

ANALISIS KLASIFIKASI FITUR RGB-GRAyscale CIFAR-10 PADA ALGORITMA CONVOLUTIONAL NEURAL NETWORK DAN SUPPORT VECTOR MACHINE

Bayu Adhi Nugroho¹, Umar Abdul Azis²

^{1,2} UIN Sunan Ampel Surabaya

bayu.lecture@gmail.com

ABSTRAK

Perkembangan teknologi kecerdasan buatan telah mendorong peningkatan signifikan dalam akurasi sistem pengenalan objek visual otomatis. Penelitian ini bertujuan untuk melakukan analisis mendalam mengenai performa klasifikasi fitur pada dataset gambar CIFAR-10 menggunakan dua pendekatan algoritma yang berbeda, yaitu *Convolutional Neural Network* (CNN) berbasis deep learning dan *Support Vector Machine* (SVM) berbasis machine learning konvensional. Dataset CIFAR-10 yang mencakup 60.000 citra berwarna dalam 10 kelas objek berbeda menjadi tolok ukur utama dalam pengujian ini. Metodologi penelitian diawali dengan tahap pra-pemrosesan citra, ekstraksi fitur, diikuti oleh pelatihan model dan evaluasi performa menggunakan matriks akurasi, precision, recall, dan F1-score. Perkembangan CNN sebagai algoritma klasifikasi diikuti oleh tuntutan sumber daya komputasi yang tinggi. Salah satu mekanisme untuk menekan kebutuhan sumber daya komputasi adalah menggunakan hasil ekstraksi fitur CNN dari citra dalam proses klasifikasi. Penelitian ini bertujuan untuk melakukan analisis kuantitatif kinerja klasifikasi antara algoritma CNN dan algoritma Support Vector Machine (SVM). Kontribusi utama dari penelitian ini adalah penyajian perbandingan komparatif yang komprehensif mengenai efisiensi komputasi dan akurasi antara model *CNN end-to-end* dan klasifikasi tradisional SVM. Kesimpulan hasil uji kinerja kedua algoritma dengan menggunakan fitur ekstraksi RGB dan *grayscale* adalah bahwa keduanya sangat kompetitif.

Kata kunci : Klasifikasi, Fitur, CNN, SVM, CIFAR-10

ABSTRACT

Advancements in artificial intelligence technology have driven significant improvements in the accuracy of automated visual object recognition systems. This study aims to conduct an in-depth analysis of feature classification performance on the CIFAR-10 image dataset using two distinct algorithmic approaches: deep learning-based Convolutional Neural Networks (CNNs) and conventional machine learning-based Support Vector Machines (SVMs). The CIFAR-10 dataset, comprising 60,000 colour images across 10 distinct object classes, serves as the primary benchmark for this evaluation. The research methodology encompasses image preprocessing and feature extraction, followed by model training and performance assessment using accuracy, precision, recall, and F1-score metrics. While CNNs are powerful classification algorithms, they require substantial computational resources. One strategy to mitigate these computational requirements is to utilise image features extracted by CNNs for classification. This study aims to quantitatively analyse the classification performance of CNN and SVM algorithms. Its primary contribution is a comprehensive comparison of computational efficiency and accuracy between end-to-end CNN models and traditional SVM classification. Performance testing using both RGB and grayscale extracted features reveals that the two algorithms are highly competitive.

Keywords: Classification, Features, Cnn, Svm, Cifar-10

Pendahuluan

Klasifikasi citra merupakan salah satu cabang ilmu komputer yang sangat penting dan terus berkembang pesat. Salah satu dataset yang sering dijadikan standar (benchmark) dalam pengenalan objek adalah CIFAR-10 [1] yang terdiri dari citra berwarna berukuran 32×32 piksel dengan 10 kelas berbeda. Secara umum, citra pada CIFAR-10 memiliki representasi warna *Red-Green-Blue (RGB)* (3 kanal). Meskipun memberikan informasi yang kaya, citra berwarna membutuhkan beban komputasi dan memori yang sangat besar dalam proses ekstraksi fitur dan pelatihan model. Oleh karena itu, teknik reduksi dimensi seperti mengubah citra RGB menjadi *grayscale* (satu kanal) [2] sering kali diterapkan untuk merampingkan data input, mempercepat waktu komputasi, dan mengurangi kompleksitas komputasi. Namun, konversi ini berisiko menghilangkan detail informasi warna [3] yang mungkin penting untuk membedakan antarobjek.

