

Kidney Stone Disease Diagnosis Using Shifted-Windows Transformer (Swin Transformer)

Diagnosis Penyakit Batu Ginjal Menggunakan *Shifted-Windows Transformer (Swin Transformer)*

Alfian Bima Prastyo, Fetty Tri Anggraeny, Retno Mumpuni

**Prodi Informatika, Fakultas Ilmu Komputer
Universitas Pembangunan Nasional Veteran Jawa Timur**

E-mail: 121081010101@student.upnjatim.ac.id

Abstract - Kidney stones are a prevalent urological condition that, if undiagnosed that can lead to serious complications. Traditional diagnostic methods, such as manual ultrasound interpretation, are error-prone and time consuming, especially in areas with limited access to healthcare professionals. This research proposes the use of the Shifted Windows Transformer (Swin Transformer), a state-of-the-art deep learning model, to improve the classification of kidney stones in ultrasound images. The model is trained on a dataset of 9,396 kidney ultrasound images, categorized into two classes normal kidneys and kidneys with stones, sourced from a publicly available on Mendeley data Kidney dataset. The results demonstrate that the Swin Transformer achieves an impressive accuracy of 99.57%, surpassing others models like Convolutional Neural Networks (CNN) and Vision Transformers (ViT) by efficiently capturing both local and global features in high-resolution images. Practical implications include faster, more accurate diagnoses, particularly in regions lacking specialized radiologists. However, limitations of this model include its dependence on high-quality ultrasound images, which may not always be available in less-resourced settings. Additionally, the model's performance may vary depending on the diversity of the dataset, limiting its generalizability in certain clinical environments. The need for substantial computational resources may also restrict the model's applicability in some healthcare settings. Despite these limitations, the Swin Transformer shows great promise as an automated tool for kidney stone detection, offering potential solutions for early diagnosis in remote and underdeveloped areas.

Keywords: Deep Learning, Image Classification, Kidney Stone, Medical AI, Swin Transformer

Abstrak - Batu ginjal adalah kondisi urologis yang umum terjadi dan, jika tidak didiagnosis, dapat menyebabkan komplikasi serius. Metode diagnostik tradisional, seperti interpretasi ultrasonografi manual, rentan terhadap kesalahan dan memakan waktu, terutama di daerah dengan akses terbatas ke tenaga medis. Studi ini mengusulkan penggunaan *Shifted Windows Transformer (Swin Transformer)*, model *deep learning*, untuk meningkatkan klasifikasi batu ginjal pada citra ultrasonografi. Model ini dilatih pada kumpulan data 9.396 gambar ultrasonografi ginjal, dikategorikan menjadi dua kelas, yaitu ginjal normal dan ginjal dengan batu, yang bersumber dari kumpulan data ginjal yang tersedia untuk umum di Mendeley. Hasilnya menunjukkan bahwa *Swin Transformer* mencapai akurasi yang mengesankan sebesar 99,57%, melampaui model lain seperti CNN dan *Vision Transformers (ViT)* dengan secara efisien menangkap fitur lokal dan global dalam gambar beresolusi tinggi. Implikasi praktisnya termasuk diagnosis yang lebih cepat dan akurat, terutama di wilayah yang kekurangan radiolog. Namun, model ini tergantung pada gambar ultrasonografi berkualitas tinggi, yang mungkin tidak selalu tersedia di lingkungan dengan sumber daya terbatas. Selain itu, kinerja model dapat bervariasi tergantung pada keragaman kumpulan data, membatasi kemampuan generalisasinya di lingkungan klinis tertentu. Kebutuhan akan sumber daya komputasi yang besar juga dapat membatasi penerapan model di beberapa lingkungan perawatan kesehatan. Meskipun ada keterbatasan ini, *Swin Transformer* menunjukkan potensi besar sebagai alat otomatis untuk deteksi batu ginjal, menawarkan solusi potensial untuk diagnosis dini di daerah terpencil dan kurang berkembang.

