PROPOSAL TESIS

DETEKSI DINI PENYAKIT KANKER PAYUDARA BERDASARKAN HASIL CITRA MAMOGRAFI MENGGUNAKAN FUZZY NEURAL NETWORK

OLEH :

KHAIRUNNAZRI

Tesis diajukan sebagai salah satu syarat

untuk memperoleh gelar Magister

Komputer

PROGRAM PASCASARJANA

MAGISTER TEKNIK INFORMATIKA

UNIVERSITAS DIAN NUSWANTORO

SEMARANG

2016

BAB I

PENDAHULUAN

# Latar Belakang

* + 1. **General**

Kanker payudara merupakan tumor ganas yang dimulai dalam sel-sel payudara [1] yang merupakan salah satu penyebab kematian utama di seluruh dunia setiap tahunnya [2].

* + 1. **Domain**

Domain dalam penelitian ini adalah Image Processing dan Image Segmentasi hasil Mamografi.

* + 1. **Masalah**

Masalah yang mendasari penelitian ini adalah karna besarnya tingkat kematian yang disebabkan oleh panyakit kanker payudara.

* + 1. **Diskusi**

Dalam bidang kedokteran, Radiolog sering mengalami kesulitan dalam mengamati citra mentah hasil Mammografi, karena citra yang dihasilkan mempunyai derajat ke abu-abuan sehingga sulit melihat jelas area pengenalan bagian kanker. Untuk membantu mengatasi hal tersebut perlu dilakukan operasi Pengolahan Citra. Salah satu operasi Pengolahan Citra yaitu dengan Mensegmentasi citra hasil Mamografi. Segmentasi Mamografi pada citra tersebut merupakan proses memperjelas dan mempertajam ciri dari citra agar citra lebih mudah dianalisa.

* + 1. **Variabel/ Parameter/ Indikator/ Faktor/ Aspek**

Dane Kurnia Putra Jurusan Teknik Elektro UNDIP. Dalam penelitiannya yang berjudul mengidentifikasi keberadaan kanker payudara pada Citra Mammografi menggunakan metode Wavelet Haar dapat menghasilkan kualitas gambar yang mampu mengidentifikasi citra sebesar 86% dengan rasio kesalahan 14% [3].

Lestari Handayani Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Pekanbaru. Dalam penelitiannya yang berjudul Segmentasi Mamografi Kanker Payudara Dengan Algoritma Expectation Maximization Segmentation (Em-Segmentation) dapat diambil kesimpulan bahwa dengan menggunakan Algoritma Expectation Maximization Segmentation bisa membagi citra menjadi beberapa Cluster yang dapat menunjukkan area kanker pada citra mammografi [4].

Fourina Ayu Novianti Jurusan Statistika, Fakultas MIPA, Institut Teknologi Sepuluh Nopember (ITS). Dalam penelitiannya yang berjudul Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi. Dapat mendiagnosa letak abnormal kanker payudara dengan ketepatan klasifikasi sebesar 94,34% [5].

* + 1. **Spesifik**

Pencitraan medis berhubungan dengan teknik dan proses untuk membuat citra dari tubuh manusia untuk keperluan medis menemukan, memeriksa, dan mendiagnosis penyakit. Dalam konteks yang sempit, pencitraan medis sering kali disamakan dengan Radiologi. Salah satu bagian dari Radiologi adalah Mamografi, yang mempelajari mengenai proses pencitraan Payudara secara Radiologi kemudian pemeriksaan citra tersebut digunakan untuk mendiagnosis penyakit Kanker Payudara. Mammografi merupakan proses pemeriksaan [Payudara](https://id.wikipedia.org/wiki/Payudara) manusia menggunakan [sinar-X](https://id.wikipedia.org/wiki/Sinar-X) dosis rendah [6] yang memberikan gambaran tentang jaringan lunak pada Payudara. Pemeriksaan ini berguna untuk membantu mendeteksi masalah atau penyakit pada Payudara. Setelah dilakukan Mamografi terhadap pasien, hasil Mamografi tersebut harus dilakukan pengujian laboraturium untuk mendeteksi pengenalan area Kanker.

# Rumusan Masalah

Berdasarkan latar belakang masalah diatas, maka dapat dirumuskan permasalahan dalam penelitian ini adalah segmentasi citra hasil Mamografi untuk mendeteksi Kanker Payudara.

# Tujuan Penelitian

Tujuan dari penelitian ini adalah untuk mengurangi angka kematian penyebab kanker payudara dengan cara deteksi dini Kanker Payudara melalui citra dari hasil Mamografi.

