



Analisis Pola Fitur Statistik Citra Digital untuk Klasifikasi Uang Kertas Menggunakan Metode K-Nearest Neighbor

Statistical Feature Pattern Analysis of Digital Images for Banknote Classification Using the K-Nearest Neighbor Method

Chandra Pratama Putra Raharja^{1*}, Hasbi Firmansyah², Wahyu Asriyani³, Rizki Prasetyo Tulodo⁴

Universitas Pancasakti Tegal

Email: sapaente30@gmail.com^{1*}, hasbifirmansyah@upstegal.ac.id²,
asriyani1409@gmail.com³, Rizki.prasetyo.tulodo@gmail.com⁴

Article Info

Article history :

Received : 30-12-2025

Revised : 01-01-2026

Accepted : 03-01-2026

Pulished : 05-01-2026

Abstract

Currency counterfeiting is a global issue significantly impacting economic stability. Manual detection of banknote authenticity is often inefficient and prone to human error, necessitating a precise automated detection system. This study aims to evaluate the performance of the K-Nearest Neighbor (K-NN) algorithm in classifying banknote authenticity based on statistical features extracted from digital images. The dataset used originates from the UCI Machine Learning Repository, consisting of 1,372 sample data with four main feature attributes: Variance, Skewness, Kurtosis, and Entropy generated from Wavelet Transformation. Testing was conducted using the 10-fold Cross Validation method to ensure model consistency. Experimental results demonstrated extremely high performance with an accuracy rate reaching **99.93%**. Confusion matrix analysis revealed that the model encountered classification error in only 1 sample out of the total test data, with a Recall value for the genuine currency class reaching 100%. This study concludes that statistical features extracted from banknote images possess robust data patterns, and the K-NN algorithm is highly effective for application in counterfeit currency detection forensic systems.

Keywords: *K-Nearest Neighbor, Banknote Classification, Statistical Features*

Abstrak

Pemalsuan mata uang merupakan masalah global yang berdampak signifikan terhadap stabilitas ekonomi. Deteksi keaslian uang kertas secara manual seringkali tidak efisien dan rentan terhadap kesalahan manusia, sehingga diperlukan sistem deteksi otomatis yang presisi. Penelitian ini bertujuan untuk mengevaluasi performa algoritma *K-Nearest Neighbor* (K-NN) dalam mengklasifikasikan keaslian uang kertas berdasarkan ekstraksi fitur statistik citra digital. Dataset yang digunakan bersumber dari *UCI Machine Learning Repository* yang terdiri dari 1.372 sampel data dengan empat atribut fitur utama: *Variance*, *Skewness*, *Kurtosis*, dan *Entropy* yang dihasilkan dari Transformasi Wavelet. Pengujian dilakukan menggunakan metode validasi silang 10-lipatan (*10-fold Cross Validation*) untuk memastikan konsistensi model. Hasil eksperimen menunjukkan performa yang sangat tinggi dengan tingkat akurasi mencapai **99,93%**. Analisis *confusion matrix* memperlihatkan bahwa model hanya mengalami kesalahan klasifikasi pada 1 data sampel dari total keseluruhan data uji, dengan nilai *Recall* untuk kelas uang asli mencapai 100%. Penelitian ini menyimpulkan bahwa fitur statistik yang diekstraksi dari citra uang kertas memiliki pola data



yang sangat kuat dan algoritma K-NN sangat efektif untuk diterapkan dalam sistem forensik deteksi uang palsu.

Kata Kunci: K-Nearest Neighbor, Klasifikasi Uang Kertas, Fitur Statistik

PENDAHULUAN

Integritas mata uang fisik memegang peranan vital dalam menjaga stabilitas moneter suatu negara. Peredaran uang palsu tidak hanya merugikan masyarakat secara individu, tetapi juga berpotensi mendestabilisasi ekonomi makro dengan memicu inflasi yang tidak terkendali. Meskipun fitur keamanan pada uang kertas modern terus ditingkatkan—mulai dari penggunaan *watermark*, benang pengaman, hingga tinta optik—para pemalsu uang juga terus mengadaptasi teknik mereka, sehingga perbedaan antara uang asli dan palsu secara kasat mata menjadi semakin sulit dan rentan terhadap kesalahan manusia (*human error*). Oleh karena itu, kebutuhan akan sistem otentikasi otomatis yang cepat, objektif, dan berakurasi tinggi menjadi tak terelakkan, terutama bagi institusi perbankan dan sektor ritel. (Mo et al., 2019)

