

Analysis Of Artificial Intelligence Deepfake Risks Regarding The Public's Ability To Distinguish Between Real And Ai-Generated Content Using A Quantitative Descriptive Method

Muhammad Rizki Akbar ¹⁾; Pahrizal ²⁾; Yulia Darmi ³⁾; Rozali Toyib ⁴⁾

^{1,2,3,4)}Program Studi Teknik Informatika, Fakultas Teknik, Universitas Muhammadiyah Bengkulu

Email: ¹⁾ muhammadrrizkiakbarr@gmail.com

How to Cite :

Akbar, M. R., Pahrizal, Darmi, Y., Toyib, R. (2026). Analysis of Artificial Intelligence Deepfake Risks Regarding The Public's Ability to Distinguish Between Real and AI-Generated Content Using A Quantitative Descriptive Method. Jurnal Media Computer Science, 5 (4).

ARTICLE HISTORY

Received [31 Agustus 2026]

Revised [25 Oktober 2026]

Accepted [28 Oktober 2026]

KEYWORDS

Deepfake, Artificial Intelligence, Detection Ability, Knowledge, Quantitative Descriptive.

This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license



ABSTRAK

Perkembangan teknologi *Artificial Intelligence* (AI) melahirkan teknologi *Deepfake* yang mampu menghasilkan konten visual menyerupai kenyataan, sehingga menimbulkan risiko disinformasi dan menyulitkan masyarakat membedakan konten nyata dan buatan AI. Penelitian ini bertujuan menganalisis tingkat kemampuan masyarakat membedakan konten tersebut serta hubungan antara tingkat pengetahuan masyarakat tentang *Deepfake* AI (X1) dengan kemampuan mendeteksi konten *Deepfake* (Y1). Penelitian menggunakan metode deskriptif kuantitatif dengan 100 responden pengguna media sosial usia 17-30 tahun, dikumpulkan daring melalui *Google Forms* pada 24 Juni-24 Juli 2026. Instrumen terdiri atas lima pernyataan Likert untuk mengukur X1 dan sepuluh pasangan foto asli-*Deepfake* untuk mengukur Y1. Hasil menunjukkan kemampuan deteksi tinggi (rata-rata 77,1%) dan pengetahuan tinggi (rata-rata 3,87 skala 1-5). Kedua instrumen valid dan reliabel (Cronbach Alpha X1=0,834; KR-20 Y1=0,712). Uji korelasi Pearson menunjukkan tidak terdapat hubungan signifikan antara X1 dan Y1 ($r=-0,028$; $p=0,784>0,05$), dikonfirmasi regresi linear sederhana ($R\text{ Square}=0,001$), sehingga H_0 diterima; pengetahuan konseptual terbukti tidak otomatis meningkatkan kemampuan deteksi. Penelitian merekomendasikan peningkatan literasi digital, regulasi watermarking konten AI, dan penyempurnaan instrumen deteksi pada penelitian selanjutnya.

ABSTRACT

The development of Artificial Intelligence (AI) technology has given rise to Deepfake technology capable of producing visual content resembling reality, creating a risk of disinformation and making it difficult for the public to distinguish real from AI-generated content. This study aims to analyze the public's ability to distinguish such content and the relationship between the public's knowledge of Deepfake AI (X1) and their ability to detect Deepfake content (Y1). This study uses a quantitative descriptive method with 100 social media user respondents aged 17-30 years, collected online via Google Forms from June 24 to July 24, 2026. The instrument consisted of five Likert-scale statements measuring X1 and ten pairs of real-Deepfake photos measuring Y1. Results show that detection ability is high (average 77.1%) and knowledge level is also high (average 3.87 on a 1-5 scale). Both instruments were valid and reliable (Cronbach's Alpha for X1 = 0.834; KR-20 for Y1 = 0.712). A Pearson correlation test showed no significant relationship between X1 and Y1 ($r = -0.028$; $p = 0.784 > 0.05$), confirmed by simple linear regression ($R\text{ Square} = 0.001$), so the null hypothesis (H_0) is accepted; conceptual knowledge was shown not to automatically improve actual detection ability. This study recommends improving digital literacy, AI content watermarking regulations, and refining the detection instrument in future research.

PENDAHULUAN

Perkembangan teknologi kecerdasan buatan atau *Artificial Intelligence (AI)* telah membawa perubahan besar dalam berbagai aspek kehidupan manusia, terutama pada bidang komunikasi digital dan media informasi. Salah satu perkembangan *AI* yang saat ini banyak dibahas adalah teknologi *Deepfake*. Teknologi *Deepfake* merupakan teknologi berbasis *Deep learning* yang mampu memanipulasi gambar, audio, maupun video sehingga terlihat seperti asli. Teknologi ini dapat digunakan untuk membuat konten palsu yang sangat realistis dengan cara mengganti wajah, suara, atau gerakan seseorang menggunakan algoritma *AI* (Payne, 2026).

