



JOURNAL CERITA:

**Creative Education of Research in Information
Technology and Artificial Informatics**

Vol. 12 No. 1 (2026) 99-108

e-ISSN: 2655 - 2574

Penerapan Swin Transformer Pada Klasifikasi Bentuk Wajah Untuk Rekomendasi Bentuk Kacamata

Alif Maulana¹, Henni Endah Wahanani², Muhammad Muharrom Al Haromainy³

^{1,2,3}Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional Veteran Jawa
Timur, Surabaya, Indonesia

Email: alifmaul687@gmail.com^{*1}; henniendah@upnjatim.ac.id²; muhammad.muharrom.if@upnjatim.ac.id³

Maulana, A., Wahanani, H. E., & Al Haromainy, M. M. (2026). Penerapan Swin Transformer Pada Klasifikasi Bentuk Wajah Untuk Rekomendasi Bentuk Kacamata. *Journal Cerita: Creative Education of Research in Information Technology and Artificial Informatics*, 12(1), 99-108

DOI: <https://doi.org/10.33050/2x8m4373>

ABSTRAK

Bentuk wajah berperan sebagai penanda identitas yang mencerminkan karakteristik fisik unik pada setiap individu. Secara umum, bentuk wajah dikategorikan ke dalam lima jenis utama, yaitu square, round, oval, heart, dan oblong. Informasi mengenai bentuk wajah dapat dimanfaatkan dalam berbagai aspek personalisasi, termasuk pemilihan model kacamata yang sesuai. Seiring berkembangnya industri fashion, kacamata tidak lagi hanya berfungsi sebagai alat bantu penglihatan. Kacamata dapat berperan sebagai elemen estetika yang mendukung penampilan. Oleh karena itu, pemilihan model kacamata berdasarkan bentuk wajah menjadi faktor penting dalam meningkatkan daya tarik visual individu. Pada jual beli daring pelanggan sering mengalami kesulitan dalam menentukan model kacamata yang sesuai, sehingga diperlukan sistem otomatis berbasis kecerdasan buatan untuk memberikan rekomendasi yang optimal. Penelitian ini mengimplementasikan model Swin Transformer untuk melakukan klasifikasi bentuk wajah sebagai dasar rekomendasi model kacamata. Swin Transformer merupakan salah satu arsitektur Deep Learning yang menggunakan pendekatan pembagian gambar menjadi patch hierarkis dan memproses citra melalui serangkaian transformer block secara bertahap. Hasil penelitian ini menunjukkan bahwa model bekerja dengan baik dalam melakukan klasifikasi. Dengan konfigurasi learning rate 0.0001, batch size 32, dan 64 epoch model memperoleh akurasi pelatihan 99.15%, validasi 98.47%, dan pengujian 97.63%. Selisih akurasi yang kecil (0.68%) menunjukkan bahwa model tidak mengalami overfitting signifikan. Selain itu, nilai loss yang rendah dan konsisten di seluruh subset data menunjukkan stabilitas model. Hasil penelitian ini menunjukkan bahwa sistem yang dikembangkan mampu mengklasifikasikan bentuk wajah dengan baik dan memberikan rekomendasi kacamata yang sesuai.

Kata kunci: Kacamata, Bentuk Wajah, Swin Transformer

ABSTRACT

Facial shape serves as an identity marker that reflects the unique physical characteristics of each individual. In general, facial shapes are categorized into five main types: square, round, oval, heart, and oblong. Information regarding facial shapes can be utilized in various aspects of personalization, including the selection of suitable eyeglass models. As the fashion industry evolves, eyeglasses are no longer solely functional as vision aids but also serve as aesthetic elements that enhance one's appearance. Therefore, selecting eyeglass models based on facial shape is a crucial factor in improving an individual's visual appeal. In online shopping, customers often face difficulties in determining the most suitable eyeglass model. Thus, an AI-based automated system is required to provide optimal recommendations. This study implements the Swin Transformer model to classify facial shapes as a basis for eyeglass model recommendations. Swin Transformer is a Deep Learning architecture that employs a hierarchical patch-based image partitioning approach and processes images through a series of transformer blocks in a progressive manner. The results of this study demonstrate that the model performs well in classification tasks. With a learning rate of 0.0001, a batch size of 32, and 64 epochs, the model achieved a training accuracy of 99.15%, a validation accuracy of 98.47%, and a test accuracy of 97.63%. The small accuracy difference (0.68%) indicates that the model does not exhibit significant overfitting. Furthermore, the consistently low loss values across all data subsets indicate the model's stability. These findings confirm that the developed system is capable of accurately classifying facial shapes and providing appropriate eyeglass recommendations.

Keywords: Eyeglass, Face Shape, Swin Transformer

I. PENDAHULUAN

Bagian tubuh manusia yang memiliki peran penting untuk menentukan identitas seseorang adalah wajah. Wajah memiliki berbagai elemen dan ciri – ciri tertentu yang berbeda serta melekat pada setiap individu, seperti hidung, alis, mata, mulut, dan juga bentuk wajah [1]. Perbedaan bentuk wajah dapat diklasifikasikan ke dalam 5 bentuk, yaitu *square*, *round*, *oval*, *heart*, dan *oblong* [2]. Bentuk wajah tidak hanya menjadi penanda identitas bagi seseorang, tetapi juga mencerminkan karakteristik fisik yang unik bagi setiap individu. Menurut [3] informasi yang terdapat pada bentuk wajah dapat memberikan rekomendasi signifikan untuk melakukan personalisasi gaya busana, gaya rambut dan juga bentuk kacamata

Fungsi penggunaan kacamata telah berkembang dari sekedar alat bantu pengelihatian menjadi salah satu kebutuhan *fashion* [4]. Oleh karena itu, diperlukan kecocokan dalam pemilihan model bentuk kacamata agar dapat meningkatkan kesan estetika bagi penggunanya. Secara umum, pemilihan bentuk kacamata dapat dilakukan secara langsung, namun bagi pelanggan jual beli *online* sulit untuk mendapatkan model yang sesuai [5]. Berdasarkan ratusan bentuk kacamata yang tersedia, pemilihan bentuk kacamata yang sesuai dapat dilakukan dengan menggunakan bentuk wajah sebagai objek klasifikasi [6].

