

**KLASIFIKASI KANKER PARU BERDASARKAN CITRA  
HISTOPATOLOGI DENGAN MENGGUNAKAN METODE CNN MODEL  
GOOGLENET**

**SKRIPSI**



**UIN SUNAN AMPEL  
S U R A B A Y A**

Disusun Oleh  
**SITI NUR FADILAH**  
**H92219061**

**PROGRAM STUDI MATEMATIKA  
FAKULTAS SAINS DAN TEKNOLOGI  
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL  
SURABAYA**

**2022**

## PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : SITI NUR FADILAH  
NIM : H92219061  
Program Studi : Matematika  
Angkatan : 2019

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "KLASIFIKASI KANKER PARU BERDASARKAN CITRA HISTOPATOLOGI DENGAN MENGGUNAKAN METODE CNN MODEL GOOGLINET". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 12 September 2022

Yang menyatakan,

  
SITI NUR FADILAH  
NIM. H92219061


## LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh


Nama : SITI NUR FADILAH  
NIM : H92219061  
Judul Skripsi : KLASIFIKASI KANKER PARU BERDASARKAN  
CITRA HISTOPATOLOGI DENGAN MENGGUNAKAN  
METODE CNN MODEL GOOGLNET

telah diperiksa dan disetujui untuk diujikan.


Pembimbing I

  
Dian Candia Rini Novitasari, M. Kom  
NIP. 198511242014032001

Pembimbing II

  
Lutfi Hakim, M.Ag  
NIP. 197312252006041001

Mengetahui,  
Ketua Program Studi Matematika  
UIN Sunan Ampel Surabaya

  
Yunlar Farida, M.T  
NIP. 197905272014032002

## PENGESAHAN TIM PENGUJI SKRIPSI


Skripsi oleh

Nama : SITI NUR FADILAH  
NIM : H92219061  
Judul Skripsi : KLASIFIKASI KANKER PARU  
BERDASARKAN CITRA  
HISTOPATOLOGI DENGAN  
MENGUNAKAN METODE CNN  
MODEL GOOGLNET

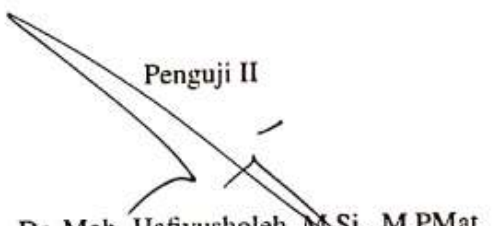
Telah dipertahankan di depan Tim Penguji  
pada tanggal 13 Januari 2023

Mengesahkan,  
Tim Penguji

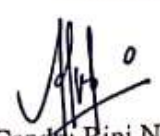
Penguji I

  
Nurissaldah Ulunruha, M.Kom  
NIP. 198002042014031001

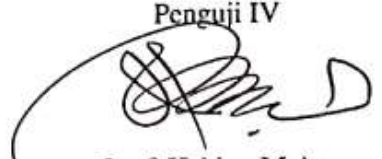
Penguji II

  
Dr. Moh. Hafiyusholeh, M.Si., M.PMat.  
NIP. 198002042014031001

Penguji III

  
Dian Candu Rini Novitasari, M.Kom  
NIP. 198511242014032001

Penguji IV

  
Lutfi Hakim, M.Ag  
NIP. 197312252006041001.

Mengetahui,

Dekan Fakultas Sains dan Teknologi  
Universitas Islam Ampel Surabaya

  
  
Abdul Hamdani, M.Pd.  
NIP. 196507312000031002



KEMENTERIAN AGAMA  
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA  
PERPUSTAKAAN

Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300  
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI  
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama : SITI NUR FADILAH  
NIM : H92219061  
Fakultas/Jurusan : SAINTEK / MATEMATIKA  
E-mail address : Sitiunurfadilah576@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :

Skripsi  Tesis  Desertasi  Lain-lain (.....)

yang berjudul :

KLASIFIKASI KANKER PARU BERDASARKAN CITRA  
HISTOPATOLOGI DENGAN MENGGUNAKAN METODE CNN  
MODEL GOOLENET

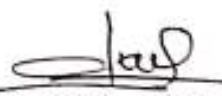
beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-eksklusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara *fulltext* untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 25 Januari 2023

Penulis

  
SITI NUR FADILAH,

*nama terang dan tanda tangan*

## ABSTRAK

### KLASIFIKASI KANKER PARU BERDASARKAN CITRA HISTOPATOLOGI DENGAN MENGGUNAKAN METODE CNN MODEL GOOGLENET

Kanker paru merupakan jenis kanker dengan jumlah kasus kematian tertinggi pertama di dunia. Kanker paru sendiri memiliki beberapa kelas yakni diantaranya *lung adenocarcinoma*, *benign lung tissue*, dan *lung squamous cell carcinoma*. Adapun tujuan dari penelitian ini yaitu mengklasifikasikan citra histopatologi kanker paru kedalam masing-masing kelas. Salah satu upaya yang dapat dilakukan untuk mengklasifikasikan citra histopatologi kanker paru yaitu dengan menggunakan bantuan sistem *Computer Aided Diagnosis (CAD)* yang mana sistem ini juga telah banyak digunakan dalam bidang kedokteran. *Convolutional Neural Network (CNN)* merupakan salah satu jenis sistem CAD yang dapat digunakan untuk membantu dalam mengklasifikasikan kanker paru. CNN sendiri memiliki beberapa model arsitektur seperti GoogLeNet yang merupakan model arsitektur CNN terbaik pada ILSVRC pada tahun 2015. Pada penelitian ini metode CNN model GoogLeNet digunakan untuk membentuk sistem klasifikasi kanker paru yang memperoleh model optimal dengan persentase akurasi, sensitivitas, serta spesifisitas sebesar 99.93%, 99.93%, dan 99.97% yang membutuhkan waktu selama 190 menit 20 detik pada pembagian data 90:10 dengan nilai probabilitas *dropout* 0.7 serta *batchsize* sebesar 64.

**Kata kunci:** GoogleNet, CNN, *Deep Learning*, Kanker Paru

## ABSTRACT

### CLASSIFICATION OF LUNG CANCER BASED ON HISTOPATHOLOGICAL IMAGES USING CNN METHODE GOOGLNET MODEL

Lung cancer is a type of cancer with the highest number of deaths in the world. Lung cancer itself has several classes, namely *lung adenocarcinoma*, *benign lung tissue*, and *lung squamous cell carcinoma*. The purpose of this study is to classify histopathological images of lung cancer into each class. One effort that can be made to classify histopathological images of lung cancer is by using the *Computer Aided Diagnosis (CAD)* system which has also been widely used in the medical field. *Convolutional Neural Network (CNN)* is a type of CAD system that can be used to assist in classifying lung cancer. CNN itself has several architectural models such as GoogLeNet which is the best CNN architectural model at ILSVRC in 2015. In this study the GoogLeNet model CNN method was used to form a lung cancer classification system which obtained an optimal model with a percentage of accuracy, sensitivity and specificity of 99.93% , 99.93%, and 99.97% which takes 190 minutes 20 seconds on a 90:10 data division with a probability value of *dropout* 0.7 and *batchsize* of 64.

**Keywords:** GoogLeNet, CNN, Deep Learning, Lung Cancer

## DAFTAR ISI

<b>HALAMAN JUDUL</b>	i
<b>LEMBAR PERSETUJUAN PEMBIMBING</b>	ii
<b>PENGESAHAN TIM PENGUJI SKRIPSI</b>	iii
<b>HALAMAN PERNYATAAN KEASLIAN</b>	iv
<b>LEMBAR PERSETUJUAN PUBLIKASI</b>	v
<b>MOTTO</b>	vi
<b>HALAMAN PERSEMBAHAN</b>	vii
<b>KATA PENGANTAR</b>	viii
<b>DAFTAR ISI</b>	x
<b>DAFTAR GAMBAR</b>	xiii
<b>ABSTRAK</b>	xv
<b>ABSTRACT</b>	xvi
<b>I PENDAHULUAN</b>	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	6
1.3. Tujuan Penelitian	7
1.4. Manfaat Penelitian	7
1.5. Batasan Masalah	8
1.6. Sistematika Penulisan	8
<b>II TINJAUAN PUSTAKA</b>	10
2.1. Kanker Paru	10
2.2. Citra Digital	12
2.2.1. Citra <i>Binary</i>	12
2.2.2. Citra <i>Grayscale</i>	13
2.2.3. Citra RGB ( <i>Red, Green, Blue</i> )	14
2.3. <i>Resizing</i>	15
2.3.1. <i>Bicubic Interpolation</i>	15



2.4. Convolutional Neural Network	16
2.4.1. Convolutional Layer	17
2.4.2. Activation Function	18
2.4.3. Pooling Layer	19
2.4.4. Cross Channel Normalization	20
2.4.5. Dropout	20
2.4.6. Fully Connected Layer	21
2.4.7. Softmax Layer	21
2.5. GoogLeNet	22
2.6. Confusion Matrix	23
2.7. Integrasi Keislaman	25
2.7.1. Penyakit dalam Perspektif Islam	25
2.7.2. Kewajiban Orang Sakit	26
<b>III METODOLOGI PENELITIAN</b>	<b>30</b>
3.1. Jenis Penelitian	30
3.2. Sumber Data	30
3.3. Kerangka Penelitian	31
<b>IV HASIL DAN PEMBAHASAN</b>	<b>33</b>
4.1. Deskripsi Data	33
4.2. Convolutional Neural Network	35
4.2.1. Convolution Layer	35
4.2.2. ReLu Layer	38
4.2.3. Pooling Layer	40
4.2.4. Cross Channel Normalization Layer	41
4.2.5. Inception Modules	44
4.2.6. Dropout Layer	44
4.2.7. Fully Connected Layer	45
4.2.8. Softmax Layer	46
4.3. Uji Coba Model dan Evaluasi aplikasi	47
4.3.1. Pembagian Data 60:40	48

4.3.2. Pembagian Data 70:30	50
4.3.3. Pembagian Data 80:20	53
4.3.4. Pembagian Data 90:10	56
4.4. Aplikasi Klasifikasi Kanker Paru Menggunakan CNN Model	
GoogLeNet	63
4.5. Integrasi Keislaman	65
<b>V PENUTUP</b>	<b>68</b>
5.1. Kesimpulan	68
5.2. Saran	68
<b>DAFTAR PUSTAKA</b>	<b>69</b>



UIN SUNAN AMPEL  
S U R A B A Y A

## DAFTAR GAMBAR

2.1 (a) <i>X-Ray</i> (b) <i>CT Scan</i> (c) <i>Histopatologi</i> . . . . .	12
2.2 Contoh Citra <i>Binary</i> . . . . .	13
2.3 Contoh Citra <i>Grayscale</i> . . . . .	13
2.4 Contoh Citra <i>RGB</i> . . . . .	14
2.5 Ilustrasi hasil <i>resizing</i> . . . . .	15
2.6 Arsitektur <i>CNN</i> . . . . .	16
2.7 Ilustrasi <i>Convolution Layer</i> . . . . .	18
2.8 Grafik Fungsi <i>ReLU</i> . . . . .	19
2.9 Ilustrasi <i>Pooling Layer</i> . . . . .	19
2.10 Ilustrasi <i>Dropout</i> . . . . .	21
2.11 Arsitektur <i>GoogLeNet</i> . . . . .	23
2.12 <i>Confusion Matrix Multi Class</i> . . . . .	24
3.1 (a) <i>Lung adenocarcinoma</i> (b) <i>Benign lung tissue</i> (c) <i>Lung squamous cell carcinoma</i> . . . . .	30
3.2 Alur Penelitian . . . . .	31
4.1 Lapisan citra <i>RGB</i> kanker paru . . . . .	33
4.2 (a)Lapisan <i>Red</i> (b)Lapisan <i>Green</i> (c) Lapisan <i>Blue</i> . . . . .	34
4.3 Ilustrasi Proses <i>Resizing</i> . . . . .	34
4.4 (a)Lapisan <i>Red</i> (b)Lapisan <i>Green</i> (c) Lapisan <i>Blue</i> . . . . .	35
4.5 Ilustrasi Operasi <i>Konvolusi</i> . . . . .	36
4.6 Visualisasi <i>Feature Map</i> Pada <i>Convolution Layer</i> . . . . .	38
4.7 Visualisasi <i>Feature Map</i> Pada <i>ReLU Layer</i> . . . . .	40
4.8 Visualisasi <i>Feature Map</i> Pada <i>Pooling Layer</i> . . . . .	41
4.9 Visualisasi <i>Feature Map</i> Pada <i>Cross Channel Normalization Layer</i> . . . . .	43

4.10 <i>Training Progress</i> Pada Pembagian 60:40 . . . . .	49
4.11 <i>Confusion Matrix</i> Pada Pembagian 60:40 . . . . .	50
4.12 <i>Training Progress</i> Pada Pembagian 70:30 . . . . .	52
4.13 <i>Confusion Matrix</i> Pada Pembagian 70:30 . . . . .	52
4.14 <i>Training Progress</i> Pada Pembagian 80:20 . . . . .	55
4.15 <i>Confusion Matrix</i> Pada Pembagian 80:20 . . . . .	55
4.16 <i>Training Progress</i> Pada Pembagian 90:10 . . . . .	57
4.17 <i>Confusion Matrix</i> Pada Pembagian 90:10 . . . . .	58
4.18 Grafik Pengaruh Pembagian Data Terhadap Waktu <i>Training</i> . . . . .	59
4.19 Grafik Pengaruh Pembagian Data Terhadap Akurasi . . . . .	59
4.20 Grafik Pengaruh <i>Dropout</i> Terhadap Waktu <i>Training</i> . . . . .	60
4.21 Grafik Pengaruh <i>Dropout</i> Terhadap Akurasi . . . . .	61
4.22 Grafik Pengaruh Jumlah <i>Batchsize</i> Terhadap Waktu <i>Training</i> . . . . .	61
4.23 Grafik Pengaruh Jumlah <i>Batchsize</i> Terhadap Akurasi . . . . .	62
4.24 Halaman Awal Aplikasi . . . . .	63
4.25 Halaman Utama Aplikasi . . . . .	64
4.26 Contoh Citra Histopatologi Setelah di <i>Resize</i> . . . . .	64
4.27 Contoh Hasil Diagnosis . . . . .	65

UIN SUNAN AMPEL  
S U R A B A Y A

# BAB I

## PENDAHULUAN

### 1.1. Latar Belakang Masalah

Kanker merupakan masalah kesehatan dan dapat berakibat kematian. Pada tahun 2018, sebanyak 9 juta kasus kanker menyebabkan kematian. WHO juga memperkirakan hingga tahun 2030 kematian yang disebabkan kanker dapat mencapai angka 11 juta jiwa. Hal ini dikarenakan kanker dapat menyerang semua kelompok (Mutiarasari, 2019). Kanker terjadi karena adanya proses mutasi gen yang kemudian tumbuh secara tidak terkendali dan dapat menginvasi serta bermetastasis (Matthews et al., 2022). Proses tersebut menyebabkan adanya perubahan fungsi protein yang nantinya akan adanya transformasi dari sel normal menjadi sel kanker serta menyebabkan adanya penurunan apoptosis serta peningkatan aktivitas proliferasi yang nantinya akan membentuk massa tumor (Rahmawati, 2021). Saat tumor tersebut berkembang dan menyebar dengan cepat maka disebut sel kanker ganas. Kanker sendiri dikelompokkan ke dalam empat jenis yaitu *leukemia*, *limpoma*, *sarcoma*, serta *carcinoma* (Shrem et al., 2022).

*Carcinoma* merupakan salah satu jenis kanker dimana sel kanker tumbuh serta berkembang di sel epitel, salah satu jenisnya yaitu kanker paru. Kanker paru adalah jenis kanker dimana sel kanker tersebut tumbuh dan berkembang pada epitel bronkus (*bronchogenic carcinoma*) dan terdapat mutasi gen yang ditandai dengan adanya perubahan urutan kromosom dalam sel (Putriani et al., 2019). Selain itu, kanker paru juga termasuk jenis kanker dengan jumlah tertinggi pertama di seluruh

dunia yakni dengan jumlah kasus sebanyak 2 juta kasus (Sibero et al., 2021). Pada tahun 2020, jumlah kasus kanker mencapai 19.2 juta kasus dengan sebanyak 18% dari total kasus merupakan kanker paru. Secara umum, kanker paru dikelompokkan dalam 2 kategori yaitu *Small Cell Lung Cancer* (SCLC) dan *Non-Small Cell Lung Cancer* (NSCLC) yang mana kedua kategori tersebut dibedakan berdasarkan pertumbuhan serta penyebarannya (Raso et al., 2021). NSCLC merupakan kategori yang umum diderita dengan persentase 85.00% dari keseluruhan penderita kanker paru. WHO menyatakan bahwa kanker paru adalah jenis kanker yang mayoritas penderitanya adalah laki-laki dengan rentan usia di atas 40 tahun (Riwanto, 2020). Oleh karena itu, angka kematian akibat kanker paru yang terjadi pada laki-laki mengalami peningkatan setiap tahunnya.

Jumlah angka kematian dari kanker paru yang semakin meningkat tidak lepas dari banyak faktor risiko yang ada. Faktor-faktor tersebut dapat menjadi penyebab kanker paru seperti genetik, kekebalan tubuh, dan kebiasaan merokok (Tahir, 2020). Kanker paru tidak hanya dapat terjadi pada perokok aktif, melainkan pula pada perokok pasif. Hal ini dikarenakan penyebabnya berada pada asap rokok tersebut, dimana asap rokok tersebut mengandung zat kimia yang bersifat karsinogen yang kemudian dihirup oleh perokok pasif dan dapat menyebabkan sesak nafas (Sandhi, 2019). Sesak nafas merupakan salah satu gejala yang ada pada penderita kanker paru. Selain itu, terdapat pula gejala lain seperti nyeri dada, batuk darah, dan nafas pendek (Aktalina, 2022). Jika telah merasakan salah satu gejala tersebut, maka diperlukan penanganan serta tindakan untuk mendiagnosis penyakit. Hal ini, dikarenakan jika sudah terdiagnosis terkena kanker paru maka membutuhkan pengobatan yang harus dilakukan dalam rentan waktu yang lama seperti kemoterapi. Kemoterapi adalah pengobatan yang dapat dilakukan dimana pengobatan ini membutuhkan waktu yang

lama serta memiliki banyak efek samping bagi penderita (Binotto et al., 2020). Oleh karena itu, kesabaran sangat dibutuhkan dalam menjalani pengobatan kanker sebagaimana dijelaskan dalam Al-Qur'an Surah Al-Baqarah ayat 153 yang berbunyi:

يَا أَيُّهَا الَّذِينَ ءَامَنُوا اسْتَعِينُوا بِالصَّبْرِ وَالصَّلَاةِ إِنَّ اللَّهَ مَعَ الصَّابِرِينَ

artinya: "Wahai orang-orang yang beriman! Mohonlah pertolongan (kepada Allah) dengan sabar dan salat. Sungguh, Allah beserta orang-orang yang sabar."(QS. Al-Baqarah:153).

Berdasarkan ayat diatas, dijelaskan bahwa memohon pertolongan harus senantiasa untuk sabar karena Allah SWT selalu bersama orang sabar. Jika senantiasa bersabar maka Allah SWT juga akan senantiasa memberikan pertolongan kepada hamba-Nya. Jadi, jika kita tertimpa suatu penyakit bersabar adalah suatu kunci utama dalam proses penyembuhan. Selain bersabar, berikhtiar juga merupakan yang harus dilakukan jika sedang tertimpa penyakit. Ikhtiar atau berusaha dengan bersungguh-sungguh sangat membantu dalam proses penyembuhan karena dengan berikhtiar artinya semangat untuk ingin sembuh dari penyakit tersebut masih ada dan berharap akan memperoleh hasil terbaik atas ikhtiar yang telah dilakukannya. Hal ini dijelaskan dalam Qur'an Surah An-Najm ayat 39-42 yang berbunyi:

وَأَن لَّيْسَ لِلْإِنسَانِ إِلَّا مَا سَعَىٰ وَأَنَّ سَعْيَهُ سَوْفَ يُرَىٰ ثُمَّ يُجْزَاهُ الْجَزَاءَ الْأَوْفَىٰ  
وَأَنَّ إِلَىٰ رَبِّكَ الْمُنتَهَىٰ

artinya: "Dan bahwa manusia hanya memperoleh apa yang telah diusahakannya, dan sesungguhnya usahanya itu kelak akan diperlihatkan (kepadanya), kemudian akan diberi balasan kepadanya dengan balasan yang paling sempurna, dan sesungguhnya kepada Tuhanmulah kesudahannya (segala sesuatu)."(QS. An-Najm: 39-42).