# Kontribusi Penelitian

Kontribusi dalam penelitian ini adalah untuk Ilmu Pengetahuan khususnya di bidang Kedokteran yang menangani Radiologi.

BAB II

TINJAUAN PUSTAKA

# Penelitian Yang Relevan

Banyak penelitian yang sudah dilakukan oleh peneliti dalam pendeteksian kanker payudara salah satunya adalah Supriyadi Pascasarjana Teknik Informatika Universitas Dian Nuswantoro Semarang yang memprediksi kanker payudara menggunakan Artificial Neural Network. Dalam penelitiannya Supriyadi hanya menguji keakuratan Artificial Neural Network untuk memprediksi kanker payudara dari penelitian sebelumnya yang menggunakan C4.5 dan hasilnya Artificial Neural Network bisa meningkatkan akurasi prediksi sekitar 1,42%. Dane Kurnia Putra Jurusan Teknik Elektro UNDIP. Juga melakukan penelitian yang berjudul mengidentifikasi keberadaan kanker payudara pada Citra Mammografi menggunakan metode Wavelet Haar dapat menghasilkan kualitas gambar yang mampu mengidentifikasi citra sebesar 86% dengan rasio kesalahan 14%. Lestari Handayani Jurusan Teknik Informatika Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Pekanbaru. Dalam penelitiannya yang berjudul Segmentasi Mamografi Kanker Payudara Dengan Algoritma Expectation Maximization Segmentation (Em-Segmentation) dapat diambil kesimpulan bahwa dengan menggunakan Algoritma Expectation Maximization Segmentation bisa membagi citra menjadi beberapa Cluster yang dapat menunjukkan area kanker pada citra mammografi. Fourina Ayu Novianti Jurusan Statistika, Fakultas MIPA, Institut Teknologi Sepuluh Nopember (ITS). Dalam penelitiannya yang berjudul Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi. Dapat mendiagnosa letak abnormal kanker payudara dengan ketepatan klasifikasi sebesar 94,34%.

|  |  |  |  |
| --- | --- | --- | --- |
| **Penulis/Judul/Tahun** | **Masalah/Research Problem** | **Pendekatan** | **Hasil** |
| Supriyadi “Prediksi Penyakit Kanker Payudara Menggunakan Artificial Neural Network” 2015. | Akurasi Algoritma C4.5 Masih dirasa kurang maksimal | Artificial Neural Network | 95,57% |
| Dane “Identifikasi Keberadaan Kanker Pada Citra Mammografi Menggunakan Metode Wavelet Haar”. | Mengidentifikasi keberadaan kanker payudara pada citra mammografi. | Wavelet Haar | 86% |
| Lestari “Segmentasi Mamografi Kanker Payudara Dengan Algoritma Expectation Maximization Segmentation (Em Segmentation)” 2014. | Segmentasi Citra Mammografi Kanker Payudara | Expectation Maximization Segmentation |  |
| Fourina ”Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi” 2012. | Analisis Diagnosis Pasien Kanker Payudara Berdasarkan Hasil Mamografi | Support Vector Machine (SVM) | 94,34% |

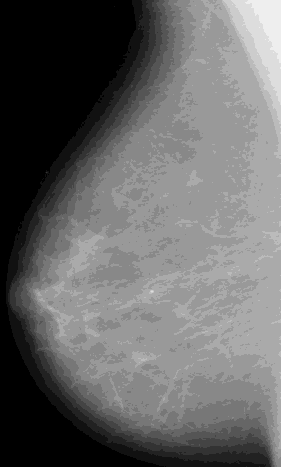
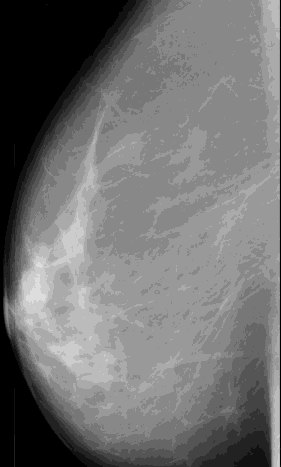
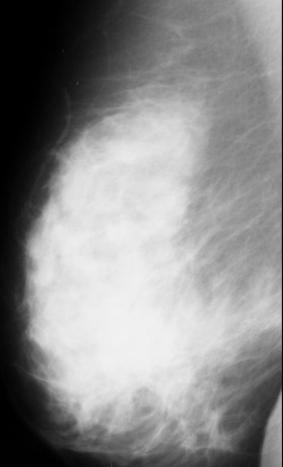
# Landasan Teori

**2.2.1 Fuzzy Neural Network**

Pada tahun 1992, Pal dan Mitra memperkenalkan klasifikasi fuzzy ke dalam algoritma backpropagation yang disebut Fuzzy Neural Network (FNN). Pada FNN, parameter-parameter yang dimiliki Artificial Neural Network (ANN) yang biasanya disajikan secara crisp, dapat diganti dengan parameter-parameter fuzzy [7]. Target pembelajaran pada ANN berupa nilai crisp, tetapi pada FNN target pembelajaran crisp diubah menjadi target fuzzy melalui fuzzifikasi.

