Journal of Information Technology and Computer Science (INTECOMS)
Volume 9 Nomor 5, Tahun 2026
e-ISSN : 2614-1574
p-ISSN : 2621-3249

DETEKSI GAMBAR AI MULTIKELAS MENGGUNAKAN MODEL HYBRID CNN-ViT DAN BROAD LEARNING SYSTEM (BLS) BERBASIS WEB

WEB-BASED MULTI-CLASS AI IMAGE DETECTION USING A HYBRID CNN-VIT MODEL AND BROAD LEARNING SYSTEM (BLS)

Adinda Karen Mulyono[1], A Taqwa Martadinata[2], Fido Rizki[3], Asep Toyip Hidayat[4]
Fakultas Ilmu Teknik, Program Studi Informatika, Universitas Bina Insan, Lubuklinggau[1,2,3,4]
karendinda0610@gmail.co[1], taqwa@univbinainsan.ac.id [2], fidorizki@univbinainsan.ac.id[3], asepfighter@gmail.com[4]

ABSTRACT

The development of artificial intelligence technology has enabled the creation of highly realistic synthetic images, making the verification of digital image authenticity increasingly challenging. This study aims to develop a web-based multi-class AI image detection system using a hybrid Convolutional Neural Network (CNN) and Vision Transformer (ViT) model combined with a Broad Learning System (BLS). The system is designed to classify images into six categories: real animals, AI-generated animals, real humans, AI-generated humans, real objects, and AI-generated objects. The hybrid model utilizes EfficientNet-B0 as a local feature extractor and Vision Transformer as a global feature extractor. The extracted features are then combined through a feature fusion mechanism and classified using the Broad Learning System to efficiently generate multi-class predictions. The system is implemented as a web application using Streamlit as the user interface and FastAPI as the backend. The test results show that the hybrid CNN–ViT–BLS model achieves high classification accuracy with minimal error. Furthermore, the web-based system operates stably and responsively, indicating its potential as a supporting tool for verifying the authenticity of digital images.

Keywords: *AI Image Detection, CNN-ViT, Broad Learning, Multiclass Classification.*

ABSTRAK

Perkembangan teknologi kecerdasan buatan memungkinkan pembuatan gambar sintetis yang sangat realistis sehingga menyulitkan proses verifikasi keaslian citra digital. Penelitian ini bertujuan untuk mengembangkan sistem deteksi gambar AI multikelas berbasis web menggunakan model hybrid Convolutional Neural Network (CNN) dan Vision Transformer (ViT) yang dikombinasikan dengan Broad Learning System (BLS). Sistem ini dirancang untuk mengklasifikasikan enam kategori citra yaitu hewan asli, hewan AI, manusia asli, manusia AI, benda asli, dan benda AI. Model hybrid memanfaatkan EfficientNet-B0 sebagai ekstraktor fitur lokal dan Vision Transformer sebagai ekstraktor fitur global. Fitur yang diperoleh kemudian digabungkan melalui mekanisme feature fusion dan diklasifikasikan menggunakan Broad Learning System untuk menghasilkan prediksi multikelas secara efisien. Sistem diimplementasikan dalam bentuk aplikasi web menggunakan Streamlit sebagai antarmuka pengguna dan FastAPI sebagai backend. Hasil pengujian menunjukkan bahwa model hybrid CNN–ViT–BLS mampu mencapai akurasi klasifikasi yang tinggi dengan tingkat kesalahan yang minimal. Selain itu, sistem berbasis web dapat berjalan dengan stabil dan responsif sehingga berpotensi digunakan sebagai alat bantu verifikasi keaslian citra digital.

Kata Kunci: AI Image Detection, CNN-ViT, Broad Learning, Multiclass Classification

PENDAHULUAN

Perkembangan teknologi kecerdasan buatan (*Artificial Intelligence*) khususnya pada bidang *computer vision* telah memberikan kontribusi besar dalam pengolahan dan analisis citra digital. Teknologi *deep learning* banyak digunakan karena mampu mengekstraksi fitur visual secara otomatis dari citra tanpa memerlukan proses rekayasa fitur secara manual [1][2]. Penerapannya mencakup berbagai bidang, seperti klasifikasi citra, deteksi objek, dan pengenalan pola visual. Salah satu metode *deep learning* yang banyak digunakan dalam pengolahan citra adalah *Convolutional Neural Network* (CNN). CNN mampu mengekstraksi fitur visual secara hierarkis melalui lapisan konvolusi sehingga dapat mengenali pola, tekstur, dan bentuk objek pada citra [3][4].

