

# Sistem Perhitungan Jumlah Sendok dan Garpu Menggunakan Segmentasi Citra dan CNN

Daffa Romyz Aufo / 13520162  
Program Studi Teknik Informatika  
Sekolah Teknik Elektro dan Informatika  
Institut Teknologi Bandung, Jalan Ganesha 10 Bandung  
E-mail (gmail): 13520162@std.stei.itb.ac.id

**Abstract**— Konsumsi pangan merupakan kebutuhan yang sangat penting. Sendok dan garpu telah menjadi alat yang paling populer untuk membantu manusia dalam memfasilitasi konsumsi makanan. Peralatan seperti sendok dan garpu merupakan peralatan penting di industri seperti restoran. Kelancaran bisnis sebuah restoran sangat bergantung pada pemeliharaan inventaris peralatan makan. Pengelolaan inventaris sendok dan garpu di restoran masih dilakukan secara manual sehingga memakan waktu dan membuka kemungkinan terjadinya kesalahan penghitungan. Makalah ini mengatasi tantangan penghitungan peralatan manual melalui segmentasi gambar dan *Convolutional Neural Network* (CNN) untuk mengotomatiskan proses penghitungan sendok dan garpu.

**Kata kunci**—*sendok; garpu; perhitungan objek; segmentasi citra; CNN*

## I. PENDAHULUAN

Dalam kehidupan manusia, konsumsi pangan merupakan kebutuhan yang sangat penting. Manusia telah berevolusi tidak hanya dalam preferensi kuliner tetapi juga dalam alat yang digunakan untuk memfasilitasi konsumsi makanan. Sendok dan garpu telah menjadi alat yang paling populer. Alat makan tersebut mengubah proses makan menjadi lancar dan efisien.

Peralatan seperti sendok dan garpu tidak hanya sekedar untuk konsumsi individu, tetapi juga peralatan penting di industri seperti restoran, dimana permintaan akan peralatan penting ini tidak ada habisnya. Kelancaran bisnis sebuah restoran sangat bergantung pada pemeliharaan inventaris peralatan yang akurat, terutama mengingat banyaknya jumlah yang dibutuhkan untuk melayani banyaknya pelanggan.

Pengelolaan inventaris sendok dan garpu di restoran masih dilakukan secara manual dan memakan waktu. Menghitung setiap peralatan satu per satu tidak hanya menghabiskan sumber daya manusia yang berharga namun juga membuka kemungkinan terjadinya kesalahan penghitungan. Dalam dunia industri restoran yang bergerak cepat, dimana presisi dan kecepatan adalah hal yang terpenting, kebutuhan akan sistem manajemen inventaris yang lebih efisien dan akurat menjadi hal yang sangat penting.

Makalah ini memperkenalkan pendekatan baru untuk mengatasi tantangan yang terkait dengan penghitungan peralatan manual di restoran melalui segmentasi gambar dan

*Convolutional Neural Network* (CNN). Sistem ini bertujuan untuk mengotomatiskan proses penghitungan sendok dan garpu dalam gambar, sehingga menyederhanakan alur kerja manajemen inventaris untuk restoran. Sistem ini tidak hanya menjanjikan penghematan waktu yang berharga namun juga memastikan tingkat akurasi yang lebih tinggi dalam penghitungan peralatan makan, sehingga berkontribusi terhadap efisiensi operasional restoran secara keseluruhan.

## II. LANDASAN TEORI

### A. Segmentasi Citra

Segmentasi citra adalah operasi mempartisi citra menjadi sejumlah region atau objek, setiap region terdiri dari sekumpulan pixel yang terhubung satu sama lain. Segmentasi citra menjadi sejumlah region bertujuan untuk:

1. Membagi citra menjadi segmen-segmen atau objek-objek yang berbeda
2. Memisahkan objek dengan latar belakang

Dengan membagi citra menjadi sejumlah segmen, kita dapat memproses hanya segmen penting atau segmen tertentu di dalam citra daripada memproses seluruh bagian citra.

Tujuan segmentasi citra adalah menemukan bagian citra yang koheren atau objek spesifik. Citra disegmentasi berdasarkan properti yang dipilih seperti kecerahan, warna, tekstur, dan sebagainya. Segmentasi membagi citra menjadi sejumlah segmen yang terhubung, tiap segmen bersifat homogen berdasarkan properti yang dipilih.

