



Tersedia Online : <http://e-journals.unmul.ac.id/>

ADOPSI TEKNOLOGI DAN SISTEM INFORMASI (ATASI)

Alamat Jurnal : <http://e-journals2.unmul.ac.id/index.php/atasi/index>



Klasifikasi Batu Ginjal Berbasis MobileNetv2 Graph Neural Network dengan Perbandingan Fungsi Aktivasi

Akhmad Irsyad ^{1)*}, Muhammad Rivani ²⁾, Muhammad Fawaz Saputra ³⁾, Muhammad Zulfariansyah ⁴⁾, Muhammad Yusril Arjulio Prayitno ⁵⁾, M. Ari Prayogo⁶⁾

^{1,2,3,4,5)} Sistem Informasi, Universitas Mulawarman

⁶⁾ Pendidikan Ilmu Komputer, Universitas Mulawarman

E-Mail : akhmadirsyad@ft.unmul.ac.id ¹⁾; rivani@unmul.ac.id ²⁾; muhammadfawaz@unmul.ac.id ³⁾; zulfariansyah@unmul.ac.id ⁴⁾; myusril785@gmail.com ⁵⁾; m.ariprayogo@unmul.ac.id ⁶⁾

ARTICLE INFO

Article history:

Received : 19/07/2026

Revised : 26/08/2026

Accepted : 27/08/2026

Available online : 31/08/2026

Keywords:

kidney stone

MobileNetV2

graph neural network

activation function

CT image classification

ABSTRACT

Kidney stones are a common urological disorder and can cause serious complications if not detected properly. Manual interpretation of computed tomography (CT) images still faces challenges due to small stone size, low contrast, and image variation between patients. This study proposes a kidney stone classification model on CT images that combines MobileNetV2 as a feature extraction backbone with a GraphConvLayer-based graph module to learn the relationships between features. Images undergo preprocessing and data augmentation before being classified using GlobalAveragePooling2D, dense layers, and a classifier head. This study also compares two activation functions, ReLU and Swish, to determine their effects on model performance. The test results show that ReLU provides the best performance with 96.59% accuracy, 96.94% precision, 96.59% recall, and 96.67% F1 score, while Swish achieves the highest AUC-ROC of 99.74%. These findings confirm that the choice of activation function has a significant effect on the performance of the MobileNetV2-Graph architecture in kidney stone classification, with ReLU being the most stable and effective configuration.

ABSTRAK

Batu ginjal merupakan gangguan urologi yang umum dan dapat menimbulkan komplikasi serius apabila tidak terdeteksi secara tepat. Interpretasi citra computed tomography (CT) secara manual masih menghadapi kendala akibat ukuran batu yang kecil, kontras rendah, serta variasi citra antar pasien. Penelitian ini mengusulkan model klasifikasi batu ginjal pada citra CT yang menggabungkan MobileNetV2 sebagai backbone ekstraksi fitur dengan graph module berbasis GraphConvLayer untuk mempelajari relasi antarfitur. Citra melalui tahap *preprocessing* dan augmentasi data sebelum diklasifikasikan melalui GlobalAveragePooling2D, dense layer, dan classifier head. Penelitian ini juga membandingkan dua fungsi aktivasi, yaitu ReLU dan Swish, guna mengetahui pengaruhnya terhadap kinerja model. Hasil pengujian menunjukkan bahwa ReLU memberikan performa terbaik dengan akurasi 96,59%, presisi 96,94%, recall 96,59%, dan skor F1 96,67%, sedangkan Swish mencapai AUC-ROC tertinggi sebesar 99,74%. Temuan ini menegaskan bahwa pemilihan fungsi aktivasi berpengaruh signifikan terhadap kinerja arsitektur MobileNetV2-Graph dalam klasifikasi batu ginjal, dengan ReLU sebagai konfigurasi paling stabil dan efektif.

Kata Kunci :

batu ginjal

MobileNetV2

graph neural network

fungsi aktivasi

klasifikasi citra CT

2026 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

1. PENDAHULUAN

Batu ginjal merupakan salah satu gangguan urologi yang umum terjadi dan dapat menimbulkan nyeri akut, obstruksi saluran kemih, infeksi, serta penurunan fungsi ginjal apabila tidak terdeteksi secara tepat (Akram et al., 2024; Fontenelle & Sarti, 2019). Menurut data epidemiologi, prevalensi batu ginjal terus meningkat dalam dua dekade terakhir dan menjadi salah satu penyebab utama kunjungan pasien pada layanan urologi di berbagai negara, sehingga deteksi dini menjadi krusial untuk mencegah komplikasi lanjutan. Dalam praktik klinis, computed

*) Corresponding Author

<https://doi.org/10.30872/atasi.v5i2.5017>

2026 Adopsi Teknologi dan Sistem Informasi (ATASI) with CC BY SA license.

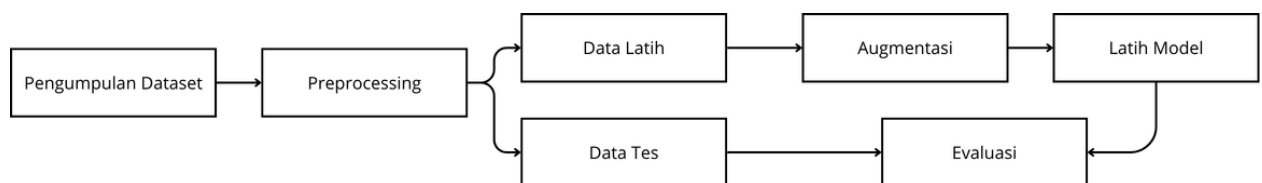
tomography (CT) menjadi modalitas pencitraan utama untuk mendeteksi batu ginjal karena mampu menampilkan struktur anatomi dan karakteristik lesi secara detail (Lu et al., 2025; Vazirinejad et al., 2026). Meskipun demikian, interpretasi citra CT secara manual masih menghadapi berbagai kendala, terutama pada kasus dengan ukuran batu yang kecil, kontras yang rendah, kemiripan intensitas dengan jaringan di sekitarnya, dan tingginya variasi tampilan citra antar pasien (Batheja & Summers, 2025; Yang et al., 2026). Kondisi tersebut dapat meningkatkan risiko kesalahan identifikasi, baik dalam bentuk false positive maupun false negative, yang pada akhirnya dapat memengaruhi ketepatan diagnosis dan pengambilan keputusan klinis (Mohammad et al., 2024; Zuo et al., 2026). Oleh karena itu, dibutuhkan sistem bantu diagnosis yang mampu mengklasifikasikan citra CT secara otomatis, konsisten, dan akurat.

