

Analisis Pengaruh Transformasi Geometri terhadap Kinerja Klasifikasi Citra Menggunakan Convolutional Neural Network

Dzakwan Muhammad Khairan Putra Purnama and 13524145^{1,2}

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

¹13524145@std.stei.itb.ac.id, ²dzakwan.mkpp@gmail.com

Abstrak — Dalam proses pembelajaran geometri dasar, anak-anak sering kali menghadapi kesulitan untuk memvalidasi apakah bentuk bangun datar yang mereka gambar sudah benar. Di sisi lain, ketersediaan pengajar untuk memberikan umpan balik secara *real-time* sangat terbatas. Solusi berbasis kecerdasan buatan, khususnya *Convolutional Neural Networks* (CNN), dapat dimanfaatkan sebagai asisten belajar mandiri. Namun, tantangan utama muncul dari karakteristik motorik anak; hasil gambar mereka cenderung tidak sempurna, miring (terotasi), tidak proporsional (*scaling*), atau penyok (*shearing*). Secara matematis, ketidaksempurnaan ini merepresentasikan variasi transformasi geometri yang sering kali gagal dikenali oleh model CNN standar yang dilatih pada data ideal. Penelitian ini bertujuan menganalisis dampak penerapan transformasi geometri terhadap akurasi klasifikasi CNN. Melalui dua skenario eksperimen, tanpa augmentasi dan dengan augmentasi data geometri, penelitian ini mengevaluasi kemampuan model dalam mengenali citra yang mengalami distorsi spasial. Hasil penelitian menunjukkan bahwa integrasi operasi transformasi geometri ke dalam proses pelatihan secara signifikan meningkatkan ketahanan model, menjadikannya solusi yang layak untuk aplikasi edukasi anak.

Kata Kunci — Convolutional Neural Networks, Transformasi Geometri, Klasifikasi Citra, Edukasi Anak, Augmentasi Data.

I. PENDAHULUAN

Pendidikan usia dini sering kali melibatkan pengenalan bentuk-bentuk geometri dasar seperti persegi, segitiga, dan lingkaran. Bayangkan sebuah skenario di mana seorang anak sedang belajar menggambar bentuk-bentuk tersebut secara mandiri di rumah. Tanpa kehadiran guru yang mendampingi, anak tersebut membutuhkan sarana untuk mengetahui apakah gambar yang ia buat sudah dapat dikenali sebagai bentuk yang dimaksud. Di sinilah teknologi *Computer Vision* dapat berperan sebagai "guru virtual" untuk memverifikasi hasil karya anak tersebut.

Implementasi teknologi ini menghadapi tantangan teknis yang unik. Berbeda dengan bentuk geometri sempurna yang dihasilkan komputer, gambar tangan anak-anak memiliki variasi yang sangat tinggi. Hal

ini dikarenakan kemampuan motorik halus yang belum matang, seorang anak mungkin menggambar persegi yang miring 30 derajat, atau segitiga yang tergeser ke pojok kertas [1]. Bagi manusia, bentuk tersebut tetap mudah dikenali. Namun, bagi sebuah model *Deep Learning* seperti *Convolutional Neural Networks* (CNN), perubahan orientasi dan posisi ini dapat dianggap sebagai fitur yang benar-benar berbeda jika model tidak dilatih dengan benar [2]. Dalam kacamata Aljabar Geometri, fenomena "gambar anak yang tidak sempurna" ini dapat dijelaskan melalui konsep transformasi geometri. Sebuah gambar yang miring adalah hasil dari operasi **Rotasi**, gambar yang terlalu kecil atau besar adalah hasil **Penskalaan** (**Scaling**), dan gambar yang tidak simetris bisa jadi akibat operasi **Geser** (**Shearing**) atau **Translasi**. Masalah mendasar pada arsitektur CNN standar adalah kurangnya sifat *geometric invariance*; artinya, model yang hanya melihat contoh persegi tegak lurus cenderung akan gagal mengklasifikasikan persegi yang miring sebagai objek yang sama [4].

Penelitian ini bertujuan untuk mengatasi kesenjangan tersebut dengan menganalisis pengaruh transformasi geometri terhadap performa klasifikasi CNN. Penulis merancang dua skenario eksperimen menggunakan dataset bentuk (*shape dataset*). **Skenario 1** mensimulasikan model yang "kaku", dilatih tanpa variasi geometri, yang merepresentasikan keterbatasan sistem saat menghadapi gambar anak yang acak. **Skenario 2** menerapkan augmentasi data menggunakan operasi matriks transformasi untuk memperkaya variasi latih, sebuah teknik yang terbukti efektif dalam meningkatkan generalisasi model [3], [5]. Dengan membandingkan kedua skenario ini, makalah ini akan membuktikan secara empiris bahwa pemahaman dan penerapan aljabar geometri dalam tahap *preprocessing* data adalah kunci utama untuk menciptakan sistem edukasi cerdas yang adaptif dan ramah terhadap ketidaksempurnaan pengguna.