Pendekatan komputasi untuk mendeteksi keaslian uang kertas umumnya melibatkan pemrosesan citra digital. Namun, mengolah data piksel mentah (*raw image*) seringkali memakan sumber daya komputasi yang besar. Solusi yang lebih efisien adalah dengan melakukan ekstraksi fitur menggunakan teknik matematika seperti Transformasi Wavelet (*Wavelet Transform*). Teknik ini mampu membedah struktur citra uang kertas menjadi komponen-komponen statistik yang lebih sederhana namun representatif, yaitu *Variance* (varians), *Skewness* (kemencengan), *Curtosis* (kurtosis), dan *Entropy* (entropi). Keempat atribut ini merepresentasikan pola distribusi intensitas piksel yang unik pada uang asli yang sulit direplikasi secara sempurna pada uang palsu. (Nabuasa et al., 2019)

Dalam ranah *data mining* dan pembelajaran mesin (*machine learning*), pemilihan algoritma klasifikasi sangat menentukan keberhasilan sistem deteksi. Salah satu metode yang dikenal efektif untuk pengenalan pola berbasis fitur statistik adalah *K-Nearest Neighbor* (K-NN). Berbeda dengan model kompleks seperti *Neural Networks* yang membutuhkan waktu pelatihan lama, K-NN merupakan algoritma *lazy learner* yang mengklasifikasikan data baru berdasarkan kemiripan jarak (*distance similarity*) terhadap data tetangga terdekatnya. Kesederhanaan K-NN justru menjadi keunggulan dalam menangani data numerik yang memiliki pemisahan kelas yang jelas, namun performanya sangat bergantung pada kualitas fitur data yang digunakan. (Science, 2024)

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis dan mengevaluasi performa algoritma K-NN dalam mengklasifikasikan keaslian uang kertas. Dataset yang digunakan bersumber dari repositori publik UCI Machine Learning yang telah melalui proses ekstraksi fitur Wavelet Transform. Penelitian ini secara spesifik berfokus pada pengujian akurasi model menggunakan metode validasi silang 10-lipatan (*10-fold Cross Validation*) untuk memastikan konsistensi dan keandalan sistem dalam membedakan pola fitur statistik antara uang asli dan uang palsu. Hasil dari penelitian ini diharapkan dapat memberikan kontribusi empiris



mengenai efektivitas penggunaan fitur statistik sederhana yang dipadukan dengan algoritma jarak untuk keperluan forensik digital. (*Resolution_preserving.Pdf*, n.d.).

METODE PENELITIAN

1. Sumber Data dan Karakteristik Fitur

Data yang digunakan dalam penelitian ini adalah dataset *Banknote Authentication* yang diperoleh dari repositori publik *UCI Machine Learning*. Dataset ini terdiri dari total 1.372 instans data yang merepresentasikan citra uang kertas. Tidak ada data yang hilang (*missing values*) dalam dataset ini, sehingga tidak diperlukan imputasi data.

Setiap instans data memiliki lima atribut, yang terdiri dari empat fitur numerik hasil ekstraksi *Wavelet Transform* dan satu label kelas. Rincian atribut tersebut adalah sebagai berikut:

- a. **Variance (Varians):** Mengukur sebaran distribusi intensitas piksel pada citra.
- b. **Skewness (Kemencengan):** Mengukur ketidaksimetrisan distribusi data piksel.
- c. **Curtosis (Kurtosis):** Mengukur keruncingan distribusi data dibandingkan dengan distribusi normal.
- d. **Entropy (Entropi):** Mengukur ketidakaturan atau kompleksitas tekstur pada citra.
- e. **Class (Kelas):** Variabel target biner, di mana nilai '0' merepresentasikan uang palsu dan nilai '1' merepresentasikan uang asli.

