

**KLASIFIKASI SPERMA MANUSIA BERDASARKAN MORFOLOGI  
DENGAN MENGGUNAKAN DEEP LEARNING  
CONVOLUTIONAL NEURAL NETWORK (CNN)**

**(Skripsi)**

**Oleh**

**MUHAMMAD NUR ASHIDDIQI  
NPM 1857051014**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

**KLASIFIKASI SPERMA MANUSIA BERDASARKAN MORFOLOGI  
DENGAN MENGGUNAKAN DEEP LEARNING  
CONVOLUTIONAL NEURAL NETWORK (CNN)**

**Oleh**

**MUHAMMAD NUR ASHIDDIQI**

**Skripsi**

**Sebagai Salah Satu Syarat untuk Mencapai Gelar  
SARJANA KOMPUTER**

**Pada**

**Jurusan Ilmu Komputer  
Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Lampung**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
UNIVERSITAS LAMPUNG  
BANDAR LAMPUNG  
2024**

## ABSTRAK

### KLASIFIKASI SPERMA MANUSIA BERDASARKAN MORFOLOGI DENGAN MENGGUNAKAN DEEP LEARNING CONVOLUTIONAL NEURAL NETWORK (CNN)

Oleh

**Muhammad Nur Ashiddiqi**

Reproduksi merupakan elemen vital untuk kelangsungan hidup dan menjadi fokus penelitian ini terkait infertilitas. Kondisi ketidaksuburan yang muncul setelah pasangan menikah dalam jangka waktu yang cukup lama tanpa tanda-tanda kehamilan. Dampak infertilitas dapat dirasakan baik oleh perempuan maupun pria, dengan salah satu penyebab pada pria adalah kualitas sperma yang *suboptimal* atau adanya kelainan pada sperma. Dalam era teknologi saat ini *Computer Aided Sperm Analysis* (CASA) membuka peluang baru untuk menilai kualitas sperma dengan memanfaatkan kecerdasan buatan. Penelitian ini membandingkan dua model *Convolutional Neural Network* (CNN) yakni *MobileNetV2* dan *EfficientNetB3* untuk melakukan analisis morfologi sperma menggunakan 3000 citra dengan tiga kelas yaitu sperma cacat (*abnormal sperm*), bukan sperma (*non sperm*), dan sperma normal (*normal sperm*). Hasil pengembangan model pembelajaran *MobileNetV2* dan *EfficientNetB3* mencapai *accuracy*, *precision*, *recall* dan *f1score* dengan hasil yang cukup memuaskan.

Kata kunci: CNN, *EfficientNetB3*, *MobileNetV2*, Reproduksi, *Sperm Morphology Image Data Set* (SMIDS).

## **ABSTRACT**

### **CLASSIFICATION OF HUMAN SPERM BASED ON MORPHOLOGY USING DEEP LEARNING CONVOLUTIONAL NEURAL NETWORK (CNN)**

**By**

**Muhammad Nur Ashiddiqi**

Reproduction as a vital element for survival, is the focus of study related to infertility, a condition of infertility can be felt by both women and men, with one of the causes in men being suboptimal sperm quality or sperm abnormalities. In the current technological era, computer aided sperm analysis (CASA) opens up new opportunities to assess sperm quality by utilizing artificial intelligence. This study compared two convolutional neural network models, namely mobileNetV2 and EfficientNetB3, to perform sperm morphology analysis using 3000 images with three main classification classes : abnormal Sperm, non-sperm, and normal sperm. The result of the learning model development showed positive performance, especially on the mobilenetv2 and efficientnetb3 architectures, with both models achieving accuracy, precision, recall, and f1score with satisfactory results.

**Keywords:** CNN, EfficientNetB3, MobileNetV2, Reproduction, Sperm Morphology Image Data Set (SMIDS).

**Judul Skripsi** : **KLASIFIKASI SPERMA MANUSIA  
BERDASARKAN MORFOLOGI DENGAN  
MENGUNAKAN DEEP LEARNING  
CONVOLUTIONAL NEURAL NETWORK  
(CNN)**

