

Perbandingan Model Deep Learning MobileNetV2, VGG19, InceptionV3, dan ResNet50 untuk Identifikasi Gambar Lukis Ciptaan Manusia dan AI

Comparison of Deep Learning Models MobileNetV2, VGG19, InceptionV3, and ResNet50 for the Identification of Human-Created and AI-Generated Paintings

M. Bahrul Subkhi*¹, Akbar Warisqianto²

Universitas PGRI Jombang, Jombang

*E-mail : bahruls27@gmail.com*¹, akbarwarisqia@gmail.com²*

**Corresponding author*

Received 5 August 2025; Revised 7 April 2026; Accepted 15 May 2026

Abstrak - Penelitian ini bertujuan untuk membandingkan efektivitas empat model deep learning—MobileNetV2, VGG19, InceptionV3, dan ResNet50—dalam membedakan antara gambar yang dihasilkan oleh kecerdasan buatan (AI) dan karya seni manusia. Keempat model ini diuji menggunakan dataset gambar yang terdiri dari dua kategori: AI-generated images dan human-generated images. Masing-masing model dilatih selama 30 epoch, dengan evaluasi menggunakan metrik akurasi dan kerugian untuk menilai kinerja pelatihan dan kemampuan generalisasi pada data validasi. Hasil penelitian menunjukkan bahwa MobileNetV2 mencapai akurasi pelatihan tertinggi (91,01%) dengan efisiensi waktu pelatihan yang lebih cepat, sementara InceptionV3 memberikan hasil akurasi yang hampir setara (90,13%) meskipun memerlukan lebih banyak waktu pelatihan. VGG19 dan ResNet50, meskipun memberikan hasil yang solid, mengalami konvergensi yang lebih lambat dan memerlukan lebih banyak sumber daya komputasi. Penurunan kerugian yang signifikan pada MobileNetV2 dan InceptionV3 menunjukkan kemampuan optimasi yang lebih baik dibandingkan dengan VGG19 dan ResNet50. Berdasarkan hasil tersebut, MobileNetV2 terbukti sebagai model yang paling efisien untuk aplikasi dengan keterbatasan sumber daya komputasi, sementara InceptionV3 lebih cocok untuk aplikasi yang memprioritaskan akurasi tinggi. Penelitian ini memberikan kontribusi terhadap pengembangan teknologi deteksi karya seni berbasis AI yang efisien dan akurat, serta membuka peluang untuk penelitian lanjutan dalam optimasi model deep learning untuk aplikasi yang lebih luas.

Kata kunci - Deteksi seni buatan AI , MobileNetV2, VGG19, InceptionV3, dan ResNet50

Abstract — *This study aims to compare the effectiveness of four deep learning models—MobileNetV2, VGG19, InceptionV3, and ResNet50—in distinguishing between images generated by artificial intelligence (AI) and human-created artworks. These four models were tested using an image dataset consisting of two categories: AI-generated images and human-generated images. Each model was trained for 30 epochs, with evaluation conducted using accuracy and loss metrics to assess training performance and generalization capability on validation data. The results show that MobileNetV2 achieved the highest training accuracy (91.01%) with faster training time efficiency, while InceptionV3 produced nearly comparable accuracy results (90.13%) despite requiring longer training time. VGG19 and ResNet50, although delivering solid results, experienced slower convergence and required more computational resources. The significant loss reduction observed in MobileNetV2 and InceptionV3 indicates better optimization capability compared to VGG19 and ResNet50. Based on these findings, MobileNetV2 proved to be the most efficient model for applications with limited computational resources, while InceptionV3 is more suitable for applications prioritizing high accuracy. This study contributes to the development of efficient and accurate AI-based artwork detection technology and opens opportunities for further research in optimizing deep learning models for broader applications.*

Keywords — *AI-generated art detection, MobileNetV2, VGG19, InceptionV3, and ResNet50.*

1. INTRODUCTION

Perkembangan pesat kecerdasan buatan (AI) telah membawa dampak signifikan dalam berbagai bidang, termasuk seni. Salah satu aplikasi AI yang paling menarik adalah kemampuannya untuk menghasilkan karya seni yang hampir tidak dapat dibedakan dari karya seni yang diciptakan oleh manusia. Fenomena ini, yang dikenal sebagai seni yang dihasilkan oleh AI, menimbulkan tantangan baru dalam mendeteksi apakah suatu gambar merupakan hasil ciptaan manusia atau AI. Meningkatnya kualitas gambar yang dihasilkan oleh AI, ditambah dengan kemudahan akses teknologi ini, memperburuk permasalahan terkait identifikasi dan keaslian karya seni [1] [2].

Beberapa penelitian telah dilakukan untuk menangani permasalahan ini dengan menggunakan berbagai pendekatan, baik yang melibatkan manusia maupun otomatis. Misalnya, Jeong Ha et al. menunjukkan bahwa kombinasi detektor otomatis dan manusia dapat memberikan akurasi yang lebih baik dalam membedakan seni manusia dan seni AI. Meskipun detektor otomatis dapat memiliki ketahanan yang baik terhadap gangguan adversarial, detektor manusia, terutama para ahli seni, dapat mengidentifikasi karya seni dengan akurasi tinggi namun sering kali menghasilkan lebih banyak kesalahan positif [1].

Penerapan teknik deep learning untuk identifikasi seni AI semakin berkembang, dengan beberapa model seperti CNN (Convolutional Neural Network) digunakan untuk menganalisis gambar seni. Xu [3] mengemukakan penggunaan CNN dalam mengenali gaya seni dan mengidentifikasi seniman berdasarkan karyanya. Model-model yang dilatih menggunakan arsitektur pre-trained, seperti ResNet50, terbukti efektif dalam mencapai akurasi tinggi dalam pengenalan gaya seni, meskipun tantangan yang ada termasuk keberagaman gaya artistik yang sangat tinggi dalam dataset seni yang digunakan.

