

PENGENALAN POLA KAIN TENUN DI SULAWESI TENGGARA DENGAN MENGGUNAKAN HISTOGRAM OF ORIENTED GRADIENT (HOG), K-NEAREST NEIGHBOR (K-NN) DAN SYNTHETIC MINORITY OVER-SAMPLING TECHNIQUE (SMOTE)

Marsita Harim, Adha Mashur Sajiah, Rizal Adi Saputra

Program Teknik Informatika, Fakultas Teknik Informatika Universitas Halu Oleo
Jl Kampus Hijau Bumi Tridharma, Anduonohu, Kec. Kambu, Kota Kendari, Sulawesi Tenggara, Indonesia
marsitaharim.id@gmail.com

ABSTRAK

Provinsi Sulawesi Tenggara memiliki kekayaan kain tenun yang beragam, namun keterbatasan informasi menyebabkan sebagian masyarakat, khususnya generasi muda, kurang mengenal keunikan daerahnya. Penelitian ini menerapkan kecerdasan buatan dalam pengenalan pola kain tenun dengan tiga tahapan utama. Pertama, ekstraksi fitur dilakukan menggunakan *Histogram of Oriented Gradient* (HOG) untuk menangkap informasi bentuk dan tekstur dari citra kain. Kedua, proses klasifikasi dilakukan menggunakan algoritma *K-Nearest Neighbor* (K-NN) untuk mengidentifikasi asal daerah kain berdasarkan kemiripan fitur. Ketiga, untuk mengatasi ketidakseimbangan jumlah antar kelas, digunakan *Synthetic Minority Over-Sampling Technique* (SMOTE) pada data latih agar dapat menghasilkan sampel sintesis dari kelas minoritas. Tujuan dari penelitian ini adalah mengevaluasi efektivitas kombinasi metode tersebut dalam mengenali asal daerah kain tenun. Hasil pengujian menunjukkan bahwa sistem berhasil mengenali kain tenun dengan baik. Penerapan SMOTE meningkat *recall* pada kelas minoritas, seperti Bombana (dari 61% menjadi 83%), Kolaka (80% menjadi 90%), dan Muna (76% menjadi 91%). Meskipun akurasi keseluruhan menurun dari 78% menjadi 68%, penggunaan SMOTE menghasilkan distribusi prediksi yang lebih merata. Hasil penelitian ini menunjukkan bahwa kombinasi metode HOG, K-NN, dan SMOTE efektif dalam mengklasifikasikan pola kain tenun, serta telah berhasil diimplementasikan dalam sebuah sistem klasifikasi berbasis GUI yang dirancang agar adil dan representatif

Kata kunci : *Pengenalan Pola, Kain Tenun, Histogram of Oriented Gradient, K-Nearest Neighbor, Synthetic Minority Over-Sampling Technique*

1. PENDAHULUAN

Indonesia telah dikenal sebagai negara yang kaya akan seni dan budaya sejak berabad-abad lalu. Salah satu warisan budaya yang paling mencolok adalah seni kerajinan tenun. Seni tenun tidak hanya mencerminkan keterampilan teknis, tetapi juga menyimpan nilai estetika, identitas budaya, dan sejarah lokal yang kuat. Berbagai daerah di Indonesia memiliki motif dan corak khas yang mencerminkan pengaruh lokal maupun luar, seperti budaya Arab, India, dan Cina [1]. Kreativitas yang tinggi dari masyarakat Indonesia menjadikan kain tenun sebagai salah satu jenis kain paling berwarna dan bermotif di dunia.

Sulawesi Tenggara merupakan salah satu provinsi yang kaya akan ragam motif kain tenun, yang mencerminkan keunikan budaya daerah seperti Muna, Buton, dan Tolaki. Keanekaragaman ini, meskipun menjadi kekayaan budaya, menghadirkan tantangan tersendiri dalam hal pengenalan dan pelestariannya. Penduduk lokal, termasuk generasi muda, kerap kali tidak memiliki pemahaman yang cukup terhadap motif kain tenun di wilayahnya sendiri akibat keterbatasan akses informasi dan dokumentasi.

Seiring dengan kemajuan teknologi, khususnya di bidang kecerdasan buatan (*Artificial Intelligence*), solusi digital mulai diadopsi untuk membantu pelestarian budaya. Salah satu cabang AI, yakni pengenalan pola (*pattern recognition*), memungkinkan

sistem untuk mengidentifikasasi dan mengklasifikasikan motif tenun secara otomatis berdasarkan ciri-ciri visualnya. Melalui pendekatan ini, teknologi dapat menjadi jembatan yang menghubungkan warisan budaya dengan generasi masa kini dan mendatang, serta memperluas akses masyarakat terhadap informasi tenun secara cepat dan akurat.

Penelitian ini menggunakan metode *Histogram of Oriented Gradient* (HOG) untuk ekstraksi fitur, *K-Nearest Neighbor* (K-NN) sebagai metode klasifikasi, serta *Synthetic Minority Over-Sampling Technique* (SMOTE) untuk menangani ketidakseimbangan data. Bidang pengenalan pola ini memungkinkan sistem untuk memahami dan mengklasifikasikan motif-motif kain tenun secara otomatis berdasarkan karakteristik unik dari setiap daerah. Dengan demikian, teknologi ini dapat menjadi jembatan agar kain tenun yang ada di Sulawesi Tenggara dapat dikenali asal daerahnya oleh masyarakat.