Namun, CNN lebih berfokus pada informasi lokal sehingga memiliki keterbatasan dalam memahami hubungan global antarbagian citra.

Perkembangan selanjutnya menghasilkan *Vision Transformer* (ViT), yaitu metode yang mengadaptasi arsitektur *Transformer* untuk pengolahan citra. ViT menggunakan mekanisme self-attention untuk memahami hubungan antarbagian citra secara global sehingga mampu menangkap informasi visual secara lebih menyeluruh [5]. Kemampuan tersebut menjadikan ViT sebagai salah satu pendekatan yang banyak digunakan

dalam penelitian *computer vision modern* [6].

Penggunaan satu arsitektur saja terkadang belum optimal ketika menghadapi citra dengan karakteristik visual yang kompleks. Oleh karena itu, pendekatan *hybrid* model yang menggabungkan CNN dan ViT dapat digunakan untuk memanfaatkan keunggulan kedua metode tersebut. CNN dapat digunakan untuk mengekstraksi fitur lokal seperti tekstur dan pola, sedangkan ViT mampu menangkap hubungan global antarbagian citra [7]. Kombinasi tersebut diharapkan dapat menghasilkan representasi fitur yang lebih informatif untuk proses klasifikasi [8].

Selain pemilihan arsitektur, metode klasifikasi juga berpengaruh terhadap efisiensi sistem. *Broad Learning System* (BLS) merupakan metode pembelajaran yang memperluas jaringan secara horizontal sehingga proses pelatihan dapat dilakukan dengan lebih cepat dan efisien dibandingkan pendekatan deep learning yang memperdalam jaringan secara vertikal. Beberapa penelitian menunjukkan bahwa BLS memiliki kemampuan yang baik dalam berbagai tugas klasifikasi dan pengenalan pola [9].

Berdasarkan permasalahan tersebut, penelitian ini mengusulkan pengembangan sistem deteksi gambar AI multikelas menggunakan *hybrid* CNN–ViT yang dikombinasikan dengan BLS sebagai classifier [10]. Pendekatan ini memanfaatkan kemampuan CNN dalam mengekstraksi fitur lokal, ViT dalam memahami fitur global, serta BLS dalam melakukan klasifikasi secara efisien [11]. Sistem yang dikembangkan ditujukan untuk membedakan citra asli dan citra hasil generasi AI pada beberapa kategori objek.

Tujuan penelitian ini adalah mengembangkan sistem deteksi gambar AI multikelas yang mampu mengidentifikasi citra asli dan citra hasil generasi AI secara akurat. Sistem juga diimplementasikan dalam bentuk aplikasi berbasis *web* agar pengguna dapat melakukan proses deteksi citra secara praktis dan *real-time*. Dengan adanya sistem ini, diharapkan proses verifikasi keaslian citra digital dapat dilakukan dengan lebih mudah dan efisien.

METODE

Penelitian ini menggunakan pendekatan eksperimen untuk mengembangkan model deteksi gambar AI multikelas berbasis deep learning. Model dibangun menggunakan kombinasi beberapa metode pembelajaran mesin dan dievaluasi berdasarkan metrik yang telah ditentukan. Pendekatan ini digunakan untuk menguji performa model secara sistematis pada dataset yang digunakan.

Dataset Penelitian

Dataset yang digunakan dalam penelitian ini terdiri dari citra digital yang terbagi menjadi beberapa kategori, yaitu citra asli dan citra hasil generasi AI. Dataset digunakan untuk melatih dan menguji model deteksi gambar AI multikelas agar mampu mengenali karakteristik visual dari setiap kategori. Dataset kemudian dibagi menjadi data pelatihan (*training* data) yang digunakan untuk mempelajari pola visual citra dan data pengujian (*testing* data) yang digunakan untuk mengevaluasi kemampuan model dalam mengklasifikasikan citra yang belum digunakan selama proses pelatihan.

Preprocessing Data

Tahap preprocessing dilakukan untuk meningkatkan kualitas dan mempersiapkan data sebelum proses pelatihan model. Tahapan ini bertujuan agar seluruh citra memiliki format dan ukuran yang seragam sehingga dapat diproses dengan baik oleh model deep learning. Proses preprocessing meliputi resizing gambar, yaitu mengubah seluruh citra ke ukuran yang sesuai dengan input model CNN, kemudian dilakukan normalisasi dengan menyesuaikan nilai piksel ke dalam rentang tertentu untuk membantu mempercepat proses pelatihan. Selanjutnya, dataset dibagi menjadi data latih (training data) dan data uji (testing data) dengan perbandingan yang telah ditentukan untuk mengevaluasi performa model dan mengurangi risiko overfitting.