Berdasarkan ayat diatas, dijelaskan bahwasanya apapun hasilnya, kewajiban kita sebagai manusia hanyalah berusaha dan untuk hasilnya hanya Allah SWT yang menentukan. Sekeras apapun sakit yang diderita jika kita sudah berikhtiar serta berdoa dan Allah telah berkehendak untuk sembuh, maka kita akan memperoleh kesembuhan tersebut. Hal ini dikarenakan Allah SWT mengetahui apa yang terbaik bagi kita. Jika Allah SWT sudah menentukan kesembuhan atas suatu penyakit, maka akan sembuhlah selaras dengan hadits berikut:

اللَّهُمَّ رَبَّ النَّاسِ أَذْهَبِ الْبَاسَ وَاشْفِهِ وَأَنْتَ الشَّافِي لَا شِفَاءَ إِلَّا شِفَاؤُكَ شِفَاءً لَا يُغَادِرُ سَقَمًا

artinya: "Ya Allah Rabb manusia, dzat yang menghilangkan rasa sakit, sembuhkanlah sesungguhnya Engkau Dzat yang Maha menyembuhkan, tidak ada kesembuhan melainkan dari kesembuhan-Mu, yaitu kesembuhan yang tidak menyisakan rasa sakit."(HR Bukhari No. 5302).

Berdasarkan hadits diatas, dijelaskan bahwa sesungguhnya tidak ada kesembuhan atas segala penyakit kecuali kesembuhan dari-Nya. Jika Allah SWT sudah menghendaki sesuatu maka sesuatu itu akan terjadi. Hal ini selaras dengan sekeras apapun penyakit yang diderita namun jika Allah SWT sudah berkehendak untuk sembuh maka hal tersebut akan terjadi. Walaupun demikian, tidak ada salahnya untuk berusaha menjaga kesehatan seperti melakukan pemeriksaan. Pemeriksaan awal pendeteksian kanker paru dapat dilakukan dengan foto *thorax* AP. Jika terdapat tumor yang memiliki kemungkinan akan menjadi tumor ganas, maka dilanjutkan dengan pemeriksaan *CT-Scan thorax*. Kemudian, dilanjutkan dengan melakukan pemeriksaan bronkoskopi untuk mendapatkan spesimen kemudian untuk mengklasifikasikan jenisnya dengan melakukan pemeriksaan histopatologi (Zulfa, 2022).

Citra yang dihasilkan pada pemeriksaan histopatologi akan dilakukan proses



selanjutnya yaitu akan mengklasifikasikannya. Salah satu upaya yang dapat dilakukan dalam mengklasifikasikan dengan memanfaatkan adanya kemajuan teknologi yaitu dengan memanfaatkan adanya teknologi *Artificial Intelligence* (AI) (Kirman et al., 2019). Adanya pemanfaatan teknologi ini yaitu untuk meminimalisir adanya kesalahan diagnosis atau kesalahan dalam mengklasifikasikan kelas dari kanker tersebut. Selain itu, dalam bidang kedokteran juga dapat membantu mengurangi waktu yang digunakan dalam mendiagnosis atau mengklasifikasikan karena dengan memanfaatkan adanya teknologi waktu yang dibutuhkan relatif sedikit (Nugraha et al., 2019). Oleh karena itu, untuk mengantisipasi terjadinya hal ini yaitu dapat memanfaatkan salah satunya sistem *Computer Aided Diagnosis* (CAD) dimana salah satu metode yang dapat digunakan yaitu *Convolutional Neural Network* (CNN) (Putri, 2022).

CNN merupakan suatu metode *deep learning* yang dirancang untuk mengolah data dua dimensi dengan lebih efisien karena tidak memerlukan proses ekstraksi fitur. Berbagai penelitian telah menggunakan metode ini seperti penelitian yang dilakukan oleh (Azizah, 2021) yang mengklasifikasikan penyakit *diabetic retinopathy* dengan menggunakan metode CNN model ResNet dan memperoleh akurasi sebesar 90.18%. Lalu, terdapat penelitian yang mengklasifikasikan kanker paru dengan menggunakan metode CNN model AlexNet memperoleh nilai akurasi sebesar 99.00% (Zulfa, 2022). Penelitian yang dilakukan oleh (Novitasari et al., 2020) yang mendeteksi Covid-19 berdasarkan citra X-ray paru memperoleh akurasi sebesar 98.00% juga menggunakan metode CNN. Berdasarkan penelitian-penelitian tersebut dengan nilai akurasi diatas 90.00% menunjukkan bahwa metode ini mampu melakukan klasifikasi dengan baik.

Berdasarkan penelitian terdahulu dapat diketahui bahwa metode CNN memiliki beberapa model yang dapat digunakan dalam penelitian dengan

menggunakan data dua dimensi. Terdapat model ResNet, AlexNet, VGGNet, GoogLeNet, dan lain sebagainya. GoogLeNet adalah model CNN yang memiliki kelebihan dimana *inception modules* yang terdiri atas sejumlah *convolution* kecil yang berguna untuk melakukan reduksi. Selain itu, GoogLeNet memiliki proses komputasi yang lebih cepat. Hal ini dibuktikan dengan penelitian sebelumnya yakni penelitian yang dilakukan oleh (Ali et al., 2022) yang mendeteksi tumor 97.00% dimana dijelaskan bahwa waktu dengan menggunakan lebih cepat dan efisien. Kemudian, penelitian yang dilakukan oleh (Bhanumathi and Sangeetha, 2019) yang mengklasifikasikan tumor otak memperoleh nilai akurasi sebesar 97.10% dengan waktu 56 detik. Selain itu, juga terdapat penelitian (Mohamed Loey, Florentin Smarandache, 2020) yang mendeteksi Covid-19 menggunakan data citra X-ray memperoleh akurasi sebesar 99.90%. Berdasarkan nilai akurasi yang diperoleh pada penelitian-penelitian sebelumnya maka model GoogLeNet sangat baik untuk digunakan dalam penelitian ini.

Berdasarkan penelitian terdahulu, metode CNN sangat baik digunakan untuk klasifikasi karena akurasi yang diperoleh sangat tinggi. Terlebih untuk metode CNN model GoogLeNet yang memiliki akurasi rata-rata diatas 90.00%. Oleh karena itu, pada penelitian ini digunakan metode CNN model GoogLeNet untuk melakukan klasifikasi kanker paru berdasarkan citra histopatologi.

## 1.2. Rumusan Masalah

Berdasarkan pemaparan masalah pada latar belakang, diperoleh rumusan masalah sebagai berikut:

1. Bagaimana model optimal yang diperoleh pada sistem klasifikasi kanker paru dengan menggunakan metode CNN model GoogLeNet berdasarkan uji coba

parameter *dropout* dan *mini batchsize*?

2. Bagaimana hasil kinerja dari metode CNN model GoogLeNet untuk melakukan klasifikasi kanker paru menggunakan data citra histopatologi?

### 1.3. Tujuan Penelitian

Berdasarkan dari beberapa rumusan masalah yang telah dipaparkan, maka tujuan yang ingin dicapai yaitu:

1. Memperoleh model optimal pada sistem diagnosis kanker paru yang terbaik dan tercepat dengan menggunakan metode CNN model GoogLeNet berdasarkan hasil uji coba parameter *dropout* dan *mini batchsize*.
2. Memperoleh hasil kinerja dari metode CNN model GoogLeNet untuk melakukan klasifikasi kanker paru menggunakan data citra histopatologi.

### 1.4. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat kepada banyak pihak.

#### 1. Manfaat Teoritis

##### (a) Bagi Penulis

Manfaat bagi penulis yakni dapat digunakan sebagai salah satu proses pembelajaran terkait penulisan dan penyelesaian masalah secara sistematis serta dapat menambah keilmuan terkait kanker paru dan metode CNN model GoogLeNet.

##### (b) Bagi Pembaca

Manfaat bagi pembaca yakni dapat digunakan sebagai referensi keilmuan

di bidang kedokteran dan menambah wawasan mengenai kanker paru dan metode CNN model GoogLeNet.

## 2. Manfaat Praktis

Dalam dunia medis, penelitian ini digunakan oleh tenaga medis dalam membangun diagnosis kanker paru menggunakan metode CNN model GoogLeNet.

### 1.5. Batasan Masalah

Penelitian ini diberikan batasan masalah dikarenakan ruang lingkup permasalahan yang sangat luas, diantaranya yaitu:

1. Data yang digunakan merupakan data citra histopatologi kanker paru yang terdiri atas tiga kelas yaitu *lung adenocarcinoma*, *benign lung tissue*, dan *lung squamous cell carcinoma*.
2. Pada tahap *preprocessing* dilakukan *resizing* dan pembagian data.
3. Metode yang digunakan yaitu CNN model GoogLeNet.
4. Parameter yang digunakan pada metode CNN yaitu *dropout* dan *mini batchsize*.

### 1.6. Sistematika Penulisan

Sistematika pada penelitian ini dibagi menjadi beberapa bagian berikut:

#### 1. BAB I PENDAHULUAN

Berisi latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian dan sistematika penulisan.

#### 2. BAB II TINJAUAN PUSTAKA

Berisi landasan teori meliputi kanker paru, citra digital, *resizing*, metode CNN, GoogLeNet dan *confusion matrix*.

### 3. BAB III METODOLOGI PENELITIAN

Berisi jenis penelitian, jenis dan sumber data serta tahap-tahap penelitian yang dilakukan.

### 4. BAB IV HASIL DAN PEMBAHASAN

Berisi hasil dari proses *resizing*, uji coba menggunakan metode CNN model GoogLeNet berdasarkan uji coba *dropout* dan *mini batchsize* serta menganalisis hasil dari penerapan tersebut.

### 5. BAB V PENUTUP

Berisi kesimpulan dari hasil penelitian yang telah dilakukan dan saran untuk penelitian selanjutnya.



UIN SUNAN AMPEL  
S U R A B A Y A

## BAB II

### TINJAUAN PUSTAKA

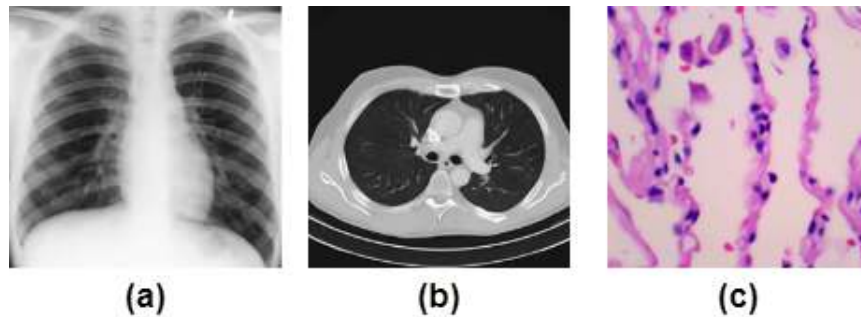
#### 2.1. Kanker Paru

Kanker paru merupakan neoplasma ganas yang berasal dari epitel bronkus yakni sel epitelium respirasi. Sel kanker ini mengalami invasi serta metastasis yang umumnya mengalami metastasis pada organ yang jauh seperti otak, tulang, serta kelenjar adrenal. Berdasarkan jenisnya sel tersebut terbagi menjadi dua yaitu *Small Cell Lung Cancer* (SCLC) dan *Non-Small Cell Lung Cancer* (NSCLC) (Hynds et al., 2021). SCLC merupakan jenis sel kecil yang diisi mukus dan berkumpul diantara pembuluh darah halus. NSCLC merupakan jenis sel kanker paru yang umum dimana hampir semua kasus kanker paru 70 hingga 80% adalah jenis NSCLC (Putriani et al., 2019). Jenis sel kanker NSCLC ini terbagi menjadi tiga subjenis yaitu *adenocarcinoma*, *squamous cell carcinoma*, dan *large cell carcinoma* (Ichsan et al., 2020). *Adenocarcinoma* adalah jenis sel kanker NSCLC yang ditemukan di bagian perifer paru dengan pola pertumbuhan *acinar*, *papillary*, *micropapillary*, *lepidic*, dan *solid* dengan jumlah kasus 35 hingga 40% dari seluruh penderita kanker paru jenis sel NSCLC. *Squamous cell carcinoma* memiliki ciri khas adanya pembentukan jembatan intraseluler dan tumbuh di bronkus utama yang kemudian menyebar ke kelenjar getah bening dengan jumlah kasus 20 hingga 25%, dan *large cell carcinoma* adalah jenis kanker paru yang massa tumornya ditemukan di area perifer paru dimana massa tumornya berbentuk poligonal dan berukuran besar. *Large cell carcinoma* sering dijumpai dari semua jenis kanker paru dan banyak terjadi pada perokok aktif.

Faktor risiko penyebab kanker paru dapat berasal dari polusi dan kebiasaan merokok. Hal ini dikarenakan terdapat zat pemicu kanker pada rokok tersebut. Perokok pasif tidak menutup kemungkinan terkena kanker paru. Sedangkan, bagi perokok aktif mempunyai nilai kemungkinan terkena kanker paru lebih besar dibandingkan dengan perokok pasif (Satriawan, 2022). Oleh karena itu, penderita kanker paru umumnya berjenis kelamin laki-laki. Umur juga termasuk salah satu faktor penyebab kanker paru dimana umumnya penderita kanker rata-rata berumur 65 tahun. Selain itu, pengaruh genetik dan status imunologi, predisposisi karena penyakit lain, dan zat karsinogen yang berada di lingkungan sekitar juga termasuk salah satu faktor (Rifai, 2020). Berdasarkan faktor-faktor tersebut maka perlu dilakukan pemeriksaan awal untuk yang mengetahui apakah seseorang menderita kanker paru.

Pemeriksaan terkait deteksi kanker paru salah satunya *X-Ray* dimana hampir semua rumah sakit memiliki alat untuk pemeriksaan ini dan cara kerja pemeriksaan ini yaitu gelombang radiasi sinar X akan dikirimkan melalui mesin. Selain itu, terdapat juga pemeriksaan lain yang dapat dilakukan *CT-Scan*, *Magnetic Resonance Imaging* (MRI), pemeriksaan bronkoskopi, dan pemeriksaan histopatologi. *CT-Scan* merupakan metode pemeriksaan dengan mengambil citra menggunakan sinar-X (Akmalovna, 2022). Pemeriksaan MRI yaitu dengan menggunakan teknologi magnet dan gelombang radio dan bisa dilakukan dalam jangka waktu yang berdekatan (Khalid et al., 2020). Pemeriksaan bronkoskopi adalah prosedur pemeriksaan yang dilakukan dengan memasukkan alat bronkoskop pada saluran pernapasan dan paru-paru. Kemudian, terdapat pemeriksaan histopatologi yang merupakan salah satu jenis pemeriksaan patologi anatomi yang bertujuan memeriksa jaringan tubuh dengan menggunakan bantuan mikroskop untuk mengidentifikasi

komponen jaringan tersebut.



Gambar 2.1 (a) X-Ray (b) CT Scan (c) Histopatologi

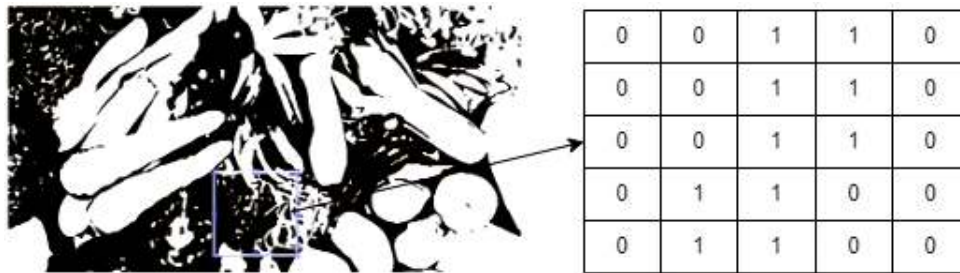
## 2.2. Citra Digital

Citra adalah sebuah gambar dua dimensi yang kemudian dibagi menjadi N baris dan M kolom sehingga menghasilkan gambar diskrit (Foady, 2019). Citra didefinisikan menggunakan fungsi  $f(x, y)$  yang merupakan intensitas keabuan dari suatu citra dengan  $x$  dan  $y$  merupakan koordinat spasial. Dalam hal ini, intensitas dari citra merupakan tingkat keabuan dari suatu citra (Prasetio, 2021). Berdasarkan jenis warna, citra digital terbagi menjadi beberapa macam diantaranya:

### 2.2.1. Citra *Binary*

Citra *binary* atau citra yang memiliki dua kemungkinan yaitu hitam dan putih pada setiap pikselnya sehingga dapat juga disebut dengan citra monokrom. Citra ini membutuhkan 1 bit yang mewakili nilai setiap pikselnya sama halnya dengan citra *grayscale* (Saputra et al., 2022). Umumnya, citra ini didapatkan melalui proses pengolahan citra seperti segmentasi yang kemudian diproses ulang untuk mendeteksi lokasi suatu objek dengan nilai piksel kecil. Berikut contoh citra *binary* ditunjukkan oleh Gambar 2.2.





Gambar 2.2 Contoh Citra *Binary*

Berdasarkan Gambar 2.2 diketahui bahwa matriks pada citra tersebut hanya bernilai antara 0 yang menunjukkan hitam atau 1 menunjukkan putih.

### 2.2.2. Citra *Grayscale*

Citra *grayscale* merupakan suatu citra digital yang memiliki warna hitam, keabuan, serta putih. Umumnya, citra ini direpresentasikan ke dalam bentuk *array* dua dimensi dimana elemen tersebut menunjukkan intensitas dari suatu citra yang berada pada koordinat yang bersesuaian dengan jumlah bit yang digunakan (Hsu et al., 2021). Berikut contoh citra *grayscale* ditunjukkan oleh Gambar 2.3.

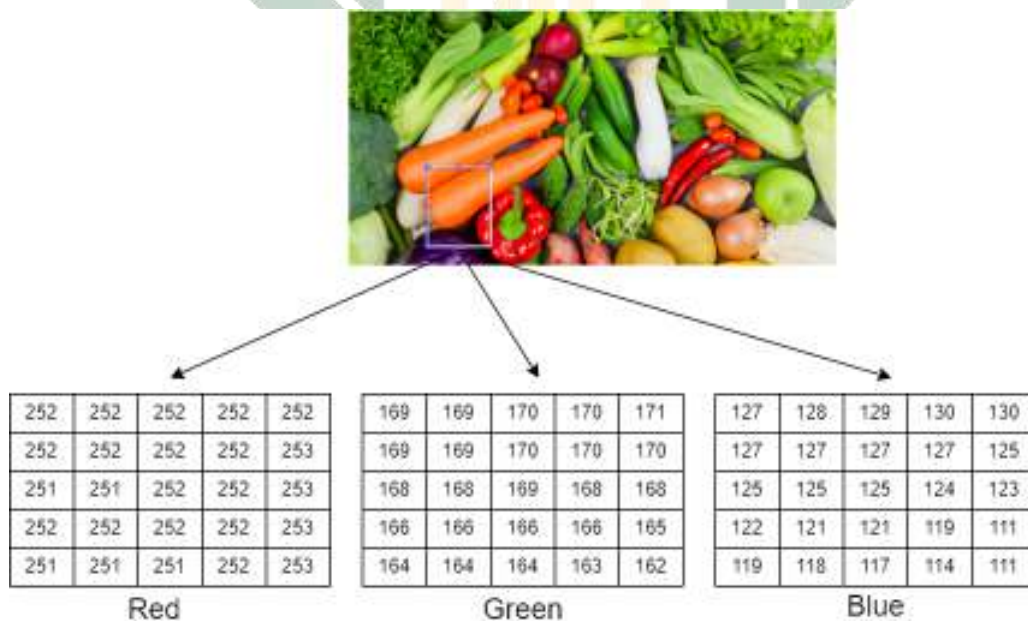


Gambar 2.3 Contoh Citra *Grayscale*

Pada Gambar 2.3 diketahui nilai-nilai piksel dari gambar tersebut merupakan nilai skala keabuan dari citra tersebut. Skala keabuan pada citra ini berada pada interval 0 hingga 255 dimana semakin tinggi nilai skala keabuan suatu citra maka semakin terang pula derajat keabuannya.

### 2.2.3. Citra RGB (*Red, Green, Blue*)

Citra RGB juga dikenal citra warna merupakan suatu citra yang terdiri dari tiga warna dasar yakni merah, hijau dan biru dimana setiap warna dasar pada citra ini masing-masing memiliki intensitas sebesar 8 bit yakni berada pada interval 0 hingga 255. Jika setiap warna dasar memiliki nilai sebesar 255 maka memerlukan media untuk penyimpanan sebesar 3 *byte* untuk tiap piksel. Selain itu, juga akan mendapatkan sebanyak 16 juta jenis warna yang diperoleh dari masing-masing nilai maksimum setiap warna dasar (Theilmann, 2022). Berikut contoh citra RGB ditunjukkan oleh Gambar 2.4.



