## KLASIFIKASI KANKER BERDASARKAN DATA RNA MENGGUNAKAN DILATED CONVOLUTIONAL NEURAL NETWORK

### **SKRIPSI**



# Disusun Oleh MOHAMMAD ANANG SYARIFUDIN H02217009

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

2021

### PERNYATAAN KEASLIAN

Saya yang bertanda tangan di bawah ini,

Nama

: Mohammad Anang Syarifudin

NIM

: H02217009

Program Studi : Matematika

Angkatan

: 2017

Menyatakan bahwa saya tidak melakukan plagiat dalam penulisan skripsi saya yang berjudul "Klasifikasi Kanker Berdasarkan Data RNA Menggunakan Dilated Convolutional Neural Network". 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, 23 Januari 2021

Yang menyatakan,

Mohammad Anang Syarifudin

NIM. H02217009

#### LEMBAR PERSETUJUAN PEMBIMBING

#### Skripsi oleh

Nama : Mohammad Anang Syarifudin

NIM : H02217009

Judul Skripsi : Klasifikasi Kanker Berdasarkan Data RNA Menggunakan

Dilated Convolutional Neural Network

telah diperiksa dan disetujui untuk diujikan.

Surabaya, 23 Januari 2021

Pembimbing I

Dian C. Rini Novitasari, M. Kom

NIP. 198511242014032001

Pembimbing II

Lutfi Hakim, M.Ag

NIP. 197312252006041001

Mengetahui,

Ketua Program Studi Matematika

UIN Sunan Ampel Surabaya

Ari Fanani, M.Kom

NIP. 198701272014031002

## PENGESAHAN TIM PENGUJI SKRIPSI

#### Skripsi oleh

Nama : Mohammad Anang Syarifudin

NIM : H02217009

Judul Skripsi : Klasifikasi Penyakit Kanker Berdasarkan Data RNA

Menggunakan Dilated Convolutional Neural Network

Telah dipertahankan di depan Tim Penguji pada tanggal 26 Januari 2021

> Mengesahkan, Tim Penguji

Penguji I

Yuniar Farida, MT

NIP. 197905272014032002

Penguji II

Nurissaidah Winnuna, M. Kom

NIP. 199011022014032004

Penguji III

Dian C. Rini Novitasari, M. Kom

NIP. 198511242014032001

Penguji IV

Lutfi Hakim, M.Ag

NIP. 197312252006041001

Mengetahui,

eRan Fakultas Sains dan Teknologi

UIN Sunan Ampel Surabaya

Rusydiyah, M.Ag

MR 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:

MOHAMMAD ANANG STARIFUDIN
: H02217009
: SAINTEK / MATEMATIKA
mohammadanang egmail. com
gan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan I Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:  Tesis  Desertasi  Lain-lain ()  KANKER BERDASARKAN DATA RNA MENGGUNAKAN
VOLUTIONAL NEURAL NETWORK
yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Non-

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, 22 FEBRUARI 2021

Penulis

(MOHAMMAD ANANG STARIFUDIN )

#### **ABSTRAK**

#### Klasifikasi Kanker Berdasarkan Data RNA Menggunakan Dilated

#### Convolutional Neural Network

Berdasarkan statistik dari WHO, kanker merupakan salah satu penyakit paling mengancam di dunia. Sebanyak 13% kematian di dunia disebabkan oleh kanker dan terus meningkat setiap tahunnya. Karsinoma merupakan kanker yang paling sering ditemukan pada manusia yaitu 90% dari seluruh kasus kanker di dunia. Salah satu parameter yang dapat digunakan dalam mendeteksi kanker karsinoma yaitu ekspresi RNA. Pada penelitian ini akan dilakukan deteksi kanker berdasarkan ekspresi RNA ke dalam 5 kelas yaitu Breast Carcinoma (BRCA), Kidney Renal clear cell Carcinoma (KIRC), Uterine Corpus Endometrial Carcinoma (UCEC), dan Lung Adenocarcinoma (LUAD) menggunakan metode Dilated Convolutional Neural Network. Beberapa percobaan dilakukan untuk mendapatkan hasil yang terbaik, diantaranya yaitu merubah nilai faktor dilasi dan membandingkan model yang menggunakan augmentasi data atau tanpa menggunakan augmentasi data. Hasil akurasi terbaik yang didapatkan yaitu 99.24% dengan sensitifitas sebesar 98.70%.

Kata kunci: Dilated CNN, Kanker, Deep Learning, RNA

#### **ABSTRACT**

## Classification of Cancer Based on RNA Using Dilated Convolutional Neural Network

Based on WHOs statistics, cancer is one of the most threatening diseases in the world. As many as 13% of deaths in the world are caused by cancer and it increase every year. Carcinoma cancer is most often found in human, which is 90% of all cases in the world. One of the parameters that can be used in detecting carcinoma cancer is the expression of RNA. In this study, cancer detection based on RNA expression will be carried out into 5 classes, there are Breast Carcinoma (BRCA), Kidney Renal clear cell Carcinoma (KIRC), Uterine Corpus Endometrial Carcinoma (UCEC), and Lung Adenocarcinoma (LUAD) using Dilated Convolutional Neural Network method. Several experiments were carried out to get best results, including changing the value of the dilation factor and comparing models using data augmentation or without using data augmentation. The best accuracy obtained is 99.24% and sensitivity about 98.70%.

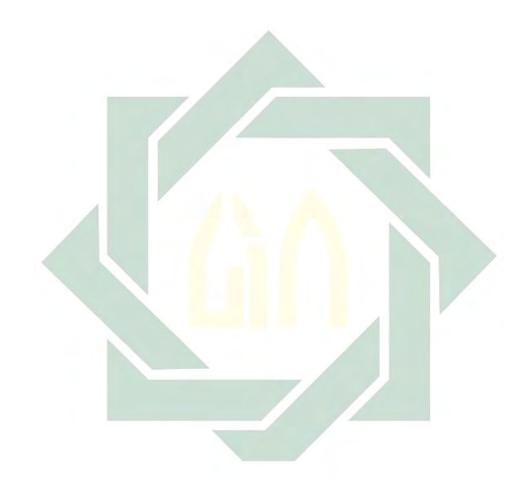
Keywords: Dilated CNN, Cancer, Deep Learning, RNA

## **DAFTAR ISI**

HALAMAN JUDUL
HALAMAN PERNYATAAN KEASLIAN
LEMBAR PERSETUJUAN PEMBIMBING ii
PENGESAHAN TIM PENGUJI SKRIPSI i
LEMBAR PERSETUJUAN PUBLIKASI
ABSTRAK
ABSTRACT vi
DAFTAR ISI
DAFTAR TABEL
DAFTAR GAMBAR
I PENDAHULUAN
1.1. Latar Belakang Masalah
1.2. Rumusan Masalah
1.3. Tujuan Penelitian
1.4. Manfaat Penelitian
1.5. Batasan Masalah
1.6. Sistematika Penulisan
II TINJAUAN PUSTAKA
2.1. Kanker
2.1.1. Kanker Paru-Paru
2.1.2. Kanker Payudara
2.1.3. Kanker Ginjal
2.1.4. Kanker Rahim
2.2. Ribonucleic Acid (RNA)
2.3. 2D Image Creation
2.4. Augmentasi Data
2.5. k-Fold Cross Validation

	2.6.	Kecerdasan Buatan	.7
	2.7.	Deep Learning	8
	2.8.	Convolutional Neural Network	8
		2.8.1. Convolution Layer	9
		2.8.2. Pooling Layer	24
		2.8.3. Activation Layer	25
		2.8.4. Softmax Layer	25
		2.8.5. Flatten	26
		2.8.6. Dropout Layer	27
		2.8.7. Fully Connected Layer	27
		2.8.8. Cross Entropy	28
	2.9.	Dilated Convolutional Neural Network	29
	2.10.	Matriks Konfusi	0
	2.11.	Integrasi Keilmuan	3
Ш	MET	ODE PENELITI <mark>an</mark>	86
	3.1.		86
	3.2.	Jenis Data	86
	3.3.	Teknik Analisis Data	86
IV	HAS	L DAN PEMBAHASAN 4	ŀ0
	4.1.	Preprocessing	10
	4.2.	Pelatihan Model	12
		4.2.1. Augmentasi Data	12
		4.2.2. Convolution Layer	19
		4.2.3. Activation Layer	53
		4.2.4. <i>Pooling Layer</i>	54
		4.2.5. <i>Dropout Layer</i>	55
		4.2.6. Fully Connected Layer	6
	4.3.	Pengujian dan Evaluasi Model	71
	4.4.	Integrasi Keilmuan	76
$\mathbf{V}$	PEN	THP 7	79

٨	FII T	FR CON	JVOI	רוד	TI (	N	T /	v	FR	•														9	25
D	AFTA:	R PUSTA	KA	•		•		•		•	•		•	•	•	 •	•	•	•	•	•	•	•	 8	80
	5.2.	Saran .																						 8	80
	5.1.	Kesimpu	ılan	•		•		•		•	•		•			 •	•	•	•		•	•	•	 •	79



## DAFTAR TABEL

2.1	Contoh Faktor Sekunder dan Persensate Penyebab Kanker	8
2.2	Fungsi Aktivasi	25
4.1	Hasil Klasifikasi Model Tanpa Augmentasi Data	71
4.2	Hasil Klasifikasi Model Dengan Augmentasi Data	72
4.3	Perbandingan Hasil sensitivitas Pada Masing-masing Kelas	76

## DAFTAR GAMBAR

2.1	Hasil Rontgen Dada Penderita Kanker Paru-Paru	9
2.2	Hasil Tes Mamografi Penderita Kanker Payudara	10
2.3	Hasil CT Scan Penderita Kanker Ginjal	11
2.4	Hasil Biopsi Penderita Kanker Rahim	12
2.5	Letak RNA	13
2.6	Pencerminan Terhadap Sumbu $x$	14
2.7	Pencerminan Terhadap Sumbu $y \ldots \ldots \ldots \ldots$	15
2.8	Rotasi Sebesar 45°	16
2.9	k-Fold Cross Validation	16
2.10	(a) Simple Neural Network; (b) Deep Learning	18
2.11	Contoh Input Pada Convolution Layer	19
2.12	Ilustrasi Convolution Layer Dengan Filter $5 \times 5 \dots \dots$	20
2.13	Contoh Padding	20
2.14	Ilustrasi Convolution Layer Dengan Stride 1	21
2.15	Arsitektur Convolutional 1D	23
2.16	Arsitektur Convolutional 2D	23
	Max Pooling dan Average Pooling	24
2.18	Contoh Flatten	26
2.19	Contoh Neural Network: (a) Tanpa Dropout; (b) Menggunakan	
	Dropout	27
2.20	Fully Connected Layer	28
2.21	(Kiri) Convolutional Neural Network dengan $d=1$ ; (Kanan)	
	Dilated Convolutional Neural Network dengan $d=2$	
	Matriks Konfusi	30
2.23	Matriks Konfusi Pada Permasalahan n Kelas	33
3.1	Diagram Alir Sistem Klasifikasi	37
3.2	Arsitektur Convolutional Neural Network	39

4.1	Ilustrasi 2D Image Creation	40
4.2	Hasil 2D Image Creation	41
4.3	Contoh Citra Input	43
4.4	Contoh Nilai Piksel Citra Input	43
4.5	Citra Hasil Refleksi Terhadap Sumbu $x \ \dots \dots \dots$	45
4.6	Citra Hasil Refleksi Terhadap Sumbu $y$	47
4.7	Citra Hasil Rotasi Sebesar 180°	49
4.8	Contoh Piksel dengan Padding	50
4.9	Hasil Convolution Layer	52
4.10	Hasil Activation Layer ReLU	53
4.11	Contoh Penerapan Average Pooling	54
4.12	Contoh Penerapan Flatten	55
4.13	Contoh Proses <i>Dropout</i>	56
4.14	Arsitektur Fully Connected Layer	56
4.15	(kiri) Model Tanpa Augmentasi Data; (kanan) Model Dengan	
	Augmentasi Data	72

#### **BABI**

#### **PENDAHULUAN**

#### 1.1. Latar Belakang Masalah

Kanker merupakan istilah umum yang digunakan untuk menggambarkan penyakit kelainan pertumbuhan pada sel (Khalifa et al., 2020). Berdasarkan statistik yang diberikan oleh *World Health Organization* (WHO), kanker menjadi penyebab 13% kematian di seluruh dunia dengan 8.2 juta jiwa meninggal setiap tahunnya (Zhang et al., 2017). WHO juga telah mengestimasikan bahwa pada tahun 2005-2015 sebanyak 84 juta orang meninggal karena kanker (Nugraha et al., 2019). Sedangkan pada tahun 2018, kanker telah menjadi penyebab kematian lebih dari 9 juta orang di seluruh dunia dengan kasus terbesar disebabkan oleh kanker paru-paru (>1.7 juta jiwa) (Bray et al., 2018). Berdasarkan data tersebut, kanker memiliki dampak yang cukup besar pada kehidupan manusia sesuai dengan firman Allah SWT pada Q.S. Al-Baqarah/2:155 mengenai cobaan yang diberikan Allah SWT kepada setiap hamba-Nya.

وَلَنَبُلُوَنَّكُمْ بِشَىءٍ مِّنَ الْخَوْفِ وَالْجُوْعِ وَنَقُصٍ مِّنَ الْاَمُوَالِ وَالْنَبُلُونَّكُمْ بِشَىء مِّنَ الْاَمُوالِ وَالْاَنْفُسِ وَالثَّمَرِتُ وَبَشِّرِ الصَّبِرِيْنَ

"Dan sungguh Kami akan menguji kalian dengan sedikit rasa takut, lapar dan kekurangan harta, jiwa dan buah-buahan, dan sampaikan berita gembira pada mereka yang sabar"

Dalam ayat tersebut dijelaskan bahwa Allah SWT akan menguji setiap hamba-Nya baik itu berupa kesedihan, kelaparan, kekurangan harta, maupun penyakit. Hal itu bertujuan untuk mengetahui kesabaran hamba-Nya.

Hingga saat ini, lebih dari 100 jenis kanker yang telah teridentifikasi (Zhang et al., 2017). Namun, secara umum kanker dapat dibagi menjadi 3 kelompok yaitu karsinoma, sarkoma, dan leukemia (limfoma). Dari kasus kanker yang ada di dunia, 90% dari kasus tersebut merupakan jenis kanker karsinoma (Khalifa et al., 2020). Cara umum yang dilakukan untuk mendeteksi kanker yaitu tes laboratorium (urin, darah, dan lain-lain), prosedur pencitraan (*CT scan, x-ray*, dan lain-lain), dan biopsi (berdasarkan jaringan pada tubuh). Beberapa cara tersebut membutuhkan biaya yang mahal dan proses yang lama karena dokter akan memeriksa pasien secara manual (Zhang et al., 2017).

Penelitian tentang kanker sudah banyak dilakukan untuk dapat mendeteksi ataupun mengklasifikasi kanker. Salah satu proyek penelitian mengenai kanker yaitu *The Cancer Genome Atlas* (TCGA) yang dikelola oleh *National Cancer Institute* (NCI) dan *National Human Genome Research Institute* (NHGRI). Data yang dihasilkan dari penelitian tersebut yaitu ekspresi gen, DNA metilasi, ekspresi protein, dan RNA (Khalifa et al., 2020). Data yang sering digunakan untuk klasifikasi kanker adalah RNA (Hsu and Si, 2018)(Kabir and Ludwig, 2019)(Lyu and Haque, 2018) karena pada penelitian yang dilakukan oleh (Kresno, 2011), RNA dapat berfungsi sebagai onkogen yaitu sel yang dapat mempercepat pertumbuhan kanker maupun *tumor suppressor gene* yaitu sel yang dapat menghambat pertumbuhan kanker, sehingga apabila terjadi kelainan pada RNA dapat digunakan sebagai acuan dalam mendiagnosa kanker. Akan tetapi, data yang didapat susah dibaca secara manual karena banyaknya parameter yang tersedia

sehingga dibutuhkan metode yang cocok untuk melakukan pembelajaran data dalam skala besar (Kurniawan and Ivandari, 2017).

Penelitian mengenai klasifikasi kanker berdasarkan data RNA juga sudah banyak dilakukan, penelitian oleh (Lyu and Haque, 2018) menggunakan salah satu metode deep learning yaitu Convolutional Neural Network mendapatkan hasil yang baik dengan akurasi sebesar 95.59% dan penelitian lain yang dilakukan oleh (Khalifa et al., 2020) menggunakan metode Convolutional Neural Network dan Binary Particle Swarm Optimization with Decision Tree (BPSO-DT) sebagai metode seleksi fitur mendapatkan hasil akurasi sebesar 96.90%, akan tetapi kedua penelitian tersebut membutuhkan proses komputasi yang cukup berat. Oleh karena itu, pada penelitian ini digunakan metode *Dilated Convolutional Neural Network* sebagai metode klasifikasi. Metode tersebut dipilih karena pada penelitian sebelumnya yang dilakukan oleh (Lei et al., 2019) (Kudo and Aoki, 2017) mendapatkan hasil bahwa metode Dilated Convolutional Neural Network mendapatkan hasil lebih baik dan waktu lebih singkat dalam proses klasifikasi citra dibandingkan dengan metode Convolutional Neural Network. Pada penelitian lain yang dilakukan oleh (Zhen et al., 2019), metode Dilated Convolutional Neural Network tidak hanya baik digunakan pada data 2 dimensi tetapi juga mendapatkan hasil yang baik ketika digunakan pada data 1 dimensi dibandingkan dengan metode lainnya seperti Long Short Term Memory (LSTM), Gated Recurrent Unit (GRU), dan lain-lain. Metode Dilated Convolutional Neural Network merupakan bagian dari deep learning, oleh karena itu proses augmentasi data akan digunakan untuk menghindari overfitting (Shorten and Khoshgoftaar, 2019). Selain itu, dari penelitian yang dilakukan oleh (Al-Dhabyani et al., 2019) pada klasifikasi kanker payudara berdasarkan hasil USG menggunakan deep learning diperoleh hasil

bahwa augmentasi data dapat berpengaruh pada akurasi klasifikasi.

