

***CLUSTERING* PARU-PARU TUBERKULOSIS DAN NORMAL
BERDASARKAN CITRA X-RAY MENGGUNAKAN
SEGMENTASI CITRA *CHAN-VESE* DAN *CANNY* DENGAN *K-
MEANS***

SKRIPSI

Diajukan untuk Memenuhi Sebagian Syarat Guna Memperoleh
Gelar Sarjana Strata Satu (S-1) dalam Ilmu Teknologi
Informasi



Oleh :

Fayza Nayla Riyana Putri

NIM. 1908096020

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI WALISONGO
SEMARANG
2023**

PERNYATAAN KEASLIAN

Yang bertandatangan dibawah ini:

Nama : Fayza Nayla Riyana Putri

NIM : 1908096020

Jurusan : Teknologi Informasi

Menyatakan bahwa skripsi yang berjudul:

***CLUSTERING PARU-PARU TUBERKULOSIS DAN NORMAL
BERDASARKAN CITRA X-RAY MENGGUNAKAN
SEGMENTASI CITRA *CHAN-VESE* DAN *CANNY* DENGAN *K-
MEANS****

Secara keseluruhan adalah hasil penelitian/karya saya sendiri,
kecuali bagian tertentu yang dirujuk sumbernya.

Semarang, 10 April 2023

Pembuat Pernyataan,



Handwritten signature of Fayza Nayla Riyana Putri.

Fayza Nayla Riyana Putri

NIM : 1908096020



KEMENTERIAN AGAMA R.I.
UNIVERSITAS ISLAM NEGERI WALISONGO
FAKULTAS SAINS & TEKNOLOGI
Jl. Prof. Dr. Hamka (Kampus II) Ngaliyan Semarang
Telp. 024-7601295 Fax.7615387

LEMBAR PENGESAHAN

Naskah proposal berikut ini :

Judul : **CLUSTERING PARU-PARU TUBERKULOSIS DAN NORMAL BERDASARKAN CITRA X-RAY MENGGUNAKAN SEGMENTASI CITRA *CHAN-VESE* DAN *CANNY* DENGAN *K-MEANS***

Penulis : Fayza Nayla Riyana Putri

NIM : 1908096020

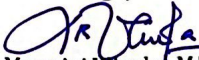
Jurusan : Teknologi Informasi

Telah diujikan dalam sidang *munaqasyah* oleh Dewan Penguji Fakultas Sains dan Teknologi UIN Walisongo Semarang dan dapat diterima sebagai salah satu syarat memperoleh gelar sarjana dalam Teknologi Informasi.

Semarang, 10 April 2023

DEWAN PENGUJI

Penguji I,


Masy Ari Ulinuha, M.T.

NIP. 198108122011011007


Penguji II,



Hery Mustofa, M.Kom.


NIP. 198703172019031007

Penguji III,


Wenty Dwi, S.Pd., M.Kom.

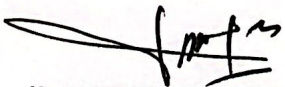
NIP. 197706222006042005

Penguji IV,


Mukhammad Ikhl Mustofa, M.Kom.

NIP. 198808072019031010

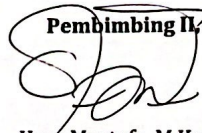
Pembimbing I,



Nur Cahyo H.W, S.T., M.Kom

NIP. 19731222 200604 1 001

Pembimbing II,



Hery Mustofa, M.Kom.

NIP. 198703172019031007

NOTA PEMBIMBING

Semarang, 10 April 2023`

Yth. Ketua Program Studi Teknologi Informasi

Fakultas Sains dan Teknologi

UIN Walisongo Semarang

Assalamu'alaikum. Wr. Wb.

Dengan ini diberitahukan bahwa saya telah melakukan bimbingan, arahan, dan koreksi naskah skripsi dengan :

Judul : *CLUSTERING* PARU-PARU TUBERKULOSIS DAN NORMAL BERDASARKAN CITRA X-RAY MENGGUNAKAN SEGMENTASI CITRA *CHAN-VESE* DAN *CANNY* DENGAN *K-MEANS*

Nama : Fayza Nayla Riyana Putri

NIM : 1908096020

Jurusan : Teknologi Informasi

Saya memandang bahwa naskah skripsi tersebut sudah dapat diajukan kepada Fakultas Sains dan Teknologi UIN Walisongo untuk diujikan dalam Sidang *Munaqasyah*

Wassalamu'alaikum. Wr. Wb.

Pembimbing I,



Nur Cahyo H.W, S.T., M.Kom

NIP. 19731222 200604 1 001

NOTA PEMBIMBING

Semarang, 10 April 2023

Yth. Ketua Program Studi Teknologi Informasi

Fakultas Sains dan Teknologi

UIN Walisongo Semarang

Assalamu'alaikum. Wr. Wb.

Dengan ini diberitahukan bahwa saya telah melakukan bimbingan, arahan, dan koreksi naskah skripsi dengan :

Judul : *CLUSTERING* PARU-PARU TUBERKULOSIS DAN NORMAL BERDASARKAN CITRA X-RAY MENGGUNAKAN SEGMENTASI CITRA *CHAN-VESE* DAN *CANNY* DENGAN *K-MEANS*

Nama : Fayza Nayla Riyana Putri

NIM : 1908096020

Jurusan : Teknologi Informasi

Saya memandang bahwa naskah skripsi tersebut sudah dapat diajukan kepada Fakultas Sains dan Teknologi UIN Walisongo untuk diujikan dalam Sidang *Munaqasyah*

Wassalamu'alaikum. Wr. Wb.

Pembimbing II,



Hery Mustofa, M.Kom.

NIP. 198703172019031007

LEMBAR PENGESAHAN

Dengan mengucapkan Puji Syukur Alhamdulillah, penulis dapat menyelesaikan karya tulis sebagai laporan tugas akhir skripsi. Karya tulis ini penulis persembahkan untuk :

1. Bapak Joko Riyono dan Ibu Retno Wulansari selaku orangtua penulis.
2. Aina Latifa Riyana Putri selaku kakak penulis.
3. Afnan Abyansyah Putra Riyono selaku adik penulis.
4. Dinda Triana selaku sahabat penulis.
5. Seluruh dosen Jurusan Teknologi Informasi.
6. Sahabat dan teman-teman seperjuangan khususnya Jurusan Teknologi Informasi 2019 Universitas Islam Negeri Walisongo Semarang.
7. Almameter Jurusan Teknologi Informasi Universitas Islam Negeri Walisongo Semarang.

MOTTO

"Karena sesungguhnya sesudah kesulitan itu ada kemudahan"

Q.S Al-Insyirah: 5

"What good are wings without the courage to fly?"

-Atticus Flinch

***CLUSTERING* PARU-PARU TUBERKULOSIS DAN NORMAL
BERDASARKAN CITRA X-RAY MENGGUNAKAN
SEGMENTASI CITRA *CHAN-VESE* DAN *CANNY* DENGAN *K-
MEANS***

Oleh :

Fayza Nayla Riyana Putri

NIM. 1908096020

ABSTRAK

Paru-paru adalah organ vital dalam tubuh manusia. Jika terjadi gangguan pada fungsi paru-paru, maka kesehatan tubuh manusia secara keseluruhan dapat terganggu. Pemeriksaan oleh tenaga medis perlu dilakukan bila terjadi gangguan pada fungsi paru. Pemeriksaan ini biasanya dapat dilakukan dengan berbagai cara, salah satunya melalui prosedur pemeriksaan rontgen dada. Penerapan *Artificial Intelligence* berkembang pesat di bidang medis, terutama dalam diagnostik dan manajemen pengobatan. Kecerdasan buatan dalam dunia medis juga dapat diterapkan dalam mengolah data citra di bidang radiologi untuk menganalisis hasil rontgen sebagai penunjang informasi diagnostik. Operator Chan-Vese dan Canny merupakan dua operator deteksi tepi dalam pengolahan citra digital dalam upaya mendapatkan informasi yang diperlukan berdasarkan bentuk dan ukuran objek. Penelitian ini dilakukan untuk

pengelompokan kondisi paru normal dan tuberkulosis berdasarkan hasil segmentasi citra rontgen dada dari Chan-Vese dan Canny menggunakan K-Means Clustering. Hasil clustering menggunakan K-Means diperoleh nilai akurasi sebesar 73%, Sensitivitas (*Recall*) sebesar 54%, Presisi sebesar 97%, dan *F1 Score* sebesar 69%.

Kata Kunci : Deteksi Tepi *Canny*, Metode Segmentasi *Chan-Vese*, *Pneumonia*, *Tuberculosis*.

KATA PENGANTAR

Segala puji penulis panjatkan kehadirat Allah SWT Sang Maha pencipta dan pemberi rahmat, yang telah melimpahkan rahmat dan hidayah-Nya kepada penulis serta telah memberikan kemudahan sehingga Tugas Akhir sebagai syarat kelulusan dalam menempuh Pendidikan di Progam Studi S1 Teknologi Informasi pada Fakultas Sains dan Teknologi Universitas Islam Negeri Walisongo Semarang telah dapat penulis selesaikan dengan judul "*CLUSTERING* PARU-PARU TUBERKULOSIS DAN NORMAL BERDASARKAN CITRA X-RAY MENGGUNAKAN SEGMENTASI CITRA *CHAN-VESE* DAN *CANNY* DENGAN *K-MEANS*".

Dalam kesempatan ini penulis ingin menyampaikan ucapan terima kasih yang tak terhingga kepada beberapa pihak yang telah banyak memberikan bantuan, bimbingan, dan pengerahan sehingga terselesaikannya Tugas Akhir ini. Ucapan terima kasih penulis sampaikan kepada:

1. Prof. Dr. H. Imam Taufiq, M.Ag., Rektor Universitas Islam Negeri Walisongo.
2. Dr. H. Ismail, M.Ag., Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Walisongo Semarang.

3. Nur Cahyo Hendro Wibowo, S.T., M.Kom., Ketua Jurusan Teknologi Informasi Universitas Islam Negeri Walisongo Semarang
4. Hery Mustofa, M.Kom., Dosen Pembimbing Skripsi Universitas Islam Negeri Walisongo Semarang.
5. Seluruh staff pengajar Fakultas Sains dan Teknologi Universitas Islam Negeri Walisongo Semarang yang telah memberikan ilmu pengetahuan yang tak ternilai.
6. Orang tua, keluarga, sahabat dan teman serta rekan seperjuangan yang selama ini menemani saya dalam menyelesaikan skripsi ini.
7. Semua pihak yang tidak disebutkan dan telah membantu Penyusunan Tugas Akhir ini.

Penulis berharap Tugas Akhir ini dapat memberikan manfaat bagi semua pihak dan pembaca pada umumnya dan kritikan yang bersifat membangun sangat penulis harapkan guna penyempurnaan di masa mendatang.

Semarang, 10 April 2023



Penulis

DAFTAR ISI

HALAMAN JUDUL.....	i
PERNYATAAN KEASLIAN	iii
LEMBAR PENGESAHAN.....	v
NOTA PEMBIMBING.....	ix
MOTTO	xiii
ABSTRAK.....	xv
DAFTAR GAMBAR	xxiii
DAFTAR TABEL	xxv
DAFTAR LAMPIRAN.....	xxvii
BAB I PENDAHULUAN	1
1.1. Latar Belakang Masalah.....	1
1.2. Identifikasi Masalah.....	7
1.3. Perumusan Masalah.....	8
2.1. Tujuan Penelitian.....	8
2.2. Manfaat Penelitian.....	9
2.3. Pembatasan Masalah	10
1.7. Sistematika Penulisan	10
BAB II TINJAUAN PUSTAKA	13

2.1.	Kajian Teori	13
2.1.1.	Penyakit Paru-Paru	13
2.1.2.	Citra X-Ray	18
2.1.3.	Pengolahan Citra.....	19
2.1.4.	Citra Grayscale	23
2.1.5.	Segmentasi.....	24
2.1.6.	Edge-Based Segmentation	25
2.1.7.	Metode <i>Chan-Vese</i>	27
2.1.8.	Metode Deteksi Tepi Canny	27
2.1.9.	Metode <i>K-Means Clustering</i>	31
2.2.	Kajian Penelitian yang Relevan.....	35
BAB III METODE PENELITIAN		37
3.1.	Teknik Pengumpulan Data	37
3.2.	Kebutuhan Perangkat Penelitian	38
3.3.	Metode Penelitian	39
3.3.1.	Pengumpulan Data	40
3.3.2.	<i>Preprocessing</i> Citra.....	41
3.3.3.	Proses Implementasi Metode Segmentasi Chan-Vese dan Canny.....	42
3.3.4.	Proses Pengolahan Data Analisa	43
3.3.5.	Proses Perhitungan Rasio (%).....	43
3.3.6.	Proses <i>Clustering</i> Data.....	44
3.3.7.	Proses Pengujian	44

3.3.8.	Uji Performa	45
BAB IV	HASIL DAN PEMBAHASAN	47
4.1.	Pengumpulan Dataset Penelitian	47
4.2.	Proses Input Citra Awal	48
4.3.	Proses Preprocessing Citra.....	48
4.4.	Proses Perhitungan Luas Jumlah Pixel Putih ...	51
4.5.	Proses Perhitungan Jumlah Keliling Pixel Putih 54	
4.6.	Proses Perhitungan Rasio	55
4.7.	Proses K-Means <i>Clustering</i>	58
4.8.	Proses Pengujian Model.....	60
4.9.	Hasil Pengujian Performa	61
BAB V	KESIMPULAN DAN SARAN	63
5.1.	Kesimpulan	63
5.2.	Saran.....	63
DAFTAR PUSTAKA		65
LAMPIRAN		73
DAFTAR LAMPIRAN.....		74

DAFTAR GAMBAR

Gambar 2.1 Citra X-Ray Pneumonia Virus.....	15
Gambar 2. 2 Citra X-Ray Normal	15
Gambar 2.3 Citra X-Ray Tuberculosis	18
Gambar 2. 4 Citra Deteksi Tepi.....	29
Gambar 2.5 Flowchart Metode K-Means	32
Gambar 3. 1 Desain Penelitian.....	39
Gambar 4.1 Source Code Input Citra.....	48
Gambar 4.2 Citra Hasil Resize.....	49
Gambar 4. 3 Citra Hasil Grayscale.....	50
Gambar 4.4 Citra Grayscale Sebelum Filtering.....	50
Gambar 4. 5 Citra Hasil Filtering	51
Gambar 4.6 Citra Grayscale dan Citra Chan-veese	51
Gambar 4. 7 Area Citra Segmentasi Chan-Vese	52
Gambar 4.8 Citra Chan-veese dan Citra Canny.....	54

DAFTAR TABEL

Tabel 3.1 Kebutuhan Perangkat Keras.....	38
Tabel 3.2 Kebutuhan Perangkat Lunak.....	38
Tabel 3.3 Jumlah Data.....	40
Tabel 3.4 Confusion Matrix.....	45
Tabel 4.1 Struktur Dataset	47
Tabel 4. 2 Data Luas Pixel Putih Tuberculosis	53
Tabel 4. 3 Data Luas Pixel Putih Normal.....	53
Tabel 4.4 Data Keliling Pixel Putih Tuberculosis	55
Tabel 4. 5 Data Keliling Pixel Putih Normal.....	55
Tabel 4. 6 Hasil Perhitungan Rasio Citra Tuberculosis .	56
Tabel 4. 7 Hasil Perhitungan Rasio Citra Normal.....	57
Tabel 4. 8 Nilai Rata-rata Rasio Citra	58
Tabel 4. 9 Nilai Centroid Cluster.....	59
Tabel 4. 10 Hasil Cluster K-Means.....	59
Tabel 4. 11 Jumlah Data Clustered Class dan Labeled Class	61

DAFTAR LAMPIRAN

Lampiran 1:Lembar Persetujuan Pembimbing.....	74
Lampiran 2 : Lembar Pengesahan Proposal.....	75
Lampiran 3 : Lembar Bimbingan Tugas Akhir	76
Lampiran 4 : Source Code Chan-Vese Demo	79
Lampiran 5 : Source Code Fungsi Segmentasi.....	80
Lampiran 6 : Source Code Luas Area Chan-Vese.....	84
Lampiran 7 : Source Code Canny dan Keliling Citra	85
Lampiran 8 : Sampel Perhitungan Segmentasi Data Tuberculosis.....	86
Lampiran 9 : Sampel Perhitungan Segmentasi Data Normal	90
Lampiran 10 : Source Code Clustering K-Means.....	93

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Paru-paru adalah organ vital dalam tubuh manusia. Organ ini merupakan organ pernapasan yang bertanggung jawab terhadap sistem pernapasan dan peredaran darah manusia. Organ yang termasuk dalam sistem pernapasan ini berfungsi sebagai media transmisi pertukaran antara (O_2) oksigen dari udara dan (CO_2) karbon dioksida dan (H_2O) uap air dari darah. Proses ini disebut respirasi eksternal atau pernapasan (Wulan et al., 2023). Paru-paru adalah organ tubuh manusia yang berfungsi dalam mengatur udara yang masuk dan memilah O_2 (Oksigen) dan CO_2 (Karbon Dioksida). Fungsi utama paru-paru adalah mengubah oksigen dari udara dengan karbon dioksida dari darah. Jika terjadi gangguan pada fungsi paru-paru, maka kesehatan tubuh manusia secara keseluruhan dapat terganggu. Pemeriksaan tenaga medis perlu dilakukan bila ada gangguan pada fungsi paru. Beberapa penyakit paru-paru memiliki gejala yang mirip seperti demam, batuk berdarah, penurunan berat badan, kehilangan nafsu makan, dan bau mulut, oleh karena itu perlu dilakukan pemeriksaan tambahan (Rahouma et al., 2021). Pemeriksaan ini biasanya dapat dilakukan dengan berbagai cara, salah satunya melalui prosedur pemeriksaan rontgen dada.