Berbagai penelitian terdahulu telah mengembangkan pendekatan berbasis kecerdasan buatan (artificial intelligence), khususnya deep learning, untuk mendukung analisis citra medis secara otomatis dan lebih konsisten. Sejumlah penelitian menunjukkan bahwa pendekatan berbasis AI mampu meningkatkan performa deteksi batu ginjal pada citra CT, baik dalam aspek akurasi klasifikasi maupun efisiensi analisis (Asif et al., 2024; Kumar et al., 2024; Vasudeva et al., 2025). Di antara berbagai arsitektur deep learning, convolutional neural network (CNN) banyak digunakan karena efektif dalam mengekstraksi fitur visual secara hierarkis (Houssein et al., 2024; Zhao et al., 2024). Beberapa penelitian lain memodifikasi arsitektur CNN dengan menambahkan mekanisme attention atau transfer learning untuk meningkatkan sensitivitas deteksi pada objek berukuran kecil (Panchbhai & Lanjewar, 2026). Namun, CNN konvensional pada umumnya memproses fitur secara independen antar-region sehingga masih memiliki keterbatasan dalam memodelkan hubungan antarfitur yang kompleks. Pendekatan berbasis graph neural network (GNN) mulai dikembangkan untuk mengatasi keterbatasan tersebut karena mampu merepresentasikan keterkaitan antarfitur secara lebih eksplisit melalui struktur graf (Chowdhury et al., 2024; Fan et al., 2026). Meskipun potensi integrasi CNN dan graph module telah ditunjukkan pada beberapa domain citra medis lain, penerapannya secara khusus pada klasifikasi batu ginjal, terutama dengan mempertimbangkan pemilihan fungsi aktivasi yang tepat masih belum banyak dieksplorasi. Celah penelitian inilah yang mendasari penelitian ini: (1) minimnya model klasifikasi batu ginjal yang mengakomodasi relasi antarfitur secara eksplisit, dan (2) belum tersedianya evaluasi sistematis mengenai pengaruh fungsi aktivasi terhadap kinerja arsitektur hibrida CNN-graph pada citra medis.

Untuk menjawab celah penelitian tersebut, penelitian ini memberikan kontribusi sebagai berikut: (1) mengusulkan model klasifikasi batu ginjal pada citra CT yang menggabungkan MobileNetV2 sebagai backbone ekstraksi fitur dengan graph module berbasis GraphConvLayer untuk mempelajari relasi antarrepresentasi fitur; (2) melakukan analisis komparatif terhadap dua fungsi aktivasi, yaitu ReLU dan Swish, yang merepresentasikan karakteristik berbeda dalam mengelola nonlinieritas dan aliran gradien pada jaringan saraf (Bhuvanya et al., 2024; Morozov et al., 2024); dan (3) mengidentifikasi konfigurasi aktivasi yang paling sesuai untuk arsitektur MobileNetV2-Graph pada tugas klasifikasi batu ginjal. ReLU dipilih sebagai baseline karena sederhana dan luas digunakan, Swish dipertimbangkan karena memiliki sifat halus dan non-monotonik. Dengan demikian, tujuan penelitian ini adalah menganalisis pengaruh perbedaan fungsi aktivasi terhadap kinerja model berbasis MobileNetV2-Graph dalam klasifikasi batu ginjal, serta mengidentifikasi konfigurasi aktivasi yang paling sesuai untuk tugas tersebut. Sistematika penulisan artikel ini disusun sebagai berikut: Bagian 2 memaparkan dasar teori, material, dan metodologi penelitian; Bagian 3 menyajikan hasil dan pembahasan; dan Bagian 4 memuat kesimpulan serta saran pengembangan untuk penelitian selanjutnya.

2. METODE PENELITIAN

Proses penelitian dapat di lihat pada Gambar 1, dimulai dari pengumpulan dataset, kemudian dilanjutkan dengan preprocessing untuk membersihkan dan menyiapkan data agar layak digunakan. Setelah itu, data dibagi menjadi data latih dan data tes, di mana data latih digunakan untuk proses augmentasi dan pelatihan model, sedangkan data tes digunakan untuk menguji performa model yang telah dibangun. Melalui tahapan ini, model dapat dipelajari secara optimal dan dievaluasi secara objektif berdasarkan kemampuannya dalam memprediksi data yang belum pernah dilihat sebelumnya.



Gambar 1. Tahapan Penelitian

A. Dataset

Dataset yang digunakan dalam penelitian ini merupakan kumpulan citra CT scan ginjal yang diperoleh dari platform Kaggle dan mencakup dua kelas, yaitu ginjal normal dan ginjal dengan batu ginjal (Hridoy, 2025). Berdasarkan sumber yang digunakan, total data sebanyak 2.642 citra, dengan distribusi 1.101 citra untuk kelas normal dan 1.101 citra untuk kelas batu ginjal pada data latih, serta 110 citra untuk masing-masing kelas pada data uji, sehingga dataset ini disusun secara seimbang untuk mendukung proses klasifikasi yang objektif dan terukur.