Dalam studi yang dilakukan oleh Duan [4], model CNN yang dilatih untuk membedakan antara lukisan ciptaan manusia dan AI, seperti yang dihasilkan oleh model CycleGAN, menunjukkan akurasi mencapai 95.26%. Angka ini lebih tinggi dibandingkan dengan akurasi yang diperoleh oleh ahli seni manusia, yang hanya mencapai 82.59%. Temuan ini mengindikasikan bahwa model CNN lebih efektif dalam mengenali perbedaan antara lukisan manusia dan lukisan AI, bahkan ketika dibandingkan dengan keahlian manusia. Hal ini menunjukkan potensi besar deep learning dalam deteksi seni buatan AI yang semakin sulit dibedakan dari seni manusia.

Namun, meskipun metode otomatis telah terbukti efektif, tantangan etis tetap ada, terutama terkait dengan hak cipta dan kepemilikan karya seni yang dihasilkan oleh AI. Li et al. [5] mencatat bahwa generasi gambar oleh AI dapat digunakan untuk penipuan atau penyebaran informasi yang salah, yang mengharuskan adanya alat yang efektif untuk mendeteksi dan membedakan antara gambar asli dan buatan AI.

Berdasarkan temuan-temuan tersebut, penelitian ini bertujuan untuk membandingkan penerapan deep learning dalam membedakan gambar lukisan ciptaan manusia atau AI, serta untuk memahami efektivitas berbagai metode yang ada dalam menghadapi tantangan ini. Beberapa penelitian sebelumnya telah mengkaji deteksi gambar berbasis AI dari berbagai sudut pandang. Ha et al. [1] menunjukkan bahwa kombinasi detektor otomatis dan manusia memberikan akurasi terbaik dalam membedakan seni manusia dan seni AI. Castellano dan Vessio [2] memberikan tinjauan komprehensif tentang pendekatan deep learning untuk ekstraksi pola dalam lukisan dan gambar, menekankan pentingnya arsitektur pre-trained. Xu [3] menggunakan CNN dan transfer learning untuk identifikasi seniman secara otomatis dengan akurasi tinggi. Duan [4] membuktikan model CNN mencapai 95,26% dalam membedakan lukisan manusia dan AI. Li et al. [5] menyajikan studi sistematis tentang deteksi gambar AI dalam skenario adversarial menggunakan dataset ARIA. Narne et al. [6] menerapkan pendekatan deep learning canggih untuk mendeteksi

gambar nyata dan palsu. Yanamadala et al. [7] menggunakan deep learning untuk pengenalan kualitas sketsa karakter anime. Hayathunnisa et al. [8] menggunakan CNN kustom dan VGG19 untuk identifikasi gambar nyata versus palsu dan mencapai akurasi sangat baik. Masilamani dan Venkatesh [9] mengaplikasikan transfer learning dengan PyTorch untuk klasifikasi seni. López-Rubio et al. [10] mengklasifikasi gambar sebagai foto atau lukisan menggunakan CNN. Kumar et al. [11] membangun classifier berbasis CNN untuk membedakan gambar yang dihasilkan AI dan gambar nyata. Ye et al. [12] mengkaji framework komputasional untuk membedakan karya seni manusia dan AI. Bird dan Lotfi [13] mengembangkan dataset CIFAKE untuk mengklasifikasi gambar sintetik dari gambar nyata menggunakan ResNet-50. De Goma et al. [14] membandingkan lima arsitektur CNN untuk membedakan karya seni digital manusia dan AI. Chinta et al. [15] menganalisis klasifikasi gambar seni AI versus manusia dan menemukan bahwa augmentasi data meningkatkan akurasi secara signifikan.

Meskipun berbagai penelitian di atas telah memberikan kontribusi signifikan, terdapat research gap yang belum ditangani: belum ada studi yang secara sistematis membandingkan performa MobileNetV2, VGG19, InceptionV3, dan ResNet50 secara bersamaan dalam kondisi eksperimen identik pada dataset seni AI versus seni manusia dengan mempertimbangkan keseimbangan antara efisiensi komputasi dan akurasi. Penelitian ini mengisi gap tersebut dengan melakukan perbandingan kuantitatif keempat model pada dataset yang sama, kondisi pelatihan identik (30 epoch, augmentasi seragam), serta mengevaluasi secara simultan akurasi, loss, dan kecepatan konvergensi.