II. LANDASAN TEORI

A. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN) adalah kelas

arsitektur *Deep Learning* yang dirancang khusus untuk memproses data dengan struktur topologi *grid*, seperti citra digital. Arsitektur ini pertama kali dipopulerkan oleh LeCun et al. untuk pengenalan dokumen tulisan tangan [1]. Berbeda dengan jaringan saraf tiruan tradisional (*fully-connected*), CNN memanfaatkan tiga prinsip utama untuk menangani variabilitas citra:

- **Local Receptive Fields:** Koneksi lokal antar neuron untuk mendeteksi fitur spasial (tepi, sudut).
- **Shared Weights:** Penggunaan bobot yang sama (filter) pada seluruh posisi citra, yang mengurangi jumlah parameter secara drastis.
- **Pooling:** Sub-sampling spasial untuk mengurangi dimensi representasi fitur [4].

Meskipun CNN memiliki sifat translation-equivariant (fitur tetap terdeteksi meskipun objek bergeser), arsitektur standar CNN tidak memiliki sifat rotation-invariant atau scale-invariant secara intrinsik. Sebuah objek yang diputar atau diubah ukurannya akan menghasilkan aktivasi fitur yang berbeda di lapisan konvolusi, yang sering kali menyebabkan model gagal mengklasifikasikan objek yang sama jika orientasinya berubah [2].

B. Representasi Matematis Transformasi Geometri

Untuk memodelkan variasi bentuk pada gambar tangan anak (seperti miring, penyok, atau kurang proporsional), digunakan konsep utama *Transformasi Affine*. Dalam pengolahan citra, transformasi ini memetakan lokasi piksel dari koordinat asal (x, y) ke koordinat baru (x', y') . Agar operasi translasi (penjumlahan vektor) dapat digabungkan dengan operasi linear (rotasi/skala) dalam satu perkalian matriks tunggal, digunakan representasi **Koordinat Homogen**. Dalam sistem ini, titik 2D (x, y) direpresentasikan sebagai vektor kolom 3D $p = [x, y, 1]^T$.

Persamaan umum *transformasi affine* dinyatakan sebagai Persamaan 1 dan Persamaan 2:

$$p' = M \cdot p \quad (1)$$

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2)$$

Dimana M adalah matriks transformasi 3x3. Berikut adalah rincian matriks untuk setiap jenis distorsi yang relevan dalam penelitian ini:

1) Translasi (Pergeseran)

Digunakan untuk memodelkan objek yang tidak berada di tengah kanvas (offset, Persamaan 3).

$$M_{trans}(t_x, t_y) = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

2) Rotasi (Perputaran)

Digunakan untuk memodelkan kemiringan goresan tangan. Rotasi sebesar sudut θ (berlawanan arah

jarum jam, Persamaan 4) dengan pusat rotasi di titik asal $(0, 0)$:

$$M_{rot}(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

3) Penskalaan (Scaling)

Digunakan untuk memodelkan variasi ukuran gambar (Persamaan 5).

$$M_{scale}(s_x, s_y) = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

4) Shearing (Geseran)

Digunakan untuk memodelkan distorsi perspektif (Persamaan 6) atau bentuk yang "penyok" (misalnya persegi menjadi jajaran genjang).

$$M_{shear}(sh_x, sh_y) = \begin{bmatrix} 1 & sh_x & 0 \\ sh_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (6)$$

C. Interpolasi Citra

Ketika transformasi geometri (terutama rotasi dan *shearing*) diterapkan pada citra digital diskrit, koordinat hasil transformasi (x', y') sering kali jatuh pada nilai bilangan riil (*floating point*) yang tidak bertepatan dengan grid piksel integer. Jika dipetakan langsung, hal ini akan menimbulkan "lubang" atau artefak pada citra output.

Untuk mengatasi hal ini, digunakan teknik interpolasi intensitas. Penelitian saya ini menggunakan Interpolasi Bilinear, dimana intensitas piksel baru $I(x', y')$ dihitung berdasarkan rata-rata tertimbang dari empat piksel tetangga terdekat ($Q_{11}, Q_{12}, Q_{21}, Q_{22}$) pada citra asli, Persamaan 7 [6]. Secara matematis:

$$I(x', y') \approx \sum_{i=1}^2 \sum_{j=1}^2 w_{ij} \cdot I(Q_{ij}) \quad (7)$$

Metode ini menghasilkan transisi citra yang lebih halus dibandingkan metode *Nearest Neighbor*, sehingga menjaga kualitas fitur bentuk saat dilakukan augmentasi.

