



CLUSTERING PENYAKIT PARU-PARU BERDASARKAN RONTGEN DADA MENGGUNAKAN DIMENSI FRAKTAL BOX COUNTING DAN K-MEDOIDS

ELSHA HANY PRATIWI^{1*}, DWI JUNIATI²

^{1,2}Universitas Negeri Surabaya

*elsha.18004@mhs.unesa.ac.id

ABSTRAK

Paru-paru merupakan salah satu organ pernapasan pada manusia yang rentan terhadap penyakit karena mudah terinfeksi. Saluran pernapasan merupakan organ bagian dalam manusia yang diserang oleh virus SARS-Cov-2 yang menyebabkan penyakit corona (Covid-19). Selain itu, Lung Opacity dan Viral Pneumonia juga merupakan penyakit yang terjadi pada paru-paru. Penyakit pada paru-paru dapat dideteksi menggunakan citra rontgen dada (*chest x-ray*), tetapi untuk mendiagnosis dari hasil citra rontgen dada tersebut tidak mudah untuk dianalisis, selain dibutuhkan hasil rontgen yang cukup baik juga dibutuhkan pengetahuan ahli untuk menentukan jenis penyakit. Metode matematika yang banyak mempelajari bentuk tak teratur adalah geometri fraktal dan dimensi fraktal telah banyak diaplikasikan dalam bidang kedokteran dalam menentukan jenis penyakit. Dalam penelitian ini dilakukan klasifikasi penyakit paru berdasarkan dimensi fraktal dari citra rontgen dada. Terdapat 100 citra rontgen dada yang akan diolah menggunakan segmentasi yang menghasilkan wilayah (*region*) paru-paru. *Region* tersebut digunakan untuk mengetahui bercak dari jenis penyakit paru-paru menggunakan deteksi tepi *Canny*. Kemudian dengan menggunakan fraktal *box counting* akan dihitung nilai dimensi fraktal untuk digunakan dalam pengklasteran. Hasil dari percobaan menggunakan metode *Clustering K-Medoids* dengan empat *cluster* yaitu Covid-19, Lung Opacity, Normal Lung, dan Viral Pneumonia memiliki akurasi sebesar 87%.

Kata Kunci: Rontgen dada, *Box Counting*, *K-Medoids*.

ABSTRACT

Lungs are one of the respiratory organs in humans that are susceptible to disease because they are easily infected. The respiratory tract is an internal human organ that is attacked by the SARS-Cov-2 virus that causes the corona disease (Covid-19). In addition, Lung Opacity and Viral Pneumonia are also diseases that occur in the lungs. Diseases of the lungs can be detected using chest x-ray images, but to diagnose from the results of the chest x-ray image it is not easy to analyze, in addition to the need for a fairly good x-ray result expert knowledge is also needed to determine the type of disease. The mathematical method that studies irregular shapes is fractal geometry and fractal dimensions have been widely applied in the medical field in determining the type of disease. In this research, lung diseases will be classified based on chest x-ray images. There are 100 chest x-ray images that will be processed using segmentation that result a region of the lungs. The region is used to determine the spots of the type of lung disease using Canny edge detection. Then by using fractal box counting, the fractal dimension value will be calculated for use in clustering. The results of the experiment using the K-Medoids clustering method with four clusters, that is Covid-19, Lung Opacity, Normal Lung, and Viral Pneumonia have an accuracy of 87%.

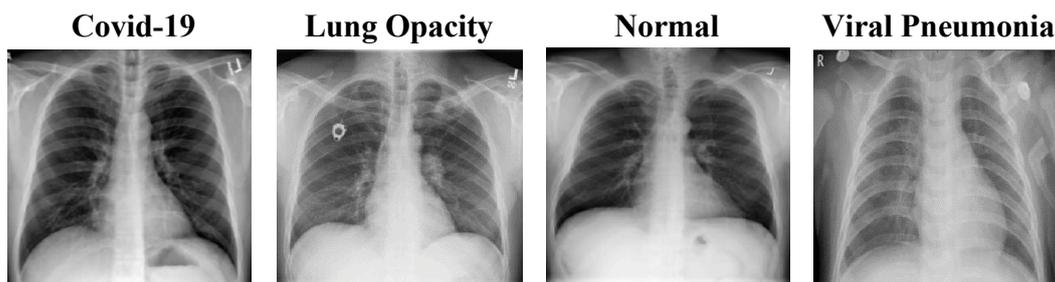
Keywords: *Chest X-ray, Box Counting, K-Medoids.*

1 Pendahuluan

Sistem pernapasan yang ada pada manusia berfungsi sebagai tempat bertukarnya gas dalam darah. Paru-paru merupakan salah satu saluran pernapasan yang ada pada manusia. Di dalam paru-paru terjadi pertukaran oksigen dengan karbondioksida. Paru-paru merupakan organ vital yang rentan terhadap penyakit karena mudah terinfeksi oleh virus, bakteri, maupun jamur. Beberapa contoh penyakit pada paru-paru adalah Covid-19, Lung Opacity, dan Viral Pneumonia. Covid-19 merupakan penyakit yang disebabkan oleh virus SARS-Cov-2, hasil rontgen paru-paru pasien Covid-19 tidak beda jauh dengan pasien Pneumonia, oleh karena itu para ahli masih meneliti lebih jauh cairan atau flek paru-paru pada pasien covid-19.