Dalam algoritma klasifikasi citra/gambar, terdapat dua pendekatan yang cukup menonjol kebutuhan untuk mengevaluasi secara mendalam pengaruh perbedaan representasi fitur warna (RGB vs Grayscale) [9] terhadap kinerja klasifikasi kedua algoritma tersebut pada dataset CIFAR-10. Penelitian ini dilakukan untuk menjawab tantangan optimalisasi pengenalan objek, menyeimbangkan antara tingkat keakuratan dan efisiensi komputasi, serta melihat sejauh mana hilangnya informasi warna memengaruhi keputusan akhir klasifikasi pada CNN dan SVM.

Rumusan masalah yang hendak dijawab oleh penelitian ini adalah: pertama, bagaimana perbandingan performa klasifikasi antara penggunaan citra asli dari dataset CIFAR-10 dengan CNN dan penggunaan fitur RGB serta fitur grayscale pada SVM? Kedua, apakah reduksi dimensi warna (dari RGB ke *grayscale*) menyebabkan penurunan akurasi yang signifikan pada model CNN dan SVM? Ketiga, apakah penggabungan fitur RGB dan Grayscale dapat meningkatkan kinerja klasifikasi? Tujuan akhir dari penelitian ini

dan memiliki karakteristik yang berbeda secara mendasar. Pertama adalah *Convolutional Neural Network (CNN)* [4]: CNN merupakan algoritma Deep Learning yang sangat dominan dalam pengolahan data spasial seperti citra. CNN memiliki keunggulan berupa kemampuan ekstraksi fitur [5] secara otomatis (tanpa rekayasa fitur manual) melalui lapisan konvolusi. Walaupun akurasinya sangat tinggi, arsitektur CNN memerlukan sumber daya komputasi yang masif serta waktu pelatihan yang lama [6]. Kedua adalah *Support Vector Machine (SVM)* [7]: Berbeda dengan CNN, SVM adalah algoritma *Machine Learning* konvensional yang terkenal efektif untuk klasifikasi data berdimensi tinggi dengan mencari *hyperplane* terbaik sebagai pemisah antarkelas [8]. SVM tidak mengekstraksi fitur secara otomatis, sehingga membutuhkan ekstraksi fitur manual dari algoritma CNN. SVM lebih hemat komputasi dibandingkan CNN, namun akurasinya sangat bergantung pada metode ekstraksi fitur yang digunakan sebelumnya.

Berdasarkan uraian tersebut, terdapat adalah: pertama, menganalisis kinerja klasifikasi algoritma CNN dan SVM pada dataset CIFAR-10. Kedua, mengevaluasi pengaruh penggunaan citra asli dibandingkan dengan fitur RGB dan Grayscale terhadap performa klasifikasi citra. Ketiga, mengetahui performa penggunaan fitur gabungan antara citra grayscale dan RGB pada model *machine learning* konvensional (SVM).

Materi dan Metode

Metode penelitian dirancang melalui pendekatan komparatif eksperimental untuk menguji kinerja klasifikasi citra. Berikut adalah kerangka metodologi yang digunakan. Tahap ini bertujuan untuk merekonstruksi seluruh varian data yang diperlukan guna melakukan evaluasi dalam satu lingkungan yang terkontrol. Proses persiapan data dibagi menjadi tiga kategori utama:

1. Rekonstruksi Data Citra Uji (*Image Data*): Dataset mentah CIFAR-10 dimuat ulang dan diproses menjadi tiga format

representasi visual: RGB, Grayscale Average, dan Grayscale Luma. Langkah ini diperlukan untuk menguji kembali kinerja model CNN asli (*End-to-End Deep Learning*) sebagai baseline pembandingan.