Berdasarkan latar belakang yang telah dijelaskan, pada penelitian akan dilakukan klasifikasi kanker berdasarkan data RNA menggunakan metode *Dilated Convolutional Neural Network* dengan beberapa percobaan yang diharapkan bisa mendapatkan hasil yang lebih baik.

#### 1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah dijelaskan, maka peneliti merumuskan permasalahan yang akan diteliti yaitu:

- 1. Bagaimana proses klasifikasi kanker berdasarkan data RNA menggunakan metode *Dilated Convolutional Neural Network*?
- 2. Bagaimana hasil dari metode *Dilated Convolutional Neural Network* dibandingkan metode *Convolutional Neural Network* dalam klasifikasi kanker berdasarkan data RNA?
- 3. Bagaimana hasil dari model menggunakan augmentasi data dibandingkan dengan model tanpa augmentasi data dalam sistem klasifikasi kanker berdasarkan data RNA menggunakan metode *Dilated Convolutional Neural Network*?

#### 1.3. Tujuan Penelitian

Berdasarkan latar belakang yang telah dijelaskan, maka tujuan dari penelitian ini yaitu:

 Mengetahui bagaimana proses klasifikasi kanker berdasarkan data RNA menggunakan metode *Dilated Convolutional Neural Network*.

- 2. Mengetahui bagaimana hasil dari metode *Dilated Convolutional Neural*Network dibandingkan metode Convolutional Neural Network dalam klasifikasi kanker berdasarkan data RNA.
- 3. Mengetahui dan menganalisis bagaimana hasil dari model menggunakan augmentasi data dibandingkan dengan model tanpa augmentasi data dalam sistem klasifikasi kanker berdasarkan data RNA menggunakan metode Dilated Convolutional Neural Network.

#### 1.4. Manfaat Penelitian

Hasil dari penelitian yang dilakukan, diharapkan memiliki manfaat sebagai berikut:

- 1. Manfaat Teoritis
  - a. Bagi Universitas Islam Negeri Sunan Ampel Surabaya
     Dapat menambah koleksi literatur bagi Universitas Islam Negeri
     Sunan Ampel Surabaya dan program studi Matematika Fakultas
     Sains dan Teknologi.
  - b. Bagi Penulis

Dapat menambah wawasan dan memperdalam ilmu pengetahuan mengenai metode-metode yang dapat digunakan sebagai klasifikasi. Serta menambah pengalaman penulis dalam menghadapi suatu permasalahan dan bagaimana mendapatkan solusinya.

c. Bagi Pembaca

Dapat menambah pengetahuan tentang kanker dan metode yang

digunakan dalam proses klasifikasi serta dapat digunakan sebagai acuan peneliti lainnya.

#### 2. Manfaat Praktis

Dapat menjadi acuan bagi dunia medis dalam klasifikasi jenis kanker berdasarkan data RNA menggunakan salah satu metode *deep learning* yaitu *dilated convolutional neural network*.

#### 1.5. Batasan Masalah

Ruang lingkup permasalahan terkait dengan penulisan penelitian ini yaitu:

- 1. Data yang digunakan merupakan data RNA yang telah dipublikasikan secara umum pada Mendeley DAta (Ferles et al., 2018).
- 2. Kanker yang diklasifikasikan pada penelitian ini yaitu BRCA (kanker payudara), KIRC (kanker ginjal), UCEC (kanker rahim), dan LUAD (kanker paru-paru).

#### 1.6. Sistematika Penulisan

Dalam penulisan penelitian ini, penulis menggunakan sistematika penulisan yang terdiri dari 5 bab, yaitu:

#### 1. Pendahuluan

Pada bab ini terdiri dari latar belakang, rumusan masalah, tujuan penelitian, manfaat penelitian, batasan masalah dan sistematika penulisan.

#### 2. Tinjauan Pustaka

Pada bab ini terdiri dari tinjauan pustaka mengenai konsep pembahasan yang menjadi masalah yang diangkat yaitu kanker, RNA, augmentasi data, convolutional neural network, dilated convolutional neural network, matriks konfusi.

#### 3. Metode Penelitian

Pada bab ini terdiri dari deskripsi data dan tahap-tahap pada penelitian.

#### 4. Hasil dan Pembahasan

Pada bab ini terdiri dari penjelasan mengenai praproses data serta hasil klasifikasi dari dilated convolutional neural network.

#### 5. Penutup

Pada bab ini terdiri atas kesimpulan dari permasalahan yang ada dan memuat saran-saran yang berkaitan dengan hasil penelitian.

#### **BAB II**

#### TINJAUAN PUSTAKA

#### 2.1. Kanker

Kanker merupakan istilah yang digunakan untuk berbagai penyakit tumor ganas. Tumor merupakan pembengkakan yang diakibatkan oleh pembelahan sel, baik itu kanker (tumor ganas) atau bukan kanker (tumor jinak). Faktor dari terjadinya kanker pada manusia dibagi menjadi dua yaitu faktor primer (keturunan) dan faktor sekunder seperti yang dijelaskan pada Tabel 2.1 (De Jong, 2005).

Tabel 2.1 Contoh Faktor Sekunder dan Persensate Penyebab Kanker

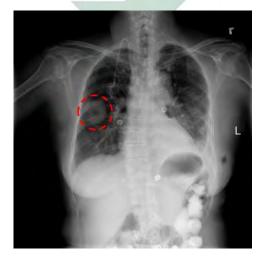
Faktor Sekun <mark>de</mark> r	Persentase Penyebab Kanker (%)
Gizi	35
Merokok	30
Infeksi	10
Faktor hormonal	6
Pekerjaan	4
Alkohol	3
Radiasi sinar	3
Pengotoran lingkungan	2
Obat-obatan/tindakan medis	1
Kebiasaan seksual	1
Produk industri	< 1
Makanan	< 1

Kanker merupakan penyakit paling mengancam di dunia. Berdasarkan statistik yang diberikan oleh *World Health Organization* (WHO), kanker menyumbang 13% kematian di seluruh dunia dengan 8.2 juta jiwa meninggal setiap tahunnya (Zhang et al., 2017). Pada tahun 2018, kanker telah menjadi penyebab kematian lebih dari 9 juta orang di seluruh dunia dengan kasus terbesar

disebabkan oleh kanker paru-paru (>1.7 juta jiwa). Beberapa jenis kanker yang menjadi penyebab kematian terbanyak yaitu kanker paru-paru (1,761,007 juta jiwa), kanker payudara (626,679 juta jiwa), kanker ginjal (175,098 juta jiwa), dan kanker rahim (89,929 juta jiwa) (Bray et al., 2018). Cara yang sudah berkembang sejak lama untuk mendeteksi kanker yaitu tes laboratorium (urin, darah, dan lain-lain), prosedur pencitraan (*CT scan*, *X-ray*, dan lain-lain), dan biopsi (berdasarkan jaringan pada tubuh) (Zhang et al., 2017).

#### 2.1.1. Kanker Paru-Paru

Kanker paru-paru adalah bertumbuhnya sel yang tidak terkendali dalam jaringan paru-paru yang dapat disebabkan oleh sejumlah karsinogen lingkungan. Penyebab utama dari penyakit kanker paru-paru adalah asap rokok, namun ada sebagian kanker paru-paru disebabkan oleh sel dari dalam kanker maupun dari kanker di bagian tubuh lain yang menjalar ke paru-paru. Kanker paru-paru umumnya dideteksi berdasarkan hasil dari foto rontgen dada pasien (Suryo, 2010). Contoh hasil rontgen dada dari pasien kanker paru-paru dapat dilihat pada Gambar 2.1.



Gambar 2.1 Hasil Rontgen Dada Penderita Kanker Paru-Paru

(Sumber: www.nih.gov)

#### 2.1.2. Kanker Payudara

Kanker payudara merupakan tumbuhnya sel ganas pada jaringan payudara yang dapat berasal dari epitel duktus maupun lobulus (Nasional, 2015). Kanker ini bisa mulai tumbuh di dalam kelenjar payudara, saluran payudara, jaringan lemak maupun jaringan ikat pada payudara. Gejala awal yang dialami penderita umumnya berupa benjolan yang berbeda dari sekitarnya yang biasanya menimbulkan nyeri dan memiliki pinggiran yang tidak teratur (Mufida et al., 2015). Tes yang umum digunakan sebagai acuan diagnosa yaitu mamografi. Contoh hasil tes mamografi dari penderita kanker payudara dapat dilihat pada Gambar 2.2.



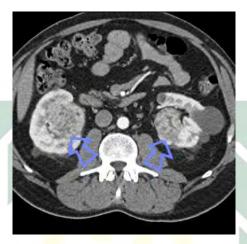
Gambar 2.2 Hasil Tes Mamografi Penderita Kanker Payudara

(Sumber: www.breastcancer.org)

#### 2.1.3. Kanker Ginjal

Kanker ginjal merupakan bertumbuhnya sel ganas yang tidak terkendali pada ginjal. Kanker ini dapat berasal dari tumor primer di ginjal maupun tumor sekunder yang berasal dari organ lain (Melisa et al., 2016). Gejala umum yang dialami oleh penderita yaitu nyeri pinggang, air seni atau *urine* menjadi kemerahan

(gross hematuria, dan benjolan pada daerah punggung dekat ginjal (Made et al., 2018). Tes umum yang dilakukan untuk mendiagnosis penderita yaitu tes darah, tes *urine*, dan *CT Scan* atau MRI (Hamid et al., 2019). Contoh hasil *CT scan* dari penderita kanker ginjal dapat dilihat pada Gambar 2.3.

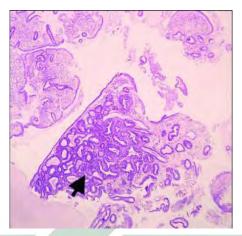


Gambar 2.3 Hasil CT Scan Penderita Kanker Ginjal

(Sumber: medicineworld.org)

#### 2.1.4. Kanker Rahim

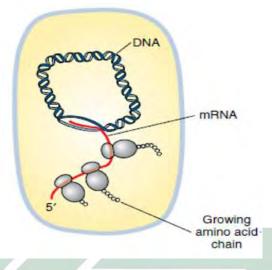
Kanker rahim merupakan tumor yang berkembang di area sekitar rahim. Kanker ini dikenal sebagai tumor ganas dengan prognosis buruk. Beberapa faktor penyebab berkembangnya kanker rahim adalah *menopause*, obesitas, dan diabetes. Gejala yang dialami oleh penderita umumnya adalah nyeri pada panggul, nyeri saat buang air kecil, dan pendarahan abnormal. Salah satu cara untuk mendiagnosa penderita yaitu dengan melakukan biopsi atau pengambilan sampel jaringan untuk diteliti (Rajanbabu, 2015). Contoh jaringan yang diambil dari penderita kanker rahim dapat dilihat pada Gambar 2.4.



Gambar 2.4 Hasil Biopsi Penderita Kanker Rahim
(Sumber: www.researchgate.net)

#### 2.2. Ribonucleic Acid (RNA)

Ribonucleic Acid (RNA) atau asam ribonukleat merupakan makromolekul yang berfungsi sebagai penyimpan dan penyalur informasi genetik. Dengan kata lain, RNA dapat mengkonversi informasi genetik dari gen (DNA) ke dalam untaian asam amino protein (Sumbono, 2019) seperti pada Gambar 2.5. Satuan yang dapat digunakan untuk menggambarkan kecepatan RNA dalam membaca gen yang ada pada DNA yaitu Reads Per Kilobase Million (RPKM). Selain itu, RNA juga dapat berfungsi sebagai onkogen maupun tumor suppressor gene. Onkogen merupakan gen yang dapat meningkatkan keganasan sel tumor, sedangkan tumor suppressor gene yaitu gen yang bisa mencegah perkembangan tumor. Terjadinya perubahan ekspresi gen pada RNA dapat menjadi indikasi bahwa terdapat kanker pada tubuh manusia (Kresno, 2011). Terdapat beberapa teknik dalam pengambilan RNA dimana Masing-masing teknik memiliki ketentuan dan hasil susunan RNA yang berbeda-beda (Salser, 1974). Pada penelitian ini digunakan data RNA yang diambil menggunakan alat yang bernama Illumina Hiseq yang disusun menggunakan metode pada alat tersebut (Clark et al., 2019).



Gambar 2.5 Letak RNA

(Sumber: Griffiths et al., 2000)

#### 2.3. 2D Image Creation

2D image creation merupakan teknik yang dilakukan untuk mengubah data vektor dengan skala [a,b] menjadi data citra dengan ukuran  $m \times n$  dengan skala [0,255]. Apabila a merupakan nilai minimal dari data vektor, b merupakan nilai maksimal dari data vektor, dan x merupakan nilai input, maka perubahan skala dapat dilakukan menggunakan Persamaan 2.1 (Khalifa et al., 2020).

$$PixelValue(x) = round\left(\frac{(x)(255)}{b-a}\right)$$
 (2.1)

#### 2.4. Augmentasi Data

Augmentasi data merupakan teknik yang dilakukan untuk memperbanyak set data yang dimiliki. Tujuan dilakukannya augmentasi data adalah untuk mengatasi ketidakseimbangan data sehingga model yang dibangun tidak terlalu condong kearah kelas yang menjadi mayoritas (Shorten and Khoshgoftaar, 2019). Metode yang umum digunakan sebagai teknik augmentasi data adalah modifikasi

warna dan transformasi geometrik (Mikolajczyk and Grochowski, 2018). Teknik modifikasi warna berhubungan dengan nilai piksel dari citra, sedangkan transformasi geometrik berhubungan dengan posisi piksel dari citra. Berikut merupakan beberapa teknik transformasi geometrik dasar yang umum digunakan pada augmentasi data.

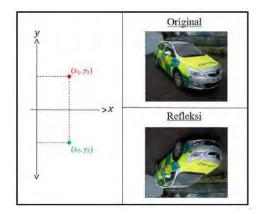
#### 1. Refleksi (Pencerminan)

Refleksi atau pencerminan merupakan perubahan posisi dari nilai piksel berdasarkan koordinat awal  $(x_1,y_1)$  menuju koordinat akhir  $(x_2,y_2)$  pada citra sesuai dengan posisi pencerminan, dimana terdapat dua jenis posisi pencerminan yang umum diaplikasikan yaitu pencerminan terhadap sumbu x dan pencerminan terhadap sumbu y.

#### a. Pencerminan Terhadap Sumbu x

Pencerminan terhadap sumbu x dapat dilihat pada Gambar 2.6 dan dilakukan menggunakan Persamaan 2.2.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$
 (2.2)



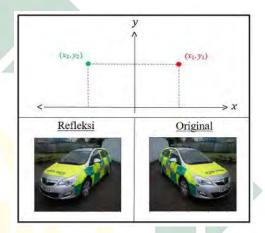
Gambar 2.6 Pencerminan Terhadap Sumbu  $\boldsymbol{x}$ 

(Sumber: www.analyticsvidhya.com)

#### b. Pencerminan Terhadap Sumbu y

Pencerminan terhadap sumbu y dapat dilihat pada Gambar 2.7 dan dilakukan menggunakan Persamaan 2.3.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$
 (2.3)



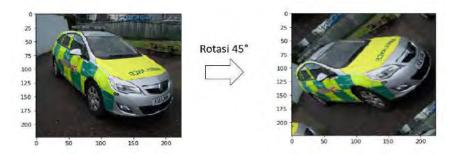
Gambar 2.7 Pencerminan Terhadap Sumbu y

(Sumber: www.analyticsvidhya.com)

#### 2. Rotasi

Rotasi merupakan salah satu transformasi geometrik dimana nilai-nilai piksel dari citra akan mengalami perubahan posisi dari yang didasarkan pada nilai variabel rotasi sebesar  $\theta$ ° terhadap posisi titik pusat rotasi. Rotasi dilakukan menggunakan Persamaan 2.4 dan dapat dilihat pada Gambar 2.8.

$$\begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos \theta^{\circ} & -\sin \theta^{\circ} \\ \sin \theta^{\circ} & \cos \theta^{\circ} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$
 (2.4)

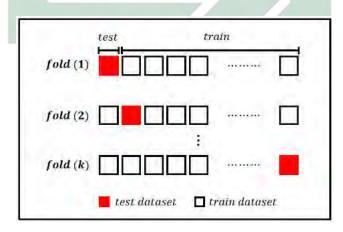


Gambar 2.8 Rotasi Sebesar  $45^{\circ}$ 

(Sumber: www.analyticsvidhya.com)

#### 2.5. k-Fold Cross Validation

k-Fold Cross Validation merupakan sebuah teknik yang digunakan untuk mengukur performa suatu model yang dibangun. Pengukuran dilakukan dengan mengambil sample secara acak untuk digunakan sebagai set data pengujian. Data yang dimiliki akan di partisi sebanyak k dengan ukuran yang kira-kira sama. Model kemudian dilatih menggunakan partisi sebanyak k-1 dan pengujian dilakukan dengan sisa partisi (Berrar, 2019). Ilustrasi dari teknik ini dapat dilihat pada Gambar 2.9.



