

KLASIFIKASI PENYAKIT GINJAL MENGGUNAKAN ALGORITMA HIBRIDA CNN-ELM

Ahmad Hasby Bik, Fetty Tri Anggraeny, Eva Yulia Puspaningrum

Informatika, Universitas Pembangunan Nasional "Veteran" Jawa Timur

Jalan Raya Rungkut Madya, Gunung Anyar, Surabaya, Jawa Timur, Indonesia, kode pos: 60294

fettyanggraeny.if@upnjatim.ac.id

ABSTRAK

Penyakit ginjal adalah masalah serius yang memerlukan deteksi dini. Studi ini menjelajahi model hybrid CNN-ELM untuk mengklasifikasikan gambar CT penyakit ginjal, menyoroti pentingnya pemilihan fungsi aktivasi. Dengan fokus pada gambar CT, pendekatan ini menjanjikan diagnosis yang akurat dengan akurasi tinggi, mendukung praktik klinis sehari-hari. Melalui percobaan jumlah filter dalam CNN dan neuron tersembunyi dalam ELM, performa model dapat ditingkatkan. ReLU mencapai akurasi tertinggi (0.9963), sedangkan Tanh (0.8419). Hasil ini memberikan panduan penting untuk mengoptimalkan konfigurasi model dalam mendiagnosis penyakit ginjal secara efisien. Dengan akurasi yang memuaskan, pendekatan ini berpotensi menjadi alat bantu berharga dalam praktik medis, membantu praktisi dalam membuat keputusan yang lebih baik.

Kata kunci : ELM, CNN, ReLu, Tanh, hybrid

1. PENDAHULUAN

Penyakit ginjal merupakan tantangan kesehatan global yang signifikan, memengaruhi jutaan orang di seluruh dunia. Kebutuhan akan sistem visi berbasis kecerdasan buatan (AI) dan teknik transfer learning yang dapat secara otomatis mendeteksi penyakit ginjal seperti batu ginjal, kista, dan tumor dari gambar CT-radiografi semakin meningkat. Hal ini disebabkan oleh meningkatnya prevalensi penyakit ginjal meskipun upaya kontrol yang ada, dan jumlah terbatas nephrologist dan radiologist di seluruh dunia, khususnya di Asia Selatan [1].

Teknologi medis, khususnya citra tomografi terkomputerisasi (CT), telah menjadi salah satu cara paling efektif untuk memvisualisasikan organ dalam tubuh manusia, termasuk ginjal [2]. Namun, analisis manual citra CT oleh dokter atau ahli radiologi membutuhkan waktu yang lama dan rentan terhadap kesalahan manusia.

Dalam beberapa tahun terakhir, kecerdasan buatan (AI), terutama dalam bentuk jaringan saraf konvolusi (CNN), telah semakin umum. CNN dirancang khusus untuk memproses dan menganalisis data citra (Li et al., 2022). Beberapa penelitian sebelumnya yang menggunakan CNN meliputi deteksi tumor ginjal [4], deteksi penyakit COVID pada citra sinar-X [5], dan deteksi kanker kulit [6].

Namun, CNN memiliki keterbatasan, termasuk kebutuhan akan dataset pelatihan yang besar, rentan terhadap overfitting pada dataset terbatas, dan tingkat akurasi yang belum optimal. Untuk mengatasi kelemahan ini, pendekatan hibrida yang menggabungkan CNN dan Extreme Learning Machine (ELM) menawarkan solusi efisien. Hibrida CNN-ELM meningkatkan efisiensi pelatihan dengan kemampuan pelatihan ELM yang cepat dan mengurangi risiko overfitting, bahkan pada dataset terbatas. Integrasi ELM juga meningkatkan kemampuan generalisasi model, membuatnya lebih

dapat diandalkan pada data baru. Selain itu, hasil klasifikasi Hybrid CNN-ELM menunjukkan tingkat akurasi yang lebih tinggi, kontribusi dalam meningkatkan presisi, recall, dan F1 score. Kelebihan ini membuatnya lebih efektif dalam menangani perubahan skala, rotasi objek, dan mengurangi risiko overfitting. Dengan gabungan fitur CNN dan klasifikasi ELM, pendekatan hibrida ini menjadi pilihan yang kokoh untuk meningkatkan kinerja model dalam pengenalan pola pada data medis [7].

Dalam konteks ini, penelitian yang bertujuan untuk menggabungkan teknologi CNN dengan ELM untuk mengklasifikasikan penyakit ginjal dalam dataset citra CT menjadi relevan dan penting untuk kemajuan diagnosa medis yang lebih baik [8]. Dengan memanfaatkan kekuatan AI dalam bidang kesehatan, penelitian ini berpotensi memberikan kontribusi besar dalam meningkatkan diagnosa secara daring dan pengobatan penyakit ginjal secara lebih efektif.

2. TINJAUAN PUSTAKA

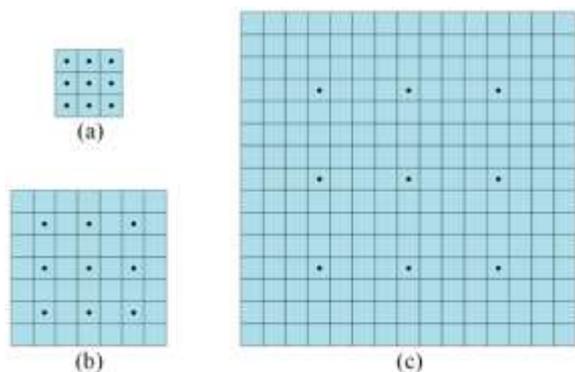
2.1. Convolutional Neural Network

CNN merupakan singkatan dari Convolutional Neural Network, adalah jenis arsitektur jaringan saraf tiruan yang digunakan untuk memproses data visual, termasuk citra. Proses kerja CNN dimulai dengan memasukkan citra ke dalam jaringan. Pada lapisan konvolusi, citra dilewatkan melalui filter atau kernel untuk mengekstrak fitur-fitur penting. Setelah itu, hasilnya melewati lapisan pooling yang mengurangi ukuran citra tetapi tetap mempertahankan informasi penting. Proses ini terus berulang pada beberapa lapisan konvolusi dan pooling.