Kemajuan yang sangat pesat dalam bidang kecerdasan buatan (*Artificial Intelligence/AI*) saat ini telah berhasil menghadirkan suatu teknologi digital yang semakin canggih dan sulit untuk dibedakan dari konten-konten asli (Raharjo, 2023). Teknologi digital ini kemudian dikenal dengan istilah *Deepfake*. Perkembangan teknologi *AI* yang kian maju ini telah memicu munculnya ancaman-ancaman baru yang sangat serius bagi keamanan digital secara global, khususnya dalam konteks permasalahan penipuan identitas digital. Keberadaan *Deepfake* yang semakin sulit untuk dideteksi dan dibedakan dari konten asli telah menjadi salah satu tantangan utama yang harus dihadapi dalam upaya menjaga keamanan serta integritas identitas digital di tengah era teknologi yang terus berkembang pesat saat ini. penyalahgunaan teknologi *Deepfake* ini telah menimbulkan berbagai macam tantangan hukum dan etika yang cukup serius. Sehingga memunculkan banyak pertanyaan mengenai batasan-batasan hukum yang harus diterapkan dalam menangani kasus-kasus penipuan identitas digital yang disebabkan oleh kehadiran teknologi *Deepfake*. (Puspita, L., & Firdaus, H. S. 2025).

Dalam konteks media sosial ini, berita palsu semakin meluas karena kemajuan teknologi informasi yang memungkinkan siapa pun untuk dengan mudah berbagi informasi. Media sosial menyediakan platform terbuka untuk pertukaran informasi, tetapi tidak ada pengawasan atau penyaringan yang ketat. Oleh karena itu, media sosial telah menjadi saluran utama penyebaran berita palsu di dalam komunitas digital. (Sabela, I. N. D., Saputra, A., & Kuncoro, W., 2026).

Menurut konsep dasarnya, *Deepfake* merupakan media sintetis yang dihasilkan menggunakan teknologi *Artificial Intelligence (AI)* untuk menampilkan sesuatu yang sebenarnya tidak pernah terjadi. Teknologi ini umumnya dibuat menggunakan metode *Generative Adversarial Network (GAN)*, yaitu dua model *AI* yang saling bekerja untuk menghasilkan konten palsu yang semakin realistis dan sulit dibedakan dari konten asli. Perkembangan teknologi *Deepfake* membuat manipulasi gambar, audio, dan video menjadi semakin canggih sehingga berpotensi menimbulkan penyebaran disinformasi dan hoaks di media digital (Altuncu, Franqueira, dan Li, 2024).

AI juga merupakan bidang ilmu komputer yang bertujuan untuk menciptakan mesin atau program komputer yang dapat melakukan tugas yang memerlukan kecerdasan manusia. (Juditha, C., 2025).

Teknologi *Deepfake* memungkinkan manipulasi konten visual dan audio untuk menciptakan media yang sangat realistis namun palsu, seperti video atau gambar yang menampilkan seseorang melakukan atau mengatakan sesuatu yang sebenarnya tidak pernah terjadi. Teknologi ini memanfaatkan algoritma *AI* untuk menggantikan wajah atau suara seseorang dengan orang lain, menghasilkan konten yang sulit dibedakan dari yang asli oleh pengamat manusia. Meskipun awalnya digunakan dalam konteks hiburan dan seni, *Deepfake* kini menjadi perhatian serius karena potensi penyalahgunaannya dalam menyebarkan informasi palsu dan propaganda (Darmawan, Junaidi, dan Khaerudin, 2025).

Deepfake memanfaatkan *Generative Adversarial Network (GAN)*, yang bekerja dengan jaringan saraf untuk menganalisis sejumlah besar data guna meniru ekspresi wajah, perilaku, suara, manusia. Istilah *Deepfake* berasal dari gabungan kata *Deep learning*, yang merujuk pada teknologi pembelajaran mesin yang mendalam, dan *fake*, yang berarti palsu. Kemunculan *Deepfake* pertama kali digunakan untuk memberikan kemudahan pada manusia dalam menciptakan hasil foto atau

video yang lebih nyata dalam konten hiburan seperti film ataupun game. Namun seiring berjalannya waktu, teknologi ini mulai disalah gunakan untuk berbagai kepentingan negatif (Seveney, Wicaksono, dan Soetijono, 2025).

Teknologi ini dapat digunakan untuk membuat video palsu yang memperlihatkan tokoh mengatakan sesuatu yang sebenarnya tidak pernah mereka ucapkan. Pengguna dapat dengan mudah menciptakan video palsu yang membuat tokoh terkenal terlihat seperti berkata atau melakukan sesuatu yang sebenarnya tidak pernah terjadi. Selain itu, teknologi ini juga bisa dimanfaatkan untuk menyebarkan disinformasi di media sosial seperti Facebook, Instagram, dan WhatsApp, serta digunakan oleh oknum-oknum untuk mempengaruhi atau memanipulasi konten yang tersebar secara meluas sehingga dapat mengakibatkan ancaman kepada masyarakat, mulai dari tindakan kriminalitas seperti penipuan, pencemaran nama baik, menimbulkan perpecahan, dan lain-lain (Fadhilah dan Retnoningsih, 2024).

LANDASAN TEORI

Kecerdasan Buatan (*Artificial Intelligence*)

Artificial Intelligence (AI) merupakan bidang ilmu komputer yang bertujuan untuk menciptakan mesin atau program komputer yang dapat melakukan tugas yang memerlukan kecerdasan manusia, seperti pengenalan pola, pengambilan keputusan, dan pemrosesan bahasa alami (Juditha, 2025). *AI* telah diterapkan secara luas pada berbagai bidang, mulai dari peramalan cuaca, sistem penerjemahan, sistem rekomendasi, hingga pembuatan konten digital.