Lung Opacity biasanya ditandai dengan hasil citra rontgen pada area tertentu akan terlihat lebih buram dari pada area di sekitarnya, Lung Opacity berbeda dengan Pneumonia, kekeruhan pada Lung Opacity tidak jelas terdapat awan putih kabur pada paru-paru, oleh karena itu diagnosis Lung Opacity sangat luas seperti infeksi pada paru-paru, pendarahan paru, maupun edema paru. Viral Pneumonia disebabkan oleh virus, bakteri, atau jamur yang mengakibatkan peradangan pada paru-paru, meskipun resiko tertinggi pada Covid-19 adalah Pneumonia namun tidak semua Pneumonia adalah Covid-19, karena Pneumonia pada Covid-19 disebabkan oleh virus SARS-Cov-2.

Pada gambar 1 merupakan perbedaan masing-masing gambar citra yang dihasilkan dari proses rontgen dada.



Gambar 1: Citra Rontgen Masing-masing Penyakit Paru

Untuk hasil citra rontgen Covid-19 terlihat mirip dengan hasil citra rontgen Pneumonia, namun pada citra Pneumonia terlihat lebih memanjang dan lebih terlihat kabur atau kurang tajam. Untuk hasil citra rontgen Lung Opacity terlihat terdapat area yang berawan putih dan terlihat kabur pada area tertentu saja berbeda dengan Viral Pneumonia terlihat hampir keseluruhan area paru-paru terlihat kabur, bentuk dari citra Lung Opacity sedikit melebar. Untuk hasil rontgen citra normal terlihat tidak ada bintik putih atau area berawan putih hanya saja terlihat ada garis-garis putih yaitu ketika *x-ray* mengenai tulang rusuk.

Infeksi pada paru-paru akan mengganggu berlangsungnya proses pernapasan seperti kesulitan dalam bernafas, kesulitan dalam beraktivitas, kekurangan oksigen, bahkan gangguan tersebut dapat menyebabkan kematian jika tidak segera ditangani dengan cepat [1]. Setiap jenis penyakit paru membutuhkan penanganan dan pengobatan yang berbeda, meskipun seringkali tanda-tanda yang muncul hampir mirip. Oleh karena itu analisis jenis penyakit paru sangat menentukan untuk tindakan selanjutnya sehingga pasien bisa disembuhkan. Maka dari itu identifikasi awal dalam mendiagnosis gangguan pada paru-paru sangat penting agar dapat segera ditangani dengan tepat. Ada beberapa cara untuk mengidentifikasi gangguan yang terjadi pada paru-paru dengan pemeriksaan pada gejala klinis, pemeriksaan secara medis, atau dengan menganalisis gambar citra yaitu citra rontgen (*x-ray*), citra MRI, dan citra CT scan [2].

Citra rontgen merupakan citra yang membutuhkan biaya lebih terjangkau dibandingkan dengan dua citra lainnya yang cukup mahal. Untuk menganalisa hasil citra rontgen lebih lanjut untuk mengetahui gangguan apa yang terjadi pada paru-paru, dibutuhkan pengetahuan khusus dan pengamat tenaga ahli seperti dokter dan tenaga medis lainnya. Tak jarang hasil dari citra

rontgen terlihat kabur atau kurang kontras, sehingga sulit bagi pengamat untuk menganalisis hasil citra [3].

Dengan adanya kemajuan teknologi, kualitas dari hasil rontgen dapat diperbaiki. Saat ini pengolahan citra digital sangat berkembang pesat terutama pada dunia medis. Pengolahan citra digital dapat digunakan untuk menganalisis hasil citra rontgen yang kabur sehingga dapat terlihat lebih jelas. Analisis rontgen selain dengan cara mengamati hasil rontgen oleh ahli, banyak ahli selain bidang kedokteran yang berusaha membantu untuk analisis dengan menggunakan teknologi dan bermacam teori termasuk diantaranya teori matematika. Salah satu teori yang telah banyak digunakan dalam analisis jenis penyakit adalah dimensi fraktal. Geometri fraktal merupakan bidang ilmu matematika yang mempelajari bentuk-bentuk yang tidak beraturan sehingga cocok untuk diaplikasikan dalam analisa rontgen. Beberapa hasil aplikasi dimensi fraktal dalam pengelompokan jenis penyakit antara lain: mengelompokkan jenis tumor kulit berdasarkan dimensi fractal [4], mengelompokkan retinopati diabetik berdasarkan dimensi fraktal [5], mengelompokkan jenis penyakit jantung berdasarkan dimensi fraktal dari suara jantung [6], klasifikasi penyakit paru berdasarkan dimensi fraktal menggunakan citra *x-ray* [7], klasifikasi suara tangisan bayi menggunakan fraktal Higuchi [8]. Pada penelitian sebelumnya yang dilakukan oleh Fani & Juniati pada artikel yang berjudul “Analisis Jenis Penyakit Paru-Paru Berdasarkan *Chest X-Ray* Menggunakan Metode Fuzzy *C-Means*”, menghasilkan akurasi 86% [9].

Dalam penelitian ini, dimensi fraktal akan digunakan untuk mengelompokkan penyakit paru yaitu Covid-19, Lung Opacity, Normal Lung, dan Viral Pneumonia dan menggunakan metode yang berbeda, yaitu *Clustering K-Medoids*. Menggunakan metode *K-Medoids* karena metode ini memiliki keunggulan dalam mengatasi kekurangan metode *K-Means* yaitu sensitif terhadap noise dan outlier ketika objek bernilai besar memungkinkan untuk menyimpang dari distribusi data [16]. Pada penelitian ini juga menghasilkan akurasi lebih baik yaitu 87% dibandingkan dengan akurasi pada penelitian sebelumnya yaitu menggunakan metode *K-Means* 86% dan metode K-NN 79,65%. Keberhasilan pengelompokan jenis penyakit paru berdasarkan dimensi fraktal tentunya akan membawa dampak yang positif dalam membantu bidang kesehatan terutama dengan kondisi saat ini dimana masih belum hilangnya kondisi pandemi yang diakibatkan penyakit covid-19. Sehingga diharapkan hasil penelitian ini nantinya bisa dilanjutkan untuk penelitian lebih dalam dalam menentukan seseorang terkena penyakit paru-paru karena covid-19 atau tidak.