Kata Kunci: Deep Learning, Klasifikasi Citra, Batu Ginjal, AI Medis, Swin Transformer

1. PENDAHULUAN

Batu ginjal, atau nefrolitiasis, adalah kondisi urologis umum yang dapat menyebabkan komplikasi parah, seperti gagal ginjal, jika tidak didiagnosis dan diobati. Insiden batu ginjal telah

meningkat secara global, dengan dehidrasi, kebiasaan makan yang buruk, dan predisposisi genetik berkontribusi terhadap peningkatan tersebut. Di negara-negara seperti Indonesia, batu ginjal adalah salah satu penyebab utama

penyakit terkait ginjal, terutama di daerah pedesaan di mana akses ke perawatan khusus terbatas. Diagnosis yang tepat waktu sangat penting, karena batu ginjal yang tidak diobati dapat menyebabkan nyeri yang signifikan, penyumbatan saluran kemih, dan kerusakan ginjal jangka panjang [1].

Pencitraan ultrasonografi telah menjadi alat diagnostik penting untuk batu ginjal karena sifatnya yang non-invasif dan tidak adanya radiasi berbahaya. Namun, interpretasi ultrasonografi membutuhkan radiolog yang terampil, sehingga rentan terhadap penilaian subjektif dan kesalahan diagnosis, terutama di daerah dengan tenaga medis yang terbatas. Proses ini juga bisa memakan waktu, semakin menunda diagnosis dan pengobatan. Tantangan-tantangan ini menyoroti kebutuhan akan alat diagnostik otomatis yang dapat membantu mendeteksi batu ginjal dengan cepat dan akurat, terutama di wilayah yang kurang berkembang [2].

Shifted Windows Transformer (Swin Transformer) adalah model pembelajaran mendalam baru yang memperkenalkan struktur hierarkis dan mekanisme jendela bergeser untuk memproses gambar besar secara efisien sambil menangkap pola lokal dan global. Berbeda dengan CNN tradisional dan bahkan model CNN-Transformer hibrida, *Swin Transformer* dirancang untuk meminimalkan kompleksitas komputasi sambil mempertahankan akurasi tinggi dalam memproses gambar kompleks. Model ini telah menunjukkan kinerja unggul dalam beberapa tugas penglihatan komputer, termasuk klasifikasi gambar, segmentasi, dan deteksi objek. Kemampuan *Swin Transformer* untuk menangkap fitur lokal berbutir halus serta informasi kontekstual global menjadikannya kandidat ideal untuk klasifikasi citra medis, terutama untuk citra ultrasonografi beresolusi tinggi yang digunakan dalam deteksi batu ginjal [3].

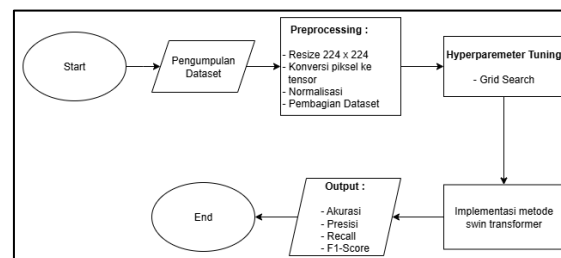
Visi komputer telah muncul sebagai teknologi yang ampuh dalam mengatasi tantangan terkait dengan analisis gambar medis manual. Dengan menggabungkan pemrosesan citra digital, pengenalan pola, dan algoritma AI, sistem penglihatan komputer dapat secara otomatis menganalisis dan menafsirkan citra medis. Teknologi ini tidak hanya meningkatkan kecepatan dan efisiensi analisis gambar, tetapi juga mengurangi kesalahan manusia, memastikan hasil yang lebih konsisten dan andal. Integrasi penglihatan komputer dengan AI telah menjadi sangat penting di daerah dengan akses terbatas ke tenaga medis, di mana model AI dapat membantu memberikan diagnosis yang tepat waktu dan akurat. Studi telah menunjukkan

bahwa penglihatan komputer, dikombinasikan dengan model pembelajaran mesin canggih, dapat secara signifikan meningkatkan kecepatan dan keandalan diagnosis berbagai kondisi medis, termasuk batu ginjal, dengan mengotomatiskan analisis gambar dan meminimalkan kesalahan manusia [4].