Deepfake

Deepfake merupakan media sintesis yang dihasilkan menggunakan teknologi *AI* untuk menampilkan sesuatu yang sebenarnya tidak pernah terjadi. Istilah *Deepfake* berasal dari gabungan kata *Deep learning*, yang merujuk pada teknologi pembelajaran mesin yang mendalam, dan *fake* yang berarti palsu. Kemunculan *Deepfake* pertama kali digunakan untuk memberikan kemudahan bagi manusia dalam menciptakan hasil foto atau video yang lebih nyata dalam konten hiburan seperti film atau gim, namun seiring berjalannya waktu teknologi ini mulai disalah gunakan untuk berbagai kepentingan negatif (Seveney, Wicaksono, & Soetijono, 2025).

Generative Adversarial Network (GAN)

GAN merupakan metode utama yang digunakan dalam pembuatan konten *Deepfake*, yang bekerja dengan menggunakan dua model jaringan saraf tiruan yang saling bersaing, yaitu generator yang menghasilkan konten palsu dan discriminator yang berupaya membedakan konten palsu dengan konten asli. Proses ini berlangsung berulang kali sehingga menghasilkan konten yang semakin realistis dan sulit dibedakan dari konten asli (Altuncu, Franqueira, & Li, 2024)

Disinformasi dan Hoaks Digital

Disinformasi merupakan informasi yang sengaja dibuat salah dengan tujuan menyesatkan atau memanipulasi opini publik. Media sosial menjadi saluran utama penyebaran hoaks karena keterbukaan platform dalam pertukaran informasi tanpa pengawasan dan penyaringan yang ketat (Sabela, Saputra, & Kuncoro, 2026)

Literasi Digital

Literasi digital merupakan kemampuan seseorang dalam memahami, mengevaluasi, dan menggunakan informasi digital secara kritis. Tingkat literasi digital menjadi faktor penentu kemampuan seseorang dalam mengenali dan menghindari dampak negatif dari konten manipulatif, termasuk *Deepfake* (Payne, 2026).

Metode Deskriptif Kuantitatif

Metode penelitian kuantitatif merupakan pendekatan ilmiah yang bertujuan untuk mengukur, menjelaskan, dan menguji hubungan antar variabel melalui data numerik yang dianalisis secara statistik, sehingga memungkinkan peneliti menarik kesimpulan yang objektif, teruji, dan dapat di generalisasikan ke populasi yang lebih luas (Hair et al., 2022). Metode deskriptif kuantitatif secara khusus bertujuan untuk menggambarkan dan menganalisis fenomena secara sistematis berdasarkan data numerik yang dikumpulkan dari lapangan.

Kuesioner sebagai Instrumen Penelitian

Kuesioner merupakan instrumen pengumpulan data yang berisi serangkaian pertanyaan atau pernyataan yang disusun secara sistematis untuk memperoleh informasi dari responden. Dalam penelitian ini, kuesioner disusun dengan menyertakan materi stimulus berupa pasangan konten (asli dan *Deepfake*) yang harus dinilai oleh responden, disertai pertanyaan mengenai pengetahuan dan pengalaman responden terkait teknologi *Deepfake* (Mamuaya et al., 2025).

Pengetahuan tentang Teknologi *Deepfake*

Pengetahuan tentang teknologi *Deepfake* mengacu pada sejauh mana seseorang memahami konsep, cara kerja, serta ciri-ciri konten yang dihasilkan oleh teknologi *Deepfake*. Pemahaman ini mencakup kesadaran akan keberadaan teknologi *Deepfake*, pengetahuan bahwa konten tersebut dihasilkan melalui algoritma AI seperti *Generative Adversarial Network (GAN)*, serta kemampuan mengenali indikator visual maupun audio yang membedakan konten asli dari konten manipulasi (Payne, 2026; Juditha, 2025). Semakin tinggi tingkat pengetahuan seseorang mengenai karakteristik dan mekanisme kerja *Deepfake*, semakin besar pula potensi mereka untuk lebih waspada dan kritis ketika mengonsumsi konten digital. Dalam penelitian ini, pengetahuan tentang *Deepfake* diposisikan sebagai variabel bebas (X1) yang diduga berkontribusi terhadap kemampuan deteksi konten *Deepfake*.

Kemampuan Deteksi Konten Digital

Kemampuan deteksi konten digital merupakan kesanggupan seseorang untuk mengidentifikasi dan membedakan konten asli dari konten hasil manipulasi atau rekayasa digital. Kemampuan ini dipengaruhi oleh berbagai faktor, termasuk tingkat literasi digital, pengalaman terhadap teknologi, serta pengetahuan mengenai ciri-ciri teknis konten manipulatif (Altuncu, Franqueira, & Li, 2024). Semakin realistis hasil manipulasi yang dihasilkan oleh teknologi AI, semakin sulit pula bagi individu untuk mendeteksinya secara visual tanpa bantuan pengetahuan atau alat bantu teknis (Payne, 2026). Dalam penelitian ini, kemampuan deteksi diposisikan sebagai variabel terikat (Y1) yang diukur secara empiris melalui uji stimulus, yaitu dengan meminta responden mengidentifikasi konten asli dari pasangan konten asli dan *Deepfake* yang disajikan dalam kuesioner.

METODE PENELITIAN

Metode Penelitian

Penelitian menggunakan metode deskriptif kuantitatif dengan 100 responden pengguna media sosial usia 17-30 tahun, dikumpulkan daring melalui *Google Forms* pada 24 Juni-24 Juli 2026. Instrumen terdiri atas lima pernyataan Likert untuk mengukur X1 dan sepuluh pasangan foto asli-*Deepfake* untuk mengukur Y1.