2 Tinjauan Pustaka

2.1 Citra Rontgen

Citra Rontgen atau yang biasa disebut dengan *X-Ray* merupakan salah satu metode yang digunakan untuk pemeriksaan pada bagian dalam tubuh manusia. Hal tersebut biasanya dilakukan untuk keperluan medis dalam mendiagnosa penyakit dalam dengan tepat dan akurat. Citra rontgen memanfaatkan sinar radiasi elektromagnetik untuk memperoleh gambar bagian dalam tubuh manusia salah satunya pada paru-paru. Proses tersebut menggunakan mesin untuk mengantarkan sinar radiasi elektromagnetik ke dalam tubuh manusia. Citra rontgen akan menghasilkan warna pada gambar yang berbeda-beda seperti hitam, abu-abu, dan putih, hal tersebut dikarenakan sinar radiasi yang diterima oleh masing-masing bagian tubuh akan berbeda-beda. Warna hitam biasanya menunjukkan jika *x-ray* mengenai udara, warna abu-abu menunjukkan jika *x-ray* mengenai cairan, lemak, dan otot, dan bagian tubuh yang padat seperti tulang atau logam akan menghasilkan warna putih. *X-Ray* memang sangat bermanfaat untuk mendiagnosis lebih lanjut pada penyakit dalam tubuh manusia tetapi karena pemeriksaan ini menggunakan sinar radiasi maka akan beresiko pada kesehatan bahkan dapat muncul adanya

kanker jika dilakukan terlalu sering. Oleh karena itu, dalam dunia medis tidak dianjurkan untuk dilakukan secara berlebihan pada tubuh manusia [10].

2.2 Covid-19

Covid-19 atau penyakit virus corona merupakan penyakit yang saat ini sedang dihadapi oleh dunia. Covid-19 merupakan penyakit yang terjadi pada saluran pernapasan manusia yang diserang oleh virus SARS-Cov-2. Pada sebagian orang yang terserang penyakit covid-19 dapat pulih dengan sendirinya dan ada juga yang butuh penanganan khusus karena mengalami gejala berat. Seperti diketahui banyak pasien yang mengalami gejala serius dan dampak yang cukup berat pada paru-paru ketika terinfeksi covid-19 sehingga tidak sedikit yang mengakibatkan kematian. Penyebaran virus covid-19 terjadi karena percikan droplet (cairan yang dihasilkan oleh saluran pernapasan) saat orang bersin, berbicara, atau bernapas. Virus ini juga mudah menyebar di tempat yang ramai dan dalam ruangan yang tertutup.

2.3 Lung Opacity

Opacity adalah area pada paru-paru yang melemahkan sinar *x-ray* sehingga akan tampak lebih buram dari area sekitarnya. Penyakit Lung Opacity biasanya ditandai dengan hasil citra rontgen pada area tertentu akan terlihat lebih buram dari pada area di sekitarnya. Diagnosis lung opacity sangat luas seperti infeksi pada paru-paru, pendarahan paru, maupun edema paru.

2.4 Normal Lung

Normal lung merupakan kondisi dimana paru-paru tidak terdeksi gangguan atau penyakit apapun. Hasil yang diperoleh dari citra rontgen menunjukkan paru-paru yang bersih tanpa area abnormal pada gambar, dimana warna yang dihasilkan sesuai dengan yang seharusnya. Seperti pada tulang akan berwarna putih, udara yang berada pada paru-paru akan berwarna hitam dan warna abu-abu pada cairan, lemak, dan otot.

2.5 Viral Pneumonia

Pneumonia merupakan penyakit pada paru-paru yang disebabkan oleh virus, bakteri, atau jamur yang mengakibatkan peradangan dimana adanya cairan atau nanah pada kantung udara di dalam paru-paru. Penyakit ini ditandai dengan adanya bintik putih pada hasil citra rontgen. Pneumonia menyebabkan kesulitan dalam bernapas, batuk berdahak, nyeri pada dada, atau bahkan dapat menyebabkan demam. Pneumonia dibagi menjadi beberapa jenis, salah satunya yaitu viral pneumonia. Dibandingkan dengan jenis yang lain viral pneumonia biasanya terjadi lebih singkat serta mengalami gejala yang lebih ringan. Akan tetapi viral pneumonia dapat beresiko tinggi pada orang yang memiliki daya tahan tubuh yang tidak kuat.

2.6 Histogram

Histogram *equalization* merupakan suatu proses pengolahan pada citra digital untuk meningkatkan kualitas citra [3]. Pengolahan ini dilakukan dengan cara meningkatkan kontras gambar dengan mendistribusikan skala keabuan pada citra agar merata. Dengan pengolahan ini, kontras rendah dapat menjadi tinggi karena intensitas didistribusikan secara merata pada citra. Kualitas gambar ditingkatkan dengan histogram *equalization* diperlukan karena citra *x-ray* yang akan diolah kurang jelas, seperti gambar yang gelap, terlalu terang maupun kurang tajam.

