

# Identifikasi Penyakit Kanker Paru-Paru Dengan Menggunakan Metode GLCM dan Convolutional Neural Network

Karimi

Fakultas Ilmu Komputer dan Teknologi Informasi, Program Studi Teknik Informatika, Universitas Budi Darma, Medan, Indonesia  
Email: kariminosra135@gmail.com  
Email Penulis Korespondensi: kariminosra135@gmail.com

**Abstrak**—Kanker adalah pertumbuhan dan penyebaran yang tidak terkontrol dari sel. Salah satu jenis kanker yang paling mematikan adalah kanker paru-paru. Kanker paru-paru adalah pertumbuhan sel kanker yang tidak terkontrol dalam jaringan paru. Kanker paru-paru merupakan penyebab kematian utama dari semua kematian akibat kanker baik pada pria maupun wanita. Kanker paru di diagnosa melalui *chest X-ray* (CXR) atau lebih dikenal dengan foto rontgen. Pemeriksaan radiologi CXR ini sangat membantu proses diagnosis dan identifikasi medis pada penyakit paru. Pengetahuan masyarakat yang masih minim dalam membaca hasil foto rontgen menyebabkan dibutuhkan tenaga ahli seperti dokter atau tenaga medis lain untuk membacanya secara manual. Membaca secara manual masih memungkinkan terjadinya kesalahan dalam pendiagnosaan penyakit. Penegakan diagnosis penyakit ini membutuhkan keterampilan dan sarana yang baik. Karena apabila tidak segera ditangani, bisa menyebar dan bermetastasis dan akhirnya meningkatkan derajat keparahan. Metode yang diajukan pada penelitian ini adalah GLCM dan *Convolutional Neural Network* untuk mengidentifikasi citra adalah normal atau kanker. Tahapan yang dilakukan pada identifikasi ini adalah *pre-processing*, *segmentation* dengan *k-means clustering*, *feature extraction* dengan *Gray Level Co-Occurrence Matrix* dan *identification*. Pada penelitian ini ditunjukkan bahwa metode yang diajukan mampu mengidentifikasi kanker dengan akurasi 75%, sensitifitas 75% dan spesifisitas 75%.

**Kata Kunci:** Kanker Paru; Gray Level Co-Occurrence Matrix; K-Means Clustering; Convolutional Neural Network .

**Abstract**—Cancer is the uncontrolled growth and spread of cells. One of the deadliest types of cancer is lung cancer. Lung cancer is the uncontrolled growth of cancer cells in lung tissue. Lung cancer is the leading cause of death of all cancer deaths in both men and women. Lung cancer is diagnosed through a chest X-ray (CXR) or better known as an X-ray. This CXR radiological examination is very helpful in the process of medical diagnosis and identification of lung disease. The lack of public knowledge in reading X-ray results requires experts such as doctors or other medical personnel to read them manually. Reading manually still allows for errors in diagnosing diseases. Diagnosing this disease requires good skills and facilities. Because if it is not treated immediately, it can spread and metastasize and eventually increase the degree of severity. The methods proposed in this study are GLCM and Convolutional Neural Networks to identify whether an image is normal or cancerous. The steps involved in this identification are pre-processing, segmentation with k-means clustering, feature extraction with Gray Level Co-Occurrence Matrix and identification. In this study it was shown that the proposed method was able to identify cancer with an accuracy of 75%, a sensitivity of 75% and a specificity of 75%.

**Keywords:** Lung Cancer; Gray Level Co-Occurrence Matrix; K-Means Clustering; Convolutional Neural Networks.

## 1. PENDAHULUAN

Pada saat sekarang ini ilmu pengetahuan dan teknologi informasi semakin mengalami perkembangan sangat pesat dan menjadi salah satu bagian penting dalam mengatasi masalah yang terjadi di masyarakat. Teknologi informasi telah diterapkan di berbagai bidang seperti e-learning pada Pendidikan, e-commerce di bidang bisnis, ATM dibidang perbankan, e-government pada pemerintahan dan lain sebagainya. Pada bidang kesehatan teknologi informasi telah diterapkan pada pengolahan citra digital telah dilakukan sejak ditemukan Tomografi Terkomputerisasi (Computerized tomography/CT) pada tahun 1970-an. Pengolahan citra digital dapat digunakan untuk deteksi tumor atau kanker rahim, identifikasi penyakit hati, identifikasi penyakit paru-paru, identifikasi penyakit tulang, klasifikasi gigi, dan analisis citra mikroskopis.