**2.2.2 Pengolahan Citra Digital**

Citra digital merupakan citra yang telah disimpan dalam bentuk file sehingga dapat diolah menggunakan komputer [8]. Salah satu jenis citra digital adalah citra grayscale. Citra grayscale yang digunakan pada tugas akhir ini adalah citra mammografi. Citra mammografi yang dianalisis pada tugas akhir ini digolongkan menjadi tiga kelas, yaitu: citra mammografi normal (Gambar 2.2.2 (a)), citra mammografi tumor jinak (Gambar 2.2.2 (b)), dan citra mammografi tumor ganas (Gambar 2.2.2 (c)).

(a) (b) (c)

# Kerangka Pemikiran

**2.3.1 Extraksi Gambar**

Ektraksi gambar adalah suatu teknik pengambilan parameter statistic dari sebuah gambar. Proses ekstraksi citra dapat dilakukan dengan Gray Level Coocurrence Matrix (GLCM). Fungsi bawaan pada MATLAB dapat digunakan untuk menganalisa entropy, kontras, korelasi, energi, dan homogenitas.

**2.3.2 Menentukan Variabel Input dan Variabel Output**

Variabel input model FNN adalah parameter-parameter statistik dari hasil ektraksi citra. Banyaknya variabel input menentukan banyaknya neuron pada lapisan input. Variabel output model FNN adalah klasifikasi dari citra. Banyaknya variabel output menentukan banyaknya neuron pada lapisan output.

**2.3.3 Pembagian Data**

Data yang diperoleh dari hasil ektraksi citra dibagi menjadi 2 yaitu data training dan data testing. Data training digunakan untuk mencari model terbaik, sedangkan data testing digunakan untuk menguji ketepatan model hasil data training.

**2.3.4 Fuzzifikasi**

Sebelum dilakukan pembelajaran, terlebih dahulu dilakukan pengubahan nilai target crisp menjadi nilai target fuzzy. Nilai target fuzzy ini digunakan sebagai target pembelajaran dalam jaringan feedforward. Proses ini disebut fuzzifikasi.