B. Preprocessing

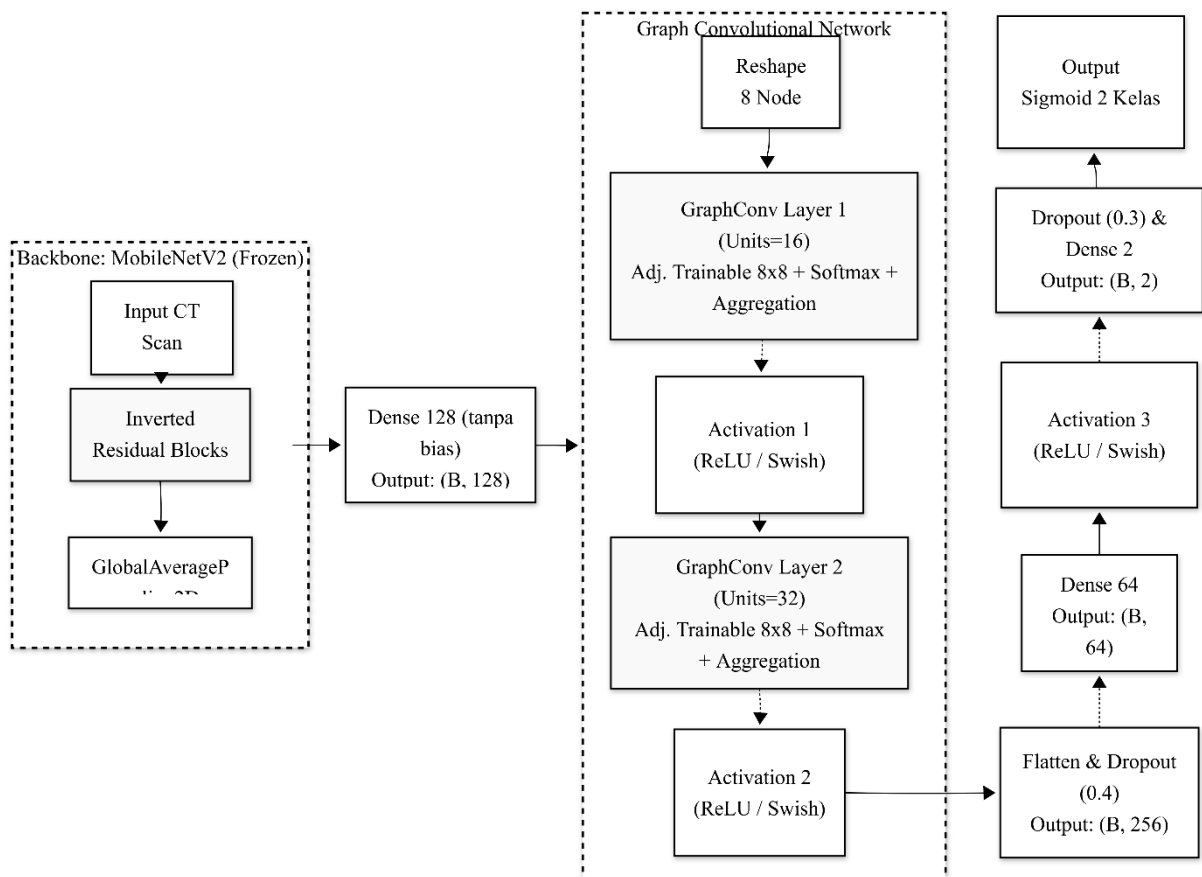
Preprocessing dilakukan untuk menyiapkan citra CT scan ginjal agar layak digunakan dalam proses pelatihan model, meliputi penyeragaman ukuran citra menjadi 224×224 piksel (Khan et al., 2022) serta pembagian dataset ke dalam data latih sebesar 80 persen, data validasi sebesar 10 persen, dan data uji sebesar 10 persen. Pembagian ini bertujuan agar model dapat dipelajari menggunakan data latih, disesuaikan melalui data validasi, dan dievaluasi secara objektif menggunakan data uji yang belum pernah dilihat sebelumnya, sehingga hasil penelitian menjadi lebih akurat dan dapat digeneralisasikan dengan baik (Bichri et al., 2024).

C. Augmentasi

Augmentasi dilakukan pada data latih untuk menambah keragaman dataset melalui transformasi citra seperti rotasi, pergeseran, pembalikan, atau perubahan skala secara acak. Langkah ini bertujuan untuk meningkatkan kemampuan generalisasi model, mengurangi risiko *overfitting*, serta membuat model lebih robust dalam mengenali pola citra ginjal meskipun terdapat variasi pada data baru (Zhang et al., 2024).

D. Proposed Architecture

Pada penelitian ini mengembangkan arsitektur dimulai dari backbone MobileNetV2, kemudian fitur visualnya diringkas melalui GlobalAveragePooling2D dan diproyeksikan ke vektor 128 dimensi sebelum diubah menjadi representasi graf dengan 8 node berukuran 16 fitur per node. Representasi ini selanjutnya diproses oleh dua lapisan Graph Convolution untuk menangkap hubungan antarfitur, diikuti aktivasi ReLU atau Swish, lalu diratakan dengan Flatten, diberi dropout, dan diteruskan ke dense layer hingga menghasilkan output sigmoid untuk klasifikasi normal dan batu ginjal, untuk lebih detail dapat di lihat pada Gambar 2.



Gambar 2. Arsitektur MobileNetV2 Graph Neural Network

MobileNetV2 merupakan arsitektur convolutional neural network yang dirancang untuk efisiensi komputasi tanpa mengorbankan kemampuan ekstraksi fitur secara signifikan. Arsitektur ini banyak digunakan pada citra medis karena memiliki jumlah parameter yang relatif kecil, sehingga cocok untuk pengembangan model yang ringan namun tetap akurat. MobileNetV2 mengandalkan konsep depthwise separable convolution dan inverted residual block untuk mengurangi kompleksitas komputasi (Sandler et al., 2018; Sultana et al., 2025). Pada penelitian ini, MobileNetV2 berperan sebagai backbone utama untuk mengekstraksi fitur visual dari citra CT batu ginjal sebelum fitur tersebut diteruskan ke modul graf. Rumus utama depthwise separable convolution dapat ditulis sebagai Persamaan (1) dan (2).

$$F_{dw}(x, y, k) = \sum_{i=1}^K \sum_{j=1}^K D_{i,j,k} X_{x+i,y+j,k} \dots\dots\dots(1)$$

$$F_{pw}(x, y, m) = \sum_{k=1}^M P_{k,m} F_{dw}(x, y, k) \dots\dots\dots(2)$$

Pada rumus di atas, F_{dw} adalah hasil depthwise convolution, F_{pw} adalah hasil pointwise convolution, X adalah input feature map, D adalah kernel depthwise, dan P adalah kernel pointwise. Depthwise convolution mengekstraksi fitur pada setiap kanal secara terpisah, sedangkan pointwise convolution menggabungkan informasi antar kanal. Kombinasi keduanya membuat MobileNetV2 lebih efisien dibandingkan convolution biasa, sehingga cocok untuk proses ekstraksi fitur pada dataset citra medis.