Gambar 2.4 Contoh Citra RGB

Pada Gambar 2.4 diketahui bahwa nilai setiap warna dasar yakni merah, hijau, dan blue berada interval 0 hingga 255 pada setiap masing-masing warna dasar.

### 2.3. Resizing

*Resizing* merupakan teknik dalam pengolahan citra digital dimana teknik ini digunakan untuk merubah ukuran citra menjadi ukuran tertentu. Umumnya, teknik ini dilakukan agar citra yang akan diinput memiliki ukuran yang sama. Teknik ini juga disebut dengan proses normalisasi suatu citra (Susim and Darujati, 2021). Pada penelitian ini, metode *resizing* yang digunakan yaitu *bicubic interpolation*.

#### 2.3.1. Bicubic Interpolation

Interpolasi pada citra digital digunakan untuk memperoleh perkiraan nilai piksel warna serta nilai intensitas citra berdasarkan nilai piksel-piksel disekitarnya. Salah satu jenis interpolasi yang sering digunakan dalam pengolahan citra digital yaitu *bicubic interpolation*. *Bicubic interpolation* memperoleh informasi dengan menggunakan  $4 \times 4$  piksel ketetanggan sehingga akan memperoleh citra yang lebih tajam. Persamaan yang digunakan untuk melakukan proses *resizing* ditunjukkan oleh Persamaan 2.1 serta contoh citra yang telah dilakukan *resizing* yang awalnya berukuran  $326 \times 620$  menjadi citra yang berukuran  $224 \times 224$  pada Gambar 4.3.

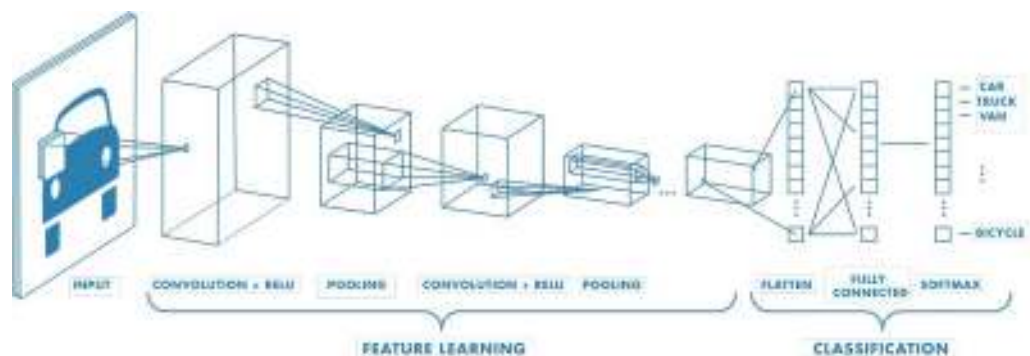
$$p(x, y) = \sum_{i=0}^3 \sum_{j=0}^3 a_{i,j} x^i y^j \quad (2.1)$$



Gambar 2.5 Ilustrasi hasil *resizing*

## 2.4. Convolutional Neural Network

*Convolutional Neural Network* (CNN) merupakan jenis *deep learning* yang dirancang untuk mengolah data citra karena memiliki hasil yang sangat signifikan dalam pengenalan citra (Ghulam et al., 2022). Oleh karena itu, CNN dirancang khusus untuk mengklasifikasikan suatu citra karena kedalaman jaringan pada CNN tergolong tinggi dan berusaha untuk mencontoh sistem pengenalan citra yang terdapat pada *visual cortex* manusia. CNN juga memiliki kelemahan yaitu membutuhkan waktu yang lama untuk proses pelatihan model (Wu et al., 2018). Namun, kini kelemahan tersebut dapat teratasi dengan menggunakan teknologi *General Purpose Graphical Processing Unit* (GPGPU). Pada *deep learning* ini terdapat beberapa komponen utama seperti *convolution layer*, *pooling layer*, *dropout*, dan lain sebagainya dimana komponen-komponen tersebut ditunjukkan oleh Gambar 2.6 yang merupakan arsitektur dari CNN.



Gambar 2.6 Arsitektur CNN

Sumber: Mathworks

Berdasarkan gambar diatas, terdapat beberapa komponen utama pada arsitektur CNN diantaranya:

### 2.4.1. Convolutional Layer

*Convolution layer* adalah *layer* pertama dan komponen utama yang mendasar yang berada pada CNN dimana *layer* ini bertugas sebagai penerima *input* citra melakukan ekstraksi fitur pada citra tersebut dan mengalami proses konvolusi. Umumnya, pada *convolution layer* dilakukan suatu operasi yakni operasi kombinasi linear filter terhadap daerah lokal dengan filter yang ukuran panjang, lebar, dan tinggi sesuai dengan data citra yang diterima (Kartikasari, 2020). Rumus yang digunakan pada operasi konvolusi ditunjukkan pada Persamaan 2.2.

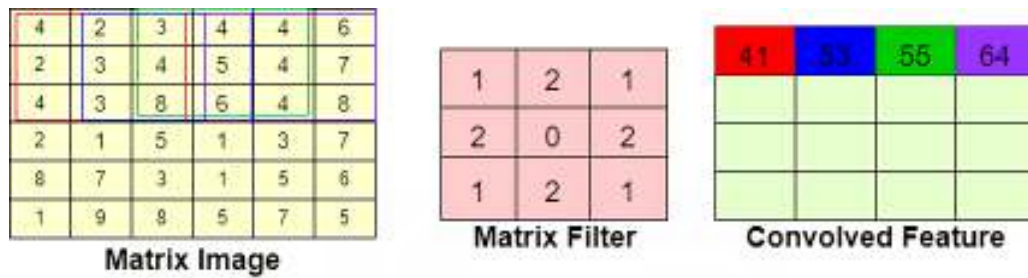
$$Q_{(i,j)} = \left( \sum_{i=1}^N I_{(i,j)} \cdot F_{(i,j)} \right) + B_j \quad (2.2)$$

Berdasarkan rumus di atas diketahui bahwa  $Q(i, j)$  merupakan hasil operasi *dot product* antara matriks input yaitu  $I(i, j)$  dengan matriks filter yaitu  $F(i, j)$  yang kemudian ditambahkan dengan nilai bias dari matriks filter tersebut. Pada Persamaan 2.2  $i$  merupakan baris pada matriks, dan  $j$  merupakan kolom pada matriks. Terdapat beberapa parameter yang digunakan pada proses ini, yaitu *stride* dan *padding*. Parameter-parameter ini nantinya digunakan untuk menghitung dimensi *output* dari *feature map* ditunjukkan oleh Persamaan 2.3.

$$H_Q = \frac{I - F + 2p}{s} + 1 \quad (2.3)$$

Berdasarkan rumus di atas diketahui bahwa  $H$  adalah ukuran dari matriks  $Q$ ,  $I$  merupakan matriks *input*,  $F$  merupakan matriks filter,  $p$  merupakan *padding* bertujuan menentukan jumlah piksel yang bernilai 0 serta mengatur dimensi *output* tetap sama agar dapat dilakukan proses ekstraksi fitur yang mendalam, dan  $s$  merupakan *stride* yang digunakan untuk menentukan jumlah pergeseran pada suatu

matriks, dimana semakin kecil *stride* komputasi yang dibutuhkan lebih besar (Sadewa et al., 2019). Ilustrasi proses pada *convolution layer* ditunjukkan oleh Gambar 2.7.



Gambar 2.7 Ilustrasi *Convolution Layer*

Berdasarkan Gambar 2.7 dilakukan perkalian matriks  $3 \times 3$  dari matriks citra dengan matriks filter. Kemudian hasil dari operasi *dot product* tersebut berada pada baris 1 kolom 1 matriks baru.

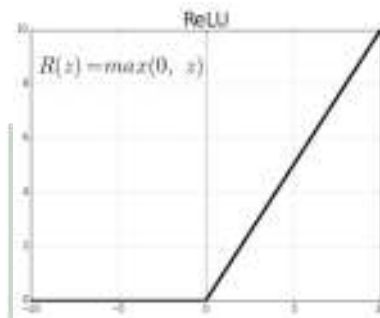
#### 2.4.2. Activation Function

Fungsi aktivasi merupakan suatu fungsi non-linier yang digunakan pada nilai hasil operasi konvolusi. Terdapat beberapa macam fungsi aktivasi, salah satunya fungsi aktivasi ReLU. *Rectified Linear Unit* atau ReLU digunakan untuk menghilangkan nilai negatif yang ada pada citra dengan cara menggantinya dengan nilai 0 (Haq, 2021). Fungsi aktivasi ReLU ini dapat membantu untuk mengurangi waktu *training* dan *testing* dengan menggunakan Persamaan 2.4 dimana  $z$  menunjukkan nilai yang diperoleh dari layer sebelumnya serta grafik yang ditunjukkan oleh Gambar 2.8.

$$R(z) = \max(0, z) \quad (2.4)$$

$$(2.5)$$

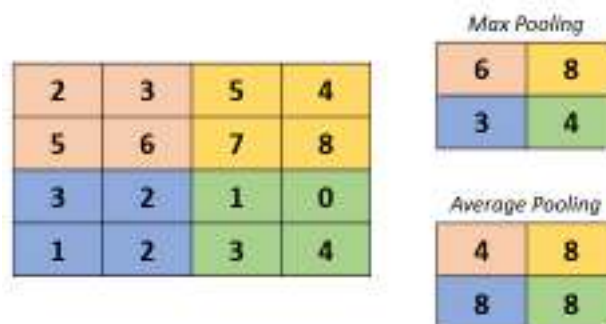
$$R(z) = \begin{cases} 0, & z \leq 0 \\ z & z > 0 \end{cases}$$



Gambar 2.8 Grafik Fungsi ReLU

### 2.4.3. Pooling Layer

*Pooling layer* merupakan suatu *layer* yang digunakan untuk mempercepat komputasi dengan teknik *downsampling* dan menghasilkan dimensi *feature map* yang lebih kecil. *Pooling layer* terbagi menjadi 2 jenis yaitu *max pooling* yang mengambil nilai maksimal dan *average pooling* yang mengambil nilai rata-rata.



Gambar 2.9 Ilustrasi Pooling Layer

#### 2.4.4. Cross Channel Normalization

*Cross channel normalization* digunakan untuk meningkatkan kemampuan generalisasi dimana mengubah nilai setiap elemen dengan nilai normalisasi yang berdasarkan nilai elemen dari fitur terdekat baik dari *feature map* yang sama ataupun tidak (Phiadelvira et al., 2021). Perhitungan yang digunakan ditunjukkan oleh Persamaan 2.6.

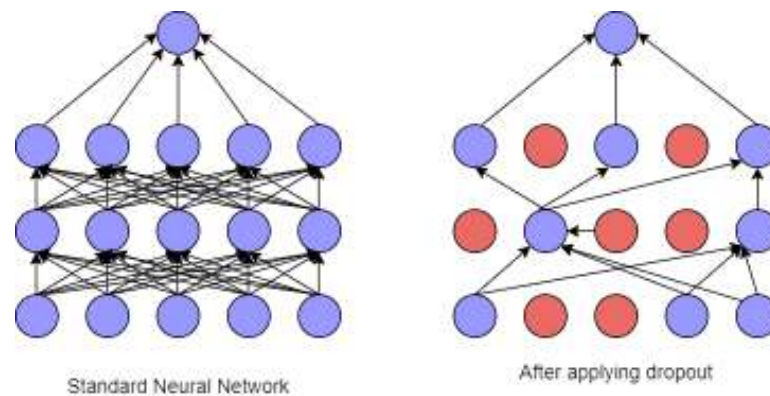
$$x_{i,j}^p = \frac{a_{i,j}^p}{\left( k + \frac{\alpha \sum_{q=\max(1, p-\frac{n}{2})}^{\min(N-1, p+\frac{n}{2})} (a_{i,j}^p)^2}{n} \right)^{\beta}} \quad (2.6)$$

Dimana  $x(i, j)^p$  merupakan *output* dari *feature map* yang diperoleh dari filter ke- $p$ ,  $a(i, j)^p$  merupakan *feature map* sebelum melakukan perhitungan pada baris ke-  $i$  dan kolom ke- $j$  serta beberapa konstanta yang digunakan seperti  $k$ ,  $\alpha$ , dan  $\beta$ . Dimana konstanta  $k$  merupakan konstanta yang digunakan untuk menghindari adanya pembagian nol,  $\alpha$  merupakan konstanta normalisasi,  $n$  adalah banyak nilai piksel berurutan, dan  $\beta$  merupakan konstanta kontras.

#### 2.4.5. Dropout

*Dropout* merupakan suatu teknik dalam algoritma jaringan saraf yang digunakan untuk mengantisipasi adanya *overfitting* dengan menggabungkan beberapa jenis arsitektur. Jika semakin tinggi nilai bobot maka akan proses *training* akan membutuhkan waktu yang semakin lama (Virkhansa et al., 2019). Oleh karena itu, jika semakin tinggi tingkat *dropout* yang digunakan, maka semakin *general* pula model yang diperoleh.





Gambar 2.10 Ilustrasi Dropout

#### 2.4.6. Fully Connected Layer

Pada *layer* ini setiap *neuron* yang memiliki keterkaitan hingga menjadi *multi layer perceptron*. *Layer* ini merupakan *layer* terakhir setelah *convolutional layer* dan *pooling layer* yang digunakan sebagai *layer* penyatuan akhir yang melakukan klasifikasi ke dalam kelas pada data *training* (Nasrulloh, 2020). *Fully connected layer* ini didefinisikan oleh Persamaan 2.7.

$$y_j = b_j + \sum_i w_{i,j} x_i \quad (2.7)$$

Berdasarkan Persamaan 2.7 dapat diketahui bahwa  $y_j$  merupakan output dari *fully connected layer*,  $b_j$  merupakan bias,  $w_{i,j}$  merupakan bobot, dan  $x_i$  merupakan *input* yang didapat dari *feature learning*. Pada Persamaan sebelumnya  $j$  menunjukkan kolom.

#### 2.4.7. Softmax Layer

*Softmax* merupakan fungsi aktivasi untuk mengubah nilai *vector* serta untuk memperoleh probabilitas tiap data dengan interval 0 hingga 1 (Hasby, 2022).

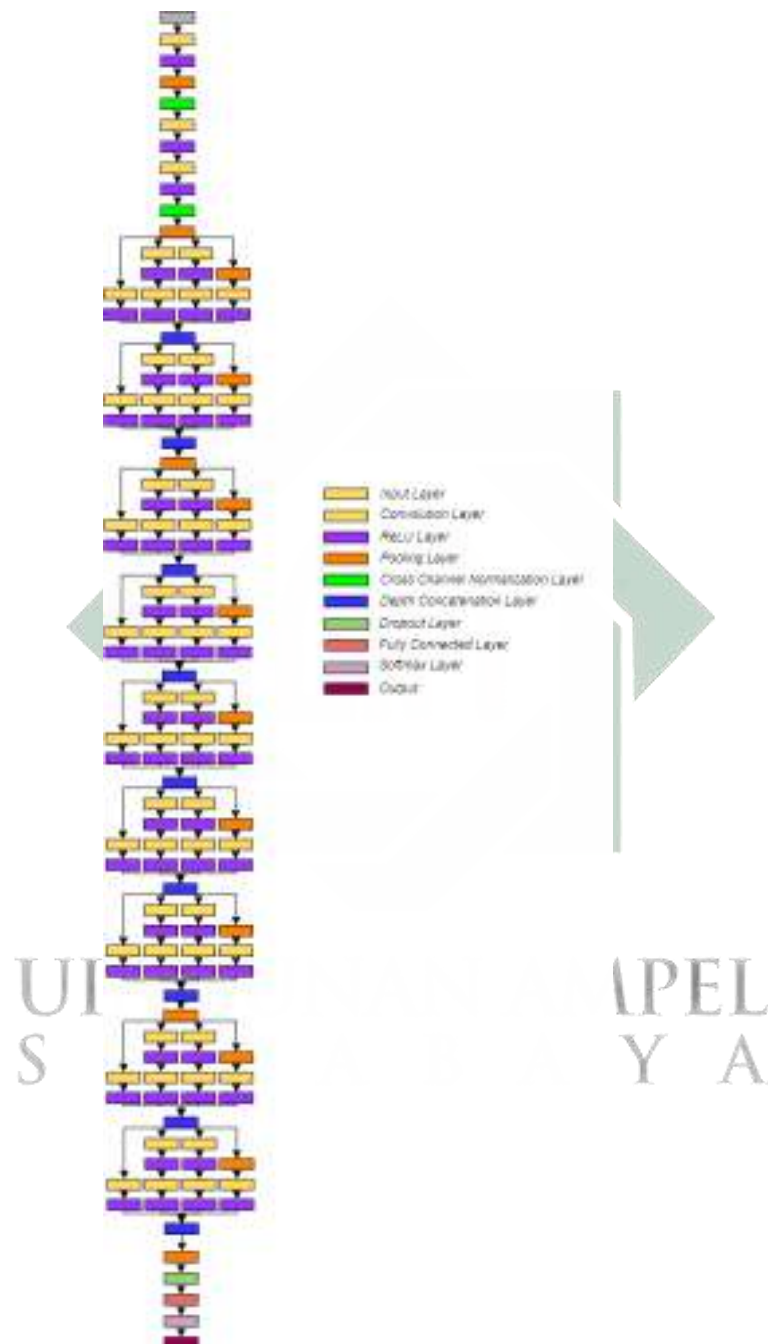
Perhitungan *softmax layer* ini ditunjukkan oleh Persamaan 2.8.

$$p_j(y) = \frac{e^{y_j}}{\sum_{i=1}^k e^{y_i}} \quad (2.8)$$

Berdasarkan Persamaan 2.8 diketahui  $p_j(y)$  merupakan probabilitas dari  $y_j$  atau kelas yang dicari yang diperoleh dari exponential  $y_j$  dibagi dengan total penjumlahan exponential  $y_i$  probabilitas semua kelas  $k$ .

## 2.5. GoogLeNet

GoogLeNet merupakan model CNN yang pertama kali dikenalkan oleh Google pada tahun 2014. CNN model ini memiliki *inception modules* yang terdiri atas sejumlah *convolution* kecil (Fattah, 2021). Akurasi atas model ini akan tinggi jika kinerja pada jaringan sarafnya meningkat yaitu dengan memperdalam *layer*. Algoritma CNN model GoogLeNet memiliki 144 layer dimana terdiri atas *convolutional layer*, *ReLU layer* *pooling layer*, serta klasifikasi (Jawahar et al., 2022). Ilustrasi arsitektur pada CNN model GoogLeNet ditunjukkan oleh Gambar 2.11.



Gambar 2.11 Arsitektur GoogLeNet

## 2.6. Confusion Matrix

Proses evaluasi dalam melakukan klasifikasi suatu hal merupakan salah satu proses penting (Normawati and Prayogi, 2021). Umumnya, proses ini menggunakan

*confusion matrix* yang digunakan untuk melihat hasil klasifikasi data yang teruji benar dan salah. Selain itu, *confusion matrix* juga dapat digunakan untuk menentukan nilai persentase keakuratan, sensitivitas, serta spesifisitas suatu model dengan menggunakan 4 istilah yang mempresentasikan hasil klasifikasinya (Endah and Encis, 2021). Istilah yang digunakan yaitu *True Positive* (TP) dimana sejumlah data yang terklasifikasi benar pada kedua kelas, *True Negative* (TN) dimana sejumlah data yang terklasifikasi salah pada kedua kelas, *False Positive* (FP) dimana sejumlah data yang terklasifikasi salah pada kelas aktual dan benar pada kelas prediksi, dan *False Negative* (FN) dimana sejumlah data yang terklasifikasi benar pada kelas aktual dan salah pada kelas prediksi (Chicco et al., 2021). Pada penelitian ini, menggunakan *confusion matrix multi class* dimana terdapat kelas *benign lung tissue*, *lung adenocarcinoma*, dan *lung squamous cell carcinoma*. Tabel *confusion matrix multi class* ditunjukkan pada Gambar 2.12.

The figure displays five confusion matrix tables arranged in two rows. Each table compares 'Kelas Aktual' (Actual Class) on the y-axis against 'Kelas Target' (Target Class) on the x-axis. The target classes are Adenocarcinoma, Benign, and Squamous Cell. The actual classes are Adenocarcinoma, Benign, and Squamous Cell. The cells in the matrices contain numerical values representing the count of data points for each combination of actual and target classes.

