

## Klasifikasi *Pneumonia* Menggunakan Metode *Support Vector Machine*

Risha Ambar Wati\*<sup>1</sup>, Hafiz Irsyad<sup>2</sup>, M Ezar Al Rivan<sup>3</sup>

<sup>1,2</sup>STMIK GI MDP; Jl. Rajawali No 14, +62(711)376400/373630

<sup>3</sup>Program Studi Teknik Informatika, STMIK GI MDP, Palembang

e-mail: \*<sup>1</sup>rishaambarwati@mhs.mdp.ac.id, <sup>2</sup>hafizirsyad@mdp.ac.id, <sup>3</sup>meedzhar@mdp.ac.id

### Abstrak

*Pneumonia* adalah jenis penyakit paru-paru yang disebabkan oleh bakteri, virus, jamur, atau parasit. Salah satu cara untuk mengetahui *pneumonia* adalah dengan rontgen. Hasil rontgen akan dianalisis untuk mengetahui ada tidaknya *pneumonia*. Penelitian ini bertujuan untuk mengklasifikasikan hasil x-ray apakah terdapat *pneumonia* atau tidak pada hasil x-ray tersebut. Metode klasifikasi yang digunakan dalam penelitian ini adalah *Support Vector Machine (SVM)* dan *Gray Level Co-Occurrence (GLCM)* untuk metode ekstraksi. Ada beberapa tahapan sebelum klasifikasi yaitu *cropping*, *resizing*, *contrast stretching*, dan *thresholding* kemudian diekstraksi menggunakan *GLCM* dan diklasifikasikan menggunakan *SVM*. Hasil penelitian menunjukkan bahwa akurasi terbaik sebesar 62.66%.

**Kata kunci**— *Pneumonia*, *Thresholding*, *GLCM*, *SVM*

### Abstract

*Pneumonia* is a type of lung disease caused by bacteria, viruses, fungi, or parasites. One way to find out *pneumonia* is by x-ray. X-rays will be analyzed to determine whether there is *pneumonia* or not. This study aims to classify the x-ray results whether there is *pneumonia* or not on the x-ray results. The classification method used in this study were *Support Vector Machine (SVM)* and *Gray Level Co-Occurrence (GLCM)* for the extraction method. There are several stages before classification, namely *cropping*, *resizing*, *contrast stretching*, and *thresholding* then extracted using *GLCM* and classified using *SVM*. The results showed that the best accuracy of 62.66%.

**Keywords**— *Pneumonia*, *Thresholding*, *GLCM*, *SVM*

## 1. PENDAHULUAN

Paru-paru merupakan salah satu organ pernapasan yang sangat mempengaruhi keluar masuknya oksigen dan karbondioksida, fungsi utama paru-paru adalah mengirimkan atau mentransfer oksigen dari udara ke darah dan melepaskan karbondioksida dari darah ke udara. dalam proses pernapasan, udara memasuki mulut atau hidung dan melewati *trakea* (tenggorokan), *bronkus*, serta *bronkiolus* hingga sampai ke *alveoli*. *Aveoli* menyerap oksigen dari udara dan menyebarkannya ke dalam darah untuk diedarkan kesekitar tubuh kemudian karbondioksida dikeluarkan setelah meninggalkan darah dan menuju ke *alveoli*.

Terdapat banyak jenis penyakit yang berasal dari paru-paru apabila paru-paru tidak memiliki kesehatan yang baik, salah satunya adalah radang paru-paru yang dapat menyebabkan gangguan pernapasan, salah satu penyakit peradangan paru-paru adalah *Pneumonia*, *Pneumonia* merupakan salah satu peradangan pada organ paru-paru yang disebabkan oleh bakteri, virus,

jamur, ataupun *parasite*. *Pneumonia* merupakan penyakit paru dimana terdapat infeksi yang terjadi pada kantung-kantung udara dalam paru-paru yang disebabkan oleh berbagai mikroorganisme termasuk bakteri jamur, mikrobakteri dan virus *pneumonia* diklasifikasikan sebagai *pneumonia* yang dikontrol komunitas, *pneumonia* yang didapat dirumah sakit (*nosocomial*), *pneumonia* pada *hospes immunocompromised*, dan *pneumonia* aspirasi [1].

Salah satu cara mengetahui penyakit *pneumonia* adalah dengan *rontgen* atau *x-ray*, dengan pemeriksaan ini didapatkan gambaran kondisi paru-paru yang mengalami *pneumonia*, diagnosa tersebut dianalisis oleh dokter ahli agar diketahui apakah terdapat *pneumonia* atau tidak. dalam beberapa tahun terakhir terdapat beberapa penelitian untuk mendeteksi *pneumonia*. Penelitian untuk mengidentifikasi *pneumonia* berdasarkan citra *rontgen* yang di *thresholding* dan menggunakan metode *Convolutional Neural Network* untuk proses klasifikasi, penelitian ini memiliki tingkat akurasi 83,3% [2].

