

`KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA X-RAY THORAX MENGGUNAKAN
METODE FRAKTAL BOX COUNTING**Beny My Sapata**Jurusan Matematika, FMIPA, Universitas Negeri Surabaya
e-mail : benysapata@mhs.unesa.ac.id**Dwi Juniati**Jurusan Matematika, FMIPA, Universitas Negeri Surabaya
e-mail : dwi_juniati@yahoo.com**Abstrak**

Paru-paru merupakan organ penting dan vital yang mudah terjadi infeksi. Infeksi pada paru dapat diperiksa melalui pemeriksaan pada citra *x-ray*, tetapi diagnosis klinis dari hasil citra *x-ray* itu sulit. Dalam mempermudah menganalisa diagnosa dari hasil citra *x-ray* tersebut, maka pada penelitian ini dilakukan proses pengklasifikasian citra *x-ray* berdasarkan jenis penyakit. Citra yang diolah ada 120 (seratus dua puluh) citra *x-ray thorax* dengan melakukan segmentasi terhadap region paru. Kemudian dari region paru ini dilakukan deteksi tepi canny untuk mengambil bercak dari penyakit paru. Dari deteksi tepi canny akan dihitung nilai dimensi fraktal dengan menggunakan metode *box counting*, sehingga dapat dilakukan pengklasifikasian. Hasil dari eksperimen dengan menggunakan metode *k-nearest neighbor* (K-NN), dimana hasil dengan akurasi terbesar ditunjukkan pada nilai $K = 5$ yaitu 79.65% dan akurasi terendah pada nilai $K = 7$ yaitu 71.28%.

Kata kunci: *X-Ray Thorax, Box Counting, K-Nearest Neighbor* (K-NN)**Abstract**

The lungs are important and vital organs that are easily infected. Lung examination can be done through examination of x-ray images, but the clinical diagnosis of the results of x-ray images is difficult. In making it easier to analyze the diagnosis of the results of the x-ray image, this research carried out the process of classifying x-ray images based on the type of disease.. The image that was processed was 120 (one hundred and twenty) chest x-ray images by segmenting the lung area. From this lung area, then canny edge detection is done to take spots from lung disease. From smart fractal edge detection values are obtained, which will be calculated using the box calculation method, so classification can be done. The results of the experiment using the nearest k-neighbor method (K-NN), where the results with the greatest accuracy are shown by the value $K = 5$ which is 79.65% and the lowest accuracy at the value $K = 7$ is 71.28%.

Keywords: *X-Ray Thorax, Box Counting, K-Nearest Neighbor* (K-NN)**1. PENDAHULUAN**

Program pemanfaatan sinar-x dalam bidang medis merupakan salah satu cara untuk meningkatkan kesehatan masyarakat. Pemeriksaan dengan pengambilan citra *x-ray* pada tubuh manusia dilakukan tanpa perlu melakukan operasi bedah. Bahkan pengambilan citra *x-ray* merupakan salah satu pemeriksaan pencitraan medis yang paling sering dan hemat biaya (Wang, Peng, Lu, Lu, Bagher, & Summers, 2017). Mesin sinar-x (foto radiologi konvensional) memiliki prinsip untuk menembuskan sebuah gelombang elektromagnetik ke dalam tubuh sehingga menghasilkan gambar yang berupa citra tubuh yaitu citra *x-ray* pada sebuah pelat film.

Paru-paru merupakan organ penting dan vital yang dapat diperiksa melalui pemeriksaan pada citra *x-ray*. Fungsi paru-paru sebagai tempat respirasi yaitu menukar oksigen (O_2) dari udara dengan karbon dioksida (CO_2) dari darah. Paru-paru terletak di dalam rongga dada (*thorax*) diliputi dua dinding lapisan yang dipisahkan oleh lapisan udara yang disebut sebagai rongga plural (Pearce, 2013). Kelainan pada paru-paru terjadi karena paru-paru terpengaruh oleh faktor-faktor tertentu seperti merokok, infeksi, dan terpapar zat berbahaya yang akan membuat kondisi paru-paru menjadi abnormal. Contoh penyakit pada paru-paru adalah *effusion, pneumonia, pneumothorax*, *tuberculosis* dan lain-lain.

Cara mendeteksi penyakit paru salah satunya yaitu dengan melakukan pemeriksaan *x-ray* pada bagian dada

KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA X-RAY THORAX MENGGUNAKAN METODE FRAKTAL BOX COUNTING

yang menghasilkan citra *x-ray thorax*. Teknik pemeriksaan *x-ray thorax* dilakukan beberapa macam cara yaitu proyeksi PA (*postero-anterior*) dimana pemeriksaan sinar-x diarahkan dari belakang pasien dan Proyeksi LAT (*lateral*) dimana pemeriksaan sinar-x diarahkan dari samping kanan atau kiri pasien. Dalam diagnosis klinis *x-ray* dada bisa sulit, dan kadang-kadang diyakini lebih sulit dari pada diagnosis melalui pencitraan CT dada (Wang, Peng, Lu, Lu, Bagher, & Summers, 2017), sehingga masyarakat pada umumnya sulit untuk memahami hasil *x-ray* maka dibutuhkan bantuan tenaga ahli seperti dokter atau tenaga medis lain agar bisa membacanya. (Joseph & Singh, 2014).

