

**KLASIFIKASI KANKER PARU BERDASARKAN CITRA
HISTOPATOLOGI MENGGUNAKAN METODE *CONVOLUTIONAL
NEURAL NETWORK* (CNN) MODEL ALEXNET**

SKRIPSI



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

Disusun Oleh
IRKHANA INDAKA ZULFA
H72218023

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

2022

PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama : IRKHANA INDAKA ZULFA

NIM : H72218023

Program Studi : Matematika

Angkatan : 2018

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "KLASIFIKASI KANKER PARU BERDASARKAN CITRA HISTOPATOLOGI MENGGUNAKAN METODE *CONVOLUTIONAL NEURAL NETWORK* (CNN) MODEL ALEXNET". 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, 17 Januari 2022

Yang menyatakan,



IRKHANA INDAKA ZULFA
NIM. H72218023

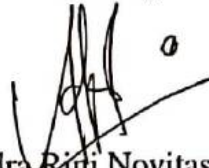
LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

Nama : IRKHANA INDAKA ZULFA
NIM : H72218023
Judul Skripsi : KLASIFIKASI KANKER PARU BERDASARKAN
CITRA HISTOPATOLOGI MENGGUNAKAN METODE
CONVOLUTIONAL NEURAL NETWORK (CNN) MODEL
ALEXNET

telah diperiksa dan disetujui untuk diujikan.

Pembimbing I



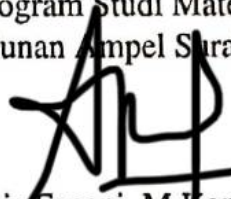
Dian Candra Riri Novitasari, M. Kom
NIP. 198511242014032001

Pembimbing II



Lutfi Hakim, M.Ag
NIP. 197312252006041001

Mengetahui,
Ketua Program Studi Matematika
UIN Sunan Ampel Surabaya



Aris Fanani, M.Kom
NIP. 198701272014031002

PENGESAHAN TIM PENGUJI SKRIPSI

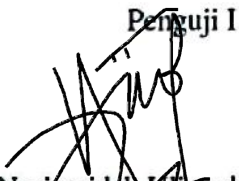
Skripsi oleh

Nama : IRKHANA INDAKA ZULFA
NIM : H72218023
Judul Skripsi : KLASIFIKASI KANKER PARU BERDASARKAN
CITRA HISTOPATOLOGI MENGGUNAKAN METODE
CONVOLUTIONAL NEURAL NETWORK (CNN) MODEL
ALEXNET

Telah dipertahankan di depan Tim Penguji
pada tanggal 17 Januari 2022

Mengesahkan,
Tim Penguji

Penguji I


Nurissaidah Ulumaha, M.Kom
NIP. 199011022014032004

Penguji II


Putros Keumala Lutan, M.Si
NIP. 198805282018012001

Penguji III


Dian Candra Rini Novitasari, M.Kom
NIP. 198511242014032001

Penguji IV


Lutfi Hakim, M.Ag
NIP. 197312252006041001



Mengetahui,

Dekan Fakultas Sains dan Teknologi
JIN Sunan Ampel Surabaya


Prof. Dr. Eyid Falsanatur Rusydiyah, M.Ag
NIP. 197312272005012003



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 : IRKHANA INDAKA ZULFA
NIM : H72218023
Fakultas/Jurusan : SAINTEK / MATEMATIKA
E-mail address : irkhanaindakazulfa@gmail.com

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

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

yang berjudul :

KLASIFIKASI KANKER PARU BERDASARKAN CITRA HISTOPATOLOGI
MENGUNAKAN METODE CONVOLUTIONAL NEURAL NETWORK (CNN)
MODEL ALEXNET

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, 14 FEBRUARI 2022

Penulis

(IRKHANA INDAKA Z)

ABSTRAK

KLASIFIKASI KANKER PARU BERDASARKAN CITRA HISTOPATOLOGI MENGGUNAKAN METODE *CONVOLUTIONAL NEURAL NETWORK* (CNN) MODEL ALEXNET

Kanker paru menempati urutan pertama sebagai penyebab kematian tertinggi di seluruh dunia maupun di Indonesia karena diagnosis kanker. Upaya untuk mengatasi masalah ini perlu dilakukan pemeriksaan dini, salah satunya dengan prosedur bronkoskopi. Saat ini, sistem *Computer-Aided Diagnosis* (CAD) banyak digunakan pada bidang kedokteran untuk membantu mendiagnosis penyakit kanker. Pada penelitian ini digunakan metode *Convolutional Neural Networks* (CNN) model AlexNet untuk membentuk sistem klasifikasi kanker paru berdasarkan citra histopatologi. CNN merupakan salah satu algoritma *Deep Learning* dengan kinerja yang bagus untuk klasifikasi data citra, sedangkan AlexNet merupakan model jaringan CNN yang memenangkan kompetisi ILSVRC 2012. Pada penelitian ini menggunakan 3 kelas, yaitu *benign lung tissue*, *lung adenocarcinoma* dan *lung squamous cell carcinoma*. Pada penelitian ini didapatkan model yang paling optimal dengan presentase akurasi, sensitivitas dan spesifitas sebesar 99%, 99% dan 99% pada pembagian data 80:20, *batch size* sebesar 64 serta nilai probabilitas *dropout* 0.5.

Kata kunci: AlexNet, CNN, *Deep Learning*, Kanker paru

ABSTRACT

CLASSIFICATION OF LUNG CANCER BASED ON HISTOPATHOLOGICAL IMAGES USING THE *CONVOLUTIONAL* *NEURAL NETWORK* (CNN) ALEXNET MODEL

Lung cancer is the leading cause of cancer death in worldwide and Indonesian. Lung cancer can be immediately treated with early examination efforts, one of which is bronchoscopy procedures. Currently, computer-aided detection or diagnosis (CAD) systems are widely used in the medical field to help diagnose cancer. In this study, the AlexNet model Convolutional Neural Networks (CNN) method was used to form a lung cancer classification system based on histopathological images. CNN is a Deep Learning algorithm with good performance for image data classification, while AlexNet is a CNN network model that won the 2012 ILSVRC competition. In this study, 3 classes were used, namely benign lung tissue, lung adenocarcinoma and lung squamous cell carcinoma. In this study, the most optimal model with accuracy, sensitivity and specificity of 99%, 99% and 99% was obtained with 80:20 data division, 64 batch size and the dropout probability value of 0.5.

Keywords: AlexNet, CNN, Deep Learning, Lung cancer

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KEASLIAN	ii
LEMBAR PERSETUJUAN PEMBIMBING	iii
PENGESAHAN TIM PENGUJI SKRIPSI	iv
LEMBAR PERSETUJUAN PUBLIKASI	v
DAFTAR ISI	vi
DAFTAR TABEL	ix
DAFTAR GAMBAR	x
ABSTRAK	xiii
ABSTRACT	xiv
I PENDAHULUAN	1
1.1. Latar Belakang Masalah	1
1.2. Rumusan Masalah	7
1.3. Tujuan Penelitian	7
1.4. Manfaat Penelitian	8
1.5. Batasan Masalah	9
1.6. Sistematika Penulisan	9
II TINJAUAN PUSTAKA	11
2.1. Integrasi Keislaman	11
2.2. Kanker	14
2.3. Kanker Paru	16
2.4. Citra Digital	17
2.4.1. Citra Biner	18
2.4.2. Citra grayscale	18
2.4.3. Citra RGB (<i>Red, Green, Blue</i>)	19
2.5. <i>Convolutional Neural Network</i> (CNN)	19
2.5.1. <i>Convolution Layer</i>	21

2.5.2.	<i>Rectified linear Unit(ReLU)</i>	22
2.5.3.	<i>Cross Channel Normalization Layer</i>	24
2.5.4.	Pooling Layer	24
2.5.5.	<i>Dropout Layer</i>	25
2.5.6.	Fully Connected Layer	26
2.6.	AlexNet	27
2.6.1.	Menghitung Parameter Pada <i>Convolution Layer</i>	29
2.6.2.	Menghitung Parameter Pada <i>Fully Connected Layer</i>	30
2.7.	Perbedaan Parameter dengan Hyperparameter	32
2.8.	<i>Glorot Inizialitation</i>	33
2.9.	<i>Confusion Matrix</i>	34
III	METODE PENELITIAN	37
3.1.	Jenis Penelitian	37
3.2.	Jenis dan Sumber Data	37
3.3.	Kerangka Penelitian	38
IV	HASIL DAN PEMBAHASAN	43
4.1.	<i>Preprocessing</i>	43
4.2.	Pembentukan Sistem Klasifikasi	46
4.2.1.	<i>Convolution Layer</i>	46
4.2.2.	<i>ReLU Layer</i>	50
4.2.3.	<i>Cross Channel Normalization Layer</i>	51
4.2.4.	Max Pooling Layer	54
4.2.5.	<i>Dropout Layer</i>	56
4.2.6.	Fully Connected Layer	57
4.2.7.	Softmax Layer	59
4.3.	Pengujian dan Evaluasi Model	60
4.3.1.	Pembagian Data 60:40	60
4.3.2.	Pembagian Data 70:30	62
4.3.3.	Pembagian Data 80:20	64
4.3.4.	Pembagian Data 90:10	66

4.4. Aplikasi klasifikasi Kanker Paru Citra Histopatologi	73
4.5. Integrasi Keislaman	77
V PENUTUP	79
5.1. Kesimpulan	79
5.2. Saran	80
DAFTAR PUSTAKA	80



UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR TABEL

2.1	Jumlah Parameter Model AlexNet	32
3.1	Sampel Data Citra Histopatologi Kanker Paru	38
3.2	Uji Coba pada Sistem Klasifikasi	42
4.1	Hasil Percobaan Pembagian Data 60:40	61
4.2	Hasil Percobaan Pembagian Data 70:30	63
4.3	Hasil Percobaan Pembagian Data 80:20	65
4.4	Hasil Percobaan Pembagian Data 90:10	67
4.5	Sampel Data Citra Histopatologi Kanker Paru	69

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR GAMBAR

2.1	Perbedaan tumor jinak dan ganas dapat dilihat	15
2.2	(a) <i>CT scan</i> , (b) <i>Magnetic Resonance Imaging (MRI) scan</i> , (c) <i>Mammography</i> , (d) <i>bone scan</i> , (e) <i>X-rays</i> , (f) <i>biopsy</i>	16
2.3	Citra Biner beserta Matriks Pikselnya	18
2.4	Citra <i>Grayscale</i> beserta Matriks Pikselnya	18
2.5	Citra RGB beserta Matriks Pikselnya	19
2.6	Arsitektur <i>Convolutional Neural Network (CNN)</i>	20
2.7	Ilustrasi Proses Konvolusi	21
2.8	Proses Konvolusi pada citra RGB	22
2.9	Perbandingan Grafik Fungsi Aktivasi Tanh (<i>Dashed Line</i>) dan ReLU (<i>Solid Line</i>) Sumber: (Krizhevsky et al., 2012)	23
2.10	Ilustrasi ReLU layer	23
2.11	Proses <i>Max Pooling</i> dan <i>Average Pooling</i>	25
2.12	Proses <i>Dropout</i>	26
2.13	Arsitektur AlexNet	28
2.14	Confusion Matrix Multi Class (a) kelas A (b) kelas B (c) kelas C	34
2.15	Contoh Perhitungan Confusion Matrix Multi Class	35
3.1	Flowchart Penelitian	39
3.2	Flowchart Sistem Klasifikasi	40
3.3	Arsitektur AlexNet Sumber: (Krizhevsky et al., 2012)	41
4.1	Ilustrasi Citra RGB	43
4.2	Nilai Piksel Citra pada <i>Channel Red</i>	44
4.3	Nilai Piksel Citra pada <i>Channel Green</i>	44
4.4	Nilai Piksel Citra pada <i>Channel Blue</i>	44
4.5	Ilustrasi Proses <i>Resize</i> Citra	45
4.6	Nilai Piksel Citra pada <i>Channel Red</i>	45

4.7	Nilai Piksel Citra pada <i>Channel Green</i>	45
4.8	Nilai Piksel Citra pada <i>Channel Blue</i>	45
4.9	Ilustrasi Perhitungan pada <i>Convolution Layer</i>	46
4.10	Ilustrasi Perhitungan <i>Convolution Layer channel Red</i>	47
4.11	Ilustrasi Perhitungan <i>Convolution Layer channel Green</i>	48
4.12	Ilustrasi Perhitungan <i>Convolution Layer channel Blue</i>	48
4.13	Visualisasi <i>Feature Map</i> pada <i>Convolution Layer</i>	50
4.14	Visualisasi <i>Feature Map</i> pada <i>ReLU Layer</i>	51
4.15	Visualisasi <i>Feature Map</i> pada <i>CCN Layer</i>	54
4.16	Visualisasi <i>Feature Map</i> pada <i>Max Pooling Layer</i>	55
4.17	Ilustrasi Proses <i>Flatten channel Blue</i>	56
4.18	Perhitungan <i>dropout Layer</i>	57
4.19	<i>Training Progress</i> Model Optimal pada Pembagian Data 60:40 . . .	61
4.20	<i>Confusion Matrix</i> Model Optimal pada Pembagian Data 60:40 . . .	62
4.21	<i>Training Progress</i> Model Optimal pada Pembagian Data 70:30 . . .	63
4.22	<i>Confusion Matrix</i> Model Optimal pada Pembagian Data 70:30 . . .	64
4.23	<i>Training Progress</i> Model Optimal pada Pembagian Data 80:20 . . .	65
4.24	<i>Confusion Matrix</i> Model Optimal pada Pembagian Data 80:20 . . .	66
4.25	<i>Training Progress</i> Model Optimal pada Pembagian Data 90:10 . . .	67
4.26	<i>Confusion Matrix</i> Model Optimal pada Pembagian Data 90:10 . . .	68
4.27	Grafik Pengaruh Pembagian Data terhadap Waktu <i>Training</i>	69
4.28	Grafik Pengaruh Pembagian Data terhadap Presentase Akurasi . . .	70
4.29	Grafik Pengaruh <i>Batch Size</i> terhadap Waktu <i>Training</i>	70
4.30	Grafik Pengaruh <i>Batch Size</i> terhadap Presentase Akurasi	71
4.31	Grafik Pengaruh Probabilitas <i>Dropout</i> terhadap Waktu <i>Training</i> . . .	72
4.32	Grafik Pengaruh Probabilitas <i>Dropout</i> terhadap Presentase Akurasi .	73
4.33	Tampilan awal Aplikasi Sistem Klasifikasi	74
4.34	Tampilan Utama Aplikasi Sistem Klasifikasi	74
4.35	Contoh Hasil Klasifikasi pada Kelas <i>Benign Lung Tissue</i>	75
4.36	Contoh Hasil Klasifikasi pada Kelas <i>Lung Adenocarcinoma</i>	76

4.37 Contoh Hasil Klasifikasi pada Kelas *Squamous Lung Cell Carcinoma* 76



UIN SUNAN AMPEL
S U R A B A Y A

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Pada tubuh manusia terdapat proses di mana ketika sel-sel sudah menjadi tua, maka sel tersebut akan mati dan digantikan oleh sel-sel baru. Namun jika terdapat sekelompok sel-sel abnormal yang terbentuk dari hasil proses pembelahan sel dalam jaringan tubuh secara berlebihan serta terus menyebar di luar kendali, maka sel-sel tersebut dapat membentuk gumpalan jaringan atau tumor (Masud et al., 2021). Berdasarkan pertumbuhannya tumor dapat dibedakan menjadi tumor ganas (*malignant*) dan tumor jinak (*benign*). Tumor ganas dapat disebut sebagai kanker yang dapat menyebar lalu berakhir di kelenjar getah bening atau organ tubuh lainnya sehingga menyebabkan masalah fungsi normal pada tubuh (Braithwaite, 2020). Menurut the *global burden of disease*, menyatakan bahwa penyakit kanker menjadi penyakit penyebab kematian tertinggi ke dua di dunia setelah penyakit kardiovaskular, dengan kasus meninggal sebesar 9.56 juta jiwa (Roth et al., 2018). Berdasarkan data dari Globocan tahun 2020, pada pria sering diderita penyakit kanker paru-paru, prostat, kolorektal, lambung dan hati. Sedangkan pada wanita sering diderita penyakit kanker payudara, kolorektal, paru-paru, serviks dan tiroid (Sung et al., 2021).

Kanker paru terjadi ketika terdapat pembelahan sel tidak normal pada dinding saluran organ paru dan terjadi tidak terkontrol hingga membentuk suatu benjolan pada jaringan. Kasus meninggal akibat kanker paru mencapai 18% dari

diagnosis penyakit kanker. Hal ini menyebabkan kanker paru menempati urutan pertama sebagai penyebab kematian tertinggi di seluruh dunia yang disebabkan karena kanker (Ferlay et al., 2021). Sedangkan di Indonesia kasus meninggal akibat kanker paru mencapai 13.2% dari diagnosis penyakit kanker. Hal ini menjadikan kanker paru menempati urutan pertama sebagai penyebab kematian tertinggi di Indonesia karena penyakit kanker (World Health Organization, 2020).

Kanker paru umumnya dibagi menjadi dua kategori besar, yakni *small cell lung cancer* (SCLC) dan *non-small cell lung cancer* (NSCLC). Kategori NSCLC terbagi lagi menjadi *adenocarcinoma* dan *squamous cell carcinoma* (McIntyre and Ganti, 2017). Sekitar 80% sampai 85% kasus kanker paru merupakan NSCLC. Faktor risiko penyebab berkembangnya kanker paru, ada yang tidak dapat dikendalikan seperti, usia, jenis kelamin, dan genetik. Sedangkan faktor risiko yang dapat dikendalikan adalah seperti, merokok, radon, polutan indoor, polutan outdoor, penyakit paru kronik, dan karsinogen di lingkungan kerja (Kemenkes RI, 2018). Sehingga, wajib bagi manusia untuk menjaga serta menjauhkan diri dari kebiasaan buruk yang dapat memperburuk kesehatan. Sebagaimana firman Allah pada surah al-Ma'idah ayat 88,

وَكُلُوا مِمَّا رَزَقَكُمُ اللَّهُ حَلَالًا طَيِّبًا وَاتَّقُوا اللَّهَ الَّذِي أَنْتُمْ بِهِ مُؤْمِنُونَ ﴿٨٨﴾

Artinya: “Dan makanlah makanan yang halal lagi baik dari apa yang Allah telah rezekikan kepadamu, dan bertakwalah kepada Allah yang kamu beriman kepada-Nya”. Berdasarkan ayat tersebut, Allah memerintahkan kepada manusia agar menjaga kesehatan melihat dari segi makanan, dimana dijelaskan agar manusia makan dari rezeki yang halal dan baik. Halal di sini mengandung arti halal bendanya dan cara memperolehnya. Sedangkan arti baik adalah dari segi kemanfaatannya, yaitu yang mengandung manfaat serta maslahat bagi tubuh,

seperti mengandung gizi, vitamin, protein dan sebagainya. Ayat ini merupakan peringatan bagi manusia untuk memperhatikan pola hidup sehat melalui makanan yang dikonsumsi karena dapat berpengaruh terhadap kesehatan jasmani serta rohani.

Kanker paru dapat segera diatasi dengan upaya pemeriksaan secara dini. Pemeriksaan sejak dini umumnya lebih disarankan untuk orang-orang yang memiliki faktor genetik yang dapat menyebabkan resiko terjadinya kanker paru. Meskipun memiliki riwayat keluarga yang pernah mengidap penyakit kanker paru, tetap tidak boleh putus asa dan terus berusaha dengan menjalani serangkaian pemeriksaan dini, hal ini sejalan dengan firman Allah dalam surah at-Tagabun ayat 11,

مَا أَصَابَ مِنْ مُصِيبَةٍ إِلَّا بِإِذْنِ اللَّهِ وَمَنْ يُؤْمِنْ بِاللَّهِ يَهْدِ اللَّهُ قَلْبَهُ وَاللَّهُ بِكُلِّ شَيْءٍ عَلِيمٌ ﴿١١﴾

Artinya: “Tidak ada suatu musibah pun yang menimpa seseorang kecuali dengan ijin Allah; dan barangsiapa yang beriman kepada Allah niscaya Dia akan memberi petunjuk kepada hatinya. Dan Allah Maha Mengetahui segala sesuatu.”

Berdasarkan ayat tersebut, segala musibah yang menimpa manusia semua terjadi karena ketentuan Allah. Kemudian dijelaskan juga barangsiapa yang beriman kepada Allah dengan istiqamah, niscaya Allah akan memberi petunjuk kepada hatinya dengan memantapkan imannya. Ayat ini memperingatkan kita dalam berusaha keras, manusia hendaknya tidak menyesal dan merasa kecewa apabila menemui hal-hal yang tidak sesuai dengan usaha dan keinginannya. Sejalan dengan masalah penderita kanker paru, yang seharusnya dilakukan adalah tetap berusaha dengan melakukan pemeriksaan dini dan mendekatkan diri kepada Allah yang Maha Mengetahui segala sesuatu.

Pemeriksaan melalui foto toraks AP/lateral merupakan tahap awal

pemeriksaan kanker paru secara dini. Kemudian jika ditemukan lesi ganas pada foto toraks, maka untuk mengevaluasi lesi tersebut dilakukan pemeriksaan *Computed Tomography (CT) scan* toraks. Selanjutnya prosedur utama untuk mendiagnosis kanker paru adalah pemeriksaan bronkoskopi. Pada prosedur ini didapatkan spesimen untuk pemeriksaan histopatologi, sehingga klasifikasi jenis kanker paru dapat ditentukan (Kemenkes RI, 2017).

Pemeriksaan kanker paru dilakukan secara *operator dependent* atau hasil diagnosis tergantung pada ahli medis yang melakukan pemeriksaan. Hal ini dapat menyebabkan munculnya salah diagnosis, di mana hasil lesi terdiagnosis sebagai non-lesi, atau kanker lesi ganas terdiagnosis lesi jinak (Li et al., 2002). Menurut data, tingkat kesalahan diagnosis yang disebabkan oleh manusia dalam analisis citra medis mencapai 10-30% (Kerlikowske et al., 2000). Upaya untuk mengurangi kemungkinan adanya human error, dapat menggunakan sistem *Computer-Aided Diagnosis (CAD)*. Saat ini, dalam bidang kedokteran banyak yang memanfaatkan teknologi *Artificial Intelligence (AI)* untuk membantu mendiagnosis kanker. Terdapat penelitian yang membandingkan performa CNN dengan ahli medis (Esteva et al., 2017) dalam mengklasifikasikan citra sebagai lesi jinak, ganas atau non-neoplastik. Hasilnya CNN lebih unggul dibanding dengan 2 ahli medis, dengan akurasi sebesar 72%, 65% dan 66% secara berturut-turut.

Beberapa penelitian sebelumnya, telah membuktikan bahwa teknologi AI dapat diimplementasikan untuk mengklasifikasi kanker serta model yang dihasilkan memiliki akurasi yang baik. Penelitian mengenai klasifikasi kanker payudara dilakukan menggunakan metode *Convolutional neural networks (CNN)*. Hasilnya metode CNN memiliki akurasi sebesar 97%. Peneliti menyatakan bahwa AI dapat mengurangi kemungkinan *human error* hingga 25% (Liu et al., 2017).