Kelas Aktual	Kelas Target		
	Adenocarcinoma	Benign	Squamous Cell
Adenocarcinoma	17	16	17
Benign	17	17	17
Squamous Cell	17	16	17

Kelas Aktual	Kelas Target		
	Adenocarcinoma	Benign	Squamous Cell
Adenocarcinoma	17	16	16
Benign	16	17	16
Squamous Cell	16	16	17

Kelas Aktual	Kelas Target		
	Adenocarcinoma	Benign	Squamous Cell
Adenocarcinoma	16	16	16
Benign	16	17	16
Squamous Cell	16	16	17

Kelas Aktual	Kelas Target		
	Adenocarcinoma	Benign	Squamous Cell
Adenocarcinoma	16	16	16
Benign	16	16	16
Squamous Cell	16	16	16

Kelas Aktual	Kelas Target		
	Adenocarcinoma	Benign	Squamous Cell
Adenocarcinoma	16	16	16
Benign	16	16	16
Squamous Cell	16	16	16

Gambar 2.12 *Confusion Matrix Multi Class*

Berdasarkan gambar di atas, maka dapat dilakukan perhitungan nilai akurasi, sensitivitas, serta spesifisitas atas hasil klasifikasi. Akurasi digunakan untuk mengukur tingkat keakuratan suatu sistem. Sensitivitas digunakan untuk memperoleh jumlah data yang terklasifikasi benar ke dalam masing-masing kelas kanker paru. Sedangkan, spesifisitas digunakan untuk memperoleh jumlah data yang terklasifikasi benar pada kelas kanker paru yang lain. Perhitungan nilai akurasi, <http://digilib.uinsa.ac.id/> <http://digilib.uinsa.ac.id/> <http://digilib.uinsa.ac.id/>

sensitivitas, serta spesifisitas ditunjukkan oleh persamaan [2.9](#), [2.10](#), [2.11](#).

$$Akurasi = \frac{TP_{all}}{n_{all}} \times 100\% \quad (2.9)$$

$$Sensitivitas = \frac{\sum \left( \frac{TP}{TP+FN} \right)}{n} \times 100\% \quad (2.10)$$

$$Spesifisitas = \frac{\sum \left( \frac{TN}{TN+FP} \right)}{n} \times 100\% \quad (2.11)$$

Pada Persamaan [2.9](#)  $TP_{all}$  merupakan total keseluruhan semua data yang termasuk kedalam *True Positive*, dan  $n_{all}$  merupakan total keseluruhan data yang digunakan, sedangkan  $n$  merupakan total kelas yang ada.

## 2.7. Integrasi Keislaman

### 2.7.1. Penyakit dalam Perspektif Islam

Manusia senantiasa dipertemukan pada tiga situasi dalam hidupnya, yakni situasi sehat, sakit, serta mati. Pada situasi sehat atau dalam bahasa arab disebut dengan *as-shihah* terkadang membuat manusia lupa untuk menjaganya. Sedangkan, situasi sakit yang dalam bahasa arab disebut *maradl* atau *as-saqham* manusia menganggapnya bahwa itu adalah cobaan yang sangat berat. Namun, manusia terkadang tidak menyadari bahwa Allah SWT memberinya penyakit bukan hanya untuk mengujinya saja tetapi juga untuk mendekatkan hamba-Nya pada-Nya serta sebagai salah satu cara menggugurkan dosa hamba-Nya. Hal ini juga dijelaskan dalam hadits Bukhari nomor 5660 yang berbunyi:

مَا مِنْ مُسْلِمٍ يَصِيبُهُ أَدَى مِنْ مَرَضٍ فَمَا سِوَاهُ إِلَّا حَطَّ اللَّهُ بِهِ سَيِّئَاتِهِ كَمَا تَحُطُّ الشَّجَرَةُ وَرَقَهَا

artinya : ”Tidaklah seorang muslim tertimpa suatu penyakit dan sejenisnya, melainkan Allah akan menggugurkan bersamanya dosa-dosanya seperti pohon yang

menggugurkan daun-daunnya.”(HR Bukhari No. 5660)

Pada hadits diatas dijelaskan bahwasanya saat seorang muslim yang tertimpa penyakit maka saat itulah Allah SWT juga akan menggugurkan dosa-dosanya. Selain itu, menurut persepektif islam penyakit sendiri adalah pemberian Allah SWT kepada hamba-Nya guna menguji tingkat keimanannya. Cobaan berupa penyakit ini supaya kita senantiasa mengingat Allah SWT. Selain mengingat Allah SWT, kita juga harus ingat bahwasanya segala penyakit pasti ada obatnya. Hal ini selaras dengan hadits Bukhari nomor 5246 yang berbunyi:

حَدَّثَنَا مُحَمَّدُ بْنُ الْمُثَنَّى حَدَّثَنَا أَبُو أَحْمَدَ الزُّبَيْرِيُّ حَدَّثَنَا عُمَرُ بْنُ سَعِيدٍ بْنُ أَبِي حَسَنٍ قَالَ حَدَّثَنِي عَطَاءُ بْنُ أَبِي رِيحٍ  
عَنْ أَبِي هُرَيْرَةَ رَضِيَ اللَّهُ عَنْهُ عَنِ النَّبِيِّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ مَا أَنْزَلَ اللَّهُ دَاءً إِلَّا أَنْزَلَ لَهُ شِفَاءً

artinya: ”Telah menceritakan kepada kami Muhammad bin Al Mutsanna telah menceritakan kepada kami Abu Ahmad Az Zubairi telah menceritakan kepada kami 'Umar bin Sa'id bin Abu Husain dia berkata: telah menceritakan kepadaku 'Atha' bin Abu Rabah dari Abu Hurairah radliallahu 'anhu dari Nabi shallallahu 'alaihi wasallam beliau bersabda: ”Allah tidak akan menurunkan penyakit melainkan menurunkan obatnya juga.”(HR Bukhari No.5246)

Sebagai muslim kita dianjurkan untuk berusaha mengobati penyakit yang diderita karena pada hadits diatas telah dijelaskan bahwa Allah SWT menurunkan penyakit maka diturunkan pula obatnya.

### 2.7.2. Kewajiban Orang Sakit

Cobaan yang Allah SWT berikan pasti mengandung hikmah, sekalipun cobaan tersebut berupa penyakit. Cobaan berupa penyakit ini dapat menimpa siapa saja dan dimana saja. Namun, kita sebagai manusia harus bisa menerima dan menjalaninya dengan lapang dada. Hal ini selaras dengan kewajiban orang yang

sakit dimana mereka setidaknya harus menjalani 4 kewajiban berikut yakni ikhlas, sabar, berikhtiar, serta berdoa.

### 1. Ikhlas

Jika kita mendapat cobaan dari Allah SWT baik cobaan tersebut berat ataupun tidak, yang pertama harus kita lakukan yaitu berusaha untuk ikhlas. Hal ini dikarenakan kita harus percaya kepada Allah SWT atas semua yang diberikan-Nya kepada kita karena itu adalah hal terbaik untuk kita. Pada Qur'an Surah At-Taghobun ayat 11 yang berbunyi:

مَا أَصَابَ مِنْ مُصِيبَةٍ إِلَّا بِإِذْنِ اللَّهِ وَمَنْ يُؤْمِنْ بِاللَّهِ يَهْدِ اللَّهُ قَلْبَهُ  
وَاللَّهُ بِكُلِّ شَيْءٍ عَلِيمٌ

artinya: "Tidak ada sesuatu musibah yang menimpa (seseorang), kecuali dengan izin Allah, dan barang siapa beriman kepada Allah, niscaya Allah akan memberi petunjuk kepada hatinya. Dan Allah Maha Mengetahui segala sesuatu."(QS. At-Taghobun:11)

Berdasarkan ayat di atas dapat disimpulkan bahwa segala sesuatu yang terjadi pada hidup kita itu semua atas kehendak serta izin Allah SWT. Walaupun berupa cobaan sesungguhnya Allah SWT tidak akan memberikan cobaan kepada hamba-Nya diluar batas kemampuannya. Maka, dihibau untuk senantiasa beriman kepada Allah SWT karena orang-orang yang beriman akan selalu diberikan pertolongan.

### 2. Sabar

Selain ikhlas, sabar merupakan salah satu kewajiban yang harus dilakukan bagi orang yang sedang tertimpa penyakit. Sabar juga merupakan salah satu

kunci atas kesembuhan, karena dalam proses penyembuhan membutuhkan waktu. Oleh karena itu, jika kita tidak sabar kita akan merasa bahwa Allah SWT memberikan cobaan yang sangat berat kepada kita. Sikap sabar telah dijelaskan dalam hadits Tirmidzi dan Ibnu Majah yang berbunyi:

إِنَّ عَظَمَ الْجَزَاءِ مَعَ عَظَمِ الْبَلَاءِ وَإِنَّ اللَّهَ إِذَا أَحَبَّ قَوْمًا ابْتَلَاهُمْ فَمَنْ رَضِيَ قَلْبَهُ الرِّضَا وَمَنْ سَخَطَ قَلْبَهُ السَّخَطَ

artinya: "Sesungguhnya besarnya pahala itu sesuai dengan besarnya ujian, dan bahwa Allah, apabila menyayangi atau mencintai suatu kaum, maka Allah akan mengujinya, dan bagi siapa saja ridha, maka baginya keridhaan dari Allah, dan barang siapa yang membenci, maka baginya kebencian dari Allah Swt."(HR Tirmidzi No. 2320 dan Ibnu Majah No. 4021)

Pada hadits tersebut dijelaskan bahwasanya Allah SWT menguji hamba-Nya karena Allah SWT menyayangi atau mencintainya. Oleh karena itu, sabar sangat penting untuk dilakukan bagi kita yang sedang sakit karena itu berarti Allah SWT menyayangi kita.

### 3. Ikhtiar

Pada dasarnya segala penyakit yang Allah SWT turunkan untuk hamba-Nya semuanya pasti ada obatnya. Oleh karena itu, tugas kita sebagai hamba-Nya yang sedang tertimpa sakit yaitu dengan berikhtiar atau berusaha untuk mencari obat tersebut. Hal ini dikarenakan Allah SWT memberikan kesembuhan kepada hamba-Nya melalui berbagai perantara, salah satunya dokter. Berikhtiar juga dijelaskan dalam Qur'an Surah Asy-Syu'ara ayat 80 yang berbunyi:

وَإِذَا مَرِضْتُ فَهُوَ يَشْفِينِ



artinya: “Dan apabila aku sakit, Dialah Yang menyembuhkan aku.” (QS Asy-Syu’ara: 80)

Ayat diatas menjelaskan bahwasanya Allah SWT yang menyembuhkan atas segala penyakit kita. Oleh karena itu, kita harus berikhtiar semaksimal mungkin, dan percaya bahwa hasil dari ikhtiar yang kita lakukan akan mendapatkan yang terbaik.

#### 4. Berdoa

Berdoa merupakan kewajiban terakhir yang harus dilakukan setelah ikhlas, sabar, dan ikhtiar. Saat tertimpa musibah, sebagai hamba-Nya kita senantiasa harus menerimanya dengan ikhlas, menjalani dengan sabar, selalu berikhtiar untuk memperoleh hasil terbaik, dan berdoa agar diberikan hasil yang terbaik atas ikhtiar yang telah dilakukan. Hal ini sebagaimana dijelaskan dalam Qur’an Surah Ghafir ayat 60 yang berbunyi:

U وَقَالَ رَبُّكُمْ ادْعُونِي أَسْتَجِبْ لَكُمْ إِنَّ الَّذِينَ يَسْتَكْبِرُونَ عَنْ  
S عِبَادَتِي سَيَدْخُلُونَ جَهَنَّمَ دَاخِرِينَ ؕ

artinya: “Dan Tuhanmu berfirman, “Berdoalah kepada-Ku, niscaya akan Aku perkenankan bagimu. Sesungguhnya orang-orang yang sombong tidak mau menyembah-Ku akan masuk neraka Jahanam dalam keadaan hina dina.” (QS Ghafir: 60)

Pada Qur’an Surah Ghafir ayat 60 menjelaskan bahwasanya Allah SWT senantiasa akan membantu hamba-Nya yang meminta kepada-Nya. Oleh karena itu, maka doa merupakan kewajiban terakhir yang harus senantiasa dilakukan karena sesungguhnya kesembuhan datangnya dari Allah SWT.

## BAB III

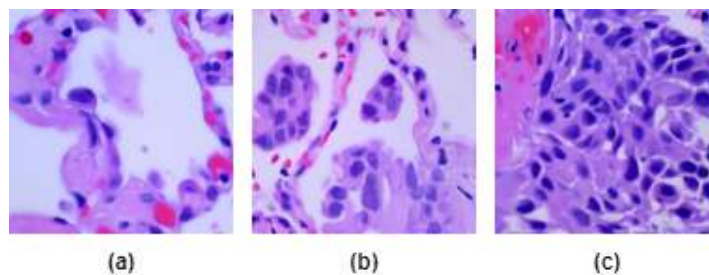
### METODOLOGI PENELITIAN

#### 3.1. Jenis Penelitian

Penelitian ini dikategorikan ke dalam jenis penelitian deskriptif kuantitatif. Pada penelitian ini data yang digunakan merupakan data citra histopatologi kanker paru. Berdasarkan data tersebut kemudian dilakukan pengolahan sehingga menghasilkan sebuah matriks, dimana pada matriks tersebut berisi nilai-nilai piksel yang ada pada citra tersebut. Kemudian, dilakukan proses perhitungan dengan menggunakan matriks tersebut hingga memperoleh hasil diagnosis pada masing-masing citra.

#### 3.2. Sumber Data

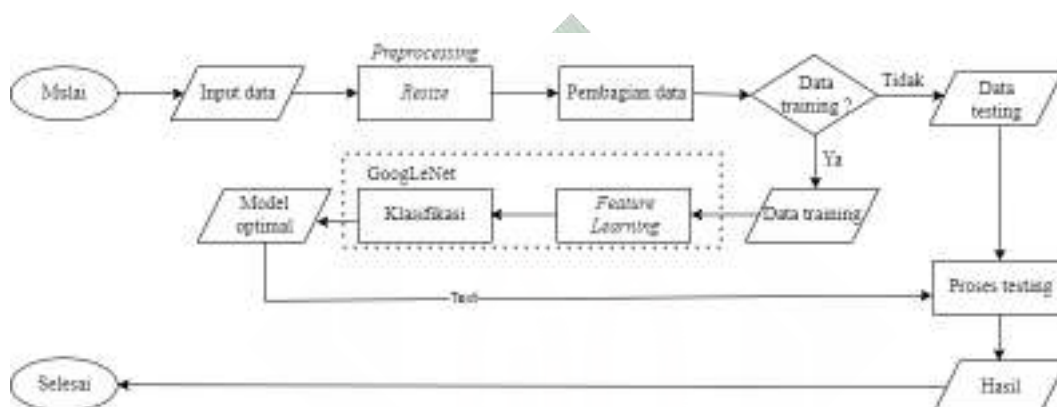
Pada penelitian ini, data yang digunakan diperoleh dari penelitian (Borkowski et al., 2019) dengan jumlah data sebanyak 15000 data yang terbagi menjadi 3 kelas yaitu *benign lung tissue*, *lung adenocarcinoma*, dan *lung squamous cell carcinoma* dimana masing-masing kelas berisi 5000 data. Sampel data citra kanker paru ditunjukkan oleh Gambar 3.1.



Gambar 3.1 (a) *Lung adenocarcinoma* (b) *Benign lung tissue* (c) *Lung squamous cell carcinoma*

### 3.3. Kerangka Penelitian

Pada penelitian ini terdapat beberapa proses penelitian dalam melakukan klasifikasi kanker paru dengan menggunakan CNN model GoogLeNet. Proses yang dilakukan ditunjukkan oleh Gambar 3.2.



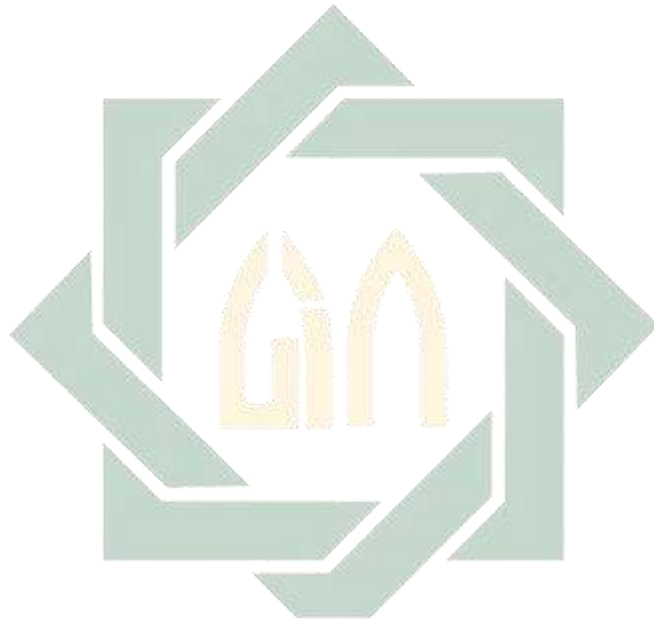
Gambar 3.2 Alur Penelitian

Pada penelitian ini terdapat beberapa tahapan yang dilakukan, berikut tahapannya yaitu:

1. Tahap *preprocessing* dilakukan proses *resizing* pada data citra histopatologi kanker paru yang semula berukuran  $768 \times 768$  piksel menjadi  $224 \times 224$  menyesuaikan dengan ukuran inputan data pada model GoogLeNet. Kemudian, pembagian data dengan beberapa persentase data *training* yang kemudian akan diuji coba dengan sistem yaitu pembagian data 60:40, pembagian data 70:30, pembagian data 80:20, serta pembagian data 90:10.
2. Tahap selanjutnya dengan menggunakan metode CNN model GoogLeNet yaitu terdapat proses *feature learning* yang digunakan untuk memperoleh nilai dari masing-masing fitur pada citra kemudian terdapat klasifikasi.
3. Tahap selanjutnya yaitu dilakukan uji coba parameter dengan beberapa variasi

*dropout* dan *mini batchsize*. Setelah itu, dilanjutkan dengan proses klasifikasi.

4. Tahap selanjutnya yaitu evaluasi hasil dengan menggunakan *confusion matrix* untuk memperoleh nilai akurasi, sensitivitas, dan spesifitas suatu model.



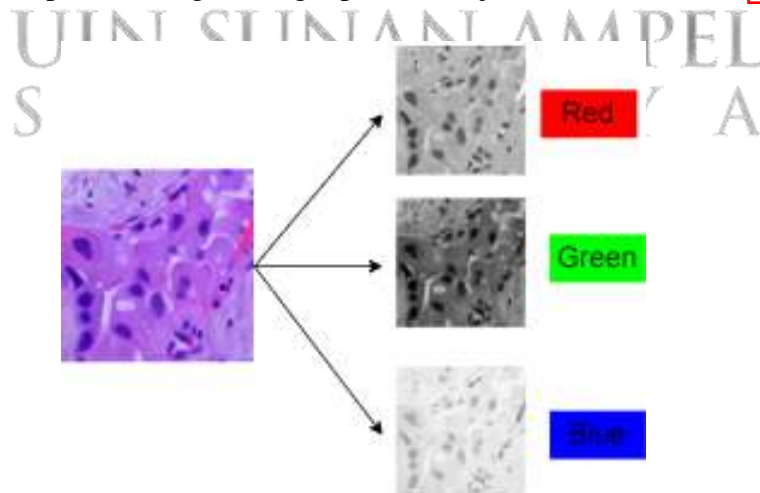
UIN SUNAN AMPEL  
S U R A B A Y A

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1. Deskripsi Data

Penelitian ini menggunakan data citra histopatologi kanker paru yang sejumlah 15000 data dan terdiri dari 3 kelas yaitu kelas *Adenocarcinoma*, *Benign*, dan *Lung Squamous Cell Carcinoma* dimana masing-masing kelas berisi 5000 data. Data yang digunakan dalam penelitian ini diperoleh dari web *Cornell University* dengan nama *LC25000 dataset* dari negara Amerika. Pada penelitian ini, data citra histopatologi kanker paru merupakan suatu citra RGB yakni suatu citra yang berukuran  $768 \times 768$  piksel dengan 3 *channel* warna yaitu *Red* (R), *Green* (G), *Blue* (B). Data ggcitra pada masing-masing lapisan ditunjukkan oleh Gambar 4.1.



