

TESIS

**MODEL HIBRIDA *CONVOLUTIONAL NEURAL NETWORK*
DAN *TRANSFORMER* UNTUK KLASIFIKASI MULTI-KELAS
CITRA CT SCAN PENYAKIT GINJAL**



Oleh:
Zishwa Muhammad Jauhar Nafis
NIM : 23206051015

**PROGRAM STUDI INFORMATIKA
PROGRAM MAGISTER FAKULTAS SAINS DAN
TEKNOLOGI
UIN SUNAN KALIJAGA**

**YOGYAKARTA
2025**



KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN KALIJAGA
FAKULTAS SAINS DAN TEKNOLOGI
Jl. Marsda Adisucipto Telp. (0274) 540971 Fax. (0274) 519739 Yogyakarta 55281

PENGESAHAN TUGAS AKHIR

Nomor : B-1593/Un.02/DST/PP.00.9/08/2025

Tugas Akhir dengan judul : MODEL HIBRIDA CONVOLUTIONAL NEURAL NETWORK DAN TRANSFORMER UNTUK KLASIFIKASI MULTI-KELAS CITRA CT SCAN PENYAKIT GINJAL

yang dipersiapkan dan disusun oleh:

Nama : ZISHWA MUHAMMAD JAUHAR NAFIS, S.T.
Nomor Induk Mahasiswa : 23206051015
Telah diujikan pada : Selasa, 29 Juli 2025
Nilai ujian Tugas Akhir : A

dinyatakan telah diterima oleh Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta

TIM UJIAN TUGAS AKHIR



Ketua Sidang

Prof. Dr. Ir. Shofwatul 'Uyun, S.T., M.Kom., IPM., ASEAN Eng.
SIGNED

Valid ID: 6891b54c5ae1d



Penguji I

Ir. Maria Ulfah Siregar, S.Kom., MIT., Ph.D.
SIGNED

Valid ID: 6891a1de99e53



Penguji II

Dr. Ir. Bambang Sugiantoro, M.T., IPU.,
ASEAN Eng.
SIGNED

Valid ID: 68888c1187132



Yogyakarta, 29 Juli 2025
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

Prof. Dr. Dra. Hj. Khurul Wardati, M.Si.
SIGNED

Valid ID: 6891c3de46726

PERNYATAAN KEASLIAN

Yang bertanda tangan dibawah ini:

Nama : Zishwa Muhammad Jauhar Nafis
NIM : 23206051015
Jenjang : Magister
Program Studi : Informatika

menyatakan bahwa naskah tesis ini secara keseluruhan adalah hasil penelitian/ karya saya sendiri, kecuali pada bagian-bagian yang dirujuk sumbernya.

Yogyakarta, 6 Juli 2025

Saya yang menyatakan,



Zishwa Muhammad Jauhar Nafis
NIM: 23206051015

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan dibawah ini:

Nama : Zishwa Muhammad Jauhar Nafis

NIM : 23206051015

Jenjang : Magister

Program Studi : Informatika

menyatakan bahwa naskah tesis ini secara keseluruhan benar-benar bebas dari plagiasi. Jika di kemudian hari terbukti melakukan plagiasi, maka saya siap ditindak sesuai ketentuan hukum yang berlaku.

Yogyakarta, 6 Juli 2025

Saya yang menyatakan,



Zishwa Muhammad Jauhar Nafis

NIM: 23206051015

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

NOTA DINAS PEMBIMBING

Kepada Yth,
Dekan Fakultas Sains dan Teknologi
UIN Sunan Kalijaga
Yogyakarta

Assalamu'alaikum wr. wb.

Setelah melakukan bimbingan, arahan, dan koreksi terhadap penulisan tesis yang berjudul:

**MODEL HIBRIDA CONVOLUTIONAL NEURAL NETWORK DAN
TRANSFORMER UNTUK KLASIFIKASI MULTI-KELAS CITRA CT SCAN
PENYAKIT GINJAL**

Yang ditulis oleh:

Nama : Zishwa Muhammad Jauhar Nafis
NIM : 23206051015
Jenjang : Magister
Program Studi : Informatika

Saya berpendapat bahwa tesis tersebut sudah dapat diajukan kepada Magister Informatika UIN Sunan Kalijaga untuk diujikan dalam rangka memperoleh gelar Magister Informatika.

Wassalamu'alaikum wr. wb.

Yogyakarta, 06 Juli 2025
Pembimbing,


Prof. Dr. Ir. Shofwatul 'Uyun, S.T.,
M.Kom., IPM., ASEAN Eng.
NIP: 19820511 200604 2-002

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

ABSTRAK

Penyakit ginjal merupakan salah satu permasalahan kesehatan global dengan angka prevalensi yang terus meningkat setiap tahunnya. Kondisi ini berpotensi menimbulkan komplikasi serius apabila tidak terdeteksi dan ditangani secara dini. Dalam menunjang diagnosis klinis yang cepat, akurat, dan konsisten, dibutuhkan sistem otomatis yang mampu mendukung pengambilan keputusan medis. Salah satu pendekatan yang dapat digunakan adalah analisis berbasis citra medis. *Computed Tomography* (CT) merupakan salah satu modalitas pencitraan yang sering digunakan dalam evaluasi morfologi dan struktur ginjal. Namun demikian, proses interpretasi citra CT secara manual masih sangat bergantung pada keahlian radiolog, sehingga proses ini cenderung memakan waktu, bersifat subjektif, dan berisiko menghasilkan kesalahan diagnostik.

