

TESIS

**KLASIFIKASI CITRA MAMOGRAFI MENGGUNAKAN
DEEP TRANSFER LEARNING BERDASARKAN
FINE-TUNING DAN XGBOOST**



Oleh:

Nur Faridah

22206051018

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA
PROGRAM STUDI INFORMATIKA

**PROGRAM MAGISTER FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI SUNAN KALIJAGA
YOGYAKARTA**

2024



PENGESAHAN TUGAS AKHIR

Nomor : B-1589/Un.02/DST/PP.00.9/08/2024

Tugas Akhir dengan judul : **KLASIFIKASI CITRA MAMOGRAFI MENGGUNAKAN DEEP TRANSFER LEARNING BERDASARKAN FINE-TUNING DAN XGBOOST**

yang dipersiapkan dan disusun oleh:

Nama : NUR FARIDAH, S.Mat
Nomor Induk Mahasiswa : 22206051018
Telah diujikan pada : Senin, 12 Agustus 2024
Nilai ujian Tugas Akhir : A-

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

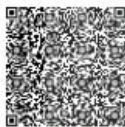
TIM UJIAN TUGAS AKHIR



Ketua Sidang

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

Valid ID: 66c8c3aca79aa



Penguji I

Dr. Ir. Sumarsono, S.T., M.Kom.
SIGNED

Valid ID: 66c9857751436



Penguji II

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

Valid ID: 66c68452d5282



Yogyakarta, 12 Agustus 2024
UIN Sunan Kalijaga
Dekan Fakultas Sains dan Teknologi

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

Valid ID: 66c87eb12282

SURAT PERNYATAAN KEASLIAN

Yang bertanda tangan di bawah ini:

Nama : Nur Faridah
NIM : 22206051018
Jenjang : Magister
Program Studi : Informatika

Menyatakan bahwa naskah tesis ini dengan judul “**KLASIFIKASI CITRA MAMOGRAFI MENGGUNAKAN *DEEP TRANSFER LEARNING* BERDASARKAN *FINE-TUNING* DAN *XGBOOST***” tidak terdapat pada karya yang pernah diajukan untuk memperoleh gelar Magister di suatu Perguruan Tinggi, dan sepengetahuan saya tidak terdapat karya atau pendapat yang pernah ditulis oleh orang lain, kecuali pada bagian-bagian yang dirujuk sumbernya.

Yogyakarta, 26 Juli 2024

Saya yang menyatakan,



Nur Faridah

NIM: 22206051018

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

SURAT PERNYATAAN BEBAS PLAGIASI

Yang bertanda tangan di bawah ini:

Nama : Nur Faridah
NIM : 22206051018
Jenjang : Magister
Program Studi : Informatika

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

Yogyakarta, 26 Juli 2024

Saya yang menyatakan,



Nur Faridah

NIM: 22206051018

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA



Universitas Islam Negeri Sunan Kalijaga

SURAT PERSETUJUAN TESIS/TUGAS AKHIR

Kepada:

Yth. Dekan Fakultas Sains dan Teknologi

UIN Sunan Kalijaga Yogyakarta

Di Yogyakarta

Assalamualaikum wr.wb.

Setelah membaca, meneliti, memberikan petunjuk dan mengoreksi serta mengadakan perbaikan seperlunya, maka saya selaku pembimbing berpendapat bahwa tesis Saudara:

Nama : Nur Faridah

NIM : 22206051018

Judul Skripsi : Klasifikasi Citra Mamografi Menggunakan *Deep Transfer Learning*
Berdasarkan *Fine-tuning* dan XGBoost

Sudah dapat diajukan kepada Program Studi Magister Informatika UIN Sunan Kalijaga sebagai salah satu syarat untuk memperoleh gelar Magister Informatika.

Waasalamualaikum wr.wb.

Yogyakarta, 25 Juli 2024

Pembimbing,

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

Prof. Dr. Ir. Shofwatul Uyun, S.T., M.Kom

NIP. 198205112006042002

ABSTRAK

Kanker payudara merupakan kanker non-kulit yang biasa menyerang wanita. Kanker ini merupakan jenis kanker dengan kasus terbanyak kedua setelah kanker paru-paru, dengan angka kematian mencapai 6,8%. Diagnosis dini kanker payudara memiliki peran penting dalam proses pengobatan. Salah satu media dalam mendiagnosis kanker payudara yaitu dengan menggunakan citra mamografi. Citra ini mempunyai tingkat kesulitan yang cukup tinggi jika diinterpretasikan secara visual, sehingga diperlukan proses diagnosis secara komputasi dengan menggunakan *deep learning*. Penelitian bertujuan untuk menawarkan metode klasifikasi otomatis kanker payudara yang tepat menggunakan teknik *deep transfer learning* dengan arsitektur CNN VGG-16 sebagai *feature extractor* dan menggunakan dua teknik pemodelan yaitu *fine-tuning* dan *Extreme Gradient Boosting* (XGBoost). Dataset yang digunakan berasal dari *Mammographic Image Analysis Society* (MIAS) yang digunakan untuk menguji keefektifan metode yang disarankan untuk klasifikasi kanker payudara menjadi tiga kelas yaitu *benign*, *malignant*, dan normal. Hasil pengujian menunjukkan bahwa akurasi terbaik pada teknik *fine-tuning* sebesar 96.02% dengan *optimizer* RMSprop, *batch size* 64, *learning rate* 0.01, dan *epoch* 100. Sedangkan pada teknik XGBoost diperoleh akurasi sebesar 93.03% dengan *max depth* 8, *learning rate* 0.4, dan *num round* 500.