2.1.1. Convolutional Layer

Konvolusi layer merupakan komponen penting dalam jaringan saraf konvolusional (CNN) untuk ekstraksi fitur. Ini menerapkan operasi konvolusi pada data masukan menggunakan filter atau kernel yang

dapat dipelajari untuk menghasilkan peta fitur. Proses konvolusi melibatkan pengaturan kernel konvolusi dengan ukuran tertentu, penerapan padding untuk menyesuaikan ukuran masukan, dan penggunaan langkah untuk mengontrol kepadatan konvolusi. Peta fitur yang dihasilkan oleh proses konvolusi dapat rentan menyebabkan masalah overfitting karena jumlah fitur yang besar. Untuk mengatasi hal ini, teknik pooling seperti max pooling dan average pooling digunakan untuk mengurangi redundansi [3].



Gambar 1. Perbandingan antara kernel konvolusi umum dan kernel konvolusi dilatasi [3].

Selain itu, konvolusi dilatasi, seperti yang ditunjukkan dalam Gambar 1, diperkenalkan untuk memungkinkan kernel konvolusi melihat area yang lebih besar dengan memasukkan nilai kosong di antara setiap titik kernel konvolusi. Teknik ini memungkinkan jaringan memiliki bidang reseptif yang lebih besar tanpa meningkatkan jumlah parameter.

Secara keseluruhan, lapisan konvolusi dalam CNN memainkan peran penting dalam ekstraksi fitur melalui penerapan operasi konvolusi, dan jurnal ini memberikan wawasan tentang berbagai teknik seperti konvolusi dilatasi dan konvolusi kelompok untuk meningkatkan efektivitas lapisan ini. Rumus 1 konvolusi pada setiap langkah dapat dirumuskan sebagai:

$$Y_{ij} = \sum_{m=0}^{F-1} \sum_{n=0}^{F-1} X_{(i+m)(j+n)} * K_{mn} + B \quad (1)$$

Keterangan:

- Y_{ij} adalah elemen ke-(i, j) dari peta fitur keluaran.
- $X_{(i+m)(j+n)}$ adalah elemen gambar masukan yang sedang dioperasikan oleh filter pada langkah konvolusi.
- K_{mn} adalah elemen ke-(m, n) dari filter.
- B adalah bias.

2.1.2. Pooling Layer

Lapisan Pooling berfungsi untuk mengurangi dimensi peta fitur dari lapisan sebelumnya dan mengurangi beban komputasi dengan mempertahankan informasi yang paling penting. Max pooling adalah salah satu jenis pooling yang paling

umum, di mana dapat dipilih nilai maksimum dalam jendela atau area tertentu dari peta fitur [9].

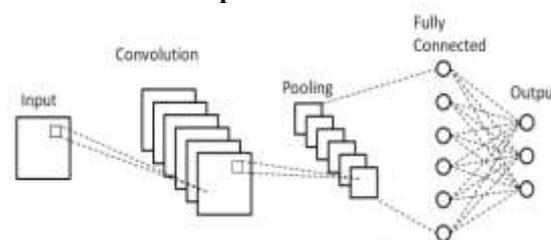
Misalnya, max pooling dengan jendela 2x2 dan stride 2 akan mengambil nilai maksimum dari setiap area 2x2 pada peta fitur dan mempertahankannya sebagai nilai representatif. Rumus 2.4 untuk max pooling dalam sebuah jendela dapat direpresentasikan sebagai:

$$Y_{ij} = \max(X_{(2i,2j)}, X_{(2i,2j+1)}, X_{(2i+1,2j)}, X_{(2i+1,2j+1)}) \quad (2)$$

Keterangan:

- Y_{ij} adalah elemen hasil dari jendela 2x2 pada peta fitur keluaran.
- $X_{(2i,2j)}$ hingga $X_{(2i+1,2j+1)}$ adalah nilai-nilai pada jendela tersebut pada peta fitur masukan.

2.1.3. Mekanisme operasi CNN

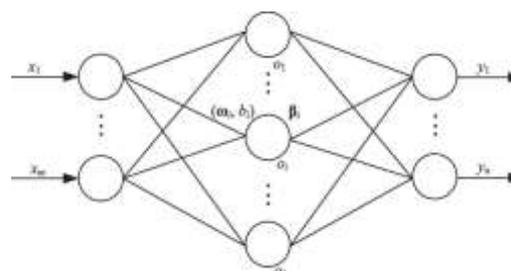


Gambar 2. Arsitektur CNN

Setelah lapisan konvolusi dan pooling, hasilnya dilewatkan melalui lapisan tersembunyi (fully connected layers) untuk klasifikasi. Proses ini melibatkan perhitungan matematis yang kompleks untuk meminimalkan galat antara hasil prediksi dan nilai yang seharusnya. Dengan memahami pola-pola dalam data visual, CNN dapat mengidentifikasi objek dalam citra dengan akurasi tinggi.

2.2. Extreme Learning Machine

Extreme Learning Machine (ELM) adalah inovasi dalam dunia jaringan saraf tiruan yang lahir pada tahun 2006. Metode ini diperkenalkan oleh Huang, Zhu, dan Siew pada sebuah makalah penelitian yang merombak paradigma pelatihan jaringan saraf tiruan. ELM dirancang sebagai alternatif untuk mempercepat proses pelatihan jaringan saraf tiruan dengan menghilangkan kompleksitas yang terkait dengan penyetulan bobot pada lapisan tersembunyi. Pendekatan inovatif ini mengubah landasan dari pembelajaran mesin tradisional.



Gambar 3. Ilustrasi arsitektur Extreme Learning Machine [7]

ELM memperkenalkan gagasan bahwa lapisan tersembunyi dalam jaringan saraf tiruan dapat diinisialisasi secara acak dengan bobot, dan tuning dari bobot ini tidak diperlukan. Konsep dasar dari ELM adalah bahwa inialisasi bobot pada lapisan tersembunyi dapat dilakukan secara acak dan pendekatan pelatihan diterapkan hanya pada output layer, tanpa modifikasi pada lapisan tersembunyi. Dengan menggunakan fungsi aktivasi non-linear seperti sigmoid atau Gaussian pada lapisan tersembunyi, ELM menghasilkan pemetaan fitur yang optimal dari input ke output dengan cepat. Rumus matematis dari pendekatan Extreme Learning Machine (ELM) adalah sebagai berikut:

Pertama, seperti pada rumus 3, output dari lapisan tersembunyi H dihitung menggunakan fungsi aktivasi $h(x)$:

$$H = h(x) = \begin{bmatrix} h_1(x) \\ h_2(x) \\ \vdots \\ h_L(x) \end{bmatrix} \quad (3)$$

Dimana L adalah jumlah neuron pada lapisan tersembunyi, $h_i(x)$ adalah aktivasi neuron ke- i terhadap input x . Selanjutnya, diberikan matriks bobot yang menghubungkan lapisan tersembunyi ke output seperti pada rumus 4, dinotasikan sebagai β :

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_L \end{bmatrix} \quad (4)$$

Output dari ELM diperoleh dengan mengalikan hasil dari lapisan tersembunyi dengan matriks bobot seperti pada rumus 5 berikut:

$$f(x) = H^T \beta \quad (5)$$

Di mana $f(x)$ adalah output dari model ELM untuk input x . Proses pelatihan ELM melibatkan perhitungan matriks pseudoinvers generalisasi dari H (dinotasikan sebagai H^+) untuk mendapatkan matriks bobot β yang optimal seperti pada rumus 6 berikut:

$$\beta = H^+ T \quad (6)$$

Di mana T adalah matriks target dari data pelatihan yang diinginkan untuk diprediksi. Matriks pseudoinvers generalisasi H^+ adalah solusi dari persamaan seperti pada rumus 7 berikut:

$$H^+ = (H^T H)^{-1} H^T \quad (7)$$

Rumus ini memungkinkan ELM untuk menghasilkan bobot output β secara efisien, yang pada gilirannya digunakan untuk menghasilkan prediksi dari model pada data yang belum dilihat sebelumnya

2.3. Confusion Matrix dan Kurva ROC

Langkah selanjutnya adalah pembuatan dan analisis konfusi matriks. Konfusi matriks memberikan pandangan yang lebih mendalam terkait kecenderungan model dalam melakukan klasifikasi

pada setiap kelas yang ada dalam dataset uji [10]. Ini membantu untuk melihat di mana model cenderung mengalami kesalahan dalam memprediksi kelas tertentu atau apakah terdapat kecenderungan tertentu pada hasil prediksinya [11].

Konfusi matriks adalah alat yang digunakan untuk mengevaluasi kinerja model klasifikasi dengan membandingkan prediksi yang dihasilkan oleh model dengan label yang sebenarnya. Matriks ini berbentuk tabel yang mencantumkan jumlah prediksi benar dan salah yang dibuat oleh model untuk setiap kelas. Dalam matriks konfusi, baris mewakili kelas sebenarnya, sementara kolom mewakili kelas yang diprediksi oleh model [12].

Matriks ini terdiri dari beberapa elemen utama: True Positive (TP), di mana model dengan benar memprediksi kelas positif; True Negative (TN), di mana model dengan benar memprediksi kelas negatif; False Positive (FP), di mana model salah memprediksi kelas positif; dan False Negative (FN), di mana model salah memprediksi kelas negatif. Dengan menggunakan konfusi matriks, berbagai metrik evaluasi seperti akurasi, presisi, recall, dan F1-score dapat dihitung, memberikan wawasan mendalam tentang kinerja model, termasuk bagaimana model menangani ketidakseimbangan kelas dan kemampuannya dalam mengidentifikasi setiap kelas secara akurat.

Tabel 1. Konfusi matriks

	Prediksi Positif	Prediksi Negatif
Aktual positif	True Positive (TP)	False Negative (FN)
Aktual Negatif	False Positive (FP)	True Negative (TN)

Metrik Evaluasi dari Konfusi Matriks:

- Akurasi (Accuracy): $(TP + TN) / (TP + TN + FP + FN)$ - Jumlah prediksi yang benar dibagi dengan total prediksi.
- Presisi (Precision): $TP / (TP + FP)$ - Kemampuan model dalam mengidentifikasi dengan benar kelas positif dari semua prediksi yang dipilih sebagai kelas positif.
- Recall atau Sensitivitas (Recall/Sensitivity): $TP / (TP + FN)$ - Kemampuan model dalam menemukan kembali semua kejadian positif.
- F1-Score: $2 * (Presisi * Recall) / (Presisi + Recall)$ - Rata-rata harmonik dari Presisi dan Recall.

Terakhir, kurva karakteristik operasi receiver (ROC) digunakan sebagai alat evaluasi yang mengukur kinerja model dalam membedakan antara kelas-kelas pada data uji. Melalui kurva ROC, seberapa baik model mampu membedakan antara kelas-kelas yang ada. Proses evaluasi ini memberikan wawasan yang komprehensif terkait dengan

kemampuan dan performa model ELM dalam mengklasifikasikan data uji.

Kurva Karakteristik Operasi Receiver (ROC) adalah alat visual yang membantu dalam mengevaluasi kinerja model klasifikasi untuk berbagai ambang batas (threshold) pemilihan kelas. Ini digunakan untuk memvisualisasikan hubungan antara tingkat True Positive Rate (TPR) dan tingkat False Positive Rate (FPR) pada berbagai nilai ambang batas [13].

Komponen Utama dari Kurva ROC:

- True Positive Rate (TPR): Jumlah positif yang benar diprediksi dibagi dengan total jumlah positif sebenarnya. TPR juga dikenal sebagai sensitivitas atau recall seperti pada rumus 8.

$$TPR = \frac{TP}{TP + FN} \tag{8}$$

- False Positive Rate (FPR): Jumlah negatif yang salah diprediksi sebagai positif dibagi dengan total jumlah negatif sebenarnya seperti pada rumus 9.

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

Pembentukan Kurva ROC:

1. Perhitungan TPR dan FPR: Pada setiap ambang batas, perhitungkan TPR dan FPR dari model.
2. Grafik TPR vs. FPR: Plot nilai TPR pada sumbu vertikal (Y-axis) dan nilai FPR pada sumbu horizontal (X-axis).

Titik Operasi:

- Titik (0,0): Mewakili model yang tidak dapat membedakan antara kelas positif dan negatif.
- Titik (1,1): Menandakan model yang sempurna dengan TPR maksimum dan FPR minimum pada setiap ambang batas.
- Garis Diagonal: Mewakili performa acak dari model.

