IDENTIFIKASI PENYAKIT KANKER PAYUDARA MENGGUNAKAN METODE GLRLM DAN SVM DENGAN REDUKSI FITUR

SKRIPSI



Disusun Oleh
RIZQA FIANY ZEIN
H92218051

PROGRAM STUDI MATEMATIKA
FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL
SURABAYA

2021

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama

: Rizqa Fiany Zein

NIM

: H92218051

Program Studi : Matematika

Angkatan

: 2016

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya PAYUDARA KANKER "IDENTIFIKASI PENYAKIT berjudul yang MENGGUNAKAN METODE GLRLM DAN SVM DENGAN REDUKSI FITUR". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 18 November 2021

Yang menyatakan,

zqa fiany Zein

NIM. H92218051

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama

: Rizqa Fiany Zein

NIM

: H92218051

Judul Skripsi : IDENTIFIKASI PENYAKIT KANKER PAYUDARA

MENGGUNAKAN METODE GLRLM DAN SVM

DENGAN REDUKSI FITUR

telah diperiksa dan disetujui untuk diujikan.

Pembimbing I

Dian Candra Kun Novitasari, M. Kom

NIP. 198511242014032001

Pembimbing II

Lutfi Hakim, M.Ag

NIP. 197312252006041001

Mengetahui,

Ketua Program Studi Matematika

UIN Sunan mpel Surabaya

NIP. 198701272014031002

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi oleh

Nama : Rizqa Fiany Zein NIM : H92218051

Judul Skripsi : IDENTIFIKASI PENYAKIT KANKER PAYUDARA

MENGGUNAKAN METODE GLRLM DAN SVM

DENGAN REDUKSI FITUR

Telah dipertahankan di depan Tim Penguji pada tanggal

Mengesahkan, Tim Penguji

Penguji I

Dian Candra Rini Novitasari, M. Kom

NIP. 198511242014032001

Penguji II

Lutfi Hakim, M.Ag

NIP. 197312252006041001

Penguji IV

Penguji III

Dr. Moh. Hafiyusholeh, M.Si., M.PMat

NIP. 198002042014031001

07

Putrge Keumala Intan, M.Si

NIP. 198805282018012001

Mengetahui, kan Fakuhas Sains dan Teknologi UIN Sunan Ampel Surabaya

rof. Dr. Fa Fann aur Rusydiyah, M.Ag

2272005012003



KEMENTERIAN AGAMA UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300 E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:
Nama : RIZQA FIANY ZEIN
NIM : 492218051
Fakultas/Jurusan : SAINTEK / MATEMATIKA
E-mail address : RIZKAVIANYP @ GMAIL. COM
Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UJN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah: Sekripsi Tesis Desertasi Lain-lain () yang berjudul: IDENTIFIKASI PENYARIT KANKER PAYUDARA MENGGUNAKAN
METODE GLALM DAN SVM DENGAN REDUKSI FITUR
Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara fulltext untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan. Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.
Demikian pernyataan ini yang saya buat dengan sebenarnya.
Surabaya,
Penulis
Rusus_

Dipindai dengan CamScanner

RIZQA FIANY ZEIN)
nama terang dan tanda tangan

ABSTRAK

IDENTIFIKASI PENYAKIT KANKER PAYUDARA MENGGUNAKAN METODE GLRLM DAN SVM DENGAN REDUKSI FITUR

Salah satu penyebab utama dari kasus kematian di dunia yaitu penyakit kanker. Di Indonesia, kanker payudara termasuk salah satu penyakit yang banyak diderita dengan angka mencapai 28.7% jumlah penderita penyakit kanker payudara. Oleh karena itu penting dilakukan pencegahan dini untuk mengurasi angka penderita kanker payudara. Dalam dunia medis, saat ini telah banyak cara untuk mendeteksi penyakit kanker payudara, yaitu salah satunya pemanfaatan citra digital (mammogram). Citra mammogram akan dilakukan beberapa langkah untuk sampai pada sistem klasifikasi. Pertama, pengumpulan data citra Mammogram untuk setiap kelas. Kedua, preprocessing yaitu digunakan untuk memperbaiki kualitas dari setiap data citra dengan menggunakan metode CLAHE dan Median Filter. Ketiga, ekstraksi fitur dan reduksi yaitu untuk pengambilan ciri penting pada setiap citra dengan menggunakan metode GLRLM dan PCA. Terakhir, sistem klasifikasi yaitu menggunakan metode SVM untuk mencari model terbaik klasfikasi. Pada penelitian ini diperoleh peningkatan evaluasi ketika dilakukan reduksi fitur, pada sistem klasifikasi Normal-Upnormal diperoleh hasil terbaik yaitu pada sudut 90° yaitu akurasi sebesar 71.875% mengalami peningkatan sebanyak 8.239%. Pada sistem klasifikasi Jinak-Ganas diperoleh hasil terbaik yaitu pada sudut 0° yaitu akurasi sebesar 81.818% mengalami peningkatan sebanyak Hasil penelitian ini diharapkan dapat membantu dalam proses 9.091%. mengidentifikasi penyakit kanker payudara.

Kata kunci: Citra Mammogram, Breast Cancer, GLRLM, PCA, SVM

ABSTRACT

IDENTIFICATION OF BREAST CANCER USING GLRLM AND SVM METHODS WITH REDUCTION FEATURE

One of the main causes of death in the world is cancer. In Indonesia, breast cancer is one of the most common diseases with a figure reaching 28. % of the number of breast cancer sufferers. Therefore, it is important to do early prevention to reduce the number of breast cancer sufferers. In the medical world, currently there are many ways to detect breast cancer, one of which is the use of digital images (mammogram). The mammogram image will be carried out in several steps to arrive at the classification system. First, the collection of mammogram image data for each class. Second, preprocessing is used to improve the quality of each image data using CLAHE and Median Filter methods. Third, feature extraction and reduction, namely to retrieve important features in each image using the GLRLM and PCA methods. Finally, the classification system uses the SVM method to find the best classification model. In this study, the evaluation was In the Normal-Upnormal obtained when feature reduction was performed. classification system, the best results were obtained, namely at an angle of 90°, namely an accuracy of 71.875%, an increase of 8.239%. In the Benign- Malignant classification system, the best results are obtained, namely at the angle of 0°, which is an accuracy of 81.818%, an increase of 9.091%. The results of this study are expected to help in the process of identifying breast cancer.

Keywords: Mammogram Image, Breast Cancer, GLRLM, PCA, SVM

DAFTAR ISI

HALAMAN JUDUL	 •	i
LEMBAR PERSETUJUAN PEMBIMBING		ii
PENGESAHAN TIM PENGUJI SKRIPSI	 •	iii
HALAMAN PERNYATAAN KEASLIAN	 •	iv
MOTTO		v
HALAMAN PERSEMBAHAN		vi
KATA PENGANTAR		vii
DAFTAR ISI		viii
DAFTAR TABEL		хi
DAFTAR GAMBAR		xii
ABSTRAK		xiii
ABSTRACT		xiv
I PENDAHULUAN		1
1.1. Latar Belakang Masalah		1
1.2. Rumusan Masalah		
1.3. Tujuan Penelitian		
1.4. Manfaat Penelitian		7
1.5. Batasan Masalah		7
1.6. Sistematika Penulisan		8
II TINJAUAN PUSTAKA	 •	10
2.1. Kanker Payudara		10
2.2. CLAHE (Contrast-Limited Adaptive Histogram Equalization)		12
2.3. Median Filter		14
2.4. Gray Level Run Length Matrix (GLRLM)		16
2.4.1. Short Run Emphasis (SRE)		17
2.4.2. Long Run Emphasi (LRE)		18
2.4.3. Gray-Level Nonuniformity (GLN)		18

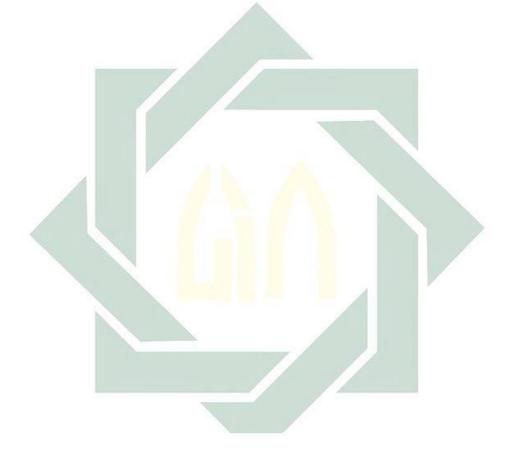
	2.4.4.	Run Length Nonuniformity (RLN)	18
	2.4.5.	Run Percentage (RP)	19
	2.4.6.	Low Gray-Level Run Emphasis (LGRE)	19
	2.4.7.	High Gray-Level Run Emphasis (HGRE)	20
	2.4.8.	Short Run Low Gray-Level Emphasis (SRHGE)	20
	2.4.9.	Short Run High Gray-Level Emphasis (SRLGE)	20
	2.4.10.	Long Run Low Gray-Level Emphasis (LRLGE)	21
	2.4.11.	Long Run High Gray-Level Emphasis (LRHGE)	21
2.5.	Princip	pal Component Analysis (PCA)	22
2.6.		Cross Validation	
2.7.	Suppor	t Vector Machine (SVM)	27
2.8.	Confus	ion Matrix <mark></mark>	29
	2.8.1.	True Positive (TP)	30
	2.8.2.	False Positive (FP)	30
	2.8.3.	False Negative (FN)	31
	2.8.4.	True Negative (TN)	
	2.8.5.	Akurasi	31
	2.8.6.	Sensitivitas	32
	2.8.7.	Spesifisitas	32
2.9.	Integra	si Keislaman	32
III MET	ODE P	ENELITIAN	35
3.1.	Jenis P	enelitian	35
3.2.	Jenis d	an Sumber Data	35
3.3.	Langka	ah-Langkah	36
IV HAS	IL DAN	N PEMBAHASAN	38
4.1.	Prepro	cessing	38
4.2.	Ekstrak	ksi Fitur	40
4.3.	Klasifil	kasi SVM	47
	4.3.1.	Klasifikasi SVM dengan GLRLM	48
4.4.	Klasifil	kasi SVM dengan Fitur GLRLM-PCA	50

	4.5.	Integrasi Keislaman	5
V	PEN	UTUP	60
	5.1.	Kesimpulan	60
	5.2.	Saran	61
DA	ETA 1	D DIICTA KA	61



DAFTAR TABEL

- 4.1 Ekstraksi Fitur Kelas Normal-Abnormal pada Sudut Orientasi 0° . . 45
- 4.2 Ekstraksi Fitur Kelas Benign-Malignant pada Sudut Orientasi 0° . . 45



DAFTAR GAMBAR

2.1	(a) Mammographic Image Analysis Society (MIAS), (b) Ultrasonography (USG) (Al-Dhabyani et al., 2020), (c) Magnetic	
	Resonance Imaging (MRI) (Morrow et al., 2011), (d) Positron Emission Tomography (PET) scan (Hao et al., 2019)	12
2.2	Contrast-Limited Adaptive Histogram Equalization (Ma'arif and	
	Arifin, 2017)	13
2.3	Contrast-Limited Adaptive Histogram Equalization (Erwin and	
	Ningsih, 2020)	13
2.4	Median Filter (Erkan et al., 2018)	16
2.5	Ilustrasi Sudut Orientasi pada GLRLM	17
2.6	Ilustrasi K-Fold Cross Validation	26
2.7	Klasifikasi SVM dua kelas	27
2.8	Confusion Matrix dengan Tiga Kelas Klasifikasi	30
3.1	(a) Normal, (b) Benign, (c) Malignant	36
3.2	Diagram Alir	36
4.1	Sampel citra Mammogram	38
4.2	Sampel pada CLAHE	39
4.3	Sampel pada Median Filter	40
4.4	Tabel Uji Coba GLRLM pada Kelas Normal-Abnormal	48
4.5	Tabel Uji Coba GLRLM pada Kelas Benign-Malignant	49
4.6	Matriks Korelasi	51
4.7	Hasil Loading Factor	52
4.8	Tabel Uji Coba GLRLM-PCA pada Kelas Normal-Abnormal	53
4.9	Perbandingan Hasil Terbaik Normal-Abnormal	54
4.10	Tabel Uji Coba GLRLM-PCA pada Kelas Benign-Malignant	55
4.11	Perbandingan Hasil Terbaik Benign-Malignant	56
4.12	Guide untuk Identifikasi Penyakit Kanker Payudara	57

BABI

PENDAHULUAN

1.1. Latar Belakang Masalah

Salah satu penyebab utama dari kasus kematian di dunia yaitu penyakit kanker (Solikhah, 2019). Pada 2012, berdasarkan data *World Health Organization* (WHO) memperkirakan bahwa terdapat 14,1 juta korban dari kasus baru terkait kanker dan kasus kematian sebanyak 8,2 juta korban akibat kanker di dunia. Insiden terkait kanker meningkat jika ditinjau dari tahun 2008, yang semula 12,7 juta korban dari kasus baru terkait kanker dan 7,6 juta korban meninggal akibat kanker (Utami and Mustikasari, 2017). Selain itu, WHO melaporkan bahwa terdapat lima jenis kanker yang sering ditemukan pada tubuh laki-laki, yaitu kanker hati, paru-paru, kolorektum, prostat, dan perut. Kanker perut atau stomach cancer ini juga sering ditemukan pada tubuh perempuan. Selain kanker perut yang sering juga ditemukan pada tubuh wanita, yaitu kanker serviks dan kanker payudara (Dewi, 2017). Di Indonesia, kanker payudara termasuk salah satu penyakit yang banyak diderita dengan angka mencapai 28,7% jumlah penderita penyakit kanker payudara (Rasmi, 2020).

Data terkait kanker payudara di Indonesia memperkirakan bahwa terdapat sekitar 100 dari 100 ribu orang penderita setiap tahunnya (Wahyuni et al., 2018). Hal ini menunjukkan bahwa ada sekitar 237 ribu penderita kanker payudara baru, jika jumlah penduduk saat ini yaitu 237 juta penduduk. Berdasarkan Data Buletin Jendela dan Informasi Kesehatan, sebesar 43,3% penduduk wanita menderita

kanker payudara pada kasus baru dan kasus kematian mencapai 12,9% dari penduduk wanita di Indonesia (Irawan, 2018). Sebagian penderita dari penyakit kanker payudara ditemukan pada tubuh wanita yang berusia muda, dan banyak remaja berusia 14 tahun menderita kanker payudara. Hal ini akan berbahaya jika tidak diketahui lebih awal, karena kanker yang awalnya kecil berpotensi menjadi ganas (Rahman, 2019). Melihat banyak kasus yang telah disebabkan oleh kanker payudara, maka dari itu diperlukannya pencegahan dini untuk mengurangi angka penderita dari kanker payudara.

Dalam dunia medis, saat ini telah banyak cara untuk mendeteksi penyakit kanker payudara, yaitu salah satunya pemanfaatan citra digital. Citra digital pada dunia medis disebut juga medical imaging, dimana teknik ini dilakukan dengan mengambil citra dari organ tubuh, salah satunya yaitu *mammogram*. *Mammogram* merupakan metode yang dikatakan efektif untuk memeriksa adanya kanker payudara (Liantoni and Santoso, 2020). Banyak penelitian sebelumnya telah digunakan juga citra mammogram untuk klasifikasi penyakit kanker payudara. Dalam mengklasifikasikan penyakit kanker payudara terdapat beberapa tahapan diantaranya pre-processing, ekstraksi fitur, dan masuk kepada sistem klasifikasi.

Pada penelitian yang menguji kualitas citra juga mengusulkan untuk menggunakan Median filter dalam melakukan pre-processing. Berdasarkan metode yang diusulkan yaitu menggunakan dataset STARE dengan mengambil citra retina dalam format JPG. Dari citra tersebut dilakukan perbaikan kualitas citra dengan menggunakan tiga kontras yaitu contrast stretching, *Histogram Equalization* (HE) dan *Contrast Limited Adaptive Histogram Equalization* (CLAHE) yang difilter menggunakan Median filter . Dari hasil pengujian dengan tiga metode yang berbeda dan berdasarkan tiga parameter yang berbeda didapatkan hasil MSE dan

PSNR terbaik yaitu 9,14 dan 42,13 dB. Dengan demikian peregangan kontras lebih baik dalam meningkatkan kualitas gambar (Erwin and Ningsih, 2020). Selain itu pada penelitian terkait citra *mammogram* juga melakukan langkah pre-processing, di mana pada penelitian tersebut dilakukan pengolahan citra digital dengan memanfaatkan Median filter dan Histogram Equalization. Hasil dari pre-processing ini akan diambil ciri ekstraksi dengan *Histogram of Oriented Gradient* (HOG), yang selanjutnya diklasifikasikan menggunakan Deep Rule Based (DRB) Classifier. Penelitian tersebut dilakukan dengan membagi data 90% untuk data latih dan 10% untuk data uji. Data diklasifikasikan menjadi tiga kelas, terdiri dari normal, jinak, dan ganas. Akurasi terbaik yang didapat adalah 92% (Novitasari et al., 2019b). Hal ini membuktikan bahwa pre-processing untuk citra mammogram perlu dilakukan untuk membantu memperbaiki kualitas dari citra.