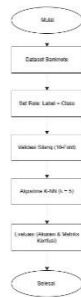
2. Alur Eksperimen

Penelitian ini dilakukan dengan menggunakan pendekatan eksperimental berbasis *data mining*. Proses komputasi dijalankan menggunakan perangkat lunak RapidMiner Studio. Tahapan penelitian dirancang sebagai berikut:

- a. **Pra-pemrosesan Data (*Data Preprocessing*):** Tahap awal melibatkan pendefinisian peran atribut (*Set Role*). Atribut 'Class' ditetapkan sebagai label (target prediksi), sementara keempat atribut fitur statistik ditetapkan sebagai variabel independen yang akan diolah oleh algoritma.
- b. **Pemodelan K-Nearest Neighbor (K-NN):** Algoritma K-NN diterapkan untuk mempelajari pola fitur. Pada penelitian ini, parameter jumlah tetangga terdekat ditetapkan pada $k=5$ (nilai *default* yang umum untuk menghindari *overfitting*). Perhitungan jarak antar data menggunakan metode **Jarak Euclidean** (*Euclidean Distance*) karena seluruh fitur bertipe numerik (real).
- c. **Validasi Model:** Untuk menghindari bias pada hasil pengujian, penelitian ini menerapkan metode **Validasi Silang 10-Lipatan** (*10-Fold Cross Validation*). Dataset dibagi menjadi 10 bagian (*folds*) acak yang saling lepas; 9 bagian digunakan untuk pelatihan (*training*)



dan 1 bagian untuk pengujian (*testing*). Proses ini diulang sebanyak 10 kali sehingga setiap data mendapat kesempatan yang sama untuk menjadi data latih maupun data uji.



Gambar 1. Diagram Alir Tahapan Penelitian.

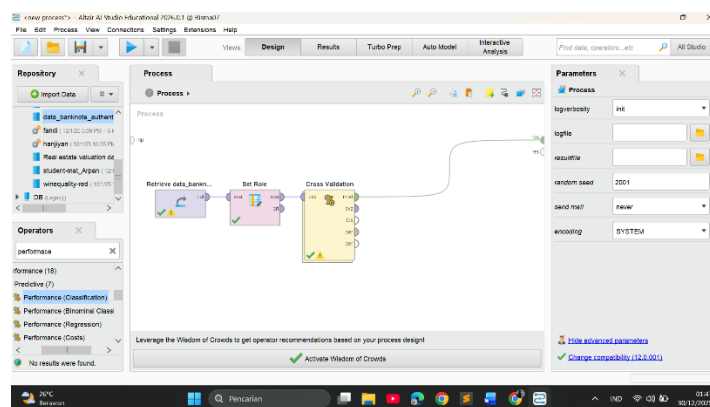
3. Evaluasi Performa

Kinerja model klasifikasi dievaluasi menggunakan *Confusion Matrix* untuk memetakan prediksi benar dan salah. Metrik utama yang menjadi acuan keberhasilan penelitian ini adalah:

- Akurasi (Accuracy):** Persentase total prediksi yang benar (baik asli maupun palsu) terhadap keseluruhan data.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$

- Recall (Sensitivitas):** Kemampuan sistem dalam mengenali kelas positif (uang asli) dengan tepat, untuk meminimalkan risiko uang asli terdeteksi sebagai palsu.



Gambar 2. Desain Proses Eksperimen Menggunakan RapidMiner.



HASIL DAN PEMBAHASAN

1. Implementasi dan Pengujian

Eksperimen dilakukan menggunakan perangkat lunak RapidMiner Studio dengan menerapkan validasi silang 10-lipatan (*10-fold Cross Validation*) pada dataset *Banknote Authentication*. Model K-Nearest Neighbor (K-NN) dikonfigurasi menggunakan nilai parameter $k = 5$ dan ukuran jarak *Euclidean*. Total data yang diuji berjumlah 1.372 sampel, yang terdiri dari dua kelas: kelas '0' (uang palsu) dan kelas '1' (uang asli). (Eyes et al., 2019)

2. Analisis Confusion Matrix

Berdasarkan hasil eksekusi model, diperoleh distribusi prediksi seperti yang ditampilkan dalam Tabel 1. *Confusion Matrix* di bawah ini merepresentasikan perbandingan antara kelas aktual dengan kelas yang diprediksi oleh sistem. (Barat, n.d.)