Area Under Curve (AUC):

- AUC: Mengukur keseluruhan kinerja model. Semakin tinggi nilai AUC, semakin baik kinerja model dalam membedakan antara kelas positif dan negatif.
- Nilai AUC: Diukur dari area di bawah kurva ROC. Jika nilainya mendekati 1, menunjukkan bahwa model memiliki kinerja yang sangat baik dalam membedakan kelas. Jika mendekati 0.5, menunjukkan bahwa model tidak lebih baik daripada klasifikasi acak.

Kurva ROC dan nilai AUC memberikan wawasan yang berguna dalam memahami sejauh mana model klasifikasi dapat membedakan antara kelas yang berbeda, terutama dalam situasi di mana distribusi kelas tidak seimbang [14].

3. METODE PENELITIAN

Dalam rangka menyusun sebuah penelitian, langkah-langkah metodologi yang terstruktur sangat penting untuk memberikan arahan kepada peneliti agar

dapat berhasil menyelesaikan penelitian dengan efektif.

Hal ini bertujuan untuk mengatasi pertanyaan penelitian dan mencapai tujuan yang telah dirumuskan sebelumnya. Skema atau diagram yang memvisualisasikan langkah-langkah penelitian dapat dilihat pada gambar dibawah:



Gambar 4. Tahapan Dalam Penelitian

Studi dimulai dengan ekstraksi dataset CT-KIDNEY-DATASET-Reduced dari arsip ZIP, diikuti oleh visualisasi untuk memahami distribusi kelas penyakit ginjal. Dataset kemudian dibagi proporsional menjadi subset training, validasi, dan test. Tahap pra-pemrosesan mencakup augmentasi data untuk memperkaya variasi dalam dataset, sementara desain model Hybrid CNN-ELM dibangun secara cermat, mengintegrasikan lapisan CNN untuk ekstraksi fitur dengan ELM untuk klasifikasi. Proses pelatihan melibatkan penyesuaian parameter model, sambil meminimalkan overfitting dengan evaluasi terus-menerus pada dataset validasi.

Evaluasi model dilakukan menggunakan dataset test, dengan fokus pada akurasi, loss, dan visualisasi hasil klasifikasi. Model kemudian diuji dalam konteks praktis untuk mengevaluasi kinerja dalam situasi nyata. Laporan penelitian akan merinci semua langkah dan hasil studi, sesuai dengan standar penulisan ilmiah, memberikan dasar untuk penelitian lebih lanjut berdasarkan dataset dan model yang dikembangkan.

3.1. Pengumpulan Data

Data dikumpulkan dari Kaggle, dengan sumber dari PACS rumah sakit di Dhaka, Bangladesh, yang mencakup pasien dengan diagnosis tumor ginjal, kista, normal, atau batu ginjal. Potongan gambar DICOM dipilih secara hati-hati untuk setiap diagnosis, dikonversi ke format JPG, dan diverifikasi ulang oleh radiologis dan teknolog medis. Dataset asli terdiri dari 12,446 data unik, dengan distribusi: 3,709 sampel untuk 'Cyst', 5,077 sampel untuk 'Normal', 1,377 sampel untuk 'Stone', dan 2,283 sampel untuk 'Tumor'. Namun, untuk menjaga keseimbangan antara jumlah sampel dalam setiap kelas yang relevan untuk klasifikasi penyakit ginjal.

Dalam dataset aslinya dari Kaggle, terdapat perbedaan yang signifikan antara jumlah sampel dalam setiap kelas. Sebagai contoh, kelas 'Cyst' memiliki 5,077 sampel, sementara kelas 'Normal' hanya memiliki 1,377 sampel. Ketidakseimbangan ini dapat menyebabkan model yang dilatih cenderung menjadi bias terhadap kelas dengan jumlah sampel yang lebih besar, yang pada gilirannya dapat mempengaruhi performa model dalam mengklasifikasikan kelas-kelas yang kurang representatif. Sejumlah sampel dipilih untuk menghasilkan dataset yang lebih terkelola dan

berukuran lebih kecil, penyesuaian dataset juga membantu memperbaiki kualitas data dengan menghilangkan sampel-sampel yang mungkin tidak representatif atau bermasalah.

Hasilnya adalah dataset yang terdiri dari 2,723 file gambar yang terbagi dalam empat kelas, yaitu 'Cyst', 'Normal', 'Stone', dan 'Tumor'. Setiap kelas dalam dataset yang disesuaikan ini memiliki jumlah sampel yang telah disesuaikan menjadi lebih terkendali, dengan 'Cyst' memiliki 608 sampel, 'Normal' sebanyak 878 sampel, 'Stone' dengan 615 sampel, dan 'Tumor' memiliki 622 sampel. Dengan total 2,723 sampel, dataset yang dihasilkan tetap mempertahankan representasi yang seimbang antar kelas-kelas yang relevan dalam konteks klasifikasi penyakit ginjal.

3.2. Praproses Data

Penelitian yang dilakukan penulis mengenai klasifikasi penyakit ginjal berbasis arsitektur CNN-ELM merupakan proses yang melibatkan serangkaian tahapan yang sangat penting untuk mempersiapkan data secara tepat. Pertama-tama, proses ini dimulai dengan pengambilan dataset dari Google Drive, dataset tersebut merupakan dataset gambar kondisi ginjal ('Cyst', 'Normal', 'Stone', dan 'Tumor'). Setelah mengakses dataset tersebut, langkah berikutnya adalah memastikan integritas dan keakuratan setiap sampel data, memastikan bahwa setiap gambar terkait dengan kategori yang sesuai.

Setelah verifikasi dataset, tahap selanjutnya adalah augmentasi data. Pada tahap ini, dilakukan proses augmentasi seperti rotasi, pembalikan, dan pergeseran gambar untuk memperluas variasi gambar dalam dataset [4]. Hal ini dilakukan untuk membantu model mempelajari berbagai variasi dari data yang ada, meningkatkan kemampuan model dalam mengenali berbagai kondisi dalam gambar.

Dataset yang telah diverifikasi dan diperluas kemudian dibagi menjadi subset untuk proses pelatihan, validasi, dan pengujian model. Pembagian dataset ini penting untuk memastikan model yang dikembangkan dapat diuji dengan baik dan dapat diandalkan ketika diterapkan pada data baru.

