DETEKSI KANKER PAYUDARA BERDASARKAN SINTESIS CITRA ULTRASOUND MENGGUNAKAN WASSERSTEIN GENERATIVE ADVERSARIAL NETWORKS

SKRIPSI

Oleh:

SAYYIDAH HUMAIRAH

20081010047

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR

2024

DETEKSI KANKER PAYUDARA BERDASARKAN SINTESIS CITRA ULTRASOUND MENGGUNAKAN WASSERSTEIN GENERATIVE ADVERSARIAL NETWORKS

SKRIPSI

Diajukan untuk memenuhi persyaratan dalam memperoleh gelar Sarjana Komputer Program Studi Informatika

Oleh:

SAYYIDAH HUMAIRAH

20081010047

PROGRAM STUDI INFORMATIKA

FAKULTAS ILMU KOMPUTER

UNIVERSITAS PEMBANGUNAN NASIONAL "VETERAN" JAWA TIMUR

2024

LEMBAR PENGESAHAN SKRIPSI

JUDUL: DETEKSI KANKER PAYUDARA BERDASARKAN SINTESIS CITRA ULTRASOUND MENGGUNAKAN WASSERSTEIN GENERATIVE ADVERSARIAL NETWORKS.

Oleh : SAYYIDAH HUMAIRAH

NPM : 20081010047

Telah Diseminarkan Dalam Ujian Skripsi Pada:

Hari Senin, Tanggal 20 Mei 2024

Mengetahui

Dosen Pembimbing

1. **Dr. I Gede Susrama Mas Diyasa, ST., MT.**
   NIP. 19700619 2021211 009
2. **Eva Yulia Puspaningrum, S.Kom., M.Kom**
   NIP. 19890705 2021212 002

Dosen Penguji

1. **Eka Prakarsa Mandyartha, ST., M.Kom**
   NIP. 19880525 2018031 001
2. **Wahyu Syaifullah JS., S.Kom., M.Kom.**
   NIP. 19860825 2021211 003

Menyetujui

Dekan
Fakultas Ilmu Komputer

Prof. Dr. Ir. Novirina Hendrasarie, MT
NIP. 19681126 199403 2 001

Koordinator Program Studi
Informatika

Fetty Tri Anggraeny, S.Kom., M.Kom
NIP. 19820211 2021212 005

Dipindai dengan CamScanner

SURAT PERNYATAAN BEBAS DARI PLAGIASI

Saya, mahasiswa Program Studi Sarjana Informatika Universitas Pembangunan Nasional "Veteran" Jawa Timur, yang bertanda tangan di bawah ini:

Nama : Sayyidah Humairah

NPM : 20081010047

Menyatakan dengan sesungguhnya bahwa Skripsi/Tugas Akhir yang saya kerjakan berjudul:

"DETEKSI KANKER PAYUDARA BERBASIS SINTESIS CITRA ULTRASOUND MENGGUNAKAN WASSERSTEIN GENERATIVE ADVERSARIAL NETWORKS"

bukan merupakan plagiasi sebagian atau keseluruhan dari Skripsi/Tugas Akhir/Penelitian orang lain dari juga bukan merupakan produk dan software yang saya beli dari pihak lain. Saya juga menyatakan bahwa Skripsi/Tugas Akhir ini secara keseluruhan adalah pekerjaan Saya sendiri, kecuali yang dinyatakan dalam Daftar Pustaka dan tidak pernah diajukan untuk syarat memperoleh gelar di Universitas Pembangunan Nasional "Veteran" Jawa Timur maupun di Institut Pendidikan lain. Bukti hasil pengecekan plagiasi dokumen ini dapat ditelusuri melalui QR Code di bawah.

Apabila di kemudian hari terbukti bahwa dokumen ini merupakan plagiasi karya orang lain, saya sanggup menerima sanksi sesuai aturan yang berlaku.

Demikian atas perhatiannya disampaikan terima kasih.

Surabaya, 30 Mei 2024

Hormat saya,

Sayyidah Humairah

NPM. 20081010047

Dipindai dengan CamScanner

DETEKSI KANKER PAYUDARA BERBASIS SINTESIS CITRA ULTRASOUND MENGGUNAKAN WASSERSTEIN GENERATIVE ADVERSARIAL NETWORKS

Nama Mahasiswa : Sayyidah Humairah

NPM : 20081010047

Program Studi : Informatika

Dosen Pembimbing : Dr. I Gede Susrama Mas Diyasa, S.T., M.T., IPU

Eva Yulia Puspaningrum, S.Kom., M.Kom.