Pada penelitian sebelumnya, pendekatan yang diusulkan untuk mengembangkan sistem klasifikasi dengan bantuan komputer untuk deteksi kanker dari citra mammogram. Citra ini dilakukan ekstraksi fitur yang digunakan sebagai fitur masukan sistem klasifikasi, yaitu dengan GLCM dan GLRLM. Matriks fitur GLCM dan GLRLM membantu sistem pengklasifikasi mendapatkan hasil terbaik dengan memberikan akurasi dan spesifisitas klasifikasi 97,03% dan 98,44% (Erkan et al., 2018). Hal ini membuktikan bahwa metode GLRLM memiliki sistem yang baik dalam sistem klasifikasi, namun metode ini memiliki dimensi yang tinggi sehingga membuat rumit perhitungan dari komputasi.

Ekstraksi fitur dari metode GLRLM memiliki dimensi yang tinggi, oleh karena itu untuk mengurangi beratnya perhitungan komputasi perlu dilakukan reduksi fitur. Reduksi fitur telah dilakukan pada penelitian sebelumnya terkait klasifikasi penyakit tumor payudara. Penelitian tersebut membandingkan antara

tiga metode reduksi fitur yaitu metode PCA, metode RFE, dan metode RFECV. Data DNA akan dilakukan klasifikasi dengan metode K Nearest Neighbor (KNN), di mana masukan dari metode ini yaitu data sebelum dan setelah melalui proses reduksi fitur. Hasil dari sistem klasifikasi menunjukkan bahwa akurasi terbaik dicapai oleh reduksi fitur metode PCA dengan akurasi sebesar 97,36%. Selain itu waktu komputasi mengalami penurunan saat menggunakan reduksi fitur PCA yaitu menjadi 1,231 detik (Wibawa and Novianti, 2017). Reduksi fitur juga telah dilakukan pada penelitian terkait penyakit malaria, di mana metode PCA digunakan untuk mereduksi fitur dari data RNA-seq vektor malaria yang memiliki berdimensi tinggi. Dalam penelitian tersebut model pengurangan dimensi menggunakan metode PCA dapat membantu meningkatkan hasil klasifikasi dari metode KNN. Hasil eksperimen mencapai kinerja yang dapat dikatakan relevan dengan akurasi klasifikasi KNN dan Decision Tree masing-masing, yaitu 86,7% dan 83,3% (Arowolo et al., 2020). Hal ini membuktikan bahwa metode PCA baik dilakukan untuk reduksi fitur, dan dapat diimplementasikan pada hasil ekstraksi fitur dari metode GLRLM.

Selain itu perbandingan hasil pada metode ekstraksi fitur GLCM, GLRLM, dan GLDM dalam mendiagnosis penyakit kanker payudara. Kanker payudara dilakukan pengklasifikasian dengan menggunakan citra mammogram pada penelitian tahun 2019. Penelitian tersebut menggunakan fitur-fitur yang didapatkan dari ekstraksi fitur digunakan sebagai masukan untuk klasifikasi SVM dengan perbandingan tiga kernel, yaitu kernel polinomial, RBF, dan linear. Hasil penelitian menunjukkan bahwa kernel terbaik adalah kernel polinomial dengan fitur yang dihasilkan dari GLRLM yaitu akurasi sebesar 93.98% (Novitasari et al., 2019a). Hal ini menunjukkan bahwa metode dari sistem klasifikasi juga berperan

penting dan metode SVM menunjukkan kinerja yang baik dalam pengklasifikasian.

Metode SVM untuk klasifikasi juga diusulkan pada penelitian sebelumnya terkait kanker serviks, di mana metode yang digunakan untuk mengidentifikasi adanya kanker serviks adalah dengan melakukan observasi visual yang menghasilkan data citra. Pada penelitian tersebut metode observasi berbasis komputer yang digunakan adalah *pre-processing*, dilanjutkan dengan proses ekstraksi ciri menggunakan *Gray Level Co-occurrence Matrix* (GLCM) dan *Support Vector Machine* (SVM) sebagai metode klasifikasi. Hasil klasifikasi SVM terbaik menggunakan kernel polinomial dan ekstraksi fitur GLCM dengan sudut 45° dan akurasi yang diperoleh mencapai 90% (Thohir et al., 2020).

Dari penjelasan penelitian sebelumnya, maka penelitian ini menggunakan citra mammogram untuk pengklasifikasian kanker payudara. Citra mammogram akan dilakukan pre-processing sebelum masuk ke proses ekstraksi fitur, dengan menggunakan metode Median filter dan CLAHE untuk memperbaiki kualitas dari citra. Ekstraksi fitur yang menggunakan metode GLRLM dan reduksi fitur metode PCA akan dijadikan sebagai masukan dari sistem klasifikasi menggunakan metode SVM. Hasil akhir dari penelitian ini yaitu evaluasi dari sistem klasifikasi untuk mengetahui hasil dari akurasi, sensitifitas dan spesifisitas. Berdasarkan dari kondisi serta penjelasan penelitian sebelumnya yang telah dipaparkan, maka penelitian akan dilakukan dengan judul "Identifikasi Penyakit Kanker Payudara Menggunakan Metode GLRLM dan SVM Dengan Reduksi Fitur". Hasil penelitian ini diharapkan dapat membantu dalam proses mengidentifikasi penyakit kanker payudara.

1.2. Rumusan Masalah

Berdasarkan permasalahan yang telah dijelaskan pada latar belakang, adapun beberapa rumusan masalah dari penelitian tersebut yaitu sebagai berikut:

- Bagaimana hasil dari pre-processing dengan menggunakan metode CLAHE dan Median filter pada citra mammogram?
- 2. Bagaimana hasil reduksi fitur PCA dari ekstraksi fitur GLRLM terhadap hasil klasifikasi penyakit kanker payudara dengan menggunakan citra mammogram?
- 3. Bagaimana hasil dari klasifikasi citra *mammogram* dengan menggunakan metode SVM?

1.3. Tujuan Penelitian

Berdasarkan beberapa rumusan masalah yang telah dijelaskan, adapun tujuan dari penelitian ini yaitu sebagai berikut.

- 1. Mengetahui hasil ekstraksi fitur dari GLRLM untuk pengklasifikasian penyakit kanker payudara dengan menggunakan citra *mammogram*.
- Mengetahui pengaruh reduksi fitur PCA dari ekstraksi fitur GLRLM terhadap hasil klasifikasi penyakit kanker payudara dengan menggunakan citra mammogram.
- Mengetahui model terbaik dan hasil dari klasifikasi citra mammogram dengan menggunakan metode SVM.

1.4. Manfaat Penelitian

Manfaat secara teoritis yaitu sebagai berikut.

- Penelitian ini diharapkan dapat memperluas dan memperkaya literature di bidang ilmu komputer sains, khususnya tentang pengelolaan dan klasifikasi citra.
- 2. Penelitian ini diharapkan dapat membantu untuk memahami metode PCA yang digunakan untuk reduksi fitur dari ekstraksi GLRLM pada pengklasifikasian citra *mammogram*.

Manfaat bagi peneliti dan pembaca yaitu sebagai berikut.

- 1. Penelitian ini diharapkan dapat menambah wawasan dalam ilmu pengetahuan tentang penyelesaian suatu masalah
- 2. Penelitian ini diharapkan dapat memahamkan tentang sistem dan kinerja dari metode yang digunakan

Manfaat di bidang kedokteran yaitu sebagai berikut.

- Penelitian ini diharapkan dapat membantu untuk mendiagnosis pasien terkait penyakit kanker payudara
- 2. Penelitian ini diharapkan dalam menambah wawasan tentang sistem deteksi penyakit kanker menggunakan teknologi komputer.

1.5. Batasan Masalah

Ruang lingkup yang dimiliki penelitian ini begitu luas, oleh karena itu penelitian ini terdapat beberapa batasan masalah, antara lain:

- Data yang akan digunakan pada penelitian ini yaitu data citra mammogram yang diambil dari *Mammographic Image Analysis Society* (MIAS) sebanyak 322 citra.
- Sistem klasfikasi pada penelitian ini yaitu dua kelas yang dilakukan secara dua kali, pertama klasifikasi dari 207 citra normal dan 115 citra upnormal, kedua klasfikasi dari 64 citra dari pasien tumor jinak, dan 51 citra dari pasien tumor ganas.
- 3. Reduksi fitur yang digunakan pada penelitian ini yaitu metode PCA.
- 4. Parameter yang digunakan pada sistem klasifikasi penelitian ini yaitu kernel *Radial Basis Function* (RBF), kernel *Linear*, dan kernel *Polynomial*
- 5. Penelitian ini merupakan salah satu penelitian dari komputer sains, di mana spesifikasi dari komputer yang digunakan processor intel core i7 dan RAM 8 GB.

1.6. Sistematika Penulisan

Penyusunan dari penelitian tersebut dijelaskan pada sistematika penulisan yaitu sebagai berikut.

- BAB I PENDAHULUAN Berisi tentang penjelasan-penjelasan dari latar belakang penelitian, rumusan masalah dari permasalahan yang ada, tujuan dari penelitian, manfaat dari penelitian, batasan masalah yang digunakan, serta sistematika dari penulisan.
- BAB II TINJAUAN PUSTAKA Berisi tentang teori-teori terkait kanker, serta kanker payudara, metode GLRLM untuk fitur ekstraksi, metode PCA untuk reduksi fitur, metode K-Fold untuk membagi data, metode SVM untuk

sistem klasifikasi, dan *confusion matrix* untuk metode evaluasi dari sistem klasifikasi.

- 3. BAB III METODE PENELITIAN Berisi tentang penjelasan dari diagram alir atau alur dari penelitian ini, selain itu penjelasan terkait jenis penelitian dan pengumpulan dari data *mammogram*.
- 4. BAB IV HASIL DAN PEMBAHASAN Berisi tentang uraian hasil dari preprocessing fitur ekstraksi GLRLM, dan reduksi fitur PCA. Selain itu terdapat evaluasi dari hasil klasifikasi kanker payudara menggunakan metode SVM pada citra mammogram.
- 5. BAB V PENUTUP Berisi tentang kesimpulan dari hasil pengujian penelitian ini, serta saran untuk penelitian selanjutnya.

BAB II

TINJAUAN PUSTAKA

2.1. Kanker Payudara

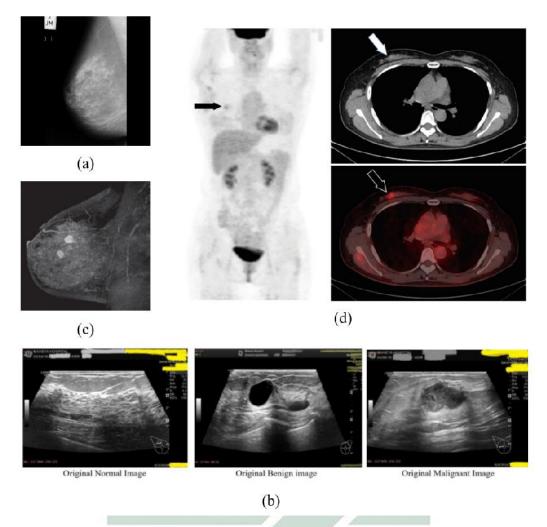
Kanker bukan penyakit yang dapat dianggap ringan, melainkan penyakit yang mematikan. Penyakit ini menyerang jaringan sel tertentu dan akan menyebar pada organ tubuh yang lain. Pada pengertian yang sederhana kanker adalah pertumbuhan dari sel yang dapat dikatakan tidak normal atau tidak terkendali (Liang et al., 2019). Pertumbuhan sel ini akan terus berkembang biak meskipun pertumbuhan ini tidak diperlukan oleh tubuh, selain itu penyebab dari pertumbuhan sel ini tidak dapat diketahui dengan pasti. Jika sel-sel kanker mengalami pertumbuhan yang tidak terkoordinasi dengan sel jaringan lainnya, maka hal ini akan berbahaya bagi tubuh. Penyebaran dari sel kanker ini dapat melalui darah, saraf, jaringan ikat, dan organ tubuh yang mengalami penyebaran ini akan mengalami pertumbuhan organ tubuh yang lambat (Devin et al., 2019).

Kejadian kanker dapat ditemukan karena penyebab tertentu, yaitu salah satunya perubahan pola makan. Contohnya pada tahun 1930-an di Negara Amerika, kanker lambung sering ditemukan dibandingkan dengan sekarang. Kanker lambung sering ditemukan karena banyak orang yang mengonsumsi makanan dengan cara diasap atau diasamkan, di mana hal ini mempengaruhi sel lambung. Kemudian pada tahun 1930-an, kanker mulut sering ditemukan dan terus meningkat hingga tahun 1990-an. Hal ini terjadi karena meningkatnya pengguna rokok sigaret, dari contoh-contoh ini membuktikan bahwa gaya hidup dan pola

makanan merupakan faktor dari terjadinya kanker (Wang et al., 2019). Sebesar 30% dari penderita kanker ini disebabkan oleh gaya hidup dan pola makan, antara lain kurangnya mengonsumsi sayur dan buah-buahan, dan kurangnya aktivitas fisik seperti olahraga. Sel kanker telah terdapat pada tubuh setiap orang, oleh karena itu tidak salah jika penyakit ini dapat menyerang kepada semua umur tanpa mengenal batasan usia dan jenis kelamin, baik pria maupun wanita. Berdasarkan data yang telah tercatat, wanita yang rentan terjangkit kanker sekitar 60% dari penderita kanker di negara Indonesia. Jenis kanker yang sangat sering diderita oleh wanita, yaitu kanker serviks, kanker ovarium, serta kanker payudara (Rahayuwati et al., 2020).

Kanker payudara merupakan salah satu dari jenis kanker yang sering diderita oleh wanita, selain itu kanker payudara merupakan penyebab dari kematian wanita pada urutan pertama di Indonesia. Penyakit ini banyak menyerang wanita-wanita yang berumur 40 tahun ke atas, tetapi tidak ada jaminan bahwa wanita yang berusia muda tidak dapat terjangkit penyakit kanker payudara (Wibisono and Hermawan, 2019). Kanker payudara dapat dideteksi dengan beberapa cara yaitu Mammographic Image Analysis Society (MIAS), Ultrasonography (USG), Magnetic Resonance Imaging (MRI), dan Positron Emission Tomography (PET) scan. Mammogram adalah salah satu alternatif yang berfungsi untuk mendeteksi adanya kanker payudara. Mammogram dapat mengetahui kondisi dari payudara dengan menggunakan amplitudo yang lebih rendah jika dibandingkan dengan X-Ray. Sistem deteksi yang dilakukan mammogram untuk deteksi awal dari kanker payudara, yaitu dengan melihat karakteristik dan bentuk jika ditemukan benjolan (Yulianti et al., 2016). Menurut WHO, mammogram dapat mengurangi faktor risiko dari kanker payudara, salah

satunya yaitu kematian akibat kanker payudara. (Ardayani et al., 2020)



Gambar 2.1 (a) Mammographic Image Analysis Society (MIAS), (b) Ultrasonography (USG) (Al-Dhabyani et al., 2020), (c) Magnetic Resonance Imaging (MRI) (Morrow et al., 2011), (d) Positron Emission Tomography (PET) scan (Hao et al., 2019)

2.2. CLAHE (Contrast-Limited Adaptive Histogram Equalization)

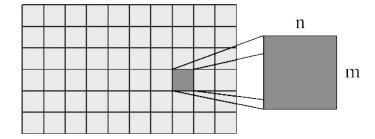
Metode CLAHE merupakan salah satu metode yang digunakan dalam pre-processing. Keunggulan dari CLAHE yaitu citra yang dihasilkan memiliki kualitas yang lebih baik jika dibandingkan dengan citra sesungguhnya (Siddhartha and Santra, 2020), dapat dilihat pada Gambar 2.2. Keunggulan ini didapatkan karena metode CLAHE digunakan untuk meningkatkan kontras dari suatu citra.



Gambar 2.2 Contrast-Limited Adaptive Histogram Equalization (Ma'arif and Arifin, 2017)

Selain itu metode ini merupakan perbaikan dari Adaptive Histogram Equalization (AHE), di mana algoritma CLAHE memiliki proses dengan menerapkan Histogram Equalization (HE) pada setiap blok dengan berisi berapa piksel yang telah dibagi dari citra aslinya (Erwin and Ningsih, 2020). Algoritma dari CLAHE yaitu sebagai berikut.

1. Selain itu metode ini merupakan perbaikan dari *Adaptive Histogram Equalization* (AHE), di mana algoritma CLAHE memiliki proses dengan menerapkan *Histogram Equalization* (HE) pada setiap block dengan berisiberapa piksel yang telah dibagi dari citra aslinya (Erwin and Ningsih, 2020).



Gambar 2.3 Contrast-Limited Adaptive Histogram Equalization (Erwin and Ningsih, 2020)

2. Menghitung histogram pada setiap block yang diproses dari nilai keabu-abuan

pada citra, dengan rumus pada persamaan 2.1 sebagai berikut.

$$S_k = \sum_{i=0}^k p(r_i); 0 \le r_i \le 1; k = 0, 1, 2, ..., L - 1$$
 (2.1)

Di mana S_k merupakan nilai keabu-abuan dari citra, L merupakan derajat keabu-abuan dari citra, dan $p(r_i)$ merupakan jumlah piksel dari citra dengan intensitas nilai r_i . Citra akan memiliki warna hitam jika nilai dari $r_i = 0$ dan akan memiliki warna putih jika nilai dari $r_i = 1$.