Gambar 4.1 Lapisan citra RGB kanker paru

Pada masing-masing lapisan pada citra tersebut memiliki nilai piksel masing-masing. Nilai piksel pada masing-masing lapisan tersebut berada pada rentang 0-255.

Nilai-nilai tersebut ditunjukkan oleh Gambar 4.2.

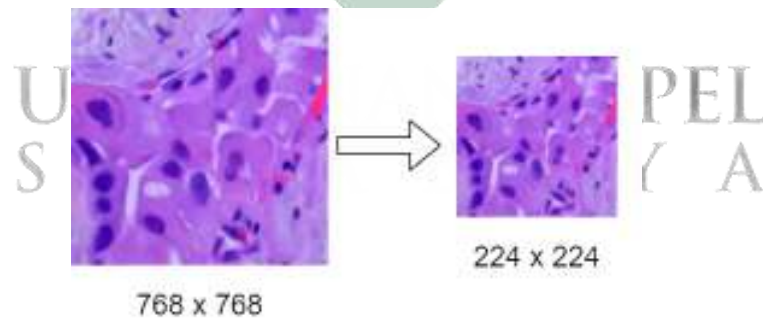
	1	2	3	4	5	...	768	767	768
1	188	185	184	185	186	...	152	154	156
2	187	184	186	188	187	...	154	155	156
3	188	186	186	186	187	...	155	157	157
4	190	187	187	187	188	...	155	156	155
5	190	187	187	187	188	...	155	154	153
...	...	...	...	...	...	...	...	...	...
766	81	79	83	82	85	...	171	178	179
767	76	74	76	76	75	...	171	177	181
768	75	73	71	70	72	...	172	178	182

	1	2	3	4	5	...	768	767	768
1	195	195	195	196	197	...	88	86	88
2	193	195	196	196	197	...	88	87	88
3	194	196	196	196	198	...	87	88	88
4	195	195	196	197	198	...	87	85	88
5	196	196	197	197	198	...	88	88	82
...	...	...	...	...	...	...	...	...	...
766	107	108	107	107	101	...	138	143	146
767	98	100	100	99	97	...	138	144	148
768	97	97	88	88	85	...	139	145	149

	1	2	3	4	5	...	768	767	768
1	240	240	240	241	241	...	223	225	227
2	240	240	241	241	241	...	225	226	227
3	240	241	241	241	242	...	226	228	228
4	241	241	241	242	242	...	226	227	226
5	241	241	242	242	242	...	226	225	224
...	...	...	...	...	...	...	...	...	...
766	215	216	216	216	213	...	229	232	235
767	214	215	213	213	212	...	227	233	237
768	212	212	211	211	210	...	228	234	238

Gambar 4.2 (a)Lapisan Red (b)Lapisan Green (c) Lapisan Blue

Pada penelitian menggunakan metode CNN model GoogLeNet yang memiliki ketentuan untuk ukuran citra yang diinputkan yaitu citra berukuran  $224 \times 224 \times 3$ . Hal ini mengakibatkan citra perlu melakukan proses *resize*, dimana pada penelitian ini metode *resize* yang digunakan yaitu dengan metode *bicubic interpolation*. Ilustrasi perubahan citra yang telah melakukan proses *resize* ditunjukkan oleh Gambar 4.3.



Gambar 4.3 Ilustrasi Proses Resizing

Perubahan ukuran pada citra juga berpengaruh pada nilai piksel dari citra tersebut. Perubahan nilai piksel citra yang telah dilakukan proses *resize* ditunjukkan oleh Gambar 4.4.

	1	2	3	4	5	...	222	223	224
1	200	210	210	211	215	...	252	252	258
2	205	211	212	212	215	...	248	249	254
3	208	212	214	215	214	...	252	253	252
4	200	210	215	215	212	...	267	262	250
5	200	214	216	211	207	...	284	275	270
...	1	1	1	1	1	...	1	1	1
222	185	178	176	180	187	...	157	161	166
223	183	178	172	176	182	...	157	160	172
224	179	174	171	171	177	...	161	164	176

	1	2	3	4	5	...	122	123	124
1	199	200	197	198	202	...	79	80	86
2	198	200	199	199	201	...	80	81	85
3	198	201	201	200	200	...	80	80	81
4	199	202	202	199	198	...	113	109	106
5	197	202	202	197	192	...	140	134	129
...	1	1	1	1	1	...	1	1	1
122	114	107	104	110	120	...	123	127	135
123	108	105	98	105	114	...	121	124	135
124	102	98	94	97	106	...	125	126	138

	1	2	3	4	5	...	222	223	224
1	236	237	240	242	248	...	221	226	231
2	238	240	242	243	237	...	237	230	227
3	240	243	245	246	251	...	237	238	220
4	241	245	248	248	251	...	237	233	220
5	242	248	251	248	246	...	237	230	225
...	1	1	1	1	1	...	1	1	1
222	216	213	214	225	216	...	238	237	230
223	211	208	209	218	251	...	239	224	255
224	206	204	206	252	224	...	225	229	241

Gambar 4.4 (a) Lapisan Red

(b) Lapisan Green

(c) Lapisan Blue

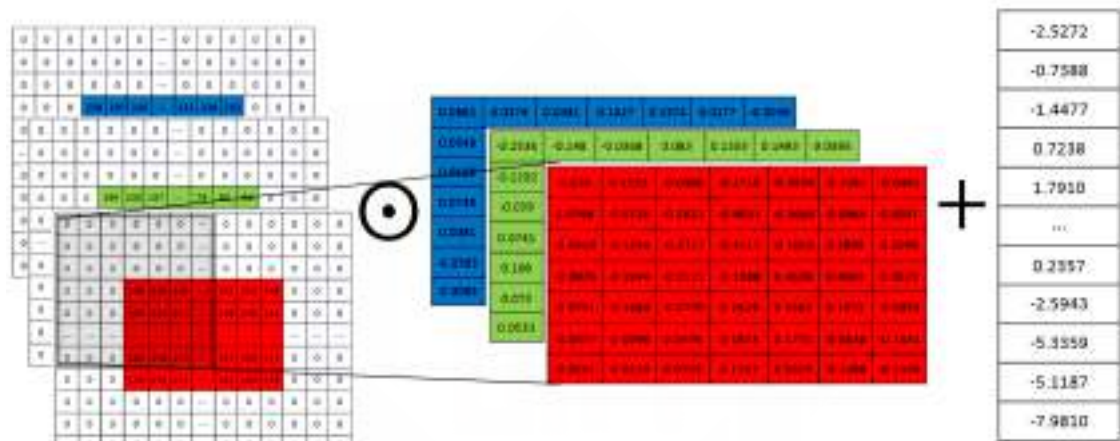
## 4.2. Convolutional Neural Network

Penelitian ini menggunakan metode CNN model GoogLeNet untuk mengklasifikasikan kanker paru berdasarkan citra histopatologi. GoogLeNet merupakan salah satu jenis arsitektur pada CNN yang memiliki ketentuan ukuran *input* citra sebesar  $224 \times 224 \times 3$ . CNN sendiri memiliki 3 komponen utama yaitu *input*, *feature learning*, dan *classification*. Pada komponen *feature learning* digunakan untuk mempelajari mengenai citra secara mendalam. Pembelajaran citra secara mendalam ini dengan beberapa tahapan diantaranya yaitu *Convolution Layer*, *ReLU Layer*, *Pooling Layer*, *Cross Channel Normalization Layer*, serta *Dropout Layer*. Lalu, pada komponen *classification* terdapat beberapa tahapan yaitu *Fully Connected Layer*, dan *Softmax Layer*. Berikut layer-layer yang ada pada GoogLeNet.

### 4.2.1. Convolution Layer

*Convolution layer* merupakan layer yang bertugas sebagai penerima input citra serta untuk mempelajari citra secara lebih mendalam. Pada layer ini dilakukan operasi konvolusi dengan menggunakan Persamaan 2.2. Operasi konvolusi yang dilakukan yaitu antara input citra yang berukuran  $224 \times 224 \times 3$  dengan *padding*

sejumlah 3 pada masing-masing sisi citra tersebut dengan sejumlah 64 filter yang berukuran  $7 \times 7 \times 3$  dengan nilai masing-masing filter random dengan rentang 0-1. Kemudian, ditambahkan dengan bias sejumlah filter yakni 64. Ilustrasi operasi konvolusi pada *convolution layer* ditunjukkan oleh Gambar 4.5.



Gambar 4.5 Ilustrasi Operasi Konvolusi

Berdasarkan Gambar 4.5 jumlah *feature map* yang akan dihasilkan dapat dihitung dengan menggunakan Persamaan 2.3.

$$H_Q = \frac{I - F + 2p}{s} + 1$$

$$H_Q = \frac{224 - 7 + 2(3)}{2} + 1$$

$$H_Q = 112$$

Berdasarkan perhitungan dengan menggunakan Persamaan 2.3 maka dapat diketahui bahwa pada operasi konvolusi yang dilakukan pada *convolution layer* akan menghasilkan *feature map* sejumlah 112. Kemudian, untuk contoh perhitungan operasi konvolusi pada *feature map* dengan menggunakan Persamaan 2.2 sebagai



berikut:

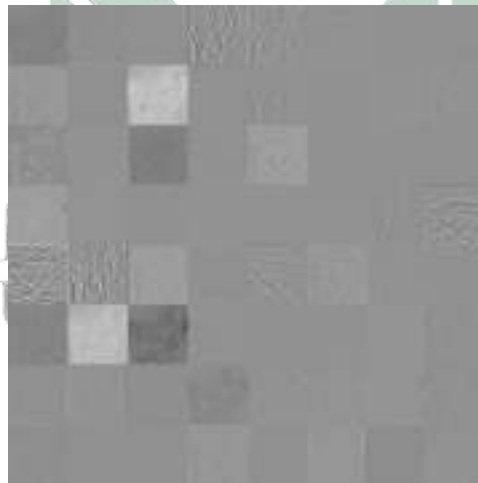
$$\begin{aligned}
 Q_{(i,j)} &= \left( \sum_{i=1}^N I_{(i,j)} \odot F_{(i,j)} \right) + B_j \\
 Q_{(1,1)} &= (I_{1,1}(1,1) \times F_{1,1}(1,1) + I_{1,1}(1,2) \times F_{1,1}(1,2) + \dots + I_{1,1}(7,7) \times F_{1,1}(7,7)) + \\
 &\quad (I_{2,1}(1,1) \times F_{2,1}(1,1) + I_{2,1}(1,2) \times F_{2,1}(1,2) + \dots + I_{2,1}(7,7) \times F_{2,1}(7,7)) + \\
 &\quad (I_{3,1}(1,1) \times F_{3,1}(1,1) + I_{3,1}(1,2) \times F_{3,1}(1,2) + \dots + I_{3,1}(7,7) \times F_{3,1}(7,7)) + B_1 \\
 &= (0 \times 0.2290) + (0 \times 0.1553) + \dots + (213 \times -0.1103) + (0 \times -0.2536) + \\
 &\quad (0 \times -0.1480) + \dots + (199 \times -0.3033) + (0 \times 0.0882) + \\
 &\quad (0 \times 0.0374) + \dots + (248 \times 0.4041) + (-2.5272) \\
 Q_{(1,1)} &= -1.4305 \\
 Q_{(1,2)} &= (I_{1,1}(1,3) \times F_{1,1}(1,1) + I_{1,1}(1,4) \times F_{1,1}(1,2) + \dots + I_{1,1}(10,7) \times F_{1,1}(7,7)) + \\
 &\quad (I_{2,1}(1,3) \times F_{2,1}(1,1) + I_{2,1}(1,5) \times F_{2,1}(1,2) + \dots + I_{2,1}(10,7) \times F_{2,1}(7,7)) + \\
 &\quad (I_{3,1}(1,3) \times F_{3,1}(1,1) + I_{3,1}(1,5) \times F_{3,1}(1,2) + \dots + I_{3,1}(10,7) \times F_{3,1}(7,7)) + B_1 \\
 &= (0 \times 0.2290) + (0 \times 0.1553) + \dots + (210 \times -0.1103) + (0 \times -0.2536) + \\
 &\quad (0 \times -0.1480) + \dots + (195 \times -0.3033) + (0 \times 0.0882) + \\
 &\quad (0 \times 0.0374) + \dots + (250 \times 0.4041) + (-2.5272) \\
 Q_{(1,2)} &= 19.0630
 \end{aligned}$$

Perhitungan yang telah dilakukan akan menghasilkan satu *feature map* dengan ukuran  $112 \times 112$ . *Feature map* yang dihasilkan ditunjukkan oleh matriks berikut:

$$Q_{(:, :, 1)} =$$

$$\begin{bmatrix} -1.4305 & 19.0630 & 25.8385 & \dots & -45.8200 & -57.5566 & -108.3773 \\ 45.5056 & 57.7826 & 53.7228 & \dots & -151.26910 & -156.7180 & -114.0224 \\ 45.4052 & 54.1394 & 48.3805 & \dots & -113.1859 & -105.2100 & -23.6661 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 16.3462 & -29.8521 & -67.1338 & \dots & 24.3220 & 39.8632 & 13.0504 \\ 6.6718 & -55.9484 & -82.0257 & \dots & 20.5665 & 24.4892 & -7.2761 \\ -55.6758 & -130.4140 & -129.6125 & \dots & -13.9220 & -30.2208 & -18.2891 \end{bmatrix}$$

Perhitungan ini dilakukan sebanyak 64 kali yakni sesuai dengan jumlah filter. Jika perhitungan sudah dilakukan sebanyak 64 kali atau sejumlah filter, maka hasil perhitungan tersebut akan divisualisasikan seperti yang ditunjukkan oleh Gambar 4.6.



Gambar 4.6 Visualisasi *Feature Map* Pada *Convolution Layer*

#### 4.2.2. *ReLu Layer*

Pada layer ini hasil dari proses pada layer *convolution layer* kemudian diaktivasi. Fungsi aktivasi yang digunakan pada metode CNN model GoogLeNet yaitu fungsi aktivasi ReLU. Fungsi aktivasi ReLU ini berfungsi untuk menghilangkan negatif yang terdapat pada hasil dari proses *convolution layer*.

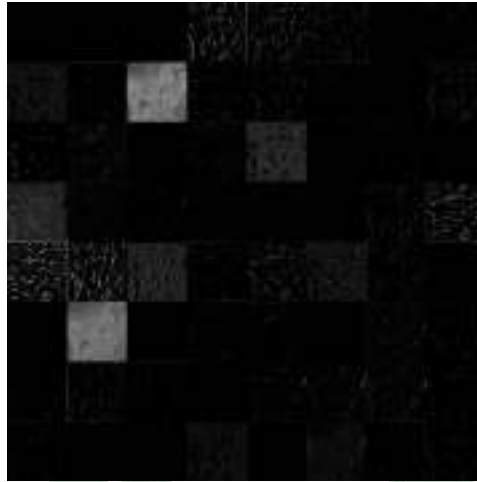
Penghilangan nilai negatif pada layer ini juga bertujuan untuk menghindari adanya *vanishing gradien* atau suatu kondisi dimana sebuah model yang tidak dapat mempelajari suatu pola baru karena memiliki nilai gradien yang semakin kecil. Perhitungan pada layer ini yaitu dengan menggunakan Persamaan 2.4. Berikut contoh perhitungan pada ReLU layer dengan  $z$  merupakan nilai yang didapat dari proses *convolution layer*.

$$\begin{aligned}
 R(z) &= \max(0, z) \\
 R(-1.4305) &= \max(0, -1.4305) = 0 \\
 R(25.8385) &= \max(0, 25.8385) = 25.8385 \\
 R(-30.2208) &= \max(0, -30.2208) = 0 \\
 R(-18.2891) &= \max(0, -18.2891) = 0
 \end{aligned}$$

Perhitungan tersebut dilakukan pada semua nilai yang didapatkan dari proses *convolution layer*. Hasil dari ReLU layer ditunjukkan oleh matriks berikut.

$$R(Q_{(:,1)}) = \begin{bmatrix} 0 & 19.0603 & 25.8385 & \dots & 0 & 0 & 0 \\ 45.5056 & 57.7826 & 53.7228 & \dots & 0 & 0 & 0 \\ 45.4052 & 54.1394 & 48.3805 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 16.3462 & 0 & 0 & \dots & 24.3220 & 39.8632 & 13.0504 \\ 6.6718 & 0 & 0 & \dots & 20.5665 & 24.4892 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

Perhitungan ini dilakukan sebanyak 64 kali yakni sesuai dengan jumlah lapisan yang dihasilkan pada layer sebelumnya. Visualisasi *feature map* dari ReLU layer ditunjukkan oleh Gambar 4.7.



Gambar 4.7 Visualisasi *Feature Map* Pada ReLU Layer

#### 4.2.3. Pooling Layer

*Pooling layer* adalah suatu layer yang digunakan untuk mempercepat waktu komputasi dengan menggunakan teknik *downsampling*. Pada CNN model GoogLeNet, layer ini berada pada urutan keempat. Pada layer ini yang merupakan layer yang berada setelah ReLU layer maka jenis *pooling* yang digunakan yakni *maxpooling* dengan *stride 2* dan *padding 0*. Berikut contoh perhitungan pada proses *pooling layer*.

$$Q_{1,1} = \max(0, 19.0603, 25.8385, 45.5056, 57.7826, 53.7228, 45.4052, 54.1394, 48.3805)$$

$$Q_{1,56} = \max(0, 0, 0, 0, 0, 0, 0, 0, 0)$$

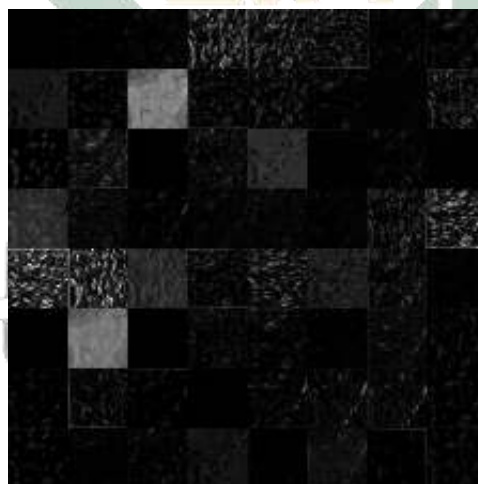
$$Q_{56,1} = \max(16.3462, 0, 0, 6.6718, 0, 0, 0, 0, 0)$$

$$Q_{56,56} = \max(24.3220, 39.8632, 13.0504, 20.5665, 24.4892, 0, 0, 0, 0)$$

Hasil perhitungan pada *pooling layer* ditunjukkan oleh matriks berikut.

$$(Q_{(:, :, 1)}) = \begin{bmatrix} 57.7825 & 53.7228 & 68.3575 & \dots & 0 & 0 & 0 \\ 54.1393 & 48.3805 & 62.0838 & \dots & 0 & 31.5559 & 66.4840 \\ 47.6417 & 48.6969 & 48.6969 & \dots & 0.3260 & 47.3643 & 66.4840 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 14.8293 & 0 & 0 & \dots & 0 & 38.5232 & 39.5537 \\ 16.3461 & 0 & 0 & \dots & 6.1899 & 39.8632 & 39.8632 \\ 6.6718 & 21.2253 & 14.7424 & \dots & 6.1899 & 24.4892 & 24.4892 \end{bmatrix}$$

Pada proses ini memperoleh sebanyak 64 *feature map* dimana satu *feature map* berukuran  $56 \times 56$ . *Feature map* yang dihasilkan ditunjukkan oleh Gambar yang merupakan visualisasi dari *pooling layer* dengan jenis *maxpooling*.



Gambar 4.8 Visualisasi *Feature Map* Pada *Pooling Layer*

#### 4.2.4. *Cross Channel Normalization Layer*

Layer ini merupakan layer yang digunakan untuk melakukan normalisasi terhadap nilai pada *feature map*. Selain itu, layer ini juga digunakan untuk membatasi nilai yang tak hingga yang diperoleh dari ReLu layer. Pada layer ini terdapat beberapa *hyperparameter* yang digunakan yaitu  $k$ ,  $\alpha$ ,  $\beta$ , dan  $n$ . Inisialisasi *hyperparameter*

$k, \alpha, \beta$  dan  $n$  berturut-turut yaitu  $(1, 0, 0001, 0.75, 5)$ . Perhitungan pada layer ini dengan menggunakan Persamaan [2.6](#).