Pemanfaatan sistem diagnosis berbantuan komputer yang memanfaatkan teknologi *Artificial Intelligence* (AI), khususnya pendekatan *Deep Learning*, merupakan salah satu solusi potensial dalam meningkatkan akurasi serta efisiensi proses identifikasi penyakit ginjal oleh tenaga medis. Dalam penelitian ini mengusulkan sebuah model hibrida dengan mengintegrasikan *Convolutional Neural Networks* (CNN) dalam *Vision Transformer* (ViT) untuk klasifikasi citra CT ginjal ke dalam empat kategori: normal, kista, tumor, dan batu ginjal. Model ini dirancang dengan arsitektur secara berurutan dimulai dengan blok residual CNN untuk ekstraksi fitur lokal, *backbone* ViT untuk menangkap konteks global, CNN klasik untuk memperkuat fitur keluaran ViT, dan blok *classifier feed-forward* untuk prediksi kelas akhir.

Evaluasi dilakukan menggunakan dataset publik CT Kidney dan divalidasi melalui metode Hold-Out dan 7-Fold *Stratified Cross Validation*. Model yang diusulkan mencapai akurasi 99.95% pada validasi dengan teknik Hold-Out dan 99.92% pada 7-Fold Cross-Validation, melampaui kinerja model *baseline* ResNet50 dan ViT murni, serta pendekatan terbaru lainnya. Pengujian generalisasi pada dua dataset eksternal, dengan sumber dan organ yang berbeda juga menunjukkan performa tinggi setelah dilakukan pelatihan ulang pada masing-masing dataset guna menyesuaikan karakteristik data yang berbeda. Selain itu, studi ablasi menunjukkan kontribusi signifikan dari

tiap komponen arsitektur terhadap kinerja model keseluruhan. Hasil ini menegaskan bahwa integrasi arsitektur CNN dan Transformer merupakan pendekatan menjanjikan untuk mendukung diagnosis penyakit ginjal secara otomatis, akurat, dan dapat diandalkan di lingkungan klinis.



MOTTO

Cintai prosesnya:

From Allah, For Allah, To Allah

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

HALAMAN PERSEMBAHAN

Tesis ini penulis persembahkan untuk:

Almamater Tercinta

Prodi Magister Informatika

Fakultas Sains dan Teknologi

Universitas Islam Negeri Sunan Kalijaga Yogyakarta



KATA PENGANTAR

Assalamu'alaikum Wr. Wb.

Segala puji penulis panjatkan ke hadirat Allah SWT atas limpahan rahmat, nikmat, dan petunjuk-Nya, yang telah memberikan kekuatan dan kemudahan sehingga penulis dapat menyelesaikan tesis ini. Shalawat dan salam semoga senantiasa tercurah kepada Nabi Muhammad SAW, teladan utama dalam kehidupan dan keilmuan. Tesis ini, yang berjudul “Model Hibrida CNN Dan Transformer Untuk Klasifikasi Multi-Kelas Citra CT Scan Penyakit Ginjal”, disusun sebagai bagian dari pemenuhan syarat untuk memperoleh gelar Magister Informatika pada Fakultas Sains dan Teknologi, Universitas Islam Negeri Sunan Kalijaga Yogyakarta. Penulis menyadari bahwa pencapaian ini tidak terlepas dari dukungan, bimbingan, serta doa dari berbagai pihak. Oleh karena itu, dengan penuh rasa hormat dan terima kasih, penulis ingin menyampaikan penghargaan yang sebesar-besarnya kepada:

1. Keluarga penulis, terkhusus Ibunda tercinta Mahfudhoh, Almarhum Bapak Muhammad Sa'i, Kakak Thoha Ahmad Aufaddin Zaka, Mbak Khoiratul Maslahah, dan Adek Ula Taliyah Thoyyibah yang senantiasa mendampingi dalam doa dan menjadi sumber semangat di setiap langkah.
2. Prof. Noorhaidi Hasan. Selaku Rektor di Universitas Islam Negeri Sunan Kalijaga Yogyakarta yang telah memberikan kesempatan untuk menempuh studi.
3. Prof. Dr. Dra. Hj. Khurul Wardati, M.Si. Selaku dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sunan Kalijaga

Yogyakarta yang telah memberikan kesempatan untuk menempuh Pendidikan di Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.

4. Dr. Ir. Sumarsono, S.T., M.Kom. selaku Kepala Prodi Magister Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta dan Ir. Muhammad Didik Rohmad Wahyudi, S.T., M.T. Selaku Sekretaris Jurusan Prodi Informatika Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
5. Ibu Prof. Dr. Ir. Shofwatul 'Uyun, S.T., M.Kom., IPM., ASEAN Eng. selaku pembimbing yang telah berkenan merelakan waktu, tenaga dan ilmunya guna memberikan bimbingan kepada penulis dalam menyelesaikan tesis ini, serta ucapan terimakasih banyak dan penghargaan yang setinggi-tingginya yang dengan penuh kesabaran dan kearifan telah memberikan bimbingan, arahan dan dorongan di sela-sela kesibukannya.
6. Bapak dan Ibu Dosen, TU dan tenaga lain Prodi Informatika S2 UIN Sunan Kalijaga Yogyakarta, khususnya yang memberi kuliah, yang sudah memberikan banyak ilmu pengetahuan sehingga penulis dapat melaksanakan penelitian dan Menyusun hasil penelitian tersebut menjadi tesis ini.
7. Teman-teman Ciblon yaitu Muhammad Fahson Hakim, Muh. Nur Aslam, Supardi Atisina, Annida Rizki Luthfi Astuti dan Rahmatun Nazila, serta penghuni PPIIB 2, Terima kasih sudah menjadi bagian perjalanan hidup selama masa perkuliahan.
8. Teman-teman mahasiswa Prodi Informatika S2 UIN Sunan Kalijaga Yogyakarta dan PBSB 2023 UIN Sunan Kalijaga Yogyakarta

9. Kolega di MA Matholi'ul Anwar Lamongan, terima kasih atas dukungan selama menjalani perkuliahan.
10. Ucapan terima kasih saya sampaikan kepada semua pihak yang tidak mungkin saya sebutkan satu demi satu, yang telah banyak memberikan bantuan dan dukungan selama penyusunan Tesis ini.



DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PENGESAHAN	ii
SURAT PERNYATAAN KEASLIAN	iii
SURAT PERNYATAAN BEBAS PLAGIASI	iv
NOTA DINAS PEMBIMBING.....	v
ABSTRAK.....	vi
MOTTO.....	viii
HALAMAN PERSEMBAHAN	ix
KATA PENGANTAR	x
DAFTAR ISI	xiii
DAFTAR TABEL.....	xvi
DAFTAR GAMBAR.....	xvii
DAFTAR LAMPIRAN.....	xviii
DAFTAR SINGKATAN	xix
BAB I PENDAHULUAN	1
A. Latar Belakang.....	1
B. Rumusan Masalah	4
C. Batasan Masalah	5
D. Tujuan.....	5
E. Manfaat Penelitian.....	6

BAB II KAJIAN PUSTAKA DAN LANDASAN TEORI	7
A. Kajian Pustaka	7
B. Landasan Teori.....	18
1. Deep Learning	18
2. Convolutional Neural Networks.....	19
3. Blok Residual	21
4. Vision Transformer.....	22
5. Arsitektur Hibrida.....	24
6. Computed Tomography	24
7. Penyakit Ginjal	25
8. Metrik Evaluasi	27
BAB III METODE PENELITIAN.....	30
A. Data Penelitian.....	30
B. Alat Dan Lingkungan	31
C. Tahapan Penelitian.....	31
1. Pra-pemrosesan.....	32
2. Desain Arsitektur Hibrida.....	33
3. Pelatihan Model.....	39
4. Evaluasi Kinerja	39
BAB IV HASIL DAN PEMBAHASAN	44
A. Pra-pemrosesan.....	44

B. Pelatihan Model.....	46
C. Evaluasi Performa Model.....	48
1. Evaluasi dengan Validasi <i>Hold-Out</i>	48
2. Evaluasi dengan Validasi Silang K-Fold	54
D. Pengujian dengan Dataset Lain	57
E. Pengujian Studi Ablasi	58
F. Perbandingan dengan Studi Sebelumnya	60
BAB V PENUTUP	68
A. Kesimpulan.....	68
B. Saran	69
DAFTAR PUSTAKA.....	70
LAMPIRAN-LAMPIRAN.....	79
DAFTAR RIWAYAT HIDUP.....	81

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
 YOGYAKARTA

DAFTAR TABEL

2.1 Ringkasan Kajian Pustaka	12
3.1 Detail Parameter Vision Transformer Tiny	36
3.2 Pembagian Dataset Untuk Validasi Holdout	40
3.3 Detail Distribusi Dataset Untuk Pengujian Dengan Dataset Lain	42
4.1 Contoh Perhitungan TP, FP, FN dan TN	49
4.2 Contoh Proses Perhitungan Metrik Evaluasi	51
4.3 Hasil Pengujian Performa Model	52
4.4 Hasil Pengujian Validasi Silang 7-Fold	55
4.5 Proses Perhitungan Standar Deviasi Pada Metrik Akurasi	56
4.6 Hasil Pengujian Dengan Dataset Lain	58
4.7 Hasil Pengujian Studi Ablasi	59
4.8 Perbandingan Model Usulan Dengan Penelitian Sebelumnya	63
4.9 Perbandingan Mdoel Usulan Dengan Penelitian Sebelumnya Pada Pengujian Dataset Lain	66

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR GAMBAR

- Gambar 1 Arsitektur Dasar CNN, 21
- Gambar 2 Blok Residual, 21
- Gambar 3 Arsitektur Vision Transformer, 22
- Gambar 4 Potongan Sudut Pandang Pengambilan Citra CT, 25
- Gambar 5 Persentasi Distribusi Data Tiap Kelas, 30
- Gambar 6 Contoh Citra CT Scan Ginjal, 31
- Gambar 7 Alur Tahapan Penelitian, 32
- Gambar 8 Desain Arsitektur Hibrida CNN-Transformer, 33
- Gambar 9 Desain Blok CNN Residual, 35
- Gambar 10 Desain Blok CNN Klasik, 37
- Gambar 11 Desain Arsitektur Blok Classifier, 38
- Gambar 12 Citra Untuk Pengujian Dataset Lain, 43
- Gambar 13 Hasil *Resizing* Ke Ukuran 224x224, 44
- Gambar 14 Contoh Hasil Augmentasi Citra CT Ginjal, 44
- Gambar 15 Perbandingan Histogram Dan Rentang Intesitas Sebelum dan Setelah Normalisasi, 45
- Gambar 16 Grafik Proses Pelatihan Usulan Model, 47
- Gambar 17 Grafik Proses Pelatihan Resnet, 47
- Gambar 18 Grafik Proses Pelatihan ViT, 47
- Gambar 19 Confusion Matrix (A) Usulan Model, (B) Resnet, (C) ViT, 50
- Gambar 20 Grafik Perbandingan Model Usulan Dengan Baseline, 53
- Gambar 21 Grafik Perbandingan Performa Berdasarkan Studi Ablasi, 60

DAFTAR LAMPIRAN

Lampiran 1 Confusion Matrix 7-Fold Cross Validation	79
Lampiran 2 Confusion Matrix Pengujian Dataset Lain dan Studi Ablasi	80



DAFTAR SINGKATAN

AI	: Artificial Intellegence
CAD	: Computer Aided Diagnosis
CNN	: Convolutional Neural Network
CT	: Computed Tomography
CV	: Cross Validation
DL	: Deep Learning
FNR	: False Negative Rate
MRI	: Magnetic Resonance Imaging
NPV	: Negative Predictive Value
PGK	: Penyakit Ginjal Kronis
SVM	: Support Vector Machine
ViT	: Vision Transformer



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

BAB I

PENDAHULUAN

A. Latar Belakang

Penyakit ginjal merupakan masalah kesehatan global yang serius, dengan jumlah penderita yang terus meningkat di berbagai belahan dunia (Bello et al., 2024). Gangguan fungsi ginjal dapat berkembang menjadi Penyakit Ginjal Kronis (PGK) atau bahkan gagal ginjal. Gagal ginjal sendiri merupakan salah satu penyebab utama terhadap angka kematian di masa kini, dengan faktor-faktor risiko seperti merokok, konsumsi alkohol berlebihan, dan kadar kolesterol tinggi memainkan peran penting dalam perkembangannya (Sharen et al., 2024). Keberhasilan penanganan dan pengobatan sangat bergantung pada deteksi dini dan klasifikasi akurat terhadap kelainan ginjal seperti kista, tumor, dan batu ginjal (Owida et al., 2024). Kelainan-kelainan tersebut termasuk yang paling umum terjadi dan dapat secara signifikan mengganggu fungsi ginjal. Tanpa penanganan yang tepat, kondisi ini dapat menimbulkan nyeri hebat dan menurunkan kualitas hidup pasien (Bingol et al., 2023).