	True 0 (Palsu)	True 1 (Asli)	Class Precision
Prediksi 0	761	0	100%
Prediksi 1	1	610	99.84%
Class Recall	99.87%	100%	

Tabel 1. Matriks Konfusi K-NN (K=5)

Dari tabel di atas, terlihat bahwa algoritma mampu mengenali pola uang kertas dengan sangat presisi. Berikut adalah rincian analisisnya:

- Prediksi Benar (True Positive & True Negative):** Sistem berhasil mengklasifikasikan 761 data uang palsu dengan tepat sebagai 'Palsu', dan 610 data uang asli dengan tepat sebagai 'Asli'.
- Prediksi Salah (False Positive/Negative):** Terdapat 1 kejadian *misclassification*, di mana satu lembar uang yang seharusnya berlabel '0' (palsu) diprediksi oleh sistem sebagai '1' (asli). Sementara itu, tidak ada satupun uang asli yang salah diprediksi sebagai uang palsu.

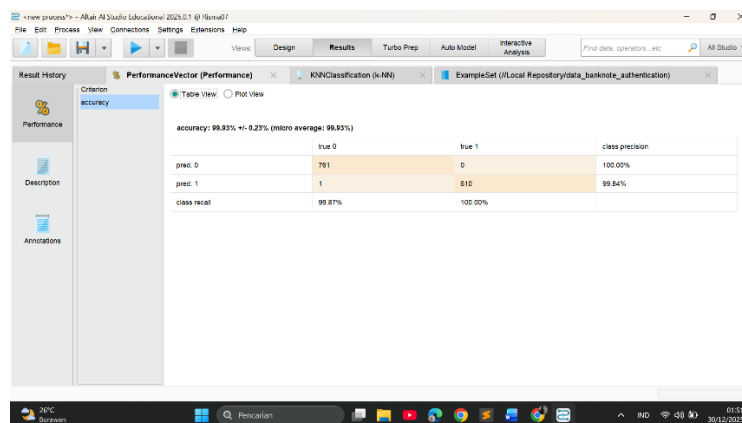
3. Evaluasi Performa Model

Secara keseluruhan, model K-NN dengan $k = 5$ menghasilkan performa yang sangat signifikan. Rangkuman metrik evaluasi dapat dilihat pada poin-poin berikut:

- Akurasi (Accuracy):** Mencapai **99,93%**. Angka ini menunjukkan bahwa variabel statistik (*Variance*, *Skewness*, *Curtosis*, *Entropy*) memiliki korelasi yang sangat kuat dalam membedakan entitas uang kertas. (Journal et al., 2011)
- Recall (Sensitivitas):** Pada kelas 1 (uang asli), nilai *Recall* mencapai **100%**. Dalam konteks sistem keamanan perbankan, hasil ini sangat krusial karena menjamin bahwa tidak ada uang asli yang akan ditolak oleh sistem (tidak ada *False Negative* pada kelas asli).



- c. **Presisi (Precision):** Nilai presisi untuk prediksi kelas 0 mencapai 100%, yang berarti setiap kali sistem menyatakan sebuah uang itu palsu, maka prediksi tersebut sepenuhnya benar.



Gambar 3. Hasil perhitungan dengan rapid miner

4. Evaluasi Performa Model

Tingginya akurasi yang diperoleh membuktikan bahwa Transformasi Wavelet sangat efektif dalam memisahkan fitur citra antara uang asli dan palsu. Algoritma K-NN, meskipun tergolong metode *lazy learning* yang sederhana, terbukti mampu bekerja optimal pada dataset ini karena fitur-fitur numerik yang dihasilkan memiliki jarak *Euclidean* yang tegas antar kelasnya. (Faustina et al., 2019)

Kesalahan tunggal yang terjadi pada satu sampel data kemungkinan disebabkan oleh adanya *outlier* atau noise pada fitur citra uang palsu tersebut yang memiliki karakteristik statistik (seperti varians atau kemencengan) yang sangat mirip dengan uang asli, sehingga "menipu" perhitungan jarak tetangga terdekat. Namun, dengan tingkat kesalahan hanya 0,07% (1 dari 1.372 data), model ini dapat dikatakan sangat *robust* (tangguh) dan layak diimplementasikan sebagai mesin verifikasi otomatis. (Wonges & Fa, 2023).