Segmentasi citra umumnya dikelompokkan berdasarkan dua pendekatan:

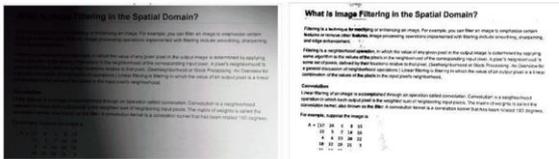
1. Diskontinuitas mempartisi citra berdasarkan perubahan nilai intensitas pixel yang cepat seperti tepi.
2. Similarity mempartisi citra berdasarkan kemiripan area menurut properti yang ditentukan.

Diskontinuitas dapat menggunakan metode seperti operator Sobel, Prewitt, Canny, Roberts, Laplacian, dan LoG. Similarity menggunakan metode pendekatan pengambangan (*thresholding*), *region growing*, *split and merge*, serta *clustering*.

Pengambangan (*thresholding*) dapat digunakan untuk memisahkan objek dengan latar belakangnya. Pengambangan dilakukan berdasarkan pada nilai intensitas pixel-pixel dan sebuah nilai ambang T. Setiap pixel pada posisi (i, j) pada citra di mana  $f(i, j) > T$  disebut titik objek, jika tidak maka akan disebut latar belakang.

$$f_B(i, j) = \begin{cases} A, & f_g(i, j) \leq T \\ B, & \text{lainnya} \end{cases}$$

Teknik pengambangan dibagi menjadi pengambangan global, lokal, dan adaptif. Pengambangan global menggunakan nilai ambang T untuk seluruh gambar. Pengambangan lokal membagi gambar menjadi beberapa wilayah yang memiliki nilai ambang masing-masing. Pengambangan adaptif mengubah nilai ambang secara dinamis pada citra.



Gambar 2.1 Pengambangan pada citra teks

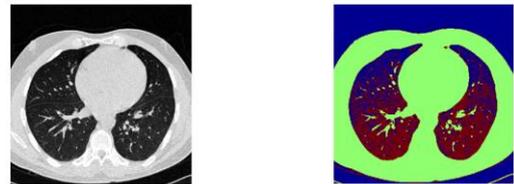
*Region growing* dilakukan dengan mengelompokkan pixel atau sub-region yang tumbuh menjadi region yang lebih besar. *Region growing* dimulai dengan umpan yang berisi himpunan beranggota satu atau lebih pixel dari *region* yang potensial, dan dari sini region berkembang dengan menambahkan pada umpan pixel-pixel tetangga yang memiliki properti yang mirip dengan umpan, lalu berhenti jika pixel-pixel tetangga tidak mirip lagi.

*Split and merge* menggunakan algoritma *divide and conquer*. Citra awalnya dibagi (*split*) menjadi sejumlah region yang disjoint. Kemudian region-region tetangga yang mirip akan digabung (*merge*).



Gambar 2.2 *Split and merge* pada citra lena

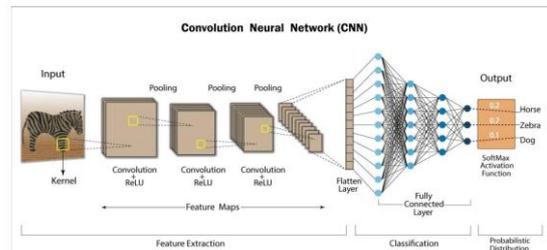
*Clustering* mengelompokkan titik-titik yang mirip dalam kelompok yang sama. K-means *clustering* merupakan algoritma *clustering* yang paling populer. K-means berawal dengan memilih k *centroid* secara acak. Kemudian, menentukan tiap titik ke centroid terdekat. Lalu, dihitung rata-rata koordinat semua titik. Koordinat akan menjadi centroid baru. Langkah-diulang hingga kelompok tidak berubah.



Gambar 2.3 K-means pada citra medis

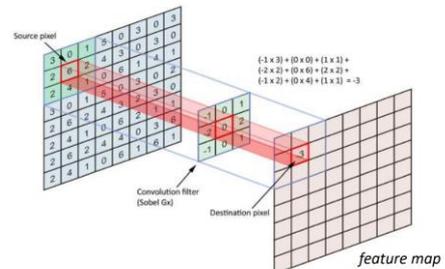
### B. Convolutional Neural Network

*Convolutional Neural Network* (CNN atau ConvNet) adalah algoritma pembelajaran mendalam yang populer, umumnya digunakan untuk memproses data yang memiliki topologi seperti grid seperti citra. CNN merupakan arsitektur jaringan untuk pembelajaran mendalam yang belajar langsung dari data, dengan menghilangkan kebutuhan untuk melakukan ekstraksi fitur secara manual. CNN terdiri dari 3 lapisan utama yaitu convolutional layer, pooling layer, dan fully connected layer.