Dalam mendukung diagnosis dan klasifikasi dini terhadap berbagai kelainan ginjal, beberapa modalitas pencitraan medis seperti sinar-X, *Computed Tomography* (CT), Ultrasonografi (USG), dan *Magnetic Resonance Imaging* (MRI) sering digunakan (Islam et al., 2022). Di antara berbagai modalitas pencitraan tersebut, CT scan dikenal sebagai salah satu metode yang sangat efektif karena kemampuannya dalam menghasilkan citra penampang tubuh yang jelas, sehingga memungkinkan visualisasi struktur internal ginjal secara lebih detail dan

akurat (Kulkarni et al., 2025). Namun, interpretasi citra CT secara manual oleh radiolog sangat bergantung pada tingkat keahlian dan pengalaman klinis, serta bersifat subjektif. Hal ini dapat menimbulkan risiko kesalahan dalam diagnosis dan berpotensi menyebabkan keterlambatan dalam pengambilan keputusan medis yang tepat (Rehman Khan, 2025). Keterbatasan tersebut mendorong pengembangan sistem Diagnosis Berbantuan Komputer (CAD) dan pemanfaatan teknologi Kecerdasan Buatan (AI) sebagai solusi yang potensial untuk meningkatkan akurasi dan efisiensi dalam proses deteksi serta diagnosis penyakit ginjal oleh tenaga medis, khususnya metode (DL) (Majid et al., 2023a).

Sejalan dengan perkembangan tersebut, berbagai penelitian telah dilakukan dengan memanfaatkan arsitektur dasar Convolutional Neural Networks (CNN) yang menunjukkan performa sangat baik dalam analisis citra medis, termasuk dalam deteksi penyakit ginjal. Sebagai contoh, XResNet-50 berhasil mencapai akurasi sebesar 96,82% untuk deteksi batu ginjal (Yildirim et al., 2021), dan model CNN dengan 6 lapisan melaporkan akurasi sebesar 97% untuk klasifikasi tumor ginjal (Alzu'bi et al., 2022). Beberapa studi juga telah mengeksplorasi berbagai pendekatan untuk klasifikasi penyakit ginjal multikelas, seperti Hypergraph Neural Networks (Sasikaladevi & Revathi, 2024), ResNet-50 (Owida et al., 2024), KidneyNet (Almuayqil et al., 2024a), CNN ringan (Bhandari et al., 2023a), YOLOv8n (Pande & Agarwal, 2024), DenseNet121 yang dimodifikasi (Kausalya et al., 2024), dan VGG16 dengan transfer learning (Prasher et al., 2024), yang semuanya menunjukkan kinerja yang kompetitif.

Model berbasis Convolutional Neural Network (CNN) telah menunjukkan kinerja yang baik dalam analisis citra medis. Namun, pendekatan ini umumnya hanya mampu mengekstraksi fitur spasial lokal, seperti tepi dan tekstur, serta memiliki keterbatasan dalam menangkap informasi kontekstual secara global maupun hubungan jangka panjang dalam citra (Zeynali et al., 2024). Untuk mengatasi keterbatasan tersebut, model berbasis Transformer, yang awalnya dikembangkan untuk tugas pemrosesan bahasa alami, telah diadaptasi dalam bidang visi komputer dan memberikan hasil yang menjanjikan. Penerapan arsitektur Transformer dalam klasifikasi citra ginjal saat ini masih terus berkembang, dengan berbagai penelitian yang berupaya mengeksplorasi dan mengoptimalkan efektivitasnya. Beberapa studi seperti, (Islam et al., 2022) mengevaluasi tiga model Transformer, yaitu EANet, CCT, dan Swin Transformer dalam tugas klasifikasi multikelas penyakit ginjal. Hasil penelitian menunjukkan bahwa Swin Transformer memberikan performa terbaik dengan akurasi mencapai 99,30%. Sementara itu, studi lain melaporkan bahwa modifikasi terhadap model ViT-base-16-224 mampu meningkatkan akurasi klasifikasi hingga 99,64% (Subedi et al., 2023).

Selain penggunaan arsitektur tunggal, berbagai pendekatan hibrida telah diusulkan untuk mengintegrasikan beberapa model dan mengekstraksi fitur-fitur komplementer secara lebih efektif. Beberapa contoh yang menonjol antara lain kombinasi antara CNN dengan Support Vector Machines (SVM) melalui pendekatan *mutual learning* (Chowdhury et al., 2023), penggabungan DenseNet-201 dengan klasifikasi Random Forest (Qadir & Abd, 2023), serta model *ensemble* yang mengintegrasikan MobileNetV2, ResNet50, EfficientNetB0, dan

jaringan Bi-LSTM (saber et al., 2025). Studi lainnya juga telah mengeksplorasi integrasi antara Inception-ConvLSTM (Rehman Khan, 2025) dan CNN-Deep Belief Networks (DBN) yang dioptimalkan menggunakan algoritma Crow-Cuckoo Search (Babu et al., 2025). Berbagai pendekatan ini menunjukkan potensi besar dari model hibrida dalam meningkatkan efektivitas sistem diagnosis penyakit ginjal.