Klasifikasi *Support Vector Machine* (SVM) merupakan salah satu metode *machine learning*. Keuntungan lain menggunakan SVM adalah metode ini dapat dianalisis secara teoritis menggunakan konsep teori pembelajaran komputasi, SVM telah digunakan pada penelitian [3] yang menggunakan SVM sebagai klasifikasi dan *Histogram of Oriented Gradients* (HOG) sebagai ekstraksi fitur yang membahas mengenai Perbandingan Jarak Potret Dan Resolusi Kamera Pada Tingkat Akurasi Pengenalan Angka Kwh Meter. Angka meteran listrik dianalisis melalui citra menggunakan pengenalan pola dalam pemrosesan citra. Hasil menggunakan 100 train-set dan test-set untuk setiap kombinasi kategori menunjukkan bahwa resolusi terbaik adalah 10 MP dan 14 MP dan jarak pengambilan gambar 30 cm dan resolusi 10 cm dengan akurasi 73,33% untuk setiap gambar dan 86,67% untuk setiap nomor dan hasil *confusion matrix* seluruh nomor adalah 75,48% [3].

Dalam pengolahan citra digital, tekstur adalah salah satu fitur yang bisa diekstraksi untuk identifikasi citra. Ekstraksi fitur dilakukan untuk mengambil informasi suatu data untuk dipakai pada proses selanjutnya. Salah satu metode untuk mengekstraksi fitur suatu citra yaitu dengan GLCM. Metode GLCM ini menghasilkan *output* berupa fitur. Seperti fitur *Contras*, *Co-occurrence*, *Energy*, dan *Homogeneity*. Untuk mengekstraksi fitur dengan metode GLCM, sebuah citra RGB dan dijadikan *grayscale*. *Gray Level Co-occurrence Matrix* (GLCM) adalah suatu metode yang digunakan untuk analisis tekstur atau ekstraksi ciri. GLCM merupakan suatu matriks yang menggambarkan frekuensi munculnya pasangan dua piksel dengan intensitas tertentu dalam jarak dan arah tertentu dalam citra. Koordinat pasangan piksel memiliki jarak dan orientasi sudut. Jarak direpresentasikan dalam piksel dan sudut direpresentasikan dalam derajat. Orientasi sudut terbentuk berdasarkan empat arah sudut yaitu, 0°, 45°, 90° dan 135°, dan jarak antar piksel sebesar 1 piksel, Inputan GLCM ini berupa citra yang akan mengeluarkan *output* berupa *feature vector*. Penelitian menggunakan GLCM telah dilakukan dalam Klasifikasi Kacang-kacangan berdasarkan tekstur menggunakan jaringan syaraf tiruan, klasifikasi kacang-kacangan dilakukan terhadap beberapa jenis kacang yaitu kacang merah, kacang hijau dan kacang tanah. Fitur tekstur yang digunakan adalah GLCM. Ekperimen yang digunakan dengan 3 jumlah neuron yang berbeda pada *hidden layer*, *training function* yang digunakan ada 17 jenis. Setiap skenario eksperimen diulang sebanyak 5 kali, berdasarkan skenario eksperimen yang dilakukan hasil yang terbaik yaitu 99,8% [4].

Pada penelitian sebelumnya, penelitian yang menggunakan citra *rontgen* paru-paru normal, kanker paru, dan penyakit paru lainnya, tujuan dari penelitian ini untuk mencari karakteristik tekstur dari citra paru-paru, citra paru-paru akan dilakukan *cropping* menjadi ukuran 640x640 piksel dengan tujuan untuk memotong citra pada daerah paru dan merubah dimensi citra menjadi lebih kecil dari ukuran asli, selanjutnya citra yang telah di *cropping* dengan ukuran 640x640 piksel dilakukan ekstraksi menggunakan GLCM menggunakan empat sudut yaitu sudut 0°, 45°, 90°, dan 135° kemudian dicari lima fitur GLCM yang digunakan yaitu *angular second moment* (ASM), kontras, *inverse different moment* (IDM), *entropi*, dan *korelasi*. Berdasarkan lima fitur GLCM tersebut terdapat tiga fitur GLCM yang memiliki perbedaan nilai yaitu IDM, kontras dan *entropi* [5]. Penelitian lainnya dilakukan ekstraksi dan klasifikasi menggunakan metode GLCM dan SVM pada citra mammogram untuk identifikasi kanker

