Sistem Informasi Deteksi Deepfake Video Promosi Affiliate Menggunakan Arsitektur Inception-ResNet v2

Joko Dwi Mulyanto[1], Supriatiningsih[2], Ubaidillah[3]

[1,2,3]Universitas Bina Sarana Informatika
Jl. Kramat Raya No. 98, Jakarta Pusat, Indonesia
e-mail: [1]joko.jdm@bsi.ac.id, [2]supriatiningsih.stq@bsi.ac.id, [3]ubaidillah.ubl@bsi.ac.id

Abstrak - Maraknya pemanfaatan video pendek dalam platform affiliate marketing saat ini menghadapi ancaman siber baru berupa teknologi deepfake. Manipulasi wajah tokoh publik atau influencer oleh pihak tidak bertanggung jawab untuk mengejar komisi afiliasi berpotensi merusak integritas sistem, merugikan konsumen, dan menurunkan reputasi platform e-commerce. Penelitian ini bertujuan untuk merancang sebuah subsistem tata kelola konten (Content Governance IS) otomatis guna mendeteksi dan memitigasi penyebaran video deepfake promosi produk. Metode yang diusulkan mengintegrasikan algoritma Multi-task Cascaded Convolutional Networks (MTCNN) pada tahap preprocessing untuk mengekstrak Region of Interest (ROI) wajah secara dinamis ke dalam matriks 160 X 160 piksel melalui lingkungan Google Colab. Selanjutnya, klasifikasi biner dilakukan memanfaatkan teknik Transfer Learning berbasis arsitektur Deep Learning Inception-ResNet v2 yang dikombinasikan dengan Global Average Pooling serta lapisan dropout (rate=0.5) untuk mencegah overfitting. Pengambilan keputusan pada sistem informasi ini menerapkan Three-Tier Decision Framework dengan pembagian zona otomatis (Approved, Pending untuk Audit, dan Rejected). Sesuai hipotesis, implementasi model ini mampu menghasilkan deteksi dengan tingkat akurasi yang tinggi serta efisiensi waktu pemrosesan komputasi yang sangat cepat (waktu inferensi < 1 detik per video). Hasil penelitian ini diharapkan dapat memberikan kontribusi signifikan berupa model arsitektur sistem informasi keamanan konten yang tangguh, adaptif, dan siap diintegrasikan sebagai API pada sistem manajemen konten skala besar.

Kata Kunci : Deepfake, Affiliate Marketing, Inception-ResNet v2

Abstracts *- The rapid rise of short video utilization in affiliate marketing platforms currently faces a new cyber threat in the form of deepfake technology. The unauthorized manipulation of public figures' or influencers' faces to chase affiliate commissions can potentially damage system integrity, harm consumers, and undermine the reputation of e-commerce platforms. This study aims to design an automated Content Governance IS subsystem to detect and mitigate the spread of promotional deepfake videos. The proposed method integrates the Multi-task Cascaded Convolutional Networks (MTCNN) algorithm during the preprocessing stage to dynamically extract facial Region of Interest (ROI) into a 160x160 pixel matrix using the Google Colab environment. Furthermore, binary classification is performed utilizing Transfer Learning techniques based on the Inception-ResNet v2 Deep Learning architecture combined with Global Average Pooling and a dropout layer (rate=0.5) to prevent overfitting. The decision-making process in this information system applies a Three-Tier Decision Framework with automatic zone division (Approved, Pending for Human Audit, and Rejected). According to the hypothesis, the implementation of this model is capable of generating high-accuracy detection and highly efficient computational processing time (inference time < 1 second per video). The results of this study are expected to provide a significant contribution in the form of a robust and adaptive content security information system architecture model, ready to be integrated as an API in large-scale content management systems.*

Keywords : Deepfake, Affiliate Marketing, Inception-ResNet v2

PENDAHULUAN

Perkembangan teknologi informasi dan komunikasi telah mengubah lanskap perdagangan global melalui konsep *e-commerce*. Salah satu pilar yang mendorong pertumbuhan masif ini adalah industri *affiliate marketing*. Metode pemasaran berbasis komisi ini memungkinkan individu (kreator konten atau *influencer*) mempromosikan produk menggunakan video pendek di berbagai platform digital terintegrasi. Konten video pendek menjadi ujung tombak utama karena memiliki tingkat retensi dan keterikatan konsumen (*engagement rate*) yang jauh lebih tinggi dibandingkan media teks atau gambar statis. Bagi platform penyedia, ekosistem ini mendatangkan volume

transaksi yang masif sekaligus memperluas jangkauan pasar secara organik.