Dengan permasalahan tersebut, peneliti mengolah dan menganalisis hasil citra *x-ray thorax* yang dapat digunakan untuk mengetahui keadaan paru-paru yang tergolong penderita *effusion*, *pnemounia*, atau *no finding*. Data diambil dari dataset NIH Chest X-ray Dataset of 14 Common Thorax Disease Categories. Dalam database tersebut, disediakan 15 kategori penyakit beserta hasil gambar. NIH Chest X-ray Dataset diekstraksi dari database PACS klinis di National Institutes of Health Clinical Center dan terdiri dari 60% dari semua *x-ray* dada di rumah sakit. Oleh karena itu diharapkan dataset tersebut secara signifikan sudah mewakili distribusi populasi pasien dan memberikan diagnosis klinis secara umum (Wang, Peng, Lu, Lu, Bagher, & Summers, 2017). Dari dataset NIH Chest X-ray Dataset tersebut akan diambil sampel data sebanyak 40 subjek diagnosis berupa gambar citra *x-rays* pada dada yang ditampilkan dengan format greyscale untuk tujuan penelitian ini. Citra *x-rays* tersebut digunakan untuk mengklasifikasikan keadaan *thorax* yang didiagnosis terdapat penyakit atau tidak pada paru-paru.

Metode yang digunakan pada proposal ini yaitu metode fraktal *box counting*. Metode *box counting* baik untuk menentukan dimensi citra *x-rays thorax* karena metode *box counting* biasa digunakan untuk menghitung dimensi pada citra gambar. Perhitungan dimensi suatu citra dengan menggunakan metode *box counting*, dilakukan dengan menutupi area obyek dengan kotak-kotak persegi (*box*) dengan ukuran yang bervariasi (Juniati & Budayasa, 2006).

Dari beberapa latar belakang diatas, pada proposal ini mengusulkan penelitian untuk : Klasifikasi Citra X-Ray Thorax Pada Penderita Paru-Paru Menggunakan Metode Fraktal Box Counting. Dataset yang digunakan yaitu NIH Chest Dataset dengan mengambil 40 subjek tiap kelas (Wang, Peng, Lu, Lu, Bagher, & Summers, 2017). Dalam penelitian ini, metode yang digunakan dalam mengolah *image detection* adalah metode Fraktal *Box Counting*. Metode Fraktal *Box Counting* merupakan salah satu metode fraktal dalam pengolahan *image detection*.

2. KAJIAN TEORI

Citra Medis

Pencitraan medis (*medical image*) adalah sebuah proses dan teknik untuk menggambarkan kondisi tubuh pada sebuah citra untuk tujuan klinis, yang merupakan aturan medis untuk mendiagnosis keadaan fisiologi dan

anatomi tubuh (memeriksa atau mendiagnosis penyakit). Pencitraan medis merupakan teknik perekaman dan proses pengukuran sehingga menghasilkan gambar dan data yang dapat diproyeksikan sebagai peta (yaitu berisi informasi posisi), seperti *ultrasonografi* (USG), *computerized tomography scan* (CT Scan), *magnetoencephalography* (MEG), dan lain sebagainya (Setiawidayat & Putri, 2009).

Citra X-Ray

X-ray merupakan suatu bentuk radiasi berbentuk elektromagnetik. *X-ray* mampu menembus berbagai objek termasuk tubuh manusia. Mesin *x-ray* memancarkan radiasi kecil yang dapat menembus tubuh, sehingga dapat merekam citra digital pada plat khusus atau film. Organ tubuh yang dilewati sinar *x-ray* akan menerima radiasi dengan dosis yang berbeda. Organ tubuh yang padat akan menerima dosis lebih besar dibandingkan organ lain yang lebih lunak. Sebagai hasilnya, organ yang padat akan berwarna putih pada *x-ray*, organ yang lunak berwarna abu-abu dan udara yang berwarna hitam. Pada *x-ray thorax*, tulang atau jaringan organ selain paru-paru akan menerima radiasi yang besar sehingga terlihat abu-abu atau putih atau terang pada citra sedangkan paru-paru yang sebagian besar berisi udara akan menerima radiasi lebih kecil sehingga terlihat lebih gelap pada citra.

Manfaat *x-ray* pada bidang medis, yaitu *x-ray* dapat digunakan untuk melihat kondisi organ tubuh manusia seperti tulang, paru-paru, jantung, dan organ lain tanpa perlu melakukan pembedahan langsung pada tubuh pasien. *X-ray* juga berbahaya bagi tubuh karena memiliki efek samping jika dilakukan secara berlebihan. Bangkan dapat sampai menimbulkan kanker. Oleh sebab itu, para ahli medis tidak menganjurkan terlalu sering mengambil citra *x-ray* pada tubuh secara berlebihan (C. Gabriel, 1996).