3. Menghitung clipped histogram dari setiap citra, dengan rumus pada persamaan 2.2 sebagai berikut.

$$\beta = \frac{M}{N} \left(1 + \frac{\alpha}{100} (S_m a x - 1) \right) \tag{2.2}$$

Di mana M merupakan luas dari block, N merupakan nilai dari graygcale dengan range dari 0 hingga 256, dan α merupakan nilai clipfactor dengan range antara 0 hingga 100, nilai ini menyatakan penambahan batas pada suatu histogram (Matondang, 2018).

2.3. Median Filter

Metode yang umumnya digunakan untuk mengurangi atau menghilangkan noise salt & paper pada citra yaitu Median filter . Selain itu Median filter juga berfungsi untuk memperhalus kualitas citra (Noor et al., 2020). Median filter merupakan salah satu metode non linear yang berbasis data statistik, dan merupakan metode pengelolaan dari citra digital yang dihitung dari nilai keabu-abuan citra. Median filter dihitung dari matriks *filtering* atau disebut window yang berukuran $n \times n$, nilai median dari window akan dicari untuk

menggantikan setiap piksel pada citra (Kosasih et al., 2021). Berikut merupakan tahapan proses dari Median filter.

- 1. Proses penambahan tingkat intensitas noise yang dilakukan manual, proses ini disebut juga pra pengolahan.
- Citra hasil dari pra pengolahan akan diproses dengan filter yaitu metode
 Median Filter , rumus dari Median filter dapat dilihat pada persamaan 2.3.

$$g(x,y) = median\{f(x-i,y-i)\}; i,j \in W$$
 (2.3)

dengan

$$DataGanjil \rightarrow median = \frac{W_{(i \times j)+1}}{2}$$

$$DataGenap \rightarrow median = \frac{W_{\frac{i \times j}{2}} + W_{\frac{i \times j}{2}+1}}{2}$$
(2.4)

Di mana f(x,y) adalah piksel-piksel dari citra dan g(x,y) adalah hasil dari perhitungan Median Filter , yang dapat dilihat pada Gambar 2.4. Selain itu i, j merupakan baris dan kolom dari W, dimana W adalah mask 2 dimensi atau matriks filtering yang berukuran $n \times n$ (Fadillah and Gunawan, 2019).

Metode pengelolaan citra selain Median filter, yaitu metode mean filter. Metode ini memiliki proses yang sama tetapi menggunakan rumus rata-rata untuk menghasilkan citra yang baru. Selain itu Median filter lebih diunggulkan jika dibandingkan dengan mean filter, berikut keunggulan Median filter.

 Median filter dikatakan lebih baik karena jika ada satu piksel yang tidak representatif dengan piksel-piksel pada citra maka pada hasil dari mean filter tidak akan menunjukkan perbaikan yang signifikan. Nilai-nilai yang dihasilkan dari Median filter merupakan nilai dari piksel itu sendiri, sehingga hasil yang didapatkan lebih terlihat realistis jika dibandingkan dengan mean filter (Foeady, 2019).

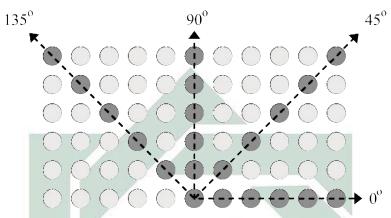


Gambar 2.4 Median Filter (Erkan et al., 2018)

2.4. Gray Level Run Length Matrix (GLRLM)

Metode GLRLM merupakan salah satu dari metode analisis tekstur yang dapat digunakan untuk fitur ekstraksi yang biasanya digunakan untuk masukan dari sistem klasifikasi. Analisis tekstur ini dapat dilakukan pada nilai keabu-abuan citra dalam arah tertentu. Pada fitur GLRLM dibentuk dari orientasi empat arah, yaitu 0°, 45°, 90°, dan 135°. Arah-arah yang berorientasi ini dapat dilihat pada ilustrasi Gambar 2.5. Nilai fitur yang didapatkan dari metode GLRLM memiliki 11 fitur yang digabungkan dari tiga peneliti, yaitu *Short Run Emphasis* (SRE), *Long Run Emphasi* (LRE), *Gray-Level Nonuniformity* (GLN), *Run Length Nonuniformity* (RLN), *Run Percentage* (RP), *Low Gray-Level Run Emphasis* (LGRE), *High Gray-Level Run Emphasis* (HGRE), *Short Run Low Gray-Level Emphasis* (SRHGE), *Short Run High Gray-Level Emphasis* (SRLGE), *Long Run Low Gray-Level Emphasis* (LRLGE), dan *Long Run High Gray-Level Emphasis* (LRHGE). Menurut Galloway, pada GLRLM terdapat 5 fitur yaitu SRE, GLN, LRE, RP, dan RLN (Durgamahanthi et al., 2021). Menurut Chu, pada GLRLM

terdapat 2 fitur yaitu LGRE dan HGRE (Vaishali et al., 2021). Menurut Dasarathy dan Holder, pada GLRLM terdapat 4 fitur yaitu SRHGE, SRLGE, LRLGE, dan LRHGE (Rahim and Foozy, 2020).



Gambar 2.5 Ilustrasi Sudut Orientasi pada GLRLM

2.4.1. Short Run Emphasis (SRE)

Nilai yang dihasilkan fitur SRE ini bergantung pada jumlah dari *short run*. Oleh karena itu nilai SRE akan berbanding terbalik dengan *run length*, maka dari itu jumlah setiap elemen dari *run length* yang dibagi dengan letak kolomnya. Pada tekstur kasar hasil yang diharapkan yaitu bernilai kecil, dan pada tekstur halus hasil yang diharapkan yaitu bernilai besar (Sari et al., 2021).

$$SRE = \frac{1}{s} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{p(i,j)}{j^2}$$
 (2.5)

Di mana P(i,j) adalah piksel-piksel dari hasil matriks run length yang berukuran $M \times N$, dengan i,j merupakan baris dan kolom. Sedangkan s adalah jumlah nilai dari run length yang didapatkan pada arah tertentu.

2.4.2. Long Run Emphasi (LRE)

Nilai yang dihasilkan fitur LRE ini bergantung pada jumlah dari *long run*. Oleh karena itu nilai LRE akan berbanding lurus dengan *run length*, maka dari itu jumlah setiap elemen dari *run length* yang dikali dengan letak kolomnya. Pada tekstur kasar hasil yang diharapkan yaitu bernilai besar, dan pada tekstur halus hasil yang diharapkan yaitu bernilai kecil (Sari et al., 2021).

$$LRE = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} j^{2} P(i, j)$$
 (2.6)

2.4.3. Gray-Level Nonuniformity (GLN)

Nilai yang dihasilkan fitur GLN ini bergantung pada nilai keabu-abuan suatu citra dan hasil yang diharapkan memiliki nilai yang kecil jika nilai dari setiap derajat keabu-abuan serupa pada seluruh piksel pada citra (Astiningrum et al., 2020). *Nonuniformity* atau ketidakseragaman dalam statistik yaitu varian atau standar deviasi, dimana untuk mengukur keragaman dari data. Pada varian dibutuhkan rata-rata data untuk mengetahui keragaman data, namun pada fitur GLN menggunakan jumlah dari setiap baris matriks *run length* yang merupakan representasi keragaman nilai *grayscale* dari citra.

$$GLN = \frac{1}{s} \sum_{i=1}^{M} \left(\sum_{j=1}^{N} P(i,j) \right)^{2}$$
 (2.7)

2.4.4. Run Length Nonuniformity (RLN)

Nilai yang dihasilkan fitur RLN ini bergantung pada nilai run length suatu citra dan hasil yang diharapkan memiliki nilai yang kecil jika nilai dari setiap

elemen run length serupa pada seluruh piksel pada citra (Astiningrum et al., 2020). *Nonuniformity* atau ketidakseragaman dalam statistik yaitu varian atau standar deviasi, dimana untuk mengukur keragaman dari data. Pada varian dibutuhkan rata-rata data untuk mengetahui keragaman data, namun pada fitur RLN menggunakan jumlah dari setiap kolom matriks *run length* yang merupakan representasi keragaman nilai *run* dari *grayscale* citra.

$$RLN = \frac{1}{s} \sum_{j=1}^{N} \left(\sum_{i=1}^{M} P(i,j) \right)^{2}$$
 (2.8)

2.4.5. Run Percentage (RP)

Nilai yang dihasilkan fitur RP ini bergantung pada kesamaan dan pembagian run pada arah tertentu dari suatu citra dan hasil yang diharapkan bernilai besar jika pada derajat keabu-abuan pada arah tertentu memiliki run length yang bernilai 1 (Sari et al., 2021).

$$RP = \frac{s}{\sum_{i=1}^{M} \sum_{j=1}^{N} j^2 P(i,j)}$$
 (2.9)

2.4.6. Low Gray-Level Run Emphasis (LGRE)

Nilai yang dihasilkan fitur LGRE ini bergantung pada distribusi dari derajat keabu-abuan suatu citra yang rendah, di mana hasil akan bernilai semakin besar jika semakin rendah nilai dari derajat keabu-abuan suatu citra (Mu'jizah, 2021). Fitur LGRE menunjukkan nilai dari tingkat rendahnya Grayscale dari citra, maka dari itu jumlah setiap elemen dari *run length* yang dibagi dengan letak barisnya.

$$LGRE = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)}{i^2}$$
 (2.10)

2.4.7. High Gray-Level Run Emphasis (HGRE)

Nilai yang dihasilkan fitur HGRE ini bergantung pada distribusi dari derajat keabu-abuan suatu citra yang tinggi, di mana hasil akan bernilai semakin besar jika semakin tinggi nilai dari derajat keabu-abuan suatu citra (Mu'jizah, 2021). Fitur HGRE menunjukkan nilai dari tingkat tingginya Grayscale dari citra, maka dari itu jumlah setiap elemen dari *run length* yang dikali dengan letak barisnya.

$$HGRE = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} i^{2} P(i, j)$$
 (2.11)

2.4.8. Short Run Low Gray-Level Emphasis (SRHGE)

Nilai yang dihasilkan fitur SRHGE ini bergantung pada distribusi relatif yang didapatkan dari short run dan derajat keabu-abuan suatu citra yang rendah. Hasil akan bernilai besar jika banyak dari short run memiliki nilai derajat keabu-abuan yang rendah (Jumhuriyah, 2021). Fitur SRHGE menunjukkan nilai dari short run yang memiliki tingkat rendahnya Grayscale citra, dimana fitur ini menggabungkan antara perhitungan pada fitur SRE dan LGRE. Oleh karena itu jumlah setiap elemen dari *run length* yang dibagi dengan letak baris dan kolomnya.

$$SRHGE = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)}{i^2 j^2}$$
 (2.12)

2.4.9. Short Run High Gray-Level Emphasis (SRLGE)

Nilai yang dihasilkan fitur SRLGE ini bergantung pada distribusi relatif yang didapatkan dari short run dan derajat keabu-abuan suatu citra yang tinggi. Hasil akan bernilai besar jika banyak dari short run memiliki nilai derajat keabu-abuan

yang tinggi (Jumhuriyah, 2021). Fitur SRHGE menunjukkan nilai dari short run yang memiliki tingkat tingginya Grayscale citra, dimana fitur ini menggabungkan antara perhitungan pada fitur SRE dan HGRE. Oleh karena itu jumlah setiap elemen dari *run length* yang dikali dengan letak baris dan dibagi dengan letak kolomnya.

$$SRLGE = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)i^{2}}{j^{2}}$$
 (2.13)

2.4.10. Long Run Low Gray-Level Emphasis (LRLGE)

Nilai yang dihasilkan fitur LRLGE ini bergantung pada distribusi relatif yang didapatkan dari long run dan derajat keabu-abuan suatu citra yang rendah. Hasil akan bernilai besar jika banyak dari long run memiliki nilai derajat keabu-abuan yang rendah (Jumhuriyah, 2021). Fitur LRLGE menunjukkan nilai dari long run yang memiliki tingkat rendahnya Grayscale citra, dimana fitur ini menggabungkan antara perhitungan pada fitur LRE dan LGRE. Oleh karena itu jumlah setiap elemen dari *run length* yang dikali dengan letak kolom dan dibagi dengan letak barisnya.

$$LRLGE = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)j^{2}}{i^{2}}$$
 (2.14)

2.4.11. Long Run High Gray-Level Emphasis (LRHGE)

Nilai yang dihasilkan fitur LRHGE ini bergantung pada distribusi relatif yang didapatkan dari long run dan derajat keabu-abuan suatu citra yang tinggi. Hasil akan bernilai besar jika banyak dari long run memiliki nilai derajat keabu-abuan yang tinggi (Jumhuriyah, 2021). Fitur LRHGE menunjukkan nilai dari long run yang memiliki tingkat tingginya Grayscale citra, dimana fitur ini

menggabungkan antara perhitungan pada fitur SRE dan LGRE. Oleh karena itu jumlah setiap elemen dari *run length* yang dibagi dengan letak baris dan kolomnya.

$$LRHGE = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} (P(i,j)) i^{2} j^{2}$$
 (2.15)

2.5. Principal Component Analysis (PCA)

Principal Component Analysis merupakan salah satu dari metode reduksi fitur biasanya digunakan untuk menyederhanakan suatu data, di mana data akan ditransformasikan secara linear dan membentuk nilai-nilai koordinat baru dengan varians yang maksimum (Rahayu et al., 2017). Keunggulan dari metode PCA dibanding dengan metode seleksi, di mana PCA tidak mengurangi jumlah dari fitur yang dianggap tidak penting melainkan metode ini dengan cara mengombinasikan inti yang terdapat dalam fitur dengan membentuk hasil subset dari fitur yang berukuran lebih kecil (Wibawa and Novianti, 2017). Langkah-langkah dari metode PCA sebagai berikut.

1. Siapkan data fitur X yang berukuran matriks $m \times n$, dimana m merupakan jumlah dari data yang digunakan, dan n merupakan jumlah fitur setiap datanya (Rosyani, 2017).

$$X(m,n) = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & x_{m,3} & \cdots & x_{m,n} \end{bmatrix}$$
(2.16)

2. Hitung hasil rata-rata setiap fitur, tahap ini akan menghasilkan matriks $m \times 1$

$$X(\bar{m}, 1) = \begin{bmatrix} \frac{x_{1,1} + x_{1,2} + x_{1,3} + \dots + x_{1,n}}{n} \\ \frac{x_{2,1} + x_{2,2} + x_{2,3} + \dots + x_{2,n}}{n} \\ \vdots \\ \frac{x_{m,1} + x_{m,2} + x_{m,3} + \dots + x_{m,n}}{n} \end{bmatrix} = \begin{bmatrix} \overline{x}_1 \\ \overline{x}_2 \\ \vdots \\ \overline{x}_m \end{bmatrix}$$
(2.17)

3. Lakukan teknik perhitungan zore-mean, hal ini dilakukan agar data memiliki distribusi normal. Jika data memiliki kedekatan dengan jumlah populasi maka dapat dikatakan data tersebut semakin dekat dengan distribusi normalnya. Rumus dari teknik perhitungan zero-mean dapat dilihat pada persamaan 2.18 sebagai berikut.

$$\hat{X} = X - \bar{X} = \begin{bmatrix} x_{1,1} - \overline{x_1} & x_{1,2} - \overline{x_1} & \cdots & x_{1,n} - \overline{x_1} \\ x_{2,1} - \overline{x_2} & x_{2,2} - \overline{x_2} & \cdots & x_{2,n} - \overline{x_2} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} - \overline{x_m} & x_{m,2} - \overline{x_m} & \cdots & x_{m,n} - \overline{x_m} \end{bmatrix}$$
(2.18)

Di mana \bar{X} merupakan hasil dari perhitungan zero-mean, X merupakan data fitur berukuran $m \times n$, dan \bar{X} merupakan nilai rata-rata dari setiap kolom X.

4. Mencari nilai *covarians* dari hasil yang didapatkan dari perhitungan zeromean, rumus dari covarians dapat dilihat pada persamaan 2.19.

$$C = \hat{X}^{T}.\hat{X} = \begin{bmatrix} \hat{x}_{1,1} & \hat{x}_{1,2} & \cdots & \hat{x}_{1,m} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{x}_{n,1} & \hat{x}_{n,2} & \cdots & \hat{x}_{n,m} \end{bmatrix} \cdot \begin{bmatrix} \hat{x}_{1,1} & \cdots & \hat{x}_{1,n} \\ \hat{x}_{2,1} & \cdots & \hat{x}_{2,n} \\ \vdots & \ddots & \vdots \\ \hat{x}_{m,1} & \cdots & \hat{x}_{m,n} \end{bmatrix}$$
(2.19)

Di mana C merupakan hasil dari covarians, dan \hat{X}^T merupakan matriks transpose dari matriks \hat{X} .

5. Mencari hasil *eigen value* dan *eigen vector* dengan menggunakan nilai yang didapatkan dari *covarians*, dimana matriks C berukuran $n \times n$. Untuk mencari nilai *eigen vector* dan *eigen value* diperlukan setiap skalar λ yang memenuhi persamaan 2.20

$$Cv = \lambda v \tag{2.20}$$

Vektor v menyatakan eigen vector dan skalar λ menyatakan eigen value, nilainilai ini dinyatakan sebagai persamaan karakteristik roots dan vektor dapat dilihat pada persamaan 2.21 sebagai berikut.