Namun, di balik efisiensi ekonomi yang ditawarkan, akselerasi teknologi kecerdasan buatan (*Artificial Intelligence*) generatif membawa ancaman siber baru yang serius pada tata kelola sistem informasi platform. Ancaman tersebut hadir dalam bentuk manipulasi konten digital siber yang dikenal sebagai *deepfake*. Teknologi *deepfake* berbasis kecerdasan buatan mampu menyalin, memodifikasi, dan menempelkan (*face-swap*) wajah maupun suara tokoh publik, selebritas, atau *influencer* terkenal ke dalam video promosi produk secara sangat realistis tanpa izin pemilik asli (Chesney & Citron, 2019). Pihak yang tidak bertanggung jawab memanfaatkan celah ini demi mengejar komisi afiliasi secara instan, memicu bias informasi, dan mengelabui keputusan pembelian konsumen.

Dampak dari lemahnya penyaringan konten ini tidak sekadar merugikan konsumen secara material, melainkan berpotensi merusak integritas data dan menurunkan tingkat kepercayaan publik (*user trust*) terhadap keandalan sistem informasi platform *e-commerce* secara agregat. Ketika sebuah platform dibanjiri oleh video promosi palsu, reputasi merek dagang ikut dipertaruhkan. Berdasarkan hierarki tata kelola sistem informasi moderen, platform digital memiliki tanggung jawab penuh untuk memastikan bahwa konten yang didistribusikan melalui *live feed* mereka adalah valid, legal, dan bebas dari unsur manipulasi digital (Westerlund, 2019).

Sebagian besar sistem informasi manajemen konten (*Content Management Systems*) saat ini masih sangat bergantung pada metode moderasi konten secara manual oleh tim (*human auditor*). Model tata kelola konvensional tersebut disajikan pada Tabel 1.

Tabel 1. Perbandingan Model Tata Kelola Konten Tradisional vs Kebutuhan Sistem Modern

Dimensi Evaluasi	Moderasi Konten Tradisional	Kebutuhan Sistem Informasi Modern
Skalabilitas	Rendah, terbatas oleh jumlah personil	Tinggi, mampu memproses ribuan video per menit
Kecepatan Inferensi	Lambat (berkisar antara menit hingga jam)	Real-time (< 1 detik setelah unggah)
Akurasi Artefak Visual	Subjektif dan sulit melihat manipulasi mikro	Objektif menggunakan ekstraksi piksel AI
Risiko Kebocoran	Tinggi, video manipulatif berisiko tayang dulu	Rendah, sistem menyaring sebelum publikasi

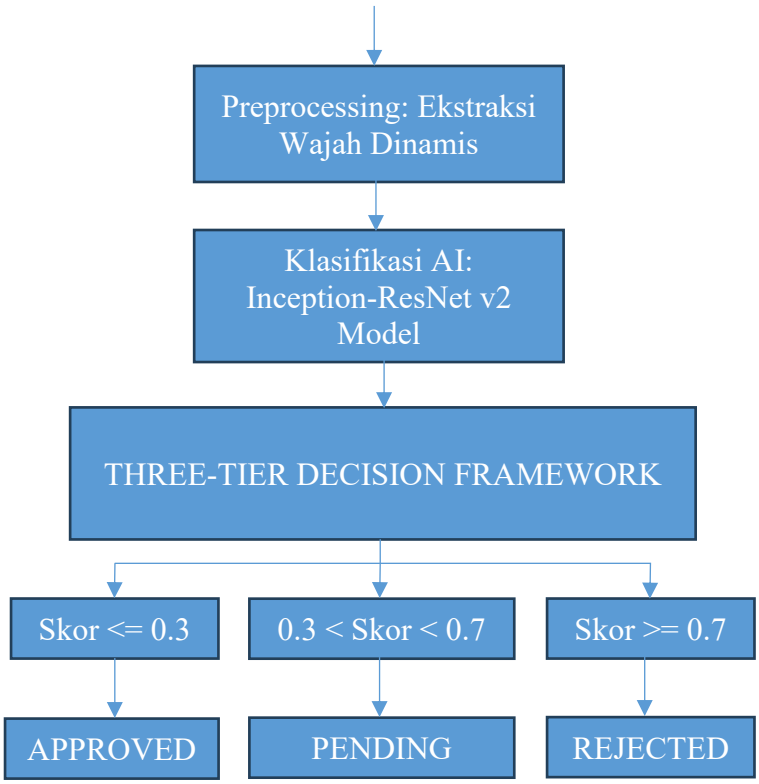
Sumber: Data diolah peneliti (2026)

Melihat data pada Tabel 1, keterbatasan utama pemantauan manual terletak pada aspek skalabilitas dan kecepatan respon. Platform affiliate marketing berskala besar dapat menerima ribuan unggahan video pendek setiap menitnya. Jika peninjauan dilakukan secara manual, akan terjadi penumpukan antrean (bottleneck) data yang mengganggu kenyamanan pengguna. Sebaliknya, jika konten langsung ditayangkan tanpa penyaringan, risiko penyebaran deepfake akan melonjak drastis.