Paru-Paru

Paru-paru merupakan salah satu organ yang sangat penting dan termasuk organ vital pada tubuh manusia. Fungsi paru-paru sebagai tempat respirasi yaitu menukar oksigen (O_2) dari udara dengan karbon dioksida (CO_2) dari darah. Letak paru-paru berada pada rongga dada (*thorax*), yang dilindungi oleh tulang rusuk dan dikelilingi oleh dua lapis dinding *pleura*. Kedua lapis dinding tersebut dipisahkan dengan lapisan udara yang dikenal sebagai rongga *pluera* yang berisi cairan *pluera* (Pearce, 2013). Kelainan pada paru-paru terjadi karena paru-paru terpengaruh oleh faktor-faktor tertentu seperti merokok, infeksi, dan terpapar zat berbahaya yang akan membuat kondisi paru-paru menjadi abnormal. Contoh penyakit pada paru-paru adalah *effusion*, *pneumonia*, *pneumothorax*, *tuberculosis* dan lain-lain.

Pneumonia

Pneumonia adalah infeksi jaringan paru-paru (*alveoli*) yang bersifat akut. *Pneumonia* terjadi karena terjadi radang pada parenkim paru dimana asinus berisi cairan dan sel radang, dengan atau tanpa disertai infiltrasi sel radang ke dalam dinding alveoli dan rongga

intestinum. Sebagian besar *pneumonia* disebabkan oleh virus, bakteri, jamur, paparan bahan kimia atau kerusakan fisik dari paru-paru. Disisi lain juga bisa karena pengaruh dari penyakit lain. Virus yang menyebabkan pneumonia adalah *influenza virus*, *respiratory syncytial virus* (RSV), *rhinovirus*, *adenoviruses*, dan lain sebagainya, sedangkan bakteri yang menyebabkan *pneumonia* adalah *mycoplasma pneumoniae* dan *streptococcus*.

Effusion

Efusi pleura (*effusion*) terjadi karena penumpukan cairan di antara lapisan *pleura*. Hal ini disebabkan karena di dalam membran *pleura* memproduksi cairan yang berlebihan atau penyerapan yang berkurang (Tobing & Widirahardjo, 2009). Cairan pada lapisan pleura ini sebenarnya berfungsi untuk pelumas agar dinding paru-paru dan dinding dada bagian dalam tidak terjadi gesekan, namun akan menjadi masalah jika cairan pada pleura menumpuk dan bertambah menjadi banyak.

No Finding

No finding dapat diartikan "Tidak Ada Temuan" dalam menganalisa jenis penyakit dalam citra *x-rays thorax*. Maksudnya dalam mengidentifikasi jenis penyakit dalam citra *x-rays* tersebut, citra teridentifikasi tanpa temuan patologi (Wang, Peng, Lu, Lu, Bagher, & Summers, 2017).

Histogram

Histogram *equalization* adalah suatu proses pemerataan piksel dari citra agar citra lebih baik. Histogram digunakan untuk peningkatan kontras gambar skala abu-abu dengan cara mendistribusi skala keabuan suatu citra yang dibuat merata. Teknik Histogram terkenal dalam literatur pengolahan gambar. Tujuan digunakan histogram agar citra dapat dilihat nilai *gray level* sehingga citra dapat diperbaiki kualitasnya dengan cara pemerataan histogram. Pemerataan histogram dapat dicapai dengan menggunakan Fungsi Distribusi Kumulatif (CDF) yang merupakan perhitungan kumulatif dari sebuah histogram, dan versi diskritnya adalah sebagai berikut:

$$f(k) = \frac{(N-1)}{M} \sum_{k=0}^n h(k); \quad n = 1, 2, 3, \dots, N - 1$$

$f(k)$ merupakan fungsi CDF dari pemerataan histogram, M menyatakan ukuran citra (piksel) $m \times n$ sedangkan N menyatakan *grayscale* atau jumlah level keabuan biasanya bernilai 256, dan $h(k)$ menyatakan besar nilai histogram pada suatu nilai k . Persamaan tersebut secara seragam mendistribusikan populasi piksel suatu gambar ke semua tingkat abu-abu yang tersedia dari perangkat visualisasi, dan dengan demikian memaksimalkan kontras. Ekualisasi histogram, awalnya dirancang untuk gambar skala abu-abu diperpanjang untuk meningkatkan gambar warna. Untuk menghindari masalah terkait perubahan warna, ide-ide lain secara merata menyebarkan histogram di bawah kendala yang menjaga rona setiap piksel tidak berubah. Dengan kata lain, proses pemerataan hanya diterapkan pada komponen pencahayaan dan saturasi gambar (Duan & Qiu, 2004).

Segmentasi

Segmentasi adalah proses pemisahan wilayah pada citra, dimana wilayah tersebut memiliki homogenitas antara warna, tekstur, dan intensitasnya. Segmentasi bertujuan untuk memisahkan antara objek dan latar belakang sehingga objek lebih mudah dikenali dan dianalisa.

Thresholds adalah salah satu metode yang banyak digunakan untuk segmentasi gambar. Ini berguna dalam membedakan latar depan dari latar belakang. Gambar tingkat abu-abu dapat dikonversi menjadi gambar biner bila memilih nilai *thresholds* (T) yang memadai. Cara paling umum untuk mengkonversi gambar tingkat abu-abu ke gambar biner adalah memilih satu nilai *thresholds* (T). Maka semua nilai tingkat abu-abu di bawah T ini akan diklasifikasikan sebagai warna hitam (0), dan yang di atas T akan diklasifikasikan sebagai warna putih (1). Menurut Al-amri, Kalyankar dan Khamitkar (2010) teknik *thresholds* ini dapat dinyatakan sebagai:

$$T = T[x, y, p(x, y), f(x, y)]$$

Dimana:

T adalah nilai ambang.

x, y adalah koordinat dari titik nilai ambang batas.

