



Perbandingan Akurasi Arsitektur EfficientNet-B0, VGG16, dan Inception V3 Dalam Deteksi Tumor Ginjal Pada Citra CT-Scan

Muhammad Fahri, Febi Yanto*, Fadhilah Syafria, Rahmad Abdillah

Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Islam Negeri Sultan Syarif Kasim Riau, Pekanbaru, Indonesia

Email: ¹11950113429@students.uin-suska.ac.id, ^{2,*}febiyanto@uin-suska.ac.id, ³fadhilahsyafria@uin-suska.ac.id,

⁴rahmadabdillah@uin-suska.ac.id

Email Penulis Korespondensi: febiyanto@uin-suska.ac.id

Abstrak—Gangguan pada fungsi ginjal dapat memicu munculnya berbagai penyakit, termasuk tumor ginjal. Deteksi dini tumor ginjal sangat penting untuk meningkatkan efektivitas pengobatan dan peluang pemulihan pasien. Pemanfaatan teknologi *deep learning* dalam klasifikasi citra medis menjadi pendekatan menjanjikan, terutama untuk mendeteksi kelainan pada organ ginjal melalui gambar CT-Scan. Penelitian ini membandingkan performa tiga arsitektur Convolutional Neural Network (CNN), yakni EfficientNet-B0, Inception-V3, dan VGG16, dalam mendeteksi tumor ginjal. Dataset yang digunakan diperoleh dari situs Kaggle berupa gambar CT-scan dengan kelas normal dan tumor, lalu dibagi dalam rasio data latih dan data uji 80:20. Hyperparameter yang digunakan adalah Stochastic Gradient Descent (SGD) dengan learning rate 0,001 dan 0,0001. Evaluasi dilakukan menggunakan confusion matrix dengan metrik *accuracy*, *precision*, *recall*, dan *F1-score*. Berdasarkan hasil pengujian, model VGG16 yang dikonfigurasi dengan learning rate 0,001 mencapai performa klasifikasi tertinggi, dengan *accuracy*, *recall*, dan *F1-score* masing-masing sebesar 99,46%.

Kata Kunci: EfficientNet-B0; Inception-V3; VGG16; Tumor Ginjal; CT-Scan

Abstract—Kidney dysfunction can trigger the development of various diseases, including kidney tumors. Early detection of kidney tumors is very important to increase the effectiveness of treatment and the chances of patient recovery. The use of *deep learning* technology in medical image classification has become a promising approach, especially in detecting abnormalities in the kidney organ through CT-Scan images. This study compares the performance of three Convolutional Neural Network (CNN) architectures, namely EfficientNet-B0, Inception-V3, and VGG16, in detecting kidney tumors. The dataset used was obtained from the kaggle website, namely CT-scan images with normal and tumor classes and divided by a ratio of *training* data and test data of 80:20. The hyperparameter used is Stochastic Gradient Descent (SGD) with a learning rate of 0.001 and 0.0001. The evaluation was carried out using a confusion matrix with metrics of *accuracy*, *precision*, *recall*, and *F1-score*. According to the test outcomes, the VGG16 model configured with a 0.001 learning rate achieved the highest classification performance, recording 99.46% *accuracy*, *precision*, *recall*, and *F1-score*.

Keywords: EfficientNet-B0; Inception-V3; VGG16; Kidney Tumor; CT-Scan

1. PENDAHULUAN

Ginjal adalah salah satu organ vital dalam tubuh manusia yang berperan utama dalam membuang sisa hasil metabolisme seperti ureum, kreatinin, dan asam urat [1]. Gangguan pada fungsi ginjal dapat memicu munculnya berbagai penyakit, termasuk tumor ginjal. Tumor sendiri merupakan pertumbuhan jaringan atau sel yang berlangsung secara terus-menerus dan tidak terkontrol [2]. Gharaibeh dan rekan-rekannya menyatakan bahwa tumor ginjal menempati urutan ke-10 sebagai jenis tumor yang paling sering terjadi pada pria maupun wanita di seluruh dunia [3]. Oleh sebab itu, deteksi dini terhadap tumor ginjal (KT) menjadi langkah krusial dalam mencegah perkembangan penyakit yang lebih serius [4].

CT-scan, singkatan dari *computed tomography*, dianggap sebagai salah satu teknik pencitraan paling unggul dan telah menjadi acuan utama dalam proses analisis serta evaluasi tumor pada ginjal [5]. Metode ini dikenal sangat akurat dalam mendeteksi massa ginjal, dengan tingkat sensitivitas mencapai 100% dan spesifisitas sebesar 95% [6]. Kecerdasan buatan (AI) telah memberikan dampak signifikan dalam bidang layanan kesehatan. Kemajuan dalam pencitraan medis belakangan ini didorong oleh meningkatnya kapasitas penyimpanan big data serta integrasi dengan teknologi *machine learning* dan *deep learning*. Hal ini memungkinkan pengembangan metode inovatif untuk deteksi dini tumor ginjal [3]. *Convolutional Neural Networks* (CNN) merupakan jenis jaringan saraf dalam pembelajaran mendalam yang dirancang khusus untuk keperluan analisis citra dan pengolahan visual [7].

Berbagai penelitian sebelumnya terkait CNN telah menunjukkan hasil yang memuaskan. Studi yang dilakukan oleh Wahyuni Rizky Perdani dan tim pada tahun 2022 berhasil mengembangkan sistem klasifikasi penyakit glaukoma menjadi lima kelas menggunakan model CNN EfficientNet-B0. Penelitian tersebut mencapai akurasi, presisi, *recall*, dan *F1-score* masing-masing sebesar 1,0000, dengan pengaturan optimizer AdaMax, learning rate 0,001, dan batch size 32 [8]. Sementara itu, penelitian oleh Nabila Asryani Sundari dan rekan-rekan pada tahun yang sama mengenai klasifikasi jenis kulit wajah menemukan bahwa sistem terbaik untuk membedakan empat tipe kulit—normal, kering, berminyak, dan kombinasi—mampu mencapai akurasi 100% tanpa mengalami *overfitting*. Penelitian ini menggunakan optimizer Adam, learning rate 0,0001, sebanyak 50 epoch, dropout 0,3, dan batch size 16, serta proporsi data *training* 80% dan data *testing* 20%. Hasil evaluasi menunjukkan nilai loss sebesar 0,001 dengan *precision*, *recall*, dan *F1-score* masing-masing mencapai 100% [9]. Pada tahun 2024, Alfataniah Nur Fajrina bersama tim melakukan penelitian mengenai klasifikasi Acute Lymphoblastic Leukemia menggunakan arsitektur EfficientNet-B0. Penelitian ini menguji beberapa kombinasi hyperparameter, termasuk jumlah epoch (20, 30, 50), learning rate (0,0001, 0,0003, 0,001, 0,003), serta dua jenis optimizer, yaitu Adam dan RMSProp. Hasil paling optimal diperoleh pada konfigurasi learning rate 0,0001, optimizer Adam, dan 30 epoch, dengan akurasi validasi sebesar 97,84% dan akurasi pengujian sebesar 98,48% [10].



