dan Marcel 2018)	Des 2018	YouTube	-	-	620	68k
Celeb-DF (Y. Li dkk. 2019)	Sep 2019	YouTube	408	-	795	-
Celeb-DFv2 (Y. Li dkk. 2020)	Nov 2019	YouTube	590	230k	5.639	2.199k
DFDC preview (Dolhansky dkk. 2019)	Okt 2019	Partisipan berbayar	1.131	-	4.119	-
DFDC (Dolhansky dkk. 2020)	Jun 2020	Partisipan berbayar	23.654	-	104.500	-
DeeperForensics-1.0 (Jiang dkk. 2020)	Jan 2020	Berbayar & YouTube	50.000	12,6M	10.000	5M
WildDeepfake (Zi dkk. No date)	Okt 2020	Koleksi daring	3.805	440k	3.509	739k
KoDF (Kwon dkk. 2021)	Mar 2021	Partisipan sukarela	62.166	-	175.776	-
ForgeryNet (He dkk. 2021)	Jul 2021	-	99.630	1,4M	121.617	1,4M

Tabel II.3 menyajikan informasi fundamental mengenai *dataset-dataset* deteksi video *deepfake* yang sudah ada. *Dataset-dataset* ini, yang dirilis antara Maret 2018 hingga Juli 2021, digunakan secara luas dalam penelitian deteksi *deepfake*. Sumber video untuk *dataset* ini bervariasi, termasuk YouTube, partisipan berbayar, video yang dikumpulkan dari internet secara umum, dan partisipan sukarela. (Sohan, Solaiman, dan Hasan 2023). Berikut adalah daftar *dataset* di atas beserta detail-detailnya:

1. *FaceForensics*: Dirilis Maret 2018, bersumber dari *YouTube*, berisi 1.004 video asli dan 1.004 video palsu. (Rössler dkk. 2018).
2. *FaceForensics++*: Dirilis Januari 2019, bersumber dari *YouTube*, berisi 1.000 video asli dan 4.000 video palsu. (Rossler dkk. 2019).
3. *DeepFakeDetection*: Dirilis September 2019, bersumber dari partisipan berbayar, berisi 363 video asli dan 3.068 video palsu. (Dufour dan Gully 2019).
4. *UADFV*: Dirilis November 2018, bersumber dari *YouTube*, berisi 49 video asli dan 49 video palsu. (M. Li dkk. 2021).
5. *DeepfakeTIMIT*: Dirilis Desember 2018, berisi 620 video palsu. Jumlah video asli tidak disebutkan dalam tabel ini. (Korshunov dan Marcel 2018).
6. *Celeb-DF*: Dirilis September 2019, bersumber dari *YouTube*, berisi 408 video asli dan 795 video palsu. (Y. Li dkk. 2019).
7. *Celeb-DFv2*: Dirilis November 2019, bersumber dari *YouTube*, berisi 590 video asli dan 5.639 video palsu. (Y. Li dkk. 2020).
8. *DFDC preview*: Dirilis Oktober 2019, bersumber dari partisipan berbayar, berisi 1.131 video asli dan 4.119 video palsu. (Dolhansky dkk. 2019).
9. *DFDC*: Dirilis Juni 2020, bersumber dari partisipan berbayar, berisi 23.654 video asli dan 104.500 video palsu. (Dolhansky dkk. 2020).

10. *DeeperForensics-1.0*: Dirilis Januari 2020, bersumber dari partisipan berbayar dan *YouTube*, berisi 50.000 video asli dan 10.000 video palsu. (Jiang dkk. 2020).
11. *WildDeepfake*: Dirilis Oktober 2020, bersumber dari koleksi online, berisi 3.805 video asli dan 3.509 video palsu. (Zi dkk. No date).
12. *KoDF*: Dirilis Maret 2021, bersumber dari partisipan sukarela, berisi 62.166 video asli dan 175.776 video palsu. (Kwon dkk. 2021).
13. *ForgeryNet*: Dirilis Juli 2021, berisi 99.630 video asli dan 121.617 video palsu. (He dkk. 2021)

## II.7 *Deepfake* Pada Dunia Nyata

Kehadiran video yang dihasilkan oleh kecerdasan buatan (*AI*) di media sosial menimbulkan tantangan baru bagi deteksi *deepfake*. Detektor yang dilatih seringkali gagal melakukan generalisasi ke skenario dunia nyata. Salah satu faktor kunci di balik kesenjangan ini adalah kompresi agresif dan *proprietary* yang diterapkan oleh platform seperti *YouTube* dan *Facebook*. Proses ini secara signifikan merusak fitur forensik tingkat rendah yang penting untuk membedakan konten asli dari media yang dimanipulasi. Meskipun *State-of-the-Art (SoA)* metode deteksi *deepfake*, seperti yang berbasis *CNN (ResNet50, DenseNet, EfficientNet, XceptionNet)* dan *Vision Transformer (ViT)*, menunjukkan kinerja sangat baik di bawah kondisi laboratorium, kinerjanya menurun drastis ketika diterapkan pada *deepfake* yang dibagikan di media sosial.

Kompresi dan pengubahan ukuran pada media sosial dilakukan untuk mengatasi kendala *bandwidth* dan penyimpanan yang sangat relevan di Indonesia. Meskipun langkah-langkah ini mengurangi ukuran file, mereka juga secara signifikan menurunkan kualitas fitur forensik yang penting untuk deteksi. Kompresi ini dapat menghilangkan atau mengaburkan artefak halus yang diperkenalkan selama proses pembuatan *deepfake*, sehingga menyulitkan detektor untuk mengidentifikasi manipulasi secara akurat. Fenomena ini telah diamati secara konsisten dalam penelitian sebelumnya.

Untuk menjembatani kesenjangan dengan aplikasi dunia nyata, beberapa upaya telah dilakukan, terutama dalam domain gambar. Pendekatan seperti membuat *dataset* dengan gambar yang sudah dikompresi atau menggunakan *fine-tuning* untuk mengadaptasi model pada data terkompresi telah diusulkan. Namun, replikasi transformasi kompresi platform untuk video dalam skala besar tetap menjadi tantangan

karena keterbatasan *API* dan kendala berbagi data. Artikel ini mengusulkan kerangka kerja emulasi untuk mereplikasi artefak kompresi dan *resizing* video platform secara lokal, sebagai solusi untuk menghasilkan data pelatihan yang lebih realistis tanpa memerlukan akses *API* langsung (Montibeller dkk. 2025).

## II.8 *Benchmarking FaceForensics++*

Tujuan utama FaceForensics++ adalah untuk menyediakan *dataset* berskala besar dan benchmark standar untuk melatih dan mengevaluasi metode deteksi manipulasi wajah secara adil dan terukur.

Metodologi mereka dapat dibagi menjadi tiga langkah utama:

### 1. Pengumpulan (*Collection*)

Sumber data dikumpulkan 1.000 video asli dari YouTube. Untuk memastikan video-video tersebut dapat dimanipulasi dengan baik oleh *tool* otomatis, mereka melakukan penyaringan berlapis:

- a. Menggunakan tag seperti "newscaster" (pembawa berita) atau "interview" (wawancara).
- b. Hanya mengunduh video dengan resolusi 480p atau lebih tinggi.
- c. Mereka menjalankan detektor wajah (Dlib) pada video untuk memastikan wajah subjek dapat dilacak, yang sangat penting untuk tool manipulasi.
- d. Tim peneliti kemudian secara manual menyaring klip-klip tersebut untuk memastikan kualitasnya tinggi, wajah subjek sebagian besar menghadap ke depan, dan tidak ada oklusi (wajah tertutup) yang signifikan.

### 2. Manipulasi (*Manipulation*)

Mereka secara sistematis membuat video palsu dari 1.000 video asli tersebut menggunakan empat metode manipulasi terkemuka pada saat itu:

- a. *Facial Reenactmen*: Mengubah ekspresi target agar sesuai dengan sumber.
  - i. *Face2Face*: Pendekatan klasik berbasis grafis komputer untuk transfer ekspresi.
  - ii. *NeuralTextures*: Pendekatan *deep learning* yang mempelajari tekstur wajah untuk melakukan *reenactment*.
- b. *Identity Swap*: Menukar wajah target dengan wajah sumber.
  - i. *FaceSwap*: *Tool* berbasis grafis komputer yang populer untuk menukar wajah.
  - ii. *DeepFakes*: Pendekatan berbasis *deep learning* yang menggunak-

an jaringan saraf untuk menukar wajah.

### 3. Kompresi (*Post-processing*)

Setelah membuat video palsu, mereka membuat tiga kualitas untuk setiap video (asli dan palsu):

- a. *Raw*: Video asli yang belum dikompresi.
- b. *High Quality (HQ)*: Video dikompresi dengan rasio 23 (*bitrate* sekitar 17 Mbit/s) menggunakan *codec* H.264.
- c. *Low Quality (LQ)*: Video dikompresi dengan rasio 40 (*bitrate* sekitar 7 Mbit/s) menggunakan *codec* H.264.

