

# Implementasi Convolutional Neural Network dan Large Language Model untuk Klasifikasi dan Deskripsi Pengenalan Pola Crochet

Hernanda Lilih Kusumaningtyas, Dedi Gunawan\*

Teknik Informatika, Fakultas Komunikasi dan Informatika, Universitas Muhammadiyah Surakarta, Surakarta, Indonesia

Email: <sup>1</sup>1200220143@student.ums.ac.id, <sup>2\*</sup>dedi.gunawan@ums.ac.id

Email Penulis Korespondensi: dedi.gunawan@ums.ac.id\*

Submitted: 14/04/2026; Accepted: 30/04/2026; Published: 30/06/2026

**Abstrak**—*Crochet* merupakan kerajinan tekstil dengan beragam pola yang sulit diidentifikasi secara manual karena memerlukan pengalaman dan waktu. Penelitian ini bertujuan mengembangkan sistem klasifikasi pola *crochet* berbasis *Convolutional Neural Network* (CNN) serta mengintegrasikannya dengan *Large Language Model* (LLM) untuk menghasilkan deskripsi pola secara otomatis dalam bahasa Indonesia. Dataset yang digunakan terdiri dari 830 citra yang dibagi ke dalam delapan kelas. Model CNN digunakan untuk ekstraksi fitur dan klasifikasi citra, kemudian hasil prediksi diteruskan ke LLM untuk menghasilkan deskripsi tekstual. Evaluasi dilakukan menggunakan metrik *accuracy*, *precision*, *recall*, *F1-score*, dan *confusion matrix*. Hasil pengujian menunjukkan bahwa model mencapai akurasi sebesar 93% dengan nilai *macro average* dan *weighted average* sebesar 0,93, yang mengindikasikan performa yang seimbang pada seluruh kelas, meskipun masih terdapat kesalahan pada pola dengan kemiripan tekstur. Kontribusi utama penelitian ini terletak pada integrasi CNN dan LLM dalam satu sistem multimodal yang tidak hanya menghasilkan klasifikasi, tetapi juga deskripsi pola secara otomatis. Sistem diimplementasikan dalam aplikasi berbasis web yang memungkinkan pengguna memperoleh hasil klasifikasi, *confidence score*, dan deskripsi secara langsung. Pendekatan ini diharapkan dapat meningkatkan efisiensi proses identifikasi pola *crochet* serta mendukung pengembangan sistem multimodal pada domain tekstil.

**Kata Kunci:** Pengenalan Pola; *Convolutional Neural Network*; *Large Language Model*; *Computer Vision*; Pengenalan Pola Crochet

**Abstract**—*Crochet* is a textile craft with various patterns that are often difficult to identify manually due to the need for experience and time. This study aims to develop a *crochet* pattern classification system using a *Convolutional Neural Network* (CNN) and integrate it with a *Large Language Model* (LLM) to automatically generate pattern descriptions in Indonesian. The dataset consists of 830 images divided into eight classes. The CNN model is used for feature extraction and image classification, and the prediction results are then forwarded to the LLM to generate textual descriptions. The evaluation is conducted using *accuracy*, *precision*, *recall*, *F1-score*, and *confusion matrix* metrics. The results show that the model achieved an accuracy of 93%, with *macro average* and *weighted average* values of 0.93, indicating balanced performance across all classes, although misclassifications still occur among patterns with similar textures. The main contribution of this study lies in integrating CNN and LLM into a multimodal system that provides both classification results and automatic descriptive explanations. The system is implemented as a web-based application that allows users to obtain classification results, confidence scores, and generated descriptions. This approach is expected to improve the efficiency of *crochet* pattern identification and support the development of multimodal systems in the textile domain.

**Keywords:** Pattern Recognition; *Convolutional Neural Network*; *Large Language Model*; *Computer Vision*; *Crochet Pattern Recognition*

## 1. PENDAHULUAN

Kerajinan *crochet* atau yang dikenal dengan rajut telah berkembang menjadi salah satu bentuk seni tekstil yang banyak diaplikasikan dalam pembuatan pakaian, aksesoris, maupun produk dekoratif [1]. Setiap pola *crochet* terbentuk dari kombinasi berbagai jenis jahitan (*stitch*) yang menghasilkan karakteristik visual unik berupa variasi bentuk, warna, dan tekstur. Namun, kompleksitas visual dan kemiripan antar pola menjadi tantangan dalam proses identifikasi, terutama pada tekstil dengan struktur pola berulang dan variasi tekstur tinggi [2]. Permasalahan serupa juga ditemukan pada penelitian pengenalan motif tekstil lainnya, seperti batik, di mana variasi dan kemiripan pola visual dapat memengaruhi kinerja model dan menurunkan akurasi klasifikasi [3]. Kondisi ini menggarisbawahi urgensi pengembangan metode komputasi yang mampu mengenali dan mengklasifikasikan pola secara lebih tepat dan konsisten.

Kemajuan yang dicapai dalam *computer vision* dan *deep learning* memberikan peluang yang luar biasa untuk menangani beragam permasalahan terkait klasifikasi citra, khususnya pada ranah tekstil. Di antara berbagai pendekatan yang ada, *Convolutional Neural Network* (CNN) menjadi pilihan utama karena kemampuannya mengklasifikasikan citra dengan mengekstraksi fitur visual secara otomatis dan terstruktur secara hierarkis [4][5]. CNN adalah arsitektur *deep learning* yang dirancang untuk mempelajari representasi fitur citra secara bertingkat, berawal dari karakteristik dasar seperti garis tepi hingga karakteristik kompleks seperti struktur objek, tanpa bergantung pada proses ekstraksi fitur secara manual [6][7]. Penelitian sebelumnya menunjukkan bahwa CNN mampu mengklasifikasikan motif kain tenun dengan performa yang baik, dengan tingkat akurasi mencapai sekitar 85%–90% [8]. Penelitian lain oleh Herwingsyah & Yuswanto Jaya juga mengungkapkan bahwa CNN terbukti andal dalam mengidentifikasi pakaian tradisional berdasarkan karakteristik visualnya [9]. Selain itu, penelitian Mufizar

et al. mengkaji klasifikasi motif batik dan menemukan bahwa kemiripan pola menjadi tantangan dalam meningkatkan akurasi model [3]. Penelitian Rizki et al. mengkaji penggunaan Faster R-CNN untuk mengenali dan mengkategorikan motif kain tenun yang mampu mengenali struktur visual kompleks pada citra tekstil, sehingga sesuai dengan karakteristik pola yang berulang dan memiliki variasi tinggi [10]. Selain itu, penelitian terbaru juga mengungkapkan bahwa Convolutional Neural Network (CNN) terbukti menghasilkan kinerja yang sangat baik dalam deteksi dan klasifikasi tekstil. Sebuah studi mengusulkan model CNN yang dirancang untuk mendeteksi cacat kain dengan kompleksitas komputasi yang rendah dan mampu mencapai akurasi di atas 96%, serta menghasilkan tingkat *precision*, *recall*, dan *F1-score* yang sangat baik melalui pemanfaatan arsitektur berbasis ResNet, *data augmentation*, dan *weighted loss function* [11]. Selain itu, CNN juga banyak digunakan dalam klasifikasi tekstur citra, khususnya pada domain tekstil, di mana model mampu mengenali pola berulang dan karakteristik tekstur secara efektif melalui pembelajaran fitur mendalam, sehingga sangat relevan untuk diterapkan pada pola *crochet* yang memiliki karakteristik tekstur berbasis jahitan dan pola berulang [12].

Meskipun berbagai penelitian telah berhasil menerapkan CNN pada klasifikasi citra tekstil, sebagian besar masih berfokus pada objek seperti batik, kain tenun, dan pakaian tradisional. *Crochet* sebagai salah satu bentuk tekstil dengan karakteristik pola berbasis jahitan yang unik masih relatif kurang dieksplorasi dalam konteks konteks pengenalan citra berbasis *deep learning*. Penelitian terkait *crochet* sendiri mulai berkembang, namun sebagian besar masih berfokus pada pengembangan alat bantu perancangan pola, bahasa visual *crochet*, pemanfaatan LLM untuk translasi pola, serta eksplorasi kreativitas melalui *generative AI* [13][14]. Penelitian-penelitian tersebut belum membahas pengenalan pola *crochet* berbasis citra menggunakan pendekatan *computer vision*. Selain itu, belum terdapat penelitian yang secara khusus mengintegrasikan proses klasifikasi pola *crochet* berbasis citra dengan kemampuan generasi deskripsi berbasis bahasa dalam satu sistem terpadu. Oleh karena itu, integrasi antara model CNN dan *Large Language Model* (LLM) menjadi pendekatan yang potensial dan relevan dalam pengembangan sistem cerdas berbasis *computer vision*, khususnya dalam kerangka *vision-language model*.