Definisi Operasional Variabel

Berdasarkan hipotesis penelitian, terdapat dua variabel yang dioperasionalkan dalam penelitian ini, yaitu variabel bebas (X1) dan variabel terikat (Y1). Definisi operasional masing-masing variabel disajikan pada Tabel 1. berikut.

Tabel 1 variabel bebas (X1) dan variabel terikat (Y1)

Variabel	Definisi Operasional	Indikator	Skala	Jumlah Item
X1 - Pengetahuan tentang Deepfake AI	Tingkat pemahaman responden mengenai konsep, cara kerja, dan ciri-ciri konten <i>Deepfake</i>	PT1-PT5	Likert 1-5	5
Y1 - Kemampuan Deteksi Konten Deepfake	Ketepatan responden mengidentifikasi konten asli vs. konten <i>Deepfake</i> pada 10 pasangan stimulus	S1-S10 (skor benar/salah), K1-K10 (tingkat keyakinan)	Nominal + Ordinal	10

Skor variabel X1 dihitung dari total skor lima pernyataan Likert (PT1-PT5). Skor variabel Y1 dihitung dari jumlah jawaban benar pada sepuluh item stimulus (S1-S10) dibagi sepuluh, dikalikan seratus, sehingga diperoleh persentase tingkat kemampuan deteksi tiap responden.

Teknik Pengumpulan Data

Kuesioner disebarakan secara daring melalui platform *Google Forms* dengan menyertakan materi stimulus berupa pasangan konten (Asli dan *Deepfake*) yang harus dinilai oleh responden. Responden diminta untuk mengidentifikasi mana konten asli dan mana yang merupakan hasil *Deepfake*, serta menjawab pertanyaan tentang pengetahuan dan pengalaman mereka terkait teknologi *Deepfake*.

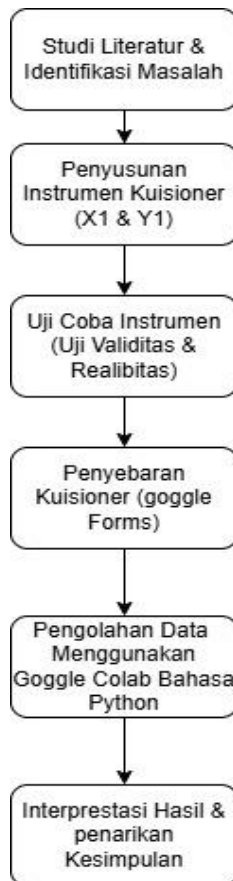
Teknik Analisis Data

Data yang terkumpul diolah menggunakan platform *Google Colab* dengan Menggunakan Bahasa *Python*. Analisis yang dilakukan meliputi analisis frekuensi dan persentase untuk menggambarkan karakteristik responden dan tingkat kemampuan deteksi *Deepfake*, analisis deskriptif untuk mengidentifikasi pola dan tren, serta analisis korelasi untuk mengidentifikasi hubungan antara variabel pengetahuan masyarakat tentang *Deepfake AI* dengan kemampuan deteksi *Deepfake*.

Sebelum uji korelasi dilakukan, data terlebih dahulu diuji normalitasnya menggunakan uji Kolmogorov-Smirnov atau Shapiro-Wilk untuk menentukan jenis uji korelasi yang sesuai. Apabila data berdistribusi normal, digunakan uji korelasi *Pearson Product Moment*; apabila tidak berdistribusi normal, digunakan uji korelasi Rank Spearman. Selanjutnya, dilakukan uji regresi linear sederhana untuk mengukur besarnya kontribusi variabel pengetahuan (X1) terhadap variabel kemampuan deteksi (Y1). Pengujian hipotesis dilakukan pada taraf signifikansi 5% ($\alpha = 0,05$), dengan kriteria H_0 ditolak apabila nilai signifikansi (p-value) kurang dari 0,05.

Interpretasi Hasil dan Penarikan Kesimpulan

Hasil analisis di interpretasikan secara komprehensif untuk menjawab pertanyaan penelitian, kemudian dirumuskan kesimpulan dan rekomendasi yang dapat menjadi masukan bagi pengembangan literasi digital masyarakat dan kebijakan penanganan risiko *Deepfake* di Indonesia.



Gambar 1 Tahapan Alur Metode Pelaksanaan

HASIL DAN PEMBAHASAN

Hasil

Hasil menunjukkan kemampuan deteksi tinggi (rata-rata 77,1%) dan pengetahuan tinggi (rata-rata 3,87 skala 1-5). Kedua instrumen valid dan reliabel (Cronbach Alpha $X1=0,834$; KR-20 $Y1=0,712$). Uji korelasi Pearson menunjukkan tidak terdapat hubungan signifikan antara $X1$ dan $Y1$ ($r=-0,028$; $p=0,784>0,05$), dikonfirmasi regresi linear sederhana ($R\text{ Square}=0,001$), sehingga H_0 diterima; pengetahuan konseptual terbukti tidak otomatis meningkatkan kemampuan deteksi. Penelitian merekomendasikan peningkatan literasi digital, regulasi watermarking konten AI, dan penyempurnaan instrumen deteksi pada penelitian selanjutnya.

Pembahasan

Sub bab ini menyajikan pembahasan secara komprehensif mengenai makna implikasi dari temuan empiris penelitian. Pembahasan dialurkan berdasarkan hasil analisis statistik yang telah dilakukan, kaitan dengan landasan teori serta penelitian terdahulu, serta implikasi praktis bagi berbagai pihak terkait.