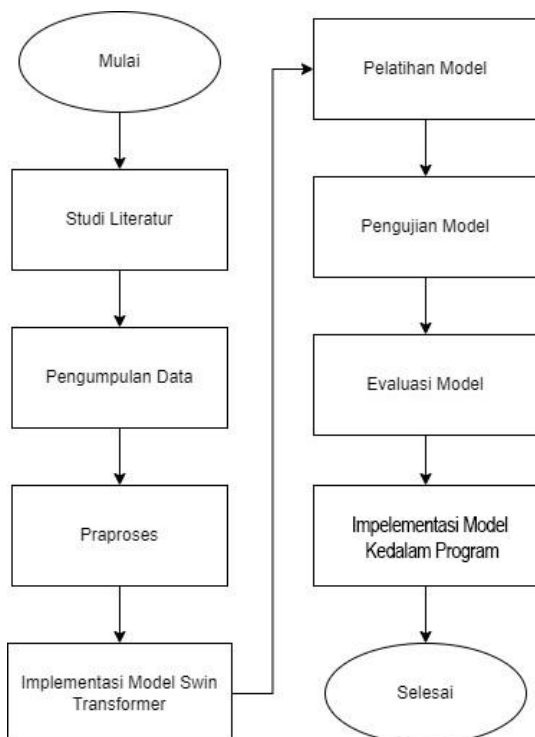
Klasifikasi merupakan aspek krusial dalam analisis citra. Proses ini bekerja dengan mengelompokkan citra ke dalam kategori tertentu berdasarkan fitur dan karakteristik yang dimiliki. Klasifikasi dapat mengidentifikasi pola dalam suatu citra untuk digunakan pada berbagai aplikasi, seperti deteksi objek, dan pengenalan wajah[7]. Dengan memanfaatkan klasifikasi pada visi komputer, pelanggan dapat mengetahui rekomendasi bentuk kacamata yang sesuai secara efektif dan efisien berdasarkan model kacamata yang telah dispesifikasikan oleh pabrikan melalui bentuk wajah. Penggunaan visi komputer dalam melakukan klasifikasi pada gambar telah diuji di berbagai kasus pada penelitian sebelumnya. Tujuan utama dari visi komputer adalah membuat sebuah model agar dapat membaca informasi dan data pada sebuah gambar [2]. *Transformer* pada awal diciptakan digunakan untuk menerjemahkan mesin pada ranah *Natural Language Processing (NLP)*. Perkembangan metode *transformer-based* memberikan performa inovatif pada berbagai tugas, salah satunya adalah visi komputer. [8]. *Swin Transformer* dapat melakukan pemrosesan gambar dengan lebih efisien dan terstruktur dengan membagi gambar menjadi beberapa potongan kecil lalu memproses potongan – potongan tersebut secara hirarkis. Setiap potongan akan diproses oleh satu set lapisan *transformator*. Setiap set lapisan *transformator* ini kemudian akan membentuk sebuah set peta fitur. Set peta fitur akan dibagi menjadi lebih

kecil untuk diproses oleh set lapisan transformator lainnya. Iterasi ini akan terus berlanjut hingga memenuhi hasil akhir. [8].

II. METODE PENELITIAN

A. Jenis Penelitian

Penelitian yang dilakukan kali ini memerlukan serangkaian langkah yang harus diikuti untuk memastikan bahwa setiap aspek penelitian dapat dilaksanakan dengan efektif dan efisien, mulai dari pengumpulan data hingga analisis hasil. Langkah-langkah penelitian dapat dilihat pada gambar 1 :



Gambar 1. Alur Penelitian

Penelitian ini menggunakan teknik mengumpulkan data yang diperlukan untuk proses pelatihan dan pengujian model kemudian dilanjutkan implementasi *Swin Trannsformer* lalu menguji model tersebut sesuai data uji.

Pengambilan data diambil dari situs *Kaggle*. Situs ini dapat diakses melalui internet secara terbuka dan menyediaka berbagai dataset yang diunggah oleh pengguna.

B. Studi Pustaka

Dalam proses studi pustaka, peneliti berupaya untuk mengumpulkan, menelaah, dan memahami berbagai sumber informasi terbuka yang tersedia dalam internet dalam

bentuk artikel ilmiah, makalah, dan penelitian terdahulu yang sesuai dengan tema penelitian. Tujuan utama dari kegiatan ini adalah mengumpulkan landasan teoretis untuk mendukung pelaksanaan penelitian, dimulai dengan mengumpulkan informasi mengenai bentuk – bentuk wajah pada manusia, citra digital, *deep learning*, arsitektur *transformer*, dan model *Swin Transformer*. Referensi yang digunakan pada penelitian ini akan dicantumkan dalam daftar pustaka yang terdapat pada akhir laporan skripsi, sebagai bukti bacaan dan studi yang telah dilakukan.

C. Pengumpulan Data

Penelitian ini menggunakan *dataset* yang diambil dari situs *Kaggle*. Situs ini dapat diakses melalui internet secara terbuka dan menyediakan berbagai *dataset* yang diunggah oleh pengguna. Gambar 3. merupakan contoh dari sampel data yang digunakan



Gambar 2. Sampel Dataset







Dataset terdiri dari 10.000 gambar selebriti wanita dari seluruh dunia yang dikategorikan menurut bentuk wajah mereka, yaitu persegi, lingkaran, oval, hati, dan persegi panjang dengan masing – masing kategori memiliki 2000 gambar. Struktur ini memungkinkan pengelolaan dan pemrosesan data yang lebih mudah, terutama dalam hal pemisahan data per kelas.

