

Klasifikasi *Pneumonia* Menggunakan Metode *K-Nearest Neighbor* Dengan Ekstraksi GLCM

Chandra Wijaya^{*1}, Hafiz Irsyad², Wijang Widhiarso³

^{1,2}STMIK GI MDP; Jl. Rajawali No 14, +62(711)376400/373630

³Program Studi Teknik Informatika, STMIK GI MDP, Palembang

e-mail: ^{*1}chandrawijaia@mhs.mdp.ac.id, ²hafizirsyad@mdp.ac.id, ³wijang@mdp.ac.id

Abstrak

Pneumonia merupakan penyakit parenkim inflamasi yang disebabkan oleh berbagai mikroorganisme, antara lain bakteri, mikro bakteri, jamur, dan virus. Penelitian ini menggunakan X-ray untuk mengetahui ada tidaknya pneumonia. Tujuan dari penelitian ini adalah untuk mengklasifikasikan hasil rontgen ada atau tidaknya pneumonia secara cepat dan tepat melalui suatu program sehingga menghasilkan akurasi yang baik. Metode klasifikasi yang digunakan dalam penelitian ini adalah K-Nearest Neighbor (KNN) dan Gray Level Co-Occurrence (GLCM) untuk metode ekstraksi. Ada beberapa tahapan sebelum diklasifikasikan yaitu cropping, resizing, contrast stretching, dan thresholding. Hasil penelitian menunjukkan bahwa akurasi terbaik per kelas adalah 66,20% untuk K = 5.

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

Abstract

Pneumonia is an inflammatory parenchymal disease caused by various microorganisms, including bacteria, micro bacteria, fungi, and viruses. This study used an X-ray to find out whether or not there was pneumonia. The objective of this study was to classify the X-ray results whether or not there was pneumonia in a fast and precise way through a program to produce good accuracy. The classification method used in this study were K-Nearest Neighbor (KNN) and Gray Level Co-Occurrence (GLCM) for the extraction method. There are several stages before being classified, namely cropping, resizing, contrast stretching, and thresholding. The results showed that the best accuracy per class was 66.20% for K = 5.

Keywords— *Pneumonia, Thresholding, GLCM, KNN*

1. PENDAHULUAN

Paru-paru adalah salah satu organ yang berada pada sistem pernapasan yang fungsinya sebagai tempat bertukarnya oksigen dengan karbondioksida didalam darah. Dalam proses pernapasan udara memasuki mulut atau hidung dan melewati trakea (tenggorokan), bronkus, serta bronkiolus hingga sampai ke alveoli. Pertukaran oksigen dan karbondioksida terjadi di alveoli. Alveoli menyerap oksigen dari udara dan menyebarkannya ke dalam darah untuk kemudian diedarkan ke sekitar tubuh. Peradangan pada paru-paru dapat mengakibatkan gangguan pada pernapasan. *Pneumonia* merupakan salah satu penyakit paru-paru ketika seseorang mengalami infeksi yang terjadi pada kantong-kantong udara dalam paru-paru. Infeksi yang ditimbulkan *pneumonia* bisa terjadi pada salah satu sisi paru-paru maupun keduanya. *Pneumonia* adalah suatu radang parenkim paru yang dapat disebabkan oleh berbagai mikroorganisme, termasuk bakteri, mikrobakteri, jamur, dan virus. *Pneumonia* diklasifikasikan sebagai *pneumonia*

yang dikontrol komunitas (CAP), *pneumonia* yang didapat di rumah sakit (*nosocomial*), *pneumonia* pada *hospes immunocompromised*, dan *pneumonia* aspirasi [1].

Seiring perkembangan teknologi yang semakin pesat, *pneumonia* dapat diidentifikasi melalui analisis citra rontgen atau x-ray. Dalam beberapa tahun terakhir sudah dilakukan penelitian untuk mendeteksi *pneumonia*. Penelitian yang melakukan perbaikan kualitas citra menggunakan median filter dan erosi pada citra yang dapat mempermudah proses ekstraksi fitur pada citra yang kemudian diklasifikasi menggunakan metode *K-Nearest Neighbor* (KNN) yang mendapatkan nilai akurasi sebesar 86,67% [2].

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 (4) sudut yaitu sudut 0o, 45o, 90o, dan 135o kemudian dicari lima (5) fitur GLCM yang digunakan yaitu angular second moment (ASM), kontras, inverse different moment (IDM), entropi, dan korelasi. Berdasarkan lima (5) fitur GLCM tersebut terdapat tiga (3) fitur GLCM yang memiliki perbedaan nilai yaitu IDM, kontras dan entropi [3].