Gambar 2.9 k-Fold Cross Validation

(Sumber: Berrar, 2019)

#### 2.6. Kecerdasan Buatan

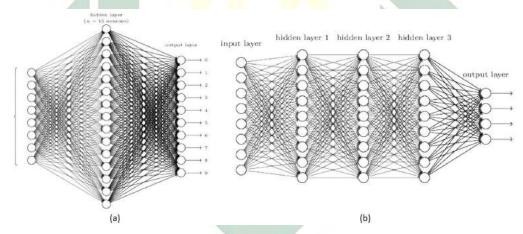
Kecerdasan buatan atau bisa juga disebut *Artificial Intelligence* (AI) merupakan pemodelan suatu sistem yang dapat mempelajari masalah layaknya jaringan saraf otak manusia. Beberapa metode yang sedang dikembangkan dalam kecerdasan buatan adalah *Fuzzy Logic* (Logika Fuzzy), *Evolutionary Computing* (Skema Evolusi), dan *Machine Learning* (Pembelajaran Mesin) (Ahmad, 2017).

Machine Learning atau pembelajaran mesin merupakan model pendekatan suatu sistem agar bisa bekerja semirip mungkin dengan jaringan saraf otak manusia. Metode ini menjadi paling populer karena mampu mempelajari dan menggeneralisasi suatu masalah layaknya otak manusia. Beberapa pengaplikasian metode ini yaitu prediksi dan klasifikasi. Ciri khas dari machine learning yaitu adanya proses pembelajaran/pelatihan dan pengujian. Salah satu metode dalam machine learning yang sering digunakan adalah Neural Network atau Jaringan Saraf Tiruan.

Neural network merupakan pemodelan suatu sistem menyerupai karakteristik jaringan saraf biologis yang dimiliki otak manusia agar dapat memproses informasi dengan baik (Wuryandari and Afrianto, 2012). Model ini terdiri dari satu atau beberapa lapisan tersembunyi yang didalamnya terdapat beberapa unit neuron. Semakin banyak lapisan tersembunyi maka akan semakin baik kemampuan dari neural network tersebut (Ahmad, 2017). Metode dalam neural network yang menggunakan banyak lapisan tersembunyi adalah Deep Learning atau pembelajaran mendalam.

#### 2.7. Deep Learning

Deep Learning merupakan bagian dari machine learning yang memiliki banyak lapisan sehingga membentuk tumpukan. Lapisan-lapisan tersebut merupakan sebuah algoritma atau metode yang digunakan untuk memproses data sehingga mendapatkan hasil sebaik mungkin (Nurfita and Ariyanto, 2018). Metode ini cocok digunakan dalam mengatasi permasalahan yang kompleks atau permasalahan yang menggunakan data dalam skala besar. Akibatnya, proses pelatihan dalam metode deep learning akan membutuhkan waktu yang lama. Perbedaan deep learning neural network dengan simple neural network dapat dilihat pada Gambar 2.10.



Gambar 2.10 (a) Simple Neural Network; (b) Deep Learning

(Sumber: Nielsen, 2015)

#### 2.8. Convolutional Neural Network

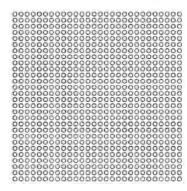
Convolutional Neural Network (CNN) adalah bagian dari deep learning yang merupakan pengembangan dari Multilayer Perceptron (MLP). Convolutional neural network terinspirasi dari jaringan saraf tiruan manusia yang dilakukan oleh Hubel dan Wiesel dalam penelitian visual cortex pada indera penglihatan kucing. Pada convolutional neural network terdapat suatu lapisan yang memiliki susunan

3D (lebar, tinggi, kedalaman) dimana lebar dan tinggi yang merupakan ukuran lapisan sedangkan kedalaman mengacu pada jumlah lapisan. Menurut jenis lapisan, secara umum convolutional neural network dibedakan menjadi 2 yaitu feature learning layer dan fully connected layer (Zufar and Setiyono, 2016). Berikut merupakan penjelasan dari masing-masing lapisan yang terdapat pada convolutional neural network.

#### 2.8.1. Convolution Layer

Konvolusi didefinisikan sebagai operasi pada dua fungsi. Fungsi yang pertama merupakan nilai *input* (misalnya nilai piksel) fungsi kedua adalah *filter* (kernel atau bobot). Operasi yang dilakukan yaitu perkalian titik antara dua fungsi sehingga menghasilkan matriks baru yang disebut *feature map*. *Filter* kemudian digeser ke posisi berikutnya sebesar *s*, besarnya pergeseran *filter* disebut sebagai *stride*. Perhitungan dilakukan hingga seluruh nilai *input* tertutup oleh *filter* (Ker et al., 2018).

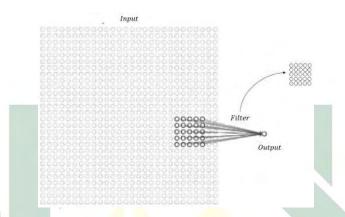
Convolution layer dapat dibagi menjadi 3 bagian yaitu nilai masukan (input), bobot (filter), dan nilai keluaran (feature map). Misal akan digunakan nilai input berupa citra berukuran  $28 \times 28$  seperti pada Gambar 2.11.



Gambar 2.11 Contoh Input Pada Convolution Layer

(Sumber: Nielsen, 2015)

Selanjutnya, beberapa nilai *input* akan dihubungkan dengan *feature map*. Nilai yang dihubungkan hanya sebagian bergantung dengan ukuran dari *filter* yang ditentukan. Misalnya digunakan *filter* dengan ukuran  $5 \times 5$ , sehingga didapatkan ilustrasi seperti pada Gambar 2.12.



Gambar 2.12 Ilustrasi *Convolution Layer* Dengan *Filter*  $5 \times 5$ 

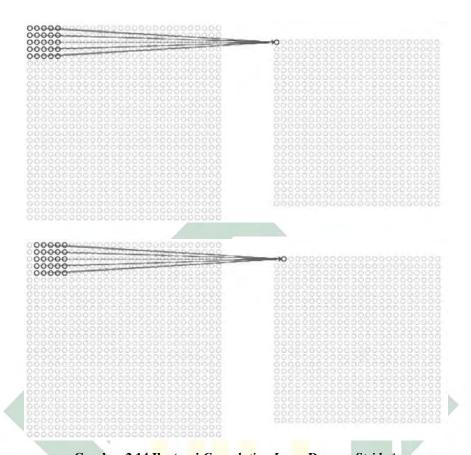
(Sumber: Nielsen, 2015)

Seperti yang sudah dijelaskan sebelumnya, *filter* akan digeser sebesar *stride* dan dimulai dari nilai *output* yang pertama. Misalnya digunakan *stride* sebesar 1, maka didapatkan ilustrasi seperti pada Gambar 2.14. Selain itu terdapat istilah *padding* yang merupakan bentuk manipulasi pada citra *input* dengan menambahkan piksel berisi nilai 0 di setiap tepi citra seperti pada Gambar 2.13.

0	0	0	0	0	0	0
0	$\mathbb{H}$					0
0	ĦŞ	15	l p	ξġ		0
0	E					0
0	+					0
0						0
0	0	0	0	0	0	0

Gambar 2.13 Contoh Padding

(Sumber: medium.com)



Gambar 2.14 Ilustrasi Convolution Layer Dengan Stride 1

(Sumber: Nielsen, 2015)

Apabila digunakan nilai *input* berukuran  $28 \times 28$ , *filter* berukuran  $5 \times 5$ , *padding* sebesar 0 (*no padding*), dan *stride* sebesar 1, maka akan didapat nilai *output* berukuran  $24 \times 24$ . Hal itu dikarenakan *filter* hanya bisa bergeser sebanyak 23 kali (Nielsen, 2015). Perhitungan ukuran *output* dapat dilakukan menggunakan Persamaan 2.5.

$$u = \frac{m - (d+f) + 2p}{s} + 1 \tag{2.5}$$

Dimana u merupakan ukuran output, d merupakan faktor dilasi, m merupakan ukuran input, f merupakan ukuran filter, p merupakan ukuran padding, dan s merupakan stride.

Pada proses *convolution*, nilai *output* didapatkan dari perkalian nilai *input* dengan *filter*. Operasi pada proses *convolution* ini dapat dinyatakan pada Persamaan 2.6 (Nielsen, 2015).

$$z_i = \sum_{a=0}^{f-1} w_a i_{i+da} \quad , d > 0$$
 (2.6)

Dimana **Z** merupakan nilai *output*, **I** merupakan matriks *input*, **W** merupakan matriks *filter*, i merupakan perhitungan ke-i, f merupakan ukuran *filter* yang digunakan, dan d merupakan faktor dilasi. *Convolution* pada umumnya menggunakan d=1 sehingga didapatkan Persamaan 2.7.

$$z_i = \sum_{a=0}^{f-1} w_a i_{i+a} \tag{2.7}$$

Persamaan di atas diasumsikan sebagai persamaan operasi konvolusi untuk satu dimensi, sehingga untuk operasi konvolusi dua dimensi dinyatakan dengan data input  $\mathbf{I}_{(m,n)}$  dan filter  $\mathbf{W}_{(a,b)}$  sehingga didapatkan Persamaan 2.8 (Nielsen, 2015).

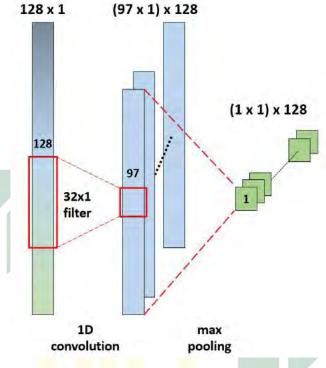
$$z_{(m,n)} = \sum_{a=0}^{f-1} \sum_{b=0}^{f-1} w_{(a,b)} i_{(m+da,n+db)} , d > 0$$
 (2.8)

Metode *convolutional neural network* dapat digunakan untuk memproses data 1 dimensi maupun 2 dimensi. Akan tetapi dalam penggunaannya, metode CNN 2D mendapatkan hasil lebih baik dibandingkan dengan metode CNN 1D (Zhao et al., 2018) (Chong et al., 2018).

#### 1. Convolutional 1 Dimension (CNN 1D)

Convolutional 1 dimension digunakan pada data yang berbentuk vektor seperti sinyal, time series, dan lain-lain (Kim, 2014). Arsitektur dari convolutional 1 dimension dapat dilihat pada Gambar 2.15. Perhitungan

nilai *output* pada *convolutional 1 dimension* dapat dilakukan menggunakan Persamaan 2.7.

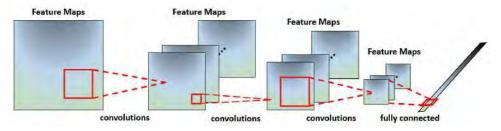


Gambar 2.15 Arsitektur Convolutional 1D

(Sumber: Novitasari et al., 2020)

## 2. Convolutional 2 Dimension (CNN 2D)

Convolutional 2 dimension digunakan pada data yang berbentuk dua dimensi (matriks), misalnya citra dan lain-lain. Arsitektur dari convolutional 2 dimension dapat dilihat pada Gambar 2.16. Perhitungan nilai output pada convolutional 2 dimension dapat dilakukan menggunakan Persamaan 2.8.

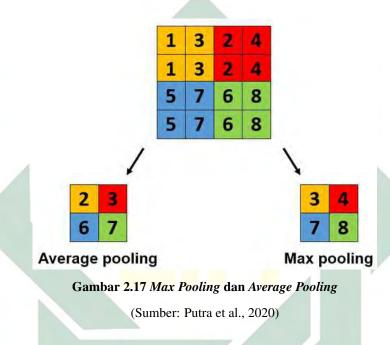


Gambar 2.16 Arsitektur Convolutional 2D

(Sumber: Novitasari et al., 2020)

## 2.8.2. Pooling Layer

Pooling layer atau subsampling merupakan pengurangan ukuran matriks. Ada dua macam pooling layer yakni max pooling dan average pooling (Arrofiqoh and Harintaka, 2018). Pada Gambar 2.17 ditunjukkan proses pooling dari input berukuran  $4 \times 4$  dengan ukuran filter  $2 \times 2$  dan stride sebesar 2.



Perhitungan ukuran *output* dari *pooling layer* dapat menggunakan Persamaan 2.9 (Dumoulin and Visin, 2016).

$$u = \frac{m-k}{s} + 1 \tag{2.9}$$

Dimana u merupakan ukuran output, m merupakan ukuran input, k merupakan ukuran filter, dan s merupakan ukuran stride.

## 2.8.3. Activation Layer

Activation layer merupakan lapisan yang berisi beberapa fungsi aktivasi yang berfungsi sebagai penentu apakah nilai dari jaringan sebelumnya diaktifkan atau tidak (Suhardi, 2007). Beberapa fungsi aktivasi yang sering digunakan dapat dilihat Tabel 2.2 atau Persamaan 2.10 hingga 2.15 (Munir et al., 2019).

Tabel 2.2 Fungsi Aktivasi

Fungsi Aktivasi	Persamaan	
Sigmoid	$y = \frac{1}{1 + e^{-x}}$	(2.10)
	$\frac{\partial y}{\partial x} = y[1 - y]$	(2.11)
Linier	y = x	(2.12)
	$\frac{\partial y}{\partial x} = 1$	(2.13)
ReLU	y = max(0, x)	(2.14)
	$\frac{\partial y}{\partial x} = \begin{cases} 0, & x \le 0 \\ 1, & x > 0 \end{cases}$	(2.15)

## 2.8.4. Softmax Layer

Fungsi *softmax* merupakan salah satu jenis fungsi aktivasi lain yang digunakan dalam arsitektur *convolutional neural network*. Fungsi ini digunakan

untuk menghitung probabilitas dari suatu vektor yang menghasilkan keluaran berupa nilai dalam rentang antara 0 dan 1, dengan jumlah probabilitas sama dengan 1. Fungsi softmax dapat dihitung menggunakan Persamaan 2.16 (Nwankpa et al., 2018).

$$y_i = \frac{e^{x_i}}{\sum_{k=1}^c e^{x_k}} \tag{2.16}$$

$$y_{i} = \frac{e^{x_{i}}}{\sum_{k=1}^{c} e^{x_{k}}}$$

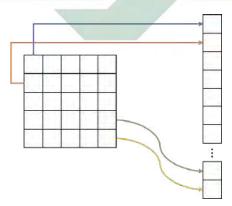
$$\frac{\partial y_{i}}{\partial x_{j}} = \begin{cases} y_{i}(1-y_{i}) &, i=j\\ -(y_{i})(y_{j}) &, i\neq j \end{cases}$$

$$(2.16)$$

Dimana x merupakan nilai dari vektor *input* dan c merupakan panjang dari vektor input.

# 2.8.5. *Flatten*

Flatten merupakan teknik yang digunakan untuk merubah hasil keluaran multidimensi dari proses ekstraksi fitur menjadi vektor. Proses tersebut dilakukan karena proses klasifikasi pada fully connected layer membutuhkan input berupa vektor (Ng et al., 2019). Ilustrasi dari flatten dapat dilihat pada Gambar 2.18.

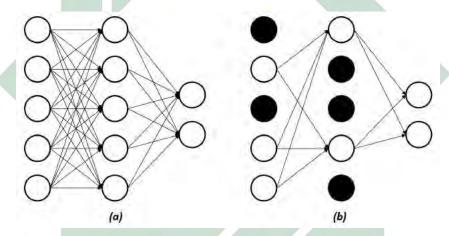


Gambar 2.18 Contoh Flatten

(Sumber: Guillemot et al., 2020)

## 2.8.6. Dropout Layer

Dropout pertama kali diusulkan pada tahun 2012 yang merupakan suatu teknik untuk menghindari overfitting. Teknik ini sangat mudah diimplementasikan pada model convolutional neural network dan akan berdampak pada performa model dalam proses pelatihan (Hinton et al., 2012). Selama proses iterasi pada tahap pelatihan, beberapa neuron akan dihilangkan dari jaringan dengan probabilitas sebesar p. Pemilihan neuron yang akan dihilangkan bersifat acak atau random. Probabilitas yang digunakan umumnya bernilai 0.5 (Labach et al., 2019). Ilustrasi dari proses dropout dapat dilihat pada Gambar 2.19.

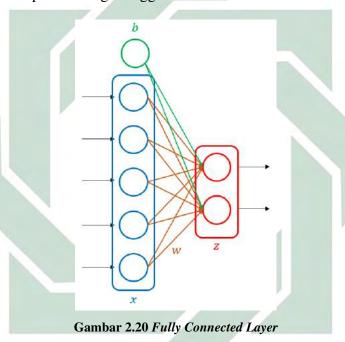


Gambar 2.19 Contoh *Neural Network*: (a) Tanpa *Dropout*; (b) Menggunakan *Dropout* (Sumber: Labach et al., 2019)

## 2.8.7. Fully Connected Layer

Fully connected layer terdiri dari beberapa lapisan yang didalamnya tersusun atas beberapa node/neuron yang terkoneksi penuh dengan layer sebelumnya. Metode yang digunakan pada lapisan ini yaitu multilayer perceptron. Lapisan ini berfungsi mengolah nilai yang didapatkan dari proses sebelumnya sehingga didapatkan hasil sesuai dengan yang diharapkan. Pada intinya, metode yang digunakan pada fully connected layer akan melihat kombinasi fitur-fitur yang

didapat dari lapisan sebelumnya untuk dipelajari, sehingga akan menunjukkan bahwa rangkaian fitur-fitur tersebut termasuk ke dalam kelas tertentu. Umumnya, metode *multilayer perceptron* yang digunakan pada *fully connected layer* yaitu *backpropagation* (Ker et al., 2018) dimana algoritma tersebut terdiri dua fase yaitu propagasi maju dan propagasi mundur (Rumelhart et al., 1986). Ilustrasi *fully connected layer* dapat dilihat pada Gambar 2.20 dan nilai keluaran pada *fully connected layer* dapat dihitung menggunakan Persamaan 2.18.



$$z_i = b_i + \sum_{j=1}^{c} x_j w_{ij}$$
 (2.18)

Dimana  $\mathbf{X}$  merupakan nilai input,  $\mathbf{Z}$  merupakan nilai keluaran,  $\mathbf{b}$  merupakan bias, c merupakan panjang dari vektor input, dan  $\mathbf{V}$  merupakan bobot yang digunakan.