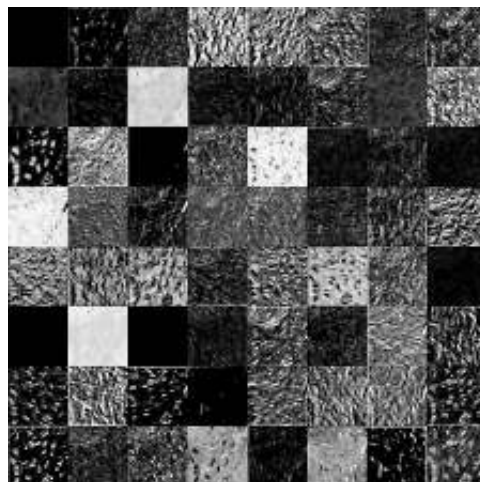
$$\begin{aligned}
 (Q_{(:,1)}) &= \begin{bmatrix} 57.7825 & 53.7228 & 68.3575 & \dots & 0 & 0 & 0 \\ 54.1393 & 48.3805 & 62.0838 & \dots & 0 & 31.5559 & 66.4840 \\ 47.6417 & 48.6969 & 48.6969 & \dots & 0.3260 & 47.3643 & 66.4840 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 14.8293 & 0 & 0 & \dots & 0 & 38.5232 & 39.5537 \\ 16.3461 & 0 & 0 & \dots & 6.1899 & 39.8632 & 39.8632 \\ 6.6718 & 21.2253 & 14.7424 & \dots & 6.1899 & 24.4892 & 24.4892 \end{bmatrix} \\
 (Q_{(:,2)}) &= \begin{bmatrix} 50.9192 & 54.9788 & 47.1248 & \dots & 67.2044 & 67.7770 & 59.3410 \\ 0 & 0 & 0 & \dots & 38.4041 & 34.7406 & 52.0562 \\ 0 & 0 & 5.4136 & \dots & 36.5751 & 14.6484 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 15.9885 & 41.9019 & 41.9019 & \dots & 7.9520 & 27.9025 & 27.9025 \\ 23.9600 & 32.9916 & 32.9916 & \dots & 9.3970 & 15.2525 & 0 \\ 39.2645 & 39.2645 & 32.9916 & \dots & 16.3244 & 15.2525 & 13.5097 \end{bmatrix} \\
 (Q_{(:,64)}) &= \begin{bmatrix} 0 & 0 & 0 & \dots & 38.5842 & 65.8662 & 65.8662 \\ 0 & 0 & 0 & \dots & 23.1992 & 9.3508 & 8.0863 \\ 0 & 0 & 4.9813 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 31.2152 & 31.2152 & 20.0410 & \dots & 20.174351 & 27.9025 & 27.9025 \\ 17.1042 & 43.2081 & 0 & \dots & 9.5824 & 9.5824 & 1.2385 \\ 30.5961 & 43.2081 & 0 & \dots & 9.5824 & 19.9038 & 19.9038 \end{bmatrix}
 \end{aligned}$$

$$\begin{aligned}
 x_{1,2}^1 &= \frac{a_{1,2}^1}{\left(1 + \frac{0.0001 \sum_{q=\max(1,1-\frac{5}{2})}^{\min(3-1,1+\frac{5}{2})} (a_{1,2}^1)^2}{5}\right)^{(0.75)}} \\
 &= \frac{57.7825}{\left(1 + \frac{0.0001((57.7825)^2 + (50.9192)^2 + (60.1883)^2)}{5}\right)^{(0.75)}} \\
 &= 50.6804
 \end{aligned}$$

Kemudian, proses tersebut dilakukan berulang sejumlah 64 *feature map*.

$$(Q_{(:, :, 64)}) = \begin{bmatrix} 50.6804 & 49.0610 & 58.5505 & \dots & 0 & 0 & 0 \\ 51.8749 & 46.2533 & 57.7819 & \dots & 0 & 30.2807 & 59.9481 \\ 46.0815 & 43.9664 & 43.9495 & \dots & 0.3085 & 44.9304 & 61.3910 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 13.9022 & 0 & 0 & \dots & 0 & 35.8918 & 36.6568 \\ 15.3033 & 0 & 0 & \dots & 6.1421 & 38.6814 & 38.8678 \\ 6.1830 & 0 & 20.2202 & \dots & 6.1183 & 23.8800 & 23.8974 \end{bmatrix}$$

Hasil dari proses ini yaitu sejumlah 64 *feature map* yang kemudian direpresentasikan kedalam suatu citra seperti yang ditunjukkan oleh Gambar 4.9.



Gambar 4.9 Visualisasi *Feature Map* Pada Cross Channel Normalization Layer

#### 4.2.5. Inception Modules

Pada CNN model GoogLeNet, salah satu kelebihan yang dimiliki yaitu adanya *inception modules*. *Inception modules* dapat mempercepat waktu komputasi karena cara kerjanya yang bersifat paralel. Pada GoogLeNet terdapat sembilan *inception* dengan masing-masing *inception* terbagi menjadi 4 bagian. Pada masing-masing bagian terdapat *convolution layer*, dan *ReLU layer* serta *pooling layer*. Pada *inception modules* terdapat *convolution reduce layer* yang dapat mengurangi kompleksitas suatu model. Sehingga, total keseluruhan terdapat 13 layer pada *inception modules*.

#### 4.2.6. Dropout Layer

Pada *dropout layer* nilai yang telah diperoleh pada layer sebelumnya digunakan sebagai input pada layer ini yang kemudian diberikan nilai probabilitas yang nilainya random namun dalam rentang 0-1. Nilai probabilitas yang diperoleh ini nantinya akan menentukan apakah node tersebut akan digunakan atau tidak. Menentukan node tersebut akan digunakan jika nilai probabilitas pada input tersebut berada di atas nilai probabilitas dari *dropout* yang diperoleh dari inisialisasi. Namun, jika nilainya berada di bawah nilai probabilitas dari probabilitas *dropout* maka node tersebut tidak digunakan. *Dropout* layer ini bertujuan untuk mengurangi adanya *overfitting* karena node yang digunakan terlalu banyak. Berikut contoh perhitungan pada layer ini dengan nilai probabilitas *dropout* sebesar 0.5.



$$\begin{aligned}
 (Q_{(:,64)}) = \begin{bmatrix} 50.6804 \\ 51.8749 \\ 46.0815 \\ \vdots \\ 13.9022 \\ 15.3033 \\ 6.1830 \end{bmatrix} &\rightarrow (Q_{(:,64)}) = \begin{bmatrix} 0.4833 \\ 0.8501 \\ 0.4112 \\ \vdots \\ 0.2951 \\ 0.5440 \\ 0.899 \end{bmatrix} &(Q_{(:,64)}) = \begin{bmatrix} 0 \\ 51.8749 \\ 0 \\ \vdots \\ 0 \\ 15.3033 \\ 6.1830 \end{bmatrix}
 \end{aligned}$$

#### 4.2.7. Fully Connected Layer

Layer ini merupakan layer terakhir yang digunakan sebagai layer penyatuan. Pada layer ini terdapat proses klasifikasi suatu citra ke dalam suatu kelas. Input yang digunakan pada layer ini yaitu diperoleh dari layer sebelumnya yang berukuran  $1 \times 1024$ . Lalu, bobot yang digunakan berukuran  $3 \times 1024$  dan bias yang berukuran  $3 \times 1$ . Berikut contoh perhitungan pada *fully connected layer* dengan menggunakan Persamaan 2.7.

$$Q = \begin{bmatrix} 0 & 51.8749 & 0 & \dots & 0 & 15.3033 & 6.1830 \end{bmatrix}$$

$$W = \begin{bmatrix} -0.0163 & 0.0711 & -0.0173 \\ -0.0482 & 0.0025 & -0.0766 \\ \vdots & \vdots & \vdots \\ 0.0593 & -0.0308 & -0.1007 \\ -0.0305 & 0.0544 & 0.0030 \end{bmatrix}$$

$$b = \begin{bmatrix} 0.0118 & -0.0130 & 0.0011 \end{bmatrix}$$

$$y_1 = 0.0118 + ((-0.0163 \times 0) + (-0.0482 \times 51.8749) + \dots + (0.0593 \times 0) + (-0.0305 \times 6.1830))$$

$$= 5.8705$$

$$y_2 = -0.0130 + ((0.0711 \times 0) + (0.0025 \times 51.8749) + \dots + (-0.0308 \times 0) + (0.0544 \times 6.1830))$$

$$= -6.6073$$

$$y_3 = 0.0011 + ((-0.0173 \times 0) + (-0.0766 \times 51.8749) + \dots + (-0.1007 \times 0) + (0.0030 \times 6.1830))$$

$$= 2.5289$$

$$y = \begin{bmatrix} 5.8705 & -6.6073 & 2.5289 \end{bmatrix}$$

#### 4.2.8. Softmax Layer

Softmx merupakan salah satu jenis fungsi aktivasi yang digunakan untuk menentukan keputusan akan suatu aplikasi. Penentuan keputusan tersebut didasari dengan perhitungan nilai probabilitas suatu citra pada setiap kelasnya. Penentuan keputusan suatu aplikasi berdasarkan nilai probabilitas tertinggi yang diperoleh dari perhitungan softmax dengan menggunakan Persamaan [2.8](#).

$$p_1(x) = \frac{\exp^{5.8705}}{\exp^{5.8705} + \exp^{-6.6073} + \exp^{2.5289}}$$

$$= 0.9658$$

$$p_2(x) = \frac{\exp^{-6.6073}}{\exp^{5.8705} + \exp^{-6.6073} + \exp^{2.5289}}$$

$$= 3.6800e - 06$$

$$p_3(x) = \frac{\exp^{2.5289}}{\exp^{5.8705} + \exp^{-6.6073} + \exp^{2.5289}}$$

$$= 0.0342$$

Berdasarkan hasil perhitungan diatas dapat dinyatakan bahwa nilai  $p_1$  menyatakan probabilitas kelas *lung adenocarcinoma*, nilai  $p_2$  menyatakan probabilitas kelas *benign lung tissue*, serta nilai  $p_3$  menyatakan probabilitas kelas *lung squamous cell carcinoma*. Pada hasil perhitungan menunjukkan bahwa data citra histopatologi yang telah diinputkan terklasifikasi ke dalam kelas *lung adenocarcinoma*.

#### 4.3. Uji Coba Model dan Evaluasi aplikasi

Pada penelitian ini dilakukan uji coba pembagian data, *dropout*, serta *batchsize* untuk memperoleh model klasifikasi yang terbaik yang nantinya akan digunakan dalam pembentukan aplikasi klasifikasi. Pada penelitian digunakan sebanyak empat kali uji coba pembagian data yaitu pembagian data 60:40, 70:30, 80:20, serta 90:10. Kemudian, juga terdapat uji coba *dropout* dimana nilai probabilitas *dropout* yang digunakan yaitu 0.6, 0.7, 0.8, dan 0.9. Selain itu, jumlah *batchsize* yang digunakan dalam uji coba yaitu sebesar 8, 16, 32, dan 64. Model klasifikasi terbaik yang diperoleh kemudian dilakukan evaluasi dengan menggunakan *confusion matrix* yang nantinya akan memperoleh nilai akurasi, sensitivitas, serta spesifisitas terbaik. Berikut hasil uji coba yang telah dilakukan.

### 4.3.1. Pembagian Data 60:40

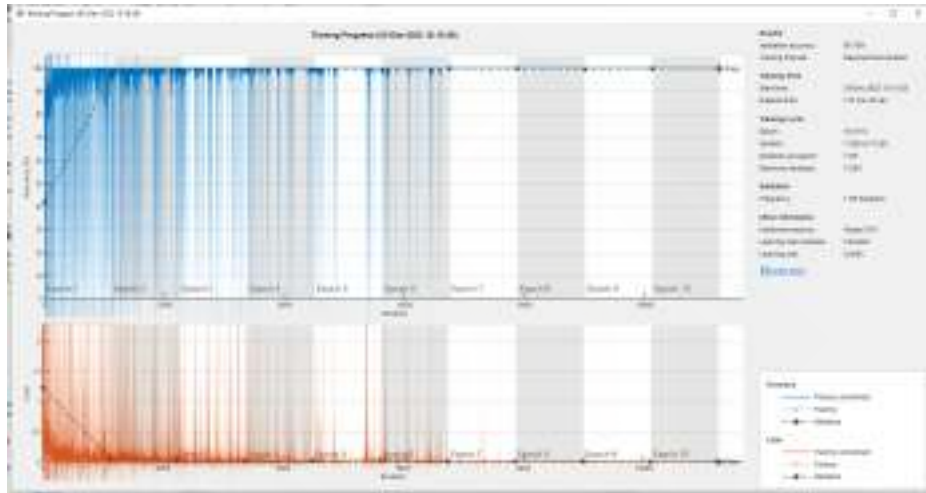
Berikut hasil uji coba pada pembagian data 60:40 ditunjukkan oleh Tabel 4.1.

**Tabel 4.1 Hasil Uji Coba Pembagian Data 60:40**

Dropout	Batchsize	Akurasi(%)	Sensitivitas(%)	Spesifisitas (%)	Waktu (min;sec)
0.6	8	99.30	99.31	99.65	177 m 34 s
	16	99.58	99.59	99.79	159 m 13 s
	32	99.55	99.55	99.78	142 m 52 s
	64	99.15	99.13	99.57	129 m 32 s
0.7	8	99.75	99.75	99.88	186 m 33 s
	16	99.68	99.68	99.84	174 m 13 s
	32	99.57	99.57	99.78	131 m 21 s
	64	99.33	99.34	99.67	121 m 45 s
0.8	8	99.78	99.78	99.89	172 m 20 s
	16	99.43	99.44	99.72	154 m 16 s
	32	99.57	99.57	99.78	152 m 57 s
	64	99.30	99.30	99.65	125 m 49 s
0.9	8	99.72	99.72	99.86	163 m 51 s
	16	99.77	99.77	99.88	157 m 44 s
	32	99.27	99.27	99.63	153 m 51 s
	64	99.10	99.11	99.55	145 m 32 s

Berdasarkan Tabel 4.1 dapat diketahui bahwa pada uji coba pembagian data 60:40 model optimal diperoleh pada nilai probabilitas *dropout* 0.8 dan *batchsize* 8 dengan nilai akurasi, sensitivitas, dan spesifisitas secara berturut-turut yaitu sebesar 99.78%, 99.78%, dan 99.89% serta membutuhkan waktu training selama 172 menit

20 detik. Pada Tabel 4.1 dapat diketahui bahwa nilai sensitivitas yang tinggi menunjukkan bahwa model yang dihasilkan dapat mengenali data citra histopatologi dengan baik. Berikut ditunjukkan plot *training progress* uji coba pembagian data 60:40, *dropout* 0.8, dan *batchsize* 8 oleh Gambar 4.10.



Gambar 4.10 *Training Progress* Pada Pembagian 60:40

Pada Gambar 4.10 merupakan *training progress* yang menghasilkan suatu model optimal pada pembagian data 60:40 yang kemudian dievaluasi dengan *confusion matrix*. Hasil evaluasi ditunjukkan oleh Gambar 4.11.

**Confusion Matrix**

Output Class	Adenocarcinoma	1994 33.2%	0 0.0%	6 0.1%	99.7% 0.3%
	Benign	0 0.0%	2000 33.3%	0 0.0%	100% 0.0%
	Squamous	7 0.1%	0 0.0%	1994 33.2%	99.7% 0.3%
		99.7% 0.3%	100% 0.0%	99.7% 0.3%	99.6% 0.2%
	Target Class				
	Adenocarcinoma	Benign	Squamous		

**Gambar 4.11 Confusion Matrix Pada Pembagian 60:40**

Pada Gambar 4.11 diketahui bahwa dari 2000 data yang dilakukan uji coba terdapat 6 data pada kelas *lung squamous cell carcinoma* yang terklasifikasi pada kelas *lung adenocarcinoma*. Kemudian, terdapat pula 7 data pada kelas *lung adenocarcinoma* yang terklasifikasi pada kelas *lung squamous cell carcinoma*.

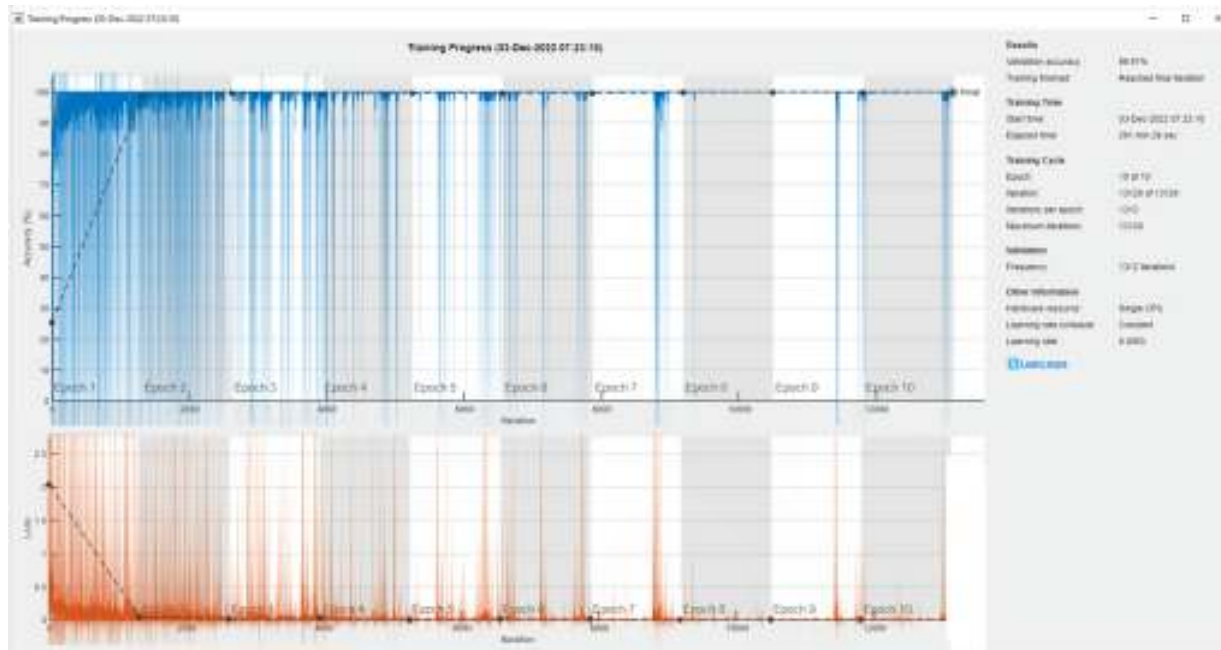
#### 4.3.2. Pembagian Data 70:30

Berikut hasil uji coba pada pembagian data 70:30 ditunjukkan oleh Tabel 4.2.

Tabel 4.2 Hasil Uji Coba Pembagian Data 70:30

Dropout	Batchsize	Akurasi(%)	Sensitivitas(%)	Spesifisitas (%)	Waktu (min;sec)
0.6	8	99.91	99.91	99.96	201 m 29 s
	16	99.82	99.82	99.91	172 m 1 s
	32	99.71	99.71	99.86	171 m 30 s
	64	99.38	99.38	99.69	170 m 48 s
0.7	8	99.76	99.76	99.88	169 m 26 s
	16	99.47	99.47	99.73	152 m 0 s
	32	99.42	99.43	99.71	148 m 40 s
	64	99.22	99.22	99.61	148 m 19 s
0.8	8	99.82	99.82	99.91	181 m 34 s
	16	99.78	99.78	99.89	165 m 27 s
	32	99.73	99.73	99.87	151 m 31 s
	64	98.80	98.83	99.41	149 m 59 s
0.9	8	99.76	99.76	99.88	169 m 28 s
	16	99.73	99.73	99.87	164 m 29 s
	32	99.47	99.48	99.74	159 m 29 s
	64	99.38	99.38	99.69	141 m 49 s

Berdasarkan Tabel 4.2 dapat diketahui bahwa pada uji coba pembagian data 70:30 model optimal diperoleh pada nilai probabilitas *dropout* 0.6 dan *batchsize* 8 dengan nilai akurasi, sensitivitas, dan spesifisitas secara berturut-turut yaitu sebesar 99.91%, 99.91%, dan 99.96% serta membutuhkan waktu training selama 201 menit 29 detik. Berikut ditunjukkan plot *training progress* uji coba pembagian data 70:30, *dropout* 0.6, dan *batchsize* 8 oleh Gambar 4.12



Gambar 4.12 *Training Progress* Pada Pembagian 70:30

Pada Gambar 4.12 merupakan *training progress* yang menghasilkan suatu model optimal pada pembagian data 70:30 yang kemudian dievaluasi dengan *confusion matrix*. Hasil evaluasi ditunjukkan oleh Gambar 4.13.