D. Mapping Bentuk Kacamata

Pemilihan bentuk kacamata yang digunakan pada penelitian ini dibatasi pada bentuk - bentuk yang umum ditemukan, yaitu square, oval, octagon, pilot / aviator, cat eye, dan round. Penentuan bentuk ini hanya didasarkan pada karakteristik desain frame tanpa mempertimbangkan preferensi lain seperti bahan, gender, maupun warna. Pembatasan ini dilakukan untuk menjaga konsistensi data dan menghindari keragaman bentuk frame yang terlalu luas sehingga penelitian dapat lebih terfokus pada analisis kecocokan antara bentuk kacamata dan bentuk wajah. Adapun jenis kacamata yang memiliki fungsi khusus, seperti kacamata

olahraga, kacamata keselamatan, atau kacamata medis, tidak termasuk dalam cakupan penelitian ini.

Tabel 1. Mapping Kecocokan Bentuk Kacamata Dengan Bentuk Wajah

| Bentuk Kacamata | Nama | Bentuk Wajah | | | | |
|---|-----------------|--------------|-------|------|-------|--------|
| | | Square | Round | Oval | Heart | Oblong |
|  | square | 0 | 1 | 0 | 0 | 1 |
|  | oval | 1 | 0 | 1 | 1 | 0 |
|  | octagon | 0 | 1 | 0 | 0 | 0 |
|  | pilot (aviator) | 0 | 0 | 1 | 1 | 1 |
|  | cat eye | 0 | 1 | 1 | 1 | 1 |
|  | round | 1 | 0 | 1 | 1 | 0 |

Peneliti mengambil gambar kacamata melalui katalog yang terdapat pada situs glassdirect.co.uk. Kriteria pengambilan gambar dilakukan dengan memilih masing – masing satu produk dari katalog produk dengan menyesuaikan bentuk frame kacamata yang telah didefinisikan

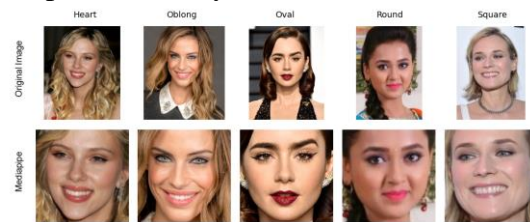
E. Persiapan Dataset

Pada persiapan dataset, dilakukan beberapa tahap untuk mempersiapkan gambar dari hasil pengumpulan dataset yang telah dilakukan sebelumnya. Proses persiapan dataset ditujukan untuk menghasilkan gambar yang diharapkan dapat memenuhi syarat dari model yang digunakan sehingga dapat memberikan hasil akurasi dan prediksi yang optimal.

Penelitian kali ini terdapat dua tahapan dalam melakukan persiapan dataset. Tahapan pertama dilakukan dengan mendeteksi bagian wajah dengan pustaka Mediapipe untuk melakukan seleksi area bagian wajah yang terdapat pada gambar di dataset. Pada tahap kedua dilakukan dengan oversampling dataset untuk menyelaraskan setiap kelas pada dataset.

a. Deteksi Menggunakan Mediapipe

Deteksi wajah menggunakan *mediapipe* digunakan untuk memastikan bahwa hanya area wajah yang digunakan untuk model dalam melakukan klasifikasi. Penggunaan *mediapipe* dilakukan dengan menggunakan inisiasi *min_detection_confidence* sebesar 0.5 untuk memastikan bahwa area yang terdeteksi adalah wajah. Lalu dilakukan iterasi pada setiap kelas yang terdapat pada *dataset* untuk memproses setiap gambar yang terdapat didalamnya.

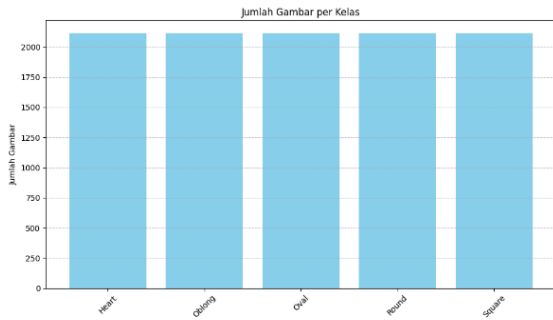


Gambar 3. Original Image dan Mediapipe

Dari setiap gambar yang telah diambil dari folder *dataset* kemudian akan dideteksi oleh *mediapipe*. Setelah area wajah terdeteksi, gambar dipotong sehingga hanya bagian wajah yang disimpan, mengurangi informasi latar belakang yang tidak relevan. Namun, jika terdapat gambar yang rusak atau tidak terdeteksi oleh *mediapipe*, maka gambar akan dihapus.

b. Oversampling Dataset

Proses oversampling dataset diperlukan setelah proses seleksi area wajah menggunakan *mediapipe*. Dari seleksi *mediapipe* sebelumnya terdapat ketidakseimbangan kelas setelah beberapa gambar dari 5 kelas yang telah dihapus karena dianggap rusak akibat tidak terdeteksi oleh *mediapipe*. Untuk memastikan bahwa setiap kelas seimbang, digunakan pustaka *imblearn* untuk melakukan random oversampling. Hal ini dilakukan agar tidak terjadi bias dari model terhadap kelas mayoritas, karena pada kelas yang memiliki bobot lebih sedikit akan dilakukan duplikasi sehingga bobot pada setiap kelas akan seimbang. Hasil akhir dari proses ini akan disimpan kedalam folder baru pada drive yang telah disiapkan.