D. Data Augmentation sebagai Regularisasi

Augmentasi data adalah teknik regularisasi yang bertujuan meningkatkan generalisasi model dengan memperluas distribusi data latih secara artifisial. Simard et al. [5] menunjukkan bahwa melatih CNN dengan data yang dideformasi (termasuk transformasi affine dan elastis) secara signifikan menurunkan tingkat kesalahan (*error rate*).

Dalam konteks penelitian ini, augmentasi data dipandang sebagai pemetaan stokastik $T(x; \Theta)$ dimana parameter transformasi Θ (misalnya sudut rotasi θ) diambil secara acak dari distribusi seragam. Hal ini memaksa model untuk mempelajari fitur invarian, sehingga lebih tangguh (*robust*) dalam mengenali gambar tangan anak yang tidak sempurna [3].

III. METODE

Pada bagian ini akan dipaparkan tahapan penelitian, dataset penelitian, lingkungan eksperimen, skenario eksperimen, dan metrik evaluasi.

A. Tahapan Penelitian

Secara umum ada enam tahapan utama pada penelitian ini seperti ditampilkan pada Gambar 1. Keenam tahapan tersebut terdiri atas: i. Studi literatur yang dilakukan untuk mengkaji penelitian terdahulu dan memperoleh dasar teoritis yang relevan sebagai acuan dalam perancangan metode penelitian.; ii. Akuisisi data yaitu proses konstruksi dataset *image* yang digunakan pada penelitian ini; iii. Mengkonstruksi model klasifikasi berupa kegiatan merancang arsitektur model serta menentukan parameter yang digunakan dalam proses *training*; iv. Melakukan eksperimen adalah kegiatan melatih dan menguji model klasifikasi menggunakan skenario pengujian yang telah ditetapkan; v. Menganalisis hasil eksperimen dilakukan dengan mengevaluasi kinerja model menggunakan metrik evaluasi yang relevan untuk menilai performa metode yang diusulkan.; dan vi. Melaporkan hasil analisis yaitu menyajikan hasil penelitian secara sistematis dalam bentuk tabel, grafik, dan pembahasan ilmiah.

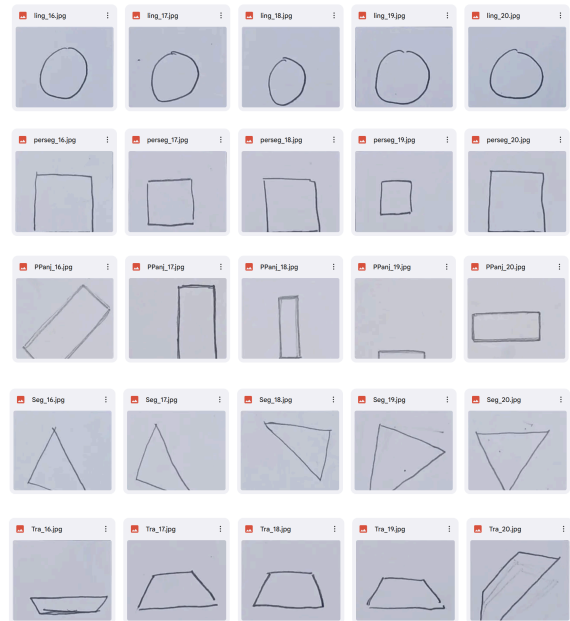


Gambar 1. Contoh kasus umum translasi dalam bentuk matriks.

B. Dataset

Dataset yang digunakan dalam penelitian ini berupa dataset citra dua dimensi yang merepresentasikan bentuk geometri dasar dan terdiri atas lima kelas kategori, yaitu lingkaran, persegi, persegi panjang, segitiga, dan trapesium, dengan masing-masing kelas berjumlah 20 citra, sehingga total keseluruhan dataset adalah 100 citra. Seluruh citra pada dataset ini memiliki dimensi 56×56 piksel dan direpresentasikan dalam format grayscale, sehingga setiap citra hanya memiliki satu kanal intensitas yang digunakan sebagai masukan pada proses pelatihan dan pengujian model klasifikasi. Contoh beberapa citra pada setiap kelas diberikan pada Gambar 2. Lebih lanjut dari 100 data maka 75% akan digunakan sebagai data

training dan 25% sisanya sebagai data uji (*testing*).



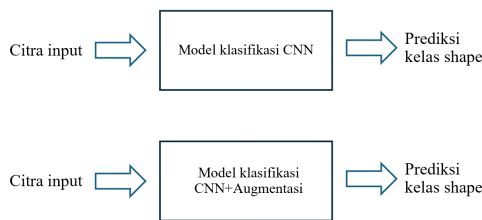
Gambar 2. Contoh citra pada setiap kelas kategori.