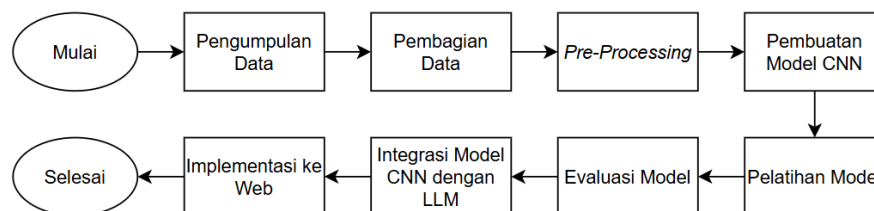
Untuk menjawab kesenjangan tersebut, pendekatan multimodal yang mengintegrasikan modalitas visual dan bahasa dalam satu sistem terpadu menjadi solusi yang relevan untuk dieksplorasi [15][16]. Pendekatan multimodal memungkinkan sistem tidak hanya mengenali objek visual, tetapi juga menghasilkan deskripsi dalam bentuk bahasa alami sehingga informasi yang diberikan menjadi lebih informatif, komunikatif, dan mudah dipahami oleh pengguna [17][18]. Integrasi ini juga memungkinkan sistem untuk menghasilkan deskripsi atau penjelasan berbasis konteks dari hasil klasifikasi, sehingga memperluas fungsi model dari sekadar prediksi menjadi sistem berbasis pengetahuan visual [19]. Hal ini menjadikan integrasi model CNN dan *Large Language Model* (LLM) sebagai salah satu solusi yang layak dipertimbangkan dalam membangun sistem cerdas berbasis visi komputer modern, khususnya dalam kerangka *vision-language model*.

Mengacu pada permasalahan yang telah diuraikan, penelitian ini dirancang untuk mengembangkan sistem pengenalan pola *crochet* berbasis kecerdasan buatan dengan memadukan *Convolutional Neural Network* (CNN) sebagai komponen klasifikasi citra dan *Large Language Model* (LLM) untuk menghasilkan deskripsi serta rekomendasi penggunaan pola. Sistem dirancang secara menyeluruh mulai dari tahap akuisisi citra hingga menghasilkan output berupa label klasifikasi dan deskripsi yang informatif. Selain itu, sistem ini diharapkan mampu memberikan nilai tambah berupa penjelasan karakteristik visual yang dapat membantu pengguna dalam memahami hasil klasifikasi secara lebih menyeluruh dan terperinci. Penelitian ini diharapkan mampu berkontribusi sebagai rujukan dalam perancangan sistem klasifikasi tekstil menggunakan *deep learning* pada objek yang memiliki tingkat kemiripan visual tinggi, sekaligus mendukung kemajuan aplikasi multimodal di bidang industri kreatif berbasis kerajinan tangan.

## 2. METODOLOGI PENELITIAN

### 2.1 Tahapan Penelitian

Penelitian ini mengadopsi pendekatan eksperimental dalam membangun sistem klasifikasi pola *crochet* yang berbasis *Convolutional Neural Network* (CNN) dan dipadukan dengan *Large Language Model* (LLM). Rangkaian tahapan penelitian meliputi pengumpulan dataset, pembagian data, *preprocessing*, perancangan dan pelatihan model, evaluasi, integrasi model, serta implementasi. Alur penelitian ditunjukkan pada Gambar 1.



Gambar 1. Alur Penelitian

### 2.2 Pengumpulan Data

Dataset dalam penelitian ini berupa citra pola *crochet* yang dikumpulkan secara mandiri dari berbagai sumber daring, seperti Google dan Pinterest. Proses pengumpulan dilakukan dengan mempertimbangkan keberagaman pola agar model dapat mengenali variasi *stitch* secara lebih baik. Dataset yang diperoleh terdiri dari delapan kelas pola *crochet*, dengan total sebanyak 830 citra. Setiap citra kemudian dikelompokkan ke dalam kelas yang sesuai berdasarkan jenis pola. Sebagian besar kelas memiliki masing-masing 100 citra, sedangkan satu kelas memiliki 130 citra. Delapan kelas pola tersebut meliputi *alpine*, *basketweave*, *chevron*, *granny square*, *ripple*, *shell*, *v*, dan *waffle*.

### 2.3 Pembagian Data

Pembagian dataset dilakukan menjadi dua subset, yakni data latih dan data uji, dengan perbandingan 80:20. Penentuan proporsi ini didasarkan pada standar yang umum digunakan dalam kajian *machine learning*, terutama pada dataset dengan jumlah terbatas, di mana porsi data pelatihan yang lebih besar sangat diperlukan guna memungkinkan model mempelajari representasi pola secara menyeluruh, adapun data pengujian berperan dalam mengevaluasi kemampuan generalisasi model pada data yang sebelumnya belum pernah dijumpai [20]. Pendekatan ini juga banyak digunakan pada penelitian-penelitian sebelumnya dalam klasifikasi citra [9][21]. Dari keseluruhan 830 citra, sejumlah 664 citra ditetapkan untuk keperluan pelatihan model dan 166 citra difungsikan untuk keperluan pengujian.

### 2.4 Pre-Processing

Data yang berhasil dikumpulkan selanjutnya dilakukan serangkaian tahapan *pre-processing* untuk memastikan konsistensi ukuran, format, dan skala piksel sebelum proses pelatihan model. Pada data uji, citra di-*resize* menjadi 256 piksel sebagai langkah standar yang lazim diterapkan pada model berbasis ResNet agar citra memiliki resolusi yang cukup sebelum dipotong, kemudian dilakukan *center crop* menjadi ukuran 224×224 yang merupakan standar input pada arsitektur CNN seperti ResNet, sehingga model dapat fokus pada area pola yang relevan dan meminimalkan distorsi pada tepi citra. Selanjutnya, citra dikonversi menjadi tensor guna mendukung proses komputasi berbasis GPU. Pada tahap berikutnya, diterapkan proses normalisasi dengan mengacu pada nilai rata-rata dan standar deviasi yang bersumber dari dataset ImageNet guna meningkatkan stabilitas pelatihan serta mempercepat konvergensi model dalam proses pembelajaran [22].

Pada data latih, selain tahapan dasar tersebut, diterapkan teknik *data augmentation* guna memperluas variasi data serta mengoptimalkan kemampuan model dalam melakukan generalisasi, khususnya pada dataset berukuran terbatas [23]. Augmentasi yang digunakan meliputi *RandomResizedCrop* dengan skala 0,7–1,0 untuk mensimulasikan variasi ukuran dan sudut pandang objek; *RandomHorizontalFlip* dan *RandomVerticalFlip* dengan probabilitas 0,5 untuk menghasilkan variasi orientasi citra secara seimbang; *RandomRotation* sebesar 10 derajat digunakan untuk menangani perubahan orientasi minor yang berpotensi muncul pada citra pola *crochet*; serta *ColorJitter* dengan pengaturan kecerahan dan kontras sebesar 0,2 untuk memperkuat ketangguhan model dalam menghadapi perbedaan kondisi cahaya. Dengan penerapan augmentasi ini, model diharapkan mampu mengenali pola dari berbagai variasi ukuran, orientasi, dan kondisi pencahayaan.

### 2.5 Pembuatan Model *Convolutional Neural Network*

*Convolutional Neural Network* (CNN) merupakan salah satu jenis arsitektur jaringan saraf tiruan yang banyak diimplementasikan dalam analisis citra dan terbukti menghasilkan kinerja yang sangat kompetitif pada sejumlah bidang seperti *computer vision*, *medical imaging*, dan *autonomous systems* [24]. CNN menjalankan prinsip kerjanya dengan menerapkan operasi konvolusi, yaitu mekanisme penyaringan citra menggunakan kernel atau filter yang berfungsi untuk menggali dan mengekstraksi representasi fitur-fitur signifikan dari suatu citra. Meskipun memiliki kemiripan dengan jaringan saraf konvensional, CNN tersusun atas neuron yang dilengkapi dengan bobot, nilai bias, dan fungsi aktivasi. Pada umumnya, CNN dibangun dari sejumlah komponen lapisan, yaitu *convolutional layer*, *pooling layer*, dan *fully connected layer*, yang masing-masing memiliki peranan tersendiri dalam rangkaian proses pembelajaran fitur dan pengklasifikasian citra.