2. TINJAUAN PUSTAKA

Dalam penelitian ini mengenai pengenalan motif kain tenun Sulawesi Tenggara, digunakan pendekatan kombinasi metode *Histogram of Oriented Gradient* (HOG), *K-Nearest Neighbor* (K-NN), dan *Synthetic Minority Over-Sampling Technique* (SMOTE). Pendekatan ini mengacu pada beberapa studi

sebelumnya yang telah membuktikan efektivitas K-NN dalam bidang klasifikasi, HOG sebagai metode ekstraksi fitur citra, serta SMOTE dalam menangani ketidak seimbangan data.

Beberapa penelitian terdahulu telah menunjukkan keberhasilan penggunaan kombinasi dua dari ketiga metode tersebut. Penelitian yang dilakukan oleh Nico Dian Nugraha dan tim pada tahun 2021 menggunakan HOG dan K-NN dalam klasifikasi citra dan mencapai akurasi tinggi sebesar 98,21%. Hal ini menunjukkan bahwa HOG mampu mengekstraksi ciri penting dari citra dengan baik, sementara K-NN dapat melakukan klasifikasi berdasarkan kedekatan fitur secara efektif [2]. Dalam penelitian lain yang dilakukan oleh A.G. Pertiwi dan rekan-rekannya mengkombinasikan SMOTE dan K-NN dalam proses klasifikasi untuk diagnosis penyakit diabetes pada data yang tidak seimbang. Hasilnya menunjukkan bahwa SMOTE mampu meningkatkan performa klasifikasi terhadap kelas minoritas [3].

2.1. Kain Tenun Tradisional Sulawesi Tenggara

Indonesia sebagai negara kepulauan memiliki kekayaan budaya tinggi, termasuk dalam seni kain tradisional seperti tenun dan batik [4]. Teknik menenun merupakan proses tradisional Menyusun benang menjadi kain berpola, dan Sulawesi Tenggara menjadi salah satu daerah penghasil kain tenun yang kaya akan motif khas. Di wilayah ini, tradisi menenun diyakini berasal dari Buton dan telah ada sejak abad ke-14, diwariskan secara turun-menurun. Kerajinan tenun Buton sering kali menggambarkan objek-objek alam sekitar, yang terwujud dalam corak dan perpaduan warna yang khas [5]. Keahlian menenun dari Buton kemudian menyebar ke daerah sekitarnya seperti Kendari, Muna, dan Konawe. Di Kendari dan Konawe, motif tenun dipengaruhi pula oleh budaya suku Bugis, sehingga memiliki kemiripan dengan motif tenun Bugis. Berdasarkan data Dinas Kebudayaan dan Pariwisata Kota Baubau, terdapat 48 motif tenun Buton yang telah memiliki Hak Atas Kekayaan Intelektual (HAKI), serta 4 motif sarung tenun dengan HAKI personal. Selain itu, daerah lain seperti Muna, Bombana, Kolaka, dan Tolaki juga memiliki motif tenun khas masing-masing.

2.2. Pengenalan Pola

Pengenalan pola merupakan proses yang dirancang untuk mengidentifikasi objek, dalam hal ini dapat berupa citra digital, berdasarkan karakteristik yang dimilikinya [6]. Dengan kata lain, pengenalan pola mencakup aktivitas klasifikasi atau representasi kuantitatif dari fitur khas suatu objek. Tujuan utamanya adalah untuk mengelompokkan atau mengkategorikan pola-pola berdasarkan atribut unik masing-masing, sehingga perbedaan antar objek dapat dikenali secara sistematis [7]. Dengan demikian, pengenalan pola tidak hanya terbatas pada proses identifikasi, tetapi juga melibatkan analisis terhadap fitur-fitur pembeda yang melekat pada objek tersebut.

2.3. Histogram of Oriented Gradient (HOG)

Ekstraksi fitur bertujuan untuk mengubah informasi dari sebuah gambar ke dalam bentuk representasi numerik berupa vektor, yang dikenal sebagai vektor fitur [8]. Salah satu metode yang umum digunakan dalam proses ekstraksi fitur adalah *Histogram of Oriented Gradient* (HOG), yang dalam berbagai studi menunjukkan kinerja yang baik dibandingkan dengan teknik ekstraksi fitur lainnya. Metode HOG bekerja berdasarkan asumsi bahwa bentuk dan struktur objek lokal dapat diidentifikasi melalui distribusi arah gradien intensitas atau kontur, tanpa memerlukan informasi posisi gradien secara eksplisit. HOG menyusun deskripsi citra berdasarkan histogram orientasi gradien yang dihitung pada area-area lokal, yang merepresentasikan frekuensi arah gradien tertentu dalam setiap wilayah gambar. Prosedur ekstraksi fitur menggunakan HOG terdiri dari beberapa tahapan utama

Adapun langkah-langkah utama dalam proses ekstraksi fitur dengan menggunakan HOG sebagai berikut.

a. Pre-processing citra

Tahap awal melibatkan konversi citra RGB menjadi *grayscale* untuk menyederhanakan informasi visual. Setelah itu, citra diubah ukurannya menjadi 128x128 piksel agar seragam dalam proses ekstraksi fitur

b. Perhitungan Gradien

Gradien citra dihitung menggunakan operator Sobel untuk memperoleh arah dan besar perubahan intensitas piksel. Perhitungan gradien horizontal dan vertikal dijelaskan pada Persamaan (1) dan (2).

$$g_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \times I \quad (1)$$

$$g_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} \times I \quad (2)$$

Keterangan :

g_x = Gradien horizontal.

g_y = Gradien vertikal.