2.7 Segmentasi

Pada proses pengolahan citra digital, tak jarang hanya menginginkan pengolahan pada wilayah (*region*) tertentu saja. Maka dari itu perlu dilakukan segmentasi untuk memisahkan *region* yang diinginkan (*foreground*) dengan *background* agar citra lebih mudah dianalisa [11]. Ada beberapa cara untuk pengolahan segmentasi, *Threshold* merupakan salah satu cara pengolahan segmentasi citra berdasarkan pada perbedaan tingkat keabuan citra. Dalam proses

ini diperlukan nilai batas yaitu nilai *threshold* (T), dimana wilayah (*region*) diubah menjadi putih (1) jika nilai intensitas gambar lebih besar atau sama dengan *threshold* (T) dan wilayah (*region*) diubah menjadi hitam (0) jika nilai intensitas gambar lebih kecil dari *threshold* (T). Metode segmentasi *threshold* dapat ditulis sebagai [12]:

$$T = T[x, y, f(x, y), g(x, y)] \quad (1)$$

dimana

T : nilai *threshold*
 x, y : koordinat titik *threshold*
 $f(x, y), g(x, y)$: titik piksel tingkat keabuan
 Sehingga segmentasi *threshold*

$$s(x, y) = \begin{cases} 1, & \text{jika } g(x, y) \geq T \\ 0, & \text{jika } g(x, y) < T \end{cases} \quad (2)$$

2.8 Deteksi tepi

Deteksi tepi merupakan metode yang menghasilkan tepian pada gambar hal ini dilakukan untuk mendeteksi ciri suatu objek pada gambar citra. Ada beberapa metode dalam deteksi tepi, yaitu *Robert*, *Prewitt*, *Sobel*, dan *Canny*. Dalam penelitian ini menggunakan deteksi tepi *Canny* karena merupakan metode deteksi tepi dengan hasil yang optimal [13].

2.9 Dimensi Fraktal *Box Counting*

Dimensi fraktal merupakan dimensi yang memperlihatkan tingkat kepadatan suatu objek. Pada penelitian ini menggunakan perhitungan nilai dimensi fraktal *box counting*. Metode tersebut dilakukan menggunakan cara menutupi objek dengan kotak-kotak (*box*) dengan ukuran yang berbeda-beda [14]. Dimensi *box counting* dapat diperoleh dari persamaan hubungan rasio persegi dengan banyaknya persegi:

$$N(r) = \frac{1}{r^D} \quad (3)$$

dimana

$N(r)$: banyak *box* yang menutupi objek
 r : rasio
 D : dimensi

Selanjutnya, persamaan (3) diubah menjadi bentuk logaritma:

$$\log(N(r)) = D \log\left(\frac{1}{r}\right) \quad (4)$$

$$D = \frac{\log N(r)}{\log\left(\frac{1}{r}\right)} \quad (5)$$

Persamaan pertama diubah menjadi bentuk logaritma untuk memperoleh nilai gradien D (α). Nilai α didapatkan dari merepresentasikan hasil pengukuran D dalam bentuk grafik, maka:

$$\alpha = \frac{n(\sum_{k=1}^n xy) - (\sum_{k=1}^n x)(\sum_{k=1}^n y)}{n(\sum_{k=1}^n x^2) - (\sum_{k=1}^n x)^2} \quad (6)$$

2.10 *K-Medoids*

K-Medoids atau algoritma *Partitioning Around Medoids* (PAM) merupakan metode *clustering* yang tidak beda jauh dengan *K-Means*. Metode *K-Medoids* menentukan sebarang objek yang mewakili (*medoid*) setiap *cluster* sebagai pusat *cluster*, sedangkan metode *K-Means* mencari nilai rata-rata (*mean*) setiap *cluster* sebagai pusat *cluster* [15]. Metode *K-Medoids* memiliki keunggulan dalam mengatasi kekurangan metode *K-Means* yaitu sensitif terhadap

noise dan outlier ketika objek bernilai besar memungkinkan untuk menyimpang dari distribusi data [16]. Metode *K-Medoids* menggunakan objek perwakilan (*medoid*) dengan metode pengelompokan partisi untuk mengelompokkan sekumpulan objek (*n*) menjadi beberapa *cluster* (*k*). Langkah – langkah dari metode *K-Medoids* [16]:

1. Inisialisasi sebanyak *k* (jumlah *cluster*) sebagai pusat *cluster*.
2. Kelompokkan setiap data (objek) ke dalam *cluster* terdekat menggunakan persamaan *Euclidian Distance*:

$$d(x_{ij}, y_{kj}) = \sqrt{\sum_{j=1}^m \sum_{i=1}^n (x_{ij} - y_{kj})^2} \tag{7}$$

dimana

- $d(x_{ij}, y_{kj})$ jarak *Euclidian Distance*
- x_{ij} objek ke-*i* pada variabel ke-*j*
- y_{kj} pusat *cluster* ke-*k* pada variabel ke-*j*
- n* banyak objek
- m* banyak variabel

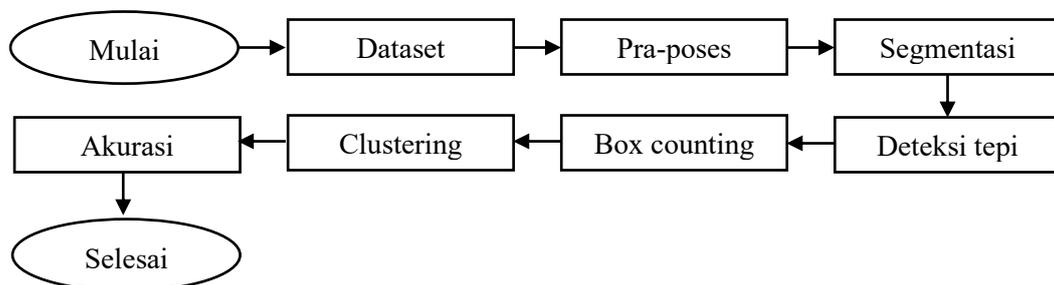
3. Pilih pusat *cluster* (*medoid*) baru secara acak pada masing-masing *cluster*.
4. Dengan kandidat *medoid* baru hitung kembali jarak setiap objek pada masing-masing *cluster* menggunakan persamaan *Euclidian Distance*.
5. Hitung nilai simpangan (*S*) dengan menghitung total *distance* baru dikurangi dengan total *distance* lama. Jika $S < 0$, maka pilih kembali *medoid* pada masing-masing *cluster* sebagai pusat *cluster* baru.
6. Ulangi langkah ke-3 sampai ke-5 hingga tidak terjadi perubahan pada *medoid*, sehingga didapatkan *cluster* dan anggota *cluster*-nya masing-masing.