Hal ini sejalan dengan penelitian yang mengimplementasikan AI dalam mendeteksi metastasis kanker payudara. Hasilnya menunjukkan bahwa metode AI dapat mengatasi *human error*, dimana hasil diagnosis ahli medis memiliki presentase akurasi sebesar 86%. Sedangkan dengan bantuan AI, didapatkan akurasi sebesar 99.5% (Wang et al., 2016). Teknologi AI juga unggul pada segi waktu komputasi dalam mengklasifikasikan citra kanker paru jika dibandingkan dengan pemeriksaan manual oleh ahli medis. Terdapat penelitian yang menggunakan teknologi AI untuk mengklasifikasikan kanker paru dalam tiga kelas, dibutuhkan waktu komputasi 2 jam sehingga didapatkan model untuk sistem klasifikasi kanker paru (Apsari et al., 2021). Sedangkan pada pemeriksaan manual oleh ahli medis dibutuhkan waktu 10-14 hari untuk mendiagnosis penyakit kanker paru (Alodokter, 2020).

Salah satu cabang pembelajaran dari AI yang saat ini sering digunakan dalam penelitian adalah algoritma *Deep Learning* (DL). Algoritma DL merupakan salah satu implementasi dari *Machine Learning* (ML), yang menggunakan *Artificial Neural Network* (ANN) pada proses klasifikasinya. Terdapat beberapa metode dalam algoritma DL, salah satunya adalah *Convolutional Neural Network* (CNN) yang dapat diimplementasikan pada sistem klasifikasi berdasarkan data citra (Hibatullah and Maliki, 2019). Pada penelitian sebelumnya metode CNN telah diimplementasikan untuk mengklasifikasikan kanker paru. Salah satunya, adalah penelitian dari Mangal dkk (Mangal et al., 2020) yang menggunakan dataset LC25000. Terdapat tiga kelas, yaitu *squamous cell carcinomas*, *adenocarcinomas* dan *benign*. Hasilnya didapatkan presentase akurasi sebesar 97%.

Pada metode CNN terdapat beberapa model yang sebelumnya telah

diimplementasikan menggunakan dataset ImageNet, seperti LeNet, AlexNet, VGG, GoogleNet dan ResNet (Shanthi and Sabeenian, 2019). Model AlexNet meraih peringkat pertama di kompetisi *ImageNet Large-Scale Visual Recognition Challenge* (ILSVRC) pada tahun 2012 (Krizhevsky et al., 2012). Model AlexNet telah digunakan pada beberapa penelitian sebelumnya. Penelitian yang dilakukan oleh Hosny dkk menggunakan model AlexNet untuk klasifikasi kanker kulit (Hosny et al., 2020). Peneliti menggunakan dataset ISIC 2018, yang terdiri dari 7 kelas. Hasilnya didapatkan presentase akurasi, sensitivitas, spesitivitas dan presisi sebesar 98.70%, 95.60%, 99.27% dan 95.06% secara berturut-turut.

Shaheen dkk melakukan penelitian mengenai klasifikasi kanker darah dengan membandingkan model LeNet dengan AlexNet (Shaheen et al., 2021). Peneliti menggunakan dataset dari Mendeley yang memiliki 3 kelas. Hasilnya menunjukkan AlexNet lebih unggul dibandingkan dengan LeNet. Berdasarkan nilai akurasi dan presisi AlexNet memiliki presentase sebesar 98.58% dan 87.40%, sedangkan LeNet memiliki presentase 96.25% dan 83.60%. Selanjutnya terdapat penelitian perbandingan metode CNN model LeNet, AlexNet, serta VGG-16 menggunakan dataset MNIST yang memiliki 10 kelas (Kulkarni and Harnoorkar, 2020). Hasilnya didapatkan nilai akurasi model LeNet, AlexNet dan VGG-16 sebesar 97.2%, 99.2% dan 98.20% secara berturut-turut.

Selanjutnya terdapat penelitian yang membandingkan waktu komputasi model AlexNet, GoogleNet, Inception V3, ResNet18, dan ResNet50 (Maeda-Gutiérrez et al., 2020), hasilnya didapatkan waktu komputasi sebesar 94 menit 17 detik, 133 menit 15 detik, 649 menit 85 detik, 147 menit 73 detik, dan 476 menit 78 detik secara berturut-turut. Sedangkan presentase akurasi diperoleh sebesar 98.93%, 99.39%, 98.65%, 99.06% dan 99.15% secara berturut-turut.

Sehingga penulis menyimpulkan bahwa model AlexNet menunjukkan performa yang bagus berdasarkan waktu komputasi tercepat dibandingkan model CNN lainnya.

Berdasarkan beberapa penelitian sebelumnya, metode CNN model AlexNet memiliki rata-rata akurasi di atas 90%, sehingga dapat dikatakan model AlexNet memiliki kinerja yang baik pada sistem klasifikasi. Oleh karena itu, pada penelitian ini digunakan metode CNN model AlexNet untuk sistem klasifikasi kanker paru berdasarkan citra histopatologi. Penelitian ini diharapkan dapat membangun sistem klasifikasi kanker paru berdasarkan citra histopatologi menggunakan metode CNN model AlexNet.

1.2. Rumusan Masalah

Berdasarkan masalah yang dijelaskan pada bab latar belakang, maka perumusan masalah penelitian ini adalah sebagai berikut:

1. Bagaimana hasil model yang optimal dalam sistem klasifikasi kanker paru citra histopatologi menggunakan metode CNN model AlexNet berdasarkan uji coba pembagian data, *batch size* dan probabilitas *dropout*?
2. Bagaimana pengaruh uji coba pembagian data, *batch size* dan probabilitas *dropout* dalam sistem klasifikasi kanker paru berdasarkan citra histopatologi menggunakan Convolutional neural network model AlexNet terhadap waktu *training* serta presentase akurasi?

1.3. Tujuan Penelitian

Berdasarkan rumusan masalah diatas, maka tujuan penelitian ini adalah sebagai berikut:

1. Dapat mengetahui serta menganalisis bagaimana model yang optimal dalam sistem klasifikasi kanker paru berdasarkan data histopatologi menggunakan metode *Convolutional neural network* model AlexNet dengan uji coba pembagian data, *batch size* dan probabilitas *dropout*.
2. Dapat mengetahui serta menganalisis bagaimana pengaruh uji coba pembagian data, *batch size* dan probabilitas *dropout* dalam sistem klasifikasi kanker paru berdasarkan citra histopatologi menggunakan *Convolutional neural network* model AlexNet terhadap waktu *training* serta presentase akurasi?

1.4. Manfaat Penelitian

Penelitian ini diharapkan bermanfaat bagi lapisan masyarakat, seperti yang tertulis sebagai berikut:

1. Manfaat Teoritis

Penelitian ini dapat dimanfaatkan sebagai salah satu acuan untuk penelitian selanjutnya dalam sistem klasifikasi menggunakan metode *Convolutional neural network* model AlexNet

2. Manfaat Praktis

(a) Bagi Penulis

Menambah pengetahuan serta wawasan mengenai sistem klasifikasi kanker paru-paru menggunakan metode CNN model AlexNet

(b) Bagi Pembaca

Menambah wawasan mengenai metode yang dapat digunakan dalam sistem klasifikasi kanker serta dapat dijadikan referensi untuk penelitian selanjutnya.

(c) Bagi Ahli Medis

Membantu ahli medis dalam mendiagnosis kanker paru serta memberikan penanganan yang tepat sehingga dapat mengurangi kasus kematian akibat kanker paru.

1.5. Batasan Masalah

Batasan masalah dari penelitian ini adalah sebagai berikut:

1. Dataset yang digunakan pada penelitian ini merupakan data kanker paru-paru yang diperoleh dari LC25000 dataset berupa citra histopatologi.
2. Dataset kanker paru-paru diklasifikasikan ke dalam tiga kelas, yaitu *benign lung tissue*, *lung adenocarcinoma*, dan *lung squamous cell carcinoma*.
3. Penelitian ini menggunakan komputer dengan spesifikasi processor Intel Core i5, RAM 4GB, dan VGA intel HD 2GB.

1.6. Sistematika Penulisan

Sistematika penyusunan pada penelitian ini adalah sebagai berikut:

1. BAB I PENDAHULUAN

Pada bab pendahuluan berisi mengenai latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan.

2. BAB II TINJAUAN PUSTAKA

Pada bab tinjauan pustaka berisi mengenai landasan teori yang digunakan dalam penelitian ini. Landasan teori tersebut meliputi kanker, kanker paru-paru, convolutional neural network (CNN), AlexNet, dan confusion matrix.

3. **BAB III METODE PENELITIAN**

Pada bab metode penelitian berisi mengenai jenis penelitian, jenis dan sumber data yang digunakan, serta langkah-langkah penelitian yang akan dilakukan.

4. **BAB IV HASIL DAN PEMBAHASAN**

Pada bab hasil dan pembahasan berisi mengenai hasil dari penerapan metode CNN model AlexNet dalam klasifikasi kanker paru serta analisis untuk menarik kesimpulan sesuai dengan tujuan penelitian.

5. **BAB V PENUTUP**

Pada bab Penutup berisi mengenai kesimpulan penelitian serta saran dari peneliti untuk penelitian selanjutnya.



UIN SUNAN AMPEL
S U R A B A Y A

BAB II

TINJAUAN PUSTAKA

2.1. Integrasi Keislaman

Di dunia yang sifatnya sementara ini, setiap manusia tentu akan mengalami ujian, baik itu orang yang beriman maupun orang kafir kepada Allah Swt. Namun, terdapat perbedaan pada keduanya yaitu, bagaimana manusia dapat memahami hakikat ujian Allah serta bagaimana mereka menyikapinya. Sebab dengan adanya ujian, manusia memiliki tingkatan derajat serta kemuliaan di hadapan Allah Swt. di mana semakin tinggi kedudukan seseorang, tentu semakin berat pula ujian yang dihadapinya. Hal ini sejalan dengan kisah Nabi Ayyub yang telah diuji dengan kekayaan harta, namun kemudian Allah uji dengan kemiskinan serta penyakit berat selama belasan tahun, namun nabi Ayyub tetap bersabar dalam menghadapi ujian. Allah Swt berfirman pada surah al Ankabut ayat 2-3.

UIN
S U

أَحْسِبَ النَّاسُ أَنْ يُتْرَكُوا أَنْ يَقُولُوا ءَامَنَّا وَهُمْ لَا يُفْتَنُونَ ﴿٢﴾
وَلَقَدْ فَتَنَّا الَّذِينَ مِنْ قَبْلِهِمْ فَلَيَعْلَمَنَّ اللَّهُ الَّذِينَ صَدَقُوا وَلَيَعْلَمَنَّ الْكٰذِبِينَ ﴿٣﴾

Artinya: “Apakah manusia itu mengira bahwa mereka dibiarkan (saja) mengatakan: ”Kami telah beriman”, sedang mereka tidak diuji lagi. Dan sesungguhnya kami telah menguji orang-orang yang sebelum mereka, maka sesungguhnya Allah mengetahui orang-orang yang benar dan sesungguhnya Dia mengetahui orang-orang yang dusta.”

Ujian juga merupakan bentuk Allah menghendaki kebaikan hamba-Nya dengan memberikan ujian sebagai hukuman atas dosa yang di lakukan serta

sebagai tanda Allah menghendaki kejelekan kepada hamba-Nya dengan membiarkan dalam kemaksiatan tanpa peringatan dari-Nya berupa musibah dan cobaan.

عَنْ أَنَسٍ رَضِيَ اللَّهُ عَنْهُ قَالَ: قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ:

إِذَا أَرَادَ اللَّهُ بِعَبْدِهِ الْخَيْرَ عَجَّلَ لَهُ الْعُقُوبَةَ فِي الدُّنْيَا وَإِذَا أَرَادَ اللَّهُ بِعَبْدِهِ الشَّرَّ أَمْسَكَ عَنْهُ بِذَنْبِهِ حَتَّى يُؤَافِيَ بِهِ يَوْمَ الْقِيَامَةِ

Diriwayatkan dari Anas ibn Malik radhiyallahu 'anhu berkata: Rasulullah shallallahu'alaihi wasallam bersabda: “Apabila Allah menghendaki kebaikan untuk hamba-Nya maka Dia akan menyegerakan untuknya hukuman di dunia. Dan apabila Allah menghendaki kejelekan untuk hamba-Nya maka Dia akan menahan darinya hukuman karena dosanya sehingga kelak di akhirat Dia akan menyempurnakan hukuman untuknya” (HR. At-Tirmidzi).

Ujian yang diberikan Allah untuk menguji iman hamba-Nya bisa melalui kenikmatan dan juga bisa berupa musibah dan kesulitan. Ujian berupa musibah dapat berupa penyakit, saat ini penyakit yang paling banyak diderita serta dapat menyebabkan kematian di Indonesia adalah kanker paru. Salah satu faktor yang menyebabkan menyebarnya sel kanker adalah kebiasaan buruk merokok sehingga menyebabkan pola hidup tidak sehat. Oleh karena itu, wajib bagi manusia untuk menjaga serta menjauhkan diri dari kebiasaan buruk yang dapat memperburuk kesehatan. Sebagaimana firman Allah pada surah al-Ma'idah ayat 88,

وَكُلُوا مِمَّا رَزَقَكُمُ اللَّهُ حَلَالًا طَيِّبًا وَاتَّقُوا اللَّهَ الَّذِي أَنْتُمْ بِهِءِ مُؤْمِنُونَ ﴿٨٨﴾

Artinya: “Dan makanlah makanan yang halal lagi baik dari apa yang Allah telah rezekikan kepadamu, dan bertakwalah kepada Allah yang kamu beriman kepada-Nya”. Berdasarkan ayat tersebut, Allah memerintahkan kepada manusia

agar menjaga kesehatan melihat dari segi makanan, di mana dijelaskan agar manusia makan dari rezeki yang halal dan baik. Halal di sini mengandung arti halal bendanya dan cara memperolehnya. Sedangkan arti baik adalah dari segi kemanfaatannya, yaitu yang mengandung manfaat serta maslahat bagi tubuh, seperti mengandung gizi, vitamin, protein dan sebagainya. Ayat ini merupakan peringatan bagi manusia untuk memperhatikan pola hidup sehat melalui makanan yang dikonsumsi karena dapat berpengaruh terhadap kesehatan jasmani serta rohani.

Di dalam islam selalu diajarkan untuk bertawaqal di mana seorang hamba berusaha terlebih dahulu kemudian menyerahkan dan mempercayakan segala urusannya hanya kepada Allah Swt, termasuk dalam hal kesembuhan atas suatu penyakit. Bentuk usaha yang dapat dilakukan adalah berobat, karena setiap penyakit pasti ada obatnya. Allah Swt telah menjamin bahwa seluruh penyakit yang menimpa seorang hamba pasti ada obatnya. Rasulullah SAW bersabda,

مَا أَنْزَلَ اللَّهُ دَاءً إِلَّا أَنْزَلَ لَهُ شِفَاءً

“Tidaklah Allah menurunkan suatu penyakit, melainkan akan menurunkan pula obat untuk penyakit tersebut” (H.R. Bukhari). Hadis tersebut menjelaskan bahwa segala penyakit yang ada pasti terdapat obatnya, namun tidak lupa bahwa kesembuhan hanya karena izin Allah Swt.

حَدَّثَنَا هَارُونُ بْنُ مَعْرُوفٍ وَأَبُو الظَّاهِرِ وَأَحْمَدُ بْنُ عِيسَى قَالُوا حَدَّثَنَا ابْنُ وَهْبٍ أَخْبَرَنِي عَمْرُو
وَهُوَ ابْنُ الْحَارِثِ عَنْ عَبْدِ رَبِّهِ بْنِ سَعِيدٍ عَنْ أَبِي الزُّبَيْرِ عَنْ جَابِرٍ عَنْ رَسُولِ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ
وَسَلَّمَ أَنَّهُ قَالَ لِكُلِّ دَاءٍ دَوَاءٌ فَإِذَا أُصِيبَ دَوَاءُ الدَّاءِ بَرَأَ بِإِذْنِ اللَّهِ عَزَّ وَجَلَّ

Telah menceritakan kepada kami Harun bin Ma'ruf dan Abu Ath Thahir serta Ahmad bin 'Isa mereka berkata; Telah menceritakan kepada kami Ibnu Wahb;

Telah mengabarkan kepadaku 'Amru, yaitu Ibnu al-Harits dari 'Abdu Rabbih bin Sa'id dari Abu Az Zubair dari Jabir dari Rasulullah shallallahu 'alaihi wasallam, beliau bersabda: "Setiap penyakit ada obatnya. Apabila ditemukan obat yang tepat untuk suatu penyakit, akan sembuhlah penyakit itu dengan izin Allah 'azza wajalla." (HR Muslim).

Hal yang perlu diingat bahwa setiap cobaan atau ujian yang diberikan Allah Swt kepada hamba-Nya selalu diselipkan solusi. Solusi tersebut disesuaikan dengan kadar tingkatan manusia itu sendiri. Allah Swt berfirman dalam Surah al-Baqarah ayat 286.

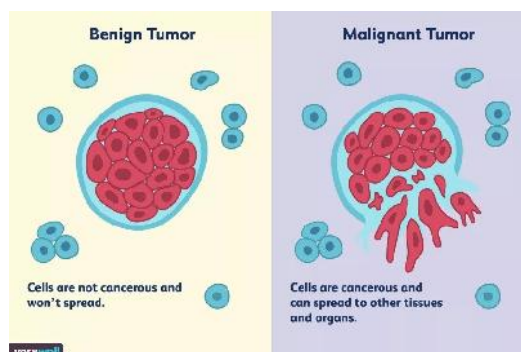
لَا يُكَلِّفُ اللَّهُ نَفْسًا إِلَّا وُسْعَهَا لَهَا مَا كَسَبَتْ وَعَلَيْهَا مَا اكْتَسَبَتْ ...

Artinya: "Allah tidak membebani seseorang melainkan sesuai dengan kesanggupannya. Ia mendapat pahala (dari kebajikan) yang diusahakannya dan ia mendapat siksa (dari kejahatan) yang dikerjakannya." Dijelaskan pada Surah al-Baqarah ayat 286 bahwa Allah Swt tidak membebani hambanya kecuali sesuai dengan kemampuannya.

2.2. Kanker

Kanker dimulai ketika terdapat sel abnormal di dalam tubuh. Hal ini disebabkan karena sel membelah diri melebihi seharusnya atau sel tidak mati ketika seharusnya mati sehingga mengganggu proses pembelahan sel dalam tubuh (Patel, 2020). Sel-sel ini mulai tumbuh secara tidak terkendali serta dapat membentuk suatu massa padat yang disebut tumor. Tumor bisa bersifat ganas (*malignant*) dan bisa bersifat jinak (*benign*). Tumor jinak bersifat non-kanker serta tidak menyebar ke jaringan tubuh lain. Sedangkan tumor ganas bersifat kanker serta dapat menyebar ke jaringan sekitarnya (Samah et al., 2017). Perbedaan tumor

jinak dan ganas dapat dilihat dilihat pada Gambar 2.1.

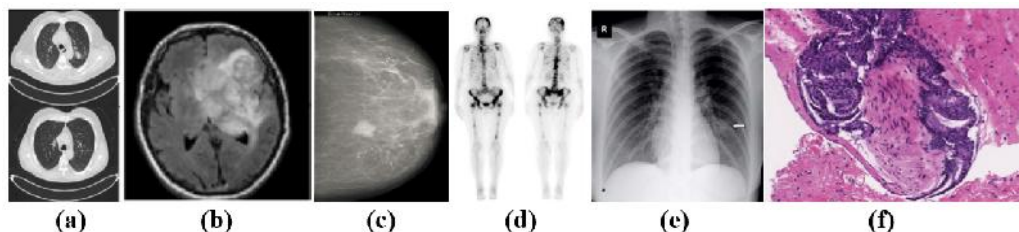


Gambar 2.1 Perbedaan tumor jinak dan ganas dapat dilihat

Sumber: (Fayed, 2021)

Beberapa sel kanker dapat berpindah melalui aliran darah atau kelenjar betah bening menuju ke jaringan lain di dalam tubuh, proses ini dikenal sebagai metastasis. Kanker dapat terjadi di berbagai bagian tubuh, seperti payudara, paru-paru, usus, organ reproduksi, darah atau kulit (Sinha, 2018). Terdapat beberapa faktor risiko penyebab berkembangnya kanker, ada yang tidak dapat dikendalikan seperti, usia dan riwayat keluarga serta ada yang dapat dikendalikan seperti kebiasaan merokok (American Cancer Society, 2018).

Pertumbuhan sel kanker dapat diagnosis sejak dini agar pasien dapat segera diatasi. Terdapat berbagai cara untuk mengetahui ada tidaknya sel kanker dalam tubuh yang dapat dilihat pada Gambar 2.2, seperti *CT scan* (Al-Yasriy et al., 2020), *Magnetic Resonance Imaging (MRI) scan* (Kang et al., 2021), *Mammography* (Dabral et al., 2021), *bone scan* (Papandrianos et al., 2020), *X-rays* (Widodo et al., 2021), dan *biopsi* (Pantanowitz et al., 2020).



Gambar 2.2 (a) *CT scan*, (b) *Magnetic Resonance Imaging (MRI) scan*, (c) *Mammography*, (d) *bone scan*, (e) *X-rays*, (f) *biopsy*

2.3. Kanker Paru

Kanker paru merupakan jenis kanker yang berada di bagian paru-paru. Kanker paru muncul ketika sel tumor pada tubuh bagian paru-paru mulai tumbuh di luar kendali (Society, 2019b). Pada umumnya, kanker paru dibagi menjadi dua kategori, yakni kanker paru non-sel kecil (*non-small cell lung cancer-NSCLC*) dan kanker paru sel kecil (*small cell lung cancer-SCLC*) (health, 2021). Sekitar 85% kasus kanker paru berasal dari kategori NSCLC (Society, 2019b). Berdasarkan awal mula tumbuhnya sel kanker, pada kategori NSCLC terbagi lagi menjadi *adenocarcinoma* dan *squamous cell carcinoma*. *Adenocarcinoma* tumbuh pada sel kelenjar bagian luar paru-paru sedangkan *squamous cell carcinoma* tumbuh pada sel lapisan bagian dalam bronkus. Pada tahun 2020 kanker paru telah merenggut korban jiwa sebanyak 1.796.144 di dunia (Globocan, 2020). Hal ini menyebabkan kanker paru menempati peringkat pertama dalam penyakit yang menyebabkan kematian karena kanker dengan presentase sebesar 18%.

Pada tahap awal, kanker paru-paru tidak menimbulkan gejala atau gejalanya disalahartikan sebagai flu biasa. Beberapa gejala awal pada kanker paru antara lain, batuk terus-menerus yang tidak hilang, dahak disertai dengan darah, nyeri dada pada saat menarik napas dalam-dalam, infeksi paru-paru, suara serak, kehilangan selera makan, penurunan berat badan, serta selalu kelelahan (Society,

2019a). Faktor risiko utama kanker paru adalah merokok. Pada laki-laki, kasus kanker mencapai 80% karena merokok, sedangkan pada perempuan mencapai 50% kasus kanker disebabkan karena merokok. Faktor lain yang menyebabkan kanker paru antara lain, kerentanan genetik, polusi udara, pajanan radon, dan pajanan industri (asbestos, silika, dan lain-lain) (Kemenkes RI, 2017).

2.4. Citra Digital

Citra merupakan representasi visual dari suatu objek, yang dapat didefinisikan sebagai fungsi $f(x, y)$, dengan x dan y merupakan koordinat spasial dan amplitudo f di titik koordinat (x, y) merupakan intensitas atau tingkat keabuan dari citra (Putra, 2010). Secara matematis, citra adalah representasi visual dari suatu objek berupa matriks dari citra dua dimensi menggunakan sejumlah elemen titik, biasa disebut piksel (Tyagi, 2018). Setiap piksel diwakili oleh nilai numerik. Persamaan 2.1 merupakan matriks yang merepresentasikan citra digital dengan ukuran $M_{column} \times N_{row}$.