Prosedur rontgen dada memberikan gambaran berdasarkan difraksi gelombang elektromagnetik yaitu sinar-X (Wulan et al., 2023). Dengan sinar ini, kondisi organ dalam tubuh manusia dapat dilihat secara visual tanpa operasi (*non-invasive procedure*). Pemeriksaan ini tentunya harus dilakukan secara hati-hati, namun pada dasarnya pembacaan hasil rontgen cenderung dilakukan melalui pengamatan secara visual, sehingga diperlukan analisis yang tepat untuk membedakan hasil rontgen dada dengan kondisi organ (Andhi et al., 2021).

Perkembangan teknologi yang kian pesat telah memberikan bantuan di berbagai aspek kehidupan manusia. Seperti halnya dalam ayat Al-Qur'an yaitu Surat Al-A'la ayat 8 yaitu :

وَنُيَسِّرُكَ لِلْيُسْرَىٰ

Arti : “Dan Kami akan memberi kamu taufik ke jalan yang mudah,”.

Dengan kata lain, Allah S.W.T telah memberi manusia berbagai jalan dalam mencapai kemudahan baik dalam perintah agama maupun dunia (Zuhaili, 1995). Salah satu kemudahan yang ada yaitu dengan implementasi perkembangan teknologi yaitu Kecerdasan Buatan atau *Artificial Intelligence* yang telah diterapkan dan membantu manusia di kehidupan sehari-hari. *Artificial Intelligence*

merupakan cabang ilmu komputer yang berkaitan dengan menciptakan mesin pintar yang dapat melakukan pekerjaan layaknya manusia. Menurut John McCarthy (1956), Kecerdasan Buatan merupakan suatu sistem komputer yang terbentuk untuk mengetahui dan memodelkan proses-proses berpikir manusia dan mendesain mesin agar dapat menirukan perilaku manusia. Teknologi Kecerdasan Buatan dalam berbagai bentuk telah digunakan dalam berbagai bidang seperti bidang Pendidikan, perbankan dan pasar keuangan, manufaktur, pertanian dan bahkan kini implementasi perkembangan teknologi kecerdasan buatan telah dijumpai dalam dunia kesehatan.

Kecerdasan Buatan dalam dunia kesehatan tidak hanya memberikan dukungan yang andal bagi praktisi dan fasilitas medis, tetapi sekaligus memaksimalkan efisiensi praktisi. Penggunaan Kecerdasan Buatan berkembang pesat di bidang medis terutama dalam diagnostik dan manajemen pengobatan. Metode Kecerdasan Buatan (*Artificial Intelligence*) dari *Machine Learning* hingga *Deep Learning* memiliki fungsi penting dalam berbagai domain terkait kesejahteraan, termasuk meningkatkan sistem klinis baru, informasi dan catatan pasien, serta pengobatan berbagai penyakit (Kumar et al., 2022).

Kecerdasan buatan dalam dunia medis juga dapat diterapkan dalam pengolahan data pencitraan pada radiologi. Dalam radiologi, dokter terlatih secara visual menilai gambar medis dan melaporkan temuan untuk mendeteksi, mengkarakterisasi, dan memantau penyakit maka dari itu dibutuhkan teknologi kecerdasan buatan untuk menganalisa hasil rontgen sebagai informasi diagnosa pendukung.

Pemanfaatan kecerdasan buatan dalam pengolahan data pencitraan telah diteliti pada penelitian sebelumnya oleh Bayu Dwi Prasetyo pada tahun 2020 Klasifikasi Citra X-Ray Paru-Paru Anak *Pneumonia* dan *Non-Pneumonia* Menggunakan Metode Segmentasi dan Deteksi Tepi dan menghasilkan hasil uji coba dari data Rumah Sakit Sardjito dengan nilai akurasi 73.3%, sensitivitas 62.5%, dan spesifisitas 77.2%. Hasil uji coba dari data kaggle dengan nilai akurasi 83.3%, sensitivitas 70%, dan spesifisitas 90% .

Metode Chan-Vese merupakan teknik pengembangan dari Edge-Based Model dan Level Set dan merupakan metode tanpa kontur yang aktif (Prasetyo, 2020). Active Contour Model merupakan model aktif dalam teknik segmentasi yang memanfaatkan batasan energi dan gaya pada citra untuk pengelompokan wilayah (Brown et al., 2012). Model kontur aktif banyak digunakan dalam berbagai aplikasi pengolah citra, khususnya pengolah citra medis, seperti citra dari X-Ray

Thorax, CT-Scan otak, atau citra MRI organ tubuh lainnya (Hemalatha et al., 2018). Model kontur aktif digunakan dalam pemrosesan citra digital dengan cara pemisahan piksel. Metode Chan-Vese digunakan untuk mendeteksi objek citra dengan energi evolusioner. Pengembangan model Mumford-Shah dan model ini didasarkan pada masalah minimisasi energi, yang dapat diformulasi ulang dalam perumusan set level, mengarah ke cara yang lebih mudah untuk menyelesaikan masalah .

Deteksi Canny adalah teknik dalam Edge-Based Segmentation yang ditemukan oleh John F. Canny pada tahun 1986 (Shihab Ahmed, 2018). Pada dasarnya metode Canny bekerja dengan menggunakan citra grayscale sebagai masukan dan menghasilkan lokasi diskontinuitas intensitas sebagai keluaran atau citra tepi (Sekehravani et al., 2020).

Dalam penerapannya terdapat langkah-langkah metode canny edge detection yaitu smoothing atau noise filtering, menghitung besarnya gradien atau arah gradien, Menerapkan non-maximum suppression, mengatur proses Hysteresis Thresholding (Shihab Ahmed, 2018). Analisis clustering adalah teknik multivariat yang bertujuan untuk mengelompokkan pengamatan yang mirip ke dalam kelompok berdasarkan nilai yang diamati dari beberapa variabel untuk setiap individu. Tujuan dalam pembuatan clustering adalah

untuk meminimalkan variasi dalam suatu cluster dan memaksimalkan variasi antar cluster (Darwis et al., 2021). Ada banyak algoritma yang umum digunakan, salah satu algoritma yang banyak digunakan adalah K-means clustering.

K-Means clustering adalah salah satu algoritma unsupervised learning yang termasuk dalam analisis cluster non-hierarchical yang digunakan untuk mengelompokkan data berdasarkan variabel atau fitur. Tujuan K-Means Clustering, seperti metode clustering lainnya, adalah untuk mendapatkan data cluster dengan memaksimalkan kesamaan karakteristik di dalam cluster dan memaksimalkan perbedaan antar cluster yang ada. Keuntungan dari algoritma K-Means adalah sebagai berikut : Relatif sederhana dan mudah diimplementasikan, Dapat diskalakan untuk kumpulan data besar, Mudah beradaptasi dengan instance baru, Biasanya diimplementasikan ke cluster dengan bentuk dan ukuran berbeda. Algoritma K-means *clustering* mengelompokkan data berdasarkan jarak antara data dengan titik centroid cluster yang diperoleh melalui proses iteratif. Kinerja algoritma pengelompokan K-means bergantung pada klaster yang sangat efisien yang dibentuknya. Sesuai dengan maksud dan tujuan penelitian ini, akan dibuat dua klaster yaitu Tuberkulosis dan Normal.

Penelitian ini dilakukan sebagai pengembangan dari metode sebelumnya yaitu *clustering* paru menjadi paru tuberkulosis dan paru normal berdasarkan rontgen dada menggunakan K-means *clustering*. Metode yang digunakan dalam penelitian ini merupakan kombinasi dari Metode Segmentasi Chan-Vese untuk segmentasi citra dan Metode Deteksi Tepi yaitu Canny Edge-Based Segmentation untuk mendeteksi tepi yang bekerja dengan menggunakan citra grayscale sebagai input dan menghasilkan lokasi diskontinuitas intensitas sebagai output atau gambar tepi. Detektor pintar merupakan pendeteksi tepi paling kuat yang dihasilkan oleh fungsi tepi (Amoako-Yirenskyi et al., 2016). Dengan metode yang digunakan, diharapkan dapat memberikan hasil *clustering* paru yang akurat meliputi paru penderita tuberkulosis dan paru normal berdasarkan citra rontgen dada.

1.2. Identifikasi Masalah

Berdasarkan latar belakang di atas, maka dapat diidentifikasi masalah sebagai berikut :

1. Penyakit paru-paru memiliki gejala yang mirip sehingga dibutuhkan pemeriksaan penunjang.
2. Hasil citra x-ray penyakit paru-paru memiliki hasil yang mirip yaitu terdapat bercak putih pada organ paru-paru.

3. Pembacaan hasil pemeriksaan penunjang berupa rontgen *thorax* cenderung dilakukan melalui pengamatan visual sehingga diperlukan sistem yang bisa membantu analisa dengan tepat sehingga diagnosa dapat dilakukan secara tepat dan akurat.

1.3. Perumusan Masalah

Berdasarkan latar belakang penelitian diatas, maka dapat dirumuskan masalah yaitu :

1. Bagaimana hasil implementasi kombinasi Metode Segmentasi Chan Vese dan Deteksi Tepi Canny untuk pengolahan citra foto rontgen penderita Tuberculosis dan paru-paru normal.
2. Bagaimana hasil uji performa *clustering* hasil pengolahan citra foto rontgen paru-paru penderita Tuberculosis dan paru-paru normal dengan Metode K-Means.

2.1. Tujuan Penelitian

Berdasarkan rumusan masalah diatas, berikut merupakan tujuan dari penelitian :

1. Menerapkan kombinasi Metode Segmentasi Chan-Vese dan Deteksi Tepi Canny untuk menganalisa citra berdasarkan foto rontgen meliputi khususnya paru-paru penderita *Tuberculosis* dan paru-paru normal.

2. Mengetahui hasil uji performa kinerja metode *K-Means Clustering* dalam mengklusterisasi hasil pengolahan citra foto rontgen paru-paru penderita Tuberculosis dan paru-paru normal.

2.2. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat secara teoritis dan praktis seperti berikut ini :

- 1) Secara teoritis
 - a. Mempermudah tenaga kerja medis dalam menegakkan diagnosa *Tuberculosis* dalam pemeriksaan penunjang berupa rontgen *thorax*.
 - b. Mengetahui kinerja Metode Segmentasi Chan-Vese dan Deteksi Tepi *Canny* dalam menganalisa citra rontgen penyakit *Tuberculosis*.
 - c. Mengetahui kinerja Metode *K-Means Clustering* dalam melakukan klusterisasi hasil pengolahan citra foto rontgen paru-paru penderita Tuberculosis dan paru-paru normal
 - d. Memberi referensi untuk penelitian selanjutnya yang terkait.
- 2) Secara praktis
 1. Sebagai penerapan implementasi bagian ilmu dari Teknologi Informasi UIN Walisongo Semarang.

2. Dapat berkontribusi menjaga nama baik Universitas melalui kemampuan mahasiswa dalam bidangnya masing-masing.

2.3. Pembatasan Masalah

Terdapat Batasan masalah pada penelitian ini, sebagai berikut ini :

1. Citra x-ray paru-paru yang digunakan pada penelitian ini merupakan data yang diperoleh dari *Kaggle.com*.
2. Format citra yang diolah adalah *Portable Network Graphics (*.png)*.
3. Data yang diperoleh merupakan data dari *Kaggle.com* dalam rentang waktu Tahun 2021.
4. Penelitian ini dilakukan untuk *clustering* kondisi paru-paru dari citra foto rontgen khususnya paru-paru penderita *Tuberculosis* dan paru-paru normal.
5. Metode Kluster yang digunakan untuk penyakit paru-paru adalah K-Means dengan mengolah data jumlah *pixel* putih pada citra hasil Analisa.

1.7. Sistematika Penulisan

Pada laporan penelitian ini terdapat beberapa bab, maka dari itu disusunlah sistematika penulisan demi memudahkan pembaca untuk mengetahui gambaran setiap bab pada laporan ini. berikut ini sistematika penulisan pada penelitian ini :

BAB I PENDAHULUAN

Pada Bab I Pendahuluan terdiri dari latar belakang masalah, identifikasi masalah, perumusan masalah, tujuan penelitian, manfaat penelitian, pembatasan masalah, dan sistematika penulisan penelitian

BAB II TINJAUAN PUSTAKA

Dalam Bab II Tinjauan Pustaka berisi kajian teori yang membahas seputar teori dasar yang berkaitan dalam pelaksanaan penelitian seperti penjelasan dasar dari penyakit paru-paru termasuk penyakit yang digunakan sebagai topik penelitian ini serta metode yang digunakan dalam penelitian ini.

BAB III METODOLOGI PENELITIAN

Pada Bab III Metodologi Penelitian terdiri dari deskripsi dataset yang membahas terkait jenis data serta jumlah data yang diolah dalam penelitian ini, metode penelitian terkait tahapan yang dilakukan dalam penelitian, dan validitas reabilitas instrumen terkait landasan pengujian validasi.

BAB IV HASIL DAN PEMBAHASAN

Bab IV Hasil dan Pembahasan membahas terkait pembahasan daalam setiap tahap pada penelitian dan hasil dari penelitian.

BAB V KESIMPULAN DAN SARAN

Bab V Kesimpulan dan Saran berisi kesimpulan dari tahap serta hasil penelitian dan uraian saran untuk penelitian selanjutnya yang berkaitan.

BAB II

TINJAUAN PUSTAKA

2.1. Kajian Teori

2.1.1. Penyakit Paru-Paru

Penyakit paru-paru adalah gangguan yang terjadi pada salah satu organ pernapasan pada manusia khususnya paru-paru. Gangguan pada paru-paru ini dapat mempengaruhi kesehatan tubuh secara menyeluruh karena salah satu fungsi pada paru-paru adalah media transmisi untuk pernapasan eksternal atau bernapas. Kesehatan paru-paru dapat dipengaruhi oleh banyak hal mulai dari kondisi lingkungan yang kotor, kebiasaan atau gaya hidup, hingga faktor keturunan.

Secara umum, infeksi pada paru-paru dapat disebabkan oleh kontak langsung dengan penderita, hal ini dapat terjadi melalui udara dari percikan penderita saat batuk, bersin atau berbicara maupun barang yang sudah terkontaminasi. Penyebab infeksi paru-paru dapat dibedakan menjadi tiga yaitu bakteri, virus dan jamur . Jika infeksi paru-paru tidak diatasi segera, dapat menyebabkan komplikasi seperti *Tuberculosis* yang disebabkan oleh bakteri dan *Covid-19* yang disebabkan oleh virus (Shah and Shah, 2022).

Sebagian besar penyakit paru-paru disebabkan oleh asap rokok. Asap rokok dapat menyebabkan gangguan pada kinerja

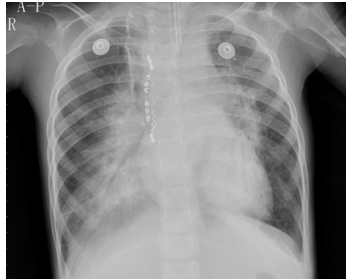
tubuh untuk menyaring udara masuk dan keluar pada paru-paru yang dapat menyebabkan penumpukan zat racun pada paru-paru. tidak hanya berimbas pada perokok aktif, tetapi perokok pasif yang terkena paparan asap rokok dapat mengalami kerusakan pada struktur paru-paru sehingga paru-paru tidak dapat bekerja dengan baik dalam menyalurkan oksigen ke darah dan menyebabkan penderita mengalami kesulitan bernapas dan gejala lainnya (Grumelli et al., 2004). Jika fungsi paru-paru sudah terganggu maka dapat menyebabkan resiko terkenanya penyakit paru-paru.