Studi ini bertujuan untuk mengevaluasi efektivitas *Swin Transformer* dalam mendeteksi batu ginjal pada gambar ultrasonografi. Dengan memanfaatkan kemampuan model untuk memproses gambar resolusi tinggi secara efisien, studi diarahkan untuk meningkatkan akurasi, keandalan, dan efisiensi deteksi batu ginjal, mengklasifikasikan gambar ultrasonografi menjadi dua kategori, yaitu ginjal normal dan ginjal dengan batu. Dengan mengatasi keterbatasan metode tradisional, *Swin Transformer* dapat menjadi alat yang berharga untuk diagnosis dini, terutama di daerah dengan akses terbatas ke radiolog spesialis. Hasil studi diharapkan dapat berkontribusi pada pengembangan alat diagnostik berbasis AI yang dapat meningkatkan hasil perawatan kesehatan dengan memberikan diagnosis yang lebih akurat dan tepat waktu di daerah terpencil dan kurang berkembang.

2. METODE PELAKSANAAN

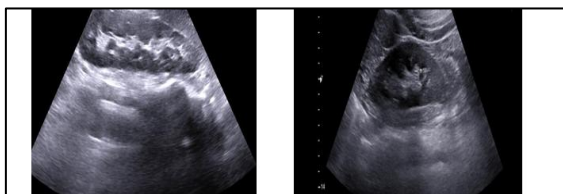
Kerangka penelitian ditunjukkan pada Gambar 1. Dataset (Gambar 2) terdiri dari 9.396 gambar ultrasonografi, dikategorikan menjadi gambar ginjal normal (4.404 gambar) dan gambar batu ginjal (4.992 gambar). Data bersumber dari repositori data Mendeley yang tersedia untuk umum dan dirancang khusus untuk deteksi batu ginjal. Angka-angka pada Gambar 2 menyoroti karakteristik visual yang berbeda antara struktur ginjal normal dan yang terkena batu ginjal, yang sangat penting untuk pelatihan dan evaluasi model [5].



Gambar 1. Alur Penelitian

Pra-pemrosesan data adalah tahap krusial sebelum pemodelan peramalan untuk memastikan data bersih, konsisten, dan sesuai format yang dibutuhkan. Proses ini mencakup *resize* ukuran gambar menjadi 224x224 *pixel* untuk

menstandarisasi ukuran *input* model. Nilai *pixel* dinormalisasi dari rentang asli 0-255 menjadi rentang -1 hingga 1 karena sebagian besar model *deep learning* mengharuskan ukuran gambar konsisten agar dapat diproses dalam *batch*. Selain itu, *resize* juga mengurangi kompleksitas komputasi, membantu menstabilkan pelatihan, mempercepat konvergensi untuk mempercepat proses pelatihan, dan meningkatkan efisiensi. Selain itu, teknik augmentasi data termasuk rotasi dan pembalikan diterapkan untuk meningkatkan generalisasi.



Gambar 2. Sampel Dataset

Setelah gambar di-*resize*, langkah selanjutnya adalah mengonversi gambar menjadi format tensor. Tensor adalah representasi numerik dari gambar yang memudahkan proses perhitungan oleh model *deep learning*. Tensor ini biasanya memiliki tiga dimensi, yaitu tinggi, lebar, dan saluran warna (RGB). Proses konversi ini sangat penting karena model *neural network* hanya dapat memproses data dalam bentuk tensor. Oleh karena itu, langkah ini memungkinkan model untuk melakukan operasi matematika yang diperlukan untuk ekstraksi fitur dan pelatihan.

Setelah *preprocessing*, dataset dibagi menjadi tiga bagian utama: data latih (*training set*), data validasi (*validation set*), dan data uji (*test set*). Pembagian ini bertujuan untuk memastikan evaluasi yang adil terhadap kinerja model. Data latih digunakan untuk melatih model, sementara data validasi digunakan untuk mengoptimalkan *hyperparameter* dan menghindari *overfitting*. Data uji digunakan untuk mengukur performa model pada data yang belum pernah dilihat sebelumnya. Dalam studi ini, dataset dibagi dengan rasio 70:15:15, yaitu 70% untuk pelatihan, 15% untuk validasi, dan 15% untuk pengujian. Pembagian ini memungkinkan evaluasi kinerja model yang objektif dan dapat memastikan model memiliki kemampuan generalisasi yang baik [6].