Contoh Penerapan Rumus dan Hasil Pengujian (Sampel)

Untuk memperjelas cara penerapan rumus-rumus statistik yang telah diuraikan pada BAB III, berikut disajikan contoh perhitungan manual pada sampel data, yaitu satu contoh penerapan yang memenuhi kriteria (benar/valid) dan satu contoh ilustrasi yang tidak memenuhi kriteria (tidak valid), sebelum disajikan tabel rekapitulasi hasil pengujian secara keseluruhan.

Contoh Penerapan pada Item yang Memenuhi Kriteria (VALID)

Diambil sampel item X1 (Pemahaman definisi teknologi *Deepfake*) dari Variabel Literasi AI. Berdasarkan hasil pengolahan data pada N = 100 responden menggunakan rumus Korelasi Pearson Product Moment (lihat BAB III), diperoleh nilai $r_{hitung} = 0,744$, sedangkan nilai r_{tabel} pada taraf signifikansi 5% dan N = 100 adalah 0,195.

Kriteria: $r_{hitung} > r_{tabel} \rightarrow 0,744 > 0,195$, maka item X1 dinyatakan VALID. Hasil ini konsisten dengan output aplikasi (*Google Colab*) pada Gambar 4.2 dan Gambar 4.3, serta rekapitulasi lengkap pada Tabel 4.2.

Contoh Ilustrasi pada Item yang Tidak Memenuhi Kriteria (TIDAK VALID)

Sebagai ilustrasi penerapan kriteria pengujian (bukan merupakan data aktual penelitian, karena seluruh item pada instrumen akhir penelitian ini dinyatakan valid), apabila terdapat suatu item hipotetis dengan hasil $r_{hitung} = 0,150$, maka penerapan kriteria pengujiannya adalah sebagai berikut:

Kriteria: $r_{hitung} < r_{tabel} \rightarrow 0,150 < 0,195$, maka item tersebut dinyatakan TIDAK VALID, sehingga perlu diperbaiki atau dihapus dari instrumen sebagaimana dijelaskan pada BAB III Sub Bab 3.7. Ilustrasi ini menunjukkan bahwa jika suatu item memperoleh r_{hitung} di bawah r_{tabel} , item tersebut tidak dapat digunakan untuk mengukur variabel penelitian.

Contoh Penerapan Rumus Reliabilitas (Cronbach's Alpha)

Pada Variabel Literasi AI (X) dengan k = 5 item, hasil pengolahan data menggunakan rumus $\alpha = [k/(k-1)] \times [1 - (\sum \sigma_i^2 / \sigma^2)]$ menghasilkan nilai Cronbach's Alpha sebesar 0,834. Karena $0,834 \geq 0,60$, maka instrumen dinyatakan RELIABEL, sebagaimana ditampilkan pada Gambar 4.4 dan Gambar 4.5.

Rekapitulasi lengkap hasil pengujian validitas dan reliabilitas untuk seluruh item disajikan pada Tabel 4.2 (Hasil Uji Validitas Indikator Literasi AI) dan uraian hasil pengujian reliabilitas pada Sub Bab 4.3.3, sedangkan rekapitulasi hasil pengujian kemampuan deteksi konten *Deepfake* untuk seluruh 10 butir soal disajikan pada Tabel 4.4 (Hasil Analisis Kemampuan Deteksi).

Tingkat Kemampuan Masyarakat dalam Membedakan Konten Nyata dan *Deepfake* AI

Hasil olah data statistik menunjukkan bahwa tingkat kemampuan responden dalam mendeteksi dan membedakan antara konten asli dan rekayasa *Deepfake* AI tergolong cukup tinggi. Hal ini ditunjukkan oleh distribusi skor variabel Kemampuan Deteksi (Y_1) yang terpusat pada angka 8 dari total skor maksimum 10, dengan nilai rata-rata estimasi (*constant*) sebesar 8,051.

Temuan ini mengindikasikan bahwa secara umum, masyarakat (responden) memiliki sensitivitas indrawi dan kepekaan perseptual dasar yang cukup baik dalam mengenali anomali visual pada media digital. Hambatan atau kegagalan visual seperti pergerakan bibir yang tidak natural (*lip-sync mismatch*), pencahayaan wajah yang tidak konsisten, hingga tekstur kulit yang terlalu mulus (*over-smoothing*) masih dapat ditangkap oleh pengamatan intuitif responden tanpa bantuan alat teknis khusus.

Hubungan dan Pengaruh Pengetahuan terhadap Kemampuan Deteksi

Secara empiris, penelitian ini menemukan bahwa Pengetahuan tentang *Deepfake* AI (X_1) tidak memiliki hubungan maupun pengaruh yang signifikan terhadap Kemampuan Deteksi (Y_1) ($r = -0,028$; $p = 0,784$; $R^2 = 0,001$).

Fenomena ini menunjukkan keberadaan *gap* atau jurang pemisah antara pemahaman kognitif/teoritis dengan keterampilan teknis/praktis (*know-do gap*):

- a. Pengetahuan Teoritis bersifat Konseptual:

Pengetahuan mengenai definisi, bahaya, sejarah, maupun etika penggunaan AI tergolong sebagai pengetahuan deklaratif (*declarative knowledge*). Seseorang dapat memahami secara mendalam apa itu *Deepfake* dan dampak buruknya, namun pemahaman tersebut tidak

otomatis melatih ketajaman penglihatan mereka saat dihadapkan pada media foto atau gambar visual yang direkayasa secara halus.

b. Deteksi Praktis bersifat Perseptual:

Kemampuan mengidentifikasi *Deepfake* merupakan bentuk keterampilan perseptual-motorik (*procedural knowledge*) yang sangat dipengaruhi oleh kecermatan pengamatan, perhatian mendetail (*visual attention*), serta kebiasaan berinteraksi dengan media digital terkini, bukan sekadar memori teoritis.