Kanker adalah pertumbuhan dan penyebaran yang tidak terkontrol dari sel (WHO, 2013). Salah satu jenis kanker yang paling mematikan adalah kanker paru-paru. Kanker paru-paru adalah pertumbuhan sel kanker yang tidak terkontrol dalam jaringan paru. Menurut World Health Organization (WHO), kanker paru merupakan penyebab kematian utama dari semua kematian akibat kanker baik pada pria maupun wanita. Menurut CDC (2010), sebanyak 205.974 orang menderita kanker paru di Amerika (110.190 pria dan 95.784 wanita) dan sebanyak 158.081 orang meninggal karena penyakit ini (87.694 pria dan 70.387 wanita) sedangkan di Indonesia, angka kejadian kanker bronkus dan paru pada pasien rawat inap sebesar 5,8% dari seluruh jenis kanker (Depkes, 2007).

Berdasarkan klasifikasi tersebut, penyakit kanker paru di diagnosa melalui chest X-ray (CXR) atau lebih dikenal dengan foto rontgen dimanfaatkan di berbagai aspek kehidupan masyarakat. Penggunaan X-ray berawal dari penemuan Wilhelm Röntgen pada tahun 1895. Pemeriksaan radiologi CXR ini sangat membantu proses diagnosis dan identifikasi medis pada penyakit paru. Akan tetapi saat membaca hasil foto rontgen pengetahuan masyarakat yang minim dalam membaca hasil rontgen, sehingga masih dibutuhkan tenaga ahli seperti dokter atau tenaga medis lain untuk membacanya, Kanker paru-paru memerlukan penanganan dan tindakan yang cepat dan terarah. Penegakan diagnosis penyakit ini membutuhkan keterampilan dan sarana yang baik. Karena apabila tidak segera ditangani, bisa menyebar dan bermetastasis dan akhirnya meningkatkan derajat keparahan. (Pedoman Diagnosis dan Penatalaksanaan di Indonesia, 2003).

Penelitian dengan memanfaatkan CXR sudah dilakukan yaitu untuk mengidentifikasi kelainan pada paru dengan judul *Detection of Lung Cancer Cells using Image Processing Techniques* dengan menggunakan metode median filtering, threshold segmentation, Watershed Algorithm dan Morphological Operations (Pratapl, 2016).

Pada tahun 2016 Poornimadevi, C.S dan Helen, S.C melakukan penelitian dengan judul *Automatic Detection of Pulmonary Tuberculosis Using Image Processing Technique*. Penelitian ini memiliki beberapa tahapan, yaitu Preprocessing, Registration Based Image Segmentation, Watershed Segmentation, Threshold dan Active Contour. Hasil dari penelitian ini memiliki tingkat akurasi sebesar 60%.

Pada penelitian yang berjudul *Hypertensive Retinopathy Identification Through Retinal Fundus Image Using Backpropagation Neural Network* pada tahun 2017 oleh Mohammad Fadly Syahputra, C. Amalia, Romi Fadillah Rahmat dan Ulfi Andayani. Dalam penelitian tersebut Backpropagation secara efisien dapat mengidentifikasi penyakit Hypertensive Retinopathy dengan akurasi sebesar 95%.

Pada tahun 2017, Amalia Rahmi menggunakan Metode Convolutional neural network dalam klasifikasi pendarahan otak. Dalam Penelitian ini klasifikasi pendarahan otak dilakukan dengan menggunakan citra CT scan otak dan menghasilkan tingkat akurasi 88%.

Berdasarkan hasil penelitian tersebut, dengan tingkat akurasi yang tinggi yang didapatkan oleh metode Backpropagation dalam melakukan identifikasi terhadap suatu objek, penulis mengajukan metode Convolutional neural network Algoritma Convolutional neural network merupakan salah satu jenis artificial neural network (ANN) dengan beberapa unit hidden layer yang berada di antara input dan output. Convolutional neural network sebelumnya juga telah dipakai untuk diagnosa retinoblastoma (widya, 2017)..