2.1.1.1. Pneumonia

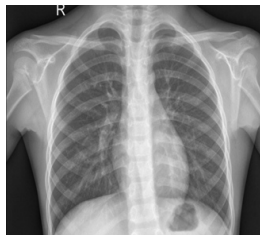
Pneumonia adalah penyakit paruparu atau infeksi yang terjadi salah satu kantung udara atau kedua paru-paru. Pneumonia dapat ditandai dengan kantung udara (alveoli) yang berisi cairan. Kondisi ini dapat menyebabkan beberapa gejala seperti batuk berdahak atau bernanah, nyeri dada, kelelahan, demam dan berkeringat, mual, muntah atau diare, dan kesulitan bernapas. Pneumonia dapat disebabkan oleh bakteri, virus, ataupun jamur.

Pada pneumonia yang disebabkan oleh bakteri biasanya hanya mempengaruhi salah satu bagian organ paru-paru. Jenis bakteri yang dapat menyebabkan pneumonia antara lain : *Streptococcus pneumoniae*, *Haemophilus influenza*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*.

Pneumonia yang disebabkan oleh virus biasanya menimbulkan gejala ringan dan singkat. Jenis virus dapat menyebabkan pneumonia seperti *espiratory syncytial virus* (RSV), *adenovirus*, virus influenza, *hantavirus*, dan *coronavirus*.



Gambar 2.1 Citra X-Ray Pneumonia Virus
(Sumber : Kaggle.com)



Gambar 2.2 Citra X-Ray Normal
(Sumber : Kaggle.com)

Pada hasil rontgen penderita yang masih mengalami infeksi ringan, maka akan terlihat gambaran buram pada satu atau kedua sisi paru-paru. Pada penderita tahap *moderate*, maka hasil rontgen akan menunjukkan bercak putih pada area paru-paru dan jika parah, hasil rontgen penderita pada tahap

severe akan menunjukkan kabut warna putih pada paru-paru secara menyeluruh (Martínez Chamorro et al., 2021).

2.1.1.2. Tuberculosis Paru

Tuberculosis merupakan radang paru-paru atau *Pneumonia* yang disebabkan oleh bakteri. Bakteri yang menyebabkan terjadinya *Tuberculosis* pada paru-paru adalah bakteri *Mycobacterium Tuberculosis* (Koch and Mizrahi, 2018). Bakteri ini pada umumnya menyerang organ paru-paru tapi bukan tidak mungkin jika bakteri ini dapat menyebar ke organ tubuh lainnya seperti tulang, sendi, kelenjar getah bening, atau selaput otak. Jika kondisi ini terjadi, maka disebut dengan *Tuberculosis* Ekstra Paru.

Menurut *World Health Organization (WHO)*, penyakit *Tuberculosis* lebih banyak terjadi di negara berkembang. Sebanyak 95% jumlah kasus *Tuberculosis* terjadi di negara berkembang salah satunya Indonesia. Jumlah kasus *Tuberculosis* di tahun 2020 sebanyak 845.000 dan sebanyak 98.300 kasus kematian yang disebabkan oleh *Tuberculosis*.

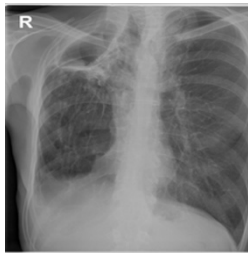
Gejala khas yang dialami oleh penderita *Tuberculosis* antara lain batuk terus-menerus yang berlangsung lebih dari 3 minggu dan biasanya mengeluarkan dahak, yang mungkin berdarah, penurunan berat badan, keringat malam, suhu tinggi, kelelahan, kehilangan selera makan, dan pembengkakan di leher. Peluang terjadinya penularan *Tuberculosis* dilihat dari

beberapa faktor risiko yaitu jumlah kasus di masyarakat sekitar, peluang kontak dengan penderita *Tuberculosis* atau sumber penularan, intensitas batuk sumber penularan, tingkat daya tahan tular dahak dari sumber penularan, durasi waktu kontak dengan sumber penularan, dan kondisi lingkungan (konsentrasi kuman di udara).

Tuberculosis merupakan penyebab kematian yang berkaitan dengan resistansi antimikroba. Jika *Tuberculosis* tidak ditangani sedari dini, kondisi serius dapat berakibat fatal untuk tubuh. Masa pengobatan *Tuberculosis* Paru memakan waktu setidaknya 6 bulan yang biasa dilakukan dengan pemberian dua antibiotik *isoniazid* dan *rifampisin* selama 6 bulan dan dua antibiotik tambahan *pirazinamid* dan *etambutol* selama dua bulan pertama (Njire et al., 2016). Pengobatan ini harus dilakukan secara teratur karena jika melewati bahkan hanya satu dosis, maka infeksi *Tuberculosis* dapat menjadi resisten terhadap antibiotik. Pemeriksaan pada *Tuberculosis* Paru dapat dilakukan melalui Tes Mantoux, Pemeriksaan Bakteriologik, Pemeriksaan IGRA, Kultur Sprutum, Pemeriksaan *Gene Xpert MTB/RIF Assay*, dan Pemeriksaan Radiologis.

Pemeriksaan Rontgen *Thorax* pada pasien *Tuberculosis* Paru dapat menghasilkan citra dengan bercak atau nodul infiltrate terutama pada lobus atas paru-paru seperti pada

Gambar 2.3 yang menunjukkan Citra *X-Ray* penderita Tuberculosis. Pemeriksaan penunjang ini juga dapat menunjukkan pembentukan kavitas, nodul klasifikasi, dan lesi nodular kecil *CT scan* dapat dilakukan untuk melihat adanya limfadenopati dan lebih *superior* dalam mengevaluasi infeksi *Tuberculosis* paru daripada rontgen *thorax* (Keliat and Abidin, 2017).



*Gambar 2.3 Citra X-Ray Tuberculosis
(Sumber : Kaggle.com)*

2.1.2. Citra X-Ray

Sinar-X adalah bentuk radiasi elektromagnetik yang memiliki energi yang lebih tinggi dan dapat menembus sebagian besar objek termasuk tubuh manusia. Dalam dunia medis, Sinar-X digunakan untuk memvisualisasi jaringan dan struktur dalam tubuh melalui hasil berupa citra gambar. Sinar-X bisa dikatakan mirip dengan cahaya tampak tetapi sinar-x dapat menembus jaringan biologis yang tidak dapat dilihat oleh cahaya tampak. Pencitraan sinar-X bergantung pada densitas bahan yang dicitrakan yaitu semakin tinggi densitas

material maka semakin tinggi penyerapan sinar-X dan penetrasi yang lebih sedikit. Dasar dari pencitraan sinar-X yaitu perbedaan densitas tulang, otot, lemak dan jaringan lunak lainnya dapat dihasilkan dari absorpsi diferensial atau koefisien atenuasi sinar-X. Perbedaan koefisien atenuasi sinar-X pada berbagai jaringan menghasilkan kontras gambar.

Sinar-X sudah diterapkan dalam dunia medis sebagai sarana pemeriksaan penunjang untuk diagnosa seperti pemeriksaan X-Ray. Pemeriksaan X-Ray memanfaatkan Sinar-X untuk menghasilkan pencitraan medis untuk melihat kondisi organ dalam tubuh. Pada prosedur X-Ray, ketika sinar-x melewati tubuh, bagian tubuh akan menyerap energi dari sinar-x dan *detector* akan mengambil sinar-x dan menghasilkan sebuah gambar citra berwarna abu-abu, hitam dan putih. Warna hitam pada citra menandakan sebagai udara, warna putih sebagai bagian tubuh seperti tulang, dan warna abu-abu menandakan otot, lemak, atau cairan (Dwi Prasetyo, 2020). Perbedaan warna ini terjadi karena jaringan tubuh dapat menangkap jumlah radiasi yang berbeda.

2.1.3. Pengolahan Citra

Citra atau *image* merupakan representasi objek dalam bentuk dua dimensi. Sebuah citra dapat direkam dengan alat optik dari pantulan cahaya pada objek tiga dimensi, sebagai contoh kamera yang menangkap objek tiga dimensi dan

diproses sehingga dapat menghasilkan citra atau image dalam dua dimensi. Secara matematis, citra atau *image* merupakan fungsi dua dimensi (x,y) dimana x dan y merupakan koordinat spasial (bidang) dan amplitudo f pada setiap pasangan koordinat (x,y) yang disebut sebagai intensitas pada citra (Putri, 2016). Jadi dapat dikatakan bahwa image direpresentasikan dalam sebuah matriks. Jika x,y dan nilai amplitudo f adalah besaran berhingga dan diskrit, maka citra tersebut dapat disebut sebagai citra digital. Pada sebuah citra digital, tersusun sejumlah elemen yang disebut sebagai piksel atau nilai tingkat keabuan. Setiap piksel memiliki lokasi dan nilai tertentu. Intensitas piksel bernilai diskrit berkisar dari 0 (warna hitam) sampai 255 (warna putih).

Pengolahan citra merupakan salah satu bidang penelitian dalam ilmu komputer. Pengolahan citra digital mencakup seluruh operasi untuk menganalisa gambar digital sehingga mendapatkan gambar yang disempurnakan dan dapat digunakan juga dalam mengekstrak informasi yang diperlukan. Pada citra analog, untuk melakukan pengolahan citra harus dikonversi terlebih dahulu menjadi citra digital (Irianto, 2016) dan dapat dilakukan dengan pengambilan *sample* matriks berukuran $m \times n$, dengan m sebagai jumlah sampling panjang dan n sebagai jumlah samping lebar pada citra. Hasil pengambilan citra dapat berbeda dengan objek asli

karena dapat disebabkan oleh berbagai faktor seperti penurunan kualitas berupa keaburan, distorsi, maupun interferensi pada alat pembuat citra. Maka dari itu, sebagai peningkatan kualitas citra dapat dilakukan dengan pengolahan citra digital.

Pengolahan citra telah diterapkan pada penelitian medis untuk menghasilkan diagnosis klinis dan rencana perawatan yang lebih efisien dan akurat, sebagai contoh pengolahan citra pada hasil pemeriksaan x-ray yang berfungsi sebagai pemeriksaan penunjang untuk mendapatkan hasil analisa yang tepat dari citra. Peningkatan citra gambar dapat dihasilkan dari peningkatan kontras pada gambar yang biasanya dilakukan pada tahap *pre-processing* sebelum melakukan teknik pemrosesan gambar lainnya. Menghasilkan citra digital dengan kecerahan, kontras, dan detail yang baik merupakan persyaratan yang kuat di beberapa bidang, salah satunya adalah analisis citra biomedis (Tuba et al., 2016). Secara garis besar, pengolahan citra digital dilakukan dengan tujuan(Kadir and Adhi, 2012):

1. Membuat gambar yang kurang kontras menjadi terlihat lebih jelas
2. Menambahkan kecerahan gambar
3. Memutar gambar
4. Memudarkan gambar

5. Menghilangkan bintik-bintik yang menodai gambar
6. Memisahkan objek dari latar belakangnya
7. Memperoleh ciri-ciri objek melalui statistika
8. Membuat objek seolah-olah dibuat dengan menggunakan pensil.

Pada dasarnya, dalam pengolahan citra terdapat beberapa tahap yang bisa dilakukan yaitu (Ratna Sulistiyanti et al., 2016):

1. Akuisisi Citra Digital

Pada tahap ini, dilakukan pengumpulan objek yang akan diubah menjadi citra dua dimensi dengan pengambilan gambar menggunakan perangkat seperti kamera, *scanner*, atau mesin x-ray.

2. Preprocessing

Pada tahap ini, citra akan diolah dan dipersiapkan terlebih dahulu dengan berbagai cara seperti :

- a. Peningkatan kualitas citra
- b. Menghilangkan *noise*
- c. Perbaiki citra (*Restoration*)
- d. Tranfromasi citra (*Transformation*)

3. Segmentasi

Pada tahap ini, akan dilakukan pemisahan bagian pada citra seperti pemisahan objek dengan latar belakang

4. Representasi

Setelah dilakukannya pemisahan objek pada tahap segmentasi, pada tahap representasi akan dilakukan klasifikasi setiap kelas objek citra berdasarkan ekstraksi ciri.

5. Pengenalan dan Interpretasi

Tahap ini bertujuan untuk memberi nama atau makna pada setiap objek dalam citra

6. Basis Pengetahuan

Tahap ini dilakukan sebagai referensi pada pengenalan pola

2.1.4. Citra Grayscale

Citra adalah matriks atau *array* dua dimensi yang memiliki baris dan kolom. Sel adalah setiap piksel dari gambar tersebut. Pada dasarnya, jenis warna citra digital terbagi menjadi tiga jenis yaitu Citra *RGB (Red, Blue, Green)*, Citra Biner, dan Citra *Grayscale* (Nafi'iyah, 2015). Citra *RGB* merupakan citra yang memiliki intensitas piksel yang tersusun dari 3 kanal yaitu *Red* (Merah), *Green* (Hijau), dan *Blue* (Biru). Sedangkan Citra Biner merupakan citra yang memiliki dua nilai intensitas yaitu 0 sebagai warna hitam dan 1 sebagai warna putih.

Perbedaan mendasar kedua citra tersebut dengan Citra *Grayscale*, yaitu pada citra ini nilai intensitas piksel didasarkan

pada derajat keabuan . Pada Citra *Grayscale* warna merah, biru, dan hijau memiliki intensitas yang sama sehingga menghasilkan warna keabuan. Pada dasarnya, warna abu-abu adalah warna ketika komponen merah, hijau dan biru memiliki intensitas yang sama dalam ruang *RGB*. Citra *Grayscale*, setiap piksel hanya berisi komponen dengan nilai intensitas berkisar antara 0 hingga 255. Dengan hitam sebagai nilai minimum dan putih sebagai nilai maksimum.

Secara matematis, konversi citra *RGB* menjadi citra *grayscale* dilakukan dengan (Dwi Prasetyo, 2020):

$$f_{grayscale}(x,y) = (0,2989 \times R(x,y)) + (0,5879 \times G(x,y)) + (0,1140 \times B(x,y)) \quad (2.1)$$

Dimana :

- $f_{grayscale}(x,y)$: Derajat keabuan pada titik (x,y)
- $R(x,y)$: Derajat keabuan citra merah pada titik (x,y)
- $G(x,y)$: Derajat keabuan citra hijau pada titik (x,y)
- $B(x,y)$: Derajat keabuan citra biru pada titik (x,y)

2.1.5. Segmentasi

Segmentasi citra merupakan sebuah metode dimana gambar digital dibagi menjadi beberapa segmen atau sub kelompok dilihat dari karakteristik piksel pada citra. Hal ini

bertujuan untuk mengurangi waktu inferensi *detector* saat memproses seluruh gambar karena dengan segmentasi, detektor tidak perlu memproses seluruh gambar melainkan hanya wilayah yang dipilih oleh algoritma segmentasi. Segmentasi citra bertujuan untuk mengklasifikasi daerah yang memiliki kesamaan sehingga juga dapat membedakan objek pada citra dengan *background* nya.

Terdapat dua teknik segmentasi yang sering digunakan dan dapat dibagi menjadi dua kategori yaitu (Joshi et al., 2014):

a. Region Segmentation

Teknik segmentasi ini digunakan untuk mencari wilayah yang memenuhi kriteria homogenitas tertentu seperti mencari nilai warna yang memiliki persamaan dengan piksel lainnya sehingga dapat dijadikan sebagai satu *region*. Contoh : *Thresholding*

b. Edge-Based Segmentation

Teknik segmentasi ini digunakan untuk mencari tepi antar wilayah dengan karakteristik yang berbeda. Contoh : *Sobel Edge Detection, Prewitt edge detection, dan Canny Edge Detection*.

2.1.6. Edge-Based Segmentation

Teknik segmentasi *Edge-Based* atau Deteksi Tepi merupakan salah satu teknik yang umum digunakan dalam analisis citra. Teknik ini digunakan dalam mencari tepi antar

wilayah citra dengan karakteristik yang berbeda. Tepi merupakan perubahan lokal dalam intensitas gambar. Tepi pada citra menandakan batas wilayah yang didasarkan pada diskontinuitas tingkat keabuan, warna, maupun tekstur dalam sebuah citra (Abubakar, 2012). Diskontinuitas merupakan perubahan pada intensitas piksel yang menggambarkan batas objek dalam sebuah citra.