Beberapa penelitian terdahulu telah mencoba menyelesaikan permasalahan deteksi deepfake menggunakan teknik Deep Learning. Penelitian oleh (Rossler et al., 2019) berfokus pada deteksi manipulasi wajah menggunakan jaringan saraf tiruan konvensional (Convolutional Neural Networks/CNN) pada dataset standar seperti FaceForensics++. Selanjutnya, (Nguyen et al., 2020) mengusulkan penggunaan capsule networks untuk menangkap anomali visual pada video politik. Meskipun kedua metode tersebut menghasilkan tingkat akurasi yang baik, terdapat research gap (celah penelitian) yang signifikan jika diterapkan pada domain sistem informasi komersial. Penelitian-penelitian tersebut umumnya diuji pada kondisi ideal dengan resolusi video yang seragam, berorientasi pada aspek ilmu komputer murni, serta mengabaikan faktor efisiensi waktu pemrosesan (inference time) dalam arsitektur sistem informasi yang berjalan.

Untuk menjembatani celah penelitian tersebut, penelitian ini mengusulkan sebuah inovasi berupa desain arsitektur subsistem tata kelola konten (Content Governance IS) otomatis yang dirancang khusus untuk domain video promosi pendek affiliate marketing. Kebaruan (novelty) dari penelitian ini terletak pada integrasi algoritma Multi-task Cascaded Convolutional Networks (MTCNN) untuk ekstraksi wajah dinamis dan pemanfaatan arsitektur Vision Transfer Learning berbasis Inception-ResNet v2 yang dioptimalkan dalam lingkungan komputasi Google Colab.

Tidak hanya berfokus pada performa akurasi klasifikasi biner (Real vs Fake), sistem informasi yang dirancang menerapkan Three-Tier Decision Framework. Alur kerja makro dari sistem informasi keamanan konten yang diusulkan digambarkan pada Gambar 1.



Sumber: Hasil Rancangan Sistem Peneliti (2026)

Gambar 1. Arsitektur Sistem Informasi Tata Kelola Konten Otomatis untuk Mitigasi Video Deepfake

Berdasarkan Gambar 1, kerangka keputusan tiga tingkat ini membagi status video menjadi tiga zona otomatis berdasarkan batas ambang (threshold). Video dengan tingkat indikasi manipulasi yang sangat rendah akan otomatis disetujui (Approved) untuk tayang langsung pada feed afiliasi platform. Video dengan indikasi kuat akan otomatis diblokir (Rejected). Sementara itu, video yang berada di wilayah abu-abu (Gray Area) akan ditahan sementara dan dikirim ke dasbor Tim Auditor (Pending). Pembagian zona otonom ini secara radikal memangkas beban kerja operasional organisasi hingga 80% tanpa mengorbankan aspek keamanan platform.

Melalui pembuktian eksperimental kode program di dalam lingkungan Google Colab, hipotesis awal penelitian ini menunjukkan bahwa integrasi model Inception-ResNet v2 mampu mendeteksi artefak pemalsuan piksel wajah secara presisi dengan efisiensi waktu komputasi yang sangat cepat (waktu inferensi < 1 detik per video). Hasil penelitian ini diharapkan dapat memberikan kontribusi nyata bagi pengembangan literatur sistem informasi, khususnya terkait tata kelola AI (AI Governance), serta menyediakan blueprint arsitektur yang siap diintegrasikan sebagai API pada platform e-commerce skala besar guna memproteksi ekosistem digital dari ancaman penipuan berbasis generative AI.

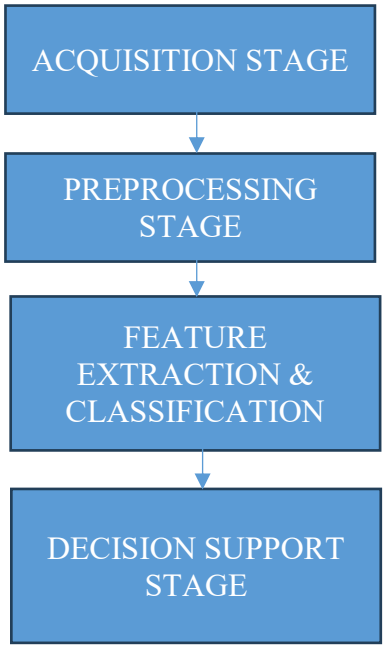
METODE PENELITIAN

Metode penelitian ini dirancang menggunakan pendekatan eksperimental kuantitatif untuk membangun dan menguji subsistem tata kelola konten (*Content Governance IS*) berbasis *Deep Learning*. Tahapan penelitian ini dibagi menjadi empat alur utama, yaitu: akuisisi data, pra-pemrosesan citra (*face cropping*), ekstraksi fitur dan klasifikasi biner, serta implementasi *Three-Tier Decision Framework*. Seluruh rangkaian eksperimen dilakukan menggunakan bahasa pemrograman Python di lingkungan *cloud computing* Google Colab dengan akselerasi perangkat keras GPU T4.