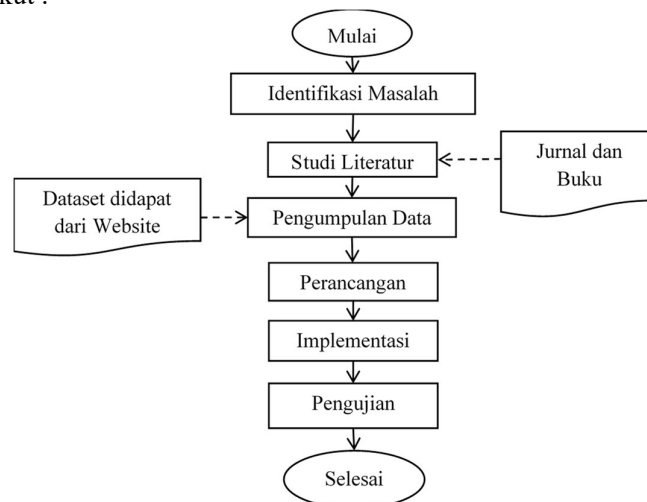
payudara, dari hasil penelitian dapat disimpulkan metode analisis tekstur menggunakan GLCM dan proses klasifikasi menggunakan SVM pada citra mammogram menghasilkan probabilitas sebesar 60% dengan pengujian yang digunakan 10 citra gabungan (5 jinak dan 5 ganas) dan data pembelajaran yang digunakan 10 data jinak dan 10 data ganas [6].

*Thresholding* adalah salah satu metode yang banyak digunakan untuk mengetahui segmentasi dan ekstraksi gambar yang berguna untuk membedakan latar depan dari *background* suatu citra. Pada penelitian membandingkan *otsu* dan *iterative adaptive thresholding* dalam *binerisasi* gigi *kaninus* foto *panoramic*, metode *binerisasi* yang digunakan adalah *Otsu* untuk memperbaiki citra menggunakan filter median, mencari nilai batas ambang atau nilai *threshold* untuk mengubah citra ke bentuk biner, secara keseluruhan uji coba mengubah citra ke bentuk biner menghasilkan gambar yang baik menggunakan metode *iterative* dan *adaptive thresholding* [7].

*Thresholding* merupakan salah satu segmentasi citra yang cukup baik sebelum citra di klasifikasi, dan ekstraksi GLCM merupakan salah satu fitur yang bisa diekstraksi untuk identifikasi citra. Ekstraksi fitur dilakukan untuk mengambil informasi suatu data untuk dipakai pada proses selanjutnya, kemudian metode SVM dapat mengelola hasil data *training* dalam jumlah yang besar dan mampu mengklasifikasi data berdasarkan kelas yang tepat, penelitian ini bertujuan untuk mengetahui hasil citra *rontgen* paru-paru normal dan *pneumonia* dengan metode klasifikasi SVM dan ekstraksi GLCM.

## 2. METODE PENELITIAN

Berikut adalah beberapa tahapan yang dilakukan dalam penelitian dapat dilihat pada Gambar 1 berikut :



**Gambar 1 Kerangka Kerja Penelitian**

### 2.1 Identifikasi Masalah

Pada tahapan ini yaitu melakukan identifikasi masalah yang mencakup pengenalan tekstur hasil rontgen, dalam hal ini adalah klasifikasi citra hasil rontgen paru-paru normal dan paru-paru *pneumonia*.

## 2.2 Studi Literatur

Melakukan tahapan pencarian jurnal yang terkait dengan klasifikasi penyakit pada paru-paru menggunakan beberapa metode yang digunakan, jurnal yang terkait mengenai metode segmentasi *Threshold*, ekstraksi GLCM dan metode yang terkait metode klasifikasi SVM.

### 2.1.1 Contrast Stretching

*Contrast Stretching* merupakan Proses yang dilakukan untuk mendapatkan RGB baru dengan kualitas kontras yang lebih baik dari sebelumnya dengan begitu peningkatan kontras citra membantu memudahkan proses selanjutnya yaitu *Thresholding*.

### 2.1.2 Thresholding

*Thresholding* adalah salah satu metode yang akan menghasilkan citra yang memiliki dua nilai tingkat keabuan yaitu hitam dan putih yang biasa disebut dengan citra biner. Proses *thresholding* akan menghasilkan citra yang memiliki dua nilai tingkat keabuan yaitu hitam dan putih atau yang disebut dengan citra biner.

### 2.1.3 Gray Level Co-Occurance (GLCM)

GLCM merupakan salah satu proses ekstraksi GLCM pada setiap citra data training dan data testing. Tahapan ekstraksi GLCM citra biner yang berukuran 300x300 piksel menghasilkan nilai ciri statistik dengan 4 arah ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  dan  $135^\circ$ ). Kemudian dibentuk matriks kookurasi dengan menghitung frekuensi kemunculan pasangan nilai keabuan piksel referensi dan piksel tetangga pada jarak dan arah yang ditentukan. Kemudian menjumlahkan semua elemen untuk menghitung probabilitas setiap elemen dengan cara membagi setiap elemen GLCM dengan total semua elemen. Setelah itu dihitung ciri statistik GLCM dari citra tersebut, ciri-ciri dari citra dikategorikan menjadi 4 ciri yaitu energi, kontras, korelasi dan homogenitas.

