

Penerapan Algoritma Residual Network 50 dan Convolutional Block Attention Module untuk Deteksi Citra Deepfake

1st Muhammad Dimas Aulia Putra
Riali

Teknologi Informasi
Telkom University
DKI Jakarta, Indonesia

dmsriali@student.telkomuniversity.ac.id

2nd Syifa Nurgaida Yutia, S.Tr., M.T.
Teknologi Informasi

Telkom University
DKI Jakarta, Indonesia

syifagaida@telkomuniversity.ac.id

3rd Demi Adidrana, ST., M.T.I.
Teknologi Informasi

Telkom University
DKI Jakarta, Indonesia

demiadidrana@telkomuniversity.ac.id

Abstrak — Deepfake merupakan teknologi berbasis kecerdasan buatan dengan teknologi *deep learning* yang mampu menciptakan atau memanipulasi wajah seseorang secara realistis. Dalam survey terbaru *VIDA Where's The Fraud – Protecting Indonesia Business from AI Generated Fraud*, menemukan lonjakan 1540% pada kasus penipuan *deepfake* di wilayah APAC tahun 2022 hingga 2023, sementara di Indonesia terdapat 1550% kasus penipuan. Peningkatan signifikan kasus *deepfake* ini menunjukkan tantangan baru dalam mendeteksi manipulasi visual menggunakan teknologi *Artificial Intelligence* (AI). Di samping manfaatnya penggunaan teknologi *deep learning* ini bisa menjadi ancaman serius dalam kasus penipuan dan pemerasan, diperlukan metode deteksi berbasis kecerdasan buatan untuk mendeteksi *deepfake* secara efektif dan efisien. Penelitian ini berfokus pada pengembangan dan penerapan metode deteksi *deepfake* berbasis *Convolutional Neural Network* (CNN) menggunakan arsitektur *Residual Network 50* (ResNet50) yang dimodifikasi dengan *Convolution Block Attention Module* (CBAM) untuk meningkatkan akurasi dalam mendeteksi pola-pola artefak yang muncul pada citra *deepfake*, sebagai kontribusi dalam upaya mitigasi penyalahgunaan teknologi *deepfake* yang semakin marak. Penelitian ini menggunakan dataset kaggle yang berjumlah 8000 dataset. Dataset yang akan digunakan untuk perbandingan keaslian data melalui tahapan normalisasi, pelatihan model, dan evaluasi performa dengan metrik akurasi yang terdiri dari gambar asli dan gambar yang sudah menggunakan *deepfake* yang dihasilkan melalui teknologi berbasis AI. Dalam studi terdahulu Penggunaan ResNet50 dalam klasifikasi menunjukkan hasil sebesar 78,87% sedangkan integrasi menggunakan *attention mechanism* pada ResNet50 dengan Long-Distance Attention Module mencapai akurasi hingga 94.30% dan AUC 98.70%. Hasil dari penelitian ini menunjukkan penggunaan CBAM pada 8000 dataset menghasilkan metrik akurasi *Accuracy* 68.44%, *Precision* 72,26%, *Recall* 53,73%, dan *F1-score* 61,63%, dan dari hasil pengamatan pada 6000 dataset ditemukan *Accuracy* 75.27%, *Precision* 67,91%, *Recall* 95,80%, dan *F1-score* 79,43%. Dengan pendekatan ini di harapkan mampu menjadi solusi yang efektif untuk meningkatkan keamanan sistem berbasis biometrik, serta mencegah penyalahgunaan teknologi *deepfake* dalam berbagai sektor.

Kata kunci— *Deepfake*, *Convolutional Neural Network*, *ResNet50*, *Convolutional Block Attention Module*, *Rectified Linear Unit*, *Sigmoid*.

I. PENDAHULUAN

Deepfake merupakan bagian dari *Deep Learning* dengan pemanfaatan teknologi *Computer Vision* dan *Artificial Intelligence* (AI). *Deepfake* berpotensi melemahkan kepercayaan publik akan informasi yang beredar dalam lingkup sektor pemerintahan maupun masyarakat[1]. Berdasarkan hasil survey *VIDA Where's The Fraud – Protecting Indonesia Business from AI Generated Fraud*, Indonesia mengalami lonjakan sebesar 1550% pada kasus penipuan *deepfake* diantara tahun 2022 hingga 2023[2]. Oleh karena itu, pengembangan metode dalam deteksi *deepfake* menjadi kebutuhan mendesak dikarenakan memberikan dampak signifikan dalam penyebaran misinformasi dan disinformasi terutama di Indonesia sendiri dimana tingginya penggunaan sosial media berdampak rentannya terhadap hoaks dan misinformasi, serta dapat berdampak kepada isu sosial dan budaya.