$p(x, y), f(x, y)$ adalah titik piksel gambar tingkat abu-abu.

Gambar ambang batas $g(x, y)$ dapat didefinisikan:

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T \end{cases}$$

Deteksi Tepi

Tepi menandai batasan dan memiliki perubahan nilai intensitas derajat keabuan yang besar (Narayana, Mohan, & Subhramanya, 2014). Tujuan deteksi tepi biasanya untuk mengelompokkan objek-objek pada sebuah citra, dan juga dapat digunakan untuk menganalisa citra lebih lanjut. Ada banyak macam metode dalam deteksi tepi, namun dalam penelitian ini digunakan deteksi tepi Canny (*Canny Edge detection*).

Fraktal Box Counting

Dimensi menurut Fraktal berbeda dengan dimensi menurut Euclid. Pada dimensi fraktal mengenal dimensi yang berupa pecahan, dimana dimensi antara 2 sampai 3 pada benda yang tidak solid dan dimensi 1 sampai 2 umumnya pada sebuah citra gambar.

Dalam penelitian ini digunakan metode *box counting* untuk menentukan dimensi citra *x-ray thorax* karena metode *box counting* adalah metode yang cocok untuk menghitung dimensi pada citra gambar. Perhitungan dimensi suatu citra dengan menggunakan metode *box counting*, dilakukan dengan menutupi area obyek dengan kotak-kotak persegi (*box*) dengan ukuran yang bervariasi (Juniati & Budayasa, 2006).

Persamaan pada dimensi dapat dicari dengan persamaan berikut:

$$D(r) = \frac{\log N(r)}{\log(\frac{1}{r})}$$

$D(r)$ merupakan besar dimensi, $N(r)$ merupakan jumlah kotak yang terisi objek, dan r adalah iterasi. Langkah-langkah pada penerapan perhitungan *box counting* sebagai berikut:

KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA X-RAY THORAX MENGGUNAKAN METODE FRAKTAL BOX COUNTING

1. Mengambil citra objek fraktal. Dari citra yang dihasilkan tersebut, kemudian dibagi menjadi r bagian dengan berbagai variasi untuk setiap iterasi. Nilai r berubah dari 1 menjadi 2^n , dengan $n = 1, 2, 3, \dots$. Jika objek berukuran $2m \times 2m$ maka nilai dari 2^n tidak lebih dari $2m$.
2. Menghitung jumlah kotak yang terisi objek ($N(r)$).
3. Menghitung nilai $\log(N(r))$ dan $\log(1/r)$.
4. Membuat garis lurus dari nilai-nilai $\log(N(r))$ dan $\log(1/r)$.
5. Menghitung kemiringan garis lurus dengan persamaan sebagai berikut :

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

Dimana $x = \log(\frac{1}{r})$, $y = \log(N(r))$, $n =$ banyak data iterasi yang digunakan, dan α adalah kemiringan (slope) yang dapat digunakan untuk menghitung nilai dimensi fraktal.

K-Nearest Neighbor (KNN)

Algoritma *k-nearest neighbor* (KNN) adalah sebuah metode yang digunakan untuk melakukan klasifikasi terhadap objek dengan didasarkan pada jumlah jarak yang paling dekat dengan objek yang dituju. Umumnya didasarkan pada jarak Euclidean terdekat antara sampel uji dan sampel pelatihan yang ditentukan (Peterson, 2009).

Tujuan dari algoritma *k-nearest neighbor* adalah mengklasifikasikan obyek baru berdasarkan sampel uji dan sampel pelatihan. Algoritma metode *k-nearest neighbor* (KNN) bekerja berdasarkan jarak terpendek dari sampel uji ke sampel pelatihan. Umumnya sampel uji dan sampel pelatihan ditentukan dengan jarak Euclidean dan dipilih jarak yang terdekat diantara keduanya (Peterson, 2009). Jarak Euclidean dapat didefinisikan sebagai berikut :

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

Dimensi D merupakan jarak titik pada data pelatihan x dan titik pada data pelatihan y yang akan diklasifikasikan, sedangkan x_i dan y_i adalah besaran skalar untuk dimensi ke- i dalam ruang vektor n dimensi.

Berikut merupakan langkah-langkah menghitung algoritma *k-nearest neighbor* (KNN) :

- a. Menentukan parameter nilai K sebagai banyak jumlah tetangga terdekat dengan objek baru.
- b. Menghitung jarak antar objek baru terhadap semua objek yang telah dilatih.
- c. Urutkan hasil perhitungan tersebut.
- d. Menentukan tetangga terdekat berdasarkan jarak minimum ke K .
- e. Menentukan kategori tetangga terdekat dengan objek baru.
- f. Gunakan kategori mayoritas sebagai klasifikasi objek/data baru.