#### 2.8.8. Cross Entropy

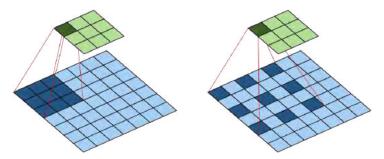
Cross entropy merupakan suatu fungsi yang digunakan untuk menyesuaikan bobot pada neural network untuk membuat model yang lebih baik.

Fungsi ini akan menghitung tingkat kesalahan antara hasil dari jaringan *neural network* dengan target yang diinginkan. Hasil tersebut kemudian akan dijadikan sebagai acuan seberapa besar bobot dari jaringan harus diperbarui. Jaringan akan terus memperbarui bobot hingga mendapatkan hasil keluaran dengan tingkat kesalahan terendah (Ho and Wookey, 2020). Apabila hasil keluaran sebanyak *h* dari jaringan *neural network* disimbolkan sebagai **y** dan target disimbolkan dengan **t**, maka fungsi *cross entropy J* dapat ditulis seperti pada Persamaan 2.19 (Murphy, 2012).

$$J = -\sum_{j=1}^{h} t_j \log y_j \tag{2.19}$$

#### 2.9. Dilated Convolutional Neural Network

Dilated Convolutional Neural Network merupakan metode yang digunakan untuk mengatasi masalah komputasi yang berat pada convolutional neural network. Prinsip dasar dari dilated convolutional neural network adalah memberikan lubang/ruang diantara titik satu dengan yang lain pada proses perkalian matriks input dengan kernel. Metode ini terbukti pada beberapa penelitian dapat membuat kernel konvolusi memperoleh lebih banyak informasi dari matriks input tanpa meningkatkan proses komputasi (Lei et al., 2019). Proses pada dilated convolutional neural network dapat dilakukan menggunakan Persamaan .2.8. Jika metode convolutional neural network menggunakan nilai d=1, maka metode dilated convolutional neural network menggunakan nilai d>1. Ilustrasi dari metode dilated convolutional neural network dapat dilihat pada Gambar 2.21.



Gambar 2.21 (Kiri) Convolutional Neural Network dengan d=1; (Kanan) Dilated Convolutional Neural Network dengan d=2

(Sumber: Lei et al., 2019)

#### 2.10. Matriks Konfusi

Matriks konfusi berisi informasi tentang hasil aktual dan prediksi yang dilakukan oleh sistem klasifikasi. Kinerja sistem dievaluasi menggunakan data dalam matriks konfusi dengan menghitung secara statistik dari nilai *True Positive* (TP) yang merupakan kelas positif teridentifikasi positif, *True Negative* (TN) yang merupakan kelas negatif teridentifikasi negatif, *False Positive* (FP) yang merupakan kelas negatif teridentifikasi positif, dan *False Negative* (FN) yang merupakan kelas positif yang teridentifikasi negatif (Navin J R and R, 2016).

	Predicted: No	Predicted: Yes
Actual: No	TN	FP
Actual: Yes	FN	TP

Gambar 2.22 Matriks Konfusi

(Sumber: Navin J R and R, 2016)

Dari Gambar 2.22 dapat dihasilkan berbagai perhitungan statistik nilai akurasi, spesifikasi, dan sensitivitas. Dari ketiga hasil tersebut, dapat menentukan baik ataupun buruknya kinerja dari sistem klasifikasi yang dibangun. Berikut penjelasan perhitungan statistik yang dilakukan.

#### 1. Akurasi

Akurasi merupakan hasil yang merepresentasikan banyaknya data yang terklasifikasi sesuai dengan data aktual. Sehingga, dapat disimpulkan bahwa semakin besar dari nilai akurasi, maka kinerja dari sistem klasifikasi dapat dikatakan baik karena banyak data yang terklasifikasi dengan benar. Perhitungan akurasi dapat dilakukan menggunakan Persamaan 2.20.

$$Akurasi = \frac{TP + TN}{TP + TN + FP + FN} \times 100\%$$
 (2.20)

## 2. Spesifisitas

Spesifisitas merupakan hasil yang merepresentasikan banyaknya data yang terklasifikasikan dengan benar pada kelas negatif. Sehingga, semakin besar nilai spesifisitas, maka semakin besar sistem klasifikasi dapat mengenali kelas negatif dengan baik. Perhitungan spesifisitas dapat dilakukan menggunakan Persamaan 2.21.

$$Spesifisitas = \frac{TN}{TN + FP} \times 100\%$$
 (2.21)

#### 3. sensitivitas

sensitivitas merupakan kebalikan dari spesifisitas, yaitu hasil yang merepresentasikan banyaknya data yang terklasifikasikan dengan benar pada kelas positif. Sehingga, semakin besar nilai sensitivitas, maka semakin besar sistem klasifikasi dapat mengenali kelas positif dengan baik. Perhitungan spesifisitas dapat dilakukan menggunakan Persamaan 2.22.

$$sensitivitas = \frac{TP}{TP + FN} \times 100\%$$
 (2.22)

Penjelasan diatas merupakan perhitungan matriks permasalahan klasifikasi 2 kelas. Apabila sistem klasifikasi yang dibangun memiliki permasalahan lebih dari 2 kelas (n kelas), maka ilustrasi matriks konfusi dapat dilihat pada Gambar 2.23. Perhitungan akurasi, spesifisitas, dan sensitivitas dapat dilakukan menggunakan Persamaan 2.23 hingga 2.29 (Manliguez, 2016).

$$Akurasi = \frac{TTP}{\sum_{i=1}^{n} \sum_{k=1}^{n} x_{ik}} \times 100\%$$
 (2.23)

$$Akurasi = \frac{TTP}{\sum_{j=1}^{n} \sum_{k=1}^{n} x_{jk}} \times 100\%$$

$$sensitivitas_{i} = \frac{TTP_{i}}{TTP_{i} + TFN_{i}} \times 100\%$$

$$Spesifisitas_{i} = \frac{TTN_{i}}{TTN_{i} + TFP_{i}} \times 100\%$$
(2.23)
$$(2.24)$$

$$Spesifisitas_{i} = \frac{TTN_{i}}{TTN_{i} + TFP_{i}} \times 100\%$$
 (2.25)

$$TTP_i = \sum_{j=1, j=i}^{n} x_{ij}$$
 (2.26)

$$TFN_i = \sum_{j=1, j \neq i}^n x_{ij} \tag{2.27}$$

$$TFP_i = \sum_{j=1, j \neq i}^n x_{ji} \tag{2.28}$$

$$TTN_i = \sum_{j=1, j \neq i}^{n} \sum_{k=1, k \neq i}^{n} x_{jk}$$
 (2.29)

		Predicted Number					
		Class 1	Class 2		Class n		
	Class 1	<i>x</i> <sub>11</sub>	<i>x</i> <sub>12</sub>	- iii	$x_{1n}$		
Actual Number	Class 2	x <sub>21</sub>	x <sub>22</sub>		$x_{2n}$		
		1	11300	-	40		
	151	5	*		0.0		
				•			
	Class $n$	$x_{n1}$	$x_{n2}$	768	$x_{nn}$		

Gambar 2.23 Matriks Konfusi Pada Permasalahan n Kelas

(Sumber: Manliguez, 2016)

## 2.11. Integrasi Keilmuan

Kanker merupakan penyakit yang sulit untuk dideteksi. Beberapa metode yang sudah ada seperti *chest x-ray*, mamografi, dan lain-lain dapat dilakukan ketika pasien memiliki gejala. Masalah yang terjadi adalah beberapa pasien yang memiliki gejala biasanya sudah menderita kanker pada stadium akhir. Hal tersebut tentunya akan menyulitkan dalam proses penyembuhan karena kanker sudah menyebar pada tubuh pasien. Oleh karena itu pada penelitian ini dilakukan klasifikasi kanker berdasarkan data RNA menggunakan metode *dilated convolutional neural network*. Data RNA digunakan karena salah satu penyebab penyakit kanker yaitu dari faktor keturunan. Hal ini sesuai dengan hadits Rasulullah SAW yang diriwayatkan oleh Imam Bukhari bahwa seorang manusia akan memiliki kemiripan dengan ayah atau ibunya.

وَأَمَّا الشَّبَهُ فِي الْوَلَدِ فَإِنَّ الرَّجُلَ إِذَا غَشِيَ الْمَرْأَةَ فَسَبَقَهَا مَاؤُهُ كَانَ الشَّبَهُ لَهُ وَإِذَا سَبَقَ مَاؤُهَا كَانَ الشَّبَهُ لَهَا "Adapun kemiripan seorang anak dengan bapaknya adalah apabila sang suami mendatangi istrinya, apabila air mani suami mendahului air mani istrinya berarti akan lahir anak yang mirip dengan bapaknya, sebaliknya apabila air mani istrinya mendahului air mani suaminya maka akan lahir anak yang mirip dengan ibunya".

[HR Bukhori]

Klasifikasi penyakit kanker merupakan salah satu cara guna melakukan pencegahan pada penderita kanker agar dapat segera ditangani. Hal ini sesuai dengan hadits Rasulullah SAW sebagai berikut.

"Sesungguhnya Allah telah menurunkan penyakit dan obat, dan menjadikan bagi setiap penyakit terdapat obatnya, maka berobatlah dan jangan berobat dengan sesuatu yang haram". [HR Abu Dawud]

Hadits diatas menjelaskan bahwa Rasulullah menganjurkan bagi setiap orang yang sakit untuk berobat atau memeriksakan penyakitnya karena setiap penyakit pasti bisa disembuhkan. Semakin cepat seseorang mendapatkan penanganan atas penyakitnya, maka semakin tinggi kemungkinan untuk disembuhkan sehingga dapat mencegah terjadinya kematian. Akan tetapi, segala macam usaha yang dilakukan manusia hanyalah bagian dari ikhtiar yang dilakukan karena dalam hal penyembuhan merupakan kuasa Allah SWT sebagaimana firman Allah dalam Q.S. Asy-Syuara/28:80 dan Q.S. Al-Anbiyaa/21:83 yang berbunyi:

# وَإِذَا مَرِضْتُ فَهُوَ يَشْفِينِ

"Dan apabila aku (Ibrahim) sakit, Dia (Allah)-lah yang menyembuhkan diriku"

# وَأَيُّوبَ إِذْ نَادَىٰ رَبَّهُ ۚ أَنِّي مَسَّنِي ٱلضُّرُّ وَأَنتَ أَرْحَمُ ٱلرَّحِمِينَ

"Dan (ingatlah kisah) Ayub, ketika ia menyeru Tuhannya: "(Ya Tuhanku), sesungguhnya aku telah ditimpa penyakit dan Engkau adalah Tuhan yang Maha Penyayang di antara semua penyayang""

Kedua ayat di atas menjelaskan bahwa Nabi Ibrahim dan Nabi Ayub memiliki keyakinan bahwa ketika suatu saat tertimpa penyakit, maka hanya Allah SWT yang beliau yakini dapat menyembuhkan. Hal tersebut adalah contoh sikap tawakal dari seorang hamba terhadap Penciptanya yang akan selalu menyayangi hamba-Nya.

Pada proses penyembuhan penyakit kanker, semua cara penyembuhan hanya merupakan usaha dan perantara, yang menentukan kesembuhan hanyalah Allah SWT. Oleh karena itu, klasifikasi penyakit kanker payudara, kanker paru-paru, kanker rahim, dan kanker ginjal merupakan bagian dari usaha manusia untuk mendapatkan kesembuhan.

#### **BAB III**

#### METODE PENELITIAN

#### 3.1. Jenis Penelitian

Penelitian ini tergolong ke dalam jenis penelitian kuantitatif karena berdasarkan tujuan yang ingin dicapai, penelitian ini mengandung unsur perhitungan serta data penelitian yang digunakan yaitu berbentuk numerik atau angka. Hasil dari penelitian yang berupa sistem klasifikasi kanker berdasarkan data RNA menggunakan metode *dilated convolutional neural network* diharapkan dapat membantu para ahli medis dalam mendiagnosis kanker pada pasien.

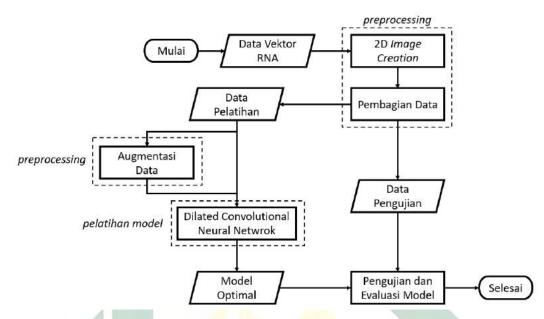
#### 3.2. Jenis Data

Data yang digunakan dalam penelitian ini merupakan data sekunder RNA yang dapat diakses pada halaman *Mendeley Data* (Ferles et al., 2018). Data tersebut merupakan data vektor berisi susunan gen dari RNA yang diambil menggunakan *Illumina HiSeq Platform* dan terdiri dari 971 kolom yang berisi tentang RPKM (*Read per Kilobase Million* dari setiap gen. Keseluruhan data berjumlah 1846 yang sudah diklasifikasikan ke dalam 4 kelas yaitu BRCA (kanker payudara), KIRC (kanker ginjal), UCEC (kanker rahim), dan LUAD (kanker paru-paru).

#### 3.3. Teknik Analisis Data

Pada penelitian ini dilakukan beberapa tahapan yaitu *preprocessing*, pelatihan model, pengujian dan evaluasi model. Tahapan pada penelitian dapat

dilihat pada Gambar 3.1.



Gambar 3.1 Diagram Alir Sistem Klasifikasi

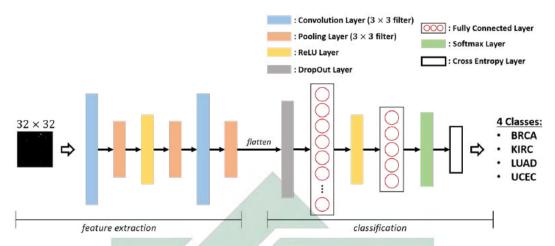
## 1. Preprocessing

Langkah pertama pada tahap ini yaitu mengubah vektor RNA yang terdiri dari 971 kolom menjadi matriks dengan jumlah baris sama dengan jumlah kolom. Berdasarkan jumlah kolom yang dimiliki, ukuran yang digunakan pada penelitian ini yakni  $32 \times 32$ . Akan tetapi, matriks dengan ukuran tersebut membutuhkan 1024 elemen, sedangkan jumlah elemen dari vektor RNA yaitu 971. Oleh karena itu akan ditambahkan nilai 0 pada vektor RNA dari kolom 972 - 1024 seperti yang dilakukan oleh (Lyu and Haque, 2018). Setelah semua vektor RNA menjadi matriks, selanjutnya dilakukan rescaling dari skala normal [0, 12620] menjadi skala image yaitu [0, 255]. Langkah kedua yaitu membagi data menjadi set data pelatihan dan set data pengujian. Pembagian data dilakukan menggunakan teknik k-fold cross validation dengan k = 4. Langkah ketiga yang dilakukan yaitu augmentasi data, hal ini dilakukan untuk mendapatkan sampel data yang lebih banyak.

Teknik augmentasi yang dilakukan yaitu refleksi terhadap sumbu x, refleksi terhadap sumbu y, dan rotasi sebesar  $180^{\circ}$ , yang masing-masing dilakukan menggunakan Persamaan 2.2 hingga 2.4. Augmentasi data hanya dilakukan pada set data pelatihan seperti pada penelitian yang dilakukan oleh (Shorten and Khoshgoftaar, 2019). Pada penelitian ini dilakukan percobaan dengan membandingkan hasil klasifikasi menggunakan augmentasi data dan tanpa menggunakan augmentasi data. Hal ini dilakukan untuk mengetahui apakah augmentasi data berpengaruh pada hasil yang diperoleh.

#### 2. Pelatihan Model

Pada tahap ini, set data pelatihan hasil dari proses sebelumnya akan dilatih menggunakan dilated convolutional neural network dengan arsitektur seperti pada Gamba<mark>r 3.2. Arsitektur</mark> tersebut dipilih karena berdasarkan (Khalifa et al., 2020), arsitektur tersebut mendapatkan hasil yang lebih baik dari arsitektur lainnya. Pada penelitian ini diharapkan metode dilated convolutional neural network dapat bekerja lebih baik dengan arsitektur Terdapat beberapa layer yang digunakan pada arsitektur yang sama. tersebut yaitu convolution layer, pooling layer, activation layer dengan ReLU sebagai fungsi aktivasi, dropout layer, softmax layer, fully connected layer, dan cross entropy layer. Pada tahap ini juga akan dilakukan percobaan dengan mengubah nilai faktor dilasi (d) menjadi 2, 4, 6, 8. Hasil yang didapatkan yaitu model dengan bobot dan filter terbaik yang nantinya akan digunakan pada proses selanjutnya. Hal ini dilakukan guna mengetahui apakah nilai faktor dilasi akan mempengaruhi hasil klasifikasi.