Kata Kunci: *Deep transfer learning*, VGG-16, *Fine-tuning*, *Extreme Gradient Boosting*, Kanker payudara.

ABSTRACT

Breast cancer is a non-skin cancer that usually attacks women. This cancer is the type of cancer with the second most cases after lung cancer, with a mortality rate of 6.8%. Early diagnosis of breast cancer has an important role in the treatment process. One medium for diagnosing breast cancer is using mammography images. This image has a fairly high level of difficulty when interpreted visually, so a computational diagnosis process using deep learning is required. The research aims to offer a precise automatic classification method for breast cancer using deep transfer learning techniques with the CNN VGG-16 architecture as a feature extractor and using two modeling techniques, namely fine-tuning and Extreme Gradient Boosting (XGBoost). The dataset used comes from the Mammographic Image Analysis Society (MIAS) which is used to test the effectiveness of the recommended method for classifying breast cancer into three classes, namely benign, malignant and normal. The test results show that the best accuracy in the fine-tuning technique is 96.02% with the RMSprop optimizer, batch size 64, learning rate 0.01, and epoch 100. Meanwhile, in the XGBoost technique, the accuracy is 93.03% with max depth 8, learning rate 0.4, and num round 500.

Keywords: *Deep transfer learning, VGG-16, Fine-tuning, Extreme Gradient Boosting, Breast cancer.*

MOTTO

“Sebaik-baik manusia adalah yang paling bermanfaat bagi orang lain.”

(Hadits Riwayat ath-Thabrani)

Boleh jadi kamu membenci sesuatu, padahal ia amat baik bagimu, dan boleh jadi (pula) kamu menyukai sesuatu, padahal ia amat buruk bagimu; Allah mengetahui, sedang kamu tidak mengetahui.

(Q.S Al-Baqarah: 216)

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

PERSEMBAHAN

Tugas akhir ini dipersembahkan:

Kepada kedua orang tua saya yang telah memberikan dukungan penuh dan mendoakan dengan segenap hatinya.

Untuk Almamater Universitas Islam Negeri Sunan Kalijaga
Yogyakarta, dan teman-teman Program Studi Magister Informatika
angkatan 2022



STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

KATA PENGANTAR

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

Puji dan syukur tak lupa penulis sampaikan kepada Allah SWT atas segala nikmat, rahmat, serta hidayah-Nya sehingga penulisan tesis yang berjudul “*Klasifikasi Citra Mamografi Menggunakan Deep Transfer Learning Berdasarkan Fine-tuning dan XGBoost*” dapat terselesaikan dengan tepat waktu. Shalawat serta salam semoga senantiasa tercurahkan kepada junjungan kita Nabi Muhammad SAW yang kita harapkan syafaatnya di yaumul akhir nanti.

Penyusunan tesis ini dimaksudkan untuk memenuhi syarat kuliah guna memperoleh gelar Magister Program Studi Informatika. Penulis menyadari dalam penulisan tesis ini tidak akan terselesaikan tanpa adanya bantuan dan bimbingan dari berbagai pihak. Oleh karena itu dengan segala kerendahan hati penyusun banyak mengucapkan terimakasih kepada:

1. Prof. Dr. Phil Al Makin, MA., selaku rektor UIN Sunan Kalijaga Yogyakarta.
2. Dr. Khurul Wardati, M.Si., selaku dekan Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
3. Dr. Ir. Bambang Sugiantoro, S.Si., M.T., selaku Ketua Program Studi Magister Informatika Fakultas Sains dan Teknologi UIN Sunan Kalijaga Yogyakarta.
4. Muhamad Taufiq Nuruzzaman, S.T., M.Eng., Ph.d., selaku Dosen Penasehat Akademik.