2. RESEARCH METHOD

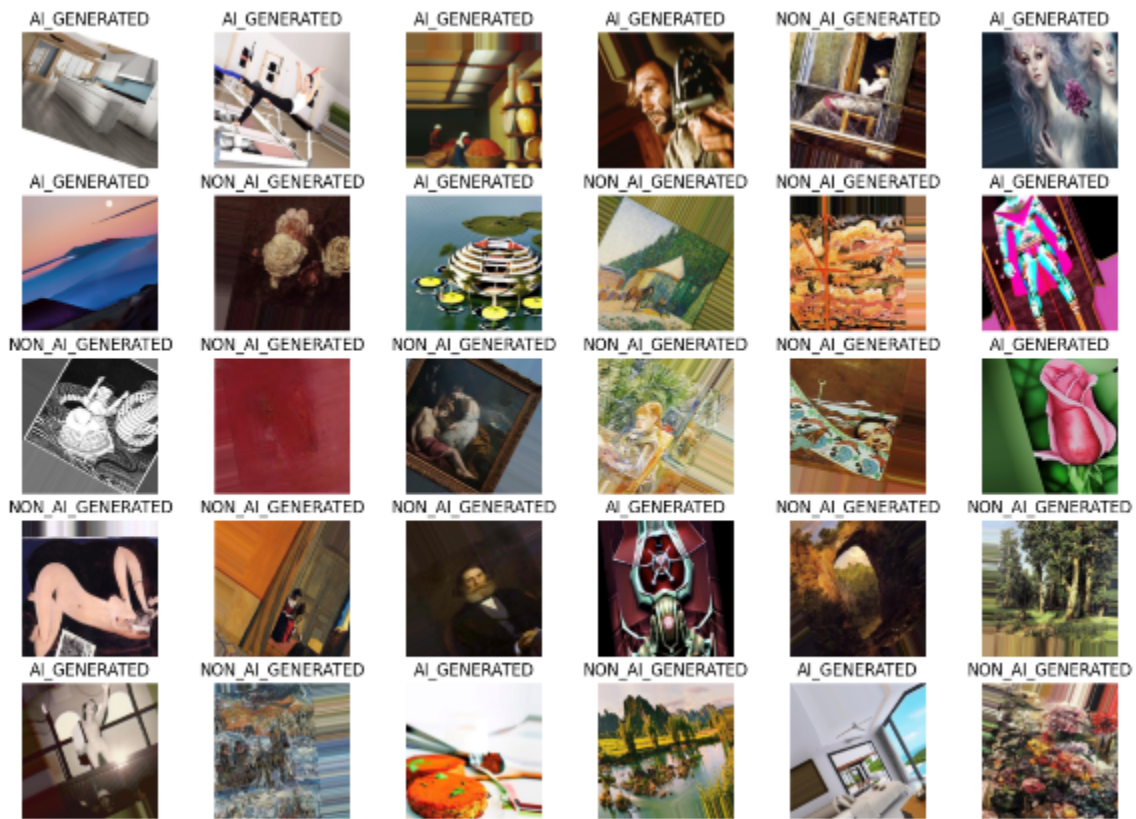
Perkembangan kecerdasan buatan (AI) yang pesat, khususnya dalam pembuatan karya seni digital, telah menimbulkan tantangan besar dalam membedakan antara karya seni yang dihasilkan oleh AI dan karya seni ciptaan manusia [1]. Seiring dengan kemampuan AI yang semakin canggih dalam menciptakan gambar yang sangat realistis, penting untuk mengembangkan metode yang dapat membedakan keduanya. [6] Penelitian ini bertujuan untuk mengeksplorasi dan membandingkan empat model deep learning—MobileNetV2, VGG19, InceptionV3, dan ResNet50—untuk mengidentifikasi gambar lukisan yang dihasilkan oleh manusia dan AI. Tujuan utama penelitian ini adalah untuk mengetahui model mana yang paling efektif dalam mendeteksi perbedaan antara karya seni AI dan karya seni manusia, serta untuk menganalisis akurasi, kecepatan pelatihan, dan ketahanan model terhadap data yang beragam.

2.1. Pengumpulan Data

Untuk penelitian ini, digunakan dataset yang berisi gambar-gambar yang dikategorikan sebagai karya seni yang dihasilkan oleh AI dan karya seni yang dihasilkan oleh manusia. Dataset menggunakan data terbuka di kaggle <https://www.kaggle.com/datasets/kausthubkannan/ai-and-human-art-classification/data>. Setiap gambar dalam dataset diproses dan diubah ukurannya menjadi 224x224 piksel untuk disesuaikan dengan input model deep learning yang digunakan. Jumlah data yang dibuat AI 8261 gambar dan 6631 gambar yang dibuat bukan AI, jadi total data yang digunakan 14892 gambar. Contoh gambar yang digunakan dalam penelitian ini dapat dilihat pada gambar 1.

Dataset ini kemudian dibagi menjadi dua subset utama:

- a. Set Pelatihan: 80% dari total dataset digunakan untuk melatih model.
- b. Set Validasi: 20% dari dataset digunakan untuk mengevaluasi kinerja model setelah pelatihan.



Gambar 1 contoh gambar yang digunakan.

Untuk meningkatkan keberagaman data dan memperkuat kemampuan model dalam menggeneralisasi, dilakukan teknik augmentasi data. Augmentasi ini mencakup rotasi gambar, pergeseran gambar secara horizontal dan vertikal, shear, zoom, dan pembalikan horizontal gambar. Selain itu, gambar juga diubah skalanya sehingga piksel berada dalam rentang [0, 1] dengan membaginya dengan 255. Augmentasi ini bertujuan untuk memperkenalkan variasi pada data, sehingga model dapat belajar dari lebih banyak contoh dan lebih tahan terhadap overfitting. [3] [2]

2.2. Arsitektur Model

Penelitian ini menggunakan empat model deep learning yang berbeda, yaitu MobileNetV2, VGG19, InceptionV3, dan ResNet50, untuk membandingkan kinerja masing-masing model dalam membedakan antara gambar karya seni manusia dan AI. Setiap model memiliki keunggulan dan kekurangan masing-masing, yang akan dieksplorasi dalam penelitian ini.