Proses *hyperparameter tuning* bertujuan untuk mengoptimalkan kinerja model dengan menemukan kombinasi parameter terbaik yang mempengaruhi proses pelatihan. Pada model *Swin Transformer*, *hyperparameter* seperti *learning rate*, *batch size*, *weight decay*, dan rasio pembagian dataset sangat mempengaruhi

kemampuan model dalam belajar dari data dan menghasilkan prediksi yang akurat.

Tahap pertama dalam *tuning* adalah mendefinisikan ruang pencarian parameter (*parameter grid*) yang mencakup variasi nilai untuk setiap parameter. Setelah itu, *Swin Transformer* diinisialisasi dan *grid search* diterapkan untuk mengevaluasi setiap kombinasi parameter yang mungkin dengan menggunakan teknik *cross-validation* untuk memastikan generalisasi model yang lebih baik.

Dalam studi ini, *grid search* digunakan untuk melakukan pencarian kombinasi parameter yang optimal. *Grid search* melakukan evaluasi terhadap semua kombinasi parameter yang ada dan mengukur performa model berdasarkan metrik evaluasi yang ditentukan, seperti *accuracy* atau *loss*. Proses ini dilakukan untuk memastikan bahwa model *Swin Transformer* mampu mempelajari pola data dengan baik dan memberikan hasil yang lebih presisi [7], parameter yang diuji ditunjukkan pada Tabel 1.

Tabel 1. Parameter Tuning

Nama	Rentang
Learning_Rate	[$1e^{-4}$, $5e^{-5}$]
Batch Size	[8, 16, 32]
Weight Decay	[$1e^{-2}$, $1e^{-3}$]
Rasio Dataset	[80:10:10, 70:15:15, 70:20:10]

Learning rate mengontrol besar perubahan yang dilakukan pada parameter model selama pelatihan. Nilai *learning rate* yang lebih kecil membuat model lebih hati-hati dalam belajar, sementara *learning rate* yang lebih besar dapat mempercepat proses pelatihan namun berisiko melewati pola penting pada data. Oleh karena itu, memilih *learning rate* yang optimal sangat penting untuk memperoleh hasil terbaik dari model.

Batch size merujuk pada jumlah sampel yang digunakan untuk memperbarui bobot model dalam satu iterasi. *Batch size* yang lebih kecil, seperti 8 atau 16, memungkinkan model untuk memperbarui parameter lebih sering, namun membutuhkan lebih banyak iterasi untuk menyelesaikan pelatihan. Sebaliknya, *batch size* yang lebih besar mengurangi jumlah iterasi yang diperlukan, tetapi dapat memperlambat proses pembelajaran karena membutuhkan lebih banyak memori.

Weight decay digunakan untuk mengurangi *overfitting* dengan menambahkan penalti terhadap besar bobot model. Nilai *weight decay* yang lebih besar, seperti $1e^{-2}$, memberikan penalti yang lebih kuat, yang membantu mencegah model terlalu menyesuaikan diri

dengan data latih. Namun, terlalu besar juga dapat mengurangi kemampuan model untuk mempelajari pola kompleks.

Rasio pembagian dataset mengacu pada cara data dibagi menjadi set pelatihan, validasi, dan pengujian. Berbagai rasio pembagian diuji untuk melihat bagaimana pembagian data ini mempengaruhi kinerja model. Pembagian yang lebih seimbang dapat meningkatkan kemampuan model untuk melakukan generalisasi pada data yang tidak terlihat sebelumnya. Dengan menguji berbagai kombinasi parameter tersebut pada data validasi, model dievaluasi untuk mengukur performanya.

Shifted Windows Transformer (*Swin Transformer*) adalah arsitektur pembelajaran mendalam inovatif yang dikembangkan oleh *Microsoft Research Asia* untuk meningkatkan efektivitas model *Transformer* dalam aplikasi visi komputer. *Swin Transformer* memperkenalkan mekanisme *multihead self-attention* berbasis jendela (W-MHSA) bersama dengan jendela yang digeser, memungkinkan pemrosesan gambar yang lebih efisien. Arsitektur ini telah menunjukkan keunggulan signifikan dalam berbagai tugas penglihatan, termasuk klasifikasi dalam citra medis [3].