Hasil akhirnya adalah *dataset* masif dengan lebih dari 1,8 juta gambar yang dimanipulasi.

## II.9 Teknik Kompresi Video

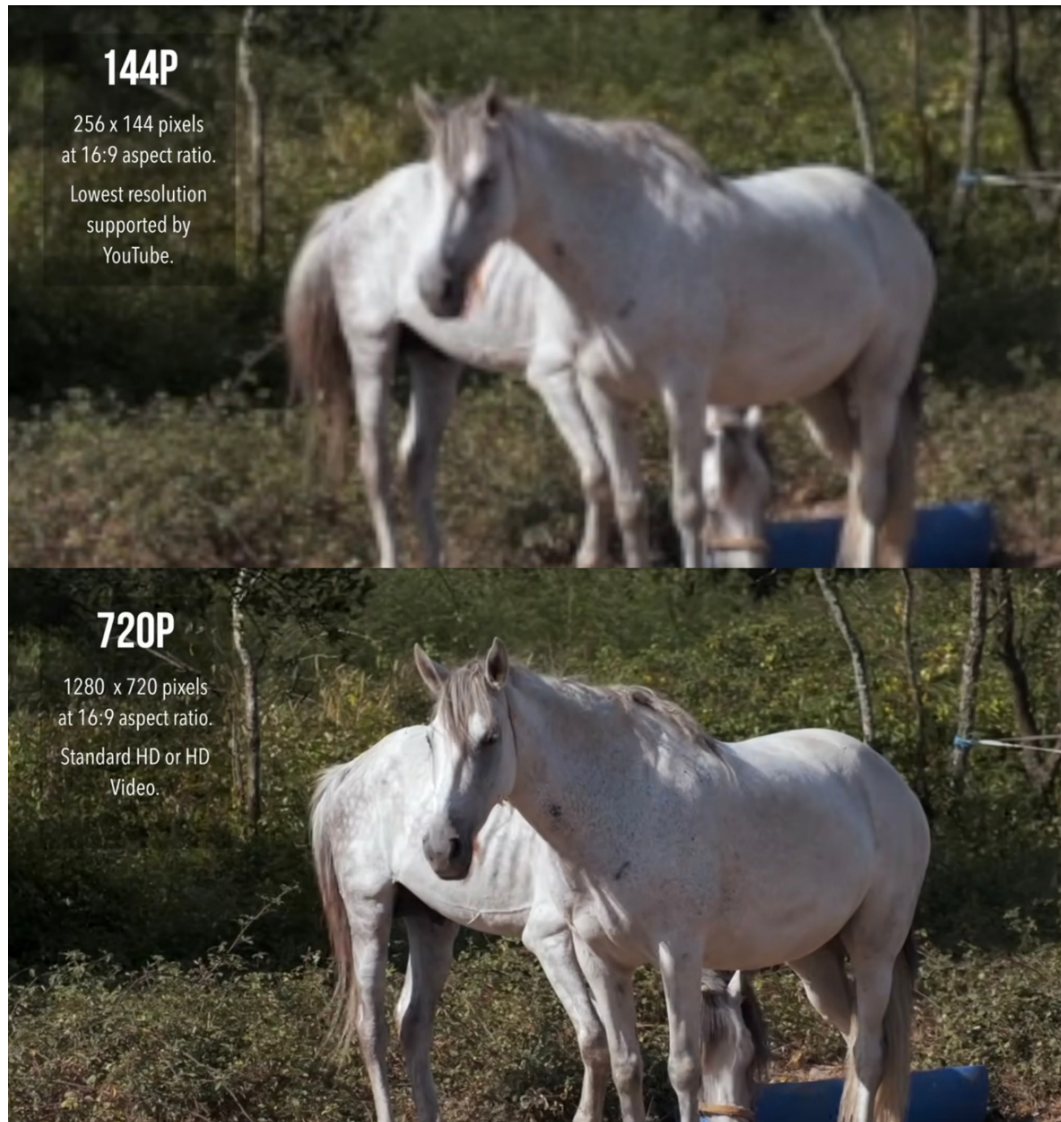
Penelitian yang dilakukan oleh (Sajati, Pujiastuti, dan Triantoro 2018) memanfaatkan FFmpeg untuk membangun sistem kompresi video berbasis web. Teknik kompresi yang digunakan adalah *Lossy Compression*, di mana data hasil kompresi tidak akan sama persis dengan data aslinya, namun dianggap sudah cukup untuk digunakan.

Proses kompresi yang diimplementasikan dalam penelitian tersebut dapat diuraikan sebagai berikut:

1. Pemisahan (*Splitting*): Sistem memecah *file* video asli (misalnya, resolusi 720p) menjadi komponen video dan audio secara terpisah.
2. Kompresi (*Compression*):
  - a. Komponen video dikompresi dan diubah resolusinya menjadi beberapa *file* keluaran (contoh: 480p, 360p, dan 240p).
  - b. Komponen audio juga dikompresi secara terpisah, namun hanya menjadi satu kualitas standar.
3. Penggabungan (*Merging*): Setelah kompresi, *file* video dan audio yang telah diproses digabungkan kembali menjadi satu *file* utuh sesuai dengan resolusi masing-masing.

Pendekatan ini terbukti menghasilkan file yang lebih kecil. Pengujian menunjukkan bahwa dengan mengompresi audio, ukuran file akhir dapat berkurang sekitar 1-2 MB lebih kecil dibandingkan jika hanya mengompresi videonya saja. Jika dibandingkan dengan ukuran file asli (misalnya dari YouTube), perbedaan ukurannya bisa mencapai 5-20 MB.

Meskipun efektif dalam mengurangi ukuran, teknik kompresi ini memiliki konsekuensi pada kualitas visual. Hasil pengujian menunjukkan bahwa kualitas gambar dari video yang dikompresi relatif lebih rendah dan tidak memenuhi standar kualitas 30dB.



Gambar II.3 Perbedaan Kualitas Video Rendah dan Kualitas Video Tinggi (BarTech TV 2022)

Perbedaan paling mencolok pada Gambar II.3 terletak pada kerapatan piksel yang berdampak langsung pada kejelasan detail visual. Sementara gambar 144p menyajikan visual yang buram dan *pixelated* sehingga detail halus seperti helaian surai kuda dan tekstur rumput melebur menjadi blok-blok warna yang tidak jelas akibat minimnya informasi data gambar 720p justru menampilkan ketajaman yang superior dengan garis batas yang tegas dan tekstur yang realistis, membuktikan bahwa pe-

ningkatan resolusi secara drastis memperkaya kualitas informasi visual yang dapat ditangkap oleh mata.

## **II.10 Kajian Pengembangan *Dataset* dan Benchmarking untuk Algoritma Deteksi**

Pengembangan algoritma deteksi *deepfake* sangat bergantung pada ketersediaan data pelatihan yang berkualitas. Menurut tinjauan sistematis oleh (Jabbar dkk. 2023), meskipun teknologi *deepfake* telah ada selama beberapa dekade, kemampuan untuk memanipulasi gambar dan video dengan cepat menggunakan *Artificial Intelligence*.

### **II.10.1 Kategorisasi *Dataset* Berdasarkan Teknik Manipulasi**

Alih-alih dikategorikan berdasarkan waktu rilisnya, evolusi *dataset deepfake* dapat dipahami lebih baik melalui teknik manipulasi yang digunakan untuk menghasilkanannya. Berdasarkan tinjauan literatur mengenai klasifikasi *deepfake*, *dataset* yang ada saat ini umumnya dibangun menggunakan pendekatan utama berikut: *Identity Swapping*, *Facial Reenactment*, dan pendekatan *Hybrid* atau campuran.

#### **1. *Dataset* Berbasis *Identity Swapping***

Metode ini adalah teknik yang paling umum ditemukan, di mana wajah subjek pada video target digantikan oleh wajah dari subjek sumber.

- a. DeepfakeTIMIT: Merupakan salah satu dataset generasi awal yang berfokus pada pertukaran wajah menggunakan algoritma berbasis GAN. *Dataset* ini memanipulasi wajah dari 32 subjek dengan resolusi yang terbatas.
- b. UADFV: *Dataset* ini juga menggunakan teknik pertukaran wajah dasar yang terdiri dari 49 video asli dan 49 video palsu, yang menjadi standar awal penelitian deteksi.
- c. Celeb-DF & Celeb-DFv2: *Dataset* ini dikembangkan untuk mengatasi rendahnya kualitas visual pada *dataset* generasi awal. Celeb-DF menggunakan algoritma pertukaran wajah yang lebih canggih untuk mengurangi artefak visual yang terlihat jelas, menghasilkan video *deepfake* berkualitas tinggi yang lebih sulit dideteksi.
- d. FaceForensics++: Dalam *dataset* berskala besar ini, terdapat subset video yang dimanipulasi menggunakan teknik *identity swapping*, yaitu menggunakan metode *DeepFakes* dan *FaceSwap*.