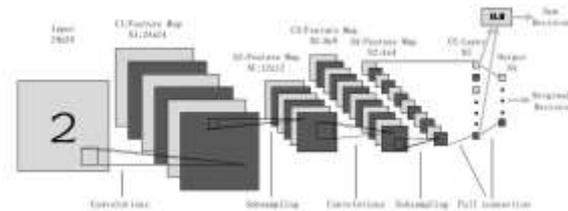
Selanjutnya, dataset dipersiapkan agar sesuai dengan format yang diharapkan oleh model. Ini termasuk mengubah ukuran gambar menjadi 128x128 piksel dan melakukan normalisasi nilai piksel agar berada dalam rentang antara 0 dan 1. Dataset yang telah diproses kemudian disimpan dalam direktori tertentu untuk memudahkan penggunaan ulang pada tahapan selanjutnya, seperti pelatihan, validasi, dan pengujian model.

Tahapan-tahapan ini memegang peranan krusial dalam memastikan kualitas dan konsistensi dataset yang digunakan dalam pengembangan model klasifikasi penyakit ginjal berbasis arsitektur CNN-ELM. Implementasi proses ini secara hati-hati membentuk landasan yang kuat bagi pengembangan model yang handal dan efektif.

3.3. Algoritma Hibrida CNN-ELM

Arsitektur hibrida CNN-ELM menggabungkan keunggulan Convolutional Neural Network (CNN) dan Extreme Learning Machine (ELM) untuk menangani diagnosa berdasarkan citra CT scan ginjal. CNN digunakan untuk mengekstraksi fitur-fitur penting dari citra, mengimplementasikan lapisan-lapisan konvolusi dan fungsi aktivasi untuk mempelajari fitur-fitur tersebut. Kekuatan CNN terletak pada kemampuannya dalam menangkap pola dan fitur berhierarki yang kompleks dari citra medis [3].

Sementara itu, ELM digunakan untuk melakukan klasifikasi berdasarkan fitur-fitur yang diekstraksi oleh CNN. ELM merupakan metode pembelajaran mesin yang cepat dan efisien dalam prosesnya. Lapisan masukan ELM menerima fitur-fitur hasil dari CNN, kemudian mengonversinya ke dalam bentuk vektor sebelum diteruskan ke lapisan tersembunyi ELM. Keunggulan ELM terletak pada kemampuannya untuk melakukan klasifikasi cepat tanpa memerlukan proses iteratif yang panjang pada pelatihannya [15].



Gambar 5. Ilustrasi Konsep Hibrida CNN-ELM

Hybrid CNN-ELM memulai prosesnya dengan citra-citra masukan yang melalui serangkaian tahapan. Pertama, setiap citra masukan diproses melalui lapisan konvolusi dari Convolutional Neural Network (CNN). Proses dimulai dengan penggunaan CNN untuk mengekstraksi fitur-fitur penting dari setiap citra masukan. CNN menggunakan serangkaian lapisan konvolusi untuk menangkap pola-pola penting dalam citra, yang kemudian dikonversi menjadi fitur-fitur numerik yang direpresentasikan dalam bentuk vector [16].

Setelah ekstraksi fitur dengan CNN, fitur-fitur yang dihasilkan digunakan sebagai masukan untuk model ELM. ELM adalah metode klasifikasi yang cepat dan efisien yang terdiri dari dua lapisan: lapisan input yang menerima fitur-fitur ekstraksi dari CNN dan lapisan output yang menghasilkan prediksi kelas. Pada tahap ini, ELM memproses fitur-fitur tersebut dan mempelajari hubungan antara fitur-fitur tersebut dan label kelas yang sesuai.

Selanjutnya, model ELM menggunakan pembelajaran satu tahap untuk menghitung bobot yang optimal antara lapisan input dan output. Bobot ini kemudian digunakan untuk mengklasifikasikan citra-citra baru berdasarkan fitur-fitur yang diperoleh dari CNN. Proses ini memungkinkan model untuk memanfaatkan keunggulan CNN dalam mengekstraksi fitur dan kemudian menggunakan pendekatan

klasifikasi yang cepat dan efisien dari ELM untuk menghasilkan prediksi akhir.

Dengan menggabungkan CNN dan ELM, hybrid CNN-ELM dapat mengatasi beberapa kelemahan yang mungkin dimiliki oleh masing-masing pendekatan secara individual. CNN mampu mengekstraksi fitur-fitur yang relevan dari citra-citra kompleks dengan tingkat abstraksi yang tinggi, sementara ELM memberikan kecepatan dan efisiensi dalam proses klasifikasi. Hasilnya, model ini dapat memberikan performa klasifikasi yang kuat dengan memanfaatkan keunggulan kedua pendekatan tersebut [17].

3.4. Skenario Pengujian

Dalam eksperimen yang dilakukan di Google Colab dengan menggunakan T4 GPU engine, terdapat serangkaian skenario yang dieksplorasi dengan memanipulasi beberapa parameter kunci pada model. Variasi tersebut meliputi ukuran filter pada lapisan konvolusi, jumlah node pada lapisan tersembunyi, dan pemilihan optimizer yang berbeda.

Pada pengaturan eksperimen ini, variasi dalam parameter kunci pada model CNN-ELM dieksplorasi dengan seksama. Salah satunya adalah aktivasi konvolusi, yang diuji menggunakan dua fungsi aktivasi berbeda: ReLU (Rectified Linear Unit) dan Tanh (Hyperbolic Tangent). Hasil eksperimen menyoroti peran penting aktivasi konvolusi dalam memengaruhi kinerja model. Selain itu, parameter lain yang dieksplorasi adalah jumlah neuron tersembunyi (hidden neurons), yang diatur dengan tiga nilai berbeda: 100, 200, dan 300. Variasi jumlah neuron tersembunyi ini memberikan wawasan tentang kompleksitas model yang dihasilkan. Semakin banyak neuron tersembunyi yang digunakan, semakin besar kapasitas representasi model terhadap data yang diberikan. Ini dapat berdampak pada kemampuan model untuk mempelajari pola yang lebih kompleks dalam data gambar dan menghasilkan representasi fitur yang lebih baik.

Pertama, eksperimen dilakukan dengan memvariasikan jumlah node pada lapisan tersembunyi, yaitu 8, 16, dan 32, dengan tujuan untuk mengevaluasi bagaimana kompleksitas model memengaruhi hasil akhirnya. Pada bagian optimizernya, beberapa optimizer populer seperti Adam, SGD, dan RMSprop digunakan dalam pengujian. Masing-masing optimizer memiliki karakteristik penyesuaian nilai bobot yang berbeda. Penggunaan T4 GPU pada Google Colab bertujuan untuk mempercepat proses pelatihan model, terutama saat menangani tugas-tugas yang membutuhkan komputasi yang intensif seperti pengolahan citra dan jaringan saraf.