Untuk setiap nilai dari skalar λ atau *eigen value* harus memenuhi jika dan hanya jika nilai determinan dari persamaan 2.22 bernilai nol (Pujianto and Rahmawati, 2019). Dimana persamaan ini akan dijabarkan menjadi fungsi polinomial yang disebut sebagai polinomial karakteristik dari C yang dapat

dilihat pada persamaan 2.23

$$\begin{vmatrix} C - \lambda I \end{vmatrix} = 0$$

$$\begin{vmatrix} \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,n} \end{vmatrix} - \lambda \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} = 0$$

$$\begin{vmatrix} \begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,n} \end{vmatrix} - \begin{vmatrix} \lambda & 0 & \cdots & 0 \\ 0 & \lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda \end{vmatrix} = 0$$

$$\begin{vmatrix} c_{1,1} - \lambda & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} - \lambda & \cdots & c_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,n} - \lambda \end{vmatrix} = 0$$

$$\begin{vmatrix} c_{1,1} - \lambda & c_{1,2} & \cdots & c_{1,n} \\ c_{2,1} & c_{2,2} - \lambda & \cdots & c_{2,n} \\ \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,n} - \lambda \end{vmatrix} = 0$$
The persamaan di atas maka utk menentukan nilai eigen dapat

Berdasarkan persamaan di atas maka utk menentukan nilai eigen dapat diperoleh dengan cara mencari determinan matriks sehingga diperoleh persamaan karakteristik 2.23.

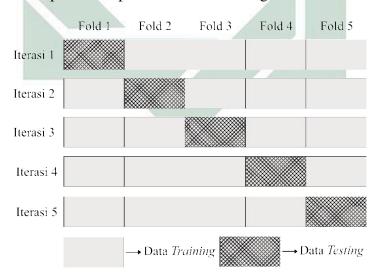
$$(c_1 - \lambda)(c_2 - \lambda) \cdots (c_n - \lambda) = 0$$

$$\lambda^n + c_1 \lambda^{n-1} + c_2 \lambda^{n-2} + \cdots + c_n = 0$$
(2.23)

6. Hasil dari eigen value akan diurutkan dari nilai yang paling kecil hingga yang paling besar, dan *eigen vector* yang berupa matriks $m \times m$ akan dibalik dari kiri ke kanan sesuai sumbu vertikal (Saputra, 2021).

2.6. K-Fold Cross Validation

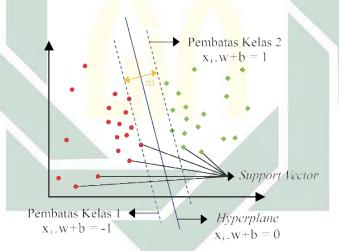
Sistem klasifikasi memerlukan data yang digunakan untuk training dan testing, di mana data training berfungsi sebagai masukan untuk mencari pola dari data dan data testing berfungsi sebagai masukan untuk mendapatkan hasil klasifikasi menggunakan pola yang telah didapatkan (Muhidin and Burhan, 2018). Salah satu metode yang digunakan untuk membagi suatu data training dan testing yaitu metode K-Fold Cross Validation. Selain itu metode ini juga berguna untuk menilai kinerja dari sistem klasifikasi dengan membagi data secara acak dan data akan dikelompokkan sebanyak nilai K, dengan nilai K lebih dari 1 dan tidak melebihi banyak dari data (Cahyanti et al., 2020; Rhomadhona and Permadi, 2019). Pengelompokan data ini dilakukan dengan partisi data sebanyak K-kali iterasi dengan hasil partisi data ke-K merupakan data testing dan sisanya merupakan data training (Supartini et al., 2017). Ilustrasi dari metode K-Fold Cross Validation dapat dilihat pada Gambar 2.6 sebagai berikut.



Gambar 2.6 Ilustrasi K-Fold Cross Validation

2.7. Support Vector Machine (SVM)

Salah satu metode yang termasuk dalam *supervised learning*, yaitu *Support Vector Machine*. Selain itu SVM juga termasuk pada kelas ANN (*Artifical Neural Network*), di mana metode ini berfungsi menganalisis suatu data untuk mengenali pola dalam pengklasifikasian suatu data (Fibrianda and Bhawiyuga, 2018). Ide dasar dari metode SVM ini pada pencarian *hyperplane* terbaik dari setiap titik koordinat data, di mana *hyperplane* ini berperan sebagai garis batas dari setiap kelas klasifikasi. Dalam klasifikasi dua kelas, *hyperplane* merupakan garis pemisah antara kelas +1 dan −1 dapat dilihat pada Gambar 2.7 (Rahutomo et al., 2018).



Gambar 2.7 Klasifikasi SVM dua kelas

Pada metode SVM, data dinotasikan sebagai $x_i \in \mathbb{R}^d$ dan label pada setiap data dinotasikan sebagai $y_i \in -1, 1$ untuk $i=1,2,3,\cdots,l$ dengan l merupakan banyak data yang digunakan. Gambar 2.7 telah mengilustrasikan *hyperplane* (Arifin and Sasongko, 2018; Hermanto et al., 2020), maka dapat diperoleh pembatas antara

dua kelas yaitu pada persamaaan 2.7.

$$x_i.w + b \ge 1, \ untuk \ y_i = 1$$

$$x_i.w + b \le -1, \ untuk \ y_i = -1$$
 (2.24)

Dimana w merupakan bobot dan b merupakan bias. Dengan bantuan persamaaan 2.24 maka pencarian hyperplane terbaik dapat diformulasikan seperti pada persamaan 2.25. Dimana c merupakan konstanta yang mengontrol jarak antara margin, dan ω merupakan nilai toleransi untuk mencari garis pemisah.

$$min\frac{1}{2}||w||^2 + c\left(\sum_{i=1}^{l}\omega_i\right)$$
 (2.25)

Dimana fungsi pemisah seperti pada persamaan 2.26

$$y_i(x_i.w + b) \ge 1 - \omega, \ \forall i = 1, 2, \cdots, l$$
 (2.26)

Berdasarkan karakteristik, metode SVM dapat dibagi menjadi SVM *linear* dan *non-linear*. Di mana SVM *linear* adalah metode yang memisahkan data secara *linear* atau dapat dikatakan memisahkan antara dua kelas klasifikasi dengan *soft margin* (Abbas, 2016). Sedangkan SVM non-linear yaitu metode yang memisahkan data yang memiliki dimensi tinggi, metode ini biasanya menggunakan fungsi kernel untuk menangani data tersebut. Macam dari fungsi kernel tersebut antara lain RBF (*Radial Basis Function*), *Linear*, dan *Polynomial*

yang dapat dilihat pada persamaan 2.27 (Chazar and Erawan, 2020).

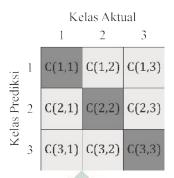
$$RBF \to f(x,y) = exp(-||x-y||^2)$$

 $Linear \to f(x,y) = x^T.y$ (2.27)
 $Polynomial \to f(x,y) = (1+x^T.y)^d$

Di mana f(x,y) merupakan nilai dari kernel yang dihasilkan, dengan x,y merupakan baris dan kolom dari vektor masukan. Selain itu x^T merupakan vektor transpose dari vektor x, dan d merupakan derajat polinomial (Jumeilah, 2017).

2.8. Confusion Matrix

Confusion Matrix merupakan metode evaluasi untuk mengetahui informasi terkait data aktual yang digunakan dan hasil prediksi dari sistem klasifikasi. Hasil yang diharapkan dari pengklasifikasian yaitu memiliki tingkat kesalahan yang rendah dan hasil prediksi dari klasifikasi yaitu tepat pada target, oleh karena itu pada sistem klasifikasi memerlukan evaluasi sistem untuk mengetahui tingkat dari keberhasilan suatu sistem klasifikasi (Rini Novitasari et al., 2020). Evaluasi sistem ini dapat dilihat hasil perhitungan dari confusion matrix yang terdiri dari akurasi, sensitivitas, dan spesifisitas. Evaluasi ini didapatkan dari menghitung *True Positive* (TP), *False Positive* (FP), *False Negative* (FN), dan *True Negative* (TN) (Foeady et al., 2018).



Gambar 2.8 Confusion Matrix dengan Tiga Kelas Klasifikasi

2.8.1. True Positive (TP)

Jika diimplementasi pada pengklasifikasian penyakit kanker payudara, *True Positive* (TP) merupakan nilai dari hasil klasifikasi pasien positif kanker yang teridentifikasi benar atau sesuai dengan targetnya yaitu positif kanker payudara. Rumus dari TP dapat dilihat pada persamaan 2.28 sebagai berikut.

$$tp = \begin{bmatrix} C(1,1) \\ C(2,2) \\ \vdots \\ C(m,m) \end{bmatrix}$$
 (2.28)

Di mana C(i,j) merupakan $confusion\ matrix$ dapat dilihat pada Gambar 2.8, dengan $i,j\in 1,2,\cdots,m$ yang merupakan baris dan kolom dari matriks C dan m adalah jumlah kelas.

2.8.2. False Positive (FP)

False Positive (FP) merupakan nilai dari hasil klasifikasi pasien negatif kanker yang teridentifikasi salah atau tidak sesuai dengan targetnya yaitu positif kanker payudara. Rumus dari FP dapat dilihat pada persamaan 2.29 sebagai

berikut.

$$tp = \begin{bmatrix} \sum_{i=1}^{m} C(i,1) \\ \sum_{i=1}^{m} C(i,2) \\ \vdots \\ \sum_{i=1}^{m} C(i,m) \end{bmatrix} - tp$$
 (2.29)

2.8.3. False Negative (FN)

False Negative (FN) merupakan nilai dari hasil klasifikasi pasien positif kanker yang teridentifikasi salah atau tidak sesuai dengan targetnya yaitu negatif kanker payudara. Rumus dari FN dapat dilihat pada persamaan 2.30 sebagai berikut.

$$tp = \begin{bmatrix} \sum_{i=1}^{m} C(1,i) \\ \sum_{i=1}^{m} C(2,i) \\ \vdots \\ \sum_{i=1}^{m} C(m,i) \end{bmatrix} - tp$$
 (2.30)

2.8.4. True Negative (TN)

True Negative (TN) merupakan nilai dari hasil klasifikasi pasien negatif kanker yang teridentifikasi benar atau sesuai dengan targetnya yaitu negatif kanker payudara (Mu'jizah, 2021). Rumus dari TN dapat dilihat pada persamaan 2.31 sebagai berikut.

$$\sum_{i=1}^{m} \sum_{j=1}^{m} C(i,j) - (tp + fp + fn)$$
 (2.31)

2.8.5. Akurasi

Akurasi merupakan nilai dari evaluasi yang menunjukkan seberapa tepat hasil dari sistem klasifikasi terhadap data yang digunakan (Maricar and Dian

Pramana, 2019). Rumus dari akurasi dapat dilihat pada persamaan 2.32 sebagai berikut.

$$Akurasi = \frac{\sum_{i=1}^{m} tp(i)}{\sum_{i=1}^{m} \sum_{j=1}^{m} C(i,j)}$$
 (2.32)

2.8.6. Sensitivitas

Sensitivitas merupakan nilai dari evaluasi yang menunjukkan hasil dari klasifikasi yaitu pasien yang dinyatakan sakit. Dalam pengklasifikasian penyakit, sensitivitas lebih diprioritaskan daripada spesifisitas. Karena jika pasien dinyatakan sakit maka tingkat kewaspadaan pasien akan meningkat dan pasien akan diarahkan untuk pemeriksaan lebih lanjut (Astuti and Adiwijaya, 2019). Rumus dari sensitivitas dapat dilihat pada persamaan 2.33 sebagai berikut.

$$Sensitifitas = \frac{\sum_{i=1}^{m} \frac{tp(i)}{tp(i) + fn(i)}}{m}$$
 (2.33)

2.8.7. Spesifisitas

Spesifisitas merupakan nilai dari evaluasi yang menunjukkan hasil dari klasifikasi yaitu pasien yang dinyatakan tidak sakit (Astuti and Adiwijaya, 2019). Rumus dari spesifitas dapat dilihat pada persamaan 2.34 sebagai berikut.

$$Sensitifitas = \frac{\sum_{i=1}^{m} \frac{tn(i)}{tn(i) + fp(i)}}{m}$$
 (2.34)

2.9. Integrasi Keislaman

Suatu penyakit merupakan salah satu pemberian dari Allah SWT kepada hamba-Nya. Penyakit juga termasuk bentuk dari cobaan yang Allah SWT berikan untuk menguji keimanan hamba-Nya. Setiap penyakit yang Allah berikan, akan

diberikan pula kesembuhan untuk hamba-Nya. Hal ini sesuai pada hadits Riwayat Bukhari, Rasulullah saw bersabda, "Setiap penyakit ada obatnya. Apabila ditemukan obat yang tepat untuk suatu penyakit, akan sembuhlah penyakit itu dengan izin Allah 'azza wajalla." (HR. Muslim).

Sebelum penyakit diturunkan oleh Allah SWT, manusia hendaklah menjaga kesehatan. Pentingnya menjaga kesehatan tertera pada hadis dari Ibnu 'Abbas radhiyallahu'anhuma. Seseorang pernah mendapat nasehat dari Rasulullah yaitu: "Manfaatkanlah lima perkara sebelum lima perkara, waktu mudamu sebelum datang waktu tuamu, waktu sehatmu sebelum datang waktu sakitmu, waktu kayamu sebelum datang waktu miskinmu, waktu luangmu sebelum datang waktu sempitmu, waktu hidupmu sebelum datang waktu matimu" (HR. Al Hakim kitab Al-Mustadrok).

Dari hadis ini dapat ditarik kesimpulan bahwa sakit bukan hal yang tidak dapat dihindari akan tetapi pencegahan dini penting dilakukan sebelum sakit lebih parah. Selain itu penting umat manusia kuat dan semangat jika diberikan sakit atau cobaan oleh Allah SWT. Hal ini sesuai dengan sabda Rasulullah shallallahu 'alaihi wasallam, beliau menyampaikan, "Bersungguh-sungguhlah untuk mendapatkan apa yang bermanfaat bagimu dan mintalah pertolongan kepada Allah ah dalam urusan serta janganlah sekali-kali engkau merasa lemah." (HR. Muslim).

إحْرِصْ عَلَى مَا يَنْفَعُكَ وَاسْتَعِنْ بِاللَّهِ وَلَا تَعْجَزْ

Hadits tersebut dapat menjadi motivasi untuk seseorang yang sakit dan yang tidak sakit, yaitu pada seseorang yang sakit harus tetap bersemangat dan tidak putus asa dengan keadaannya, sedangkan seseorang yang tidak sakit harus tetap bersemangat untuk membantu mencari pengobatan untuk seseorang yang sakit. Pengobatan yang di maksud yaitu dapat membantu meneliti suatu penyakit agar dapat teridentifikasi. Selain itu sebagai bentuk usaha yang dapat dilakukan oleh orang yang sakit mulai dari berdoa, menjaga pola makan dan melakukan pemeriksaan dini pada ahlinya. Jika semua usaha telah dilakukan Allah akan memberi kesembuhan bagi seseorang yang sakit akan sembuh dengan obat yang tepat melalui perantara dokter.

BAB III

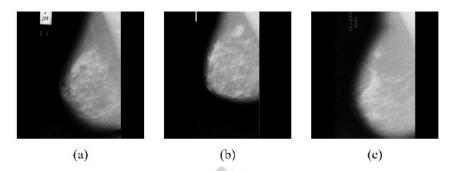
METODE PENELITIAN

3.1. Jenis Penelitian

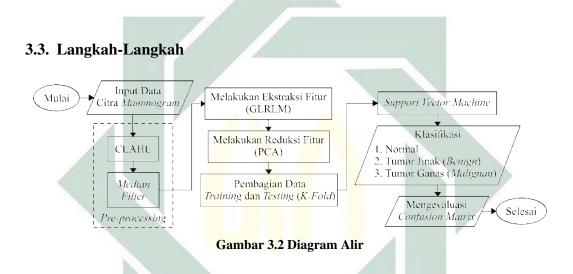
Penelitian terkait klasifikasi kanker payudara pada citra *mammogram* dengan menggunakan metode GLRLM-PCA (*Gray Level Run Length Matrix-Principal Component Analysis*) dan metode SVM (Support Vector Machine). Penelitian ini termasuk dalam jenis penelitian deskriptif kuantitatif, di mana penelitian ini memiliki unsur perhitungan dan juga analisis berdasarkan hasil evaluasi yang diperoleh. Hasil dari evaluasi sistem klasifikasi pada penelitian ini memiliki tujuan yaitu menjadi salah satu alternatif dalam pengklasifikasian penyakit kanker payudara pada citra *mammogram*.

3.2. Jenis dan Sumber Data

Penelitian ini menggunakan data citra mammogram yang berasal dari website Kaggle.com. Data terdiri dari tiga kelas data yaitu citra dari pasien normal, citra dari pasien tumor jinak (benign), dan citra dari pasien tumor ganas (malignant). Data keseluruhan berjumlah 322 citra, di mana masing-masing dari kelas yaitu citra dari pasien normal berjumlah 207, citra dari pasien tumor jinak berjumlah 64, dan citra dari pasien tumor ganas berjumlah 51.