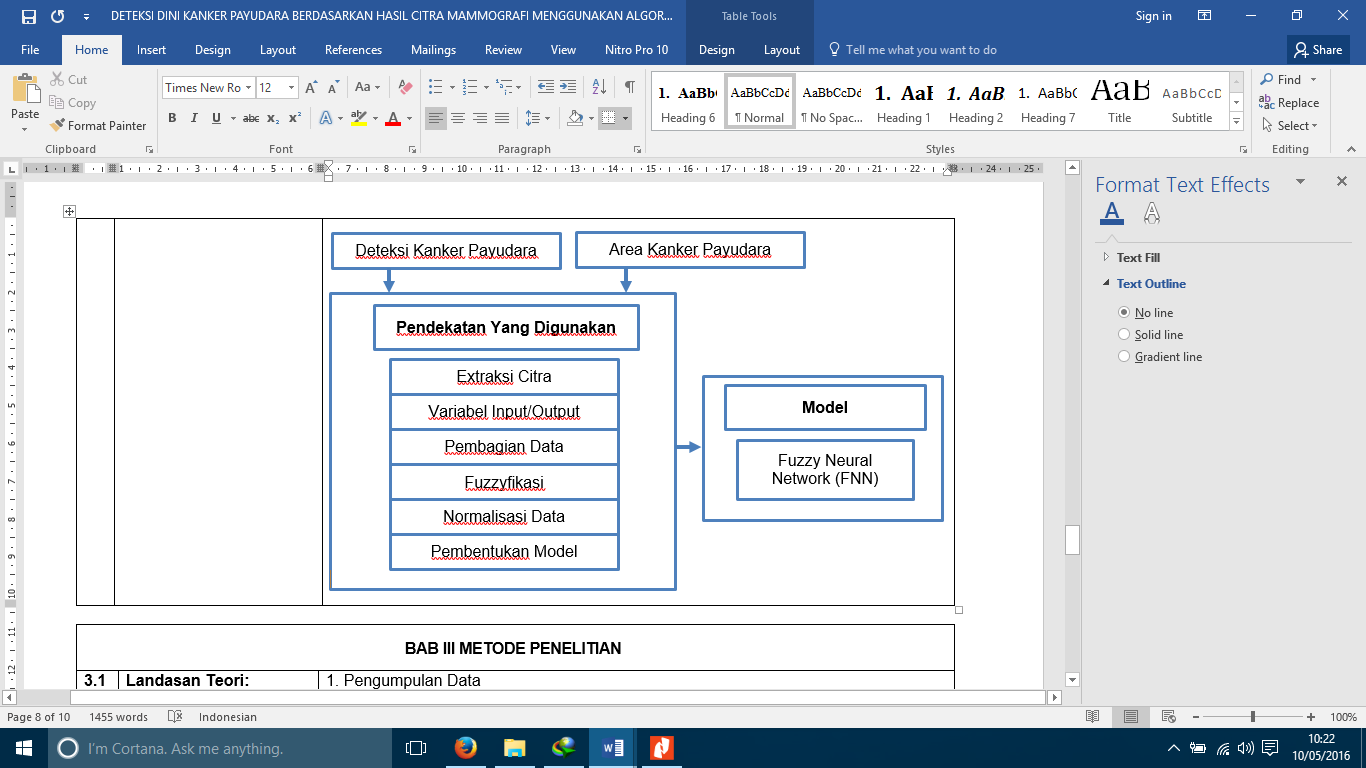
**2.3.5 Normalisasi Data**

Sebelum dilakukan proses pembelajaran, data input dan target fuzzy harus dinormalisasi terlebih dahulu. Data input dan target fuzzy dinormalisasi dengan cara membawa data ke bentuk normal yang memiliki mean=0 dan standar deviasi=1.

**2.3.6 Pembentukan Model Terbaik**

Model terbaik diperoleh dari hasil pembelajaran menggunakan algoritma backpropagation pada data training. Model terbaik ditentukan berdasarkan banyak neuron terbaik pada lapisan tersembunyi dan input yang optimal dengan melihat nilai akurasi tertinggi.

Untuk lebih jelasnya kerangka pemikiran dapat dilihat pada gambar dibawah ini :



Gambar kerangka pemikiran

BAB III

METODE PENELITIAN

# Strategi Pembuktian

**3.1.1 Pengumpulan Data**

Pada tahap ini data yang akan diproses berupa Gambar hasil mammografi dari pasien kanker payudara.

**3.1.2 Extraksi Gambar**

Ektraksi citra adalah suatu teknik pengambilan parameter statistik dari sebuah citra. Proses ekstraksi citra dapat dilakukan dengan Gray Level Coocurrence Matrix (GLCM). Fungsi bawaan pada MATLAB dapat digunakan untuk menganalisa entropy, kontras, korelasi, energi, dan homogenitas. Perintah untuk memperoleh parameter-parameter statistik tersebut adalah sebagai berikut :

g=imread('D:/gambar/1B.jpg'); {digunakan untuk membaca gambar}

h=entropy(g)

glcm=graycomatrix(g);

stats=graycoprops(glcm,'all')

Sehingga diperoleh hasil output sebagai berikut:

Entropy=7,4263

Contrast=0,0476

Correlation=0,9940

Energy= 0,1507

Homogeneity=0,9812

Namun GLCM dapat memberikan banyak parameter statistik seperti Autocorrelation, Contrast, Correlation, Cluster Prominence, Cluster Shade, Dissimilarity, Energy, Entropy, Homogeneity, Maximum probability, Sum of squares, Sum average, Sum variance, Sum entropy, Difference variance, Difference entropy, Information measure of correlation and Inverse difference normalized [9].

**3.1.3 Menentukan Variabel Input dan Variabel Output**

Variabel input model FNN adalah parameter-parameter statistik dari hasil ektraksi citra. Banyaknya variabel input menentukan banyaknya neuron pada lapisan input Variabel output model FNN adalah klasifikasi dari citra. Banyaknya variabel output menentukan banyaknya neuron pada lapisan output.

**3.1.4 Pembagian Data**

Data yang diperoleh dari hasil ektraksi citra dibagi menjadi 2 yaitu data training dan data testing. Data training digunakan untuk mencari model terbaik, sedangkan data testing digunakan untuk menguji ketepatan model hasil data training.

Terdapat beberapa komposisi data training dan data testing yang sering digunakan [10] yaitu:

1) 80% untuk data training dan 20% untuk data testing.

2) 75% untuk data training dan 25% untuk data testing.

3) 60% untuk data training dann 40% untuk data testing.