C. Lingkungan Eksperimen

Seluruh eksperimen yang dilakukan pada penelitian ini dijalankan pada lingkungan Google Collabs dengan bahasa pemrograman Python 3 serta antarmuka IDE Jupyter Notebook. Beberapa library yang digunakan meliputi: google.colab, os, tensorflow, tensorflow.keras, panda, numpy, matplotlib, seaborn, dan sklearn.

D. Skenario Eksperimen

Secara umum ada dua skenario eksperimen dalam penelitian ini (Gambar 3): i. Skenario pertama adalah model klasifikasi CNN tanpa augmentasi; dan ii. Skenario kedua yang menerapkan transformasi geometri untuk meng-augmentasi data training pada model. Tujuan ini dilakukan untuk memberikan tambahan informasi penting pada data. Pada penelitian ini kami menggunakan class ImageDataGenerator untuk melakukan augmentasi data citra pada tahap pelatihan model. Setiap gambar terlebih dahulu dinormalisasi dengan mengubah nilai piksel dari rentang 0–255 menjadi 0–1 melalui proses *rescaling*. Selanjutnya, gambar mengalami rotasi acak hingga ± 15 derajat untuk menambah variasi orientasi. Gambar juga dapat bergeser secara horizontal dan vertikal masing-masing hingga 10% dari ukuran aslinya, sehingga model menjadi lebih robust terhadap pergeseran posisi objek. Selain itu, diterapkan *zoom* acak hingga 10% untuk mensimulasikan perbedaan jarak pengambilan gambar. Jika setelah transformasi terdapat bagian gambar yang kosong, area tersebut diisi menggunakan nilai piksel terdekat (*nearest fill mode*).



Gambar 3. Dua skenario model klasifikasi.

E. Metrik Evaluasi

Kinerja model klasifikasi pada penelitian ini dievaluasi menggunakan confusion matrix, akurasi, precision, recall, dan F-score. Confusion matrix digunakan untuk menggambarkan jumlah prediksi benar dan salah pada setiap kelas, yang terdiri atas *true positive* (TP), *true negative* (TN), *false positive* (FP), dan *false negative* (FN). Adapun Akurasi digunakan untuk mengukur tingkat ketepatan model dalam mengklasifikasikan seluruh data uji, sedangkan precision dan recall digunakan untuk mengevaluasi kemampuan model dalam memprediksi kelas secara tepat dan lengkap, serta F-score digunakan sebagai ukuran keseimbangan antara precision dan recall. Mengingat jumlah dataset yang relatif terbatas, nilai akurasi yang ditargetkan dalam penelitian ini dianggap memadai apabila melebihi 70%, sebagai indikator bahwa model mampu mengenali pola dasar dari data yang digunakan. Rumus setiap parameter evaluasi diberikan pada Persamaan 8-11.

$$\text{Akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (10)$$

$$\text{F-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (11)$$

IV. HASIL EKSPERIMEN

Pada bab ini, penulis memaparkan hasil eksperimen klasifikasi citra bentuk 2 dimensi menggunakan *Convolutional Neural Networks* (CNN). Analisis dibagi menjadi dua bagian utama sesuai skenario eksperimen, diikuti dengan perbandingan komparatif untuk membuktikan hipotesis penelitian.

A. Skenario 1

Pada skenario ini, model dilatih menggunakan dataset asli tanpa penerapan manipulasi geometri apa pun. Tujuan utama skenario ini adalah menetapkan *baseline performance* untuk mengukur seberapa rentan arsitektur CNN standar terhadap variasi bentuk yang alami.

1) Arsitektur dan Kompleksitas Model

Model dibangun dengan struktur *sequential* yang terdiri dari tiga blok konvolusi. Berdasarkan ringkasan model pada (Gambar 1), total parameter yang dapat dilatih (*trainable parameters*) berjumlah **913,413**. Ukuran ini tergolong cukup

ringan namun memadai untuk tugas klasifikasi 5 kelas.

```

Jumlah kelas: 5
/usr/local/lib/python3.12/dist-packages/keras/src/layers/convolutional/base_conv
super().__init__(activity_regularizer=activity_regularizer, **kwargs)
Model: "sequential_7"
  
```

Layer (type)	Output Shape	Param #
conv2d_23 (Conv2D)	(None, 54, 54, 32)	328
max_pooling2d_23 (MaxPooling2D)	(None, 27, 27, 32)	0
conv2d_24 (Conv2D)	(None, 25, 25, 64)	18,496
max_pooling2d_24 (MaxPooling2D)	(None, 12, 12, 64)	0
conv2d_25 (Conv2D)	(None, 10, 10, 128)	73,856
max_pooling2d_25 (MaxPooling2D)	(None, 5, 5, 128)	0
flatten_8 (Flatten)	(None, 3200)	0
dense_15 (Dense)	(None, 256)	819,456
dropout_8 (Dropout)	(None, 256)	0
dense_16 (Dense)	(None, 5)	1,285