5. Prof. Dr. Ir. Shofwatul 'Uyun, S.T., M.Kom selaku Dosen Pembimbing Tesis yang telah banyak membantu dan membimbing penulis dalam menyusun tesis ini.
6. Bapak Mansur dan Ibu Aminah selaku orang tua penulis yang telah memberikan cinta dan kasih sayang kepada penulis sehingga penulis mampu untuk menyelesaikan tesis ini.
7. Semua teman Informatika angkatan 2022 yang selalu memberikan motivasi, semangat dan bantuannya selama masa-masa perkuliahan.
8. Keluarga PP Al-Munawwir kompleks Q yang telah memberikan semangat dan dukungan kepada penulis dalam menyelesaikan skripsi.
9. Pihak-pihak lainnya yang tidak bisa penulis sebutkan satu persatu yang turut memberikan dukungan kepada penulis dalam menyelesaikan tesis ini.

Penulis menyadari bahwa tesis ini masih jauh dari kesempurnaan dan masih terdapat banyak kekurangan, untuk itu penulis mengharap segala kritik dan saran yang bersifat membangun yang dapat digunakan untuk penelitian selanjutnya. Penulis juga berharap semoga tesis ini dapat bermanfaat bagi banyak orang terutama dalam dunia akademik dan ilmu pengetahuan. Aamiin ya Rabbal Aalamiin.

Yogyakarta, 25 Agustus 2024

Penulis
Nur Faridah

DAFTAR ISI

LEMBAR PENGESAHAN.....	ii
PERNYATAAN KEASLIAN.....	iii
PERNYATAAN BEBAS PLAGIASI.....	iv
SURAT PERSETUJUAN TESIS/TUGAS AKHIR.....	v
ABSTRAK.....	vi
ABSTRACT.....	vii
MOTTO.....	viii
PERSEMBAHAN.....	ix
KATA PENGANTAR.....	x
DAFTAR ISI.....	xii
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
BAB I PENDAHULUAN.....	1
A. LATAR BELAKANG.....	1
B. RUMUSAN MASALAH.....	5
C. BATASAN MASALAH.....	5
D. TUJUAN PENELITIAN.....	5
E. MANFAAT PENELITIAN.....	5
F. KEASLIAN PENELITIAN.....	6
BAB II TINJAUAN PUSTAKA DAN LANDASAN TEORI.....	7
A. TINJAUAN PUSTAKA.....	7
B. LANDASAN TEORI.....	15
1. <i>Artificial Intelligence</i>	15
2. <i>Machine Learning</i>	16
3. <i>Deep Learning</i>	17
4. Convolutional Neural Network.....	18

5.	VGG-16	22
6.	<i>Extreme Gradient Boosting (XGBoost)</i>	23
7.	Kanker Payudara	24
8.	Evaluasi Kinerja.....	27
BAB III METODE PENELITIAN		29
A.	INPUT CITRA	29
B.	<i>PREPROCESSING DATA</i>	31
C.	EKSTRAKSI FITUR	31
D.	PEMBUATAN MODEL	31
E.	UJI COBA DAN ANALISIS HASIL	36
BAB IV HASIL DAN PEMBAHASAN		37
A.	HASIL	37
1.	Teknik <i>fine-tuning</i>	37
2.	Teknik <i>Extreme Gradient Boosting</i>	47
B.	PEMBAHASAN	55
BAB V PENUTUP		59
A.	KESIMPULAN	59
B.	SARAN.....	59
DAFTAR PUSTAKA		60

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu.....	11
Tabel 2.2 <i>Confusion Matrix 3X3</i>	27
Tabel 3.1 Rincian <i>Hyperparameter Tuning</i>	33
Tabel 3.2 Parameter model XGBoost	36
Tabel 4.1 Deskripsi Dataset	38
Tabel 4.2 Parameter Model VGG-16	39
Tabel 4.3 Akurasi Pengujian Skenario pada Model VGG16	42
Tabel 4.4 Akurasi Pengujian Skenario Pertama pada Model XGBoost	49
Tabel 4.5 Akurasi Pengujian Skenario Kedua pada Model XGBoost.	50
Tabel 4.6 Perbedaan Teknik <i>Fine-tuning</i> dan XGBoost	56
Tabel 4.7 Akurasi Teknik <i>Fine-tuning</i> dan XGBoost.....	57
Tabel 4.8 Perbandingan dengan <i>State-of-The-Art</i> Model yang Diusulkan.....	58