Fokus utama dari eksperimen ini adalah untuk memahami dampak variasi parameter-parameter ini terhadap kinerja dan kualitas model. Skenario-skenario yang dieksplorasi dipilih dengan tujuan memberikan wawasan yang komprehensif tentang pengaturan arsitektur model, pemilihan optimizer

yang optimal, serta pemahaman tentang pengaruh ukuran filter dan kompleksitas lapisan tersembunyi dalam tugas klasifikasi citra, lalu diperoleh hasil dengan perbandingan parameter-parameter hasil yang berbeda.

4. HASIL DAN PEMBAHASAN

4.1. Evaluasi Metrik Pengujian

Tabel 2. Perbandingan evaluasi metrik CNN- ELM dari semua percobaan

Aktivasi Konvolusi	Filter	Jumlah Node	Metrik			
			Precision	Recall	F1 Score	Akurasi (%)
ReLU	8	100	0.8700	0.8713	0.8693	87
		300	0.9655	0.9651	0.9651	97
		600	0.9945	0.9945	0.9945	99
	16	100	0.8756	0.8732	0.8725	87
		300	0.9743	0.9743	0.9742	97
		600	0.9963	0.9963	0.9963	100
	32	100	0.8752	0.8695	0.8675	87
		300	0.9855	0.9853	0.9853	99
		600	0.9963	0.9963	0.9963	100
Tanh	8	100	0.8424	0.8438	0.8419	84
		300	0.9343	0.9338	0.9334	93
		600	0.9835	0.9835	0.9834	98
	16	100	0.8839	0.8842	0.8813	88
		300	0.9672	0.9669	0.9666	97
		600	0.9853	0.9853	0.9853	99
	32	100	0.8928	0.8934	0.8914	89
		300	0.9712	0.9706	0.9701	97
		600	0.9908	0.9908	0.9908	99

Hasil perbandingan metrik performa antara penggunaan fungsi aktivasi ReLU (Rectified Linear Unit) dan Tanh (Hyperbolic Tangent) pada tabel 1, merupakan model jaringan saraf tiruan menunjukkan perbedaan yang signifikan. Untuk model dengan fungsi aktivasi ReLU, rentang nilai metrik performa seperti presisi, recall, dan F1-score berkisar antara 0,8693 hingga 0,9963. Hal ini menandakan bahwa model dengan ReLU cenderung memberikan hasil yang konsisten dan sangat baik dalam mengklasifikasikan data. Selain itu, rentang nilai akurasi model dengan ReLU mencapai 87 hingga 100 persen untuk semua filter yang digunakan.

Di sisi lain, penggunaan fungsi aktivasi Tanh menghasilkan rentang nilai metrik performa yang sedikit lebih rendah, yaitu antara 0,8419 hingga 0,9908. Meskipun demikian, nilai-nilai ini masih menunjukkan kinerja yang baik dalam klasifikasi data. Rentang nilai akurasi untuk model dengan fungsi Tanh berkisar antara 84 hingga 99 persen pada semua filter yang digunakan.

Perbedaan dalam rentang nilai metrik performa antara ReLU dan Tanh dapat disebabkan oleh karakteristik masing-masing fungsi aktivasi. ReLU cenderung lebih efektif dalam mengatasi masalah vanishing gradient dan mempercepat proses pelatihan karena sifatnya yang non-linear dan konvergensi yang lebih cepat. Di sisi lain, Tanh memiliki rentang nilai output yang lebih terpusat di sekitar nol, sehingga

mungkin lebih cocok untuk data yang terstandarisasi dengan baik. Meskipun hasil perbandingan menunjukkan bahwa model dengan ReLU cenderung memberikan kinerja yang sedikit lebih baik, pemilihan fungsi aktivasi harus dipertimbangkan secara cermat berdasarkan karakteristik data dan tujuan dari model yang dibangun. Kedua fungsi aktivasi tersebut menawarkan kelebihan dan kekurangan masing-masing, dan pemilihan yang tepat dapat membantu meningkatkan kualitas dan performa model secara keseluruhan.

4.2. Matriks Konfusi

Matriks Konfusi adalah alat evaluasi yang penting dalam data mining untuk mengukur kinerja model klasifikasi. Matriks ini membandingkan hasil prediksi model dengan membandingkan nilai yang diprediksi dengan nilai yang sebenarnya. Dengan memperhatikan empat kemungkinan hasil dari prediksi dan label sebenarnya, yaitu True Positive (TP), True Negative (TN), False Positive (FP), dan False Negative (FN), kita dapat menghitung metrik evaluasi yang penting. Salah satu metrik yang dihasilkan dari matriks konfusi adalah akurasi, yang mengukur seberapa sering model klasifikasi benar dalam mengklasifikasikan data. Akurasi dihitung dengan membagi jumlah prediksi yang benar (TP dan TN) dengan jumlah total prediksi.

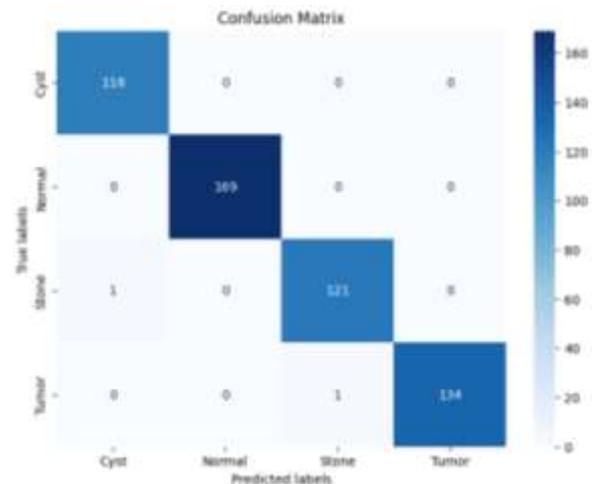
Selain akurasi, matriks konfusi juga digunakan untuk menghitung Precision, yang merupakan proporsi dari prediksi positif yang benar dibandingkan dengan total prediksi positif yang dilakukan oleh model. Recall (Sensitivity) mengukur seberapa baik model dalam menemukan kembali semua instance dari kelas positif. F-Measure adalah nilai rata-rata harmonik dari Precision dan Recall, memberikan gambaran yang lebih komprehensif tentang kinerja model dalam mengklasifikasikan data.



