

Comparative Analysis of VGG16 and ResNet50 Model Performance in Cardiac ECG Image Classification

Hanif Riaqi ^{1*}, Imam Tahyudin ²

^{1*} Program Studi Informatika, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto

² Program Studi Magister Ilmu Komputer, Fakultas Ilmu Komputer, Universitas Amikom Purwokerto
hanifrizaqii@gmail.com ¹, imam.tahyudin@amikompurwokerto.ac.id ²

Article Info

Article history:

Received 2025-03-26

Revised 2025-04-08

Accepted 2025-04-14

Keywords:

*ECG Image Classification,
Deep Learning,
VGG16,
ResNet50,
Transfer Learning.*

ABSTRACT

This study systematically evaluates and compares the effectiveness of two deep learning architectures, VGG16 and ResNet50, in automating electrocardiogram (ECG) image classification for cardiac condition diagnosis. The dataset was obtained from a public source and consists of 2,898 color ECG images converted from raw signals, categorized into four classes: Abnormal Heartbeat, Myocardial Infarction, Normal Individual, and History of Heart Attack. The data underwent preprocessing steps including resizing to 224×224 pixels, pixel normalization to a 0–1 range, label encoding, one-hot encoding, and an 80:20 split for training and testing. Transfer learning was applied using feature representations from the VGG16 and ResNet50 models, employing the Adam optimizer and categorical cross-entropy loss function. To enhance training efficiency and prevent overfitting, early stopping was implemented based on validation loss performance. Model performance was evaluated using accuracy, precision, recall, and F1-score metrics. The results showed that VGG16 achieved 95% accuracy with a loss of 0.1522, precision of 95%, recall of 94%, and F1-score of 94%. In contrast, ResNet50 attained 81% accuracy with a loss of 0.5730, precision of 82%, recall of 79%, and F1-score of 80%. These findings indicate that, within the context of this study, VGG16 consistently outperformed ResNet50 across all evaluation metrics in the ECG image classification task. Therefore, the application of transfer learning using the VGG16 model demonstrates strong potential as an effective approach for AI-based ECG image classification systems.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

I. PENDAHULUAN

Penyakit kardiovaskular merupakan salah satu penyebab utama kematian di dunia, dengan gagal jantung sebagai faktor signifikan yang berkontribusi terhadap tingginya angka tersebut[1]. Deteksi dini dan diagnosis yang akurat terhadap kondisi jantung sangat penting dalam mencegah komplikasi serius yang berpotensi mengancam jiwa. Berdasarkan data global dari *World Health Organization* (WHO) [2], penyakit kardiovaskular (CVD) menyebabkan sekitar 17,9 juta kematian setiap tahun, menjadikannya sebagai masalah kesehatan masyarakat yang membutuhkan perhatian khusus dalam dunia medis dan penelitian. Elektrokardiogram (EKG) adalah alat diagnostik *non-invasif* yang berfungsi merekam aktivitas listrik jantung serta mengidentifikasi gangguan

irama maupun kelainan struktural pada organ tersebut. Sebagai perangkat utama dalam mendiagnosis berbagai kondisi kardiovaskular, EKG memainkan peran penting dalam menilai fungsi jantung[3]. Namun, analisis sinyal EKG secara manual memerlukan keahlian khusus dan cenderung subjektif, yang dapat memengaruhi konsistensi hasil diagnosis. Oleh karena itu, diperlukan sistem klasifikasi otomatis yang handal agar dapat memberikan informasi tentang penyakit, guna meningkatkan ketepatan dan efisiensi proses analisis serta deteksi dini gangguan jantung.

Dalam beberapa tahun terakhir, pembelajaran mendalam (*deep learning*) telah menunjukkan potensi besar dalam analisis citra medis, termasuk klasifikasi citra EKG[4]. Salah satu metode yang banyak dikembangkan dalam *deep learning* adalah *Convolutional Neural Network* (CNN), yang mampu

mengekstraksi fitur penting dari citra medis untuk membantu proses diagnosis secara otomatis dan akurat. Di antara berbagai model CNN yang tersedia, VGGNet dan ResNet menjadi dua arsitektur yang sering digunakan karena kemampuannya dalam menangani data citra yang kompleks[5][6]. Kedua model ini merupakan contoh *pre-trained* CNN, yaitu model yang telah dilatih pada dataset besar dan dapat diterapkan pada dataset yang lebih kecil dengan tugas spesifik melalui pendekatan *transfer learning*. Pendekatan ini memungkinkan pengurangan kebutuhan data pelatihan dalam jumlah besar serta meningkatkan efisiensi waktu dan penggunaan sumber daya komputasi[7].

Pemilihan model VGG16 dan ResNet50 dalam penelitian ini didasarkan pada kinerja kedua arsitektur ini yang cukup baik dalam klasifikasi citra medis. Beberapa penelitian sebelumnya telah banyak mengeksplorasi penggunaan *transfer learning* yang menggunakan kedua model arsitektur ini. Misalnya, sebuah penelitian yang dilakukan oleh Azizah et al. membahas tentang deteksi dini cardiomegaly pada citra chest X-ray dada dengan memanfaatkan model arsitektur VGG16[8]. Penelitian ini bertujuan untuk mengevaluasi kinerja model VGG16 dalam mengklasifikasikan citra X-ray dengan menggunakan dataset open-source yang terdiri dari 13.880 gambar yang diklasifikasikan ke dalam dua kategori: cardiomegaly dan non-cardiomegaly. Hasil penelitian menunjukkan bahwa model VGG16 mampu mendeteksi cardiomegaly dengan akurasi sebesar 78,75% dan nilai ROC-AUC 0,83, yang mengindikasikan performa model yang cukup baik.

Disisi lain, penelitian oleh Yin et al. [9] berfokus pada pengembangan model prediksi berbasis *deep learning radiomics* dengan memanfaatkan ResNet50 untuk mengekstrak fitur citra dari CT scan, guna memprediksi respons terapi serta *Progression-Free Survival* (PFS) pada pasien dengan *unresectable hepatocellular carcinoma* (uHCC). Studi ini melibatkan 172 pasien, yang kemudian dibagi ke dalam tiga kelompok: training cohort (92 pasien), testing cohort (30 pasien), dan external validation cohort (50 pasien). Hasil evaluasi menunjukkan bahwa model ResNet50 terbukti efektif sebagai *feature extractor* utama dalam model *deep learning*, yang mampu memprediksi respons terapi serta PFS pada pasien uHCC dengan tingkat akurasi 89.5% pada training cohort. Sementara itu, pada testing cohort mendapatkan akurasi 80.4%, dan pada external validation cohort, diperoleh akurasi 79.1%.