Penelitian ini mengacu pada penelitian terdahulu yang menggabungkan ResNet50 dengan model *attention* untuk meningkatkan akurasi deteksi. Dalam penelitian tersebut, integrasi ResNet50 dengan *Long-Distance Attention Module* mencapai akurasi hingga 94.30% dan AUC 98.70%, menunjukkan bahwa integrasi mekanisme *attention* dapat meningkatkan performa klasifikasi secara signifikan[3]. Adapun juga penelitian sebelumnya yang menggabungkan model *Xception* dengan mekanisme *long-distance attention* untuk meningkatkan akurasi deteksi video *deepfake*. Dalam penelitian tersebut, model spasial-temporal yang diusulkan berhasil mencapai akurasi hingga 96.2% dan AUC sebesar 99.4% pada dataset FaceForensics++ HQ, serta menunjukkan performa unggul dalam mendeteksi berbagai jenis manipulasi wajah. Hasil ini menunjukkan bahwa integrasi mekanisme *attention* dalam domain spasial dan temporal dapat meningkatkan performa klasifikasi secara signifikan[4]. Dan juga penelitian ini juga mengacu pada studi sebelumnya yang menyoroti pentingnya *model attribution* dalam konteks video *deepfake*. Dalam penelitian tersebut, metode DMA-STA yang mengintegrasikan mekanisme *spatial* dan *temporal attention* dikembangkan untuk mengidentifikasi model pembuat *deepfake* secara akurat. Hasil eksperimen pada

dataset DFDM menunjukkan bahwa metode ini mampu mencapai akurasi hingga 71.94% pada video berkualitas tinggi, menunjukkan efektivitas pendekatan berbasis *attention* dalam membedakan model generatif secara halus[5].

Penelitian ini bertujuan untuk meningkatkan matrix akurasi dalam mendeteksi gambar *deepfake* melalui pengembangan model berbasis *deep learning* menggunakan arsitektur ResNet50 yang dikombinasikan dengan CBAM. ResNet50 dipilih sebagai *backbone* model karena kemampuannya dalam mengekstraksi fitur mendalam dari gambar secara efisien melalui mekanisme *residual learning*. Namun, untuk lebih menyoroti informasi penting dalam citra, CBAM ditambahkan sebagai modul perhatian yang bekerja secara adaptif dalam dua tahap: *channel attention* dan *spatial attention*. Penambahan CBAM memungkinkan model untuk memfokuskan perhatian pada area atau fitur yang lebih relevan, sehingga meningkatkan performa klasifikasi[6]. Dataset yang digunakan terdiri dari dua kelas utama, yaitu gambar asli (*Real*) dan gambar palsu (*Deepfake*), yang kemudian dibagi ke dalam subset pelatihan, validasi, dan pengujian. Untuk mendukung stabilitas dan generalisasi model, digunakan teknik augmentasi citra serta algoritma AdamW sebagai *optimizer*. Setelah proses pelatihan selesai, evaluasi dilakukan menggunakan metrik seperti akurasi, *presisi*, *recall*, *F1-score*, dan AUC-ROC, disertai dengan visualisasi Grad-CAM. Visualisasi Grad-CAM berfungsi untuk menampilkan area dalam gambar yang paling memengaruhi keputusan model, sehingga memberikan transparansi dan interpretasi terhadap proses prediksi. Hasil yang diperoleh, diharapkan penambahan CBAM pada ResNet50 dapat meningkatkan fokus fitur terhadap wilayah-wilayah penting pada wajah yang mengindikasikan manipulasi, serta memberikan hasil evaluasi yang kompetitif dibandingkan model baseline tanpa CBAM.