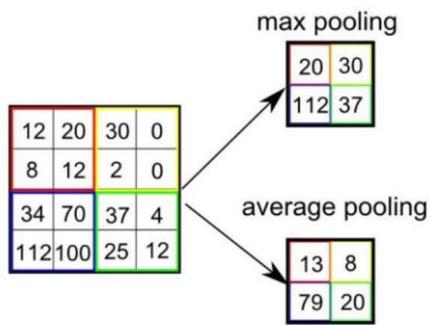
Gambar 2.4 Arsitektur CNN

*Convolutional layer* melakukan operasi konvolusi pada citra masukan dengan sejumlah penapis (*filter*). Tiap penapis menghasilkan luaran yang disebut *feature map*. Hasil konvolusi akan dimasukkan ke fungsi aktivasi ReLU. ReLU adalah layer tambahan yang memungkinkan pelatihan yang lebih cepat dan efektif dengan memetakan nilai negatif ke nol dan mempertahankan nilai positif. Pada dasarnya ReLU adalah operasi per-pixel dengan cara mengganti nilai negatif pixel di dalam *feature map* menjadi nol.



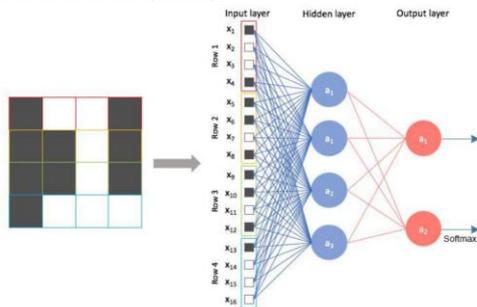
Gambar 2.5 Konvolusi

*Pooling layer* bertanggung jawab untuk mengurangi ukuran spasial dari matriks fitur hasil konvolusi. Hal ini bertujuan untuk mengurangi daya komputasi yang diperlukan untuk memproses data melalui pengurangan dimensi. Terdapat dua jenis pooling yaitu *max pooling* dan *average pooling*. *Max pooling* mengembalikan nilai maksimum dari bagian gambar yang dicakup oleh *kernel*. *Average pooling* mengembalikan rata-rata semua nilai dari bagian gambar yang dicakup oleh *kernel*.



Gambar 2.6 Pooling

fully connected layer menghasilkan vektor dimensi K, dalam hal ini K adalah jumlah kelas yang dapat diprediksi oleh jaringan. Vektor ini berisi probabilitas untuk setiap kelas dari setiap gambar yang diklasifikasikan. Lapisan terakhir dari arsitektur CNN menggunakan fungsi softmax untuk menyediakan luaran klasifikasi



Gambar 2.7 Fully-connected layer

### C. Dataset

Dataset yang digunakan untuk pelatihan model CNN berasal dari [4]. Dataset berisi 199 citra garpu dan 151 citra sendok. Dataset citra memiliki variasi latar belakang, jumlah objek, dan bentuk objek. Citra yang digunakan adalah citra yang memiliki satu objek dan latar belakang yang homogen. Setelah pemilihan gambar didapatkan 95 citra garpu dan 95 citra sendok.

### III. METODE PENYELESAIAN MASALAH

Sistem perhitungan jumlah sendok dan garpu memiliki tahapan sebagai berikut.

#### 1. Membaca citra masukan

Citra yang akan dihitung jumlah sendok dan garpunya adalah seperti berikut.

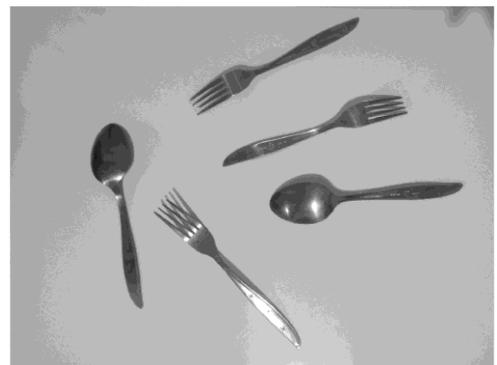


Gambar 3.1 Citra masukan

Citra memiliki beberapa objek sendok dan garpu. Objek sendok terdapat 2 buah dan objek garpu terdapat 3 buah. Latar belakang citra berwarna putih.