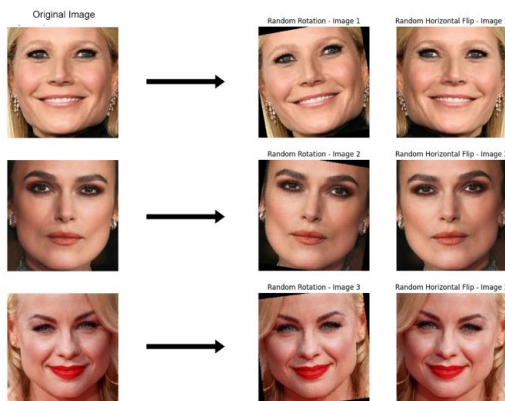


Gambar 4. Distribusi hasil dari oversampling

Pada gambar 4 merupakan hasil dari *oversampling* yang telah dilakukan. Pada setiap kelas memiliki bobot yang seimbang, yaitu sebanyak 2115 gambar. Hal ini diharapkan dapat meningkatkan kinerja model dalam tugas klasifikasi.

F. Argumentasi Data dan Transformasi Data

Augmentasi *dataset* merupakan sebuah proses memodifikasi gambar untuk mengenali segala fitur dan pola yang terdapat didalamnya. Pada penelitian kali ini, proses augmentasi data dilakukan dengan menerapkan *Flipping Horizontal* untuk memberikan variasi data dengan membalik gambar. Augmentasi data dengan *RandomRotation*, yang memberikan variasi posisi wajah melalui rotasi acak untuk meningkatkan fleksibilitas model terhadap orientasi gambar. Selain itu, *RandomHorizontalFlip* diaplikasikan untuk memberikan variasi data dengan membalik gambar secara horizontal. *CenterCrop* juga digunakan untuk memastikan bagian tengah gambar, yang biasanya memuat area wajah, tetap menjadi fokus utama dan sesuai dengan dimensi input model. Untuk membuat data lebih dikenali oleh model dilakukan penerapan *Resize* menjadi 224 x 224 piksel sesuai dengan input model *Swin Transformer*.



Gambar 5. Hasil Argumentasi Citra

Gambar 5 menunjukkan hasil augmentasi yang diterapkan pada citra. Penggunaan augmentasi ini bertujuan untuk menambah variasi dari data dan membiasakan model dengan keberagaman orientasi citra yang diterima.

G. Pembagian Dataset

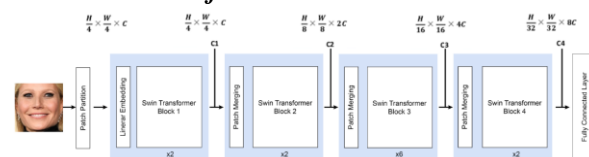
Pembagian *dataset* diperlukan untuk memastikan model dapat dilatih dan dievaluasi secara efektif dan objektif. Pada umumnya pembagian data latih mendapatkan rasio persentase terbesar daripada data validasi dan data uji. Data latih diharapkan mampu memberikan lebih banyak variasi agar model mampu belajar lebih baik. Sedangkan data validasi digunakan untuk mengetahui performa model selama proses pelatihan. Data uji digunakan untuk mengukur seberapa baik model dapat melakukan generalisasi terhadap data yang belum pernah dilihat oleh model.

Tabel 2. Pembagian dataset

| Rasio | Jenis Data | Perbandingan | Jumlah Data |
|-------------------|---------------|--------------|---------------|
| 9:1 | Data Latih | 80% | 8.461 |
| | Data Validasi | 10% | 1.057 |
| | Data Uji | 10% | 1.057 |
| Total Data | | | 10.575 |

Pada Tabel 3.2 merupakan pembagian dataset yang telah disiapkan sebelumnya. Pembagian dataset dilakukan dengan membagi dataset menjadi tiga subset yang berbeda dengan rasio 9:1. Dari total 10.575 gambar dilakukan pembagian sebesar 8.461 gambar untuk digunakan sebagai data latih, jumlah tersebut setara dengan 80% dari keseluruhan dataset. Lalu, sebanyak 1.057 gambar digunakan sebagai data validasi, jumlah tersebut setara dengan 10% dari keseluruhan dataset. Kemudian, 10% gambar yang tersisa digunakan sebagai data uji dengan jumlah 1.057 gambar. Pembagian ini bertujuan untuk menjaga validitas hasil evaluasi dengan memastikan bahwa model diuji pada data yang tidak pernah dilihat selama proses pelatihan.

H. Swin Transformer



Gambar 6. Arsitektur Swin Tranformer

Pada penelitian ini penulis menggunakan model *Swin Transformer* untuk memproses data yang telah disiapkan. *Swin Transformer* terdiri dari empat stage utama, di mana setiap stage memiliki resolusi fitur dan jumlah layer yang berbeda untuk menangkap informasi dari tingkat lokal hingga global, diantaranya:

1) Stage 1

Pada tahap awal, input citra yang telah diproses melalui Patch Partition akan memiliki resolusi lebih kecil dengan dimensi $H/4 \times W/4 \times C$, di mana C adalah jumlah dimensi awal. Tahap ini terdiri dari beberapa *Swin Transformer Block* yang menerapkan Window-based Multi-Head Self-Attention (W-MSA) dan Shifted Window Multi-Head Self-Attention (SW-MSA) untuk menangkap informasi dari wilayah lokal.

2) Stage 2

Setelah melewati Stage 1 resolusi fitur akan dikurangi setengahnya melalui Patch Merging, sementara jumlah saluran (C) akan meningkat dua kali lipat. Dengan demikian, resolusi fitur menjadi $H/8 \times W/8 \times 2C$. Pada tahap ini, model mulai menangkap pola yang lebih kompleks dari citra.

3) Stage 3

Sama seperti sebelumnya, pada Stage 3 dilakukan Patch Merging yang kembali mengurangi resolusi menjadi $H/16 \times W/16 \times 4C$. Jumlah channel kembali meningkat sehingga model dapat memahami hubungan dengan lebih baik. Tahapan ini memainkan peran penting dalam memahami struktur wajah secara lebih luas, terutama untuk klasifikasi bentuk wajah.