Penelitian ini menawarkan pendekatan baru dengan mengombinasikan arsitektur ResNet50 sebagai model klasifikasi dan CBAM untuk pemfokusan pada objek detail seperti mata dan hidung dengan metode dengan dukungan visualisasi Grad-CAM untuk meningkatkan interpretabilitas prediksi dalam deteksi citra *deepfake*. Hasil dari pelatihan menunjukkan Hasil eksperimen menggunakan model CBAM-ResNet50 menunjukkan bahwa model mampu mencapai akurasi pengujian sebesar 75.27%, dengan nilai *precision* 67.91%, *recall* 95.80%, dan *F1-score* 79.48% pada dataset yang digunakan. Nilai ini menggambarkan bahwa meskipun model cukup baik dalam mendeteksi kelas positif (*recall* tinggi), namun tingkat *presisi* yang relatif rendah menyebabkan adanya prediksi positif yang keliru dalam jumlah cukup besar. Sementara itu, hasil pelatihan pada ResNet50 baseline menghasilkan akurasi pengujian sebesar 86.40%, dengan nilai *precision* 80.69%, *recall* 95.70%, *F1-score* 87.56%, dan AUC sebesar 94.62%.

II. KAJIAN TEORI

A. Residual Network 50

Residual Network 50 merupakan bagian dari *Convolutional Neural Network* (CNN) adalah arsitektur jaringan saraf tiruan yang dirancang untuk memproses data berdimensi grid, seperti citra dua dimensi. dengan penggunaan *Convolutional layer* dan menghasilkan *feature map* dan di lanjutkan ke ReLU (*Rectified Linear Unit*) dan diteruskan ke *Convolutional Block* dengan terusan *residual learning* dengan penjumlahan input dan output dari blok konvolusi melalui *shortcut connection*, yang membantu menjaga gradien selama proses pelatihan.

B. Convolutional Block Attention Module

Convolutional Block Attention Module (CBAM) merupakan *Block Attention* yang digunakan untuk meningkatkan representasi fitur pada CNN dengan cara mengekstraksi fitur yang paling bermakna melalui dua tahap *attention* secara berurutan, yaitu *channel attention* dan *spatial attention*[11]. *Input feature* yang di terima dari layer sebelumnya masuk ke tahap *channel attention*, pada tahap ini proses *Max pooling* dan *Average pooling* secara paralel pada dimensi spasial untuk menangkap informasi penting dari setiap *channel*. Hasil dari kedua pooling ini di proses menggunakan *Shared MLP*, dan dijumlahkan sebelum dilewatkan fungsi aktivasi *sigmoid* untuk menghasilkan *channel attention map*. *Attention map* dikalikan secara *channel-wise* dengan *input feature* di proses ada tahap *Spatial Attention*, pada tahap ini dilakukan *Max pooling* dan *Average pooling* secara spasial untuk memperoleh hasil, kedua hasil di gabungkan (*concatenate*) dan di konvolusi dengan kernel 7×7 di ikuti fungsi *sigmoid* untuk mengkasikan *spatial attention map*, dan dikalikan secara *channel-wise* dengan *refined feature* sebelumnya dengan *output feature* CBAM dengan diperkuat pada aspek *channel* maupun *spasial* [12].

C. Rectified Linear Unit

Rectified Linear Unit (ReLU) adalah fungsi aktivasi yang banyak digunakan pada jaringan saraf konvolusional modern karena sifatnya yang sederhana dan efisien dalam mempercepat konvergensi pelatihan. Fungsi ini bekerja dengan memetakan semua nilai negatif menjadi nol, sementara nilai positif dibiarkan apa adanya. Secara matematis, fungsi ReLU dapat dinyatakan sebagai:

$$\mathcal{F}(x) = \max(0, \mathcal{F}(x)) \quad (1)$$

D. Sigmoid

Fungsi aktivasi *Sigmoid* mengubah nilai masukan menjadi keluaran pada rentang $[0,1]$, sehingga sering digunakan untuk merepresentasikan probabilitas atau bobot normalisasi. Secara matematis, fungsi *Sigmoid* didefinisikan sebagai:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

Channel Attention pada CBAM *Sigmoid* menghasilkan vektor *gate* $M_e \in [1,0]^c$ yang merepresentasikan tingkat kepentingan masing-masing *channel* pada *feature map* hasil ResNet50. Spatial Attention pada CBAM *Sigmoid* menghasilkan peta perhatian $M_e \in$

$[1,0]^{H \times W}$ yang menandai lokasi spasial mana yang paling relevan untuk klasifikasi.

III. METODE

Metode penelitian yang digunakan dalam tugas akhir ini merupakan *mixed method*, yaitu gabungan antara metode kualitatif dan kuantitatif. Metode kualitatif digunakan untuk mengumpulkan informasi melalui kajian literatur dan penelitian terdahulu guna memperoleh gambaran umum serta mengidentifikasi celah penelitian (*research gap*). Sementara itu, metode kuantitatif digunakan pada tahap implementasi, pelatihan, pengujian, dan evaluasi model deteksi *deepfake*.