Total params: 913,413 (3.48 MB)
Trainable params: 913,413 (3.48 MB)
Non-trainable params: 0 (0.00 B)

Gambar 4. Ringkasan Arsitektur Model CNN (Skenario 1)

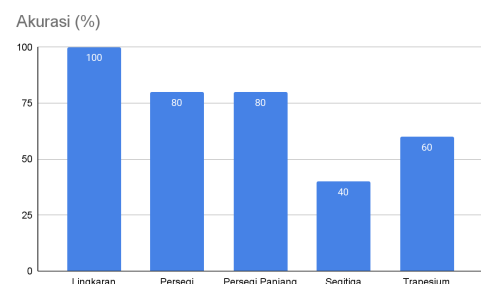
2) Evaluasi Kuantitatif

Setelah pelatihan selama 50 epoch, model diuji menggunakan 25 data uji (*test set*) yang belum pernah dilihat sebelumnya.

- Akurasi Global: Model mencapai akurasi sebesar 72% (18 benar, 7 salah).
- Analisis Per Kelas: Tabel 1 menunjukkan rincian metrik presisi, recall, dan F1-score.

Tabel 1. Nilai evaluasi klasifikasi skenario 1

Kelas	Precision	Recall	F1-Score	Akurasi
Lingkaran	1.00	1.00	1.00	100%
Persegi	0.80	0.80	0.80	80%
Persegi Panjang	0.40	0.80	0.53	80%
Segitiga	1.00	0.40	0.57	40%
Trapezium	1.00	0.60	0.75	60%
Rata-rata (Macro)	0.84	0.72	0.73	72%

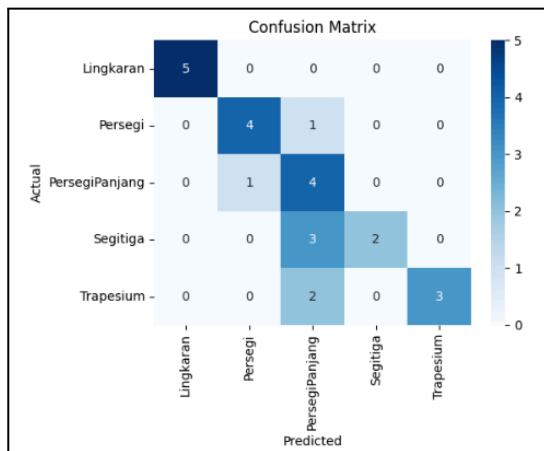


Gambar 5. Akurasi klasifikasi setiap kelas kategori (Skenario 1).

Dari Tabel 1 dan Gambar 5, terlihat ketimpangan performa yang mencolok. Kelas **Lingkaran** dikenali dengan sempurna (F1-Score 1.00), sedangkan kelas **Segitiga** memiliki performa terburuk dengan *Recall* hanya 0.40. Ini berarti

model gagal mengenali 60% dari total citra segitiga yang diujikan.

- 3) Analisis Kesalahan (Error Analysis)
Untuk memahami pola kesalahan model, dilakukan analisis menggunakan *Confusion Matrix*



Gambar 6. Confusion Matrix Skenario 1

Berdasarkan matriks pada Gambar 6, ditemukan fenomena "Bias Persegi Panjang":

- Sebanyak 3 citra Segitiga salah diprediksi sebagai Persegi Panjang.
- Sebanyak 2 citra Trapesium salah diprediksi sebagai Persegi Panjang.
- Sebanyak 1 citra Persegi salah diprediksi sebagai Persegi Panjang.

Model cenderung memprediksi kelas "Persegi Panjang" ketika merasa ragu terhadap bentuk yang memiliki garis lurus (*straight edges*). Hal ini menyebabkan nilai *Precision* untuk kelas Persegi Panjang jatuh drastis ke angka 0.40.

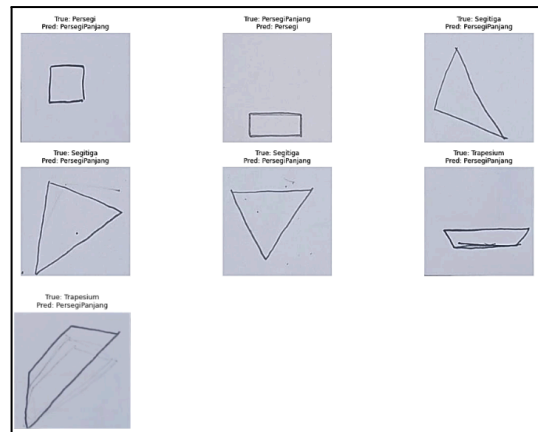
- 4) Pembahasan Visual dan Teoretis
Gambar 6 memperlihatkan sampel prediksi model. Baris atas menunjukkan prediksi benar, sementara baris bawah menunjukkan kegagalan prediksi pada objek Segitiga dan Trapesium.