STATE ISLAMIC UNIVERSITY
SUNAN KALIJAGA
YOGYAKARTA

DAFTAR GAMBAR

Gambar 2.1 Hubungan <i>Artificial Intelligence</i> , <i>Machine Learning</i> , dan <i>Deep Learning</i> (Zaharchuk et al., 2018).	16
Gambar 2.2 Struktur <i>Artificial Neural Networks</i>	18
Gambar 2.3 Arsitektur CNN (Sumber: Matworks).....	19
Gambar 2.4 Ilustrasi Proses pada Convolutional Layer.....	20
Gambar 2.5 Ilustrasi Aktivasi ReLU.....	20
Gambar 2.6 Ilustrasi Proses Pooling.....	21
Gambar 2.7 Arsitektur VGG-16.....	23
Gambar 2.8 Jumlah Kasus Baru Penyakit Kanker di Dunia	25
Gambar 2.9 Jumlah Kasus Baru Penyakit Kanker di Indonesia	26
Gambar 3.1 Tahap Penelitian	29
Gambar 3.2 Dataset MIAS.....	30
Gambar 3.3 Alur Pemrosesan Citra dengan Teknik <i>Fine-tuning</i>	32
Gambar 3.4 Skenario <i>Fine-tuning</i>	33
Gambar 3.5 Alur Pemrosesan Citra dengan Teknik <i>Extreme Gradient Boosting</i>	35
Gambar 4.1 Grafik VGG-16 proses latih dan validasi, (a) akurasi, (b) loss.....	44
Gambar 4.2 <i>Confusion Matrix</i> Model VGG-16	45
Gambar 4.3 Classification Report Model VGG-16.....	46
Gambar 4.4 Grafik Model XGBoost Proses Latih dan Validasi, (a) Akurasi, (b) Loss.	52
Gambar 4.5 <i>Confusion Matrix</i> Model XGBoost.....	53
Gambar 4.6 <i>Classification Report</i> Model XGBoost	54

BAB I

PENDAHULUAN

A. LATAR BELAKANG

Kanker merupakan penyebab utama kematian di seluruh dunia, dimana kurang lebih 10 juta kematian atau satu dari enam kematian disebabkan oleh kanker (WHO, 2022). WHO mencatat tahun 2022 kasus baru kanker mencapai 19,9 juta kasus di dunia dengan angka kematian mencapai 48,7% atau sebanyak 9,7 juta kasus. Dari jenis kanker yang ada, kanker payudara merupakan jenis kanker dengan kasus terbanyak kedua setelah kanker paru-paru, dengan angka kematian mencapai 6,8% (Globocan, 2022). Kanker payudara merupakan kanker non-kulit yang biasa menyerang wanita (Hadjiiski et al., 2006). Kanker payudara tumbuh ketika sel-sel di jaringan payudara menjadi abnormal dan berkembang secara tidak terkendali. Sel-sel ini membentuk gumpalan besar, yang menyebabkan munculnya kanker (Balakumaran et al., 2010). Kanker payudara dapat diklasifikasikan menjadi dua jenis, yaitu kanker jinak (*benign*) dan kanker ganas (*malignant*). Kanker jinak tersusun atas sel-sel yang berkembang secara lokal dan tidak menyebar. Sebaliknya kanker ganas tersusun dari sel-sel kanker yang berkembang biak tak terkendali dan dapat menyebar serta menyerang jaringan lainnya (Mehdy et al., 2017).

Diagnosis dini kanker sangat penting dalam keberhasilan pengobatan. Oleh karena itu, teknik pencitraan telah dikembangkan untuk meningkatkan kemungkinan diagnosis dini kanker payudara. Berbagai metode pencitraan, termasuk pencitraan resonansi magnetik (MRI), ultrasonografi (USG) dan mamografi, digunakan

untuk mendiagnosis kanker payudara (Mehdy et al., 2017). Di antara metode tersebut, mamografi relatif murah, sederhana, dan banyak digunakan sebagai tes skrining kanker payudara (Li et al., 2016).

Mamografi sinar-X menjadi metode yang sering digunakan oleh ahli radiologi untuk mendiagnosis kanker payudara serta menentukan adanya massa kanker dan kista. Akan tetapi, mamografi gambar sulit untuk ditafsirkan. Menurut *National Cancer Center* di Amerika Serikat, 10 sampai 30 persen massa payudara tidak terlihat oleh ahli radiologi (Wallis et al., 1991). Massa mikroklasifikasi yang merupakan partikel yang sangat kecil, merupakan tanda dan gejala kanker pada citra mamografi. Oleh karena itu sangat sulit untuk mendiagnosis gejala ini dengan benar, sehingga ketepatan diagnosis yang diberikan sangat dipengaruhi dari pengalaman dokter. Tanpa alat tambahan untuk mendiagnosis, dokter hanya memiliki tingkat akurasi antara 65-80% (Brinker et al., 2018). Dengan menggunakan bantuan *deep learning* tingkat akurasi diagnosis menjadi semakin tinggi dan akurat (Saber et al., 2021).

Pendekatan klasik seperti *pre-processing*, *segmentation*, *feature extraction*, dan *classification* telah digunakan oleh sebagian besar metode penelitian berbasis *machine learning* sebelum tahun 2016. Keahlian khusus diperlukan untuk proses pembelajaran mesin klasik yang sangat kompleks ini, terutama dalam hal proses ekstraksi dan pemilihan fitur yang membutuhkan waktu lama. Kesalahan dan hilangnya data selama proses ini juga memengaruhi kualitas klasifikasi (Brinker et al., 2018). Metode *deep learning* dapat menghasilkan representasi fitur yang lebih baik untuk menangani masalah klasifikasi citra dan lokalisasi objek, mengatasi masalah dengan metode *machine learning* klasik

(Saber et al., 2021). Algoritma *deep learning* paling populer yang diusulkan dalam literatur adalah *Convolutional Neural Network* (CNN), yang secara khusus diubah dengan struktur input gambar 2D (Lotter et al., 2019).