Tahapan penelitian diawali dengan pengumpulan teori dan studi pustaka terkait arsitektur CNN, khususnya model ResNet50 serta modul CBAM sebagai mekanisme *attention*, dan implementasi model menggunakan framework TensorFlow/Keras.

Dataset yang digunakan merupakan dataset *deepfake* dengan dua kelas, yaitu *Real* (wajah asli) dan *Deepfake* (wajah hasil manipulasi). Data diperoleh dari kaggle, kemudian dilakukan *preprocessing* berupa normalisasi citra dan augmentasi data untuk meningkatkan keragaman. Model dibangun dengan *backbone* ResNet50 dan dimodifikasi dengan CBAM untuk memperkuat pemfokusan pada fitur-fitur penting dalam ekstraksi. Selanjutnya dilakukan *training* dan *testing* menggunakan data yang telah dipisahkan sesuai proporsi 80% pelatihan, 20% pengujian.

Evaluasi performa model dilakukan menggunakan metrik kuantitatif, yaitu *accuracy*, *precision*, *recall*, *F1-score*, *Area Under Curve* (AUC) pada kurva *Receiver Operating Characteristic* (ROC), serta *confusion matrix*.

IV. HASIL DAN PEMBAHASAN

E. Skenario Percobaan

Penelitian ini dilakukan dengan dua skenario percobaan. Skenario pertama menggunakan arsitektur ResNet50 baseline tanpa penambahan modul perhatian (*attention*), sedangkan skenario kedua menggunakan arsitektur ResNet50 dengan CBAM (Convolutional Block Attention Module). Tujuan dari skenario pertama adalah mengevaluasi kinerja dasar ResNet50 dalam mendeteksi citra *deepfake*, sementara skenario kedua bertujuan untuk mengetahui pengaruh penambahan modul CBAM dalam meningkatkan kemampuan model memfokuskan perhatian pada fitur penting citra.

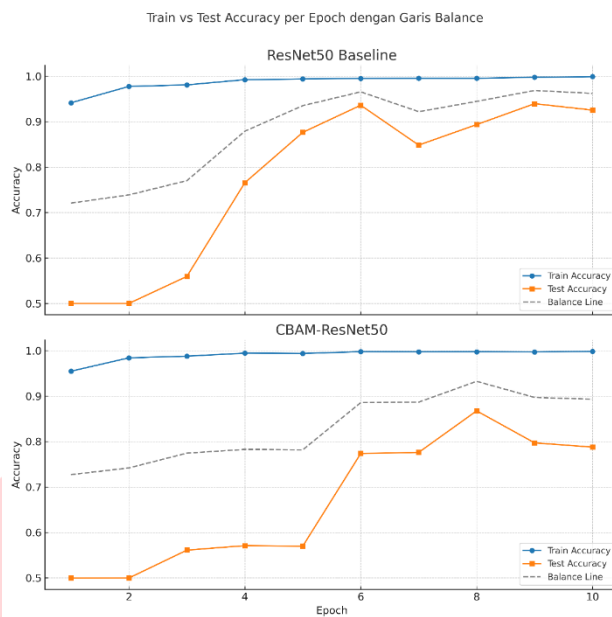
F. Konfigurasi Model

Dataset yang digunakan terdiri dari 8.000 citra yang terbagi menjadi dua kelas: *Real* (asli) dan *Deepfake* (manipulasi). Data dibagi menjadi 80% untuk pelatihan, 20% untuk validasi. Konfigurasi model ditampilkan pada Tabel 1(A).