## 2. METODOLOGI PENELITIAN

Dalam melakukan penelitian agar mendapat hasil seperti yang diharapkan, maka diperlukan tahapan penelitian. Dimana tahapan penelitian yang akan dilakukan digambarkan.

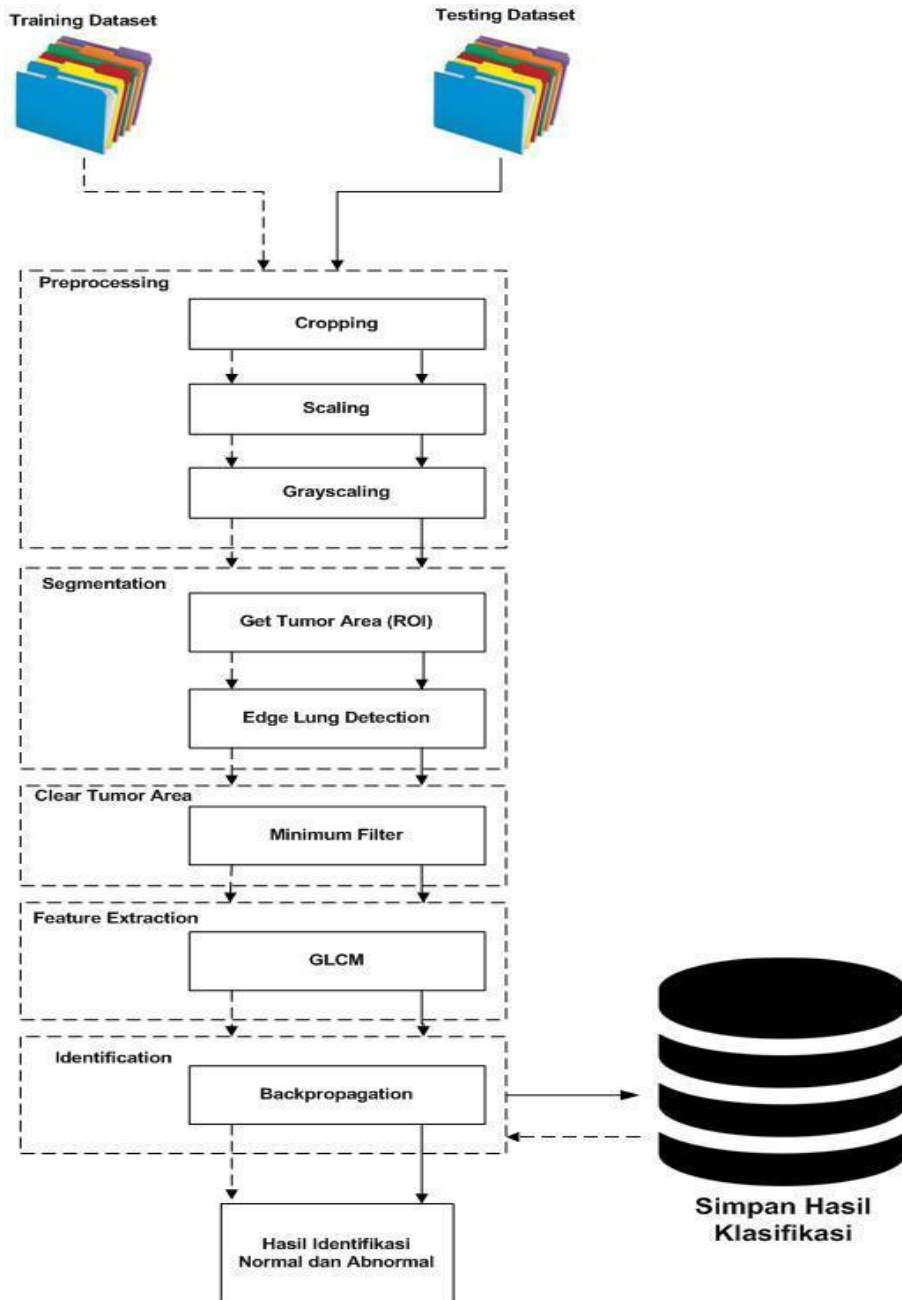
- a. Studi Lapangan  
Merupakan salah satu metode pengumpulan data dalam kualitatif yang tidak mengetahui mendalam akan literatur dari kemampuan yang digunakan tertentu dari pihak penelitian
- b. Studi Pustaka  
Penyusun mempelajari buku-buku, serta internet yang berkaitan dengan penelitian pada metode skripsi ini.
- c. Metode Pengumpulan Data  
Cara ini digunakan untuk penelitian yang dilakukan secara langsung kelapangan dengan konsultasi kepada dokter.
- d. Analisa Masalah  
Data yang diperoleh studi pustaka dan wawancara di analisis untuk merencanakan pembuatan sistem yang sesuai dengan kebutuhan.
- e. Perancangan Sistem  
Pada tahap ini dilakukan perancangan flowchart dan interface dari sistem yang akan dibuat sesuai dengan tujuan dan batasan masalah yang ada
- f. Pengujian Sistem  
Sistem akan diuji dengan memberi berbagai inputan yang sedikit berbeda dengan pola sebelumnya untuk melihat beberapa akurat sistem yang dibuat sesuai dengan metode ini.
- g. Implementasi Sistem  
Pada tahap ini sistem akan diimplementasikan dengan bahasa pemrograman.
- h. Dokumentasi  
Pada tahap ini dilakukan penyusunan laporan dari hasil analisis dan perancangan sistem dalam formal penulisan penelitian.