Selain itu, klasifikasi *K-Nearest Neighbor* (KNN) merupakan metode yang dapat digunakan untuk proses klasifikasi atau mengelompokkan data citra uji yang sudah ditentukan kelasnya pada data latih yang sudah disimpan pada datastore [4].

Penelitian untuk mengidentifikasi *pneumonia* berdasarkan dataset *public* yang di dapat dari *website* Mendeley sebanyak 74 citra, yaitu 22 citra *Bronchopneumonia*, 25 citra *Pneumonia Lobaris* dan 27 citra paru-paru normal, data tersebut digunakan sebagai data pelatihan dan data pengujian. Pada tahap *Preprocessing* citra rontgen paru-paru dilakukan *Scaling* dengan ukuran 300x300 piksel, lalu citra di-*Grayscale*, kemudian citra *grayscale* dilakukan contrast stretching dengan tujuan untuk mendapatkan citra baru dengan kontras yang lebih bagus dari citra asalnya. Tahapan setelah *preprocessing* yaitu *thresholding* yang bertujuan untuk menghasilkan citra menjadi warna hitam dan putih dengan nilai *threshold* adalah 170, dan metode yang digunakan untuk proses klasifikasi yaitu Convolutional Neural Network, penelitian ini memiliki tingkat akurasi mencapai 83,3% [5].

Penelitian untuk mengklasifikasikan jenis tumor otak dengan menggunakan penerapan metode *threshold* untuk memisahkan citra dengan objek lainnya dan menggunakan GLCM untuk mengekstraksi ciri yang terkandung dalam citra tersebut dengan menggunakan metode *K-Nearest Neighbor* yang berfungsi untuk menghitung jarak yang paling dekat berdasarkan data pembelajaran. Citra MRI diolah dengan melakukan peningkatan kualitas citra, lalu dilakukan segmentasi citra untuk memilih dan memisahkan suatu objek dari keseluruhan citra yang ada dan citra yang sudah di segmentasi akan di ekstraksi. Setelah didapat nilai ekstraksi menggunakan GLCM, maka dilakukan proses klasifikasi menggunakan metode KNN. Data *training* yang digunakan sebanyak 20 data dan data testing sebanyak 10 data dan nilai K yang digunakan adalah 3, sehingga didapat tingkat akurasi sebesar 83.33% [6].

Penelitian untuk klasifikasi kacang-kacangan berdasarkan tekstur menggunakan jaringan syaraf tiruan pada [7], 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% [7].

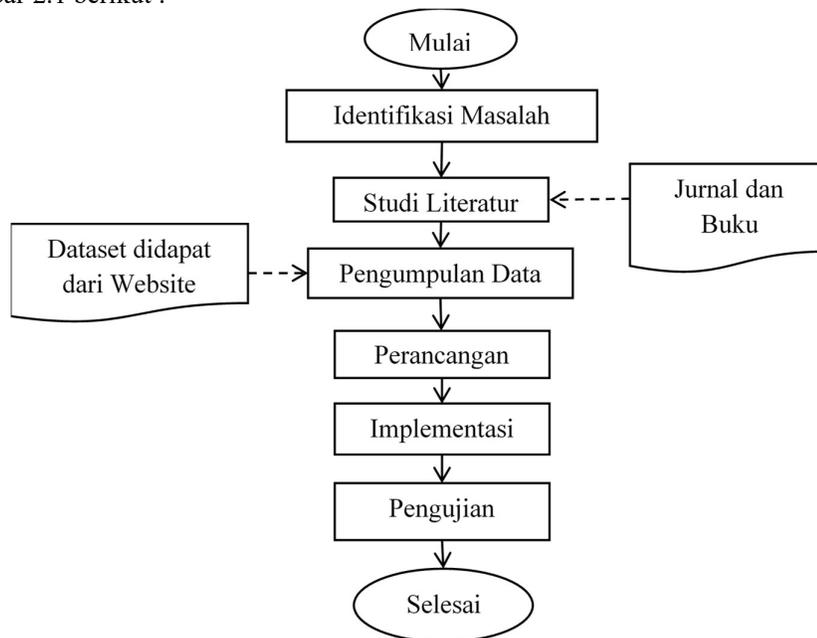
Pada penelitian terkait dengan KNN dilakukan [8] untuk melakukan pengenalan Alfabet *American Sign Language* (ASL) menggunakan KNN dengan ekstraksi fitur *Histogram of Oriented Gradients* (HOG) dengan menggunakan 3 *distance* yaitu *Euclidean Distance*, *Manhattan Distance* dan *Chebychev Distance*, hasil terbaik dari penelitian ini yaitu 0.99 dengan menggunakan *Euclidean Distance* dan *Manhattan Distance* dengan nilai K = 3 dan K = 5 [8].