Di antara beragam pendekatan model hibrida, kombinasi antara Convolutional Neural Network dan arsitektur berbasis Transformer akhir-akhir ini semakin mendapat perhatian karena kekuatan keduanya yang saling melengkapi. Meskipun model CNN-Transformer telah berhasil diterapkan dalam sejumlah tugas analisis citra medis lainnya, seperti pada kasus retinopati diabetik (Sadeghzadeh et al., 2023), kanker paru-paru (Wang et al., 2024), dan kanker payudara (Hayat et al., 2024), penerapannya pada klasifikasi citra CT ginjal masih relatif terbatas. Untuk mengisi kesenjangan tersebut, penelitian ini mengusulkan sebuah arsitektur hibrida CNN-Transformer yang secara khusus dirancang untuk mengklasifikasikan berbagai jenis penyakit ginjal berdasarkan citra CT scan. Model yang diusulkan ini mengintegrasikan keunggulan CNN dalam ekstraksi fitur spasial lokal dan kemampuan Transformer dalam melakukan pemodelan konteks global.

B. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana merancang sebuah model hibrida yang mengintegrasikan Convolutional Neural Network (CNN) dan Vision Transformer (ViT) untuk klasifikasi multi-kelas citra CT ginjal?

2. Bagaimana performan model hibrida CNN-Transformer jika dibandingkan dengan model konvensional berbasis CNN atau Transformer tunggal?
3. Bagaimana kontribusi masing-masing komponen arsitektur yang diusulkan terhadap peningkatan performa keseluruhan model?

C. Batasan Masalah

Pada penelitian ini memiliki batasan-batasan masalah yang diterapkan untuk menjaga fokus dan ruang lingkup penelitian. Batasan masalah pada penelitian ini adalah :

1. Strategi arsitektur hibrida yang digunakan dalam penelitian ini terbatas pada satu jenis kombinasi, yaitu antara CNN (ResNet dan CNN klasik) dengan Vision Transformer (ViT), tanpa melakukan eksplorasi lebih lanjut terhadap alternatif strategi hibrida lainnya.
2. Penelitian ini tidak mencakup proses pencarian atau penyesuaian *hyperparameter* secara sistematis (*hyperparameter tuning*).
3. Ruang lingkup penelitian ini tidak menerapkan penggunaan teknik Explainable AI (XAI), sehingga interpretabilitas hasil klasifikasi tidak dianalisis dalam proses evaluasi.

D. Tujuan

Tujuan dari penelitian ini adalah sebagai berikut:

1. Mengembangkan arsitektur model hibrida yang mengintegrasikan CNN dan Transformer untuk klasifikasi multikelas penyakit ginjal dari citra CT.

2. Mengevaluasi performa model hibrida yang diusulkan dan membandingkannya dengan model konvensional berbasis CNN dan Transformer.
3. Menganalisis kontribusi setiap komponen arsitektur melalui studi ablasi untuk mengetahui peran penting masing-masing komponen dalam performa akhir model.

E. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan kontribusi dalam pengembangan teknologi klasifikasi citra medis dengan menghadirkan arsitektur hibrida CNN-Transformer yang mampu meningkatkan performa dalam mendeteksi berbagai jenis kelainan ginjal melalui citra CT. Selain itu, hasil dari penelitian ini juga dapat menjadi referensi bagi studi lanjutan dalam penerapan model deep learning, khususnya dalam pengembangan arsitektur hibrida CNN-Transformer untuk diagnosis penyakit secara otomatis. Di samping itu, penelitian ini diharapkan dapat mendorong pengembangan sistem pendukung keputusan medis yang lebih andal, efisien, dan akurat dalam konteks analisis citra medis.

BAB V

PENUTUP

A. Kesimpulan

Penelitian ini mengusulkan sebuah arsitektur hibrida CNN-Transformer untuk klasifikasi citra CT ginjal, yang secara efektif menggabungkan kemampuan ekstraksi fitur lokal CNN dengan representasi global Vision Transformer. Arsitektur mode usulan secara berurutan tersusun atas blok residual sebagai fitur ekstraktor pertama kemudian masuk ke arsitektur utama ViT. Hasil output ViT akan diproses oleh CNN klasik dan akan dimasukkan ke classifier untuk menentukan jenis penyakitnya. Pada pengujian dengan baseline model dengan teknik validasi hold-out model ini diatas *baseline* model ResNet dan ViT. Selain pengujian dengan teknik validasi hold-out, dilakukan juga validasi silang 7-fold yang menunjukkan tingkat kestabilan yang sangat baik pada performa model usulan.

Evaluasi pada dataset eksternal menunjukkan bahwa model juga memiliki kemampuan generalisasi yang sangat baik terhadap variasi data dari sumber yang berbeda, meskipun pada dataset dengan organ yang berbeda performa model usulan masih berada dibawah beberapa penelitian. Studi ablasi menunjukkan bahwa tiap komponen CNN, baik blok residual maupun CNN klasik yang digabungkan pada arsitektur transformer dapat meningkatkan performa model. Didapatkan bahwa residual blok dan blok CNN klasik memiliki kontribusi signifikan dalam mendukung pembelajaran fitur pada model ini. Meskipun CNN dan Transformer masing-masing memiliki kekuatan tersendiri, penggabungannya dalam desain hibrida menghasilkan representasi fitur

yang lebih kaya dan menyeluruh. Penelitian ini menunjukkan bahwa pendekatan arsitektur hibrida ini efektif dan memiliki potensi kuat untuk diimplementasikan dalam sistem alat bantu diagnosis penyakit ginjal.