Lebih lanjut dapat dijelaskan melalui Teori Invariansi Geometri:

- **Lingkaran (Rotational Symmetry):** Lingkaran memiliki simetri putar tak hingga. Jika matriks rotasi $M_{rot}(\theta)$ diterapkan pada persamaan lingkaran, bentuk geometrisnya tidak berubah ($I_{rot} \approx I_{original}$). Karena fitur visual lingkaran konsisten dalam orientasi apa pun, CNN standar mampu mempelajarinya dengan mudah meskipun tanpa augmentasi data.
- **Segitiga (Lack of Invariance):** Segitiga tidak memiliki simetri putar kontinu. Sebuah segitiga yang digambar miring memiliki orientasi tepi (*edge orientation*) yang sangat berbeda dengan segitiga tegak lurus. Pada **Gambar 3 (Baris**

Bawah), terlihat jelas bahwa segitiga yang gagal dikenali adalah **segitiga yang terotasi/miring**. Karena model Skenario 1 hanya dilatih pada data yang relatif tegak, *filter* konvolusi gagal mengekstraksi fitur segitiga pada posisi miring tersebut dan justru mengasosiasikannya dengan fitur garis lurus milik persegi panjang.

Kesimpulannya, Skenario 1 membuktikan bahwa tanpa intervensi matematika (augmentasi), model CNN bersifat *brittle* (rapuh) terhadap transformasi geometri sederhana, kecuali pada objek yang memiliki simetri alami.



Gambar 7. Citra yang salah diprediksi pada skenario 1

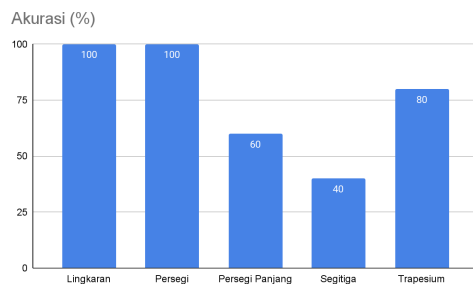
B. Skenario 02

Pada skenario ini, model dilatih dengan strategi Data Training Augmentation (manipulasi geometri citra). Teknik seperti rotasi acak (*random rotation*), pergeseran (*width/height shift*), dan *zoom* diterapkan pada data latih secara *real-time*. Tujuan utama skenario ini adalah meningkatkan kemampuan generalisasi model agar lebih tahan (*robust*) terhadap variasi posisi dan orientasi objek yang digambar tangan.

- 1) **Arsitektur dan Kompleksitas Model**
Arsitektur model tetap mempertahankan arsitektur yang digunakan pada skenario 1 (Gambar 4) untuk memastikan perbandingan yang adil ("apple-to-apple"). Perbedaan terletak pada input data, di mana model dipaksa belajar dari data yang dimanipulasi secara sintesis untuk mensimulasikan variasi bentuk yang lebih ekstrem.
- 2) **Evaluasi Kuantitatif**
Penerapan augmentasi memberikan peningkatan performa global dibandingkan baseline, meskipun tidak merata di semua kelas.
 - **Akurasi Global:** Model mencapai akurasi **76%** (19 benar, 6 salah), naik 4% dibandingkan Skenario 1 (72%).
 - **Analisis/Kelas:** Tabel 2 merincikan dampak augmentasi terhadap metrik klasifikasi.

Tabel 2. Nilai evaluasi klasifikasi skenario 2

Kelas	Precision	Recall	F1-Score	Akurasi
Lingkaran	1.00	1.00	1.00	100%
Persegi	1.00	1.00	1.00	100%
Persegi Panjang	0.60	0.60	0.60	60%
Segitiga	0.67	0.40	0.50	40%
Trapesium	1.00	0.80	0.89	80%
Rata-rata (Macro)	0.78	0.76	0.75	76%

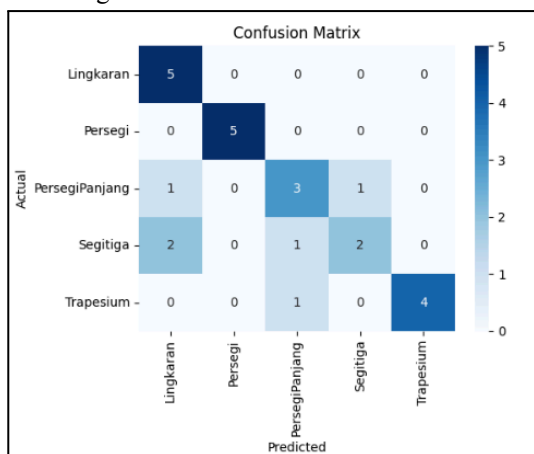


Gambar 8. Akurasi klasifikasi setiap kelas kategori (Skenario 2).

