

**KLASIFIKASI KANKER PAYUDARA BERDASARKAN CITRA
ULTRASOUND MENGGUNAKAN ALGORITMA *SUPPORT
VECTOR MACHINE (SVM)***

SKRIPSI

Diajukan untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

NURAZMI APRILIA

19.11.2884

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

**KLASIFIKASI KANKER PAYUDARA BERDASARKAN CITRA
ULTRASOUND MENGGUNAKAN ALGORITMA *SUPPORT*
*VECTOR MACHINE (SVM)***

SKRIPSI

untuk memenuhi salah satu syarat mencapai derajat Sarjana
Program Studi Informatika



disusun oleh

NURAZMI APRILIA

19.11.2884

Kepada

**FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA**

2023

HALAMAN PERSETUJUAN

SKRIPSI

**KLASIFIKASI KANKER PAYUDARA BERDASARKAN CITRA
ULTRASOUND MENGGUNAKAN ALGORITMA *SUPPORT VECTOR
MACHINE* (SVM)**

yang disusun dan diajukan oleh

Nurazmi Aprilia

19.11.2884

telah disetujui oleh Dosen Pembimbing Skripsi
pada tanggal 15 Agustus 2023

Dosen Pembimbing,



Rumini, M. Kom

NIK. 190302246

HALAMAN PENGESAHAN

SKRIPSI

**KLASIFIKASI KANKER PAYUDARA BERDASARKAN CITRA
ULTRASOUND MENGGUNAKAN ALGORITMA *SUPPORT VECTOR
MACHINE (SVM)***

yang disusun dan diajukan oleh

Nurazmi Aprilia

19.11.2884

Telah dipertahankan di depan Dewan Penguji
pada tanggal 15 Agustus 2023.

Susunan Dewan Penguji

Nama Penguji

Anna Baita, M.Kom

NIK. 190302290

Uyock Anggoro Saputro, M.Kom

NIK. 190302419

Rumini, M.Kom

NIK. 190302246

Tanda Tangan



Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal 15 Agustus 2023

DEKAN FAKULTAS ILMU KOMPUTER



Hanif Al Fatta, S.Kom., M.Kom.

NIK. 190302096

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Yang bertandatangan di bawah ini,

Nama mahasiswa : Nurazmi Aprilia
NIM : 19.11.2884

Menyatakan bahwa Skripsi dengan judul berikut:

**Klasifikasi Kanker Payudara Berdasarkan Citra *Ultrasound* Menggunakan
Algoritma *Support Vector Machine* (SVM)**

Dosen Pembimbing : Rumini, M.Kom

1. Karya tulis ini adalah benar-benar ASLI dan BELUM PERNAH diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya.
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian SAYA sendiri, tanpa bantuan pihak lain kecuali arahan dari Dosen Pembimbing.
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini.
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab SAYA, bukan tanggung jawab Universitas AMIKOM Yogyakarta.
5. Pernyataan ini SAYA buat dengan sesungguhnya, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka SAYA bersedia menerima SANKSI AKADEMIK dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi.