Arsitektur *Swin Transformer*, dibagi empat fase utama. Prosedur awal adalah *partisi patch*, dimana gambar berdimensi $H \times W \times 3$ disegmentasi menjadi beberapa *patch* kecil, masing-masing diwakili sebagai token dalam lapisan penyematan linier. Token-token tersebut diproses lebih lanjut melalui serangkaian *Block Swin Transformer* yang menggunakan W-MHSA untuk pemrosesan informasi intra-jendela dan SW-MHSA untuk komunikasi antar-jendela. Seiring berjalannya tahapan, penggabungan *patch* dilakukan untuk mengurangi jumlah token dan meningkatkan dimensi fitur, memungkinkan model menangkap informasi hierarkis yang lebih mendalam [8].

Representasi fitur hierarkis adalah salah satu karakteristik utama yang membedakan *Swin Transformer* dari *Vision Transformer* (ViT). Alih-alih mempertahankan ukuran *patch* tetap di semua lapisan seperti ViT, *Swin Transformer* menerapkan struktur hierarkis yang secara bertahap mengurangi jumlah token di setiap tahap. Pendekatan ini memungkinkan model untuk mengekstraksi karakteristik pada berbagai skala, meningkatkan efisiensi pemrosesan gambar beresolusi tinggi sekaligus mengurangi kompleksitas komputasi [9].

Mekanisme W-MHSA adalah komponen inti dari *Swin Transformer* yang dirancang untuk meningkatkan efisiensi komputasi saat mem-

proses gambar beresolusi tinggi. Dalam model *Transformer* konvensional, *self-attention* memiliki kompleksitas komputasi $O(N^2)$, di mana N adalah jumlah token dalam gambar. Hal ini menjadi masalah ketika berhadapan dengan gambar beresolusi tinggi yang berisi banyak token. Untuk mengatasi hal ini, *Swin Transformer* membagi penglihatan menjadi banyak jendela lokal dan menerapkan *self-attention* secara eksklusif di dalam jendela-jendela ini. Metode ini secara substansial mengurangi beban komputasi sambil tetap mempertahankan efektivitas. Secara matematis, seperti yang diwakili dalam persamaan, operasi W-MHSA dirumuskan sebagai berikut:

$$\text{Attention}(Q, K, V) = \text{SoftMax}\left(\frac{QK^T}{\sqrt{d} + B}\right)V \quad 1)$$

dimana Q mewakili *query*, K mewakili *key*, dan V mewakili *value*, masing-masing mewakili token dalam jendela, d adalah dimensi kunci, dan B adalah suku bias. Hasilnya dilewatkan melalui fungsi *Softmax* sebelum dikalikan dengan V untuk menghasilkan *output* akhir. Setelah proses perhatian, setiap token dari W-MHSA diproses oleh *Feed Forward Network* (FFN). Komponen ini meningkatkan representasi fitur yang diekstraksi. FFN dalam *Swin Transformer* terdiri dari dua lapisan linier yang diselingi dengan fungsi aktivasi, GELU, yang memberikan non-linearitas pada proses pembelajaran [10]. Operasi ini secara matematis diwakili dalam persamaan berikut:

$$\text{FFN}(x) = (\text{GELU}(Z_{w_1} + b_1))W_2 + b_2 \quad 2)$$

dimana x adalah *token input*, w adalah bobot, dan b adalah bias untuk lapisan linier. Seperti yang ditunjukkan pada persamaan diatas, operasi ini meningkatkan representasi fitur yang dipelajari selama pelatihan.

Mekanisme *shifted windows* diperkenalkan untuk mengatasi keterbatasan mekanisme W-MHSA, yang hanya beroperasi dalam jendela tetap. Teknik ini meningkatkan komunikasi informasi antara jendela yang sebelumnya terisolasi. Gambar awalnya disegmentasi menjadi jendela tetap, *Multi-Head Self-Attention* (MSA) diterapkan di dalam setiap jendela. Setelah itu, jendela-jendela digeser secara siklis, memastikan komunikasi antar-jendela. Setelah pergeseran, perhatian dihitung ulang dengan *Masked Multi-Head Self-Attention* (MSA) untuk menjaga batas antar jendela. Langkah terakhir melibatkan pembalikan pergeseran, memungkinkan model untuk menangkap hubungan antara area yang

sebelumnya terisolasi dalam jendela yang berbeda.