Model Hybrid CNN-Vision Transformer

Penelitian ini menggunakan model hybrid CNN–ViT, di mana CNN berfungsi mengekstraksi fitur lokal seperti tekstur, warna, dan pola visual, sedangkan ViT menangkap hubungan global antarbagian citra melalui mekanisme self-attention. Kombinasi keduanya menghasilkan representasi fitur yang lebih komprehensif untuk meningkatkan performa klasifikasi citra.

Broad Learning System

Broad Learning System (BLS) digunakan sebagai classifier untuk mengklasifikasikan fitur hasil ekstraksi CNN–ViT. BLS memperluas jaringan secara horizontal sehingga proses pelatihan dapat dilakukan lebih cepat dan efisien, serta mampu menangani fitur berdimensi tinggi.

Evaluasi Model

Evaluasi model dilakukan untuk mengetahui performa model klasifikasi citra yang telah dikembangkan. Pada penelitian ini digunakan beberapa metrik evaluasi yaitu *accuracy, precision, recall*, dan *F1-score* untuk mengukur tingkat keberhasilan model dalam mengklasifikasikan citra ke dalam kategori yang benar. Selain itu, digunakan juga *confusion matrix* untuk menganalisis kesalahan klasifikasi yang terjadi serta melihat distribusi prediksi model pada setiap kelas sehingga dapat memberikan gambaran yang lebih jelas mengenai performa model yang dihasilkan.

HASIL DAN PEMBAHASAN

Implementasi Model Hybrid CNN-ViT dan BLS Implementasi Aplikasi Web

Aplikasi *web* dikembangkan menggunakan *Streamlit* sebagai antarmuka pengguna dan *FastAPI* sebagai *backend* inferensi model. Pengguna dapat mengunggah gambar melalui halaman utama aplikasi, kemudian sistem akan memproses gambar tersebut dan menampilkan hasil klasifikasi secara real-time. Antarmuka aplikasi dirancang dalam satu halaman terpadu yang menampilkan deskripsi sistem, fitur unggah gambar, serta hasil klasifikasi berupa label prediksi dan nilai probabilitas atau tingkat kepercayaan model.



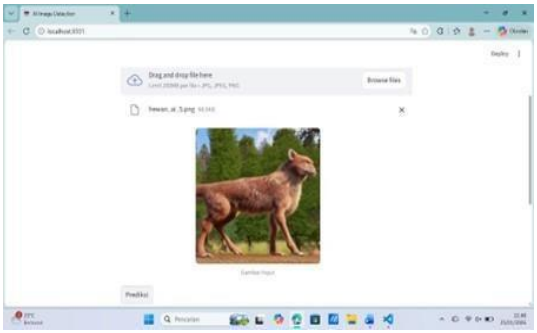
Gambar 2. Halaman Utama Aplikasi Web Deteksi Gambar AI Multikelas

Model yang dikembangkan dalam penelitian ini terdiri dari tiga komponen utama yaitu EfficientNet-B0 sebagai ekstraktor fitur lokal, Vision Transformer (ViT-Base) sebagai ekstraktor fitur global, dan Broad Learning System (BLS) sebagai classifier multikelas. EfficientNet-B0 digunakan untuk menangkap pola lokal seperti tekstur dan artefak pada gambar, sedangkan *ViT-Base* memahami hubungan global antarbagian gambar melalui mekanisme *self-attention*. Fitur dari kedua model kemudian digabungkan (*feature*

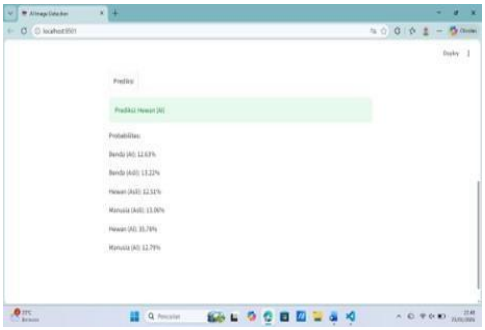
fusion) dan digunakan sebagai input bagi BLS untuk melakukan klasifikasi secara efisien, sehingga sistem mampu menghasilkan performa klasifikasi yang baik dengan waktu pelatihan yang relatif singkat.