### 2.1.4 Support vector machine (SVM)

Pada tahap ini merupakan tahapan dalam menerapkan metode klasifikasi *SVM* untuk mendapatkan *classifier* yang berbentuk *feature vector* untuk mendapatkan prediksi *testing*. Setelah melalui tahap-tahap sebelumnya, hasil ekstraksi diproses untuk menghasilkan model klasifikasi SVM. Vektor yang sudah dipetakan akan dihitung jaraknya. Jarak terjauh akan digunakan sebagai pemisah kelas dari vektor. Kemudian diberikan sebuah *hyperplane* untuk memisahkan dua kelas. Hal terpenting yang dibutuhkan untuk membuat model klasifikasi SVM adalah dengan mengkonversi dokumen dalam bentuk vektor. Variasi nilai digunakan untuk menemukan nilai untuk mencapai tingkat akurasi terbaik. Proses ini diperlukan untuk mengkonversi dokumen uji menjadi vektor. Proses selanjutnya adalah dengan mengambil data kata dokumen uji yang biasa disebut vektor untuk kemudian dimasukkan ke dalam model SVM yang telah dibuat sebelumnya. Persamaan *hyperplane* merupakan persamaan klasifikasi ditunjukkan oleh persamaan (1) dengan parameter klasifikasinya adalah nilai  $w$  dan  $b$  sebagai nilai bobot dan bias yang ditunjukkan oleh persamaan (2) dan persamaan (3) secara berturut-turut.

$$f_{svm}(x) = w \cdot x + b \quad (1)$$

$$w = \sum_{i=1}^N a_i y_i x_i \quad (2)$$

$$b = -\frac{1}{2}(w \cdot x^+ + w \cdot x^-) \quad (3)$$

Dimana  $N$  ialah banyaknya data,  $a_i$  nilai koefisien bobot setiap titik pasang data dan label ( $x_i$ ). SVM juga baik mengelola *dataset* dengan memanfaatkan *kernel trik* sebagai upaya menemukan *hyperplane* salah satunya yaitu *Linear Kernel* yang ditunjukkan oleh persamaan pada (4).

$$K(x_i, x_j) = x_i \cdot x_j \quad (4)$$

Dari hasil *kernel* yang didapatkan SVM membuat sebuah persamaan klasifikasi yang disesuaikan dengan *kernel* yang digunakan seperti persamaan (5) dengan nilai klasifikasi kelas didasarkan pada persamaan (6).

$$f_{svm}(x) = \sum \alpha_i y_i K(x_i, x) + b \quad i \in N \quad (5)$$

$$class = \begin{cases} 1, & f_{svm}(x) \geq 0 \\ -1, & \text{selainnya} \end{cases} \quad (6)$$

Penentuan nilai  $\alpha_i$  dilakukan dengan penyelesaian permasalahan optimasi LD (*Dualitas Lagrange Multiplier*), yang ditunjukkan oleh persamaan (7) dan dengan konstrainnya pada persamaan (8)

$$Ld = \max_{\alpha_i \in R} \sum_{i \in N} \alpha_i - \frac{1}{2} \sum_{i \in N} \sum_{j \in N} \alpha_i \alpha_j y_i y_j \quad (7)$$

$$0 \leq \alpha_i \leq C \text{ dan } \sum_{i \in N} \alpha_i y_i = 0 \quad (8)$$

Dimana  $C$  ialah nilai konstanta dengan,  $Ld(x_i)$  fungsi *kernel* yang digunakan.

### 2.3 Pengumpulan Data

Pengumpulan data menggunakan *dataset* yang didapat dari website berjumlah 5.853 citra rontgen paru-paru, yang diambil pada penelitian ini sebanyak 3140 citra rontgen lebih banyak dari penelitian sebelumnya yang menggunakan 74 citra rontgen. Gambar-gambar dibagi menjadi 2 jenis citra rontgen diantaranya paru-paru normal, paru-paru *pneumonia*. Citra rontgen paru-paru dilakukan penyesuaian ukuran yaitu menjadi 300x300 piksel merujuk pada penelitian [2], yang bertujuan membuang bagian yang tidak penting yang bias mempengaruhi proses. Setiap citra rontgen memiliki ukuran berbeda dikarenakan data diambil dari rentang usia 1-5 tahun. Berikut adalah jumlah data yang terdapat pada setiap citra rontgen dapat dilihat pada Tabel 1 dan Tabel 2:

**Tabel 1 Jumlah Dataset testing 600**

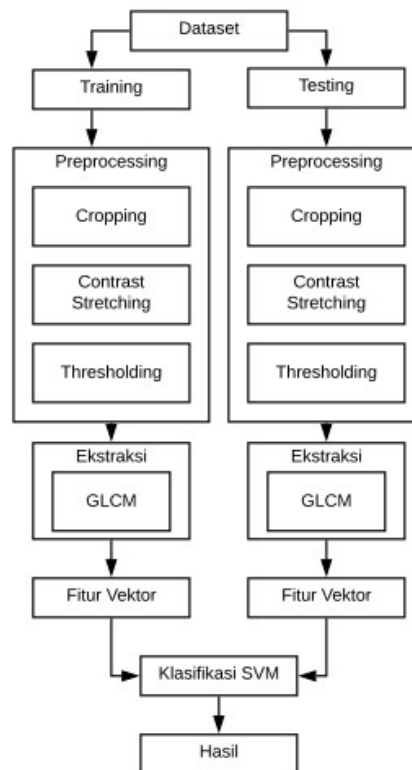
Jenis – jenis Citra	Jumlah data yang dipilih	Jumlah data Training	Jumlah data Testing
Normal	1500 gambar	1200 gambar	300 gambar
<i>Pneumonia</i>	1500 gambar	1200 gambar	300 gambar
<b>Total</b>	3000 gambar	2400 gambar	600 gambar