- a. MobileNetV2: Model ini dirancang untuk efisiensi komputasi dan digunakan untuk tugas klasifikasi gambar yang membutuhkan sumber daya komputasi yang lebih rendah. Dengan menggunakan transfer learning, MobileNetV2 yang telah dilatih pada dataset ImageNet digunakan sebagai model dasar dan disesuaikan untuk klasifikasi biner (AI-generated vs human-generated). Arsitektur MobileNetV2 terdiri dari 53 layer dengan sekitar 3,4 juta parameter dan menggunakan blok inverted residual with linear bottleneck. Secara matematis, output blok ini adalah $y = x + F(x)$ jika dimensi sama, atau $y = F(x)$

- jika berbeda. Fungsi aktivasi yang digunakan adalah ReLU6: $f(x) = \min(\max(0,x), 6)$, dengan faktor ekspansi $t=6$ sebelum depthwise separable convolution [6]. [6]
- b. VGG19: Merupakan model yang lebih dalam dengan lebih banyak parameter dibandingkan MobileNetV2. VGG19 dikenal karena struktur arsitekturnya yang dalam dan kesederhanaannya, yang memungkinkan untuk menangkap detail tinggi dalam gambar seni, namun dengan biaya komputasi yang lebih besar. VGG19 memiliki 16 convolutional layer dan 3 fully-connected layer dengan total sekitar 143 juta parameter. Setiap convolutional layer menggunakan filter 3×3 dengan fungsi aktivasi ReLU: $f(x) = \max(0,x)$, dan max-pooling 2×2 untuk downsampling. Output layer diadaptasi menggunakan sigmoid: $\sigma(z) = 1/(1+e^{(-z)})$ untuk klasifikasi biner [5]. [5]
 - c. InceptionV3: Model ini memiliki arsitektur yang lebih kompleks dengan berbagai tingkat kedalaman dan menggunakan teknik yang lebih canggih dalam menangani data yang lebih besar. InceptionV3 sering digunakan untuk pengenalan gambar yang lebih rumit dan dapat memberikan hasil yang lebih baik untuk tugas-tugas klasifikasi gambar yang membutuhkan analisis lebih mendalam. InceptionV3 memiliki sekitar 48 layer dengan 23,8 juta parameter dan menggunakan modul Inception yang memproses input secara paralel: $h(x) = \text{concat}[f_{1 \times 1}(x), f_{3 \times 3}(x), f_{5 \times 5}(x), \text{pool}(x)]$. Teknik factorized convolution mengurangi kompleksitas, misalnya konvolusi 5×5 difaktorkan menjadi dua konvolusi 3×3 berurutan. Batch Normalization diterapkan setelah setiap layer untuk menstabilkan pelatihan.
 - d. ResNet50: Dikenal karena arsitektur residual yang memungkinkan model untuk sangat dalam tanpa mengalami masalah hilangnya gradien. ResNet50 cocok untuk tugas klasifikasi gambar yang kompleks dan diharapkan memberikan akurasi yang sangat baik, terutama pada data yang besar dan beragam. ResNet50 memiliki 50 layer dengan sekitar 25,6 juta parameter menggunakan skip connection: $y = F(x, \{W_i\}) + x$. Apabila dimensi berbeda digunakan proyeksi linear: $y = F(x, \{W_i\}) + W_s * x$. Setiap blok residual menggunakan struktur bottleneck tiga konvolusi ($1 \times 1 \rightarrow 3 \times 3 \rightarrow 1 \times 1$) dengan Batch Normalization dan ReLU [6]. [6]

Keempat model ini dilatih dengan menggunakan class weights untuk menangani ketidakseimbangan kelas yang mungkin ada dalam dataset, di mana satu kelas (misalnya, gambar yang dihasilkan oleh AI) lebih banyak dibandingkan kelas lainnya (gambar yang dihasilkan oleh manusia). [5] [6]

2.3. Pelatihan Model

Keempat model dilatih menggunakan ImageDataGenerator, yang bertanggung jawab untuk augmentasi data dan pembagian dataset menjadi dua subset: pelatihan dan validasi. Setiap model dilatih selama 30 epoch untuk memastikan model memiliki waktu yang cukup untuk mempelajari fitur dari data.

Selain itu, beberapa callback digunakan untuk meningkatkan kinerja dan efisiensi pelatihan:

- a. EarlyStopping: Memonitor kerugian validasi dan menghentikan pelatihan jika tidak ada perbaikan dalam beberapa epoch berturut-turut. Hal ini untuk mencegah overfitting dan memastikan model tidak melatih terlalu lama. [4]
- b. ReduceLROnPlateau: Mengurangi laju pembelajaran jika kerugian validasi tidak membaik dalam beberapa epoch berturut-turut, yang dapat membantu model untuk berkonvergensi lebih cepat.
- c. ModelCheckpoint: Menyimpan model terbaik berdasarkan akurasi validasi untuk memastikan model yang optimal dipilih.

2.4. Evaluasi Kinerja Model

Setelah pelatihan selesai, keempat model dievaluasi menggunakan dataset validasi untuk mengukur akurasi, presisi, recall, dan F1-score. Confusion matrix juga digunakan untuk menganalisis distribusi prediksi positif sejati, negatif sejati, positif palsu, dan negatif palsu yang memberikan wawasan lebih dalam mengenai kinerja model. Dengan ini, dapat dianalisis seberapa baik model dapat membedakan antara gambar AI dan gambar manusia dalam berbagai kondisi. [7]

Selain itu, dibandingkan pula kecepatan pelatihan dan ketahanan model terhadap augmentasi data, untuk mengevaluasi efisiensi masing-masing model. Sebagai contoh, MobileNetV2 diharapkan lebih cepat dalam pelatihan, sementara VGG19 dan ResNet50 diharapkan memiliki akurasi yang lebih tinggi, meskipun dengan waktu pelatihan yang lebih lama. [8]

2.5. Perbandingan Kuantitatif

Perbandingan kuantitatif dilakukan untuk membandingkan hasil kinerja masing-masing model berdasarkan [5]:

- a. Akurasi: Kemampuan model dalam mengklasifikasikan gambar dengan benar.
- b. Kecepatan Pelatihan: Waktu yang diperlukan oleh model untuk melatih dataset.
- c. Ketahanan terhadap Augmentasi Data: Kemampuan model untuk mengatasi variasi data yang dihasilkan oleh teknik augmentasi.