Gambar 3.2 Arsitektur Convolutional Neural Network

(Sumber: Khalifa et al., 2020)

## 3. Pengujian dan Evaluasi Model

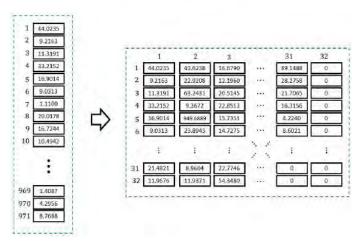
Setelah didapatkan bobot dan *filter* pada proses sebelumnya, langkah selanjutnya yaitu melakukan pengujian terhadap model yang dibangun dengan memberikan set data yang berbeda dengan set data pada proses pelatihan. Set data yang diberikan yaitu set data pengujian hasil dari pembagian menggunakan teknik *k-fold cross validation* pada tahap *preprocessing*. Model akan mengklasifikasikan set data pengujian yang nantinya hasil klasifikasi akan dievaluasi menggunakan konfusi matriks.

#### **BAB IV**

## HASIL DAN PEMBAHASAN

## 4.1. Preprocessing

Data yang digunakan pada penelitian ini yaitu data vektor RNA dengan 971 kolom yang berisi tentang RPKM dari masing-masing gen. Arsitektur *dilated convolutional neural network* yang digunakan menggunakan *input layer* dengan masukan berupa matriks, oleh karena itu langkah pertama pada tahap *preprocessing* yaitu mengubah vektor RNA yang terdiri dari 971 kolom menjadi matriks dengan jumlah baris sama dengan jumlah kolom. Berdasarkan jumlah kolom yang dimiliki, ukuran yang digunakan pada penelitian ini yakni 32 × 32. Akan tetapi, matriks dengan ukuran tersebut membutuhkan 1024 elemen, sedangkan jumlah elemen dari vektor RNA yaitu 971. Oleh karena itu, akan ditambahkan nilai 0 pada vektor RNA dari kolom 972 — 1024. Gambar 4.1 merupakan ilustrasi dari proses *2d image creation* yang dilakukan pada data RNA dari kanker payudara.



Gambar 4.1 Ilustrasi 2D Image Creation

Setelah diubah menjadi matriks, langkah selanjutnya yaitu merubah skala matriks dari [0,12620] menjadi skala *image* yaitu [0,255] menggunakan Persamaan 2.1. Sehingga didapatkan data gambar hasil *2d image creation* seperti pada Gambar 4.2.

$$PixelValue(44.0235) = round\left(\frac{(44.0235)(255)}{(12620 - 0)}\right)$$

$$= 1$$

$$PixelValue(9.2163) = round\left(\frac{(9.2163)(255)}{(12620 - 0)}\right)$$

$$= 0$$

$$PixelValue(11.3191) = round\left(\frac{(11.3191)(255)}{(12620 - 0)}\right)$$

$$= 0$$

$$\vdots$$

$$PixelValue(0) = round\left(\frac{(0)(255)}{(12620 - 0)}\right)$$

$$= 0$$

BRCA KIRC LUAD UCEC

Gambar 4.2 Hasil 2D Image Creation

Hasil dari proses sebelumnya yaitu data citra dengan ukuran  $32 \times 32$ . Data tersebut kemudian dibagi menjadi set data pelatihan dan set data pengujian menggunakan k-fold cross validation dengan k=4.

#### 4.2. Pelatihan Model

Proses pelatihan model dilakukan dengan membandingkan beberapa percobaan untuk mendapatkan model yang terbaik. Percobaan yang dilakukan yaitu merubah nilai d pada  $convolution\ layer$ , dan membandingkan hasil model yang menggunakan augmentasi data dan yang tidak menggunakan augmentasi data. Nilai d yang digunakan yaitu 2,4,6, dan 8, sedangkan proses augmentasi data yang dilakukan yaitu refleksi terhadap sumbu x, refleksi terhadap sumbu y, serta rotasi sebesar  $180^{\circ}$ . Arsitektur yang digunakan pada proses pelatihan model dapat dilihat pada Gambar 3.2. Data yang digunakan pada proses pelatihan model sebanyak 1383 - 1386. Berikut merupakan Langkah-langkah yang dilakukan pada proses pelatihan model klasifikasi RNA menggunakan  $dilated\ convolutional\ neural\ network$ . Sistem yang dibangun akan mengklasifikan kanker menjadi 4 kelas yaitu BRCA dengan kode label [1000], KIRC [0100], LUAD [0010] dan UCEC [0001].

#### 4.2.1. Augmentasi Data

Proses augmentasi dilakukan pada data citra hasil dari proses 2D image creation yang sudah dipartisi sebanyak k. Contoh data tersebut dapat dilihat pada Gambar 4.3. Citra tersebut tersusun dari piksel yang memiliki rentang nilai [0, 255] seperti pada Gambar 4.4. Augmentasi data yang dilakukan yaitu refleksi terhadap sumbu x, refleksi terhadap sumbu y, dan rotasi sebesar  $180^{\circ}$ .



Gambar 4.3 Contoh Citra Input

1	1	0	1	0	1	1	•••	0	2	0
0	0	0	0	2	0	0	•••	0	1	0
0	1	0	0	0	6	3	•••	1	0	0
1	0	0	0	0	0	1	•••	0	0	0
0	19	0	0	0	1	1	•••	0	0	0
0	0	0	0	0	0	4	•••	0	0	0
0	2	0	1	0	1	0	•••	1	0	0
:	:	:	:	:	:	:		:	:	:
•	•	•	•	•	•	•		•	•	•
0	0	1	0	0	0	0	•••	1	0	0
0	0	0	0	0	0	0	•••	0	0	0
0	0	1	0	0	1	0	•••	0	0	0
	0 0 1 0 0 0	0 0 0 1 1 1 0 0 19 0 0 0 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0 0 0 0	0 0 0 0 0 0 0 1 0 0 0 0 0 0 0 0 0 0 0 0	0       0       0       0       2         0       1       0       0       0         1       0       0       0       0         0       19       0       0       0         0       0       0       0       0         0       2       0       1       0         0       0       1       0       0         0       0       0       0       0	0       0       0       0       2       0         0       1       0       0       0       6         1       0       0       0       0       0         0       19       0       0       0       1         0       0       0       0       0       0         0       2       0       1       0       1         1       1       0       0       0       0         0       0       1       0       0       0         0       0       0       0       0       0	0       0       0       0       2       0       0         0       1       0       0       0       0       6       3         1       0       0       0       0       0       1       1         0       19       0       0       0       1       1         0       0       0       0       0       0       4         0       2       0       1       0       1       0         1       0       0       0       0       0       0         0       0       1       0       0       0       0       0         0       0       0       0       0       0       0       0       0	0       0       0       0       2       0       0         0       1       0       0       0       6       3         1       0       0       0       0       0       1         0       19       0       0       0       1       1         0       0       0       0       0       4          0       2       0       1       0       1       0          0       0       1       0       0       0       0          0       0       0       0       0       0       0	0       0       0       0       2       0       0         0       1       0       0       0       6       3       1         1       0       0       0       0       0       1       0         0       19       0       0       0       1       1       0         0       0       0       0       0       4       0       0         0       2       0       1       0       1       0       1         1       0       0       0       0       0       0       1         0       0       1       0       0       0       0       0       0         0       0       0       0       0       0       0       0       0	0       0       0       0       2       0       0       1       0       1       0       1       0       1       0       1       0       0       0       1       0

Gambar 4.4 Contoh Nilai Piksel Citra Input

# 1. Refleksi Terhadap Sumbu x

Refleksi terhadap sumbu x dilakukan menggunakan Persamaan 2.2. Sebagai contoh Gambar 4.3 akan dilakukan transformasi refleksi terhadap sumbu x.

$$(x,y) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$(0,0) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$(0,1) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\vdots$$

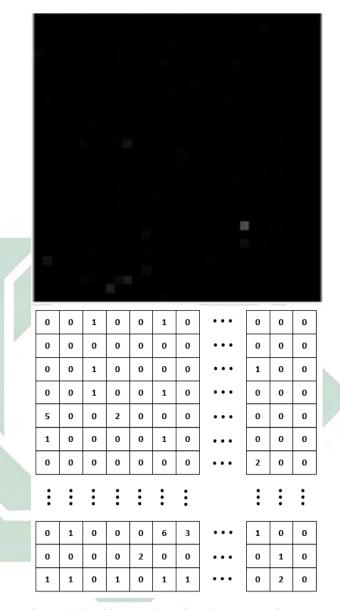
$$(31,30) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 32 \\ 31 \end{bmatrix}$$

$$= \begin{bmatrix} 32 \\ -31 \end{bmatrix}$$

$$(31,31) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 32 \\ 32 \end{bmatrix}$$

$$= \begin{bmatrix} 32 \\ -32 \end{bmatrix}$$

Dari hasil tersebut didapatkan citra baru dengan nilai piksel seperti pada Gambar 4.5.



Gambar 4.5 Citra Hasil Refleksi Terhadap Sumbu  $\boldsymbol{x}$ 

# 2. Refleksi Terhadap Sumbu y

Refleksi terhadap sumbu y dilakukan menggunakan Persamaan 2.3. Sebagai contoh Gambar 4.3 akan dilakukan transformasi refleksi terhadap sumbu y, sehingga didapatkan citra baru dengan nilai piksel seperti pada Gambar 4.6.

$$(x,y) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$(0,0) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

$$(0,1) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

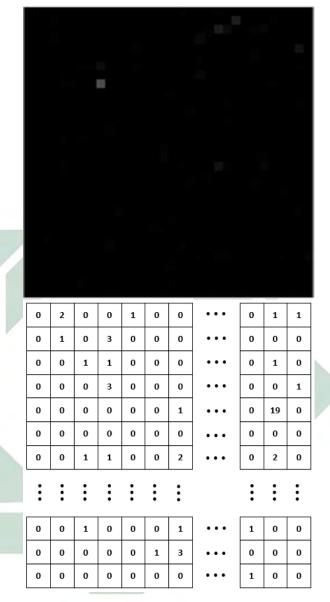
$$\vdots$$

$$(31,30) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 32 \\ 31 \end{bmatrix}$$

$$= \begin{bmatrix} -32 \\ 31 \end{bmatrix}$$

$$(31,31) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 32 \\ 32 \end{bmatrix}$$

$$= \begin{bmatrix} -32 \\ 32 \end{bmatrix}$$



Gambar 4.6 Citra Hasil Refleksi Terhadap Sumbu y

## 3. Rotasi

Rotasi dilakukan menggunakan Persamaan 2.4. Sebagai contoh Gambar 4.3 akan dilakukan transformasi rotasi sebesar 180°.

$$(x,y) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} \cos 180^{\circ} & -\sin 180^{\circ} \\ \sin 180^{\circ} & \cos 180^{\circ} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

$$(0,0) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

$$(0,1) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

$$= \begin{bmatrix} -1 \\ -2 \end{bmatrix}$$

$$\vdots$$

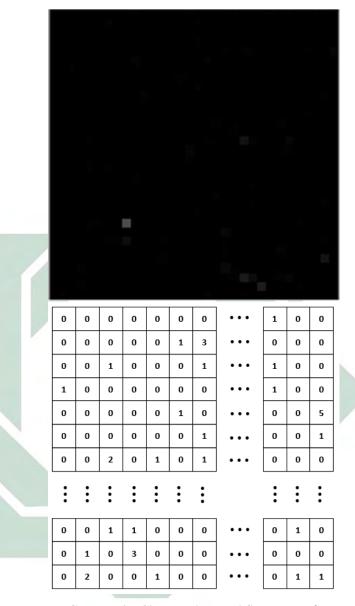
$$(31,30) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 32 \\ 31 \end{bmatrix}$$

$$= \begin{bmatrix} -32 \\ -31 \end{bmatrix}$$

$$(31,31) \Rightarrow \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 32 \\ 32 \end{bmatrix}$$

$$= \begin{bmatrix} -32 \\ -32 \end{bmatrix}$$

Dari hasil tersebut didapatkan citra baru dengan nilai piksel seperti pada Gambar 4.7.



Gambar 4.7 Citra Hasil Rotasi Sebesar  $180^{\circ}$ 

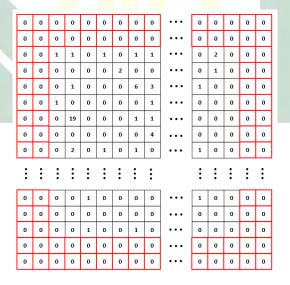
## 4.2.2. Convolution Layer

Pada penelitian ini dilakukan dua kali proses *convolution* seperti arsitektur yang dapat dilihat pada Gambar 3.2. Proses pada *Convolution layer* dilakukan menggunakan *filter* berukuran  $3 \times 3$ , *stride* (s) sebesar 1, *padding* (p) sama dengan faktor dilasi (d), dan percobaan nilai faktor dilasi sebesar 2, 4, 6, 8. Perbedaan *convolution layer* yang pertama dan kedua terdapat pada jumlah *filter* yang

digunakan yaitu 16 dan 32. Berikut ini merupakan contoh perhitungan pada *convolution layer* yang akan dilakukan menggunakan Gambar 4.3 dengan faktor dilasi sebesar 2. Ukuran *output* yang didapatkan dapat dihitung menggunakan Persamaan 2.5.

$$u = \frac{m - (d+f) + 2p}{s} + 1$$
$$= \frac{32 - (2+3) + 2(2)}{1} + 1$$
$$= 32$$

Perhitungan pada *convolution layer* dilakukan menggunakan Persamaan 2.8. Contoh perhitungan akan dilakukan menggunakan piksel pada Gambar 4.4 dengan nilai *padding* sama dengan faktor dilasi yaitu sebesar 2, sehingga nilai piksel dengan tambahan *padding* dapat dilihat pada Gambar 4.8.



Gambar 4.8 Contoh Piksel dengan Padding

Perhitungan pada convolution layer dilakukan menggunakan filter (**W**) berukuran  $3 \times 3$ . Nilai filter didapat secara acak dalam rentang nilai [-1,1]. Filter yang digunakan pada penelitian ini dapat dilihat pada Lampiran A. Salah satu contoh

filter dan akan digunakan pada contoh perhitungan yaitu:

$$\mathbf{W} = \begin{bmatrix} -0.0820 & 0.1306 & -0.1955 \\ 0.0525 & 0.0363 & -0.0091 \\ 0.2203 & 0.1783 & -0.0140 \end{bmatrix}$$

$$z_{(0,0)} = \sum_{a=0}^{3-1} \sum_{b=0}^{3-1} w_{(a,b)} i_{(0+2a,0+2b)}$$

$$= \left( (w_{(0,0)} i_{(0,0)}) + (w_{(0,1)} i_{(0,2)}) + (w_{(0,2)} i_{(0,4)}) \right) + \left( (w_{(1,0)} i_{(2,0)}) + (w_{(1,1)} i_{(2,2)}) + (w_{(1,2)} i_{(2,4)}) \right) + \left( (w_{(2,0)} i_{(4,0)}) + (w_{(2,1)} i_{(4,2)}) + (w_{(2,2)} i_{(4,4)}) \right)$$

$$= \left( (-0.0820 \times 0) + (0.1306 \times 0) + (-0.1955 \times 0) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 0) \right) + \left( (0.2203 \times 0) + (0.1783 \times 0) + (-0.0140 \times 0) \right)$$

$$= 0.0363$$

$$z_{(0,1)} = \sum_{a=0}^{3-1} \sum_{b=0}^{3-1} w_{(a,b)} i_{(0+2a,1+2b)}$$

$$= \left( (w_{(0,0)} i_{(0,1)}) + (w_{(0,1)} i_{(0,3)}) + (w_{(0,2)} i_{(0,5)}) \right) + \left( (w_{(1,0)} i_{(2,1)}) + (w_{(1,1)} i_{(2,3)}) + (w_{(1,2)} i_{(2,5)}) \right) + \left( (w_{(2,0)} i_{(4,1)}) + (w_{(2,1)} i_{(4,3)}) + (w_{(2,2)} i_{(4,5)}) \right)$$

$$= \left( (-0.0820 \times 0) + (0.1306 \times 0) + (-0.1955 \times 0) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0) + (0.0363 \times 1) + (-0.0091 \times 1) \right) + \left( (0.0525 \times 0)$$

$$\left( (0.2203 \times 0) + (0.1783 \times 1) + (-0.0140 \times 0) \right)$$

$$= 0.2055$$

$$\vdots$$

$$z_{(31,31)} = \sum_{a=0}^{3-1} \sum_{b=0}^{3-1} w_{(a,b)} i_{(31+2a,31+2b)}$$

$$= \left( (w_{(0,0)} i_{(31,31)}) + (w_{(0,1)} i_{(31,33)}) + (w_{(31,35)} i_{(1,5)}) \right) +$$

$$\left( (w_{(1,0)} i_{(33,31)}) + (w_{(1,1)} i_{(33,33)}) + (w_{(33,35)} i_{(3,5)}) \right) +$$

$$\left( (w_{(2,0)} i_{(35,31)}) + (w_{(2,1)} i_{(35,33)}) + (w_{(35,35)} i_{(5,5)}) \right)$$

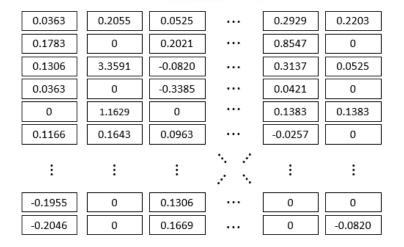
$$= \left( (-0.0820 \times 1) + (0.1306 \times 0) + (-0.1955 \times 0) \right) +$$

$$\left( (0.0525 \times 0) + (0.0363 \times 0) + (-0.0091 \times 0) \right) +$$

$$\left( (0.2203 \times 0) + (0.1783 \times 0) + (-0.0140 \times 0) \right)$$

$$= -0.0820$$

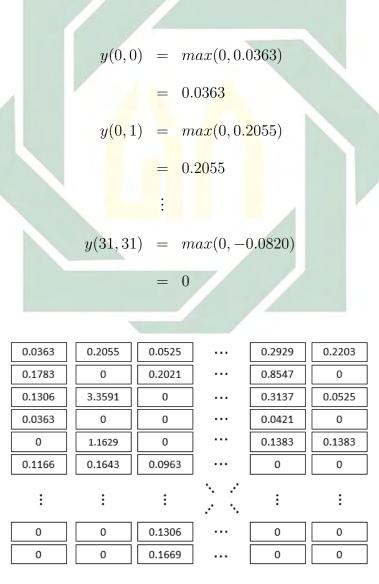
Perhitungan pada *convolution layer* dilakukan pada masing-masing *filter* sebanyak jumlah *filter* yang digunakan, sehingga setiap *filter* akan mendapatkan hasil yang berbeda-beda.