3 Metode Penelitian

3.1 Sumber Data

Pada penelitian ini, data citra rontgen dada yang akan dikelompokkan sebanyak 100 citra yang terdiri dari 25 citra Covid-19, 25 citra Lung Opacity, 25 citra Normal Lung, dan 25 citra Viral Pneumonia. Data tersebut diambil dari COVID-19 Radiography Database diambil dari sumber data pada <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>.

3.2 Langkah-langkah Penelitian



Gambar 2: Diagram Langkah Penelitian

3.2.1 Dataset

Dataset yang digunakan adalah data sekunder yang diambil dari Covid-19 Radiography Database. Data diambil seperlunya yaitu 25 citra Covid-19, 25 citra Lung Opacity, 25 citra Normal, dan 25 citra Viral Pneumonia. Data citra rontgen dada akan diolah dan diproses satu persatu.

3.2.2 Pra-Proses

Pada tahapan ini citra akan diproses melalui dua tahap:

1. Citra keabuan (*Greyscale*)

Kebanyakan dari gambar citra rontgen sudah berwarna keabuan. Namun bisa saja ada citra rontgen yang bukan berwarna keabuan. Maka dari itu, perlu adanya proses greyscale agar citra yang akan diolah memiliki format yang sama. Proses *greyscale* dilakukan dengan mengambil setiap piksel dari warna Red, Green, dan Blue menggunakan persamaan dibawah ini [3]:

$$greyscale = \frac{R + G + B}{3} \quad (8)$$

2. Histogram

Proses histogram menunjukkan distribusi intensitas piksel suatu citra. Selain itu proses histogram berfungsi untuk mengurangi noise dari proses sebelumnya yaitu proses greyscale agar kualitas citra lebih baik. Noise tersebut berupa seperti citra yang kabur, terlalu gelap, terlalu terang, maupun kurang tajam. Proses histogram dilakukan dengan menentukan variabel histogram untuk menampung hasil histogram yang dibuat, selanjutnya derajat keabuan dari tiap piksel akan diuji dan dikelompokkan. Kemudian jumlah piksel yang termasuk dalam tingkat keabuan tersebut dihitung [3].

3.2.3 Segmentasi

Pada tahap ini proses segmentasi menggunakan metode *threshold*. Proses segmentasi ini bertujuan untuk memisahkan objek dengan background agar citra lebih mudah dianalisa. Dengan nilai *threshold* nilai intensitas citra akan diubah menjadi bentuk biner. Pada perhitungan kali ini menggunakan nilai $T = 33$.

3.2.4 Deteksi Tepi *Canny*

Pada proses ini deteksi tepi yang digunakan adalah deteksi tepi *canny*. Dengan menggunakan deteksi tepi, bercak yang ada pada paru-paru akan terdeteksi. Tepi gambar dapat dianggap sebagai tempat piksel dengan perbedaan yang berlebihan dalam nilai intensitas citra. Deteksi tepi dilakukan dengan menggunakan suatu kernel matriks untuk proses konvolusi matriks. Langkah – langkah metode deteksi tepi *canny* [4]: 1) *Smoothing* atau *noise filtering*; 2) Menghitung magnitudo gradient dan arah gradient; 3) Menerapkan *non-maximum suppression*; 4) Menetapkan proses *Hysteresis Thresholding*

3.2.5 Fraktal *Box Counting*

Dengan menggunakan metode fraktal *box counting* akan dihitung nilai dimensi fraktal untuk digunakan dalam proses pengklasteran. Langkah – langkah memperoleh nilai dimensi fraktal box counting: 1) Menutupi gambar citra dengan kotak-kotak persegi dengan rasio yang sama yang ditentukan sebelumnya; 2) Kemudian menghitung kotak-kotak yang menutupi objek; 3) Selanjutnya menghitung nilai dimensi menggunakan persamaan (5).

3.2.6 *Clustering*

Setelah menghitung nilai dimensi dari proses sebelumnya, selanjutnya citra akan diklasifikasikan berdasarkan nilai dimensi fraktal dengan menggunakan metode *clustering K-Medoids*. Pada tahap ini akan dikelompokkan sebanyak 100 objek menjadi empat *cluster* yaitu Covid-19, Lung Opacity, Normal Lung, dan Viral Pneumonia.