Berbagai arsitektur CNN diterapkan sesuai kebutuhan setiap penelitian. Sihotang et al. (2017) meneliti klasifikasi ekspresi wajah menggunakan VGG-Net, dan menemukan bahwa konfigurasi optimal—learning rate 0,0001, optimizer Adam, serta 100 epoch memberikan akurasi validasi tertinggi 66% (0,658) pada VGG16 setelah augmentasi data dan 83 epoch [11]. Penelitian yang dilakukan oleh Ferani Christy dan rekan-rekannya pada tahun 2022 menunjukkan bahwa penerapan teknik CLAHE pada model VGG16 berhasil mencapai tingkat akurasi yang tinggi, yakni 96,84%, dengan nilai *precision*, *recall*, dan *F1-score* masing-masing sebesar 96,83%. Parameter yang digunakan dalam studi ini mencakup batch size sebesar 5, learning rate 0,00001, sebanyak 20 epoch, serta dropout rate sebesar 0,4 [12]. Prasetyo (2023), dalam “Deteksi SARS-CoV-2 dari Citra CT- Scan Paru Menggunakan Fine Tuning pada Model Deep-CNN Terlatih”, mencatat akurasi *testing* tertinggi 97,59% pada VGG16 hasil fine tuning [13]. Terakhir, Idawati (2024) mengembangkan model VGG16 untuk melakukan klasifikasi kanker payudara ke dalam tiga kategori, yaitu normal, benign, dan malignant, dengan melalui tahapan pre-processing serta pengujian hyperparameter seperti optimizer, learning rate, epoch, dan batch size. Model ini berhasil mencapai akurasi sebesar 78,87%, *precision* 75,69%, *recall* 79,85%, dan *F1-score* sebesar 74,67% [14].

Inception-V3 juga kerap digunakan dalam perancangan model CNN. Sebagai contoh, pada tahun 2022, Mujiibur-Rahman Kapa dan rekan-rekannya menerapkan arsitektur Inception-V3 untuk klasifikasi citra leukemia, dengan hasil akurasi, sensitivitas, dan spesifisitas masing-masing mencapai 100%. Kombinasi hyperparameter terbaik menggunakan RMSProp dengan momentum 0,9 (mini-batch 8, learning rate 0,001), serta RMSProp dengan momentum 0,99 dan 0,999 (mini-batch 16, learning rate 0,0001). Meskipun begitu, akurasi tertinggi pada tahap pengujian—baik dengan maupun tanpa augmentasi—adalah 98,96%, dan variasi ukuran mini-batch tidak memberikan pengaruh signifikan terhadap kinerja model [15]. Pada 2023, Muhammad Ahnaf Amrullah dan Mohammad Isa Irawan menunjukkan efektivitas Inception-V3 untuk deteksi katarak pada gambar funduskopi digital, dengan hasil terbaik menggunakan CLAHE dan fine-tuning, mencapai akurasi 98,33% [16]. Tahun berikutnya Andi Nuridin dkk. menggunakan Inception-V3 untuk klasifikasi penyakit daun tomat, hasil klasifikasi menggunakan optimizer Adam, SGD, dan RMS Prop menghasilkan bahwa optimizer Adam memiliki hasil yang paling baik, dibuktikan dengan performa model yang stabil dan akurasi yang paling tinggi dengan jumlah 93,8%. [17]. Terbaru, pada 2024 Novia Wahyu Wulansari dan Muslih menerapkan Inception-V3 untuk mendeteksi kematangan buah pisang menggunakan 919 citra (641 *training*, 94 *testing*, 184 validasi), menghasilkan akurasi *training* 95% dan pada confusion matrix *precision* 93%, *recall* 94%, serta akurasi 94% [18].

Berbagai penelitian pada CNN yang menggunakan dataset medis khusus nya citra ct-scan telah dilakukan. Penelitian yang dilakukan oleh Alfitra Salam dan rekan-rekannya pada tahun 2023 membandingkan performa klasifikasi citra CT-Scan kanker paru-paru menggunakan teknik contrast stretching pada arsitektur CNN EfficientNet-B0. Dataset yang digunakan mencakup 1.097 citra CT paru-paru yang dibagi menjadi dua kelompok: data asli dan data dengan kontras yang telah ditingkatkan. Pembagian data dilakukan dengan rasio 90:10 untuk pelatihan dan validasi, tanpa ada duplikasi gambar. Hasil penelitian menunjukkan bahwa penerapan contrast stretching menghasilkan performa yang lebih baik dibandingkan dengan data asli, dengan hasil tertinggi diperoleh saat menggunakan optimizer ADAM dan learning rate 0,01, yang memberikan akurasi 72,48%, presisi 71,52%, *recall* 64,2%, dan *F1-score* 64,76% [19]. Pada tahun yang sama, Dzaky Abdillah Salafy dan tim melakukan penelitian terkait klasifikasi citra CT-Scan kanker paru-paru dengan memanfaatkan teknik image enhancement CLAHE pada arsitektur EfficientNet-B0. Penelitian ini mengklasifikasikan citra menjadi dua kelas, yaitu normal dan kanker, menggunakan total 1.537 gambar yang terbagi menjadi dua jenis: citra asli dan citra hasil CLAHE, dengan rasio data 90:10. Dataset terdiri atas 749 citra normal dan 788 citra kanker. Hasil eksperimen terbaik diperoleh dari citra yang diproses dengan CLAHE menggunakan konfigurasi clip limit 2 dan tile grid 8×8, serta pengaturan optimizer Adam dan learning rate 0,1. Kombinasi ini menghasilkan *F1-score* sebesar 84,45% dan akurasi 84,71% [20].