Kaitan dengan Teori dan Penelitian Terdahulu

Temuan penelitian ini memberikan konfirmasi dan pengayaan terhadap kerangka teori serta literatur terdahulu yang dirujuk pada BAB II:

a. Mendukung Konsep *Literacy Paradox / Overconfidence Bias*:

Temuan ini sejalan dengan penelitian terdahulu yang menyoroti adanya *overconfidence bias* atau efek *Illusion of Explanatory Depth* pada literasi digital. Individu yang merasa memiliki pengetahuan AI yang tinggi cenderung terlalu percaya diri dengan persepsinya sendiri, sehingga saat melakukan pengujian praktis, mereka kurang teliti dalam menginspeksi artefak visual yang samar.

b. Perkembangan Teknologi Generatif AI yang Melampaui Persepsi Manusia:

Model *Generative Adversarial Networks (GAN)* dan *Diffusion Models* berkembang pesat secara eksponensial. Kualitas rekayasa *Deepfake* modern saat ini telah mencapai tingkat realisme tinggi (*hyper-realism*) yang dirancang untuk mengelabui persepsi visual manusia. Oleh karena itu, peningkatan pengetahuan teoritis manusia tidak mampu mengimbangi kompleksitas algoritma rekayasa yang kian sempurna.

Implikasi Praktis

Temuan bahwa pengetahuan teoritis tidak menjamin ketajaman deteksi praktis memberikan implikasi strategis bagi berbagai pemangku kepentingan:

Bagi Masyarakat

1. Menghindari Sikap Terlalu Percaya Diri (*Overconfidence*): Masyarakat diimbau untuk tidak hanya mengandalkan prasangka intuitif atau merasa aman karena sudah mengetahui konsep *Deepfake*.
2. Mengadopsi Sikap Verifikasi Objektif: Diperlukan kebiasaan verifikasi silang (*lateral reading*), seperti memeriksa sumber berita resmi, menggunakan alat pembanding citra (*reverse image search*), dan tidak langsung menyebarkan konten visual yang berpotensi memicu emosi.

Bagi Pemerintah dan Pembuat Kebijakan

1. Redesain Program Literasi Digital: Kementerian/lembaga terkait perlu menggeser fokus kampanye literasi digital dari sekadar sosialisasi teoritis menuju pelatihan simulasi praktis (*hands-on detection training*).
2. Penyediaan Infrastruktur Deteksi Publik: Pemerintah perlu mendorong penyediaan alat/peranti lunak pendeteksi *Deepfake* berbasis AI yang dapat diakses secara terbuka oleh masyarakat umum dan insan pers untuk memverifikasi keabsahan media secara instan.

Bagi Akademisi dan Peneliti Selanjutnya

1. Pengembangan Model Penelitian Multi-Faktor: Peneliti selanjutnya disarankan untuk mengeksplorasi variabel-variabel lain yang lebih relevan dalam menjelaskan kemampuan deteksi, seperti *focal attention*, tingkat skeptisisme kritis, durasi paparan media, maupun penggunaan *AI-based detection tools*.

2. Pendekatan Eksperimental:

Menggunakan metode pengujian berbasis eksperimen laboratorium (*experimental design*) dengan mengontrol jenis dan kualitas kecanggihan *Deepfake* (misalnya membedakan tingkat kualitas *low-tier* hingga *high-tier Deepfake*) guna memperoleh analisis perseptual yang lebih mendalam.

KESIMPULAN DAN SARAN

Kesimpulan

Berdasarkan hasil analisis data dan pembahasan yang telah dilakukan terhadap sampel penelitian sebanyak 100 responden pengguna media sosial usia 17–30 tahun, dengan data uji berupa sepuluh pasangan konten stimulus foto (10 foto asli dan 10 foto hasil rekayasa *Deepfake AI*), mengenai hubungan antara Pengetahuan tentang *Deepfake AI* (X1) dengan Kemampuan Deteksi (Y1), dapat ditarik kesimpulan sebagai berikut:

Tingkat Kemampuan Masyarakat dalam Membedakan Konten Nyata dan *Deepfake AI*

Tingkat kemampuan masyarakat (responden) dalam membedakan antara konten asli dan rekayasa *Deepfake AI* berada pada kategori cukup tinggi, dengan tingkat akurasi rata-rata keseluruhan sebesar 77,1% dari 10 butir soal uji. Distribusi skor variabel Kemampuan Deteksi (Y1) memuncak pada skor 8 dari total skor maksimum 10 (dengan frekuensi terbanyak sekitar 35 responden), serta nilai estimasi rata-rata (constant) regresi sebesar 8,051. Temuan ini menandakan bahwa secara visual dan intuitif, masyarakat memiliki ketajaman persepsi dasar dalam mengenali hal yang tidak biasa atau kejanggalan pada media digital.