TABEL 1
(A) KONFIGURASI MODEL

Model	Datase t	Resolu si Gamb ar	Epoc h	Batc h Size	Activati on	Optimiz er
ResNet 50 Baselin e	Real & Deepfa ke	224×2 24	10	32	ReLU	Adam
ResNet 50 , CBAM	Real & Deepfa ke	224×2 24	20	32	ReLU	AdamW

G. Hasil Percobaan



GAMBAR 1

(A) Perbandingan Nilai Akurasi

Berdasarkan grafik Gambar (A) akurasi pelatihan, validasi, dan pengujian, terlihat bahwa ResNet50 *baseline* memiliki performa yang cukup stabil dengan tingkat generalisasi yang baik. Akurasi pelatihan mencapai 99,9%, akurasi validasi sebesar 93,9%, dan akurasi pengujian sebesar 86,4%. Gap antara akurasi pelatihan dan validasi relatif kecil, yaitu sekitar 6%, sedangkan gap antara akurasi validasi dan pengujian sekitar 7,5%. Pola ini menunjukkan bahwa model mengalami sedikit *overfitting*, namun tidak signifikan, karena performa pada data uji masih tergolong tinggi. Dengan AUC sebesar 94,62% dan nilai F1-score mencapai 87,56%, dapat disimpulkan bahwa ResNet50 *baseline* masih mampu melakukan generalisasi dengan baik terhadap data baru, meskipun terjadi penurunan akurasi pada data uji.

Sebaliknya, ResNet50 dengan CBAM menunjukkan tren akurasi yang berbeda. Akurasi pelatihan mencapai hampir 99,8%, sedangkan akurasi validasi hanya sekitar 77,8%, dan akurasi pengujian turun menjadi 75,27%. Gap antara akurasi pelatihan dan validasi cukup besar, yakni sekitar 22%, sementara gap antara validasi dan pengujian sebesar 2,5%. Pola ini jelas mengindikasikan terjadinya *overfitting* yang signifikan, karena model sangat mampu menghafal data pelatihan (train mendekati sempurna), tetapi gagal mempertahankan performa pada data validasi dan pengujian. Hal ini juga tercermin dari nilai metrik evaluasi, di mana *precision* hanya mencapai 67,91%, sedangkan *recall* sangat tinggi yaitu 95,80%, dengan F1-score 79,48%. Tingginya *recall* menunjukkan model sangat sensitif dalam mendeteksi citra *deepfake*, tetapi rendahnya *precision* mengindikasikan peningkatan jumlah *false positive*, sehingga banyak citra *real* yang salah terklasifikasi sebagai *deepfake*.

Secara keseluruhan, dapat disimpulkan bahwa ResNet50 *baseline* lebih seimbang dan stabil dibandingkan dengan ResNet50 + CBAM, yang rentan mengalami *overfitting*. Namun, CBAM masih memberikan keunggulan dalam aspek sensitivitas (*recall* tinggi), sehingga dapat dipertimbangkan penggunaannya apabila tujuan utama adalah memaksimalkan deteksi *deepfake* meskipun dengan risiko meningkatnya *false positive*.

H. Alur Pengelolaan Data

TABEL 2
(B) TABEL PERHITUNGAN

	Gambar	Rata-rata Mean	Median Mean	Rata-rata Std	Seluruh Min	Seluruh Max
conv5_block3_out	10	0.1998	0.1995	0.8363	0	19.611
Cbam_Channel	10	0.0997	0.0998	0.5408	0	16.9438
Cbam_Spatial	10	0.0419	0.0424	0.298	0	16.4069
Dense_feat	10	0.1117	0.1026	0.1627	0	1.1209
Classifier Output Sigmoid	10	0.3594	0.3029	0	0.059	Prob range 0.059â€”0.9539

Data citra masukan berukuran $224 \times 224 \times 3$ diproses melalui arsitektur ResNet50. Setiap konvolusi diakhiri dengan fungsi aktivasi ReLU yang didefinisikan sebagai:

$$f(x) = \max(0, x) \quad (3)$$

ReLU membuang nilai negatif dan hanya mempertahankan aktivasi positif. Hasil kalkulasi aktual pada layer terakhir backbone conv5_block3_out menunjukkan rata-rata nilai aktivasi sebesar 0.1998, standar deviasi 0.8363, serta nilai maksimum mencapai 19.61. Hal ini memperlihatkan bahwa backbone berhasil mengekstraksi fitur dengan rentang nilai tinggi, namun distribusinya relatif jarang aktif (*sparse*). Hasil dari backbone selanjutnya masuk ke modul *Channel Attention* pada CBAM. Mekanismenya menghitung *global pooling* per kanal, lalu melewati *Multi-Layer Perceptron* (MLP) dan fungsi aktivasi *Sigmoid*:

$$Mc(F) = \sigma(W_1(\delta(W_0(F_{avg}))) + W_1(\delta(W_0(F_{max})))) \quad (4)$$