Gambar 1. Arsitektur Model Hybrid CNN-ViT BLS



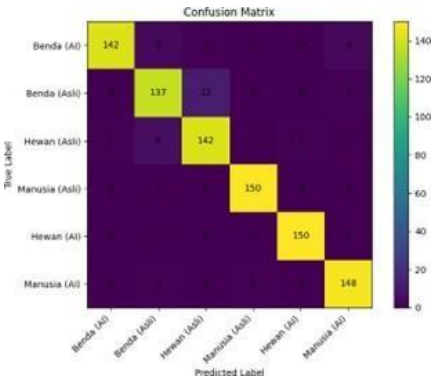
Gambar 3. Unggah Gambar ke Aplikasi Web



Gambar 4. Tampilan Hasil Deteksi Gambar

Hasil Confusion Matrix

Confusion matrix digunakan untuk mengevaluasi kinerja model klasifikasi dalam membedakan enam kelas citra, yaitu Benda (AI), Benda (Asli), Hewan (Asli), Manusia (Asli), Hewan (AI), dan Manusia (AI). *Matriks* ini menunjukkan hubungan antara label aktual (*true* label) dan label prediksi *predicted label*, sehingga dapat dianalisis tingkat ketepatan serta pola kesalahan klasifikasi yang terjadi



Gambar 5. Hasil Confusion Matrix

Gambar tersebut menunjukkan confusion matrix yang menggambarkan kinerja model dalam mengklasifikasikan data. Sebagian besar nilai berada pada diagonal yang menandakan prediksi model sudah benar, dengan hanya sedikit kesalahan klasifikasi. Hal ini menunjukkan bahwa model memiliki performa yang baik.

Hasil Evaluasi Matrix Kinerja

Evaluasi kinerja model dilakukan menggunakan metrik akurasi, *precision*, *recall*, dan *F1-score* yang diperoleh dari classification report. Pengujian dilakukan pada data validasi yang terdiri dari 900 citra dengan 6 kelas, masing-masing kelas memiliki 150 sampel.

```
=== HASIL EVALUASI BLS ===
Akurasi Validasi : 0.9655555555555555

Classification Report:
              precision    recall  f1-score   support

           0       0.99      0.95      0.97       150
           1       0.93      0.91      0.92       150
           2       0.92      0.95      0.93       150
           3       0.99      1.00      1.00       150
           4       0.99      1.00      1.00       150
           5       0.97      0.99      0.98       150

    accuracy                           0.97       900
   macro avg       0.97      0.97      0.97       900
weighted avg       0.97      0.97      0.97       900
```

Gambar 6. Hasil Classification Report

Gambar tersebut menampilkan hasil evaluasi model BLS dengan akurasi validasi sebesar 0,9666 atau sekitar 96,66%. Laporan klasifikasi menunjukkan nilai *precision*, *recall*, dan *f1-score* yang tinggi pada setiap kelas sekitar 0,92–1,00, sehingga dapat disimpulkan bahwa model memiliki kinerja yang sangat baik dalam mengklasifikasikan data.

Pengujian Fungsional (Black-Box Testing)

Pengujian fungsional sistem web

dilakukan menggunakan metode *black-box testing* untuk memastikan seluruh fitur berjalan sesuai kebutuhan. Pengujian meliputi proses unggah gambar, inferensi model, dan penampilan hasil klasifikasi. Hasilnya menunjukkan bahwa sistem dapat menerima file gambar, memprosesnya melalui API *FastAPI*, serta menampilkan hasil klasifikasi pada antarmuka *Streamlit* tanpa *error*.

Analisis Respons dan Stabilitas Sistem

Hasil pengujian menunjukkan bahwa sistem memiliki waktu respons yang cepat dalam memproses gambar sehingga cocok untuk penggunaan real-time berbasis *web*. Penggunaan BLS sebagai *classifier* meningkatkan efisiensi komputasi, dan sistem juga menunjukkan stabilitas yang baik saat digunakan dengan berbagai jenis citra, menandakan integrasi model CNN–ViT–BLS dengan *FastAPI–Streamlit* berjalan dengan baik.

Analisis Performa Model CNN ViT BLS

Hasil pengujian menunjukkan bahwa arsitektur *hybrid CNN–ViT* mampu meningkatkan performa deteksi gambar AI multikelas, di mana CNN menangkap fitur lokal dan ViT memahami konteks global gambar. Fitur yang digabungkan dan diklasifikasikan menggunakan Broad Learning System (BLS) menghasilkan klasifikasi yang lebih akurat dan efisien dibandingkan model tunggal, serta mudah diintegrasikan ke dalam aplikasi *web*.