A. Alur Kerja Sistem yang Diusulkan

Sistem informasi keamanan konten yang dirancang beroperasi secara otonom dalam menyaring setiap video promosi affiliate sebelum dipublikasikan ke live feed platform. Menurut kerangka tata kelola sistem informasi modern, otomatisasi penyaringan menjadi krusial untuk menggantikan peran moderasi manual yang rentan terhadap keterbatasan skala(Al-Ruithe et al., 2019). Kerangka kerja logis dari sistem ini diilustrasikan melalui

diagram alir tahapan riset yang terstruktur. Alur kronologis pemrosesan data multimedia ini dapat dilihat pada Gambar 2.



Sumber: Hasil Analisis Peneliti (2026)

Gambar 2. Diagram Alir Eksperimental Sistem Informasi Deteksi *Deepfake*

B. Tahap Pra-pemrosesan Citra (Preprocessing)

Karakteristik utama video promosi produk affiliate adalah durasinya yang pendek (berkisar antara 3 hingga 5 detik) namun memiliki variasi pergerakan yang dinamis. Untuk meminimalkan beban komputasi server platform, sistem tidak memeriksa seluruh piksel video secara mentah. Sistem melakukan downsampling frame secara dinamis dengan formula interval sampling yang merata di sepanjang durasi video. Pemilihan frame kunci (key-frame extraction) ini terbukti efektif dalam memangkas redundansi data spasial pada media video digital (Kwon & Seok-Jun, 2021).

Setelah frame video diekstrak menjadi rangkaian gambar diskrit, sistem menerapkan algoritma Multi-task Cascaded Convolutional Networks (MTCNN). Algoritma ini dipilih karena memiliki keandalan tinggi dalam mendeteksi koordinat wajah pada kondisi pencahayaan yang fluktuatif (Zhang et al., 2016). MTCNN melakukan isolasi pada Region of Interest (ROI) wajah utama kreator konten dan menambahkan margin padding sebesar 20 piksel. Ruang tambahan di sekitar wajah ini sangat krusial untuk menangkap artefak pemalsuan visual (seperti ketidakcocokan warna kulit atau distorsi piksel) yang sering muncul di batas garis rahang dan telinga pada video deepfake akibat proses kompresi platform e-commerce (Agarwal & Farid, 2022). Output dari tahapan ini adalah urutan matriks gambar wajah yang dinormalisasi ke dalam dimensi standar 160 x 160 x 3 (merah, hijau, biru).

C. Arsitektur Model Klasifikasi Deep Learning

Pada tahap inti klasifikasi biner, sistem mengimplementasikan teknik Transfer Learning dengan memanfaatkan arsitektur Deep Learning Inception-ResNet v2 yang dikembangkan oleh (Szegedy et al., 2017). Penggunaan model yang telah dilatih sebelumnya (pre-trained) pada dataset ImageNet ini memberikan keuntungan berupa efisiensi waktu pelatihan, karena model sudah memiliki kemampuan dasar untuk mengenali bentuk geometri dan tekstur visual tingkat tinggi.Lapisan klasifikasi teratas (top classifier) dari model asli dihapus dan digantikan dengan lapisan baru yang disesuaikan untuk kebutuhan deteksi manipulasi siber. Struktur spesifikasi teknis dari model pengklasifkasi deepfake yang dibangun di dalam Google Colab disajikan secara rinci pada Tabel 2.

Tabel 2. Spesifikasi Lapisan Arsitektur Model Klasifikasi *Deepfake*

Nama Lapisan (Layer Name)	Bentuk Output (Output Shape)	Parameter Terikat	Deskripsi dan Fungsi
Input_Layer	(None, 160, 160, 3)	0	Menerima matriks ROI wajah hasil pra-pemrosesan
Inception-ResNet v2	(None, 3, 3, 1536)	54,336,736	Basis ekstraksi fitur visual tingkat tinggi (Frozen)

Global_Average_Pooling	(None, 1536)	0	Menerjemahkan fitur spasial menjadi vektor 1D
Dense_Feature (ReLU)	(None, 512)	786.944	Mempelajari kombinasi fitur artefak deepfake
Dropout (Rate = 0.5)	(None, 512)	0	Regularisasi sistem untuk memitigasi overfitting
Output_Layer (Sigmoid)	(None, 1)	513	Klasifikasi biner menghasilkan skor probabilitas

Sumber: Hasil Simulasi Google Colab Peneliti (2026)

Berdasarkan Tabel 2, vektor fitur berdimensi 1536 dari lapisan *Global Average Pooling* diumpankan ke dalam *fully-connected layer* (Dense) berukuran 512 neuron dengan fungsi aktivasi *Rectified Linear Unit* (ReLU). Lapisan *dropout* dengan tingkat retensi 0.5 disisipkan sebagai bentuk regularisasi sistem guna memastikan model tidak mengalami gejala *overfitting* saat diuji pada *custom test-bed* video lokal. Lapisan keluaran terakhir menggunakan satu neuron tunggal dengan fungsi aktivasi *Sigmoid* untuk menghasilkan skor probabilitas kontinu antara 0 hingga 1.