**Tabel 2 Jumlah Dataset testing 750**

Jenis – jenis Citra	Jumlah data yang dipilih	Jumlah data Training	Jumlah data Testing
Normal	1575 gambar	1200 gambar	375 gambar
<i>Pneumonia</i>	1575 gambar	1200 gambar	375 gambar
<b>Total</b>	3150 gambar	2400 gambar	750 gambar

#### 2.4 Perancangan

Tahap ini melakukan perancangan skema dalam penelitian, mula – mula melakukan preprocessing terhadap citra paru-paru yaitu *Cropping*, *Resize*, *Contrast Stretching* dan dilakukan segmentasi menggunakan *threshold* selanjutnya citra akan diekstraksi GLCM dan dilanjutkan dengan proses klasifikasi menggunakan SVM. Berikut adalah gambar dari skema dalam penelitian dapat dilihat pada Gambar 2 berikut.

**Gambar 2 Perancangan Skema**

## 2.5 Implementasi

Setelah melakukan proses perancangan, pada tahap ini dilakukan implementasi terhadap rancangan yang telah dibuat. Tahapan pertama dilakukan *cropping* pada citra dengan ukuran 300x300 piksel, selanjutnya citra dilakukan peningkatan kontras dengan *Contrast Stretching*, Kemudian citra yang telah dilakukan peningkatan kontras akan di-*thresholding* dengan hasil berupa citra biner, tahapan selanjutnya dilakukan ekstraksi GLCM untuk mendapatkan empat ciri dari citra itu sendiri yaitu energi, kontras, korelasi dan homogenitas. Setelah didapat nilai ciri dari citra, maka proses selanjutnya yaitu klasifikasi menggunakan *SVM*.

## 2.6 Pengujian

Setelah proses implementasi dilanjutkan dengan pengujian citra *training* dan citra *testing* secara acak sebanyak 3 kali untuk 1 jenis citra rontgen dengan *cross validation*. Hasil dari pengujian yang didapatkan berupa nilai akurasi yang berasal dari perhitungan rata-rata sebanyak 3 kali pada pengujian tersebut.

Selanjutnya, hasil klasifikasi menggunakan SVM dihitung untuk mendapatkan tingkat keberhasilan dari metode yang digunakan, dengan menggunakan metode *Confussion Matrix* yang dimana akan menghitung nilai *Precision*, *Recall*, dan *Accuracy*. Cara penghitungan *Confussion Matrix* tersebut bisa dilihat pada Tabel 3 dan persamaan (1), (2), (3).

**Tabel 3 Confussion Matrix**

		TRUE VALUES	
		TRUE	FALSE
PREDICTIO N	TRUE	TP Correct Result	FP Unexpected Result
	FALSE	FN Missing Result	TN Correct Absence of Result

$$Precision = \frac{TP}{TP+FP} \times 100\% \quad (1)$$

$$Recall = \frac{TP}{TP+FN} \times 100\% \quad (2)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (3)$$

Keterangan :

TP = Jumlah data positif citra jenis paru-paru yang terklasifikasi dengan benar oleh sistem.

TN = Jumlah data negatif citra jenis paru-paru yang terklasifikasi dengan benar oleh sistem.

FN = Jumlah data negatif citra jenis paru-paru namun terklasifikasi salah oleh sistem.

FP = Jumlah data positif citra jenis paru-paru namun terklasifikasi salah oleh sistem.

Cara penghitungan *Confussion Matrix* tersebut bisa dilihat pada Tabel 3 dan persamaan (1), (2), (2).

## 3. HASIL DAN PEMBAHASAN

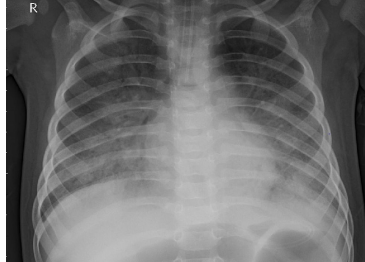
Pada tahap ini membahas tentang implementasi dimana terdapat penjelasan mengenai implementasi pada tahap *preprocessing* yang mana terdapat tiga proses diantaranya *cropping*, *Contrast Stretching* dan *Thresholding*, implementasi ekstraksi GLCM, implementasi klasifikasi menggunakan SVM, serta membahas hasil pengujian yang menjelaskan tingkat akurasi dari hasil pengujian klasifikasi *pneumonia*.