dengan $\sigma(x) = \frac{1}{1+e^{-x}}$

Sigmoid mengubah nilai ke rentang $[0,1]$ sehingga setiap kanal memiliki bobot “penting” atau “tidak penting”. Berdasarkan hasil pelatihan, rata-rata nilai aktivasi turun menjadi 0.0997, standar deviasi 0.5408, dan maksimum 18.15. Artinya, banyak kanal ditekan nilai kecil, sementara hanya beberapa kanal penting yang dipertahankan. Fitur hasil CBAM selanjutnya dilewatkan ke lapisan *fully connected* dengan fungsi aktivasi ReLU, lalu masuk ke layer *Sigmoid* akhir untuk menentukan probabilitas kelas:

$$\hat{y} = \sigma(Wx + b) = \frac{1}{1 + e^{-(Wx+b)}} \quad (5)$$

Distribusi nilai aktual pada tahap ini menunjukkan rata-rata 0.1117, dengan maksimum mencapai 1.1209, dan output akhir probabilitas berada di rentang 0.059 – 0.9539. Nilai inilah yang menghasilkan akurasi latih = 83%, validasi = 77.1%, dan uji = 68.4% sesuai hasil pelatihan.

V. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai deteksi citra deepfake menggunakan arsitektur ResNet50 dan ResNet50 yang dimodifikasi dengan, dapat disimpulkan bahwa penelitian ini berhasil membangun model deteksi deepfake berbasis kedua arsitektur tersebut sesuai dengan tujuan penelitian. Implementasi CBAM pada ResNet50 dapat dilakukan dengan baik dan memberikan kemampuan tambahan dalam pemfokusan area penting pada citra sehingga model lebih terarah dalam mempelajari fitur yang relevan. Hasil evaluasi menunjukkan bahwa ResNet50 baseline memiliki performa yang lebih stabil dengan akurasi pengujian sebesar 86,40%, precision 80,69%, recall 95,70%, F1-score 87,56%, dan AUC 94,62%. Performa ini didukung oleh gap akurasi antara data pelatihan, validasi, dan pengujian yang relatif kecil, sehingga model hanya mengalami sedikit overfitting dan masih mampu melakukan generalisasi dengan baik. Sebaliknya, ResNet50 yang dimodifikasi dengan CBAM menghasilkan akurasi pengujian sebesar 75,27%, dengan precision 67,91%, recall 95,80%, F1-score 79,48%, dan AUC 89%. Meskipun recall yang tinggi menunjukkan sensitivitas yang lebih baik dalam mendeteksi citra deepfake, gap akurasi antara pelatihan dan pengujian yang cukup besar menandakan adanya overfitting signifikan, sehingga performa keseluruhan masih berada di bawah ResNet50 baseline. Dengan demikian, penelitian ini menyimpulkan bahwa ResNet50 baseline lebih seimbang dalam hal generalisasi, sedangkan ResNet50 dengan CBAM lebih unggul dari segi sensitivitas deteksi namun dengan konsekuensi penurunan akurasi keseluruhan.

REFERENSI

Electronic References

• Journal

- [1] Nurdin, W., & Nugraha, F. (2025). Ancaman Deepfake dan Disinformasi Berbasis AI: Implikasi Terhadap Keamanan Siber dan Stabilitas Nasional Indonesia. *JIMR: Journal Of International Multidisciplinary Research*. <https://doi.org/10.62668/jimr.v4i01.1551>
- [2] Borade, S., Jain, N., Patel, B., Kumar, V., Godhrawala, M., Kolaskar, S., Nagare, Y., Shah, P., & Shah, J. (n.d.). Advancements in Video Deepfake Detection: Integration of ResNet50, EfficientNetB7, and Efficient NetAutoAtt B4 Models. In *Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE* (Vol. 2024, Issue 3). www.ijisae.org
- [3] Lu, W., Liu, L., Zhang, B., Luo, J., Zhao, X., Zhou, Y., & Huang, J. (2024). Detection of Deepfake Videos Using Long-Distance Attention. *IEEE Transactions on Neural Networks and Learning Systems*, 35(7), 9366–9379. <https://doi.org/10.1109/TNNLS.2022.3233063>
- [4] Jia, S., Li, X., & Lyu, S. (2022). Model Attribution Of Face-Swap Deepfake Videos. *Proceedings - International Conference on Image Processing, ICIP*, 2356–2360. <https://doi.org/10.1109/ICIP46576.2022.9897972>