	Adenocarcinoma	Benign	Squamous	
Adenocarcinoma	1499 33.3%	0 0.0%	3 0.1%	99.8% 0.2%
Benign	0 0.0%	1500 33.3%	0 0.0%	100% 0.0%
Squamous	1 0.0%	0 0.0%	1497 33.3%	99.9% 0.1%
	99.9% 0.1%	100% 0.0%	99.8% 0.2%	99.9% 0.1%
	Adenocarcinoma	Benign	Squamous	
Output Class	Adenocarcinoma	Benign	Squamous	Target Class

Gambar 4.13 *Confusion Matrix* Pada Pembagian 70:30



Pada Gambar 4.13 diketahui bahwa dari 1500 data yang dilakukan uji coba terdapat 3 data pada kelas *lung squamous cell carcinoma* yang terklasifikasi pada kelas *lung adenocarcinoma*. Kemudian, terdapat pula 1 data pada kelas *lung adenocarcinoma* yang terklasifikasi pada kelas *lung squamous cell carcinoma*. Berdasarkan Gambar 4.13 dapat diketahui bahwa walaupun membutuhkan waktu *training* yang cukup lama tetapi nilai akurasi serta sensitivitas yang dihasilkan cukup tinggi. Nilai akurasi dan sensitivitas yang cukup tinggi membuktikan bahwa model dapat mengenali data dengan baik.

#### 4.3.3. Pembagian Data 80:20

Berikut hasil uji coba pada pembagian data 80:20 ditunjukkan oleh Tabel 4.3.



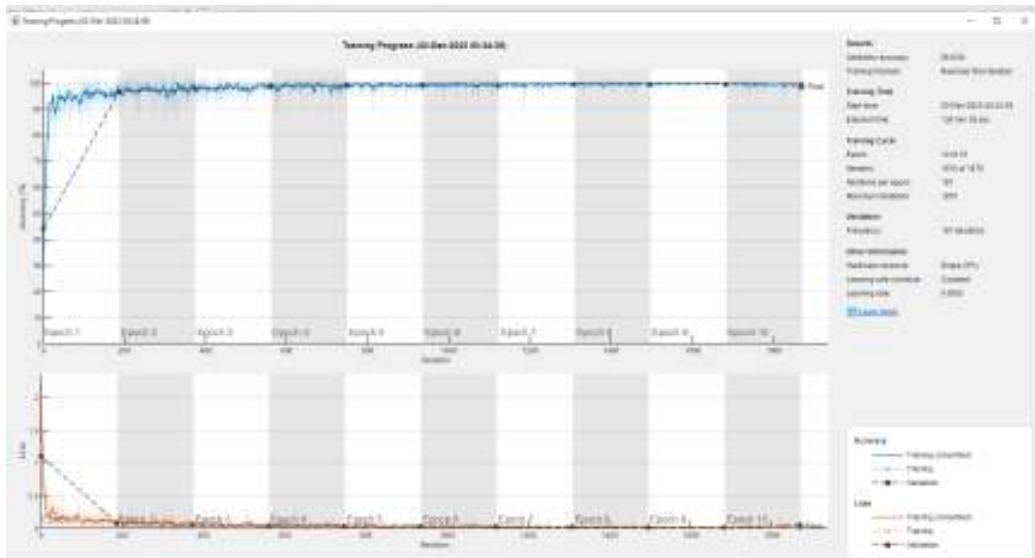
UIN SUNAN AMPEL  
S U R A B A Y A

Tabel 4.3 Hasil Uji Coba Pembagian Data 80:20

Dropout	Batchsize	Akurasi(%)	Sensitivitas(%)	Spesifisitas (%)	Waktu (min;sec)
0.6	8	99.77	99.77	99.88	181 m 3 s
	16	99.33	99.34	99.67	176 m 46 s
	32	99.60	99.60	99.80	162 m 38 s
	64	98.63	98.67	99.33	162 m 6 s
0.7	8	99.80	99.80	99.90	191 m 47 s
	16	99.70	99.70	99.85	158 m 48 s
	32	99.00	99.02	99.51	150 m 44 s
	64	99.87	98.89	99.44	148 m 28 s
0.8	8	99.77	99.77	99.88	189 m 4 s
	16	99.33	99.33	99.67	171 m 8 s
	32	99.60	99.60	99.80	163 m 27 s
	64	98.63	98.63	99.32	139 m 38 s
0.9	8	99.87	99.87	99.93	198 m 27 s
	16	99.80	99.80	99.90	175 m 27 s
	32	99.40	99.40	99.70	162 m 22 s
	64	98.43	98.43	99.72	160 m 56 s

Berdasarkan Tabel 4.3 dapat diketahui bahwa pada uji coba pembagian data 80:20 model optimal diperoleh pada nilai probabilitas *dropout* 0.7 dan *batchsize* 64 dengan nilai akurasi, sensitivitas, dan spesifisitas secara berturut-turut yaitu sebesar 99.87%, 99.87%, dan 99.43% serta membutuhkan waktu training selama 148 menit 28 detik. Pada Tabel 4.3 dapat diketahui bahwa nilai sensitivitas yang tinggi menunjukkan bahwa model yang dihasilkan dapat mengenali data citra histopatologi dengan baik. Berikut ditunjukkan plot *training progress* uji coba pembagian data <http://digilib.uinsa.ac.id/> <http://digilib.uinsa.ac.id/> <http://digilib.uinsa.ac.id/>

80:20, *dropout* 0.7, dan *batchsize* 64 oleh Gambar 4.14.



Gambar 4.14 *Training Progress* Pada Pembagian 80:20

Pada Gambar 4.14 merupakan *training progress* yang menghasilkan suatu model optimal pada pembagian data 80:20 yang kemudian dievaluasi dengan *confusion matrix*. Hasil evaluasi ditunjukkan oleh Gambar 4.15.

		Target Class				
		Adenocarcinoma	Benign	Squamous		
Output Class	Adenocarcinoma	967 32.1%	0 0.0%	3 0.1%	96.7%	0.2%
	Benign	0 0.0%	1000 33.3%	0 0.0%	100%	0.0%
	Squamous	38 1.3%	0 0.0%	997 33.2%	96.3%	3.7%
		96.2%	100%	99.7%	96.6%	3.8%
		0.0%	0.0%	0.3%	1.4%	

Gambar 4.15 *Confusion Matrix* Pada Pembagian 80:20

Pada Gambar 4.15 diketahui bahwa dari 1000 data yang dilakukan uji coba terdapat 3 data pada kelas *lung squamous cell carcinoma* yang terklasifikasi pada kelas *lung adenocarcinoma*. Kemudian, terdapat pula 38 data pada kelas *lung adenocarcinoma* yang terklasifikasi pada kelas *lung squamous cell carcinoma*.

#### 4.3.4. Pembagian Data 90:10

Berikut hasil uji coba pada pembagian data 90:10 ditunjukkan oleh Tabel 4.4.

**Tabel 4.4 Hasil Uji Coba Pembagian Data 90:10**

Dropout	Batchsize	Akurasi(%)	Sensitivitas(%)	Spesifisitas (%)	Waktu (min;sec)
0.6	8	99.80	99.80	99.90	219 m 49 s
	16	99.87	99.87	99.93	187 m 59 s
	32	99.53	99.53	99.77	188 m 45 s
	64	99.60	99.60	99.80	173 m 10 s
0.7	8	99.93	99.93	99.97	190 m 20 s
	16	99.73	99.73	99.87	174 m 11 s
	32	99.73	99.73	99.87	168 m 46 s
	64	99.33	99.33	99.67	153 m 26 s
0.8	8	99.87	99.87	99.93	208 m 9 s
	16	99.80	99.80	99.90	184 m 55 s
	32	99.73	99.73	99.87	172 m 8 s
	64	99.47	99.47	99.73	163 m 53 s
0.9	8	99.87	99.87	99.93	246 m 21 s
	16	99.87	99.87	99.93	196 m 31 s
	32	99.53	99.53	99.77	194 m 54 s
	64	99.80	99.80	88.90	176 m 22 s



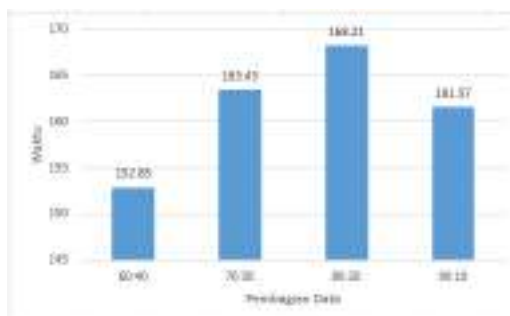
**Confusion Matrix**

		Target Class			
		Adenocarcinoma	Benign	Squamous	
Output Class	Adenocarcinoma	499 33.3%	0 0.0%	0 0.0%	100% 0.0%
	Benign	0 0.0%	500 33.3%	0 0.0%	100% 0.0%
	Squamous	1 0.1%	0 0.0%	500 33.3%	99.9% 0.2%
		99.9% 0.2%	100% 0.0%	100% 0.0%	99.9% 0.1%

**Gambar 4.17 Confusion Matrix Pada Pembagian 90:10**

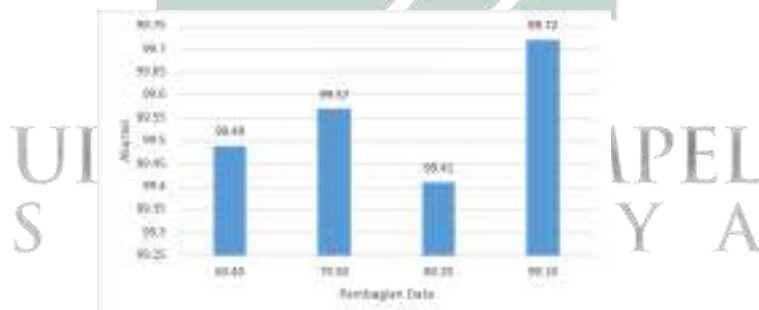
Pada Gambar 4.17 diketahui bahwa dari 500 data yang dilakukan uji coba hanya terdapat 1 data pada kelas *lung adenocarcinoma* yang terklasifikasi pada kelas *lung squamous cell carcinoma*. Berdasarkan Tabel 4.4 dapat diketahui bahwa walaupun membutuhkan waktu *training* yang cukup lama tetapi nilai akurasi serta sensitivitas yang dihasilkan cukup tinggi. Nilai akurasi dan sensitivitas yang cukup tinggi membuktikan bahwa model dapat mengenali data dengan baik.

Pada penelitian ini dapat diketahui bahwa pembagian data, probabilitas *dropout*, serta jumlah *batchsize* sangat mempengaruhi nilai akurasi serta waktu *training*. Pembagian data sangat berpengaruh pada waktu *training* serta akurasi suatu model. Hal ini ditunjukkan oleh Gambar 4.18.



**Gambar 4.18 Grafik Pengaruh Pembagian Data Terhadap Waktu *Training***

Berdasarkan Gambar 4.18 dapat diketahui bahwa pada pembagian data 60:40 rata-rata waktu untuk melakukan proses *training* cukup cepat yaitu selama 152 menit 85 detik jika dibandingkan dengan pembagian data 80:20 yang memiliki rata-rata waktu sebesar 168 menit 21 detik. Kemudian, pembagian data juga mempengaruhi nilai akurasi suatu model. Hal ini ditunjukkan oleh Gambar 4.19.

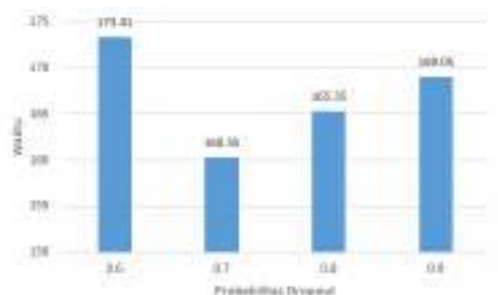


**Gambar 4.19 Grafik Pengaruh Pembagian Data Terhadap Akurasi**

Berdasarkan Gambar 4.19 dapat diketahui bahwa pada pembagian data 90:10 rata-rata akurasi yang diperoleh sangat baik yaitu 99.72% sedangkan pada pembagian data 80:20 rata-rata akurasi yang diperoleh yaitu 99.41%.

Penelitian ini juga melakukan uji coba pada nilai probabilitas *dropout* dimana nilai probabilitas *dropout* yang digunakan yaitu 0.6, 0.7, 0.8, dan 0.9. Nilai probabilitas *dropout* juga memiliki pengaruh pada rata-rata waktu saat melakukan

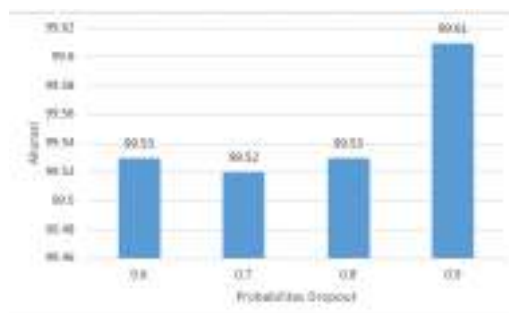
proses *training* serta berpengaruh pada nilai akurasi suatu model. Pada Gambar menunjukkan grafik pengaruh nilai probabilitas *dropout* terhadap rata-rata waktu saat proses *training*.



**Gambar 4.20 Grafik Pengaruh *Dropout* Terhadap Waktu *Training***

Berdasarkan Gambar 4.20 diketahui bahwa pada probabilitas *dropout* 0.6 rata-rata waktu yang dibutuhkan untuk melakukan proses *training* yaitu sebesar 173 menit 41 detik. Pada probabilitas *dropout* 0.7 rata-rata waktu yang dibutuhkan jauh lebih sedikit dibandingkan dengan probabilitas *dropout* 0.6 yaitu sebesar 160 menit 35 detik. Sedangkan, pada probabilitas *dropout* 0.8 dibutuhkan waktu sebesar 165 menit 35 detik, dan pada probabilitas *dropout* 0.9 rata-rata waktu yang dibutuhkan untuk melakukan proses *training* yaitu sebesar 169 menit 05 detik. Oleh karena itu, dapat disimpulkan bahwasanya berdasarkan nilai probabilitas *dropout* yang digunakan yaitu pada probabilitas 0.6 yakni nilai probabilitas yang kecil membutuhkan rata-rata waktu yang sangat lama untuk melakukan proses *training*. Nilai probabilitas *dropout* juga berpengaruh pada nilai akurasi yang diperoleh. Hal ini ditunjukkan oleh Gambar 4.21.

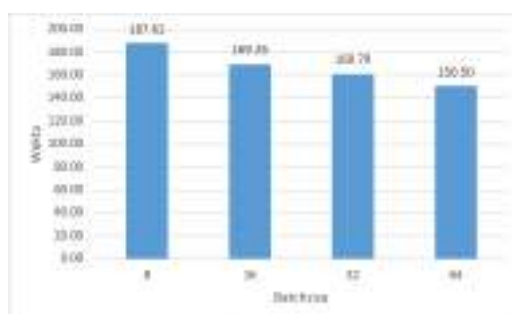




**Gambar 4.21 Grafik Pengaruh *Dropout* Terhadap Akurasi**

Pada Gambar 4.21 diketahui bahwa pada nilai probabilitas *dropout* tertinggi yaitu 0.9 memperoleh rata-rata nilai akurasi yang tinggi jika dibandingkan dengan nilai probabilitas yang lain yakni sebesar 99.61%. Sedangkan, pada nilai probabilitas *dropout* 0.6 dan 0.8 memperoleh rata-rata nilai akurasi yang sama yaitu sebesar 99.53% dan pada nilai probabilitas *dropout* 0.7 memperoleh rata-rata nilai akurasi yang lebih kecil dibandingkan dengan yang lain yakni sebesar 99.52%.

Selain itu, pada penelitian ini juga melakukan uji coba pada jumlah *batchsize*. Jumlah *batchsize* juga berpengaruh pada waktu saat melakukan proses *training* serta nilai akurasi yang diperoleh. Pengaruh jumlah *batchsize* terhadap waktu *training* ditunjukkan oleh Gambar 4.22.



**Gambar 4.22 Grafik Pengaruh Jumlah *Batchsize* Terhadap Waktu *Training***

Pada Gambar 4.22 diketahui bahwa pada *batchsize* sejumlah 8 waktu untuk

melakukan proses *training* membutuhkan rata-rata waktu sebesar 187 menit 62 detik. Sedangkan, pada *batchsize* sejumlah 64 membutuhkan waktu *training* sebesar 150 menit 50 detik. Hal ini, dapat disimpulkan bahwasanya semakin besar jumlah *batchsize* maka semakin cepat pula proses *training* yang dilakukan. Jumlah *batchsize* tidak hanya mempengaruhi waktu *training* tetapi juga dapat mempengaruhi nilai akurasi yang diperoleh. Pengaruh jumlah *batchsize* terhadap nilai akurasi ditunjukkan oleh Gambar 4.23.



Gambar 4.23 Grafik Pengaruh Jumlah *Batchsize* Terhadap Akurasi

Berdasarkan Gambar 4.23 diketahui bahwa pada jumlah *batchsize* sebanyak 8 rata-rata nilai akurasi yang diperoleh yaitu sangat tinggi sebesar 99.78%. Pada jumlah *batchsize* 16 nilai akurasi yang diperoleh sedikit lebih rendah dibandingkan jumlah *batchsize* 8 yaitu sebesar 99.67%. Hal ini, juga terjadi pada jumlah *batchsize* 32 dimana rata-rata nilai akurasi sedikit lebih rendah dibandingkan jumlah *batchsize* 16 yakni memperoleh rata-rata akurasi sebesar 99.53%. Sedangkan, pada jumlah *batchsize* 64 rata-rata nilai akurasi yang diperoleh termasuk cukup rendah jika dibandingkan dengan jumlah *batchsize* 8 yaitu sebesar 99.21%. Oleh karena itu, berdasarkan Gambar 4.23 dapat disimpulkan bahwasanya semakin besar jumlah *batchsize* maka rata-rata nilai akurasi yang diperoleh semakin menurun. Hal ini, berbanding terbalik dengan waktu *training*.

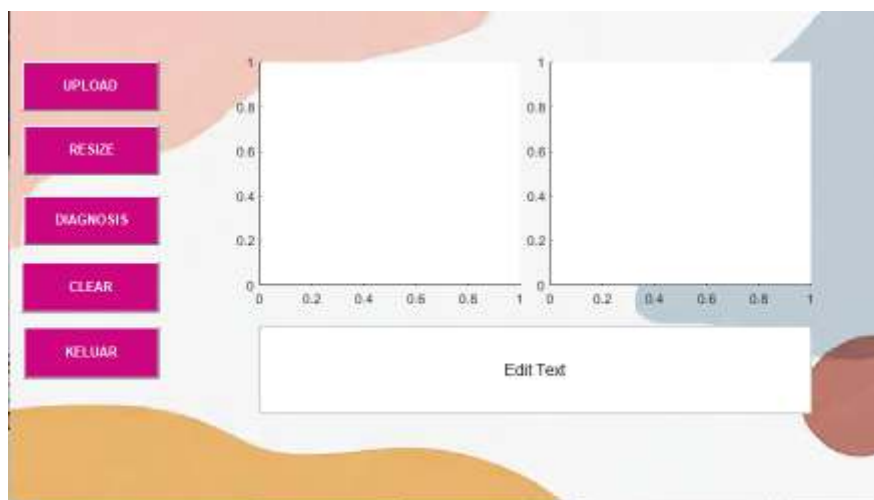
#### 4.4. Aplikasi Klasifikasi Kanker Paru Menggunakan CNN Model GoogLeNet

Pada penelitian ini, dilakukan beberapa uji coba sehingga memperoleh hasil yang optimal. Hasil yang diperoleh ini kemudian dibuat dalam bentuk aplikasi atau aplikasi. Pembentukan aplikasi ini digunakan untuk membantu tenaga medis dalam mengklasifikasikan citra histopatologi kanker paru kedalam kelasnya yakni terdapat tiga kelas yaitu *lung adenocarcinoma*, *benign lung tissue*, serta *squamous lung cell carcinoma*. Tampilan awal aplikasi yang dibentuk ditunjukkan oleh Gambar 4.24.



Gambar 4.24 Halaman Awal Aplikasi

Pada Gambar 4.24 menunjukkan halaman dari aplikasi yang telah dibentuk. Pada halaman awal aplikasi terdapat logo UIN Sunan Ampel Surabaya, judul aplikasi, tombol mulai dan tutup, serta nama pembuat aplikasi. Kemudian, lanjut pada halaman utama ditunjukkan oleh Gambar 4.25.