## 3. HASIL DAN PEMBAHASAN

### 3.1 Pembahasan

Penelitian ini akan membahas tentang analisis dan perancangan dalam aplikasi identifikasi penyakit kanker paru-paru. Tahap pertama yaitu analisis data yang digunakan, analisis dengan menggunakan beberapa tahapan pengolahan citra yang digunakan, kemudian implementasi metode GLCM dan Convolutional neural network. Pada tahapan selanjutnya yaitu dilakukan perancangan tampilan antarmuka sistem. Metode yang diajukan untuk mengidentifikasi penyakit kanker paru-paru yang disebabkan pertumbuhan sel yang tidak terkendali yang ditandai berupa bercak berwarna putih terdiri dari beberapa tahapan. Tahapan-tahapan tersebut dimulai dari pengumpulan data citra Normal, dan kanker paru-paru (Abnormal) yang akan digunakan untuk citra latih dan dan citra uji, tahap preprocessing yang terdiri atas cropping untuk memotong citra pada daerah paru dan merubah ukuran citra menjadi berdimensi lebih kecil agar memudahkan proses pengolahan citra selanjutnya, Scaling yang digunakan untuk menentukan ukuran piksel yang digunakan dalam pemrosesan citra paru. Kemudian dilakukan penyeragaman gambar keabuan dengan menggunakan Grayscale. Tahapan selanjutnya yaitu proses segmentasi menggunakan region of interest dimana daerah yang merupakan kanker

akan didapatkan kemudian proses edge detection dilakukan untuk mendapatkan tepi dari bagian kanker yang sudah didapatkan pada proses region of interest (ROI). setelah itu dilakukan proses erosi dan minimum filter untuk memperjelas area tumor, kemudian ekstraksi fitur dilakukan dengan menggunakan metode GLCM, hingga pada akhirnya implementasi metode Convolutional neural network dilakukan untuk mengidentifikasi kanker paru-paru. Adapun tahapan-tahapan diatas dapat dilihat dalam bentuk arsitektur umum pada Gambar 1.



**Gambar1.** Proses Identifikasi kanker paru-paru

### 3.2 Dataset

Data citra yang digunakan dalam penelitian ini adalah citra Chest Xray (CXR) yang diperoleh dari <http://www.jsrt.or.jp> JSRT (Japanese society of radiology). Tujuan dari Japanese Society of Radiological Technology (JSRT) adalah untuk berkontribusi dalam pengembangan dan peningkatan ilmu pengetahuan dan teknologi radiologi dalam hal aktivitas penelitian, pertukaran intelektual, dan komunikasi ilmiah dengan pihak-pihak terkait lainnya. Data citra yang diperoleh adalah citra paru-paru dada normal dan citra kanker paru-paru. Data tersebut akan digunakan sebagai data pelatihan dan data pengujian.