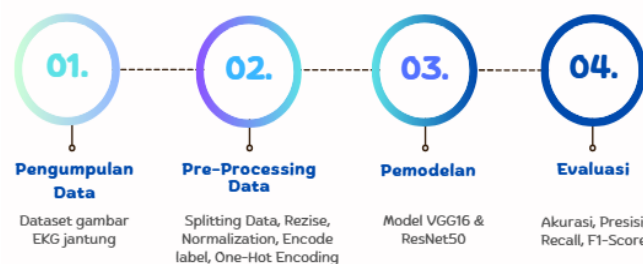
Kemudian, penelitian yang dilakukan oleh Berliani et al.[10] membandingkan performa model VGG-16 dan ResNet-50 dalam klasifikasi citra X-ray paru-paru. Penelitian ini bertujuan untuk mengevaluasi kemampuan kedua model dalam mengidentifikasi empat kategori, yaitu normal, Covid-19, viral pneumonia, dan lung opacity. Dalam studi ini, digunakan dua dataset publik yang diperoleh dari University of Montreal dan Qatar University. Dataset tersebut kemudian dibagi menjadi *training set* dan *testing set* untuk menguji performa model. Hasil penelitian menunjukkan bahwa baik VGG-16 maupun ResNet-50 mampu mencapai akurasi tinggi

di atas 90% dalam klasifikasi citra X-ray paru-paru. Namun, terdapat beberapa perbedaan utama antara kedua model. ResNet-50 memiliki akurasi yang lebih tinggi dibandingkan VGG-16, tetapi waktu inferensinya lebih lambat. Sebaliknya, VGG-16 lebih unggul dalam kecepatan pemrosesan, meskipun akurasinya sedikit lebih rendah dibandingkan ResNet-50.

Meskipun terdapat berbagai penelitian yang menerapkan *transfer learning* pada citra medis, masih terdapat tantangan dalam menentukan arsitektur model yang paling optimal untuk tugas ini. Perbedaan dalam struktur model, jumlah parameter, serta metode ekstraksi fitur dapat memengaruhi akurasi dan efisiensi komputasi. Oleh karena itu, penelitian ini bertujuan mengembangkan sistem klasifikasi citra *Elektrokardiogram* (EKG) berbasis *transfer learning* dengan memanfaatkan arsitektur VGG16 dan ResNet50. Evaluasi komparatif dilakukan untuk menganalisis kinerja kedua model dalam mendeteksi kelainan jantung menggunakan metrik akurasi, presisi, recall, dan F1-score yang diturunkan dari confusion matrix. Hasil penelitian diharapkan dapat menjadi dasar pengembangan sistem diagnosis otomatis yang lebih akurat dan efisien, sehingga dapat membantu dalam deteksi dini dan analisis gangguan jantung secara lebih objektif.

II. METODE

Metodologi penelitian ini terdiri dari beberapa tahap utama yang disusun secara sistematis untuk memperoleh hasil klasifikasi citra EKG yang optimal. Proses ini mencakup pengumpulan data, pre-processing, pembuatan model, serta evaluasi. Setiap tahap memiliki peran penting dalam memastikan bahwa model *deep learning* yang digunakan, yakni VGG16 dan ResNet50, mampu mengklasifikasikan kondisi kesehatan jantung berdasarkan citra EKG. Alur tahapan penelitian ini dapat dilihat pada gambar 1.

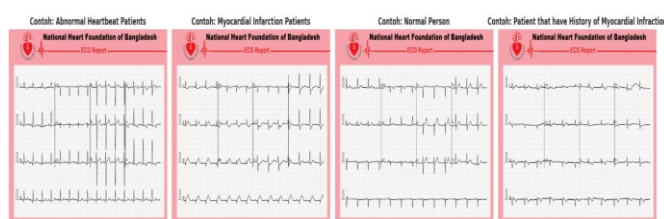


Gambar 1 Alur Tahapan Penelitian

A. Pengumpulan Data

Tahap awal dalam penelitian ini adalah pengumpulan dataset citra medis *Elektrokardiogram* (EKG) yang berkaitan dengan kondisi kesehatan jantung. Dataset yang digunakan diperoleh dari sumber terbuka (open source) melalui platform Kaggle: <https://www.kaggle.com/datasets/drkhaledmohsin/national-heart-foundation-2023-ecg-dataset>. Dataset ini merupakan kontribusi dari National Heart Foundation of

Bangladesh dan terdiri dari 2.898 citra EKG berwarna hasil konversi sinyal EKG mentah ke dalam format visual. Data diklasifikasikan ke dalam empat kelas utama, yaitu: Detak Jantung Abnormal, Infark Miokard (Myocardial Infarction), Normal, dan Riwayat Infark Miokard (History of Myocardial Infarction). Kategori Infark Miokard mencakup pasien yang mengalami serangan jantung akut dengan indikasi pola EKG seperti elevasi segmen ST, inversi gelombang T, atau gelombang Q patologis. Sementara itu, kelas Riwayat Infark Miokard berisi citra EKG dari pasien yang sebelumnya pernah mengalami serangan jantung dan dapat menunjukkan perubahan jangka panjang akibat kerusakan miokard yang telah terjadi. Pemisahan kedua kategori ini bertujuan untuk mempertahankan konteks klinis yang berbeda, yang penting dalam mendukung proses klasifikasi berbasis citra. Contoh visual dari citra EKG yang digunakan dalam penelitian ini ditampilkan pada Gambar 2.



Gambar 2 Contoh dataset citra EKG

B. Pre-processing Data

Pra-pemrosesan data menjadi komponen penting dalam analisis citra medis guna untuk memastikan kesiapan data serta meningkatkan kualitasnya sebelum proses pelatihan model. Tahap pre-processing dilakukan bertujuan untuk meminimalkan noise, menyelaraskan ukuran gambar, dan memperjelas kontras, sehingga model dapat mengekstraksi fitur dengan lebih representatif terhadap pola yang terdapat dalam data [11]. Pada proses pre-processing kali ini meliputi beberapa tahap seperti pembagian data (*splitting*), penyesuaian ukuran (*resizing*), normalisasi, label encode, one-hot encoding. Dalam penerapan *deep learning* klasifikasi citra medis, normalisasi dan transformasi data memiliki peran penting dalam memastikan model dapat belajar secara optimal serta meminimalkan bias akibat perbedaan skala antar data. Selain itu, teknik encoding label dan one-hot encoding diterapkan untuk membantu model mengenali kategori dengan lebih terstruktur tanpa membentuk hubungan hierarkis yang tidak diinginkan. Dengan menerapkan pre-processing yang optimal, akurasi model dalam mengklasifikasikan citra medis dapat meningkat secara signifikan [12][13].