Gambar 6. Matriks Konfusi terendah pada percobaan yang dilakukan penulis

Pada gambar 3 merupakan hasil konfusi matriks terendah dari percobaan aktivasi konvolusi Tanh

dengan filter 8 dan lapisan tersembunyi sebesar 100. Terdapat 2006 data yang bernilai benar dan terdapat 170 data bernilai salah. Dapat disimpulkan bahwa nilai True Positive (TP) ialah 459, True Negative (TN) adalah 1547, False Positive (FP) yaitu 84, dan False Negative (FN) berjumlah 86.



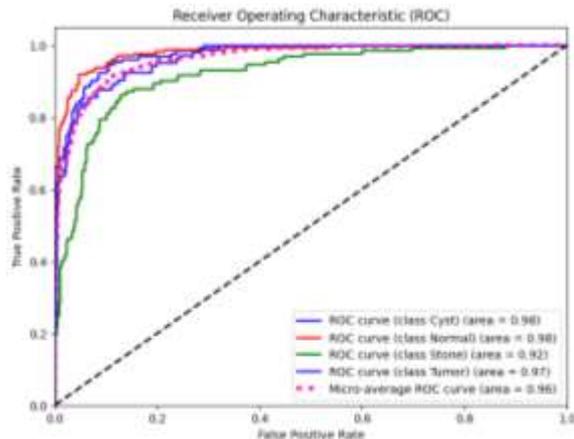
Gambar 7. Matriks Konfusi tertinggi pada percobaan yang dilakukan penulis

Pada gambar 4 merupakan hasil konfusi matriks tertinggi dari percobaan aktivasi konvolusi ReLU dengan filter 32 dan lapisan tersembunyi sebesar 600. Terdapat 2172 data yang bernilai benar dan terdapat 4 data bernilai salah. Dapat disimpulkan bahwa nilai True Positive (TP) ialah 542, True Negative (TN) adalah 1630, False Positive (FP) yaitu 1, dan False Negative (FN) berjumlah 3.

4.3. Matriks Konfusi

Kurva ROC (Receiver Operating Characteristic) adalah alat evaluasi yang penting dalam analisis kinerja model klasifikasi, khususnya dalam konteks pengukuran sensitivitas dan spesifisitas. Kurva ROC digunakan untuk memvisualisasikan hubungan antara tingkat sensitivitas (True Positive Rate) dan tingkat spesifisitas (False Positive Rate) dari suatu model klasifikasi yang berbeda.

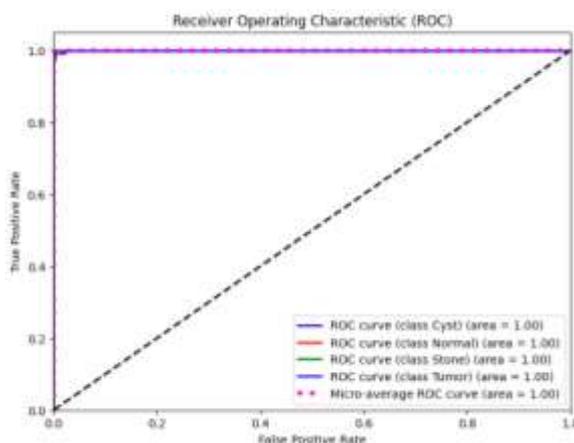
Dengan menggambar kurva ROC, kita dapat memperoleh pemahaman yang mendalam tentang seberapa baik model kita dalam membedakan antara kelas positif dan negatif. Titik-titik di kurva ROC mewakili nilai sensitivitas dan spesifisitas pada berbagai titik ambang yang berbeda yang digunakan untuk mengklasifikasikan data.



Gambar 8. Kurva ROC terendah pada percobaan yang dilakukan penulis

Pada gambar 5 menyajikan evaluasi kinerja model klasifikasi untuk empat kelas yang berbeda. Setiap kurva ROC menggambarkan tingkat sensitivitas (True Positive Rate) terhadap tingkat spesifisitas (False Positive Rate) pada berbagai ambang pengklasifikasi.

Pertama, kelas cyst memiliki nilai AUC sebesar 0.98, yang menunjukkan bahwa model memiliki tingkat sensitivitas dan spesifisitas yang sangat tinggi dalam membedakan antara sampel yang termasuk dalam kelas cyst dan bukan. Kedua, kelas normal juga memiliki nilai AUC sebesar 0.98, menunjukkan bahwa model memiliki kinerja yang sangat baik dalam mengklasifikasikan sampel sebagai normal atau bukan. Ketiga, kelas stone memiliki nilai AUC sebesar 0.92, yang masih merupakan nilai yang tinggi, tetapi sedikit lebih rendah dibandingkan dengan kelas cyst dan normal. Ini menunjukkan bahwa model memiliki sensitivitas dan spesifisitas yang tinggi dalam mengidentifikasi sampel sebagai stone, meskipun ada sedikit penurunan dalam kinerja dibandingkan dengan kelas lainnya. Keempat, kelas tumor memiliki nilai AUC sebesar 0.97, menunjukkan bahwa model juga memiliki kinerja yang sangat baik dalam membedakan antara sampel tumor dan bukan tumor.



Gambar 9. Kurva ROC tertinggi pada percobaan yang dilakukan penulis

Secara keseluruhan, nilai micro average sebesar 0.96 menunjukkan bahwa model memiliki kinerja yang sangat baik secara keseluruhan dalam mengklasifikasikan semua kelas. Ini adalah hasil yang menggembirakan dan menunjukkan bahwa model ini efektif dalam mengidentifikasi berbagai jenis penyakit ginjal dari gambar CT.

Pada gambar 6 menyajikan evaluasi kinerja model klasifikasi untuk empat kelas yang berbeda. Gambar kurva ROC ini menunjukkan evaluasi kinerja model klasifikasi yang luar biasa tinggi.