Graph Module digunakan untuk memodelkan hubungan antarfitur dalam bentuk graf sehingga informasi relasional antar node dapat dipelajari dengan lebih baik (Verma & Singh, 2025). Dalam konteks penelitian ini, fitur hasil ekstraksi MobileNetV2 diubah menjadi representasi graf dan diproses menggunakan Graph Convolutional Layer. Pendekatan ini memungkinkan model menangkap keterkaitan antarbagian fitur yang tidak selalu dapat dipelajari secara optimal oleh CNN biasa. Dengan demikian, Graph Module berfungsi memperkaya representasi fitur sebelum masuk ke classifier head. Rumus umum graph convolution dapat dituliskan sebagai Persamaan (3) dan (4).

$$H^{(l+1)} = \sigma(\hat{A}H^{(l)}W^{(l)}) \dots\dots\dots(3)$$

Dengan

$$\hat{A} = D^{-\frac{1}{2}}(A + I)D^{-\frac{1}{2}} \dots\dots\dots(4)$$

Pada Persamaan (3), $H^{(l+1)}$ adalah fitur pada layer ke-l, $W^{(l)}$ adalah matriks bobot yang dipelajari, A adalah adjacency matrix, I adalah matriks identitas, dan D adalah degree matrix. Persamaan (4) adalah Matriks $\hat{A}$ merupakan adjacency matrix yang telah dinormalisasi agar proses propagasi fitur lebih stabil. Fungsi aktivasi σ digunakan untuk menambahkan nonlinieritas, sehingga graph module mampu menghasilkan representasi fitur yang lebih informatif.

E. Fungsi Aktivasi

ReLU atau Rectified Linear Unit merupakan fungsi aktivasi yang sangat umum digunakan pada jaringan saraf dalam karena sederhana dan efisien. Fungsi ini mempertahankan nilai positif secara langsung, sedangkan nilai negatif diubah menjadi nol. Karakteristik ini membuat ReLU mampu mempercepat proses pelatihan dan mengurangi permasalahan vanishing gradient pada bagian tertentu dari jaringan. Dalam penelitian ini, ReLU digunakan sebagai salah satu fungsi aktivasi pembandingan karena sifatnya yang stabil dan sering memberikan performa yang baik pada model deep learning. Rumus ReLU dapat di lihat pada persamaan (5).

$$f(x) = \max(0, x) \dots\dots\dots(5)$$

Fungsi ini menghasilkan keluaran nol untuk input negatif dan mempertahankan input positif tanpa perubahan. Secara praktis, ReLU membantu model belajar fitur-fitur yang paling relevan dengan cara yang sederhana dan efisien. Karena bentuknya yang tidak kompleks, ReLU sering menjadi baseline dalam berbagai eksperimen klasifikasi citra.

Swish merupakan fungsi aktivasi non-monotonik yang dirancang untuk memberikan fleksibilitas lebih besar dibandingkan ReLU. Aktivasi ini sering digunakan karena mampu mempertahankan informasi pada nilai negatif dalam bentuk terkontrol, sehingga proses pembelajaran menjadi lebih halus. Swish dianggap bermanfaat pada model dengan representasi fitur yang kompleks karena tidak langsung mematikan neuron bernilai negatif seperti ReLU. Dalam penelitian ini, Swish digunakan untuk melihat apakah aktivasi yang lebih smooth dapat meningkatkan performa klasifikasi batu ginjal.

$$f(x) = x \cdot \sigma(x) \dots\dots\dots(6)$$

Dengan

$$\sigma(x) = \frac{1}{1+e^{-x}} \dots\dots\dots(7)$$

Pada rumus persamaan (7), $\sigma(x)$ adalah fungsi sigmoid. Persamaan (6) adalah Swish $f(x)$ mengalikan input dengan nilai sigmoid dari input tersebut, sehingga keluaran yang dihasilkan tetap bergantung pada besar kecilnya masukan. Sifat ini membuat Swish mampu memberikan transisi aktivasi yang lebih halus dibandingkan ReLU.

F. Evaluasi

Evaluasi dilakukan untuk menilai kemampuan model dalam mengklasifikasikan data uji secara tepat dengan menggunakan beberapa metrik, yaitu akurasi, Presisi, recall, dan skor-F1. Akurasi terdapat pada persamaan (8) dan digunakan untuk menunjukkan proporsi prediksi yang benar terhadap seluruh data, Presisi terdapat pada persamaan (9) untuk melihat ketepatan prediksi kelas positif, recall terdapat pada persamaan (10) untuk mengukur kemampuan model dalam menemukan seluruh data positif, sedangkan skor-F1 terdapat pada persamaan (11) sebagai nilai

harmonik dari precision dan recall agar performa model dapat dinilai secara lebih seimbang. Pada rumus tersebut, TP (True Positive) merupakan jumlah data positif yang diprediksi benar sebagai positif, TN (True Negative) adalah jumlah data negatif yang diprediksi benar sebagai negatif, FP (False Positive) adalah jumlah data negatif yang salah diprediksi sebagai positif, dan FN (False Negative) adalah jumlah data positif yang salah diprediksi sebagai negatif. Melalui keempat metrik tersebut, kualitas model dapat diketahui secara lebih komprehensif, baik dari sisi ketepatan maupun kemampuan model dalam mengenali kelas dengan benar (Manjunath et al., 2024).

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

$$P = \frac{TP}{TP+FP} \dots\dots\dots(9)$$

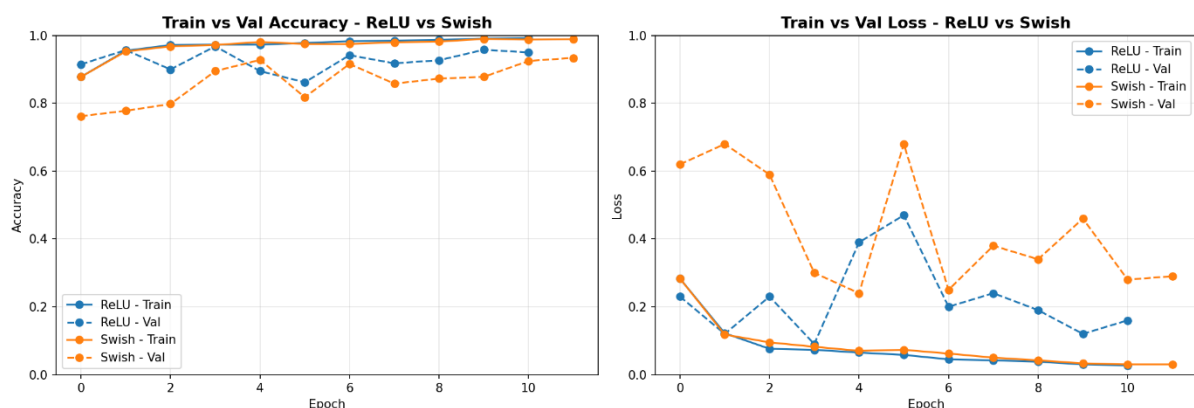
$$R = \frac{TP}{TP+FN} \dots\dots\dots(10)$$