Abdelrahman Elmaddah (2024) memanfaatkan dataset citra CT-scan ginjal yang terbagi ke dalam empat kategori Normal, Cyst, Tumor, dan Stone dengan setiap kelas disimpan dalam folder terpisah lalu diekstrak ke dalam DataFrame yang memuat path dan label gambar. Seluruh citra diubah ukurannya menjadi 224×224 piksel, kemudian dataset dibagi secara stratifikasi menjadi 80% data latih, 10% validasi, dan 10% uji agar proporsi tiap kelas tetap seimbang. Untuk memasukkan data ke model, digunakan ImageDataGenerator dari Keras tanpa augmentasi tambahan, sehingga batch processing berjalan efisien. Model yang dirancang adalah CNN sederhana, terdiri dari dua lapisan Conv2D (32 filter, ReLU) masing-masing diikuti MaxPooling2D, lalu lapisan Flatten, satu lapisan Dense (128 unit, ReLU), dan output layer Softmax dengan empat neuron untuk klasifikasi. Kompilasi model memakai optimizer Adam dan loss function categorical crossentropy, dengan metrik evaluasi akurasi. Setelah dilatih selama 10 epoch, model ini cepat konvergen akurasi validasi sudah mencapai 100% sejak epoch kedua dan bertahan sempurna di data latih, validasi, dan uji, dengan loss masing-masing 0,000012, 0,000194, dan 0,000137. Analisis confusion matrix mengonfirmasi bahwa ke-1.245 citra uji terklasifikasi tanpa kesalahan, dan *precision*, *recall*, serta *F1-score* untuk semua kelas mencapai 1,00, menunjukkan model ini sangat akurat dan seimbang dalam mengenali tiap kondisi ginjal, sehingga sangat potensial dikembangkan menjadi sistem pendukung diagnosis otomatis berbasis *deep learning* pada citra CT ginjal [21]

Pemilihan arsitektur EfficientNet-B0, Inception-V3, dan VGG16 didasarkan pada performa dan tingkat adopsinya yang tinggi dalam bidang klasifikasi citra medis, khususnya citra CT-Scan. VGG16 merupakan arsitektur klasik yang banyak dijadikan baseline karena kesederhanaan strukturnya dan performa yang stabil [22]. Inception-V3 menawarkan efisiensi komputasi yang baik melalui mekanisme convolusi multi-skala, sehingga efektif dalam menangkap variasi tekstur dan bentuk dalam citra medis [23]. Sementara itu, EfficientNet-B0 merupakan arsitektur modern yang



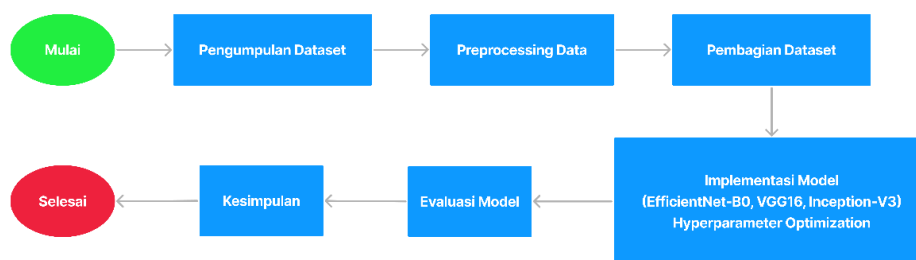
mengoptimalkan skala model secara sistematis dan mampu menghasilkan akurasi tinggi dengan jumlah parameter yang jauh lebih kecil dibandingkan arsitektur sebelumnya [24]. Oleh karena itu, membandingkan ketiga arsitektur ini penting untuk mengidentifikasi model terbaik dalam mendeteksi tumor ginjal pada citra CT-Scan, baik dari segi akurasi maupun efisiensi.

Berdasarkan latar belakang dan permasalahan di atas, penelitian ini bertujuan membandingkan kinerja tiga arsitektur CNN EfficientNet-B0, Inception-V3, dan VGG16 dalam mendeteksi tumor ginjal pada citra CT-Scan. Evaluasi dilakukan menggunakan metrik akurasi, presisi, recall, dan F1-score guna menentukan model yang paling optimal dan efisien untuk keperluan klasifikasi di bidang medis. Diharapkan, hasil penelitian ini dapat memberikan kontribusi yang berarti dalam pengembangan sistem deteksi dini tumor ginjal berbasis deep learning, serta menjadi acuan bagi peneliti dan praktisi dalam memilih arsitektur CNN yang paling sesuai untuk tugas klasifikasi citra medis, khususnya pada gambar CT-Scan.

2. METODOLOGI PENELITIAN

2.1 Tahapan Penelitian

Penelitian ini dilakukan melalui enam tahap utama, yaitu pengumpulan dataset, *preprocessing* data, pembagian dataset, implementasi model, serta penarikan kesimpulan, sebagaimana ditunjukkan di Gambar 1.



Gambar 1. Tahapan Penelitian

2.2 Pengumpulan Dataset

Pada bagian ini dataset yang digunakan diperoleh dari situs Kaggle melalui tautan berikut: <https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone>. Dataset ini terdiri dari 12.446 citra CT-Scan ginjal yang telah melalui proses verifikasi oleh tenaga medis. Dataset ini dikumpulkan dari sistem PACS (Picture Archiving and Communication System) di berbagai rumah sakit di Dhaka, Bangladesh. Dataset mencakup empat kelas yaitu: normal (5.077 citra), kista (3.709 citra), batu ginjal (1.377 citra), dan tumor (2.283 citra). Namun, dalam penelitian ini hanya dua kelas yang digunakan, yaitu kelas normal dan kelas tumor.

2.3 Preprocessing Data

Pada langkah ini, data citra CT-Scan ginjal dengan resolusi awal 512×512 piksel dipra-pemrosesan. Tahap pertama meliputi pemotongan otomatis (*cropping*) untuk menyesuaikan area fokus gambar, yang diimplementasikan menggunakan Python dan pustaka OpenCV. OpenCV digunakan karena mendukung deteksi kontur dan tepi objek secara efisien, yang penting dalam mengisolasi organ ginjal dari latar belakang.

2.4 Pembagian Dataset

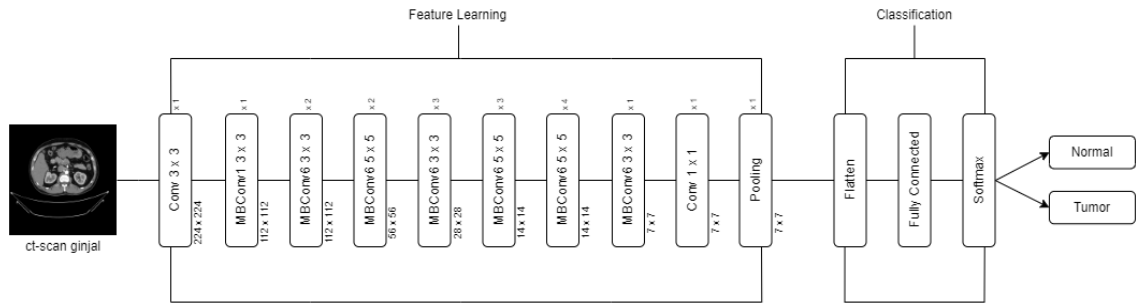
Pada langkah ini, dataset dibagi menjadi dua bagian untuk pelatihan dan pengujian model dengan perbandingan 80:20. Sebanyak 80% data dialokasikan untuk pelatihan dan validasi, sedangkan 20% sisanya dipakai sebagai data pengujian. Data pelatihan dan validasi berperan dalam proses pembelajaran model untuk menyesuaikan bobot parameter, sementara data pengujian digunakan untuk menilai kinerja model yang sudah dilatih. Rasio 80:20 dipilih karena rasio ini memberikan performa stabil dalam kasus klasifikasi [25], dengan hasil akurasi dan metrik evaluasi lainnya yang unggul dibandingkan rasio pembagian data lain, sebagaimana ditunjukkan dalam penelitian klasifikasi tingkat obesitas menggunakan LightGBM [26].