Gambar 3.1 (a) Normal, (b) Benign, (c) Malignant



Pada diagram alir yang tercantum pada Gambar 3.2, tahapan-tahapan terkait diagram alir akan diperjelas sebagai berikut.

- Penelitian ini menggunakan 322 data citra *Mammogram*, di mana data ini diakses dari website Kaggle.com. Data akan bedakan menjadi dua kelompok dan diberi label dengan dua kelas, yaitu kelompok pertama 207 data normal dan 115 data Abnormal, dan kelompok keddua 64 data *benign*, dan 51 data *malignant*.
- Citra-citra yang berlabel akan dilakukan pre-processing yaitu CLAHE dan Median filter, dimana CLAHE berfungsi untuk meningkatkan kualitas citra dengan menambah kontras dari citra, dan Median filter berfungsi untuk

memperhalus citra setelah CLAHE dilakukan. CLAHE akan dilakukan dengan langkah-langkah yang telah dijelaskan pada subbab 2.4 dan Median filter akan dilakukan dengan langkah-langkah seperti yang telah dijelaskan pada subbab 2.5.

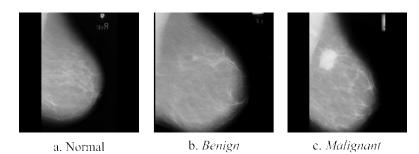
- 3. Ekstraksi fitur diambil dari hasil pre-processing dengan menggunakan metode GLRLM, di mana ada 11 fitur yang dihasilkan dengan 4 arah orientasi (0°, 45°, 90°, 135°). Masing-masing dari fitur dihitung dengan menggunakan persamaan 2.4 hingga persamaan 2.14, perhitungan ini dilakukan untuk setiap citra.
- 4. Hasil 11 fitur GLRLM akan dilakukan reduksi fitur dengan tujuan mengurangi fitur, yaitu menggunakan metode PCA.
- 5. Fitur yang dihasilkan dari reduksi fitur akan dilakukan pembagian data yaitu dengan menggunakan metode K-Fold seperti pada Gambar 2.6.
- 6. Klasifikasi akan dilakukan dengan memasukkan data yang telah dibagi menjadi 10 fold. Klasifikasi pada penelitian ini menggunakan metode SVM untuk mencari model terbaik dari masing-masing kelas.
- 7. Evaluasi dari hasil klasifikasi dilakukan menggunakan metode confusion matrix, hasil yang didapatkan yaitu akurasi, sensitivitas, dan spesifisitas. Hasil ini didapatkan dengan menggunakan rumus pada persamaan 2.27 hingga persamaan 2.28. Sebelum itu, untuk mendapatkan hasil evaluasi perlu dihitung nilai dari TP, FP, FN, dan TN pada rumus persamaan 2.23 hingga persamaan 2.26.

BAB IV

HASIL DAN PEMBAHASAN

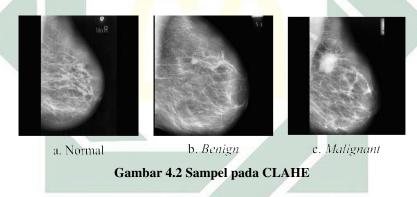
4.1. Preprocessing

Penelitian ini menggunakan data Mammogram dengan total yaitu 332 data, akan dibagi menjadi dua kelompok dimana data pengklasifikasian. Pengklasifikasian pertama, data akan diberi label dengan dua kelas yaitu 207 data normal dan 115 data abnormal. Data abnormal yang digunakan ini merupakan gabungan dari data benign (tumor jinak) dan data malignant (tumor ganas). Pengklasifikasian kedua, data akan diberi label dengan dua kelas yaitu 64 data benign dan 51 data malignant. Sebelum melakukan klasifikasi citra Mammrogram menggunakan metode SVM terlebih dahulu akan dilakukan tahap preprocessing. Tahap ini dilakukan untuk memperbaiki kualitas dari citra agar citra dapat memberikan banyak informasi. Pada penelitian ini citra Mammogram sudah memiliki ukuran yang sama yaitu setiap citra berukuran 1024x1024 piksel. Gambar 4.1 menunjukkan sampel data dari data normal, data benign, dan data malignant.

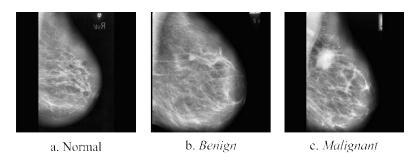


Gambar 4.1 Sampel citra Mammogram

Hal yang menyebabkan kualitas citra lemah yaitu pada intensitas cahaya yang bervariasi pada citra *Mammogram* sehingga citra memiliki kontras dan pencahayaan yang lemah. Peningkatan kontras dapat dilakukan dengan menggunakan metode Contrast Limited Adaptive Histogram Equalization (CLAHE). Meotode CLAHE dapat mempertajam warna-warna dari citra, dimana hal ini dapat memperjelas objek dari citra. Cara kerja metode ini yaitu dengan membagi citra menjadi beberapa piksel persegi, dimana piksel persegi dapat disebut blok. Setiap blok akan dihitung histogram untuk meratakan histogram yang terlalu tinggi, apabila terdapat histigram yang terlalu tinggi maka akan dibatasi dengan memberikan nilai batas yang dapat disebut *crip limit*. Hasil yang didapatkan dari metode ini dapat dilihat pada Gambar 4.2.



Selain itu, *noise* pada citra juga penyebab kualitas citra menjadi lemah maka perlu menghilangkan *noise* dengan menggunakan metode Median Filter. Metode ini akan berkerja dari pembagian piksel persegi dari setiap citra, yang kemudian akan diurutkan nilai-nilai dari piksel dan diambil nilai tengah dari piksel. Hasil yang didapatkan dari metode ini dapat dilihat pada Gambar 4.3.



Gambar 4.3 Sampel pada Median Filter

4.2. Ekstraksi Fitur

Ekstraksi Fitur merupakan salah satu proses yang berguna untuk mengambil ciri pada citra. dilakukan untuk mendapatkan Hal ini informasi-informasi penting untuk sistem klasifikasi. Ekstraksi fitur pada citra mammogram menggunakan GLRLM, dimana metode ini akan menganalisis tekstur dari citra dengan menggunakan matriks run length. Matriks run length dapat merepresentasikan total kemunculan panjang run setiap piksel dan tingkat keabuan terhadap setiap sudut orientasi. Penelitian ini menggunakan empat sudut orientasi untuk membentuk matriks run length yaitu sudut 0°, 45°, 90°, dan 1350°. Hasil dari matriks run length akan dilakukan perhitungan untuk 11 fitur pada GLRLM yaitu SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, dan LRHGE. Berikut proses yang dilakukan dalam ekstrkasi fitur pada citra mammogram menggunakan GLRLM. Contoh perhitungan akan dimulai dari mencari matriks run length pada setiap sudut orientasi, dimana nilai kembar dari piksel akan dihitung tergantung arah dari sudut orientasi. Berikut salah satu matriks run length yaitu pada sudut orientasi 0° dengan ukuran 255 × 1024, dimana baris merupakan nilai maksimal grayscale dan kolom merupakan panjang run.

$$RL_0 = \begin{bmatrix} 76 & 33 & 47 & \cdots & 0 & 0 & 0 \\ 315 & 165 & 137 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 12 & 5 & 6 & \cdots & 0 & 0 & 0 \\ 12 & 3 & 2 & \cdots & 0 & 0 & 0 \end{bmatrix}$$

Selanjutnya fitur-fitur pada GLRLM dihitung, perhitungan dengan sudut orientasi 0° sebagai berikut.

$$SRE_{0} = \frac{1}{s} \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{p(i,j)}{j^{2}}$$

$$= \frac{1}{261902} \left[\left(\frac{76}{1^{2}} + \frac{33}{2^{2}} + \frac{47}{3^{2}} + \dots + \frac{0}{1024^{2}} \right) + \left(\frac{315}{1^{2}} + \frac{165}{2^{2}} + \frac{137}{3^{2}} + \dots + \frac{0}{1024^{2}} \right) + \dots + \left(\frac{12}{1^{2}} + \frac{5}{2^{2}} + \frac{6}{3^{2}} + \dots + \frac{0}{1024^{2}} \right) + \left(\frac{12}{1^{2}} + \frac{3}{2^{2}} + \frac{2}{3^{2}} + \dots + \frac{0}{1024^{2}} \right) \right]$$

$$= 0.00204$$

$$LRE_{0} = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} j^{2} P(i,j)$$

$$= \frac{1}{261902} \left[\left((76 \times 1^{2}) + (33 \times 2^{2}) + (47 \times 3^{2}) + \dots + (0 \times 1024^{2}) \right) + \left((315 \times 1^{2}) + (117 \times 2^{2}) + (135 \times 3^{2}) + \dots + (0 \times 1024^{2}) \right) + \dots + \left((12 \times 1^{2}) + (5 \times 2^{2}) + (6 \times 3^{2}) + \dots + (0 \times 1024^{2}) \right) + \left((12 \times 1^{2}) + (2 \times 3^{2}) + \dots + (0 \times 1024^{2}) \right) \right]$$

$$= 17626.28950$$

$$GLN_0 = \frac{1}{s} \sum_{i=1}^{M} \left(\sum_{j=1}^{N} P(i,j) \right)^2$$

$$= \frac{1}{261902} \left[(76 + 33 + 47 + \dots + 0)^2 + (315 + 117 + 135 + \dots + 0)^2 + \dots + (12 + 5 + 6 + \dots + 0)^2 + (12 + 3 + 2 + \dots + 0)^2 \right]$$

$$= 129402.82195$$

$$RLN_0 = \frac{1}{s} \sum_{j=1}^{N} \left(\sum_{i=1}^{M} P(i,j) \right)^2$$

$$= \frac{1}{261902} \left[(76 + 315 + \dots + 12 + 12)^2 + (33 + 165 + \dots + 5 + 3)^2 + (47 + 137 + \dots + 6 + 2)^2 + \dots + (0 + 0 + \dots + 0 + 0)^2 \right]$$

$$= 1254.70673$$

$$RP_0 = \frac{s}{\sum_{i=1}^{M} \sum_{j=1}^{N} j^2 P(i,j)}$$

$$= \frac{s}{\left[((76 \times 1^2) + (33 \times 2^2) + (47 \times 3^2) + \dots + (0 \times 1024^2)) \right]}$$

$$= 1.00355$$

$$LGRE_0 = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)}{i^2}$$

$$= \frac{1}{261902} \left[\left(\frac{76}{4^2} + \frac{33}{4^2} + \frac{47}{4^2} + \dots + \frac{0}{4^2} \right) + \left(\frac{316}{5^2} + \frac{117}{5^2} + \frac{135}{5^2} + \dots + \frac{0}{1024^2} \right) + \dots + \left(\frac{12}{250^2} + \frac{5}{250^2} + \frac{6}{250^2} + \dots + \frac{0}{250^2} \right)$$

$$+ \left(\frac{12}{251^2} + \frac{3}{251^2} + \frac{2}{251^2} + \dots + \frac{0}{251^2} \right) \right]$$

$$= 0.73024$$

$$HGRE_{0} = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} i^{2}P(i,j)$$

$$= \frac{1}{261902} \left[\left((76 \times 4^{2}) + (33 \times 4^{2}) + (47 \times 4^{2}) + \dots + (0 \times 4^{2}) \right) + \left((315 \times 5^{2}) + (117 \times 5^{2}) + (135 \times 5^{2}) + \dots + (0 \times 5^{2}) \right) + \dots + \left((12 \times 250^{2}) + (5 \times 250^{2}) + (6 \times 250^{2}) + \dots + (0 \times 250^{2}) \right) + \left((12 \times 251^{2}) + (3 \times 251^{2}) + (2 \times 251^{2}) + \dots + (0 \times 251^{2}) \right) \right]$$

$$= 626.37570$$

$$SRHGE_{0} = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)}{i^{2}j^{2}}$$

$$= \frac{1}{261902} \left[\left(\frac{76}{4^{2} \times 1^{2}} + \frac{33}{4^{2} \times 2^{2}} + \frac{47}{4^{2} \times 3^{2}} + \dots + \frac{0}{4^{2} \times 1024^{2}} \right) + \dots + \left(\frac{12}{250^{2} \times 1^{2}} + \frac{5}{250^{2} \times 2^{2}} + \frac{135}{5^{2} \times 3^{2}} + \dots + \frac{0}{250^{2} \times 1024^{2}} \right) + \dots + \left(\frac{12}{251^{2} \times 1^{2}} + \frac{3}{251^{2} \times 2^{2}} + \frac{2}{251^{2} \times 3^{2}} + \dots + \frac{0}{250^{2} \times 1024^{2}} \right) + \dots + \left(\frac{12}{251^{2} \times 1^{2}} + \frac{3}{251^{2} \times 2^{2}} + \frac{2}{251^{2} \times 3^{2}} + \dots + \frac{0}{251^{2} \times 1024^{2}} \right) \right]$$

$$= 0.00056$$

$$SRLGE_{0} = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)i^{2}}{j^{2}}$$

$$= \frac{1}{261902} \left[\left(\frac{76 \times 4^{2}}{1^{2}} + \frac{33 \times 4^{2}}{2^{2}} + \frac{47 \times 4^{2}}{3^{2}} + \dots + \frac{0 \times 4^{2}}{1024^{2}} \right) + \dots + \left(\frac{12 \times 250^{2}}{1^{2}} + \frac{117 \times 5^{2}}{2^{2}} + \frac{135 \times 5^{2}}{3^{2}} + \dots + \frac{0 \times 4^{2}}{1024^{2}} \right) + \dots + \left(\frac{12 \times 250^{2}}{1^{2}} + \frac{5 \times 250^{2}}{2^{2}} + \frac{6 \times 250^{2}}{3^{2}} + \dots + \frac{0 \times 250^{2}}{1024^{2}} \right) + \dots + \left(\frac{12 \times 250^{2}}{1^{2}} + \frac{5 \times 250^{2}}{2^{2}} + \frac{6 \times 250^{2}}{3^{2}} + \dots + \frac{0 \times 250^{2}}{1024^{2}} \right) + \dots + \left(\frac{12 \times 251^{2}}{1024^{2}} + \frac{3 \times 251^{2}}{2^{2}} + \frac{2 \times 251^{2}}{3^{2}} + \dots + \frac{0 \times 251^{2}}{1024^{2}} \right) \right]$$

$$= 24.88397$$

$$LRLGE_{0} = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} \frac{P(i,j)j^{2}}{i^{2}}$$

$$= \frac{1}{261902} \left[\left(\frac{76 \times 1^{2}}{4^{2}} + \frac{33 \times 2^{2}}{4^{2}} + \frac{47 \times 3^{2}}{4^{2}} + \dots + \frac{0 \times 1024^{2}}{4^{2}} \right) + \left(\frac{316 \times 1^{2}}{5^{2}} + \frac{117 \times 2^{2}}{5^{2}} + \frac{135 \times 3^{2}}{5^{2}} + \dots + \frac{0 \times 1024^{2}}{5^{2}} \right) + \dots + \left(\frac{12 \times 1^{2}}{250^{2}} + \frac{5 \times 2^{2}}{250^{2}} + \frac{6 \times 3^{2}}{250^{2}} + \dots + \frac{0 \times 1024^{2}}{250^{2}} \right) + \left(\frac{12 \times 1^{2}}{251^{2}} + \frac{3 \times 2^{2}}{251^{2}} + \frac{2 \times 3^{2}}{251^{2}} + \dots + \frac{0 \times 1024^{2}}{251^{2}} \right) \right]$$

$$= 1.36818$$

$$LRHGE_{0} = \frac{1}{s} \sum_{i=1}^{M} \sum_{j=1}^{N} (P(i,j)) i^{2} j^{2}$$

$$= \frac{1}{261902} \left[\left(\left(76 \times 4^{2} \times 1^{2} \right) + \left(33 \times 4^{2} \times 2^{2} \right) + \left(47 \times 4^{2} \times 3^{2} \right) + \dots + \left(0 \times 4^{2} \times 1024^{2} \right) \right) + \left(\left(315 \times 5^{2} \times 1^{2} \right) + \left(117 \times 5^{2} \times 2^{2} \right) + \left(135 \times 5^{2} \times 3^{2} \right) + \dots + \left(0 \times 5^{2} \times 1024^{2} \right) + \dots + \left(\left(12 \times 250^{2} \times 1^{2} \right) + \left(\left(12 \times 250^{2} \times 1^{2} \right) + \left(3 \times 251^{2} \times 2^{2} \right) + \left(2 \times 251^{2} \times 3^{2} \right) + \dots + \left(0 \times 251^{2} \times 1024^{2} \right) \right) \right]$$

$$= 64325.33521$$

Perhitungan ini akan dilakukan juga pada setiap data citra *mammogram*, dan pada setiap sudut orientasi 0°, 45°, 90°, dan 1350°. Tabel 4.1 adalah hasil ekstraksi fitur GLRLM dengan sudut orientasi 0° pada kelas Normal-Abnormal, dan Tabel 4.2 adalah ekstraksi fitur GLRLM dengan sudut orientasi 0° pada kelas Benign-Malignant. Berdasarkan analisis terhadap 11 fitur pada GLRLM terdapat ciri khusus pada setiap fitur. Pada fitur SRE pada kelas normal memiliki nilai yang