I = Matriks intensitas piksel citra input

c. Magnitude dan Orientation Gradien

Setelah gradien horizontal dan vertikal diperoleh, tahap selanjutnya adalah menghitung *magnitude* dan *orientation* gradien pada setiap piksel. Perhitungan ini dijelaskan pada Persamaan (3) dan (4)

$$M = \sqrt{g_x^2 + g_y^2} \quad (3)$$

$$\theta = \left(\arctan 2 \frac{g_y}{g_x} \right) \times \frac{180}{\pi} \quad (4)$$

Keterangan

M = Gradient magnitude.

g_x = Operasi matriks horizontal.

g_y = Operasi matriks vertikal.

θ = Gradient orientation

Selanjutnya dilakukan proses *polarity change* untuk menghasilkan *descriptor* yang *invariant* terhadap perubahan antar *foreground* dan *background*. Proses ini mengubah nilai orientasi menjadi positif, sebagaimana dijelaskan pada Persamaan (5).

$$\theta'(x, y) = \begin{cases} \theta(x, y) & ; \theta(x, y) \geq 0 \\ \theta(x, y) + 180 & ; \theta(x, y) < 0 \end{cases} \quad (5)$$

d. Histogram Gradien per Cell

Setelah nilai *magnitude* dan *orientation* diperoleh, langkah berikutnya adalah menghitung histogram orientasi gradien untuk setiap *cell* berukuran 8x8 piksel. Setiap piksel memberikan kontribusi terhadap histogram berdasarkan nilai *magnitude*-nya. Distribusi ini dihitung dengan metode *interpolasi bilinear*, yang memungkinkan setiap piksel memengaruhi dua *bin* terdekat secara proporsional.

Sudut orientasi (0° - 180°) dibagi menjadi sejumlah *bin*, umumnya 9 *bin* dengan rentang 20° per *bin*. Nilai *magnitude* dari setiap piksel didistribusikan ke dua *bin* terdekat berdasarkan jarak relatifnya terhadap masing-masing *bin*. Proses ini dirumuskan dalam Persamaan (6) dan (7).

$$Vote_1 = \left(1 - \frac{\Delta R_i}{H}\right) \times M_i \quad (6)$$

$$Vote_2 = \frac{\Delta R_i}{H} \times M_i \quad (7)$$

Pendekatan ini memastikan bahwa informasi arah gradien tidak hanya bergantung pada satu *bin*, melainkan tersebar secara lebih halus dan akurat, sehingga menghasilkan deskriptor yang lebih representatif.

e. Normalisasi

Normalisasi dilakukan untuk mengatasi sensitivitas nilai gradien terhadap variasi pencahayaan global. Jika intensitas gambar berubah, misalnya menjadi lebih gelap, maka *magnitude* gradien dan nilai histogram juga akan berubah secara proporsional. Oleh karena itu, diperlukan normalisasi agar representasi fitur menjadi invarian terhadap pencahayaan.

Proses ini dilakukan dengan menghitung norma dari vektor histogram (biasanya normal L_2), sebagaimana ditunjukkan pada Persamaan 2.8 berikut.

$$Norma L_2 = \sqrt{(V_0)^2 + (V_1)^2 + \dots + (V_n)^2} \quad (8)$$

Selanjutnya setiap nilai pada histogram dinormalisasi terhadap nilai norma tersebut menggunakan formula pada Persamaan 2.9 berikut.

$$Normalisasi = \frac{\text{Vektor Histogram}}{\text{Norma } L_2} \quad (9)$$

Pada gambar berukuran 128x128 piksel, citra dibagi menjadi *cell* berukuran 8x8 piksel sehingga terbentuk 16x16 *cell*. Blok dibentuk dari 2x2 *cell* dengan 50% *overlapping*, menghasilkan 15x15=225 blok. Setiap blok memuat 36 fitur (9 orientasi x 4 *cell*), sehingga total diperoleh 8.100 fitur untuk satu gambar.

2.4. K-Nearest Neighbor (K-NN)

K-Nearest Neighbor (K-NN) merupakan salah satu algoritma *machine learning* yang bersifat non-parametrik dan digunakan untuk melakukan klasifikasi berdasarkan kedekatan data uji dengan data pelatihan. Pada proses klasifikasinya, K-NN mencari sejumlah k tetangga terdekat dari data uji, lalu menentukan kelas berdasarkan mayoritas label dari tetangga tersebut [3]. Algoritma ini tidak melakukan proses pelatihan secara eksplisit, melainkan menyimpan seluruh data latih dalam bentuk vektor fitur beserta labelnya. Identifikasi kelas dilakukan dengan menghitung kedekatan antar data, yang umumnya menggunakan jarak *Euclidean* sebagai metrik pengukuran [9].