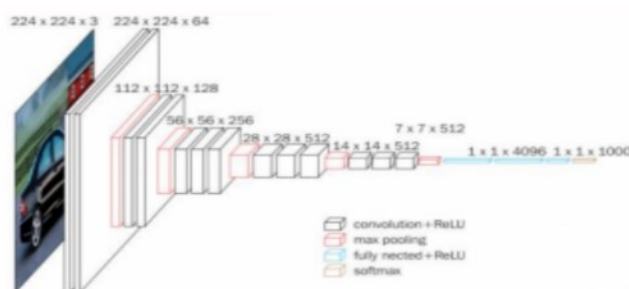
C. Pemodelan

Setelah proses pengumpulan data dan pre-processing selesai, tahap berikutnya dalam penelitian ini adalah merancang dan melatih model *deep learning* untuk klasifikasi gambar EKG. Penelitian ini menggunakan dua arsitektur, yaitu VGG16 dan ResNet50, yang dipilih berdasarkan rekam

jeaknya dalam berbagai studi terkait klasifikasi gambar medis. Selain itu, fitur *transfer learning* yang dimiliki oleh kedua arsitektur ini memberikan keunggulan dalam meningkatkan akurasi model. VGG16 memiliki arsitektur yang lebih sederhana dan mudah dipahami, sementara ResNet50 mengandalkan residual connections untuk memungkinkan pelatihan model yang lebih dalam tanpa mengalami kendala vanishing gradient [14].

Proses pemodelan mencakup penggunaan model *pre-trained*, dengan melakukan penyesuaian pada lapisan output agar selaras dengan jumlah kelas dalam dataset, serta pelatihan model menggunakan dataset yang telah diproses sebelumnya dengan menerapkan teknik *early stopping*. *Early stopping* merupakan teknik yang secara otomatis menghentikan proses pelatihan model ketika kinerja validasi mulai menurun. Tujuannya adalah untuk mencegah *overfitting* dan memastikan model dapat melakukan generalisasi dengan baik terhadap data baru. Metode ini bekerja dengan memantau metrik validasi, seperti akurasi atau loss, di setiap *epoch*. Jika dalam beberapa *epoch* berturut-turut tidak terjadi peningkatan, pelatihan akan dihentikan untuk menghindari pelatihan yang berlebihan [15].

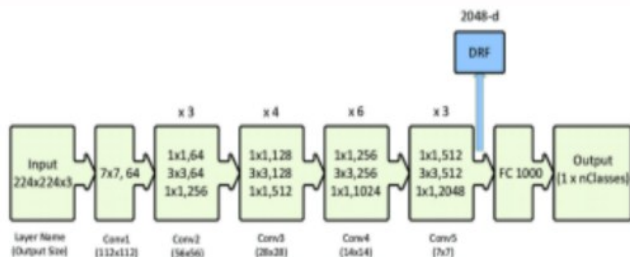
1) **VGG16:** VGG16 merupakan arsitektur model *pre-trained* CNN (Convolutional Neural Network) yang diperkenalkan oleh Simonyan dan Zisserman pada tahun 2014 di Visual Geometry Group (VGG), University of Oxford. Model ini memiliki jaringan yang cukup dalam dengan total 16 lapisan utama, terdiri dari 13 lapisan konvolusi dan 3 lapisan fully connected. Berbeda dengan model sebelumnya seperti AlexNet, VGG16 menggantikan kernel berukuran besar dengan kernel konvolusi kecil (3×3) yang diterapkan secara bertahap, sehingga dapat mengekstraksi fitur yang lebih kompleks tanpa secara drastis meningkatkan jumlah parameter [10].



Gambar 3 Arsitektur VGG16 [16]

Selain itu, model ini menggunakan max pooling 3×3 dengan stride 2 untuk mengurangi dimensi spasial tanpa kehilangan informasi penting dalam citra. Salah satu keunggulan utama VGG16 adalah konsistensinya dalam penggunaan filter dan padding, yang menjadikannya lebih fleksibel untuk diterapkan pada berbagai tugas klasifikasi. Dengan sekitar 138 juta parameter, model ini tetap menjadi pilihan yang banyak digunakan dalam penelitian *deep learning*, terutama dalam klasifikasi citra medis [8]. Arsitektur model VGG16 dapat dilihat pada gambar 3.

2) *ResNet50*: ResNet50 juga merupakan salah satu arsitektur Convolutional Neural Network (CNN) yang dikembangkan oleh He et al. (2016) [17] dan telah banyak digunakan dalam berbagai tugas klasifikasi citra. Model ini memiliki total 50 lapisan dan diperkuat dengan konsep residual learning, yang menjadi pembeda utama dibandingkan arsitektur sebelumnya. Dengan memanfaatkan residual connections atau skip connections, ResNet50 mampu mengatasi permasalahan vanishing gradient, sehingga tetap dapat beroperasi secara optimal meskipun memiliki kedalaman jaringan yang signifikan[10]. Struktur arsitektur ResNet50 terdiri dari 5 tahap utama, yang masing-masing mencakup convolutional block dan identity block untuk meningkatkan proses ekstraksi fitur. Model ini memiliki lebih dari 23 juta parameter yang telah dilatih menggunakan dataset ImageNet, sehingga mampu mencapai tingkat akurasi tinggi dalam klasifikasi citra. Keunggulannya dalam mengekstraksi fitur mendalam serta menjaga aliran informasi antar lapisan menjadikannya model yang banyak diterapkan dalam pengenalan pola dan klasifikasi citra medis[18] Arsitektur model ResNet50 dapat dilihat pada gambar 4.



Gambar 4 Arsitektur ResNet50 [16]

Transfer learning diterapkan menggunakan arsitektur VGG16 dan ResNet50 yang telah dilatih sebelumnya pada dataset ImageNet. Dalam implementasinya, penelitian ini hanya memanfaatkan representasi fitur dari model (*feature extraction*) tanpa melakukan fine-tuning terhadap bobot *pre-trained*. Pendekatan ini dipilih karena keterbatasan sumber daya komputasi. Fine-tuning membutuhkan proses pelatihan ulang sebagian atau seluruh lapisan model, yang memerlukan waktu lebih lama dan perangkat keras dengan kapasitas tinggi. Dengan hanya melakukan ekstraksi fitur, pelatihan menjadi lebih ringan dan efisien, serta tetap memanfaatkan kekuatan representasi dari model *pre-trained*. Meskipun model berkemungkinan tidak sepenuhnya beradaptasi dengan karakteristik spesifik citra EKG, pendekatan ini tetap efektif dalam menghindari *overfitting* dan mampu menghasilkan performa klasifikasi yang kompetitif. Menurut Kornblith et al. (2019), ekstraksi fitur dengan model *pre-trained* masih dapat memberikan performa yang baik dalam berbagai tugas klasifikasi, meskipun tanpa dilakukan fine-tuning[19].