Hasil Uji Korelasi antara Pengetahuan dan Kemampuan Deteksi

Pengujian Korelasi Pearson Product Moment menunjukkan nilai koefisien korelasi $r = -0,028$ dengan nilai signifikansi $p\text{-value} = 0,784$ ($p > 0,05$). Hasil ini membuktikan bahwa tidak terdapat hubungan yang signifikan antara Pengetahuan tentang *Deepfake AI* (X1) dengan Kemampuan Deteksi (Y1), dengan tingkat kekuatan hubungan yang tergolong sangat lemah.

Hasil Uji Regresi dan Pengujian Hipotesis

Model regresi $\hat{Y} = 8,051 - 0,017X_1$ terbukti tidak layak (not fit) dengan nilai $F = 0,075$ ($p = 0,784$) dan nilai koefisien determinasi (R^2) sebesar 0,001, yang berarti Pengetahuan tentang *Deepfake AI* (X1) hanya memberikan kontribusi sebesar 0,1% terhadap Kemampuan Deteksi (Y1), sedangkan 99,9% sisanya dipengaruhi oleh faktor lain di luar model. Berdasarkan nilai signifikansi $p = 0,784 > 0,05$, disimpulkan bahwa H_0 DITERIMA dan H_a DITOLAK, artinya pengetahuan teoritis mengenai *Deepfake AI* yang dimiliki seseorang tidak secara otomatis meningkatkan keterampilan praktisnya dalam mendeteksi konten rekayasa *Deepfake*.

Kelebihan Metode/Pendekatan yang Digunakan

Pendekatan deskriptif kuantitatif dengan kuesioner daring (*Google Forms*) yang digunakan dalam penelitian ini memiliki beberapa kelebihan. Dari segi objektivitas, hasil diperoleh dari perhitungan statistik sehingga meminimalkan bias subjektif peneliti. Dari segi efisiensi, pendekatan ini mampu menjangkau 100 responden dalam waktu relatif singkat tanpa perlu mendatangi lokasi responden secara langsung. Dari segi kemudahan pengolahan dan komunikasi hasil, data dapat diolah secara cepat menggunakan bahasa pemrograman *Python* dan disajikan dalam bentuk tabel maupun grafik yang mudah dipahami, seperti pada Tabel 4.2, Tabel 4.4, dan Gambar 4.10. Selain itu, pendekatan ini juga dapat direplikasi oleh peneliti lain menggunakan prosedur pengambilan data yang sama, sehingga memperkuat tingkat kepercayaan terhadap hasil.

Kelemahan Metode/Pendekatan yang Digunakan

Di sisi lain, pendekatan ini juga memiliki beberapa kelemahan. Kurangnya kedalaman eksplorasi menjadi salah satu kelemahan, karena data yang dihasilkan berupa angka sehingga kurang mampu menggali alasan atau proses berpikir responden secara mendalam saat menilai konten *Deepfake*. Selain itu, terdapat potensi bias responden dan rendahnya kontrol atas kondisi pengisian kuesioner daring, seperti kesungguhan menjawab, gangguan lingkungan, maupun keterbatasan pemahaman responden terhadap pertanyaan. Keterbatasan stimulus penelitian juga menjadi kelemahan, karena hanya menggunakan 10 sampel foto sehingga belum mampu merepresentasikan seluruh variasi kualitas dan jenis konten *Deepfake* (video maupun audio) yang ada di dunia nyata. Kelemahan terakhir adalah sifat data yang dikumpulkan pada satu titik waktu (*cross-sectional*), sehingga tidak dapat menggambarkan perubahan kemampuan deteksi responden dari waktu ke waktu.

Saran

Berdasarkan temuan dan kesimpulan penelitian ini, termasuk kelemahan pendekatan yang telah diuraikan pada Sub Bab 5.1 poin 5, diajukan beberapa saran sebagai berikut:

Solusi untuk Mengatasi Kelemahan Metode/Pendekatan Penelitian

Untuk mengatasi keterbatasan kedalaman eksplorasi, penelitian selanjutnya disarankan mengombinasikan kuesioner kuantitatif dengan wawancara atau *focus group discussion* (pendekatan mixed methods) agar dapat menggali alasan dan proses berpikir responden secara lebih mendalam. Untuk mengurangi potensi bias responden pada kuesioner daring, disarankan menambahkan pertanyaan kontrol (*attention check*) serta memperjelas instruksi pengisian. Untuk mengatasi keterbatasan jumlah dan jenis stimulus, disarankan memperbanyak dan memvariasikan sampel uji (foto, audio, dan video) dengan tingkat kecanggihan *Deepfake* yang beragam. Terakhir, untuk mengatasi sifat data *cross-sectional*, disarankan menggunakan desain penelitian longitudinal agar perkembangan kemampuan deteksi masyarakat dapat diamati dari waktu ke waktu.

Bagi Masyarakat

Masyarakat diimbau untuk tidak hanya berpuas diri pada pemahaman konseptual/teoritis mengenai *Deepfake AI*, melainkan melatih diri dengan keterampilan verifikasi praktis. Masyarakat perlu membiasakan tindakan verifikasi silang (*fact-checking*), memvalidasi sumber berita resmi, menggunakan alat penelusuran gambar terbalik (*reverse image search*), serta bersikap kritis dan skeptis terhadap media visual yang berpotensi memicu sensasi emosional.

Bagi Pemerintah dan Pembuat Kebijakan

Pemerintah melalui kementerian dan lembaga terkait (seperti Kominfo) disarankan untuk mereformulasi strategi dan kurikulum literasi digital. Kampanye literasi digital hendaknya tidak lagi berfokus pada sosialisasi teori semata, melainkan dialihkan ke pelatihan simulasi praktis (*hands-on training*) berbasis kasus nyata. Selain itu, pemerintah perlu memperkuat regulasi tata kelola kecerdasan buatan serta menyediakan infrastruktur publik berupa perangkat lunak pendeteksi *Deepfake* gratis yang mudah diakses oleh masyarakat umum.