2.5 Implementasi Model

Setelah selesai melakukan pembagian dataset, tahapan selanjutnya adalah implementasi *deep learning* dengan melakukan pelatihan terhadap masing-masing arsitektur yaitu *EfficientNet-B0*, *VGG16*, dan *Inception-V3*.

a. EfficientNet-B0

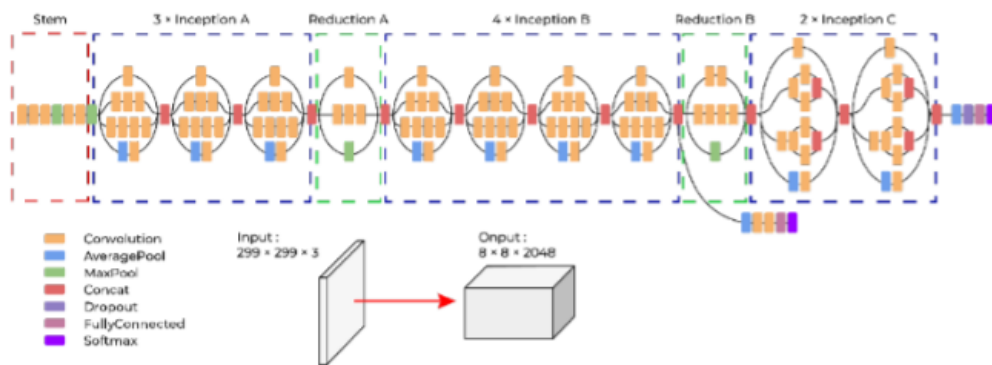
EfficientNet-B0 adalah salah satu arsitektur model pre-trained yang mampu menghasilkan akurasi tinggi dan meningkatkan kinerja model melalui pengoptimalan parameter [27]. Gambar 2 menunjukkan ilustrasi arsitektur EfficientNet-B0.



Gambar 2. Arsitektur EfficientNet-B0

b. Inception-V3

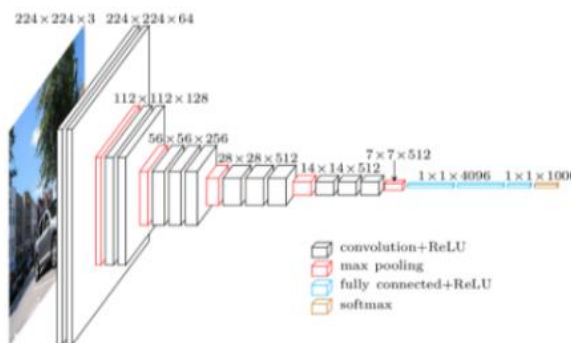
Inception-V3 adalah versi pengembangan dari Inception-V1 (GoogLeNet). Arsitektur ini dirancang di Keras dan telah melalui proses pelatihan awal pada dataset ImageNet. Ukuran input standar untuk citra adalah 299×299 piksel dengan tiga saluran warna [16]. Berikut adalah ilustrasi arsitektur Inception-V3 pada gambar 3.



Gambar 3. Arsitektur Inception-V3

c. VGG16

VGG16 merupakan salah satu arsitektur *Convolutional Neural Network* (CNN) yang populer berkat kesederhanaannya dalam desain dan pemakaian kernel berukuran kecil. Arsitektur ini diperkenalkan pada tahun 2014 dan menggunakan kernel 3×3 , dipadukan dengan Max Pooling dan stride sebesar 1 dalam proses ekstraksi fitur. Di bagian akhirnya, VGG16 memiliki tiga fully connected layer. Penggunaan kernel kecil memungkinkan penambahan kedalaman jaringan tanpa meningkatkan kompleksitas secara signifikan, sehingga dapat meningkatkan akurasi dibandingkan dengan beberapa arsitektur lain. Setiap lapisan konvolusi juga dilengkapi dengan fungsi aktivasi ReLU untuk membantu mengurangi risiko *overfitting* [12]. Berikut adalah gambar 5 menunjukkan ilustrasi arsitektur VGG16.



Gambar 4. Arsitektur VGG16

d. Hyperparameter Optimization

Dalam penelitian ini, optimasi hyperparameter akan menggunakan Stochastic Gradient Descent (SGD). SGD adalah salah satu algoritma paling umum untuk mengoptimalkan jaringan saraf. Metode gradient descent bekerja dengan meminimalkan fungsi objektif $J(\theta)$ dengan parameter model $\theta \in \mathbb{R}$ dengan memperbarui nilai parameter ke arah yang berlawanan dari gradien $\nabla J(\theta)$ terhadap parameter tersebut [29]. Optimizer ini dipilih karena bersifat sederhana dan sangat efisien untuk klasifikasi linier. Selain itu, implementasi SGD relatif lebih mudah dibanding algoritma optimasi lainnya [30].



2.6 Evaluasi Model

Tahap evaluasi digunakan untuk menilai sejauh mana ketepatan model yang telah dibangun. Pengukuran akurasi dilakukan menggunakan confusion matrix dengan menghitung metrik seperti akurasi, presisi, *recall*, dan *F1-score*.

a. Akurasi

Digunakan untuk menghitung rasio kelas prediksi yang benar dengan jumlah total sampel dievaluasi.

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

Keterangan:

TP (True Positive): Jumlah data positif yang diprediksi dengan benar oleh model.

TN (True Negative): Jumlah data negatif yang diprediksi dengan benar oleh model.

FP (False Positive): Jumlah data negatif yang salah diprediksi sebagai positif oleh model.

FN (False Negative): Jumlah data positif yang salah diprediksi sebagai negatif oleh model.

b. Presisi

Digunakan untuk menghitung pola positif yang diprediksi dengan benar oleh semua pola prediksi di kelas positif.

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

c. Recall

Digunakan untuk menghitung pecahan dari pola positif yang diklasifikasikan dengan benar.