Penelitian ini mengimplementasikan arsitektur ResNet-50, salah satu varian CNN yang dipilih karena kemampuannya mengatasi permasalahan *vanishing gradient* melalui konsep *residual block* dan *shortcut connections*, sehingga memungkinkan proses pelatihan jaringan yang lebih dalam berlangsung secara stabil dan efisien [25]. Dibandingkan dengan arsitektur CNN lain seperti VGG dan AlexNet, ResNet-50 menawarkan keseimbangan yang baik antara kedalaman jaringan, akurasi, dan efisiensi komputasi, sehingga menjadikannya pilihan yang tepat untuk tugas klasifikasi citra dengan dataset berukuran terbatas [26]. Setiap *residual block* tersusun atas beberapa *convolutional layer* yang diikuti oleh *batch normalization* dan *activation layer* (ReLU). Selain itu, *pooling layer* berfungsi untuk melakukan pengurangan dimensi fitur sembari tetap melestarikan informasi yang relevan, sedangkan *fully connected layer* pada bagian akhir bertugas menghasilkan prediksi kelas.

Model menggunakan *pre-trained weights* dari ImageNet sebagai bagian dari pendekatan transfer learning untuk memanfaatkan fitur yang telah dipelajari sebelumnya [27]. Pada tahap awal, seluruh *layer* dibekukan (*freeze*) untuk menjaga fitur-fitur generik dan fundamental yang telah diperoleh melalui proses pembelajaran pada dataset berskala besar. Selanjutnya, dilakukan *fine-tuning* dengan membuka (*unfreeze*) parameter pada blok *layer3* dan

*layer4* karena lapisan-lapisan akhir pada ResNet cenderung mempelajari fitur yang lebih spesifik dan tingkat tinggi, sehingga perlu diadaptasikan sesuai dengan karakteristik dataset *crochet* yang digunakan.

Pada bagian akhir model, *fully connected layer* digantikan dengan arsitektur baru yang terdiri atas *dropout* dengan nilai 0,5 sebagai upaya untuk mengurangi potensi *overfitting* pada model. *Dropout* sebesar 0,5 diterapkan pada *fully connected layer* untuk menghindari *overfitting* melalui mekanisme penghentian sementara sejumlah neuron secara acak pada tahap pelatihan, sehingga model tidak mengalami ketergantungan berlebihan terhadap fitur-fitur tertentu dan mampu menunjukkan kemampuan generalisasi yang lebih tinggi pada data yang sebelumnya belum pernah dikenali. Arsitektur tersebut juga dilengkapi dengan *linear layer* yang disesuaikan berdasarkan jumlah kelas yang ditetapkan dalam penelitian ini. Dengan pendekatan tersebut, model mampu memadukan fitur umum dari *pre-trained model* dengan karakteristik khusus dataset klasifikasi pola *crochet*.

## 2.6 Pelatihan Model CNN

Model dilatih menggunakan data *training* untuk mengenali pola pada citra *crochet* melalui proses *forward propagation* guna menghasilkan prediksi dan *backward propagation* untuk memperbarui bobot berdasarkan nilai *loss*. *Training* dilakukan dengan memanfaatkan *optimizer* Adam yang dikonfigurasi dengan *learning rate* 0.0003, *batch size* 32, serta *weight decay* 1e-4 untuk mengurangi risiko *overfitting*. *Optimizer* Adam dipilih karena secara konsisten menunjukkan performa kompetitif pada berbagai tugas *deep learning* dan tetap menjadi pilihan yang kuat meskipun metode-metode terbaru belum mampu melampauinya secara signifikan [28]. Nilai *learning rate* 0,0003 ditentukan berdasarkan eksperimen awal untuk menyeimbangkan kecepatan konvergensi dan kestabilan pelatihan. *Weight decay* 1e-4 diterapkan sebagai regularisasi L2 yang bekerja dengan memberikan penalti pada bobot yang terlalu besar sehingga mendorong model untuk mempelajari representasi yang lebih sederhana dan mengurangi risiko *overfitting* [6]. *Batch size* 32 dipilih karena nilai *batch size* yang lebih besar tidak selalu menghasilkan akurasi yang lebih tinggi, sehingga ukuran yang lebih kecil-menengah dipilih guna menjaga keseimbangan optimal antara efisiensi komputasi dan kapasitas generalisasi model [29].

Di samping itu, digunakan *scheduler* ReduceLROnPlateau yang berfungsi menyesuaikan *learning rate* secara adaptif ketika performa model tidak menunjukkan peningkatan selama sejumlah epoch tertentu [30]. Nilai *factor* 0,3 dan *patience* 3 ditentukan berdasarkan eksperimen awal, di mana *factor* 0,3 dipilih untuk menurunkan *learning rate* secara cukup signifikan saat terjadi stagnasi, sedangkan *patience* 3 ditentukan agar penurunan *learning rate* tidak terlalu cepat mengingat jumlah *epoch* pelatihan yang terbatas. Model dilatih selama 15 *epoch* yang dinilai memadai untuk model berbasis *transfer learning* pada dataset berukuran terbatas, karena *pre-trained weights* dari ImageNet memungkinkan model mencapai konvergensi lebih cepat dibandingkan pelatihan dari awal. Evaluasi dilakukan pada setiap *epoch*, dan bobot dengan akurasi terbaik pada data uji disimpan sebagai model *final*.

## 2.7 Evaluasi Model

Sesuai proses pelatihan berlangsung, dilakukan tahap evaluasi untuk menilai sejauh mana kemampuan model CNN dalam mengidentifikasi dan mengkategorikan pola *crochet* dengan tepat. Tujuan evaluasi ini adalah memastikan model tidak semata-mata unggul dalam hal prediksi, melainkan juga mampu secara konsisten mengidentifikasi setiap kategori pola yang ada [21]. Untuk itu, digunakan serangkaian metrik yang dapat menghasilkan gambaran menyeluruh dan terperinci mengenai tingkat kinerja model [31]. Pengukuran performa model dilakukan dengan mengacu pada metrik *accuracy*, *precision*, *recall*, dan *F1-score*.

Nilai *accuracy* merepresentasikan tingkat akurasi model dalam melakukan klasifikasi secara menyeluruh, yakni rasio antara jumlah prediksi yang tepat dibandingkan keseluruhan data, sebagaimana diformulasikan pada persamaan (1).

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \quad (1)$$

Nilai *recall* menggambarkan sejauh mana model mampu mengenali seluruh data positif secara akurat, sebagaimana dituliskan pada persamaan (2).

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

Nilai *precision* menggambarkan kemampuan model dalam memprediksi data positif secara tepat, sebagaimana dituliskan pada persamaan (3).

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

Nilai *F1-Score* adalah ukuran rata-rata harmonis yang menggabungkan *precision* dan *recall* guna mengukur tingkat keseimbangan antara keduanya, sebagaimana diformulasikan pada persamaan (4).

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

Keterangan:

TP ( <i>True Positive</i> )	= banyaknya data yang tepat diklasifikasikan sebagai kelas positif
TN ( <i>True Negative</i> )	= banyaknya data yang tepat diklasifikasikan sebagai kelas negatif
FP ( <i>False Positive</i> )	= banyaknya data yang keliru diklasifikasikan sebagai kelas positif
FN ( <i>False Negative</i> )	= banyaknya data yang keliru diklasifikasikan sebagai kelas negatif

## 2.8 Integrasi Model *Convolutional Neural Network* dengan *Large Language Model*

Integrasi antara *Convolutional Neural Network* (CNN) dan *Large Language Model* (LLM) dilakukan untuk menggabungkan kemampuan analisis citra dan pemrosesan bahasa dalam satu sistem. CNN terlebih dahulu memproses citra untuk menghasilkan prediksi berupa label kelas pola *crochet*. Label tersebut kemudian dikonversi menjadi *prompt* teks terstruktur yang digunakan sebagai input bagi LLM. Dengan mekanisme ini, LLM memanfaatkan label hasil klasifikasi sebagai konteks untuk menghasilkan deskripsi tekstual yang sesuai. Proses ini memungkinkan transformasi keluaran CNN yang bersifat kategorikal menjadi representasi bahasa alami yang lebih informatif tanpa memerlukan pelatihan ulang pada model LLM. LLM umumnya dibangun menggunakan arsitektur *transformer* yang menggunakan mekanisme *attention* sebagai fondasi untuk memahami konteks bahasa serta menghasilkan respons yang terstruktur dan runtut [32].