#### **2. *Dataset* Berbasis *Facial Reenactment***

Berbeda dengan *identity swapping*, metode ini memanipulasi ekspresi wajah

target (seperti gerakan mulut, kedipan mata, atau pose kepala) agar meniru ekspresi dari wajah sumber, tanpa mengubah identitas orang tersebut.

- a. FaceForensics++: *Dataset* ini menjadi *benchmark* utama untuk kategori ini karena menyertakan manipulasi berbasis *reenactment* secara sistematis. Metode yang digunakan mencakup *Face2Face*, yang merupakan pendekatan klasik berbasis grafis komputer untuk mentransfer ekspresi secara *real-time*, dan *Neural Textures*, yang menggunakan pendekatan *deep learning* untuk memodifikasi tekstur wajah saat melakukan *reenactment*.

### 3. *Dataset* Berbasis Pendekatan Hybrid

Kategori ini mencakup *dataset* yang tidak bergantung pada satu metode pembuatan saja, melainkan menggabungkan berbagai teknik atau mengumpulkan video yang sudah beredar luas di internet untuk mensimulasikan kondisi nyata.

- a. DFDC: *Dataset* ini dirancang untuk tantangan skala besar dengan melibatkan ribuan aktor berbayar. Video dalam *dataset* ini dimanipulasi menggunakan berbagai metode campuran, termasuk pertukaran wajah dan modifikasi audio-visual, untuk menciptakan keragaman artefak yang luas.
- b. *Wild Deepfake*: *Dataset* ini dikumpulkan langsung dari internet, sehingga metode pembuatannya sangat bervariasi dan sering kali tidak diketahui metadatanya secara spesifik. *Dataset* ini merepresentasikan ancaman nyata di mana video mengandung beragam jenis manipulasi yang telah melalui berbagai proses kompresi *platform*.

## II.10.2 Metodologi Pengembangan *Dataset Benchmark*

Untuk membangun *dataset* yang valid sebagai *benchmark*, metodologi yang ketat diperlukan untuk menjamin *ground truth* yang akurat. (Rössler dkk. 2018) menguraikan pendekatan *end-to-end* dalam pembuatan *dataset* FaceForensics yang mencakup tahapan berikut:

### 1. Pengumpulan Data (*Data Collection*)

Video asli dikumpulkan dari YouTube dengan kriteria resolusi tinggi (di atas 480p) dan disaring secara manual untuk memastikan wajah subjek tidak terhalang.

### 2. Manipulasi Otomatis (*Automated Manipulation*)

Proses manipulasi dilakukan menggunakan pendekatan *state-of-the-art* seperti *Face2Face*. (Rössler dkk. 2018) menjelaskan proses ini melibatkan pela-

cakan wajah (*tracking*), rekonstruksi model 3D, dan *re-rendering* ekspresi wajah dari video sumber ke video target. (Jabbar dkk. 2023) mengklasifikasikan teknik manipulasi ini ke dalam empat kategori utama: *Editing & Synthesis*, *Identity Replacement (Swap/Transfer)*, *Reenactment* (ekspresi, mulut, tatapan), dan metode lainnya.

### 3. Simulasi Kompresi (*Post-Processing*)

(Rössler dkk. 2018) menekankan pentingnya mengevaluasi algoritma pada video yang terkompresi, karena kompresi media sosial sering menghilangkan jejak manipulasi tingkat rendah. Oleh karena itu, *benchmark* standar menyediakan data dalam berbagai tingkat kompresi, misalnya menggunakan *co-dec* H.264 dengan parameter kuantisasi 23 (*light compression*) dan 40 (*strong compression*).

## II.11 Keterbatasan *Dataset Benchmark* dalam Skenario Dunia Nyata

Berdasarkan penelitian terbaru oleh (Sahu dkk. 2025), menunjukkan adanya kesenjangan yang signifikan antara kinerja model deteksi di lingkungan laboratorium yang terkontrol dibandingkan dengan skenario media sosial yang kacau. Berikut ini adalah keterbatasan *dataset benchmark* dalam skenario dunia nyata.

### 1. Kesenjangan Antara Kondisi Ideal dan Kondisi *Chaotic*

Sebagian besar alat deteksi *deepfake* saat ini dikembangkan dan diuji menggunakan *dataset* berkualitas tinggi yang diberi label dengan jelas dalam kondisi laboratorium yang sempurna. Namun, lingkungan media sosial (seperti *Instagram*, *TikTok*, dan *Facebook*) sangat berbeda karena sifatnya yang kacau, penuh dengan konten berkualitas rendah, dan terus-menerus dibagikan ulang. *Dataset* standar sering kali gagal merepresentasikan kondisi ini, di mana konten sering kali dikompresi secara agresif, memiliki resolusi rendah, dan mengandung *noise* visual yang signifikan.

### 2. Dampak Kompresi dan Artefak *Platform*

Keterbatasan utama dari *dataset benchmark* tradisional adalah ketidakmampuannya meniru degradasi kualitas yang terjadi di *platform* sosial.

- a. Kompresi Agresif: *Platform* media sosial secara rutin mengompres video dan gambar secara berat untuk menghemat penyimpanan, yang menghilangkan sinyal visual halus yang biasanya digunakan detektor untuk menemukan manipulasi.
- b. *Noise* dan Variabilitas: *Dataset benchmark* cenderung bersih, sedangkan data dunia nyata mengandung *noise*, konversi format *file*, dan variasi *frame rate* yang menutupi jejak forensik.



- c. Hilangnya Fitur Deteksi: Studi menemukan bahwa kompresi media sosial dapat menyebabkan penurunan presisi deteksi rata-rata sebesar 15-20% karena hilangnya fitur-fitur kritis yang ada pada data pelatihan berkualitas tinggi.

### 3. Bukti Empiris Kegagalan Generalisasi

Ketergantungan pada *dataset benchmark* membatasi validitas eksternal model, yang terbukti dari penurunan kinerja drastis saat model diuji pada data *in-the-wild*:

- a. Penurunan Akurasi: Model canggih seperti *CrossDF*, yang bekerja baik pada *benchmark*, mengalami penurunan presisi menjadi 71,2% dan *recall* menjadi 66,4% saat diuji pada *dataset* media sosial nyata.
- b. Kegagalan Model *Real-Time*: Model yang dioptimalkan untuk kecepatan, seperti *Tiny-LaDeDa*, juga menderita penurunan performa signifikan dengan presisi hanya mencapai 68,4% akibat kompresi berat dan pengunggahan ulang konten.
- c. Kasus Nyata: Dalam studi kasus kampanye misinformasi di *Instagram*, sistem otomatis hanya mampu mendeteksi sekitar 55% konten *deepfake*, sementara 15% sisanya lolos sepenuhnya karena resolusi rendah dan kompresi yang menyembunyikan tanda manipulasi.

### 4. Implikasi: Perlunya Pendekatan *Multimodal*

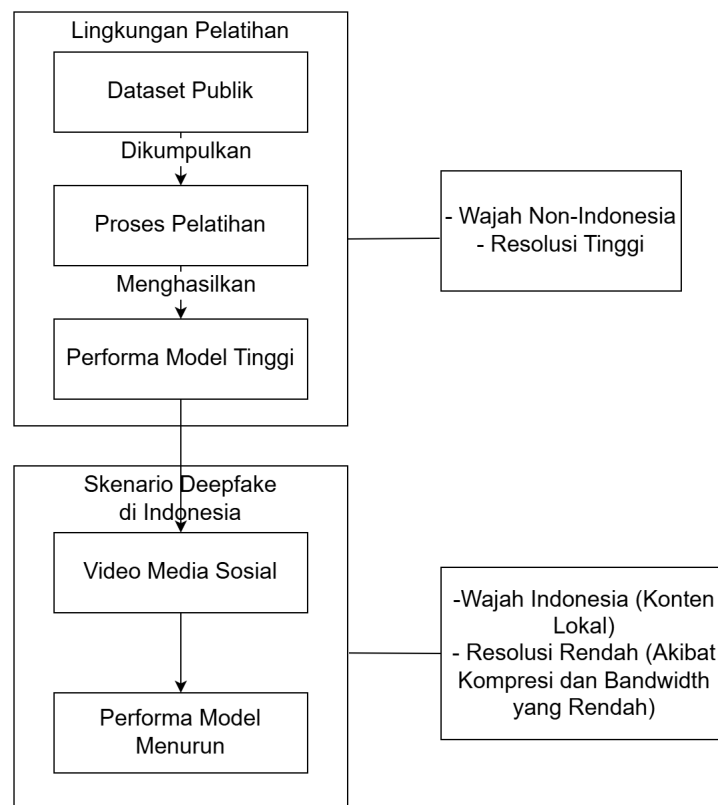
Keterbatasan *dataset* visual tunggal menegaskan bahwa mengandalkan fitur visual saja tidak lagi cukup, terutama dengan kemajuan teknik *generation* (seperti GAN dan difusi) yang meminimalkan artefak visual. Detektor masa depan harus dilatih menggunakan *dataset* yang mencakup modalitas ganda (visual, *audio*, dan temporal) untuk meningkatkan ketahanan terhadap variasi kualitas di dunia nyata.