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

d. F1-score

Menghitung rata-rata harmonik antara tingkat *recall* dan presisi.

$$F1 - Score = 2 \times \frac{Recall \times Precision}{Recall + Precision} \quad (4)$$

Keterangan:

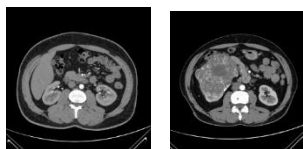
Precision mengukur seberapa banyak dari prediksi positif yang benar ($TP / (TP + FP)$)

Recall mengukur seberapa banyak data positif yang berhasil terdeteksi ($TP / (TP + FN)$)

3. HASIL DAN PEMBAHASAN

3.1 Pengumpulan Data

Dataset yang digunakan dalam penelitian ini berhasil dikumpulkan dari situs Kaggle. Dari total 12.446 citra CT-Scan ginjal yang tersedia, hanya dua kelas yang dipilih sesuai dengan tujuan penelitian, yaitu kelas normal dan kelas tumor. Setelah proses seleksi, jumlah citra yang digunakan terdiri dari 5.077 citra untuk kelas normal dan 2.283 citra untuk kelas tumor, sehingga total data yang digunakan adalah sebanyak 7.360 citra. Pemilihan dua kelas ini dilakukan untuk menyederhanakan proses klasifikasi serta memfokuskan penelitian pada perbedaan antara kondisi ginjal yang sehat dan yang mengalami gangguan berupa tumor. Gambar 5 menampilkan contoh citra dari masing-masing kelas yang digunakan dalam proses pelatihan dan pengujian model.



(a) Normal

(b) Tumor

Gambar 5. Dataset Ginjal Normal dan Tumor

3.2 Data Preprocessing

Proses ini meliputi resize citra menghasilkan dataset yang seragam, dengan ukuran yang sesuai kebutuhan masing-masing model:

- Untuk EfficientNet-B0, citra diresize menjadi ukuran 224x224 piksel, yang merupakan ukuran input standar untuk arsitektur ini.
- Untuk VGG16, citra juga diresize menjadi ukuran 224x224 piksel, karena VGG16 menerima input dengan ukuran yang sama.
- Untuk Inception-V3, citra diresize menjadi ukuran 299x299 piksel, sesuai dengan ukuran input yang dibutuhkan oleh arsitektur ini.

Proses cropping dan resizing ini bertujuan sebagai standar ukuran input citra yang akan digunakan untuk membangun model dari ketiga arsitektur



3.3 Pembagian Dataset

Tabel 1 memperlihatkan pembagian dataset yang digunakan beserta arsitektur model deep learning yang diterapkan dalam eksperimen penelitian ini untuk menguji performa klasifikasi citra CT-Scan ginjal.

Tabel 1. Pembagian Dataset

No.	Rasio Dataset	Arsitektur Model
1.	80:20	EfficientNet-B0
2.	80:20	EfficientNet-B0
3.	80:20	InceptionV3
4.	80:20	InceptionV3
5.	80:20	VGG16
6.	80:20	VGG16

3.4 Implementasi Model

3.4.1 Skenario Eksperimen

Proses eksperimen dalam penelitian ini dilakukan menggunakan platform Google Colaboratory (Google Colab) dengan bahasa pemrograman Python. Beberapa pustaka yang digunakan antara lain TensorFlow, Keras, Matplotlib, Scikit-learn, Seaborn. Total eksperimen yang dilakukan sebanyak 6 eksperimen. Berikut ini terlihat pada tabel 2 skenario eksperimen pada penelitian ini.

Tabel 2. Skenario Eksperimen

No	Arsitektur	Preprocessing Function	Ukuran Input	Learning Rate
1	EfficientNet-B0	efficientnet_preprocess_input	224 x 224	0.001
2	EfficientNet-B0	efficientnet_preprocess_input	224 x 224	0.0001
3	Inception-V3	inception_preprocess_input	224 x 224	0.001
4	Inception-V3	inception_preprocess_input	224 x 224	0.0001
5	VGG16	vgg16_preprocess_input	299 x 299	0.001
6	VGG16	vgg16_preprocess_input	299 x 299	0.0001

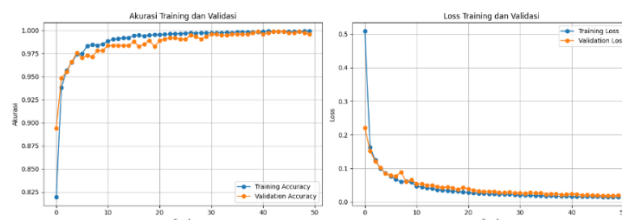
3.4.2 Pelatihan Model

Setelah menentukan skenario eksperimen, setiap model dilatih menggunakan kombinasi *preprocessing function*, ukuran input, dan learning rate yang telah ditetapkan pada Tabel 1. Proses pelatihan dilakukan selama 50 epoch dengan batch size 32, menggunakan fungsi *loss binary_crossentropy* dan fungsi aktivasi sigmoid pada output layer karena klasifikasi bersifat biner (normal dan tumor). Berikut hasil dari pelatihan model pada penelitian ini pada tabel 3.

Tabel 3. Hasil Pelatihan Model

No.	Rasio Dataset	Arsitektur Model	Learning Rate	Training Accuracy	Training Validation Accuracy
1.	80:20	EfficientNet-B0	0.001	0.9725	0.9755
2.	80:20	EfficientNet-B0	0.0001	0.8946	0.8993
3.	80:20	InceptionV3	0.001	0.9363	0.9823
4.	80:20	InceptionV3	0.0001	0.9828	0.9823
5.	80:20	VGG16	0.001	0.9992	0.9986
6.	80:20	VGG16	0.0001	0.9792	0.9673

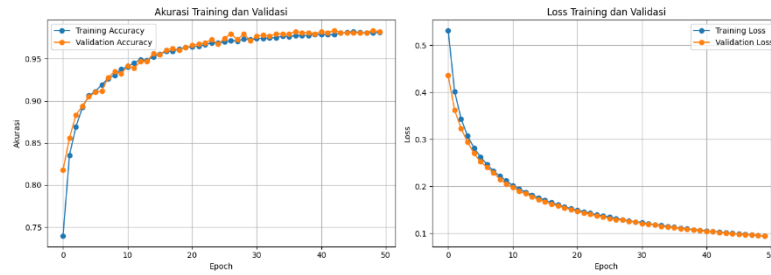
Tabel 3 menunjukkan hasil percobaan pelatihan model yang dilakukan terhadap tiga arsitektur EfficientNet-B0, InceptionV3, dan VGG16 dengan dua nilai learning rate yang berbeda, yaitu 0.001 dan 0.0001. Seluruh eksperimen menggunakan rasio pembagian data 80:20 dan dijalankan hingga 50 epoch. Dari hasil pelatihan, VGG16 dengan learning rate 0,001 meraih validasi akurasi tertinggi pada eksperimen ke-5, diikuti InceptionV3 dengan learning rate 0,0001 yang mencapai puncak pada eksperimen ke-4, serta EfficientNet-B0 dengan learning rate 0,001 yang memperoleh validasi akurasi terbaik pada eksperimen pertama.