Sebagai contoh, ResNet50 dan InceptionV3 diharapkan memberikan akurasi yang lebih tinggi dibandingkan dengan MobileNetV2, namun dengan waktu pelatihan yang lebih lama. VGG19 juga diperkirakan menunjukkan akurasi yang sangat baik, tetapi dengan penggunaan sumber daya komputasi yang lebih besar.

Penelitian ini bertujuan untuk memberikan gambaran yang lebih jelas mengenai penggunaan deep learning untuk membedakan antara gambar yang dihasilkan oleh AI dan karya seni manusia, dengan membandingkan empat model deep learning yang berbeda. Hasil penelitian ini diharapkan dapat memberikan wawasan yang lebih mendalam mengenai kekuatan dan kelemahan masing-masing model dalam hal akurasi, kecepatan, dan ketahanan terhadap variasi data, serta memberikan kontribusi terhadap pengembangan teknologi deteksi karya seni AI dalam berbagai aplikasi, seperti moderasi konten dan perlindungan hak cipta [4] [9]

3. RESULTS AND DISCUSSION

Hasil pelatihan empat model—MobileNetV2, VGG19, InceptionV3, dan ResNet50—untuk tugas klasifikasi antara gambar yang dihasilkan oleh AI dan karya seni manusia menunjukkan perbedaan yang signifikan dalam hal akurasi, kerugian, dan kecepatan konvergensi selama pelatihan. Semua model dievaluasi selama 30 epoch, dengan metrik yang meliputi akurasi pelatihan, akurasi validasi, kerugian pelatihan, dan kerugian validasi.

Tabel 1 hasil akurasi dan loss pada epoch 1

Model	Akurasi Pelatihan (Epoch 1)	Akurasi Validasi (Epoch 1)	Kerugian Pelatihan (Epoch 1)	Kerugian Validasi (Epoch 1)

MobileNetV2	76.56%	86.57%	0.5148	0.3263
VGG19	71.07%	74.81%	0.5428	0.5039
InceptionV3	69.92%	84.21%	0.8730	0.3600
ResNet50	60.96%	66.91%	0.6990	0.5989

Tabel 2 hasil akurasi dan loss pada epoch 30

Model	Akurasi Pelatihan (Epoch 30)	Akurasi Validasi (Epoch 30)	Kerugian Pelatihan (Epoch 30)	Kerugian Validasi (Epoch 30)
MobileNetV2	91.01%	88.85%	0.2252	0.2612
VGG19	85.46%	82.86%	0.3296	0.3823
InceptionV3	90.13%	85.90%	0.2314	0.3421
ResNet50	74.25%	75.18%	0.5217	0.5131

Tabel 1 menjelaskan akurasi pelatihan pada epoch pertama, di mana semua model menunjukkan akurasi pelatihan di bawah 80%. MobileNetV2 mencatatkan akurasi pelatihan sebesar 76,56%, diikuti oleh VGG19 dengan 71,07%, InceptionV3 dengan 69,92%, dan ResNet50 dengan 60,96%. Meskipun model-model ini memulai pelatihan dengan akurasi yang relatif rendah, peningkatan yang signifikan dapat diamati seiring berjalannya waktu. Namun, setelah beberapa epoch, seperti pada tabel 2 semua model menunjukkan peningkatan yang stabil dalam akurasi pelatihan. MobileNetV2 berhasil mencapai akurasi pelatihan tertinggi pada epoch ke-30 yaitu 91,01%, sementara InceptionV3 mengikuti dengan 90,13%. VGG19 mencapai 85,46%, dan ResNet50 menunjukkan 74,25% pada akhir pelatihan. Meskipun ResNet50 memiliki kinerja awal yang lebih rendah, model ini berhasil mengalami peningkatan yang cukup signifikan seiring berjalannya pelatihan, meskipun tidak mencapai tingkat akurasi model-model lainnya.

Pada Tabel 2 MobileNetV2 menunjukkan kinerja terbaik dengan akurasi pelatihan yang mencapai 91,01% pada epoch ke-30, diikuti oleh InceptionV3 dengan akurasi pelatihan 90,13%. VGG19 mengalami peningkatan yang lebih lambat, dengan akurasi pelatihan 85,46%, sementara ResNet50 memiliki akurasi pelatihan terendah pada 74,25%.

Dari segi akurasi validasi, MobileNetV2 tetap menjadi model yang unggul dengan 88,85% pada epoch terakhir, diikuti oleh InceptionV3 dengan 85,90%. VGG19 menunjukkan akurasi validasi yang stabil di sekitar 82,86%, sedangkan ResNet50 memiliki akurasi validasi 75,18%, yang lebih rendah dibandingkan model lainnya.

Dalam hal kerugian, MobileNetV2 dan InceptionV3 berhasil mengurangi kerugian pelatihan dan kerugian validasi secara signifikan, menunjukkan bahwa kedua model ini mengoptimalkan pembelajarannya dengan sangat baik. VGG19 dan ResNet50 juga menunjukkan penurunan kerugian, meskipun dengan tingkat yang sedikit lebih lambat dibandingkan dengan MobileNetV2 dan InceptionV3.