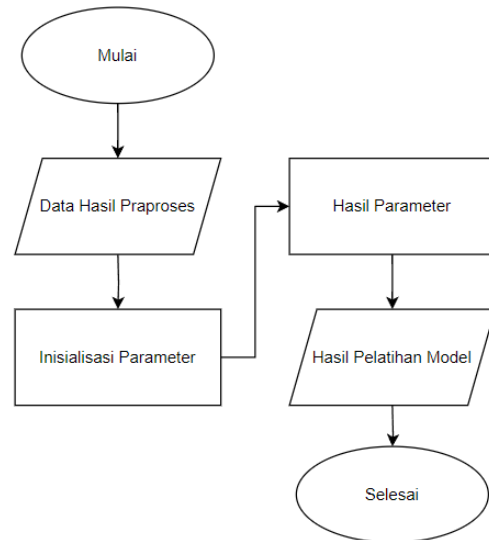
4) Stage 4

Merupakan tahap terakhir dari ekstraksi fitur. Pada tahapan ini resolusi citra semakin kecil, yaitu $H/32 \times W/32 \times 8C$. Meskipun resolusi lebih kecil, informasi yang diperoleh sudah sangat abstrak dan merepresentasikan fitur global dari citra. Hasil akhir dari tahap ini akan digunakan untuk tahap classification head pada model untuk menentukan bentuk wajah berdasarkan fitur yang telah diekstraksi.

I. Pelatihan Model

Setelah melakukan perancangan arsitektur *Swin Transformer* dan melakukan konfigurasi hyperparameter, proses penelitian dilanjutkan dengan melakukan pelatihan terhadap model. Proses pelatihan model pada penelitian ini

dilakukan dengan menggunakan dataset yang telah dipersiapkan sebelumnya pada tahap praproses. Untuk mencapai model yang optimal, proses ini akan mengikuti urutan langkah yang telah ditetapkan pada Gambar 8.



Gambar 7. Diagram Pelatihan Model

Langkah awal pada pelatihan model dilakukan dengan melakukan evaluasi dan menentukan parameter kunci yang akan digunakan selama pelatihan, termasuk pemilihan optimizer, fungsi loss, jumlah epoch, langkah per epoch, validation data, valuation steps, dan penyesuaian parameter untuk mengoptimalkan proses pelatihan dan menghindari overfitting. Pelatihan model memberikan hasil akhir berupa model baru yang telah melalui tahap pembelajaran menggunakan data latih. Model baru tersebut diharapkan telah berhasil mempelajari pola dan fitur – fitur yang terdapat pada data latih, sehingga model mampu untuk mengenali dan mengklasifikasikan data sesuai dengan kategori yang telah ditentukan.

J. Pengujian Model

Setelah melakukan pelatihan pada model, diperlukan pengujian untuk mengevaluasi performa. Pengujian model dilakukan dengan memberikan data baru yang sebelumnya belum pernah dilihat oleh model. Data baru ini digunakan untuk menguji kemampuan model dalam mengenali dan mengklasifikasikan data yang belum pernah dikenal sebelumnya. Pengujian ini penting untuk memastikan bahwa model mampu menggeneralisasi pola dan fitur yang telah dipelajari selama proses pelatihan.

K. Evaluasi Model

Setelah melakukan pelatihan model, tahap selanjutnya adalah mengevaluasi performa dari model yang telah dikembangkan. Evaluasi model dilakukan dengan analisis *confusion matrix* dan *classification report* untuk memperoleh pemahaman yang mendalam mengenai seberapa baik model dapat beroperasi dengan efektif dan akurat. *Confusion Matrix* merupakan tabel analisis yang membandingkan nilai prediksi dengan nilai sebenarnya. Dalam penelitian ini, *Confusion Matrix* digunakan untuk mengevaluasi nilai prediksi setiap kelas pada bentuk wajah berdasarkan hasil pengujian model terhadap nilai sebenarnya. *Confusion Report* merupakan matriks analisis yang digunakan untuk mengukur kinerja model klasifikasi pada setiap kelas dalam data pengujian. Laporan ini menyajikan berbagai metrik evaluasi, seperti *precision*, *recall*, dan *f1-score* untuk membantu memahami kekuatan dan kelemahan model dalam mengklasifikasikan data.

L. Skenario Penelitian

Untuk mendapatkan informasi mengenai performa terbaik dari model, maka diperlukan skenario pelatihan. Pada penelitian kali ini, skenario penelitian dilakukan dengan membandingkan 2 *learning rate*, yaitu sebesar 0.0001 dan 0.00001 yang dipadukan dengan penyesuaian parameter model. Tabel 3. menunjukkan setiap skenario yang digunakan untuk penelitian.

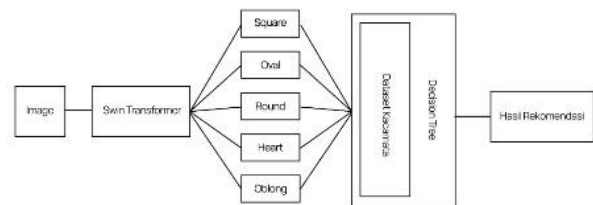
Tabel 3. Skenario pelatihan model

| Skenario | Learning Rate | Batch Size | Epoch |
|----------|---------------|------------|-------|
| 1 | 0.0001 | 32 | 32 |
| 2 | | 32 | 64 |
| 3 | | 64 | 32 |
| 4 | | 64 | 64 |
| 5 | 0.00001 | 32 | 32 |
| 6 | | 32 | 64 |
| 7 | | 64 | 32 |
| 8 | | 64 | 64 |

Pada setiap *learning rate* yang digunakan akan dilakukan perbandingan dengan menggunakan penyesuaian parameter berupa 32 epoch dengan 32 batch size, 64 epoch dengan 32 batch size, 32 epoch dengan 64 batch size, dan 64 epoch dengan 64 batch size. Dengan menerapkan skenario tersebut, peneliti akan membandingkan hasil tingkat akurasi dari setiap skenario. Perbedaan tingkat akurasi dan metrik perbandingan yang didapatkan pada setiap skenario penelitian akan ditarik kesimpulan.