CNN membutuhkan sebuah data latih (*data training*) dengan jumlah besar sebagai proses *learning* sebelum menghasilkan sebuah hasil. Akan tetapi data citra medis yang tersedia sangat terbatas, sehingga dengan menggunakan teknik *transfer learning* dari dataset imagenet dan melakukan *fine-tuning* dapat meningkatkan akurasi klasifikasi dan mempercepat proses pembelajaran (Saber et al., 2021). Dengan memanfaatkan model CNN yang telah dilatih sebelumnya, fitur dari citra mamografi dapat diekstraksi secara otomatis dan *classifier* dapat dilatih menggunakan fitur-fitur tersebut (Guan S & Loew M, 2017). Dalam penelitiannya, Guan dan Loew menggunakan tiga teknik berbeda, yaitu melatih CNN dari awal (*full training*), melatih CNN untuk ekstraksi fitur, dan melakukan *fine-tuning* pada model CNN yang telah dilatih sebelumnya. Metode kedua memberikan tingkat akurasi yang paling tinggi, dengan perbedaan nilai akurasi sebesar 0,008 dari model *fine-tuning*.

Pre-trained CNN mencakup berbagai model seperti ResNet-50, VGG-16, dan VGG-19 yang digunakan dalam penelitian klasifikasi citra histologi kanker payudara. Dalam penelitian tersebut, ketiga arsitektur CNN dibandingkan dengan dua skenario, yaitu *training scratch* dan *transfer learning*. Akurasi terbaik mencapai 92,60% dari *pre-trained* VGG-16 dengan *fine-tuning* menggunakan *classifier logistic regression* (Shallu & Mehra, 2018). Selaras dengan penelitian tersebut, penerapan *pre-trained* CNN seperti V3, VGG-16, VGG-19, ResNet50, dan

InceptionV2ResNet untuk mendeteksi dan mengklasifikasikan kanker payudara menggunakan metode *transfer learning* memperlihatkan bahwa model VGG-16 menghasilkan performa terbaik (Saber et al., 2021).

Metode deep learning lainnya yang sering digunakan adalah XGBoost (*Extreme Gradient Boosting*). XGBoost merupakan salah satu implementasi paling populer dan efisien dari *Gradient Boosted Trees*, sebuah metode *ensemble learning* yang didasarkan pada perkiraan fungsi dengan mengoptimalkan *loss function* dan menerapkan beberapa teknik regularisasi. XGBoost bekerja dengan menggabungkan beberapa pohon keputusan yang dilatih secara berurutan, di mana setiap pohon baru berfokus pada kesalahan yang dibuat oleh pohon sebelumnya (Chen & Guestrin, 2016). *Extreme Gradient Boosting* (XGBoost) untuk diagnosis kanker payudara memiliki akurasi lebih tinggi dibandingkan dengan *Support Vector Machine* (SVM), dengan akurasi mencapai 96,95% (Ravly Andryan et al., 2022).

Penelitian ini melakukan optimasi dalam klasifikasi kanker payudara menjadi tiga kelas (*benign*, *malignant*, dan *normal*) menggunakan *deep transfer learning* dengan arsitektur CNN VGG-16 sebagai *feature extractor* dan menggunakan dua metode yaitu *fine-tuning* dan XGBoost. Dataset yang digunakan berasal dari *Mammographic Image Analysis Society* (MIAS) yang digunakan untuk menguji keefektifan metode yang disarankan untuk klasifikasi kanker payudara.algoritma

B. RUMUSAN MASALAH

Berdasarkan uraian dari latar belakang, rumusan masalah pada penelitian ini adalah bagaimana melakukan klasifikasi kanker payudara dengan akurasi yang lebih optimal pada dataset *Mammographic Image Analysis Society* (MIAS) menggunakan metode *deep transfer learning* dengan membandingkan metode *fine-tuning* dan XGBoost dengan *pre-trained model* VGG-16.

C. BATASAN MASALAH

Batasan masalah pada penelitian ini adalah penggunaan dataset citra mamografi dari dataset *Mammographic Image Analysis Society* (MIAS) yang mencakup tiga jenis yaitu ganas (*malignant*), jinak (*benign*), dan normal.

D. TUJUAN PENELITIAN

Berdasarkan rumusan masalah dan batasan masalah tersebut, tujuan penelitian ini adalah melakukan optimasi klasifikasi kanker payudara pada dataset *Mammographic Image Analysis Society* (MIAS) menggunakan metode *deep transfer learning* dengan membandingkan metode *fine-tuning* dan XGBoost dengan *pre-trained model* VGG-16.