#### 2. Konversi ke citra grayscale

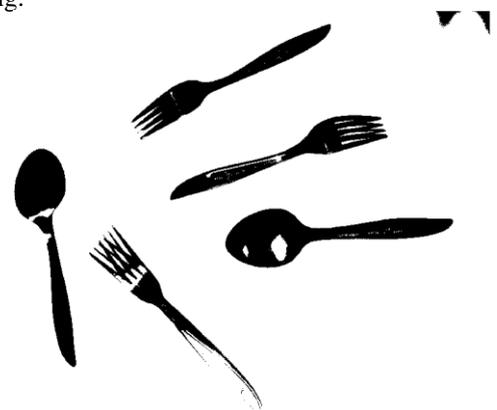
Sebelum dilakukan perhitungan, akan dilakukan beberapa pra-pemrosesan. Konversi ke citra grayscale dilakukan dengan fungsi bawaan matlab **rgb2gray**.



Gambar 3.2 Citra grayscale hasil konversi

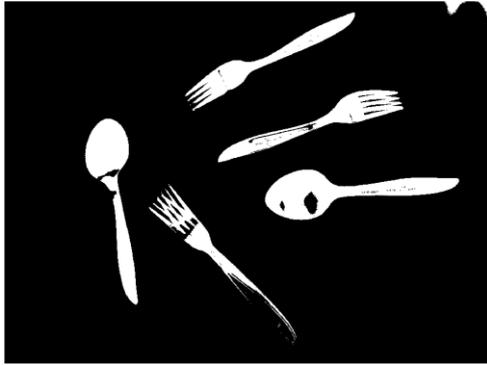
#### 3. Konversi ke citra biner

Konversi ke citra biner dilakukan dengan metode otsu menggunakan fungsi bawaan matlab **imbinarize**. Konversi ke citra biner dilakukan untuk membagi citra menjadi objek dan latar belakang.



Gambar 3.3 Citra biner hasil konversi

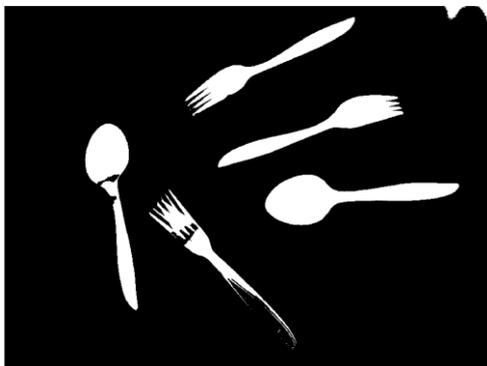
Hasil binerisasi menghasilkan citra dengan objek bernilai 0 (warna hitam) dan latar belakang bernilai 1 (warna putih). Hal ini berbeda dengan yang diinginkan sehingga perlu dilakukan operasi citra negatif.



Gambar 3.4 Citra negatif

#### 4. Menutup lubang

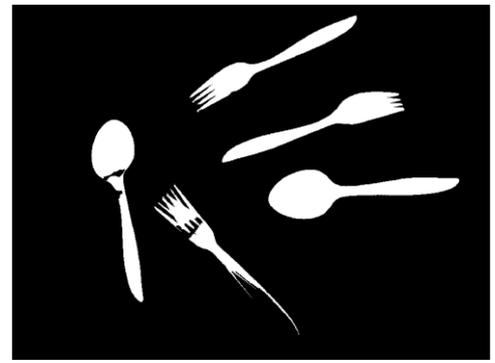
Citra hasil binerisasi memiliki beberapa lubang. Lubang-lubang ini akan mengakibatkan citra pada lubang tersebut hilang sehingga tidak dapat dikenali oleh CNN. Maka dari itu, lubang-lubang tersebut akan diisi dengan fungsi **imfill**.