### 3.3 Pre-processing

Tahapan ini merupakan tahap pengolahan citra yang bertujuan untuk menghasilkan citra yang lebih baik untuk diproses pada tahapan selanjutnya. Tahapan preprocessing ini terdiri dari, cropping, Scaling dan Grayscale.

### 3.3.1 Cropping

Tahapan pertama yang dilakukan adalah cropping. Tahapan ini bertujuan untuk memotong citra pada daerah paru-paru agar memudahkan proses pengolahan citra selanjutnya. Proses cropping dilakukan dengan manual menggunakan photoshop dari citra berukuran 2010 x 2010 piksel kemudian dipotong pada bagian daerah paru-paru menjadi berukuran 1760 x 1760 piksel.

### 3.3.2 Scaling

Proses scaling dilakukan pada citra hasil cropping yaitu dengan mengubah dimensi citra menjadi berukuran 320 x 320 piksel . dibutuhkan untuk mengatur ukuran piksel pada citra. Semakin banyak jumlah piksel maka akan semakin banyak waktu untuk proses pengolahan citra.

### 3.3.3 Grayscale

Tahapan akhir pada proses preprocessing adalah grayscale. Dimana pada tahapan ini bertujuan untuk penyeragaman warna keabuan pada citra yang akan diproses. Pada citra asli terlihat warna keabuan tidak merata. Citra Grayscale dapat dilihat pada

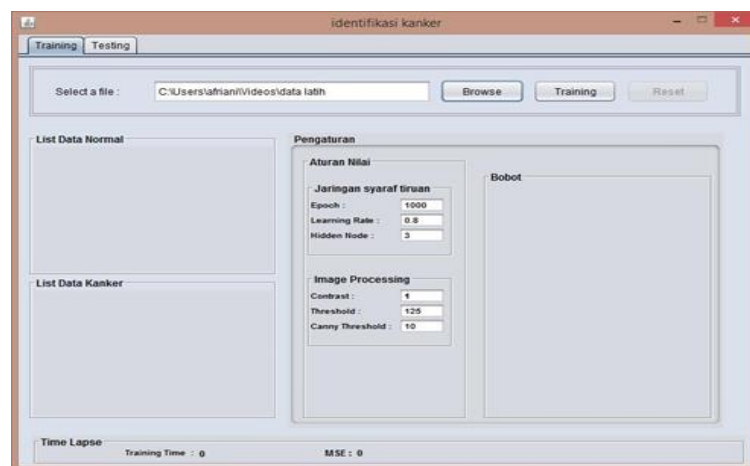


Gambar 2. citra grayscale

## 3.3 Implementasi perancangan antarmuka

Implementasi perancangan antarmuka berdasarkan rancangan yang telah dilakukan pada penelitian ini adalah sebagai berikut.

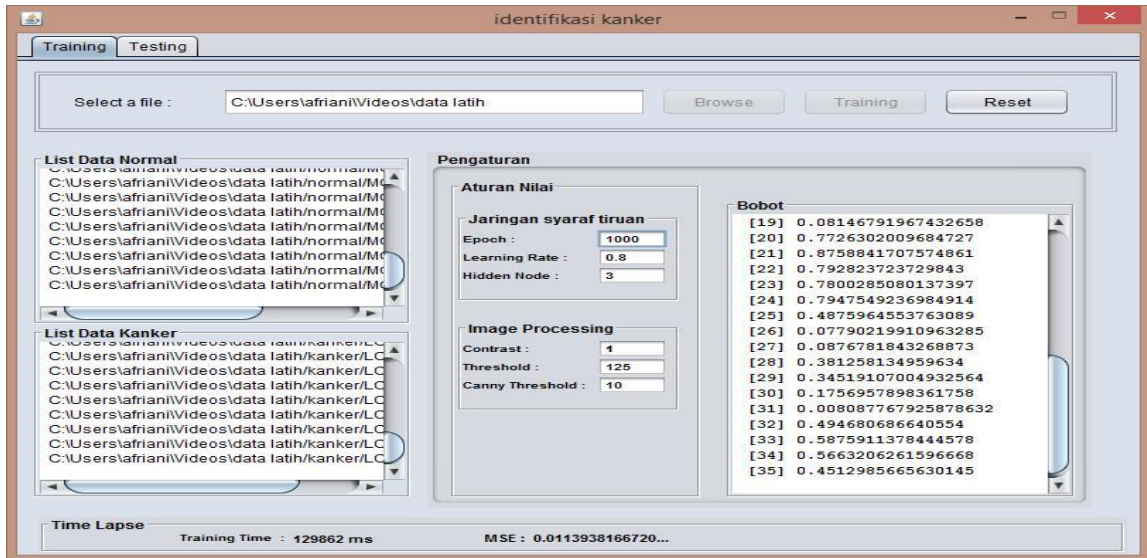
Tampilan awal yang aplikasi seperti yang ditunjukkan pada gambar 4.1. memiliki satu buah tombol, yaitu tombol “Start” yang digunakan untuk masuk ke tampilan utama dalam sistem. Tampilan utama sistem memiliki dua buah *tab* yang masing-masing memiliki fungsi tunggal untuk *training* dan *testing*. Tampilan tab *training* seperti yang ditunjukkan oleh Gambar 4.2 memiliki tiga buah tombol yaitu “Browse”, “Training”, dan “Reset”. Tombol *Browse* digunakan untuk memilih *folder* citra CXR yang akan diproses. Tampilan saat tombol “Browse” dipilih dapat dilihat pada Gambar 3