M. Implementasi Model

Proses implementasi model dilakukan setelah menyelesaikan langkah permodelan. Model yang telah dilatih pada tahap sebelumnya akan di muat ulang pada sebuah program sederhana untuk menentukan bentuk kacamata berdasarkan bentuk wajah. Program rekomendasi bentuk kacamata dibuat secara sederhana menggunakan platform *huggingface* dan *gradio*.



Gambar 8. Alur Sistem Implementasi Model

Gambar 8 Merupakan alur sistem rekomendasi bentuk kacamata berdasarkan bentuk wajah.

1) Input Gambar

Pada langkah pertama pengguna dapat mengunggah gambar berupa bentuk wajah. Gambar yang telah diunggah oleh pengguna nantinya akan diteruskan untuk diproses oleh Swin Transformer.

2) Swin Transformer

Gambar yang telah diterima akan diproses oleh model. Model Swin Transformer yang digunakan pada tahapan ini merupakan model dengan hasil akhir terbaik berdasarkan uji coba melalui beberapa skenario yang telah dilakukan pada tahap pelatihan. Model akan melakukan klasifikasi untuk mengetahui bentuk wajah pada gambar berdasarkan 5 bentuk wajah, yaitu *square*, *oval*, *round*, *heart*, dan *oblong*.

3) Decision Tree

Hasil klasifikasi yang telah didapatkan pada tahapan sebelumnya akan dijadikan sebagai

variable utama dalam menentukan kecocokan antara bentuk wajah dan bentuk frame kacamata. Pada tahap ini *decision tree* akan menentukan relasi yang sesuai antara bentuk wajah yang terdeteksi pada tahap sebelumnya dengan dataset frame kacamata untuk menentukan jenis frame yang paling sesuai berdasarkan mapping yang telah dibuat.

4) Hasil rekomendasi

Tahap akhir dari program akan menampilkan hasil rekomendasi bentuk kacamata berdasarkan hasil dari *decision tree*.

III. HASIL DAN PEMBAHASAN

Pelatihan model dilakukan dengan delapan skenario berbeda untuk mengevaluasi kombinasi *learning rate*, *batch size*, dan jumlah *epoch* terbaik.

| Learning Rate | Batch Size | Epoch | Training | | Validation | | Testing | |
|---------------|------------|-------|----------|-------|------------|-------|----------|-------|
| | | | Accuracy | Loss | Accuracy | Loss | Accuracy | Loss |
| 0.0001 | 32 | 32 | 99,56% | 0,011 | 97,52% | 0,106 | 97,16% | 0,164 |
| | 32 | 64 | 99,15% | 0,027 | 98,47% | 0,076 | 97,63% | 0,131 |
| | 64 | 32 | 98,93% | 0,031 | 97,80% | 0,108 | 97,06% | 0,193 |
| | 64 | 64 | 99,30% | 0,018 | 98,00% | 0,090 | 97,54% | 0,126 |
| 0.00001 | 32 | 32 | 99,00% | 0,032 | 97,61% | 0,099 | 96,11% | 0,138 |
| | 32 | 64 | 99,66% | 0,011 | 98,09% | 0,082 | 96,87% | 0,120 |
| | 64 | 32 | 98,56% | 0,050 | 97,14% | 0,095 | 96,11% | 0,149 |
| | 64 | 64 | 99,25% | 0,025 | 97,23% | 0,111 | 97,06% | 0,123 |

Gambar 9. Hasil Evaluasi Model

Dari data yang disajikan terlihat bahwa semua model menunjukkan performa nilai akurasi pada data training cenderung sangat tinggi dengan sebagian besar hasil berada di atas 98%. Namun untuk menentukan model terbaik perlu untuk mempertimbangkan beberapa aspek penting, seperti keseimbangan metrik antara training dan validation, serta performa pada testing set yang menunjukkan kemampuan generalisasi model.

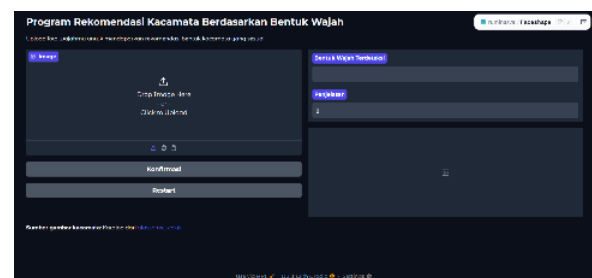
Hasil dari skenario learning rate 0.0001 dan 0.00001 pada penelitian menghasilkan kesimpulan bahwa learning rate 0.0001 secara konsisten memberikan hasil testing accuracy yang lebih baik dibandingkan learning rate 0.00001. Hal ini menunjukkan bahwa learning rate 0.00001 terlalu kecil sehingga menyebabkan proses pembelajaran berjalan terlalu lambat dan model kesulitan mencapai konvergensi optimal. Sebaliknya, pada learning rate 0.0001 memberikan keseimbangan yang baik sehingga memungkinkan model belajar secara efektif.

Model dengan learning rate 0.0001, batch size 32, dan epoch 64 menunjukkan performa terbaik secara keseluruhan. Model ini mencapai

training accuracy 99.15%, validation accuracy 98.47%, dan testing accuracy 97.63%. Skenario tersebut menghasilkan nilai testing accuracy tertinggi di antara semua model. Faktor penting lainnya yang terdapat pada skenario tersebut adalah jarak akurasi antara training dan validation yang relatif kecil, yaitu senilai 0.68%. Hal tersebut mengindikasikan model tidak mengalami overfitting yang signifikan. Jumlah loss juga konsisten dan relatif rendah di semua dataset dengan hasil training 0.027, validation 0.076, dan testing 0.131.