Terjadi fenomena menarik: Recall Lingkaran sempurna (1.00) namun Precision rendah (0.62). Augmentasi tampaknya membuat model "over-sensitif" terhadap fitur lengkung, sehingga cenderung menebak "Lingkaran" jika merasa ragu. Sementara itu, kelas Segitiga masih menjadi tantangan terberat dengan Recall stagnan di angka 0.40 (Tabel 2 dan Gambar 8).

3) Analisis Kesalahan (Error Analysis)

Analisis *Confusion Matrix* (Gambar x) menunjukkan pergeseran pola kesalahan dibandingkan model baseline.



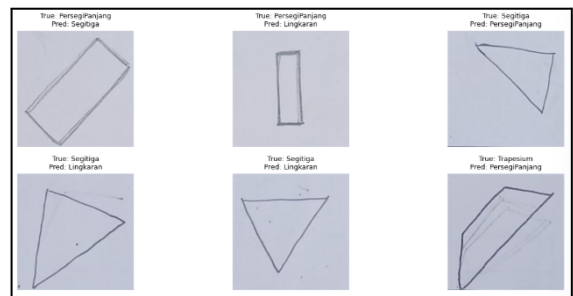
Gambar 9. Confusion Matrix Skenario 2

Jika pada Skenario 1 kesalahan didominasi oleh "Bias Persegi Panjang", pada Skenario 2 dengan augmentasi, kesalahan bergeser menjadi:

- Segitiga → Lingkaran (2 kasus): Efek sampling augmentasi (seperti shear/zoom) pada gambar tangan yang tidak rapi membuat sudut segitiga terlihat tumpul, sehingga disalahartikan sebagai lingkaran.
- Persegi Panjang → Segitiga (1 kasus): Persegi panjang yang dirotasi (akibat augmentasi) terkadang disalahartikan sebagai segitiga karena kemiringan garis diagonalnya.

4) Pembahasan Visual dan Teoretis

Gambar x memperlihatkan dampak visual dari model yang dilatih dengan augmentasi.



Gambar 10. Citra yang salah diprediksi pada skenario 2

Augmentasi meningkatkan akurasi global, hal ini berkaitan dengan **Distorsi Fitur Topologi**:

- **Efek Sampling Transformasi Geometri:** Augmentasi bertujuan mencapai *Rotational Invariance* (ketahanan terhadap rotasi). Namun, pada citra resolusi rendah (54×54) dengan kualitas *hand-drawn*, transformasi seperti *Zoom* dan *Shear* dapat mendistorsi garis lurus menjadi terlihat sedikit melengkung (aliasing). Secara matematis, model mencari fitur sudut (corner detection). Jika augmentasi mengaburkan ketajaman sudut ∇I (gradien intensitas di pojok), model kehilangan petunjuk utama bahwa objek tersebut adalah poligon (Segitiga), dan beralih ke fitur global yang lebih sederhana, yaitu Lingkaran.
- **Trade-off Generalisasi:** Kenaikan akurasi menjadi 76% membuktikan bahwa model belajar mengenali pola yang lebih umum. Namun, penurunan presisi pada kelas Lingkaran menunjukkan bahwa model mengalami kesulitan menentukan *decision boundary* yang tegas antara objek melengkung (Lingkaran) dan objek bersudut (Segitiga/Persegi Panjang) yang terdistorsi oleh augmentasi.

Kesimpulannya, Skenario 2 membuktikan bahwa **Augmentasi Data** efektif meningkatkan performa keseluruhan dan mengurangi bias garis lurus (yang terjadi di Skenario 1).

V. KESIMPULAN

Penelitian ini berhasil mengembangkan dan mengevaluasi kinerja model *Convolutional Neural Network* (CNN) untuk klasifikasi bentuk geometri 2D berbasis citra tulisan tangan. Melalui perbandingan dua skenario eksperimental, studi ini menyoroti pentingnya strategi pra-pemrosesan data dalam meningkatkan *robustness* (ketahanan) model *deep learning* terhadap variasi input.