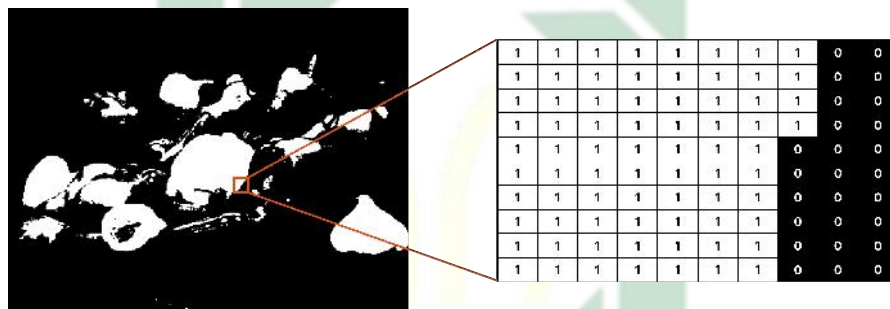
$$f(x, y) = \begin{bmatrix} f(0, 0) & f(0, 1) & \dots & f(0, M - 1) \\ f(1, 0) & f(1, 1) & \dots & f(1, M - 1) \\ \vdots & \vdots & & \vdots \\ f(N - 1, 0) & f(N - 1, 1) & \dots & f(N - 1, M - 1) \end{bmatrix} \quad (2.1)$$

Citra digital merupakan hasil dari proses konversi citra analog yang bersifat kontinu menjadi citra digital yang bersifat diskrit, proses ini disebut sebagai pendigitalan citra (Sulistiyani et al., 2016). Proses pendigitalan citra dilakukan agar citra dapat diproses menggunakan komputer citra harus dinyatakan secara numerik dan bernilai diskrit (Sutoyo et al., 2009). Berdasarkan tingkat warnanya,

citra digital dapat dibagi menjadi tiga jenis citra (Andono and Sutojo, 2017). Berikut merupakan jenis-jenis citra digital berdasarkan tingkat warnanya.

2.4.1. Citra Biner

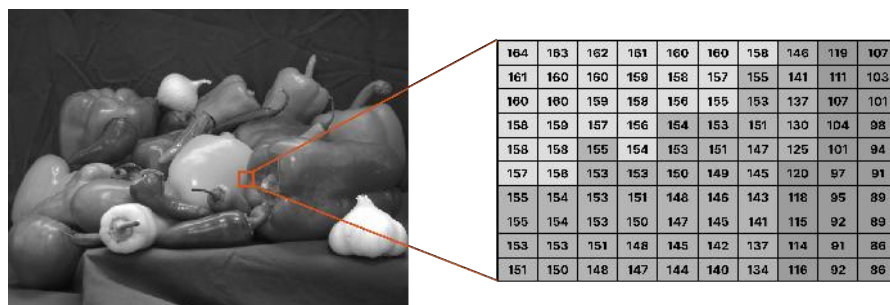
Citra biner hanya memiliki dua nilai piksel, yaitu 0 dan 1. Nilai piksel 0 merepresentasikan warna hitam sedangkan nilai piksel 1 merepresentasikan warna putih. Gambar 2.3 menunjukkan citra biner beserta dengan nilai intensitas pikselnya.



Gambar 2.3 Citra Biner beserta Matriks Pikselnya

2.4.2. Citra grayscale

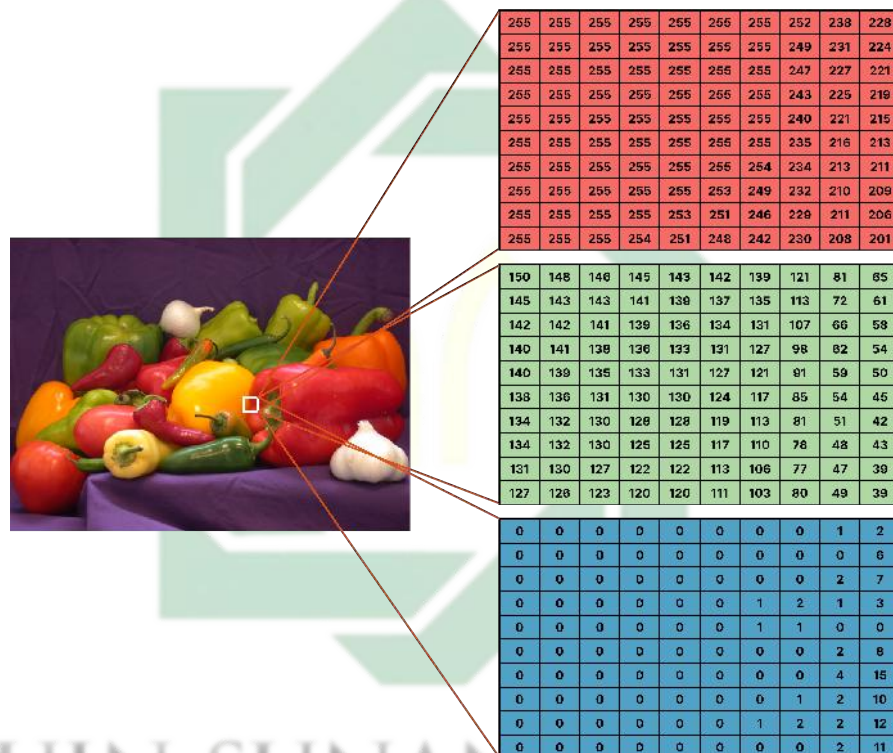
Pada citra *grayscale* memiliki nilai piksel di antara rentang 0 sampai 255. Pada setiap piksel merepresentasikan gradasi warna dari warna hitam yang bernilai 0 hingga putih yang bernilai 255. Gambar 2.4 menunjukkan citra *grayscale* beserta dengan nilai intensitas pikselnya.



Gambar 2.4 Citra *Grayscale* beserta Matriks Pikselnya

2.4.3. Citra RGB (*Red, Green, Blue*)

Pada citra RGB memiliki nilai piksel yang mewakili dari tiga lapisan komponen warna, yaitu *red* (R), *green* (G), dan *blue* (B). Pada masing-masing komponen warna tersebut memiliki nilai piksel di antara rentang 0 sampai 255. Gambar 2.5 menunjukkan citra RGB beserta dengan nilai intensitas pikselnya.

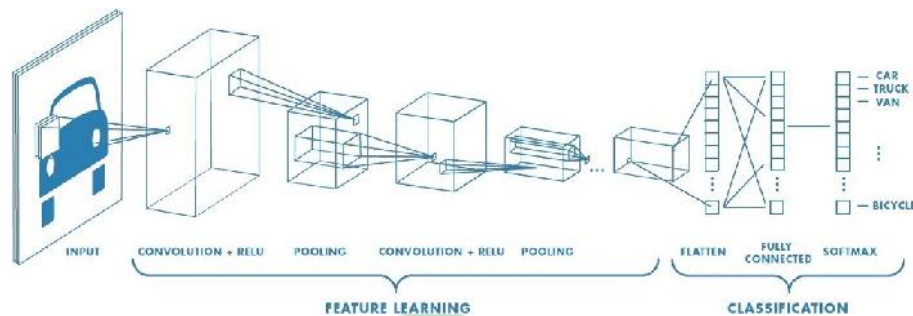


Gambar 2.5 Citra RGB beserta Matriks Pikselnya

2.5. Convolutional Neural Network (CNN)

Convolutional neural network (CNN) merupakan salah satu algoritma *Deep Learning* (DL) karena memiliki arsitektur jaringan yang mendalam, sehingga sangat memungkinkan untuk menemukan pola dalam gambar untuk mengenali objek, wajah maupun pemandangan *AI-Saffar2017*. CNN merupakan metode yang populer dari DL untuk pemrosesan data berupa citra maupun video (Masud et al., 2021). Seperti arsitektur jaringan saraf lainnya, CNN terdiri dari *input layer*,

output layer, dan banyak *hidden layer* di antaranya (Al-Saffar et al., 2017).



Gambar 2.6 Arsitektur *Convolutional Neural Network* (CNN)

Sumber: (MathWorks)

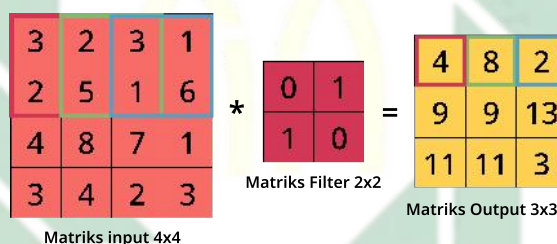
Pada Gambar 2.6, arsitektur CNN terdapat 3 bagian, yaitu input, *feature learning*, dan *classification*. Pada *feature learning* terdapat bagian medasar dari CNN yaitu *convolution layer* yang diikuti ReLU dan *pooling layer* (Alom et al., 2018). *Convolutional layer* memiliki bagian penting dari CNN yaitu filter, di mana fitur dalam citra input dapat diekstraksi (Sewak et al., 2018). Setiap filter memiliki lebar, tinggi serta jumlah tertentu pada setiap layer, sehingga menghasilkan ukuran *output* yang berbeda. ReLU memungkinkan pelatihan yang lebih cepat dan lebih efektif dengan memetakan nilai *output* negatif menjadi nol dan mempertahankan nilai positif (Ide and Kurita, 2017). *pooling layer* berguna untuk menyederhanakan ukuran *output* di setiap tumpukan *feature map* (Alwanda et al., 2020).

Sedangkan bagian *classification* merupakan hasil dari *feature map* yang berbentuk multi-dimensional, sehingga harus dilakukan proses flatten di mana matriks hasil dari *feature map* dirubah menjadi sebuah vektor agar dapat digunakan *input* dari *fully connected* (FC) *layer*. Dalam CNN, FC *layer* merupakan *Multi Layer Perceptron* (MLP), di mana terdapat *input layer*, *hidden layer*, fungsi aktivasi, *output layer* serta *loss function*. Sehingga FC *layer* menghasilkan vektor berdimensi sebanyak jumlah kelas yang dipelajari oleh jaringan. Selanjutnya FC *layer* dikenakan fungsi aktivasi softmax yang

menghasilkan nilai probabilitas pada setiap kelas (Hossain and Sajib, 2019).

2.5.1. Convolution Layer

Convolution merupakan proses operasi yang digunakan untuk ekstraksi fitur, di mana terdapat matriks kernel atau filter yang diterapkan menggunakan hadamard product ke seluruh matriks input (Yamashita et al., 2018). Filter bergeser sebanyak *stride* yang telah ditentukan pada piksel citra input. *Stride* merupakan jumlah pergeseran matriks filter terhadap matriks input baik secara horizontal dan vertikal. Ilustrasi proses konvolusi dengan *stride* sebanyak 1 dapat dilihat pada Gambar 2.7.



Gambar 2.7 Ilustrasi Proses Konvolusi

Hasil proses dari *convolutional layer* berupa *feature map* dari semua lapisan filter (Elsharif et al., 2019). Matriks output didapatkan dari hasil hadamard product antara matriks input dengan matriks filter, ditunjukkan pada persamaan 2.2

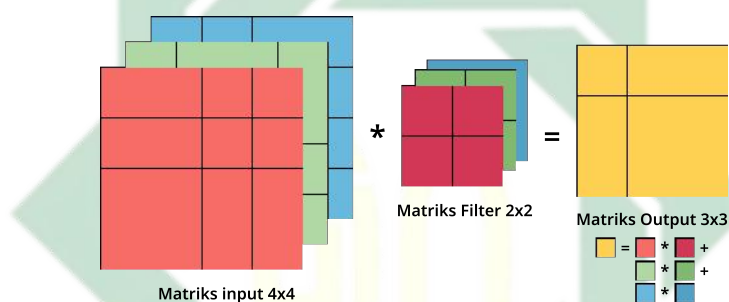
$$H(i, j) = I(i, j) * F(i, j) \quad (2.2)$$

Di mana $H(i, j)$ merupakan hasil keluaran dari convolutional layer pada baris ke- i dan kolom ke- j . Sedangkan $I(i, j)$ merupakan matriks input citra pada baris ke- i dan kolom ke- j , dan F merupakan matriks filter pada baris ke- i dan kolom ke- j . Besarnya ukuran matriks input, matriks filter, jumlah *stride* dan *padding* mempengaruhi ukuran dimensi matriks output dari feature map. Padding

merupakan parameter yang menambahkan nilai 0 pada setiap sisi matriks input. Rumus untuk mengetahui dimensi feature map dapat menggunakan persamaan 2.3.

$$h = \frac{(i - f + 2p)}{s} + 1 \quad (2.3)$$

Di mana h merupakan ukuran dimensi *feature map*, f merupakan ukuran matriks filter, p merupakan jumlah *padding*, dan s merupakan jumlah *stride*. Proses konvolusi pada citra RGB dapat dilihat pada Gambar 2.8.



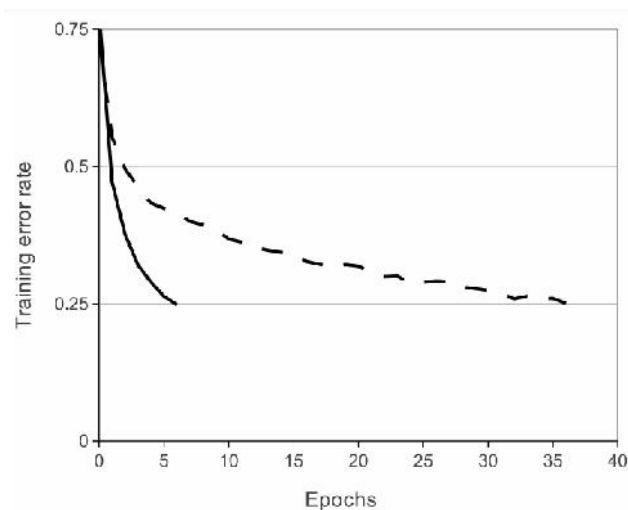
Gambar 2.8 Proses Konvolusi pada citra RGB

Di mana matriks input pada setiap *channel* warna akan diterapkan dengan filter menggunakan operasi hadamard product. Kemudian hasil pada setiap piksel di masing-masing *channel* warna dijumlahkan. Proses ini berulang dengan menerapkan beberapa filter sehingga membentuk sejumlah *feature map* yang mewakili karakteristik berbeda dari matriks input.

2.5.2. Rectified linear Unit(ReLU)

Terdapat beberapa fungsi aktivasi yang dapat digunakan dalam proses pembelajaran jaringan, seperti Tanh dan ReLU. Namun pada algoritma DL umumnya digunakan fungsi aktivasi ReLU (Ide and Kurita, 2017). Hal ini disebabkan karena ReLU mempelajari jaringan 6 kali lebih cepat dibandingkan Tanh (Glorot et al., 2011), perbandingan grafik tingkat kesalahan fungsi aktivasi

ReLU dengan Tanh dapat dilihat pada Gambar 2.9.

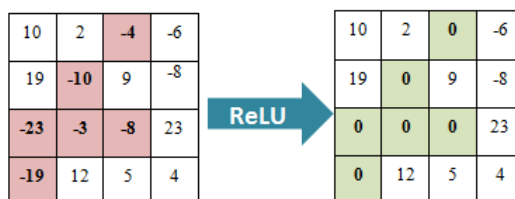


Gambar 2.9 Perbandingan Grafik Fungsi Aktivasi Tanh (*Dashed Line*) dan ReLU (*Solid Line*)
Sumber: (Krizhevsky et al., 2012)

Rumus fungsi aktivasi ReLU didefinisikan pada Persamaan 2.4, dengan x merupakan nilai input node.

$$F(x) = \max(0, x) \quad (2.4)$$

Sederhananya, fungsi aktivasi ReLU menghasilkan 0 ketika x bernilai negatif dan jika x bernilai lebih dari atau sama dengan 0 maka nilai keluarannya akan tetap seperti nilai inputannya (Agarap, 2018). Proses ReLU dilihat pada Gambar 2.10.



Gambar 2.10 Ilustrasi ReLU layer

2.5.3. Cross Channel Normalization Layer

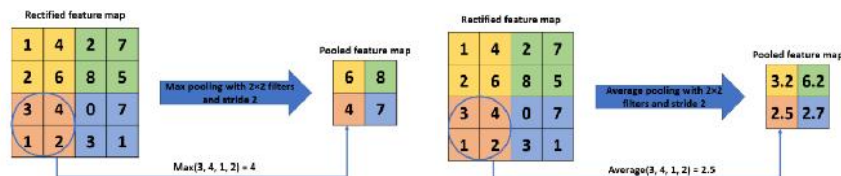
Cross Channel Normalization (CCN) layer pertama kali diterapkan pada arsitektur AlexNet dengan fungsi aktivasi yang digunakan adalah ReLU. CCN berperan penting untuk menormalisasikan nilai keluaran yang tidak terbatas dari fungsi aktivasi ReLU. Arsitektur yang mengimplementasikan CCN dalam jaringan terbukti efektif dapat mengurangi tingkat kesalahan sebesar 2% (Krizhevsky et al., 2012). CCN layer didefinisikan pada Persamaan 2.5.

$$b_{x,y}^i = \frac{a_{x,y}^i}{\left(k + \frac{\alpha \sum_{j=\max(1,1-\frac{n}{2})}^{\min(N-1,i+\frac{n}{2})} (a_{x,y}^j)^2}{n}\right)^\beta} \quad (2.5)$$

Di mana i merupakan output dari filter ke- i , $a(x, y)$ dan $b(x, y)$ merupakan nilai piksel pada (x, y) sebelum dan sesudah dinormalisasi secara berturut-turut, dan N merupakan jumlah total dari channel warna. Konstanta (k, α, β, n) merupakan parameter di mana k digunakan untuk menghindari singularitas atau pembagian dengan nol, α sebagai konstanta normalisasi dan β sebagai konstanta kontras. Sedangkan n digunakan untuk menentukan panjang piksel yang berdekatan yang perlu dipertimbangkan saat melakukan normalisasi.

2.5.4. Pooling Layer

Pooling layer berguna untuk mengurangi dimensi *feature map* (*downsampling*) sehingga parameter yang digunakan semakin sedikit namun tidak menghilangkan informasi penting didalamnya (Coskun et al., 2017). Terdapat beberapa metode yang populer dalam *pooling layer*, seperti *max pooling* dan *average pooling* (Jie and Wanda, 2020), namun yang umum digunakan adalah *max pooling* (Lee et al., 2017). Proses *pooling layer* ditunjukkan pada Gambar 2.11.



Gambar 2.11 Proses *Max Pooling* dan *Average Pooling*

Sumber: (Gholamalinezhad and Khosravi, 2020)

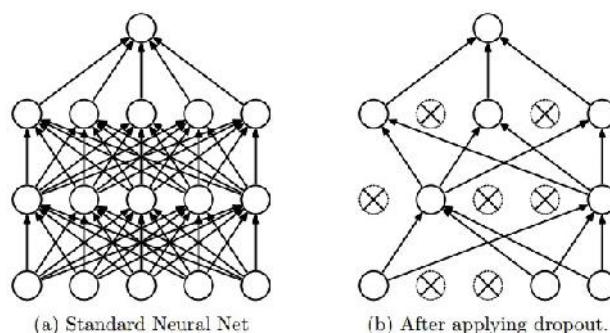
Max pooling diperoleh dari nilai piksel terbesar di setiap grid, sedangkan *average pooling* diperoleh dari rata-rata nilai piksel di setiap grid. Ukuran dimensi pada *pooling layer* sesuai dengan jumlah *stride* dan ukuran matriks filter yang digunakan. Rumus untuk mengetahui dimensi *pooling layer* dapat dilihat pada persamaan 2.6

$$O = \frac{(i - k)}{s} + 1 \quad (2.6)$$

Di mana O merupakan ukuran dimensi *pooling layer*, i merupakan ukuran dimensi matriks input citra, k merupakan ukuran dimensi matriks filter dan s merupakan jumlah *stride*. Ukuran *stride* yang umum digunakan pada *pooling layer* sebesar 2x2. Ukuran ini dapat mengurangi *feature map* hingga 75% dari ukuran aslinya (Ilahiyah and Nilogiri, 2018).

2.5.5. Dropout Layer

Dropout merupakan teknik regulasi untuk mengatasi overfitting pada model (Cai et al., 2019). Arsitektur yang mengimplementasikan *dropout* dalam jaringan terbukti efektif dapat mengurangi tingkat kesalahan mencapai hingga 3% (Srivastava et al., 2014).



Gambar 2.12 Proses *Dropout*

Sumber: (Srivastava et al., 2014)

Prinsip dari *dropout layer* adalah menghilangkan beberapa node dalam jaringan secara acak dengan jumlah probabilitas yang telah ditentukan, di mana probabilitas bernilai dari 0 sampai 1. Pada umumnya, probabilitas yang digunakan sebesar 0.5 (Srivastava et al., 2014). Hal ini dapat mengatasi masalah *overfitting*, di mana jaringan terlalu mengenali pola data input pada proses *training*, sehingga jika terdapat data baru model kesulitan dalam mengklasifikasikan data tersebut (Brownlee, 2018). Ilustrasi *dropout layer* dapat dilihat pada Gambar 2.12.

2.5.6. Fully Connected Layer

Setelah proses pada *convolutional layer* dan *pooling layer*, terdapat *FC layer* yang berguna untuk melakukan proses klasifikasi data (Ilahiyah and Nilogiri, 2018). Pada *FC layer* setiap node saling terhubung dengan node pada layer sebelumnya. Nilai *input* pada *FC layer* merupakan hasil *feature map* yang telah diubah dalam bentuk vektor dengan proses *flatten*. *FC layer* didefinisikan pada Persamaan 2.7.

$$u_j = f\left(\sum_{i=1}^n W_{i,j}x_i + b_j\right) \quad (2.7)$$

di mana x_i merupakan input, $W_{i,j}$ merupakan bobot, b_j merupakan bias dan f merupakan fungsi aktivasi (Dewa and Fadhilah, 2018). Nilai dari bobot dan bias

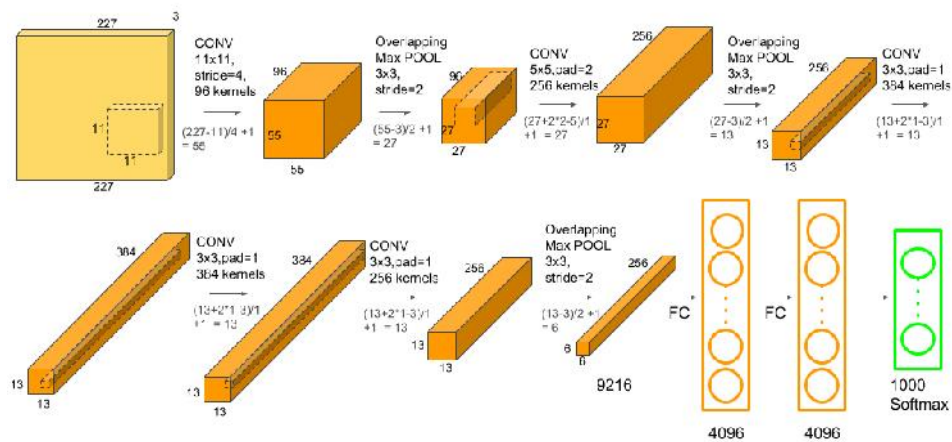
didapatkan dari perhitungan glorot *inicialitation*. FC layer diikuti fungsi aktivasi softmax di mana nilai vektor yang bernilai riil dirubah ke dalam rentang 0 sampai 1, jika semua hasil dijumlahkan bernilai satu (He et al., 2018). Fungsi aktivasi softmax merubah nilai menjadi probabilitas dengan menggunakan Persamaan 2.8

$$p_j(x) = \frac{e^{x_j}}{\sum_{i=1}^j e^{x_i}} \quad (2.8)$$

Di mana $p(x)$ merupakan nilai probabilitas di setiap kelas, dan x merupakan nilai dari vektor input. Data diklasifikasikan ke dalam kelas dengan nilai softmax yang mendekati satu atau paling tinggi (Bora et al., 2016).