Gambar 6. Grafik Pelatihan Model Arsitektur VGG16

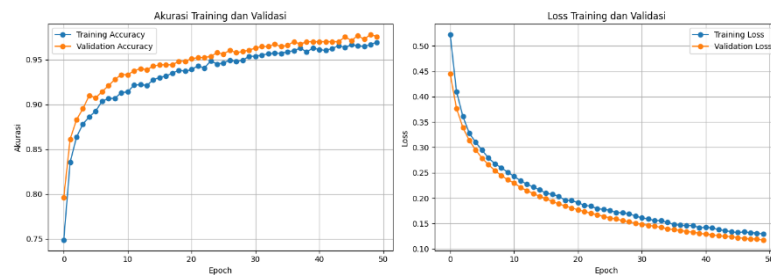


Gambar 6 menunjukkan grafik pelatihan dari model arsitektur VGG16 dengan learning rate 0.001 menghasilkan *training accuracy* tertinggi 0.9992 dan *validation accuracy* tertinggi sebesar 0.9986.



Gambar 7. Grafik Pelatihan Model Arsitektur Inception-V3

Gambar 7 menunjukkan grafik hasil pelatihan dari model arsitektur InceptionV3 dengan learning rate 0.0001 menghasilkan *training accuracy* tertinggi 0.9828 dan *validation accuracy* sebesar 0.9823.



Gambar 8. Grafik Pelatihan Model Arsitektur EfficientNet-B0

Gambar 8 menunjukkan grafik hasil pelatihan dari model arsitektur EfficientNet-B0 dengan denagn learning rate 0.001 menghasilkan *training accuracy* tertinggi 0.9725 dan *validation accuracy* 0.9755. Secara keseluruhan, arsitektur VGG16 menunjukkan performa paling konsisten dan unggul dibandingkan arsitektur lainnya pada skenario yang diuji dan dari grafik pelatihan pun dapat dilihat bahwa model tidak mengalami *overfitting* dan *underfitting*.

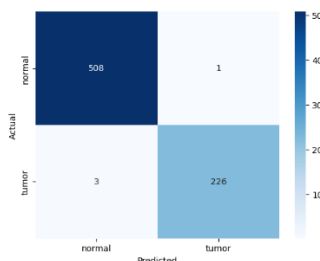
3.4.3 Hasil Pengujian Model

Setelah proses pelatihan model menggunakan data latih dan validasi diselesaikan, tahap berikutnya adalah pengujian dengan menggunakan dataset uji. Pengujian ini mencakup semua arsitektur yang telah dilatih sebelumnya, yaitu EfficientNet-B0, Inception-V3, dan VGG16. Evaluasi dilakukan dengan memanfaatkan confusion matrix untuk menghitung metrik kinerja seperti akurasi, presisi, recall, dan F1-score, serta untuk melihat hasil klasifikasi pada tiap kelas, yaitu normal dan tumor.

Tabel 4. Hasil Pengujian Model

No.	Rasio Dataset	Arsitektur Model	Learning Rate	Accuracy	Precision	Recall	F1-score
1.	80:20	EfficientNet-B0	0.001	97.56%	97.56%	97.56%	97.55%
2.	80:20	EfficientNet-B0	0.0001	89.84%	89.92%	89.84%	89.53%
3.	80:20	InceptionV3	0.001	98.24%	98.24%	98.24%	98.23%
4.	80:20	InceptionV3	0.0001	97.56%	97.56%	97.56%	97.56%
5.	80:20	VGG16	0.001	99.46%	99.46%	99.46%	99.46%
6.	80:20	VGG16	0.0001	95.66%	95.70%	95.66%	95.61%

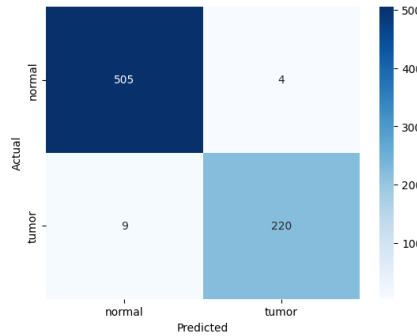
Berdasarkan hasil pengujian pada Tabel 4, VGG16 (learning rate 0,001) mencatat performa tertinggi pada eksperimen ke- 5. selanjutnya InceptionV3 terbaik pada eksperimen ke-3, dan EfficientNet-B0 unggul pada eksperimen ke-1.



Gambar 9. Confusion matrix Model Arsitektur VGG16

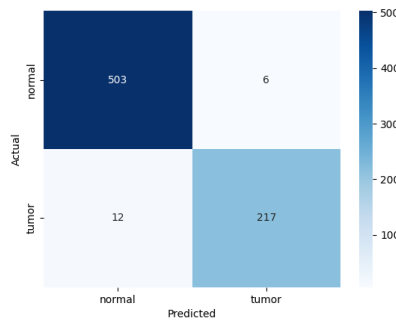


Gambar 9 merupakan Confusion matrix pada eksperimen ke-5 yaitu pada model arsitektur VGG16 yang memiliki *accuracy* 99,46%, *precision* 99,46%, *recall* 99,46%, dan *F1-score* 99,46%.



Gambar 10. Confusion matrix Model Arsitektur Inception-V3

Gambar 10 menampilkan confusion matrix untuk eksperimen ketiga pada arsitektur Inception-V3, dengan nilai *accuracy* 98,24%, *precision* 98,24%, *recall* 98,24%, dan *F1-score* 98,23%.



Gambar 11. Confusion matrix Model Arsitektur EfficientNet-B0

Gambar 11 menampilkan confusion matrix dari eksperimen pertama pada arsitektur EfficientNet-B0, dengan *accuracy* 97,56%, *precision* 97,56%, *recall* 97,56%, dan *F1-score* 97,55%.