Gambar 4.9 Hasil Convolution Layer

## 4.2.3. Activation Layer

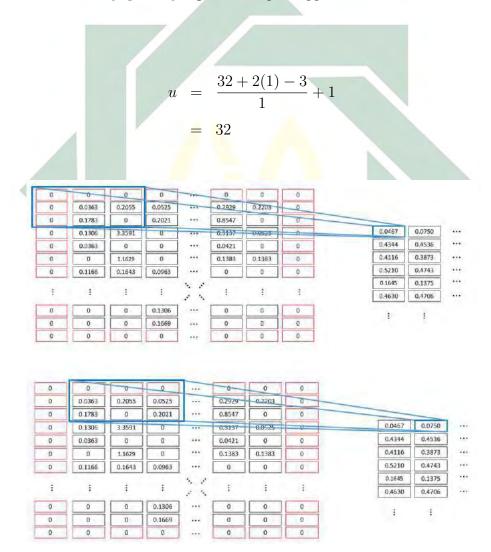
Pada penelitian ini digunakan fungsi aktivasi *Rectified Linear Unit* (ReLU) yang bertujuan untuk mengantisipasi nilai negatif pada proses sebelumnya. Hasil dari proses *convolution layer* pada Gambar 4.9 masih terdapat nilai negatif, sedangkan skala dari citra yaitu [0, 255]. Perhitungan fungsi aktivasi dilakukan menggunakan Persamaan 2.14 sehingga didapatkan hasil pada Gambar 4.10.



Gambar 4.10 Hasil Activation Layer ReLU

## 4.2.4. Pooling Layer

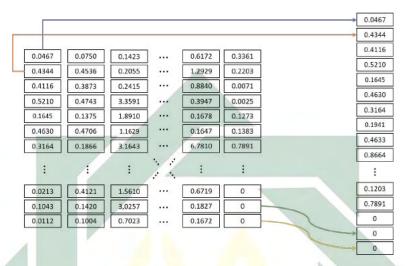
Pada arsitektur yang digunakan terdapat satu *pooling layer* dengan ukuran  $3 \times 3$  dengan *padding* sama dengan *stride* yaitu sebesar 1. Jenis *pooling layer* yang digunakan adalah *average pooling*. Gambar 4.11 merupakan contoh hasil penerapan dari ReLU *activation layer average pooling* pada Gambar 4.10. Hasil keluaran dari *average pooling* dapat dihitung menggunakan Persamaan 2.9.



Gambar 4.11 Contoh Penerapan Average Pooling

Hasil dari proses *average pooling* akan digunakan pada proses selanjutnya yaitu *dropout layer*. Akan tetapi, sebelum digunakan pada proses selanjutnya, akan

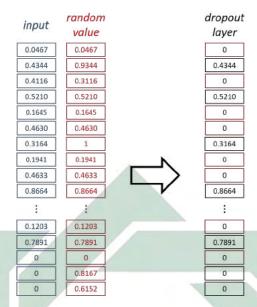
dilakukan *flatten* (meratakan) terlebih dahulu agar hasil dari proses *pooling layer* dapat digunakan pada tahap selanjutnya. Contoh penerapan *flatten* dapat dilihat pada Gambar 4.12.



Gambar 4.12 Contoh Penerapan Flatten

#### 4.2.5. Dropout Layer

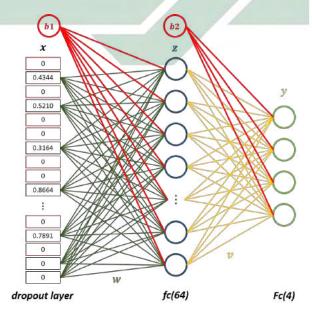
Pada proses ini akan dilakukan seleksi apakah neuron diteruskan ke lapisan selanjutnya atau tidak. Hasil dari proses sebelumnya berupa vektor yang terdiri dari 1024 elemen. Hasil tersebut akan menjadi nilai *input* pada *dropout layer* yang terdiri dari 1024 neuron. Neuron tersebut akan diseleksi dengan probabilitas masing-masing sebesar 0.5. Langkah yang dilakukan yaitu membuat vektor baru terdiri dari 1024 elemen yang berisi nilai acak antara 0 sampai 1. Apabila hasil nilai acak kurang dari nilai probabilitas, maka neuron tersebut tidak digunakan pada proses selanjutnya. Neuron yang tidak digunakan dapat diganti nilai 0 atau dilewati dalam proses perhitungan. Proses pada *dropout layer* dapat dilihat pada Gambar 4.13.



Gambar 4.13 Contoh Proses Dropout

# 4.2.6. Fully Connected Layer

Pada arsitektur yang digunakan terdapat dua *fully connected layer* yang terdiri dari 64 neuron dan 4 neuron. Arsitektur dari *fully connected layer* dapat dilihat pada Gambar 4.14. Berikut merupakan langkah-langkah yang dilakukan pada algoritma *backpropagation*.



Gambar 4.14 Arsitektur Fully Connected Layer

Inisialisasi bobot awal W dan V serta bias b1 dan b2 secara acak dalam rentang -1 hingga 1. Nilai toleransi kesalahan (e) yang digunakan pada penelitian ini yaitu 0.00001 dan laju pemahaman (α) sebesar 0.1. Ukuran dari matriks W, V, b1 dan b2 tergantung dari banyaknya neuron pada lapisan x, z, dan y. Sebagai contoh pada penelitian ini digunakan 1024 neuron pada lapisan x, 64 neuron pada lapisan z, dan 4 neuron pada lapisan y sehingga ukuran dari matriks W yaitu 64 × 1024, matriks V yaitu 4 × 64, matriks b1 yaitu 1 × 64 dan matriks b2 yaitu 1 × 4.

$$\mathbf{W} = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & 1024 \\ 0.0975 & -0.8003 & 0.9340 & -0.2769 & \cdots & 0.2238 \\ -0.2785 & -0.1419 & 0.6787 & 0.0462 & \cdots & -0.6797 \\ 0.5469 & -0.4218 & 0.7577 & 0.0971 & \cdots & 0.6551 \\ 4 & 0.9575 & 0.9157 & 0.7743 & 0.8235 & \cdots & 0.1626 \\ 5 & -0.9649 & 0.7922 & -0.3922 & 0.6948 & \cdots & -0.1190 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 64 & 0.1576 & 0.9595 & 0.6555 & 0.3171 & \cdots & -0.4984 \end{pmatrix}$$

$$\mathbf{V} = \begin{pmatrix} 0.7572 & 0.5308 & 0.0119 & -0.1656 & \cdots & 0.5688 \\ -0.7537 & 0.7792 & 0.3371 & 0.6020 & \cdots & 0.6892 \\ -0.3804 & -0.9340 & -0.1622 & -0.2630 & \cdots & -0.4694 \\ 0.5678 & 0.1299 & 0.7943 & -0.6541 & \cdots & 0.7482 \end{pmatrix}$$

$$\mathbf{b1} = \begin{pmatrix} 0.4505 & 0.0838 & -0.2290 & -0.9133 & 0.1524 & \cdots & 0.8258 \end{pmatrix}$$
  
 $\mathbf{b2} = \begin{pmatrix} -0.9961 & -0.0782 & -0.4427 & 0.1067 \end{pmatrix}$ 

#### Propagasi maju:

Menghitung nilai keluaran pada lapisan z menggunakan Persamaan 2.20.
 Perhitungan dilakukan dengan nilai x merupakan dropout layer, nilai w dan b merupakan bobot W serta b1 yang sudah diinisialisasi pada langkah pertama.

$$znet_{1} = b1_{1} + (x_{1})(w_{1,1}) + (x_{2})(w_{1,2}) + \dots + (x_{1024})(w_{1,1024})$$

$$= 0.4505 + (0)(0.0975) + (0.4344)(-0.8003) + \dots$$

$$+ (0)(0.2238)$$

$$= 3.6585$$

$$znet_{2} = b1_{2} + (x_{1})(w_{2,1}) + (x_{2})(w_{2,2}) + \dots + (x_{1024})(w_{2,1024})$$

$$= 0.0838 + (0)(-0.2785) + (0.4344)(-0.1419) + \dots$$

$$+ (0)(-0.6797)$$

$$= -0.6566$$

$$\vdots$$

$$znet_{64} = b1_{1} + (x_{1024})(w_{64,1}) + (x_{2})(w_{64,2}) + \dots$$

$$+ (x_{1024})(w_{64,1024})$$

$$= 0.8258 + (0)(0.1576) + (0.4344)(0.9595) + \dots$$

$$+ (0)(-0.4984)$$

$$= 1.4859$$

Selanjutnya yaitu menghitung nilai keluaran menggunakan fungsi aktivasi. Pada lapisan ini digunakan fungsi aktivasi ReLU seperti pada Gambar 3.2. Perhitungan dilakukan menggunakan Persamaan 2.14.

$$z_{1} = max(0, 3.6585)$$

$$= 3.6585$$

$$z_{2} = max(0, -0.6566)$$

$$= 0$$

$$\vdots$$

$$z_{64} = max(0, 1.4859)$$

$$= 1.4859$$

3. Menghitung nilai keluaran pada lapisan y menggunakan Persamaan 2.20. Perhitungan dilakukan dengan nilai x merupakan hasil keluaran dari lapisan z, nilai w dan b merupakan bobot V serta b2 yang sudah diinisialisasi pada langkah pertama.

$$ynet_1 = b2_1 + (z_1)(v_{1,1}) + (z_2)(v_{1,2}) + \dots + (z_{64})(v_{1,64})$$

$$= -0.9961 + (3.6585)(0.7572) + (0)(0.5308) + \dots$$

$$+ (1.4859)(0.5688)$$

$$= -0.8875$$

$$ynet_2 = b2_2 + (z_1)(v_{2,1}) + (z_2)(v_{2,2}) + \dots + (z_{64})(v_{2,64})$$

$$= -0.0782 + (3.6585)(-0.7573) + (0)(0.7792) + \dots$$

$$+ (1.4859)(0.6892)$$

$$= 0.7761$$

$$ynet_{3} = b2_{3} + (z_{1})(v_{3,1}) + (z_{2})(v_{3,2}) + \dots + (z_{64})(v_{3,64})$$

$$= -0.4427 + (3.6585)(-0.3804) + (0)(-0.9340) + \dots$$

$$+ (1.4859)(-0.4694)$$

$$= 0.4616$$

$$ynet_{4} = b2_{4} + (z_{1})(v_{4,1}) + (z_{2})(z_{4,2}) + \dots$$

$$+ (z_{64})(z_{4,64})$$

$$= 0.1067 + (3.6585)(0.5678) + (0)(0.1299) + \dots$$

$$+ (01.4859)(0.7482)$$

$$= 0.0851$$

Selanjutnya yaitu menghitung nilai keluaran menggunakan fungsi aktivasi. Pada lapisan ini digunakan fungsi aktivasi *softmax* seperti pada Gambar 3.2. Perhitungan dilakukan menggunakan Persamaan 2.16.

$$y_{1} = \frac{e^{-0.8875}}{e^{-0.8875} + e^{0.7761} + e^{0.4616} + e^{0.0851}}$$

$$= 0.0783$$

$$y_{2} = \frac{e^{0.7761}}{e^{-0.8875} + e^{0.7761} + e^{0.4616} + e^{0.0851}}$$

$$= 0.4131$$

$$y_{3} = \frac{e^{0.4616}}{e^{-0.8875} + e^{0.7761} + e^{0.4616} + e^{0.0851}}$$

$$= 0.3016$$

$$y_{4} = \frac{e^{0.0851}}{e^{-0.8875} + e^{0.7761} + e^{0.4616} + e^{0.0851}}$$

$$= 0.2070$$

4. Menghitung nilai *error* untuk mengetahui seberapa baik kinerja dari sistem yang dibangun. Pada sistem ini digunakan rumus *cross entropy* untuk menghitung nilai *error*. Perhitungan dilakukan menggunakan Persamaan 2.19. Variabel t merupakan label asli dari data yang digunakan, sedangkan y merupakan hasil perhitungan pada lapisan y. Pada perhitungan ini digunakan data dari kanker payudara, oleh karena itu nilai dari variabel t yaitu [1000]. Apabila nilai *error* yang didapat lebih besar dari nilai toleransi *error* yang diinisialisasi pada langkah pertama, maka dilakukan propagasi mundur hingga memperoleh nilai *error* yang diinginkan.

$$J = -\sum_{j=1} t_j \log y_j$$

$$= -(t_1 \log y_1 + t_2 \log y_2 + t_3 \log y_3 + t_4 \log y_4)$$

$$= -(1 \log 0.0783 + 0 \log 0.4131 + 0 \log 0.3016 + 0 \log 0.2070)$$

$$= -(-1.1062)$$

$$= 1.1062$$

Propagasi mundur:

5. Menghitung kesalahan pada lapisan sebelumnya  $\left(\frac{\partial J}{\partial z_i}\right)$  menggunakan persamaan berikut:

$$\frac{\partial J}{\partial z_i} = \frac{\partial J}{\partial y_i} \frac{\partial y_j}{\partial z_i}$$

Dimana  $\frac{\partial y_j}{\partial z_i}$  dapat dilihat pada Persamaan 2.17, sedangkan  $\frac{\partial J}{\partial y_j}$  sebagai berikut.

$$\frac{\partial J}{\partial y_j} = -\sum_{j=1} \frac{\partial t_j \log y_j}{\partial y_j}$$
$$= -\sum_{j=1} \frac{t_j}{y_j}$$

Sehingga didapatkan  $\left(\frac{\partial J}{\partial z_i}\right)$  sebagai berikut.

$$\frac{\partial J}{\partial z_i} = \frac{\partial J}{\partial y_j} \frac{\partial y_j}{\partial z_I}$$

$$= -\frac{t_i}{y_i} y_i (1 - y_i) + \sum_{j \neq i} \frac{t_j}{y_j} (y_i) (y_j)$$

$$= t_i y_i - t_i + y_i \sum_{j \neq i} t_j$$

$$= y_i (t_i + \sum_{j \neq i} t_j) - t_i$$

$$= y_i - t_i$$

$$\frac{\partial J}{\partial z_1} = y_1 - t_1$$

$$= 0.0783 - 1$$

$$= -0.9217$$

$$\frac{\partial J}{\partial z_2} = y_2 - t_2$$

$$= 0.4131 - 0$$

$$= 0.4131$$

$$\frac{\partial J}{\partial z_3} = y_3 - t_3$$

$$= 0.3016 - 0$$

$$= 0.3016$$

$$\frac{\partial J}{\partial z_4} = y_4 - t_4$$
$$= 0.2070 - 0$$
$$= 0.2070$$

6. Menghitung koreksi nilai bobot dan bias yang nantinya akan digunakan untuk memperbarui bobot **V** dan bias **b2**.

$$\Delta v_{jk} = \alpha \frac{\partial J}{\partial z_j} z_k$$

$$\Delta v_{1,1} = 0.1 \times -0.9217 \times 3.6585$$

$$= -0.3372$$

$$\Delta v_{1,2} = 0.1 \times -0.9217 \times 0$$

$$= 0$$

$$\vdots$$

$$\Delta v_{1,64} = 0.1 \times -0.9217 \times 1.4859$$

$$= -0.1369$$

$$\Delta v_{2,1} = 0.1 \times 0.4131 \times 3.6585$$

$$= 0.1511$$

$$\vdots$$

$$\Delta v_{4,64} = 0.1 \times 0.2070 \times 1.4859$$

$$= 0.0308$$

$$\Delta b2_{j} = \alpha \frac{\partial J}{\partial z_{j}} b_{j}$$

$$\Delta b2_{1} = 0.1 \times -0.9217 \times -0.9961$$

$$= 0.0918$$

$$\Delta b2_{2} = 0.1 \times 0.4131 \times -0.0782$$

$$= -0.0032$$

$$\Delta b2_{3} = 0.1 \times 0.3016 \times -0.4427$$

$$= -0.0134$$

$$\Delta b2_{4} = 0.1 \times 0.2070 \times 0.1067$$

$$= 0.0022$$

7. Menghitung koreksi nilai bobot dan bias yang nantinya akan digunakan untuk memperbarui bobot W dan bias b1.

$$snet_{k} = \sum_{j=1}^{4} \frac{\partial J}{\partial z_{j}} v_{jk}$$

$$snet_{1} = \frac{\partial J}{\partial z_{1}} v_{1,1} + \frac{\partial J}{\partial z_{2}} v_{2,1} + \frac{\partial J}{\partial z_{3}} v_{3,1} + \frac{\partial J}{\partial z_{4}} v_{4,1}$$

$$= -0.9217 \times 0.7572 + 0.4131 \times -0.7537$$

$$0.3016 \times -0.3804 + 0.2070 \times 0.5678$$

$$= -1.0065$$

$$snet_{2} = \frac{\partial J}{\partial z_{1}} v_{1,2} + \frac{\partial J}{\partial z_{2}} v_{2,2} + \frac{\partial J}{\partial z_{3}} v_{3,2} + \frac{\partial J}{\partial z_{4}} v_{4,2}$$

$$= -0.9217 \times 0.5308 + 0.4131 \times 0.7792$$

$$0.3016 \times -0.9340 + 0.2070 \times 0.1299$$

$$= -0.4222$$

$$\vdots$$

$$snet_{64} = \frac{\partial J}{\partial z_1} v_{1,64} + \frac{\partial J}{\partial z_2} v_{2,64} + \frac{\partial J}{\partial z_3} v_{3,64} + \frac{\partial J}{\partial z_4} v_{4,64}$$

$$= -0.9217 \times 0.5688 + 0.4131 \times 0.6892$$

$$0.3016 \times -0.4694 + 0.2070 \times 0.7482$$

$$= -0.2262$$

$$s_k = snet_k \frac{\partial z_k}{\partial x}$$

Dimana  $\frac{\partial z_j}{\partial x}$  dapat dilihat pada Persamaan 2.15.