Bagi Peneliti Selanjutnya

Memperluas cakupan wilayah geografis serta keberagaman demografi responden guna meningkatkan generalisasi hasil penelitian; menambah variasi stimulus penelitian dengan mengombinasikan berbagai bentuk konten rekayasa (*Deepfake* audio, video, serta gambar statis) dengan variasi tingkat kecanggihan algoritma; dan mempertimbangkan pendekatan *Machine Learning*, misalnya membandingkan batas kemampuan deteksi intuitif manusia (*human perception*) dengan kinerja model deteksi berbasis *Convolutional Neural Network* (CNN), untuk memperoleh gambaran komparatif yang lebih mendalam mengenai efektivitas deteksi *Deepfake*.

DAFTAR PUSTAKA

- Altuncu, E., Franqueira, V. N. L. Dan Li, S. (2024) Deepfake: Definitions, Performance Metrics And Standards, Datasets, And A Meta-Review.
- Darmawan, M. T., Junaidi, A. Dan Khaerudin, A. (2025) Penegakan Hukum Terhadap Penyalahgunaan Deepfake Pada Pornografi Anak Di Era Artificial Intelligence Di Indonesia. *Jurnal Penelitian Serambi Hukum*, 18(01), Hal. 42-54.
- Dhahir, D. F., Kenda, N., Dirgahayu, D., Syarifuddin, S., Djaffar, R., Widiastuti, R., ... & Pala, R. (2024). The Relationship Of Digital Literacy, Exposure To AI-Generated Deepfake Videos, And The Ability To Identify Deepfakes In Generation X. *Jurnal Pekommas*, 9(2), 357-368.
- Fadhilah, A. D. Z. Dan Retnoningsih, S. (2024) Perancangan Kampanye Digital Melawan Disinformasi Melalui Artificial Intelligence Dan Deepfake Di Kalangan Pra Lansia Usia 45-55 Tahun. *FAD*, 3(02).
- Juditha, C. (2025). Artificial Intelligent Dan Hoaks: Tantangan Disrupsi Bagi Literasi Digital Masyarakat Di Komunitas Online Dan Upaya Antisipasi Pemerintah. *Jurnal Ilmiah Manajemen Informasi Dan Komunikasi*, 28-41.
- Koutlis, C., Pianese, A., Cozzolino, D., Schinas, M., Mylonas, S., Verdoliva, L., & Papadopoulos, S. (2026). Video Deepfake Detection: Challenges And Recent Trends. In *Countering Disinformation In The Era Of Generative AI* (Pp. 215-246). Cham: Springer Nature Switzerland.
- Lestari, P. A., Saputri, Z. R., Jamilah, R., Febiyanti, F., Fadilah, M. F., Firnanda, M. F., & Koeswara, T. S. N. (2026). Konten Illegal Kecerdasan Buatan Berbasis Deepfake Dan Upaya Penanggulangannya Di Indonesia. *Jurnal Ilmiah Sistem Informasi*, 5(1), 26-35.
- Lin, L., Gupta, N., Zhang, Y., Ren, H., Liu, C. H., Ding, F., ... & Hu, S. (2024). Detecting Multimedia Generated By Large Ai Models: A Survey. *Arxiv Preprint Arxiv:2402.00045*.
- Mamuaya, N. C., SE, M., Wahyudi, M. P., Syah, N., CST, M. P., Arifin, M. Z., ... & Asmalinda Sy, S. K. M. (2025). *Metode Penelitian Kuantitatif*. Azzia Karya Bersama.
- Mirsky, Y., & Lee, W. (2021). The Creation And Detection Of Deepfakes: A Survey. *ACM Computing Surveys (CSUR)*, 54(1), 1-41.
- Payne, L. (2026) Deepfake. Encyclopaedia Britannica. Tersedia Pada: <https://www.britannica.com/technology/deepfake> (Diakses: 13 Mei 2026)
- Puspita, L., & Firdaus, H. S. (2025). The Role Of Law In Combating The Spread Of Deepfake Content For Digital Identity Fraud. *Uir Law Review*, 9(1), 1-14.
- Raharjo, B. (2023). *Teori Etika Dalam Kecerdasan Buatan (AI)*. Yayasan Prima Agus Teknik.
- Sabela, I. N. D., Saputra, A., & Kuncoro, W. (2026). Hoaks Berbasis Kecerdasan Buatan Ai Sebagai Tantangan Literasi Media Sosial. *Relasi: Jurnal Penelitian Komunikasi (E-Issn: 2807-6818)*, 6(01), 44-54.

- Seveney, M. C., Wicaksono, D. B. Dan Soetijono, I. K. (2025) Urgensi Regulasi Terhadap Penyalahgunaan Deepfake Berbasis AI (Artificial Intelligence) Pada Konten Pornografi. Disiplin: Majalah Civitas Akademika Sekolah Tinggi Ilmu Hukum Sumpah Pemuda, 31(2), Hal. 97-106.
- Wang, T., Liao, X., Chow, K. P., Lin, X., & Wang, Y. (2024). Deepfake Detection: A Comprehensive Survey From The Reliability Perspective. *ACM Computing Surveys*, 57(3), 1-35.