$$F1 = 2 \cdot \frac{P \cdot R}{P+R} \dots\dots\dots(11)$$

3. HASIL

Bab hasil pada penelitian ini memaparkan lingkungan pengembangan, perangkat, serta library yang digunakan dalam proses implementasi model klasifikasi citra CT scan ginjal. Seluruh eksperimen dijalankan menggunakan Google Colab berbasis Jupyter Notebook dengan dukungan GPU dan TensorFlow 2.19.0, sehingga pelatihan model dapat dilakukan secara lebih efisien. Adapun library yang digunakan meliputi NumPy untuk pengolahan numerik, Pandas untuk pengelolaan data, Matplotlib untuk visualisasi, scikit-learn untuk pembagian dataset dan evaluasi, serta TensorFlow/Keras untuk membangun dan melatih model deep learning. Selain itu, proses augmentasi dan pembangkitan data dilakukan menggunakan ImageDataGenerator, sedangkan optimasi pelatihan dibantu oleh callback seperti EarlyStopping, ReduceLROnPlateau, dan ModelCheckpoint. Dengan konfigurasi tersebut, penelitian ini disusun untuk mendukung alur kerja dari pengambilan data, preprocessing, pelatihan model, hingga evaluasi secara terstruktur dan reproduisibel.

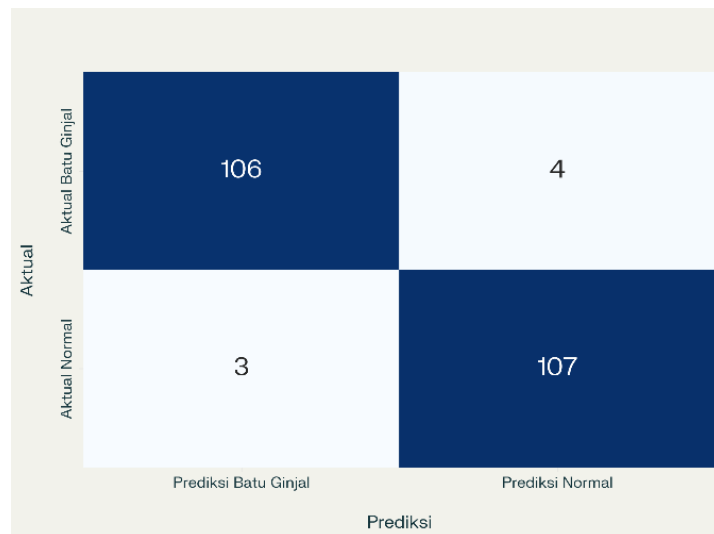
Berdasarkan Gambar 3, model Graph-MobileNet berhasil menghindari overfitting pada pelatihan dengan fungsi aktivasi ReLU dan Swish. Hal ini ditunjukkan oleh kurva akurasi data latih dan data validasi yang sama-sama meningkat hingga mencapai nilai tinggi, serta kurva loss data latih yang menurun secara konsisten sepanjang epoch. Pada ReLU, akurasi validasi cenderung stabil dan mendekati akurasi pelatihan, disertai loss validasi yang berada pada tingkat relatif rendah pada epoch akhir. Hasil ini menandakan bahwa model dapat mempelajari pola pada data latih sekaligus mempertahankan kinerja yang baik pada data validasi. Sementara itu, Swish juga menghasilkan akurasi pelatihan yang sangat tinggi dan akurasi validasi yang baik, meskipun kurva akurasi serta loss validasinya menunjukkan fluktuasi yang lebih besar dibandingkan ReLU. Fluktuasi tersebut menunjukkan bahwa proses generalisasi Swish kurang konsisten, tetapi tidak membentuk pola penurunan akurasi validasi atau peningkatan loss validasi secara berkelanjutan. Oleh karena itu, kedua fungsi aktivasi mampu mendukung konvergensi model tanpa menyebabkan overfitting, dengan ReLU memberikan kestabilan pembelajaran dan generalisasi yang lebih baik daripada Swish.



Gambar 3. Gambar hasil Test Model MobileNetv2 Graph Neural Network dengan aktivasi ReLU

Hasil *confusion matrix* yang ada pada Gambar 4, menunjukkan model klasifikasi dengan ReLU memiliki kinerja yang sangat baik dalam membedakan kelas Batu Ginjal dan Normal pada total 220 data uji. Sebanyak 106 data Batu Ginjal berhasil diprediksi secara benar sebagai Batu Ginjal, 107 data Normal berhasil diprediksi secara benar sebagai Normal, 4 data Normal salah diprediksi sebagai Batu Ginjal, dan 3 data Batu Ginjal salah diprediksi sebagai Normal. Dengan demikian, model memperoleh 213 prediksi benar dari keseluruhan data uji, sehingga menghasilkan nilai akurasi sebesar 96,82%. Pada kelas Batu Ginjal, model memperoleh nilai presisi sebesar 96,36%, *recall* sebesar 97,25%, dan F1-score sebesar 96,80%, yang menunjukkan bahwa model memiliki kemampuan sangat baik dalam mengidentifikasi kasus Batu Ginjal, dengan hanya sebagian kecil kasus positif yang keliru diklasifikasikan sebagai Normal. Sementara itu, pada kelas Normal, model memperoleh nilai presisi sebesar 97,27%, *recall* sebesar 96,40%, dan F1-score sebesar 96,83%, yang mengindikasikan bahwa prediksi terhadap kelas

Normal juga dilakukan secara akurat dan konsisten. Secara keseluruhan, nilai presisi, *recall*, dan F1-score rata-rata mencapai 96,82%, sehingga dapat disimpulkan bahwa model memiliki performa klasifikasi yang tinggi, seimbang, dan andal dalam mendeteksi kedua kelas.