3. HASIL DAN PEMBAHASAN

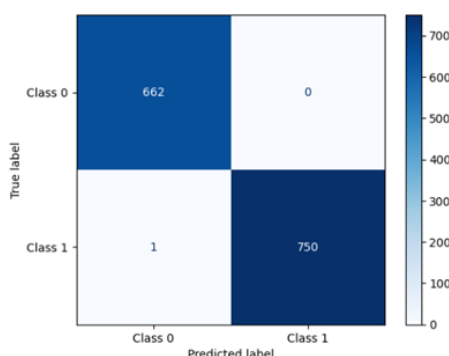
Penerapan tahapan metode memberikan hasil evaluasi yang menunjukkan bahwa model *Swin Transformer* yang telah di-*tuning* memberikan hasil klasifikasi yang lebih stabil dan akurat dibandingkan dengan model lainnya seperti CNN dan ViT.

Tabel 2 menunjukkan bahwa model *Swin Transformer* yang telah di-*tuning* mengungguli model CNN, *swin (base)*, dan ViT dalam hal akurasi, *precision*, *recall*, dan *F1-Score*, dengan hasil masing-masing 0.996, 1.000, 0.998, dan 0.999. Dibandingkan dengan model *baseline Swin* yang memiliki akurasi 0.976, *tuning hyperparameter* secara signifikan meningkatkan kemampuan model dalam mengklasifikasikan gambar. Meskipun ViT menunjukkan kinerja baik, model *Swin* yang di-*tuning* lebih konsisten, dengan performa yang lebih baik dalam *precision* dan *recall*.

Tabel 2. Hasil Metriks Evaluasi

Model	Akurasi	Precision	Recall	F1-score
<i>Swin (Tuned)</i>	0.996	1.000	0.998	0.999
<i>Swin (Base)</i>	0.976	0.981	0.974	0.978
ViT	0.963	0.939	0.996	0.967
CNN	0.909	0.875	0.968	0.919

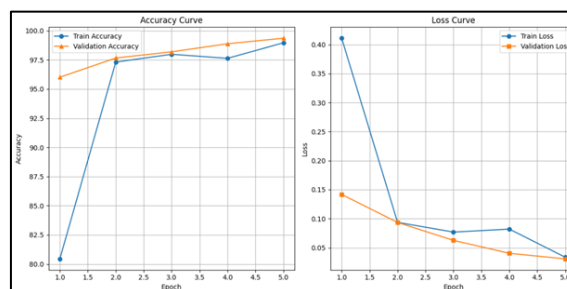
Gambar 3 menunjukkan bahwa model berhasil memprediksi hampir semua sampel dengan benar, dengan *Accuracy* 99.85%. Model tidak membuat kesalahan pada Kelas Negatif dan hanya membuat 1 kesalahan pada Kelas Positif. *Precision* untuk Kelas Positif adalah 100%, dengan *Recall* 99.87% dan *F1-Score* 99.93%, yang menunjukkan kinerja model yang sangat baik dan hampir sempurna.



Gambar 3. Confusion Matrix

Grafik 4 menunjukkan kinerja model selama pelatihan, dengan kurva *Accuracy* dan *Loss* yang menggambarkan peningkatan signifikan. Pada grafik *Accuracy*, *Train Accuracy*

meningkat tajam sejak *epoch* pertama dan mendekati 100% pada *epoch* ke-5, sementara *Validation Accuracy* juga menunjukkan peningkatan yang stabil, meskipun sedikit lebih lambat. Di sisi lain, grafik *Loss* menunjukkan penurunan drastis pada *Train Loss* dan *Validation Loss*, yang mengindikasikan bahwa model berhasil mengurangi kesalahan prediksi secara signifikan pada data latih dan data validasi.