K-NN bekerja berdasarkan asumsi bahwa data yang memiliki karakteristik serupa akan berada dalam kelas yang sama. Oleh karena itu, pemilihan nilai k yang tepat serta metrik jarak yang digunakan menjadi faktor penting yang memengaruhi performa klasifikasi. Adapun langkah-langkah penerapan algoritma K-NN dalam menentukan kelas suatu objek dapat dijelaskan sebagai berikut:

a. Menentukan Nilai Jumlah Tetangga (k)

Langkah pertama adalah menentukan jumlah tetangga terdekat (k) yang akan digunakan untuk klasifikasi. Nilai k menentukan jumlah data yang dijadikan acuan dalam memprediksi kelas data uji. Pemilihan nilai k ganjil dianjurkan untuk menghindari hasil klasifikasi yang imbang.

b. Menghitung Jarak antara Data Uji dan Data Latih

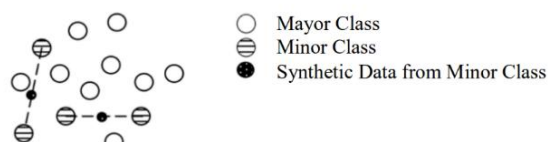
Setelah nilai k ditentukan, dilakukan penghitungan jarak antara data uji dan seluruh data latih. Dalam kasus citra, perhitungan dilakukan setelah proses ekstraksi fitur. Jarak dihitung menggunakan metode *Euclidean distance*, yang mengukur perbedaan antar fitur. Rumus perhitungan ditunjukkan pada Persamaan 10.

$$d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (10)$$

- c. Mengurutkan Semua Objek
Setelah jarak antara data uji dan seluruh data latih dihitung, langkah selanjutnya adalah mengurutkan data tersebut berdasarkan jarak terkecil hingga terbesar. Pengurutan ini bertujuan untuk memudahkan identifikasi k tetangga terdekat.
- d. Mengidentifikasi Kategori Tetangga Terdekat
Dari hasil pengurutan, diambil k data dengan jarak terpendek. Kategori atau label dari masing-masing tetangga ini dicatat untuk menentukan kecenderungan klasifikasi.
- e. Memprediksi Kelas Berdasarkan Mayoritas Tetangga
Langkah terakhir adalah menentukan kelas data uji berdasarkan kategori yang paling sering muncul di antara k tetangga tersebut. Kelas dengan jumlah terbanyak dipilih sebagai hasil prediksi, sehingga Keputusan klasifikasi didasarkan pada informasi terbanyak dari tetangga terdekat.

2.5. Synthetic Minority Over-Sampling Technique (SMOTE)

Synthetic Minority Over-Sampling Technique (SMOTE) adalah metode yang dikembangkan oleh Chawla untuk mengatasi ketidakseimbangan kelas dalam data. Berbeda dengan teknik *oversampling* konvensional yang menduplikasi data minoritas secara acak, SMOTE menghasilkan sampel sintetis dengan melakukan interpolasi antara data minoritas dan *k-nearest neighborhood*-nya yang dipilih secara acak. Teknik ini mampu memperluas representasi data minoritas dalam ruang fitur tanpa mengulang data yang sama, sehingga lebih efektif dalam mendukung kinerja model [3]. Ilustrasi hasil distribusi data setelah penerapan SMOTE ditampilkan pada Gambar 1.



Gambar 1. Ilustrasi SMOTE

Selain menghasilkan data sintetis, SMOTE juga dapat digabungkan dengan teknik *under-sampling* pada kelas mayoritas. Kombinasi ini bertujuan untuk menciptakan distribusi data yang lebih seimbang, yang pada akhirnya memungkinkan model pembelajaran membentuk batas keputusan yang lebih general. Dengan cara ini, risiko *overfitting* terhadap data minoritas dapat dikurangi, dan performa klasifikasi pada kelas yang sebelumnya terpinggirkan pun dapat meningkat [10].

3. METODE PENELITIAN

Penelitian ini bertujuan untuk mengembangkan sistem klasifikasi citra kain tenun berdasarkan

karakteristik visualnya. Metodologi yang digunakan mencakup beberapa tahapan utama, yaitu pengumpulan dan persiapan data, ekstraksi fitur dari citra menggunakan *Histogram of Oriented Gradient* (HOG), penyeimbangan data menggunakan *Synthetic Minority Over-Sampling Technique* (SMOTE), serta klasifikasi menggunakan algoritma *K-Nearest Neighbor* (K-NN). Setiap tahapan dilakukan secara berurutan guna memastikan kualitas data dan akurasi model dalam mengenali pola kain tenun secara efektif

3.1. Pengumpulan Data

Data citra kain tenun diperoleh melalui pengambilan gambar menggunakan kamera ponsel serta dokumen digital yang memuat koleksi citra kain. Proses pengumpulan dilakukan di tiga lokasi utama: Kecamatan Sulaa Kota Baubau, Dinas Pariwisata dan Kebudayaan Kota Baubau, serta Dewan Kerajinan Nasional Daerah (Dekranasda) Provinsi Sulawesi Tenggara. Selain itu, sebagian data juga dihimpun secara langsung dari masyarakat di Kota Kendari.