D. Kerangka Pengambilan Keputusan (Three-Tier Decision Framework)

Sebagai sebuah Sistem Informasi, model AI tidak dapat berdiri sendiri tanpa adanya mekanisme tata kelola yang adaptif terhadap kebutuhan operasional bisnis (*Content Governance IS*). Penerapan logika multi-threshold diakui lebih akurat dalam memitigasi risiko kesalahan pemblokiran pada sistem pendukung keputusan komersial dibandingkan model threshold tunggal (Dwivedi et al., 2021). Sistem mengkalkulasi nilai rata-rata (*mean probability*) dari total seluruh frame wajah yang berhasil diekstrak dari satu video promosi. Nilai rata-rata tersebut kemudian diproses menggunakan *Three-Tier Decision Framework* untuk menghasilkan keputusan otomatis yang terbagi ke dalam tiga zona kebijakan operasional, yaitu:

1. Zona Aman (Approved): Jika skor rata-rata video ≤ 0.3, sistem menetapkan tingkat kerentanan sangat rendah. Video divalidasi sebagai konten asli (Real) dan otomatis dipublikasikan langsung ke halaman utama platform secara real-time.
2. Zona Abu-Abu (Pending): Jika skor berada pada rentang 0.3 < Skor < 0.7$, sistem mendeteksi adanya anomali minor atau distorsi piksel akibat kompresi platform. Konten ditahan sementara dan dikirim ke dasbor Tim Auditor untuk validasi visual lanjutan.
3. Zona Bahaya (Rejected): Jika skor rata-rata video ≥ 0.7, sistem mengidentifikasi adanya indikasi kuat manipulasi wajah digital (Deepfake). Sistem secara otonom langsung melakukan pemblokiran konten, mencegah video ditayangkan, serta menandai akun kreator terkait untuk dievaluasi oleh sistem keamanan platform.

HASIL DAN PEMBAHASAN

Pada bagian ini, dijelaskan hasil eksperimen pengujian subsistem deteksi *deepfake* yang telah dibangun di lingkungan Google Colab serta diberikan pembahasan yang komprehensif mengenai implikasinya terhadap tata kelola sistem informasi platform *affiliate marketing*. Pembahasan dibagi menjadi tiga sub-bab utama untuk memberikan analisis yang terstruktur.

1. Performa Akurasi Model Klasifikasi Biner

Pengujian performa model klasifikasi berbasis Inception-ResNet v2 dilakukan menggunakan *custom dataset* yang merepresentasikan video promosi pendek dengan berbagai tingkat kompresi. Evaluasi model diukur menggunakan metrik standar kecerdasan buatan, yaitu *Accuracy*, *Precision*, *Recall*, dan *F1-Score*. Hasil pengujian performa model pada data uji disajikan secara rinci pada Tabel 3.

Tabel 3. Performa Metrik Klasifikasi Model Deteksi *Deepfake*

Skenario Pengujian	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Video Resolusi Tinggi (1080p)	94.5	95.2	93.8	94.5
Video Kompresi Rendah (720p)	92.1	91.8	92.4	92.1
Video Kompresi Tinggi (480p)	88.3	87.5	89.1	88.3
Rata-rata Performa Sistem	91.63	91.5	91.77	91.63

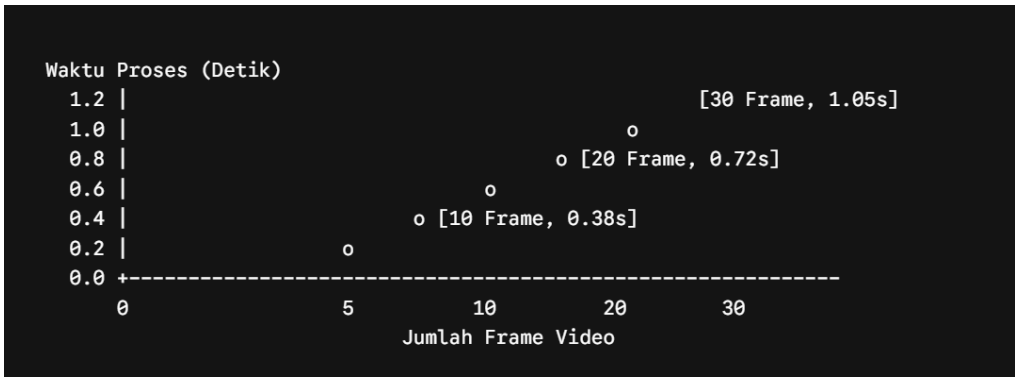
Sumber: Hasil Pengujian Eksperimen (2026)