Representasi tepi secara signifikan dapat mengurangi jumlah data yang akan diproses dan mempertahankan informasi penting pada bentuk objek dalam citra. Pada dasarnya, Metode deteksi tepi mengubah gambar asli menjadi gambar tepi dengan perubahan warna menjadi keabuan pada citra. Metode ini merupakan pendekatan yang paling dikenal untuk mendeteksi diskontinuitas yang signifikan dalam nilai intensitas (Amelia dan Marwati, 2012).

Beberapa keunggulan dari implementasi metode deteksi tepi dalam segmentasi yaitu algoritma yang tidak kompleks, tepi adalah fitur penting dalam citra untuk memisahkan region atau wilayah, dan dapat bekerja dengan baik dengan citra yang memiliki kontras yang baik antar objek dan latar belakang (Muthukrishnan R dan Radha, 2011). Terdapat banyak teknik deteksi tepi berbasis diskontinuitas, salah satunya adalah Metode *Chan-Vese* dan Deteksi Tepi *Canny*.

2.1.7. Metode *Chan-Vese*

Metode *Chan-Vese* adalah teknik perkembangan dari *Edge-Based Model* dan *Level Set* dan merupakan metode aktif kontur tanpa tepi (Dwi Prasetyo, 2020). Model Kontur Aktif merupakan salah satu model aktif dalam teknik segmentasi yang memanfaatkan kendala energi dan gaya pada gambar untuk pengelompokkan wilayah (Getreuer, 2012). Model kontur aktif biasa digunakan dalam berbagai aplikasi pengolahan citra khususnya pengolahan citra medis seperti pada gambar hasil X-Ray *Thorax*, *CT-Scan* otak, atau gambar *MRI* pada organ tubuh lainnya (Hemalatha et al., 2018a). Model kontur aktif digunakan dalam pengolahan citra digital dengan cara pemisahan piksel. Metode *Chan-Vese* digunakan dalam mendeteksi objek gambar dengan energi evolusi. Perkembangan dari model Mumford-Shah dan model ini didasarkan pada masalah minimisasi energi, yang dapat dirumuskan kembali dalam formulasi set level, yang mengarah ke cara yang lebih mudah untuk menyelesaikan masalah (Getreuer, 2012).

2.1.8. Metode Deteksi Tepi Canny

Deteksi Canny merupakan salah satu teknik dalam *Segmentasi Edge-Based* yang ditemukan oleh John F.Canny pada 1986 (Pratiwi et al., 2019). Pada dasarnya, dengan metode Canny bekerja dengan menggunakan citra abu-abu

atau *grayscale* sebagai input dan menghasilkan lokasi diskontinuitas intensitas sebagai *output* atau *image* tepi (Canny, 1986).

Dalam penerapannya, terdapat Langkah-langkah metode deteksi tepi canny, yaitu (Hermana and Juerman, 2015):

a. *Smoothing* atau *noise filtering*

Sebelum memulai deteksi tepi, terlebih dahulu dilakukan *smoothing* pada image untuk mengurangi jumlah tepi yang dideteksi dari *noise*. *Smoothing* pada dasarnya dilakukan dengan implementasi *low-pass filter* pada citra.

b. Menghitung *magnitude gradient* atau arah gradient

Pada dasarnya, tepi pada citra merupakan hasil dari perubahan cahaya dengan asumsi baha setiap objek memiliki piksel dengan cahaya yang tidak bervariasi. Citra hasil deteksi tepi ditunjukkan pada Gambar 2.4. Salah satu cara untuk mengukur seberapa besar perubahan cahaya yaitu untuk menghitung *magnitude gradient* pada setiap piksel. Pada operator sobel, operator didasarkan pada filter 3x3 yang menghitung komponen x dan y pada gradient :

$$G_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} * Img_{smooth} \quad (2.2)$$

$$G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} * \text{Img}_{smooth} \quad (2.3)$$

Dengan $G_x = \text{mask horizontal}$, $G_y = \text{mask vertikal}$. Setelah x dan y dihitung, maka magnitudo dapat diolah dengan :

$$|\nabla \text{Img}_{smooth}| = \sqrt{G_x^2 + G_y^2} \quad (2.4)$$

Hal ini bertujuan untuk mengetahui hasil gabungan dari tepi vertikal dan horizontal.



Gambar 2.4 Citra Deteksi Tepi
(Sumber : web.stanford.edu)

c. Menerapkan *non-maximum suppression*

Pada proses *smoothing*, batas objek pada citra menjadi kabur. Untuk mendapat tepi yang lebih tajam, digunakan *local maxima* pada gambar gradien. *Local maxima* dapat diperoleh dengan berbagai cara, salah satunya adalah

menganggap gradien sebagai *local maxima* jika lebih besar dari atau sama dengan tetangganya dalam arah positif dan negatif. Arah gradien dapat dihitung dengan :

$$\theta_G = \arctan\left(\frac{G_y}{G_x}\right) \quad (2.5)$$

d. Menetapkan proses *Hysteresis Thresholding*

Terdapat 3 kondisi dalam proses hysteresis thresholding (Filsa et al., 2019):

- Jika nilai gradient piksel diatas nilai “*High*”, maka Piksel merupakan *edge* piksel
- Jika nilai gradient piksel diatas nilai “*Low*”, maka Piksel bukan merupakan *edge* piksel
- Jika nilai gradient piksel diantara nilai “*High*” dan “*Low*”, maka piksel merupakan *edge* piksel jika dan hanya jika piksel tersebut terhubung dengan *edge* piksel atau melalui piksel yang berada diantara nilai “*High*” dan “*Low*”.

Terdapat kriteria yang merupakan keunggulan dalam penerapan Deteksi Tepi Canny, yakni (Hermana and Juerman, 2015):

a. Kriteria Deteksi

Deteksi Canny memiliki kemampuan deteksi tepi sesuai dengan parameter konvolusi yang ditentukan dan fleksibel dalam penentuan tingkat deteksi ketebalan tepi.

b. Kriteria Lokalisasi

Deteksi Canny memungkinkan untuk memberi jarak yang minimum antara tepi hasil deteksi dan tepi asli.

c. Kriteria Respon

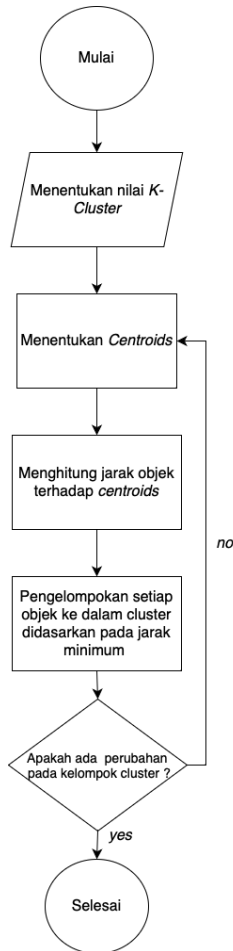
Pada setiap tepi hanya memiliki satu respon sehingga meminimalisir kesalahan pada tahap pengolahan citra selanjutnya.

2.1.9. Metode *K-Means Clustering*

Analisis *clustering* adalah teknik multivariat yang bertujuan untuk mengelompokkan pengamatan yang mirip ke dalam kelompok berdasarkan nilai yang diamati dari beberapa variabel untuk setiap individu. Tujuan dalam pembuatan *clustering* adalah untuk meminimalkan variasi dalam suatu *cluster* dan memaksimalkan variasi antar *cluster* (Kanungo et al., 2002). Ada banyak algoritma yang umum digunakan, salah satu algoritma yang banyak digunakan adalah *K-means clustering*.

K-Means Clustering adalah metode analisis data atau metode *data mining* yang melakukan proses pemodelan tanpa pengawasan dimana k adalah konstanta untuk jumlah *cluster* dan *Means* adalah nilai rata-rata dari suatu kelompok data yang didefinisikan sebagai cluster [4]. *K-Means Clustering* juga merupakan salah satu metode analisis klaster tanpa hirarki yang mengelompokkan objek berdasarkan karakteristiknya

sehingga berada dalam kluster yang sama dan objek yang memiliki karakteristik berbeda dikelompokkan bersama dengan objek sejenis pada kluster lain. Proses algoritma *K-Means* dapat digambarkan dalam diagram flowchart atau bagan alir sebagai berikut :



Gambar 2.5 Flowchart Metode *K-Means*

Berdasarkan Gambar 2.5, tahap awal penerapan algoritma *K-Means Clustering* adalah pemilihan nilai k secara acak berdasarkan jumlah *cluster* yang dibutuhkan, menentukan *centroid* atau titik pusat dari setiap *cluster*, menghitung jarak matriks objek ke titik pusat atau *centroid*, mengelompokkan setiap objek ke dalam *cluster* berdasarkan jarak minimum atau terpendek, dan ulangi setiap proses mulai dari menentukan *centroid* berdasarkan hasil akhir *cluster* sementara. Jika tidak ada perubahan anggota kelompok dari setiap *cluster*, maka iterasi dapat dihentikan dan akan terlihat hasil akhir dari *clustering* menggunakan metode *K-Means*.

Penjelasan mengenai tahapan *K-Means Clustering* dapat dirumuskan sebagai berikut (Riyono et al., 2022a):

1. Tentukan nilai k secara acak

Penentuan nilai k secara acak dilakukan berdasarkan kebutuhan atau keinginan peneliti. Nilai k digunakan sebagai jumlah *cluster* yang akan dibentuk. Penentuan nilai k dapat dilakukan melalui beberapa pertimbangan, baik konseptual maupun teoritis.

2. Tentukan titik pusat atau k centroid

Penentuan titik pusat atau centroid awal dapat dilakukan secara acak dari objek yang tersedia. Jumlah titik pusat yang ditentukan harus berdasarkan jumlah nilai k karena titik pusat yang dimaksud adalah titik

pusat dari setiap cluster. Pada proses iterasi, penentuan centroid cluster ke-i dapat dilakukan dengan persamaan :

$$v = \frac{\sum_{i=1}^n x_i}{n}; i=1,2,3,\dots,n \quad (2.6)$$

Keterangan :

v : centroid pada cluster

x_i : objek ke-i

n : jumlah objek / jumlah objek yang menjadi anggota cluster.

3. Hitung jarak setiap objek ke setiap centroid

Perhitungan jarak antara masing-masing objek dengan centroid dapat dilakukan dengan menghitung Euclidean Distance dengan persamaan sebagai berikut:

$$(x,y) = \|x-y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}, i=1,2,3,\dots,n \quad (2.7)$$

Keterangan :

x_i : objek x ke-i

y_i : pangkat ke-y

n : jumlah objek

4. Kelompokkan setiap objek ke dalam cluster

Pengelompokan setiap objek ke dalam cluster didasarkan pada jarak minimum atau jarak terpendek yang dimiliki setiap objek ke setiap titik pusat atau centroid. Pada proses iterasi, pengelompokan objek ke

dalam setiap cluster dapat dilakukan dengan Hard K-Means atau Fuzzy C-Means.

5. Melakukan iterasi

Proses iterasi dilakukan mulai dari penentuan centroid berdasarkan hasil akhir cluster sementara dan proses selanjutnya dilakukan seperti yang telah dilakukan sebelumnya. Jika pada iterasi baru tidak ada perubahan pada kelompok cluster, maka iterasi dapat dihentikan dan dapat dilihat hasil akhir cluster. Jika hasilnya sama maka algoritma analisis cluster K-Means sudah konvergen, namun jika berbeda maka belum konvergen sehingga perlu dilakukan iterasi berikutnya.

2.2. Kajian Penelitian yang Relevan

Penerapan kecerdasan buatan terbukti dapat diterapkan dalam pengolahan citra data digital pada foto rontgen.

Penelitian yang berkaitan dengan hal ini telah dilakukan oleh Yuli Sun Hariyani et al. pada tahun 2020 yaitu Deteksi Penyakit Covid-19 Berdasarkan Citra X-Ray Menggunakan *Deep Residual Network* dengan akurasi yang dicapai lebih besar 95% (Hariyani et al., 2020).

Penelitian lain dilakukan oleh M. Syaifulloh Fattah pada tahun 2021 yaitu Deteksi penyakit *Pneumonia* dan *COVID-19* menggunakan citra x-ray dengan metode *Convolutional Neural*

Network (CNN) model *GoogleNet* dengan hasil akurasi 97,22%, sensitivitas 96,71%, dan spesifisitas 98,63%(Fattah, 2021).

Penelitian lain dilakukan oleh Nur Azizah et al. pada tahun 2021 yaitu Analisis Jenis Penyakit Paru-Paru Berdasarkan *Chest X-Ray* Menggunakan Metode Fuzzy C-Means dengan hasil akurasi 86% (Nur Azizah et Al., 2021).

Penelitian lain dilakukan oleh Naibaho et al. pada tahun 2021 yaitu Identifikasi Kanker Paru pada Citra *Chest X-Ray* dengan Metode *Convolutional Neural Network* dengan hasil akurasi 87% (Naibaho, 2021).

Penelitian lain dilakukan oleh Bayu Dwi Prasetyo pada tahun 2020 Klasifikasi Citra X-Ray Paru-Paru Anak Pneumonia dan Non-Pneumonia Menggunakan Metode Segmentasi dan Deteksi Tepi dan menghasilkan hasil uji coba dari data Rumah Sakit Sardjito dengan nilai akurasi 73.3%, sensitivitas 62.5%, dan spesifisitas 77.2%. Hasil uji coba dari data kaggle dengan nilai akurasi 83.3%, sensitivitas 70%, dan spesifisitas 90% (Dwi Prasetyo, 2020).

Penelitian lain dilakukan oleh Elsha Heny Pratiwi dan Dwi Juniati pada tahun 2022 Clustering Penyakit Paru-Paru Berdasarkan Rontgen Dada Menggunakan Dimensi Fraktal Box Counting Dan K-Medoids dengan hasil akurasi sebesar 83%. (Hany Pratiwi and Juniati, 2022).

BAB III

METODE PENELITIAN

3.1. Teknik Pengumpulan Data

Pada penelitian ini, Data yang digunakan adalah jenis Data Sekunder. Data sekunder adalah data yang diperoleh atau dikumpulkan oleh orang yang melakukan penelitian dari sumber-sumber yang telah ada. Data ini digunakan untuk mendukung informasi primer yang telah diperoleh yaitu dari bahan pustaka, literatur, penelitian terdahulu, buku, dan lain sebagainya.

Data Sekunder yang digunakan dalam penelitian ini didapatkan melalui metode Studi Literatur yang merupakan suatu metode untuk pengumpulan data yang didapatkan dari sejumlah buku, majalah, atau artikel yang memiliki keterikatan terhadap masalah dan tujuan penelitian yang dimiliki (Riyono et al., 2022b). Studi Literatur juga dapat didefinisikan sebagai kegiatan dalam pengumpulan data dengan membaca, mencatat serta mengolah bahan atau *instrument* penelitian. Metode ini digunakan dengan mengumpulkan data sekunder pada penelitian yang berupa citra x-ray pada penyakit Paru-paru meliputi *Tuberculosis* dan Paru-paru Normal yang bersumber dari *kaggle.com*.

3.2. Kebutuhan Perangkat Penelitian

Pada pelaksanaan penelitian ini tentunya dibutuhkan perangkat yang menunjang keperluan penelitian. Kebutuhan perangkat penelitian dibagi menjadi dua yaitu perangkat keras dan perangkat lunak. Deskripsi Perangkat dan Spesifikasi dijabarkan pada Tabel 3.1 untuk kebutuhan perangkat keras dan Tabel 3.2 untuk kebutuhan perangkat lunak.