Penelitian lainnya yang menggunakan segmentasi citra *thresholding otsu* dan menggunakan ekstraksi fitur GLCM untuk mengklasifikasi motif batik pesisir yang menggunakan *K-Nearest Neighbor* untuk mengklasifikasikannya, hasil dari klasifikasi menunjukkan bahwa dengan menambahkan *thresholding otsu* menghasilkan akurasi yang cukup baik, dengan persentase tertinggi sebesar 85% pada $K = 1$ [9].

Berdasarkan uraian pada paragraf sebelumnya, *threshold* merupakan salah satu metode segmentasi citra yang cukup baik dalam mengsegmentasikan citra sebelum citra di ekstraksi menggunakan GLCM, serta metode KNN dapat mengelola secara efektif data *training* dalam skala besar serta mampu mengklasifikasi data berdasarkan kelas yang tepat.

2. METODE PENELITIAN

Berikut adalah beberapa tahapan yang dilakukan dalam penelitian dapat dilihat pada Gambar 2.1 berikut :



Gambar 2.1 Kerangka Kerja Penelitian

2.1 Identifikasi Masalah

Pada tahapan ini melakukan identifikasi masalah sesuai dengan bidang ilmu, dalam hal ini adalah mengenai klasifikasi *pneumonia*.

2.2 Studi Literatur

Tahapan ini merupakan pencarian beberapa jurnal yang terkait mengenai klasifikasi *pneumonia* pada citra *rontgen*, jurnal terkait yang menggunakan segmentasi citra *thresholding*, ekstraksi GLCM, dan metode yang terkait metode klasifikasi *K-Nearest Neighbor* (KNN).

2.1.1 Contrast Stretching

Contrast Stretching merupakan salah satu metode yang sangat berguna dalam meningkatkan kontras terutama terhadap citra yang memiliki kontras rendah. Implementasi *contrast stretching* dilakukan pada tahap *preprocessing*. Perbaikan kontras

citra dengan *contrast stretching* merupakan data asli yang diproses yang artinya proses ini hanya tergantung nilai intensitas (*gray level*) satu piksel dan tidak tergantung dengan piksel lain yang berada disekitarnya [11].

2.1.2 *Thresholding*

Thresholding adalah salah satu metode yang banyak digunakan untuk segmentasi citra. Ini berguna dalam membedakan latar depan dari latar belakang. Proses *thresholding* akan menghasilkan citra yang memiliki dua nilai tingkat keabuan yaitu hitam dan putih atau yang disebut dengan citra biner [3].

2.1.3 *Gray Level Co-Occurance (GLCM)*

GLCM merupakan salah satu metode yang efektif dalam proses ekstraksi ciri suatu citra, dimana metode ini dapat melakukan klasifikasi dan identifikasi. Setelah diekstraksi menggunakan GLCM, akan menghasilkan nilai ciri statistik dengan empat (4) arah (0° , 45° , 90° dan 135°). Karena mampu memberikan informasi yang detail tentang suatu citra dalam hal tekstur. Tahap ini membentuk GLCM 4 arah (0° , 45° , 90° dan 135°) dengan jarak $d=1$ yang akan menentukan koordinat arah. Selanjutnya, dibentuk matriks kookurasi dengan cara 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. Langkah terakhir yaitu menghitung ciri statistic GLCM [6].