2.6. AlexNet

Model AlexNet merupakan arsitektur pertama CNN yang populer di bidang *computer vision*. AlexNet memenangkan kompetisi ILSVRC pada tahun 2012 (Krizhevsky et al., 2012) dengan melatih dataset ImageNet yang memiliki 1000 kelas. Tingkat kesalahan yang dicapai AlexNet sebesar 15.3% sedangkan di posisi kedua diperoleh tingkat kesalahan sebesar 26.2% (Atole and Park, 2018). Arsitektur AlexNet memiliki 8 *layer* (Zulkeflie et al., 2019), dengan 5 *layer* merupakan *convolutional layer* serta 3 *layer* lainnya adalah FC *layer*. Filter yang digunakan AlexNet pada *convolutional* berukuran 11x11 pada *layer* pertama, 5x5 pada *layer* kedua, dan 3x3 pada sisa *layer* dengan ReLU sebagai fungsi aktivasi pada setiap *convolution layer*. Arsitektur AlexNet dapat dilihat pada Gambar 2.13.



Gambar 2.13 Arsitektur AlexNet

Sumber: (Nayak, 2018)

Terdapat perbedaan arsitektur AlexNet dari model CNN sebelumnya (Lecun et al., 1998) yaitu, pada AlexNet menggunakan fungsi aktivasi ReLU, CCN serta *overlapping pooling*. Fungsi aktivasi ReLU terbukti memiliki kinerja yang lebih baik dalam DL dibandingkan fungsi aktivasi Sigmoid dan Tanh (Dahl et al., 2013). Pada model AlexNet, CCN *layer* mengikuti lapisan *convolutional layer* pertama dan kedua. *Max pooling layer* mengikuti CCN *layer* serta *convolutional layer* kelima. ReLU diterapkan pada *output* dari setiap *convolutional layer* dan FC *layer*. Sehingga secara keseluruhan, model AlexNet terdiri dari 5 *convolutional layer* dan 3 FC *layer*. Node pada FC *layer* merupakan hasil dari *layer* sebelumnya. Output dari FC *layer* adalah softmax 1000 yang merupakan distribusi dari label kelas.

Pada model AlexNet terdapat *dropout layer* yang berguna untuk mengatasi masalah *overfitting* (Krizhevsky et al., 2012). Teknik *dropout* berfungsi untuk menghilangkan node pada proses *training* dengan probabilitas yang telah ditentukan. Setelah node terkena dropout maka node tersebut tidak menjadi input

untuk proses selanjutnya. Hal ini dapat mengatasi masalah *overfitting*, di mana jaringan terlalu mengenali pola data input pada proses *training*, sehingga jika terdapat data baru model kesulitan dalam mengklasifikasikan data tersebut (Brownlee, 2018). Model AlexNet memiliki 62 juta parameter (bobot dan bias), dimana pada setiap *convolution layer* dan *FC layer* memiliki jumlah parameter yang didapatkan dari perhitungan sebagai berikut: (Mallick and Nayak, 2018)

2.6.1. Menghitung Parameter Pada *Convolution Layer*

Pada metode CNN terdapat dua parameter yaitu bobot dan bias. Rumus perhitungan jumlah parameter dapat dilihat pada Persamaan 2.9.

$$\begin{aligned} W_c &= K^2 \times C \times N \\ B_c &= N \\ P_c &= W_c + B_c \end{aligned} \tag{2.9}$$

Di mana W_c merupakan jumlah bobot pada *convolution layer*, B_c merupakan jumlah bias pada *convolution layer*, P_c merupakan jumlah parameter pada *convolution layer*, K merupakan ukuran filter pada *convolution layer*, C merupakan jumlah *channel* pada input *layer* serta N merupakan jumlah filter. Berikut merupakan contoh perhitungan parameter yang terdapat pada *convolution layer* pertama, di mana filter memiliki ukuran 11×11 , jumlah *channel* pada input *layer* adalah 3 serta jumlah filter ada 96.

$$\begin{aligned} W_c &= K^2 \times C \times N \\ &= 11^2 \times 3 \times 96 \\ &= 34,848 \end{aligned}$$

$$B_c = 96$$

$$\begin{aligned} P_c &= 34,848 + 96 \\ &= 34,944 \end{aligned}$$

Sehingga jumlah parameter pada *convolution layer* ke-2, *convolution layer* ke-3, *convolution layer* ke-4, dan *convolution layer* ke-5 masing-masing adalah 614,656, 885,120, 1,327,488 dan 884,992.

2.6.2. Menghitung Parameter Pada *Fully Connected Layer*

Terdapat dua jenis layer yang terhubung pada *FC layer* di CNN. Lapisan *FC layer* pertama terhubung dari *convolution layer* terakhir, sedangkan *FC layer* selanjutnya terhubung dari *FC layer* lainnya. Berikut penjelasan perhitungan pada masing-masing layer:

Kasus 1: *FC Layer* yang terhubung dari *Convolution layer*

Rumus perhitungan jumlah parameter dapat dilihat pada Persamaan 2.10.

$$\begin{aligned} W_{cf} &= O^2 \times N \times F \\ B_{cf} &= F \\ P_{cf} &= W_{cf} + B_{cf} \end{aligned} \quad (2.10)$$

Di mana W_{cf} merupakan jumlah bobot pada *FC layer* yang terhubung dari *convolution layer*, B_{cf} merupakan jumlah bias pada *FC layer* yang terhubung dari *convolution layer*, P_{cf} merupakan jumlah parameter pada *FC layer* yang terhubung dari *convolution layer*, O merupakan ukuran output dari *convolution layer*, N merupakan jumlah filter pada *convolution layer* serta F merupakan jumlah node pada *FC layer*. Berikut merupakan contoh perhitungan parameter yang terdapat

pada FC *layer* pertama yang terhubung dengan *convolution layer* sebelumnya, di mana *output* dari *convolution layer* sebelumnya memiliki ukuran 6×6 , jumlah filter pada *input layer* adalah 256 serta jumlah node pada FC *layer* ada 4,096.

$$\begin{aligned} W_{cf} &= O^2 \times N \times F \\ &= 6^2 \times 256 \times 4,096 \\ &= 37,748,736 \end{aligned}$$

$$\begin{aligned} B_{cf} &= F \\ &= 4096 \end{aligned}$$

$$\begin{aligned} P_{cf} &= W_{cf} + B_{cf} \\ &= 37,748,736 + 4096 \end{aligned}$$

Kasus 2: FC Layer yang terhubung dari FC Layer

Rumus perhitungan jumlah parameter dapat dilihat pada Persamaan 2.11.

$$\begin{aligned} W_{ff} &= F_{-1} \times F \\ B_{ff} &= F \end{aligned} \tag{2.11}$$

$$P_{ff} = W_{ff} + B_{ff}$$

Di mana W_{ff} merupakan jumlah bobot pada FC *layer* yang terhubung dari FC *layer*, B_{ff} merupakan jumlah bias pada FC *layer* yang terhubung dari FC *layer*, P_{ff} merupakan jumlah parameter pada FC *layer* yang terhubung dari FC *layer*, F_{-1} merupakan jumlah node pada FC *layer* sebelumnya serta F merupakan jumlah node pada FC *layer*. Berikut merupakan contoh perhitungan parameter yang terdapat pada FC *layer* terakhir yang terhubung dengan FC *layer* sebelumnya, di mana jumlah node pada FC *layer* sebelumnya adalah 4096 serta jumlah node pada FC *layer* adalah 1,000.

$$W_{ff} = F_{-1} \times F$$

$$= 4,096 \times 1,000$$

$$= 4,096,000$$

$$B_{ff} = F$$

$$= 1,000$$

$$P_{ff} = W_{ff} + B_{ff}$$

$$= 4,096,000 + 1,000$$

Sehingga jumlah parameter pada FC layer ke-2 adalah 16,781,312. Total jumlah parameter yang ada pada model AlexNet disajikan pada Tabel 2.1.

Tabel 2.1 Jumlah Parameter Model AlexNet

Nama Layer	Bobot	Bias	Parameter
Convolution-1	34,848	96	34,944
Convolution-2	614,400	256	614,656
Convolution-3	884,736	384	885,120
Convolution-4	1,327,104	384	1,327,488
Convolution-5	884,736	256	884,992
FC-1	37,748,736	4,096	37,752,832
FC-2	16,777,216	4,096	16,781,312
FC-3	4,096,000	1,000	4,097,000
Total	62,367,776	10,568	62,378,344

2.7. Perbedaan Parameter dengan Hyperparameter

Pada algoritma *neural network* terdapat beberapa istilah yang memiliki arti yang tidak konsisten, hal ini dikarenakan terdapat istilah yang sama pada bidang

lain namun penggunaannya berbeda (Brownlee, 2017). Berikut merupakan penjelasan perbedaan antara parameter dengan *hyperparameter*. Parameter merupakan variabel konfigurasi dari dalam model dan nilainya dapat diperkirakan dari data, pada kasus *neural network* parameter adalah bobot dan bias (Dudek, 2019).

hyperparameter digunakan untuk konfigurasi berbagai aspek dalam suatu algoritma pembelajaran di luar model dan nilainya tidak dapat diperkirakan dari data. Nilai hyperparameter terbaik dapat diperoleh dengan melihat dari permasalahan sebelumnya, menggunakan *rule-of-thumbs* atau dengan cara *trial* dan *error* (Claesen and Moor, 2015). Pada *neural network* terdapat beberapa *hyperparameter* seperti, *learning rate*, *batch size*, dan jumlah *epoch* (Aszemi and Dominic, 2019). *Learning rate* merupakan *hyperparameter* yang mengatur seberapa banyak perubahan model berdasarkan dari nilai *error* yang diperkirakan setiap kali pemberian bobot, *batch size* adalah jumlah data training yang digunakan pada jaringan dalam satu iterasi sebelum bobot diperbarui sedangkan jumlah *epoch* menentukan berapa kali algoritma pembelajaran melalui seluruh dataset pelatihan (Aszemi and Dominic, 2019).

2.8. Glorot Inizialitation

Glorot *inizialitation* atau biasa disebut Xavier merupakan inisialisasi nilai awal parameter untuk model, yaitu nilai bobot dan bias(Mathworks). Rumus inisialisasi glorot didefinisikan pada Persamaan 2.12.

$$parameter = \left[-\sqrt{\frac{6}{N_o + N_i}}, \sqrt{\frac{6}{N_o + N_i}} \right] \quad (2.12)$$

Dimana nilai dari N_o dan N_i tergantung pada jenis layer yang digunakan. Pada *convolution layer* N_o merupakan $prod(ukuranfilter) \times jumlahfilter$ lalu N_i merupakan $prod(ukuranfilter) \times jumlahchannel$. Sedangkan pada *FC layer*

N_o merupakan jumlah node *output* lalu N_i merupakan jumlah node input.

2.9. Confusion Matrix

Proses evaluasi pada sistem klasifikasi umumnya menggunakan *confusion matrix* untuk membandingkan hasil klasifikasi dengan nilai sebenarnya. Pada *confusion matrix* terdapat 4 istilah untuk merepresentasikan hasil klasifikasi yaitu, *True Positive* (TP) merupakan jumlah data terklasifikasi benar dari kelas aktual positif, *True Negative* (TN) merupakan jumlah data terklasifikasi benar dari kelas aktual negatif, *False Positive* (FP) merupakan jumlah data terklasifikasi salah dari kelas aktual negatif dan *False Negative* (FN) merupakan jumlah data terklasifikasi salah dari kelas aktual positif. Ilustrasi *confusion matrix multi class* dapat dilihat pada 2.14, di mana kelas *benign lung tissue*, *lung adenocarcinoma* dan *lung squamous cell carcinoma* dimisalkan kelas A, B dan C secara berturut-turut.

Target Class \ Output Class	A	B	C
A	TP	FN	FN
B	FP	TN	TN
C	FP	TN	TN

(a)

Target Class \ Output Class	A	B	C
A	TN	FP	TN
B	FN	TP	FN
C	TN	FP	TN

(b)

Target Class \ Output Class	A	B	C
A	TN	TN	FP
B	TN	TN	FP
C	FN	FN	TP

(c)

Gambar 2.14 Confusion Matrix Multi Class (a) kelas A (b) kelas B (c) kelas C

Dari Gambar 2.14, didapatkan beberapa perhitungan untuk menentukan tingkat keakuratan sistem klasifikasi yang dibangun, seperti akurasi, sensitivitas dan spesifitas. Akurasi merupakan nilai yang merepresentasikan banyak data yang terklasifikasi sesuai dengan data aktual, sensitivitas merupakan nilai yang merepresentasikan banyak data yang terklasifikasi benar dalam kelas positif, sedangkan spesifitas merupakan nilai yang merepresentasikan banyak data yang terklasifikasi benar dalam kelas negatif. Perhitungan confusion matrix dapat dilihat

pada Persamaan 2.13, 2.14 dan 2.15.

$$Akurasi = \frac{TP_{all}}{n(all)} \times 100\% \quad (2.13)$$

$$Sensitivitas = \frac{\sum \frac{TP}{TP+FN}}{n} \times 100\% \quad (2.14)$$

$$Spesitivitas = \frac{\sum \frac{TN}{TN+FP}}{n} \times 100\% \quad (2.15)$$

Misalkan didapatkan *confusion matrix multi class* seperti pada Gambar 2.15 dengan data sebanyak 5000 pada setiap kelas. Sehingga didapatkan nilai akurasi, sensitivitas serta spesitivitas sebagai berikut:

Target Class	A	5000	200	200
	B	100	4500	200
	C	100	200	4500
		A	B	C

Output Class

Gambar 2.15 Contoh Perhitungan Confusion Matrix Multi Class

$$\begin{aligned}
 Akurasi &= \frac{TP_{all}}{n(all)} \times 100\% \\
 &= \frac{5000 + 4500 + 4500}{15000} \times 100\% \\
 &= 93.3\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Sensitivitas} &= \frac{\sum \frac{TP}{TP+FN}}{n} \times 100\% \\
 &= \frac{\frac{5000}{5000+400} + \frac{4500}{4500+300} + \frac{4500}{4500+300}}{3} \times 100\% \\
 &= 93.4\%
 \end{aligned}$$

$$\begin{aligned}
 \text{Spesitivitas} &= \frac{\sum \frac{TN}{TN+FP}}{n} \times 100\% \\
 &= \frac{\frac{9400}{9400+200} + \frac{9800}{9800+400} + \frac{9800}{9800+400}}{3} \times 100\% \\
 &= 96.7\%
 \end{aligned}$$



UIN SUNAN AMPEL
S U R A B A Y A

BAB III

METODE PENELITIAN

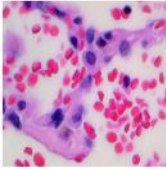
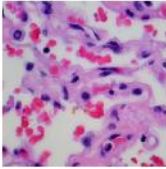
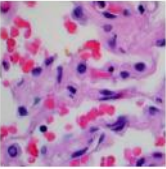
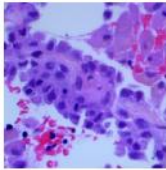
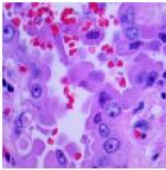
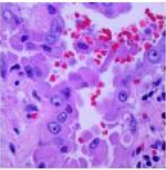
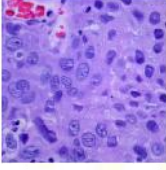
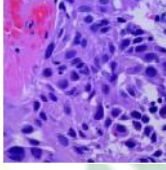
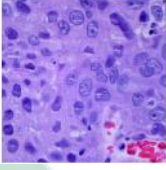
3.1. Jenis Penelitian

Pada penelitian ini menggunakan penelitian jenis kuantitatif. Data yang digunakan merupakan data kanker paru berupa citra histopatologi, yang kemudian ditransformasikan ke dalam numerik bentuk matrix. Hasil dari penelitian ini berupa sistem klasifikasi kanker paru menggunakan metode CNN model AlexNet yang diharapkan dapat membantu ahli medis dalam mendiagnosis penyakit kanker paru secara cepat dan tepat.

3.2. Jenis dan Sumber Data

Pada penelitian ini digunakan data sekunder citra histopatologi kanker paru LC25000 dataset, yang didapatkan dari Web Cornell University (Borkowski et al., 2019). Pada dataset LC25000, kanker paru dikelompokkan menjadi tiga kelas, yaitu lung adenocarcinoma, lung squamous cell carcinoma dan benign lung tissue. Masing-masing kelas memiliki data sebanyak 5000 data citra berukuran 768x768 piksel. Sampel data kanker paru pada setiap kelas ditunjukkan pada Tabel 3.1.

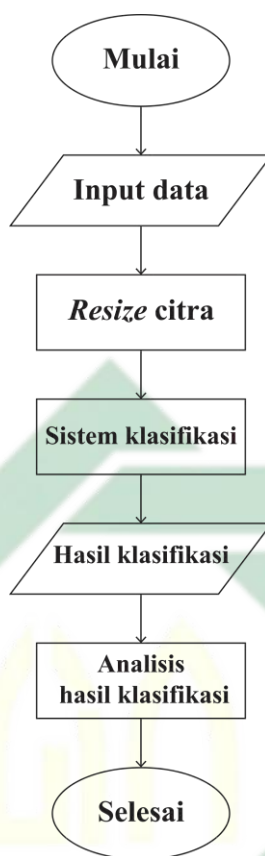
Tabel 3.1 Sampel Data Citra Histopatologi Kanker Paru

Kelas	Sampel Citra		
<i>Benign Lung Tissue</i>			
<i>Lung Adenocarcinoma</i>			
<i>Lung Squamous Cell Carcinoma</i>			

3.3. Kerangka Penelitian

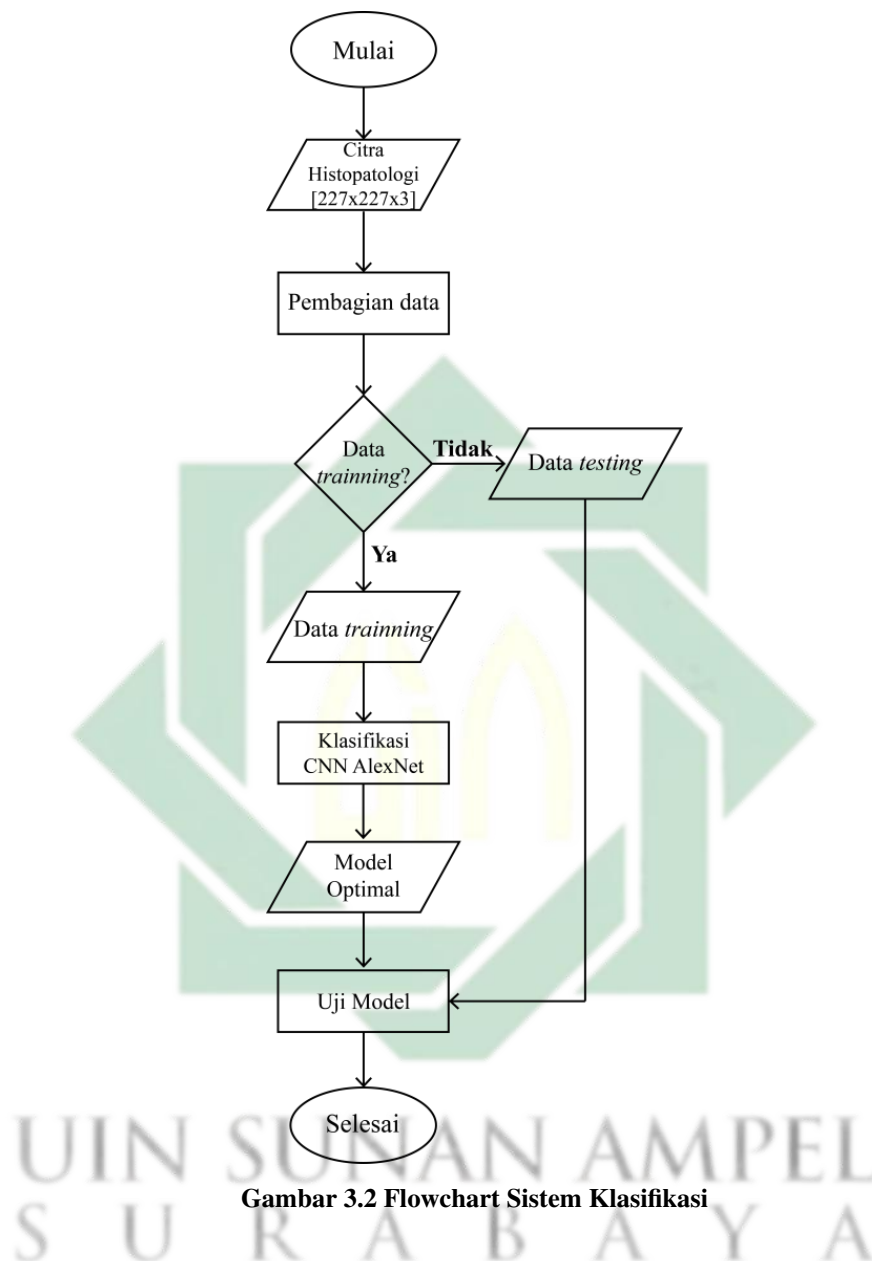
Pada penelitian ini terdapat beberapa tahap yang dilakukan dalam sistem klasifikasi kanker paru menggunakan CNN model AlexNet. Flowchart pada penelitian ini ditunjukkan pada Gambar 3.1

UIN SUNAN AMPEL
S U R A B A Y A



Gambar 3.1 Flowchart Penelitian

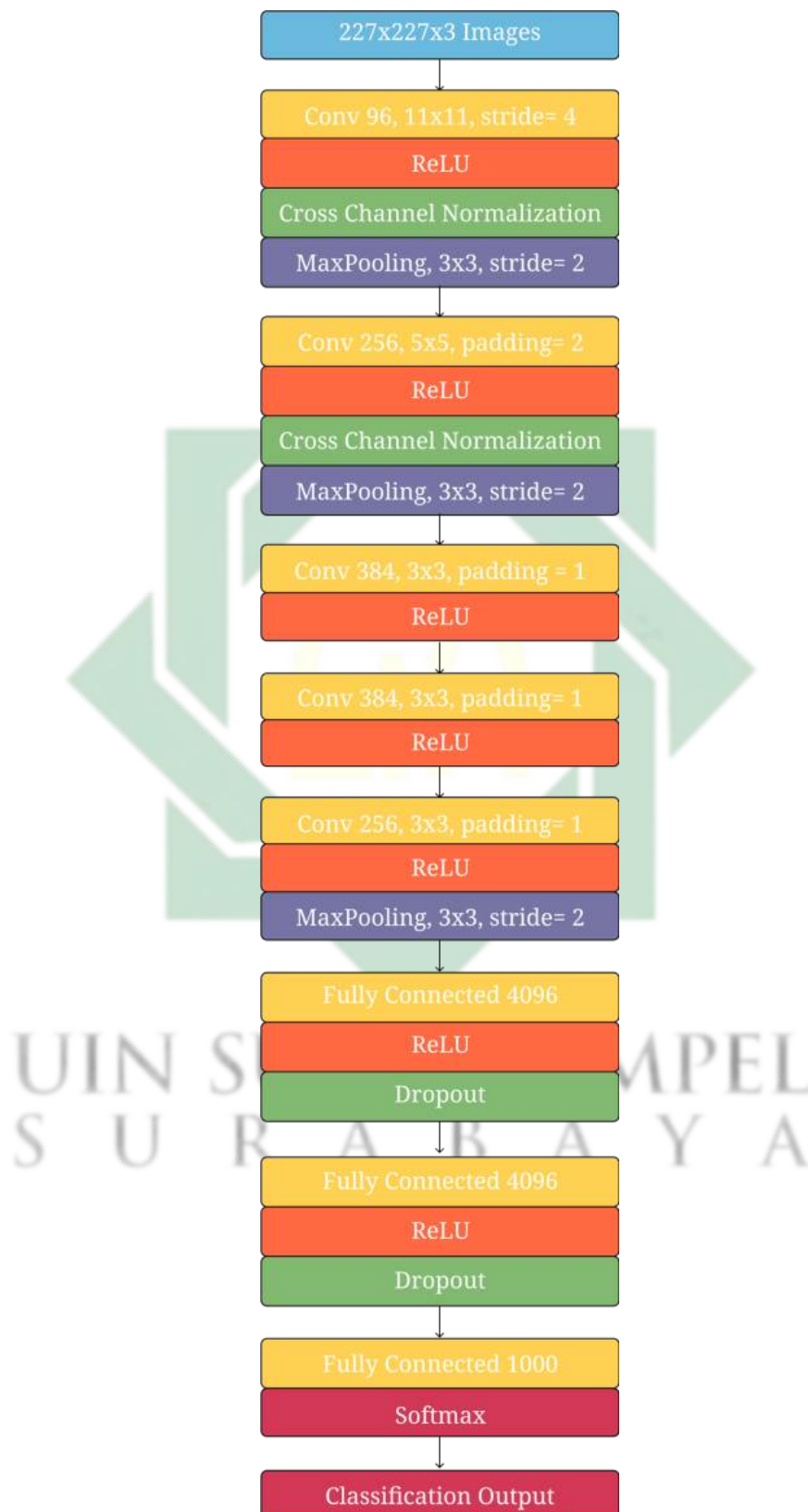
1. **Input Data** Data input berupa citra histopatologi kanker paru yang diperoleh dari Cornell University bernama LC25000 dataset. Pada dataset LC25000 terdapat tiga kelas kanker paru, yaitu *benign lung tissue*, *lung adenocarcinoma* dan *lung squamous cell carcinoma*.
2. **Resize citra** Proses resize dilakukan karena pada dataset citra histopatologi kanker paru berukuran 768x768 piksel, sedangkan pada sistem klasifikasi menggunakan CNN model AlexNet memiliki ketentuan di mana input citra berukuran 227x227 piksel.
3. **Sistem klasifikasi** Pada penelitian ini, sistem digunakan metode CNN model AlexNet untuk pembuatan sistem klasifikasi. Flowchart pada pembuatan sistem klasifikasi dapat dilihat pada Gambar 3.2.