Gambar 3. Tampilan tab *training* aplikasi setelah citra dipilih

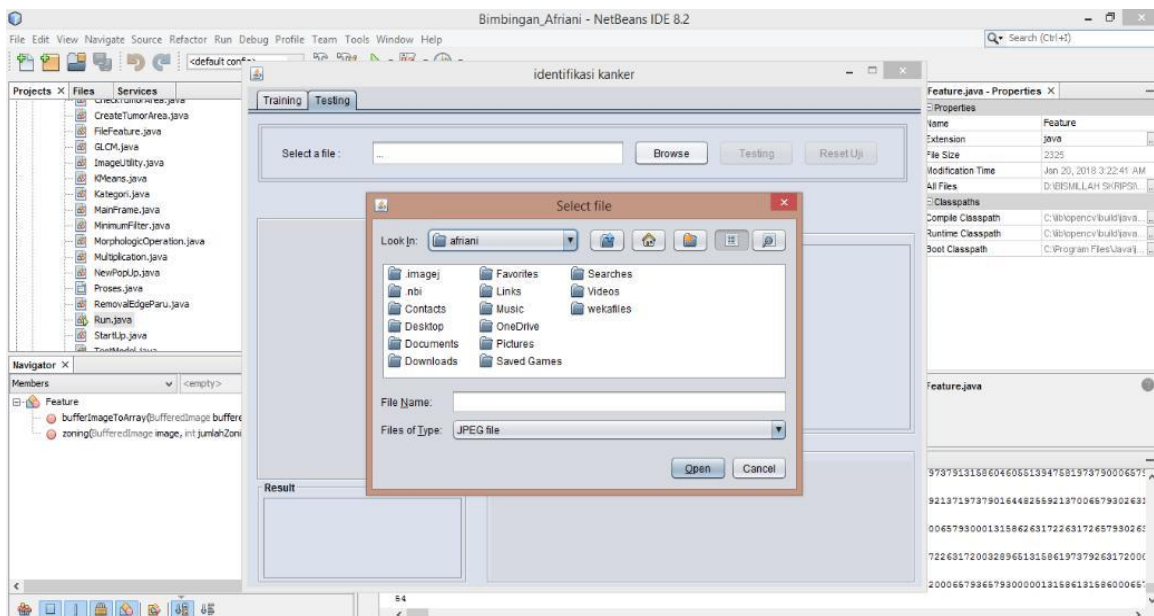
Tombol “Training” yang sudah aktif dapat digunakan untuk menjalankan proses *pre-processing*, *segmentation*, *feature extraction* dan data *training* sekaligus dengan menggunakan satu tombol. Hasil dari setiap tahapan proses disimpan dalam *folder* Identifikasi Penyakit yang akan otomatis ter-*create* pada saat proses *start* berlangsung. Hasil

pada tahapan *pre-processing* data *training* ini akan disimpan pada *sub-folder* images yang terdapat didalam *sub-folder* train. Sedangkan, hasil ekstraksi fitur yang diperoleh akan disimpan dalam bentuk file teks di dalam *sub-folder* text dengan nama train.txt. Setelah proses *training dataset* selesai dilakukan tombol “Reset” akan aktif . Tombol “Reset” dapat digunakan untuk mengembalikan sistem seperti pada saat sistem dimulai. Sehingga *user* dapat melakukan *input* data yang berbeda. Hasil yang akan ditampilkan pada sistem setelah proses *training* selesai adalah seperti yang ditunjukkan pada Gambar 4.