Gambar 3.5 Citra yang terisi

#### 5. Penapisan luas

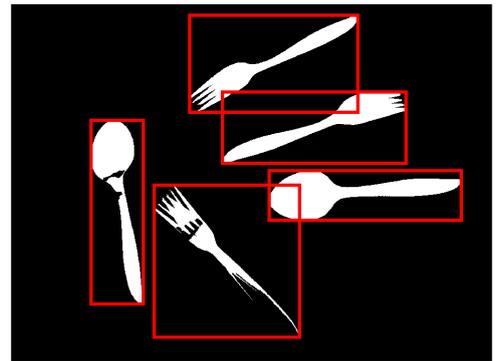
Citra biner yang telah terisi memiliki objek yang tidak diinginkan. Objek tersebut terletak di kanan atas citra. Objek ini perlu dihilangkan agar tidak dimasukkan pada tahap perhitungan. Penghapusan dilakukan dengan penapis luas **bwareopen**. Penapis luas akan menghilangkan objek pada citra biner yang memiliki luas dibawah suatu ambang batas. Ambang batas yang digunakan pada proses ini sebesar 10000.



Gambar 3.6 Citra hasil penapisan luas

#### 6. Penentuan *bounding box*

Citra biner hasil penapisan luas hanya akan memiliki objek yang ingin dihitung. Untuk melakukan perhitungan, objek-objek tersebut harus dipisahkan menjadi citra individu. Maka dari itu, diperlukan sebuah kotak batas (*bounding box*) untuk setiap objek. Penentuan *bounding box* menggunakan fungsi **regionprops**.



Gambar 3.7 Citra *bounding box*

#### 7. Segmentasi

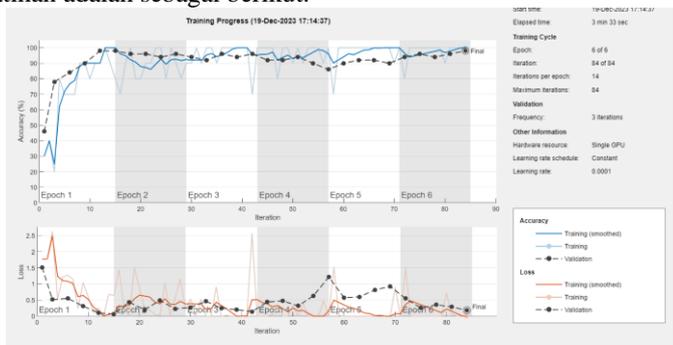
Citra kemudian disegmentasi berdasarkan *bounding box* masing-masing objek. Tahap ini akan menghasilkan beberapa citra sesuai dengan jumlah objek. Pada citra ini didapatkan lima buah citra yang terdiri dari 2 citra sendok dan 3 citra garpu.



Gambar 3.8 Citra hasil segmentasi

#### 8. Klasifikasi dengan CNN

Citra hasil segmentasi akan diklasifikasi oleh model CNN. Model CNN yang digunakan adalah AlexNet yang sudah dilatih. Model AlexNet akan dilatih ulang menggunakan dataset citra sendok dan garpu. Dataset citra memiliki variasi latar belakang, jumlah objek, dan bentuk objek. Citra yang digunakan adalah citra yang memiliki satu objek dan latar belakang yang homogen. Dataset memiliki 95 citra garpu dan 95 citra sendok. Untuk setiap kelas, data dibagi menjadi 70 data latih dan 25 data validasi. Sebelum digunakan, citra akan disegmentasi dan diubah ukurannya. Segmentasi dilakukan untuk menghilangkan latar belakang. Ukuran citra diubah menjadi 227x227 sesuai dengan lapisan masukan model. Model dilatih dengan ukuran mini batch 10, jumlah epoch 6, learning rate 1e-4, dan melakukan shuffle setiap epoch. Grafik pelatihan adalah sebagai berikut.



Gambar 3.9 Grafik pelatihan

Model hasil pelatihan memiliki akurasi sebesar 98% terhadap data validasi. Setiap citra yang telah disegmentasi akan diklasifikasi kelasnya dengan menggunakan model CNN yang telah dilatih. Hasil klasifikasi citra adalah sebagai berikut.



Gambar 3.10 Citra hasil klasifikasi

### 9. Hitung jumlah sendok dan garpu

Klasifikasi akan menghasilkan kelas sendok atau garpu. Setiap kemunculan kelas akan dihitung. Hasil hitungan jumlah kemunculan kelas akan dikeluarkan oleh sistem sebagai berikut.