KESIMPULAN

Berdasarkan hasil eksperimen dan analisis yang telah dilakukan terhadap dataset *Banknote Authentication* menggunakan algoritma *K-Nearest Neighbor* (K-NN), dapat ditarik beberapa kesimpulan sebagai berikut:

1. **Efektivitas Fitur Statistik:** Penerapan ekstraksi fitur berbasis statistik (*Variance*, *Skewness*, *Curtosis*, dan *Entropy*) terbukti sangat efektif dalam merepresentasikan karakteristik unik citra uang kertas. Pola data yang dihasilkan memiliki keterpisahan (*separability*) yang tinggi antara kelas uang asli dan uang palsu. (Debarr, n.d.)



2. **Performa Algoritma:** Algoritma K-NN dengan konfigurasi $k=5$ dan pengukuran jarak *Euclidean* menunjukkan kinerja yang sangat presisi. Melalui pengujian validasi silang 10-lipatan (*10-fold Cross Validation*), model mencapai tingkat akurasi **99,93%**.
3. **Keandalan Sistem:** Model berhasil mencapai nilai *Recall* 100% untuk kelas uang asli. Hal ini mengindikasikan bahwa sistem sangat andal untuk diterapkan dalam skenario dunia nyata, karena mampu mengenali seluruh uang asli tanpa ada yang terdeteksi salah sebagai uang palsu (*zero false negative* untuk kelas positif).

UCAPAN TERIMA KASIH

Puji syukur penulis panjatkan ke hadirat Allah SWT karena atas rahmat dan karunia-Nya penelitian ini dapat terselesaikan dengan baik. Penulis menyampaikan ucapan terima kasih yang tulus kepada:

1. **Universitas Pancasakti Tegal**, khususnya Fakultas Teknik dan Ilmu Komputer, yang telah memberikan dukungan fasilitas serta lingkungan akademik yang kondusif selama pelaksanaan penelitian ini.
2. **UCI Machine Learning Repository**, selaku penyedia dataset publik *Real Estate Valuation* yang telah memfasilitasi ketersediaan data secara terbuka untuk kepentingan riset dan pengembangan ilmu pengetahuan.
3. Rekan-rekan mahasiswa Informatika semester 5 yang telah memberikan dukungan moral serta diskusi yang membangun selama proses penyusunan jurnal ini.

Semoga hasil penelitian ini dapat memberikan manfaat bagi perkembangan teknologi informasi, khususnya dalam bidang *Data Mining* dan sistem pendukung keputusan.

DAFTAR PUSTAKA

Barat, P. J. (n.d.). *No Title*.

Debarr, D. (n.d.). *Chapter 1 Pattern Recognition in Time Series*. 1–28.

Eyes, B., Engineering, I., & Publication, S. (2019). *International Journal of Recent Technology and Engineering Website : www.ijrte.org*. 6.

Faustina, F. D. E. A., Matematika, P. S., Sains, F., Teknologi, D. A. N., Islam, U., & Sunan, N. (2019). *Identifikasi pola kerapuhan tulang berdasarkan fitur tekstur citra dental panoramic radiograph (dpr) menggunakan gray level run length matrix (glrlm) dan support vector machine (svm)*.

Journal, I., Economics, O. F., & Studies, F. (2011). *Data Mining, Capital Markets, Data Mining Methods*. 3(2), 57–67.

Mo, M., Wai, M., Aung, N. P., & Htay, L. L. (2019). *Software Implementation of Iris Recognition System using MATLAB*. 3(5), 290–295.

Nabuasa, Y. N., Komputer, J. I., Cendana, U. N., Digital, C., & Histogram, E. (2019). *PENGOLAHAN CITRA DIGITAL PERBANDINGAN METODE HISTOGRAM*



EQUALIZATION DAN. 7(1), 87–95.

resolution_preserving.pdf. (n.d.).

Science, F. (2024). *Digital Image Analysis (DIA) in Food Technology : An Overview*. 21(2), 29–34.

Wonges, R. P., & Fa, L. B. (2023). *Penerapan Algoritma A *(A-Star) untuk Mencari Jalur Terdekat (Shortest Path) dalam Kecerdasan Buatan (Studi Kasus : Game Snake)*. 12(2).