Gambar 4. Gambar hasil Test Model MobileNetv2 Graph Neural Network dengan aktivasi ReLu

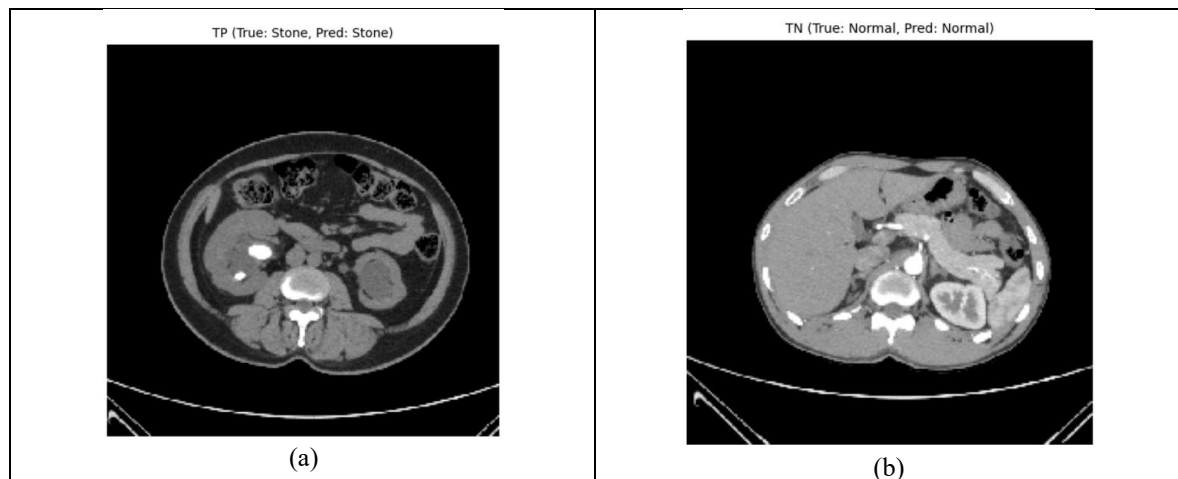
Tabel 1 menunjukkan hasil evaluasi performa model berdasarkan perbandingan dua fungsi aktivasi, yaitu ReLU dan Swish, dengan menggunakan metrik akurasi, presisi, recall, F1-score, dan AUC-ROC. Model dengan aktivasi ReLU menghasilkan akurasi sebesar 96,59%, presisi 96,94%, recall 96,59%, dan F1-score 96,67%, yang menunjukkan bahwa model memiliki kemampuan klasifikasi yang tinggi serta keseimbangan yang baik antara ketepatan prediksi dan kemampuan mendeteksi kelas positif. Sementara itu, model dengan aktivasi Swish memperoleh akurasi 92,72%, presisi 94,57%, recall 92,72%, dan F1-score 93,08%, sehingga kinerjanya masih tergolong baik, namun secara umum berada di bawah ReLU pada sebagian besar metrik utama. Meskipun demikian, Swish menghasilkan AUC-ROC sebesar 99,74%, sedikit lebih tinggi dibandingkan ReLU yang memperoleh 99,67%, yang mengindikasikan bahwa Swish memiliki kemampuan pemisahan kelas yang sangat kuat. Berdasarkan hasil tersebut, ReLU cenderung lebih unggul dalam klasifikasi akhir, sedangkan Swish menunjukkan performa yang kompetitif dalam hal kemampuan diskriminasi model.

aktivasi ReLU menunjukkan performa yang lebih baik dibandingkan Swish pada sebagian besar metrik evaluasi, yaitu akurasi, presisi, recall, dan F1-score. Hasil tersebut menunjukkan bahwa ReLU lebih efektif dalam menghasilkan klasifikasi yang tepat serta mampu menjaga keseimbangan antara kemampuan mendeteksi kelas positif dan meminimalkan kesalahan prediksi. Meskipun Swish memperoleh nilai AUC-ROC yang sedikit lebih tinggi, selisih tersebut tidak memberikan pengaruh yang signifikan terhadap performa klasifikasi secara keseluruhan. Dengan demikian, ReLU dapat dikatakan lebih unggul karena memberikan hasil prediksi yang lebih stabil, konsisten, dan andal pada data uji.

Tabel 1. Hasil Penelitian

Metriks	ReLU	Swish
Akurasi (%)	96,59	92,72
Presisi (%)	96,94	94,57
Recall (%)	96,59	92,72
Skor F1 (%)	96,67	93,08
AUC-ROC (%)	99,67	99,74

Gambar 5 (a) menunjukkan hasil true negative dengan kondisi asli normal dan prediksi model juga normal, sedangkan Gambar 5 (b) menunjukkan true positive dengan kondisi asli batu ginjal dan prediksi model juga batu ginjal. Hasil ini mengindikasikan bahwa model dengan aktivasi ReLU mampu mengenali pola citra ginjal secara tepat pada dua kelas utama, baik saat citra berada dalam kondisi normal maupun ketika terdapat batu ginjal. Secara analitis, keberhasilan pada kedua contoh tersebut mencerminkan bahwa fitur yang dipelajari model cukup representatif untuk membedakan karakteristik visual antara ginjal normal dan ginjal dengan batu, sehingga prediksi yang dihasilkan selaras dengan label sebenarnya. Dengan demikian, contoh prediksi ini mendukung hasil evaluasi sebelumnya bahwa ReLU memberikan performa klasifikasi yang baik dan stabil pada data uji.



Gambar 5. Gambar hasil Test TP dan TN Model MobileNetv2 Graph Neural Network dengan aktivasi ReLu

Model MobileNetV2 Graph Neural Network dengan fungsi aktivasi ReLU berhasil melakukan klasifikasi gagal ginjal namun masih terdapat kemungkinan kesalahan klasifikasi, hal ini disebabkan oleh batu ginjal yang masih belum terlihat jelas ataupun ketika sudut gambar yang diambil kurang baik. Subgambar 6 (a) merupakan kasus False Positive (FP), yaitu citra yang sebenarnya termasuk kelas normal, tetapi diprediksi oleh model sebagai citra batu ginjal. Kesalahan ini menunjukkan bahwa model menangkap karakteristik tertentu pada citra normal yang menyerupai pola visual batu ginjal. Sementara itu, subgambar (b) merupakan kasus False Negative (FN), yaitu citra yang sebenarnya termasuk kelas batu ginjal, tetapi diprediksi sebagai citra normal. Kesalahan tersebut dapat terjadi karena karakteristik batu ginjal pada citra kurang terlihat jelas atau memiliki kemiripan dengan struktur ginjal normal. Kedua contoh ini menunjukkan bahwa meskipun model memiliki kinerja klasifikasi yang tinggi, masih terdapat kemungkinan kesalahan prediksi yang perlu diperhatikan melalui analisis lebih lanjut terhadap kualitas citra, variasi karakteristik objek, dan proses pelatihan model, sebagaimana ditunjukkan pada Gambar 6.