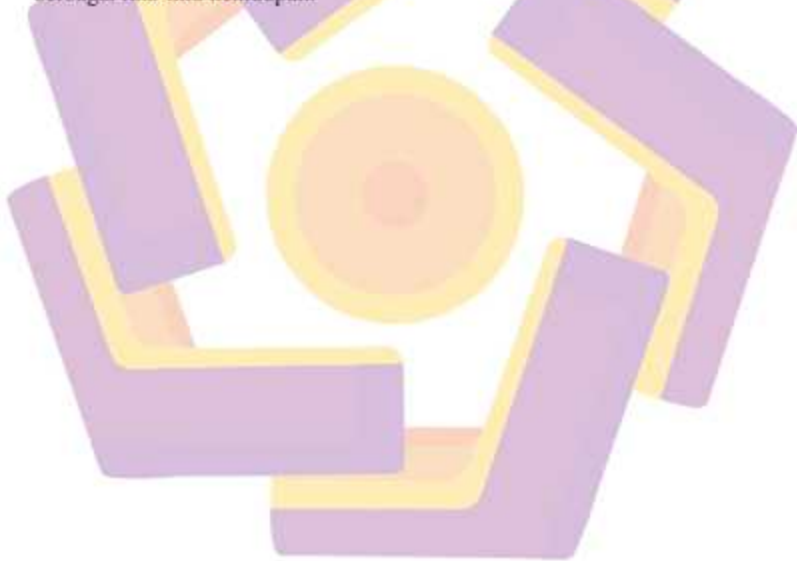
Yogyakarta, 15 Agustus 2023

Yang Menyatakan,


METERAL TEMPEL
E03AKK548847847
Nurazmi Aprilia

HALAMAN PERSEMBAHAN

Skripsi ini disusun sebagai wujud persembahan kepada orang-orang yang selama ini selalu mendampingi dan mendukung saya khususnya untuk keluarga dan terutama kepada kedua orang tua yang selalu memberikan dukungan tak terputuskan dalam perjalanan pendidikan ini. Serta dengan kerendahan hati, penulis ingin menyampaikan rasa terima kasih yang dalam kepada orang tua tercinta yang telah memberikan begitu banyak pengorbanan dalam membimbing dan mendukung pendidikan penulis hingga kini. Tanpa kehadiran dan dukungan keluarga terutama kedua orang tua, penulis tak akan mampu mencapai titik ini dalam menjalani berbagai lika-liku kehidupan.



KATA PENGANTAR

Segala puji bagi Allah SWT atas limpahan rahmat, petunjuk, dan karunia-Nya yang telah memungkinkan penulis menyelesaikan penyusunan skripsi ini berjudul "Klasifikasi Kanker Payudara Berdasarkan Citra *Ultrasound* Menggunakan Algoritma *Support Vector Machine* (SVM)". Skripsi ini disusun sebagai bagian penting dalam rangka memenuhi persyaratan kelulusan dari Program Studi Informatika Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.

Penyusunan skripsi ini tidak lepas dari bantuan, dukungan, dan bimbingan dari berbagai pihak. Oleh karena itu, penulis ingin mengucapkan terima kasih yang sebesar-besarnya kepada:

1. Ibu Rumini, M.Kom, selaku dosen pembimbing, atas dukungan, arahan, serta masukan berharga yang telah membimbing langkah-langkah penyusunan skripsi ini.
2. Keluarga tercinta, yang senantiasa memberikan dukungan moral, materi, doa, serta semangat dalam perjalanan panjang penyusunan skripsi ini.
3. Sahabat dan teman, yang telah berbagi ilmu, pengalaman, dan inspirasi sepanjang perjalanan akademik di Universitas Amikom Yogyakarta.
4. Semua individu yang turut menyumbangkan gagasan, pandangan, dan sokongan dalam berbagai bentuk, yang telah membantu kelancaran perjalanan penyelesaian skripsi ini.

Skripsi ini adalah hasil kerja keras dan dedikasi penulis, yang berusaha untuk menggali serta memahami suatu permasalahan dengan sebaik-baiknya. Akhir kata, meski kesempurnaan tak dapat dicapai, penulis berharap bahwa skripsi ini mampu memberikan kebermanfaatan dan inspirasi bagi para pembaca serta kontribusi dalam pengembangan ilmu pengetahuan di bidang *Machine Learning*.

Yogyakarta, 10 Agustus 2023

Penulis

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERSETUJUAN.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERNYATAAN KEASLIAN SKRIPSI	iv
HALAMAN PERSEMBAHAN	v
KATA PENGANTAR	vi
DAFTAR ISI.....	vii
DAFTAR TABEL	x
DAFTAR GAMBAR	xi
DAFTAR LAMPIRAN	xii
INTISARI	xiii
ABSTRACT	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	2
1.4 Tujuan Penelitian	3
1.5 Manfaat Penelitian	3
1.6 Sistematika Penulisan	3
BAB II TINJAUAN PUSTAKA	5
2.1 Studi Literatur	5
2.2 Dasar Teori	11
2.2.1. Kanker Payudara	11
2.2.2. Ekstraksi Fitur	12
2.2.3. <i>Gray Level Co-Occurrence Matrix (GLCM)</i>	12
2.2.4. <i>Synthetic Minority Over-sampling Technique (SMOTE)</i>	14