**Nama Mahasiswa** : **Muhammad Nur Ashididiqi**

**Nomor Pokok Mahasiswa** : **1857051014**

**Program Studi** : **S1 Ilmu Komputer**

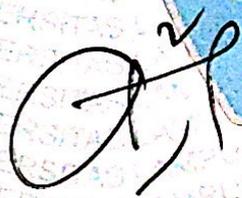
**Jurusan** : **Ilmu Komputer**

**Fakultas** : **Matematika dan Ilmu Pengetahuan Alam**



**MENYETUJUI**

**1. Komisi Pembimbing**

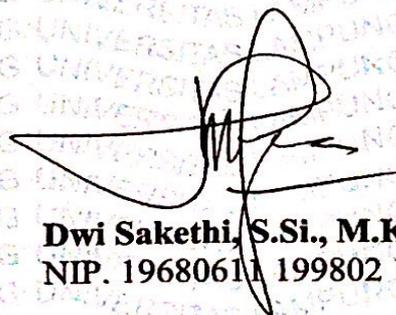


**Dr. Aristoteles, S.Si., M.Si.**  
**NIP. 19810521 200604 1 002**



**Favorisen R. Lumbanraja, Ph.D.**  
**NIP. 19830110 200812 1 002**

**2. Ketua Jurusan Ilmu Komputer**



**Dwi Sakethi, S.Si., M.Kom.**  
**NIP. 19680611 199802 1 001**

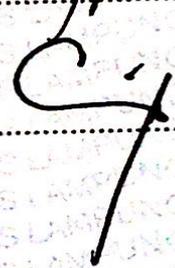
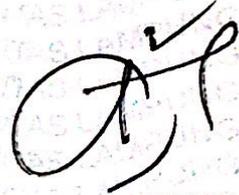
**MENGESAHKAN**

**1. Tim Penguji**

**Ketua Penguji : Dr. Aristoteles, S.Si., M.Si.**

**Sekretaris Penguji : Favorisen R. Lumbanraja, Ph.D.**

**Penguji Utama : Rico Andrian, S.Si., M.Kom.**



**2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam**



**Dr. Eng. Heri Satria, S.Si., M.Si.**  
**NIP. 19711001 200501 1 002**

**Tanggal Lulus Ujian Skripsi: 20 Mei 2024**

## PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Muhammad Nur Ashiddiqi

NPM : 1857051014

Dengan ini menyatakan bahwa skripsi saya yang berjudul **“Klasifikasi Sperma Manusia Berdasarkan Morfologi Dengan Menggunakan Deep Learning Convolutional Neural Network (CNN)”** merupakan karya saya sendiri dan bukan karya orang lain. Semua tulisan yang tertuang dalam skripsi ini telah mengikuti kaidah penulisan karya ilmiah Universitas Lampung. Apabila dikemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka bersedia menerima sanksi berupa pencabutan gelar yang telah saya terima.

Bandar Lampung, 20 Mei 2024



Muhammad Nur Ashiddiqi  
NPM. 1857051014

## RIWAYAT HIDUP



Lahir di Bandar Lampung, pada tanggal 5 April 1999, sebagai anak ketiga dari empat bersaudara. Menyelesaikan pendidikan dasar di SDS Al-Kautsar Bandar Lampung dan selesai pada tahun 2011. Pendidikan menengah pertama di SMPN 22 Bandar Lampung yang diselesaikan pada tahun 2014, lalu melanjutkan ke jenjang pendidikan menengah atas di SMKS Yadika Bandar Lampung yang diselesaikan pada tahun 2017.

Pada tahun 2018, penulis terdaftar sebagai mahasiswa Jurusan Ilmu Komputer, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur SMMPTN. Selama menjadi mahasiswa melakukan beberapa kegiatan antara lain:

1. Menjadi anggota Adapter Himpunan Mahasiswa Jurusan Ilmu Komputer pada periode 2018/2019.
2. Menjadi anggota Bidang Keilmuan Himpunan Mahasiswa Jurusan Ilmu Komputer periode 2019/2020.
3. Menjadi anggota Media Informasi Badan Eksekutif Mahasiswa Fakultas periode 2020/2021.
4. Menjadi asisten dosen Sistem Operasi Jurusan Ilmu Komputer pada Semester Genap 2019/2020.
5. Menjadi asisten dosen Keamanan Sistem Informasi Jurusan Ilmu Komputer pada Semester Ganjil Tahun Ajaran 2020/2021
6. Menjadi asisten dosen Sistem Informasi Geografis Jurusan Ilmu Komputer pada Semester Genap Tahun Ajaran 2020/2021.