D. Evaluasi

Tahap akhir dalam penelitian ini adalah mengevaluasi kinerja model untuk mengukur efektivitas VGG16 dan ResNet50 dalam klasifikasi citra EKG. Evaluasi dilakukan

untuk menentukan sejauh mana model dapat menghasilkan klasifikasi dengan tingkat akurasi yang optimal. Salah satu metode yang diterapkan adalah *confusion matrix*, yang memungkinkan perhitungan berbagai metrik, termasuk akurasi, presisi, recall, dan F1-score [20]. Analisis terhadap metrik-metrik ini tidak hanya memberikan gambaran mengenai kinerja model, tetapi juga berfungsi sebagai dasar perbandingan efisiensi kedua algoritma dalam tugas klasifikasi EKG.

Adapun metrik evaluasi yang digunakan dalam penelitian ini adalah sebagai berikut:

- 1) *Akurasi*: merupakan perhitungan untuk mengukur sejauh mana model dapat mengklasifikasikan data dengan benar secara keseluruhan. Metrik ini diperoleh dengan membandingkan jumlah prediksi yang tepat terhadap total data yang diuji. Rumus perhitungannya ditunjukkan pada persamaan (1).

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

- 2) *Presisi*: merupakan indikator yang mengevaluasi seberapa akurat prediksi positif model dengan membandingkan jumlah true positive terhadap seluruh prediksi yang dihasilkan sebagai positif. Perhitungan presisi dapat dilihat pada persamaan (2).

$$\text{Presisi} = \frac{TP}{TP + FP} \quad (2)$$

- 3) *Recall*: merupakan metrik yang mengukur seberapa banyak data positif yang berhasil ditemukan oleh model dari semua data positif yang ada. Nilai recall dapat dihitung dengan persamaan (3).

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3)$$

- 4) *F1-Score*: merupakan metrik yang mengukur keseimbangan antara presisi dan recall, dihitung menggunakan rata-rata harmonik dari kedua metrik tersebut. Nilainya dirumuskan dalam persamaan (4).

$$\text{F1-Score} = 2 * \frac{\text{presisi} * \text{recall}}{\text{presisi} + \text{recall}} \quad (4)$$

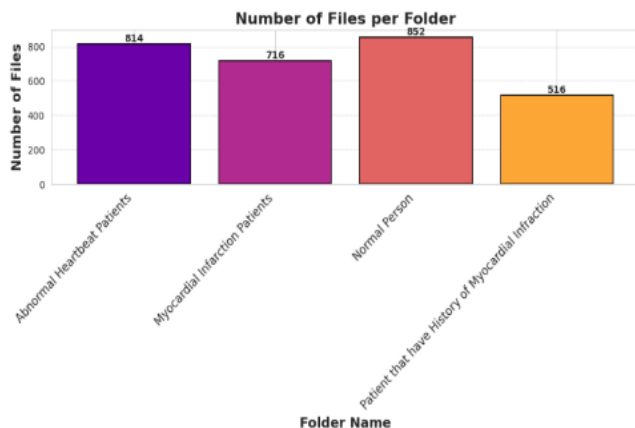
True Positive (TP) adalah kondisi ketika model berhasil mengklasifikasikan data positif dengan benar, sedangkan *True Negative* (TN) terjadi saat model secara akurat mengenali data negatif. Sebaliknya, *False Positive* (FP) merupakan kesalahan prediksi di mana data negatif diklasifikasikan sebagai positif, sementara *False Negative* (FN) terjadi ketika model salah mengidentifikasi data positif sebagai negatif[21].

III. HASIL DAN PEMBAHASAN

Hasil penelitian ini dilakukan menggunakan Google Colab sebagai platform pemodelan, dengan dua pendekatan berbeda untuk mengklasifikasikan citra EKG. Model pertama yang diuji adalah VGG16, diikuti oleh ResNet50. Hasil dari kedua model kemudian dibandingkan berdasarkan metrik evaluasi yang telah ditentukan untuk menilai performa terbaik dalam klasifikasi citra EKG.

A. Dataset

Dataset yang digunakan dalam penelitian ini merupakan kumpulan citra *Elektrokardiogram* (EKG) yang dikategorikan ke dalam empat kelas utama, yaitu Pasien dengan Detak Jantung Abnormal, Pasien Infark Miokard, Normal, serta Pasien dengan Riwayat Infark Miokard. Setiap kategori terdiri dari sejumlah gambar yang merepresentasikan pola aktivitas listrik jantung sesuai dengan kondisi pasien. Secara keseluruhan, dataset ini mencakup 2.898 gambar dengan distribusi sebagai berikut: 814 gambar untuk kategori Pasien dengan Detak Jantung Abnormal, 716 gambar untuk Pasien dengan Infark Miokard, 852 gambar untuk kategori Normal, dan 516 gambar untuk Pasien dengan Riwayat Infark Miokard. Distribusi jumlah citra data per kategori kelas dapat dilihat pada gambar 5.



Gambar 5 Distribusi citra data tiap kategori kelas

B. Pre-processing Data

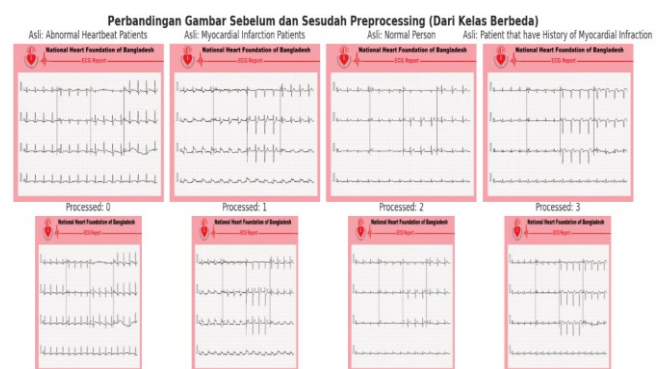
Tahap pre-processing data merupakan aspek penting dalam penelitian ini untuk memastikan data memiliki kualitas optimal sebelum digunakan dalam pelatihan model *deep learning*. Proses pre-processing yang diterapkan didalam penelitian meliputi pembagian data (*splitting*), penyesuaian ukuran gambar (*resizing*), normalisasi, encoding label, dan one-hot encoding. Langkah-langkah ini bertujuan untuk meningkatkan efisiensi pemrosesan, mengurangi potensi bias akibat perbedaan skala, serta memungkinkan model mengekstraksi fitur dengan lebih akurat. Penjelasan lebih detail terkait proses pre-processing dalam penelitian ini dijelaskan seperti dibawah ini.