B. Saran

Pengembangan lebih lanjut dari penelitian ini dapat diarahkan pada integrasi berbagai modalitas citra medis, seperti MRI atau ultrasound, selain citra CT ginjal, guna memperkaya representasi data dan meningkatkan performa klasifikasi. Strategi pendekatan hibrida masih memiliki ruang untuk dieksplorasi lebih mendalam, baik melalui pemanfaatan arsitektur model lain dari CNN atau Transformer, maupun melalui penyusunan desain hibrida alternatif, seperti arsitektur paralel atau penggabungan fitur dari berbagai blok pemrosesan. Selain itu, penggabungan dataset dari berbagai sumber sebagai satu himpunan data pelatihan juga dapat menjadi strategi potensial untuk meningkatkan keberagaman data, sehingga mendukung pengembangan model yang lebih handal dan adaptif terhadap variasi kasus klinis di dunia nyata. Eksplorasi ini diharapkan mampu menghasilkan model yang andal dan siap diterapkan dalam praktik klinis nyata untuk mendukung proses klasifikasi citra medis secara lebih akurat dan efisien.

DAFTAR PUSTAKA

- Almuayqil, S. N., Abd El-Ghany, S., Abd El-Aziz, A. A., & Elmogy, M. (2024a). KidneyNet: A Novel CNN-Based Technique for the Automated Diagnosis of Chronic Kidney Diseases from CT Scans. *Electronics*, 13(24), 4981. <https://doi.org/10.3390/electronics13244981>
- Alzu'bi, D., Abdullah, M., Hmeidi, I., AlAzab, R., Gharaibeh, M., El-Heis, M., Almotairi, K. H., Forestiero, A., Hussein, A. M., & Abualigah, L. (2022). Kidney Tumor Detection and Classification Based on Deep Learning Approaches: A New Dataset in CT Scans. *Journal of Healthcare Engineering*, 2022, 1–22. <https://doi.org/10.1155/2022/3861161>
- Ankile, L. L., Heggland, M. F., & Krange, K. (2020). *Deep Convolutional Neural Networks: A survey of the foundations, selected improvements, and some current applications*.
- Azevedo, B. F., Rocha, A. M. A. C., & Pereira, A. I. (2024). Hybrid approaches to optimization and machine learning methods: a systematic literature review. *Machine Learning*, 113(7), 4055–4097. <https://doi.org/10.1007/s10994-023-06467-x>
- Babu, G. R., Pushpalatha, N., Khekare, G., Krishnamoorthy, El-Ebiary, Y. A. B., & Devi, S. A. (2025). Enhanced Kidney Stone Detection through Hybrid Crow-Cuckoo Search Optimized Convolutional Deep Belief Network Model. *Journal of Advances in Information Technology*, 16(3), 380–395. <https://doi.org/10.12720/jait.16.3.380-395>
- Bello, A. K., Okpechi, I. G., Levin, A., Ye, F., Damster, S., Arruebo, S., Donner, J.-A., Caskey, F. J., Cho, Y., Davids, M. R., Davison, S. N., Htay, H., Jha, V., Lalji, R., Malik, C., Nangaku, M., See, E., Sozio, S. M., Tonelli, M., ... Zaidi, D. (2024). An update on the global disparities in kidney disease burden and care across world

- countries and regions. *The Lancet Global Health*, 12(3), e382–e395. [https://doi.org/10.1016/S2214-109X\(23\)00570-3](https://doi.org/10.1016/S2214-109X(23)00570-3)
- Bhandari, M., Yogarajah, P., Kavitha, M. S., & Condell, J. (2023a). Exploring the Capabilities of a Lightweight CNN Model in Accurately Identifying Renal Abnormalities: Cysts, Stones, and Tumors, Using LIME and SHAP. *Applied Sciences*, 13(5), 3125. <https://doi.org/10.3390/app13053125>
- Bingol, H., Yildirim, M., Yildirim, K., & Alatas, B. (2023). Automatic classification of kidney CT images with relief based novel hybrid deep model. *PeerJ Computer Science*, 9, e1717. <https://doi.org/10.7717/peerj-cs.1717>
- Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., & Kalinin, A. A. (2020). Albumentations: Fast and Flexible Image Augmentations. *Information*, 11(2), 125. <https://doi.org/10.3390/info11020125>
- Carmo, D., Pinheiro, G., Rodrigues, L., Abreu, T., Lotufo, R., & Rittner, L. (2023). *Automated computed tomography and magnetic resonance imaging segmentation using deep learning: a beginner's guide*.
- Chowdhury, S., Mukhopadhyay, S., & Narendra, K. S. (2023). *Mutual Learning Algorithm for Kidney Cyst, Kidney Tumor and Kidney Stone Diagnosis*. 401–410. <https://doi.org/10.15439/2023F378>
- da Cruz, L. B., Araújo, J. D. L., Ferreira, J. L., Diniz, J. O. B., Silva, A. C., de Almeida, J. D. S., de Paiva, A. C., & Gattass, M. (2020). Kidney segmentation from computed tomography images using deep neural network. *Computers in Biology and Medicine*, 123, 103906. <https://doi.org/10.1016/j.combiomed.2020.103906>
- Dong, S., Wang, P., & Abbas, K. (2021). A survey on deep learning and its applications. *Computer Science Review*, 40, 100379. <https://doi.org/10.1016/j.cosrev.2021.100379>

- Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*.
- Dwivedi, N., Jamadar, A., Mathew, S., Fields, T. A., & Rao, R. (2023). Myofibroblast depletion reduces kidney cyst growth and fibrosis in autosomal dominant polycystic kidney disease. *Kidney International*, 103(1), 144–155. <https://doi.org/10.1016/j.kint.2022.08.036>
- Farhadpour, S., Warner, T. A., & Maxwell, A. E. (2024). Selecting and Interpreting Multiclass Loss and Accuracy Assessment Metrics for Classifications with Class Imbalance: Guidance and Best Practices. *Remote Sensing*, 16(3), 533. <https://doi.org/10.3390/rs16030533>
- Ganie, S. M., Pramanik, P. K. D., & Zhao, Z. (2024). Deep Transfer Learning for Kidney Disease Detection Using CT Scan Images. *2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 5624–5629. <https://doi.org/10.1109/BIBM62325.2024.10822760>
- Gharaibeh, M., Alzu'bi, D., Abdullah, M., Hmeidi, I., Al Nasar, M. R., Abualigah, L., & Gandomi, A. H. (2022). 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*, 6(1), 29. <https://doi.org/10.3390/bdcc6010029>
- Hayat, M., Ahmad, N., Nasir, A., & Ahmad Tariq, Z. (2024). Hybrid Deep Learning EfficientNetV2 and Vision Transformer (EffNetV2-ViT) Model for Breast Cancer Histopathological Image Classification. *IEEE Access*, 12, 184119–184131. <https://doi.org/10.1109/ACCESS.2024.3503413>
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). *Deep Residual Learning for Image Recognition*.

- Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, 12(1), 5979. <https://doi.org/10.1038/s41598-022-09954-8>
- Hossain, M. M., Walid, Md. A. A., Galib, S. M. S., Azad, M. M., Rahman, W., Shafi, A. S. M., & Rahman, M. M. (2024). COVID-19 detection from chest CT images using optimized deep features and ensemble classification. *Systems and Soft Computing*, 6, 200077. <https://doi.org/https://doi.org/10.1016/j.sasc.2024.200077>
- Hossain, T., Sayed, F., & Islam, S. (2024). *Adaptive Local Binary Pattern: A Novel Feature Descriptor for Enhanced Analysis of Kidney Abnormalities in CT Scan Images using ensemble based Machine Learning Approach.*
- Islam, M. N., Hasan, M., Hossain, Md. K., Alam, Md. G. R., Uddin, M. Z., & Soyulu, A. (2022). Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography. *Scientific Reports*, 12(1), 11440. <https://doi.org/10.1038/s41598-022-15634-4>
- K, K. S., K P, B. M., Patwari, N., & D A, S. (2023). DeepKidney: Multiclass Classification of Kidney Stones, Cysts, Tumors, and Normal Cases Using Convolutional Neural Networks. *2023 4th International Conference on Communication, Computing and Industry 6.0 (C2I6)*, 1–5. <https://doi.org/10.1109/C2I659362.2023.10431043>
- Kausalya, K., Chandra, B., Kumar, S. B., & Mohammed Riyaz, A. (2024). Enhancing Chronic Kidney Disease Diagnosis Through Densenet-121 Approach. *2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN)*, 55–62. <https://doi.org/10.1109/ICPCSN62568.2024.00017>
- Khan, Z. A., Waqar, M., Khan, H. U., Chaudhary, N. I., Khan, A. T., Ishtiaq, I., Khan, F. A., & Raja, M. A. Z. (2025). Fine-tuned deep

- transfer learning: an effective strategy for the accurate chronic kidney disease classification. *PeerJ Computer Science*, 11, e2800. <https://doi.org/10.7717/peerj-cs.2800>
- Kim, J. W., Urooj Khan, A., & Banerjee, I. (2024). *Systematic Review of Hybrid Vision Transformer Architectures for Radiological Image Analysis*. <https://doi.org/10.1101/2024.06.21.24309265>
- Krichen, M. (2023). Convolutional Neural Networks: A Survey. *Computers*, 12(8), 151. <https://doi.org/10.3390/computers12080151>
- Kulkarni, S., S, S. R., Chopra, T. K., S P, Y., K J, B., & K R, S. (2025). Optimizing Kidney Stone Detection with Vision Transformers and CNNs: A Transfer Learning Strategy. *2025 International Conference on Pervasive Computational Technologies (ICPCT)*, 829–833. <https://doi.org/10.1109/ICPCT64145.2025.10940924>
- Kumar, U. (2024). A noval approach of classification of COVID-19 from chest CT-Scan images using ensemble classifier in combination with cognition based texture features. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-024-18670-6>
- Majid, M., Gulzar, Y., Ayoub, S., Khan, F., Reegu, F. A., Mir, M. S., Jaziri, W., & Soomro, A. B. (2023b). Enhanced Transfer Learning Strategies for Effective Kidney Tumor Classification with CT Imaging. *International Journal of Advanced Computer Science and Applications*, 14(8). <https://doi.org/10.14569/IJACSA.2023.0140847>
- Noor, M. H. M., & Ige, A. O. (2025). *A Survey on State-of-the-art Deep Learning Applications and Challenges*.
- Owida, H. A., Al-Nabulsi, J., Al Hawamdeh, N., Abuowaida, S., Salah, Z., & Elsoud, E. A. (2024). Deep Learning-Based Kidney Disease Classification Using ResNet50: Enhancing Diagnostic Accuracy. *2024 Second Jordanian International Biomedical Engineering*

- Conference (JIBEC)*, 126–129.
<https://doi.org/10.1109/JIBEC63210.2024.10931915>
- Özbay, E., Özbay, F. A., & Gharehchopogh, F. S. (2024). Kidney Tumor Classification on CT images using Self-supervised Learning. *Computers in Biology and Medicine*, 176, 108554.
<https://doi.org/https://doi.org/10.1016/j.compbiomed.2024.108554>
- Pande, S. D., & Agarwal, R. (2024). Multi-Class Kidney Abnormalities Detecting Novel System Through Computed Tomography. *IEEE Access*, 12, 21147–21155.
<https://doi.org/10.1109/ACCESS.2024.3351181>
- Patro, K. K., Allam, J. P., Neelapu, B. C., Tadeusiewicz, R., Acharya, U. R., Hammad, M., Yildirim, O., & Pławiak, P. (2023). Application of Kronecker convolutions in deep learning technique for automated detection of kidney stones with coronal CT images. *Information Sciences*, 640, 119005.
<https://doi.org/10.1016/j.ins.2023.119005>
- Prasher, S., Nelson, L., & Sandhya, V. (2024). VGG16 Transfer Learning Model for Diagnosing Multi-Class Kidney Disorder. *2024 5th International Conference for Emerging Technology (INCET)*, 1–5.
<https://doi.org/10.1109/INCET61516.2024.10593229>
- Qadir, A. M., & Abd, D. F. (2023). Kidney Diseases Classification using Hybrid Transfer-Learning DenseNet201-Based and Random Forest Classifier. *Kurdistan Journal of Applied Research*, 131–144.
<https://doi.org/10.24017/Science.2022.2.11>
- Rangel, G., Cuevas-Tello, J. C., Nunez-Varela, J., Puente, C., & Silva-Trujillo, A. G. (2024). A Survey on Convolutional Neural Networks and Their Performance Limitations in Image Recognition Tasks. *Journal of Sensors*, 2024(1).
<https://doi.org/10.1155/2024/2797320>