Untuk memberikan gambaran visual terhadap data yang digunakan, berikut ditampilkan beberapa citra kain tenun hasil akuisisi dari berbagai daerah di Sulawesi Tenggara. Citra-citra ini mewakili lima kelas utama daerah asal, yaitu Muna, Bombana, Kolaka, Buton, dan Tolaki, yang masing-masing memiliki motif serta karakteristik visual yang khas. Beberapa citra hasil akuisisi ditampilkan pada Gambar 2 hingga Gambar 6. Masing-masing gambar mewakili kain tenun dari 5 daerah berbeda di Sulawesi Tenggara yang digunakan dalam penelitian ini.



Gambar 2. Citra kain tenun dari daerah Bombana



Gambar 3. Citra kain tenun dari daerah Buton



Gambar 4. Citra kain tenun dari daerah Kolaka



Gambar 5. Citra kain tenun dari daerah Muna



Gambar 6. Citra kain tenun dari daerah Tolaki

3.2. Persiapan Dataset: Pre-processing dan Augmentasi Data

Sebelum data digunakan dalam pelatihan model, dilakukan tahapan *preprocessing* dan augmentasi guna memastikan kualitas serta keragaman data yang memadai. Proses ini bertujuan untuk menyeragamkan ukuran gambar dan memperkaya variasi visual, sehingga dapat meningkatkan kemampuan model dalam mengenali pola.

- a. *Pre-processing* dilakukan dengan mengubah ukuran seluruh gambar menjadi 128x128 piksel. Penyeragaman ukuran ini penting untuk memastikan bahwa *input* yang diterima oleh sistem bersifat konsisten dan tidak menimbulkan perbedaan hasil ekstraksi fitur akibat perbedaan resolusi. Gambar 7 memperlihatkan hasil citra setelah dilakukan *pre-processing*.

Gambar 7. Citra hasil *pre-processing* (resize 128x128 piksel)

- b. Augmentasi data dilakukan untuk memperkaya variasi citra seperti *flipping* horizontal dan vertikal, rotasi pada sudut 45°, 90°, 180°, 135° hingga 270°, penambahan *noise gaussian* dan *laplace*, serta penyesuaian kontras menggunakan metode *sigmoid* dan *linear*. Variasi augmentasi ini memungkinkan model mempelajari pola dari berbagai kondisi pencahayaan, orientasi, dan gangguan visual. Hasil augmentasi ditampilkan pada Gambar 8, yang memperlihatkan transformasi citra menggunakan teknik rotasi sebesar 45°, sebagai salah satu metode augmentasi yang diterapkan dalam penelitian ini.



Gambar 8. Citra hasil augmentasi dengan rotasi 45°

3.3. Pembagian Data

Dalam penelitian ini, pembagian *dataset* dilakukan dengan rasio 70:30, dimana 70% data digunakan sebagai data pelatihan (*training set*) dan 30% sisanya sebagai data pengujian (*testing set*). Total jumlah data yang digunakan adalah 5.936 citra kain tenun, sehingga diperoleh 4.095 citra untuk pelatihan dan 1.833 citra untuk pengujian. Pembagian ini bertujuan untuk menguji performa algoritma *K-Nearest Neighbor* (K-NN) dalam mengklasifikasikan citra yang belum pernah dilihat sebelumnya.

3.4. Pengenalan Gambar

Pengenalan gambar merupakan proses dimana sistem mengidentifikasi dan menganalisis citra berdasarkan model yang telah dilatih sebelumnya. Dalam penelitian ini, proses pengenalan dilakukan menggunakan algoritma *K-Nearest Neighbor* (K-NN), dengan fitur citra yang diekstraksi melalui metode *Histogram of Oriented Gradient* (HOG).

3.4.1. Histogram of Oriented Gradient (HOG)

Pada penelitian ini, *Histogram of Oriented Gradient* (HOG) digunakan sebagai metode ekstraksi fitur untuk mengidentifikasi pola visual dari citra kain tenun. Teknik ini dipilih karena mampu menangkap informasi structural seperti bentuk dan arah garis yang menjadi ciri khas motif kain. Setiap citra yang telah melalui tahap *preprocessing* diproses dengan HOG untuk menghasilkan vektor fitur yang merepresentasikan pola-pola motif secara numerik. Representasi ini kemudian digunakan sebagai *input* dalam proses klasifikasi untuk membedakan motif berdasarkan daerah asalnya.

3.4.2. K-Nearest Neighbor (K-NN)

Pada penelitian ini, algoritma *K-Nearest Neighbor* (K-NN) digunakan sebagai metode klasifikasi untuk mengenali motif kain tenun berdasarkan fitur yang telah diekstraksi sebelumnya menggunakan HOG. K-NN bekerja dengan mengklasifikasikan data uji berdasarkan kedekatan jaraknya dengan data pelatihan dalam ruang fitur. Semakin dekat suatu citra dengan tetangganya dalam data pelatihan, semakin besar kemungkinan citra tersebut termasuk ke dalam kelas yang sama. Pemilihan nilai k dilakukan dengan mempertimbangkan performa model pada data validasi untuk memperoleh hasil klasifikasi yang optimal.