2.1.4 *K-Nearest Neighbor (KNN)*

K-Nearest Neighbor (KNN) Classifier adalah sebuah metode untuk melakukan klasifikasi terhadap objek berdasarkan data pembelajaran yang jaraknya paling dekat dengan objek tersebut. Data pembelajaran diproyeksikan ke ruang berdimensi banyak, yang masing-masing dimensi merepresentasikan fitur dari data. Ruang dimensi dibagi menjadi bagian-bagian berdasarkan klasifikasi data pembelajaran. Nilai k yang terbaik untuk algoritma ini tergantung pada data, secara umum nilai k yang tinggi akan mengurangi efek *noise* pada klasifikasi, akan tetapi membuat batasan antara setiap klasifikasi menjadi lebih buram. Nilai k yang bagus dapat dipilih dengan optimis parameter, misalnya dengan menggunakan *cross-validation*. Kasus khusus untuk klasifikasi diprediksikan berdasarkan data pembelajaran yang paling dekat (dengan kata lain, $k = 1$) yang biasanya disebut algoritma *nearest neighbor* [9]. Metode *Euclidean Distance* merupakan penghitungan jarak pada algoritma KNN yang paling banyak digunakan oleh peneliti.

Rumus *Euclidean Distance* :

$$d(x,y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1)$$

untuk,

- $D(x,y)$ adalah jarak antara data x ke data y
- x_i adalah data testing ke- i
- y_i adalah data training ke- i
- n adalah dimensi data

2.3 Pengumpulan Data

Pengumpulan data menggunakan dataset yang didapat dari website kaggle yaitu <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia> dengan jumlah 5.863 citra

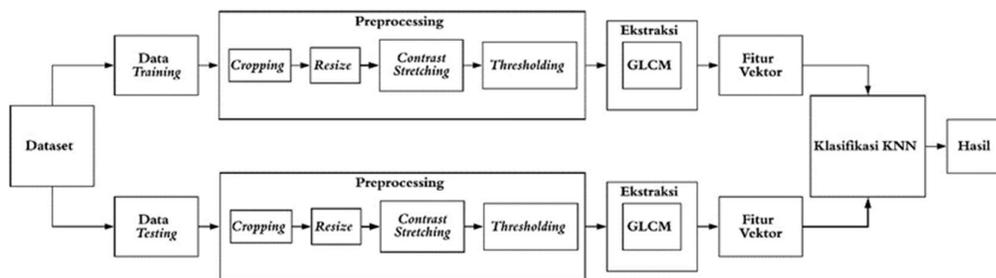
rontgen paru-paru. Gambar-gambar dibagi menjadi tiga (3) jenis citra rontgen diantaranya paru-paru normal, paru-paru pneumonia bakteri, dan paru-paru pneumonia virus. Citra rontgen paru-paru dilakukan *cropping* lalu *Resize* menjadi 300x300 piksel, dengan tujuan untuk membuang bagian yang tidak penting yang dapat mempengaruhi proses, yaitu area diluar dari rongga dada atau diluar area paru-paru. Dataset diambil dari hasil rontgen anak-anak dengan rentang usia 1-5 tahun, karena pada anak-anak dengan rentang usia tersebut daya tahannya paling rendah sehingga mudah terserang *pneumonia*. Berikut adalah jumlah data yang terdapat pada setiap citra rontgen dapat dilihat pada Tabel 2.1 :

Tabel 2.1 Jumlah Data Setiap Jenis Citra Rontgen

Jenis – jenis Citra	Jumlah data yang dipilih	Jumlah data <i>Training</i>	Jumlah data <i>Testing</i>
Normal	1400 gambar	1200 gambar	200 gambar
<i>Pneumonia</i> Bakteri	1400 gambar	1200 gambar	200 gambar
<i>Pneumonia</i> Virus	1400 gambar	1200 gambar	200 gambar
Total	4200 gambar	3600 gambar	600 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 *K-Nearest Neighbor (KNN)*. Berikut adalah gambar dari skema dalam penelitian dapat dilihat pada Gambar 2.2 berikut :



Gambar 2.2 Perancangan Skema

2.5 Implementasi

Setelah melakukan proses perancangan, pada tahap ini dilakukan implementasi terhadap rancangan yang telah dibuat ke dalam program matlab. Tahapan pertama dilakukan *cropping* pada citra dengan ukuran 300x300 piksel, selanjutnya citra dilakukan peningkatan kontras dengan *Contrast Stretching* agar didapat kualitas citra yang lebih baik. Kemudian citra yang telah dilakukan peningkatan kontras akan di-*thresholding* dengan hasil berupa citra biner. Setelah didapat citra biner, maka tahapan selanjutnya dilakukan ekstraksi GLCM untuk mendapatkan

empat (4) ciri dari citra itu sendiri yaitu energi, kontras, korelasi dan homogenitas. Setelah didapat nilai ciri dari citra, maka proses selanjutnya yaitu klasifikasi menggunakan *K-Nearest Neighbor* (KNN).