- Rehman Khan, S. U. (2025). Multi-level feature fusion network for kidney disease detection. *Computers in Biology and Medicine*, 191, 110214. <https://doi.org/10.1016/j.compbimed.2025.110214>
- saber, A., Hassan, E., Elbedwehy, S., Awad, W. A., & Emara, T. Z. (2025). Leveraging ensemble convolutional neural networks and metaheuristic strategies for advanced kidney disease screening and classification. *Scientific Reports*, 15(1), 12487. <https://doi.org/10.1038/s41598-025-93950-1>
- Sadeghzadeh, A., Junayed, M. S., Aydin, T., & Islam, M. B. (2023). Hybrid CNN+Transformer for Diabetic Retinopathy Recognition and Grading. *2023 Innovations in Intelligent Systems and Applications Conference (ASYU)*, 1–6. <https://doi.org/10.1109/ASYU58738.2023.10296789>
- Sahoo, P., Saha, S., Mondal, S., & Gowda, S. (2022). Vision Transformer Based COVID-19 Detection Using Chest CT-scan images. *2022 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, 01–04. <https://doi.org/10.1109/BHI56158.2022.9926823>
- Saputra, N. A., Riza, L. S., Setiawan, A., & Hamidah, I. (2024). A Systematic Review for Classification and Selection of Deep Learning Methods. *Decision Analytics Journal*, 12, 100489. <https://doi.org/10.1016/j.dajour.2024.100489>
- Sasikaladevi, N., & Revathi, A. (2024). Digital twin of renal system with CT-radiography for the early diagnosis of chronic kidney diseases. *Biomedical Signal Processing and Control*, 88, 105632. <https://doi.org/10.1016/j.bspc.2023.105632>
- Schlosser, T., Friedrich, M., Meyer, T., & Kowerko, D. (2024). *A Consolidated Overview of Evaluation and Performance Metrics for Machine Learning and Computer Vision*. <https://doi.org/10.13140/RG.2.2.14331.69928>

- Senan, E. M., Al-Adhaileh, M. H., Alsaade, F. W., Aldhyani, T. H. H., Alqarni, A. A., Alsharif, N., Uddin, M. I., Alahmadi, A. H., Jadhav, M. E., & Alzahrani, M. Y. (2021). Diagnosis of Chronic Kidney Disease Using Effective Classification Algorithms and Recursive Feature Elimination Techniques. *Journal of Healthcare Engineering*, 2021, 1–10. <https://doi.org/10.1155/2021/1004767>
- Sharen, H., Narendra, M., & Anbarasi, L. J. (2024). MSKd_Net: Multi-Head Attention-Based Swin Transformer for Kidney Diseases Classification. *IEEE Access*, 12, 181975–181986. <https://doi.org/10.1109/ACCESS.2024.3510634>
- Sharma, G., Anand, V., Chauhan, R., Pokhariya, H. S., Gupta, S., & Sunil, G. (2024). Revolutionizing Kidney Disease Diagnosis: A Comprehensive CNN-Based Framework for Multi-Class CT Classification. 2024 *IEEE International Conference on Information Technology, Electronics and Intelligent Communication Systems (ICITEICS)*, 1–6. <https://doi.org/10.1109/ICITEICS61368.2024.10624992>
- Soares, E., Angelov, P., Biaso, S., Froes, M. H., & Abe, D. K. (2020). SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification. *MedRxiv*, 2020.04.24.20078584. <https://doi.org/10.1101/2020.04.24.20078584>
- Subedi, R., Timilsina, S., & Adhikari, S. (2023). Kidney CT Scan Image Classification Using Modified Vision Transformer. *Journal of Engineering and Sciences*, 2(1), 24–29. <https://doi.org/10.3126/jes2.v2i1.60381>
- Szczykutowicz, T. P., Toia, G. V., Dhanantwari, A., & Nett, B. (2022). A Review of Deep Learning CT Reconstruction: Concepts, Limitations, and Promise in Clinical Practice. *Current Radiology Reports*, 10(9), 101–115. <https://doi.org/10.1007/s40134-022-00399-5>

- Wang, L., Zhang, C., & Li, J. (2024). A Hybrid CNN-Transformer Model for Predicting N Staging and Survival in Non-Small Cell Lung Cancer Patients Based on CT-Scan. *Tomography*, 10(10), 1676–1693. <https://doi.org/10.3390/tomography10100123>
- Yildirim, K., Bozdog, P. G., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. *Computers in Biology and Medicine*, 135, 104569. <https://doi.org/10.1016/j.compbiomed.2021.104569>
- Yousefpanah, K., Ebadi, M. J., Sabzekar, S., Zakaria, N. H., Osman, N. A., & Ahmadian, A. (2024). An emerging network for COVID-19 CT-scan classification using an ensemble deep transfer learning model. *Acta Tropica*, 257, 107277. <https://doi.org/https://doi.org/10.1016/j.actatropica.2024.107277>
- Zeynali, A., Tinati, M. A., & Tazehkand, B. M. (2024). Hybrid CNN-Transformer Architecture With Xception-Based Feature Enhancement for Accurate Breast Cancer Classification. *IEEE Access*, 12, 189477–189493. <https://doi.org/10.1109/ACCESS.2024.3516535>