Tabel 4.1 Ekstraksi Fitur Kelas Normal-Abnormal pada Sudut Orientasi 0°

Fitur					
	1	2	3	•••	322
SRE	0,002294456	0,001947078	0,000847757	•••	0,001993497
LRE	19119,50338	20746,69474	19893,35841	•••	18072,05346
GLN	132335,5085	154371,6327	176476,413	•••	163030,6263
RLN	1634,415757	1674,430977	2972,767301	•••	1980,382301
RP	1,103151808	1,244500613	1,465750834	•••	1,338752298
LGRE	0,704643412	0,716319151	0,706919332	•••	0,709516701
HGRE	334,2216452	278,2307456	240,7275073		221,0709862
SRLGE	0,000805785	0,000645304	0,000318008	1	0,000658955
SRHGE	12,95305726	10,82575365	9,475032033	•••	8,517771409
LRLGE	13959,82969	154 <mark>02,77</mark> 883	14373,05239	\	13695,66113
LRHGE	81943,98048	82 <mark>84</mark> 6,0 <mark>37</mark> 82	7 <mark>3690,3</mark> 9860		61023,20368

Tabel 4.2 Ekstraksi Fitur Kelas Benign-Malignant pada Sudut Orientasi 0°

Eitm		Data ke-								
Fitur	1	2	3	•••	115					
SRE	0,002008969	0,001379153	0,000779376		0,000882371					
LRE	17630,53753	18200,94682	19226,46316		18937,15500					
GLN	129124,8187	154213,6674	171297,6936		245141,5606					
RLN	1254,547436	1498,104837	2731,724477		3037,710109					
RP	1,003102022	1,205809589	1,438591452		1,898387714					
LGRE	0,729629718	0,727304465	0,700151905		0,730540109					
HGRE	626,1178941	430,7957607	246,9267314		71,68869917					
SRLGE	0,000541025	0,000460932	0,000279245		0,000295375					
SRHGE	24,87775861	16,98776455	9,714158824		2,282037945					
LRLGE	13687,51945	13651,01987	13976,08378		14031,39702					
LRHGE	64433,49716	74238,01046	70446,77489		750470,5910					

cenderung kecil, dan bernilai lebih besar pada kelas *malignant*. dikarenakan semakin ganas suatu tumor maka semakin kasar tektur dari citra, serta semakin banyak run length yang dihasilkan dan semakin kecil pula nilai SRE. Pada fitur LRE pada kelas normal memiliki nilai yang cenderung besar, dan bernilai lebih kecil pada kelas *malignant*. Hal ini dikarenakan semakin ganas suatu tumor maka semakin halus tektur dari citra, serta semakin sedikit run length yang dihasilkan dan semakin kecil pula nilai LRE. Fitur GLN pada kelas normal cenderung memiliki derajat keabuan yang serupa pada seluruh piksel pada citra, sehingga memiliki nilai yang kecil. Kemudian sebaliknya pada kelas malignant cendurung sedikit memiliki kesamaan derajat keabuan dengan seluruh piksel citra, sehingga memiliki nilai yang lebih besar. Pada kelas benign memiliki kesamaan derajat keabuan yang lebih b<mark>an</mark>yak dibanding kelas *malignant*, sehingga nilai yang dihasilkan lebih kecil dibanding kelas malignant. Fitur RLN pada kelas benign cenderung memiliki nilai yang kecil sehingga elemen run length yang serupa dengan seluruh piksel pada citra, sedangkan kelas normal cenderung memiliki nilai yang besar sehingga sedikit memiliki elemen run length yang serupa dengan seluruh piksel pada citra. Fitur RP nilai yang dihasilkan pada citra malignant cenderung banyak memiliki run length bernilai 1 untuk seluruh tingkat keabuan, sedangkan banign cenderung sedikit memiliki run length bernilai 1 untuk seluruh tingkat keabuan. Fitur LGRE pada kelas normal memiliki nilai yang cenderung kecil, dan bernilai lebih besar pada kelas malignant. Hal ini dikarenakan semakin rendah derajat keabuan dari citra maka akan semakin besar nilai LGRE. Citra pada kelas malignant cenderung memiliki derajat keabuan yang rendah, sedangkan citra pada kelas normal cenderung memiliki derajat keabuan yang tinggi. Fitur HGRE pada kelas malignant cenderung memiliki derajat keabuan yang rendah, sedangkan citra pada kelas benign cenderung memiliki derajat keabuan yang tinggi. Fitur SRLGE pada kelas malignant cenderung memiliki run length yang panjang dan memiliki derajat keabuan yang tinggi, maka dari itu nilai yang dihasilkan SRLGE cenderung kecil. Sedangkan pada kelas benign cenderung memiliki run length yang panjang dan memiliki derajat keabuan yang rendah, maka nilai dari SRLGE lebih besar. Fitur SRLGE pada kelas normal cenderung memiliki run length yang sedikit dan memiliki derajat keabuan yang tinggi, maka nilai dari SRLGE lebih besar. Fitur LRLGE dan LRHGE cenderung memiliki nilai yang lebih besar ketika tumor semakin ganas, pada kelas malignant cenderung memiliki run length yang panjang dan memiliki derajat keabuan yang rendah.

4.3. Klasifikasi SVM

Dalam penelitian ini akan membandingkan dua metode yaitu GLRLM dan GLRLM-PCA, kemudian proses klasifikasi dilakukan secara dua kali yaitu pada kelas normal-abnormal dan benign-malignant. Setiap kelas akan melalui tahap awal yaitu proses training dan testing pada setiap citra yang telah melalui ekstraksi fitur. Data untuk proses training dan testing akan dibagi menggunakan K-Fold, dengan nilai K=10. Pada tahap training, pola data akan dipelajari untuk pembentukan model dengan data pada setiap kelas klasifikasi. Sedangkan pada tahap testing, model yang telah di bentuk dari proses training akan di uji untuk mendapatkan evaluasi sistem klasifikasi. Pembagian data pada kelas klasifikasi normal-abnormal yaitu pada tahap training berjumlah kisaran 289 sampai 291 data untuk setiap K, dan pada tahap testing berjumlah kisaran 31 sampai 33 data. Sedangkan pembagian data pada kelas klasifikasi normal-abnormal yaitu pada tahap training berjumlah kisaran 103 hingga 104 data untuk setiap K, dan pada tahap testing berjumlah kisaran 11 sampai 12 data untuk dicari evaluasi sistem. Evaluasi ini dilakukan menggunakan confusion matriks, dimana hasil evaluasi

dapat dilihat pada akurasi, sensitifitas, dan spesifisitas. Kernel yang digunakan dalam sistem klasifikasi ada beberapa yaitu kernel RBF (Radial Basis Function), Linear, dan Polynomial.

4.3.1. Klasifikasi SVM dengan GLRLM

Pada klasifikasi ini menggunakan fitur GLRLM sebagai data masukan sistem klasifikasi, dimana fitur-fitur yang digunakan berjumlah 11 fitur untuk setiap sudutnya.

Sudut	Evaluasi	Kernel	K-Fold									
Suuut	Evaluasi	Kernei	1	2	3	4	5	6	7	8	9	10
_		RBF	56.25	54.55	56.25	59.38	59.38	68.75	59.38	54.55	57.58	58.06
	Akurasi	LIN	65.63	63.64	59.38	62.50	62.50	59.38	65.63	63.64	63.64	64.52
		POL	56.25	51.52	56.25	34.38	43.75	62.50	59.38	60.61	45.45	54.84
_		RBF	80.95	85.71	85.71	90.00	90.00	95.24	85.71	80.95	90.48	90.00
0 °	Sensitifitas	LIN	100.00	100.00	90.48	100.00	100.00	90.48	100.00	100.00	100.00	100.00
		POL	66.67	61.90	61.90	55.00	50.00	76.19	71.43	66.67	52.38	70.00
		RBF	9.09	0.00	0.00	8.33	8.33	18.18	9.09	8.33	0.00	0.00
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	36.36	33.33	45.45	0.00	33.33	36.36	36.36	50.00	33.33	27.27
		RBF	65.63	54.55	65.63	65.63	59.38	68.75	59.38	57.58	54.55	51.61
	Akurasi	LIN	65.63	63.64	59.38	62.50	62.50	59.38	65.63	63.64	63.64	64.52
		POL	56.25	48.48	65.63	31.25	53.13	59.38	62.50	57.58	54.55	48.39
_	Sensitifitas	RBF	90.48	85.71	95.24	95.00	90.00	95.24	90.48	85.71	85.71	75.00
45°		LIN	100.00	100.00	90.48	100.00	100.00	90.48	100.00	100.00	100.00	100.00
		POL	66.67	52.38	80.95	50.00	65.00	66.67	66.67	61.90	71.43	60.00
	Spesifisitas	RBF	18.18	0.00	9.09	16.67	8.33	18.18	0.00	8.33	0.00	9.09
		LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	36.36	41.67	36.36	0.00	33.33	45.45	54.55	50.00	25.00	27.27
		RBF	68.75	63.64	65.63	56.25	56.25	71.88	59.38	51.52	60.61	51.61
	Akurasi	LIN	65.63	63.64	65.63	62.50	62.50	65.63	65.63	63.64	63.64	64.52
		POL	59.38	45.45	50.00	40.63	40.63	46.88	53.13	60.61	48.48	48.39
_		RBF	95.24	95.24	90.48	90.00	90.00	100.00	90.48	76.19	90.48	75.00
90°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		POL	76.19	66.67	57.14	55.00	55.00	57.14	76.19	66.67	66.67	60.00
_		RBF	18.18	8.33	18.18	0.00	0.00	18.18	0.00	8.33	8.33	9.09
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	27.27	8.33	36.36	16.67	16.67	27.27	9.09	50.00	16.67	27.27
		RBF	62.50	57.58	68.75	59.38	53.13	62.50	65.63	60.61	54.55	58.06
	Akurasi	LIN	65.63	63.64	59.38	62.50	62.50	59.38	65.63	63.64	63.64	64.52
		POL	46.88	39.39	62.50	46.88	40.63	56.25	53.13	75.76	48.48	54.84
_		RBF	95.24	85.71	100.00	95.00	85.00	90.48	90.48	85.71	80.95	80.00
135°	Sensitifitas	LIN	100.00	100.00	90.48	100.00	100.00	90.48	100.00	100.00	100.00	100.00
		POL	61.90	57.14	66.67	65.00	55.00	71.43	61.90	80.95	66.67	65.00
_		RBF	0.00	8.33	9.09	0.00	0.00	9.09	18.18	16.67	8.33	18.18
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	-	POL	18.18	8.33	54.55	16.67	16.67	27.27	36.36	66.67	16.67	36.36

Gambar 4.4 Tabel Uji Coba GLRLM pada Kelas Normal-Abnormal

Berdasarkan tabel pada Gambar 4.4 telah didapatkan evaluasi setiap sudut. Evaluasi hasil akurasi terbaik yaitu sudut 135° sebesar 75.76% pada kernel *Polynomial*. Evaluasi hasil sensitifitas pada kernel *Linear* menunjukkan bahwa sebagian besar mendapatkan nilai sebesar 100%, sedangkan hasil spesifisitas menunjukkan bahwa sebagian besar mendapatkan nilai sebesar 0%. Sensitifitas merepresentasikan hasil klasifikasi ketika pasien dinyatakan sakit, dan spesifisitas merepresentasikan hasil klasifikasi ketika pasien dinyatakan tidak sakit. Hal ini disebabkan oleh data yang tidak dapat dipisahkan secara *Linear*, maka dari itu banyak data yang teridentifikasi Abnormal.

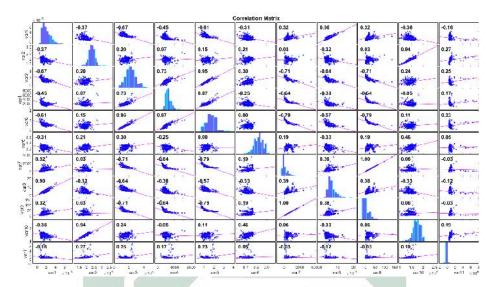
Sudut	Evaluasi	Kernel	K-Fold									
Sudui	Evaluasi	Kernei	1	2	3	4	5	6	7	8	9	10
0 °		RBF	63.64	58.33	50.00	50.00	54.55	45.45	36.36	41.67	45.45	50.00
	Akurasi	LIN	45.45	58.33	41.67	50.00	45.45	54.55	54.55	50.00	54.55	50.00
		POL	36.36	58.33	66.67	58.33	45.45	45.45	54.55	33.33	72.73	50.00
		RBF	100.00	85.71	71.43	71.43	66.67	83.33	66.67	57.14	66.67	66.67
	Sensitifitas	LIN	83.33	100.00	57.14	85.71	83.33	100.00	100.00	85.71	100.00	100.00
		POL	50.00	71.43	85.71	85.71	33.33	33.33	83.33	42.86	100.00	66.67
		RBF	20.00	20.00	20.00	20.00	40.00	0.00	0.00	20.00	20.00	33.33
	Spesifisitas	LIN	0.00	0.00	20.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	20.00	40.00	40.00	20.00	60.00	60.00	20.00	20.00	40.00	33.33
		RBF	72.73	58.33	58.33	58.33	45.45	54.55	27.27	41.67	54.55	41.67
	Akurasi	LIN	45.45	58.33	50.00	58.33	45.45	54.55	54.55	50.00	45.45	50.00
		POL	45.45	41.67	58.33	50.00	18.18	45.45	36.36	50.00	45.45	66.67
_		RBF	83.33	85.71	100.00	57.14	83.33	83.33	50.00	71.43	83.33	83.33
45°	Sensitifitas	LIN	83.33	100.00	85.71	100.00	83.33	100.00	100.00	85.71	83.33	100.00
		POL	33.33	57.14	85.71	57.14	16.67	50.00	50.00	57.14	66.67	83.33
	Spesifisitas	RBF	60.00	20.00	0.00	60.00	0.00	20.00	0.00	0.00	20.00	0.00
		LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	60.00	20.00	20.00	40.00	20.00	40.00	20.00	40.00	20.00	50.00
		RBF	45.45	41.67	58.33	75.00	45.45	54.55	36.36	58.33	45.45	50.00
	Akurasi	LIN	54.55	58.33	58.33	58.33	45.45	63.64	54.55	50.00	45.45	50.00
_		POL	36.36	33.33	50.00	50.00	54.55	72.73	9.09	50.00	27.27	50.00
_		RBF	50.00	57.14	100.00	85.71	83.33	83.33	66.67	71.43	66.67	83.33
90°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	83.33	100.00	100.00	85.71	66.67	100.00
_		POL	50.00	42.86	71.43	42.86	66.67	83.33	16.67	42.86	16.67	66.67
		RBF	40.00	20.00	0.00	60.00	0.00	20.00	0.00	40.00	20.00	16.67
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	20.00	0.00	0.00	20.00	0.00
		POL	20.00	20.00	20.00	60.00	40.00	60.00	0.00	60.00	40.00	33.33
		RBF	63.64	66.67	50.00	50.00	54.55	45.45	27.27	58.33	45.45	58.33
	Akurasi	LIN	54.55	58.33	58.33	58.33	54.55	54.55	54.55	50.00	54.55	50.00
_		POL	45.45	25.00	58.33	41.67	54.55	54.55	36.36	58.33	54.55	58.33
_		RBF	100.00	100.00	85.71	57.14	50.00	83.33	50.00	71.43	66.67	100.00
135°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	85.71	100.00	100.00
_		POL	33.33	14.29	71.43	57.14	33.33	33.33	50.00	57.14	50.00	83.33
_		RBF	20.00	20.00	0.00	40.00	60.00	0.00	0.00	40.00	20.00	16.67
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	60.00	40.00	40.00	20.00	80.00	80.00	20.00	60.00	60.00	33.33

Gambar 4.5 Tabel Uji Coba GLRLM pada Kelas Benign-Malignant

Berdasarkan tabel pada Gambar 4.5 telah didapatkan evaluasi setiap sudut. Evaluasi hasil akurasi terbaik yaitu sudut 455° sebesar 71.88% pada kernel *Polynomial*. Evaluasi hasil sensitifitas pada setiap kernel menunjukkan bahwa sebagian besar mendapatkan nilai sebesar 100%, sedangkan hasil spesifisitas menunjukkan bahwa sebagian besar mendapatkan nilai sebesar 0%. Hal ini disebabkan oleh data yang sangat susah untuk dibedakan tiap kelasnya, karena sedikitnya data yang digunakan pada kelas Benign-Malignant.

4.4. Klasifikasi SVM dengan Fitur GLRLM-PCA

Pada klasifikasi ini fitur-fitur GLRLM akan dilakukan tahap reduksi fitur terlebih dahulu, dimana fitur yang tersisa akan digunakan untuk data masukan sistem klasifikasi. Reduksi Fitur merupakan salah satu cara yang digunakan untuk mengurangi fitur atau mengambil fitur yang paling berpengaruh. Dalam penelitian ini reduksi fitur menggunakan metode PCA, dimana fitur-fitur yang digunakan yaitu hasil dari ekstraksi fitur GLRLM. Ekstraksi fitur dengan GLRLM menghasilkan sebanyak 11 fitur yaitu SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, dan LRHGE. Fitur-fitur tersebut akan menjadi masukan untuk metode PCA dengan data sebanyak 332 data dan 115 data pada setiap sudut orientasi. Perhitungan dilakukan bermulai dari menghitung matriks korelasi dengan hasil seperti pada Gambar 4.6.