3.5. Penyeimbangan Data dengan SMOTE

Ketidakseimbangan distribusi data antar kelas merupakan salah satu tantangan umum dalam pengembangan model klasifikasi. Kondisi ini dapat menyebabkan model lebih cenderung mengenali pola dari kelas mayoritas, sehingga akurasi untuk kelas minoritas menjadi rendah. Dalam penelitian ini, permasalahan tersebut diatasi menggunakan teknik *Synthetic Minority Over-sampling Technique* (SMOTE), yang berfungsi untuk menghasilkan sampel sintetis pada kelas minoritas berdasarkan kedekatan jarak antar fitur. Teknik ini tidak hanya menambah jumlah data minoritas, tetapi juga memperluas keragaman data dalam kelas tersebut.

SMOTE diterapkan setelah proses ekstraksi fitur menggunakan *Histogram of Oriented Gradient* (HOG), dan hanya digunakan pada data latih. Data uji tetap mempertahankan data asli untuk memastikan hasil evaluasi model mencerminkan performa yang realistis terhadap data baru. Dengan menyeimbangkan data latih, model *K-Nearest Neighbor* (K-NN) dapat mempelajari karakteristik dari setiap kelas secara lebih adil, sehingga meningkatkan kemampuan prediksi terutama untuk kelas-kelas dengan representasi awal yang rendah.

3.6. Implementasi Sistem Berbasis GUI

Kemudahan penggunaan sistem menjadi salah satu aspek yang diperhatikan dalam penelitian ini. Antarmuka pengguna (*Graphical User Interface/GUI*)

dirancang menggunakan pustaka PyQt5 sebagai bagian dari sistem pengenalan kain tenun yang dikembangkan. Melalui antarmuka ini, pengguna dapat menjalankan seluruh proses pengenalan secara interaktif. Kehadiran GUI ini diharapkan dapat mempermudah penggunaan sistem pada perangkat komputer oleh pengguna non-teknis, serta mendukung penerapan sistem dalam lingkungan kerja atau penelitian secara langsung. Tampilan antarmuka sistem dapat dilihat pada Gambar 9 dibawah ini.



Gambar 9. Tampilan antarmuka sistem pengenalan kain tenun berbasis GUI

4. HASIL DAN PEMBAHASAN

Hasil evaluasi model klasifikasi kain tenun disajikan berdasarkan penerapan metode *Histogram of Oriented Gradient* (HOG) untuk ekstraksi fitur dan algoritma *K-Nearest Neighbor* (K-NN) sebagai klasifier. Proses evaluasi dilakukan pada dua scenario, yaitu tanpa dan dengan penerapan *Synthetic Minority Oversampling Technique* (SMOTE) untuk menangani ketidakseimbangan kelas.

Setiap model dievaluasi menggunakan metrik *accuracy*, *precision*, *recall*, dan *f1-score*. Pembahasan difokuskan pada perbandingan kinerja model dalam kedua scenario serta analisis performa terhadap masing-masing kelas untuk memahami dampak keseimbangan data terhadap kemampuan klasifikasi model.

4.1. Evaluasi Model Tanpa SMOTE

Evaluasi awal dilakukan terhadap model K-NN yang dilatih menggunakan data asli tanpa penerapan teknik penyeimbangan. Untuk melihat performa klasifikasi model terhadap masing-masing kelas, disusun tabel evaluasi yang memuat nilai *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) untuk setiap kelas, sebagaimana ditampilkan pada Tabel 1. Nilai-nilai menjadi dasar dalam perhitungan metrik evaluasi seperti *precision*, *recall*, dan *f1-score*, yang akan digunakan untuk menilai konsistensi performa model terhadap setiap kelas.

Tabel 1. Komponen *confusion matrix* K-NN

Class	TP	FP	FN	TN
Bombana	110	14	69	1.640
Buton	857	162	156	658
Kolaka	175	29	43	1.586
Muna	215	198	67	1.353
Tolaki	68	5	73	1.687

Berdasarkan nilai TP, FP, FP, dan FN pada Tabel 1, dilakukan perhitungan metrik evaluasi berupa *precision*, *recall*, dan *f1-score* untuk masing-masing kelas. Ketiga metrik ini digunakan untuk memberikan penilaian yang lebih mendalam terhadap performa model dalam mengenali setiap kelas secara akurat dan konsisten. Hasil perhitungan tersebut disajikan dalam Tabel 2, yang juga mencantumkan nilai akurasi keseluruhan model K-NN tanpa penerapan SMOTE. Tabel 2 berikut ini menyajikan Gambaran menyeluruh mengenai kemampuan klasifikasi model pada setiap kelas.

Tabel 2. Metrik evaluasi K-NN

Class	Precision	Recall	F1-score
Bombana	89%	61%	73%
Buton	84%	85%	84%
Kolaka	86%	80%	83%
Muna	52%	76%	62%
Tolaki	93%	48%	64%
Accuracy	78%		

Berdasarkan Tabel 2, model K-NN tanpa penerapan SMOTE memperoleh akurasi sebesar 78%. Performa model tergolong baik pada kelas Kolaka dan Buton, yang masing-masing menunjukkan nilai *precision*, *recall*, dan *f1-score* diatas 80%. Sebaliknya, performa klasifikasi terhadap kelas Muna dan Tolaki kurang optimal, yang terlihat dari rendahnya *precision* pada Muna (52%) dan *recall* pada Tolaki (48%). Hasil ini mencerminkan adanya ketidakseimbangan data yang berdampak pada kemampuan model dalam mengenali semua kelas secara merata.