- 1) *Splitting Data (Membagi Data)*: Dataset dibagi menjadi dua bagian utama, yakni data latih (80%) dan data uji

(20%), dengan total 2.138 gambar untuk pelatihan dan 530 gambar untuk pengujian. Pembagian ini bertujuan agar model dapat belajar dari sebagian besar data yang tersedia serta dievaluasi menggunakan data yang belum pernah digunakan sebelumnya.

- 2) *Resize (Penyesuaian Ukuran Citra)*: Setiap gambar dalam dataset memiliki ukuran yang bervariasi. Oleh karena itu, seluruh citra diubah ukurannya menjadi 224×224 piksel agar sesuai dengan arsitektur model VGG16 dan ResNet50, yang memerlukan input gambar dengan dimensi tetap.
- 3) *Normalisasi*: Agar proses komputasi lebih efisien dan konvergensi model berlangsung lebih cepat, nilai piksel pada setiap gambar dinormalisasi ke rentang 0 hingga 1 dengan membagi setiap nilai piksel dengan 255. Langkah ini memungkinkan model bekerja dengan skala data yang lebih seragam, sehingga meningkatkan efektivitas proses pembelajaran.
- 4) *Label Encoding*: Kategori kelas pada dataset yang awalnya berbentuk teks dikonversi menjadi nilai numerik untuk memudahkan pemrosesan dalam model deep learning. Setiap kelas diberikan indeks sebagai berikut: Pasien dengan Detak Jantung Abnormal (0), Pasien dengan Infark Miokard (1), Normal (2), serta Pasien dengan Riwayat Infark Miokard (3).
- 5) *One-hot Encoding*: Setelah label dikonversi ke format numerik, one-hot encoding diterapkan agar model dapat mengenali setiap kategori secara independen tanpa mengasumsikan adanya hubungan ordinal antar kelas. Pendekatan ini membantu meningkatkan akurasi prediksi. Kategori yang digunakan adalah Pasien dengan Detak Jantung Abnormal [1. 0. 0. 0.], Pasien dengan Infark Miokard [0. 1. 0. 0.], Normal [0. 0. 1. 0.], serta Pasien dengan Riwayat Infark Miokard [0. 0. 0. 1.].

Setelah setiap semua tahap pre-processing selesai, data yang telah diproses siap digunakan untuk pelatihan model. Gambar 6 di bawah ini menampilkan hasil perbandingan contoh data sebelum dan sesudah dilakukan pre-processing.



Gambar 6 Hasil proses preprocessing

C. Implementasi Model

Tahapan ini mencakup pembangunan serta pelatihan model *deep learning* untuk klasifikasi citra EKG menggunakan arsitektur VGG16 dan ResNet50 dengan pendekatan *transfer learning*. Model dilatih menggunakan dataset yang telah melalui proses pre-processing.

1) *VGG16*: Pada tahap ini model VGG16 dilatih menggunakan arsitektur yang sudah ditentukan. Gambar 7 menunjukkan arsitektur VGG16 yang digunakan dengan pendekatan *transfer learning*. Model *pre-trained* pada ImageNet ini berfungsi sebagai ekstraktor fitur, dengan lapisan fully connected asli dihapus (`include_top=False`) dan diganti sesuai jumlah kelas dataset. Model menerima input gambar berukuran $224 \times 224 \times 3$ piksel, dengan lapisan konvolusi dibekukan (`layer.trainable=False`) untuk mempertahankan bobot *pre-trained*. Setelah ekstraksi fitur, digunakan Global Average Pooling 2D, lapisan Dense (1024 neuron, ReLU), dan Softmax untuk klasifikasi. Model dikompilasi dengan optimizer Adam dan categorical crossentropy sebagai loss function untuk klasifikasi multi-kelas.

```
# Load pre-trained VGG16 model
base_model = VGG16(weights='imagenet', include_top=False, input_shape=(224, 224, 3))

# Add custom classification layers
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation='relu')(x)
predictions = Dense(num_classes, activation='softmax')(x) # Use num_classes from previous code

# Create the final model
model_vgg16 = Model(inputs=base_model.input, outputs=predictions)

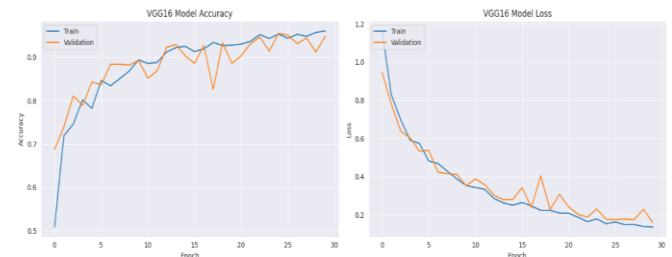
# Freeze the layers of the pre-trained model
for layer in base_model.layers:
    layer.trainable = False

# Compile the model
model_vgg16.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])

# Print the model summary
model_vgg16.summary()
```

Gambar 7 Arsitektur model VGG16

Model dilatih menggunakan dataset yang telah diproses, dengan pembagian *training set* dan *validation set* dalam rasio 80:20 menggunakan teknik stratifikasi untuk menjaga keseimbangan kelas. Early stopping diterapkan untuk mencegah *overfitting* dengan menghentikan pelatihan jika validation loss tidak membaik dalam tiga epoch berturut-turut. Pelatihan dilakukan selama maksimal 30 epoch dengan batch size 32, menggunakan categorical crossentropy sebagai loss function dan optimizer Adam. Kinerja model dipantau berdasarkan akurasi dan loss yang divisualisasikan dalam gambar 8 dibawah, menunjukkan bahwa VGG16 memiliki performa klasifikasi yang baik. Akurasi meningkat secara konsisten pada epoch pertama akurasi pada saat pelatihan sebesar 40% dengan loss 1.3005 sementara akurasi validasi sebesar 69% dan val_loss 0.9463 dan hingga mencapai 95% di akhir epoch, sementara loss menurun stabil hingga sekitar 0.1522 pada data uji. Konsistensi antara akurasi dan loss menunjukkan bahwa model mampu melakukan generalisasi dengan baik.