### 3.1 Implementasi *Preprocessing*

Tahap ini mengimplementasikan beberapa *preprocessing* data sebelum dilakukannya ekstraksi GLCM, yaitu sebagai berikut.

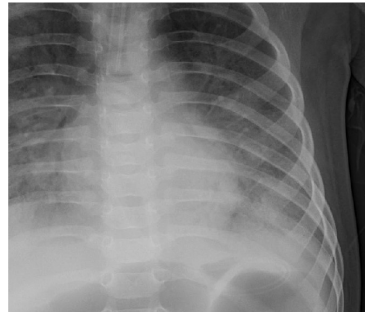
#### 3.4.1 *Cropping*

Tujuan proses *cropping* dilakukan untuk membuang bagian-bagian yang dianggap tidak penting dan mendapatkan citra yang lebih baik dan *detail*, proses ini dilakukan menggunakan fitur matlab dengan ukuran 300x300 pixel sehingga mendapatkan hasil citra yang sesuai dengan yang diperlukan saat ini. Berikut citra asli pada Gambar 3 dan citra setelah di-*cropping* pada Gambar 4 berikut.



Sumber: [kaggle.com/paultimothymooney/chest-xray-pneumonia](https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia)

**Gambar 3 Citra Asli**



**Gambar 4 Citra Setelah di-*Cropping***

#### 3.4.2 *Resize*

Proses ini dilakukan dengan tujuan menyamaratakan semua ukuran citra paru-paru normal dan citra paru-paru yang mengindap *pneumonia*. Berikut citra setelah dilakukan *thresholding* pada Gambar 5 berikut.



**Gambar 5 Citra Setelah di-*Resize***



### 3.4.1 Contrast Stretching

Proses ini dilakukan untuk mendapatkan RGB baru dengan kualitas kontras yang lebih baik dari sebelumnya dengan begitu peningkatan kontras citra membantu memudahkan proses selanjutnya yaitu *Thresholding*. Berikut citra setelah dilakukan *contrast stretching* pada Gambar 6 berikut.



**Gambar 6 Citra Setelah di-Contrast Stretching**

### 3.4.2 Thresholding

Pada proses *Thresholding* akan menghasilkan citra yang memiliki dua nilai tingkat keabuan yaitu hitam dan putih yang biasa disebut dengan citra biner. Berikut citra setelah dilakukan *thresholding* pada Gambar 7 berikut.



**Gambar 7 Citra Setelah di-Thresholding**

## 3.2 Implementasi Ekstraksi GLCM

Pada tahap ini dilakukan proses ekstraksi GLCM pada setiap citra data training dan data testing. Tahapan ekstraksi GLCM citra biner yang berukuran 300x300 piksel menghasilkan nilai ciri statistik dengan 4 arah ( $0^\circ$ ,  $45^\circ$ ,  $90^\circ$  dan  $135^\circ$ ). Kemudian dibentuk matriks kookurasi dengan menghitung frekuensi kemunculan pasangan nilai keabuan piksel referensi dan piksel tetangga pada jarak dan arah yang ditentukan. Kemudian menjumlahkan semua elemen untuk menghitung probabilitas setiap elemen dengan cara membagi setiap elemen GLCM dengan total semua elemen. Setelah itu dihitung ciri statistik GLCM dari citra tersebut, ciri-ciri dari citra dikategorikan menjadi 4 ciri yaitu energi, kontras, korelasi dan homogenitas.

### 3.3 Implementasi Klasifikasi SVM

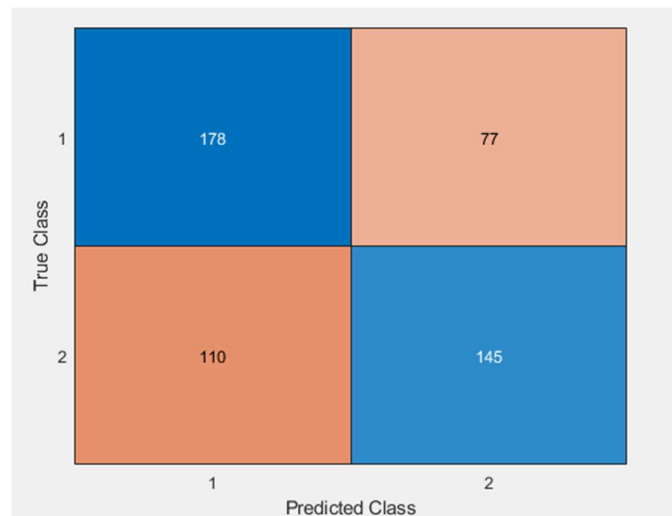
Pada tahap ini merupakan tahapan dalam menerapkan metode klasifikasi *SVM* untuk mendapatkan *classifier* yang berbentuk *feature vector* untuk mendapatkan prediksi *testing*. Setelah melalui tahap-tahap sebelumnya, hasil ekstraksi diproses untuk menghasilkan model klasifikasi *SVM*. Vektor yang sudah dipetakan akan dihitung jaraknya. Jarak terjauh akan digunakan sebagai pemisah kelas dari vektor. Kemudian diberikan sebuah *hyperplane* untuk memisahkan dua kelas. Hal terpenting yang dibutuhkan untuk membuat model klasifikasi *SVM* adalah dengan mengkonversi dokumen dalam bentuk vektor. Variasi nilai digunakan untuk menemukan nilai untuk mencapai tingkat akurasi terbaik. Proses ini diperlukan untuk mengkonversi dokumen uji menjadi vektor. Proses selanjutnya adalah dengan mengambil data kata dokumen uji yang biasa disebut vektor untuk kemudian dimasukkan ke dalam model *SVM* yang telah dibuat sebelumnya.