## BAB III

### ANALISIS MASALAH

#### III.1 Analisis Kondisi Saat Ini

Sistem deteksi *deepfake* yang ada saat ini dapat dimodelkan sebagai sebuah alur proses yang bergantung penuh pada ketersediaan data pelatihan. Gambar III.1 menggambarkan model konseptual dari alur kerja deteksi *deepfake* yang umum digunakan saat ini di Indonesia.



Gambar III.1 Model Konseptual Alur Deteksi *Deepfake* Saat Ini

Alur kerja dari model konseptual dimulai dari *dataset* publik yang telah banyak ditinjau dalam studi literatur, seperti FaceForensics , Celeb-DFv2 , dan DFDC (Sohan, Solaiman, dan Hasan 2023). Sebagian besar dataset ini memiliki karakteristik utama berupa wajah non-Indonesia dan video dengan resolusi yang masih tergolong tinggi. Dataset ini yang kemudian digunakan untuk proses pelatihan , dilanjutkan dengan arsitektur *deep learning* belajar mengenali pola dan artefak manipulasi. Hasil dari proses ini adalah sebuah model deteksi yang menunjukkan performa sangat tinggi di lingkungan laboratorium.

Namun, seperti yang telah diidentifikasi dalam studi literatur , masalah fundamental muncul apabila model tersebut digunakan untuk mendeteksi *deepfake* dalam konteks dunia nyata di Indonesia. Video-video di Indonesia berisi orang-orang lokal (Indonesia) dan, yang lebih penting, mengalami kompresi agresif akibat penyebaran di sosial media serta *bandwidth* internet yang tidak merata.

Kondisi ini sejalan dengan temuan (Montibeller dkk. 2025), yang menyatakan bahwa kompresi video secara signifikan merusak fitur forensik tingkat rendah, sehingga menyulitkan detektor. Akibat dari kesenjangan data (distribusi *out-of-distribution*) antara data latih (resolusi tinggi, non-Indonesia) dengan data di lapangan (resolusi rendah, Indonesia) , performa model yang sebelumnya tinggi dapat menurun drastis. Kondisi inilah yang menjadi justifikasi utama perlunya pengembangan *dataset* baru yang dapat menjembatani kesenjangan tersebut.

## **III.2 Analisis Kebutuhan**

Berikut ini adalah penjabaran terkait kebutuhan-kebutuhan untuk membuat *dataset* yang baru.

### **III.2.1 Identifikasi Masalah Pengguna**

Pengguna sistem ini adalah Steven (pembuat model) dan peneliti lain di bidang deteksi *deepfake*. Masalah yang dihadapi oleh para pengguna adalah:

1. Steven (pembuat model)
  - a. Steven membutuhkan data pelatihan yang memadai untuk membuat model yang *robust* terhadap video berkualitas rendah di Indonesia.
  - b. Modelnya mungkin akan *overfitting* pada video berkualitas tinggi dan gagal di skenario dunia nyata Indonesia.
2. Peneliti lainnya
  - a. Peneliti tidak memiliki *benchmark dataset* untuk menguji validitas mo-

del deteksi *deepfake* mereka dalam konteks Indonesia.

Untuk mencari solusi atas masalah-masalah tersebut, perlu disusun kebutuhan fungsional dan nonfungsional sistem yang diperlukan. Subbab berikut menjabarkan kebutuhan-kebutuhan tersebut.

### III.2.2 Kebutuhan Fungsional

Berikut ini adalah Tabel III.1 yang berisi kebutuhan-kebutuhan fungsional yang dibutuhkan untuk menjawab masalah-masalah tersebut.

Tabel III.1 Kebutuhan Fungsional *Dataset*

Kode	Kebutuhan Fungsional	Deskripsi
F-01	Konteks Lokal	<i>Dataset</i> harus berisi video asli yang menampilkan wajah orang Indonesia yang beragam.
F-02	Representasi Etnis	<i>Dataset</i> harus merepresentasikan keragaman suku bangsa utama di Indonesia.
F-03	Teknik Relevan	<i>Dataset</i> harus berisi video palsu yang dibuat menggunakan teknik <i>deepfake</i> yang umum, seperti visual <i>deepfake</i> , audio <i>deepfake</i> , maupun <i>multimodal deepfake</i> .
F-04	Simulasi Kompresi	<i>Dataset</i> harus menyertakan versi video (asli dan palsu) yang telah disimulasikan efek kompresinya untuk meniru kualitas rendah di media sosial.
F-05	Pelabelan Jelas	Setiap video dalam <i>dataset</i> harus memiliki label yang jelas.
F-06	Metadata Lengkap	<i>Dataset</i> harus disertai <i>metadata</i> yang detail untuk setiap video palsu.
F-07	Keseimbangan Data	<i>Dataset</i> harus memiliki keseimbangan yang wajar antara jumlah video asli dan palsu, serta antara video berkualitas tinggi dan rendah.

### III.2.3 Kebutuhan Nonfungsional

Berikut ini adalah Tabel III.2 yang berisi atribut-atribut mengenai bagaimana (*HOW*) sistem bekerja.

Tabel III.2 Kebutuhan Non Fungsional *Dataset*

Kode	Kebutuhan Non Fungsional	Penjelasan
NF-01	Relevansi	<i>Dataset</i> harus sangat relevan dengan masalah disinformasi <i>deepfake</i> di Indonesia.
NF-02	Keberagaman	<i>Dataset</i> harus beragam, tidak hanya dari segi etnis, tetapi juga pencahayaan, latar belakang, dan usia subjek.
NF-03	Validitas	Label harus 100% akurat. Proses validasi harus memastikan tidak ada kesalahan pelabelan.
NF-04	Realisme	Video berkualitas rendah yang dihasilkan harus secara akurat mencerminkan artefak kompresi yang ditemukan di platform nyata.

### III.3 Analisis Pemilihan Solusi

Berikut ini adalah beberapa alternatif solusi yang diusulkan beserta metode dalam penentuan solusi.

### III.4 Analisis Pemilihan Solusi

Berikut ini adalah beberapa alternatif solusi yang diusulkan beserta metode dalam penentuan solusi.

#### III.4.1 Penjabaran Alternatif Solusi

Terdapat tiga alternatif solusi utama yang diidentifikasi untuk memenuhi kebutuhan *dataset deepfake* berkualitas rendah dalam konteks Indonesia.

##### III.4.1.1 Alternatif Solusi 1: Kompresi *Dataset* Publik yang Ada

Solusi ini mengusulkan penggunaan *dataset* publik yang sudah mapan dan teruji, seperti FaceForensics++ atau Celeb-DF, sebagai basis. Pekerjaan utama dalam alternatif ini adalah mengambil video-video (asli dan palsu) dari *dataset* tersebut dan kemudian menerapkan simulasi kompresi video secara sistematis, sebagaimana yang diuraikan oleh (Montibeller dkk. 2025). Pendekatan ini secara langsung menargetkan kebutuhan F-04 (Simulasi Kompresi) dan NF-04 (Realisme Kompresi).

Kelebihan utama dari pendekatan ini adalah efisiensi waktu dan sumber daya. Proses pengumpulan data (Fase Pengumpulan) dan generasi *deepfake* (Fase Generasi) dapat dilewati seluruhnya. Mengingat proses generasi *deepfake* sangat intensif secara komputasi dan memakan waktu (Risiko 3), alternatif ini sangat menghemat

sumber daya. *Dataset* yang ada juga sudah memiliki pelabelan yang jelas (F-05) dan dibuat dengan teknik yang relevan (F-03).

Namun, solusi ini memiliki kelemahan fundamental yang tidak dapat diabaikan. *Dataset* yang dihasilkan akan gagal total dalam memenuhi kebutuhan inti dari penelitian ini, yaitu F-01 (Konteks Lokal) dan F-02 (Representasi Etnis). Model yang dilatih pada *dataset* ini (misalnya, wajah Kaukasia berkualitas rendah) kemungkinan besar akan tetap gagal melakukan generalisasi ketika dihadapkan pada video *deepfake* dengan wajah Indonesia berkualitas rendah. Kegagalan ini terjadi karena model tidak hanya belajar dari artefak kompresi, tetapi juga dari fitur-fitur wajah yang spesifik secara demografis. Selain itu, *metadata* yang ada (F-06) mungkin tidak cukup lengkap untuk tujuan penelitian, karena kita tidak memiliki kontrol atas parameter generasi *deepfake* aslinya.