3.2.7 Akurasi

Pada tahap terakhir akan dihitung berapa akurasi yang diperoleh dari hasil klasifikasi sebelumnya dengan cara berikut ini [4]:

$$Akurasi = \frac{n}{S} \times 100\% \tag{9}$$

dimana

n = jumlah data yang benar

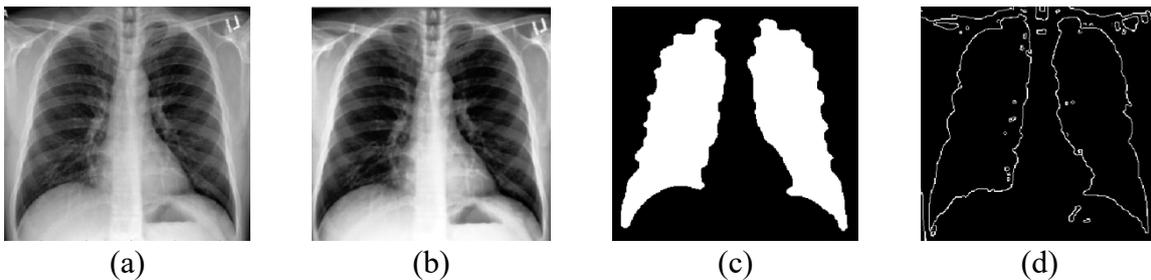
S = jumlah data keseluruhan

4 Hasil dan Pembahasan

4.1 Pengolahan Citra

Pada tahap ini menunjukkan hasil pengolahan masing-masing citra menggunakan aplikasi Matlab. Pertama dilakukan perbaikan citra dengan dilakukan proses histogram. Selanjutnya memisahkan wilayah paru-paru dengan *background* menggunakan metode segmentasi *thresholding*. Lalu mendeteksi tepi dengan metode deteksi tepi *canny*. Berikut contoh hasil pengolahan citra untuk masing - masing penyakit paru.

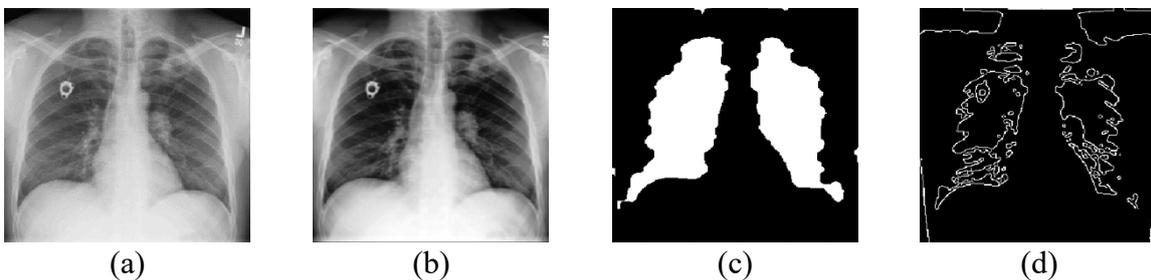
Covid-19



Gambar 3: Contoh Pengolahan Citra Covid-19

Gambar (a) merupakan citra rontgen asli Covid-19; gambar (b) menunjukkan hasil proses histogram terlihat gambar lebih tajam; gambar (c) menunjukkan hasil proses segmentasi menggunakan threshold yang menghasilkan citra biner; dan gambar (d) menunjukkan hasil dari proses deteksi tepi *Canny*.

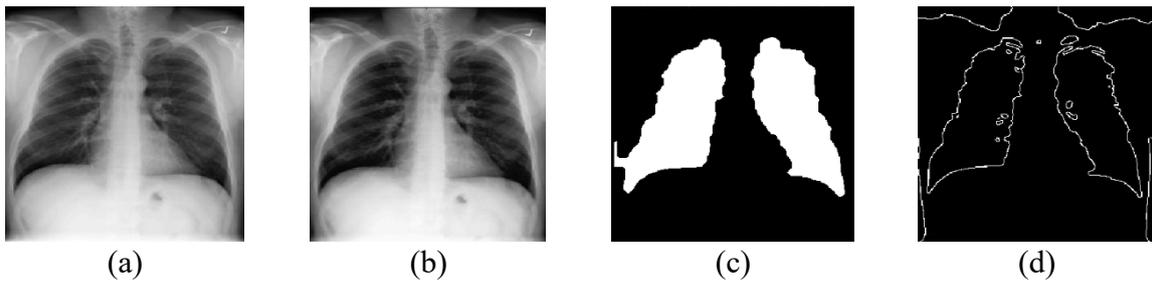
Lung Opacity



Gambar 4: Contoh Pengolahan Citra Lung Opacity

Gambar (a) merupakan citra rontgen asli Lung Opacity; gambar (b) menunjukkan hasil proses histogram terlihat gambar lebih tajam; gambar (c) menunjukkan hasil proses segmentasi menggunakan threshold yang menghasilkan citra biner; dan gambar (d) menunjukkan hasil dari proses deteksi tepi *Canny*.

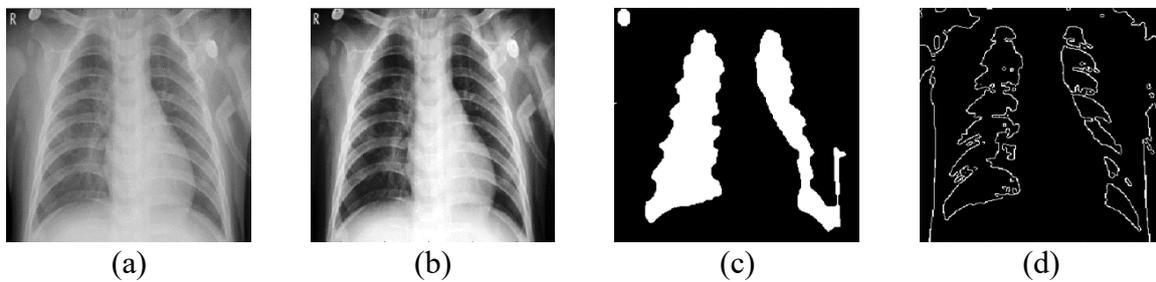
Normal Lung



Gambar 5: Contoh Pengolahan Citra Normal Lung

Gambar (a) merupakan citra rontgen asli Normal Lung; gambar (b) menunjukkan hasil proses histogram terlihat gambar lebih tajam; gambar (c) menunjukkan hasil proses segmentasi menggunakan threshold yang menghasilkan citra biner; dan gambar (d) menunjukkan hasil dari proses deteksi tepi *Canny*.