### 3.4 Hasil Klasifikasi SVM

Setelah dilakukan implementasi klasifikasi *SVM*, selanjutnya didapat hasil berupa *Confussion Matrix* dari masing-masing jumlah dataset yang berbeda.

#### 3.4.1 *Confussion Matrix* dataset *testing* 600

Tahap ini adalah tahapan *Confussion matrix* untuk mendapatkan sebuah nilai atau hasil prediksi dengan menggunakan metode klasifikasi *SVM* dan fitur *GLCM* dengan jumlah data *testing* sebanyak 600, Sehingga didapat nilai dari hasil *Confussion Matrix* pada Gambar 8.



**Gambar 8** *Confussion Matrix* dengan jumlah data *testing* 600

Gambar 8 pada setiap *cell* yang berwarna biru menunjukkan jumlah nilai dari setiap jenis paru-paru yang dapat dikenali dengan menggunakan fitur *GLCM* sedangkan *cell* yang berwarna *orange* menunjukkan jumlah nilai dari setiap jenis paru-paru yang tidak dapat dikenali.

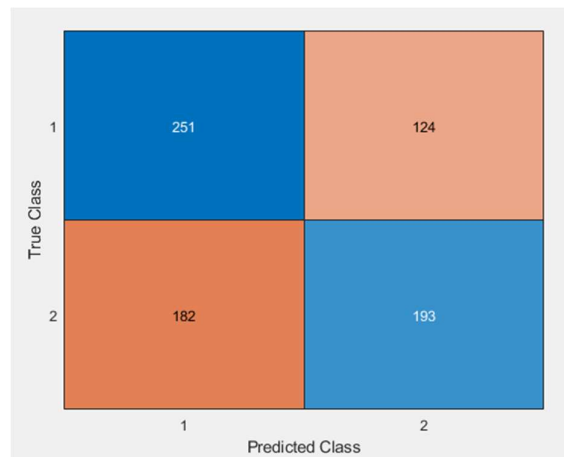
**Tabel 4** Hasil *Accuracy Perkelas* Dari Jumlah Data *Testing* 600

No	Kelas	TP	TN	FP	FN	Accuracy	Precision	Recall
1.	Normal	204	172	128	96	62,60%	61,45%	68%
2.	Pneumonia	172	204	96	128	62,66%	64,17%	57,33%

Kemampuan sistem dalam mengenali atau membedakan data paru-paru normal dan paru-paru yang mengidap *pneumonia*, Paru Normal ditunjukkan oleh nilai *precision* yaitu sebesar 61,45 %, kemampuan sistem dalam mengenali data positif atau yang disebut dengan *recall* sebesar 68% dan hasil *accuracy* yang di dapat pada kelas Paru Normal adalah 62,6% nilai tersebut didapatkan dari nilai TP,TN FP,FN dan Kemampuan sistem dalam mengenali atau membedakan pada paru *Pneumonia* ditunjukkan oleh nilai *precision* yaitu sebesar 64,17%, kemampuan sistem dalam mengenali data positif atau yang disebut dengan *recall* sebesar 57,33% dan hasil *accuracy* yang di dapat pada kelas Paru Normal adalah 62,66% nilai tersebut didapatkan dari nilai TP,TN FP,FN seperti pada Tabel 4.

### 3.4.2 Confussion Matrix dataset testing 750

Tahap ini adalah tahapan *Confussion matrix* untuk mendapatkan sebuah nilai atau hasil prediksi dengan menggunakan metode klasifikasi SVM dan fitur GLCM dengan jumlah data *testing* sebanyak 750. Sehingga didapat nilai dari hasil *Confussion Matrix* pada Gambar 9 berikut.



**Gambar 9 Confussion Matrix dataset testing 750**

Gambar 9 pada setiap *cell* yang berwarna biru menunjukkan jumlah nilai dari setiap jenis paru-paru yang dapat dikenali dengan menggunakan fitur GLCM sedangkan *cell* yang berwarna *orange* menunjukkan jumlah nilai dari setiap jenis paru-paru yang tidak dapat dikenali.