Model CT-Kidney terbagi menjadi empat kategori : Kista, Normal, Batu, dan Tumor. Deteksi untuk kelas Normal dan Tumor berjalan dengan sangat baik: dari 508 data Normal dan 228 data Tumor, tidak ada kasus yang terlewat atau salah klasifikasi, sehingga *precision*, *recall*, dan *F1-score* untuk kedua kelas ini masing-masing 1,00, dengan akurasi keseluruhan 100% [21]. Sebaliknya, model VGG16 (lr 0,001) penelitian ini digunakan untuk klasifikasi biner Normal dan Tumor mencapai akurasi 99,46 % (734/738). Pada kelas Normal didapat *precision* sekitar 99,4 % dan *recall* sekitar 99,8 % (F1 sekitar 99,6 %), sedangkan untuk kelas Tumor *precision* sekitar 99,6 % dan *recall* sekitar 98,7 % (F1 sekitar 99,1 %). Dari 509 sampel normal, satu di antaranya keliru diprediksi sebagai tumor (false positive), sementara dari 229 sampel tumor, tiga sampel salah terklasifikasi sebagai normal (false negative). Secara klinis, meskipun *precision* Tumor mencapai 0,9956 dan *recall* Tumor 0,9870 (*F1-score* 0,9913), kemungkinan terlewatnya tiga kasus tumor dapat menghambat penanganan.

4. KESIMPULAN

Penelitian ini membahas klasifikasi citra CT-Scan ginjal untuk mendeteksi keberadaan tumor ginjal dengan membandingkan performa tiga arsitektur yaitu EfficientNet-B0, InceptionV3, dan VGG16. Seluruh proses pelatihan dan pengujian dilakukan menggunakan dataset yang di peroleh dari website Kaggle, dengan rasio pembagian data sebesar 80:20, serta menggunakan dua nilai learning rate yaitu 0.001 dan 0.0001. Evaluasi dilakukan berdasarkan metrik *accuracy*, *precision*, *recall*, dan *F1-score* terhadap hasil pengujian model menggunakan data uji. Berdasarkan hasil evaluasi, model VGG16 dengan learning rate 0.001 menunjukkan performa terbaik dengan nilai *accuracy* 99,46%, *precision* 99,46%, *recall* 99,46%, dan *F1-score* 99,46%. InceptionV3 dengan learning rate 0,001 mencatat akurasi, presisi, dan *recall* sebesar 98,24%, serta *F1-score* 98,23%, sedangkan EfficientNet-B0 pada learning rate yang sama mencapai akurasi 97,56%, *precision* 97,56%, *recall* 97,56%, dan *F1-score* 97,55%. Secara keseluruhan, VGG16 terbukti paling unggul dalam mendeteksi tumor ginjal pada citra CT-Scan dibanding dua arsitektur lainnya berdasarkan akurasi dan metrik evaluasi yang diuji. Untuk penelitian selanjutnya dilakukan pengujian terhadap arsitektur CNN lainnya yang lebih kompleks atau terbaru, serta mengeksplorasi teknik augmentasi data yang lebih beragam untuk meningkatkan generalisasi model. Selain itu, dapat juga dipertimbangkan penggunaan teknik transfer learning lanjutan atau fine-tuning yang lebih mendalam, serta evaluasi menggunakan dataset yang lebih besar dan beragam agar hasil penelitian lebih representatif terhadap kondisi nyata.