ABSTRAK

Kanker payudara menjadi salah satu kanker dengan penderita paling banyak di Indonesia. Salah satu cara untuk deteksi dini kanker payudara adalah dengan memanfaatkan citra ultrasound untuk menklasifikasikan kondisi payudara menjadi normal, tumor jinak, atau kanker. Adanya teknologi deep learning memungkinkan proses analisis citra dapat dilakukan, seperti dengan menggunakan Transfer Learning Convolutional Neural Network (CNN). Pada umumnya, model CNN memerlukan dataset yang besar dan kelas yang seimbang untuk dapat memberikan performa yang baik dalam melakukan klasifikasi, namun dataset medis seperti citra ultrasound payudara cenderung terbatas dan memiliki kelas yang tidak seimbang.

Wasserstein GAN merupakan salah satu metode augmentasi data secara generatif yang dapat menghasilkan citra sintesis dengan mempelajari pola distribusi data citra asli. Dengan menggunakan WGAN, maka ukuran data citra asli dapat dikembangkan sebanyak yang diinginkan sehingga dapat mengatasi masalah ketidakseimbangan kelas dan terbatasnya data medis citra ultrasound.

Hasil implementasi WGAN pada dataset citra ultrasound memberikan dampak peningkatan akurasi pada proses klasifikasi oleh model transfer learning dengan model terbaik adalah VGG16 yang dilatih pada dataset yang telah dikembangkan menjadi 403 data citra per kelas dengan akurasi mencapai 83.33%. Selain itu,

integrasi model dengan sebuah system diagnosis memungkinkan pengguna awam dapat memanfaatkan teknologi ini untuk deteksi dini kanker payudara.

Kata kunci: *Wasserstein GAN, Kanker Payudara, Transfer Learning, CNN*

KATA PENGANTAR

Assalaamualaikum Wr. Wb.

Segala puji dan syukur kami panjatkan kepada Allah SWT. yang telah melimpahkan karunia, rahmat, dan hidayah-Nya sehingga peulis dapat menyelesaikan penyusunan skripsi yang berjudul “Deteksi Kanker Payudara Berdasarkan Sintesis Citra Ultrasound Menggunakan Wasserstein Generative Adversarial Networks”. Penyusunan skripsi ini dilakukan sebagai salah satu persyaratan dalam menyelesaikan program studi S1 Informatika, Fakultas Ilmu Komputer, Universitas Pembangunan Nasional “Veteran” Jawa Timur.

Isu keterbatasan data citra *ultrasound* di tengah potensi teknologi *deep learning* dalam melakukan prediksi klasifikasi dini kondisi payudara menjadikan augmentasi data citra sebuah urgensi. Penelitian ini bertujuan untuk mengimplementasikan teknik augmentasi generative menggunakan Wasserstein GAN untuk melakukan augmentasi data citra *ultrasound* payudara, sehingga dapat meningkatkan akurasi prediksi kanker yang dilakukan oleh *classifier CNN transfer learning*. Dengan adanya penelitian ini, diharapkan dapat membantu pembuatan sistem diagnosis kanker payudara yang lebih akurat.

Akhir kata, penulis menyadari bahwa dalam penyusunan penelitian ini masih terdapat beberapa kekurangan. Oleh karena itu, penulis mengharapkan kritik dan saran pembaca untuk perbaikan dan penyempurnaan penelitian ini. Sekian.

Terima kasih.

Wassalaamualaikum Wr. Wb.

Surabaya, 31 Mei 2024

Penulis

UCAPAN TERIMA KASIH

Alhamdulillaah, puji syukur dan terima kasih dipanjatkan kepada Allah SWT., atas anugerah dan pertolongan-Nya penulis dapat menyelesaikan skripsi yang berjudul “Deteksi Kanker Payudara Berdasarkan Sintesis Citra Ultrasound Menggunakan Wasserstein Generative Adversarial Networks”. Penulis juga menyampaikan ucapan terima kasih kepada:

1. Kedua orang tua tercinta, ibuk dan abah, yang senantiasa menyayangi, menjaga, mendoakan, mendidik, membimbing, mendukung dalam segala bentuk, mengapresiasi, memberikan pengorbanan & kerja keras, dan memotivasi penulis. Gelar sarjana ini penulis persembahkan untuk ibuk dan abah.
2. Prof. Dr. Ir. Akhmad Fauzi, M.MT selaku Rektor Universitas Pembangunan Nasional “Veteran” Jawa Timur.
3. Prof. Dr. Ir. Novirina Hendrasarie, MT., selaku Dekan Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur.
4. Fetty Tri Anggraeny, S.Kom., M.Kom., selaku koordinator program studi Informatika Universitas Pembangunan Nasional “Veteran” Jawa Timur.
5. Dr. I Gede Susrama Mas Diyasa, ST. MT., selaku dosen pembimbing pertama dan Wakil Dekan I Fakultas Ilmu Komputer Universitas Pembangunan Nasional “Veteran” Jawa Timur, atas bantuan, kritik, dan saran yang membangun dalam penyelesaian skripsi ini.
6. Eva Yulia Puspaningrum, S.Kom., M.Kom., selaku dosen pembimbing kedua dan dosen wali, yang telah memberikan kritik dan saran yang membangun dalam penyelesaian skripsi ini.
7. Eka Prakarsa Mandyartha, ST., M.Kom., selaku dosen penguji pertama yang telah memberikan masukan dan penilaian terhadap penelitian skripsi ini.
8. Wahyu Syaifullah JS, S.Kom., M.Kom., selaku dosen penguji kedua yang telah memberikan masukan dan penilaian terhadap penelitian skripsi ini.