Gambar 4.6 Matriks Korelasi

Selanjutnya nilai-nilai dari matriks korelasi ini akan dihitung *eigen vector* dan *eigen value*, dengan hasil seperti berikut.

$$eigen \, vector = \begin{bmatrix} -0,3342 & 0,2365 & 0,2802 & -0,1353 & \cdots & -0,0101 \\ 0,1336 & -0,4733 & 0,4225 & -0,3427 & \cdots & -0,0055 \\ 0,4207 & -0,0116 & -0,0891 & 0,1223 & \cdots & 0,0000 \\ 0,3511 & 0,2060 & 0,1759 & 0,0738 & \cdots & -0,0018 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0,1099 & -0,1478 & 0,5426 & 0,7392 & \cdots & 0,0025 \end{bmatrix}$$

$$eigen \, value = \begin{bmatrix} 5,0232726 & 2,6079681 & 1,1335589 & 0,8970601 & \cdots & 0,0000072 \end{bmatrix}^T$$

Kemudian untuk melihat banyak fitur yang berpengaruh perlu dilakukan perhitungan, dimana setiap elemen dari eigen value dibagi dengan keseluruhan

elemen dari eigen value dan dikali dengan 100%.

$$PC = \frac{\left[5,023273 \quad 2,607968 \quad 1,133559 \quad 0,897060 \quad \cdots \quad 0,000007\right]^{T}}{5,023273 + 2,607968 + 1,133559 + 0,897060 + \cdots + 0,0000007} \times 100\%$$

$$= \left[0,4566611 \quad 0,2370880 \quad 0,1030508 \quad 0,0815509 \quad \cdots \quad 0,0000006\right]^{T} \times 100\%$$

$$= \left[45,66611 \quad 23,70880 \quad 10,30508 \quad 8,15509 \quad \cdots \quad 0.00006\right]^{T}$$

Berdasarkan nilai-nilai PC yang telah dihitung, dilihat bahwa yang memiliki nilai lebih dari 10% yaitu ada 3 PC. Selanjutnya letak dari 3 PC tersebut akan digunakan untuk menghitung *factor loading*, dimana factor ini yang menjadi penentu untuk hasil dari reduksi fitur. Letak dari PC yang memiliki nilai lebih dari 10% yaitu pada 3 baris pertama. Kemudian perhitungan selanjutnya dilakukan dengan mengkalikan 3 baris pertama *eigen vector* dengan akar dari setiap elemen 3 baris pertama *eigen value*. Loading factor ini disebut juga component matrix, dapat dilihat pada Gambar 4.7

Component Matrix ^a										
		Component								
	1	2	3							
V1	-0.749	-0.382	0.298							
V2	0.299	0.764	0.450							
V3	0.943	0.019	-0.095							
V4	0.787	-0.333	0.187							
V5	0.953	-0.180	0.055							
V6	0.115	0.656	-0.509							
V7	-0.794	0.451	-0.057							
V8	-0.741	-0.340	0.362							
V9	-0.794	0.453	-0.061							
V10	0.273	0.850	0.254							
V11	0.246	0.239	0.578							

Extraction Method: Principal Component

a. 3 components extracted.

Gambar 4.7 Hasil Loading Factor

Berdasarkan nilai dari factor loading yang memiliki nilai lebih dari 0,5. Pada PC baris 1 yang terpenuhi yaitu fitur 3 (GLN), fitur 4 (RLN), dan fitur 5 (RP). Pada PC baris 2 tidak memiliki fitur yang terpenuhi dan pada PC baris 3 hanya ada pada fitur 11 (LRHGE). Berdasarkan banyak fitur yang terpenuhi maka fitur pada PC baris 1 yang akan digunakan menjadi masukan dari sistem klasifikasi karena memiliki paling banyak fitur yang memenuhi. Uji coba yang telah dilakukan dengan GLRLM-PCA dapat dilihat berdasarkan tabel pada Gambar 4.8.

			1 - 10		1.11		K-I	Zold				
Sudut	Evaluasi	Kernel		2	3	4	5	6	7	8	9	10
		RBF	65.63	63.64	65.63	62.50	59.38	65.63	65.63	63.64	63.64	64.52
- 0°	Akurasi	LIN	65.63	63.64	65.63	62.50	62.50	65.63	65.63	63.64	63.64	64.52
	AKUI asi	POL	65.63	63.64	65.63	62.50	59.38	65.63	65.63	60.61	60.61	64.52
		RBF	100.00	100.00	100.00	100.00	95.00	100.00	100.00	100.00	100.00	100.00
	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
ŭ	Sensitinus	POL	100.00	100.00	100.00	100.00	95.00	100.00	100.00	95.24	95.24	100.00
-		RBF	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		RBF	65.63	63.64	65.63	62.50	62.50	62.50	65.63	63.64	63.64	64.52
	Akurasi	LIN	65.63	63.64	65.63	62.50	62.50	65.63	65.63	63.64	63.64	64.52
		POL	65.63	63.64	65.63	62.50	59.38	71.88	65.63	63.64	63.64	64.52
		RBF	100.00	100.00	100.00	100.00	100.00	90.48	100.00	100.00	100.00	100.00
45°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		POL	100.00	100.00	100.00	100.00	95.00	100.00	100.00	95.24	100.00	100.00
•	Spesifisitas	RBF	0.00	0.00	0.00	0.00	0.00	9.09	0.00	0.00	0.00	0.00
		LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	0.00	0.00	0.00	18.18	0.00	8.33	0.00	0.00
		RBF	65.63	63.64	65.63	56.25	59.38	62.50	65.63	60.61	63.64	64.52
	Akurasi	LIN	65.63	63.64	65.63	62.50	62.50	65.63	65.63	63.64	63.64	64.52
		POL	65.63	63.64	65.63	59.38	59.38	68.75	65.63	60.61	60.61	64.52
		RBF	100.00	100.00	100.00	90.00	95.00	95.24	100.00	95.24	100.00	100.00
90°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		POL	100.00	100.00	100.00	95.00	95.00	95.24	100.00	95.24	95.24	100.00
		RBF	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	0.00	0.00	0.00	18.18	0.00	0.00	0.00	0.00
		RBF	65.63	63.64	65.63	62.50	59.38	56.25	65.63	63.64	63.64	64.52
	Akurasi	LIN	65.63	63.64	65.63	62.50	62.50	65.63	65.63	63.64	63.64	64.52
		POL	65.63	63.64	65.63	59.38	59.38	65.63	65.63	60.61	57.58	64.52
		RBF	100.00	100.00	100.00	100.00	95.00	80.95	100.00	100.00	100.00	100.00
135°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		POL	100.00	100.00	100.00	95.00	95.00	100.00	100.00	95.24	90.48	100.00
		RBF	0.00	0.00	0.00	0.00	0.00	9.09	0.00	0.00	0.00	0.00
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Gambar 4.8 Tabel Uji Coba GLRLM-PCA pada Kelas Normal-Abnormal



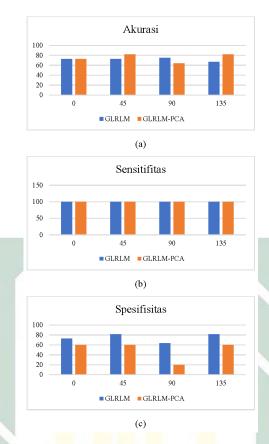
Gambar 4.9 Perbandingan Hasil Terbaik Normal-Abnormal

Uji coba yang telah dilakukan pada klasifikasi kelas normal-abnormal menunjukkan pada kernel RBF dan Linear banyak terdapat spesifisitas yang bernilai 0, dikarenakan seluruh hasil klasifikasi masuk pada kelas Normal. Selain itu uji coba tersebut membuktikan bahwa sistem klasifikasi yang dilakukan dengan fitur GLRLM mengalami peningkatan. Ketika dilakukan proses reduksi fitur dengan PCA yaitu pada kernel Polynomial pada sudut 45° dengan akurasi 71.875%, sensitifitas 100%, dan spesifisitas 18.182%. Hal ini membuktikan bahwa adanya peningkatan akurasi sebesar 3.125%, untuk melihat perbandingan hasil terbaik dapat dilihat pada Gambar 4.9.

Gambar 4.10 Tabel Uji Coba GLRLM-PCA pada Kelas Benign-Malignant

Sudut	Evaluasi	Kernel					K-I	old				
Sudut	Evaluasi	Kernei	1	2	3	4	5	6	7	8	9	10
		RBF	45.45	58.33	50.00	58.33	63.64	36.36	54.55	58.33	45.45	41.67
- 0°	Akurasi	LIN	54.55	58.33	58.33	58.33	54.55	54.55	54.55	58.33	54.55	50.00
		POL	36.36	58.33	50.00	58.33	72.73	54.55	45.45	41.67	54.55	33.33
		RBF	83.33	100.00	85.71	85.71	100.00	66.67	100.00	85.71	83.33	83.33
0°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
90° -		POL	66.67	100.00	85.71	71.43	100.00	50.00	83.33	71.43	83.33	50.00
		RBF	0.00	0.00	0.00	20.00	20.00	0.00	0.00	20.00	0.00	0.00
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	0.00	40.00	40.00	60.00	0.00	0.00	20.00	16.67
		RBF	45.45	58.33	58.33	58.33	54.55	36.36	54.55	50.00	54.55	41.67
	Akurasi	LIN	54.55	58.33	58.33	58.33	54.55	54.55	54.55	58.33	54.55	50.00
_		POL	36.36	50.00	50.00	58.33	81.82	36.36	54.55	50.00	63.64	33.33
	Sensitifitas	RBF	83.33	100.00	100.00	100.00	100.00	66.67	100.00	85.71	100.00	83.33
45°		LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
_		POL	66.67	85.71	85.71	85.71	100.00	33.33	100.00	85.71	100.00	50.00
	Spesifisitas	RBF	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	0.00	20.00	60.00	40.00	0.00	0.00	20.00	16.67
		RBF	45.45	41.67	50.00	58.33	54.55	45.45	36.36	58.33	45.45	50.00
	Akurasi	LIN	54.55	58.33	58.33	58.33	54.55	54.55	54.55	58.33	54.55	50.00
_		POL	45.45	50.00	50.00	58.33	63.64	54.55	45.45	58.33	45.45	41.67
		RBF	83.33	57.14	85.71	100.00	100.00	83.33	66.67	100.00	83.33	100.00
90°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
_		POL	83.33	85.71	85.71	100.00	100.00	100.00	83.33	100.00	83.33	83.33
		RBF	0.00	20.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	0.00	0.00	20.00	0.00	0.00	0.00	0.00	0.00
		RBF	54.55	58.33	58.33	66.67	63.64	36.36	45.45	58.33	45.45	41.67
	Akurasi	LIN	54.55	58.33	58.33	58.33	54.55	54.55	54.55	58.33	54.55	50.00
_		POL	36.36	50.00	58.33	58.33	81.82	45.45	45.45	50.00	54.55	33.33
		RBF	83.33	85.71	85.71	85.71	100.00	66.67	83.33	85.71	83.33	83.33
135°	Sensitifitas	LIN	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
_		POL	66.67	85.71	85.71	85.71	100.00	50.00	83.33	85.71	100.00	50.00
		RBF	20.00	20.00	20.00	40.00	20.00	0.00	0.00	20.00	0.00	0.00
	Spesifisitas	LIN	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		POL	0.00	0.00	20.00	20.00	60.00	40.00	0.00	0.00	0.00	16.67

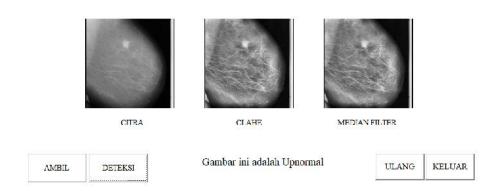
Uji coba pada klasifikasi kelas Benign-Malignant menunjukkan pada kernel RBF dan Linear banyak terdapat spesifisitas yang bernilai 0, dikarenakan seluruh hasil klasifikasi masuk pada kelas Normal. Selain itu uji coba tersebut membuktikan bahwa sistem klasifikasi yang dilakukan dengan fitur GLRLM mengalami peningkatan. Ketika dilakukan proses reduksi fitur dengan PCA yaitu pada kernel Polynomial pada sudut 0° dengan akurasi 81,82%, sensitifitas 70%, dan spesifisitas 100%.



Gambar 4.11 Perbandingan Hasil Terbaik Benign-Malignant

Berdasarkan hasil yang telah didapatkan dan uji coba yang telah dilakukan hasil menunjukkan bahwa kernel polynomial merupakan kernel yang memiliki kinerja yang baik dalam melakukan klasifikasi. Hal ini telah dibuktikan oleh (Novitasari et al., 2019a) dengan akurasi mencapai 93.98% pada data *mammogram* dan pada ekstraksi fitur GLRLM juga. Selain itu kernel polynomial juga telah dibuktikan oleh (Thohir et al., 2020) pada ekstraksi fitur GLCM dengan sudut 45° dan hasil akurasi mencapai 90%. Model yang mendapatkan hasil terbaik dijadikan model untuk sistem identifikasi pada Guide untuk mempermudah identifikasi penyakit kanker. Berikut hasil guide untuk mengidentifikasi penyakit kanker payudara yaitu pada Gambar 4.12.

IDENTIFIKASI PENYAKIT KANKER PAYUDARA MENGGUNAKAN METODE GLRLM DAN SVM DENGAN REDUKSI FITUR.



Gambar 4.12 Guide untuk Identifikasi Penyakit Kanker Payudara

Pada tampilan guide yang ditunjukkan pada Gambar 4.12 bahwa memiliki beberapa tombol yaitu "Ambil", "Deteksi", "Ulang", dan "Keluar". Tombol "Ambil" digunakan untuk mengambil citra yang akan dideteksi, tampilan setelah mengambil citra yaitu citra asli, citra hasil CLAHE, dan citra hasil Median Filter. Tombol "Deteksi" digunakan untuk mendeteksi citra yang telah dipilih, apakah termasuk kelas normal atau abnormal. Tombol "Ulang" digunakan untuk mengulang deteksi dari citra dan memilih kembali citra. Tombol "Keluar" digunakan untuk menutup program dari deteksi penyakit kanker payudara.

4.5. Integrasi Keislaman

Proses identifikasi yang menggunakan metode SVM pada penyakit Kanker Payudara yaitu memerlukan beberapa langkah untuk sampai pada sistem klasifikasi. Pertama, pengumpulan data citra Mammogram untuk setiap kelas. Kedua, preprocessing yaitu digunakan untuk memperbaiki kualitas dari setiap data citra dengan menggunakan metode CLAHE dan Median Filter. Ketiga, ekstraksi fitur dan reduksi yaitu untuk pengambilan ciri penting pada setiap citra dengan

menggunakan metode GLRLM dan PCA. Terakhir, sistem klasifikasi yaitu menggunakan metode SVM untuk mencari model terbaik klasfikasi. Model terbaik ini digunakan untuk menghasilkan akurasi terbaik juga, sehingga dari penjelasan tersebut dapat disimpulkan bahwa untuk mendapatkan hasil terbaik perlu dilakukan suatu usaha. Hal ini sesuai dengan ayat Al Qur'an yaitu surat At-Taubah ayat 105 yang berbunyi seperti berikut.

Artinya: "Bekerjalah kamu, maka Allah akan melihat pekerjaanmu, begitu juga Rasul-Nya dan orang-orang mukmin, dan kamu akan dikembalikan kepada (Allah) Yang Mengetahui yang gaib dan yang nyata, lalu diberitakan-Nya kepada kamu apa yang telah kamu kerjakan."

Berdasarkan ayat tersebut Allah telah menyuruh setiap mukmin untuk bekerja, karena setiap kesulitan Allah akan berikan kemudahan setelahnya. Hal ini sesuai dengan ayat Al Qur'an yaitu surat Asy Syarh ayat 6 yang berbunyi seperti berikut.

Artinya: "Sesungguhnya sesudah kesulitan itu ada kemudahan."

Ayat tersebut menjelaskan agar tetap berusaha jika dalam kesusahan, sebagaimana pada penelitian ini diharapkan dapat bermafaat bagi pasien Kanker Payudara untuk sistem identifikasi. Apabila pasien telah terdiagnosa menderita

Kanker Payudara, pasien harus tetap berusaha untuk sembuh dari penyakit yang diderita karena penyakit yang diberi oleh Allah maka akan diberi pula penawarnya. Allah tidak akan langsung memberikan kesembuhan jika seseorang tersebut tidak berusaha, dan Allah akan berikan kesembuhan melalui perantara dokter. Selain itu pasien tidak diperbolehkan putus asa ketika diberi penyakit, seperti perintah Allah dalam surat Az-Zumar ayat 53 sebagai berikut.

Artinya: "Wahai anak-anakku! Pergilah kamu, carilah (berita) tentang Yusuf dan saudaranya dan jangan kamu berputus asa dari rahmat Allah. Sesungguhnya yang berputus asa dari rahmat Allah, hanyalah orang-orang yang kafir."