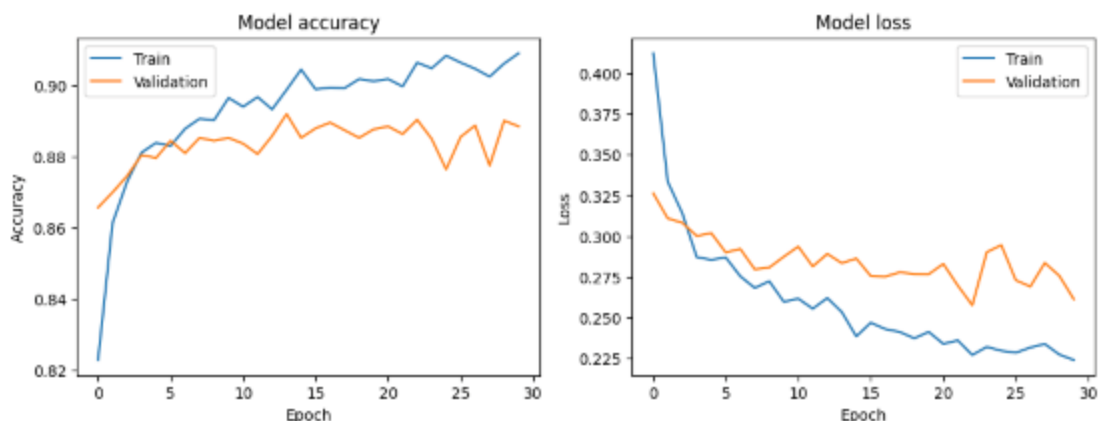
MobileNetV2 menunjukkan akurasi pelatihan tertinggi di antara keempat model, dengan kemampuan konvergensi yang cepat dan efisien. Hal ini menunjukkan bahwa MobileNetV2 dapat belajar dari data pelatihan dengan lebih cepat, dan mampu menggeneralisasi dengan baik pada data validasi, meskipun dengan sumber daya komputasi yang lebih rendah. Keunggulan utama MobileNetV2 adalah kemampuannya untuk bekerja pada perangkat dengan keterbatasan sumber daya, membuatnya sangat ideal untuk aplikasi dengan batasan perangkat keras.

InceptionV3 juga menunjukkan hasil yang sangat baik, dengan akurasi pelatihan mencapai 90,13% dan akurasi validasi mencapai 85,90%. Model ini, meskipun lebih kompleks dan memerlukan lebih banyak waktu pelatihan, dapat menangkap fitur yang lebih mendalam dalam gambar, menjadikannya pilihan yang sangat baik ketika akurasi lebih diutamakan dan sumber daya komputasi tidak menjadi kendala.

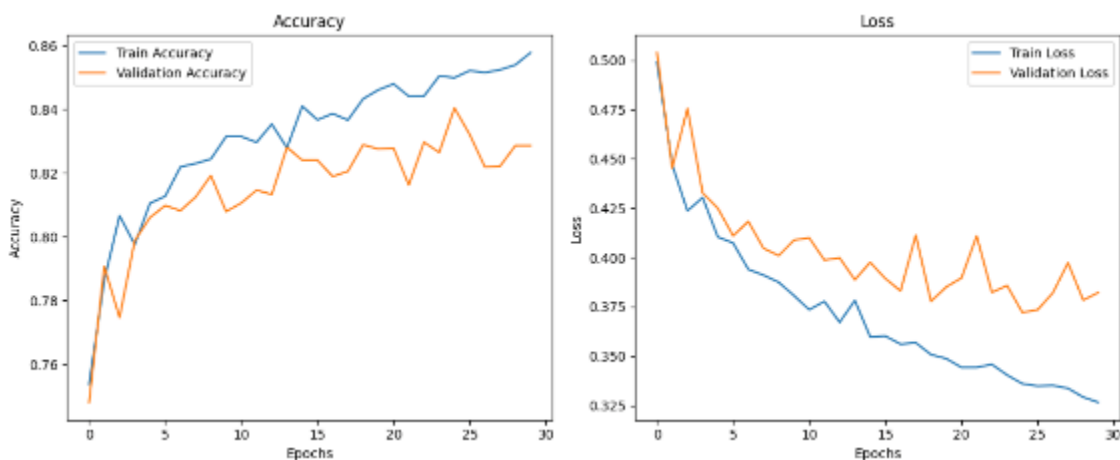
VGG19, meskipun memiliki arsitektur yang lebih dalam dan lebih banyak parameter, menunjukkan hasil yang lebih lambat dalam hal konvergensi dibandingkan dengan MobileNetV2

dan InceptionV3. VGG19 berhasil mencapai akurasi pelatihan sebesar 85,46%, namun membutuhkan waktu pelatihan yang lebih lama, yang menjadikannya kurang efisien dibandingkan dengan model yang lebih ringan.