Dalam penelitian ini digunakan model LLM Qwen, yang diperkenalkan sebagai seri pertama dari keluarga model bahasa besar yang dikembangkan oleh Alibaba Group, yang dilatih menggunakan dataset berskala besar dan menunjukkan performa superior pada berbagai tugas, termasuk model chat yang menghasilkan respons yang kompetitif dan informatif [33]. Parameter *temperature* sebesar 0,7 digunakan untuk menjaga keseimbangan antara variasi dan konsistensi respons, di mana nilai yang semakin besar cenderung menghasilkan teks yang lebih bervariasi, sementara nilai yang semakin kecil menghasilkan keluaran yang lebih deterministik. Selain itu, parameter *max\_tokens* sebesar 500 digunakan untuk membatasi panjang deskripsi agar tetap informatif namun tidak berlebihan. Nilai parameter tersebut ditentukan berdasarkan eksperimen awal dengan mempertimbangkan kualitas keluaran, khususnya dari segi kelengkapan informasi, keterbacaan, dan konsistensi deskripsi. Dengan integrasi ini, sistem tidak hanya mampu mengidentifikasi pola *crochet* dari citra, tetapi juga menyajikan penjelasan yang lebih kontekstual sehingga meningkatkan interpretabilitas hasil model.

## 2.9 Implementasi ke Web

Sistem diimplementasikan pada platform web sehingga pengguna dapat mengunggah gambar pola *crochet* dan memperoleh hasil klasifikasi beserta deskripsi pola secara otomatis. *Backend* dibangun menggunakan *Python Flask*, yang langsung memuat model CNN (*.pth*) untuk melakukan prediksi kelas pola *crochet* secara *real-time*. Untuk deskripsi pola yang dihasilkan oleh LLM, sistem menggunakan *REST API* agar *backend* dapat berkomunikasi dengan model bahasa. *Frontend* dibuat dengan memanfaatkan *HTML*, *CSS*, dan *JavaScript*, sehingga interaktif dan responsif di berbagai perangkat.

# 3. HASIL DAN PEMBAHASAN

Bagian ini menguraikan hasil penerapan sistem klasifikasi pola *crochet* berbasis web yang telah dikembangkan, serta analisis kinerja sistem melalui serangkaian proses pelatihan dan evaluasi model. Pembahasan mencakup hasil pelatihan model *Convolutional Neural Network* (CNN), evaluasi kinerja model dalam proses pengklasifikasian citra, serta implementasi sistem yang terintegrasi dengan *Large Language Model* (LLM) untuk mendukung penyajian informasi kepada pengguna.

## 3.1 Hasil Pelatihan dan Evaluasi Model CNN

### 3.1.1 Hasil Pelatihan Model CNN

Pelatihan model *Convolutional Neural Network* (CNN) yang menggunakan arsitektur ResNet-50 dilakukan dengan memanfaatkan kumpulan citra pola *crochet* yang terdiri atas 830 citra yang terbagi dalam delapan kelas, yakni *alpine*, *basketweave*, *chevron*, *granny square*, *ripple*, *shell*, *v*, dan *waffle*. Setiap kelas memiliki karakteristik visual yang berbeda, terutama pada pola jahitan, arah tekstur, dan tingkat kerapatan rajutan. Keberagaman karakteristik tersebut menjadikan proses klasifikasi cukup kompleks karena terdapat beberapa pola yang memiliki kemiripan visual. Hal ini memberikan tantangan tersendiri dalam proses pelatihan model.

Tahap pelatihan dilaksanakan selama 15 *epoch* dengan mengimplementasikan teknik *transfer learning* yang memanfaatkan bobot awal yang bersumber dari ImageNet. Selain itu, dilakukan *fine-tuning* pada layer akhir, yaitu *layer3* dan *layer4*, untuk menyesuaikan model dengan karakteristik citra pola *crochet*. Untuk meningkatkan variasi data latih, diterapkan teknik *data augmentation* yang meliputi *random resized crop*, *horizontal flip*, *vertical flip*, *rotation*, dan *color jitter*. Kombinasi pendekatan ini digunakan selama proses training model.

Selama *training*, model menunjukkan peningkatan performa pada setiap *epoch*. Di epoch pertama, akurasi pelatihan tercatat sebesar 39,46% dan akurasi pengujian sebesar 65,06%. Pada *epoch* ke-2, akurasi data uji

meningkat menjadi 86,75%. Peningkatan ini menunjukkan adanya perubahan performa model pada tahap awal pelatihan.

Performa model pada *epoch* berikutnya mengalami fluktuasi dengan nilai akurasi data uji berada pada rentang 83% hingga 92%. Nilai akurasi tertinggi diperoleh pada *epoch* ke-10 dengan akurasi data uji sebesar 92,77%. Oleh karena itu, model pada *epoch* ke-10 digunakan sebagai model terbaik. Riwayat akurasi model selama proses pelatihan ditunjukkan pada Tabel 1.

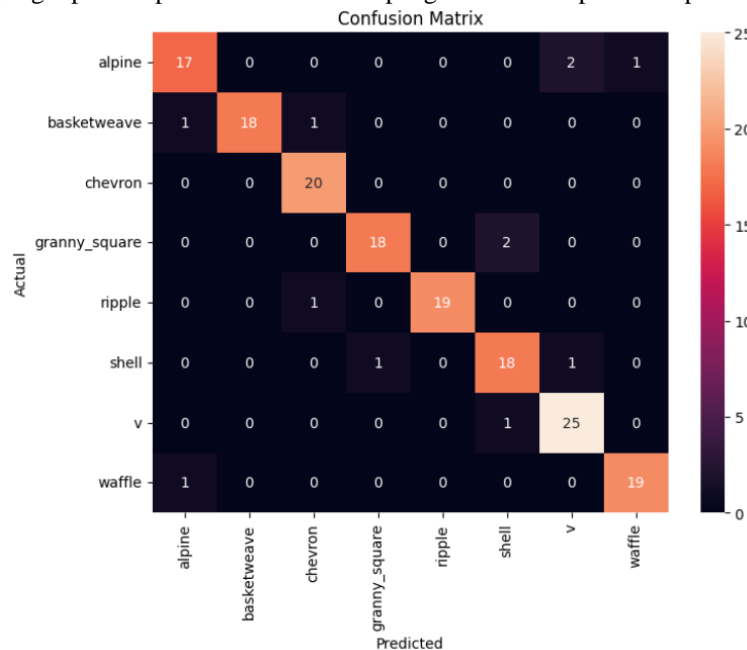
**Tabel 1.** Riwayat Akurasi Model pada setiap *Epoch*

<i>Epoch</i>	<i>Train Accuracy</i>	<i>Test Accuracy</i>
<i>Epoch 1</i>	39,46%	65,06%
<i>Epoch 2</i>	79,67%	86,75%
<i>Epoch 3</i>	90,96%	83,73%
<i>Epoch 4</i>	92,02%	84,94%
<i>Epoch 5</i>	92,62%	87,35%
<i>Epoch 6</i>	95,63%	90,36%
<i>Epoch 7</i>	95,33%	86,75%
<i>Epoch 8</i>	96,23%	85,54%
<i>Epoch 9</i>	97,89%	86,75%
<b><i>Epoch 10</i></b>	<b>95,63%</b>	<b>92,77%</b>
<i>Epoch 11</i>	97,14%	90,96%
<i>Epoch 12</i>	98,34%	90,96%
<i>Epoch 13</i>	97,74%	85,54%
<i>Epoch 14</i>	98,04%	86,75%
<i>Epoch 15</i>	98,80%	88,55%

Berdasarkan Tabel 1, akurasi data latih mengalami peningkatan hingga mencapai sekitar 98%, sedangkan akurasi data uji berada pada rentang yang relatif stabil setelah beberapa *epoch*. Nilai akurasi data uji tertinggi diperoleh pada *epoch* ke-10. Setelah *epoch* tersebut, nilai akurasi data uji mengalami fluktuasi namun tetap berada dalam rentang yang relatif stabil. Perbedaan antara akurasi data *training* dan data *testing* juga dapat diamati dari hasil pelatihan tersebut.