**3.1.5 Fuzzyfikasi**

Sebelum dilakukan pembelajaran, terlebih dahulu dilakukan pengubahan nilai target crisp menjadi nilai target fuzzy. Nilai target fuzzy ini digunakan sebagai target pembelajaran dalam jaringan feedforward. Proses ini disebut fuzzifikasi. Fuzzifikasi adalah pemetaan himpunan crisp ke himpunan fuzzy (Ross, 2004: 94). Fuzzifikasi pada FNN menggunakan fungsi keanggotaan pada Persamaan (2.19 & 2.20), kemudian digunakan operator Intensification (INT) untuk menentukan derajat keanggotaan yang digunakan sebagai target pembelajaran sehingga data input ke-h, yaitu 𝑥h akan memiliki target fuzzy ke-k sebagai berikut:

𝑑𝑘 = 𝜇𝐼𝑁𝑇 (𝑘)(𝑥h)

dengan 0 = 𝑑𝑘 = 1 untuk setiap k.

**3.1.6 Normalisasi Data**

Sebelum dilakukan proses pembelajaran, data input dan target fuzzy harus dinormalisasi terlebih dahulu. Data input dan target fuzzy dinormalisasi dengan cara membawa data ke bentuk normal yang memiliki mean = 0 dan standar deviasi = 1.

**3.1.7 Pembentukan Model Terbaik**

Model terbaik diperoleh dari hasil pembelajaran menggunakan algoritma backpropagation pada data training. Model terbaik ditentukan berdasarkan banyak neuron terbaik pada lapisan tersembunyi dan input yang optimal dengan melihat nilai akurasi tertinggi. Pembelajaran awal dilakukan dengan menentukan banyak neuron pada lapisan input. Setelah banyak neuron input ditentukan maka langkah selanjutnya menentukan banyak neuron terbaik pada lapisan tersembunyi. Penentuan banyak neuron pada lapisan tersembunyi ditentukan oleh nilai akurasi tertinggi yang diperoleh dari trial and error pada data training, sedangkan banyak neuron pada lapisan output tergantung dari banyak kelas. Setelah diperoleh banyak neuron terbaik pada lapisan tersembunyi, maka selanjutnya menentukan input yang optimal. Input yang optimal adalah input yang menghasilkan nilai akurasi tertinggi dari proses pembelajaran menggunakan algoritma backpropagation.

# Jadwal Penelitian

Adapun jadwal penelitian yang direncanakan dalam penelitian ini dapat dilihat pada tabel dibawah ini :

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Bulan | I | | | | II | | | | III | | | | IV | | | |
| Minggu | 1 2 3 4 | | | | 1 2 3 4 | | | | 1 2 3 4 | | | | 1 2 3 4 | | | |
| Pengumpulan Data |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Image Pre-Processing |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Image Segmentation |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| Pengujian Model |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**DAFTAR PUSTAKA**

[1] Supriyadi, Vincent Suhartono, Catur Supriyanto “Prediksi Penyakit Kanker Payudara Menggunakan Artificial Neural Network” Vol 11 Nomor 1, April 2015.

[2] Diakses pada tanggal 25 April 2016 dari http://www.depkes.go.id.

[3] Dane Kurnia Putra “Identifikasi Keberadaan Kanker Pada Citra Mammografi Menggunakan Metode Wavelet Haar”

[4] Lestari Handayani “Segmentasi Mamografi Kanker Payudara Dengan Algoritma Expectation Maximization Segmentation (Em Segmentation)” Vol. 11, No. 2, Juni 2014.

[5] Fourina Ayu Novianti ”Analisis Diagnosis Pasien Kanker Payudara Menggunakan Regresi Logistik dan Support Vector Machine (SVM) Berdasarkan Hasil Mamografi” Vol. 1, No. 1, Sept. 2012.

[6] Diakses pada tanggal 25 April 2016 dari https://id.wikipedia.org.

[7] Lin, C.-T. & Lee, George. Neuro Fuzzy Systems. New Jersey: PrenticeHall 1996.

[8] Gonzalez, Rafael C. & Woods, Richard E. Digital Image Processing 2 67 Nd Edition. Upper Saddle River: Prentice Hall 2002.

[9] Sharma, M. & Mukharjee, S. Fuzzy C-Means, ANFIS, and Genetic Algorithm for Segmenting Astroctyoma-A Tybe of Brain Tumor. IAES International Journal of Artificial Intelligence, Vol. 3, 2014.

[10] Hota, H.S., Shrivas, A.K., Singhai, S.K. Artificial Neural Network, Decision Tree and Statistical Techniques Applied for Designing and Developing E-mail Classifier. International Journal of Recent Technology and Engineering, Issue 6, 2013.