Gambar 8 Grafik akurasi dan loss model VGG16

2) *ResNet50*: Pada tahap ini model ResNet50 dilatih menggunakan arsitektur yang sudah ditentukan. Gambar 9 menunjukkan arsitektur ResNet50 yang diterapkan dengan transfer learning, di mana model *pre-trained* pada ImageNet digunakan sebagai *feature extractor*. Lapisan fully connected asli dihapus (`include_top=False`) dan diganti dengan lapisan klasifikasi baru sesuai jumlah kelas dataset. Model menerima input gambar berukuran $224 \times 224 \times 3$ piksel dengan lapisan konvolusi dibekukan (`layer.trainable = False`). Ekstraksi fitur dilakukan menggunakan Global Average Pooling 2D (GAP), diikuti oleh Dense 1024 neuron (ReLU) dan Softmax untuk klasifikasi empat kategori. Model dikompilasi dengan optimizer Adam dan categorical crossentropy sebagai loss function untuk klasifikasi multi-kelas.

```
# Load pre-trained ResNet50 model
base_model = ResNet50(weights='imagenet', include_top=False, input_shape=(224, 224, 3))

# Add custom classification layers
x = base_model.output
x = GlobalAveragePooling2D()(x)
x = Dense(1024, activation='relu')(x)
predictions = Dense(num_classes, activation='softmax')(x)

# Create the final model
model_resnet50 = Model(inputs=base_model.input, outputs=predictions)

# Freeze the layers of the pre-trained model
for layer in base_model.layers:
    layer.trainable = False

# Compile the model
model_resnet50.compile(optimizer='adam', loss='categorical_crossentropy', metrics=['accuracy'])

# Print the model summary
model_resnet50.summary()
```

Gambar 9 Arsitektur model ResNet50

Sama halnya dengan VGG16, model ResNet50 dilatih menggunakan dataset yang telah diproses, dengan pembagian *training set* dan *validation set* dalam rasio 80:20 menggunakan stratifikasi untuk menjaga keseimbangan kelas. Early stopping diterapkan untuk mencegah *overfitting* dengan menghentikan pelatihan jika validation loss tidak membaik dalam tiga epoch berturut-turut. Model dilatih selama maksimal 30 epoch dengan batch size 32, menggunakan categorical crossentropy sebagai loss function dan optimizer Adam. Kinerja model dipantau berdasarkan akurasi dan loss yang divisualisasikan dalam Gambar 10, menunjukkan bahwa ResNet50 memiliki performa klasifikasi yang cukup baik. Akurasi meningkat secara bertahap hingga mencapai 81% pada data uji, meskipun terdapat variasi di beberapa epoch

awal. Grafik loss menurun secara stabil hingga mencapai 0.5730, menandakan model mampu mengenali pola dengan baik.



Gambar 10 Arsitektur model ResNet50

D. Evaluasi dan Analisis Kinerja Model

Evaluasi model dilakukan untuk menilai performa VGG16 dan ResNet50 dalam mengklasifikasikan citra EKG ke dalam empat kelas. Proses evaluasi dilakukan menggunakan beberapa metrik, termasuk akurasi, loss, precision, recall, dan F1-score, yang diperoleh berdasarkan hasil prediksi pada data uji.

1) Hasil Evaluasi Model Vgg16

Evaluasi terhadap model VGG16 dilakukan untuk mengukur kemampuannya dalam mengklasifikasikan citra EKG ke dalam empat kategori, yaitu Abnormal Heartbeat Patients, Myocardial Infarction Patients, Normal Person, dan Patient that have History of Myocardial Infarction. Berdasarkan hasil evaluasi terhadap 580 citra pada data uji, model VGG16 menunjukkan performa yang sangat baik dengan akurasi mencapai 95%. Nilai precision, recall, dan F1-score rata-rata juga tergolong tinggi, yaitu masing-masing sebesar 0,95, 0,94, dan 0,94 untuk macro average, serta 0,95 pada weighted average, seperti ditunjukkan pada Gambar 11.

	precision	recall	f1-score	support
0	0.94	0.98	0.96	163
1	0.98	1.00	0.99	143
2	0.96	0.94	0.95	171
3	0.91	0.85	0.88	103
accuracy			0.95	580
macro avg	0.95	0.94	0.94	580
weighted avg	0.95	0.95	0.95	580

Gambar 11 Classification report model VGG16

Secara rinci, model mencatatkan performa terbaik pada kelas Myocardial Infarction Patients, dengan precision sebesar 0,98 dan recall mencapai 1,00, menghasilkan nilai F1-score sebesar 0,99 tanpa kesalahan klasifikasi. Kelas Abnormal Heartbeat Patients juga menunjukkan hasil yang sangat baik dengan F1-score sebesar 0,96. Pada kelas Normal Person, model meraih F1-score sebesar 0,95 dengan precision 0,96 dan recall 0,94. Sementara itu, performa terendah terjadi pada kelas Patient that have History of Myocardial Infarction dengan F1-score sebesar 0,88, mengindikasikan adanya

tantangan dalam membedakan karakteristik citra pada kelas ini.

Confusion Matrix - VGG16

	Abnormal Heartbeat Patients	Myocardial Infarction Patients	Normal Person	Patient that have History of Myocardial Infarction
Abnormal Heartbeat Patients	160	0	0	3
Myocardial Infarction Patients	0	143	0	0
Normal Person	2	3	166	6
Patient that have History of Myocardial Infarction	9	0	6	88
	Abnormal Heartbeat Patients	Myocardial Infarction Patients	Normal Person	Patient that have History of Myocardial Infarction

True Label

Predicted Label

Gambar 12 Confusion matrix VGG16

Hasil *confusion matrix* yang ditampilkan pada Gambar 12, juga menampilkan bahwa sebagian besar prediksi model berada pada diagonal utama, yang menandakan tingkat akurasi klasifikasi yang tinggi. Namun, masih terdapat beberapa kesalahan klasifikasi, terutama pada kelas Patient that have History of Myocardial Infarction, yang terkadang salah dikenali sebagai kelas Abnormal Heartbeat dan Normal Person. Hal ini menunjukkan bahwa meskipun model VGG16 memiliki performa yang sangat kompetitif, masih diperlukan pengembangan lebih lanjut, khususnya untuk meningkatkan sensitivitas model terhadap karakteristik citra pada kelas dengan performa rendah.