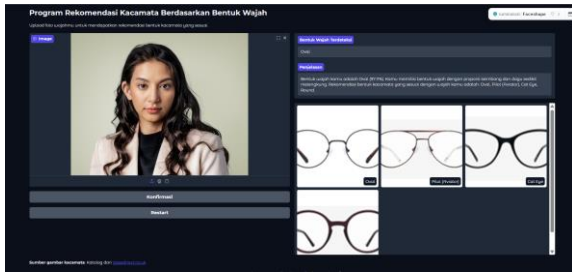
Rekomendasi bentuk kacamata setelah menjalankan beberapa skenario pelatihan dan evaluasi pada model, penelitian dilanjutkan dengan melakukan implementasi model dengan performa terbaik berdasarkan skenario pengujian ke dalam program rekomendasi kecocokan bentuk kacamata dan bentuk wajah. Proses pengembangan program dilakukan dengan memanfaatkan platform *huggingface* dan *gradio*.

1. Pengujian Program Rekomendasi Bentuk Kacamata



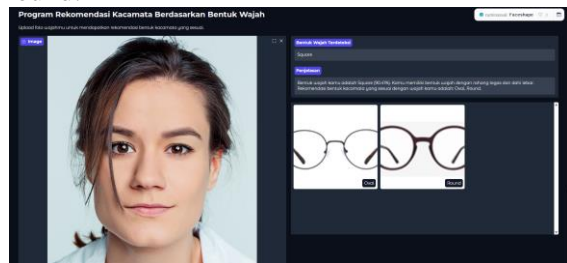
Gambar 10. Tampilan Program Rekomendasi bentuk kacamata

Gambar 10 merupakan tampilan home dari program rekomendasi bentuk wajah untuk menentukan rekomendasi bentuk kacamata. Pengguna dapat menekan box “click to upload” atau menggeser file dari penyimpanan lokal untuk melakukan unggah gambar. Lalu pengguna dapat memulai deteksi wajah dengan menekan tombol “konfirmasi” untuk mendapatkan hasil rekomendasi bentuk kacamata. Ketika pengguna salah melakukan input gambar, maka pengguna dapat menekan tombol “restart” untuk mengulang program dan menghapus data yang telah diinput.



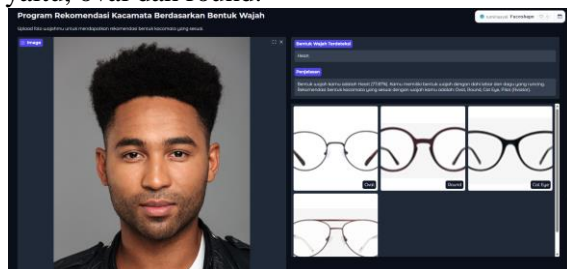
Gambar 11. Hasil prediksi wajah oval

Gambar 11 merupakan hasil untuk ditampilkan kepada pengguna ketika proses deteksi telah berhasil dilakukan oleh program. Pada bagian ini program berhasil mendeteksi bentuk wajah sebagai oval dengan tingkat *confidence* sebesar 97%. Program juga memberikan rekomendasi bentuk kacamata yang diberikan yaitu, oval, pilot (aviator), cat eye, dan round.



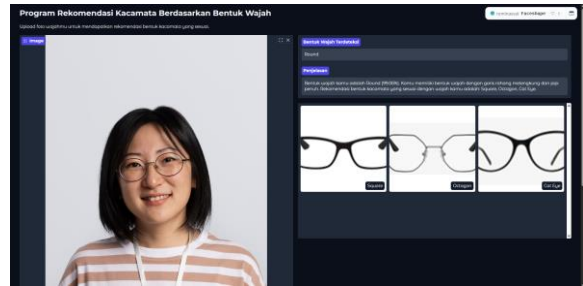
Gambar 12. Hasil prediksi wajah square

Gambar 12 adalah hasil yang ditampilkan program kepada pengguna ketika proses deteksi telah berhasil dilakukan. Pada bagian ini program berhasil mendeteksi bentuk wajah sebagai square dengan tingkat *confidence* sebesar 90,41%. Program juga memberikan rekomendasi bentuk kacamata yang diberikan yaitu, oval dan round.



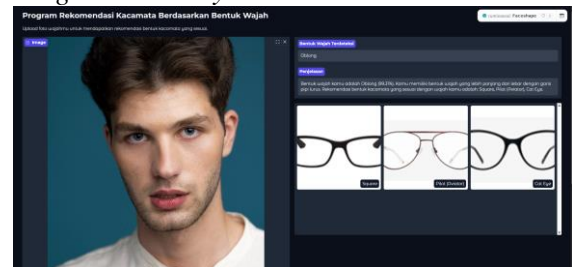
Gambar 13. Hasil prediksi wajah heart

Gambar 13 merupakan hasil yang ditampilkan program kepada pengguna ketika proses deteksi wajah telah berhasil dilakukan. Pada bagian ini program berhasil mendeteksi bentuk wajah sebagai heart dengan tingkat *confidence* sebesar 75,97%. Program juga memberikan rekomendasi bentuk kacamata yang diberikan yaitu, pilot(aviator), cat eye dan round.



Gambar 14. Hasil prediksi wajah round

Gambar 14 merupakan hasil yang ditampilkan oleh program ketika proses deteksi wajah telah berhasil dilakukan. Pada bagian ini program berhasil mendeteksi bentuk wajah sebagai *round* dengan tingkat *confidence* sebesar 99,06%. Program juga memberikan rekomendasi bentuk kacamata yang diberikan yaitu, *square*, *octagon* dan *cat eye*.