#### III.4.1.2 Alternatif Solusi 2: Pengumpulan Pasif *Deepfake* Beredar

Solusi ini berfokus pada pengumpulan video *deepfake* yang sudah beredar secara "liar" (*in-the-wild*) di platform media sosial Indonesia. Pendekatan ini memiliki keunggulan teoretis tertinggi dalam hal relevansi (NF-01). *Dataset* yang dihasilkan akan terdiri dari video-video yang memang menjadi masalah nyata di lapangan, sehingga secara otomatis memenuhi F-01 (Konteks Lokal) dan F-04 (Simulasi Kompresi), karena video tersebut adalah produk kompresi dunia nyata.

Meskipun demikian, solusi ini memiliki banyak ketidakpastian fatal yang membuatnya tidak cocok untuk penelitian sistematis.

1. *Ketiadaan Ground Truth*: Hampir tidak mungkin mendapatkan video asli (*ground truth*) untuk setiap video palsu yang ditemukan. Tanpa video asli, pelabelan (F-05) menjadi tidak terkontrol dan tidak dapat divalidasi 100%.
2. *Metadata Tidak Diketahui*: Teknik *deepfake* yang digunakan (F-03) dan *metadata* video (F-06) tidak akan diketahui. Kita tidak tahu alat apa yang digunakan, parameter apa, atau berapa kali video itu dikompresi ulang. Ini menjadikannya *dataset* "kotak hitam" (*black box*) yang buruk untuk penelitian ilmiah.
3. *Kompresi Acak*: Kompresi (F-04) bersifat acak dan tidak terukur. Satu video mungkin dikompresi oleh WhatsApp, yang lain oleh TikTok, masing-masing dengan artefak yang berbeda. Ini mustahil untuk direplikasi atau dipelajari secara sistematis.
4. *Keseimbangan Data*: Keseimbangan data (F-07) akan sangat bergantung pada apa yang berhasil ditemukan (*sampling bias*). Kemungkinan besar *dataset*

akan sangat tidak seimbang, baik antara jumlah video asli dan palsu, maupun representasi etnisnya (F-02) yang mungkin hanya terkonsentrasi pada beberapa tokoh publik saja.

#### III.4.1.3 Alternatif Solusi 3: Pengembangan *Dataset End-to-End* Konteks Lokal

Solusi ini merupakan pendekatan paling komprehensif dan satu-satunya yang selaras dengan metodologi *data-centric*. Pendekatan ini meniru metodologi *benchmark* global seperti FaceForensics++ namun menerapkannya pada konteks Indonesia. Proses dimulai dengan Fase Pengumpulan (F-01), di mana video asli lokal yang merepresentasikan wajah dan demografi Indonesia (F-02) dikumpulkan secara strategis. Selanjutnya, pada Fase Generasi (F-03), video-video ini digunakan untuk membuat *deepfake* baru secara internal. Terakhir, baik video asli maupun video palsu yang dihasilkan akan melalui proses simulasi kompresi yang sistematis dan terukur (F-04).

Kelebihan utama dari pendekatan ini adalah kontrol penuh atas seluruh *pipeline*.

1. Kontrol Pengumpulan: Memastikan F-01 (Konteks Lokal) dan F-02 (Representasi Etnis) terpenuhi sesuai strategi pengambilan sampel pada Tabel II.2.
2. Kontrol Generasi: Memastikan F-03 (Teknik Relevan) terpenuhi dan F-06 (Metadata Lengkap) tercatat dengan akurat (misal: video A dibuat dengan *face swap*, video B dengan *lip-sync*).
3. Kontrol Kompresi: Memastikan F-04 (Simulasi Kompresi) diterapkan secara konsisten, memungkinkan pembuatan beberapa versi dari video yang sama (misal: Kualitas Tinggi, Kualitas Rendah 360p, Kualitas Rendah 144p) untuk analisis yang terisolasi.
4. Kontrol Pelabelan: Karena kita memiliki *ground truth* (video asli) untuk setiap *deepfake* yang kita buat, F-05 (Pelabelan Jelas) dan NF-03 (Validitas) terjamin 100%. Ini juga memungkinkan F-07 (Keseimbangan Data) untuk diatur secara presisi.

Meskipun solusi ini paling padat karya dan membutuhkan sumber daya komputasi yang signifikan (Risiko 3), ini adalah satu-satunya alternatif yang dapat memenuhi seluruh kebutuhan fungsional yang telah diidentifikasi dan menghasilkan *dataset* yang valid secara ilmiah.

### III.4.2 Analisis Penentuan Solusi

Berikut ini adalah Tabel III.3 yang menampilkan beberapa alternatif solusi yang diajukan untuk memenuhi kebutuhan akan *dataset deepfake* berkualitas rendah di Indonesia.

Tabel III.3 Analisis Penentuan Solusi Dataset

Kebutuhan Fungsional	Alternatif Solusi 1	Alternatif Solusi 2	Alternatif Solusi 3
F-01 (Konteks Lokal)	Tidak Terpenuhi	Terpenuhi	Terpenuhi
F-02 (Representasi Etnis)	Tidak Terpenuhi	Tidak Terkontrol	Terpenuhi
F-03 (Teknik Relevan)	Terpenuhi	Tidak Diketahui	Terpenuhi
F-04 (Simulasi Kompresi)	Terpenuhi	Kompresi Acak	Terpenuhi
F-05 (Pelabelan Jelas)	Terpenuhi	Tidak Terkontrol	Terpenuhi
F-06 (Metadata Lengkap)	Tidak Terpenuhi	Tidak Diketahui	Terpenuhi
F-07 (Keseimbangan Data)	Tergantung Dataset	Tidak Seimbang	Terpenuhi

Berdasarkan Tabel III.3, dapat disimpulkan bahwa alternatif solusi 3 (Mengumpulkan video asli lokal, membuat *deepfake* sendiri, dan menerapkan kompresi secara sistematis) adalah satu-satunya yang dapat memenuhi semua kebutuhan fungsional yang diidentifikasi. Solusi ini memberikan kontrol penuh atas konteks, jenis manipulasi, tingkat kompresi, dan pelabelan, yang sangat penting untuk tujuan penelitian.



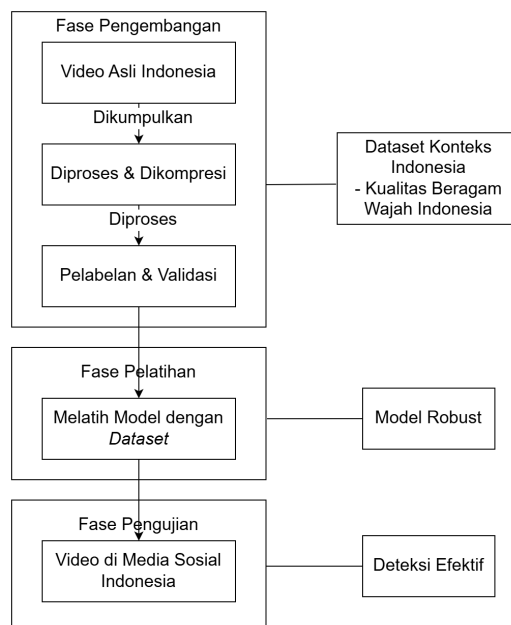
## BAB IV

### DESAIN KONSEP SOLUSI

Pada bab ini akan dijelaskan terkait konsep solusi yang ditawarkan untuk menjawab masalah yang telah dijabarkan pada bab sebelumnya.

#### IV.1 Desain Konsep Solusi

Pada bagian ini difokuskan untuk pengembangan terkait model konseptual yang dapat menjembatani antara *deepfake* yang beredar di Indonesia dengan solusi berupa *dataset* lokal yang membantu untuk mendeteksi *deepfake* tersebut. Gambar IV.1 menggambarkan model konseptual solusi yang diusulkan, kesenjangan tersebut diatasi oleh komponen-komponen yang akan dibangun dalam penelitian ini.



Gambar IV.1 Model Konseptual Solusi yang Diusulkan

Alur kerja dari model konseptual solusi yang diusulkan diawali dengan fase pengembangan. Berbeda dengan pendekatan sebelumnya yang bergantung pada dataset publik asing, proses ini dimulai dengan pengumpulan video asli Indonesia secara mandiri. Video-video tersebut kemudian melalui tahap pemrosesan dan simulasi kompresi, serta dilanjutkan dengan proses pelabelan dan validasi.

Hasil dari fase pengembangan ini adalah sebuah *dataset* konteks Indonesia yang memiliki karakteristik kualitas yang beragam (tinggi dan rendah) serta memuat representasi wajah orang Indonesia. *Dataset* inilah yang kemudian menjadi masukan utama dalam fase pelatihan untuk melatih model deteksi.