2) Hasil Evaluasi Model ResNet50

Evaluasi terhadap model ResNet50 dilakukan untuk mengukur kemampuannya dalam mengklasifikasikan citra EKG ke dalam empat kategori, yakni pasien dengan detak jantung tidak normal, pasien dengan serangan jantung, individu normal, dan pasien dengan riwayat serangan jantung. Berdasarkan hasil pengujian terhadap 580 citra pada data uji, model ResNet50 mencapai tingkat akurasi sebesar 81%. Meskipun nilai ini menunjukkan performa klasifikasi yang cukup baik, hasil tersebut masih berada di bawah capaian model VGG16. Gambar 13 menyajikan *classification report* dari model ResNet50 yang menampilkan nilai precision, recall, dan F1-score untuk masing-masing kelas. Nilai rata-rata macro untuk precision, recall, dan F1-score masing-masing sebesar 0,82, 0,79, dan 0,80, sementara nilai rata-rata tertimbang (*weighted average*) untuk ketiganya adalah 0,81. Hal ini menunjukkan bahwa model mampu memberikan prediksi yang cukup baik secara keseluruhan, meskipun terdapat variasi performa antar kelas yang menunjukkan adanya ketidakseimbangan.

	precision	recall	f1-score	support
0	0.87	0.75	0.81	163
1	0.79	0.92	0.85	143
2	0.76	0.88	0.81	171
3	0.85	0.61	0.71	103
accuracy			0.81	580
macro avg	0.82	0.79	0.80	580
weighted avg	0.81	0.81	0.80	580

Gambar 13 Clasification report model ResNet50

Secara rinci, performa terbaik ditunjukkan pada kelas Myocardial Infarction Patients dengan recall mencapai 0,92 dan F1-score sebesar 0,85, menunjukkan bahwa model cukup andal dalam mengenali kategori ini. Kelas Abnormal Heartbeat Patients dan Normal Person memiliki nilai F1-score yang sama, yaitu 0,81, meskipun precision pada Abnormal Heartbeat Patients lebih tinggi (0,87). Sementara itu, kinerja paling rendah terjadi pada kelas Patient that have History of Myocardial Infarction, yang hanya mencapai F1-score sebesar 0,71, menandakan bahwa model masih menghadapi tantangan dalam mengidentifikasi citra dari kelas tersebut.

Confusion Matrix - ResNet50				
True Label	Abnormal Heartbeat Patients	Myocardial Infarction Patients	Normal Person	Patient that have History of Myocardial Infarction
	123	13	17	30
	0	131	12	0
	3	16	151	1
Predicted Label				
	Abnormal Heartbeat Patients	Myocardial Infarction Patients	Normal Person	Patient that have History of Myocardial Infarction
	15	5	20	63

Gambar 14 Confusion matrix ResNet50

Gambar 14 menampilkan confusion matrix dari hasil klasifikasi yang dilakukan oleh model ResNet50. Meskipun sebagian besar prediksi berada pada diagonal utama, yang menandakan klasifikasi yang benar, masih terdapat sejumlah kesalahan prediksi, terutama pada kelas Patient that have History of Myocardial Infarction. Kelas ini kerap diklasifikasikan secara keliru sebagai Abnormal Heartbeat Patients maupun Normal Person. Selain itu, terjadi pula kekeliruan dalam membedakan antara kelas Myocardial Infarction Patients dan Normal Person, yang mengindikasikan adanya kemiripan pola fitur yang ditangkap oleh model dari kedua kelas tersebut. Secara umum, meskipun ResNet50 menunjukkan akurasi yang cukup baik dan mampu mengklasifikasikan beberapa kelas dengan performa yang

memadai, pengembangan lebih lanjut masih diperlukan guna meningkatkan kemampuan model dalam mengenali kelas-kelas yang memiliki kompleksitas fitur yang tinggi.

E. Perbandingan Kinerja Model

Berdasarkan hasil analisis model VGG16 dan ResNet50 tersebut kemudian dilakukan perbandingan performa kedua model dirangkum dalam Tabel 1, yang mencakup metrik akurasi, loss, precision, recall, dan F1-score.

TABEL I
PERBANDINGAN KINERJA MODEL VGG16 DAN RESNET50

Metrik Evaluasi	Model	
	VGG16	ResNet50
Akurasi	95%	81%
Loss	0.1522	0.5730
Precision	95%	82%
Recall	94%	79%
F1-Score	94%	80%

Tabel di atas menyajikan perbandingan performa antara model VGG16 dan ResNet50 berdasarkan metrik *macro average* serta nilai loss. Hasil evaluasi menunjukkan bahwa VGG16 unggul dalam semua aspek evaluasi dibandingkan ResNet50 berdasarkan analisis confusion matrix.

Model VGG16 berhasil memperoleh nilai akurasi sebesar 95% dan nilai loss yang rendah, yakni 0,1522, yang mengindikasikan proses pembelajaran yang stabil tanpa indikasi *overfitting* yang signifikan. Selain itu, nilai precision, recall, dan F1-score *macro average* masing-masing sebesar 0,95, 0,94, dan 0,94, menunjukkan kemampuan klasifikasi yang konsisten dan seimbang terhadap seluruh kelas yang diuji. Di sisi lain, ResNet50 memperoleh akurasi sebesar 81% dengan nilai loss sebesar 0,5730. Nilai precision, recall, dan F1-score *macro average* masing-masing sebesar 0,82, 0,79, dan 0,80. Hasil ini mengindikasikan bahwa meskipun ResNet50 memiliki arsitektur yang lebih dalam dan kompleks, performanya pada dataset ini masih belum sebaik VGG16. Nilai loss yang lebih tinggi menunjukkan bahwa model mengalami lebih banyak kesalahan dalam proses prediksi dan belum mampu mengenali fitur-fitur penting dalam citra EKG secara optimal.

IV. KESIMPULAN

Penelitian ini telah mengevaluasi dan membandingkan performa model VGG16 dan ResNet50 dalam mengklasifikasikan citra EKG ke dalam empat kategori, yaitu pasien dengan detak jantung abnormal, pasien dengan infark miokard, individu dengan kondisi normal, serta pasien dengan riwayat infark miokard. Hasil analisis menunjukkan bahwa VGG16 unggul dengan akurasi mencapai 95%, nilai loss yang rendah (0.1522), serta precision, recall, dan F1-score masing-masing sebesar 95%, 94%, dan 94%. Hal ini mengindikasikan kestabilan model VGG16 serta kemampuannya dalam mengenali pola citra EKG secara efektif. Sebaliknya, model ResNet50 menunjukkan performa yang lebih rendah, dengan akurasi sebesar 81%, loss sebesar 0.5730, serta precision,

recall, dan F1-score berturut-turut sebesar 82%, 79%, dan 80%, yang menandakan bahwa model ini kurang optimal dalam menangkap fitur dari citra EKG pada penelitian ini. Meskipun ResNet50 memiliki arsitektur yang lebih kompleks dengan residual connections, namun performa model dalam konteks dataset ini belum mampu menyaingi kinerja VGG16 secara menyeluruh.