1. Kebutuhan Perangkat Keras

Tabel 3.1 Kebutuhan Perangkat Keras

No.	Perangkat	Spesifikasi
1	<i>Device</i>	Macbook Air M1
2	<i>Processor</i>	Apple Chip M1
3	Memori (RAM)	8 GB
4	Monitor	13,3 inch

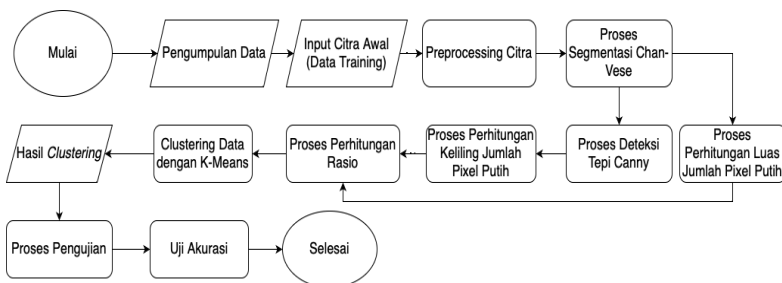
2. Kebutuhan Perangkat Lunak

Tabel 3.2 Kebutuhan Perangkat Lunak

No.	Perangkat	Spesifikasi
1	Sistem Operasi	macOS 13.1
2	Bahasa Pemograman	MATLAB, Phyton
3	<i>Word Processing</i>	Ms.Word 2019, Ms. Excel 2019
4	Peramban	Safari

3.3. Metode Penelitian

Penelitian ini merupakan jenis dari penelitian kuantitatif yaitu adalah jenis penelitian yang menghasilkan penemuan-penemuan yang dapat dicapai (diperoleh) dengan menggunakan prosedur-prosedur statistik atau cara lain dari kuantifikasi (pengukuran). Pada penelitian ini, digunakan data numerik berupa rentang nilai atau range dari jumlah piksel putih pada citra X-Ray yang selanjutnya akan diolah untuk *clustering* kondisi Paru-Paru khususnya paru-paru penderita *Tuberculosis* dan paru-paru normal. Adapun langkah-langkah pada penelitian ini direpresentasikan melalui diagram *Flowchart* pada Gambar 3.1. *Flowchart* berfungsi sebagai diagram untuk menggambarkan suatu proses urutan operasi secara visual dengan menampilkan sejumlah simbol (Yuniarti, 2019).



Gambar 3.1 Desain Penelitian

3.3.1. Pengumpulan Data

Langkah pertama dalam penelitian ini adalah mengumpulkan data terkait citra rontgen yang digunakan sebagai objek pada penelitian ini. Data berupa citra digital akan diolah dalam penelitian dalam format *Portable Network Graphics (*.png)*. Citra yang digunakan merupakan citra hasil rontgen paru-paru untuk penderita *Tuberculosis* dan paru-paru normal.

Keterangan jumlah data yang diolah dalam penelitian ini ditunjukkan pada Tabel 3.3 dibawah ini :

Tabel 3.3 Jumlah Data

	Total Data
Data Normal	1341
Data Tuberculosis	650

Keseluruhan data diperoleh melalui website *kaggle.com* yang merupakan *platform* yang menampung kumpulan dataset sebagai bahan penelitian. Pengumpulan *dataset* ini dilakukan dengan mengunduh folder *dataset* yang dibutuhkan sebagai objek penelitian. Data berupa citra ini sudah memiliki label nama sehingga sudah diketahui jenis penyakit pada setiap citra rontgen. Adapun citra tersebut merupakan hasil rontgen paru-paru penderita *Tuberculosis*

dan paru-paru normal. Dataset yang diperoleh untuk penelitian ini merupakan data yang sudah memiliki label.

3.3.2. *Preprocessing* Citra

Langkah kedua yaitu dilakukannya penginputan citra yang diperoleh untuk diolah dalam analisa penelitian. Langkah selanjutnya adalah dilakukannya *preprocessing* data sebelum dilakukannya segmentasi pada setiap citra. Langkah *Preprocessing* ini melalui tiga tahap yaitu proses Normalisasi, konversi menjadi citra *Grayscale*, dan proses *Histogram Equalization*. *Preprocessing* citra pada penelitian ini dilakukan dalam beberapa tahap, antara lain :

1. Proses Normalisasi

Proses Normalisasi pada citra bertujuan untuk mengubah ukuran resolusi citra sehingga seluruh citra memiliki ukuran resolusi yang sama. Perubahan ukuran resolusi citra akan disamakan menjadi 256 x 256 piksel. Dengan ukuran 256 x 256 piksel, maka ukuran citra dapat direpresentasikan menjadi matriks dengan 256 baris dan 256 kolom. Proses normalisasi ukuran citra dilakukan dengan fungsi *imresize* melalui *MATLAB*

2. Proses *Grayscale*

Proses *Grayscale* pada citra bertujuan untuk merubah gambar *RGB* (*Red, Green, Blue*) menjadi

gambar *grayscale* atau keabuan dengan intensitas 0 hingga 255 (Citra 8-bit). Proses *grayscale* dalam *MATLAB* dilakukan dengan fungsi *rgb2gray*.

3. Proses *Filtering*

Proses *Filtering* bertujuan untuk meningkatkan kualitas citra dengan mengurangi *noise* dan mengatur kecerahan citra. Proses *Filtering* pada penelitian menggunakan dua teknik, yakni :

1. *Histogram Equalization* yaitu proses pemerataan grafik yang merepresentasikan distribusi intensitas citra *grayscale* (Nabuasa, 2019). Proses *Filtering* pada penelitian ini menggunakan teknik *Histogram Equalization* yaitu proses pemerataan grafik yang merepresentasikan distribusi intensitas citra *grayscale* (Nabuasa, 2019). Proses ini dilakukan dengan fungsi *histeq* pada *MATLAB*.
2. *Contrast Stretching* yaitu proses meningkatkan kualitas citra dengan meningkatkan kontras pada citra (Soepomo and Jogjakarta, 2016). Proses ini dilakukan dengan fungsi *imadjust* pada *MATLAB*.

3.3.3. Proses Implementasi Metode Segmentasi Chan-Vese dan Canny

Setelah tahap *preprocessing* pada citra, maka dilakukannya segmentasi bertujuan memisahkan objek dan

background menggunakan Metode Chan-Vese dan Metode Deteksi Tepi Canny. Hasil citra sampai pada tahap ini berupa citra dengan objek yang memiliki *background* hitam. Pada penelitian ini diimplementasikan kombinasi dari Metode *Chan-Vese* dan Deteksi Tepi Canny.

Pada tahap ini dilakukannya segmentasi dengan Metode *Chan-Vese* pada citra hasil *preprocessing* lalu pada tahap selanjutnya adalah penerapan Metode Deteksi Tepi Canny dan hasil citra akan digunakan untuk perhitungan jumlah *pixel* putih.

3.3.4. Proses Pengolahan Data Analisa

Proses perhitungan luas jumlah piksel putih dilakukan pada citra hasil segmentasi Chan-Vese sedangkan untuk citra hasil segmentasi Chan-Vese dan Deteksi Tepi Canny dilakukannya proses perhitungan keliling. Kedua hasil ini akan digunakan dalam proses *clustering* dan ditunjukkan pada bagian *workspace* yang disediakan oleh *MATLAB*. Setelah itu, hasil akan diolah melalui *Ms.Excel* untuk dilakukan perhitungan pada proses *clustering*.

3.3.5. Proses Perhitungan Rasio (%)

Proses Perhitungan Rasio pada citra didasarkan pada jumlah piksel putih. Hasil dari perhitungan ini berupa nilai dari jumlah piksel putih yang akan membagi citra x-ray menjadi dua kelompok yaitu *Tuberculosis* dan Normal.

Persamaan perhitungan rasio yang diterapkan dalam penelitian ini adalah (Dwi Prasetyo, 2020) :

$$\text{Citra Rontgen} = \frac{\text{Jumlah Pixel Citra Deteksi Tepi}}{\text{Jumlah Pixel Citra Segmentasi}} \quad (3.1)$$

3.3.6. Proses *Clustering Data*

Pada proses *Clustering Data*, terdapat beberapa tahap yang harus dilakukan yaitu :

1. Menghitung rata-rata dari setiap data hasil dari perhitungan rasio pada setiap kategori penyakit yaitu *Tuberculosis* dan *Normal*.
2. Menentukan nilai *k* untuk jumlah *cluster*. Dalam kasus penelitian ini nilai $k = 2$ karena *cluster* dibagi menjadi dua *cluster* yang nantinya didefinisikan sebagai *cluster Tuberculosis* dan *Normal*.
3. Setelah mendapatkan hasil clustering yaitu terbagi menjadi dua *cluster*, maka didapatkan *centroid* dari perhitungan rata-rata setiap *cluster*.
4. Menentukan label *cluster* dengan membandingkan jarak minimum nilai rata-rata setiap *cluster* dengan rata-rata data awal sebelum di *cluster*.

3.3.7. Proses Pengujian

Semua data uji sudah memiliki label sehingga proses pengujian dapat dilakukan dengan mudah yaitu dengan cara mencocokkan apakah nilai jumlah piksel putih pada setiap

citra uji termasuk ke dalam salah satu *cluster* yang telah ditentukan. Berikut langkah-langkah dalam proses pengujian :

1. Mengelompokkan data pada setiap kategori yang termasuk dalam *cluster*. Berdasarkan titik centroid yang telah didapatkan pada tabel, kategori dapat dikelompokkan. Penentuan *cluster* dilakukan dengan cara mencocokkan nilai perhitungan jumlah piksel terdekat dengan titik centroid *cluster* yang telah ditentukan.
2. Menghitung jumlah data yang sesuai antara hasil *clustering* dengan lebel pada setiap kategori untuk dilakukan uji validasi.

3.3.8. Uji Performa

Uji performa dari implementasi metode pada penelitian ini dilakukan dalam 4 jenis penilaian dengan formula *Confusion Matrix*.

Tabel 3.4 Confusion Matrix

	Classes	Actual Class	
		A	B
Predicted Class	A	<i>TP</i>	<i>FP</i>
	B	<i>FN</i>	<i>TN</i>

Setelah mendefinisikan variabel, maka dapat dihitung melalui Tabel 3.4 yaitu Tabel *Confusion Matrix*, dapat digunakan formula dibawah ini untuk melakukan 4 jenis penilaian (Coenen, 2012) :

a. Akurasi

$$\text{Persamaan Akurasi} = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (3.2)$$

b. Sensitivitas (*Recall*)

$$\text{Persamaan Sensitivitas} = \frac{TP}{TP+FN} \times 100\% \quad (3.3)$$

c. Presisi

$$\text{Persamaan Presisi} = \frac{TP}{TP+FP} \times 100\% \quad (3.4)$$

d. *F1 Score*

$$\text{Persamaan } F1 \text{ Score} = \frac{2 \times \text{Recall} \times \text{Presisi}}{\text{Recall} + \text{Presisi}} \times 100\% \quad (3.5)$$

Keterangan : *TP (True Positive)*, *TN (True Negative)*, *FP (False Positive)*, *FN (False Negative)*.

1. *True Positive*, Data yang diprediksi positif dan data sebenarnya juga positif.
2. *True Negative*, Data yang diprediksi negatif dan data sebenarnya juga negative.
3. *False Positive*, Data negatif tetapi diprediksi sebagai data positif.
4. *False Negative*, Data positif tetapi diprediksi sebagai data negative.

BAB IV

HASIL DAN PEMBAHASAN

4.1. Pengumpulan Dataset Penelitian

Dataset diperoleh melalui website *kaggle.com*. Dataset yang digunakan menampung data yang menjadi objek penelitian yakni citra rontgen *Tuberculosis* dan Normal. Pencarian data diawali dengan memberi batasan rentang waktu yaitu data yang dimuat pada 2021. Jumlah keseluruhan data yang digunakan sebanyak 1991 seperti yang dipaparkan *pada* struktur dari *dataset* citra yang digunakan dalam dalam penelitian ini pada Gambar 4.1.

Tabel 4.1 Struktur Dataset

Struktur Data	Citra	
	<i>Tuberculosis</i>	<i>Normal</i>
Jumlah	650	1341
Format	<i>Portable Network Graphics (*.png)</i>	<i>Portable Network Graphics (*.png)</i>
Label Nama	Tuberculosis	Normal
Jenis Citra	<i>RGB dan Gray</i>	Gray
Dimensi Citra	512x512	Bervariasi

4.2. Proses Input Citra Awal

Setelah proses pengumpulan data maka tahap selanjutnya ada proses analisa yang diawali dengan input citra awal yaitu citra yang ada pada folder *Data Training*. Proses analisa pada penelitian ini menggunakan *MATLAB*. Adapun *source code* penginputan citra seperti pada Gambar 4.1.

```
clc; clear; close all;

folder_1 = ('/Users/fayzanaylariyanaputri/skripsi/dataset/train/tbc');
folder_2 = ('/Users/fayzanaylariyanaputri/skripsi/dataset/train/normal');

file_1 = dir(fullfile(folder_1, '*png'));
file_2 = dir(fullfile(folder_2, '*png'));

jumlah_file_1 = numel(file_1);
jumlah_file_2 = numel(file_2);

training_data_1 = zeros(jumlah_file_1);
training_data_2 = zeros(jumlah_file_2);
```

Gambar 4.1 Source Code Input Citra

Pada *source code* diatas terdapat variabel untuk pendefinisian *directory* dan pemanggilan seluruh file citra dalam format **png* dari folder data training yang akan diolah. Setiap hasil citra pada setiap proses akan disimpan pada sebuah folder.

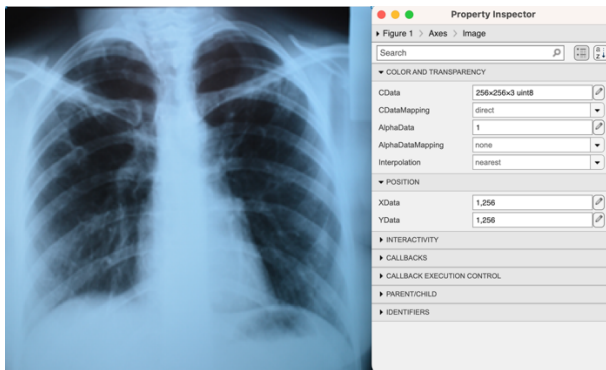
4.3. Proses Preprocessing Citra

Langkah selanjutnya setelah penginputan citra ada Preprocessing citra. Langkah ini bertujuan untuk memperbaiki citra sebelum dilakukannya analisa sehingga data mentah siap untuk diolah. Pada penelitian ini terdapat 3 tahap dalam proses preprocessing yaitu mengubah ukuran citra sehingga seluruh citra memiliki ukuran yang sama yakni 256x256 *pixel*,

mengubah citra *RGB* menjadi citra *Grayscale*, dan menerapkan filtering dengan metode *Histogram Equalization*.

1. Proses Normalisasi Ukuran Citra

Proses normalisasi yang dilakukan dengan merubah ukuran citra menjadi 256×256 *pixel* dilakukan dengan menggunakan fungsi dari *MATLAB* yakni *imresize*. Dari proses tersebut, maka dapat dihasilkan perubahan ukuran citra seperti pada Gambar 4.2.



Gambar 4.2 Citra Hasil Resize

2. Proses Grayscale

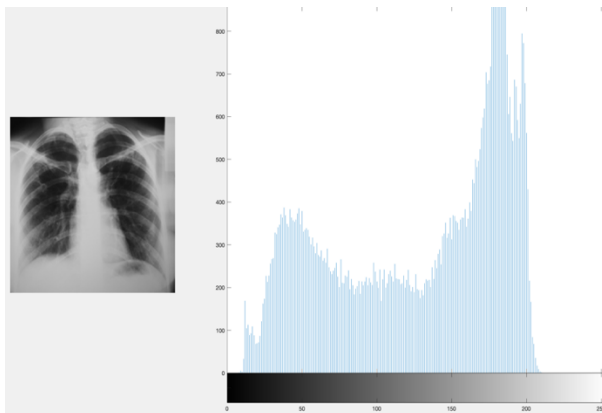
Setelah dilakukannya *Resize* pada citra, langkah selanjutnya adalah proses *Grayscale* untuk mengubah jenis citra yang sebelumnya *RGB* menjadi citra *Grayscale* dengan fungsi *rgb2gray* sehingga dapat menghasilkan citra *Grayscale* pada Gambar 4.3.



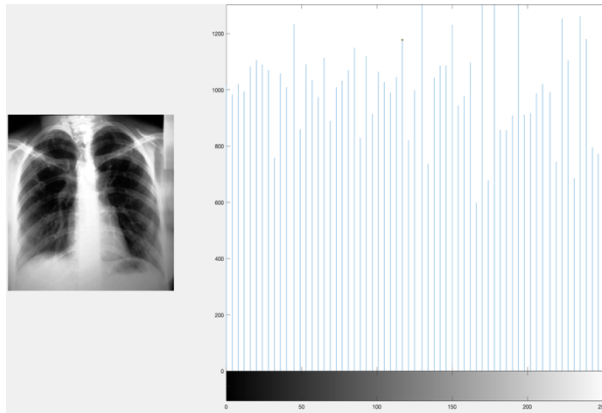
Gambar 4.3 Citra Hasil Grayscale

3. Proses *Filtering* Dengan *Histogram Equalization*

Proses *Filtering* dilakukan dengan *Histogram Equalization*. Proses ini dilakukan dengan pendistribusian seluruh piksel didalam seluruh nilai rentang secara merata sehingga citra memiliki histogram yang lebih datar.