Gambar 6. Gambar hasil Test FP dan FN Model MobileNetv2 Graph Neural Network dengan aktivasi ReLu

Penelitian ini memiliki beberapa keterbatasan. Pertama, eksperimen hanya membandingkan dua fungsi aktivasi, yaitu ReLU dan Swish, sehingga belum memberikan gambaran yang lebih luas mengenai efektivitas aktivasi lain yang mungkin memiliki karakteristik berbeda. Kedua, evaluasi dilakukan pada dataset citra CT scan ginjal dengan skema pembagian data tertentu, sehingga hasil yang diperoleh masih terbatas pada kondisi dataset tersebut dan belum tentu identik ketika diterapkan pada data eksternal. Ketiga, penelitian masih berfokus pada performa klasifikasi akhir tanpa analisis mendalam terhadap pola kesalahan prediksi atau interpretasi area citra yang paling berpengaruh terhadap keputusan model. Oleh karena itu, hasil penelitian ini masih memerlukan pengujian lanjutan untuk memastikan kekuatan generalisasi model.

Penelitian selanjutnya dapat memperluas perbandingan dengan menguji fungsi aktivasi lain agar diketahui apakah terdapat aktivasi yang memberikan performa lebih baik daripada ReLU dan Swish. Selain itu, pengujian pada dataset yang lebih besar dan lebih beragam sangat disarankan untuk meningkatkan validitas hasil dan kemampuan generalisasi model. Studi berikutnya juga dapat menambahkan analisis interpretabilitas seperti visualisasi fitur atau area prediktif agar keputusan model dapat dijelaskan secara lebih transparan. Dengan demikian, penelitian lanjutan diharapkan tidak hanya membandingkan performa numerik, tetapi juga memperkuat pemahaman terhadap karakteristik aktivasi dalam konteks klasifikasi citra medis.

4. KESIMPULAN

Berdasarkan hasil evaluasi pada penelitian ini, model dengan aktivasi ReLU menunjukkan performa yang lebih baik dibandingkan Swish dalam klasifikasi citra CT scan ginjal. ReLU menghasilkan akurasi sebesar 96,59%, presisi 96,94%, recall 96,59%, F1-score 96,67%, dan AUC-ROC 99,67%, sedangkan Swish memperoleh akurasi 92,72%, presisi 94,57%, recall 92,72%, F1-score 93,08%, dan AUC-ROC 99,74%. Hasil tersebut menunjukkan bahwa ReLU lebih unggul pada metrik klasifikasi utama, sehingga lebih efektif dalam membedakan data normal dan batu ginjal pada data uji, meskipun Swish memiliki nilai AUC-ROC yang sedikit lebih tinggi. Dengan demikian, dapat disimpulkan bahwa fungsi aktivasi ReLU memberikan hasil yang lebih stabil dan lebih baik secara keseluruhan untuk model yang digunakan dalam penelitian ini.

5. UCAPAN TERIMA KASIH

Berdasarkan hasil evaluasi pada penelitian ini, model dengan aktivasi ReLU menunjukkan performa yang lebih baik dibandingkan Swish dalam klasifikasi citra CT scan ginjal. ReLU menghasilkan akurasi sebesar 96,59%, presisi 96,94%, recall 96,59%, F1-score 96,67%, dan AUC-ROC 99,67%, sedangkan Swish memperoleh akurasi 92,72%, presisi 94,57%, recall 92,72%, F1-score 93,08%, dan AUC-ROC 99,74%. Hasil tersebut menunjukkan bahwa ReLU lebih unggul pada metrik klasifikasi utama, sehingga lebih efektif dalam membedakan data normal dan batu ginjal pada data uji, meskipun Swish memiliki nilai AUC-ROC yang sedikit lebih tinggi. Dengan demikian, dapat disimpulkan bahwa fungsi aktivasi ReLU memberikan hasil yang lebih stabil dan lebih baik secara keseluruhan untuk model yang digunakan dalam penelitian ini.