Viral Pneumonia



Gambar 6: Contoh Pengolahan Citra Viral Pneumonia

Gambar (a) merupakan citra rontgen asli Viral Pneumonia; gambar (b) menunjukkan hasil proses histogram terlihat gambar lebih tajam; gambar (c) menunjukkan hasil proses segmentasi menggunakan threshold yang menghasilkan citra biner; dan gambar (d) menunjukkan hasil dari proses deteksi tepi *Canny*.

4.2 Dimensi Fraktal *Box Counting*

Berikut merupakan nilai dimensi dari masing - masing citra rontgen dimana data 1-25 merupakan citra Covid-19, data 26-50 merupakan citra Lung Opacity, data 51-75 merupakan citra Normal, dan data 76-100 merupakan citra Viral Pneumonia, dengan metode dimensi fraktal *box counting* yang diperoleh menggunakan aplikasi Matlab, dimana nilai yang dihitung dengan persamaan (5):

Table 1: Nilai Dimensi Masing-masing Citra

No.	Dimensi	No.	Dimensi	No.	Dimensi	No.	Dimensi
1	1.3162	26	1.3932	51	1.2238	76	1.3422
2	1.3155	27	1.3659	52	1.2022	77	1.3536
3	1.2330	28	1.3891	53	1.2360	78	1.3309
4	1.3222	29	1.3653	54	1.2644	79	1.3324
5	1.2886	30	1.3741	55	1.2635	80	1.3602
6	1.2810	31	1.3589	56	1.2248	81	1.3796
7	1.3047	32	1.3877	57	1.2269	82	1.3550
8	1.2911	33	1.3749	58	1.2552	83	1.3408
9	1.2943	34	1.3571	59	1.2598	84	1.3401

10	1.2927	35	1.3802	60	1.2580	85	1.3516
11	1.3519	36	1.3699	61	1.2164	86	1.3463
12	1.3302	37	1.3415	62	1.2380	87	1.3628
13	1.3237	38	1.3850	63	1.2524	88	1.3449
14	1.3244	39	1.3861	64	1.2448	89	1.3569
15	1.3039	40	1.3649	65	1.2515	90	1.3359
16	1.3229	41	1.3814	66	1.2653	91	1.3435
17	1.2869	42	1.3822	67	1.2689	92	1.3563
18	1.2802	43	1.3966	68	1.2248	93	1.3463
19	1.2810	44	1.3796	69	1.2448	94	1.3435
20	1.2810	45	1.3949	70	1.2448	95	1.3408
21	1.2810	46	1.3621	71	1.2486	96	1.3456
22	1.2810	47	1.3867	72	1.2380	97	1.3543
23	1.2810	48	1.3716	73	1.2207	98	1.3483
24	1.2878	49	1.3966	74	1.2626	99	1.3608
25	1.2935	50	1.3710	75	1.2671	100	1.3496

Rata – rata nilai dimensi dari masing – masing penyakit paru:

Table 2: Rata-Rata Nilai Dimensi Masing-Masing Penyakit

Jenis Penyakit Paru	Rata-rata Dimensi
Covid	1.2980
Lung Opacity	1.3767
Normal	1.2441
Viral Pneumonia	1.3489

4.3 Clustering K-Medoids

Pengelompokkan dilakukan dengan metode *K-Medoids* dengan menggunakan aplikasi RapidMiner Studio. Diperoleh hasil sebagai berikut:

Table 3: Cluster Centroid

Centroid table	
cluster_0	1.294
cluster_1	1.380
cluster_2	1.350
cluster_3	1.267

Dengan *cluster* centroid diatas, data yang diproses akan dikelompokkan menjadi 4, yaitu kelompok covid-19, kelompok lung opacity, kelompok normal lung, dan kelompok viral pneumonia. Hasil pengelompokkan diperoleh sebagai berikut:

Table 4: Hasil Pengelompokkan Menggunakan Metode *K-Medoids*

No.	Cluster	No.	Cluster	No.	Cluster	No.	Cluster
1	cluster_0	26	cluster_1	51	cluster_3	76	cluster_2
2	cluster_0	27	cluster_1	52	cluster_3	77	cluster_2
3	cluster_3	28	cluster_1	53	cluster_3	78	cluster_2
4	cluster_2	29	cluster_1	54	cluster_3	79	cluster_2

5	cluster_0	30	cluster_1	55	cluster_3	80	cluster_2
6	cluster_0	31	cluster_2	56	cluster_3	81	cluster_1
7	cluster_0	32	cluster_1	57	cluster_3	82	cluster_2
8	cluster_0	33	cluster_1	58	cluster_3	83	cluster_2
9	cluster_0	34	cluster_2	59	cluster_3	84	cluster_2
10	cluster_0	35	cluster_1	60	cluster_3	85	cluster_2
11	cluster_2	36	cluster_1	61	cluster_3	86	cluster_2
12	cluster_2	37	cluster_2	62	cluster_3	87	cluster_2
13	cluster_2	38	cluster_1	63	cluster_3	88	cluster_2
14	cluster_2	39	cluster_1	64	cluster_3	89	cluster_2
15	cluster_0	40	cluster_1	65	cluster_3	90	cluster_2
16	cluster_2	41	cluster_1	66	cluster_3	91	cluster_2
17	cluster_0	42	cluster_1	67	cluster_3	92	cluster_2
18	cluster_3	43	cluster_1	68	cluster_3	93	cluster_2
19	cluster_0	44	cluster_1	69	cluster_3	94	cluster_2
20	cluster_0	45	cluster_1	70	cluster_3	95	cluster_2
21	cluster_0	46	cluster_2	71	cluster_3	96	cluster_2
22	cluster_0	47	cluster_1	72	cluster_3	97	cluster_2
23	cluster_0	48	cluster_1	73	cluster_3	98	cluster_2
24	cluster_0	49	cluster_1	74	cluster_3	99	cluster_2
25	cluster_0	50	cluster_1	75	cluster_3	100	cluster_2

Dari Tabel 7 dapat dilihat bahwa cluster_0 merupakan kelompok Covid-19, cluster_1 merupakan kelompok Lung Opacity, cluster_3 merupakan kelompok Normal Lung, dan cluster_2 merupakan kelompok Viral Pneumonia. Terdapat 13 dari 100 data yang diproses berada pada cluster yang salah atau tidak sesuai dengan dataset yang diambil dari COVID-19 Radiography Database.