Gambar 4.4 Citra Grayscale Sebelum Filtering

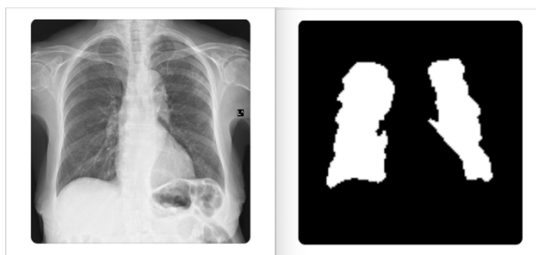


Gambar 4. 5 Citra Hasil Filtering

Dari Gambar 4.5 terlihat warna citra yang lebih kontras dan memiliki histogram yang lebih datar sehingga dapat menghasilkan citra dengan kontras yang lebih tinggi dari citra pada Gambar 4.4.

4.4. Proses Perhitungan Luas Jumlah Pixel Putih

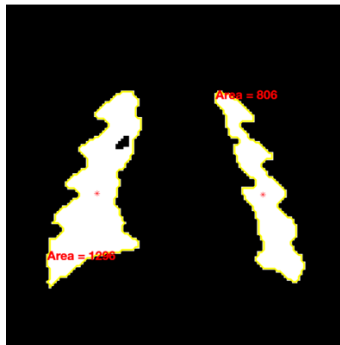
Pada citra hasil segmentasi Chan-Vese dilakukannya proses luas area pixel putih sebagai variabel untuk perhitungan *clustering* citra rontgen.



Gambar 4.6 Citra Grayscale dan Citra Chanvese

Pada proses ini digunakan fungsi *regionprops* untuk mengukur properti dari sebuah citra biner. Sebelum dilakukannya pengaplikasian fungsi tersebut, citra hasil segmentasi chan-vese sebelumnya diubah menjadi citra biner dengan fungsi *im2bw* dan dilakukannya proses pengisian daerah citra yang memiliki *holes* atau lubang dengan fungsi *imfill*.

Hasil dari perhitungan salah satu citra dapat terlihat pada Gambar 4.6. Pada Gambar 4.6, terdapat hasil luas untuk dua daerah segmentasi dalam satu citra. Setiap citra dapat memiliki jumlah daerah yang bervariasi, maka dari itu hasil penjumlahan seluruh area dilakukan sehingga dapat diterapkan pada seluruh citra mengetahui luas area *pixel* putih secara keseluruhan.



Gambar 4. 7 Area Citra Segmentasi Chan-Vese

Hasil perhitungan seluruh citra dijabarkan pada *workspace* yang disediakan dari *MATLAB* seperti pada Tabel 4.2 untuk *Tuberculosis* dan Tabel 4.3 untuk *Normal*.

Tabel 4. 2 Data Luas Pixel Putih Tuberculosis

Citra	Luas
'Tuberculosis-1.png'	4342
'Tuberculosis-10.png'	3700
'Tuberculosis-100.png'	4265
'Tuberculosis-101.png'	4426
'Tuberculosis-102.png'	3328
'Tuberculosis-103.png'	3356
'Tuberculosis-104.png'	3728

Tabel 4. 3 Data Luas Pixel Putih Normal

Citra	Luas
'IM-0115-0001.png'	2823
'IM-0117-0001.png'	2097
'IM-0119-0001.png'	2269
'IM-0122-0001.png'	2067
'IM-0125-0001.png'	1990
'IM-0127-0001.png'	2297
'IM-0128-0001.png'	2629

4.5. Proses Perhitungan Jumlah Keliling Pixel Putih

Setelah dilakukannya segmentasi Chan-Vese pada citra, maka langkah selanjutnya adalah proses deteksi tepi Canny. Proses ini diimplementasikan sehingga dapat menemukan garis tepi dan menghasilkan keliling dari citra sebagai variabel lainnya yang dibutuhkan dalam proses perhitungan *clustering*. Tepi itu sendiri ditemukan dengan mendeteksi perubahan tingkat kecerahan pada citra.

Proses Deteksi Tepi Canny dapat menggunakan fungsi *edge* dari *MATLAB*. Hasil dari operator Canny yang telah diimplementasikan terdapat pada Gambar 4.8 sebagai perbandingan citra Hasil Segmentasi Chan-Vese dan Deteksi Tepi Canny.



Gambar 4.8 Citra Chan-veve dan Citra Canny

Hasil perhitungan seluruh citra dijabarkan pada *workspace* yang disediakan dari *MATLAB* seperti pada Tabel 4.4 untuk *Tuberculosis* dan Tabel 4.5 untuk *Normal*.

Tabel 4.4 Data Keliling Pixel Putih Tuberculosis

Citra	Keliling
'Tuberculosis-1.png'	418
'Tuberculosis-10.png'	486
'Tuberculosis-100.png'	485
'Tuberculosis-101.png'	513
'Tuberculosis-102.png'	487
'Tuberculosis-103.png'	592
'Tuberculosis-104.png'	468

Tabel 4.5 Data Keliling Pixel Putih Normal

Citra	Keliling
'IM-0115-0001.png'	539
'IM-0117-0001.png'	473
'IM-0119-0001.png'	491
'IM-0122-0001.png'	375
'IM-0125-0001.png'	495
'IM-0127-0001.png'	414
'IM-0128-0001.png'	506

4.6. Proses Perhitungan Rasio

Proses *clustering* dilakukan untuk membedakan penyakit tuberculosis dengan kondisi paru normal berdasarkan jumlah piksel putih pada citra masing-masing

kategori. Proses ini dilakukan dengan menerapkan persamaan (3.1).

Berikut adalah contoh perhitungannya:

- Jumlah Deteksi Gambar Piksel dari Canny 'Tuberculosis-1.png' : 418
- Jumlah Segmentasi Citra Piksel Chan-Vese 'Tuberculosis-1.png' : 4342

Perhitungan :

$$\frac{\text{Jumlah Pixel Citra Deteksi Tepi}}{\text{Jumlah Pixel Citra Segmentasi}} = \frac{418}{4342} = 0.096269 \quad (3.3)$$

Maka dari hasil di atas dapat disimpulkan, jumlah piksel putih pada citra 'Tuberculosis-1.png' adalah 0,096269. Proses ini diterapkan di semua gambar.

Semua data perhitungan disimpan di MATLAB workspace yang kemudian diolah melalui Jupyter Notebook dengan Bahasa Pemrograman Python untuk proses K-Means *Clustering* sebagai acuan pembagian *cluster* dalam proses *Clustering*.

Tabel 4. 6 Hasil Perhitungan Rasio Citra Tuberculosis

Citra	Luas	Keliling	Rasio
'Tuberculosis-1.png'	4342	418	0,096269
'Tuberculosis-10.png'	3700	486	0,13135135

'Tuberculosis-100.png'	4265	485	0,1137163
'Tuberculosis-101.png'	4426	513	0,11590601
'Tuberculosis-102.png'	3328	487	0,14633413
'Tuberculosis-103.png'	3356	592	0,17640048
'Tuberculosis-104.png'	3728	468	0,12553648

Tabel 4. 7 Hasil Perhitungan Rasio Citra Normal

Citra	Luas	Keliling	Rasio
'IM-0115-0001.png'	2823	539	0,19093163
'IM-0117-0001.png'	2097	473	0,22556032
'IM-0119-0001.png'	2269	491	0,21639489
'IM-0122-0001.png'	2067	375	0,18142235
'IM-0125-0001.png'	1990	495	0,24874372

'IM-0127-0001.png'	2297	414	0,18023509
'IM-0128-0001.png'	2629	506	0,19246862

4.7. Proses K-Means *Clustering*

Langkah awal sebelum dilakukan *clustering* adalah menghitung rata-rata dari semua data yaitu data Tuberkulosis dan Normal. Nilai rata-rata ini merupakan nilai yang diperoleh dari proses sebelumnya yaitu hasil rasio.

Tabel 4. 8 Nilai Rata-rata Rasio Citra

Data	<i>Mean (Average)</i>
Tuberculosis	0,137366886
Normal	0,18574854

Proses *clustering* ini bertujuan untuk mengelompokkan data menjadi dua cluster yaitu *Cluster 0* dan *Cluster 1*. *Cluster 0* memiliki nilai yang paling mendekati nilai rata-rata pada data Tuberkulosis begitu juga dengan *Cluster 1* pada data Normal, sehingga dapat disimpulkan bahwa *Cluster 0* mendefinisikan *Cluster* Tuberkulosis dan *Cluster 1* untuk *Cluster* Normal. Proses *clustering* dilakukan melalui aplikasi *Jupyter Notebook* dan didapatkan titik *centroid* serta hasil sebagai berikut :

Tabel 4. 9 Nilai Centroid Cluster

Cluster Class	Centroid
Cluster 0	0,144031
Cluster 1	0,203558

Tabel 4. 10 Hasil Cluster K-Means

Citra	Hasil	Cluster
'Tuberculosis-1.png'	0,096269	0
'Tuberculosis-10.png'	0,131351	0
'Tuberculosis-100.png'	0,113716	0
'Tuberculosis-101.png'	0,115906	0
'Tuberculosis-102.png'	0,146334	0
'Tuberculosis-104.png'	0,125536	0
.....
'NORMAL2-IM-1360-0001.png'	0,190455	1
'NORMAL2-IM-1365-0001.png'	0,25779	1
'NORMAL2-IM-1371-0001.png'	0,279753	1
'NORMAL2-IM-1396-0001.png'	0,220749	1
'NORMAL2-IM-1400-0001.png'	0,176865	1

'NORMAL2-IM-1401-0001.png'	0,177606	1
'NORMAL2-IM-1406-0001.png'	0,20443	1
'NORMAL2-IM-1412-0001.png'	0,226816	1
'NORMAL2-IM-1419-0001.png'	0,182996	1
'NORMAL2-IM-1422-0001.png'	0,279835	1
'NORMAL2-IM-1423-0001.png'	0,224031	1

4.8. Proses Pengujian Model

Semua data uji sudah memiliki label sehingga proses pengujian dapat dilakukan dengan mudah yaitu dengan cara mencocokkan apakah nilai jumlah piksel putih pada setiap citra uji termasuk ke dalam salah satu *cluster* yang telah ditentukan yaitu pada Tabel. Berikut langkah-langkah dalam proses pengujian :

1. Mengelompokkan data pada setiap kategori yang termasuk dalam *cluster*. Berdasarkan titik centroid yang telah didapatkan pada tabel, kategori dapat dikelompokkan. Penentuan *cluster* dilakukan dengan

cara mencocokkan nilai perhitungan jumlah piksel terdekat dengan titik centroid cluster yang telah ditentukan.

2. Menghitung jumlah data pada setiap kategori untuk dilakukan uji performa dari implementasi model.

Setelah dilakukan pengelompokan data, langkah selanjutnya adalah menghitung jumlah data benar dan salah yang masuk ke dalam *cluster* yang telah ditentukan dengan hasil seperti pada Tabel 4.11.

Tabel 4. 11 Jumlah Data Clustered Class dan Labeled Class

	Classes	Predicted Class	
		Tuberculosis	Normal
Actual Class	Cluster 0 (Tuberculosis)	629	21
	Cluster 1 (Normal)	521	820

4.9. Hasil Pengujian Performa

Tabel 4.11 mendefinisikan jumlah data berlabel paru-paru yang termasuk dalam *cluster* paru-paru tuberkulosis terklastr sebanyak 629 dari total 650 data sedangkan jumlah data berlabel paru-paru normal yang termasuk dalam *cluster* paru-paru normal yang terklastr adalah 820 data dari total 1341 data.

Uji performa dilakukan melalui 4 penilaian dengan hasil sebagai berikut :

a. Akurasi

$$\text{Akurasi} = \frac{629+820}{629+820+21+521} \times 100\% = 73\%$$

b. Sensitivitas (*Recall*)

$$\text{Sensitivitas} = \frac{629}{629+521} \times 100\% = 54\%$$

c. Presisi

$$\text{Presisi} = \frac{629}{629+21} \times 100\% = 97\%$$

d. *F1 Score*

$$F1 \text{ Score} = \frac{2 \times \text{Recall} \times \text{Presisi}}{\text{Recall} + \text{Presisi}} = 69\%$$

BAB V

KESIMPULAN DAN SARAN

5.1. Kesimpulan

Penelitian ini dilakukan sebagai penyempurnaan dari metode sebelumnya yaitu *clustering* kondisi paru yang meliputi tuberkulosis dan paru normal berdasarkan citra rontgen dada dengan menggunakan metode K-Means *Clustering*. Dari hasil pembahasan diketahui bahwa uji performa penerapan model algoritma pada data dapat ditentukan dari jumlah data berlabel yang sesuai dengan *clustering*. Implementasi *Clustering K-Means* dari citra kombinasi Metode Segmentasi *Chan-Vese Segmentation* dan *Canny Edge-Based Segmentation* menghasilkan nilai akurasi sebesar 73%, Sensitivitas (*Recall*) sebesar 54%, Presisi sebesar 97%, dan *F1 Score* sebesar 69%. Berdasarkan analisis hasil, untuk meningkatkan akurasi hasil, perlu dilakukan pemilihan jumlah sampel data yang lebih seimbang antara foto rontgen dada paru normal dan paru tuberkulosis.

5.2. Saran

Dari penelitian yang telah dilakukan ini, diharap dapat digunakan sebagai bahan referensi untuk dilakukan penelitian selanjutnya yang berkaitan dan dapat meningkatkan nilai akurasi dari *Clustering K-Means* pada hasil implementasi

Metode *Chan-Vese* dan Metode Deteksi Tepi yaitu *Canny Edge-Based Segmentation*.

DAFTAR PUSTAKA

- Abubakar, F.M., 2012. Study Of Image Segmentation By Using Edge Detection Techniques. *International Journal of Engineering Research & Technology* 1. <https://doi.org/10.17577/IJERTV1IIS9239>
- Adln, P., Universitas, A., Skripsi, D., Kanker, P.-P., Dari, C., Foto, R., Menggunakan, J., Saraf, T., Backpropagation, T., Deviasari, W., 2012. PROGRAM STUDI TEKNOBIOMEDIK DEPARTEMEN FISIKA FAKULTAS SAINS DAN TEKNOLOGI UNIVERSITAS AIRLANGGA 2012.
- Amelia, L., Marwati, R., 2012. PERBANDINGAN METODE ROBERTS DAN SOBEL DALAM MENDETEKSI TEPI SUATU CITRA DIGITAL.
- Amoako-Yirenkyi, P., Appati, J.K., Dontwi, I.K., Amoako-Yirenkyi, P., Appati, J.K., Dontwi, I.K., 2016. Performance Analysis of Image Smoothing Techniques on a New Fractional Convolution Mask for Image Edge Detection. *Open Journal of Applied Sciences* 6, 478–488. <https://doi.org/10.4236/OJAPPS.2016.67048>
- Andhi, P., Pratama, H., Teguh, R., Sahay, A.S., Wilentine, V., Informatika, T., Teknik, F., Kampus, P., Nyaho, T., Yos, J., Palangka, S., 73112, R., Kesehatan, L., Provinsi, K., Tengah, K., Raya, P., 2021a. Deteksi COVID-19 Berdasarkan Hasil Rontgen Dada (Chest Xray) Menggunakan Python. *Journal of Information Technology and Computer Science* 1, 58–67. <https://doi.org/10.47111/JOINTECOMS.V1I1.2956>
- Andhi, P., Pratama, H., Teguh, R., Sahay, A.S., Wilentine, V., Informatika, T., Teknik, F., Kampus, P., Nyaho, T., Yos, J., Palangka, S., 73112, R., Kesehatan, L., Provinsi, K., Tengah, K., Raya, P., 2021b. Deteksi COVID-19 Berdasarkan Hasil Rontgen Dada (Chest Xray) Menggunakan Python. *Journal of Information*

- Technology and Computer Science 1, 58–67.
<https://doi.org/10.47111/JOINTECOMS.V1I1.2956>
- Brown, E.S., Chan, T.F., Bresson, X., 2012. Completely convex formulation of the Chan-Vese image segmentation model. *Int J Comput Vis* 98, 103–121.
<https://doi.org/10.1007/S11263-011-0499-Y/METRICS>
- Canny, J., 1986. A Computational Approach to Edge Detection. *IEEE Trans Pattern Anal Mach Intell PAMI-8*, 679–698.
<https://doi.org/10.1109/TPAMI.1986.4767851>
- Darwis, M., Hasibuan, L.H., Firmansyah, M., Ahady, N., Tiaharyadini, R., 2021. Implementation of K-Means clustering algorithm in mapping the groups of graduated or dropped-out students in the Management Department of the National University. *JISA(Jurnal Informatika dan Sains)* 4, 1–9. <https://doi.org/10.31326/JISA.V4I1.848>
- Dwi Prasetyo, B., 2020. Klasifikasi Citra X-Ray Paru-Paru Anak Pneumonia dan Non-Pneumonia Menggunakan Metode Segmentasi dan Deteksi Tepi [WWW Document]. URL <https://dspace.uui.ac.id/handle/123456789/28404;jsessionid=A40324F8FED71BB89D98E5C98FE00327> (accessed 9.22.22).
- Fattah, M.S., 2021. Deteksi penyakit pneumonia dan COVID-19 menggunakan citra x-ray dengan metode Convolutional Neural Network (CNN) model GoogleNet [WWW Document]. URL <http://digilib.uinsby.ac.id/49030/> (accessed 9.26.22).
- Filsa, N., Widodo, Prasetya Adhi, B., 2019. Kinerja Algoritma Canny untuk Mendeteksi Tepi dalam Mengidentifikasi Tulisan pada Citra Digital Meme. *PINTER: Jurnal Pendidikan Teknik Informatika dan Komputer* 3, 45–53.
<https://doi.org/10.21009/pinter.3.1.8>