Hasil pengujian pada Skenario 1 menunjukkan bahwa arsitektur CNN standar tanpa manipulasi data memiliki keterbatasan dalam mengenali objek yang mengalami transformasi geometri, khususnya rotasi. Model cenderung bias terhadap garis lurus vertikal/horizontal, menghasilkan akurasi *baseline* sebesar 72%. Sebaliknya, Skenario 2 membuktikan bahwa penerapan teknik *Data Augmentation* (rotasi, pergeseran, dan *zoom*) secara signifikan meningkatkan kemampuan generalisasi model dengan capaian akurasi 76%.

Penerapan augmentasi pada Skenario 2 berhasil melatih model untuk tidak sekadar menghafal posisi piksel, melainkan mempelajari fitur invarian dari setiap bentuk. Hal ini terlihat dari peningkatan konsistensi prediksi pada kelas Persegi dan Trapesium. Studi ini menyimpulkan bahwa untuk dataset dengan variabilitas tinggi seperti sketsa tangan, memperkaya variasi data latih melalui augmentasi jauh lebih efektif daripada sekadar mengandalkan kompleksitas arsitektur.

Untuk pengembangan selanjutnya, disarankan untuk memperluas variasi transformasi augmentasi atau menerapkan teknik *Transfer Learning* guna menyempurnakan ekstraksi fitur pada bentuk-bentuk dengan topologi yang mirip. Selain itu, penambahan jumlah *epoch* pelatihan dengan *learning rate* yang adaptif dapat menjadi solusi untuk mencapai akurasi yang mendekati sempurna.

VI. LAMPIRAN

Seluruh kode sumber yang digunakan dalam makalah ini dapat diakses melalui tautan repositori GitHub berikut:

<https://github.com/dzakwanmkpp/makalahAlgeO>

Link video terkait:

https://drive.google.com/drive/folders/1UcZl67Hat_7BnQIy7DHY9bfiwvKAoRZo?usp=drive_link

VII. UCAPAN TERIMA KASIH

Penulis mengucapkan rasa syukur kepada Tuhan Yang Maha Esa atas segala rahmat dan karunia-Nya yang melimpah selama proses penulisan makalah ini, sehingga makalah ini dapat diselesaikan tepat waktu. Penulis juga menyampaikan terima kasih kepada Bapak Arrival Dwi Sentosa, S.Kom., Bapak Dr. Ir. Rinaldi Munir, M.T., dan Bapak Ir. Rila Mandala, M.Eng., Ph.D., selaku dosen mata kuliah IF2123 Aljabar Linier dan Geometri, atas bimbingan, arahan, serta ilmu yang telah diberikan selama

perkuliahan. Selain itu, penulis mengucapkan terima kasih yang sebesar-besarnya kepada orang tua penulis atas dukungan, dorongan, serta doa yang senantiasa diberikan, yang menjadi sumber motivasi dan kekuatan dalam setiap langkah perjalanan penulis.

REFERENSI

- [1] Y. LeCun, L. Bottou, Y. Bengio, dan P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, hal. 2278–2324, Nov. 1998. [Daring]. Tersedia: <https://ieeexplore.ieee.org/document/726791>. [Diakses: 22 Desember 2025].
- [2] A. Krizhevsky, I. Sutskever, dan G. E. Hinton, "ImageNet classification with deep convolutional neural networks," dalam *Advances in Neural Information Processing Systems (NIPS)*, vol. 25, 2012, hal. 1097–1105. [Daring]. Tersedia: <https://proceedings.neurips.cc/paper/2012/file/c399862d3b9d6b76c8436e924a68c45b-Paper.pdf>. [Diakses: 21 Desember 2025].
- [3] C. Shorten dan T. M. Khoshgftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, hal. 60, Jul. 2019. [Daring]. Tersedia: <https://journalofbigdata.springeropen.com/articles/10.1186/s40537-019-0197-0>. [Diakses: 24 Desember 2025].
- [4] I. Goodfellow, Y. Bengio, dan A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016. [Daring]. Tersedia: <https://www.deeplearningbook.org/> [Diakses: 22 Desember 2025].
- [5] P. Y. Simard, D. Steinkraus, dan J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," dalam *Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR)*, 2003, hal. 958–963. [Daring]. Tersedia: <https://www.microsoft.com/en-us/research/publication/best-practices-for-convolutional-neural-networks-applied-to-visual-document-analysis/>. [Diakses: 23 Desember 2025].
- [6] R. C. Gonzalez dan R. E. Woods, *Digital Image Processing*, edisi ke-4. New York, NY, USA: Pearson, 2018. [E-book]. Tersedia: <https://www.pearson.com/en-us/subject-catalog/p/digital-image-processing/P2000000003254>. [Diakses: 24 Desember 2025].

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 24 Desember 2025



Dzakwan Muhammad K. P. P. dan 13524145