Pertama, kelas cyst memiliki nilai AUC sempurna, yaitu 1.00, menunjukkan bahwa model memiliki tingkat sensitivitas dan spesifisitas yang sempurna dalam membedakan antara sampel yang termasuk dalam kelas cyst dan yang tidak. Kedua, kelas normal juga memiliki nilai AUC sempurna, yaitu 1.00, menunjukkan bahwa model memiliki kinerja yang luar biasa dalam mengklasifikasikan sampel sebagai normal atau bukan. Ketiga, kelas stone juga memiliki nilai AUC sempurna, yaitu 1.00, menunjukkan bahwa model memiliki sensitivitas dan spesifisitas yang sempurna dalam mengidentifikasi sampel sebagai stone. Keempat, kelas tumor juga memiliki nilai AUC sempurna, yaitu 1.00, menunjukkan bahwa model memiliki kinerja yang luar biasa dalam membedakan antara sampel tumor dan bukan tumor.

Secara keseluruhan, nilai micro average sebesar 1.00 menunjukkan bahwa model memiliki kinerja yang sempurna secara keseluruhan dalam mengklasifikasikan semua kelas. Ini adalah hasil yang luar biasa dan menunjukkan bahwa model ini sangat efektif dalam mengidentifikasi berbagai jenis penyakit ginjal dari gambar CT.

5. KESIMPULAN DAN SARAN

Kesimpulan dari pengerjaan aplikasi ini adalah Hasil pengujian yang dilakukan oleh penulis menunjukkan bahwa kinerja terbaik dari algoritma CNN-ELM terjadi saat menggunakan sktivasi konvolusi Relu dan struktur yang terdiri dari 32 filter pada lapisan CNN serta 600 node pada lapisan tersembunyi ELM. Dengan konfigurasi tersebut, model berhasil mencapai akurasi sebesar 100%. Selain itu, nilai rata-rata presisi, F1-Score, dan recall juga mencapai 0.9963, menunjukkan kualitas model yang baik dalam mengklasifikasikan data. Penyetelan parameter ini memperlihatkan pengaruhnya yang signifikan terhadap performa keseluruhan algoritma, menegaskan pentingnya pemilihan struktur yang optimal dalam pengembangan model klasifikasi.

Pada penelitian berikutnya, disarankan untuk meningkatkan keandalan dan generalisasi model, disarankan untuk menambahkan lebih banyak data latih yang representatif. Pengumpulan data tambahan dari sumber yang beragam dapat membantu mengatasi potensi bias dan meningkatkan kemampuan model dalam mengklasifikasikan citra dengan lebih akurat.

DAFTAR PUSTAKA

- [1] M. N. Islam, M. Hasan, M. K. Hossain, M. G. R. Alam, M. Z. Uddin, and A. Soyly, "Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography," *Sci Rep*, vol. 12, no. 1, Dec. 2022, doi: 10.1038/s41598-022-15634-4.
- [2] M. Nahiduzzaman *et al.*, "Parallel CNN-ELM: A multiclass classification of chest X-ray images to identify seventeen lung diseases including COVID-19," *Expert Syst Appl*, vol. 229, Nov. 2023, doi: 10.1016/j.eswa.2023.120528.
- [3] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.
- [4] S. Abdul Hannan Babasaheb Ambedkar, P. Pal, and S. Abdul Hannan, "Mini Review," *J Neurol Neurorehab Res*, vol. 8, p. 1, 2023, doi: 10.35841/aajnnr-8.2.136.
- [5] M. Harahap, Em Manuel Laia, Lilis Suryani Sitanggang, Melda Sinaga, Daniel Franci Sihombing, and Amir Mahmud Husein, "Deteksi Penyakit Covid-19 Pada Citra X-Ray Dengan Pendekatan Convolutional Neural Network (CNN)," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 6, no. 1, pp. 70–77, Feb. 2022, doi: 10.29207/resti.v6i1.3373.
- [6] R. AGUSTINA, R. MAGDALENA, and N. K. C. PRATIWI, "Klasifikasi Kanker Kulit menggunakan Metode Convolutional Neural Network dengan Arsitektur VGG-16," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 10, no. 2, p. 446, Apr. 2022, doi: 10.26760/elkomika.v10i2.446.
- [7] Amity University, Amity University. School of Engineering and Technology, Institute of Electrical and Electronics Engineers. Uttar Pradesh Section, and Institute of Electrical and Electronics Engineers, *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN) : 22-23 February, 2018, Amity School of Engineering and Technology, Noida, India*.
- [8] S. Zhou and B. Tan, "Electrocardiogram soft computing using hybrid deep learning CNN-ELM," *Applied Soft Computing Journal*, vol. 86, Jan. 2020, doi: 10.1016/j.asoc.2019.105778.
- [9] R. T. Schirrmeyer *et al.*, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum Brain Mapp*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017, doi: 10.1002/hbm.23730.
- [10] P. Cavalin and L. Oliveira, "Confusion matrix-based building of hierarchical classification," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Springer Verlag, 2019, pp. 271–278. doi: 10.1007/978-3-030-13469-3_32.
- [11] J. T. Townsend, "Theoretical analysis of an alphabetic confusion matrix*," 1971.
- [12] D. Krstinić, M. Braović, L. Šerić, and D. Božić-Štulić, "Multi-label Classifier Performance Evaluation with Confusion Matrix," Academy and Industry Research Collaboration Center (AIRCC), Jun. 2020, pp. 01–14. doi: 10.5121/csit.2020.100801.
- [13] N. R. Cook, "Use and misuse of the receiver operating characteristic curve in risk prediction," *Circulation*, vol. 115, no. 7, pp. 928–935, Feb. 2007, doi: 10.1161/CIRCULATIONAHA.106.672402.
- [14] B. J. Mcneil and J. A. Hanley, "Statistical Approaches to the Analysis of Receiver Operating Characteristic (ROC) Curves."
- [15] S. Revathy, B. Bharathi, P. Jeyanthi, and M. Ramesh, "Chronic kidney disease prediction using machine learning models," *Int J Eng Adv Technol*, vol. 9, no. 1, pp. 6364–6367, Oct. 2019, doi: 10.35940/ijeat.A2213.109119.
- [16] S. Frangky Handono, F. Tri Anggraeny, and B. Rahmat, "IMPLEMENTASI CONVOLUTIONAL NEURAL NETWORK (CNN) UNTUK DETEKSI RETINOPATI DIABETIK," 2020.
- [17] B. Nugroho and E. Yulia, "KINERJA METODE CNN UNTUK KLASIFIKASI PNEUMONIA DENGAN VARIASI UKURAN CITRA INPUT," vol. 8, no. 3, pp. 533–538, 2021, doi: 10.25126/jtiik.202184515