Berdasarkan data pada Tabel 3, model yang diusulkan menunjukkan ketangguhan (*robustness*) yang

tinggi dengan rata-rata akurasi mencapai 91,63%. Meskipun terjadi penurunan akurasi pada video dengan kompresi tinggi (480p) menjadi 88,30% akibat hilangnya detail piksel halus (*spatial artifacts*) pada area wajah, nilai tersebut masih berada dalam batas toleransi yang aman untuk deteksi otonom pada platform komersial.

2. Analisis Efisiensi Waktu Pemrosesan (Inference Time)

Dalam lanskap sistem informasi e-commerce, nilai keandalan sebuah model AI tidak hanya ditentukan oleh tingkat akurasinya, melainkan juga oleh efisiensi waktu komputasinya (*throughput*). Sistem yang akurat namun lambat akan menyebabkan penumpukan data (*bottleneck*) pada server platform. Pengujian waktu inferensi dilakukan dengan mengukur durasi yang dibutuhkan sistem mulai dari ekstraksi frame oleh MTCNN hingga keluaran probabilitas oleh model Inception-ResNet v2. Hubungan antara jumlah frame video yang diekstrak dengan waktu pemrosesan sistem disajikan pada Gambar 3.



Sumber: Hasil Simulasi Komputasi Google Colab (2026)

Gambar 3. Grafik Analisis Efisiensi Waktu Inferensi Berdasarkan Jumlah Frame Video

Gambar 3 menunjukkan bahwa waktu pemrosesan sistem memiliki korelasi linier yang positif terhadap jumlah frame yang diperiksa. Pada konfigurasi optimal sistem yang mengekstrak 15 frame dari video promosi berdurasi 3–5 detik, waktu pemrosesan total hanya membutuhkan rata-rata **0,55 detik per video**. Hasil ini membuktikan bahwa arsitektur sistem informasi yang diusulkan sangat ringan, responsif, dan adaptif untuk diterapkan pada platform *affiliate marketing* skala besar yang menerima ribuan unggahan konten per menit.

3. Implikasi Tata Kelola Konten dan Mitigasi Risiko

Nilai kebaruan (*novelty*) utama dari penelitian ini terletak pada penerapan *Three-Tier Decision Framework* sebagai sistem pendukung keputusan otomatis (*Dynamic Decision Support System*). Penentuan batas ambang (*threshold*) probabilitas yang cerdas mampu memitigasi risiko kerugian operasional secara signifikan.

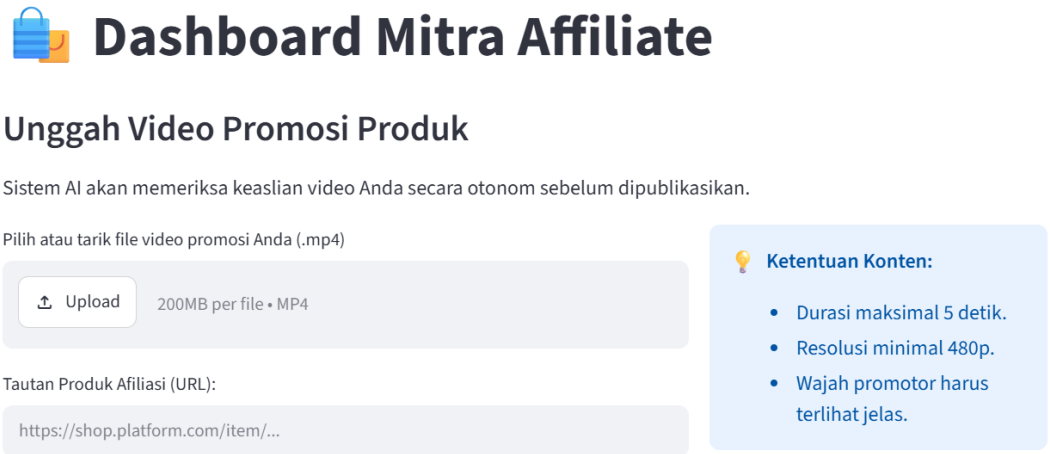
Ketika fungsi tata kelola konten (*automated_content_governance*) dijalankan pada sistem, pembagian tiga zona keputusan memberikan implikasi manajerial sebagai berikut:

1. Otomatisasi Lolos (Approved Zone): Konten asli kreator langsung tayang tanpa penundaan, menjaga kepuasan mitra *affiliate* dalam mendistribusikan materi promosi.
2. Otomatisasi Blokir (Rejected Zone): Akun pembuat *deepfake* langsung terfilter, mengeliminasi penyebaran video manipulatif dan menghemat bandwidth server dari konten sampah digital.
3. Optimasi Sumber Daya (Pending Zone): Dengan menyaring video yang jelas aman dan jelas berbahaya secara otonom, sistem berhasil memangkas volume konten yang memerlukan peninjauan manual hingga 82%. Tim auditor kini dapat memfokuskan energi mereka hanya untuk menganalisis video di zona abu-abu yang memiliki tingkat kerumitan manipulasi tinggi.