4.2. Evaluasi Model dengan SMOTE

Evaluasi selanjutnya dilakukan terhadap model K-NN yang dilatih menggunakan data yang telah diseimbangkan dengan teknik SMOTE (*Synthetic Minority Over-sampling Technique*). Tujuan dari penerapan SMOTE adalah untuk menyeimbangkan jumlah data di setiap kelas dengan menambahkan sampel buatan pada kelas yang datanya lebih sedikit. Dengan distribusi yang lebih seimbang, diharapkan model dapat mengenali pola dari seluruh kelas secara lebih adil dan konsisten.

Untuk menilai performa klasifikasi setelah penerapan SMOTE, disusun *confusion matrix* dalam bentuk tabel sebagaimana ditampilkan pada Tabel 3. Tabel ini memuat jumlah prediksi benar maupun salah melalui *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN) untuk masing-masing kelas, dan menjadi dasar dalam perhitungan metrik evaluasi pada tahap selanjutnya.

Tabel 3. Komponen *confusion matrix* K-NN

Class	TP	FP	FN	TN
Bombana	149	88	30	1.733
Buton	559	12	454	975
Kolaka	196	89	22	1.693
Muna	258	341	24	1.377
Tolaki	91	50	57	1.811

Berdasarkan nilai TP, FP, TN, dan FN pada Tabel 3, dilakukan perhitungan metrik evaluasi berupa *precision*, *recall*, dan *f1-score* untuk masing-masing kelas. Metrik-metrik ini digunakan untuk mengevaluasi sejauh mana model mampu mengklasifikasikan setiap kelas dengan akurat dan seimbang setelah data pelatihan diseimbangkan menggunakan SMOTE. Hasil perhitungan tersebut dapat dilihat pada Tabel 4, yang juga menampilkan nilai akurasi keseluruhan dari model K-NN setelah penerapan SMOTE.

Tabel 4. Metrik evaluasi K-NN+SMOTE

Class	Precision	Recall	F1-score
Bombana	63%	83%	72%
Buton	98%	55%	71%
Kolaka	69%	90%	78%
Muna	43%	91%	59%
Tolaki	65%	65%	65%
Accuracy	68%		

Setelah penerapan SMOTE, performa model menunjukkan perubahan yang signifikan. Meskipun akurasi keseluruhan menurun menjadi 68%, peningkatan *recall* pada sebagian besar kelas seperti kelas Muna dan Kolaka yang mencapai 91% dan 90% yang menunjukkan bahwa model menjadi lebih sensitive dalam mengenali kelas-kelas yang sebelumnya kurang terdeteksi. Hal ini mencerminkan efektivitas SMOTE dalam memperbaiki representasi kelas minoritas. Meski terdapat penurunan *precision* pada beberapa kelas akibat meningkatnya *false positive*, secara umum penerapan SMOTE berhasil meningkatkan keseimbangan performa klasifikasi antar kelas dan memberikan pemahaman yang lebih merata terhadap pola data.

4.3. Analisis Pengaruh SMOTE terhadap Performa Model

Model K-NN dievaluasi menggunakan dua scenario yaitu sebelum dan sesudah penerapan SMOTE, untuk mengamati pengaruhnya terhadap performa klasifikasi. Sebelum SMOTE diterapkan, *dataset* menunjukkan ketidakseimbangan kelas yang cukup signifikan, dimana kelas mayoritas seperti Buton memiliki jumlah citra jauh lebih banyak dibandingkan kelas minoritas. Setelah SMOTE diterapkan pada data pelatihan, distribusi antar kelas menjadi seimbang, dengan masing-masing kelas memiliki 2.241 citra.

Penerapan SMOTE memberikan dampak yang cukup mencolok terhadap performa model. Meskipun akurasi menurun dari 78% menjadi 68%, *recall*

meningkat secara signifikan pada kelas-kelas minoritas. Misalnya, *recall* kelas Bombana meningkat dari 61% menjadi 83%, dan Muna dari 76% menjadi 91% menunjukkan peningkatan sensitivitas model terhadap kelas yang sebelumnya sulit dikenali.

Namun, peningkatan *recall* tersebut disertai penurunan *precision* pada sebagian besar kelas, seperti Bombana (89% → 63%) dan Tolaki (93% → 65%). Selain itu, kelas mayoritas seperti Buton mengalami penurunan *recall* dari 85% menjadi 55%, meskipun *precision*-nya meningkat dari 84% menjadi 98%. Temuan ini mengindikasikan adanya redistribusi prediksi akibat keseimbangan baru dalam data pelatihan.