Konversi citra RGB dari dataset CIFAR-10 ke grayscale menggunakan dua algoritma yang berbeda, yaitu Average [10] dan Luma [10].

- Grayscale Average (Metode Rata-Rata) mengambil nilai dari ketiga kanal warna RGB (merah, hijau, biru) dan membaginya secara merata. $\text{Grayscale} = (R+G+B)/3$
- Grayscale Luma (Metode Luminositas/Persepsi) menyesuaikan konversi dengan sensitivitas mata manusia yang paling peka terhadap warna hijau dan kurang peka terhadap warna biru. $\text{Grayscale} = (0.299 \times R) + (0.587 \times G) + (0.114 \times B)$

Dataset CIFAR-10 dalam bentuk citra RGB, citra Grayscale Average, dan citra Grayscale Luma digunakan sebagai input oleh 3 (tiga) CNN, yaitu: CNN RGB, CNN Gray. Average dan CNN Gray. Luma dalam bentuk citra asli kemudian diekstraksi fitur yang sesuai pada masing-masing arsitektur CNN.

- Pemuatan Vektor Fitur (*Feature Vectors*):** Berkas fitur yang telah diekstraksi dan disimpan pada eksperimen sebelumnya dimuat kembali. Ini mencakup fitur dari *domain* RGB, Grayscale Average, dan Grayscale Luma, untuk membentuk dataset bagi model SVM RGB, SVM Gray. AVG dan SVM Gray. Luma.
- Konstruksi Fitur Gabungan (*Fusion Features*):** Vektor-vektor fitur tunggal digabungkan kembali (*concatenated*) untuk membentuk dataset bagi model SVM kombinasi. Merupakan gabungan

fitur dari domain RGB, Grayscale Average, dan Grayscale Luma.

Metrik Evaluasi yang digunakan untuk Mengukur dan menganalisis performa tiap model menggunakan parameter berikut:

- Akurasi (*accuracy*), rasio prediksi benar dibandingkan dengan keseluruhan data.
- Presisi (*precision*) mengukur ketepatan model dalam memprediksi kelas positif dari seluruh data yang diprediksi positif.
- Recall (sensitivitas) mengukur kemampuan model untuk menemukan kembali seluruh data positif dari total data yang sebenarnya positif. Parameter ini sangat penting untuk melihat apakah ada objek yang terlewat oleh model.
- Skor F1 (*F1-score*): Rata-rata harmonik dari presisi dan recall. Metrik ini memberikan satu nilai tunggal yang menyeimbangkan ketepatan (presisi) dan sensitivitas (recall) model.

Lapisan arsitektur CNN untuk ekstraksi fitur RGB adalah sebagaimana ditunjukkan pada Gambar 1 dan dilatih sebanyak 10 epoch. Sedangkan untuk fitur Grayscale, ubah input layer dari 3 (tiga) dimensi menjadi 1 (satu) dimensi.

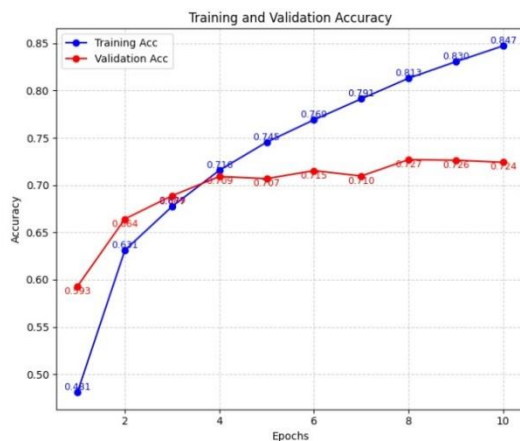
Layer (type)	Output Shape	Param #
input_layer_1 (InputLayer)	(None, 32, 32, 3)	0
conv2d_2 (Conv2D)	(None, 32, 32, 64)	1,792
max_pooling2d_2 (MaxPooling2D)	(None, 16, 16, 64)	0
conv2d_3 (Conv2D)	(None, 16, 16, 96)	55,392
max_pooling2d_3 (MaxPooling2D)	(None, 8, 8, 96)	0
flatten_1 (Flatten)	(None, 6144)	0
layer_fitur_rgb (Dense)	(None, 192)	1,179,840
dropout_1 (Dropout)	(None, 192)	0
dense_1 (Dense)	(None, 10)	1,930