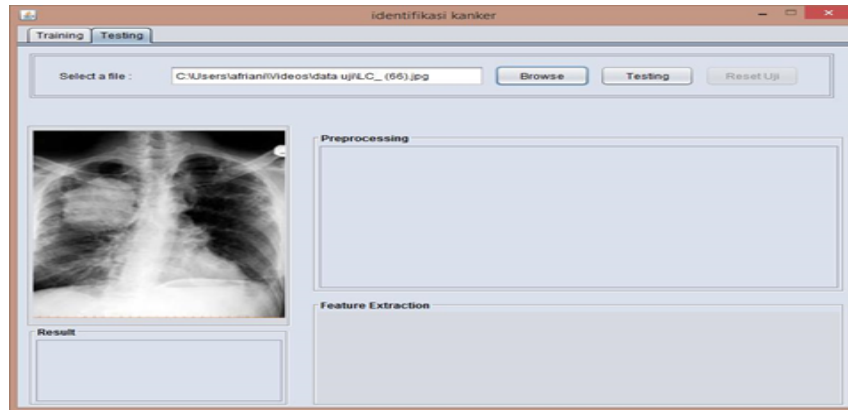
**Gambar 4.** Tampilan awal aplikasi pada hasil *training dataset*

Selanjutnya, setelah proses data *training* selesai maka tombol “Browse” pada *tab* testing akan aktif. Tombol “Browse” pada *tab testing* memiliki fungsi yang sama dengan tombol “Browse” pada *tab training*. Tombol ini digunakan untuk mencari *file* input citra CXR dengan ekstensi .jpg atau jpeg. Tampilan pada saat tombol “Browse” dipilih ditunjukkan pada Gambar 5.



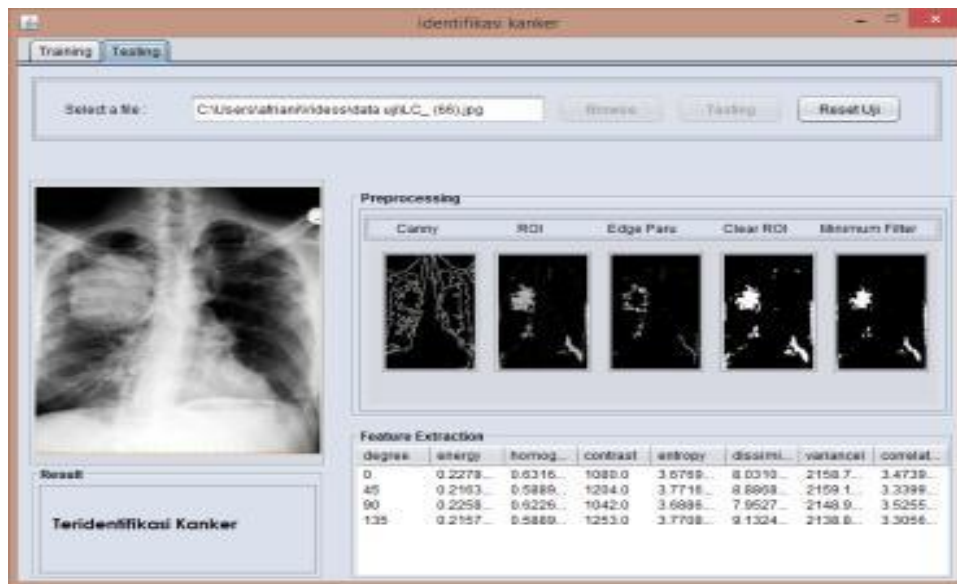
**Gambar 5.** Tampilan pada *tab testing* saat tombol *browse* dipilih

*File* yang telah dipilih akan ditampilkan pada panel images untuk memberikan gambaran perbedaan antara input, proses dan hasil yang diperoleh. Tampilan setelah *file* input dipilih ditunjukkan pada Gambar 6. Pada tahap ini tombol “testing” akan diaktifkan.