Spoon = 2

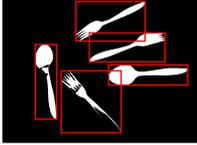
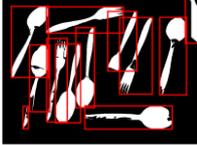
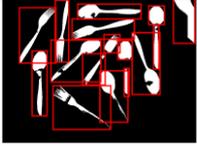
Fork = 3

Gambar 3.11 Citra hasil hitungan

#### IV. HASIL EKSPERIMEN

Eksperimen dilakukan dengan beberapa citra yang memiliki beberapa objek sendok dan garpu. Citra diambil sendiri oleh penulis. Hasil eksperimen adalah sebagai berikut.

Tabel 4.1 Tabel hasil eksperimen

Citra Awal	Citra Bounding Box	Hasil Penghitungan
		Spoon = 2 Fork = 3
		Spoon = 4 Fork = 4
		Spoon = 4 Fork = 8
		Spoon = 5 Fork = 11

#### V. KESIMPULAN

Penggunaan segmentasi citra dan *convolutional neural network* memungkinkan penghitungan jumlah sendok dan garpu secara otomatis. Sistem ini dapat membantu pemeliharaan inventaris peralatan makanan sebuah restoran. Berdasarkan hasil eksperimen, sistem mampu menghitung jumlah sendok dan garpu dengan baik ketika sistem mampu melakukan segmentasi citra dengan baik.

Walaupun begitu, sistem masih memiliki beberapa keterbatasan. Citra masukan sistem harus memiliki latar belakang yang homogen. Jika latar belakang tidak homogen, maka latar belakang akan dianggap sebagai objek saat proses binerisasi. Kendala ini bermasalah karena pada skenario asli tidak selalu akan mendapatkan kondisi yang ideal. Selain itu, tantangan muncul ketika berhadapan dengan objek yang bertumpuk. Objek-objek yang seharusnya disegmentasi menjadi citra masing-masing akan menghasilkan citra yang

berisi beberapa objek sehingga mengakibatkan salah satu objek tidak terhitung. Proses binerisasi dapat secara tidak sengaja memecah suatu objek menjadi beberapa bagian. Kendala ini berpotensi menyebabkan ketidakakuratan dalam proses penghitungan. Objek-objek yang berdekatan akan mengganggu klasifikasi karena objek lain akan ikut masuk dalam citra hasil segmentasi.

#### PRANALA KODE PROGRAM

[https://github.com/DaffaRomyz/IF4073\\_Makalah](https://github.com/DaffaRomyz/IF4073_Makalah)

#### UCAPAN TERIMA KASIH

Penulis mengucapkan syukur kepada Tuhan yang Maha Esa atas segala rahmat yang diberikan sehingga penulis dapat menyelesaikan makalah ini. Secara khusus, penulis mengucapkan terima kasih kepada Bapak Dr. Ir. Rinaldi Munir, M.T. selaku dosen mata kuliah IF4073 Interpretasi dan Pengolahan Citra. Penulis mengucapkan terima kasih keluarga dan semua pihak yang telah memberi dukungan dalam pengerjaan makalah ini.

#### REFERENSI

- [1] Munir, R. (2023). *Segmentasi Citra (Bagian 1)*. <https://informatika.stei.itb.ac.id/~rinaldi.munir/Citra/2023-2024/22-Segmentasi-Citra-Bagian1-2023.pdf>
- [2] Munir, R. (2023). *Segmentasi Citra (Bagian 2)*. <https://informatika.stei.itb.ac.id/~rinaldi.munir/Citra/2023-2024/23-Segmentasi-Citra-Bagian2-2023.pdf>
- [3] Munir, R. (2023). *Convolutional Neural Network*. <https://informatika.stei.itb.ac.id/~rinaldi.munir/Citra/2023-2024/21-CNN-2023.pdf>
- [4] Mikhail. (2019). *spoon-vs-fork*. Diakses 19 Desember 2023. <https://www.kaggle.com/datasets/kilianovski/spoonvsfork/>
- [5] The MathWorks. (2023). *Transfer Learning Using AlexNet*. Diakses 19 Desember 2023. <https://www.mathworks.com/help/deeplearning/ug/transfer-learning-using-alexnet.html>

#### 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, 19 November 2023



Daffa Romyz Aufa 13520162