9. Bapak/Ibu dosen yang telah membagikan ilmu pengetahuan dan pengalamannya kepada penulis selama menempuh pendidikan Sarjana di Universitas Pembangunan Nasional “Veteran” Jawa Timur.
10. Adik-adik tercinta Sany, Habibie, dan Yasmin, atas canda tawa & suka duka yang kita saling bagi selama pengerjaan skripsi ini. Tak lupa, keluarga besar lainnya yang selalu memberikan doa, bantuan, dan dukungan kepada penulis.
11. Teman-teman seperjuangan di lingkungan kampus Universitas Pembangunan Nasional “Veteran” Jawa Timur, khususnya Abik, Endin, & Icha, yang membantu dan mendukung penulis dalam penyelesaian skripsi ini.
12. Teman-teman seperjuangan di luar lingkup kampus, khususnya Diaz dan Ayu, yang telah mendukung, membantu, dan direpotkan selama penyelesaian skripsi ini.
13. Teman-teman *awardees* IISMA Twente 2022, atas kebaikan & pengalaman tak terlupakan selama IISMA. Memori itu telah menguatkan dan memotivasi penulis selama penyelesaian skripsi ini.
14. Pihak PT. Rezeki Mekanik Indonesia selaku mitra MSIB batch 4 & yang terlibat di dalamnya: Bapak Octav, kak Rizky dan yang lainnya. Terima kasih telah memberikan kesempatan bagi penulis untuk mempelajari topik GAN sehingga penulis dapat mengembangkannya menjadi penelitian skripsi ini. Tak lupa, teman-teman magang yang telah menemani penulis.
15. Pihak-pihak lain yang tidak dapat disebutkan satu per satu, yang telah membantu dalam penyelesaian skripsi ini. Terima kasih banyak.

Semoga Allah SWT memberikan balasa yang berlipat atas bantuan & kebaikan yang telah diberikan kepada penulis dalam menyelesaikan skripsi ini.

Surabaya, 31 Mei 2024

DAFTAR ISI

ABSTRAK iv
KATA PENGANTAR vi
UCAPAN TERIMA KASIH vii
DAFTAR ISI ix
DAFTAR GAMBAR xi
DAFTAR TABEL xiii
DAFTAR KODE PROGRAM xiv
PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 3
1.3 Batasan Masalah 3
1.4 Tujuan Penelitian 3
1.5 Manfaat Penelitian 4
TINJAUAN PUSTAKA 5
1.6 Penelitian Sebelumnya 5
1.7 Kanker Payudara dan Citra Ultrasound 7
1.8 Pengolahan Citra Digital 7
1.9 Convolutional Neural Network 8
1.10 Transfer Learning 10
1.11 Augmentasi Data 11
1.12 Wasserstein Generative Adversarial Networks 12
METODOLOGI PENELITIAN 16
1.13 Tahapan Penelitian 16
1.14 Pengumpulan Data 18
1.15 Preprocessing Data 18
1.16 Perancangan & Pelatihan Model Wasserstein GAN 19
1.17 Pelatihan Model CNN Transfer Learning untuk Klasifikasi 21
1.18 Pengujian & Evaluasi Model CNN Transfer Learning Berdasarkan Data Sintesis 22
1.19 Skenario Pengujian 24
1.20 Integrasi Model pada Aplikasi 25
HASIL DAN PEMBAHASAN 26
1.21 Implementasi Program 26

1.21.1 Pengunduhan dan Persiapan Dataset.....26
1.21.2 Model Wasserstein GAN.....29
1.21.3 Augmentasi Data dengan Wasserstein GAN.....41
1.21.4 Klasifikasi dengan CNN Transfer Learning.....46
1.22 Hasil Pengujian dan Pembahasan.....47
1.23 Penerapan Aplikasi pada Sistem Diagnosis.....62
KESIMPULAN DAN SARAN.....65
1.24 Kesimpulan.....65
1.25 Saran.....66
DAFTAR PUSTAKA.....67