Secara keseluruhan, VGG16 terbukti lebih unggul dalam klasifikasi citra EKG pada penelitian ini. Dalam implementasi praktis, kedua model memiliki potensi untuk diterapkan dalam sistem klasifikasi citra EKG. Namun, untuk meningkatkan kinerja model, untuk penelitian lebih lanjut disarankan mengeksplorasi arsitektur lain yang lebih adaptif, baik melalui modifikasi lapisan pada VGG16 dan ResNet50 maupun dengan memanfaatkan model pretrained lainnya. Selain itu, peningkatan kualitas dataset, seperti penyeimbangan jumlah sampel antar kelas dan penerapan teknik preprocessing lanjutan, dapat menjadi strategi yang efektif guna meningkatkan akurasi serta kemampuan generalisasi model dalam klasifikasi citra EKG.

DAFTAR PUSTAKA

- [1] V. Artanti, M. Faisal, and F. Kurniawan, "Klasifikasi Cardiovascular Diseases Menggunakan Algoritma K-Nearest Neighbors (KNN) Classification of Cardiovascular Diseases using K-Nearest Neighbors (KNN) Algorithm," *Techno.COM*, vol. 23, no. 2, pp. 467–479, May 2024.
- [2] World Health Organization (WHO), "Cardiovascular diseases (CVDs)," [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)).
- [3] I. Wijayanto, A. Humairani, A. Rizal, And S. Hadiyoso, "Klasifikasi Sinyal EKG menggunakan Ciri Statistik dan Parameter Hjorth dengan SVM dan k-NN," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 10, no. 1, p. 132, Jan. 2022, doi: 10.26760/elkomika.v10i1.132.
- [4] D. Shen, G. Wu, and H.-I. Suk, "Deep Learning in Medical Image Analysis," *Annu Rev Biomed Eng*, vol. 42, p. 46, 2017, doi: 10.1146/annurev-bioeng-071516.
- [5] D. Gunawan and H. Setiawan, "Convolutional Neural Network dalam Analisis Citra Medis," *KONSTELASI: Konvergensi Teknologi dan Sistem Informasi*, vol. 2, no. 2, pp. 376–390, Dec. 2022.
- [6] Y. K. Bintang and H. Imaduddin, "Pengembangan Model Deep Learning Untuk Deteksi Retinopati Diabetik Menggunakan Metode Transfer Learning," *JIPi (Jurnal Ilmiah Penelitian dan Pembelajaran Informatika)*, vol. 9, no. 3, pp. 1442–1455, Aug. 2024, doi: 10.29100/jipi.v9i3.5588.
- [7] Q. Li and M. Q. Yang, "Comparison of machine learning approaches for enhancing Alzheimer's disease classification," *PeerJ*, vol. 9, Feb. 2021, doi: 10.7717/peerj.10549.
- [8] R. N. Azizah, M. M. Huda, V. A. Tricahyo, and A. A. Septarina, "Implementasi Arsitektur Visual Geometry Group 16 (VGG16) untuk Deteksi Cardiomegaly pada Chest X-Ray," *Jurnal Teknik Elektro dan Komputer TRLAC*, vol. 11, no. 1, pp. 31–36, Jun. 2024.
- [9] L. Yin, R. Liu, W. Li, S. Li, and X. Hou, "Deep learning-based CT radiomics predicts prognosis of unresectable hepatocellular carcinoma treated with TACE-HAIC combined with PD-1 inhibitors and tyrosine kinase inhibitors," *BMC Gastroenterol*, vol. 25, no. 1, p. 24, Dec. 2025, doi: 10.1186/s12876-024-03555-7.
- [10] T. Berliani, E. Rahardja, and L. Septiana, "Perbandingan Kemampuan Klasifikasi Citra X-ray Paru-paru menggunakan Transfer Learning ResNet-50 dan VGG-16," *Journal of Medicine and Health*, vol. 5, no. 2, pp. 123–135, Aug. 2023, doi: 10.28932/jmh.v5i2.6116.
- [11] D. S. Uplaonkar* and Dr. B. Amarapur, "A hybrid neural filter (HNF) based on adaptive median and weiner techniques for reducing speckle noise of ultrasound liver tumor images," *International Journal of Recent Technology and Engineering*, vol. 8, no. 3, pp. 2243–2250, Sep. 2019, doi: 10.35940/ijrte.B3252.098319.
- [12] Rio Subandi, Herman, and Anton Yudhana, "Pre-Processing Pada Klasifikasi Citra Medis Pneumonia," *Decode: Jurnal Pendidikan Teknologi Informasi*, vol. 4, no. 1, pp. 86–93, Nov. 2023, doi: 10.51454/decode.v4i1.198.
- [13] H. Zhang and Y. Qie, "Applying Deep Learning to Medical Imaging: A Review," *Applied Sciences*, vol. 13, no. 18, p. 10521, Sep. 2023, doi: 10.3390/app131810521.
- [14] F. Muhammad, M. A. Arimurthy, and D. Chahyati, "Transfer learning pada Network VGG16 dan ResNet50," *The Indonesian Journal of Computer Science*, vol. 12, no. 1, p. 361, Feb. 2023.
- [15] J. Li, T. V. Nguyen, C. Hegde, and R. K. W. Wong, "Implicit Sparse Regularization: The Impact of Depth and Early Stopping," *Adv Neural Inf Process Syst*, no. 34, pp. 28298–28309, Aug. 2021, [Online]. Available: <http://arxiv.org/abs/2108.05574>
- [16] N. Khasanah, "Komparasi Arsitektur Resnet50 Dan Vgg16 Untuk Klasifikasi Citra Tanda Tangan," *JSI: Jurnal Sistem Informasi (E-Journal)*, vol. 14, no. 1, pp. 2611–2621, Apr. 2022, [Online]. Available: <http://ejournal.unsri.ac.id/index.php/jsi/index>
- [17] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016, [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>
- [18] E. H. Rachmawanto, D. Hermanto, Z. Pratama, and C. A. Sari, "Performa Convolutional Neural Network Dalam Deep Layers Resnet-50 Untuk Klasifikasi Mri Tumor Otak," *Semnas Ristek (Seminar Nasional Riset dan Inovasi Teknologi)*, vol. 8, no. 01, Jan. 2024.
- [19] S. Kornblith, J. Shlens, and Q. V. Le, "Do better imagenet models transfer better?," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Jun. 2019, pp. 2656–2666. doi: 10.1109/CVPR.2019.00277.
- [20] M. Fahmy Amin and F. Amin, "Confusion Matrix in Binary Classification Problems: A Step-by-Step Tutorial," *Journal of Engineering Research*, vol. 6, no. 5, 2022.
- [21] C. Kaope and Y. Pristyanto, "The Effect of Class Imbalance Handling on Datasets Toward Classification Algorithm Performance," *MATRIK : Jurnal Manajemen, Teknik Informatika dan Rekayasa Komputer*, vol. 22, no. 2, pp. 227–238, Mar. 2023, doi: 10.30812/matrik.v22i2.2515.