Gambar 3.2 Flowchart Sistem Klasifikasi

Pada proses klasifikasi, data dibagi menjadi data training dan data testing.

Pada penelitian ini digunakan metode CNN model AlexNet proses training, arsitektur AlexNet dapat dilihat pada Gambar 3.3



Gambar 3.3 Arsitektur AlexNet Sumber: (Krizhevsky et al., 2012)

Proses training dilakukan dengan beberapa uji coba, pada penelitian ini uji coba yang dilakukan adalah jumlah pembagian data, jumlah batchsize dan probabilitas dropout, penjabaran lebih lanjut dapat dilihat pada Tabel 3.2. Uji coba dilakukan untuk mendapat model yang paling optimal, yang selanjutnya diterapkan pada data testing.

Tabel 3.2 Uji Coba pada Sistem Klasifikasi

Pembagian data (<i>training</i> : <i>testing</i>)	60 : 40	70 : 30	80 : 20	90 : 10
<i>Batchsize</i>	8	16	32	64
Probabilitas <i>Dropout</i>	0.5	0.6	0.7	0.8

- Analisa hasil klasifikasi Setelah didapatkan hasil klasifikasi, model dievaluasi menggunakan confusion matrix. Berdasarkan nilai akurasi, sensitivitas, serta spesifisitas, dapat diketahui apakah metode CNN model AlexNet dapat diimplementasikan pada sistem klasifikasi kanker paru berdasarkan citra histopatologi. Kemudian berdasarkan uji coba yang telah dilakukan didapatkan model yang paling optimal untuk sistem klasifikasi kanker paru berdasarkan citra histopatologi.

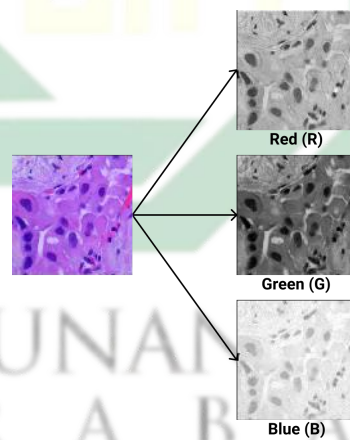
UIN SUNAN AMPEL
S U R A B A Y A

BAB IV

HASIL DAN PEMBAHASAN

4.1. *Preprocessing*

Pada penelitian ini, dilakukan klasifikasi kanker paru berdasarkan citra histopatologi menggunakan CNN model AlexNet. Citra histopatologi kanker paru didapatkan dari web Cornell University yang bernama LC25000 dataset. Data citra histopatologi kanker paru ini merupakan citra RGB, dimana dalam satu data citra memiliki tiga *channel* warna yaitu merah (R), hijau (G), dan biru (B). Ilustrasi citra RGB dapat dilihat pada Gambar 4.1



Gambar 4.1 Ilustrasi Citra RGB

Pada masing-masing *channel* warna tersebut memiliki nilai piksel di antara rentang 0 sampai 255. Nilai piksel pada masing-masing *channel* warna ditunjukkan pada Gambar 4.2, 4.3, 4.4

x/y	1	2	3	4	5	...	764	765	766	767	768
1	211	211	211	212	210	...	149	151	152	154	156
2	211	211	212	212	210	...	151	153	154	155	156
3	211	212	212	212	211	...	154	155	155	157	157
4	212	212	212	213	211	...	154	155	155	156	155
5	212	212	213	213	211	...	154	155	155	154	153
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
764	177	178	178	178	178	...	166	169	172	174	175
765	174	174	174	176	175	...	164	168	172	175	178
766	170	171	172	172	173	...	162	166	171	176	179
767	166	167	168	169	170	...	161	165	171	177	181
768	164	164	166	166	168	...	161	165	172	178	182

Gambar 4.2 Nilai Pixel Citra pada *Channel Red*

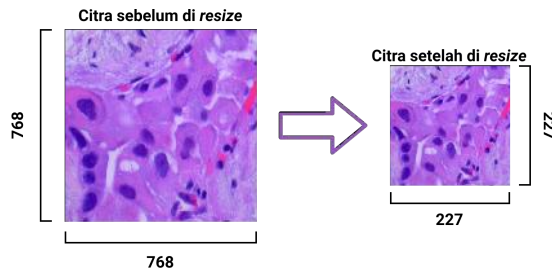
x/y	1	2	3	4	5	...	764	765	766	767	768
1	195	195	195	196	197	...	82	83	84	86	88
2	195	195	196	196	197	...	84	85	86	87	88
3	195	196	196	196	198	...	86	87	87	86	86
4	196	196	196	197	198	...	86	87	87	85	84
5	196	196	197	197	198	...	83	84	84	83	82
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
764	105	106	106	104	101	...	132	135	138	140	142
765	104	104	104	104	101	...	131	135	139	142	145
766	102	103	102	102	101	...	129	133	138	143	146
767	99	100	100	99	97	...	128	132	138	144	148
768	97	97	98	98	95	...	128	132	139	145	149

Gambar 4.3 Nilai Pixel Citra pada *Channel Green*

x/y	1	2	3	4	5	...	764	765	766	767	768
1	240	240	240	241	241	...	221	222	223	225	227
2	240	240	241	241	241	...	223	224	225	226	227
3	240	241	241	241	242	...	225	226	226	228	228
4	241	241	241	242	242	...	225	226	226	227	226
5	241	241	242	242	242	...	223	226	226	225	224
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
764	217	218	218	217	215	...	226	229	232	234	233
765	218	218	218	216	214	...	222	226	230	233	234
766	215	216	216	216	213	...	220	224	229	232	235
767	214	215	213	213	212	...	217	221	227	233	237
768	212	212	211	211	210	...	217	221	228	234	238

Gambar 4.4 Nilai Pixel Citra pada *Channel Blue*

Pada dataset LC25000, kanker paru dikelompokkan menjadi tiga kelas, yaitu *benign lung tissue*, *lung adenocarcinoma*, dan *lung squamous cell carcinoma*. Masing-masing kelas memiliki data sebanyak 5000 data citra yang berukuran $768 \times 768 \times 3$ piksel. Pada model AlexNet memiliki ketentuan input data citra berukuran $227 \times 227 \times 3$ piksel. Sehingga diperlukan proses *resize* pada citra awal berukuran $768 \times 768 \times 3$ piksel menjadi citra yang berukuran $227 \times 227 \times 3$ piksel. Pada penelitian ini digunakan metode *Bicubic interpolation* untuk proses *resize*. Ilustrasi proses *resize* citra ditunjukkan pada Gambar 4.5.



Gambar 4.5 Ilustrasi Proses *Resize* Citra

Nilai piksel citra yang telah di *resize* pada masing-masing *channel* warna ditunjukkan pada Gambar 4.6, 4.7, 4.8

x/y	1	2	3	4	5	...	223	224	225	226	227
1	211	211	210	210	209	...	163	151	150	151	155
2	213	212	212	211	208	...	164	156	151	152	154
3	213	213	213	211	209	...	169	160	153	154	156
4	213	213	213	212	208	...	179	171	164	165	164
5	213	211	210	209	210	...	190	182	179	178	173
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
223	185	182	179	180	167	...	164	173	168	171	178
224	186	184	181	181	176	...	152	162	163	168	178
225	187	184	180	181	185	...	159	156	151	157	163
226	180	179	176	177	180	...	169	161	156	162	171
227	168	169	167	171	175	...	167	171	161	161	176

Gambar 4.6 Nilai Piksel Citra pada *Channel Red*

x/y	1	2	3	4	5	...	223	224	225	226	227
1	195	197	198	201	202	...	97	89	86	84	86
2	197	199	201	202	201	...	98	91	83	81	83
3	197	199	201	202	201	...	107	97	88	86	88
4	197	199	201	203	199	...	123	118	108	106	103
5	198	198	200	200	200	...	144	139	135	131	124
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
223	120	115	114	122	121	...	138	145	140	141	147
224	114	110	108	115	125	...	126	133	133	136	145
225	111	105	101	107	126	...	132	126	120	124	129
226	107	101	97	102	117	...	142	131	124	129	137
227	100	96	91	98	110	...	141	142	131	128	143

Gambar 4.7 Nilai Piksel Citra pada *Channel Green*

x/y	1	2	3	4	5	...	223	224	225	226	227
1	240	242	242	243	246	...	227	221	220	222	226
2	242	243	244	245	245	...	229	226	220	221	225
3	242	243	245	244	245	...	233	227	221	221	225
4	242	243	245	245	242	...	241	236	229	230	229
5	245	245	246	245	245	...	249	241	239	240	235
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
223	216	216	218	226	224	...	225	229	221	223	230
224	215	214	214	220	226	...	217	224	223	224	232
225	220	215	211	215	229	...	226	223	215	218	222
226	219	213	208	210	221	...	236	228	219	223	230
227	214	210	202	206	215	...	232	234	223	218	232

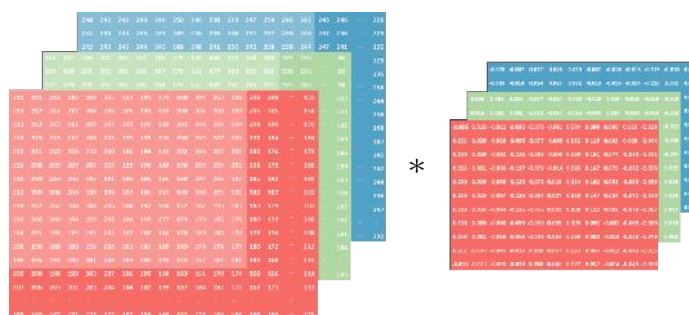
Gambar 4.8 Nilai Piksel Citra pada *Channel Blue*

4.2. Pembentukan Sistem Klasifikasi

Pada penelitian ini, digunakan metode CNN model AlexNet untuk pembentukan sistem klasifikasi. Secara garis besar, proses pada CNN terbagi menjadi dua bagian yaitu, *feature learning* dan *classification*. Pada *feature learning* berguna untuk mempelajari pola yang ada pada data citra menggunakan filter-filter yang terdapat pada *convolution layer*. Sedangkan pada *classification*, terdapat proses *flatten* dimana matiks hasil *feature learning* di ubah menjadi sebuah vektor, kemudian di proses dalam *fully connected layer* yang akan menghasilkan vektor berdimensi sebanyak jumlah kelas yang ada. Berikut merupakan penjelasan *layer-layer* pada model AlexNet.

4.2.1. Convolution Layer

Tahap pertama dalam pembentukan sistem klasifikasi adalah menginputkan data citra RGB histopatologi kanker paru yang telah di *resize* dengan ukuran $227 \times 227 \times 3$ piksel. Selanjutnya pada bagian *feature learning* terdapat *convolution layer*. Pada model AlexNet *channel* pertama merupakan *convolution layer* yang memiliki filter sebanyak 96 dengan ukuran $11 \times 11 \times 3$ piksel, *padding* sebesar 0 dan *stride* sebanyak 4. Berikut merupakan ilustrasi proses *convolution layer* yang ditunjukkan pada Gambar 4.9.



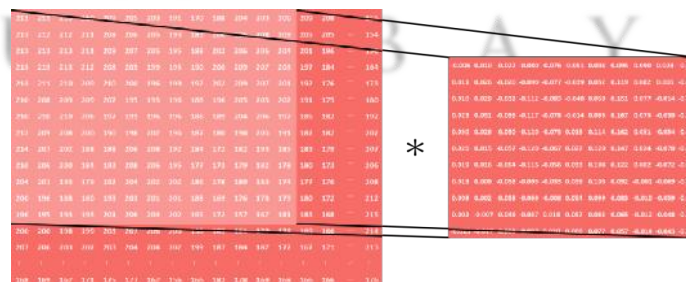
Gambar 4.9 Ilustrasi Perhitungan pada Convolution Layer

Pada Gambar 4.9, dapat dilihat bahwa matriks citra yang terdiri dari 3

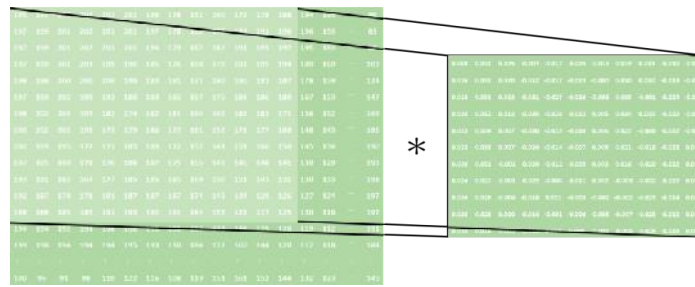
channel kemudian dioperasikan dengan matriks filter menggunakan hadamard product. Matriks filter bergeser sebanyak *stride* yang telah ditentukan pada piksel citra input, secara horizontal dan vertical. Pada model AlexNet terdapat *convolution layer* dengan filter berukuran $11 \times 11 \times 3$ piksel serta *padding* sebesar 0 dan *stride* sebanyak 4. Ukuran hasil *feature map* pada *convolution layer* dapat dihitung menggunakan Persamaan 2.3.

$$\begin{aligned}
 h &= \frac{(i - k + 2p)}{s} + 1 \\
 &= \frac{(227 - 11 + 0)}{4} + 1 \\
 &= \frac{(216)}{4} + 1 \\
 &= 54 + 1 \\
 &= 55
 \end{aligned}$$

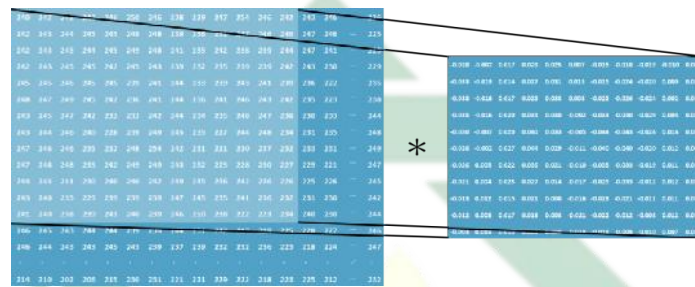
Berdasarkan perhitungan diatas, ukuran *feature map* yang dihasilkan dari *convolution layer* dengan filter berukuran $11 \times 11 \times 3$ piksel serta *padding* sebesar 0 dan *stride* sebanyak 4 adalah 55×55 . Perhitungan konvolusi pada setiap *channel* dapat menggunakan persamaan 2.2. Berikut ilustrasi perhitungan konvolusi di setiap *channel* yang ditunjukkan pada Gambar 4.10, 4.11, 4.12.



Gambar 4.10 Ilustrasi Perhitungan *Convolution Layer channel Red*



Gambar 4.11 Ilustrasi Perhitungan *Convolution Layer* channel Green



Gambar 4.12 Ilustrasi Perhitungan *Convolution Layer* channel Blue

Perhitungan konvolusi menggunakan operasi perkalian hadamard product, dimana setiap elemen matriks input dikalikan dengan setiap elemen matriks filter kemudian hasil perkalian tersebut dijumlahkan semuanya. Hasil perhitungan dari konvolusi di setiap *channel* kemudian dijumlahkan sehingga menghasilkan *feature map* sesuai dengan koordinat pikselnya. Berikut merupakan perhitungan konvolusi di setiap *channel*.

UIN SUNAN AMPEL
S U R A B A Y A

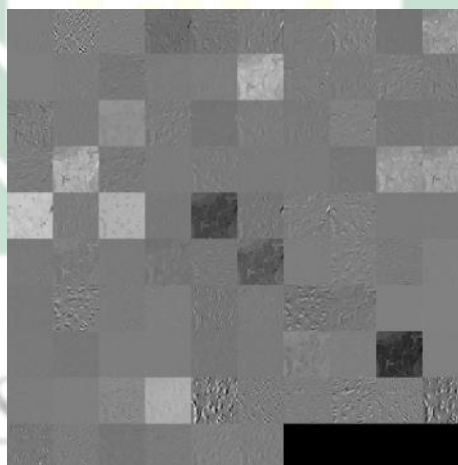
$$\begin{aligned}
H_{(1,1)} &= (I_r(1,1) \times F_r(1,1) + I_r(1,2) \times F_r(1,2) + I_r(1,3) \times F_r(1,3) + \dots + \\
&\quad I_r(11,11) \times F_r(11,11)) + (I_g(1,1) \times F_g(1,1) + I_g(1,2) \times F_g(1,2) + \\
&\quad I_g(1,3) \times F_g(1,3) + \dots + I_g(11,11) \times F_g(11,11)) + (I_b(1,1) \times F_b(1,1) + \\
&\quad I_b(1,2) \times F_b(1,2) + I_b(1,3) \times F_b(1,3) + \dots + I_b(11,11) \times F_b(11,11)) \\
&= (211 \times (-0.006) + 211 \times 0.010 + 210 \times (-0.022)) + \dots + 181 \times (-0.038)) \\
&\quad + (195 \times 0.008 + 197 \times 0.031 + 198 \times 0.025 + \dots + 143 \times 0.002) + \\
&\quad (240 \times (-0.028)) + 242 \times (-0.007) + 242 \times 0.0017) + \dots + 232 \\
&\quad \times 0.021)
\end{aligned}$$

$$\begin{aligned}
H_{(1,2)} &= (I_r(1,5) \times F_r(1,1) + I_r(1,6) \times F_r(1,2) + I_r(1,7) \times F_r(1,3) + \dots + \\
&\quad I_r(11,15) \times F_r(11,11)) + (I_g(1,5) \times F_g(1,1) + I_g(1,6) \times F_g(1,2) + \\
&\quad I_g(1,7) \times F_g(1,3) + \dots + I_g(11,15) \times F_g(11,11)) + (I_b(1,5) \times F_b(1,1) + \\
&\quad I_b(1,6) \times F_b(1,2) + I_b(1,7) \times F_b(1,3) + \dots + I_b(11,15) \times F_b(11,11)) \\
&= (209 \times (-0.006) + 205 \times 0.010 + 203 \times (-0.022)) + \dots + 176 \times \\
&\quad (-0.038)) + (202 \times 0.008 + 196 \times 0.031 + 178 \times 0.025 + \dots + \\
&\quad 118 \times 0.002) + (246 \times (-0.028)) + 250 \times (-0.007) + 246 \times 0.0017) \\
&\quad + \dots + 224 \times 0.021)
\end{aligned}$$

Berdasarkan perhitungan diatas, berikut merupakan hasil *feature map* dengan ukuran 55x55.

$$H_{(:, :, 1)} = \begin{bmatrix} -35.43 & 12.46 & -59.34 & -25.44 & -10.59 & -17.92 & \dots & -42.71 \\ 2.09 & -8.92 & -30.89 & -33.75 & -5.64 & -33.12 & \dots & -22.811 \\ 6.36 & -37.69 & -15.17 & -8.38 & -20.34 & -114.01 & \dots & -11.74 \\ -9.16 & -35.58 & -6.63 & -31.08 & 0.80 & -163.12 & \dots & -17.36 \\ -13.68 & -9.58 & -22.07 & -36.24 & -10.02 & -67.07 & \dots & -7.86 \\ \dots & \dots & \dots & \dots & \dots & \dots & \ddots & \dots \\ -41.12 & -9.00 & -20.57 & -0.93 & -27.73 & -8.00 & \dots & -12.47 \end{bmatrix}$$

Perhitungan diatas berulang sesuai dengan jumlah filter, pada contoh ini filter sebanyak 96. Visualisasi hasil feature map di convolution layer ditunjukkan pada Gambar 4.13.