DAFTAR GAMBAR

Gambar 2.1 Arsitektur GAN (Pan et al., 2020) 13
Gambar 3.1 Diagram Tahapan Penelitian 17
Gambar 3.2 Sampel Citra Tiap Kelas 18
Gambar 3.3 Sampel Citra Tiap Kelas setelah Preprocessing 19
Gambar 3.5 Layer dalam Model WGAN 20
Gambar 3.6 Alur Pelatihan Model CNN 22
Gambar 3.7 Alur Pengujian dan Evaluasi Model CNN 23
Gambar 4.1 Arsitektur Lapisan pada Generator 31
Gambar 4.2 Arsitektur Lapisan pada Discriminator 32
Gambar 4.3 Training Loss Data Class 0 36
Gambar 4.4 Training Loss Data Class 1 37
Gambar 4.5 Training Loss Data Class 2 37
Gambar 4.6 Citra Sintesis Kelas 0 Epoch ke-0 39
Gambar 4.7 Citra Sintesis Kelas 0 Epoch ke-2500 39
Gambar 4.8 Citra Sintesis Kelas 0 Epoch ke-5000 39
Gambar 4.9 Citra Sintesis Kelas 1 Epoch ke-0 39
Gambar 4.10 Citra Sintesis Kelas 1 Epoch ke-2500 40
Gambar 4.11 Citra Sintesis Kelas 1 Epoch ke-5000 40
Gambar 4.12 Citra Sintesis Kelas 2 Epoch ke-0 40
Gambar 4.13 Citra Sintesis Kelas 2 Epoch ke-2500 40
Gambar 4.14 Citra Sintesis Kelas 2 Epoch ke-5000 41
Gambar 4.15 Sampel Citra Sintesis Kelas 0 (Jinak) 44
Gambar 4.16 Sampel Citra Sintesis Kelas 1 (Normal) 45
Gambar 4.17 Sampel Citra Sintesis Kelas 2 (Ganas) 45
Gambar 4.18 Confusion Matrix VGG16 Dataset Original 48
Gambar 4.19 Confusion Matrix VGG16 Dataset Balanced 48
Gambar 4.20 Confusion Matrix VGG16 Dataset Expand 49
Gambar 4.21 Confusion Matrix ResNet50 Dataset Original 53
Gambar 4.22 Confusion Matrix ResNet50 Dataset Balanced 53
Gambar 4.23 Confusion Matrix ResNet50 Dataset Expand 54
Gambar 4.24 Confusion Matrix MobileNetV2 Dataset Original 57
Gambar 4.25 Confusion Matrix MobileNetV2 Dataset Balanced 58

Gambar 4.26 Confusion Matrix MobileNetV2 Dataset Expand.....58
Gambar 4.27 Halaman Utama Sistem Diagnosis.....63
Gambar 4.28 Hasil Prediksi pada Sistem Diagnosis.....64

DAFTAR TABEL

Tabel 3.1 Algoritma WGAN (Arjovsky et al., 2017) 19
Tabel 3.1 Confusion Matrix 23
Tabel 3.2 Performa Model CNN X dengan Data Training Berbeda 24
Tabel 4.1 Perbandingan Ukuran Dataset 46
Tabel 4.2 Penjelasan Confusion Matrix Gambar 4.18 48
Tabel 4.3 Penjelasan Confusion Matrix Gambar 4.19 49
Tabel 4.4 Penjelasan Confusion Matrix Gambar 4.20 49
Tabel 4.5 Performa Model VGG16 50
Tabel 4.6 Penjelasan Confusion Matrix Gambar 4.21 53
Tabel 4.7 Penjelasan Confusion Matrix Gambar 4.22 54
Tabel 4.8 Penjelasan Confusion Matrix Gambar 4.23 54
Tabel 4.9 Performa Model ResNet50 55
Tabel 4.10 Penjelasan Confusion Matrix Gambar 4.24 57
Tabel 4.11 Penjelasan Confusion Matrix Gambar 4.25 58
Tabel 4.12 Penjelasan Confusion Matrix Gambar 4.26 59
Tabel 4.13 Performa Model MobileNetV2 59

DAFTAR KODE PROGRAM

Kode Program 4.1 Persiapan Dataset.....27
Kode Program 4.2 Model Generator & Discriminator WGAN 30
Kode Program 4.3 Training WGAN 34
Kode Program 4.4 Generator WGAN.....42
Kode Program 4.5 Augmentasi Data Menggunakan WGAN 44
Kode Program 4.6 Klasifikasi dengan Transfer Learning.....47