**Gambar 4.25 Halaman Utama Aplikasi**

Pada Gambar 4.25 terdapat 5 tombol yaitu yang pertama terdapat tombol *upload* yang digunakan untuk mencari citra histopatologi kanker paru yang akan dilakukan klasifikasi. Kemudian terdapat tombol *resize* digunakan untuk melakukan *resizing* ukuran citra histopatologi menyesuaikan dengan ketentuan ukuran dari *googlenet* yaitu  $224 \times 224$ . Jika diklik tombol *resize* maka akan memunculkan hasil *resize* dari citra yang diupload sebelumnya. Tampilan setelah diklik tombol *upload* ditunjukkan oleh Gambar 4.26.



**Gambar 4.26 Contoh Citra Histopatologi Setelah di *Resize***

Lalu, ada tombol diagnosis dimana jikat tombol ini diklik maka akan muncul hasil klasifikasi dari citra histopatologi yang telah *upload*. Tampilan jika diklik tombol diagnosis ditunjukkan oleh Gambar [4.27](#)



**Gambar 4.27 Contoh Hasil Diagnosis**

Kemudian, terdapat tombol *clear* yang digunakan untuk menghapus citra yang telah *upload* sebelumnya jika ingin mengganti dengan citra histopatologi yang lain serta tombol keluar jika ingin kembali ke halaman awal.

#### 4.5. Integrasi Keislaman

Penelitian ini memiliki tujuan yaitu melakukan klasifikasi suatu penyakit yakni kanker paru. Kanker paru merupakan salah satu jenis kanker yang memiliki angka kematian tertinggi. Oleh karena itu, tingginya angka kematian yang disebabkan oleh kanker paru dengan penyebab terbanyak merokok. Merokok merupakan salah satu kegiatan yang tidak bermanfaat namun sering dilakukan tanpa sadar bahwa akibat kebiasaan merokok sangatlah berbahaya. Oleh karena itu, dalam hadits riwayat Tirmidzi no. 2317 dan Ibnu Majah no. 3976 yang berbunyi:

مِنْ حُسْنِ إِسْلَامِ الْمَرْءِ تَرْكُهُ مَا لَا يَغْنِيهِ

artinya: "Di antara kebaikan islam seseorang adalah meninggalkan hal yang tidak bermanfaat." (HR Tirmidzi no. 2317 dan Ibnu Majah no. 3976)

Sebagaimana dijelaskan pada hadits diatas bahwa meninggalkan hal yang tidak bermanfaat merupakan salah satu kebaikan islam seseorang. Jadi, meninggalkan kebiasaan merokok yang mana kebiasaan tersebut merupakan hal yang tidak bermanfaat adalah salah satu cara yang secara tidak langsung juga dapat menjaga kesehatan. Sesungguhnya banyak manusia telah melupakan akan pentingnya menjaga kesehatan sebagaimana dijelaskan dalam hadits riwayat Al-Bukhari Nomor 6412, at-Tirmidzi nomor 2304, Ibnu Majah nomor 4170, yang berbunyi:

تَمَتُّانِ مَغْبُورٌ فِيهَا كَثِيرٌ مِنَ النَّاسِ الصِّحَّةُ وَالْفَرَاغُ

artinya: "Dua kenikmatan yang sering dilupakan oleh kebanyakan manusia adalah kesehatan dan waktu luang." (Al-Bukhari Nomor 6412, at-Tirmidzi nomor 2304, Ibnu Majah nomor 4170)

Pada hadits dijelaskan bahwasanya kesehatan merupakan kenikmatan yang sering dilupakan. Sedangkan, pada pemeriksaan mengenai kanker paru membutuhkan waktu yang cukup lama yang dapat menyebabkan adanya keterlambatan penanganan karena dalam proses mendiagnosis atau mengklasifikasikan membutuhkan waktu yang cukup lama. Oleh karena itu, penelitian ini menghasilkan suatu aplikasi yang dapat membantu tenaga medis dalam mengklasifikasikan kanker paru dengan waktu yang lebih cepat. Hal ini selaras dengan yang telah dijelaskan dalam Qur'an Surah Al-Maidah ayat 2 yang berbunyi:

يَا أَيُّهَا الَّذِينَ آمَنُوا لَا تَحِلُّوا شَعَائِرَ اللَّهِ وَلَا الشَّهْرَ الْحَرَامَ وَلَا الْهَدْيَ وَلَا  
 الْقَلَائِدَ وَلَا آمِينَ الْبَيْتِ الْحَرَامِ يَبْتَغُونَ فَضْلًا مِنْ رَبِّهِمْ وَرِضْوَانًا وَإِذَا حَلَلْتُمْ  
 فَاصْطَادُوا وَلَا يَجْرِمَنَّكُمْ شَنَا نُ قَوْمٍ أَنْ صَدُّوكُمْ عَنِ الْمَسْجِدِ الْحَرَامِ أَنْ  
 تَعْتَدُوا وَتَعَاوَنُوا عَلَى الْبِرِّ وَالتَّقْوَىٰ وَلَا تَعَاوَنُوا عَلَى الْإِثْمِ وَالْعُدْوَانِ وَاتَّقُوا  
 اللَّهَ إِنَّ اللَّهَ شَدِيدُ الْعِقَابِ

artinya: “Hai orang-orang yang beriman, janganlah kamu melanggar syi’ar-syi’ar Allah, dan jangan melanggar kehormatan bulan-bulan haram, jangan (mengganggu) binatang-binatang had-ya, dan binatang-binatang qalaa-id, dan jangan (pula) mengganggu orang-orang yang mengunjungi Baitullah sedang mereka mencari kurnia dan keridhaan dari Tuhannya dan apabila kamu telah menyelesaikan ibadah haji, maka bolehlah berburu. Dan janganlah sekali-kali kebencian(mu) kepada sesuatu kaum karena mereka menghalang-halangi kamu dari Masjidilharam, mendorongmu berbuat aniaya (kepada mereka). Dan tolong-menolonglah kamu dalam (mengerjakan) kebajikan dan takwa, dan jangan tolong-menolong dalam berbuat dosa dan pelanggaran. Dan bertakwalah kamu kepada Allah, sesungguhnya Allah amat berat siksa-Nya.”(QS Al-Maidah:2)

Pada ayat diatas dijelaskan bahwasanya harus senantiasa tolong menolong dalam kebaikan, dimana dalam penelitian adanya aplikasi yang telah diperoleh telah mencerminkan adanya tolong menolong antar sesama. Aplikasi yang dihasilkan merupakan bentuk tolong menolong non materil yakni berupa ilmu pengetahuan yang dimiliki peneliti dan kemudian di implementasikan menjadi sebuah aplikasi.

## BAB V

### PENUTUP

#### 5.1. Kesimpulan

Berdasarkan hasil yang diperoleh dari penelitian ini yakni penelitian terkait klasifikasi kanker paru berdasarkan citra histopatologi menggunakan metode CNN model GoogLeNet, maka dapat diambil beberapa kesimpulan, yaitu:

1. Berdasarkan beberapa uji coba yang telah dilakukan berdasarkan pembagian data, probabilitas *dropout*, *batchsize* maka diperoleh model yang optimal yaitu dengan persentase nilai akurasi, sensitivitas serta spesifitas berturut-turut sebesar 99.93%, 99.93%, dan 99.97% dan membutuhkan waktu *training* selama 190 menit 20 detik pada pembagian data 90:10, probabilitas *dropout* 0.7 dan *batchsize* 64.
2. Berdasarkan beberapa uji coba yang telah dilakukan bahwasanya berdasarkan nilai probabilitas dropout yang digunakan yaitu pada probabilitas 0.6 yakni nilai probabilitas yang kecil membutuhkan rata-rata waktu yang sangat lama untuk melakukan proses *training* dan semakin besar jumlah *batchsize* maka semakin cepat pula proses *training* yang dilakukan.

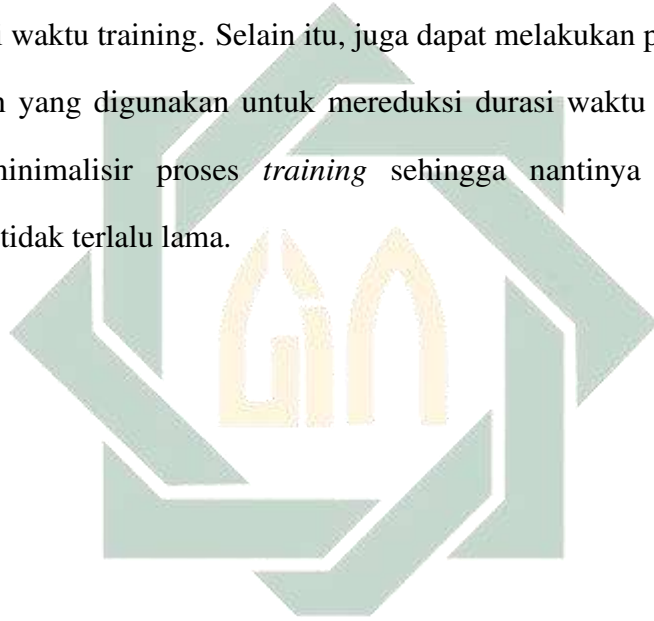
#### 5.2. Saran

Pada penelitian ini yakni terkait klasifikasi kanker paru berdasarkan data citra histopatologi menggunakan CNN model GoogLeNet perlu dilakukan perbaikan serta inovasi yang baru supaya memperoleh sistem klasifikasi yang lebih baik lagi.



Berikut beberapa saran dari penulis untuk penelitian selanjutnya antara lain:

1. Pada penelitian ini dengan menggunakan metode CNN model GoogLeNet membutuhkan waktu training yang cukup lama. Pada penelitian selanjutnya, dapat menggunakan metode lain atau model CNN yang lain untuk mengurangi waktu training. Selain itu, juga dapat melakukan penggabungan metode lain yang digunakan untuk mereduksi durasi waktu *training* agar dapat meminimalisir proses *training* sehingga nantinya waktu yang dibutuhkan tidak terlalu lama.



UIN SUNAN AMPEL  
S U R A B A Y A

## DAFTAR PUSTAKA

Akmalovna, B. S. (2022). X-RAY EXAMINATION IN THE DIAGNOSIS OF CANCER. *Galaxy International Interdisciplinary Research Journal*, 10(2):256–258.

Aktalina, L. (2022). Edukasi Merokok sebagai Faktor Risiko terjadinya Kanker Paru: merokok sebagai faktor risiko kanker paru. *Pengabdian Deli Sumatera*, 1(2).

Ali, F., Khan, S., Abbas, A. W., Shah, B., Hussain, T., Song, D., El-Sappagh, S., and Singh, J. (2022). A Two-Tier Framework Based on GoogLeNet and YOLOv3 Models for Tumor Detection in MRI. *Computers, Materials and Continua*, 72:73.

Azizah, A. N. (2021). Klasifikasi penyakit Diabetic Retinopathy menggunakan Metode Convolutional Neural Network (CNN) Model Deep Residual Network (Resnet).

Bhanumathi, V. and Sangeetha, R. (2019). CNN based training and classification of MRI brain images. In *2019 5th International Conference on Advanced Computing Communication Systems (ICACCS)*, pages 129–133. IEEE.

Binotto, M., Reinert, T., Werutsky, G., Zaffaroni, F., and Schwartzmann, G. (2020). Health-related quality of life before and during chemotherapy in patients with early-stage breast cancer. *Pengabdian Deli Sumatera*, 6(1):200247.

Borkowski, A. A., Bui, M. M., Thomas, L. B., Wilson, C. P., DeLand, L. A., and Mastorides, S. M. (2019). Lung and colon cancer histopathological image dataset (1c25000). *arXiv preprint arXiv:1912.12142*.

Chicco, D., Tötsch, N., and Jurman, G. (2021). The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData mining*, 14(1):1–22.

Endah, F. and Encis, I. S. (2021). Evaluasi Dan Prediksi Penguasaan Bahasa Inggris Maritim Menggunakan Metode Decision Tree Dan Confusion Matrix (Studi Kasus Di Universitas Maritim Amni). *Prosiding Kemaritiman 2021*.

Fattah, M. S. (2021). Deteksi penyakit pneumonia dan COVID-19 menggunakan citra x-ray dengan metode Convolutional Neural Network (CNN) model GoogleNet.

Foeady, A. Z. (2019). *Sistem klasifikasi kanker kulit berdasarkan data citra dermoscopic dengan menggunakan metode Deep Extreme Learning Machine*. PhD thesis.

Ghulam, A., Ali, F., Sikander, R., Ahmad, A., Ahmed, A., and Patil, S. (2022). ACP-2DCNN: Deep learning-based model for improving prediction of anticancer peptides using two-dimensional convolutional neural network. *Chemometrics and Intelligent Laboratory Systems*, 226:104589.

Haq, D. Z. (2021). Klasifikasi Citra Kanker Kulit Menggunakan Convolutional Neural Network Model Googlenet.

Hasby, A. N. F. A. (2022). *Sistem diagnosa citra histopatologi kanker ginjal menggunakan metode cnn model googlenet*. PhD thesis.

Hsu, C.-M., Hsu, C.-C., Hsu, Z.-M., Shih, F.-Y., Chang, M.-L., and Chen, T.-H. (2021). Colorectal polyp image detection and classification through grayscale images and deep learning. *Sensors*, 21(18):5995.

Hynds, R. E., Frese, K. K., Pearce, D. R., Grönroos, E., Dive, C., and Swanton, C. (2021). Progress towards non-small-cell lung cancer models that represent clinical evolutionary trajectories. *Open biology*, 11(1):200247.

Ichsan, M. I. N., Bakhtiar, R., Moerad, E. B., and Irawiraman, H. (2020). Hubungan Antara Kadar Carcinoembryonic Antigen (Cea) Dengan Subtipe Histologi Kanker Paru Di Rsud Abdul Wahab Sjahrane Samarinda. *Jurnal Kedokteran Mulawarman*, 7(3):23–33.

Jawahar, M., Sharen, H., and Gandomi, A. H. (2022). ALNett: A cluster layer deep convolutional neural network for acute lymphoblastic leukemia classification. *Computers in Biology and Medicine*, 148:105894.

Kartikasari, M. D. (2020). Implementasi Deep Learning Object Detection Rambu K3 pada Video Menggunakan Metode Convolutional Neural Network (CNN) dengan Tensorflow (Studi Kasus: Rambu Kesehatan dan Keselamatan Kerja (K3) Jalur Evakuasi dan Alat Pemadam Api pada Gedung FMIPA UII). Tu.

Khalid, H., Hussain, M., Ghamdi, M. A. A., Khalid, T., Khalid, K., Khan, M. A., Fatima, K., Masood, K., Almotiri, S. H., Farooq, M. S., and Ahmed, A. (2020). A comparative systematic literature review on knee bone reports from mri, x-rays and ct scans using deep learning and machine learning methodologies. *Diagnostics*, 10(8):518.

Kirman, K., Saputra, A., and Sukmana, J. (2019). Sistem Pakar Untuk Mendiagnosis Penyakit Lambung Dan Penanganannya Menggunakan Metode Dempster Shafer. *Pseudocode*, 6(1):58–66.

Matthews, H. K., Bertoli, C., and de Bruin, R. A. M. (2022). Cell cycle control in cancer. *Nature Reviews Molecular Cell Biology*, 23(1):74–88.

Mohamed Loey, Florentin Smarandache, N. E. M. K. (2020). SS symmetry Within the Lack of Chest COVID-19 X-ray Dataset : A Novel Detection Model Based on GAN and Deep.

Mutiarasari, D. (2019). Ischemic stroke: symptoms, risk factors, and prevention. *Medika Tadulako: Jurnal Ilmiah Kedokteran Fakultas Kedokteran dan Ilmu Kesehatan*, 6(1):60–73.

Nasrulloh, N. Z. Z. (2020). Ta: Perbandingan Layer Arsitektur Residual Network Untuk Deteksi Penggunaan Sabuk Pengaman Pada Pengemudi Mobil.

Normawati, D. and Prayogi, S. A. (2021). Implementasi Naïve Bayes Classifier Dan Confusion Matrix Pada Analisis Sentimen Berbasis Teks Pada Twitter. *J-SAKTI (Jurnal Sains Komputer dan Informatika)*, 5(2):697–711.

Novitasari, D. C. R., Hendradi, R., Caraka, R. E., Rachmawati, Y., Fanani, N. Z., Syarifudin, A., Toharudin, T., and Chen, R. C. (2020). Detection of COVID-19 chest X-ray using support vector machine and convolutional neural network. *Commun. Math. Biol. Neurosci.*, 2020:Article-ID.

Nugraha, F. S., Shidiq, M. J., and Rahayu, S. (2019). Analisis Algoritma Klasifikasi Neural Network Untuk Diagnosis Penyakit Kanker Payudara. *Jurnal Pilar Nusa Mandiri*, 15(2):149–156.

Phiadelvira, B. Y., Matematika, P. S., Sains, F., Teknologi, D. A. N., Islam, U., and Sunan, N. (2021). Klasifikasi kanker serviks berdasarkan citra kolposkopi menggunakan convolutional neural network (cnn) model alexnet.

Prasetio, A. (2021). *Citra Digital Dan Algoritma Penerapannya.*” PhD thesis.

- Putri, A. R. W. (2022). Classification of Breast Cancer Using the Digital Mammogram Method. *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, 9(4):2752–2761.
- Putriani, F. A., Kholis, F. N., and Purwoko, Y. (2019). Perbedaan Faktor Risiko Penderita Adenokarsinoma Paru Dengan Mutasi Egfr Dan Non Mutasi Egfr.
- Rahmawati, S. (2021). Peran Onkogen dan Tumor Suppressor Gene pada Karsinogenesis. *JK Unila*, 5(1):61–68.
- Raso, M. G., Bota-Rabassedas, N., and Wistuba, I. I. (2021). Pathology and Classification of SCLC. *Cancers*, 13(4):820.
- Rifai, Y. (2020). Analisis Faktor-Faktor Penyebab Kanker Paru-Paru Pada Pasien Di RSUD BATARA Guru Belopa.
- Riwanto, A. C. (2020). Hubungan Dukungan Keluarga Dengan Tingkat Depresi Pada Pasien Kanker Yang Mengikuti Kemoterapi di RS Pku Muhammadiyah Yogyakarta.
- Sadewa, R. P., Irawan, B., and Setianingsih, C. (2019). Fire detection using image processing techniques with convolutional neural networks. In *2019 International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, pages 290–295. IEEE.
- Sandhi, S. I. (2019). Studi Fenomenologi: Kesadaran Diri (Self Awareness) Perokok Aktif yang mempunyai Anak Balita dalam Perilaku Merokok di Tempat Umum di Kelurahan Pegulon Kabupaten Kendal. *Jurnal Kebidanan Harapan Ibu Pekalongan*, 6:237–243.

- Saputra, S., Yudhana, A., and Umar, R. (2022). Implementation of Naïve Bayes for Fish Freshness Identification Based on Image Processing. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 6(3):412–420.
- Satriawan, D. (2022). yun. *Jurnal Litbang Sukowati: Media Penelitian dan Pengembangan*, 5(2):51–58.
- Shrem, G., Azani, L., Feferkorn, I., Listovsky, T., Hussaini, S., Farber, B., Dahan, M. H., and Salmon-Divon, M. (2022). Effect of Malignancy on Semen Parameters. *Life*, 12(6):922.
- Sibero, J. T., Siregar, A. P., and Fitria, A. (2021). Peningkatan Pengetahuan Remaja untuk Deteksi Dini Kanker Payudara dengan Edukasi dan Pelaksanaan Pemeriksaan Payudara Sendiri (Sadari) di Yayasan Perguruan Budi Agung Medan Tahun 2021. *Jurnal Abdidas*, 2(3):705–712.
- Susim, T. and Darujati, C. (2021). Pengolahan Citra untuk Pengenalan Wajah (Face Recognition) Menggunakan OpenCV. *Jurnal Syntax Admiration*, 2(3):534–545.
- Tahir, M. (2020). Karakteristik Penderita Kanker Paru Primer Periode Januari 2017–Desember 2019.
- Theilmann, F. (2022). Hands-on-experiments for additive colour mixing. *Physics Education*, 57(5):55031.
- Virkhansa, C. F., Setiawan, B. D., and Dewi, C. (2019). Klasifikasi Status Gunung Berapi dengan Metode Learning Vector Quantization (LVQ). *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer e-ISSN*, 2548:964X.
- Wu, H., Huang, Q., Wang, D., and Gao, L. (2018). A CNN-SVM combined model

for pattern recognition of knee motion using mechanomyography signals. *Journal of Electromyography and Kinesiology*, 42:136–142.

Zulfa, I. I. (2022). *Klasifikasi kanker paru berdasarkan citra histopatologi menggunakan metode Convolutional Neural Network (CNN) model AlexNet*. PhD thesis.



UIN SUNAN AMPEL  
S U R A B A Y A