### 3.1.2 Evaluasi Model CNN

Setelah model terbaik diperoleh dari proses pelatihan, dilakukan evaluasi untuk melihat kinerja model dalam mengklasifikasikan citra pola *crochet*. Evaluasi tidak semata-mata bertumpu pada nilai akurasi, melainkan juga mempertimbangkan distribusi hasil prediksi terhadap label yang sesungguhnya. Salah satu pendekatan yang diterapkan dalam proses evaluasi ini adalah *confusion matrix*. Metode tersebut diterapkan guna memetakan banyaknya prediksi yang tepat maupun kesalahan dalam pengklasifikasian pada setiap kategori kelas.



**Gambar 3.** *Confusion Matrix* Hasil Klasifikasi Model

Hasil evaluasi melalui *confusion matrix* ditampilkan pada Gambar 3. Berdasarkan gambar tersebut, sebagian besar nilai terkonsentrasi pada diagonal utama yang mencerminkan banyaknya prediksi yang tepat pada setiap kategori kelas. Angka-angka yang terletak di luar diagonal mengindikasikan adanya ketidakepatan

klasifikasi pada beberapa kategori kelas. Pola distribusi nilai ini mampu merepresentasikan gambaran yang komprehensif terkait performa model pada setiap kelas.

Beberapa kelas seperti *chevron*, *ripple*, dan *waffle* memperlihatkan jumlah prediksi benar yang cukup tinggi. Selain itu, kelas *basketweave*, *granny square*, dan *shell* juga menunjukkan jumlah prediksi benar yang dominan. Hal ini tercermin dari nilai pada diagonal utama yang lebih tinggi apabila dibandingkan dengan nilai yang berada di posisi lainnya. Dengan demikian, model terbukti memiliki kemampuan dalam mengklasifikasikan mayoritas data dengan tingkat ketepatan yang baik.

Namun demikian, masih terdapat sejumlah kekeliruan klasifikasi pada kelas-kelas tertentu. Kesalahan tersebut terlihat pada nilai di luar diagonal utama, seperti pada kelas *alpine* dan *granny\_square*. Nilai tersebut menunjukkan adanya data yang diprediksi ke kelas lain. Distribusi kesalahan ini dapat diamati secara langsung pada *confusion matrix*.

### 3.1.3 Perhitungan Metrik Evaluasi

Untuk memperoleh gambaran yang lebih kuantitatif, dilakukan perhitungan berbagai metrik pengukuran evaluasi yang mencakup *accuracy*, *precision*, *recall*, dan *F1-score*. Sejumlah metrik tersebut dihitung berdasarkan nilai pada *confusion matrix* sehingga mampu menghadirkan analisis yang lebih terperinci. Dengan memanfaatkan metrik tersebut, performa model dapat dinilai baik secara keseluruhan maupun pada tingkat setiap kelas secara terpisah. Hasil perhitungan dirangkum dalam Tabel 2.

**Tabel 2.** Hasil Evaluasi Model CNN

Kelas Pola	<i>Precision</i>	<i>Recall</i>	<i>F1-Score</i>	<i>Support</i>
<i>alpine</i>	0,89	0,85	0,87	20
<i>basketweave</i>	1,00	0,90	0,95	20
<i>chevron</i>	0,91	1,00	0,95	20
<i>granny_square</i>	0,95	0,90	0,92	20
<i>ripple</i>	1,00	0,95	0,97	20
<i>shell</i>	0,86	0,90	0,88	20
<i>v</i>	0,89	0,96	0,93	26
<i>waffle</i>	0,95	0,95	0,95	20
<i>Accuracy</i>	-	-	0,93	166
<i>Macro AVG</i>	0,93	0,93	0,93	166
<i>Weighted AVG</i>	0,93	0,93	0,93	166

Sebagai bentuk validasi terhadap hasil pada Tabel 2, dilakukan contoh perhitungan metrik evaluasi pada beberapa kelas berdasarkan nilai pada *confusion matrix*. Perhitungan ini bertujuan untuk menunjukkan proses perolehan nilai *precision*, *recall*, dan *F1-score*.

- a. Perhitungan pada Kelas *Alpine*

$$Precision = \frac{TP}{TP + FP} = \frac{17}{17 + 2} = \frac{17}{19} = 0,8947 \approx \mathbf{0,89}$$

$$Recall = \frac{TP}{TP + FN} = \frac{17}{17 + 3} = \frac{17}{20} = \mathbf{0,85}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{0,8947 \times 0,85}{0,8947 + 0,85} = 0,8718 \approx \mathbf{0,87}$$

- b. Perhitungan pada Kelas *Ripple*

$$Precision = \frac{TP}{TP + FP} = \frac{19}{19 + 0} = \frac{19}{19} = \mathbf{1,00}$$

$$Recall = \frac{TP}{TP + FN} = \frac{19}{19 + 1} = \frac{19}{20} = \mathbf{0,95}$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} = 2 \times \frac{1,00 \times 0,95}{1,00 + 0,95} = 0,9744 \approx \mathbf{0,97}$$

Berdasarkan perhitungan tersebut, hasil *precision*, *recall*, dan *F1-score* yang diperoleh sesuai dengan nilai pada Tabel 2. Perolehan tertinggi tercatat pada kelas *ripple* dengan *F1-score* sebesar 0,97, sementara perolehan terendah dijumpai pada kelas *alpine* dan *shell*. Secara keseluruhan, model memperoleh nilai *accuracy* sebesar 0,93. Nilai *macro average* dan *weighted average* keduanya menunjukkan angka 0,93 untuk seluruh metrik.

### 3.2 Implementasi dan Pengujian Sistem

#### 3.2.1 Implementasi Sistem

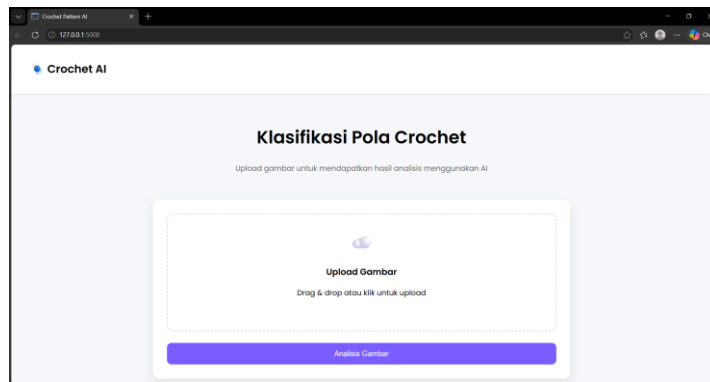
Model *Convolutional Neural Network* (CNN) yang sudah selesai dilatih selanjutnya diintegrasikan ke dalam sistem berbasis web agar dapat diakses oleh pengguna. Sistem ini dirancang untuk memfasilitasi pengguna dalam mengunggah citra pola *crochet* dan mendapatkan keluaran klasifikasi secara langsung dan otomatis. Penerapan sistem ini dimaksudkan untuk menyediakan media yang mempermudah proses identifikasi pola tanpa memerlukan pemahaman teknis mengenai model yang digunakan.

Setelah citra diunggah, sistem akan memproses gambar menggunakan model CNN untuk menentukan kelas pola *crochet*. Hasil prediksi ditampilkan dalam bentuk nama kelas sebagai keluaran utama dari sistem. Selain itu, sistem juga menampilkan nilai *confidence score* sebagai representasi tingkat kepercayaan model dalam menghasilkan prediksi yang diperoleh.

Untuk melengkapi informasi yang diberikan, sistem dihubungkan dengan modul *Large Language Model* (LLM) yang mampu memproduksi deskripsi secara otomatis berdasarkan hasil klasifikasi. Integrasi dilakukan melalui pemanggilan API menggunakan metode *HTTP request*. Sistem mengirimkan nama kelas hasil prediksi sebagai input, kemudian model bahasa menghasilkan deskripsi dalam bentuk teks.