- Getreuer, P., 2012. Chan-Vese Segmentation. <https://doi.org/10.5201/ipol.2012.g-cv>
- Grumelli, S., Corry, D.B., Song, L.Z., Song, L., Green, L., Huh, J., Hacken, J., Espada, R., Bag, R., Lewis, D.E., Kheradmand, F., 2004. An immune basis for lung parenchymal destruction in chronic obstructive pulmonary disease and emphysema. *PLoS Med* 1, 075–083. <https://doi.org/10.1371/JOURNAL.PMED.0010008>
- Hany Pratiwi, E., Juniati, D., 2022. Clustering Penyakit Paru-Paru Berdasarkan Rontgen Dada Menggunakan Dimensi Fraktal Box Counting Dan K-Medoids. *J. Ris. & Ap. Mat* 06, 1–12.
- Hariyani, Y., Hadiyoso, S., Siadari, T.S., 2020. Deteksi Penyakit Covid-19 Berdasarkan Citra X-Ray Menggunakan Deep Residual Network. *ELKOMIKA: Jurnal Teknik Energi Elektrik, Teknik Telekomunikasi, & Teknik Elektronika* 8, 443. <https://doi.org/10.26760/elkomika.v8i2.443>
- Hemalatha, R.J., Thamizhvani, T.R., Dhivya, A.J.A., Joseph, J.E., Babu, B., Chandrasekaran, R., 2018a. Active Contour Based Segmentation Techniques for Medical Image Analysis. *Medical and Biological Image Analysis*. <https://doi.org/10.5772/INTECHOPEN.74576>
- Hemalatha, R.J., Thamizhvani, T.R., Dhivya, A.J.A., Joseph, J.E., Babu, B., Chandrasekaran, R., Hemalatha, R.J., Thamizhvani, T.R., Dhivya, A.J.A., Joseph, J.E., Babu, B., Chandrasekaran, R., 2018b. Active Contour Based Segmentation Techniques for Medical Image Analysis. *Medical and Biological Image Analysis*. <https://doi.org/10.5772/INTECHOPEN.74576>

- Hermana, A.N., Juerman, S., 2015. Implementasi Algoritma Canny dan Backpropagation dalam Pengenalan Pola Rumah Adat.
- Irianto, S., 2016. Analisa Citra Digital dan Content Based Image Retrieval.
- Joshi, R., Singh, A.K., Ha, P., Perumal, K., 2014. A Study Analysis on the Different Image Segmentation Techniques Role of Segment at ion Technique in Digital Image Processing (DIP).
- Kadir, A., Adhi, S., 2012. Ekstraksi Fitur Tekstur. Teori dan Aplikasi Pengolahan Citra 575–640.
- Kanungo, T., Mount, D.M., Netanyahu, N.S., Piatko, C.D., Silverman, R., Wu, A.Y., 2002. An efficient k-means clustering algorithms: Analysis and implementation. *IEEE Trans Pattern Anal Mach Intell* 24, 881–892. <https://doi.org/10.1109/TPAMI.2002.1017616>
- Keliat, E.N., Abidin, A., 2017. DIAGNOSIS TUBERKULOSIS.
- Koch, A., Mizrahi, V., 2018. Mycobacterium tuberculosis. *Trends Microbiol* 26, 555–556. <https://doi.org/10.1016/J.TIM.2018.02.012>
- Kumar, Y., Koul, A., Singla, R., Ijaz, M.F., 2022. Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda. *J Ambient Intell Humaniz Comput* 1, 1. <https://doi.org/10.1007/S12652-021-03612-Z>
- Martínez Chamorro, E., Díez Tascón, A., Ibáñez Sanz, L., Ossaba Vélez, S., Borruel Nacenta, S., 2021. Radiologic diagnosis of patients with COVID-19. *Radiología (English Edition)* 63, 56–73. <https://doi.org/10.1016/J.RXENG.2020.11.001>
- Muthukrishnan R, Radha, M., 2011. EDGE DETECTION TECHNIQUES FOR IMAGE SEGMENTATION. *International Journal of Computer Science &*

- Information Technology (IJCSIT) 3.
<https://doi.org/10.5121/ijcsit.2011.3620>
- Nabuasa, Y., 2019. PENGOLAHAN CITRA DIGITAL PERBANDINGAN METODE HISTOGRAM EQUALIZATION DAN SPESIFIKATION PADA CITRA ABU-ABU. J-ICON 7, 87–95.
- Nafi'iyah, N., 2015. Algoritma Kohonen dalam Mengubah Citra Graylevel Menjadi Citra Biner [WWW Document]. URL <https://www.jurnal.stmikasia.ac.id/index.php/jitika/article/view/125/146> (accessed 9.22.22).
- Naibaho, Y.N., 2021. Identifikasi Kanker Paru pada Citra Chest X-Ray dengan Metode Convolutional Neural Network [WWW Document]. URL <https://repositori.usu.ac.id/handle/123456789/47406> (accessed 9.22.22).
- Njire, M., Tan, Y., Mugweru, J., Wang, C., Guo, J., Yew, W.W., Tan, S., Zhang, T., 2016. Pyrazinamide resistance in Mycobacterium tuberculosis: Review and update. Adv Med Sci 61, 63–71. <https://doi.org/10.1016/J.ADVMS.2015.09.007>
- Nur Azizah, F., Juniati, D., 2021. ANALISIS JENIS PENYAKIT PARU-PARU BERDASARKAN CHEST X-RAY MENGGUNAKAN METODE FUZZY C-MEANS. Jurnal Ilmiah Matematika.
- Prasetyo, D., 2020. Klasifikasi Citra X-Ray Paru-Paru Anak Pneumonia dan Non-Pneumonia Menggunakan Metode Segmentasi dan Deteksi Tepi.
- Pratiwi, S.M., Mulyana, I., Thaheer, H., 2019. PENGOLAHAN CITRA DIGITAL MENGGUNAKAN METODE CANNY DAN EUCLIDEAN DISTANCE UNTUK MENGUKUR TINGKAT KEMIRIPAN SKETSA WAJAH.

- Rahouma, K.H., Mabrouk, S.M., Aouf, M., 2021. Lung Cancer Diagnosis Based on Chan-Vese Active Contour and Polynomial Neural Network. *Procedia Comput Sci* 194, 22–31.
<https://doi.org/10.1016/J.PROCS.2021.10.056>
- Ratna Sulistiyanti, S., Sentyawan, A., Komarudin, M., 2016. PENGOLAHAN CITRA; DASAR DAN CONTOH PENERAPANNYA.
- Riani Putri Jurusan Pendidikan Teknologi Informasi, A., PGRI Tulungagung Jl Mayor Sujadi Timur no, S., 2016. PENGOLAHAN CITRA DENGAN MENGGUNAKAN WEB CAM PADA KENDARAAN BERGERAK DI JALAN RAYA. *Jurnal Ilmiah Pendidikan Informatika* 1, 1–6.
- Riyono, J., Puspa, S.D., Pujiastuti, C.E., 2022a. Simulasi Clustering Provinsi di Indonesia dalam Penyebaran Covid-19 Berdasarkan Indikator Kesehatan Masyarakat Menggunakan Algoritma Gaussian Mixture Model. *MAJAMATH: Jurnal Matematika dan Pendidikan Matematika* 5, 43–60.
<https://doi.org/10.36815/MAJAMATH.V5I1.1699>
- Riyono, J., Puspa, S.D., Pujiastuti, C.E., 2022b. Simulasi Clustering Provinsi di Indonesia dalam Penyebaran Covid-19 Berdasarkan Indikator Kesehatan Masyarakat Menggunakan Algoritma Gaussian Mixture Model. *MAJAMATH: Jurnal Matematika dan Pendidikan Matematika* 5, 43–60.
<https://doi.org/10.36815/MAJAMATH.V5I1.1699>
- Sekehravani, E.A., Babulak, E., Masoodi, M., 2020. Implementing canny edge detection algorithm for noisy image. *Bulletin of Electrical Engineering and Informatics* 9, 1404–1410.
<https://doi.org/10.11591/EEI.V9I4.1837>

- Shah, A., Shah, M., 2022. Advancement of Deep Learning in Pneumonia/Covid-19 Classification and Localization: A Qualitative and Quantitative Analysis.
- Shihab Ahmed, A., 2018. COMPARATIVE STUDY AMONG SOBEL, PREWITT AND CANNY EDGE DETECTION OPERATORS USED IN IMAGE PROCESSING. *J Theor Appl Inf Technol* 15, 19.
- Soepomo, J., Jogjakarta, J., 2016. PERBAIKAN KUALITAS CITRA MENGGUNAKAN HISTOGRAM LINEAR CONTRAST STRETCHING PADA CITRA SKALA KEABUAN Murinto Program Studi Teknik Informatika Universitas Ahmad Dahlan Kampus III UAD.
- Tuba, M., Jordanski, M., Arsic, A., 2016. Improved weighted thresholded histogram equalization algorithm for digital image contrast enhancement using the bat algorithm. *Bio-Inspired Computation and Applications in Image Processing* 61–86. <https://doi.org/10.1016/B978-0-12-804536-7.00004-1>
- Wulan, T.D., Purwanti, E., Yasin, M., 2023. Deteksi Kanker Paru-Paru Dari Citra Foto Rontgen Menggunakan Jaringan Saraf Tiruan Backpropagation [WWW Document]. UNAIR REPOSITORY. URL <https://repository.unair.ac.id/25693/> (accessed 2.11.23).
- Yuniarti, W.D., 2019. Dasar-dasar pemrograman dengan Python [WWW Document]. Deepublish. URL <https://opac.perpusnas.go.id/DetailOpac.aspx?id=1279704> (accessed 11.23.22).
- Zuhaili, W., 1995. At-Tafsir Al-Wajiz Wa Mu'jam Ma'ani Al-Qur'an Al-Aziz.

LAMPIRAN

DAFTAR LAMPIRAN**Lampiran 1: Lembar Persetujuan Pembimbing****PERSETUJUAN PEMBIMBING.**

Proposal Skripsi ini telah disetujui oleh Pembimbing antara
dan disahkan Disetujui pada

Hari Kamis

Tanggal 20 Oktober 2022

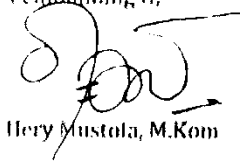
Pembimbing I,



Nur Cahyo Hendro Wibowo,
S.T., M.Kom

NIP. 19731222 200604 1 001

Pembimbing II,




Hery Mustala, M.Kom

NIP. 19870317 201903 1 007

Mengetahui,

Ketua Jurusan Teknologi Informasi



Nur Cahyo Hendro Wibowo, S.T., M.Kom

NIP. 19731222 200604 1 001

Lampiran 2 : Lembar Pengesahan Proposal

PENGESAHAN

Naskah proposal berikut ini :

Judul : KLASIFIKASI PENYAKIT PARU-PARU BERDASARKAN
CITRA X-RAY DENGAN METODE SEGMENTASI CHAN-VESE DAN DETEKSI
TEPI CANNY

Penulis : Fayza Nayla Riyana Putri

NIM : 1908096020

Jurusan : Teknologi Informasi

Telah diujikan dalam sidang komprehensif oleh Dewan Penguji Fakultas
Sains dan Teknologi UIN Walisongo Semarang dan dapat diterima sebagai
salah satu syarat memperoleh gelar sarjana dalam Teknologi Informasi.

Semarang, 30 November 2022

DEWAN PENGUJI

Penguji I,



Masy Ari Ulinuha, M.T.

NIP. 198108122011011007

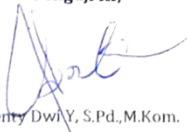
Penguji II,



Hery Mustofa, M.Kom.

NIP. 198703172019031007

Penguji III,



Wenny Dwi Y. S.Pd., M.Kom.

NIP. 197706222006042005

Penguji IV,



Mokhammad Ikhlil Mustofa, M.Kom.

NIP. 198808072019031010

Lampiran 3 : Lembar Bimbingan Tugas Akhir

LEMBAR BIMBINGAN TUGAS AKHIR



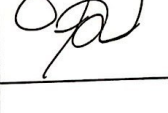
Nama : Fayza Nayla Riyana Putri
 NIM : 1908096020
 Judul : Clustering Paru-Paru Tuberkulosis Dan Normal Berdasarkan Citra X-Ray Menggunakan Segmentasi Citra Chan-Vese Dan Canny Dengan K-Means.
 Dosen Pembimbing I : Nur Cahyo Hendro Wibowo S.T.,M.Kom



No	Tanggal Bimbingan	Rincian Kegiatan Bimbingan	Tanda Tangan Pembimbing
1	10-10-2022	Bimbingan BAB I, BAB II, BAB III	
2	19-10-2022	Bimbingan BAB I, BAB II, BAB III	
3	28-11-2022	Bimbingan Analisa	
4	21-02-2023	Bimbingan Analisa, BAB IV	
5	14-03-2023	Bimbingan Analisa, BAB IV, BAB V	

LEMBAR BIMBINGAN TUGAS AKHIR

Nama : Fayza Nayla Riyana Putri
NIM : 1908096020
Judul : Clustering Paru-Paru Tuberkulosis Dan Normal Berdasarkan Citra X-Ray Menggunakan Segmentasi Citra Chan-Vese Dan Canny Dengan K-Means.

Dosen Pembimbing II : Hery Mustofa, M.Kom.