Gambar 4.13 Visualisasi *Feature Map* pada *Convolution Layer*

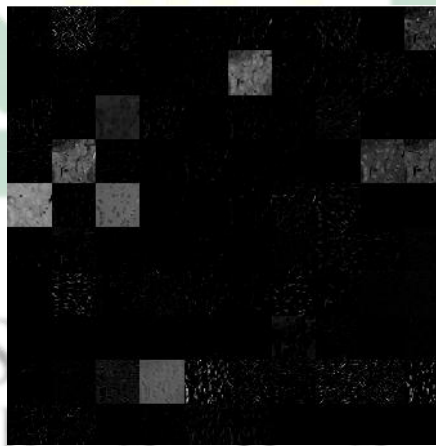
4.2.2. *ReLU Layer*

Hasil *feature map* pada *Convolution Layer* selanjutnya dikenakan fungsi aktivasi ReLU yang berguna untuk menggantikan nilai negatif menjadi nol. Fungsi aktivasi dapat dihitung menggunakan Persamaan 2.4. Nilai input yang bernilai negatif pada *feature map* dapat menyebabkan kecenderungan nilai gradien mendekati nol, sehingga model tidak dapat mempelajari pola dikarenakan nilai

gradien yang semakin kecil, situasi ini disebut dengan *vanishing gradien*. Berikut merupakan perhitungan pada ReLU Layer.

$$f(H_{(:, :, 1)}) = \begin{bmatrix} 0 & 12.46 & 0 & 0 & 0 & 0 & \dots & 0 \\ 2.09 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 6.36 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Perhitungan diatas berulang sesuai dengan jumlah *feature map*. Visualisasi hasil *feature map* pada *convolution layer* ditunjukkan pada Gambar 4.14



Gambar 4.14 Visualisasi *Feature Map* pada ReLU Layer

4.2.3. Cross Channel Normalization Layer

Hasil *feature map* yang telah diaktivasi pada ReLU Layer, selanjutnya di normalisasi pada cross channel normalization yang berguna untuk membatasi nilai pada *feature map*. Hal ini bertujuan untuk meningkatkan kemampuan generalisasi model. Pada CCN memiliki beberapa parameter, antara lain (k, α, β, n) . Perhitungan pada CCN dapat dihitung menggunakan Persamaan 2.5. Berikut

perhitungan pada CCN *layer* dengan inisialisasi hyperparameter (1, 0.0001, 0.75, 5).

$$f(H_{(:, :, 1)}) = \begin{bmatrix} 0 & 12.46 & 0 & 0 & 0 & 0 & \dots & 0 \\ 2.09 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 6.36 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

$$f(H_{(:, :, 2)}) = \begin{bmatrix} 0 & 128.50 & 40.59 & 0 & 0 & 28.82 & \dots & 0 \\ 0 & 215.85 & 0 & 0 & 0 & 460.68 & \dots & 0 \\ 24.70 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

UIN SUNAN AMPEL
S U R A B A Y A

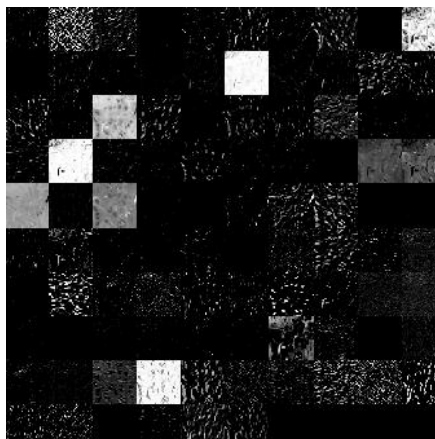
$$f(H_{(:, :, 96)}) = \begin{bmatrix} 7.99 & 0 & 30.25 & 0 & 0 & 7.56 & \dots & 0 \\ 5.92 & 42.83 & 23.89 & 0 & 0 & 56.78 & \dots & 0 \\ 11.38 & 50.29 & 10.54 & 0 & 0 & 77.56 & \dots & 0 \\ 10.68 & 15.78 & 16.62 & 0 & 0 & 61.24 & \dots & 0 \\ 0 & 0 & 0 & 8.13 & 4.80 & 0 & \dots & 15.36 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 51.84 & 66.30 & 16.95 & 0 & \dots & 0 \end{bmatrix}$$

$$\begin{aligned} b_{1,2}^i &= \frac{a_{x,y}^1}{\left(1 + \frac{0.0001 \sum_{j=\max(1, a_{1,2}^j)}^3 (a_{1,2}^j)^2}{5}\right)^{0.75}} \\ &= \frac{12.46}{\left(1 + \frac{0.0001 \times (12.46^2 + 128.50^2 + 0^2)}{5}\right)^{0.75}} \\ &= 10.04 \end{aligned}$$

Perhitungan diatas berulang sesuai dengan jumlah elemen dalam *feature map*.

$$b_{(:, :, 1)} = \begin{bmatrix} 0 & 10.04 & 0 & 0 & 0 & 0 & \dots & 0 \\ 2.09 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ 11.38 & 6.30 & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0.78 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{bmatrix}$$

Visualisasi hasil *feature map* pada *cross channel normalization layer* ditunjukkan pada Gambar 4.15



Gambar 4.15 Visualisasi Feature Map pada CCN Layer

4.2.4. Max Pooling Layer

Hasil *feature map* yang telah di normalisasi selanjutnya dilakukan proses *max pooling* yang berguna untuk mengurangi ukuran *feature map* sehingga dapat mempercepat proses pada *feature learning* tanpa menghilangkan informasi penting didalamnya. Hasil *max pooling* diperoleh dari nilai piksel terbesar pada setiap *grid filter*. Pada model AlexNet, terdapat *max pooling layer* dengan ukuran filter sebesar 3×3 dan *stride* sebanyak 2. Ukuran *feature map* hasil pada *max pooling layer* ini dapat dihitung menggunakan persamaan 2.6

$$\begin{aligned}
 O &= \frac{(i - k)}{s} + 1 \\
 &= \frac{(55 - 3)}{2} + 1 \\
 &= \frac{(55)}{2} + 1 \\
 &= 26 + 1 \\
 &= 27
 \end{aligned}$$

Berdasarkan perhitungan diatas, ukuran *feature map* yang dihasilkan dari *max pooling layer* dengan filter berukuran 3×3 dan *stride* sebanyak 2 adalah 27×27 .

Berikut perhitungan pada *max pooling layer*.

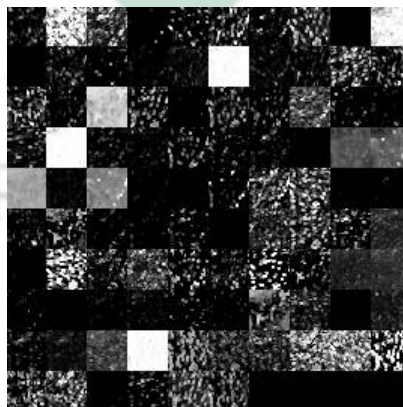
$$Y_{1,1} = \max(0, 10.04, 0, 2.09, 0, 0, 6.30, 0, 0) = 10.04$$

$$Y_{1,3} = \max(0, 0, 0, 0, 0, 0, 0, 0, 0, 68.05) = 68.05$$

$$Y_{1,5} = \max(0, 0, 0, 0, 0, 0, 0, 2.65, 0, 0) = 2.65$$

$$Y_{(:, :, 1)} = \begin{bmatrix} 10.04 & 0 & 68.05 & 68.5 & 2.65 & 0 \dots & 0 \\ 6.30 & 0.78 & 71.25 & 71.25 & 10.79 & 0 & \dots & 0 \\ 0 & 4.88 & 81.24 & 81.24 & 55.45 & 1.02 & \dots & 5.73 \\ 0 & 4.88 & 81.24 & 81.24 & 55.45 & 0 & \dots & 0 \\ 0 & 0.27 & 0 & 6.61 & 6.61 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0.88 & 0 & 7.73 & 0 & 7.56 & \dots & 0 \end{bmatrix}$$

Perhitungan diatas berulang sesuai dengan jumlah feature map. Visualisasi hasil *feature map* pada *max pooling layer* ditunjukkan pada Gambar 4.16



Gambar 4.16 Visualisasi Feature Map pada *Max Pooling Layer*

Hasil feature map pada max pooling layer ini perlu dilakukan proses flatten, dimana matriks feature map akan dirubah dalam bentuk vektor yang selanjutnya digunakan dalam perhitungan pada layer setelah max pooling layer. Berikut contoh

proses perhitungan flatten yang ditunjukkan pada Gambar 4.17.

10.04	0.00	68.05	68.05	2.65	...	13.18	0.00	4.96	0.00	0.00
6.30	0.78	71.25	71.25	10.79	...	13.18	0.00	0.00	0.00	0.00
0.00	4.88	81.24	81.24	55.45	...	0.00	0.00	0.00	0.00	5.73
0.00	4.88	81.24	81.24	55.45	...	0.00	5.25	5.25	0.00	0.00
0.00	0.27	0.00	6.61	6.61	...	0.00	5.25	5.25	0.00	0.00
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.00	4.08	4.08	10.58	0.00	...	0.00	49.13	42.47	0.00	0.00
0.00	4.71	4.71	31.66	31.66	...	0.00	0.00	12.57	3.13	0.00
0.00	9.50	0.00	56.10	56.10	...	0.00	0.00	0.00	0.00	0.00
0.00	9.50	0.00	32.67	32.67	...	0.00	0.00	0.00	0.00	0.00
0.00	0.88	0.00	7.73	0.00	...	3.25	0.00	0.00	0.00	0.00

Gambar 4.17 Ilustrasi Proses *Flatten channel Blue*

4.2.5. Dropout Layer

Hasil dari proses *flatten* kemudian menjadi inputan pada *dropout* layer. Nilai dari setiap node tersebut dilakukan seleksi apakah akan diteruskan ke *layer* selanjutnya atau tidak. Node tersebut diseleksi berdasarkan dari inisialisasi vektor yang nilainya acak dengan range antara 0 sampai 1. Sehingga jika nilai inisialisasi vektor kurang dari nilai probabilitas yang telah ditentukan, maka nilai dari node tersebut diganti menjadi 0 dengan kata lain node tersebut tidak digunakan pada proses selanjutnya. Perhitungan pada *dropout layer* dengan probabilitas sebesar 0.5 yang ditunjukkan pada Gambar 4.18

INPUT	INISIALISASI RANDOM	OUTPUT
10.04	0.483	0.00
6.30	0.805	6.30
0.00	0.411	0.00
⋮	⋮	⋮
0.88	0.295	0.00
68.05	0.544	68.05
71.25	0.897	71.25
81.24	0.646	81.24
81.24	0.807	81.24
6.61	0.098	0.00
⋮	⋮	⋮
0.00	0.885	0.00
0.00	0.885	0.00

Gambar 4.18 Perhitungan *dropout Layer*

4.2.6. Fully Connected Layer

Pada FC *layer* berguna untuk mengklasifikasikan data sesuai dengan jumlah kelas yang ada. Pada proses ini setiap node saling terhubung pada node di layer sebelumnya. Hasil dari proses *flatten* akan menjadi inputan (x) yang selanjutnya dikalikan dengan bobot (W) dan dijumlah dengan bias (b). Perhitungan pada FC *layer* dapat dihitung menggunakan persamaan 2.7. Berikut perhitungan pada FC *layer* dengan input berukuran 1×4096 , bobot dan bias dengan ukuran masing-masing 4096×3 dan 1×3 . Nilai bobot dan bias didapatkan dari perhitungan Glorot *Inizialitation*, namun pada perhitungan dibawah ini menggunakan nilai bobot dan bias yang sudah optimal.

$$x = \begin{bmatrix} 0 & 6.30 & 0 & 0 & 0 & 68.05 & \dots & 0 & 0 \end{bmatrix}$$

$$W = \begin{bmatrix} 0.08 & 0.43 & -0.42 \\ -0.57 & 0.01 & 0.25 \\ -0.03 & -0.05 & 0.17 \\ -0.20 & 0.55 & -0.53 \\ \vdots & \vdots & \vdots \\ 0.99 & -0.34 & 0.80 \end{bmatrix}$$

$$b = \begin{bmatrix} 0.02 & -0.01 & -0.06 \end{bmatrix}$$

Perhitungan pada FC *layer* dapat dihitung menggunakan persamaan 2.7

$$\begin{aligned} y_1 &= \sum_{i=1}^n W_{i,1} + b_1 \\ &= (0.08 \times 0 + (-0.57) \times 6.30 + (-0.03) \times 0 + \dots + 0.99 \times 0) + 0.022 \\ &= -7.05 \end{aligned}$$

$$\begin{aligned} y_2 &= \sum_{i=1}^n W_{i,2} + b_2 \\ &= (0.43 \times 0 + 0.01 \times 6.30 + (-0.05) \times 0 + \dots + (-0.34) \times 0) + (-0.01) \\ &= 3.95 \end{aligned}$$

$$\begin{aligned} y_3 &= \sum_{i=1}^n W_{i,3} + b_3 \\ &= (-0.42 \times 0 + 0.25 \times 6.30 + 0.17 \times 0 + \dots + 0.80 \times 0) + (-0.06) \\ &= -1.32 \end{aligned}$$

$$y = \begin{bmatrix} -7.05 & 3.95 & -1.32 \end{bmatrix}$$

4.2.7. Softmax Layer

Hasil dari FC layer kemudian diaktivasi menggunakan fungsi aktivasi softmax yang berguna untuk mengubah nilai riil menjadi bentuk probabilitas dengan *range* antara 0 sampai 1. Hasil klasifikasi dapat dilihat dari nilai probabilitas yang tertinggi. Berikut perhitungan pada Softmax *layer* yang dapat dihitung menggunakan persamaan 2.8

$$\begin{aligned} p_1(x) &= \frac{e^{x_1}}{\sum_{i=1}^j e^{x_i}} \\ &= \frac{e^{-7.05}}{e^{-7.05} + e^{3.95} + e^{-1.32}} \\ &= 0.000 \end{aligned}$$

$$\begin{aligned} p_2(x) &= \frac{e^{x_2}}{\sum_{i=1}^j e^{x_i}} \\ &= \frac{e^{3.95}}{e^{-7.05} + e^{3.95} + e^{-1.32}} \\ &= 0.995 \end{aligned}$$

$$\begin{aligned} p_3(x) &= \frac{e^{x_3}}{\sum_{i=1}^j e^{x_i}} \\ &= \frac{e^{-1.32}}{e^{-7.05} + e^{3.95} + e^{-1.32}} \\ &= 0.000 \end{aligned}$$

$$p = \begin{bmatrix} 0.000 & 0.995 & 0.005 \end{bmatrix}$$

Pada p(1) menyatakan nilai probabilitas pada kelas *benign lung tissue*, p(2) menyatakan nilai probabilitas pada kelas *lung adenocarcinoma* sedangkan pada p(3) menyatakan nilai probabilitas pada kelas *lung squamous cell carcinoma*. Sehingga hasil pada perhitungan pembentukan sistem klasifikasi ini data citra

terklasifikasikan pada kelas *lung adenocarcinoma*.

4.3. Pengujian dan Evaluasi Model

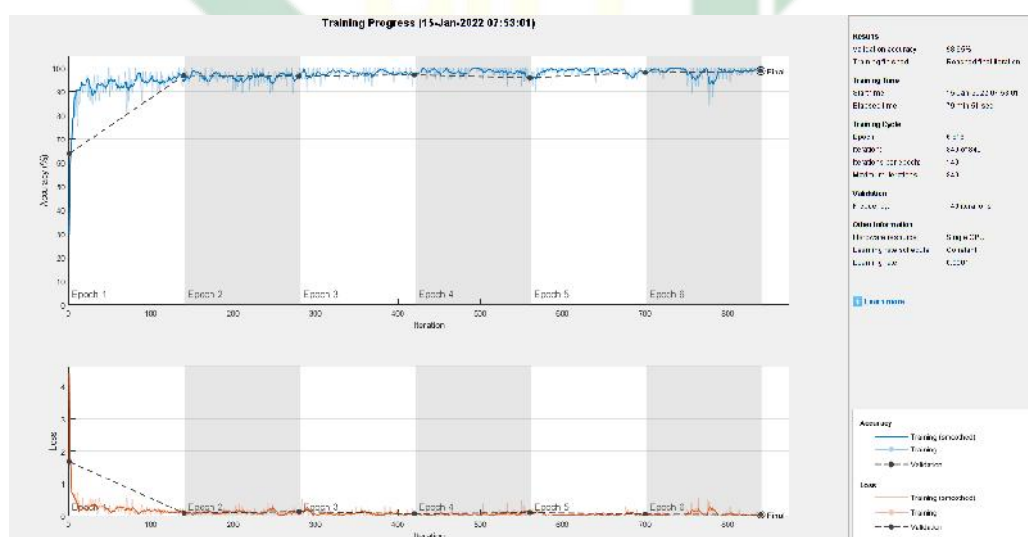
Pada penelitian ini pembentukan sistem klasifikasi dilakukan beberapa uji coba berdasarkan pembagian data, *batch size* serta probabilitas *dropout*. Hal ini dilakukan untuk mendapatkan model yang paling optimal. Dalam penelitian ini pembagian data sebanyak 60:40, 70:30, 80:20 dan 90:10. Jumlah *batch size* yang digunakan sebesar 8, 16, 32, dan 64. Probabilitas *dropout* yang di uji coba sebesar 0.5, 0.6, 0.7 dan 0.8. Hyperparameter lainnya yang digunakan adalah epoch sebanyak 6 serta inisialisasi *learning rate* sebesar 10^{-4} . Evaluasi model dilakukan menggunakan *Confusion matrix* berdasarkan nilai presentase akurasi, sensitivitas serta spesitifitas. Berikut hasil dari beberapa uji coba yang telah dilakukan.

4.3.1. Pembagian Data 60:40

Berdasarkan Tabel 4.1, pada percobaan pembagian data 60:40 didapatkan model yang optimal pada *batch size* sebesar 64 dan nilai probabilitas *dropout* 0.6 dengan presentase akurasi, sensitivitas dan spesitifitas sebesar 98.7%, 98.7%, dan 99.3% secara berturut-turut. Pada 4.19 menunjukkan plot *training progress*, dimana pada percobaan ini diperlukan waktu *training* selama 79 menit 51 detik.

Tabel 4.1 Hasil Percobaan Pembagian Data 60:40

Pembagian Data	Batch Size	Probabilitas Dropout	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)	Waktu (min;sec)
60:40	8	0.5	96.5	96.7	98.3	222;3
		0.6	97.6	97.6	98.8	232;0
		0.7	96.5	96.5	98.2	217;51
		0.8	93.9	93.9	96.9	219;39
	16	0.5	97.2	97.0	97.5	135;11
		0.6	96.3	96.3	98.1	136;31
		0.7	95.1	95.1	97.5	138;26
		0.8	94.4	94.4	97.2	142;8
	32	0.5	97.2	97.2	98.6	100;17
		0.6	97.2	98.5	99.3	98;25
		0.7	97.7	97.7	98.8	94;30
		0.8	97.3	96.8	97.2	102;44
64	0.5	97.4	97.4	98.7	101;0	
	0.6	98.6	98.6	99.3	83;14	
	0.7	98.7	98.7	99.3	79;51	
	0.8	97.0	97.0	98.5	105;27	



Gambar 4.19 Training Progress Model Optimal pada Pembagian Data 60:40

Confusion matrix pada model optimal pada percobaan pembagian data 60:40 ditunjukkan pada Gambar 4.20

Confusion Matrix

Target Class	Benign lung tissue	2000	0	0	100%
	Lung Adenocarcinomas	0	1970	51	97.5%
	Lung squamous cell carcinomas	0	30	1949	98.5%
		100%	98.5%	97.5%	98.7%
		Output Class			
		Benign lung tissue	Lung Adenocarcinomas	Lung squamous cell carcinomas	

Gambar 4.20 Confusion Matrix Model Optimal pada Pembagian Data 60:40

Berdasarkan Gambar 4.20 kelas aktual (*Target class*) direpresentasikan pada setiap baris, sedangkan hasil klasifikasi (*Output class*) direpresentasikan pada setiap kolom. Pada percobaan ini terdapat hasil klasifikasi yang tidak sesuai dengan kelas aktualnya. Terdapat 51 citra *lung adenocarcinoma* yang terklasifikasikan pada kelas *lung squamous cell carcinoma* dan 30 citra *lung squamous cell carcinoma* yang terklasifikasikan sebagai *lung adenocarcinoma*.

4.3.2. Pembagian Data 70:30

Berdasarkan Tabel ??, pada percobaan pembagian data 70:30 didapatkan model yang optimal pada *batch size* sebesar 32 dan nilai probabilitas *dropout* 0.5 dengan presentase akurasi, sensitivitas dan spesifitas sebesar 98.5%, 98.5%, dan 99.3% secara berturut-turut. Pada Gambar 4.21 menunjukkan plot *training progress*, dimana pada percobaan ini diperlukan waktu *training* selama 108 menit 13 detik.

Tabel 4.2 Hasil Percobaan Pembagian Data 70:30

Pembagian Data	Batch Size	Probabilitas Dropout	Akurasi (%)	Sensitivitas (%)	Spesitiftas (%)	Waktu (min;sec)
70:30	8	0.5	97.7	97.7	98.8	244;57
		0.6	97.7	97.7	98.9	240;43
		0.7	95.8	95.8	97.9	248;16
		0.8	95.3	95.3	97.7	263;28
	16	0.5	97.8	97.8	98.9	160;5
		0.6	98.3	98.3	99.1	151;26
		0.7	97.5	97.5	98.8	150;46
		0.8	96.0	96.0	98.0	172;30
	32	0.5	98.5	98.5	99.3	108;13
		0.6	98.3	98.3	99.2	104;5
		0.7	98.2	98.2	99.1	103;2
		0.8	91.4	97.0	98.0	105;35
	64	0.5	97.5	97.5	98.7	90;55
		0.6	96.8	96.7	98.4	123;9
		0.7	96.3	96.3	98.2	87;16
		0.8	97.8	98.3	98.5	101;3



Gambar 4.21 Training Progress Model Optimal pada Pembagian Data 70:30

Confusion matrix pada model optimal pada percobaan pembagian data 70:30 ditunjukkan pada Gambar 4.22

Confusion Matrix

Target Class	Output Class			
	Benign lung tissue	Lung Adenocarcinomas	Lung squamous cell carcinomas	
Benign lung tissue	1500	0	0	99.9%
Lung Adenocarcinomas	0	1444	56	99.6%
Lung squamous cell carcinomas	0	10	1490	97.3%
	100%	97.1%	98.5%	98.5%

Gambar 4.22 Confusion Matrix Model Optimal pada Pembagian Data 70:30

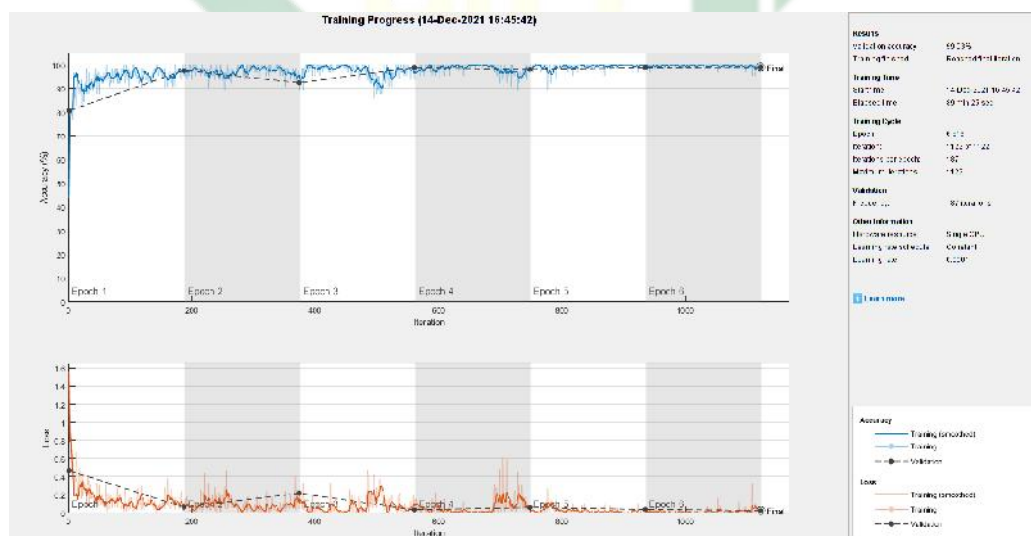
Berdasarkan Gambar 4.22 kelas aktual (*Target class*) direpresentasikan pada setiap baris, sedangkan hasil klasifikasi (*Output class*) direpresentasikan pada setiap kolom. Pada percobaan ini terdapat hasil klasifikasi yang tidak sesuai dengan kelas aktualnya. Terdapat 56 citra *lung adenocarcinoma* yang terklasifikasikan pada kelas *lung squamous cell carcinoma* dan 10 citra *lung squamous cell carcinoma* yang terklasifikasikan sebagai *lung adenocarcinoma*.