$$s_1 = snet_1 \frac{\partial z_1}{\partial x}$$

$$= -1.0065 \times 1$$

$$= -1.0065$$

$$s_2 = snet_2 \frac{\partial z_2}{\partial x}$$

$$= -0.4222 \times 0$$

$$= 0$$

i

$$s_{64} = snet_{64} \frac{\partial z_{64}}{\partial x}$$
$$= -0.2262 \times 1$$
$$= -0.2262$$

$$\Delta w_{ki} = \alpha s_k x_i$$

$$\Delta w_{1,1} = 0.1 \times -1.0065 \times 0$$

$$= 0$$

$$\Delta w_{1,2} = 0.1 \times -1.0065 \times 0.4344$$

$$= -0.4372$$

$$\vdots$$

$$\Delta w_{1,1024} = 0.1 \times -1.0065 \times 0$$

$$= 0$$

$$\vdots$$

$$\Delta w_{2,1} = 0.1 \times 0 \times 0$$

$$= 0$$

$$\vdots$$

$$\Delta w_{64,1024} = 0.1 \times -0.2262 \times 0$$

$$= 0$$

$$\Delta b1_k = \alpha s_k b1_k$$

$$\Delta b1_1 = \alpha s_1 b1_1$$

$$= 0.1 \times -1.0065 \times 0.4505$$

$$= -0.0453$$

$$\Delta b1_2 = \alpha s_2 b1_2$$

$$= 0.1 \times 0 \times 0.0838$$

$$= 0$$

$$\vdots$$

$$\Delta b1_{64} = \alpha s_{64}b1_{64}$$

$$= 0.1 \times -0.2262 \times 0.8258$$

$$= -0.0187$$

8. Memperbarui bobot dan bias pada masing-masing lapisan menggunakan nilai pada proses sebelumnya.

$$w_{ki} = w_{ki} + \Delta w_{ki}$$

$$w_{1,1} = w_{1,1} + \Delta w_{1,1}$$

$$= 0.0975 + 0$$

$$= 0.0975$$

$$w_{1,2} = w_{1,2} + \Delta w_{1,2}$$

$$= -0.8003 + (-0.4372)$$

$$= -1.2375$$

$$\vdots$$

$$w_{1,1024} = w_{1,1024} + \Delta w_{1,1024}$$

$$= 0.2238 + 0$$

$$= 0.2238$$

$$w_{2,1} = w_{2,1} + \Delta w_{2,1}$$

$$= -0.2785 + 0$$

$$= -0.2785$$

$$\vdots$$

$$w_{64,1024} = w_{64,1024} + \Delta w_{64,1024}$$
$$= -0.4984 + 0$$
$$= -0.4984$$

$$b1_{1} = b1_{1} + \Delta b1_{1}$$

$$= 0.4505 + (-0.0453)$$

$$= 0.4052$$

$$b1_{2} = b1_{2} + \Delta b1_{2}$$

$$= 0.0838 + 0$$

$$= 0.0838$$

$$\vdots$$

$$b1_{64} = b1_{64} + \Delta b1_{64}$$

$$= 0.8258 + (-0.0187)$$

$$= 0.8071$$

$$v_{jk} = v_{jk} + \Delta v_{jk}$$

$$v_{1,1} = v_{1,1} + \Delta v_{1,1}$$

$$= 0.7572 + (-0.3372)$$

$$= 0.4200$$

$$v_{1,2} = v_{1,2} + \Delta v_{1,2}$$

$$= 0.5308 + 0$$

$$= 0.5308$$

$$\vdots$$

$$v_{1,64} = v_{1,64} + \Delta v_{1,64}$$

$$= 0.5688 + (-0.1369)$$

$$= 0.4319$$

$$v_{2,1} = v_{2,1} + \Delta v_{2,1}$$

$$= -0.7537 + 0.1511$$

$$= -0.6026$$

$$\vdots$$

$$v_{4,64} = v_{4,64} + \Delta v_{4,64}$$
$$= 0.7482 + 0.0308$$
$$= 0.7790$$

$$b2_1 = b2_1 + \Delta b2_1$$
$$= 0.0918 + (-0.9961)$$
$$= -0.9043$$

$$b2_2 = b2_2 + \Delta b2_2$$
$$= -0.0032 + (-0.0782)$$
$$= -0.0812$$

$$b2_3 = b2_3 + \Delta b2_3$$
$$= -0.0134 + (-0.4427)$$
$$= -0.4561$$

$$b2_4 = b2_4 + \Delta b2_4$$
$$= 0.0022 + 0.1067$$
$$= 0.1089$$

Sehingga didapatkan bobot dan bias baru dari masing-masing lapisan sebagai berikut.

$$\mathbf{W} = \begin{pmatrix} 1 & 2 & 3 & 4 & \cdots & 1024 \\ 0.0975 & -1.2375 & 0.9340 & -0.1367 & \cdots & 0.2238 \\ -0.2785 & 0.0176 & 0.6787 & 0.1262 & \cdots & -0.6797 \\ 0.5469 & -0.5162 & 0.7577 & -0.0216 & \cdots & 0.6551 \\ 0.9575 & 0.9317 & 0.7743 & 0.8841 & \cdots & 0.1626 \\ 5 & -0.9649 & 0.7112 & -0.3922 & 0.6741 & \cdots & -0.1190 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 64 & 0.1576 & 0.8116 & 0.6555 & 0.3242 & \cdots & -0.4984 \\ & & & & & & & & & & & & & & & & \\ 0.4200 & 0.5308 & 0.0271 & -0.0614 & \cdots & 0.4319 \\ -0.7160 & 0.7792 & 0.4401 & 0.7021 & \cdots & 0.6934 \\ -0.4640 & -0.9340 & -0.1122 & -0.2211 & \cdots & -0.5274 \\ 0.5077 & 0.1299 & 0.7741 & -0.7256 & \cdots & 0.7790 \end{pmatrix}$$

$$\mathbf{b1} = \begin{pmatrix} 0.4052 & 0.0838 & -0.1271 & -0.9133 & 0.1458 & \cdots & 0.8071 \end{pmatrix}$$

$$\mathbf{b2} = \begin{pmatrix} -0.9034 & -0.0812 & -0.4561 & 0.1089 \end{pmatrix}$$

Perhitungan diulang kembali dari langkah 2 menggunakan bobot dan bias baru yang sudah didapatkan. Perhitungan dilakukan hingga nilai *error* pada langkah 4 lebih kecil dari nilai *error* yang diinisialisasi pada langkah pertama.

## 4.3. Pengujian dan Evaluasi Model

Pengujian pada model yang telah optimal dilakukan menggunakan 460-463 data. Uji coba yang dilakukan pada penelitian ini yaitu mengubah nilai faktor dilasi sebesar 2,4,6, dan 8. Selain itu juga akan dilakukan perbandingan hasil antara model yang menggunakan augmentasi data dan model tanpa menggunakan augmentasi data. Hasil uji coba yang dilakukan dapat dilihat pada Tabel 4.1 dan Tabel 4.2.

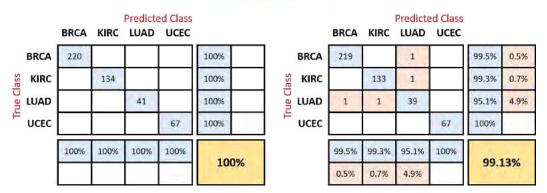
Tabel 4.1 Hasil Klasifikasi Model Tanpa Augmentasi Data

Faktor Dilasi	k-fold	Data Latih	Data Uji	Akurasi
2	1	1385	461	98.70%
	2	1384	462	99.13%
	3	1383	463	99.35%
	4	1386	460	99.13%
	Average		7//	99.08%
	1	1383	463	98.70%
4	2	1385	461	99.35%
	3	1384	462	100.0%
	4	1386	460	98.91%
Average				99.24%
	1	1385	461	99.35%
6	2	1384	462	98.70%
U	3	1384		98.92%
	4	1385	461	99.13%
Average				99.03%
	1	1383	463	98.70%
8	2	1384	462	99.13%
	3	1385	461	99.57%
	4	1386	460	99.13%
	99.13%			

Tabel 4.2 Hasil Klasifikasi Model Dengan Augmentasi Data

Fak	tor Dilasi	k-fold	Data Latih	Data Uji	Akurasi
2		1	1385	461	98.26%
	2	2	1384	462	98.70%
	Z	3	1383	463	98.92%
		4	1386	460	98.48%
Average					98.59%
4		1	1383	463	98.06%
	1	2	1385	461	98.48%
	4	3	1384	462	99.13%
		4	1386	460	98.91%
Average				98.65%	
6	7/	1	1385	461	98.70%
	6	2	1384	462	98.05%
	O	3	1384	462	98.05%
		4	1385	461	98.92%
- 4		Average			98.43%
8		1	1383	463	97.84%
	O	2	1384	462	98.48%
	8	3	1385	461	98.92%
		4	1386	460	98.48%
1	1	Ave	rage		98.43%
7.1					

Pada tabel tersebut dapat dilihat bahwa hasil akurasi tertinggi dari masing-masing model yang diuji yaitu 99.24% dan 98.65%. Hasil tersebut didapatkan dengan nilai faktor dilasi sebesar 4. Matriks konfusi dari masing-masing model dapat dilihat pada Gambar 4.15.



Gambar 4.15 (kiri) Model Tanpa Augmentasi Data; (kanan) Model Dengan Augmentasi Data

Model tanpa menggunakan augmentasi data dapat mengklasifikan kanker dengan sangat baik dengan akurasi, sensitivitas dan spesifisitas yang didapatkan masing-masing sebesar 100%. Sedangkan model yang menggunakan augmentasi dapat mengklasifikasikan kanker dengan baik dengan akurasi, sensitivitas, dan spesifisitas masing-masing sebesar 99.13%, 98.48%, dan 98.48%. Perhitungan akurasi, sensitivitas, dan spesifisitas dilakukan menggunakan Persamaan 2.23 hingga 2.29.

$$TTP_{1} = \sum_{j=1,j=i}^{4} x_{1,j} \qquad TTP_{3} = \sum_{j=1,j=i}^{4} x_{3,j}$$

$$= x_{1,1} \qquad = x_{3,3}$$

$$= 219 \qquad = 39$$

$$TTP_{2} = \sum_{j=1,j=i}^{4} x_{2,j} \qquad TTP_{4} = \sum_{j=1,j=i}^{4} x_{4,j}$$

$$= x_{2,2} \qquad = x_{4,4}$$

$$= 133 \qquad = 67$$

$$akurasi = \frac{TTP}{\sum_{j=1}^{4} \sum_{k=1}^{4} x_{jk}} \times 100\%$$

$$= \frac{TTP_1 + TTP_2 + TTP_3 + TTP_4}{x_{1,1} + x_{1,2} + x_{1,3} + x_{1,4} + x_{2,1} + x_{2,2} + \dots + x_{4,4}} \times 100\%$$

$$= \frac{219 + 133 + 39 + 67}{219 + 0 + 1 + 0 + 0 + 133 + 1 + \dots + 67} \times 100\%$$

$$= 99.13\%$$

$$TFN_1 = \sum_{j=1,j\neq 1}^4 x_{1,j}$$
  $TFN_2 = \sum_{j=1,j\neq 1}^4 x_{2,j}$   
=  $x_{1,2} + x_{1,3} + x_{1,4}$  =  $x_{2,1} + x_{2,3} + x_{2,4}$   
= 1

$$TFN_3 = \sum_{j=1, j\neq 1}^4 x_{3,j} \qquad TFN_4 = \sum_{j=1, j\neq 1}^4 x_{4,j}$$
$$= x_{3,1} + x_{3,2} + x_{3,4} \qquad = x_{4,1} + x_{4,2} + x_{4,3}$$
$$= 2 \qquad = 0$$

$$sens_1 = \frac{TTP_1}{TTP_1 + TFN_1} \times 100\% \qquad sens_3 = \frac{TTP_3}{TTP_3 + TFN_3} \times 100\%$$

$$= \frac{219}{219 + 1}100\% \qquad = \frac{39}{39 + 2}100\%$$

$$= 99.5\% \qquad = 95.1\%$$

$$sens_2 = \frac{TTP_2}{TTP_2 + TFN_2} \times 100\% \qquad sens_4 = \frac{TTP_4}{TTP_4 + TFN_4} \times 100\%$$

$$= \frac{133}{133 + 1}100\% \qquad = \frac{67}{67 + 0}100\%$$

$$= 99.3\% \qquad = 100\%$$

$$sensitivitas = \frac{\sum_{i=1}^{4} sens_i}{4}$$

$$= \frac{99.5\% + 99.3\% + 95.1\% + 1000\%}{4}$$

$$= 98.48\%$$

$$TFP_{1} = \sum_{j=1, j\neq 1}^{4} x_{j,1} \qquad TFP_{3} = \sum_{j=1, j\neq 1}^{4} x_{j,3}$$

$$= x_{2,1} + x_{3,1} + x_{4,1} \qquad = x_{1,3} + x_{1,3} + x_{4,3}$$

$$= 1 \qquad = 2$$

$$TFP_{2} = \sum_{j=1, j\neq 1}^{4} x_{j,2} \qquad TFP_{4} = \sum_{j=1, j\neq 1}^{4} x_{j,4}$$

$$= x_{1,2} + x_{3,2} + x_{4,2} \qquad = x_{1,4} + x_{2,4} + x_{3,4}$$

$$= 1 \qquad = 0$$

$$TTN_{1} = \sum_{j=1, j\neq 1}^{4} \sum_{k=1, k\neq 1}^{4} x_{j,k} \qquad TTN_{3} = \sum_{j=1, j\neq 3}^{4} \sum_{k=1, k\neq 3}^{4} x_{j,k}$$

$$= x_{2,2} + x_{2,3} + x_{2,4} + \dots + x_{4,4} + = x_{1,1} + x_{1,2} + \dots + x_{4,4} +$$

$$= 241 \qquad = 419$$

$$TTN_{2} = \sum_{j=1, j\neq 2}^{4} \sum_{k=1, k\neq 2}^{4} x_{j,k} \qquad TTN_{4} = \sum_{j=1, j\neq 4}^{4} \sum_{k=1, k\neq 4}^{4} x_{j,k}$$

$$= x_{1,1} + x_{1,3} + \dots + x_{4,4} + = x_{1,1} + x_{1,2} + \dots + x_{3,3} +$$

$$= 387 \qquad = 395$$

$$spes_{1} = \frac{TTN_{1}}{TTN_{1} + TFP_{1}} \times 100\% \qquad spes_{3} = \frac{TTN_{3}}{TTN_{3} + TFP_{3}} \times 100\%$$

$$= \frac{241}{241 + 1} 100\% \qquad = \frac{39}{39 + 2} 100\%$$

$$= 99.5\% \qquad spes_{4} = \frac{TTN_{4}}{TTN_{4} + TFP_{4}} \times 100\%$$

$$= 99.3\% \qquad spes_{4} = \frac{67}{67 + 0} 100\%$$

$$= 99.3\% \qquad spes_{4} = \frac{67}{67 + 0} 100\%$$

$$= 100\%$$

$$spesifisitas = \frac{\sum_{i=1}^{4} spes_{i}}{4}$$

$$= 99.5\% + 99.3\% + 95.1\% + 1000\%$$

$$= 99.48\%$$

Berdasarkan hasil tersebut, model tanpa menggunakan augmentasi data mendapatkan hasil yang lebih baik karena proses augmentasi data akan mempengaruhi susunan data yang dimiliki. Data yang digunakan merupakan data urutan RNA dimana proses penyusunannya menggunakan ketentuan dari metode yang digunakan. Metode yang berbeda akan memberikan urutan yang berbeda dan

akan merepresentasikan yang berbeda juga. Proses augmentasi data bisa merubah susunan tersebut sehingga model akan lebih sulit untuk mengenali data yang digunakan. Metode *dilated convolutional neural network* memberikan hasil klasifikasi lebih baik dibandingkan dengan metode *convolutional neural network* menggunakan arsitektur yang sama. Perbandingan dari kedua metode tersebut dapat dilihat pada Tabel 4.3.