Luaran dari proses pelatihan tersebut adalah sebuah model *robust*, yang tidak hanya mengenali manipulasi wajah tetapi juga adaptif terhadap kualitas video yang rendah. Akhirnya, pada fase pengujian saat model dihadapkan pada skenario dunia nyata yaitu Video di media sosial Indonesia, model mampu melakukan generalisasi dengan baik sehingga menghasilkan deteksi efektif.

Model konseptual solusi yang diusulkan memperkenalkan satu baru di awal alur kerja yaitu fase pengembangan, pada fase ini dilakukan pengembangan *dataset* dengan konteks Indonesia yang berisi video-video dengan wajah orang Indonesia yang dilengkapi dengan berbagai kualitas. Berikut ini adalah tabel yang akan membandingkan fase-fase yang ada di kondisi saat ini dengan kondisi yang ada pada solusi yang diusulkan.

Tabel IV.1 Perbandingan Solusi dengan Kondisi Saat Ini

Fase	Kondisi Saat Ini	Kondisi Solusi
Pengembangan	<i>Dataset</i> Publik (kurang relevan dengan konteks demografis Indonesia)	<i>Dataset</i> Indonesia (wajah orang Indonesia dan kualitas menyesuaikan)
Pelatihan	Dilatih dengan <i>dataset</i> berkualitas tinggi	Dilatih dengan <i>dataset</i> dengan berbagai kualitas
Pengujian	Kurang relevan dengan kondisi nyata di Indonesia	Lebih relevan dengan kondisi nyata di Indonesia

Tabel IV.1 merangkum pergeseran pendekatan yang ditawarkan penelitian ini dibandingkan kondisi eksisting dalam tiga fase utama, yaitu pengembangan, pelatihan, dan pengujian. Pada fase pengembangan, kondisi saat ini sangat bergantung pada *dataset* publik global yang dinilai kurang relevan dengan karakteristik demografis Indonesia, sehingga solusi yang diusulkan adalah membangun *dataset* baru

yang secara spesifik memuat wajah orang Indonesia dengan kualitas yang disesuaikan. Implikasinya berlanjut ke fase pelatihan, di mana model tidak lagi hanya dilatih menggunakan data berkualitas tinggi, melainkan dipapar dengan variasi kualitas video untuk meningkatkan ketangguhan model. Akhirnya, pada fase pengujian, pendekatan ini memastikan hasil evaluasi yang jauh lebih relevan dengan skenario nyata di media sosial Indonesia dibandingkan metode terdahulu yang sering kali gagal merepresentasikan kondisi lapangan.

## BAB V

### RENCANA SELANJUTNYA

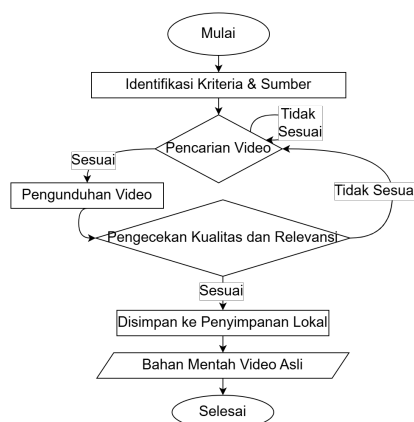
Pengerjaan tugas akhir ini merupakan kolaborasi tim, fokus penulis (Harry Truman Suhalim) adalah pada pengembangan *dataset*. *Dataset* yang dihasilkan kemudian akan menjadi masukan krusial bagi Steven Adrian Corne untuk pengembangan model deteksi, dan model tersebut selanjutnya akan diintegrasikan ke dalam pengembangan *browser extension* oleh Alvin Fadhilah Akmal.

#### V.1 Rencana Implementasi

Dikarenakan pada tugas akhir ini fokus utama saya pada *dataset*, maka fase pengembangan adalah fase yang difokuskan dalam implementasinya. Fase pengembangan *dataset* terdiri dari 4 tahapan utama yaitu:

##### V.1.1 Fase Pengumpulan *Collection*

Fase ini bertujuan untuk mengumpulkan bahan video asli yang merepresentasikan konten lokal Indonesia, sesuai dengan metodologi yang dijelaskan pada Bab I.



Gambar V.1 Fase Pengumpulan *Collection*

Pada Gambar V.1 dijelaskan terkait fase *collection* yang akan berfokus penuh pada identifikasi, pengunduhan, dan kurasi video asli. Alat bantu utama yang akan digunakan adalah untuk pengunduhan video berkualitas tinggi dan *spreadsheet* untuk melacak sumber data. Proses pengecekan kualitas dan relevansi juga mencakup validasi manual. Video akan diperiksa secara visual untuk memastikan subjek merepresentasikan keragaman etnis Indonesia seperti yang telah didefinisikan (termasuk suku-suku pada Tabel II.2). Video yang subjeknya tidak dapat diidentifikasi secara jelas atau berada di luar cakupan demografis akan dieksklusi.

Proses validasi sumber data video ini dilakukan sebagai berikut:

1. Validasi Konteks Lokal (Visual)

Video diperiksa secara visual untuk memastikan subjek utama adalah orang yang merepresentasikan demografi Indonesia, sesuai kebutuhan. Video yang subjeknya tidak jelas atau di luar konteks demografi Indonesia akan dieksklusi.

2. Validasi Representasi Etnis (Fenotipe)

Peneliti akan melakukan pelabelan kategori dominan. Subjek dalam video akan dicocokkan secara visual (fenotipe) dengan salah satu dari 13 kelompok suku bangsa yang telah diidentifikasi pada Tabel II.2.

3. Validasi Relevansi Audio (Logat)

Terkait data multimodal dan audio, saluran audio akan divalidasi. Video harus memiliki audio ucapan yang jelas dan diutamakan yang memiliki logat (aksen) kedaerahan Indonesia yang dapat diidentifikasi.

4. Validasi Kualitas Teknis

Video harus memiliki kualitas resolusi awal yang tinggi dan bebas dari artefak kompresi yang signifikan agar dapat dianggap sebagai bahan mentah yang valid.

Hanya video yang lolos keempat kriteria validasi ini yang akan diunduh dan diproses ke fase selanjutnya.

Dalam konteks penelitian ini, Video Mentah (disebut juga *raw video*, *pristine video*, atau video sumber) merujuk pada klip video digital yang otentik, asli, dan belum mengalami manipulasi digital, khususnya manipulasi *deepfake*.

Video-video ini berfungsi sebagai data pembanding atau *ground truth*. Peran utamanya dalam penelitian ini ada dua:

1. Sebagai Sampel Kelas Asli

Video ini merepresentasikan data asli dalam *dataset*. Model deteksi akan dilatih untuk mengenali video-video ini sebagai non-manipulasi.

## 2. Sebagai Sumber Pembuatan *Deepfake*

Video mentah ini digunakan sebagai bahan dasar untuk proses sintesis *deepfake*. Wajah dalam video mentah ini akan diekstraksi dan digantikan untuk menghasilkan video palsu yang akan menjadi bagian dari *dataset*.

Pengambilan klip video dari *platform* publik seperti YouTube untuk penyusunan *dataset* penelitian ini didasarkan pada prinsip *fair use*, sebagaimana diakui dalam hukum hak cipta AS dan dijelaskan dalam pedoman Google/YouTube.

Menurut UU RI No. 28 Tahun 2014 tentang Hak Cipta, penggunaan materi berhak cipta untuk penyusunan Tugas Akhir ini dapat dibenarkan dan tidak termasuk dalam pelanggaran Hak Cipta.

Dasar hukum utama untuk justifikasi ini terdapat pada Pasal 43 (a), yang menyatakan bahwa pengambilan Ciptaan untuk "kepentingan: pendidikan, penelitian, penulisan karya ilmiah..." tidak dianggap sebagai pelanggaran Hak Cipta, selama tidak merugikan kepentingan yang wajar dari Pencipta.

Hal ini diperkuat oleh Pasal 44 ayat (1), yang secara eksplisit membolehkan penggunaan tersebut dengan syarat:

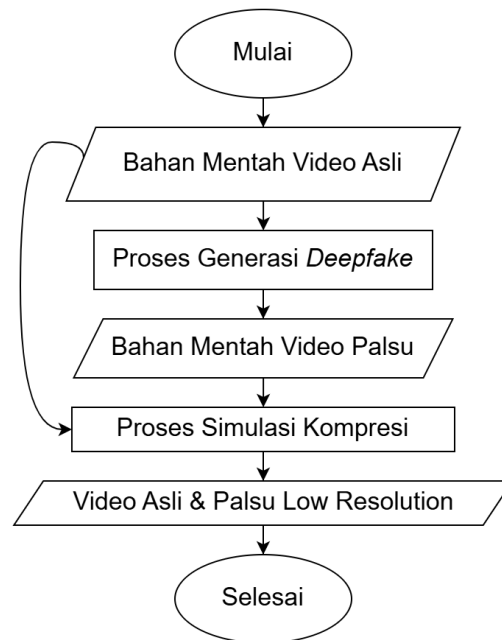
1. Penggunaannya tidak komersial; dan
2. Tidak merugikan kepentingan yang wajar dari Pencipta.