4.3.3. Pembagian Data 80:20

Berdasarkan Tabel 4.3, pada percobaan pembagian data 80:20 didapatkan model yang optimal pada *batch size* sebesar 64 dan nilai probabilitas *dropout* 0.5 dengan presentase akurasi, sensitivitas dan spesitifitas sebesar 99%, 99%, dan 99% secara berturut-turut. Pada Gambar 4.23 menunjukkan plot *training progress*, dimana pada percobaan ini diperlukan waktu *training* selama 89 menit 25 detik.

Tabel 4.3 Hasil Percobaan Pembagian Data 80:20

Pembagian Data	Batch Size	Probabilitas Dropout	Akurasi (%)	Sensitivitas (%)	Spesifisitas (%)	Waktu (min;sec)
80:20	8	0.5	92.2	92.2	96.1	283;57
		0.6	97.9	96.9	98.5	261;57
		0.7	92.7	92.7	96.4	261;21
		0.8	96.8	96.8	98.4	259;16
	16	0.5	98.7	98.7	99.3	103;59
		0.6	98.2	98.2	99.1	157;59
		0.7	97.5	97.5	98.8	157;36
		0.8	95.5	97.8	98.0	160;26
	32	0.5	98.5	98.5	99.2	114;24
		0.6	95.2	97.5	98.7	109;49
		0.7	98.9	98.9	99.4	111;27
		0.8	98.8	98.8	99.4	117;13
	64	0.5	99.0	99.0	99.0	89;25
		0.6	97.4	97.7	98.7	115;47
		0.7	96.8	96.8	98.4	114;4
		0.8	98.2	98.2	99.1	100;41



Gambar 4.23 Training Progress Model Optimal pada Pembagian Data 80:20

Confusion matrix pada model optimal pada percobaan pembagian data 80:20 ditunjukkan pada Gambar 4.24

Confusion Matrix

Target Class	Benign lung tissue	1000	0	0	99.9%
	Lung Adenocarcinomas	0	978	22	99.6%
	Lung squamous cell carcinomas	0	7	993	97.3%
		100%	97.1%	98.5%	99.0%
		Benign lung tissue	Lung Adenocarcinomas	Lung squamous cell carcinomas	
		Output Class			

Gambar 4.24 Confusion Matrix Model Optimal pada Pembagian Data 80:20

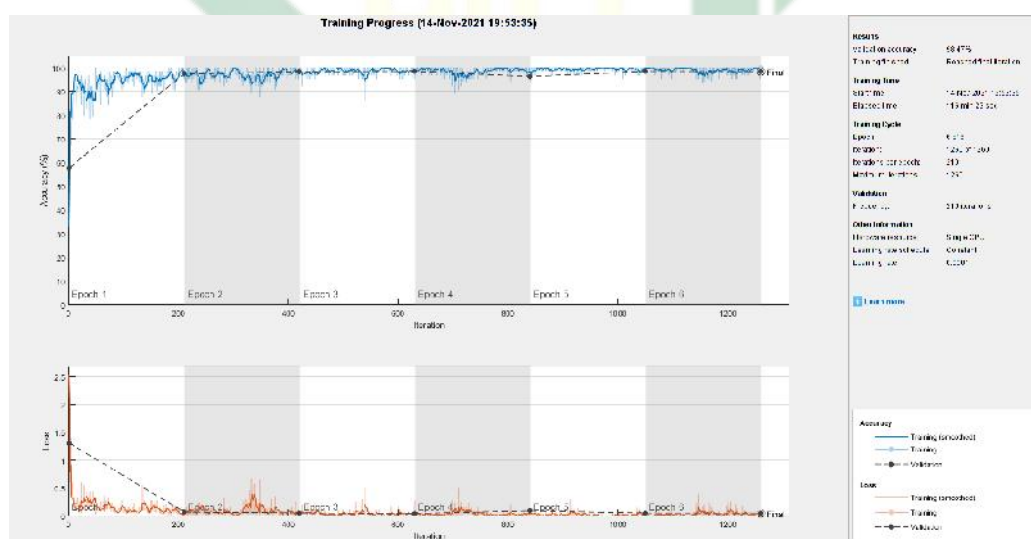
Berdasarkan Gambar 4.24 kelas aktual (*Target class*) direpresentasikan pada setiap baris, sedangkan hasil klasifikasi (*Output class*) direpresentasikan pada setiap kolom. Pada percobaan ini terdapat hasil klasifikasi yang tidak sesuai dengan kelas aktualnya. Terdapat 22 citra *lung adenocarcinoma* yang terklasifikasikan pada kelas *lung squamous cell carcinoma* dan 7 citra *lung squamous cell carcinoma* yang terklasifikasikan sebagai *lung adenocarcinoma*.

4.3.4. Pembagian Data 90:10

Berdasarkan Tabel 4.4, pada percobaan pembagian data 90:10 didapatkan model yang optimal pada *batch size* sebesar 64 dan nilai probabilitas *dropout* 0.5 dengan presentase akurasi, sensitivitas dan spesitifitas sebesar 98.7%, 98.7%, dan 99.4% secara berturut-turut. Pada Gambar 4.25 menunjukkan plot *training progress*, dimana pada percobaan ini diperlukan waktu *training* selama 116 menit 23 detik.

Tabel 4.4 Hasil Percobaan Pembagian Data 90:10

Pembagian Data	Batch Size	Probabilitas Dropout	Akurasi (%)	Sensitivitas (%)	Spesitivitas (%)	Waktu (min:sec)
90:10	8	0.5	97.2	97.2	98.6	282;50
		0.6	97.7	97.7	98.9	273;15
		0.7	96.4	96.4	98.2	280;9
		0.8	96.6	96.6	98.3	282;34
	16	0.5	96.9	96.9	98.5	170;57
		0.6	96.6	90.5	95.3	169;21
		0.7	94.9	94.9	97.5	169;1
		0.8	93.9	93.9	97.0	174;31
	32	0.5	96.7	96.7	98.3	121;26
		0.6	95.1	95.1	97.6	124;51
		0.7	97.4	97.4	98.7	115;20
		0.8	98.5	98.5	99.3	124;3
	64	0.5	98.7	98.7	99.4	116;23
		0.6	96.3	96.3	98.1	114;27
		0.7	98.3	98.3	99.2	100;18
		0.8	97.8	97.8	98.9	112;11



Gambar 4.25 Training Progress Model Optimal pada Pembagian Data 90:10

Confusion matrix pada model optimal pada percobaan pembagian data 90:10 ditunjukkan pada Gambar 4.26

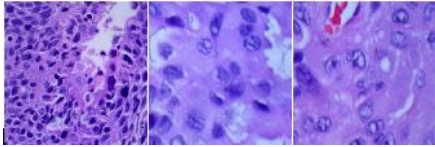
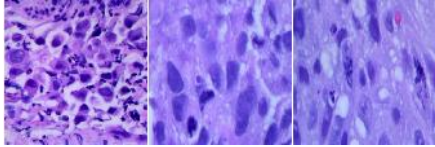
Confusion Matrix

	Benign lung tissue	Lung Adenocarcinomas	Lung squamous cell carcinomas		
Target Class	Benign lung tissue	500	0	0	100%
	Lung Adenocarcinomas	0	493	12	97.6%
	Lung squamous cell carcinomas	0	7	488	98.6%
		100%	98.6%	97.6%	98.7%
	Benign lung tissue	Lung Adenocarcinomas	Lung squamous cell carcinomas	Output Class	

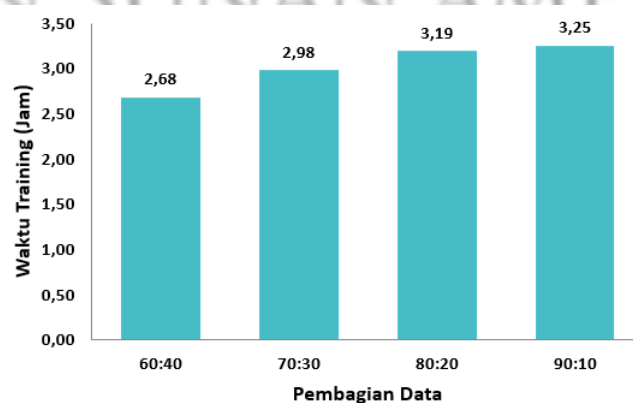
Gambar 4.26 Confusion Matrix Model Optimal pada Pembagian Data 90:10

Berdasarkan Gambar 4.26 kelas aktual (*Target class*) direpresentasikan pada setiap baris, sedangkan hasil klasifikasi (*Output class*) direpresentasikan pada setiap kolom. Pada percobaan ini terdapat hasil klasifikasi yang tidak sesuai dengan kelas aktualnya. Terdapat 12 citra *lung adenocarcinoma* yang terklasifikasikan pada kelas *lung squamous cell carcinoma* dan 7 citra *lung squamous cell carcinoma* yang terklasifikasikan sebagai *lung adenocarcinoma*. Setelah dilakukan beberapa uji coba di atas, didapatkan model yang optimal pada percobaan pembagian data 80:20, *batch size* sebesar 64 dan probabilitas *dropout* bernilai 0.5 dengan akurasi 99%, sensitivitas 99% dan spesifitas 99% serta waktu *training* selama 89 menit 25 detik. Berdasarkan *Confusion matrix* pada Gambar 4.24 terdapat kelas aktual *lung adenocarcinoma*, namun terklasifikasi sebagai *lung squamous cell carcinomas* sebanyak 22 citra. Kemudian terdapat kelas aktual *lung squamous cell carcinoma* terklasifikasi, namun terklasifikasi sebagai *lung adenocarcinoma* sebanyak 7 citra. Sampel data citra histopatologi yang terklasifikasi tidak sesuai dengan kelas aktualnya dapat dilihat pada Tabel 4.5

Tabel 4.5 Sampel Data Citra Histopatologi Kanker Paru

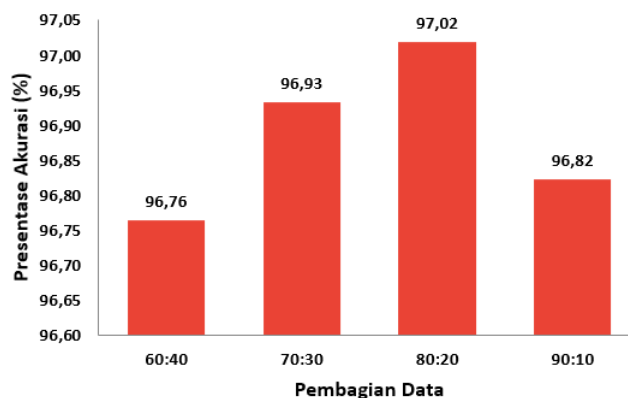
Kelas <i>lung adenocarcinoma</i> terklasifikasi sebagai <i>lung squamous cell carcinoma</i>	
Kelas <i>lung squamous cell carcinoma</i> terklasifikasi sebagai <i>lung adenocarcinoma</i>	

Pada Tabel 4.5, dapat dilihat bahwa terdapat citra histopatologi kelas *lung adenocarcinoma* yang serupa pada citra histopatologi kelas *lung squamous cell carcinoma*, begitu pula sebaliknya. Hal ini yang menyebabkan model salah mengklasifikasikan citra histopatologi tersebut. Pada penelitian ini mencoba untuk melakukan analisis pengaruh pembagian data, jumlah *batch size* serta nilai probabilitas *dropout* terhadap waktu *training* dan presentase akurasi. Perhitungan dilakukan dengan menghitung rata-rata waktu *training* dan presentase akurasi pada setiap parameter uji coba. Pengaruh pembagian data terhadap waktu *training* disajikan pada 4.27

Gambar 4.27 Grafik Pengaruh Pembagian Data terhadap Waktu *Training*

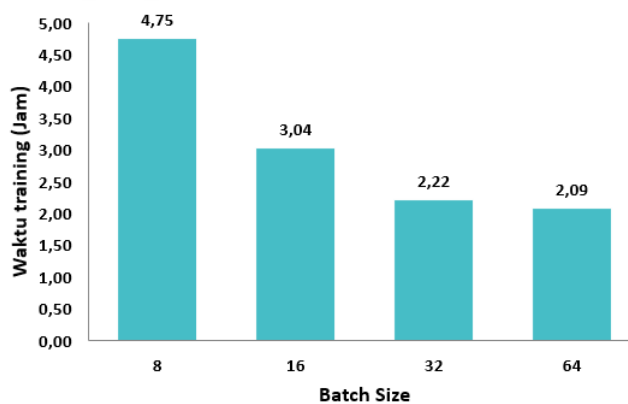
Berdasarkan Gambar 4.27, dapat disimpulkan bahwa semakin banyak data

training yang digunakan maka waktu yang dibutuhkan untuk *training* juga semakin lama. Pengaruh pembagian data terhadap presentase akurasi disajikan pada Gambar 4.28



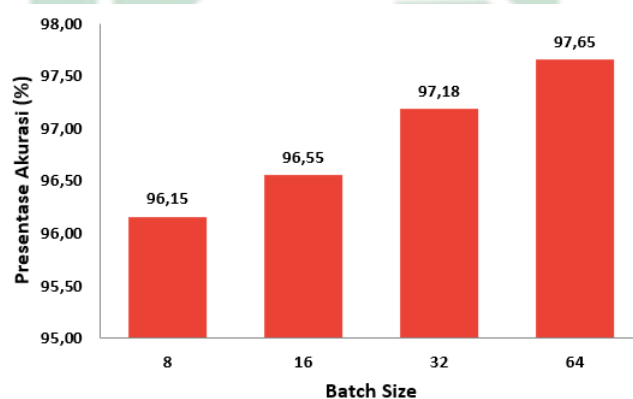
Gambar 4.28 Grafik Pengaruh Pembagian Data terhadap Presentase Akurasi

Berdasarkan Gambar 4.28, dapat disimpulkan bahwa terlalu sedikit maupun terlalu banyak data *training* akan menyebabkan presentase akurasi menurun. Hal ini dikarenakan terlalu sedikit data *training* maka model cenderung sedikit mempelajari pola yang ada pada data citra histopatologi (Althnian et al., 2021). Sebaliknya, terlalu banyak data *training* maka model dapat mengalami overfitting (Koroteev, 2018). Pengaruh jumlah *batch size* terhadap waktu *training* disajikan pada Gambar 4.29



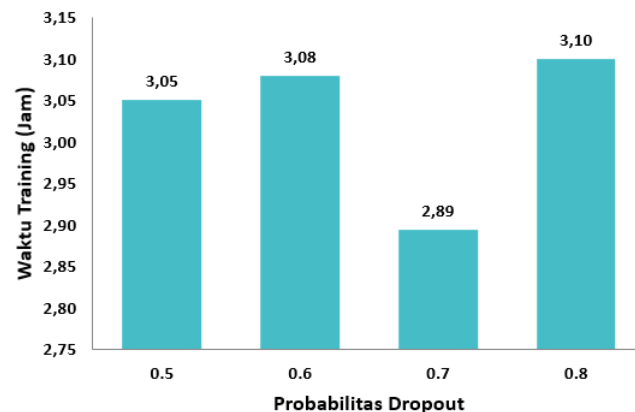
Gambar 4.29 Grafik Pengaruh *Batch Size* terhadap Waktu *Training*

Berdasarkan Gambar 4.29, dapat disimpulkan bahwa semakin besar nilai *batch size* maka waktu yang dibutuhkan untuk *training* semakin cepat. Hal ini disebabkan karena semakin banyak sampel data yang digunakan maka iterasi setiap *epoch* yang di jalankan pada proses *training* akan semakin sedikit sehingga waktu komputasi yang dibutuhkan semakin cepat. Namun semakin besar nilai *batch size* maka dibutuhkan memori komputer yang lebih besar juga (Wilson, 2019). Selanjutnya pengaruh jumlah *batch size* terhadap presentase akurasi disajikan pada Gambar 4.30



Gambar 4.30 Grafik Pengaruh *Batch Size* terhadap Presentase Akurasi

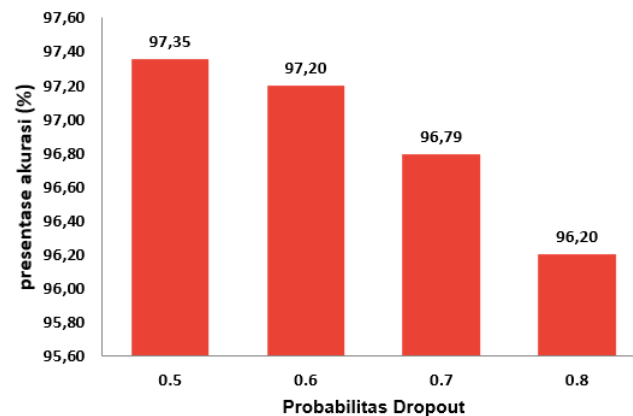
Berdasarkan Gambar 4.30, dapat disimpulkan bahwa semakin besar nilai *batch size* maka presentase akurasi yang didapatkan semakin bagus. Hal ini sesuai dengan pernyataan dari Ibrahem Kandel yang menyatakan semakin besar nilai *batch size* maka semakin bagus presenase akurasi yang didapatkan (Kandel and Castelli, 2020). Selanjutnya pengaruh nilai probabilitas *dropout* terhadap waktu *training* disajikan pada Gambar 4.31



Gambar 4.31 Grafik Pengaruh Probabilitas *Dropout* terhadap Waktu *Training*

Berdasarkan Gambar 4.31, dapat dilihat pada probabilitas *dropout* sebesar 0.5, 0.6 dan 0.8 waktu komputasi yang dibutuhkan relatif semakin lama. Namun pada probabilitas 0.7 waktu komputasi yang didapatkan lebih cepat dibandingkan probabilitas lainnya. Hal ini disebabkan karena cara kerja pada *dropout* layer tergantung pada inisialisasi vektor yang bernilai acak antara 0 sampai 1. Sehingga waktu komputasi yang dibutuhkan pada proses training tidak tergantung berdasarkan nilai probabilitas *dropout*.

Namun terdapat pendapat dari penelitian sebelumnya (Poernomo and Kang, 2018) yang menyatakan jika jaringan menggunakan *dropout* secara signifikan akan meningkatkan waktu pelatihan karena membutuhkan waktu lebih lama untuk menuju konvergen dibandingkan jaringan yang tidak menggunakan *dropout* dengan arsitektur yang sama. Sehingga perlu dilakukan uji coba nilai probabilitas *dropout* untuk menemukan nilai yang paling optimal dalam suatu dataset.



Gambar 4.32 Grafik Pengaruh Probabilitas *Dropout* terhadap Presentase Akurasi

Berdasarkan Gambar 4.32, dapat dilihat bahwa semakin besar nilai probabilitas *dropout* maka semakin kecil presentase akurasi yang didapatkan. Hal ini sejalan dengan penelitian sebelumnya, dimana peneliti menyimpulkan bahwa semakin besar nilai probabilitas *dropout* maka model semakin lama mencapai konvergen sehingga menghasilkan akurasi yang jelek (Piotrowski et al., 2020).

4.4. Aplikasi klasifikasi Kanker Paru Citra Histopatologi

Pada penelitian ini, model yang optimal dari beberapa uji coba selanjutnya di buat dalam bentuk aplikasi. Hal ini dilakukan guna untuk mempermudah ahli medis untuk memasukkan data citra histopatologi yang kemudian akan didapatkan hasil klasifikasi. Pada aplikasi ini terdapat tiga kelas citra histopatologi kanker paru, yaitu *benign lung tissue*, *lung adenocarcinoma* dan *squamous lung cell carcinoma*.



Gambar 4.33 Tampilan awal Aplikasi Sistem Klasifikasi

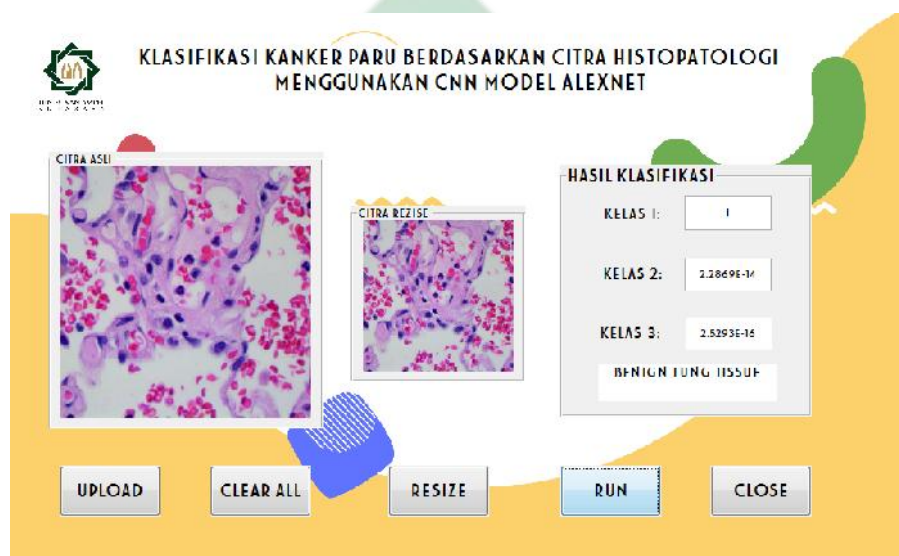
Pada Gambar 4.33, merupakan tampilan awal pada aplikasi sistem klasifikasi. Terdapat judul aplikasi, logo UIN Sunan Ampel Surabaya, tombol mulai yang berguna untuk menuju pada tampilan utama proses klasifikasi terakhir terdapat nama pembuat aplikasi serta dosen pembimbing.



Gambar 4.34 Tampilan Utama Aplikasi Sistem Klasifikasi

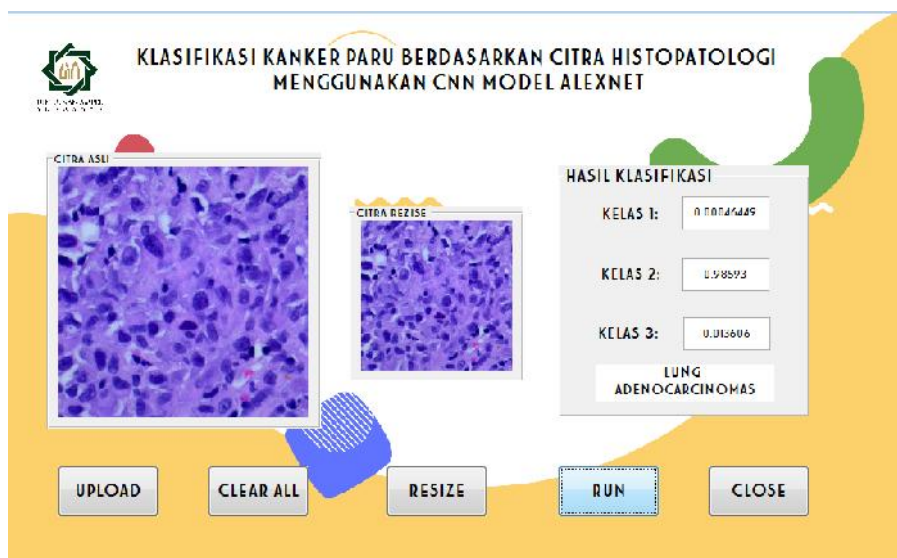
Selanjutnya pada tampilan utama terdapat 5 tombol yang dapat dilihat pada Gambar 4.34. Tombol *upload* berguna untuk mencari citra histopatologi yang akan digunakan sebagai inputan pada sistem klasifikasi. Tombol *clear all* berguna untuk

menghapus semua konten yang ada pada tampilan utama. Tombol *resize* berguna untuk menyesuaikan ukuran citra input berdasarkan ketentuan model AlexNet, yaitu $227 \times 227 \times 3$. Tombol *run* berguna untuk menjalankan proses klasifikasi yang kemudian akan menghasilkan probabilitas pada masing-masing kelas. Hasil klasifikasi ditentukan berdasarkan nilai probabilitas yang paling tinggi. Tombol *close* berguna untuk menutup tampilan utama.