Gambar 1. Lapisan Arsitektur CNN RGB

Hasil dan Pembahasan

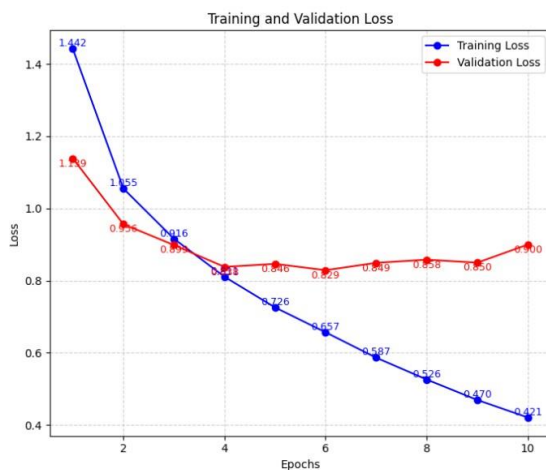
Proses pelatihan CNN RGB menunjukkan *overfitting* mulai *epoch* ke-4, sebagaimana

diperlihatkan pada Gambar 2 yang memuat akurasi selama proses pelatihan. Gambar 3 yang menunjukkan selisih nilai *loss* yang besar setelah epoch 4 mengindikasikan saturasi pada proses pelatihan.



Gambar 2. Akurasi Proses Pelatihan CNN RGB

Model SVM yang digunakan adalah SVM dengan kernel *Radial Basis Function* (RBF). Pemilihan tipe kernel SVM ini didasarkan pada teori bahwa kernel RBF mampu membentuk *decision boundary* nonlinier. SVM yang dikombinasikan dengan fitur gabungan dari RGB dan grayscale AVG serta grayscale Luma menunjukkan kinerja yang lebih baik dibandingkan dengan model lain.



Gambar 3. Fungsi *Loss* Pelatihan CNN RGB

Tabel 1. Hasil Perbandingan Komparatif Model

Model	Akurasi	Presisi	Recall	F1-Score
CNN RGB	72.41%	72.72%	72.41%	72.19%
CNN Gray. AVG	69.77%	69.87%	69.77%	69.68%
CNN Gray. Luma	67.94%	68.60%	67.94%	67.94%
SVM RGB	73.05%	73.13%	73.05%	73.08%
SVM Gray. AVG	69.63%	69.74%	69.63%	69.67%
SVM Gray. Luma	69.54%	69.57%	69.54%	69.54%
SVM Kombinasi	76.21%	76.38%	76.21%	76.27%

Berbagai teknik penggabungan atau penggunaan fitur CNN telah digunakan dalam kasus dan dataset lain [11], [12], [13]. Dalam eksperimen penelitian ini digunakan kombinasi dua macam fitur grayscale yang digabungkan dengan fitur RGB dari CNN tunggal. Eksperimen menunjukkan bahwa penggunaan gabungan fitur untuk klasifikasi menggunakan SVM menghasilkan hasil yang lebih baik daripada klasifikasi menggunakan citra asli.

Kesimpulan

Kesimpulan yang dapat dibuat adalah, pertama: performa klasifikasi antara penggunaan citra asli dari dataset CIFAR-10 dengan CNN dibandingkan dengan penggunaan fitur RGB dan fitur grayscale pada SVM menunjukkan kinerja maksimal ketika digunakan fitur gabungan pada SVM. Kedua, reduksi dimensi warna (dari RGB ke grayscale) menurunkan kinerja klasifikasi, tetapi penggunaan citra asli bukan variabel penentu kinerja klasifikasi maksimal. Ketiga, penggunaan fitur gabungan antara fitur Grayscale dan RGB memberikan kinerja maksimal pada eksperimen ini dengan algoritma SVM kernel RBF.