Tabel 4.3 Perbandingan Hasil sensitivitas Pada Masing-masing Kelas

	(Lyu and Haque, 2018)	(Khalifa et al., 2020)	Penelitian Ini	
Dataset	Cancer Genom Atlas	Tumor Gene Expression		
Arsitektur	23 layers	14 layers		
<b>BRCA</b>	99.0%	98.3%	99.7%	
KIRC	95.0%	98.2%	99.3%	
LUAD	95.0%	84.8%	96.3%	
UCEC	96.0 <mark>%</mark>	96.4%	99.6%	
Rata-rata	96.25%	94.43%	98.70%	

Pada penelitian ini didapatkan hasil sensitivitas sebesar 98.70% dengan model tanpa menggunakan augmentasi data. Metode dilated convolutional neural network yang digunakan mendapatkan hasil yang lebih baik dibandingkan dengan beberapa penelitian sebelumnya yaitu oleh (Lyu and Haque, 2018) dengan hasil rata-rata sensitivitas sebesar 96.25% dan penelitian (Khalifa et al., 2020) menggunakan arsitektur yang sama dengan penelitian ini mendapatkan hasil rata-rata sensitivitas sebesar 94.43%.

## 4.4. Integrasi Keilmuan

Pada penelitian ini didapatkan hasil bahwa metode *dilated convolutional* neural network dapat bekerja dengan baik dalam klasifikasi kanker berdasarkan data RNA. Hasil yang didapatkan yaitu 99.24% untuk model tanpa menggunakan augmentasi data, dan 98.65% untuk model tanpa menggunakan augmentasi data. Selain itu, pada penelitian ini juga didapatkan hasil bahwa model tanpa

menggunakan augmentasi data mendapatkan hasil yang lebih baik dibandingkan model dengan augmentasi data. Hal ini sesuai dengan firman Allah SWT dalam Q.S. Ar-Rum/28:80 yang berbunyi:

## وَمِنْ اٰیٰتِهٖ خَلْقُ السَّمٰوٰتِ وَالْاَرْضِ وَاخْتِلَافُ اَلْسِنَتِکُمْ وَاَلْوَانِکُمْۗ اِنَّ فِيْ ذَٰلِكَ لَاٰیْتٍ لِّلْعٰلِمِیْنَ

"Dan di antara tanda-tanda kekuasaan-Nya ialah menciptakan langit dan bumi dan berlain-lainan bahasamu dan warna kulitmu (yaitu kulit yang putih, hitam, merah, kuning, dan kehijauan, padahal kalian merupakan keturunan dari satu bapak dan satu ibu yang sama, dan kalian merupakan jenis yang sama yaitu jenis manusia; bahkan setiap diri kalian memiliki perbedaan dari orang lain). Sesungguhnya pada yang demikian itu benar-benar terdapat tanda-tanda bagi orang-orang yang mengetahui".

Maksud dari ayat tersebut bahwa manusia diciptakan oleh Allah SWT berbeda-beda. Setiap manusia akan selalu berbeda dengan manusia lainnya meskipun memiliki garis keturunan yang sama. Hal tersebut berlaku pada DNA manusia dimana setiap makhluk hidup pasti memiliki susunan DNA yang berbeda-berbeda. Apabila dalam suatu susunan DNA dilakukan perubahan, maka representasi dari susunan tersebut juga bisa berubah.

Hasil dari penelitian juga bisa digunakan sebagai alat untuk mengetahui jenis kanker yang diderita oleh pasien sehingga bisa mendapatkan obat dan penanganan yang tepat. Hal itu merupakan sebuah bentuk ikhtiar yang dilakukan oleh seorang hamba dalam proses penyembuhan penyakit yang diderita. Akan tetapi, dalam proses penyembuhan merupakan kuasa Allah SWT yang maha menyembuhkan segala macam penyakit sebagaimana firman Allah SWT dalam

Q.S. Asy-Syuara/28:80 dan Q.S. Al-Anbiyaa/21:83 yang berbunyi:

# وَإِذَا مَرِضْتُ فَهُوَ يَشْفِينِ

"Dan apabila aku (Ibrahim) sakit, Dia (Allah)-lah yang menyembuhkan diriku"

"Dan (ingatlah kisah) Ayub, ketika ia menyeru Tuhannya: "(Ya Tuhanku), sesungguhnya aku telah ditimpa penyakit dan Engkau adalah Tuhan yang Maha Penyayang di antara semua penyayang""

Kedua ayat di atas menjelaskan bahwa Nabi Ibrahim dan Nabi Ayub memiliki keyakinan bahwa ketika suatu saat tertimpa penyakit, maka hanya Allah SWT yang beliau yakini dapat menyembuhkan. Hal tersebut adalah contoh sikap tawakal dari seorang hamba terhadap Penciptanya yang akan selalu menyayangi hamba-Nya. Pada proses penyembuhan penyakit kanker, semua cara penyembuhan hanya merupakan usaha dan perantara, yang menentukan kesembuhan hanyalah Allah SWT. Oleh karena itu, klasifikasi penyakit kanker payudara, kanker paru-paru, kanker rahim, dan kanker ginjal merupakan bagian dari usaha manusia untuk mendapatkan kesembuhan.

### **BAB V**

## **PENUTUP**

## 5.1. Kesimpulan

Berdasarkan hasil yang didapatkan dari implementasi metode *dilated* convolutional neural network dalam klasifikasi kanker berdasarkan data RNA, kesimpulan yang dapat diambil adalah:

- 1. Metode yang digunakan merupakan termasuk kedalam metode *deep learning* dimana proses ekstraksi fitur dan klasifikasi terdapat dalam satu arsitektur. Metode tersebut dapat bekerja dengan baik dalam klasifikasi kanker dengan akurasi dan sensitivitas sebesar 99.24% dan 98.70%,.
- 2. Metode *dilated convolutional neural network* dapat bekerja lebih baik dibandingkan dengan metode *convolutional neural network* dengan arsitektur yang sama. Rata-rata akurasi yang didapatkan yaitu sebesar 99.24% dengan nilai faktor dilasi sebesar 4.
- 3. Metode augmentasi data dapat mempengaruhi hasil klasifikasi yang didapatkan. Pada Penelitian ini model tanpa augmentasi data dapat memberikan hasil akurasi yang lebih baik dibandingkan dengan model dengan augmentasi data. Hasil yang didapatkan dari masing-masing model yaitu 99.24% untuk model tanpa augmentasi data dan 98.65% untuk model dengan augmentasi data.

#### 5.2. Saran

Pada penelitian mengenai klasifikasi kanker berdasarkan RNA menggunakan dilated convolutional neural network masih memiliki beberapa kekurangan sehingga perlu adanya perbaikan. Beberapa hal yang disarankan untuk penelitian mendatang yaitu:

- Mengimplementasikan metode ini menggunakan data RNA lain yang memiliki kelas normal sehingga model yang dibangun dapat digunakan sebagai deteksi dini kanker berdasarkan data RNA.
- 2. Mengumpulkan lebih banyak data dari kelas yang sudah ada atau dari kelas lain agar model yang dibangun dapat mengklasifikasikan kanker kedalam kelas yang lebih banyak.

#### DAFTAR PUSTAKA

- Ahmad, A. (2017). Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *Yayasan Cahaya Islam, Jurnal Teknologi Indonesia*, (October).
- Al-Dhabyani, W., Fahmy, A., Gomaa, M., and Khaled, H. (2019). Deep learning approaches for data augmentation and classification of breast masses using ultrasound images. *International Journal of Advanced Computer Science and Applications*, 10(5):618–627.
- Arrofiqoh, E. N. and Harintaka, H. (2018). Implementasi Metode Convolutional Neural Network Untuk Klasifikasi Tanaman Pada Citra Resolusi Tinggi. *Geomatika*, 24(2):61.
- Berrar, D. (2019). Cross-Validation. In *Encyclopedia of Bioinformatics and Computational Biology*, volume 1, pages 542–545. Elsevier.
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., and Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6):394–424.
- Chong, P., Tan, Y. X. M., Guarnizo, J., Elovici, Y., and Binder, A. (2018). Mouse authentication without the temporal aspect What does a 2D-CNN learn? Proceedings - 2018 IEEE Symposium on Security and Privacy Workshops, SPW 2018, pages 15–21.

- Clark, D. P., Pasdernik, N. J., and McGehee, M. R. (2019). *Molecular Biology, Third Edition*. Academic Press.
- De Jong, W. (2005). Kanker, Apakah Itu? EGC.
- Dumoulin, V. and Visin, F. (2016). A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*, pages 1–31.
- Ferles, C., P, Y., and Naidoo, K. (2018). Cancer Types: RNA Sequencing Values from Tumor Samples/Tissues.
- Griffiths, A. J. F., Miller, J. H. ., Suzuki, D. T., Lewontin, R. C., and Gelbart, W. M. (2000). *An Introduction to Genetic Analysis*. WH Freeman and Company.
- Guillemot, M., Heusele, C., Korichi, R., Schnebert, S., and Chen, L. (2020). Breaking batch normalization for better explainability of deep neural networks through layer-wise relevance propagation. *arXiv*.
- Hamid, A. R. A., Umbas, R., Oka, A. A. G., Mochtar, C. A., Djatisoesanto, W., adi Soedarso, M., Safriadi, F., Danarto, Warli, S. M., Ali, A. Z., Seputra, K. P., Sihombing, A. T., and Hakim, L. (2019). *IAUI Pedoman Tatalaksana Kanker Ginjal 2019*. Ikatan Ahli Urologi Indonesia (IAUI).
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. R. (2012). Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580, pages 1–18.
- Ho, Y. and Wookey, S. (2020). The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling. *IEEE Access*, 8:4806–4813.
- Hsu, Y. H. and Si, D. (2018). Cancer Type Prediction and Classification Based on RNA-sequencing Data. *Proceedings of the Annual International Conference of*

- the IEEE Engineering in Medicine and Biology Society, EMBS, 2018-July:5374–5377.
- Kabir, M. F. and Ludwig, S. A. (2019). Classification Models and Survival Analysis for Prostate Cancer Using RNA Sequencing and Clinical Data. *Proceedings* 2019 IEEE International Conference on Big Data, Big Data 2019, pages 2736–2745.
- Ker, J., Wang, L., Rao, J., and Lim, T. (2018). Deep Learning Applications in Medical Image Analysis. *IEEE Access*, 6:9375–9389.
- Khalifa, N. E. M., Taha, M. H. N., Ezzat Ali, D., Slowik, A., and Hassanien, A. E. (2020). Artificial Intelligence Technique for Gene Expression by Tumor RNA-Seq Data: A Novel Optimized Deep Learning Approach. *IEEE Access*, 8:22874–22883.
- Kim, Y. (2014). Convolutional neural networks for sentence classification. *EMNLP* 2014 2014 Conference on Empirical Methods in Natural Language Processing, Proceedings of the Conference, pages 1746–1751.
- Kresno, S. B. (2011). MicroRNA dan implikasinya pada kanker. *Indonesian Journal of Cancer*, 5(3):119–27.
- Kudo, Y. and Aoki, Y. (2017). Dilated convolutions for image classification and object localization. *Proceedings of the 15th IAPR International Conference on Machine Vision Applications, MVA 2017*, pages 452–455.
- Kurniawan, F. and Ivandari (2017). Komparasi Algoritma Data Mining Untuk Klasifikasi Penyakit Kanker Payudara. *Jurnal Stmik*, XII(1):1–8.

- Labach, A., Salehinejad, H., and Valaee, S. (2019). Survey of Dropout Methods for Deep Neural Networks. *arXiv preprint arXiv:1904.13310*.
- Lei, X., Pan, H., and Huang, X. (2019). A dilated cnn model for image classification. *IEEE Access*, 7:124087–124095.
- Lyu, B. and Haque, A. (2018). Deep Learning Based Tumor Type Classification Using Gene Expression Data. ACM-BCB 2018 - Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, (December):89–96.
- Made, N., Adnyani, D., and Widiana, I. G. R. (2018). Diagnosis dan tatalaksana pasien karsinoma sel renal. *Jpdunud*, 2(2):23–27.
- Manliguez, C. (2016). Generalized Confusion Matrix for Multiple Classes.

  Machine Learning, (November):4–6.
- Melisa, J., Monoarfa, A., and Tjandra, F. (2016). Profil penderita karsinoma sel ginjal ( renal cell carcinoma ). *jurnal e-Clinic (eCI)*, 4(2).
- Mikolajczyk, A. and Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem. 2018 International Interdisciplinary PhD Workshop (IIPhDW), pages 117–122.
- Mufida, N., Mudatsir, M., and Hasbalah, K. (2015). Pengaruh Pelatihan Model Pengajaran Langsung Terhadap Pengetahuan dan Keterampilan Kader Kesehatan Tentang Pemeriksaan Payudara Sendiri (SADARI) Di Kecamatan Delima Kabupaten Pidie. *Jurnal Ilmu Keperawatan*, 3(2).
- Munir, K., Elahi, H., Ayub, A., Frezza, F., and Rizzi, A. (2019). Cancer Diagnosis Using Deep Learning: A. *Cancers*, 11(9):1235.

- Murphy, K. P. (2012). A Probabilistic Perspective. MIT Press.
- Nasional, K. P. K. (2015). Panduan Penatalaksanaan Kanker Payudara. Kementerian Kesehatan Republik Indonesia, pages 1, 12–4, 24–26, 45.
- Navin J R, M. and R, P. (2016). Performance Analysis of Text Classification Algorithms using Confusion Matrix. *International Journal of Engineering and Technical Research (IJETR)*, 6(4).
- Ng, W., Minasny, B., Montazerolghaem, M., Padarian, J., Ferguson, R., Bailey, S., and McBratney, A. B. (2019). Convolutional neural network for simultaneous prediction of several soil properties using visible/near-infrared, mid-infrared, and their combined spectra. *Geoderma*, 352(January):251–267.
- Nielsen, M. (2015). *Neural Networks and Deep Learning*. Determination Press.
- Novitasari, D. C. R., Hendradi, R., Caraka, R. E., Rachmawati, Y., Fanani, N. Z., Syarifudin, A., Toharudin, T., and Chen, R. C. (2020). Detection of COVID-19 chest x-ray using support vector machine and convolutional neural network. *Communications in Mathematical Biology and Neuroscience*, 2020(July):1–19.
- Nugraha, F. S., Shidiq, M. J., and Rahayu, S. (2019). Analisis Algoritma Klasifikasi Neural Network Untuk Diagnosis Penyakit Kanker Payudara. *Jurnal Pilar Nusa Mandiri*, 15(2):149–156.
- Nurfita, R. D. and Ariyanto, G. (2018). Implementasi Deep Learning Berbasis Tensorflow Untuk Pengenalan Sidik Jari. *Emitor: Jurnal Teknik Elektro*, 18(01):22–27.
- Nwankpa, C., Ijomah, W., Gachagan, A., and Marshall, S. (2018). Activation

- Functions: Comparison of trends in Practice and Research for Deep Learning. *arXiv preprint arXiv:1811.03378*, pages 1–20.
- Putra, R. E., Tjandrasa, H., and Suciati, N. (2020). Severity classification of non-proliferative diabetic retinopathy using convolutional support vector machine. International Journal of Intelligent Engineering and Systems, 13(4):156–170.
- Rajanbabu, A. (2015). Molecular Pathology and Cytogenetics of Endometrial Carcinoma, Carcinosarcoma, and Uterine Sarcomas. In *Uterine Cancer*, pages 13–23. Springer India, New Delhi.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088):533–536.
- Salser, W. A. (1974). DNA Sequencing Techniques. Annual Review of Biochemistry, 43(1):923–965.
- Shorten, C. and Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, 6(1).
- Suhardi, I. (2007). Evaluasi Pegaruh Fungsi Aktivasi Dan Parameter Kemiringannya Terhadap Unjukkerja Pengenalan (Studi Kasus pada Pengenalan Karakter Angka Tulisan Tangan). *Jurnal Ilmiah Semesta Teknika*, 10(1):53–68.
- Sumbono, A. (2019). Biomolekul. Deepublish.
- Suryo, J. (2010). Herbal Penyembuh Gangguan Sistem Pernapasan. Penerbit B First.
- Wuryandari, M. D. and Afrianto, I. (2012). Perbandingan Metode Jaringan Syaraf Tiruan Backpropagation Dan Learning Vector Quantization Pada Pengenalan Wajah. *Komputa*, 1(1):45–51.

- Zhang, Y. H., Huang, T., Chen, L., Xu, Y. C., Hu, Y., Hu, L. D., Cai, Y., and Kong, X. (2017). Identifying and analyzing different cancer subtypes using RNAseq data of blood platelets. *Oncotarget*, 8(50):87494–87511.
- Zhao, J., Mao, X., and Chen, L. (2018). Learning deep features to recognise speech emotion using merged deep CNN. *IET Signal Processing*, 12(6):713–721.
- Zhen, X., Chakraborty, R., Vogt, N., Bendlin, B., and Singh, V. (2019). Dilated Convolutional Neural Networks for Sequential Manifold-Valued Data. In 2019 IEEE/CVF International Conference on Computer Vision (ICCV), volume 176, pages 10620–10630. IEEE.
- Zufar, M. and Setiyono, B. (2016). Convolutional Neural Networks Untuk Pengenalan Wajah Secara Real-time. *Jurnal Sains dan Seni ITS*, 5(2):128862.