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil penelitian dari Identifikasi Penyakit Kanker Payudara Menggunakan Metode GLRLM dan SVM dengan Reduksi Fitur dapat disimpulkan bahwa:

- 1. Preprocessing dengan menggunakan metode CLAHE dan Median Filter dalam pengambilan ciri pada ekstraksi fitur pada citra mammogram mampu memperbaiki citra agar lebih jelas dan terang. Citra semula memiliki kontras yang lemah dan memiliki banyak noise, setelah dilakukan preprocessing citra menjadi memiliki kontras yang lebih jelas dan noise pada citra telah dihilangkan.
- 2. Ekstraksi fitur pada GLRLM menghasilkan 11 fitur diantaranya SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, dan LRHGE. Fitur-fitur ini dilakukan proses reduksi fitur, dimana fitur yang tersisa yaitu GLN, RLN, dan RP. Fitur ini merepresentasikan dari citra mammogram, pada citra normal cenderung memiliki tekstur lebih halus, pada citra benign cenderung memiliki tekstur lebih kasar jika dibandingkan citra normal, dan pada citra malignant cenderung memiliki tekstur yang kasar.
- 3. Identifikasi penyakit kanker payudara diperoleh peningkatan hasil akurasi ketika memakai reduksi fitur. Pada sistem klasifikasi Normal-Abnormal

diperoleh hasil terbaik yaitu pada sudut 90° yaitu akurasi sebesar 71.875% mengalami peningkatan sebanyak 8.239%. Pada sistem klasifikasi Jinak-Ganas diperoleh hasil terbaik yaitu pada sudut 0° yaitu akurasi sebesar 81.818% mengalami peningkatan sebanyak 9.091%.

5.2. Saran

Pada penelitian ini yang membahas terkait Identifikasi Penyakit Kanker Payudara Menggunakan Metode GLRLM dan SVM dengan Reduksi Fitur masih memiliki banyak kekurangan. Sehingga terdapat beberapa saran yang diajukan oleh peneliti untuk penelitian mendatang diantaranya yaitu:

1. Pada penelitian ini memiliki data awal 322 data yang terdiri dari 207 data normal, 64 data benign, dan 51 data malignant. Hal ini dapat mempengaruhi proses pembelajaran karena data pada kelas benign dan malignant memiliki jumlah yang jauh lebih sedikit jika dibandingkan dengan data normal. Sehingga akan lebih baik jika data yang digunakan memiliki perbandingan yang seimbang dan lebih banyak data.

DAFTAR PUSTAKA

- Abbas, I. (2016). Penerapan Metode Moving Average (MA) Berbasis Algoritma Support Vector Machine (SVM) untuk Membandingkan Pola Kurva dengan Trend Kurva pada Trading Forex Online. *ILKOM Jurnal Ilmiah*, 8(1):37–43.
- Al-Dhabyani, W., Gomaa, M., Khaled, H., and Fahmy, A. (2020). Dataset of breast ultrasound images. *Data in Brief*, 28:104863.
- Ardayani, T., Fauziah, L., and Sitorus, N. (2020). Deteksi Dini Kanker Payudara Melalui Pemeriksaan Sadanis (Pemeriksaan Payudara Dengan Tenaga Medis) Di Desa Babakan Kecamatan Ciparay Kabupaten Bandung. *JPKMI (Jurnal Pengabdian Kepada Masyarakat Indonesia*), 1(1):14–17.
- Arifin, O. and Sasongko, T. B. (2018). Analisa perbandingan tingkat performansi metode support vector machine dan naïve bayes classifier. *Seminar Nasional Teknologi Informasi dan Multimedia* 2018, 6(1):67–72.
- Arowolo, M. O., Adebiyi, M., Adebiyi, A., and Okesola, O. (2020). PCA Model for RNA-Seq Malaria Vector Data Classification Using KNN and Decision
 Tree Algorithm. 2020 International Conference in Mathematics, Computer Engineering and Computer Science, ICMCECS 2020.
- Astiningrum, M., Syulistyo, A. R., and Zakariya, M. A. (2020). Identifikasi Kualitas Biji Jagung Manis Layak Jual dari Warna dan Tekstur Menggunakan HSV Dan Gray Level Run Length Matrix (GLRLM). *JIP (Jurnal Informatika Polinema)*, 7(1):37–44.

- Astuti, W. and Adiwijaya, A. (2019). Principal Component Analysis sebagai Ekstraksi Fitur Data Microarray untuk Deteksi Kanker Berbasis Linear Discriminant Analysis. *Jurnal Media Informatika Budidarma*, 3(2):72.
- Cahyanti, D., Rahmayani, A., and Husniar, S. A. (2020). Analisis Performa Metode KNN pada Dataset Pasien Pengidap Kanker Payudara. *Indonesian Journal of Data and Science*, 1(2):39–43.
- Chazar, C. and Erawan, B. (2020). Machine Learning Diagnosis Kanker Payudara Menggunakan Algoritma Support Vector Machine. *INFORMASI* (Jurnal Informatika dan Sistem Informasi), 12(1):67–80.
- Devin, J. L., Hill, M. M., Mourtzakis, M., Quadrilatero, J., Jenkins, D. G., and Skinner, T. L. (2019). Acute high intensity interval exercise reduces colon cancer cell growth. *The Journal of Physiology*, 597(8):2177–2184.
- Dewi, M. (2017). Sebaran Kanker di Indonesia, Riset Kesehatan Dasar 2007. Indonesian Journal of Cancer, 11(1):1–8.
- Durgamahanthi, V., Anita Christaline, J., and Shirly Edward, A. (2021). GLCM and GLRLM Based Texture Analysis: Application to Brain Cancer Diagnosis Using Histopathology Images. pages 691–706.
- Erkan, U., Gökrem, L., and Enginoğlu, S. (2018). Different applied median filter in salt and pepper noise. *Computers & Electrical Engineering*, 70:789–798.
- Erwin and Ningsih, D. R. (2020). Improving Retinal Image Quality Using the Contrast Stretching, Histogram Equalization, and CLAHE Methods with Median Filters. *International Journal of Image, Graphics and Signal Processing*, 12(2):30–41.

- Fadillah, N. and Gunawan, C. R. (2019). Mendeteksi Keakuratan Metode Noise Salt and Pepper dengan Median Filter. *Jurnal Informatika*, 6(1):91–95.
- Fibrianda, M. F. and Bhawiyuga, A. (2018). Analisis Perbandingan Akurasi Deteksi Serangan Pada Jaringan Komputer Dengan Metode Naïve Bayes Dan Support Vector Machine (SVM). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, II(9):3112–3123.
- Foeady, A. Z. (2019). Sistem Klasifikasi Kanker Kulit Berdasarkan Data Citra Dermoscopic dengan Menggunakan Metode Deep Extreme Learning Machine. PhD thesis, UIN Sunan Ampel Surabaya.
- Foeady, A. Z., Rini Novitasari, D. C., Asyhar, A. H., and Firmansjah, M. (2018).

 Automated Diagnosis System of Diabetic Retinopathy Using GLCM Method and SVM Classifier. *Proceeding of the Electrical Engineering Computer Science and Informatics*, 5(5).
- Hao, X., Meng, X., and Wu, Z. (2019). Clinical Utility of PET/CT in Breast CancerManagement and Targeted Therapy. *Nuclear Medicine in Oncology*, pages 11–24.
- Hermanto, H., Mustopa, A., and Kuntoro, A. Y. (2020). ALGORITMA KLASIFIKASI NAIVE BAYES DAN SUPPORT VECTOR MACHINE DALAM LAYANAN KOMPLAIN MAHASISWA. *JITK* (*Jurnal Ilmu Pengetahuan dan Teknologi Komputer*), 5(2):211–220.
- Irawan, E. (2018). Faktor-Faktor Pelaksanaan Sadari/Breast Self Examination (BSE) Kanker Payudara (Literature Review). *Jurnal Keperawatan BSI*, 6(1).
- Jumeilah, F. S. (2017). Penerapan Support Vector Machine (SVM) untuk

- Pengkategorian Penelitian. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, 1(1):19–25.
- Jumhuriyah, L. (2021). Klasifikasi Penyakit Tuberkulosis Berdasarkan Citra X-Ray Menggunakan Metode Elman Recurrent Neural Network. PhD thesis, UIN Sunan Ampel Surabaya.
- Kosasih, R., Informatika, F. T., Gunadarma, U., Margonda, J., No, R., and Cina, P. (2021). Pendeteksian Kendaraan Menggunakan Metode Median Filter. *Jurnal Ilmiah Komputasi*, 20(1):53–58.
- Liang, M., Huang, G., Liu, Z., Wang, Q., Yu, Z., Liu, Z., Lin, H., Li, M., Zhou, X., and Zheng, Y. (2019). Elevated levels of hsa_circ_006100 in gastric cancer promote cell growth and metastasis via miR-195/GPRC5A signalling. *Cell Proliferation*, 52(5):1–11.
- Liantoni, F. and Santoso, A. (2020). Perbaikan Kontras Citra Mammogram pada Klasifikasi Kanker Payudara Berdasarkan Fitur Gray-Level Co-Occurrence Matrix. SINTECH (Science and Information Technology) Journal, 3(1):46–51.
- Ma'arif, F. and Arifin, T. (2017). Optimasi Fitur Menggunakan Backward Elimination dan Algoritma SVM untuk Klasifikasi Kanker Payudara. *Jurnal Informatika*, 4(1):46–53.
- Maricar, M. A. and Dian Pramana (2019). Perbandingan Akurasi Naïve Bayes dan K-Nearest Neighbor pada Klasifikasi untuk Meramalkan Status Pekerjaan Alumni ITB STIKOM Bali. *Jurnal Sistem dan Informatika (JSI)*, 14(1):16–22.
- Matondang, Z. A. (2018). Penerapan Metode Contrast Limited Adaptive Histogram

- Equalization (CLAHE) Pada Citra Digital Untuk Memperbaiki Gambar X-Ray. Publikasi Ilmiah Teknologi Informasi Neumann (PITIN), 3(2):107–112.
- Morrow, M., Waters, J., and Morris, E. (2011). MRI for breast cancer screening, diagnosis, and treatment. *The Lancet*, 378(9805):1804–1811.
- Muhidin, A. and Burhan, A. (2018). Klasifikasi Tingkat Produksi Tanaman Padi Kabupaten Karawang Menggunakan Metode Naive Bayes dan K-Fold Cross Validation. *Jurnal Teknologi Pelita Bangsa- SIGMA*, 8(2):2407–3903.
- Mu'jizah, H. (2021). Klasifikasi Kanker Usus Besar Berdasarkan Citra Histopatologi Menggunakan Metode Gray Level Run Length Matrix-Extreme Learning Machine. PhD thesis, UIN Sunan Ampel Surabaya.
- Noor, A., Zhao, Y., Khan, R., Wu, L., and Abdalla, F. Y. (2020). Median filters combined with denoising convolutional neural network for Gaussian and impulse noises. *Multimedia Tools and Applications*, 79(25-26):18553–18568.
- Novitasari, D. C. R., Lubab, A., Sawiji, A., and Asyhar, A. H. (2019a). Application of Feature Extraction for Breast Cancer using One Order Statistic, GLCM, GLRLM, and GLDM. *Advances in Science, Technology and Engineering Systems Journal*, 4(4):115–120.
- Novitasari, D. C. R., Wahyuni, D., Munir, M., Hidayati, I., Amin, F. M., and Oktafianto, K. (2019b). Automatic Detection of Breast Cancer in Mammographic Image Using the Histogram Oriented Gradient (HOG) Descriptor and Deep Rule Based (DRB) Classifier Method. In 2019 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA), pages 185–190. IEEE.

- Pujianto, R. and Rahmawati, A. A. (2019). Analisis Ekstraksi Fitur Principle Component Analysis pada Klasifikasi Microarray Data Menggunakan Classification And Regression Trees. 6(1):2368–2379.
- Rahayu, G., Data, L., Program, M., Sistem, S., Islam, U., Sultan, N., Kasim, S., and
 Analysis, P. C. (2017). Principal Component Analysis untuk Dimensi Reduksi
 Data Clustering Sebagai Pemetaan Persentase Sertifikasi Guru di Indonesia.
 pages 18–19.
- Rahayuwati, L., Rizal, Iqbal, A., Lukman, M., and Juniarti, N. (2020). Pendidikan Kesehatan tentang Pencegahan Penyakit Kanker dan Menjaga Kualitas Kesehatan Pendahuluan Indonesia mencatat kanker sebagai penyebab kematian nomor tujuh di Indonesia dan menjadi penyebab kematian nomor dua di dunia. Angka penderita kanker selal. *Media Karya Kesehatan*, 3(1):59–69.
- Rahim, N. and Foozy, C. F. M. (2020). Source Camera Identification for Online Social Network Images Using Texture Feature. pages 283–296.
- Rahman, M. A. (2019). Sistem Deteksi Dini Kanker Payudara Pada Citra Mammogram Menggunakan Metode Convolutional Neural Network (CNN). PhD thesis, UIN Sunan Ampel Surabaya.
- Rahutomo, F., Saputra, P. Y., and Fidyawan, M. A. (2018). Implementasi Twitter
 Sentiment Analysis Untuk Review Film MenggunaRahutomo, F., Saputra, P.
 Y. and Fidyawan, M. A. (2018) 'Implementasi Twitter Sentiment Analysis
 Untuk Review Film Menggunakan Algoritma Support Vector Machine', Jurnal
 Informatika Polinema. *Jurnal Informatika Polinema*, 4(2):93.
- Rasmi, R. P. (2020). Peningkatan Hasil Diagnosis Kanker Payudara dari Hasil

- Citra Mammogram Menggunakan Metode Ekstraksi Ciri dan Klasifikasi. PhD thesis, Universitas Islam Indonesia.
- Rhomadhona, H. and Permadi, J. (2019). Klasifikasi Berita Kriminal Menggunakan Naïve Bayes Classifier (NBC) dengan Pengujian K-Fold Cross Validation. *Jurnal Sains dan Informatika*, 5(2):108–117.
- Rini Novitasari, D. C., Foeady, A. Z., Thohir, M., Arifin, A. Z., Niam, K., and Asyhar, A. H. (2020). Automatic Approach for Cervical Cancer Detection Based on Deep Belief Network (DBN) Using Colposcopy Data. 2020 International Conference on Artificial Intelligence in Information and Communication, ICAIIC 2020, pages 415–420.
- Rosyani, P. (2017). Pengenalan Wajah Menggunakan Metode Principal Component

 Analysis (PCA) Dan Canberra Distance. *Jurnal Informatika Universitas*Pamulang, 2(2).
- Saputra, R. W. (2021). Kombinasi Metode Smote Dan Pca Untuk Meningkatkan Performa Klasifikasi Fibrosis Hati Data Pasien Hepatitis C Berbasis Extreme Learning Machine. PhD thesis, UIN Sunan Ampel Surabaya.
- Sari, J. P., Erlansari, A., and Purwandari, E. P. (2021). Identifikasi Citra Digital Kura-Kura Sumatera Dengan Perbandingan Ekstraksi Fitur GLCM Dan GLRLM Berbasis Web. *Pseudocode*, 8(1):66–75.
- Solikhah, S. (2019). Skrining Kanker Payudara pada Wanita di Indonesia Breast Cancer Screening among Indonesian Women. *Jurnal MKMI*, 15(1):1–9.
- Supartini, I. A. M., Sukarsa, I. K. G., and Srinadi, I. G. A. M. (2017). Analisis

- Diskriminan Pada Klasifikasi Desa Di Kabupaten Tabanan Menggunakan Metode K-Fold Cross Validation. *E-Jurnal Matematika*, 6(2):106.
- Thohir, M., Foeady, A. Z., Novitasari, D. C. R., Arifin, A. Z., Phiadelvira, B. Y., and Asyhar, A. H. (2020). Classification of Colposcopy Data Using GLCM-SVM on Cervical Cancer. In 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), pages 373–378. IEEE.
- Utami, S. S. and Mustikasari, M. (2017). Aspek Psikososial Pada Penderita Kanker Payudara: Studi Pendahuluan. *Jurnal Keperawatan Indonesia*, 20(2):65–74.
- Vaishali, D., Priya, P. V., Govind, N., and Prabha, K. V. R. (2021). Higher Order Statistical Analysis in Multiresolution Domain - Application to Breast Cancer Histopathology. pages 495–508.
- Wahyuni, F. S., Windrasari, W., and Khambri, D. (2018). Evaluasi Terapi Adjuvant Hormonal dan Hubungannya Terhadap Outcome Klinis Pasien Kanker Payudara Stadium Dini di Kota Padang. *Jurnal Sains Farmasi & Klinis*, 5(3):176–184.
- Wang, S., Su, X., Xu, M., Xiao, X., Li, X., Li, H., Keating, A., and Zhao, R. C. (2019). Exosomes secreted by mesenchymal stromal/stem cell-derived adipocytes promote breast cancer cell growth via activation of Hippo signaling pathway. *Stem Cell Research and Therapy*, 10(1):1–12.
- Wibawa, M. S. and Novianti, K. D. P. (2017). Reduksi Fitur untuk Optimalisasi Klasifikasi Tumor Payudara Berdasarkan Data Citra FNA. *Konferensi Nasional Sistem & Informatika*, pages 73–78.
- Wibisono, G. and Hermawan, A. (2019). Faktor-Faktor Penentu Gejala Penyakit

Kanker Payudara Dengan Pendekatan Jaringan Saraf Tiruan. *Jurnal Jasiek*, 1(1):1–6.

Yulianti, I., Santoso, H., and Sutinigsih, D. (2016). Faktor-Faktor Risiko Kanker Payudara (Studi Kasus Pada Rumah Sakit Ken Saras Semarang). *Jurnal Kesehatan Masyarakat Universitas Diponegoro*, 4(4):401–409.