**Gambar 6.** Tampilan setelah *file input* dipilih

Beberapa proses seperti *pre-processing*, *segmentation*, *feature extraction* dan *testing* data input dilakukan secara bersamaan saat tombol “*Testing*” dipilih. Hasil dari proses *pre-processing* ditampilkan pada panel Preprocessing dan hasil dari *feature extraction* diletakkan pada panel Feature Extraction yang ditampilkan dalam bentuk tabel seperti yang ditunjukkan pada Gambar 7. Setiap citra yang berada pada panel Preprocessing memiliki fasilitas *zoom* yang berfungsi untuk memperbesar tampilan citra.



**Gambar 7.** Tampilan hasil proses *testing*

## 4. KESIMPULAN

Kesimpulan yang dapat diambil berdasarkan hasil pengujian sistem Identifikasi penyakit kanker paru-paru dengan menggunakan GLCM dan Convolutional Neural Network. Dimana metode Convolutional Neural Network mampu melakukan Identifikasi penyakit kanker paru-paru melalui citra CXR dengan tingkat akurasi 75%, sensitivity 75% dan specificity 75%. Hidden Layer sangat mempengaruhi akurasi. Setelah melalui beberapa pengujian, semakin kecil nilai hidden Layer maka semakin kecil pula tingkat akurasi yang didapatkan. sebaliknya semakin besar nilai Hidden Layer maka semakin besar tingkat akurasi yang didapatkan. Kesalahan identifikasi dipengaruhi oleh fitur yang didapatkan pada hasil segmentasi citra yang mirip antara dua kategori citra CXR. Kesimpulan yang dapat diambil berdasarkan hasil pengujian sistem Identifikasi penyakit kanker paru-paru dengan menggunakan GLCM dan Convolutional Neural Network. Dimana metode Convolutional Neural Network mampu melakukan Identifikasi penyakit kanker paru-paru melalui citra CXR dengan tingkat akurasi 75%, sensitivity 75% dan specificity 75%. Hidden Layer sangat mempengaruhi akurasi. Setelah melalui beberapa pengujian, semakin kecil nilai hidden Layer maka semakin kecil pula tingkat akurasi yang didapatkan. sebaliknya semakin besar nilai Hidden Layer maka semakin besar tingkat akurasi yang didapatkan. Kesalahan identifikasi dipengaruhi oleh fitur yang didapatkan pada hasil segmentasi citra yang mirip antara dua kategori citra CXR.

## REFERENCES

[1] Centers for Disease Control and Prevention, 2010. Lung Cancer Statistics. Available

- at:<http://www.cdc.gov/cancer/lung/statistics/index.html>. (Accessed on 29 April 2018).
- [2] Charles S. Dela Cruz, Lynn T.Tannoue, Richard A. Matthay 2011. Lung cancer; Epidemiology, Etiology and Prevention. *Clinic Chest Med.*2011 Dec; 32(4)
  - [3] Churg M. A., Myers L. J., et al, 2005. Etiology of Lung Cancer. *Thurlbeck's*
  - [4] Pathology of the Lung. Published by Thieme Medical Publishers, Inc, 415 . Diananda, Rama. 2009. *Mengenal Seluk Beluk Kanker*. Penerbit KataHati:Yogyakarta.
  - [5] Iliis, H., 2006. *Clinical Anatomy A Revision and applied anatomy for clinical students.* 11th ed. Blackwell publishing.
  - [6] Gonzales, R.C. & Woods, R.E. 2008. *Digital Image Processing*. Prentice Hall : New Jersey.
  - [7] Haralick, R.M., Shanmugam, K. & Dinstein, I. 1973. Textural features for image classification. *IEEE Trans on Systems, Man and Cybernetics Vol.SMC-3*, pp. 610 – 621.
  - [8] Khobragade, S., Tiwari, A, Pati. C.Y. & Narke, V.2016. *Automatic Detection of Major Lung Diseases Using Chest Radiographs and Classification by Feed-forward Artificial Neural Network*, 2007.