Secara keseluruhan, penerapan SMOTE membantu memperbaiki kemampuan model dalam mengenali kelas-kelas minoritas dan meningkatkan keseimbangan prediksi antar kelas. Namun, hal ini datang dengan *trade-off* berupa penurunan ketepatan klasifikasi dan akurasi secara umum. Temuan ini menunjukkan bahwa penggunaan SMOTE efektif untuk meningkatkan sensitivitas model, tetapi perlu diimbangi dengan pertimbangan potensi *overfitting* dan efek terhadap kelas mayoritas.

5. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengembangkan sistem pengenalan kain tenun Sulawesi Tenggara dengan mengombinasikan metode *Histogram of Oriented Gradient* (HOG), *K-Nearest Neighbor* (K-NN), dan *Synthetic Minority Over-Sampling Technique* (SMOTE). HOG efektif dalam mengekstraksi fitur bentuk, pola, dan tekstur dari citra kain, yang berperan penting dalam proses klasifikasi. K-NN menunjukkan performa yang cukup baik, dan penerapan SMOTE berhasil meningkatkan *recall* pada kelas minoritas seperti Bombana (61% menjadi 83%), Kolaka (80% menjadi 90%), dan Muna (76% menjadi 91%). Namun, peningkatan ini disertai penurunan *precision* pada beberapa kelas, serta penurunan akurasi keseluruhan dari 78% menjadi 68%, yang menunjukkan perlunya evaluasi menyeluruh terhadap performa model. Untuk pengembangan selanjutnya, disarankan mencoba metode *balancing data* lain seperti ADASYN atau *Borderline-SMOTE*, serta augmentasi yang lebih kompleks seperti GAN atau teknik augmentasi berbasis pola kain. Sistem ini juga berpotensi untuk dikembangkan menjadi aplikasi *mobile* agar lebih mudah diakses oleh pengguna secara langsung.

DAFTAR PUSTAKA

- [1] N. Indriani, "Perkembangan Industri Sarung Tenun pada Masyarakat di Kelurahan Sulaa Kecamatan Betoambari Kota Baubau," *Jurnal Penelitian Pendidikan Sejarah*, vol. 5, no. 1, pp. 57–64, 2020, [Online]. Available: http://ojs.uho.ac.id/index.php/p_sejarah_uho
- [2] N. D. Nugraha, F. Utaminingrum, and H. Fitriyah, "Alat Pendeteksi Uang untuk Tunanetra Menggunakan Metode Histogram of Oriented Gradients dan K-Nearest Neighbor," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 4, pp. 1569–1577, 2021, [Online]. Available: <http://j-ptiik.ub.ac.id>
- [3] A. G. Pertiwi, N. Bachtiar, R. Kusumaningrum, I. Waspada, and A. Wibowo, "Comparison of Performance of K-Nearest Neighbor Algorithm Using Smote and K-Nearest Neighbor Algorithm Without Smote in Diagnosis of Diabetes Disease in Balanced Data," *J Phys Conf Ser*, vol. 1524, no. 1, pp. 1–8, Jun. 2020, doi: 10.1088/1742-6596/1524/1/012048.
- [4] I. A. Winarno, "Analisis Motif Kain Tradisional Indonesia: Pemaknaan Visualisasi Abstrak hingga Naturalis," *Jurnal Pengabdian Seni*, vol. 1, no. 1, pp. 89–102, 2022.
- [5] A. Slamet, "Corak Motif Flora Sarung Tenun Buton Sebagai Pembelajaran Berbasis Lingkungan," in *Seminar Nasional Pendidikan Biologi dan Saintek II*, 2017, pp. 571–577.
- [6] A. Septiarini, "Pengenalan Pola Pada Citra Digital Dengan Fitur Momen Invariant," *Jurnal Informatika Mulawarman*, vol. 7, no. 1, pp. 8–11, 2012.
- [7] J. W. Yodha and A. W. Kurniawan, "Perbandingan Penggunaan Deteksi Tepi dengan Metode Laplace, Sobel dan Prewitt dan Canny pada Pengenalan Pola," *Techno.COM*, vol. 13, no. 3, pp. 189–197, 2014.
- [8] N. Hidayat, A. Tjahyanto, X. Z. Loviano, M. Yunus, and D. Adzkiya, "Individual Recognition Using Histogram of Oriented Gradients," in *The 2018 International Conference on Information Technology, Engineering, Science, and its Applications (ITES 2018)*, 2018, pp. 190–194. [Online]. Available: <https://ssrn.com/abstract=3268935Electroniccopyavailableat:https://ssrn.com/abstract=3248449;http://ssrn.com/link/ITES-2018.html>
- [9] M. H. Ramdani, I. G. P. S. Wijaya, and R. Dwiyanaputra, "Optimalisasi Pengenalan Wajah Berbasis Linear Discriminant Analysis dan K-Nearest Neighbor Menggunakan Particle Swarm Optimization," *Jurnal Teknologi Informasi, Komputer dan Aplikasinya (JTika)*, vol. 4, no. 1, pp. 40–51, 2022, [Online]. Available: <http://jtika.if.unram.ac.id/index.php/JTika/>
- [10] F. Y. Pamuji and S. D. A. Putri, "Komparasi Metode Smote dan Adasyn untuk Penanganan Data Tidak Seimbang Multiclass," *JIP (Jurnal Informatika Polinema)*, vol. 9, no. 3, pp. 331–338, 2023.