Gambar 15. Hasil prediksi wajah oblong

Gambar 15 merupakan hasil yang ditampilkan oleh program ketika proses deteksi wajah telah berhasil dilakukan. Pada bagian ini program berhasil mendeteksi bentuk wajah sebagai *oblong* dengan tingkat *confidence* sebesar 99,31%. Program juga memberikan rekomendasi bentuk kacamata yang diberikan yaitu, *square*, *oval*, *pilot/aviator* dan *cat eye*.

IV. KESIMPULAN

Model Swin Transformer berhasil diterapkan dalam proses klasifikasi bentuk wajah dengan menggunakan metode ekstraksi fitur berbasis shifted window. Proses persiapan dataset dengan menggunakan Mediapipe untuk melakukan seleksi area wajah serta penerapan augmentasi data dapat berkontribusi dalam meningkatkan kualitas input untuk model. Hal ini memungkinkan model untuk memahami karakteristik wajah secara efektif.

Berdasarkan penelitian yang telah dilakukan dapat disimpulkan bahwa, menggunakan kombinasi ratio pada dataset dan melakukan konfigurasi jumlah epoch, batch size, serta learning rate yang bervariasi dapat mempengaruhi kinerja model. Pelatihan dengan

skenario terbaik diperoleh menggunakan konfigurasi learning rate 0.0001, batch size 32, dan 64 epoch. Model memperoleh akurasi pelatihan 99.15%, validasi 98.47%, dan pengujian 97.63% dengan selisih akurasi yang kecil yaitu 0.68%.

Model yang telah melalui proses pelatihan berhasil diimplementasikan dalam program pendeteksi wajah dengan performa yang baik, sehingga dapat memprediksi data baru yang sebelumnya tidak dikenali saat pelatihan. Program ini mampu mengklasifikasikan bentuk wajah ke dalam lima kategori, yaitu square, heart, oblong, oval, dan round. Berdasarkan klasifikasi ini, program dapat memberikan rekomendasi bentuk kacamata yang sesuai dengan mencocokkan bentuk wajah dan kacamata menggunakan skema pemetaan yang telah ditentukan sebelumnya

DAFTAR PUSTAKA

- S. Bahri dan H. Kusindaryadi, "Rancang Bangun Pemantauan Absensi Mahasiswa dengan Menggunakan Sidik Wajah secara Simultan Melalui CCTV Ruang Kelas," *Resist. (elektRONika kEndali Telekomun. tenaga List. kOMputeR)*, vol. 3, no. 1, hal. 37, 2020, doi: 10.24853/resistor.3.1.37-44.
- T. Hidayat, I. A. Astuti, A. Yaqin, A. P. Tjilen, dan T. Arifianto, "Grouping of Image Patterns Using Inceptionv3 For Face Shape Classification," *JOIV Int. J. Informatics Vis.*, vol. 7, no. 4, hal. 2336–2343, 2023, doi: 10.30630/joiv.7.4.1743.
- J. Duan, X. Su, J. Ren, dan L. Xie, "Face Shape Classification Based on Bilinear Network with Attention Mechanism," *J. Phys. Conf. Ser.*, vol. 2278, no. 1, 2022, doi: 10.1088/1742-6596/2278/1/012041.
- E. D. Madyatmadja, L. Kusumawati, S. P. Jamil, W. Kusumawardhana, S. Informasi, dan U. B. Nusantara, "Infotech: journal of technology I nformation," *Raden Ario Damar*, vol. 7, no. 1, hal. 55–62, 2021.
- Siti Khotimatul Wildah, S. Agustiani, Ali Mustopa, Nanik Wuryani, Hendri Mahmud Nawawi, dan Rizky Ade Safitri, "Pengenalan Wajah Menggunakan Pembelajaran Mesin Berdasarkan Ekstraksi Fitur Pada Gambar Wajah Berkualitas Rendah," *INFOTECH J. Inform. Teknol.*, vol. 2, no. 2, hal. 95–103, 2021, doi: 10.37373/infotech.v2i2.189.
- A. A. Retno Hapsari, R. Gernowo, dan C. E. Widodo, "Penggunaan Algoritma CART untuk Pemilihan Bingkai Kacamata dengan Penerapan Model Morfologi Indeks Wajah untuk Identifikasi Bentuk Wajah," *J. Sist. Inf. Bisnis*, vol. 10, no. 1, hal. 1–9, 2019, doi: 10.21456/vol10iss1pp1-9.
- A. Naufal Hilmi dkk., "Implementasi Algoritma K-Nearest Neighbor (KNN) untuk Identifikasi Penyakit pada Tanaman Jeruk Berdasarkan Citra Daun," no. 2, hal. 107–117, 2024, [Daring]. Tersedia pada: <https://doi.org/10.62951/router.v2i 2.78>.
- R. A. Dihin, E. AlShemmary, dan W. Al-Jawher, "Diabetic Retinopathy Classification Using Swin Transformer with Multi Wavelet," *J. Kufa Math. Comput.*, vol. 10, no. 2, hal. 167–172, 2023, doi: 10.31642/jokmc/2018/100225. <https://doi.org/10.24127/emteka.v5i1.5247>
- Septiani, K. R. E., & Irsyadi, F. Y. A. (2020). GAME EDUKASI TARI TRADISIONAL INDONESIA UNTUK SISWA TUNARUNGU KELAS VI SEKOLAH DASAR. *Jurnal Teknik Informatika (Jutif)*, 1(1), 7. <https://doi.org/10.20884/1.jutif.2020.1.1.11>
- Wulandari, Y. R. (2019). FINGER PAINTING BERMEDIA TIGA DIMENSI TERHADAP KEMAMPUAN MOTORIK HALUS ANAK DENGAN SPEKTRUM AUTIS. *Jurnal Pendidikan Khusus*, 14(4).