Gambar 4. Grafik Acc & Loss

Secara keseluruhan model menunjukkan kemampuan belajar yang baik dengan meningkatnya *Accuracy* dan menurunnya *Loss*, menandakan bahwa model semakin optimal dalam mempelajari pola data. Hasil evaluasi ini menunjukkan bahwa model *Swin Transformer* yang telah di-*tuning* berhasil mencapai performa yang sangat baik dengan peningkatan akurasi yang signifikan dan efisiensi komputasi yang lebih baik. Model ini tidak hanya menunjukkan akurasi tinggi pada data latih dan validasi, tetapi juga mampu meminimalkan kesalahan prediksi dengan stabil. Dengan hasil yang memuaskan dalam hal *Accuracy*, *Precision*, dan *Recall*, serta performa yang efisien dalam penggunaan memori dan waktu komputasi, model ini sangat cocok untuk diterapkan pada tugas-tugas klasifikasi gambar ultrasonografi, dimana akurasi dan kecepatan sangat penting.

4. PENUTUP

Hasil studi menunjukkan efektivitas *Swin Transformer* dalam mengklasifikasikan gambar ultrasonografi ginjal untuk deteksi batu, mencapai akurasi 99,58%, presisi 99,61%, dan *recall* 99,58%. Meskipun hasil ini menjanjikan, kinerja model didasarkan pada kumpulan data tertentu, dan kemampuan generalisasinya di berbagai lingkungan klinis, termasuk variasi demografi pasien dan peralatan ultrasonografi, masih perlu dievaluasi sepenuhnya.

DAFTAR PUSTAKA

- [1]. B. Kebijakan Pembangunan, K. Kementerian, and K. Ri, "Dalam Angka Tim Penyusun Ski

- 2023 Dalam Angka Kementerian Kesehatan Republik Indonesia.”
- [2]. N. Larasati Kartika Sari, I. Ermina Br Barus, B. Santoso, I. Kusuma, and D. Mulyati, “Aplikasi Image Enhancement untuk Peningkatan Kualitas Citra Ultrasonografi Ginjal,” *Jurnal Ilmiah GIGA*, vol. 25, no. 1, pp. 2022–2023, Jun. 2022, doi: 10.47313/jig.v%vi%i.1627.
- [3]. Z. Liu *et al.*, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” Aug. 2021, doi: <https://doi.org/10.48550/arXiv.2111.09883>.
- [4]. F. T. Anggraeny, Y. V. Via, R. Mumpuni, H. Rahmania Hatta, N. E. Putri, and J. Bastian, “Neural Network Classification in Javanese Handwriting Recognition using Projection Profile Histogram and Local Binary Pattern Histogram,” Apr. 2023, doi: <https://doi.org/10.47577/technium.v16i.9970>.
- [5]. R. Baskaran and M. B. Reuben, “Automated Detection of kidney Stone through Neural Networks,” Mar. 2024, doi: 10.17577/IJERTCONV12IS01010.
- [6]. M. Sivakumar, S. Parthasarathy, and T. Padmapriya, “Trade-off between training and testing ratio in machine learning for medical image processing,” *PeerJ Comput Sci*, vol. 10, Aug. 2024, doi: 10.7717/PEERJ-CS.2245.
- [7]. M. O. Dimgba and R. Andreas, “Optimal Hyperparameter Search Strategies: Benchmarking Grid Search, Random Search, and Genetic Algorithms across Regression, Classification, and Clustering Tasks,” Jun. 2024, doi: 10.13140/RG.2.2.24957.37603.
- [8]. R. Sun, Y. Pang, and W. Li, “Efficient Lung Cancer Image Classification and Segmentation Algorithm Based on an Improved Swin Transformer,” *Electronics (Switzerland)*, vol. 12, no. 4, Feb. 2023, doi: 10.3390/electronics12041024.
- [9]. H. Herath, N. Madusanka, and B. Il Lee, “A Systematic Review of Medical Image Quality Assessment,” Apr. 01, 2025, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/jimaging11040100.
- [10]. N. Shazeer, “GLU Variants Improve Transformer,” Feb. 2020, [Online]. Available: <http://arxiv.org/abs/2002.05202>