2.2.5.	<i>Mutual Information</i>	14
2.2.6.	<i>Machine Learning</i>	15
2.2.7.	Klasifikasi Citra Digital	15
2.2.8.	<i>Support Vector Machine (SVM)</i>	16
2.2.9.	Pengukuran Kinerja Model Klasifikasi	18
2.2.5.1.	<i>K-Fold Cross Validation</i>	18
2.2.5.2.	<i>Confusion Matrix</i>	19
BAB III METODE PENELITIAN		22
3.1.	Alur Penelitian	22
3.1.1.	Akuisisi Data	23
3.1.2.	Ekstraksi Fitur	23
3.1.3.	<i>Pre-Processing</i>	23
3.1.3.1.	<i>Balancing Data</i>	24
3.1.3.2.	Normalisasi Data	24
3.1.3.3.	Seleksi Fitur	24
3.1.4.	<i>Split Data</i>	25
3.1.5.	Klasifikasi	25
3.1.6.	Evaluasi Kinerja Model Klasifikasi	25
3.1.6.1.	<i>K-Fold Cross Validation</i>	26
3.1.6.2.	<i>Confusion Matrix</i>	26
3.1.7.	Analisis	26
3.2.	Alat	27
BAB IV HASIL DAN PEMBAHASAN		29
4.1.	Akuisisi Data	29
4.2.	Ekstraksi Fitur Menggunakan GLCM	29

4.3	<i>Pre-Processing Data</i>	33
4.3.1.	<i>SMOTE Balancing Data</i>	33
4.3.2.	<i>Min-Max Normalization</i>	35
4.3.3.	Seleksi Fitur <i>Mutual Information</i> dengan <i>SelectKBest</i>	35
4.4	<i>Train/Test Split</i>	37
4.5	Klasifikasi <i>Support Vector Machine (SVM)</i>	37
4.6	Evaluasi Kinerja Klasifikasi dengan <i>K-Fold Cross Validation</i>	39
4.7	Evaluasi Kinerja Klasifikasi dengan <i>Confusion Matrix</i>	41
4.8	Analisis.....	42
BAB V PENUTUP.....		44
5.1	Kesimpulan.....	44
5.2	Saran.....	44
REFERENSI.....		45
LAMPIRAN.....		49

DAFTAR TABEL

Tabel 2.1 Keaslian Penelitian	7
Tabel 4.1 Hasil Ekstraksi Citra <i>Benign</i>	30
Tabel 4.2 Hasil <i>Labelling</i>	31
Tabel 4.3 Hasil <i>SMOTE</i>	34
Tabel 4.4 <i>Train/test Split</i>	37
Tabel 4.5 Perbandingan Hasil Evaluasi Model Klasifikasi	38
Tabel 4.6 Perbandingan Hasil Evaluasi Menggunakan <i>K-Fold Cross Validation</i>	39
Tabel 4.7 Perbandingan Jenis Pengujian	40
Tabel 4.8 Perbandingan Penelitian	42