2.6 Pengujian

Setelah melakukan proses implementasi dilanjutkan dengan proses pengujian pada citra *training* dan citra *testing* secara acak sebanyak tiga (3) kali untuk satu (1) jenis citra rontgen dengan *cross validation*. Hasil dari pengujian yang didapatkan berupa nilai akurasi yang berasal dari perhitungan rata-rata sebanyak tiga (3) kali pada pengujian tersebut. Selanjutnya, hasil klasifikasi menggunakan *K-Nearest Neighbor* (KNN) dihitung untuk mendapatkan tingkat keberhasilan dari metode yang digunakan, dengan menggunakan metode *Confusion Matrix* yang akan menghitung nilai *Precision*, *Recall*, dan *Accuracy*. Cara penghitungan *Confusion Matrix* tersebut bisa dilihat pada Tabel 2.2 dan persamaan (2), (3), (4), dan (5) berikut :

Tabel 2.2 *Confusion Matrix*

		True Values	
		Positive	Negative
Prediction	Positive	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

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

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

$$Precision = \frac{TP}{(TP+FP)} * 100\% \quad (4)$$

$$Overall Accuracy = \frac{\sum Data Uji Benar}{\sum Total Uji} * 100\% \quad (5)$$

Keterangan :

TP = Jumlah data positif citra rontgen yang terklasifikasi dengan benar oleh sistem.

TN = Jumlah data negatif citra rontgen yang terklasifikasi dengan benar oleh sistem.

FN = Jumlah data negatif citra rontgen namun terklasifikasi salah oleh sistem.

FP = Jumlah data positif citra rontgen namun terklasifikasi salah oleh sistem.

3. HASIL DAN PEMBAHASAN

Pada tahap ini membahas tentang implementasi dimana terdapat penjelasan mengenai implementasi pada tahap *preprocessing* yang mana terdapat tiga (3) proses diantaranya *cropping*, *Contrast Stretching* dan *Thresholding*, implementasi ekstraksi GLCM, implementasi klasifikasi menggunakan KNN, 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*

Cropping dilakukan untuk mendapatkan citra yang lebih detail dikarenakan akan membuang bagian bagian yang dianggap tidak penting atau mengganggu sehingga didapat citra dengan bagian dada dan paru-paru saja. Berikut citra sebelum di-*cropping* yang memiliki ukuran asli dan citra sesudah di-*cropping* :



Sumber: kaggle.com/paultimothymooney/chest-xray-pneumonia

Gambar 3.1 Citra Asli



Gambar 3.2 Citra Setelah di-*Cropping*

3.4.2 *Resize*

Resize dilakukan untuk mengatur ukuran piksel menjadi 300x300 piksel pada citra yang digunakan. Semakin banyak jumlah piksel pada citra maka akan semakin banyak waktu untuk proses pengolahan citra. Berikut gambar sesudah dilakukan *Resize* :

Gambar 3.3 Citra Setelah di-*Resize*

3.4.3 *Contrast Stretching*

Contrast Stretching dilakukan untuk mendapat RGB baru dengan kualitas kontras yang lebih baik. Dengan citra yang lebih baik pada citra dapat meningkatkan ketajaman warna objek pada citra. Objek yang terlihat jelas pada citra dapat membantu pada proses selanjutnya yaitu *thresholding*. Berikut gambar sesudah dilakukan *contrast stretching* :

Gambar 3.4 Citra Setelah di-*Contrast Stretching*

3.4.4 *Thresholding*

Pada proses *thresholding* akan menghasilkan citra yang memiliki dua nilai tingkat keabuan yaitu hitam dan putih atau yang disebut dengan citra biner. Parameter nilai yang digunakan sebagai nilai *threshold* adalah 150. Berikut gambar sesudah dilakukan *thresholding* :



Gambar 0.5 Citra Setelah di-*Thresholding*

3.2 Implementasi Ekstraksi GLCM

Pada tahap ekstraksi GLCM, citra biner yang berukuran 300x300 piksel menghasilkan nilai ciri statistik dengan empat (4) arah (0° , 45° , 90° dan 135°). Selanjutnya, dibentuk matriks kookurasi dengan cara 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 statistic GLCM dari citra tersebut, ciri-ciri dari citra dikategorikan menjadi 4 nilai ciri yaitu energi, kontras, korelasi dan homogenitas. Hasil dari ekstraksi GLCM berupa nilai *vector* dari matriks berukuran 3600 x 5.