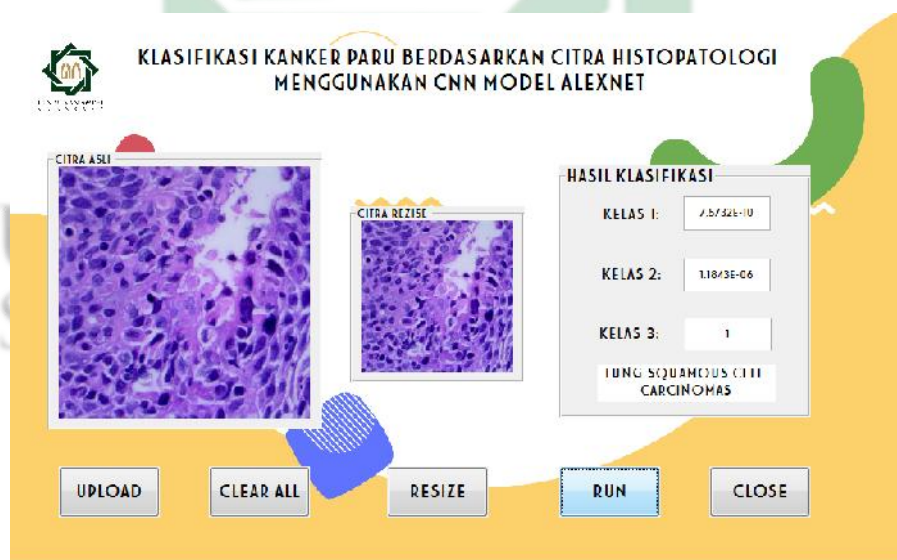
Gambar 4.35 Contoh Hasil Klasifikasi pada Kelas *Benign Lung Tissue*

Pada Gambar 4.35 merupakan contoh hasil klasifikasi kelas *benign lung tissue*. Dapat dilihat pada hasil klasifikasi probabilitas tertinggi ada pada kelas 1, yaitu *benign lung tissue*. Sehingga pada contoh ini citra input terklasifikasikan benar sesuai dengan kelas aktualnya.



Gambar 4.36 Contoh Hasil Klasifikasi pada Kelas *Lung Adenocarcinoma*

Pada Gambar 4.36 merupakan contoh hasil klasifikasi kelas *lung adenocarcinoma*. Dapat dilihat pada hasil klasifikasi probabilitas tertinggi ada pada kelas 2, yaitu *lung adenocarcinoma*. Sehingga pada contoh ini citra input terklasifikasikan benar sesuai dengan kelas aktualnya.



Gambar 4.37 Contoh Hasil Klasifikasi pada Kelas *Squamous Lung Cell Carcinoma*

Pada Gambar 4.37 merupakan contoh hasil klasifikasi kelas *squamous lung cell carcinoma*. Dapat dilihat pada hasil klasifikasi probabilitas tertinggi ada pada kelas 3, yaitu *squamous lung cell carcinoma*. Sehingga pada contoh ini citra input

terklasifikasikan benar sesuai dengan kelas aktualnya.

4.5. Integrasi Keislaman

Pada penelitian ini mencoba untuk mengklasifikasikan penyakit kanker paru berdasarkan citra histopatologi. Dalam masalah ini terdapat tiga kelas, yaitu *benign lung tissue*, *lung adenocarcinoma*, dan *lung squamous cell carcinomas*. Kanker paru merupakan penyakit dengan kasus kematian tertinggi baik di dunia maupun di Indonesia yang disebabkan karena diagnosis kanker. Namun sebagai umat muslim kita tidak boleh berpasrah diri tanpa adanya usaha, karena jika Allah SWT menurunkan suatu penyakit, maka pasti ada obatnya kecuali penyakit pikun/tua. Sebagaimana hadist nabi yang diriwayatkan oleh Abu Dawud, Tirmidzi, Nasa'I dan Ibnu Majah.

عَنْ أَسَامَةَ بْنِ شَرِيكٍ أَنَّ رَسُولَ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ قَالَ: تَدَاوُوا فَإِنَّ اللَّهَ عَزَّ وَجَلَّ لَمْ يَضَعْ دَاءً إِلَّا وَضَعَ لَهُ دَوَاءً غَيْرَ دَاءٍ وَاحِدٍ: الْهَرَمُ

“Berobatlah, karena Allah tidak menjadikan penyakit kecuali menjadikan pula obatnya, kecuali satu penyakit yaitu pikun (tua)”. HR. Abu Dawud, Tirmidzi, Nasa'I dan Ibnu Majah. Hadits tersebut menerangkan bahwa semua penyakit pasti ada obatnya. Maka dari itu, jika seorang manusia mendapatkan ujian berupa suatu penyakit harus senantiasa mencari tahu obat yang dapat menyembuhkan penyakit tersebut, namun perlu diingat bahwa kesembuhan suatu penyakit hanya karena izin Allah SWT semata. Hadist tersebut memberikan dorongan kepada para peneliti untuk terus melakukan penelitian agar segera menemukan obat penyakit berat yang belum diketahui obatnya. Allah SWT telah memerintahkan hambanya agar terus mempelajari ilmu baru, salah satunya di bidang sains dan teknologi. Ilmu pada bidang sains dan teknologi sangat berguna untuk berbagai aspek kehidupan sehari-hari. Perintah ini sesuai dengan hadist diriwayatkan oleh Ibnu Majah.

قَالَ رَسُولُ اللَّهِ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ: طَلَبُ الْعِلْمِ فَرِيضَةٌ عَلَى كُلِّ مُسْلِمٍ وَ مُسْلِمَةٍ

Rasulullah SAW bersabda: “Menuntut ilmu itu hukumnya wajib, bagi muslim laki-laki dan muslim perempuan”. (HR. Ibnu Majah). Pada hadist tersebut dijelaskan bahwa penting bagi umat muslim untuk mempelajari banyak hal. Sejalan dengan pernyataan “mempelajari banyak hal”, suatu sistem klasifikasi juga melakukan pembelajaran pada proses feature learning dari data citra histopatologi. Pada penelitian ini telah dilakukan beberapa uji coba yang digunakan untuk menemukan model optimal dalam klasifikasi kanker paru. Dalam hal ini, diperlukan adanya ikhtiar serta diimbangi dengan doa kepada Allah SWT agar mendapatkan kemudahan dalam mencapai hasil yang optimal. Hal ini sejalan dengan firman Allah SWT pada QS. al-Ghafir ayat 60, yang berbunyi:

وَقَالَ رَبُّكُمْ ادْعُونِي أَسْتَجِبْ لَكُمْ إِنَّ الَّذِينَ يَسْتَكْبِرُونَ عَنْ عِبَادَتِي سَيَدْخُلُونَ جَهَنَّمَ دَاخِرِينَ ٦٠

Artinya : “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. al-Ghafir ayat 60). Pada ayat tersebut dijelaskan bahwa jika seorang hamba berdoa kepada Allah, maka doa tersebut akan dikabulkan-Nya. Pembentukan sistem klasifikasi kanker paru pada penelitian ini diperlukan waktu *training* yang cukup lama hingga akhirnya model dapat mempelajari data citra histopatologi dengan hasil yang optimal.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan hasil yang didapatkan dari klasifikasi kanker paru berdasarkan citra histopatologi menggunakan metode CNN model AlexNet, kesimpulan yang dapat di ambil yaitu:

1. Berdasarkan beberapa uji coba yang telah dilakukan berdasarkan pembagian data, batch size dan nilai probabilitas dropout didapatkan model yang paling optimal berdasarkan presentase nilai akurasi, sensitivitas serta spesifitas sebesar 99%, 99% serta 99% secara berturut-turut. Model tersebut menggunakan pembagian data 80:20, batch size sebesar 64 serta probabilitas dropout bernilai 0.5 dengan waktu komputasi 89 menit 25 detik.
2. Pada beberapa percobaan yang telah dilakukan didapatkan pengaruh pada setiap parameter uji coba terhadap waktu komputasi. Pada penelitian ini semakin banyak data yang digunakan pada proses training maka semakin lama waktu komputasi yang dibutuhkan. Namun jika data training terlalu sedikit maupun terlalu banyak maka presentase akurasi yang dihasilkan semakin menurun. Semakin besar nilai batch size maka semakin cepat waktu komputasi yang dibutuhkan serta akurasi presentase yang dihasilkan semakin bagus namun diperlukan memory komputer yang lebih besar. Selanjutnya didapatkan waktu komputasi tercepat pada probabilitas 0.7 sedangkan pada nilai probabilitas lainnya semakin besar nilai probabilitas

maka semakin lama waktu komputasi yang dibutuhkan serta presentase akurasi yang didapatkan semakin menurun.

5.2. Saran

Pada penelitian mengenai klasifikasi kanker paru berdasarkan data citra histopatologi menggunakan metode CNN model AlexNet perlu dilakukan perbaikan dan inovasi baru agar didapatkan sistem klasifikasi yang lebih baik lagi. Beberapa saran dari penulis untuk penelitian selanjutnya antara lain:

1. Melakukan beberapa uji coba dengan hyperparameter lainnya, seperti banyaknya epoch, inisialisasi *learning rate* dan lain sebagainya.
2. Menggunakan penggabungan dari beberapa dataset citra histopatologi kanker paru, sehingga model dapat mempelajari pola citra histopatologi kanker paru lebih banyak. Hal ini dilakukan agar model dapat belajar lebih banyak pola citra histopatologi kanker paru yang diharapkan dapat menghasilkan presentase akurasi yang lebih baik dari penelitian ini.

UIN SUNAN AMPEL
S U R A B A Y A

DAFTAR PUSTAKA

- Agarap, A. F. (2018). Deep Learning using Rectified Linear Units (ReLU). (1):2–8.
- Al-Saffar, A. A. M., Tao, H., and Talab, M. A. (2017). Review of deep convolution neural network in image classification. *International Conference on Radar, Antenna, Microwave, Electronics and Telecommunications*, pages 26–31.
- Al-Yasriy, H. F., Al-Husieny, M. S., Mohsen, F. Y., Khalil, E. A., and Hassan, Z. S. (2020). Diagnosis of Lung Cancer Based on CT Scans Using CNN. *IOP Conference Series: Materials Science and Engineering*, 928(2).
- Alodokter (2020). Mengenal Laporan Patologi Kanker.
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Van Esesn, B. C., Awwal, A. A. S., and Asari, V. K. (2018). The history began from alexnet: A comprehensive survey on deep learning approaches. *arXiv preprint arXiv:1803.01164*.
- Althnian, A., AlSaeed, D., Al-Baity, H., Samha, A., Dris, A. B., Alzakari, N., Abou Elwafa, A., and Kurdi, H. (2021). Impact of dataset size on classification performance: An empirical evaluation in the medical domain. *Applied Sciences (Switzerland)*, 11(2):1–18.
- Alwanda, M. R., Ramadhan, R. P. K., and Alamsyah, D. (2020). Implementasi Metode Convolutional Neural Network Menggunakan Arsitektur LeNet-5 untuk Pengenalan Doodle. *Jurnal Algoritme*, 1(1):45–56.

- American Cancer Society (2018). Cancer of Unknown Primary Risk Factors. *cancer.org*, pages 1–6.
- Andono, P. N. and Sutojo, T. (2017). *Pengolahan citra digital*. Penerbit Andi.
- Apsari, R., Aditya, Y. N., Purwanti, E., and Arof, H. (2021). Development of lung cancer classification system for computed tomography images using artificial neural network. In *AIP Conference Proceedings*, volume 2329.
- Aszemi, N. M. and Dominic, P. D. (2019). Hyperparameter optimization in convolutional neural network using genetic algorithms. *International Journal of Advanced Computer Science and Applications*, 10.
- Atole, R. R. and Park, D. (2018). A multiclass deep convolutional neural network classifier for detection of common rice plant anomalies. *International Journal of Advanced Computer Science and Applications*, 9(1):67–70.
- Bora, K., Chowdhury, M., Mahanta, L. B., Kundu, M. K., and Das, A. K. (2016). Pap smear image classification using convolutional neural network. In *Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing*, pages 1–8.
- 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 (LC25000). pages 1–2.
- Braithwaite, J. (2020). What Is Cancer? *The Lancet*.
- Brownlee, J. (2017). What is the difference between a parameter and a hyperparameter?

- Brownlee, J. (2018). A gentle introduction to dropout for regularizing deep neural networks.
- Cai, S., Shu, Y., Chen, G., Ooi, B. C., Wang, W., and Zhang, M. (2019). Effective and Efficient Dropout for Deep Convolutional Neural Networks. pages 1–12.
- Claesen, M. and Moor, B. L. D. (2015). Hyperparameter search in machine learning.
- Coskun, M., Ucar, A., Yildirim, O., and Demir, Y. (2017). Face recognition based on convolutional neural network. *Proceedings of the International Conference on Modern Electrical and Energy Systems, MEES 2017*, 2018-Janua:376–379.
- Dabral, I., Singh, M., and Kumar, K. (2021). Cancer Detection Using Convolutional Neural Network. *Lecture Notes in Networks and Systems*, 175(July):290–298.
- Dahl, G. E., Sainath, T. N., and Hinton, G. E. (2013). Improving deep neural networks for LVCSR using rectified linear units and dropout. In *2013 IEEE international conference on acoustics, speech and signal processing*, pages 8609–8613. IEEE.
- Dewa, C. K. and Fadhilah, A. L. (2018). Convolutional Neural Networks for Handwritten Javanese Character Recognition. 12(1):83–94.
- Dudek, G. (2019). Generating random weights and biases in feedforward neural networks with random hidden nodes. *Information Sciences*, 481.
- Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115–118.
- Fayed, L. (2021). Differences between a malignant and benign tumor.

Ferlay, J., Colombet, M., Soerjomataram, I., Parkin, D. M., Piñeros, M., Znaor, A., and Bray, F. (2021). Cancer statistics for the year 2020: An overview. *International journal of cancer*.

Gholamalinezhad, H. and Khosravi, H. (2020). Pooling Methods in Deep Neural Networks, a Review.

Globocan (2020). Globocan 2020: New global cancer data.

Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323. JMLR Workshop and Conference Proceedings.

He, Y.-L., Zhang, X.-L., Ao, W., and Huang, J. Z. (2018). Determining the optimal temperature parameter for Softmax function in reinforcement learning. *Applied Soft Computing*, 70:80–85.

health, V. (2021). An overview of lung cancer stages.

Hibatullah, A. and Maliki, I. (2019). Penerapan Metode Convolutional Neural Network Pada Pengenalan Pola Citra Sandi Rumput.

Hosny, K. M., Kassem, M. A., and Fouad, M. M. (2020). Classification of skin lesions into seven classes using transfer learning with AlexNet. *Journal of digital imaging*, 33(5):1325–1334.

Hossain, M. A. and Sajib, M. S. A. (2019). Classification of image using convolutional neural network (CNN). *Global Journal of Computer Science and Technology*.

- Ide, H. and Kurita, T. (2017). Improvement of learning for CNN with ReLU activation by sparse regularization. *Proceedings of the International Joint Conference on Neural Networks*.
- Ilahiyah, S. and Nilogiri, A. (2018). Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network. *JUSTINDO (Jurnal Sistem dan Teknologi Informasi Indonesia)*, 3(2):49–56.
- Jie, H. J. and Wanda, P. (2020). Runpool: A dynamic pooling layer for convolution neural network. *International Journal of Computational Intelligence Systems*, 13(1):66–76.
- Kandel, I. and Castelli, M. (2020). The effect of batch size on the generalizability of the convolutional neural networks on a histopathology dataset. *ICT Express*, 6(4):312–315.
- Kang, J., Ullah, Z., and Gwak, J. (2021). Mri-based brain tumor classification using ensemble of deep features and machine learning classifiers. *Sensors*, 21(6):1–21.
- Kemenkes RI (2017). *Pedoman Nasional Pelayanan Kanker Paru*.
- Kemenkes RI (2018). *Pedoman Pengendalian Faktor Risiko Kanker Paru*.
- Kerlikowske, K., Carney, P. A., Geller, B., Mandelson, M. T., Taplin, S. H., Malvin, K., Ernster, V., Urban, N., Cutter, G., Rosenberg, R., and Ballard-Barbash, R. (2000). Performance of screening mammography among women with and without a first-degree relative with breast cancer. *Annals of Internal Medicine*, 133(11):855–863.

- Koroteev, M. (2018). Machine Learning Models Overfitting and Generalization in Very Big Datasets. 10:75–78.
- Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25:1097–1105.
- Kulkarni, S. and Harnoorkar, S. (2020). Comparative Analysis of CNN Architectures. 7(June 6):1459–1464.
- Lecun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based Learning Applied to Document Recognition. *proc. OF THE IEEE*.
- Lee, K. B., Cheon, S., and Kim, C. O. (2017). A convolutional neural network for fault classification and diagnosis in semiconductor manufacturing processes. *IEEE Transactions on Semiconductor Manufacturing*, 30(2):135–142.
- Li, F., Sone, S., Abe, H., MacMahon, H., Armato, S. G., and Doi, K. (2002). Lung cancers missed at low-dose helical CT screening in a general population: Comparison of clinical, histopathologic, and imaging findings. *Radiology*, 225(3):673–683.
- Liu, Y., Gadepalli, K., Norouzi, M., Dahl, G. E., Kohlberger, T., Boyko, A., Venugopalan, S., Timofeev, A., Nelson, P. Q., Corrado, G. S., Hipp, J. D., Peng, L., and Stumpe, M. C. (2017). Detecting Cancer Metastases on Gigapixel Pathology Images. pages 1–13.
- Maeda-Gutiérrez, V., Galván-Tejada, C. E., Zanella-Calzada, L. A., Celaya-Padilla, J. M., Galván-Tejada, J. I., Gamboa-Rosales, H., Luna-García, H., Magallanes-Quintanar, R., Méndez, C. A. G., and Olvera-Olvera, C. A. (2020). Comparison

- of convolutional neural network architectures for classification of tomato plant diseases. *Applied Sciences (Switzerland)*, 10.
- Mallick, S. and Nayak, S. (2018). Number of parameters and tensor sizes in a convolutional neural network (cnn).
- Mangal, S., Chaurasia, A., and Khajanchi, A. (2020). Convolution Neural Networks for diagnosing colon and lung cancer histopathological images.
- Masud, M., Sikder, N., Nahid, A.-a., and Bairagi, A. K. (2021). A Machine Learning Approach to Diagnosing Lung and Colon Cancer Using a Deep Learning-Based Classification Framework. pages 1–20.
- Mathworks. Initialize learnable parameters for model function.
- MathWorks. What is a convolutional neural network.
- McIntyre, A. and Ganti, A. K. (2017). Lung cancer—A global perspective. *Journal of Surgical Oncology*, 115(5):550–554.
- Nayak, S. (2018). Understanding alexnet.
- Pantanowitz, L., Quiroga-Garza, G. M., Bien, L., Heled, R., Laifenfeld, D., Linhart, C., Sandbank, J., Albrecht Shach, A., Shalev, V., Vecsler, M., Michelow, P., Hazelhurst, S., and Dhir, R. (2020). An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: a blinded clinical validation and deployment study. *The Lancet Digital Health*, 2(8):e407–e416.
- Papandrianos, N., Papageorgiou, E., Anagnostis, A., and Feleki, A. (2020). A deep-learning approach for diagnosis of metastatic breast cancer in bones from whole-body scans. *Applied Sciences (Switzerland)*, 10(3).

- Patel, A. (2020). Benign vs Malignant Tumors. *JAMA Oncology*, 6(9):1488.
- Piotrowski, A. P., Napiorkowski, J. J., and Piotrowska, A. E. (2020). Impact of deep learning-based dropout on shallow neural networks applied to stream temperature modelling. *Earth-Science Reviews*, 201(August 2019):103076.
- Poernomo, A. and Kang, D. K. (2018). Biased dropout and crossmap dropout: Learning towards effective dropout regularization in convolutional neural network. *Neural networks : the official journal of the International Neural Network Society*, 104:60–67.
- Putra, D. (2010). *Pengolahan citra digital*. Penerbit Andi.
- Roth, G. A., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., et al. (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, 392(10159):1736–1788.
- Samah, A. A., Fauzi, M. F. A., and Mansor, S. (2017). Classification of benign and malignant tumors in histopathology images. *Proceedings of the 2017 IEEE International Conference on Signal and Image Processing Applications, ICSIPA 2017*, pages 102–106.
- Sewak, M., Karim, M. R., and Pujari, P. (2018). *Practical convolutional neural networks: implement advanced deep learning models using Python*. Packt Publishing Ltd.
- Shaheen, M., Khan, R., Biswal, R. R., Ullah, M., Khan, A., Uddin, M. I., Zareei,

- M., and Waheed, A. (2021). Acute Myeloid Leukemia (AML) Detection Using AlexNet Model. *Complexity*.
- Shanthi, T. and Sabeenian, R. S. (2019). Modified Alexnet architecture for classification of diabetic retinopathy images. *Computers & Electrical Engineering*, 76:56–64.
- Sinha, T. (2018). Tumors: Benign and Malignant. *Cancer therapy & Oncology International Journal*, 10(3):1–3.
- Society, A. C. (2019a). Signs and symptoms of lung cancer.
- Society, A. C. (2019b). What is lung cancer.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Machine Learning Research*, pages 1929–1958.
- Sulistiyani, S. R., Setyawan, A., and Komarudin, M. (2016). *Pengolahan Citra: dasar dan contoh penerapannya*. TEKNOSAIN.
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., and Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: A Cancer Journal for Clinicians*.
- Sutoyo, T. d., Mulyanto, E., Suhartono, V., and Nurhayati, O. D. (2009). Teori pengolahan citra digital. *Yogyakarta: Andi*.
- Tyagi, V. (2018). Understanding Digital Image Processing. *Understanding Digital Image Processing*, (November).

- Wang, D., Khosla, A., Gargeya, R., Irshad, H., and Beck, A. H. (2016). Deep Learning for Identifying Metastatic Breast Cancer. pages 1–6.
- Widodo, S., Faculty, H. S., Java, C., Yudistira, I. A., Faculty, H. S., Java, C., Sulistiyanti, A., Faculty, H. S., Java, C., Faculty, H. S., Surakarta, U. A., and Java, C. (2021). Detection of Covid-19 on X-Ray Images Using a Deep Learning Convolution Neural Network. pages 255–259.
- Wilson, J. (2019). Does batch size affect GPU usage?
- World Health Organization (2020). Cancer Incident in Indonesia. *International Agency for Research on Cancer*, 858:1–2.
- Yamashita, R., Nishio, M., Kinh, R., Do, G., and Togashi, K. (2018). Convolutional neural networks : an overview and application in radiology. pages 611–629.
- Zulkeflie, S. A., Fammy, F. A., Ibrahim, Z., and Sabri, N. (2019). Evaluation of Basic Convolutional Neural Network, AlexNet and Bag of Features for Indoor Object Recognition. *International Journal of Machine Learning and Computing*, 9(6):801–806.

UIN SUNAN AMPEL
S U R A B A Y A