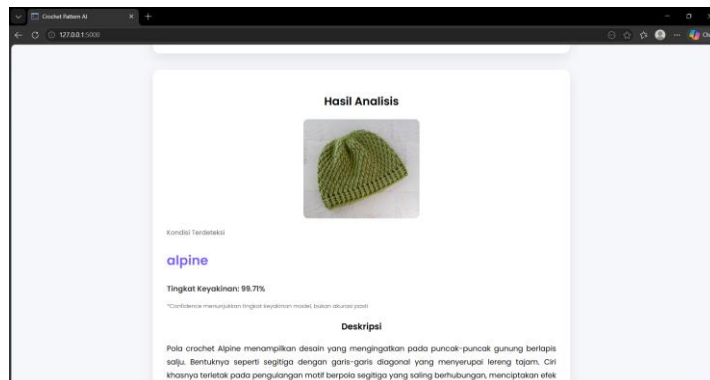
Model bahasa yang digunakan adalah Qwen2.5-VL-72B-Instruct yang diakses melalui layanan OpenRouter API. Parameter *temperature* sebesar 0,7 digunakan dalam proses generasi teks, sedangkan *max tokens* dibatasi hingga 500 token. Hasil keluaran dari LLM kemudian diproses untuk memastikan format teks rapi sebelum ditampilkan kepada pengguna.

Untuk memperjelas gambaran mengenai realisasi sistem yang telah dikembangkan, ditampilkan visualisasi tampilan antarmuka pengguna yang dimanfaatkan dalam proses klasifikasi. Tampilan ini mencakup halaman utama untuk mengunggah citra, hasil prediksi dari model CNN, serta deskripsi pola yang dihasilkan oleh modul LLM.



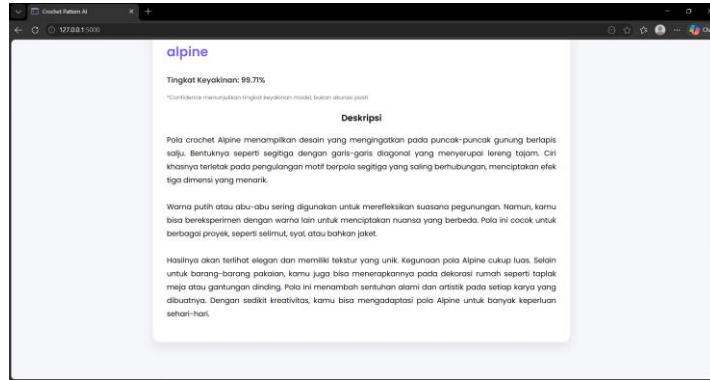
**Gambar 4.** Tampilan Antarmuka Sistem Web

Gambar 4. menampilkan halaman utama sistem yang berfungsi sebagai media pengunggahan citra pola *crochet*. Pengguna dapat memilih berkas gambar dan menjalankan proses klasifikasi melalui tombol yang tersedia. Antarmuka ini dirancang untuk mempermudah interaksi pengguna dengan sistem tanpa memerlukan pengetahuan teknis.



**Gambar 5.** Hasil Prediksi Kelas Pola *Crochet* oleh Model CNN

Tampilan pada Gambar 5. menampilkan keluaran sistem berupa nama pola hasil klasifikasi citra menggunakan model CNN yang disertai dengan tingkat keyakinan (*confidence score*) terhadap hasil prediksi berdasarkan gambar yang diunggah pengguna.



**Gambar 6.** Deskripsi Pola *Crochet* yang Dihasilkan oleh LLM







Gambar 6. memperlihatkan deskripsi tekstual pola *crochet* yang dihasilkan secara otomatis oleh modul LLM berdasarkan hasil klasifikasi model CNN.



Berdasarkan capaian implementasi yang telah dijalankan, sistem identifikasi pola *crochet* berbasis web berhasil berjalan dengan baik sesuai fungsi yang dirancang. Sistem mampu menjalankan proses klasifikasi citra secara otomatis serta menampilkan hasil prediksi beserta tingkat keyakinan model. Selain itu, integrasi dengan *Large Language Model* (LLM) memungkinkan sistem menghasilkan deskripsi pola secara otomatis. Dengan demikian, sistem yang dikembangkan tidak hanya menjalankan fungsi sebagai instrumen klasifikasi, namun sekaligus berperan sebagai rujukan informasi yang membantu pengguna dalam memahami karakteristik pola *crochet*.

### 3.2.2 Pengujian Sistem

Tahap pengujian dilaksanakan guna memverifikasi bahwa aplikasi yang telah dikembangkan dapat beroperasi dengan baik sesuai fungsinya. Pengujian dilakukan dengan mengunggah beberapa sampel citra pola *crochet* dari setiap kelas ke dalam sistem. Setiap citra yang diunggah selanjutnya diproses oleh model CNN guna menghasilkan keluaran prediksi kelas. Keluaran prediksi tersebut selanjutnya dicocokkan dengan label yang sesungguhnya guna mengetahui tingkat kesesuaian hasil klasifikasi.

**Tabel 3.** Hasil Validasi Sistem Web

Gambar	Label Sebenarnya	Prediksi Sistem	Kesesuaian
	<i>alpine</i>	<i>alpine</i>	Sesuai
	<i>basketweave</i>	<i>basketweave</i>	Sesuai
	<i>chevron</i>	<i>chevron</i>	Sesuai
	<i>granny_square</i>	<i>granny_square</i>	Sesuai
	<i>ripple</i>	<i>ripple</i>	Sesuai
	<i>shell</i>	<i>shell</i>	Sesuai

Gambar	Label Sebenarnya	Prediksi Sistem	Kesesuaian
	v	v	Sesuai
	waffle	waffle	Sesuai

Hasil pengujian sistem ditunjukkan pada Tabel 3. Berdasarkan hasil tersebut, sistem mampu menghasilkan prediksi yang sesuai dengan citra yang diunggah pengguna. Selain itu, sistem juga berhasil menampilkan *confidence score* serta deskripsi pola yang dihasilkan oleh modul *Large Language Model* (LLM). Hal tersebut membuktikan bahwa seluruh fungsi utama sistem bekerja dengan baik tanpa kendala.

Seluruh sampel pengujian menunjukkan kesesuaian antara label asli dan keluaran prediksi sistem. Kondisi ini menunjukkan bahwa model CNN yang diimplementasikan memiliki kapabilitas yang memadai dalam melakukan klasifikasi citra. Selain itu, integrasi antara model CNN dan LLM juga berjalan secara optimal dalam menghasilkan keluaran yang informatif. Secara keseluruhan, sistem mampu difungsikan sebagai sarana pendukung dalam mengidentifikasi pola *crochet* secara otomatis sekaligus menyajikan informasi tambahan kepada pengguna.

### 3.3 Pembahasan

Hasil penelitian membuktikan bahwa model *Convolutional Neural Network* (CNN) berbasis arsitektur ResNet-50 mampu mengklasifikasikan pola *crochet* dengan tingkat akurasi yang tinggi. Model berhasil meraih nilai *accuracy* sebesar 93%, dengan *macro average* dan *weighted average* sebesar 0,93 pada keseluruhan metrik evaluasi, yang mengindikasikan bahwa model menunjukkan kinerja yang seimbang dan stabil pada setiap kelas, tidak hanya pada kelas mayoritas. Perolehan *precision*, *recall*, dan *F1-score* yang relatif tinggi pada sebagian besar kelas turut membuktikan bahwa model menunjukkan kinerja yang konsisten dan stabil. Hal ini membuktikan bahwa pendekatan CNN terbukti mampu menangani permasalahan klasifikasi citra dengan keragaman karakteristik visual yang cukup tinggi.