Daftar Pustaka

- [1] N. Hernández, R. Morales, and L. A. Pineda, 'Remembering CIFAR-10 images with the entropic associative memory', *Pattern Recognit.*, vol. 172, p.

- 112639, 2026, doi: <https://doi.org/10.1016/j.patcog.2025.112639>.
- [2] Z. N. Khudhair *et al.*, 'Color to Grayscale Image Conversion Based on Singular Value Decomposition', *IEEE Access*, vol. 11, pp. 54629–54638, 2023, doi: 10.1109/ACCESS.2023.3279734.
- [3] Q. Liang and S. Xiang, 'Invertible Color-to-Grayscale Conversion Using Lossy Compression and High-Capacity Data Hiding', *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 11, pp. 7373–7385, 2022, doi: 10.1109/TCSVT.2022.3184949.
- [4] P. Purwono, A. Ma'arif, W. Rahmانيar, H. I. K. Fathurrahman, A. Z. K. Frisky, and Q. M. ul Haq, 'Understanding of Convolutional Neural Network (CNN): A Review', *Int. J. Robot. Control Syst.*, vol. 2, no. 4, pp. 739–748, 2023, doi: 10.31763/ijrcs.v2i4.888.
- [5] A. A. Barbhuiya, R. K. Karsh, and R. Jain, 'CNN based feature extraction and classification for sign language', *Multimed. Tools Appl.*, vol. 80, no. 2, pp. 3051–3069, Jan. 2021, doi: 10.1007/s11042-020-09829-y.
- [6] L. Ning, H. Guan, and X. Shen, 'Adaptive Deep Reuse: Accelerating CNN Training on the Fly', in *2019 IEEE 35th International Conference on Data Engineering (ICDE)*, 2019, pp. 1538–1549. doi: 10.1109/ICDE.2019.00138.
- [7] S. Suthaharan, 'Support Vector Machine', in *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning*, Boston, MA: Springer US, 2016, pp. 207–235. doi: 10.1007/978-1-4899-7641-3_9.
- [8] P. Birzhandi, K. T. Kim, B. Lee, and H. Y. Youn, 'Reduction of Training Data Using Parallel Hyperplane for Support Vector Machine', *Appl. Artif. Intell.*, vol. 33, no. 6, pp. 497–516, 2019, doi: 10.1080/08839514.2019.1583449.
- [9] N. Srisermphoak, P. Amornphimoltham, R. Chaisuparat, P. Achararit, and T. Fuangrod, 'A systematic evaluation of grayscale conversion methods for mitigating color variation in deep learning-based histopathological image analysis', *J. Pathol. Inform.*, vol. 21, p. 100647, Apr. 2026, doi: 10.1016/j.jpi.2026.100647.
- [10] R. M. H. Nguyen and M. S. Brown, 'Why You Should Forget Luminance Conversion and Do Something Better', in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017.
- [11] W. Gong *et al.*, 'A Novel Deep Learning Method for Intelligent Fault Diagnosis of Rotating Machinery Based on Improved CNN-SVM and Multichannel Data Fusion', *Sensors*, vol. 19, no. 7, 2019, doi: 10.3390/s19071693.
- [12] R. Kuttala, R. Subramanian, and V. R. M. Oruganti, 'Multimodal Hierarchical CNN Feature Fusion for Stress Detection', *IEEE Access*, vol. 11, pp. 6867–6878, 2023, doi: 10.1109/ACCESS.2023.3237545.
- [13] A. D. Jia, B. Z. Li, and C. C. Zhang, 'Detection of cervical cancer cells based on strong feature CNN-SVM network', *Neurocomputing*, vol. 411, pp. 112–127, 2020, doi: <https://doi.org/10.1016/j.neucom.2020.06.006>.