No	Tanggal Bimbingan	Rincian Kegiatan Bimbingan	Tanda Tangan Pembimbing
1	03-10-2022	Bimbingan Gambaran Umum Penelitian	
2	05-10-2022	Bimbingan BAB I, BAB II	
3	07-10-2022	Bimbingan BAB I, BAB II	
4	10-10-2022	Bimbingan BAB I, BAB II, BAB III	
5	19-10-2022	Bimbingan BAB I, BAB II, BAB III	
6	28-11-2022	Bimbingan Analisa	

7	21-02-2023	Bimbingan Analisa, BAB IV	
8	14-03-2023	Bimbingan Analisa, BAB IV, BAB V	

No	Tanggal	Rincian Kegiatan Bimbingan	Tanda Tangan
1	07-10-2022	Bimbingan Pembinaan Bab IV	
2	08-10-2022	Bimbingan BAB I, BAB II	
3	07-10-2022	Bimbingan BAB I, BAB II	
4	10-10-2022	Bimbingan BAB I, BAB II, BAB III	
5	19-10-2022	Bimbingan BAB I, BAB II, BAB III	
6	28-11-2022	Bimbingan Analisa	

Lampiran 4 : Source Code Chan-Vese Demo

```

for k = 1:jumlah_file_2
    image_read = imread(fullfile(folder_2, file_2(k).name));
    image_resize = imresize(image_read, [256 256]);
    Iinitial = image_resize;
    [rows, columns, numberOfColorChannels]= size(Iinitial);
    if numberOfColorChannels > 1
        image_gray = rgb2gray(Iinitial);
    else
        image_gray = Iinitial;
    end
    image_histeq = histeq(image_gray);
    I = imadjust(image_histeq);

    %-- load citra
    m = zeros(size(I,1),size(I,2));%-- membuat initial mask
    m(90:170,50:90) = 1;
    m(90:170,160:200) = 1;

    I = imresize(I,.5); %-- resize citra
    m = imresize(m,.5);

    subplot(2,2,1); imshow(I); title('Input Image');
    subplot(2,2,2); imshow(m); title('Initialization');
    subplot(2,2,3); title('Segmentation');

    seg = region_seg(I, m, 50); %-- Run segmentasi

    subplot(2,2,4); imshow(seg); title('Global Region-Based Segmentation');
    imwrite(seg,fullfile(folder_newdit, file_2(k).name));
end

```

Lampiran 5 : Source Code Fungsi Segmentasi

```

function seg = region_seg(I,init_mask,max_its,alpha,display)

%-- nilai default parameter alpha : .1
if(~exist('alpha','var'))
    alpha = .2;
end
%-- default behavior untuk menampilkan intermediate outputs
if(~exist('display','var'))
    display = true;
end
%-- memastikan image 2D double matrix
I = im2graydouble(I);

%-- menciptakan Signed Distance Map (SDF) dan mask
phi = mask2phi(init_mask);

%--main loop
for its = 1:max_its

    idx = find(phi <= 1.2 & phi >= -1.2); %Mendapatkan narrow band kurva

    %-- find interior and exterior mean
    upts = find(phi<=0); % interior points
    vpts = find(phi>0); % exterior points
    u = sum(I(upts))/(length(upts)+eps); % interior mean
    v = sum(I(vpts))/(length(vpts)+eps); % exterior mean

    F = (I(idx)-u).^2-(I(idx)-v).^2;
    curvature = get_curvature(phi,idx);

    dphidt = F./max(abs(F)) + alpha*curvature; % penurunan gradien untuk meminimalkan energi

    %-- mempertahankan kondisi CFL
    dt = .45/(max(dphidt)+eps);

    %-- mengembangkan kurva
    phi(idx) = phi(idx) + dt.*dphidt;

```

```

%-- Keep SDF smooth
phi = sussman(phi, .5);

%-- intermediate output
if((display>0)&&(mod(its,20) == 0))
    showCurveAndPhi(I,phi,its);
end
end

%-- final output
if(display)
    showCurveAndPhi(I,phi,its);
end

%-- membuat mask dari SDF
seg = phi<=0; %-- Get mask from levelset

%-----

%-- Menampilkan gambar dengan kurva
function showCurveAndPhi(I, phi, i)
    imshow(I,'initialmagnification',200,'displayrange',[0 255]); hold on;
    contour(phi, [0 0], 'g','LineWidth',4);
    contour(phi, [0 0], 'k','LineWidth',2);
    hold off; title([num2str(i) ' Iterations']); drawnow;

%-- convert mask ke SDF
function phi = mask2phi(init_a)
    phi=bwdist(init_a)-bwdist(1-init_a)+im2double(init_a)-.5;

%-- menghitung curvature sepanjang SDF
function curvature = get_curvature(phi,idx)
    [dimy, dimx] = size(phi);
    [y x] = ind2sub([dimy,dimx],idx);

%-- menghubungkan neighbors
ym1 = y-1; xm1 = x-1; yp1 = y+1; xp1 = x+1;

```

```

%-- cek batas
ym1(ym1<1) = 1; xm1(xm1<1) = 1;
yp1(yp1>dimy)=dimy; xp1(xp1>dimx) = dimx;

%-- index untuk 8 neighbors
idup = sub2ind(size(phi),yp1,x);
iddn = sub2ind(size(phi),ym1,x);
idlt = sub2ind(size(phi),y,xm1);
idrt = sub2ind(size(phi),y,xp1);
idul = sub2ind(size(phi),yp1,xm1);
idur = sub2ind(size(phi),yp1,xp1);
iddl = sub2ind(size(phi),ym1,xm1);
idrr = sub2ind(size(phi),ym1,xp1);

%-- turunan sentral dari SDF at x,y
phi_x = -phi(idlt)+phi(idrt);
phi_y = -phi(iddn)+phi(idup);
phi_xx = phi(idlt)-2*phi(idx)+phi(idrt);
phi_yy = phi(iddn)-2*phi(idx)+phi(idup);
phi_xy = -0.25*phi(iddl)-0.25*phi(idur)...
        +0.25*phi(idrr)+0.25*phi(idul);
phi_x2 = phi_x.^2;
phi_y2 = phi_y.^2;

%-- menghitung curvature (Kappa)
curvature = ((phi_x2.*phi_yy + phi_y2.*phi_xx - 2*phi_x.*phi_y.*phi_xy)./...
            (phi_x2 + phi_y2 +eps).^(3/2)).*(phi_x2 + phi_y2).^(1/2);

%-- Convert citra menjadi grayscale
function img = im2graydouble(img)
[dimy, dimx, c] = size(img);
if(isfloat(img)) % image : double
    if(c==3)
        img = rgb2gray(uint8(img));
    end
else % image : int
    if(c==3)
        img = rgb2gray(img);
    end
end

```

```

    img = double(img);
end

%-- level set reinitialization dengan metode sussman
function D = sussman(D, dt)
    % forward/backward
    a = D - shiftR(D); % backward
    b = shiftL(D) - D; % forward
    c = D - shiftD(D); % backward
    d = shiftU(D) - D; % forward

    a_p = a; a_n = a; % a+ and a-
    b_p = b; b_n = b;
    c_p = c; c_n = c;
    d_p = d; d_n = d;

    a_p(a < 0) = 0;
    a_n(a > 0) = 0;
    b_p(b < 0) = 0;
    b_n(b > 0) = 0;
    c_p(c < 0) = 0;
    c_n(c > 0) = 0;
    d_p(d < 0) = 0;
    d_n(d > 0) = 0;

    dD = zeros(size(D));
    D_neg_ind = find(D < 0);
    D_pos_ind = find(D > 0);
    dD(D_pos_ind) = sqrt(max(a_p(D_pos_ind).^2, b_n(D_pos_ind).^2) ...
        + max(c_p(D_pos_ind).^2, d_n(D_pos_ind).^2)) - 1;
    dD(D_neg_ind) = sqrt(max(a_n(D_neg_ind).^2, b_p(D_neg_ind).^2) ...
        + max(c_n(D_neg_ind).^2, d_p(D_neg_ind).^2)) - 1;

    D = D - dt .* sussman_sign(D) .* dD;

%-- turunan seluruh matrix
function shift = shiftD(M)
    shift = shiftR(M)';

```

```

function shift = shiftL(M)
    shift = [ M(:,2:size(M,2)) M(:,size(M,2)) ];

function shift = shiftR(M)
    shift = [ M(:,1) M(:,1:size(M,2)-1) ];

function shift = shiftU(M)
    shift = shiftL(M)';

function S = sussman_sign(D)
    S = D ./ sqrt(D.^2 + 1);

```

Lampiran 6 : Source Code Luas Area Chan-Vese

```

calc = [];
for l = 1:jumlah_file_nrm1
    I = imread(fullfile(folder_nrm1, file_nrm1(l).name));

    %Wfigure, imshow(I)

    bw1 = im2bw(I);
    bw2 = imfill(bw1, 'holes');

    s = regionprops(bw2, 'centroid', 'area');
    centroids = cat(1, s.Centroid);
    area = cat(1, s.Area);
    sum_area = sum(area);

    hold on
    plot(centroids(:, 1), centroids(:,2), 'r*')

    [B,L] = bwboundaries(bw2, 'noholes');

    [~,jumlah_file_nrm1] = bwlabel(bw2,8);

    for k = 1:jumlah_file_nrm1
        boundary = B{k};
        plot(boundary(:,2), boundary(:,1), 'y', 'LineWidth', 2)
        text(boundary(1,2), boundary(1,1), strcat(['Area = ',num2str(area(k))]), 'Color', 'r', 'FontSize',15, 'FontWeight', 'bold');
    end
    calc = [calc;sum_area];
    cell_areas(l,:) = {file_nrm1(l).name, sum_area};
    hold off
end
end

```


Lampiran 7 : Source Code Canny dan Keliling Citra

```
count = [];  
  
for k = 1:jumlah_file_tbc  
    I = imread(fullfile(folder_tbc, file_tbc(k).name));  
    BW1 = edge(I, 'canny');  
    imshow(BW1);  
    [m,n]= size(BW1);  
    calc = 0;  
  
    for i = 1 : m  
        for j = 1 : n  
            if BW1(i,j) == 1  
                calc = calc+1;  
            %%else  
            end  
        end  
    end  
    count = [count;calc];  
    cell_count(k,:) = {file_tbc(k).name, calc};  
    imwrite(BW1,fullfile(folder_newdit, file_tbc(k).name));  
end
```

Lampiran 8 : Sampel Perhitungan Segmentasi Data Tuberculosis

Citra	Luas	Keliling	Rasio
'Tuberculosis-1.png'	4342	418	0,096269
'Tuberculosis-10.png'	3700	486	0,13135135
'Tuberculosis-100.png'	4265	485	0,1137163
'Tuberculosis-101.png'	4426	513	0,11590601
'Tuberculosis-102.png'	3328	487	0,14633413
'Tuberculosis-103.png'	3356	592	0,17640048
'Tuberculosis-104.png'	3728	468	0,12553648
'Tuberculosis-105.png'	3576	452	0,12639821
'Tuberculosis-106.png'	1712	198	0,11565421
'Tuberculosis-107.png'	2172	300	0,13812155
'Tuberculosis-108.png'	2575	433	0,16815534

'Tuberculosis-109.png'	2158	399	0,18489342
'Tuberculosis-11.png'	2696	410	0,15207715
'Tuberculosis-110.png'	2061	329	0,15963125
'Tuberculosis-111.png'	2303	364	0,15805471
'Tuberculosis-112.png'	2248	351	0,15613879
'Tuberculosis-113.png'	2380	373	0,15672269
'Tuberculosis-114.png'	1877	309	0,1646244
'Tuberculosis-115.png'	3709	480	0,12941494
'Tuberculosis-116.png'	4302	481	0,11180846
'Tuberculosis-117.png'	4235	446	0,10531287
.....
'Tuberculosis-635.png'	2763	383	0,13861744

'Tuberculosis-636.png'	1904	369	0,19380252
'Tuberculosis-637.png'	1389	277	0,19942405
'Tuberculosis-638.png'	2551	334	0,13092905
'Tuberculosis-639.png'	3243	394	0,12149245
'Tuberculosis-64.png'	3315	426	0,12850679
'Tuberculosis-640.png'	2657	386	0,14527663
'Tuberculosis-641.png'	2946	435	0,14765784
'Tuberculosis-642.png'	2804	450	0,16048502
'Tuberculosis-643.png'	3031	409	0,13493896
'Tuberculosis-644.png'	3347	448	0,13385121
'Tuberculosis-645.png'	2630	390	0,14828897
'Tuberculosis-646.png'	2760	401	0,14528986

'Tuberculosis-647.png'	2422	372	0,15359207
'Tuberculosis-648.png'	2615	387	0,14799235
'Tuberculosis-649.png'	2587	383	0,14804793
'Tuberculosis-65.png'	3560	497	0,13960674
'Tuberculosis-650.png'	2764	407	0,14725036

Lampiran 9 : Sampel Perhitungan Segmentasi Data Normal

Citra	Luas	Keliling	Rasio
'IM-0115-0001.png'	539	2823	0,19093163
'IM-0117-0001.png'	473	2097	0,22556032
'IM-0119-0001.png'	491	2269	0,21639489
'IM-0122-0001.png'	375	2067	0,18142235
'IM-0125-0001.png'	495	1990	0,24874372
'IM-0127-0001.png'	414	2297	0,18023509
'IM-0128-0001.png'	506	2629	0,19246862
'IM-0129-0001.png'	507	2346	0,21611253
'IM-0131-0001.png'	415	1797	0,23094046
'IM-0133-0001.png'	378	1837	0,20577028
'IM-0135-0001.png'	475	2810	0,16903915
'IM-0137-0001.png'	455	1975	0,23037975
'IM-0140-0001.png'	503	3270	0,15382263
'IM-0141-0001.png'	434	2137	0,20308844
'IM-0143-0001.png'	455	2562	0,17759563
'IM-0145-0001.png'	459	2145	0,21398601
'IM-0147-0001.png'	540	3063	0,17629775
'IM-0149-0001.png'	431	2542	0,16955153
'IM-0151-0001.png'	380	1925	0,1974026
'IM-0152-0001.png'	432	1956	0,2208589
'IM-0154-0001.png'	414	2513	0,16474333

'IM-0156-0001.png'	432	2504	0,17252396
'IM-0158-0001.png'	439	1740	0,25229885
'IM-0160-0001.png'	417	2103	0,19828816
'IM-0162-0001.png'	418	2481	0,16848045
'IM-0164-0001.png'	586	2830	0,20706714
'IM-0166-0001.png'	473	2402	0,19691923
.....
'NORMAL2-IM-1362-0001.png'	496	3138	0,15806246
'NORMAL2-IM-1365-0001.png'	455	1765	0,25779037
'NORMAL2-IM-1371-0001.png'	362	1294	0,2797527
'NORMAL2-IM-1376-0001.png'	424	2474	0,17138238
'NORMAL2-IM-1379-0001.png'	433	3074	0,14085882
'NORMAL2-IM-1385-0001.png'	466	3680	0,12663043
'NORMAL2-IM-1396-0001.png'	383	1735	0,22074928
'NORMAL2-IM-1400-0001.png'	448	2533	0,17686538

'NORMAL2-IM-1401-0001.png'	414	2331	0,17760618
'NORMAL2-IM-1406-0001.png'	443	2167	0,20443009
'NORMAL2-IM-1412-0001.png'	406	1790	0,22681564
'NORMAL2-IM-1419-0001.png'	452	2470	0,18299595

Lampiran 10 : Source Code Clustering K-Means

```

## Import Library
import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline
from sklearn.cluster import KMeans
from sklearn.preprocessing import StandardScaler

```

```

df=pd.read_excel("citra.xlsx")
df

```

	Images	Result
0	'Tuberculosis-1.png'	0.096269
1	'Tuberculosis-10.png'	0.131351
2	'Tuberculosis-100.png'	0.113716
3	'Tuberculosis-101.png'	0.115906
4	'Tuberculosis-102.png'	0.146334
...
1986	'NORMAL2-IM-1406-0001.png'	0.204430
1987	'NORMAL2-IM-1412-0001.png'	0.226816
1988	'NORMAL2-IM-1419-0001.png'	0.182996
1989	'NORMAL2-IM-1422-0001.png'	0.279835
1990	'NORMAL2-IM-1423-0001.png'	0.224031

1991 rows x 2 columns

```
dfa = df[['Result']]
print('Total Row : ', len(dfa))
dfa
```

Total Row : 1991

	Result
0	0.096269
1	0.131351
2	0.113716
3	0.115906
4	0.146334
...	...
1986	0.204430
1987	0.226816
1988	0.182996
1989	0.279835
1990	0.224031

1991 rows × 1 columns

```

## Feature Scaling P(0), P(1), P(2)
sc_dfa = StandardScaler()
dfa_std = sc_dfa.fit_transform(dfa.astype(float))
dfa_std

array([[ -2.02571313],
       [ -1.06123829],
       [ -1.5460565 ],
       ...,
       [  0.35856163],
       [  3.02084634],
       [  1.48669865]])

```

```

dfa_df = pd.DataFrame(dfa_std, columns = ['Result' ])
dfa_df

```

	Result
0	-2.025713
1	-1.061238
2	-1.546056
3	-1.485858
4	-0.649336
...	...
1986	0.947823
1987	1.563241
1988	0.358562
1989	3.020846
1990	1.486699

1991 rows x 1 columns

```

import sklearn.cluster as cluster
kmeans = cluster.KMeans(n_clusters=2, init="k-means++")
kmeans = kmeans.fit(dfa_df[['Result']])

```

```
df['Clusters'] = kmeans.labels_
df
```

	Images	Result	Clusters
0	'Tuberculosis-1.png'	0.096269	0
1	'Tuberculosis-10.png'	0.131351	0
2	'Tuberculosis-100.png'	0.113716	0
3	'Tuberculosis-101.png'	0.115906	0
4	'Tuberculosis-102.png'	0.146334	0
...
1986	'NORMAL2-IM-1406-0001.png'	0.204430	1
1987	'NORMAL2-IM-1412-0001.png'	0.226816	1
1988	'NORMAL2-IM-1419-0001.png'	0.182996	1
1989	'NORMAL2-IM-1422-0001.png'	0.279835	1
1990	'NORMAL2-IM-1423-0001.png'	0.224031	1

1991 rows x 3 columns

Lampiran 10 : Riwayat Hidup

RIWAYAT HIDUP**A. Identitas Diri**

1. Nama Lengkap : Fayza Nayla Riyana Putri
2. Tempat & Tanggal Lahir : Bekasi, 24 November 2001
3. Alamat Rumah : Villa Gading Baru C1/23
4. HP : 085867853396
5. Email : fayzanaylaa@gmail.com

B. Riwayat Pendidikan

1. Sekolah Dasar Negeri Harapan Baru III Kota Bekasi
2. Sekolah Menengah Pertama Negeri I Kota Bekasi
3. Sekolah Menengah Atas Negeri II Kota Bekasi