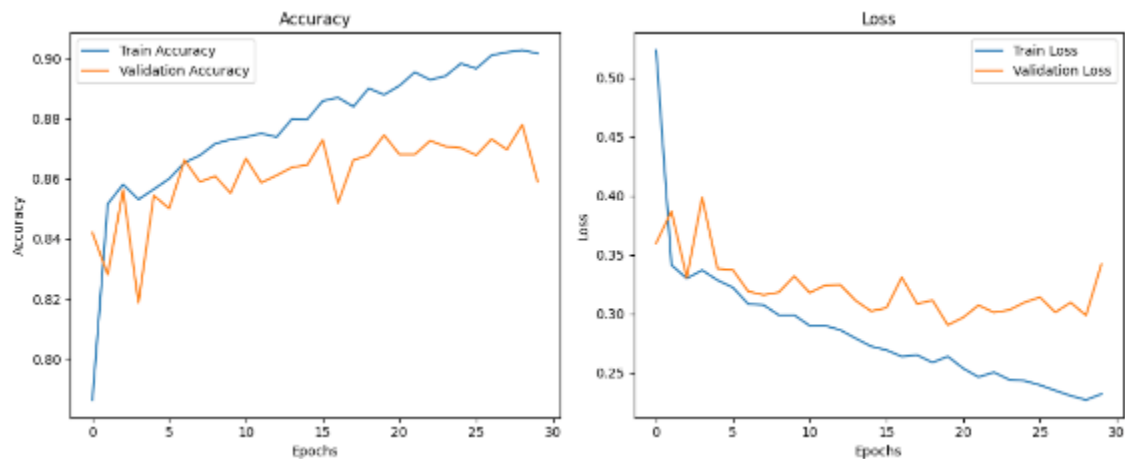
ResNet50 menunjukkan akurasi pelatihan terendah, yaitu 74,25%, dengan fluktuasi yang lebih besar selama pelatihan. Model ini, yang dirancang untuk arsitektur yang lebih dalam, mengalami kesulitan dalam mencapai konvergensi yang cepat. ResNet50 juga memiliki akurasi validasi yang lebih rendah, yaitu 75,18%, yang menunjukkan bahwa meskipun model ini memiliki kemampuan untuk menangani data yang lebih besar, ia tidak dapat menggeneralisasi dengan baik pada data yang tidak terlihat sebelumnya.



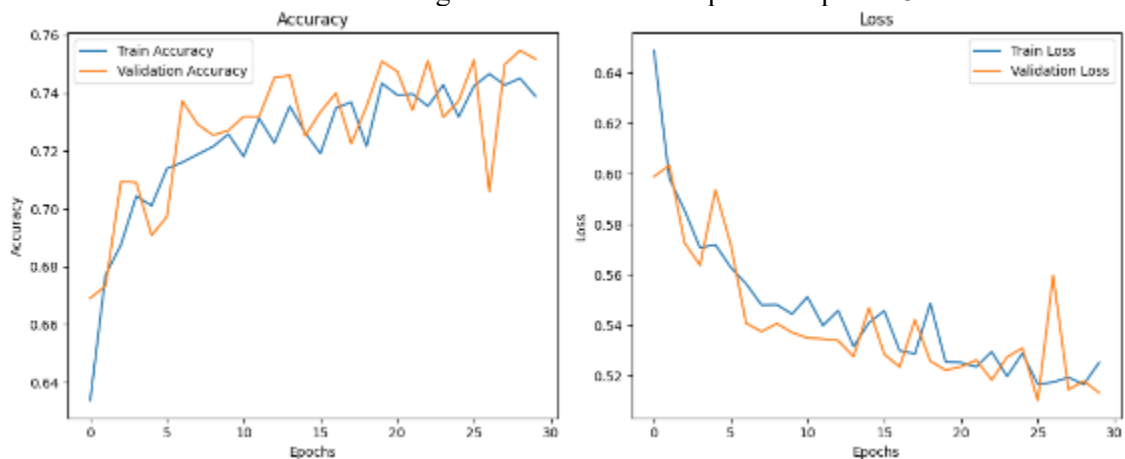
Gambar 2 grafik akurasi dan loss pada mobileNetV2



Gambar 3 grafik akurasi dan loss pada VGG19



Gambar 4 grafik akurasi dan loss pada InceptionV3



Gambar 5 grafik akurasi dan loss pada RestNet50

Grafik akurasi pelatihan dan akurasi validasi menunjukkan bahwa MobileNetV2 dan InceptionV3 pada gambar 2 dan gambar 4 memiliki kenaikan yang lebih cepat dalam akurasi pelatihan, sementara VGG19 dan ResNet50 pada gambar 3 dan 5 cenderung lebih lambat. MobileNetV2 menunjukkan peningkatan yang hampir linear dalam akurasi, dengan sedikit fluktuasi, menunjukkan konsistensi model ini dalam mempelajari fitur dari data pelatihan dan validasi. InceptionV3 juga menunjukkan hasil yang baik, meskipun memerlukan lebih banyak waktu pelatihan per epoch.

Pada grafik kerugian, MobileNetV2 dan InceptionV3 pada gambar 2 dan gambar 4 menunjukkan penurunan yang lebih stabil dan signifikan dalam kerugian pelatihan dan kerugian validasi. Sebaliknya, VGG19 dan ResNet50 pada gambar 3 dan 5 menunjukkan penurunan yang lebih lambat dan fluktuasi yang lebih besar, yang mungkin menunjukkan kesulitan dalam optimisasi dan generalisasi.

4. CONCLUSION

Dalam bab Pendahuluan, telah diungkapkan bahwa tujuan utama dari penelitian ini adalah untuk membandingkan efektivitas beberapa model deep learning, yaitu MobileNetV2, VGG19, InceptionV3, dan ResNet50, dalam membedakan gambar yang dihasilkan oleh AI dan karya seni manusia. Berdasarkan hasil yang diperoleh di bab Hasil dan Diskusi, yang menunjukkan perbandingan akurasi dan kerugian masing-masing model, dapat disimpulkan

bahwa MobileNetV2 menunjukkan hasil yang paling optimal dalam hal akurasi pelatihan dan validasi, dengan waktu pelatihan yang efisien. InceptionV3 juga menunjukkan kinerja yang sangat baik, meskipun memerlukan waktu pelatihan yang lebih lama. VGG19 dan ResNet50, meskipun memberikan hasil yang solid, tidak seefisien MobileNetV2 dan InceptionV3 dalam hal kecepatan pelatihan dan efisiensi komputasi.