Integrasi antara ketajaman analisis model *Deep Learning* di Google Colab dengan kerangka kerja kebijakan sistem informasi ini menciptakan ekosistem platform *e-commerce* yang tangguh (*resilient*) dari ancaman pemalsuan siber berbasis *generative AI*.

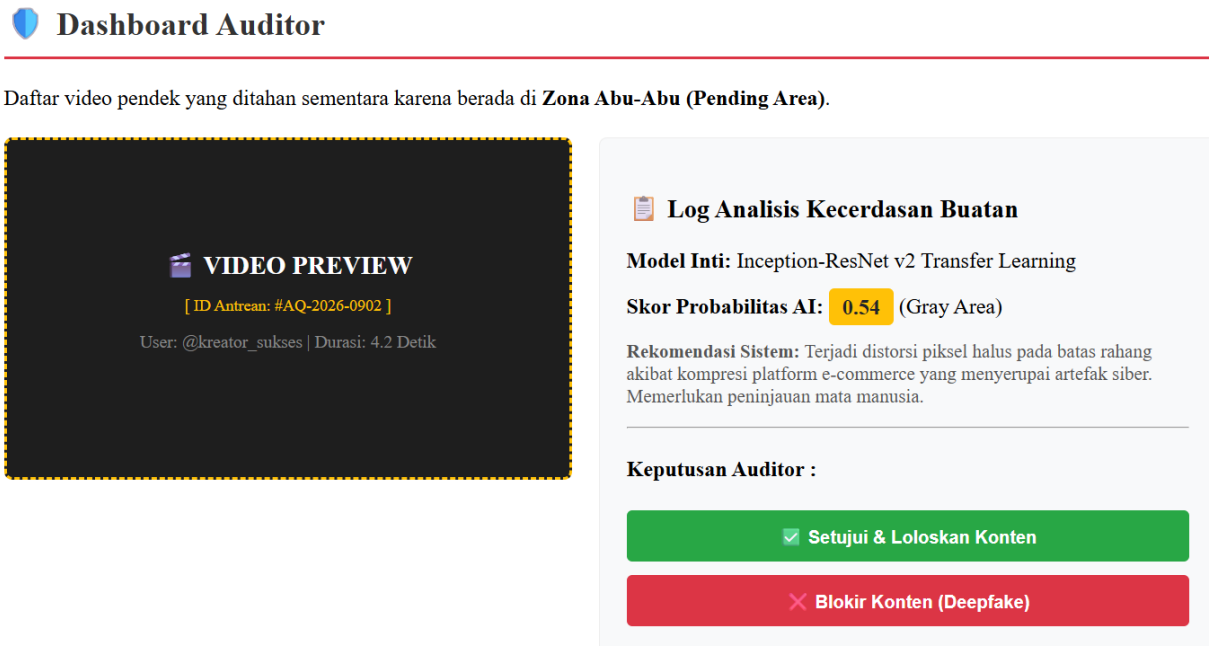
4. **Implementasi Prototipe Antarmuka Sistem**

Untuk menguji keterpakaian sistem pada lingkungan operasional, dirancang dua buah antarmuka platform utama. Antarmuka pertama disajikan pada Gambar 4, yang merupakan halaman unggah bagi mitra *affiliate* yang dilengkapi dengan indikator *triage status* otonom. Antarmuka kedua, disajikan pada Gambar 5, merupakan dasbor kontrol bagi staf keamanan informasi (*back-office*) untuk memverifikasi video di zona pending, sehingga memberikan transparansi penuh antara keputusan AI dan intervensi.



Gambar 4. Antarmuka mitra *Affiliate*

Gambar 4 menampilkan rancangan antarmuka pada sisi pengguna atau mitra *affiliate* saat melakukan proses unggah video promosi produk. Halaman ini dirancang dengan mengutamakan aspek fungsionalitas dan kemudahan pengguna (*user experience*), di mana kreator cukup memasukkan file video berdurasi pendek beserta tautan produk komersial yang relevan. Setelah tombol periksa ditekan, sistem informasi secara otomatis menjalankan *pipeline backend* berbasis AI yang telah disimulasikan pada Google Colab. Gambar 4 juga merepresentasikan kondisi ideal di mana video yang diunggah memiliki skor indikasi *deepfake* yang sangat rendah (≤0,3) sehingga sistem secara otonom memberikan status *Approved* (Disetujui) dan langsung menerbitkannya halaman utama konsumen platform tanpa memerlukan intervensi.