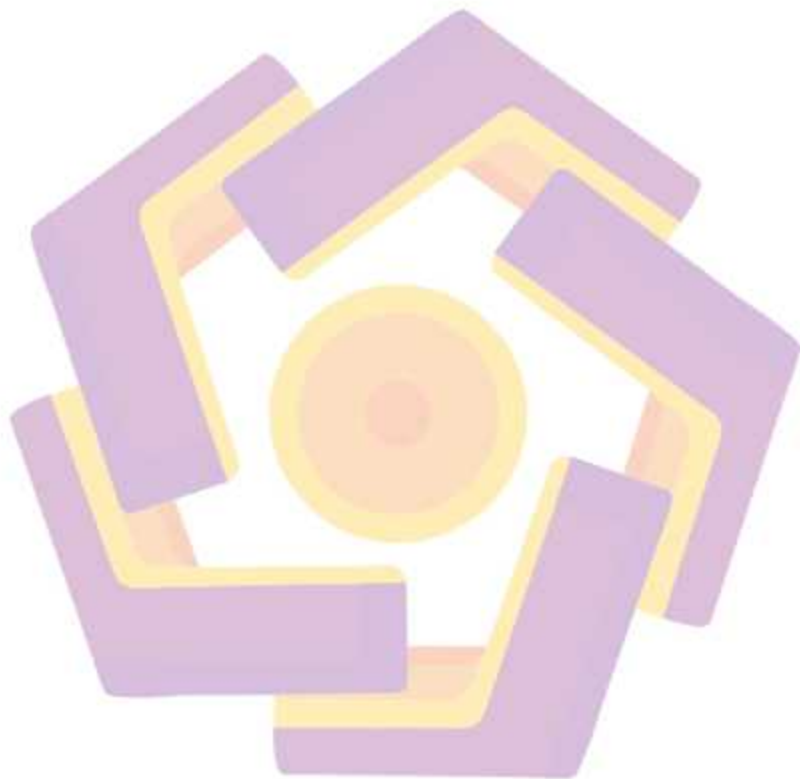


DAFTAR GAMBAR

Gambar 2.1 Bentuk <i>Confusion Matrix</i> 2 Kelas	19
Gambar 3.1 Alur Penelitian	22
Gambar 4.1 Citra <i>Benign</i>	29
Gambar 4.2 Hasil Ekstraksi Fitur GLCM	31
Gambar 4.3 Visualisasi Keseluruhan Data	32
Gambar 4.4 Visualisasi Fitur Disimilaritas 0° Dengan Energi 135°	33
Gambar 4.5 Hasil <i>Balancing</i> Kelas Dataset	34
Gambar 4.6 Hasil <i>Min-Max Normalization</i> Data Seimbang	35
Gambar 4.7 Skor <i>Mutual Information</i> Data Seimbang	36
Gambar 4.8 Hasil Seleksi 6 Fitur Terbaik Pada Data Seimbang	37
Gambar 4.9 Hasil <i>Classification Report</i>	41
Gambar 4.10 Hasil <i>Confusion Matrix</i>	41

DAFTAR LAMPIRAN

Lampiran 1. Hasil Ekstraksi Fitur & <i>Labeling</i> Dataset Kanker Payudara	49
Lampiran 2. Visualisasi Hubungan dan Pola Data Antar Kombinasi Pasangan Fitur Dataset Kanker Payudara	51



INTISARI

Kanker, terutama kanker payudara, menjadi isu kesehatan serius di Indonesia dengan tercatat ada 68.858 kasus baru dan 22.000 kematian pada tahun 2020 yang dilansir oleh data statistik dari *Global Burden of Cancer Study* (Globocon) *World Health Organization* (WHO). Deteksi dini memiliki peran krusial untuk mengurangi angka kematian akibat kanker payudara dan teknologi ultrasonografi (USG) diakui salah satu yang potensial dalam mendukung ini. Penelitian ini fokus pada pengklasifikasian citra ultrasonografi/*ultrasound* menggunakan algoritma *Support Vector Machine* (SVM), ekstraksi fitur GLCM, normalisasi Min-Max, serta ekstraksi fitur *Mutual Information* dan *SelectKBest*. Dari beberapa kali percobaan menggunakan algoritma SVM dengan berbagai kombinasi nilai parameter yang telah diatur dan pengujian yang berbeda yaitu menggunakan *Train/Test Split* dengan proporsi 80/20 dan *K-Fold Cross Validation* menunjukkan bahwa algoritma SVM mampu mengklasifikasikan citra ultrasonografi kanker payudara menjadi dua kategori (*Benign/Tumor Jinak* dan *Malignant/Tumor Ganas*) dengan akurasi maksimal yang sama besar yaitu 79% setelah menerapkan teknik SMOTE maupun tanpa menggunakan teknik *Balancing Data*. Dengan demikian, algoritma *Support Vector Machine* (SVM) memiliki potensi sebagai model yang cukup bagus dalam melakukan klasifikasi citra *ultrasound* kanker payudara baik itu pada data asli yang belum dilakukan penyeimbangan maupun data dengan penggunaan teknik *Balancing Data*.

Kata kunci: Klasifikasi, Citra Ultrasound, Kanker payudara, Support Vector Machine

ABSTRACT

According to statistics from the Global Burden of Cancer Study (Globocon) of the World Health Organization (WHO), cancer, particularly breast cancer, is a severe health issue in Indonesia with 68,858 new cases and 22,000 deaths recorded in 2020. Ultrasonography (USG) technology is acknowledged as one of the potentials to support early detection, which is vital in reducing mortality from breast cancer. This study focuses on classifying ultrasound images using the Support Vector Machine (SVM) algorithm, GLCM feature extraction, Min-Max normalization, and Mutual Information with SelectKBest Feature Selection. From several experiments using the SVM algorithm with various combinations of parameter values that have been set and different texts, namely using a Train/Test Split with a proportion of 80/20 and K-Fold Cross Validation, it shows that the SVM algorithm is capable of classifying ultrasound images of breast cancer into two categories (Benign Tumor and Malignant Tumor) with the same maximum accuracy of 79% after applying the SMOTE Balancing Data technique or without using the Balancing Data technique. As a result, the Support Vector Machine (SVM) algorithm has the potential to be an effective model for identifying breast cancer ultrasound images, both on data from the original set that has not been balanced and data from the set that has been balanced.

Keyword: *Classification, Ultrasound Image, Breast Cancer, Support Vector Machine*