Hasil penelitian ini sejalan dengan ekspektasi yang telah dijelaskan dalam bab Pendahuluan, di mana diharapkan model-model yang lebih ringan seperti MobileNetV2 dapat menunjukkan kinerja yang lebih baik dalam hal efisiensi dan akurasi, sedangkan model-model yang lebih kompleks seperti VGG19 dan ResNet50 menunjukkan akurasi yang baik tetapi dengan waktu pelatihan yang lebih lama. Dengan demikian, terdapat kesesuaian antara tujuan penelitian dan hasil yang diperoleh.

Ke depan, MobileNetV2 dan InceptionV3 dapat menjadi pilihan utama dalam aplikasi yang membutuhkan keseimbangan antara akurasi dan efisiensi, khususnya untuk perangkat dengan sumber daya terbatas. Selain itu, penelitian lebih lanjut dapat mengkaji lebih dalam tentang penerapan transfer learning dan teknik augmentasi data untuk meningkatkan akurasi dan generalisasi model-model ini pada dataset yang lebih besar dan lebih beragam. Penelitian selanjutnya juga dapat mengeksplorasi penggunaan arsitektur model yang lebih dalam atau penggabungan beberapa model untuk mencapai kinerja yang lebih optimal dalam tugas klasifikasi gambar seni AI versus manusia. Dengan adanya temuan ini, diharapkan dapat memberikan kontribusi signifikan pada pengembangan aplikasi deteksi gambar seni berbasis AI yang lebih efisien dan akurat di masa depan.

REFERENCES

- [1] A. Y. J. Ha, J. Passananti, R. Bhaskar, S. Shan, R. Southen, H. Zheng, and B. Y. Zhao, "Organic or Diffused: Can We Distinguish Human Art from AI-generated Images?," *Proc. 2024 ACM SIGSAC Conf. on Computer and Communications Security (CCS)*, pp. 4822–4836, 2024. doi: 10.1145/3658644.3670306.
- [2] G. Castellano and G. Vessio, "Deep learning approaches to pattern extraction and recognition in paintings and drawings: an overview," *Neural Computing and Applications*, vol. 33, pp. 12263–12282, 2021. doi: 10.1007/s00521-021-05893-z.
- [3] K. Xu, "Automated Artist Identification Using Deep Learning and Transfer Learning Techniques," *Science and Technology of Engineering, Chemistry and Environmental Protection*, vol. 1, no. 3, 2024.
- [4] Z.-L. Duan, "Who Made the Oil Paintings: Machine or Human?," *arXiv preprint*, 2023. arXiv:2302.xxxxx.
- [5] Y. Li, Z. Liu, J. Zhao, L. Ren, F. Li, J. Luo, and B. Luo, "The Adversarial AI-Art: Understanding, Generation, Detection, and Benchmarking," in *Computer Security – ESORICS 2024*, LNCS vol. 14982, pp. 311–331, 2024. doi: 10.1007/978-3-031-70879-4_16.
- [6] V. Narne, C. D. Koduru, K. Rao, and R. Kumar, "Advanced Deep Learning Approaches for Detecting Real and Fake Images," in *2024 1st Int. Conf. on Innovations in Communications, Electrical and Computer Engineering (ICICEC)*, pp. 1–6, 2024.

- [7] I. Yanamadala, A. Das, D. Roy, and D. Raychaudhuri, “Deep Learning Based Anime Character Sketch Quality Recognition,” *Journal of Student Research*, vol. 13, no. 1, 2024.
- [8] V. Hayathunnisa, K. Pradeep, M. Anand, and A. Mohan, “Art of Detection: Custom CNN and VGG19 for Accurate Real Vs Fake Image Identification,” in *Proc. 2023 6th Int. Conf. on Recent Trends in Advance Computing (ICRTAC)*, pp. 306–312, 2024.
- [9] G. Masilamani and V. Ramasubramanian, “Art Classification with PyTorch Using Transfer Learning,” in *2021 Int. Conf. on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–6, 2021.
- [10] J. M. López-Rubio, M. A. Molina-Cabello, G. Rodríguez-Jerónimo, and E. López-Rubio, “Classification of Images as Photographs or Paintings by Using Convolutional Neural Networks,” in *Lecture Notes in Computer Science*, vol. 12861, pp. 1–12, Springer, 2021.
- [11] U. Kumar, V. Chawla, S. Shandilya, and K. Krishnamoorthy, “A Convolution Neural Network Based Classifier for Discerning AI Generated Images and Real Images,” in *2024 4th Asian Conf. on Innovation in Technology (ASIANCON)*, pp. 1–6, 2024.
- [12] Y. Ye, R. He, K. Zhang, Z. Wu, and W. Zhao, “Everyone Can Be Picasso? A Computational Framework into the Myth of Human versus AI Painting,” *arXiv preprint*, arXiv:2312.xxxxx, 2023.
- [13] J. J. Bird and A. Lotfi, “CIFAKE: Image Classification and Explainable Identification of AI-Generated Synthetic Images,” *IEEE Access*, vol. 12, pp. 15642–15650, 2024. doi: 10.1109/ACCESS.2024.3356188.
- [14] J. C. De Goma, J. B. L. Olivar, H. L. Perez, S. C. E. Velasco, and A. C. Villaluz, “Art Authentication: A Comparative Analysis of CNN Architectures for Detecting AI-Generated and Human-made Digital Artworks,” in *Proc. 2024 5th Asia Service Sciences and Software Engineering Conf. (ASSE)*, pp. 147–152, 2025. doi: 10.1145/3702138.3702147.
- [15] D. S. Chinta, S. Kamineni, R. P. Chatragadda, and S. Kamepalli, “Analyzing Image Classification on AI-Generated Art Vs Human Created Art Using Deep Learning Models,” in *Proc. 4th Workshop on Security Implications of Deepfakes and Cheapfakes*, pp. 1–6, 2024. doi: 10.1145/3709022.3736544.