Kendati demikian, masih ditemukan sejumlah ketidaktepatan klasifikasi pada kategori kelas tertentu, seperti *alpine* yang sesekali diklasifikasikan sebagai kelas lain yang memiliki kemiripan visual. Analisis terhadap *confusion matrix* mengungkapkan pola kesalahan klasifikasi yang perlu diperhatikan secara lebih mendalam. Kesalahan terbanyak ditemukan pada kelas *alpine*, di mana 2 citra keliru terklasifikasi sebagai kelas v dan 1 citra keliru terklasifikasi sebagai kelas *waffle*, sehingga hanya 17 dari 20 citra yang berhasil diklasifikasikan dengan benar, menghasilkan nilai *recall* terendah sebesar 0,85. Kesalahan ini dapat dijelaskan oleh kemiripan karakteristik visual antara pola *alpine* dengan pola v yang sama-sama memiliki elemen tekstur diagonal, serta kemiripan tingkat kerapatan rajutan dengan pola *waffle*. Selain itu, kelas *granny square* mengalami 2 kesalahan klasifikasi yang keduanya salah diprediksi sebagai kelas *shell*, kemungkinan disebabkan oleh kesamaan bentuk elemen melingkar pada kedua pola tersebut. Kelas *shell* sendiri memperoleh nilai *precision* terendah sebesar 0,86, yang mengindikasikan bahwa terdapat citra dari kelas lain yang keliru diprediksi sebagai *shell* akibat kemiripan bentuk lengkung yang dimiliki pola *shell* dengan beberapa pola lainnya. Secara keseluruhan, pola kesalahan yang terjadi secara konsisten menunjukkan bahwa misklasifikasi cenderung terjadi antarkelas yang memiliki kemiripan karakteristik visual, baik dari sisi arah tekstur, bentuk elemen rajutan, maupun tingkat kerapatan pola. Hal ini menunjukkan bahwa kemiripan tekstur dan struktur pola menjadi tantangan utama dalam proses klasifikasi. Fenomena ini sejalan dengan penelitian sebelumnya pada klasifikasi tekstil, seperti motif batik, yang juga menghadapi kendala serupa akibat kesamaan pola visual [3]. Dengan demikian, hasil penelitian ini memperkuat bahwa tingkat kemiripan visual antar kelas merupakan faktor yang memengaruhi performa model.

Apabila dibandingkan dengan penelitian-penelitian sebelumnya pada domain klasifikasi tekstil, hasil penelitian ini menunjukkan performa yang kompetitif. Beberapa penelitian terdahulu melaporkan akurasi pada rentang 82%–90%, seperti penelitian Dabbo & Bisilisin dengan akurasi 85% [8], Rizki et al. dengan akurasi 82,14% [10], serta Herwinskyah & Yuswanto Jaya dengan akurasi pengujian 90% namun memiliki keterbatasan pada data validasi [9]. Sementara itu, penelitian Mufizar et al. pada klasifikasi motif batik menunjukkan akurasi awal sebesar 75% yang meningkat hingga 100% setelah melalui proses optimasi yang sangat ekstensif [3]. Penelitian ini berhasil mencapai akurasi sebesar 93% tanpa memerlukan proses optimasi yang kompleks, melalui penerapan ResNet-50 dengan pendekatan *transfer learning* dan *fine-tuning*. Penggunaan teknik *data augmentation* juga berkontribusi dalam meningkatkan kemampuan generalisasi model meskipun jumlah dataset terbatas. Dengan demikian, pendekatan yang digunakan tidak hanya kompetitif dari sisi akurasi, tetapi juga lebih efisien dalam proses pelatihan.

Selain dari sisi klasifikasi, kontribusi signifikan penelitian ini terletak pada integrasi model CNN dengan *Large Language Model* (LLM) Qwen dalam satu sistem multimodal terpadu. Berbeda dengan penelitian-penelitian sebelumnya pada domain tekstil yang umumnya hanya menghasilkan keluaran berupa label klasifikasi, sistem dalam penelitian ini mampu menghasilkan deskripsi pola secara otomatis dalam bahasa Indonesia berdasarkan hasil prediksi CNN. Deskripsi yang dihasilkan LLM mencakup karakteristik visual pola, ciri khas, dan rekomendasi penggunaan pola *crochet*, sehingga menghadirkan informasi yang jauh lebih komprehensif dan mudah dipahami dibandingkan sekadar label kelas. Mekanisme integrasi bekerja secara sekuensial di mana label hasil prediksi CNN dikonversi menjadi *prompt* terstruktur yang kemudian diteruskan ke model Qwen untuk menghasilkan deskripsi tekstual. Pendekatan ini memperluas fungsi sistem dari sekadar alat klasifikasi menjadi sistem berbasis pengetahuan visual yang interaktif, sejalan dengan perkembangan pendekatan multimodal pada penelitian *vision-language*. Dengan adanya deskripsi tambahan, sistem menjadi lebih interpretatif dan memberikan nilai tambah nyata bagi pengguna, khususnya dalam konteks edukasi maupun industri kreatif berbasis kerajinan tangan yang membutuhkan pemahaman mendalam tentang karakteristik pola.

Hasil pengujian sistem memperlihatkan bahwa integrasi antara model CNN dan LLM mampu berjalan secara optimal dalam satu alur kerja. Sistem berhasil menerima masukan citra, menjalankan klasifikasi, serta menghasilkan deskripsi tanpa kendala berarti. Hal ini menunjukkan bahwa pendekatan multimodal yang memadukan pemrosesan citra dan bahasa alami dapat diimplementasikan secara efektif. Dengan demikian, sistem yang dibangun tidak semata-mata berfungsi sebagai perangkat klasifikasi, melainkan juga berperan sebagai media informasi berbasis kecerdasan buatan.

Secara keseluruhan, penelitian ini berkontribusi dalam pengembangan sistem klasifikasi pola tekstil berbasis *deep learning* dengan pendekatan multimodal. Sistem yang dihasilkan mampu mengatasi sebagian tantangan dalam klasifikasi pola yang memiliki kemiripan visual tinggi. Di samping itu, integrasi dengan LLM memberikan kesempatan bagi perancangan sistem yang lebih interaktif dan informatif ke depannya. Penelitian ini juga dapat dijadikan acuan bagi pengembangan aplikasi sejenis pada domain tekstil maupun bidang lainnya yang memanfaatkan kombinasi *computer vision* dan *natural language processing*.

## 4. KESIMPULAN

Hasil penelitian ini membuktikan bahwa model *Convolutional Neural Network* (CNN) yang menggunakan arsitektur ResNet-50 dengan penerapan *transfer learning* dan *data augmentation* mampu mengklasifikasikan pola *crochet* secara otomatis dengan akurasi mencapai 93% serta performa yang stabil pada seluruh kelas, meskipun kesalahan masih muncul pada pola yang memiliki kemiripan visual tinggi. Temuan ini menegaskan efektivitas *deep learning* dalam menggantikan proses identifikasi manual yang sebelumnya memerlukan waktu lama dan rentan terhadap subjektivitas manusia. Kontribusi utama penelitian tidak semata-mata terletak pada peningkatan performa klasifikasi, namun juga meliputi implementasi model ke dalam platform berbasis web yang memfasilitasi pengguna dalam menjalankan klasifikasi citra secara otomatis sekaligus memperoleh *confidence score* pada setiap prediksi sehingga hasil menjadi lebih transparan, terukur, dan mudah dievaluasi. Selain itu, integrasi *Large Language Model* (LLM) memungkinkan sistem menghasilkan deskripsi pola secara otomatis, yang meningkatkan interpretabilitas hasil serta membantu pengguna memahami karakteristik pola yang terdeteksi tanpa memerlukan keahlian teknis khusus. Pendekatan ini membuktikan bahwa kombinasi CNN dan LLM dapat menghasilkan solusi klasifikasi yang akurat, informatif, dan aplikatif dalam konteks industri maupun edukasi tekstil. Namun, keterbatasan jumlah dataset serta kemiripan visual antarkelas masih menjadi tantangan utama, sehingga penelitian selanjutnya disarankan berfokus pada perluasan dataset, peningkatan variasi data, optimasi arsitektur model, serta pengujian pada skenario penggunaan nyata untuk meningkatkan akurasi dan kemampuan generalisasi sistem.

## REFERENCES

- [1] E. Ji, "Introduction to the Application of Hand Crochet in Tweed Clothing — An Example of the Design of the 'White Velvet Shiyun Shang'," *J. Theory Pract. Humanit. Sci.*, vol. 1, no. 5, pp. 13–19, 2024, [Online]. Available: <https://woodyinternational.com/index.php/jtphss/article/view/78>
- [2] H. Qu, Y. Zhou, P. Y. Mok, G. Flatz, and L. Li, "Efficient and Effective Detection of Repeated Pattern from Fronto-Parallel Images with Unknown Visual Contents," *Signals*, vol. 6, no. 1, p. 4, 2025, doi: 10.3390/signals6010004.
- [3] T. Mufizar, A. Sudiarjo, E. D. S. Mulyani, A. A. Wakih, M. A. Kasyfurrahman, and L. A. Mahbub, "Classification of Tasikmalaya batik motifs using convolutional neural networks," *IAES Int. J. Artif. Intell.*, vol. 14, no. 4, pp. 3287–3299, 2025, doi: 10.11591/ijai.v14.i4.pp3287-3299.
- [4] X. Zhao, L. Wang, Y. Zhang, X. Han, M. Deveci, and M. Parmar, "A review of convolutional neural networks in computer vision," *Artif. Intell. Rev.*, vol. 57, no. 4, p. 99, 2024, doi: 10.1007/s10462-024-10721-6.
- [5] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3, p. 52, 2023, doi: 10.3390/computation11030052.