- [5] Ghizlane, H., Jamal, R., Mahraz, M. A., Ali, Y., & Hamid, T. (2022). Spam image detection based on convolutional block attention module. *2022 International Conference on Intelligent Systems and Computer Vision, ISCV 2022*. <https://doi.org/10.1109/ISCV54655.2022.9806065>
- [6] Yesugade, K., & Jadhav, R. (2024). Implementation of Deep Learning Techniques for Deepfake Classification: A comparative study using ResNet-50 and VGG16. *2024 IEEE Pune Section International Conference, PuneCon 2024*. <https://doi.org/10.1109/PuneCon63413.2024.10894868>
- [7] Jannu, O., Sekar, V., Padhy, T., & Padalkar, P. (2024). Comparative Analysis of Deepfake Detection Models. *2024 IEEE 9th International Conference for Convergence in Technology, I2CT 2024*. <https://doi.org/10.1109/I2CT61223.2024.10543823>
- [8] Ritter, P., Lucian, D., Anderies, & Chowanda, A. (2023). Comparative Analysis and Evaluation of CNN Models for Deepfake Detection. *2023 4th International Conference on Artificial Intelligence and Data Sciences: Discovering Technological Advancement in Artificial Intelligence and Data Science, AiDAS 2023 - Proceedings*, 250–255. <https://doi.org/10.1109/AiDAS60501.2023.10284611>
- [9] Yee, L. K., Hamid, I. R. A., Chaiwen, C., Abdullah, Z., Kipli, K., & Foozy, C. F. M. (2024). Deepfake Image Detection Using ResNet50 Model. *1st International Conference on Cyber Security and Computing 2024, CyberComp 2024*, 80–87. <https://doi.org/10.1109/CyberComp60759.2024.10913843>
- [10] Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. (2018). CBAM: Convolutional block attention module. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 11211 LNCS, 3–19. https://doi.org/10.1007/978-3-030-01234-2_1
- [11] Nafi'an, M. H. I., Bachtiar, F. A., & Setiawan, B. D. (2023). Modified MobileNetV2 with Convolutional Block Attention Module for Facial Expression Recognition. *IWAIIIP 2023 - Conference Proceeding: International Workshop on Artificial Intelligence and Image Processing*, 22–27. <https://doi.org/10.1109/IWAIIIP58158.2023.10462765>
- [12] Mudegol, N. L., & Urunkar, A. (2025). Supervised Learning Techniques for Deepfake Detection: Integrating ResNet50 and LSTM. *2025 1st International Conference on AIML-Applications for Engineering and Technology, ICAET 2025*. <https://doi.org/10.1109/ICAET63349.2025.10932283>
- [13] Sharma, J., Sharma, S., Kumar, V., Hussein, H. S., & Alshazly, H. (2022). Deepfakes Classification of Faces Using Convolutional Neural Networks. *Traitement Du Signal*, 39(3), 1027–1037. <https://doi.org/10.18280/ts.390330>
- [14] Arini, A., Bahaweres, R. B., & al Haq, J. (2022). Quick Classification of Xception And Resnet-50 Models on Deepfake Video Using Local Binary Pattern. *2021 International Seminar on Machine Learning, Optimization, and Data Science, ISMODE 2021*, 254–259. <https://doi.org/10.1109/ISMODE53584.2022.9742852>
- [15] Bommareddy, S., Samyal, T., & Dahiya, S. (2023). Implementation of a Deepfake Detection System using Convolutional Neural Networks and Adversarial Training. *2023 3rd International Conference on Intelligent Technologies, CONIT 2023*. <https://doi.org/10.1109/CONIT59222.2023.10205614>
- [16] Xu, Z., Wen, X., Zhong, G., & Fang, Q. (2025). Public perception towards deepfake through topic modelling and sentiment analysis of social media data. *Social Network Analysis and Mining*, 15(1). <https://doi.org/10.1007/s13278-025-01445-8>
- [17] Dhahir, D. F., Kenda, N., & Dirgahayu, D. (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. <https://doi.org/10.56873/jpkm.v9i2.5873>
- World Wide Web
- [1] “Deepfake Fraud in Indonesia Surges by 1550%: Here’s How VIDA Fights It,” VIDA Press Release, Oct. 28, 2024. [Online]. Available: <https://vida.id/id/pressrelease/penipuan-deepfake-indonesia-melonjak-1550-begini-cara-vida-memeranginya>. Accessed: Jul. 21, 2025.