3. METODE

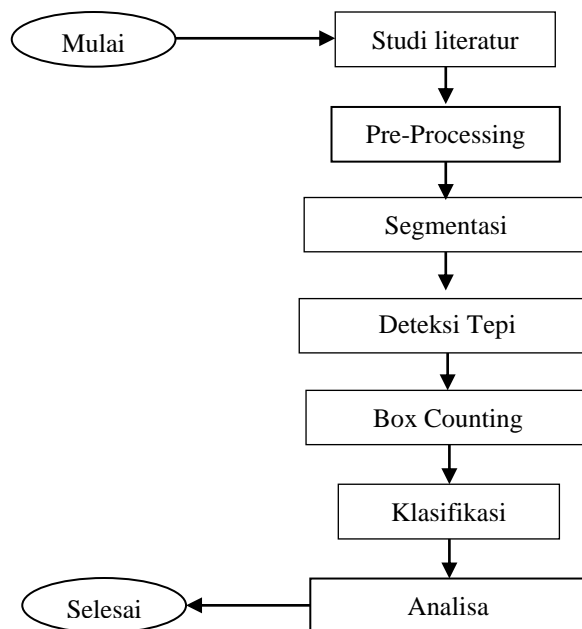
Dataset

Data citra *x-ray* yang digunakan pada penelitian ini baik untuk data *training* dan data *testing* adalah data sekunder yang diambil dari kaggle dataset dengan nama file NIH Chest X-ray Dataset of 14 Common Thorax Disease Categories. Data yang diambil hanya data yang diperlukan untuk penelitian ini yaitu empat puluh data penyakit *effusion*, empat puluh data penyakit *pneumonia*, dan empat puluh data *no finding*.

Dalam database NIH Chest X-ray Dataset of 14 Common Thorax Disease Categories terdapat citra *x-ray thorax* baru, dengan nama "ChestX-ray8", yang terdiri dari 112.120 tampilan depan gambar *x-ray* dari 30.805 pasien dengan membagi menjadi empat belas kelas penyakit dan satu kelas *no finding*, dari *associated radiological reports* menggunakan *natural language processing*. 112.120 gambar *x-ray thorax* tersebut ditampilkan dengan format PNG dalam resolusi $1024 * 1024$.

Proses pengambilan data pada penelitian ini hanya mengambil data citra *x-ray thorax* yang diperlukan yaitu *effusion*, *pneumonia*, dan *no finding* dari lima belas data *x-ray thorax* yang ada pada database NIH Chest X-ray Dataset of 14 Common Thorax Disease Categories. Data yang digunakan untuk penelitian ini berjumlah 140 data citra *x-ray thorax* yang merupakan 40 citra *x-ray* penyakit *effusion*, 40 citra *x-ray* penyakit *pneumonia*, dan 40 citra *x-ray no finding*.

Tahap-Tahap Penelitian



Gambar 1. Diagram blok tahap penelitian

1. Studi literatur

Langkah pertama yang dilakukan dalam penelitian ini adalah studi literatur pada jurnal yang membahas tentang klasifikasi data citra *x-ray*, keadaan *thorax*, dan metode fraktal *box counting*. Data yang didapat dari studi

literatur akan digunakan sebagai acuan untuk membuat latar belakang dan membantu penyusunan penelitian ini.

2. Pre-Processing

Pada fase ini, data *pre-processed* akan di sampling dan diproses dengan menggunakan beberapa metode untuk mempermudah proses komputasi. Metode-metode tersebut adalah sebagai berikut :

a. Greyscale

Citra yang diinputkan berupa citra digital gambar *x-ray thorax* yang format sebenarnya sudah *greyscale*, namun jika ada format dari citra bukan *greyscale* akan dikonversi kedalam citra *greyscale*. Citra *greyscale* bertujuan agar citra yang semula berwarna menjadi citra dalam bentuk keabuan.

b. Histogram

Setelah proses pengolahan data citra *x-ray* yang berupa *greyscale*, biasanya citra masih memiliki masalah seperti noise, citra kurang tajam, terlalu gelap atau terang, kabur, atau masalah-masalah lainnya sehingga diperlukan histogram. Fungsi histogram menunjukkan sebaran intensitas piksel dari suatu citra. Dengan adanya histogram citra akan terlihat kualitasnya, tergolong citra yang terang atau yang gelap sehingga citra yang kurang bagus dapat diperbaiki kualitasnya.

3. Segmentasi

Segmentasi adalah proses pemisahan wilayah pada citra, dimana wilayah tersebut memiliki homogenitas antara warna, tekstur, dan intensitasnya. Segmentasi bertujuan untuk memisahkan antara objek dan latar belakang sehingga objek lebih mudah dikenali dan dianalisa. Dalam penelitian ini memisahkan bentuk paru dengan organ lainnya.

Threshold merupakan teknik yang digunakan untuk mensegmentasi citra dengan analisa kemiripan (similaritas). Dengan menentukan nilai ambang batas T , *thresholding* akan merubah citra *greyscale* kedalam bentuk biner. Nilai batas yang digunakan sebesar 33. Apabila nilai intensitas T berada dibawah atau diatas 33 citra paru tidak terbentuk.

4. Deteksi Tepi

Pada penelitian ini metode deteksi tepi yang digunakan adalah deteksi tepi canny, karena hasil dari deteksi tepi canny terlihat lebih jelas dan optimal. Deteksi tepi canny bertujuan untuk menentukan bercak yang ada pada paru-paru sehingga nantinya dapat dihitung nilai dengan menggunakan metode *box counting*.