E. MANFAAT PENELITIAN

Penelitian ini berguna untuk menyempurnakan metode yang sudah ada untuk klasifikasi kanker payudara, sehingga dapat menjadi salah satu acuan bagi penelitian-penelitian selanjutnya.

F. KEASLIAN PENELITIAN

Berdasarkan tinjauan Pustaka, penelitian terkait klasifikasi kanker payudara menggunakan dataset *Mammographic Image Analysis Society* (MIAS) dengan metode *deep transfer learning* dan arsitektur VGG-16 serta dengan membandingkan teknik *fine-tuning* dan XGBoost belum pernah dilakukan.



BAB V

PENUTUP

A. KESIMPULAN

Berdasarkan hasil penelitian dalam tesis ini, penggunaan metode *deep transfer learning* dengan memanfaatkan arsitektur CNN VGG-16 menunjukkan hasil yang signifikan untuk mengklasifikasikan citra mamografi ke dalam tiga kelas (*benign*, *malignant*, dan *normal*). Dua teknik pemodelan utama yang diterapkan yaitu *fine-tuning* dan XGBoost telah diuji menggunakan dataset *Mammographic Image Analysis Society*. Hasil pengujian menunjukkan bahwa akurasi terbaik pada teknik *fine-tuning* sebesar 96.02% dengan *optimizer* RMSprop, *batch size* 64, *learning rate* 0.01, dan *epoch* 100. Sedangkan pada teknik XGBoost diperoleh akurasi sebesar 93.03% dengan *max depth* 8, *learning rate* 0.4, dan *num round* 500. Penambahan layer pada teknik *fine-tuning* memungkinkan model lebih spesifik terhadap dataset baru tanpa kehilangan kemampuan generalisasi, sehingga hasil akurasinya lebih unggul dari teknik XGBoost.

B. SARAN

Untuk penelitian selanjutnya, disarankan agar dilakukan eksplorasi lebih lanjut terhadap berbagai arsitektur CNN selain VGG-16, seperti ResNet, Inception dan yang lainnya untuk membandingkan dengan VGG-16 atau mungkin meningkatkan akurasi klasifikasi kanker payudara. Selain itu, dapat mempertimbangkan penggunaan teknik augmentasi data yang lebih ekstensif, penyeimbangan data serta variasi parameter yang lebih luas dalam proses *fine-tuning* dan XGBoost untuk memperoleh hasil yang lebih optimal.

DAFTAR PUSTAKA

- ACS. (2022). *Breast Cancer*. American Cancer Society. <https://www.cancer.org/cancer/breast-cancer.html>
- Alanazi, S. A., Kamruzzaman, M. M., Islam Sarker, M. N., Alruwaili, M., Alhwaiti, Y., Alshammari, N., & Siddiqi, M. H. (2021). Boosting Breast Cancer Detection Using Convolutional Neural Network. *Journal of Healthcare Engineering*, 2021. <https://doi.org/10.1155/2021/5528622>
- Alzubi, J., Nayyar, A., & Kumar, A. (2018). Machine Learning from Theory to Algorithms: An Overview. *Journal of Physics: Conference Series*, 1142(1). <https://doi.org/10.1088/1742-6596/1142/1/012012>
- Anand, V., Gupta, S., Altameem, A., Nayak, S. R., Poonia, R. C., & Saudagar, A. K. J. (2022). An Enhanced Transfer Learning Based Classification for Diagnosis of Skin Cancer. *Diagnostics*, 12(7). <https://doi.org/10.3390/diagnostics12071628>
- Balakumaran, T., Vennila, I., & Shankar, C. G. (2010). Detection of Microcalcification in Mammograms Using Wavelet Transform and Fuzzy Shell Clustering. In *IJCSIS) International Journal of Computer Science and Information Security* (Vol. 7, Issue 1). <http://sites.google.com/site/ijcsis/>
- Bokade, A., & Shah, A. (2021). Breast Mass Classification with Deep Transfer Feature Extractor Model and Random Forest Classifier. *2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, 634–641. <https://doi.org/10.1109/RTEICT52294.2021.9573909>
- Brinker, T. J., Hekler, A., Utikal, J. S., Grabe, N., Schadendorf, D., Klode, J., Berking, C., Steeb, T., Enk, A. H., & Von Kalle, C. (2018). Skin cancer classification using convolutional neural networks: Systematic review. In *Journal of Medical Internet Research* (Vol. 20, Issue 10). JMIR Publications Inc. <https://doi.org/10.2196/11936>
- Cai, X., Li, X., Razmjooy, N., & Ghadimi, N. (2021). Breast Cancer Diagnosis by Convolutional Neural Network and Advanced Thermal Exchange Optimization Algorithm. *Computational and*