Aspek Etika dan Perlindungan Data Pribadi "Selain aspek hak cipta, penelitian ini juga memperhatikan aspek privasi subjek data sesuai dengan UU No. 27 Tahun 2022 tentang Perlindungan Data Pribadi (UU PDP). Wajah merupakan data biometrik yang bersifat sensitif. Oleh karena itu, langkah mitigasi etika yang dilakukan meliputi:

1. Penggunaan Data Publik: Video dikumpulkan hanya dari tokoh publik atau konten kreator yang telah mempublikasikan wajahnya secara sadar di domain publik (YouTube), bukan dari akun privat.
2. Tujuan Non-Destruktif: Manipulasi *deepfake* dilakukan semata-mata untuk pelatihan algoritma deteksi (defensif), bukan untuk pencemaran nama baik atau penipuan.
3. Transparansi *Dataset*: *Dataset* yang dihasilkan tidak akan disebarluaskan secara bebas tanpa *Data Use Agreement* yang ketat."

### V.1.2 Fase Generasi dan Pemrosesan (*Construction*)

Fase ini merupakan inti teknis dari pengembangan *dataset*. Pada tahap ini, video *deepfake* akan dibuat dan simulasi kompresi akan diterapkan untuk menghasilkan video berkualitas rendah.

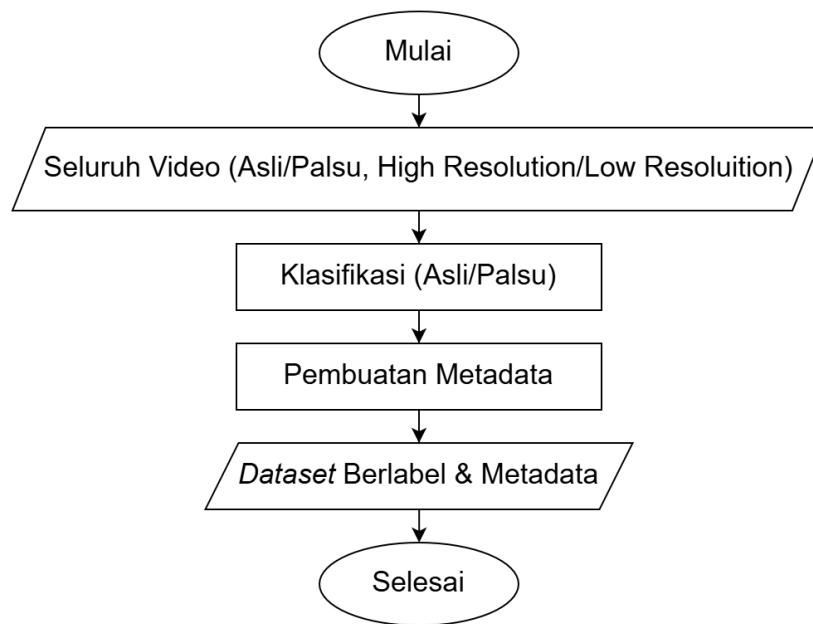


Gambar V.2 Fase Generasi dan Pemrosesan (*Construction*)

Pada Gambar V.2 dijelaskan terkait fase *construction* yang akan dilakukan generasi video-video *deepfake* dari video yang telah dikumpulkan pada fase sebelumnya. Setelah itu, video *deepfake* dan video asli akan dilakukan kompresi untuk mendapatkan video *low quality*. Video berkualitas rendah di sini adalah video dengan kualitas minimal 144p dalam konteks YouTube dan kualitas-kualitas di atasnya yang tetap merepresentasikan video berkualitas rendah.

### V.1.3 Fase Pelabelan(*Labelling*)

Pada fase ini, setiap data yang telah diproses akan diberi label informatif (asli/palsu) dan metadata yang detail untuk memberikan konteks bagi model *machine learning*.



Gambar V.3 Fase Pelabelan (*Labelling*)

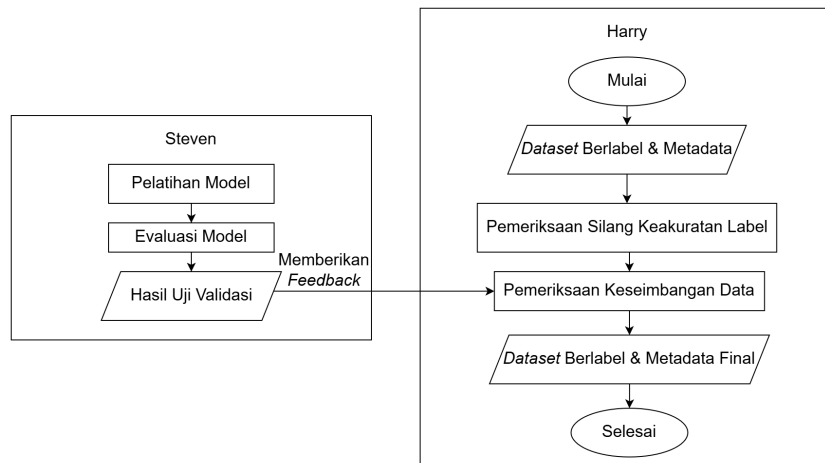
Pada Gambar V.3 dijelaskan terkait fase *labelling* yang akan dibuat struktur folder yang sistematis dan file metadata yang mencatat informasi detail untuk setiap file video.

Untuk memenuhi kebutuhan Representasi Etnis dan menjaga konsistensi pelabelan, penelitian ini menggunakan pendekatan Pelabelan Kategori Dominan. Setiap video akan dievaluasi secara visual dan dimasukkan ke dalam satu kategori tunggal dari kelompok suku bangsa yang telah diidentifikasi pada Tabel II.2. Kategori yang dipilih adalah yang paling merepresentasikan fitur visual (fenotipe) dominan dari subjek dalam video.

#### V.1.4 Fase Validasi (*Transition*)

Fase terakhir ini bertujuan untuk memastikan kualitas, integritas, dan keseimbangan dataset sebelum diserahkan untuk pelatihan model. Fase ini mirip dengan fase *transition* pada pengembangan perangkat lunak, di mana produk diuji oleh pengguna akhirnya.





Gambar V.4 Fase Validasi (*Transition*)

Pada Gambar V.4 dijelaskan terkait fase *transition* yang akan dilakukan pengujian akhir pada *dataset* untuk memastikan kualitas dan keseimbangan data sebelum diserahkan untuk pelatihan model deteksi dengan validasi internal dan validasi eksternal sebagai berikut.

1. Validasi internal (dilakukan oleh peneliti *dataset*/Harry):
  - a. Pemeriksaan keakuratan label: dilakukan *cross-check* untuk memastikan video ‘Asli’ benar-benar asli dan ‘Palsu’ benar-benar palsu.
  - b. Pemeriksaan keseimbangan data: diperiksa bahwa distribusi data seimbang, baik antar kelas (asli vs palsu) maupun antar kualitas (tinggi vs rendah).
2. Validasi eksternal (dilakukan oleh pengguna/Steven):
  - a. Validasi umpan balik model: *dataset* diserahkan kepada Steven untuk melatih model.
  - b. Kinerja model sebagai tolak ukur: hasil uji validasi dari performa model Steven akan menjadi *feedback* langsung atas kualitas *dataset*. Jika model yang dilatih Steven menunjukkan performa yang baik dan *robust*, maka *dataset* Anda dapat divalidasi sebagai berkualitas dan berhasil.

## V.2 Analisis Risiko dan Mitigasi

Pengerjaan tugas akhir selanjutnya tentunya tidak akan lepas dari risiko-risiko yang ada, baik dalam aspek sumber daya manusia, teknologi, proses, maupun finansial. Daftar risiko dalam pengerjaan tugas akhir ini beserta mitigasinya dijelaskan lebih

lanjut pada Tabel V.1 berikut ini.