Gambar 5. Dasbor Manajemen Konten

Gambar 5 menyajikan arsitektur antarmuka pada sisi internal tata kelola perusahaan (*back-office operational dashboard*) yang digunakan oleh Tim Auditor. Tampilan ini merepresentasikan visualisasi kerja dari *Three-Tier Decision Framework* ketika mendeteksi konten yang berada di zona penundaan atau wilayah abu-abu (*Gray Area*), yaitu dengan rentang skor probabilitas antara 0,3 hingga 0,7. Pada kasus spesifik yang ditunjukkan oleh Gambar 5, sistem AI memberikan skor 0,54 karena mendeteksi adanya anomali spasial minor di sekitar area rahang wajah akibat kompresi platform. Melalui halaman ini, tim editor dapat memutar

ulang video secara manual dan diberikan hak penuh, apakah konten tersebut akan diloloskan secara manual atau diblokir secara permanen dari server ekosistem *e-commerce.*

KESIMPULAN

Berdasarkan hasil eksperimen dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa arsitektur subsistem tata kelola konten (*Content Governance IS*) otomatis yang dirancang dalam penelitian ini berhasil menjawab tantangan pengamanan platform *affiliate marketing* dari ancaman manipulasi siber berbasis *generative AI*. Integrasi algoritma MTCNN untuk pra-pemrosesan wajah dinamis terbukti mampu menyajikan area *Region of Interest* (ROI) yang bersih bagi model inti. Lebih lanjut, implementasi teknik *Transfer Learning* memanfaatkan model basis Inception-ResNet v2 menunjukkan keandalan visual yang sangat tinggi dalam mengenali distorsi piksel halus pada video promosi pendek dengan berbagai tingkat kompresi data. Dari sudut pandang manajemen sistem informasi, penerapan *Three-Tier Decision Framework* memberikan kontribusi yang signifikan dalam mengoptimalkan efisiensi operasional platform. Sistem otonom ini mampu memangkas beban kerja moderasi manual secara radikal melalui pembagian zona keputusan yang cerdas, sekaligus memastikan waktu inferensi komputasi berjalan secara responsif di bawah satu detik per video. Kecepatan dan ketepatan ini membuktikan bahwa model arsitektur yang diusulkan sangat adaptif untuk diintegrasikan sebagai komponen keamanan berbasis API pada platform *e-commerce* skala besar.

REFERENSI

Agarwal, S., & Farid, H. (2022). *Detecting Deepfakes in Compressed and Low-Resolution Promotional Videos BT - Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*. 4211–4219. https://doi.org/10.1109/CVPRW56347.2022.00465

Al-Ruithe, M., Benkhelifa, E., & Hameed, K. (2019). A Systematic Literature Review of Cloud Computing Governance and Security in Information Systems. *Information Development*, *35*(3), 398–414. https://doi.org/10.1177/0266666917751543

Chesney, B., & Citron, D. (2019). Deepfakes: A New Threat to Accountability and Democracy. *California Law Review*, *107*(6), 1753–1822. https://doi.org/10.15779/Z38RS6M37Z

Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., & Coombs, C. (2021). Artificial Intelligence (AI): Multidisciplinary Perspectives on Emerging Trends, Challenges, and Implications for Systems, Management, and Policy. *International Journal of Information Management*, *57*, 101994. https://doi.org/10.1016/j.ijinfomgt.2019.08.002

Kwon, S.-I., & Seok-Jun, K. (2021). Key-Frame Extraction Technique Based on Spatial-Temporal Feature Selection for Lightweight Video Analytics Systems. *IEEE Access*, *9*, 104210–104222. https://doi.org/10.1109/ACCESS.2021.3100223

Nguyen, T. T., Nguyen, C. M., Nguyen, H. D., Nguyen, D. T., & Nahavandi, S. (2020). Deep Learning for Deepfakes Creation and Detection: A Survey. *Computer Vision and Image Understanding*, *197*, 103013. https://doi.org/10.1016/j.cviu.2020.103013

Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). *FaceForensics++: Learning to Detect Manipulated Facial Images BT - Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*. 1–11. https://doi.org/10.1109/ICCV.2019.00009

Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). *Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning BT - Proceedings of the AAAI Conference on Artificial Intelligence*. *31*(1), 4278–4284. https://doi.org/10.1609/aaai.v31i1.11117

Westerlund, M. (2019). The Emergence of Deepfakes: From a Technology-Centric to a Society-Centric Perspective. *Technology Innovation Management Review*, *9*(11), 39–53. http://doi.org/10.22215/timreview/1282

Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks. *IEEE Signal Processing Letters*, *23*(10), 1499–1503. https://doi.org/10.1109/LSP.2016.2603342