Analisis Kesalahan Klasifikasi

Hasil pengujian menggunakan *confusion matrix* menunjukkan masih terdapat beberapa kesalahan klasifikasi pada kelas yang memiliki karakteristik visual serupa, seperti antara gambar asli dan gambar AI dengan tingkat realisme tinggi. Kesalahan ini juga dipengaruhi oleh variasi pencahayaan, sudut pengambilan gambar, dan kompleksitas latar belakang. Namun, jumlah kesalahan relatif kecil sehingga tidak berdampak signifikan terhadap performa keseluruhan sistem.

Implikasi dan Keterbatasan Sistem

Hasil penelitian menunjukkan bahwa sistem deteksi gambar AI multikelas berbasis *hybrid CNN–ViT–BLS* berpotensi digunakan untuk mengidentifikasi citra hasil generasi AI secara otomatis, seperti pada verifikasi konten digital dan penyaringan media visual. Namun, penelitian ini masih memiliki keterbatasan pada variasi dataset dan pengujian sistem yang belum dilakukan pada skala pengguna besar, sehingga diperlukan pengembangan lebih lanjut untuk meningkatkan performa dan keandalan sistem.

SIMPULAN

Berdasarkan hasil perancangan, implementasi, dan pengujian, sistem deteksi gambar AI multikelas berbasis *hybrid CNN–ViT–BLS* berhasil dikembangkan dengan baik. Kombinasi *EfficientNet-B0* sebagai ekstraktor fitur lokal dan *Vision Transformer* sebagai ekstraktor fitur global mampu menghasilkan representasi fitur yang efektif dalam membedakan citra asli dan citra AI. Hasil evaluasi menunjukkan tingkat akurasi yang tinggi dengan kesalahan klasifikasi yang minimal pada enam kelas citra. Selain itu, sistem *web* yang dibangun menggunakan *Streamlit* dan *FastAPI* dapat berjalan stabil dan menampilkan hasil klasifikasi secara *real-time*, sehingga tujuan penelitian untuk membangun sistem deteksi gambar AI yang akurat dan efisien telah tercapai.

DAFTAR PUSTAKA

[1] dan R. H. S. Wahyuni, A. Prasetyo, “Klasifikasi penyakit daun tanaman menggunakan metode Convolutional Neural Network berbasis citra digital,” *urnal RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 7, pp. 654–661, 2023.

[2] dan Y. S. M. F. Ramadhan, D. Kurniawan, “Penerapan Deep Learning untuk deteksi penyakit daun

tanaman menggunakan arsitektur CNN," *urnal Teknol. Inf. dan Ilmu Komput.*, vol. 11, pp. 125–134, 2024.

[3] R. Setiawan dan D. Prabowo, "Implementasi Convolutional Neural Network untuk klasifikasi citra digital," *J. Inform. Mulawarman*, vol. 18, pp. 45–52, 2023.

[4] P. H. U. Octavia Devi Safitri Sunanto, "Implementasi Deep Learning Dengan Convolutional Neural Network Untuk Klasifikasi Gambar Sampah Organik Dan Anorganik," *Pattimura Proceeding Conf. Sci. Technol.*, vol. 1, no. 2, pp. 335–340, 2022.

[5] N. Sari dan B. Nugroho, "Analisis performa Vision Transformer pada klasifikasi citra digital," *J. Ilm. Inform.*, vol. 11, pp. 25–33, 2024.

[6] T. U. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, H. Zhai, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," *Int. Conf. Learn. Represent.*, 2020, doi: 10.48550/arXiv.2010.11929.

[7] dan N. R. A. Nugraha, R. Saputra, "Perbandingan metode ensemble learning pada klasifikasi citra penyakit tanaman," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 12, pp. 210–218, 2023.

[8] A. P. dan R. Setiawan, "Klasifikasi citra menggunakan metode deep learning berbasis computer vision," *J. Inform. Indones.*, vol. 9, pp. 120–128, 2024.

[9] S. Rahman dan A. Wijaya, "Analisis metode machine learning pada pengolahan citra digital," *J. Inform. dan Sist. Inf.*, vol. 9, pp. 55–63, 2022.

[10] dan M. R. N. Zufar Faiil Haq, Mufti Ari Bianto, Afifah Agustin, "Implementasi Convolutional Neural Network (CNN) untuk Klasifikasi Citra Batik Nusantara," *J. Tek. Inform. dan Teknol. Inf.*, 2024, doi:https://doi.org/10.55606/jutiti.v5i1.5421.

[11] Dede Wulan Rahayu, David Setiadi, dan Dwi Yuniarto, "Penerapan Deep Learning untuk Klasifikasi Buah Berdasarkan Citra dengan Metode Convolutional Neural Networks," *Indones. J. Educ. Comput. Sci.*, 2023, doi: https://doi.org/10.60076/indotech.v2i3.9 78