- [6] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016. [Online]. Available: <http://www.deeplearningbook.org>
- [7] I. Hermawan, M. Agustin, and D. Arnaldy, "Rice Seedling Image Classification Using Light Convolutional Neural Network," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 9, no. 2, pp. 111–119, 2023, doi: 10.23917/khif.v9i2.20401.
- [8] P. Dabbo and F. Y. Bisilisin, "Klasifikasi Motif Kain Tenun Sabu Raijua Menggunakan Convolutional Neural Network (CNN) Berbasis Citra," *KETIK J. Inform.*, vol. 1, no. 6, pp. 11–18, 2024, doi: 10.70404/ketik.v1i06.88.
- [9] H. Herwingsyah and D. Yuswanto Jaya, "Accuracy potential of the Convolutional Neural Network (CNN) in recognizing traditional clothing," *IT J. Res. Dev.*, vol. 8, no. 2, pp. 95–106, 2023, doi: 10.25299/itjrd.2023.12690.
- [10] Y. Rizki, R. M. Taufiq, H. Mukhtar, and D. Putri, "Klasifikasi Pola Kain Tenun Melayu Menggunakan Faster R-CNN," *IT J. Res. Dev.*, vol. 5, no. 2, pp. 215–225, 2021, doi: 10.25299/itjrd.2021.vol5(2).5831.
- [11] S. S. Mohammed, H. G. Clarke, and S. N. Mahmood, "A fully convolutional neural network for fast detection, classification, and segmentation of fabric defects," *Neural Comput. Appl.*, vol. 37, no. 28, pp. 23249–23272, 2025, doi: 10.1007/s00521-025-11495-w.
- [12] L. Tan, Q. Fu, and J. Li, "An Improved Neural Network Model Based on DenseNet for Fabric Texture Recognition," *Sensors*, vol. 24, no. 23, p. 7758, 2024, doi: 10.3390/s24237758.
- [13] Z. Marciniak, P. Lertjaturaphat, and A. Bianchi, "StitchFlow: Enabling In-Situ Creative Explorations of Crochet Patterns with Stitch Tracking and Process Sharing," in *UIST 2025 - Proceedings of the 38th Annual ACM Symposium on User Interface Software and Technology*, Association for Computing Machinery, 2025. doi: 10.1145/3746059.3747715.
- [14] K. Seitz, P. Rein, J. Lincke, and R. Hirschfeld, "Digital Crochet: Toward a Visual Language for Pattern Description," in *Proceedings of the 2022 ACM SIGPLAN International Symposium on New Ideas, New Paradigms, and Reflections on Programming and Software*, Association for Computing Machinery, 2022, pp. 48–62. doi: 10.1145/3563835.3567657.
- [15] G. Singh, "Vision–Language Foundation Models and Multimodal Large Language Models: A Comprehensive Survey of Architectures, Benchmarks, and Open Challenges," 2026. doi: 10.20944/preprints202602.0467.v1.
- [16] Z. Li, X. Wu, H. Du, F. Liu, H. Nghiem, and G. Shi, "A Survey of State of the Art Large Vision Language Models: Alignment, Benchmark, Evaluations and Challenges," in *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2025, pp. 1578–1597. doi: 10.1109/CVPRW67362.2025.00147.
- [17] J. Li, D. Li, C. Xiong, and S. Hoi, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation," 2022, doi: 10.48550/arXiv.2201.12086.
- [18] Z. Gan, L. Li, C. Li, L. Wang, Z. Liu, and J. Gao, "Vision-Language Pre-Training: Basics, Recent Advances, and Future Trends," *Found. Trends Comput. Graph. Vis.*, vol. 14, no. 3–4, pp. 163–352, 2022, doi: 10.1561/0600000105.
- [19] J. Li, D. Li, S. Savarese, and S. Hoi, "BLIP-2: bootstrapping language-image pre-training with frozen image encoders and large language models," in *Proceedings of the 40th International Conference on Machine Learning*, Honolulu, Hawaii, USA: JMLR.org, 2023, pp. 19730–19742. [Online]. Available: <https://dl.acm.org/doi/10.5555/3618408.3619222>
- [20] V. R. Joseph, "Optimal ratio for data splitting," *Stat. Anal. Data Min. ASA Data Sci. J.*, vol. 15, no. 4, pp. 531–538, 2022, doi: 10.1002/sam.11583.
- [21] I. Wirabowo and I. Susilawati, "Implementasi Convolution Neural Network (CNN) untuk Deteksi Penyakit pada Daun Jagung Berbasis Citra Digital," *J. Pustaka Data*, vol. 5, no. 1, pp. 233–241, 2025, doi: 10.55382/jurnalpustakadata.v5i1.1046.
- [22] H. Imaduddin, W. Widayat, and M. S. Adhantoro, "Transfer Learning Architectures for Breast Cancer Detection in Mammographic Image Data," *J. Theor. Appl. Inf. Technol.*, vol. 103, no. 20, pp. 8336–8347, 2025.
- [23] K. Alomar, H. I. Aysel, and X. Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *J. Imaging*, vol. 9, no. 2, p. 46, 2023, doi: 10.3390/jimaging9020046.
- [24] M. H. Reza *et al.*, "A comprehensive review of convolutional neural networks: foundations, enhancements and applications," *Neural Comput. Appl.*, vol. 38, no. 4, p. 56, 2026, doi: 10.1007/s00521-025-11827-w.
- [25] A. Puspitasari, D. S. Salsabila, and D. Roliawati, "Penerapan ResNet50-CNN untuk Optimalisasi Klasifikasi pada Data Fashion," *Indones. J. Data Sci.*, vol. 3, no. 1, pp. 1–12, 2025, doi: 10.30989/ijds.v3i1.1533.
- [26] D. R. Fauzi and G. A. Haqdu D, "Comparison of CNN Models Using EfficientNetB0, MobileNetV2, and ResNet50 for Traffic Density with Transfer Learning," *J. Intell. Syst. Technol. Informatics*, vol. 1, no. 1, pp. 22–30, 2025, doi: 10.64878/jistics.v1i1.6.
- [27] H. Imaduddin, F. Y. A'la, A. Fatmawati, and B. A. Hermansyah, "Comparison of transfer learning method for COVID-19 detection using convolution neural network," *Bull. Electr. Eng. Informatics*, vol. 11, no. 2, pp. 1091–1099, 2022, doi: 10.11591/eei.v11i2.3525.
- [28] R. M. Schmidt, F. Schneider, and P. Hennig, "Descending through a Crowded Valley - Benchmarking Deep Learning Optimizers," in *Proceedings of the 38th International Conference on Machine Learning*, PMLR, 2021, pp. 9367–9376. [Online]. Available: <https://proceedings.mlr.press/v139/schmidt21a.html>
- [29] I. Kandel and M. Castelli, "The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset," *ICT Express*, vol. 6, no. 4, pp. 312–315, 2020, doi: 10.1016/j.ict.2020.04.010.
- [30] K. D. Reddy and A. Patil, "CXR-MultiTaskNet a unified deep learning framework for joint disease localization and classification in chest radiographs," *Sci. Rep.*, vol. 15, no. 1, p. 32022, 2025, doi: 10.1038/s41598-025-16669-z.
- [31] A. Zheng, *Evaluating Machine Learning Models*. O'Reilly Media, 2015.
- [32] I. D. Mienye, N. Jere, G. Obaido, O. O. Ogunraku, E. Esenogho, and C. Modisane, "Large language models: an overview of foundational architectures, recent trends, and a new taxonomy," *Discov. Appl. Sci.*, vol. 7, no. 9, p. 1027, 2025, doi: 10.1007/s42452-025-07668-w.
- [33] J. Bai *et al.*, "Qwen Technical Report," 2023. doi: 10.48550/arXiv.2309.16609.