5. Box Counting

Dalam penelitian ini digunakan metode *box counting* untuk menentukan besar dimensi suatu citra. Nilai dimensi *box counting* akan menentukan besar nilai pada deteksi tepi canny (bercak penyakit).

6. Klasifikasi

Klasifikasi dilakukan dengan menggunakan nilai dimensi dari canny. Metode klasifikasi yang digunakan pada penelitian ini adalah metode *k-nearest neighbor* (KNN), dikarenakan metode ini memiliki performa yang baik dalam menjalankan tugasnya. Metode KNN bekerja dengan cara kerja berdasarkan jarak terpendek dari sampel uji ke sampel pelatihan dan hasil dikatakan sesuai apabila dari keseluruhan data uji memiliki jarak

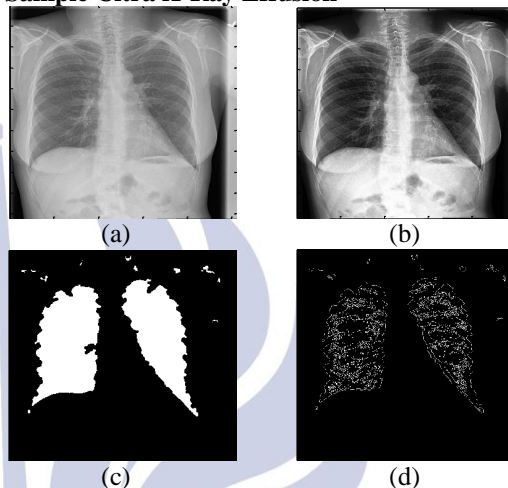
minimum. Nilai K dari metode *k-nearest neighbor* ditentukan dengan nilai ganjil agar hasil tidak berimbang. Biasanya nilai K pada *k-nearest neighbor* yang digunakan adalah 3, 5, 7, dan 9. Nilai K dengan akurasi tertinggi adalah nilai yang dipakai.

7. Analisa

Analisa dilakukan dengan cara melihat jumlah akurasi untuk mengevaluasi hasil penelitian. Setiap data uji dilihat nilai kebenaran identifikasinya.

4. PEMBAHASAN

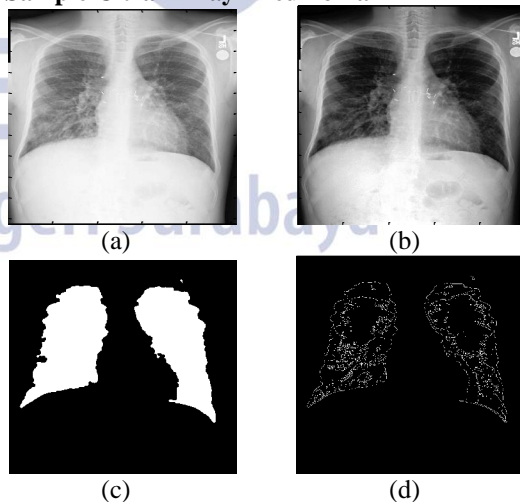
Sample Citra X-Ray Effusion



Gambar 2. Pengolahan Citra Effusion

Proses pada Gambar 2 diatas merupakan pengolahan citra pada citra *x-ray* penderita *effusion*. Gambar (a) merupakan sampel citra *x-ray thorax* yang diinputkan. Gambar (b) merupakan citra yang telah diperbaiki. Gambar (c) merupakan citra *x-ray* hasil segmentasi. Gambar (d) merupakan hasil citra rontgen setelah dilakukan deteksi tepi.

Sample Citra X-Ray Pneumonia



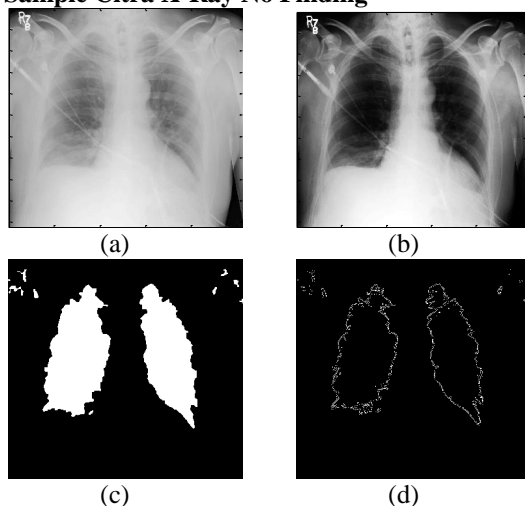
Gambar 3. Pengolahan Citra Effusion

Proses pada Gambar 3 diatas merupakan pengolahan citra pada citra *x-ray* penderita *pneumonia*. Gambar (a) merupakan sampel citra *x-ray thorax* yang diinputkan. Gambar (b) merupakan citra yang telah

KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA X-RAY THORAX MENGGUNAKAN METODE FRAKTAL BOX COUNTING

diperbaiki. Gambar (c) merupakan citra x-ray hasil segmentasi dan Gambar (d) merupakan hasil citra rontgen setelah dilakukan deteksi tepi.