3.3 Implementasi Klasifikasi KNN

Pada tahapan ini, implementasi klasifikasi KNN akan melanjutkan hasil dari ekstraksi GLCM yang mana hasil dari ekstraksi GLCM tersebut akan di proses oleh klasifikasi KNN sehingga didapat hasil berupa model. Model yang didapatkan dari klasifikasi KNN akan digunakan untuk proses pengujian (*testing*). Hasil dari pengujian (*testing*) yang didapat akan dihitung menggunakan *Euclidean Distance* dengan menggunakan nilai $k=3$, $k=5$, dan $k=7$ untuk masing-masing *distance*.

3.4 Hasil Klasifikasi KNN

Setelah dilakukan implementasi klasifikasi KNN, selanjutnya didapat hasil berupa *Confussion Matrix* dengan nilai $K = 3$, $K = 5$, dan $K = 7$.

3.4.1 *Confussion Matrix* $K = 3$

Pada tahapan ini merupakan tahapan *confussion matrix* melalui klasifikasi *K-Nearest Neighbor* (KNN) untuk $K = 3$ dengan melakukan *cropping* pada citra. Sehingga didapat nilai dari hasil *Confussion Matrix* pada Gambar 4.6 :

Pneumonia Classification Using KNN K=3

True Class	1	129	51	20	64.5%	35.5%
	2	97	41	62	20.5%	79.5%
	3	91	49	60	30.0%	70.0%
		40.7%	29.1%	42.3%		
		59.3%	70.9%	57.7%		
		1	2	3		
		Predicted Class				

Gambar 4.6 Confussion Matrix K = 3

Hasil klasifikasi dengan menggunakan *K-Nearest Neighbor* (KNN) untuk $K = 3$ untuk jenis citra rontgen paru-paru normal dapat dikenali sebanyak 129 citra, nilai *recall* sebesar 64,5% dan nilai *preccission* sebesar 40,7% dapat dikatakan bagus dalam mengklasifikasikan citra rontgen paru-paru normal, untuk jenis citra rontgen paru-paru *pneumonia bacteria* dapat mengenali 41 citra, persentase *recall* sebesar 20,5% dan persentase *preccission* sebesar 29,1% dapat dikatakan buruk dan jenis citra rontgen paru-paru *pneumonia virus* dapat mengenali 60 citra, nilai *recall* sebesar 30,0% dan nilai *preccission* sebesar 42,3% dapat dikatakan buruk dalam mengenali citra rontgen *pneumonia virus*. Hasil *confussion matrix K-Nearest Neighbor* (KNN) untuk $K = 3$ dapat dikatakan buruk karena perbandingan dari ketiga citra tersebut tidak seimbang atau bisa dikatakan hanya citra rontgen paru-paru normal saja yang cukup banyak dikenali sedangkan untuk citra rontgen paru-paru *pneumonia bacteria* dan *pneumonia virus* cukup banyak yang tidak dapat dikenali.

3.4.2 Confussion Matrix K = 5

Pada tahapan ini merupakan tahapan *confussion matrix* melalui klasifikasi *K-Nearest Neighbor* untuk $K = 5$ dengan melakukan *cropping* pada citra. Sehingga didapat nilai dari hasil *Confussion Matrix* pada Gambar 4.7 berikut :

Pneumonia Classification Using KNN K=5

True Class	1	129	54	17	64.5%	35.5%
	2	89	45	66	22.5%	77.5%
	3	89	51	60	30.0%	70.0%
		42.0%	30.0%	42.0%		
		58.0%	70.0%	58.0%		
		1	2	3		
		Predicted Class				

Gambar 4.7 Confussion Matrix K = 5

3.4.3 Confussion Matrix K = 7

Pada tahapan ini merupakan tahapan *confussion matrix* melalui klasifikasi *K-Nearest Neighbor* untuk $K = 7$ dengan melakukan *cropping* pada citra. Sehingga didapat nilai dari hasil *Confussion Matrix* pada Gambar 4.8 :

True Class	1	127	51	22	63.5%	36.5%
	2	83	39	78	19.5%	80.5%
	3	90	46	64	32.0%	68.0%
		42.3%	28.7%	39.0%		
		57.7%	71.3%	61.0%		
		1	2	3		
		Predicted Class				