- Mathematical Methods in Medicine*, 2021.
<https://doi.org/10.1155/2021/5595180>
- Charan, S., Khan, M. J., & Khurshid, K. (2018). Breast cancer detection in mammograms using convolutional neural network. *International Conference on Computing, Mathematics and Engineering Technologies (ICoMET)*.
<https://doi.org/10.1109/ICOMET.2018.8346384>
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 13-17-August-2016*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Demir, F., Sengur, A., & Bajaj, V. (2020). Convolutional neural networks based efficient approach for classification of lung diseases. *Health Information Science and Systems*, 8(1).
<https://doi.org/10.1007/s13755-019-0091-3>
- Eissa, H., Telesi, N., & Mansour, S. (2022). *Comparison study of Commonly Used Activation Functions for Deep Neural Networks*.
<https://ajbas.academy.edu.ly/ar/j/issue-articles/578/download>
- Globacon. (2022). *The Global Cancer Observatory-Indonesia*.
<https://gco.iarc.who.int/media/globocan/factsheets/populations/360-indonesia-fact-sheet.pdf>
- Globocan. (2022). *The Global Cancer Observatory-World*.
<https://gco.iarc.who.int/media/globocan/factsheets/populations/900-world-fact-sheet.pdf>
- Guan S, & Loew M. (2017). *Breast Cancer Detection Using Transfer Learning in Convolutional Neural Networks*.
- Hadjiiski, L., Sahiner, B., Helvie, M. A., Chan, H. P., Roubidoux, M. A., Paramagul, C., Blane, C., Petrick, N., Bailey, J., Klein, K., Foster, M., Patterson, S. K., Adler, D., Nees, A. V., & Shen, J. (2006). Breast masses: Computer-aided diagnosis with serial mammograms. *Radiology*, 240(2), 343–356.
<https://doi.org/10.1148/radiol.2401042099>
- Hanin, M. A., Patmasari, R., Yunendah, R., & Fu'adah, N. (2021). SISTEM KLASIFIKASI PENYAKIT KULIT MENGGUNAKAN CONVOLUTIONAL NEURAL NETWORK (CNN) SKIN DISEASE CLASSIFICATION SYSTEM USING

- CONVOLUTIONAL NEURAL NETWORK (CNN). *E-Proceeding of Engineering*, 273–281.
- Hasan, M. R., Fatemi, M. I., Monirujjaman Khan, M., Kaur, M., & Zaguia, A. (2021). Comparative Analysis of Skin Cancer (Benign vs. Malignant) Detection Using Convolutional Neural Networks. *Journal of Healthcare Engineering*, 2021. <https://doi.org/10.1155/2021/5895156>
- Hijazi, S., Kumar, R., & Rowen, C. (2015). *Using Convolutional Neural Networks for Image Recognition*. www.cadence.com
- Jadoon, M. M., Zhang, Q., Haq, I. U., Butt, S., & Jadoon, A. (2017). Three-Class Mammogram Classification Based on Descriptive CNN Features. *BioMed Research International*, 2017. <https://doi.org/10.1155/2017/3640901>
- Jafari, Z., & Karami, E. (2023). Breast Cancer Detection in Mammography Images: A CNN-Based Approach with Feature Selection. *Information (Switzerland)*, 14(7). <https://doi.org/10.3390/info14070410>
- Khasanah, N., Komarudin, R., Afni, N., Maulana, Y. I., & Salim, A. (2021). Skin Cancer Classification Using Random Forest Algorithm. *SISFOTENIKA*, 11(2), 137. <https://doi.org/10.30700/jst.v11i2.1122>
- Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Li, H., Zhang, S., Wang, Q., & Zhu, R. (2016). Clinical value of mammography in diagnosis and identification of breast mass. *Pakistan Journal of Medical Sciences*, 32(4). <https://doi.org/10.12669/pjms.324.9384>
- Lotter, W., Diab, A. R., Haslam, B., Kim, J. G., Grisot, G., Wu, E., Wu, K., Onieva, J. O., Boxerman, J. L., Wang, M., Bandler, M., Vijayaraghavan, G., & Sorensen, A. G. (2019). *Robust breast cancer detection in mammography and digital breast tomosynthesis using annotation-efficient deep learning approach*. <http://arxiv.org/abs/1912.11027>
- Mahesh, B. (2020). Machine Learning Algorithms - A Review. *International Journal of Science and Research (IJSR)*, 9(1), 381–386. <https://doi.org/10.21275/ART20203995>