6. DAFTAR PUSTAKA

- Akram, M., Jahrreiss, V., Skolarikos, A., Geraghty, R., Tzelves, L., Emilliani, E., Davis, N. F., & Somani, B. K. (2024). Urological Guidelines for Kidney Stones: Overview and Comprehensive Update. *Journal of Clinical Medicine* 2024, Vol. 13, Page 1114, 13(4), 1114. <https://doi.org/10.3390/JCM13041114>
- Asif, S., Zheng, X., & Zhu, Y. (2024). An optimized fusion of deep learning models for kidney stone detection from CT images. *Journal of King Saud University - Computer and Information Sciences*, 36(7), 102130. <https://doi.org/10.1016/J.JKSUCI.2024.102130>
- Batheja, V., & Summers, R. (2025). Progress in fully automated abdominal CT interpretation-an update over the past decade. *Abdominal Radiology* 2025 51:2, 51(2), 1016–1034. <https://doi.org/10.1007/S00261-025-05094-5>
- Bhuvanya, R., Kujani, T., Padmavathy, R., Matheswaran, P., & Punitha, P. (2024). Beyond ReLU: Unlocking Superior Plant Disease Recognition with Swish. <https://doi.org/10.1142/S0219877024410013>, 21(8). <https://doi.org/10.1142/S0219877024410013>
- Bichri, H., Chergui, A., & Hain, M. (2024). Investigating the Impact of Train / Test Split Ratio on the Performance of Pre-Trained Models with Custom Datasets. *IJACSA International Journal of Advanced Computer Science and Applications*, 15(2). www.ijacsa.thesai.org
- Chowdhury, S. T., Kumpati, N. S., & Mukhopadhyay, S. (2024). Mutual Learning for News Classification. *Lecture Notes in Networks and Systems*, 1066 LNNS, 37–54. https://doi.org/10.1007/978-3-031-66428-1_3/SAVE-RESEARCH
- Fan, J., Ma, J., & Gai, H. (2026). MAC2STI: Mamba network with autoregressive clustering for two-stage spatio-temporal imputation. *Complex and Intelligent Systems*, 12(1), 19-. <https://doi.org/10.1007/S40747-025-02130-1/FIGURES/15>
- Fontenelle, L. F., & Sarti, T. D. (2019). Kidney Stones: Treatment and Prevention. *American Family Physician*, 99(8), 490–496. <https://www.aafp.org/afp/2019/0415/p490>
- Houssein, E. H., Abdelkareem, D. A., Hu, G., Hameed, M. A., Ibrahim, I. A., & Younan, M. (2024). An effective multiclass skin cancer classification approach based on deep convolutional neural network. *Cluster Computing*, 27(9), 12799–12819. <https://doi.org/10.1007/S10586-024-04540-1/TABLES/8>
- Hridoy. (2025). *Kidney Stone Dataset*. Kaggle. <https://www.kaggle.com/datasets/harshghadiya/kidneystone>
- Khan, E., Rehman, M. Z. U., Ahmed, F., Alfouzan, F. A., Alzahrani, N. M., & Ahmad, J. (2022). Chest X-ray Classification for the Detection of COVID-19 Using Deep Learning Techniques. *Sensors*, 22(3). <https://doi.org/10.3390/s22031211>
- Kumar, Y., Brar, T. P. S., Kaur, C., & Singh, C. (2024). A Comprehensive Study of Deep Learning Methods for Kidney Tumor, Cyst, and Stone Diagnostics and Detection Using CT Images. *Archives of Computational Methods in Engineering*, 31(7), 4163–4188. <https://doi.org/10.1007/S11831-024-10112-8/METRICS>
- Lu, G., Tian, J., Shi, F., Zhang, D. G., & Wang, D. (2025). Association of cardiovascular-kidney-metabolic syndrome stages with kidney stone prevalence: a population-based analysis of NHANES 2007–2020. *BMJ Open*, 15(5), e096533. <https://doi.org/10.1136/BMJOPEN-2024-096533>
- Manjunath, R. V., Ghanshala, A., & Kwadiki, K. (2024). Deep learning algorithm performance evaluation in detection and classification of liver disease using CT images. *Multimedia Tools and Applications*, 83(1), 2773–2790. <https://doi.org/10.1007/s11042-023-15627-z>
- Mohammad, F., Sharma, G., Khandelwal, K., Arif, A., Vani, M. S., & Sathyanarayana, C. (2024). Development of a Predictive Model for Kidney Stone Formation Using Deep Learning Techniques. *Proceedings of the*

- 2024 International Conference on Artificial Intelligence and Emerging Technology, Global AI Summit 2024, 485–489. <https://doi.org/10.1109/GLOBALAISUMMIT62156.2024.10947936>
- Morozov, A. V., Levkivskiy, V. L., & Plechystyy, D. D. (2024). Activation Functions in Neural Networks: Overview and Comparison. *Scientific Notes of Taurida National V.I. Vernadsky University. Series: Technical Sciences*, 1(3), 144–151. <https://doi.org/10.32782/2663-5941/2024.3.1/22>
- Panchbhai, K. G., & Lanjewar, M. G. (2026). KiCNN: A lightweight CNN model for kidney disease detection using spatial attention mechanisms. *Multimedia Tools and Applications*, 85(3), 277-. <https://doi.org/10.1007/S11042-026-21398-0/METRICS>
- Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4510–4520. <https://doi.org/10.1109/CVPR.2018.00474>
- Sultana, J., Sen, J., Tasnim, N., & Rahman, M. M. (2025). CyStoneNet: Pseudo-Color Enhanced SE-Attention MobileNetV2 for Kidney Stone and Cyst Detection. *2025 28th International Conference on Computer and Information Technology (ICCIT)*, 4723–4728. <https://doi.org/10.1109/ICCIT68739.2025.11491240>
- Vasudeva, N., Dhaka, V. S., & Sinwar, D. (2025). Enhancing kidney stone diagnosis with AI-driven radiographic imaging: a review. *Discover Artificial Intelligence*, 5(1), 200-. <https://doi.org/10.1007/S44163-025-00434-2/TABLES/3>
- Vazirinejad, R., Ahmadiania, H., Heydari Khoormizi, S. M., Jamshidi, M., Khanamani Falahati-pour, S., Mohammadi-Shahrokhi, V., Safarian Neamatabad, M., Mohammadi Nematabad, V., & Abbasi-Dokht-Rafsanjani, M. (2026). Hypertension and incidence of kidney stones: a prospective study based on the data of the Rafsanjan cohort study. *Scientific Reports 2026*. <https://doi.org/10.1038/s41598-026-60737-x>
- Verma, S., & Singh, P. K. (2025). Automated Kidney Stone Detection and Classification utilizing Dynamic Generative Residual Graph Convolutional Neural Networks for Computed Tomography-Based Medical Imaging. *Biomedical Materials and Devices*, 1–18. <https://doi.org/10.1007/S44174-025-00525-9/METRICS>
- Yang, Z., Jiao, B., & Ren, X. (2026). RAC-Net: Interpretable Medical Small Target Segmentation Network with X-ray Radiation Attenuation Characterization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/TPAMI.2026.3705213>
- Zhang, Z., Hou, J., & Lin, Z. (2024). Exploring Integration methods for Image Data Augmentation. *2024 6th International Conference on Internet of Things, Automation and Artificial Intelligence, IoTAAI 2024*, 707–713. <https://doi.org/10.1109/IoTAAI62601.2024.10692583>
- Zhao, X., Wang, L., Zhang, Y., Han, X., Deveci, M., & Parmar, M. (2024). A review of convolutional neural networks in computer vision. *Artificial Intelligence Review*, 57(4), 99-. <https://doi.org/10.1007/S10462-024-10721-6/FIGURES/33>
- Zuo, J., Zhang, Z., Chen, J., Gou, K., Zhou, J., Wen, L., Wei, H., Li, X., Zhan, P., Chen, P., Li, H., Zhao, J., Wang, H., Fu, S., Chen, J., & Wang, J. (2026). The Application of Nanomaterials in Kidney Stone Disease: Emerging Strategies for Early Diagnosis, Targeted Therapy, and Prevention. *International Journal of Nanomedicine*, 21. <https://doi.org/10.2147/IJN.S610906>