Gambar 4.8 Confussion Matrix K = 7

4. KESIMPULAN

Berdasarkan hasil penelitian implementasi ekstraksi *Gray Level Co-Occurrence* (GLCM) dan *K-Nearest Neighbor* untuk mengklasifikasikan pneumonia dapat disimpulkan sebagai berikut :

1. Dengan melakukan proses *cropping* pada citra rontgen paru-paru dapat meningkatkan hasil dari akurasi.
2. Dengan menggunakan ekstraksi *Gray Level Co-Occurrence* (GLCM) dan klasifikasi *K-Nearest Neighbor* (KNN) untuk $K = 5$ mendapatkan *Accuracy* per *class* tertinggi sebesar 66,20%.
3. Objek citra paru-paru Virus memiliki tingkat *Accuracy* paling tinggi yaitu sebesar 72,90%.
4. Tingkat *Precision* dan *Recall* cenderung memiliki nilai yang sama dikarenakan *Gray Level Co-Occurrence* (GLCM) tidak mempresentasikan nilai yang tepat.

5. SARAN

Berdasarkan kesimpulan tersebut, untuk menghasilkan *Accuracy* yang lebih baik, berikut saran yang nantinya dapat dipertimbangkan untuk penelitian selanjutnya.

1. Menggunakan ekstraksi fitur citra dan klasifikasi yang lebih baik guna bisa meningkatkan tingkat akurasi.
2. Menggabungkan beberapa ekstraksi fitur yang sesuai sehingga nilai ciri dari citra yang akan diperoleh menjadi lebih baik.
3. Sebaiknya dataset diperoleh langsung dari lembaga yang sudah terpercaya sehingga diperoleh dataset yang valid.
4. Menggunakan citra yang jelas dan jumlah data latih lebih banyak sehingga ketika diuji akan menghasilkan akurasi yang tinggi.

DAFTAR PUSTAKA

- [1] F. A. P. Scholastica, *Asuhan Keperawatan Pada Pasien Dengan Gangguan Sistem Pernapasan*. Yogyakarta : Pustaka Baru Press, 2018.
- [2] F. Liantoni, "Klasifikasi Daun Dengan Perbaikan Fitur Citra Menggunakan Metode K-Nearest Neighbor", *Ultimatics : Jurnal Teknik Informatika*, vol. 7, no. 2, pp. 98-104, Aug. 2016.
- [3] W. Christanto, "Klasifikasi pneumonia menggunakan convolutional neural network". *Repositori Institusi Universitas Sumatera Utara*, 2018.
- [4] T. D. Wulan, "Ekstraksi fitur citra paru-paru menggunakan gray level co-occurrence matriks", *Technology Science and Engineering Journal*, vol. 1, no. 3, pp. 225–229, 2017.
- [5] E. Prasetyo, *Data Mining Mengolah Data Menggunakan Matlab*. Yogyakarta : Penerbit Andi, 2012.
- [6] J. Sofian and R. H. Laluma, "Jenis Tumor Otak Dengan Metode Image Threshold Dan Glcm Menggunakan Algoritma K-NN (Nearest Neighbor) Classifier Berbasis Web," *Jurnal Infotronik*, vol. 4, no. 2, pp. 1-7, 2019.
- [7] M. E. A. Rivan, N. Rachmat, & M. R. Ayustin, "Klasifikasi Jenis Kacang-Kacangan Berdasarkan Tekstur Menggunakan Jaringan Syaraf Tiruan", *jkt*, vol. 6, no. 1, May 2020.
- [8] M. E. A. Rivan, H. Irsyad, Kevin, A. T. Narta, "Pengenalan Alfabet American Sign Language", *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 5, no. 3, pp. 328-339, 2019.
- [9] N. Wakidah, "Perbaikan Kualitas Citra Menggunakan Metode Contrast Stretching (Improvement of image quality using a method Contrast Stretching)", *Jurnal Transformatika*, vol. 8, no. 2, pp. 78-83, 2011.
- [10] N. K. Ningrum, N. Hendriyanto dan D. Kurniawan, "Pengaruh thresholding otsu pada klasifikasi motif batik pesisir menggunakan k-nearest neighbour", *Science And Engineering National Seminar 3 (SENS 3)*, pp. 142-147, 2017.
- [11] F. Liantoni, & H. Nugroho, "Klasifikasi Daun Herbal Menggunakan Metode Naïve Bayes Classifier Dan Knearest Neighbor", *Jurnal Simantec*, vol. 5, no. 1, 2015.