- Mehdy, M. M., Ng, P. Y., Shair, E. F., Saleh, N. I. M., & Gomes, C. (2017). Artificial neural networks in image processing for early detection of breast cancer. In *Computational and Mathematical Methods in Medicine* (Vol. 2017). Hindawi Limited. <https://doi.org/10.1155/2017/2610628>
- Mohammed, M., Khan, M. B., & Bashie, E. B. M. (2017). *Machine learning: Algorithms and applications*.
- Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). Fundamentals of Artificial Neural Networks and Deep Learning. In: Multivariate Statistical Machine Learning Methods for Genomic Prediction. In *Multivariate Statistical Machine Learning Methods for Genomic Prediction*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-89010-0>
- Nugraha, W. (2021). PREDIKSI PENYAKIT JANTUNG CARDIOVASCULAR MENGGUNAKAN MODEL ALGORITMA KLASIFIKASI. *JURNAL SIGMATA*, 9(2), 78–84. <https://www.kaggle.com/andrewmvd/heart->
- Pannakkong, W., Thiwa-Anont, K., Singthong, K., Parthanadee, P., & Buddhakulsomsiri, J. (2022). Hyperparameter Tuning of Machine Learning Algorithms Using Response Surface Methodology: A Case Study of ANN, SVM, and DBN. *Mathematical Problems in Engineering*, 2022. <https://doi.org/10.1155/2022/8513719>
- Powers, D. M. W. (2020). *EVALUATION: FROM PRECISION, RECALL AND F-MEASURE TO ROC, INFORMEDNESS, MARKEDNESS & CORRELATION*.
- Ravly Andryan, M., Fajri, M., Nina Sulistyowati, dan, Informatika, T., Singaperbangsa Karawang Jl Ronggo Waluyo, U. H., Timur, T., & Barat, J. (2022). KOMPARASI KINERJA ALGORITMA XGBOOST DAN ALGORITMA SUPPORT VECTOR MACHINE (SVM) UNTUK DIAGNOSA PENYAKIT KANKER PAYUDARA. *Jurnal Informatika Dan Komputer*, 6(1), 1–5.
- Russell, S., & Norvig, P. (2010). *Artificial Intelligence: A Modern Approach* (Third).
- Saber, A., Sakr, M., Abo-Seida, O. M., Keshk, A., & Chen, H. (2021). A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning

- Technique. *IEEE Access*, 9, 71194–71209. <https://doi.org/10.1109/ACCESS.2021.3079204>
- Safutra, A. R., & Prabowo, D. W. (2016). *DIAGNOSIS PENYAKIT KANKER PAYUDARA MENGGUNAKAN METODE NAIVE BAYES BERBASIS DESKTOP*.
- Sanjaya, J., & Ayub, M. (2020). Augmentasi Data Pengenalan Citra Mobil Menggunakan Pendekatan Random Crop, Rotate, dan Mixup. *Jurnal Teknik Informatika Dan Sistem Informasi*, 6(2). <https://doi.org/10.28932/jutisi.v6i2.2688>
- Shallu, & Mehra, R. (2018). Breast cancer histology images classification: Training from scratch or transfer learning? *ICT Express*, 4(4), 247–254. <https://doi.org/10.1016/j.icte.2018.10.007>
- Sholihin, R. (2017). *Mengenal, Mencegah, Mengatasi Silent Killer "Kanker"* (Ika Wijayati S.N, Ed.). Romawi Pustaka.
- Sonar, P., Bhosle, U., & Choudhury, C. (2017). *Mammography Classification Using Modified Hybrid SVM-KNN*.
- Tamina, S. (2019). Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images. *International Journal of Scientific and Research Publications (IJSRP)*, 9(10), p9420. <https://doi.org/10.29322/ijsrp.9.10.2019.p9420>
- Taninaga, J., Nishiyama, Y., Fujibayashi, K., Gunji, T., Sasabe, N., Iijima, K., & Naito, T. (2019). Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study. *Scientific Reports*, 9(1). <https://doi.org/10.1038/s41598-019-48769-y>
- Thenmozhi, K., & Srinivasulu Reddy, U. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture*, 164. <https://doi.org/10.1016/j.compag.2019.104906>
- Ubaidillah, R., Muliadi, M., Nugrahadi, D. T., Faisal, M. R., & Herteno, R. (2022). Implementasi XGBoost Pada Keseimbangan Liver Patient Dataset dengan SMOTE dan Hyperparameter Tuning Bayesian Search. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, 6(3), 1723. <https://doi.org/10.30865/mib.v6i3.4146>

- Wallis, M. G., Walsh, M. T., & Lee, J. R. (1991). A Review of False Negative Mammography Population in a Symptomatic. In *Clinical Radiology* (Vol. 44).
- WHO. (2021). *Breast cancer*. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
- WHO. (2022). *cancer*. <https://www.who.int/news-room/fact-sheets/detail/cancer>
- Wlodarczak, P. (2020). *Machine Learning and its Applications; Edition 1*. CRC Press. <https://doi.org/10.4324/9780429448782>
- Zaharchuk, G., Gong, E., Wintermark, M., Rubin, D., & Langlotz, C. P. (2018). Deep learning in neuroradiology. In *American Journal of Neuroradiology* (Vol. 39, Issue 10, pp. 1776–1784). American Society of Neuroradiology. <https://doi.org/10.3174/ajnr.A5543>