4.4 Akurasi

Selanjutnya akan dihitung akurasi dari proses clustering:

$$Akurasi = \frac{87}{100} \times 100\% = 87\%$$

Dari perhitungan tersebut diperoleh akurasi sebesar 87%.

5 Kesimpulan

Berdasarkan hasil penelitian, menunjukkan bahwa *rontgen* dada yang diolah melalui proses *greyscale*, *histogram equalization*, segmentasi *thresholding*, dan deteksi tepi *canny* untuk mengenali ciri objek. Kemudian dihitung nilai dimensi dengan menggunakan metode dimensi fraktal *box counting*. Nilai dimensi digunakan untuk proses *clustering k-medoids* menjadi 4 *cluster* menghasilkan akurasi sebesar 87%. Hal tersebut menunjukkan bahwa *clustering* dengan metode *k-medoids* cukup baik diterapkan untuk data tersebut karena tidak ada perbedaan yang signifikan dengan dataset yang diambil dari COVID-19 Radiography Database.

Daftar Pustaka

- [1] A. Saputra, “Sistem Pakar Identifikasi Penyakit Paru-Paru Pada Manusia Menggunakan Pemrograman Visual Basic 6.0,” *Jurnal Teknologi Dan Informatika (Teknomatika)*, vol. 1, no. 3, pp. 202–222, 2011.
- [2] A. Mardiyah and A. Harjoko, “Metode Segmentasi Paru-Paru dan Jantung pada Citra X-ray Thorax,” *IJEIS*, vol. 1, no. 2, pp. 35–44, 2011.
- [3] R. Rahmadewi and R. Kurnia, “Klasifikasi Penyakit Paru Berdasarkan Citra Rontgen Dengan Metoda Segmentasi Sobel,” *Jurnal Nasional Teknik Elektro*, vol. 5, no. 1, pp. 7–12, 2016.
- [4] N. A. Isnaini and D. Juniati, “Klasifikasi Jenis Tumor Kulit Menggunakan Dimensi Fraktal Box Counting dan K-means,” *Jurnal Riset dan Aplikasi Matematika*, vol. 3, no. 2, pp. 71–77, 2019.
- [5] D. W. Safitri and D. Juniati, “Classification of diabetic retinopathy using fractal dimension analysis of eye fundus image,” *AIP Conference Proceedings*, vol. 1867 020011, 2017.
- [6] D. Juniati, C. Khotimah Sumarno, D. Dwi Wardani, and I. K. Budayasa, “Fractal dimension to classify the heart sound recordings with KNN and fuzzy c-mean clustering methods,” *Journal of Physics Conference Series*, vol. 953 012202, 2018.
- [7] M. B. Sapata and D. Juniati, “Klasifikasi Penyakit Paru Berdasarkan Citra X-Ray Thorax Menggunakan Metode Fraktal Box Counting,” *Mathunesa: Jurnal Ilmiah Matematika*, vol. 7, no. 3, pp. 228–235, 2019.
- [8] D. Widhyanti and D. Juniati, “Classification of Baby Cry Sound Using Higuchi’s Fractal Dimension with K-Nearest Neighbor and Support Vector Machine,” *Journal of Physics Conference Series*, vol. 1747 012014, 2021.
- [9] N. F. Azizah and D. Juniati, “Analisis Jenis Penyakit Paru-Paru Berdasarkan Chest X-Ray Menggunakan Metode C-Means,” *Mathunesa: Jurnal Ilmiah Matematika*, vol. 9, no. 2, pp. 322–332, Jul. 2021.
- [10] C. Gabriel, S. Gabriel, and E. Corthout, “The Dielectric Properties of Biological Tissues,” *Physics in Medicine & Biology*, vol. 41, no. 11, 1996.
- [11] R. Munir, *Pengolahan Citra Digital dengan Pendekatan Algoritmik*. Bandung: Informatika, 2004.
- [12] S. S. Al-amri, N. V. Kalyankar, and K. S.D, “Image Segmentation by Using Edge Detection,” *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 2, no. 3, pp. 804–807, 2010.
- [13] W. Abdurahman, “Object Traking Menggunakan Metode Deteksi Tepi Canny Berbasis Python Opencv,” 2019.
- [14] D. Juniati and I. K. Budayasa, *Geometri Fractal dan Aplikasinya*. Surabaya: Unesa University Press, 2016.
- [15] N. K. Kaur, U. Kaur, and D. Singh, “K-Medoid Clustering Algorithm-A Review,” *International Journal of Computer Application and Technology*, vol. 1, no. 1, pp. 42–45, 2014.
- [16] D. F. Pramesti, M. T. Furqon, and C. Dewi, “Implementasi Metode K-Medoids Clustering untuk Pengelompokkan Data Potensi Kebakaran Hutan/Lahan Berdasarkan Titik Panas (Hotspot),” *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 1, no. 9, pp. 723–732, 2017.