Sample Citra X-Ray No Finding



Gambar 4. Pengolahan Citra No Finding

Proses pada Gambar 4 diatas merupakan pengolahan citra pada citra x-ray *no finding*. Gambar (a) merupakan sampel citra *x-ray thorax* yang diinputkan. Gambar (b) merupakan citra yang telah diperbaiki. Gambar (c) merupakan citra x-ray hasil segmentasi dan Gambar (d) merupakan hasil citra rontgen setelah dilakukan deteksi tepi.

Dimensi Fraktal

Dalam proses klasifikasi dari tiga kelas yaitu *effusion*, *pneumonia*, dan *no finding* diperlukan nilai untuk diidentifikasi sifatnya. Dalam penelitian ini nilai tersebut menggunakan besaran dimensi, karena dimensi dapat mewakili sifat dari tiap-tiap kelas. Nilai dimensi yang dihitung adalah nilai dimensi bercak paru-paru yang dianggap karakteristik dari sifat penyakit paru yang berupa nilai dimensi dari deteksi tepi. Metode dimensi fraktal yang digunakan adalah dimensi fraktal *box counting*. Berikut Sampel hasil dari perhitungan dimensi fraktal dapat dilihat pada tabel berikut.

Tabel 3. Sampel Nilai Dimensi Dari Deteksi Tepi

No.	Jenis Penyakit	Nilai Dimensi
1.	Effusion	1.6196
2.	Pneumonia	1.4656
3.	No Finding	1.2873

Klasifikasi

Dengan menggunakan nilai dimensi fraktal dari deteksi tepi canny yang digunakan untuk fitur input dalam proses klasifikasi, maka dalam penelitian ini menggunakan metode *k-nearest neighbor* (K-NN) pada proses klasifikasi citra *x-ray thotax*. Proses klasifikasi ini dilakukan dengan menggunakan *software matlab*. Sebelum melakukan proses klasifikasi, data dibagi menjadi data *training* dan data *testing* dengan

menggunakan metode *k-fold cross validation* dengan cara yang telah dijelaskan sebelumnya.

Karena data di bagi menggunakan *k-fold cross validation* dengan nilai k = 5, maka data dibagi menjadi 5 partisi dan dihitung sebanyak 5 iterasi. Setiap iterasi didapatkan nilai dari *error rate* dan dihitung nilai rata-rata dari tiap iterasinya. Hasil iterasi dapat ditampilkan pada tabel berikut.

Tabel 1. Hasil Error Rate

Iterasi ke-	Nilai Error Rate
1	0.0833
2	0.1875
3	0.2361
4	0.2604
5	0.2500
Rata-rata Error Rate	0.20346

Dari tabel hasil *error rate* di atas maka nilai akurasi dari dari contoh sistem diatas adalah :

$$\text{Akurasi} = 1 - 0.20436 = 0.79654 = 79.65\%$$

Selanjutnya untuk mencari nilai akurasi terbaik, pada metode metode *k-nearest neighbor* (K-NN) dilakukan eksperimen untuk Nilai K = 3, K = 5, K = 7, dan K = 9. Berikut adalah hasil eksperimen :

Tabel 2. Hasil Eksperimen K-NN

Nilai K-NN	Correct per iterasi pada <i>k-fold cross validation</i> (dalam %)				
	Iterasi ke-1	Iterasi ke-2	Iterasi ke-3	Iterasi ke-4	Iterasi ke-5
K = 3	66.67	70.83	70.83	75.00	76.67
K = 5	91.67	81.25	76.39	73.96	75
K = 7	62.50	70.83	72.22	75	75.83
K = 9	62.50	68.75	75	76.04	79.17

Pada tabel 2 di atas menunjukkan hasil dari eksperimen *k-nearest neighbor* (K-NN), dan perhitungan rata-rata akan ditunjukan pada tabel 3 berikut ini.

Tabel 3. Hasil Eksperimen K-NN

Nilai K-NN	Rata-rata
K = 3	72
K = 5	79.65
K = 7	71.28
K = 9	72.29

Pada tabel 3 di atas menunjukkan hasil rata-rata dari eksperimen *k-nearest neighbor* (K-NN), dimana hasil

dengan akurasi terbesar ditunjukkan pada nilai $K = 5$ yaitu 79.65% dan akurasi terendah pada nilai $K = 7$ yaitu 71.28%. Eksperimen dilakukan dengan mempartisi menjadi 5 bagian dengan menggunakan metode *k-fold cross validation*. Eksperimen pertama dilakukan pada metode *k-nearest neighbor* (K-NN) pada nilai $K = 3$. Pada iterasi ke-1 eksperimen dilakukan dengan partisi ke-1 sebagai data testing dan partisi ke-2, ke-3, ke-4, ke-5 sebagai data training pada metode *k-fold cross validation*. Pada iterasi ke-2 eksperimen dilakukan dengan partisi ke-2 sebagai data testing dan partisi ke-1, ke-3, ke-4, ke-5 sebagai data training pada metode *k-fold cross validation*. Pada iterasi ke-3 eksperimen dilakukan dengan partisi ke-3 sebagai data testing dan partisi ke-1, ke-2, ke-4, ke-5 sebagai data training pada metode *k-fold cross validation*. Pada iterasi ke-4 eksperimen dilakukan dengan partisi ke-4 sebagai data testing dan partisi ke-1, ke-2, ke-3, ke-5 sebagai data training pada metode *k-fold cross validation*. Pada iterasi ke-5 eksperimen dilakukan dengan partisi ke-5 sebagai data testing dan partisi ke-1, ke-2, ke-3, ke-4 sebagai data training pada metode *k-fold cross validation*. Dengan menggunakan cara yang sama dilakukan eksperimen pada metode *k-nearest neighbor* (K-NN) dengan nilai $K = 5$, $K = 7$, dan $K = 9$ sehingga mendapatkan hasil seperti tabel 2 dan 3 seperti diatas.