**Tabel 5 Hasil Accuracy Perkelas Dari Jumlah Data Testing 750**

No	Kelas	TP	TN	FP	FN	Accuracy	Precision	Recall
1.	Normal	251	193	182	124	59,2%	57,9%	67%
2.	Pneumonia	193	251	124	182	59,2%	60,8%	51,4%

Kemampuan sistem dalam mengenali atau membedakan data paru-paru normal dan paru-paru yang mengidap *pneumonia*, Paru Normal ditunjukkan oleh nilai *precision* yaitu sebesar 57,9%, kemampuan sistem dalam mengenali data positif atau yang disebut dengan *recall* sebesar 67% dan hasil *accuracy* yang di dapat pada kelas Paru Normal adalah 59,2% nilai tersebut didapatkan dari nilai TP,TN FP,FN dan kemampuan sistem dalam mengenali atau membedakan data pada paru *Pneumonia* ditunjukkan oleh nilai *precision* yaitu sebesar 60,8%, kemampuan sistem dalam mengenali data positif atau yang disebut dengan *recall* sebesar 51,4% dan hasil *accuracy* yang di dapat pada kelas Paru Normal adalah 65,92% nilai tersebut didapatkan dari nilai TP,TN FP,FN seperti pada Tabel 5.

#### 4. KESIMPULAN

Hasil pengujian dapat disimpulkan dari penelitian ini bahwa Performa SVM dengan fitur GLCM untuk jenis citra paru-paru normal dengan jumlah dataset *testing* 600 mendapatkan nilai *precision* sebesar 61,45%, *recall* 68% dan *accuracy* 62,66%, Performa SVM dengan fitur GLCM untuk jenis citra paru-paru yang mengidap *pneumonia* dengan jumlah dataset *testing* 600 mendapatkan nilai *precision* sebesar 64,17%, *recall* 57,33% dan *accuracy* 62,66%, Performa SVM dengan fitur GLCM untuk jenis citra paru-paru normal dengan jumlah dataset *testing* 750 mendapatkan nilai *precision* sebesar 57,9%, *recall* 67% dan *accuracy* 59,2%. Performa SVM dengan fitur GLCM untuk jenis citra paru-paru yang mengidap *pneumonia* dengan jumlah dataset *testing* 750 mendapatkan nilai *precision* sebesar 60,8%, *recall* 51,4% dan *accuracy* 59,2%. SVM mampu mengklasifikasikan citra paru-paru dengan jumlah dataset *testing* sebanyak 600 menggunakan fitur GLCM dengan hasil *accuracy* keseluruhan 62,66%, SVM mampu mengklasifikasikan citra paru-paru dengan jumlah dataset *testing* sebanyak 750 menggunakan fitur GLCM dengan hasil *accuracy* keseluruhan 59,2%.

#### 5. SARAN

Berdasarkan penelitian yang telah dilakukan adapun saran yang dapat dilakukan oleh peneliti selanjutnya adalah Sebaiknya menggunakan fitur ekstraksi lain guna meningkatkan tingkat akurasi, penelitian ini menggunakan dataset publik, belum diketahui hasil akurasi yang diperoleh jika dataset yang digunakan diambil secara langsung oleh peneliti, Menggunakan dataset dengan jarak potret yang sama dan Menambahkan jumlah data latih dan data uji.

#### DAFTAR PUSTAKA

- [1] F. A. P. Scholastica, Asuhan Keperawatan Pada Pasien Dengan Gangguan Sistem Pernapasan. Yogyakarta : Pustaka Baru Press, pp. 14–99, 2018.
- [2] W. Christanto, “Klasifikasi Pneumonia Menggunakan Convolutional Neural Network”. *Repositori Institusi Universitas Sumatera Utara*, 2018.
- [3] D. Amputri, S. Nadra, Gasim, & M. E. Al Rivan, "Perbandingan Jarak Potret Dan Resolusi Kamera Pada tingkat Akurasi Pengenalan Angka Kwh Meter Menggunakan Svm", *Jurnal Ilmiah Informatika Global*, 2017.
- [4] M. E. Al Rivan, N. Rachmat, & M. R. Ayustin, “Klasifikasi Jenis Kacang-Kacangan Berdasarkan Tekstur Menggunakan Jaringan Syaraf Tiruan”, vol. 6, no. 1, 2020. <https://doi.org/doi.org/10.35143/jkt.v6i1.3546>
- [5] T. D. Wulan, “Ekstraksi Fitur Citra Paru-Paru Menggunakan Gray Level Co-ocurance matriks”, *Technology Science and Engineering Journal*, vol. 1, no. 3, pp. 225–229, 2017.
- [6] B. D. Junita, “Ekstraksi Fitur Klasifikasi Menggunakan Metode GLCM Dan SVM Citra Mammogram Untuk Identifikasi Kanker Payudara”, *Jurnal Teknologi Rekayasa*, vol. 22, no. 1, pp. 18–26, 2017.
- [7] N. Nafiyah, R. Wardhani, “Perbandingan Otsu Dan Iterative Adaptive Thresholding Dalam Binerisasi Gigi Kanius Foto Panoramik ”, *Jurnal Ilmiah Teknologi Informasi Asia*, vol. 11, no. 1, 2017.