REFERENCES

- [1] N. Syuryani, E. Arman, and G. E. Putri, "Perbedaan Kadar Ureum Sebelum Dan Sesudah Hemodialisa Pada Penderita Gagal Ginjal Kronik," *Jurnal Kesehatan Saintika Meditory*, vol. 4, no. 2, p. 117, Nov. 2021, doi: 10.30633/jsm.v4i2.1292.
- [2] A. S. Lestari and H. Harun, "Pemeriksaan Penunjang Dalam Mendiagnosis Tumor Ginjal," *Jurnal Medical Profession (Medpro)*, vol. 1, no. 2, pp. 112–117, 2019.
- [3] M. Gharaibeh *et al.*, "Radiology Imaging Scans for Early Diagnosis of Kidney Tumors: A Review of Data Analytics-Based *Machine learning* and *Deep learning* Approaches," *Big Data and Cognitive Computing*, vol. 6, no. 1, Mar. 2022, doi: 10.3390/bdcc6010029.
- [4] D. Alzu'bi *et al.*, "Kidney Tumor Detection and Classification Based on *Deep learning* Approaches: A New Dataset in CT Scans," *J Healthc Eng*, vol. 2022, 2022, doi: 10.1155/2022/3861161.
- [5] F. Yanto, N. Jannata, L. Handayani, and E. P. Cynthia, "Pengaruh Contrast Limited Adaptive Histogram Equalization dalam Klasifikasi CT-Scan Tumor Ginjal menggunakan Deep Learning," *INOVTEK Polbeng - Seri Informatika*, vol. 9, no. 1, Jun. 2024, doi: 10.35314/isi.v9i1.4235.
- [6] F. Yanto, M. I. Hatta, I. Afrianty, and L. Afriyanti, "Pengaruh Image Enhancement Contrast Stretching dalam Klasifikasi CT-Scan Tumor Ginjal menggunakan Deep Learning," *INOVTEK Polbeng - Seri Informatika*, vol. 9, no. 1, Jun. 2024, doi: 10.35314/isi.v9i1.4233.
- [7] J. Kers *et al.*, "Deep learning-based classification of kidney transplant pathology: a retrospective, multicentre, proof-of-concept study," *Lancet Digit Health*, vol. 4, no. 1, pp. e18–e26, Jan. 2022, doi: 10.1016/S2589-7500(21)00211-9.
- [8] W. R. Perdani, R. Magdalena, and N. K. Caecar Pratiwi, "Deep learning untuk Klasifikasi Glaukoma dengan menggunakan Arsitektur EfficientNet," *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika*, vol. 10, no. 2, p. 322, Apr. 2022, doi: 10.26760/elkomika.v10i2.322.
- [9] N. A. Sundari, R. Magladena, and S. Saidah, "Klasifikasi Jenis Kulit Wajah Menggunakan Metode Covolutional Neural Network (CNN) Efficientnet-B0 Skin Classification System Using Convolutional Neural Network (CNN) EfficientNet-B0," *E-Proceeding of Engineering*, vol. 8, no. 0, p. 6, 2022.
- [10] A. N. Fajrina, Z. H. Pradana, S. I. Purnama, and S. Romadhona, "Penerapan Arsitektur EfficientNet-B0 Pada Klasifikasi Leukimia Tipe Acute Lymphoblastik Leukimia," *Jurnal Riset Rekayasa Elektro*, vol. 6, no. 1, p. 59, Jun. 2024, doi: 10.30595/jrre.v6i1.22090.
- [11] E. Reihardian, "Klasifikasi Ekspresi Wajah Manusia Menggunakan Metode CNN Dengan Arsitektur Visual Geometry Group (VGG-Net)," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 8, no. 8, Aug. 2024, [Online]. Available: <https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/14081>
- [12] H. D. Hekmatyar, W. A. Saputra, and C. Ramdani, "Klasifikasi Pneumonia Dengan Deep Learning Faster Region Convolutional Neural Network Arsitektur VGG16 dan ResNet50," *InComTech : Jurnal Telekomunikasi dan Komputer*, vol. 12, no. 3, p. 186, Dec. 2022, doi: 10.22441/incomtech.v12i3.15112.
- [13] S. Y. Prasetyo, "SARS-CoV-2 Detection from Lung CT-Scan Images Using Fine Tuning Concept on Deep-CNN Pretrained Model," *CESS (Journal of Computer Engineering, System and Science)*, vol. 8, no. 1, p. 101, Jan. 2023, doi: 10.24114/cess.v8i1.40897.
- [14] I. Idawati, D. P. Rini, A. Primanita, and T. Saputra, "Klasifikasi Kanker Payudara Menggunakan Metode Convolutional Neural Network (CNN) dengan Arsitektur VGG-16," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 5, no. 3, p. 529, Apr. 2024, doi: 10.30865/json.v5i3.7553.
- [15] Rizki Firdaus Mulya, Ema Utami, and Dhani Ariatmanto, "Classification of Acute Lymphoblastic Leukemia based on White Blood Cell Images using InceptionV3 Model," *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, vol. 7, no. 4, pp. 947–952, Aug. 2023, doi: 10.29207/resti.v7i4.5182.
- [16] M. A. Amrullah and M. I. Irawan, "Implementasi Jaringan Saraf Konvolusional dengan Inception-V3 untuk Deteksi Katarak Menggunakan Gambar Digital Funduskopi," *Jurnal Sains dan Seni ITS*, vol. 12, no. 1, May 2023, doi: 10.12962/j23373520.v12i1.106807.
- [17] A. Nurdin, D. S. Yudha Kartika, and A. R. Efrat Najaf, "Klasifikasi Penyakit Daun Tomat Dengan Metode Convolutional Neural Network Menggunakan Arsitektur Inception-V3," *JURNAL ILMIAH INFORMATIKA*, vol. 12, no. 02, pp. 114–119, Sep. 2024, doi: 10.33884/jif.v12i02.9162.
- [18] N. W. Wulansari and M. Muslih, "Implementasi Metode Convolutional Neural Network Untuk Deteksi Kematangan Buah Pisang Menggunakan Inception V3," *JOINS (Journal of Information System)*, vol. 8, no. 2, pp. 147–155, Nov. 2023, doi: 10.33633/joins.v8i2.9074.
- [19] A. Salam, F. Yanto, S. Agustian, and S. Ramadhani, "Perbandingan Klasifikasi Citra CT-Scan Kanker Paru-Paru Menggunakan Contrast Stretching Pada CNN dengan EfficientNet-B0," *Media Online*, vol. 4, no. 3, pp. 1341–1351, 2023, doi: 10.30865/klik.v4i3.1448.
- [20] D. Abdillah Salafy, F. Yanto, S. Agustian, and F. Insani, "Perbandingan Klasifikasi Citra CT-Scan Kanker Paru-Paru Menggunakan Image Enhancement CLAHE Pada EfficientNet-B0," *Media Online*, vol. 4, no. 3, pp. 1651–1659, 2023, doi: 10.30865/klik.v4i3.1514.
- [21] A. Elmaddah, "ct_kidney," *Kaggle*. Accessed: Jun. 20, 2025. [Online]. Available: <https://www.kaggle.com/code/abdelrahmanelmaddah/ct-kidney>
- [22] Z.-P. Jiang, Y.-Y. Liu, Z.-E. Shao, and K.-W. Huang, "An Improved VGG16 Model for Pneumonia Image Classification," *Applied Sciences*, vol. 11, no. 23, p. 11185, Nov. 2021, doi: 10.3390/app112311185.
- [23] M. A. S. Al Husaini, M. H. Habaebi, T. S. Gunawan, M. R. Islam, E. A. A. Elsheikh, and F. M. Suliman, "Thermal-based early breast cancer detection using inception V3, inception V4 and modified inception MV4," *Neural Comput Appl*, vol. 34, no. 1, pp. 333–348, Jan. 2022, doi: 10.1007/s00521-021-06372-1.
- [24] S. Abd El-Ghany, M. A. Mahmood, and A. A. Abd El-Aziz, "Adaptive Dynamic Learning Rate Optimization Technique for Colorectal Cancer Diagnosis Based on Histopathological Image Using EfficientNet-B0 Deep Learning Model," *Electronics (Basel)*, vol. 13, no. 16, p. 3126, Aug. 2024, doi: 10.3390/electronics13163126.



- [25] B. Vrigazova, "The Proportion for Splitting Data into *Training* and Test Set for the Bootstrap in Classification Problems," *Business Systems Research*, vol. 12, no. 1, pp. 228–242, May 2021, doi: 10.2478/bsrj-2021-0015.
- [26] N. F. Fahrudin, K. R. Putra, S. Umaroh, And G. B. Lautan, "Influence of Data Scaling and Train/Test Split Ratios on LightGBM Efficacy for Obesity Rate Prediction," *MIND Journal*, vol. 9, no. 2, pp. 220–234, Dec. 2024, doi: 10.26760/mindjournal.v9i2.220-234.
- [27] H. Almirza, S. Sanjaya, L. Handayani, and F. Syafria, "Klasifikasi Daging Sapi dan Daging Babi Menggunakan Convolutional Neural Network EfficientNet-B0 dengan Augmentasi Citra," *Media Online*, vol. 3, no. 6, pp. 1013–1021, 2023, doi: 10.30865/klik.v3i6.910.
- [28] Rismiyati and Luthfiarta Ardytha, "VGG16 Transfer Learning Architecture for Salak Fruit Quality Classification," *Jurnal Informatika dan Teknologi Informasi*, vol. 18, no. 1, pp. 37–48, 2021, doi: 10.31515/telematika.v18i1.4025.
- [29] F. Masrurroh, B. Surarso, and B. Warsito, "Perbandingan Kinerja Inception- Resnetv2, Xception, Inception-v3, dan Resnet50 pada Gambar Bentuk Wajah," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 10, no. 1, pp. 11–20, Feb. 2023, doi: 10.25126/jtiik.2023104941.
- [30] A. W. Setiawan, "Perbandingan Arsitektur Convolutional Neural Network Pada Klasifikasi Pneumonia, COVID-19, Lung Opacity, dan Normal Menggunakan Citra Sinar-X Thoraks," *Jurnal Teknologi Informasi dan Ilmu Komputer*, vol. 9, no. 7, pp. 1563–1570, Dec. 2022, doi: 10.25126/jtiik.2022976742.