5. PENUTUP

Simpulan

1. Proses klasifikasi jenis penyakit paru menggunakan citra *x-ray* pada penelitian ini melalui beberapa tahapan, tahap pertama yaitu memastikan citra dalam bentuk *greyscale*. Setelah citra dipastikan dalam bentuk *greyscale*, proses selanjutnya adalah histogram. Fungsi histogram menunjukkan sebaran intensitas piksel dari suatu citra. Dengan adanya histogram nantinya akan terlihat kualitas citra lebih terang atau lebih gelap sehingga dengan mengekualisasi gray level, citra yang kurang bagus dapat diperbaiki kualitasnya. Hasil dari histogram tersebut akan digunakan untuk proses segmentasi paru. Proses segmentasi ini sangat penting karena dalam proses ini berguna untuk memisahkan antara paru-paru dengan organ lain sehingga bentuk paru-paru terlihat jelas. Setelah itu dilakukan deteksi tepi *canny* untuk melihat ciri penyakit paru. Hasil dari deteksi tepi *canny* akan digunakan untuk menghitung dimensi fraktal. Metode yang digunakan untuk menghitung dimensi fraktal adalah metode *box counting*. Dimana metode tersebut membagi citra dalam r iterasi dan menghitung jumlah kotak yang berisi bercak dengan iterasinya tersebut. Untuk menghitung tingkat akurasi dari dimensi dari penyakit tersebut, dilakukan beberapa kali eksperimen dengan menggunakan metode *k-flow cross validation* dan *k-nearest neighbor* (K-NN) dengan nilai k dari *k-flow cross validation* = k dan nilai k dari *k-nearest neighbor* (K-NN) = 3, 5, 7, 9.
2. Berdasarkan hasil penelitian diperoleh akurasi terbaik yaitu sebesar 91.67%. Akurasi terbaik ini berada pada *k-nearest neighbor* (K-NN) dengan nilai 5-NN pada nilai optimum.

Saran

Pengembangan yang dapat dilakukan untuk penelitian ini yaitu :

1. Diharapkan untuk penelitian selanjutnya dapat menerapkan metode lain untuk menghitung dimensi fraktal dari citra *x-ray thorax*, seperti metode *Higuchi*, *Eksponen Hurst*, dll.
2. Diharapkan penelitian selanjutnya lebih banyak melakukan eksperimen untuk pemilihan nilai k dari *k-flow cross validation* maupun *k-nearest neighbor* (K-NN).
3. Diharapkan penelitian selanjutnya dapat mengambil data primer, data yang diambil sendiri.
4. Diharapkan penelitian selanjutnya dapat menerapkan metode lain yang lebih efisien dari segi waktu dan cara kerjanya.

DAFTAR PUSTAKA

- Al-amri, S. S., Kalyankar, N. V., & Khamitkar, S. D. (2010). Image Segmentation by Using Thershod. *JOURNAL OF COMPUTING* .
- C. Gabriel, S. G. (1996). *The Dielectric Properties of Biological Tissues*. United Kingdom: King's College, London.
- Danil, C. (2008). Edge Detection with Algoritma Canny. 197–203.
- Duan, J., & Qiu, G. (2004). *Novel Histogram Processing for Colour Image Enhancement*. United Kingdom: School of Computer Science, The University of Nottingham.
- Joseph, R. P., & Singh, C. S. (2014). BRAIN TUMOR MRI IMAGE SEGMENTATION AND DETECTION IN IMAGE PROCESSING. 1-5.
- Juniati, D., & Budayasa, I. K. (2006). *Geometri Fraktal & Aplikasinya*. Surabaya: Unesa University Press.
- Narayana, R. D., Mohan, A. R., & Subramanya, B. (2014). Canny Edge Detection using Verilog. *INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY* .
- Pearce, E. C. (2013). *Anatomi dan Fisiologi untuk Paramedis*. Jakarta: PT Gramedia.
- Peterson, L. E. (2009). K-nearest neighbor. *Scholarpedia* , 4(2):1883.
- Setiawidayat, S., & Putri, S. I. (2009). Filtering Data Diskrit Elektrokardiogram Untuk Penentuan PQRST Dalam Satu Siklus.
- Tobing, E. M., & Widirahardjo. (2009). Characteristics of Patients with Pleural Effusion in RSUP H.Adam Malik Medan.

KLASIFIKASI PENYAKIT PARU BERDASARKAN CITRA X-RAY THORAX MENGGUNAKAN METODE FRAKTAL BOX COUNTING

Wang, X., Peng, Y., Lu, L., Lu, Z., Bagher, M., & Summers, R. M. (2017). ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. *Computer Vision Foundation* .