Tabel V.1 Analisis Risiko dan Mitigasi

No	Risiko	Penyebab	Dampak	Mitigasi
1	Kesulitan Pengumpulan Video Lokal	Keterbatasan video asli yang berkualitas tinggi di platform publik.	Dataset tidak representatif, bias terhadap demografi tertentu.	1. Memperluas sumber pencarian (arsip berita daerah, video pariwisata).
2	Kualitas Generasi Deepfake Buruk	Alat <i>open-source</i> sulit dikonfigurasi; hasil terlihat tidak realistis.	Dataset terlalu mudah dideteksi, model tidak belajar artefak yang halus.	1. Alokasi waktu di awal untuk eksperimen alat. 2. Menggunakan <i>pre-trained model</i> yang sudah ada. 3. Memecah video menjadi klip pendek untuk hasil lebih stabil.
3	Kebutuhan Komputasi (GPU) Tinggi	Proses generasi <i>deepfake</i> membutuhkan daya komputasi GPU yang besar dan waktu yang lama.	Fase Generasi terhambat, pengerjaan TA terlambat.	1. Melakukan <i>downsampling</i> video sebelum diproses jika diperlukan. 2. Menggunakan layanan <i>cloud computing</i> (Google Colab Pro, AWS) sebagai anggaran cadangan.
4	Simulasi Kompresi Tidak Realistis	Parameter tidak secara akurat mencerminkan kompresi asli.	Kesenjangan antara <i>dataset</i> dan dunia nyata tetap ada, model gagal generalisasi.	1. Melakukan studi pendahuluan ( <i>upload/download</i> sampel video) untuk menganalisis <i>bitrate</i> asli. 2. Menerapkan <i>feedback loop</i> dengan Steven (model) untuk menguji kompresi.
5	Ketidakselarasan dengan Tim Model	<i>Blocking dependency</i> : Pekerjaan Steven (Model) tidak bisa dimulai sebelum pekerjaan <i>Dataset</i> 100% selesai.	Keterlambatan penyelesaian <i>dataset</i> akan menghambat seluruh kemajuan tim.	1. Menerapkan rilis <i>dataset</i> secara inkremental. 2. Menjadwalkan sinkronisasi mingguan wajib dengan Steven untuk memvalidasi <i>batch</i> data yang sudah selesai.

## DAFTAR PUSTAKA

- Alrashoud, Mubarak. 2025. "Deepfake video detection methods, approaches, and challenges". *Alexandria Engineering Journal*.
- Anugrah, Nur Syifa, Haidar Adilah, Danny Putranto Panggabean, dan Rizky Prabowo Laksana. 2025. "ANALISIS KECEPATAN INTERNET SELULER DI INDONESIA BERDASARKAN REGIONAL". *Journal of Data Analytics, Information, and Computer Science (JDAICS)*.
- Badan Pusat Statistik. 2012. *Kewarganegaraan, Suku Bangsa, Agama, dan Bahasa Sehari-hari Penduduk Indonesia: Hasil Sensus Penduduk 2010*. Technical report.
- Banerjee, S., S. K. Yadav, A. Dhara, dan M. Ajij. 2024. "A Survey: Deepfake and Current Technologies for Solutions". Dalam *The 2024 Sixth Doctoral Symposium on Intelligence Enabled Research (DoSIER 2024)*.
- BarTech TV. 2022. *From 144P to 8K. Every Resolution Compared*. Video YouTube. [Daring]. Tersedia: <https://www.youtube.com/watch?v=bTqVNM9ztLI>.
- Bonet-Jover, A., R. Sepúlveda-Torres, E. Saquete, P. Martínez-Barco, A. Piad-Morffis, dan S. Estevez-Velarde. 2023. "Applying Human-in-the-Loop to construct a dataset for determining content reliability to combat fake news". *Engineering Applications of Artificial Intelligence*.
- Ciftci, Umur Aybars, İlke Demir, dan Lijun Yin. 2020. "FakeCatcher: Detection of Synthetic Portrait Videos using Biological Signals". *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- Croitoru, Florinel-Alin, dkk. 2024. *Deepfake Media Generation and Detection in the Generative AI Era: A Survey and Outlook*.
- Dolhansky, B., J. Bitton, B. Pflaum, J. Lu, R. Howes, M. Wang, dan C. C. Ferrer. 2020. *The deepfake detection challenge (DFDC) dataset*.

- Dolhansky, B., R. Howes, B. Pflaum, N. Baram, dan C. C. Ferrer. 2019. *The deepfake detection challenge (DFDC) preview dataset*.
- Dufour, N., dan A. Gully. 2019. *Contributing data to deepfake detection research*. Google AI Blog.
- Fernandes, Yoggy Arif, dan Yulia Fatma. 2025. "METODE DEEP LEARNING DALAM TEKNOLOGI DEEPPFAKE: SYSTEMATIC LITERATURE REVIEW". *JATI (Jurnal Mahasiswa Teknik Informatika)*.
- Firmani, D., P. Merialdo, E. Nieddu, A. Rossi, dan R. Torlone. 2020. "Crowdsourcing for Building Knowledge Graphs at Scale from the Vatican Archives (Discussion Paper)". Dalam *Proceedings of the 28th Italian Symposium on Advanced Database Systems (SEBD 2020)*.
- Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé III, dan Kate Crawford. 2021. *Datasheets for Datasets*.
- He, Y., dkk. 2021. "ForgeryNet: A versatile benchmark for comprehensive forgery analysis". Dalam *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.
- Jabbar, N., A. Saghir, A. A. Meer, S. A. Sahi, B. Hassan, dan S. M. Yasir. 2023. "Media Forensics and Deepfake - Systematic Survey". *RS Open Journal on Innovative Communication Technologies*.
- Jiang, L., R. Li, W. Wu, C. Qian, dan C. C. Loy. 2020. "DeeperForensics-1.0: A large-scale dataset for real-world face forgery detection". Dalam *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.
- Kementerian Komunikasi dan Digital. 2025. *Deepfake Naik 550%, Kemkomdigi Minta Platform Global Sediakan Fitur Cek Konten AI*. Siaran Pers. [Daring]. Tersedia: <https://www.komdigi.go.id/berita/siaran-pers/detail/deepfake-naik-550-kemkomdigi-minta-platform-global-sediakan-fitur-cek-konten-ai>.
- Korshunov, P., dan S. Marcel. 2018. *Deepfakes: A new threat to face recognition? assessment and detection*.
- Kwon, P., J. You, G. Nam, S. Park, dan G. Chae. 2021. "KODF: A large-scale Korean deepfake detection dataset". Dalam *Proceedings of the IEEE/CVF International Conference on Computer Vision*.

- Li, M., B. Liu, Y. Hu, dan Y. Wang. 2021. "Exposing deepfake videos by tracking eye movements". Dalam *2020 25th International Conference on Pattern Recognition (ICPR)*.
- Li, Y., dan S. Lyu. 2019. "Exposing DeepFake Videos By Detecting Face Warping Artifacts". Dalam *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*.
- Li, Y., X. Yang, P. Sun, H. Qi, dan S. Lyu. 2019. *Celeb-DF: A new dataset for Deepfake Forensics*.
- . 2020. "Celeb-DF: A large-scale challenging dataset for deepfake forensics". Dalam *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*.
- Montibeller, Andrea, Dasara Shullani, Davide Baracchi, Alessandro Piva, dan Giulia Boato. 2025. "Bridging the Gap: A Framework for Real-World Video Deepfake Detection via Social Network Compression Emulation". Dalam *Proceedings of the 1st Deepfake Forensics Workshop: Detection, Attribution, Recognition, and Adversarial Challenges in the Era of AI-Generated Media (DFF '25)*.
- Orr, Will, dan Kate Crawford. 2024. *Building Better Datasets: Seven Recommendations for Responsible Design from Dataset Creators*.
- Rössler, Andreas, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, dan Matthias Nießner. 2019. "FaceForensics++: Learning to detect manipulated facial images". Dalam *Proceedings of the IEEE/CVF international conference on computer vision*.
- Rössler, Andreas, Davide Cozzolino, Luisa Verdoliva, Christian Riess, Justus Thies, dan Matthias Nießner. 2018. *FaceForensics: A large-scale video dataset for forgery detection in human faces*.
- Sahu, L. N., R. K. Namdeo, S. Gupta, dan P. Singh. 2025. "Benchmarking DeepFake Detection on Social Media: Real-World Dataset and Case Study". *Research Square*.
- Sajati, Haruno, Asih Pujiastuti, dan Afrizal Triantoro. 2018. "THE AUDIO VIDEO OF WEB-BASED COMPRESSION WITH FFMPEG". *COMPILER*.

Sohan, Md Fahim, Mohammad Solaiman, dan Md Anwarul Hasan. 2023. “A survey on deepfake video detection datasets”. *Indonesian Journal of Electrical Engineering and Computer Science*.

Vital Brazil, Elizabete, Eduardo Soares, Luiz Villa Real, Lucas Azevedo, Vanesa Segura, Leonardo Zerkowski, dan Renato Cerqueira. 2023. *Position Paper on Dataset Engineering to Accelerate Science*.

Zha, Daochen, Zaid Pervaiz Bhat, Kwei-Herng Lai, Fan Yang, Zhimeng Jiang, Shaochen Zhong, dan Xia Hu. 2023. *Data-centric Artificial Intelligence: A Survey*.

Zi, B., M. Chang, J. Chen, X. Ma, dan Y.-G. Jiang. No date. “WildDeepfake: A challenging real-world dataset for deepfake detection”. Dalam *Proceedings of the 28th ACM International Conference on Multimedia*.