7. Melaksanakan Kerja Praktek pada juli periode 2020/2021 di Badan Pelaksana Kuliah Kerja Nyata Universitas Lampung.
8. Melaksanakan KKN di Desa Hajimena, Kecamatan Natar, Kabupaten Lampung Selatan pada tahun 2021.
9. Mengikuti ujian sertifikasi dan mendapat sertifikat *Junior Network Administrator* oleh Lembaga Sertifikasi Profesi Informatika pada tahun 2020.
10. Mengikuti ujian sertifikasi dan mendapat sertifikat *Junior Web Developer* oleh Lembaga Sertifikasi Profesi Informatika pada tahun 2021.

## **MOTTO**

*“Success is not final; failure is not fatal, it is the courage to continue the counts.”*

**(Winston Churchill)**

“Nakal Boleh, Bodoh Jangan.”

**(Anonim)**

*“Don’t let yesterday take up too much of today.”*

**(Will Rogers)**

## **PERSEMBAHAN**

Puji dan syukur kehadirat Allah SWT karena atas berkat dan rahmat serta karunia-Nya sehingga saya dapat menyelesaikan skripsi ini dengan baik.

Shalawat serta salam selalu saya sanjung agungkan kepada Nabi Muhammad SAW yang telah membawa umat manusia menuju zaman yang terang benderang ini dan syafaatnya senantiasa dinantikan di yaumul akhir kelak.

Kupersembahkan karya ini kepada  
**Kedua Orang Tuaku Tersayang**

Yang selalu mendukung dan memberi doa tanpa henti pada setiap langkah saya. Saya sangat berterima kasih atas segala kasih sayang, didikan, dan perjuangan yang telah diberikan hingga sekarang, yang tak akan pernah bisa saya balas sepenuhnya.

**Seluruh Keluarga Besar Ilmu Komputer 2018**  
Yang selalu memberikan semangat dan dukungan.

**Almamater Tercinta Universitas Lampung dan Jurusan Ilmu Komputer**  
Tempat mengemban semua ilmu untuk menjadi bekal hidup.

## SANWACANA

Puji syukur kehadirat Allah SWT atas berkat rahmat dan hidayah serta karunia-Nya, saya dapat menyelesaikan skripsi dengan judul **“Klasifikasi Sperma Manusia Berdasarkan Morfologi Dengan Menggunakan Deep Learning Convolutional Neural Network (CNN)”**. Dalam melaksanakan penelitian dan pembuatan skripsi ini, saya mendapat banyak bimbingan dan dukungan dari berbagai pihak, sehingga pada kesempatan ini saya ingin menyampaikan ungkapan terima kasih kepada:

1. Kedua orang tua tercinta yaitu Almarhum Bapak R. Supriyanto dan Ibu Sri Kusmawati serta kakakku yaitu Tito Riyan Wicaksana dan Tita Maulidya serta adikku tersayang Sabiluna Khoirunnisa yang selalu memberi dukungan, semangat, motivasi, kasih sayang serta doa tiada henti.
2. Bapak Dr. Aristoteles, S.Si., M.Si. selaku Dosen Pembimbing utama yang telah memberikan bimbingan kepada saya dalam menyelesaikan skripsi dengan baik.
3. Bapak Favorisen R. Lumbanraja, Ph.D. selaku Dosen Pembimbing kedua yang telah memberikan bimbingan kepada saya dalam menyelesaikan skripsi dengan baik.
4. Bapak Rico Andrian, S.Si., M.Kom. selaku Dosen Pembahas dan Dosen Pembimbing Akademik yang telah memberikan banyak masukan dalam menyelesaikan skripsi ini.
5. Bapak Dr. Eng. Heri Satria, S.Si., M.Si. selaku Dekan FMIPA Universitas Lampung.
6. Bapak Dwi Sakethi, S.Si., M.Kom. selaku Ketua Jurusan Ilmu Komputer FMIPA Universitas Lampung.

7. Ibu Anie Rose Irawati, S.T., M.Cs. selaku Sekretaris Jurusan Ilmu Komputer FMIPA Universitas Lampung.
8. Ibu Ade Nora Maela, Bang Zainuddin dan Mas Nofal yang telah membantu segala urusan administrasi penulis di Jurusan Ilmu Komputer
9. Bapak dan Ibu Dosen Jurusan Ilmu Komputer FMIPA Universitas Lampung yang membantu saya dalam perkuliahan.
10. Teman-teman grup 2B2F Arfina Shella Meilany, Arafia Isnayu Akaf, Bobby Malela Hutagalung, lalu teman-teman grup sakanigadik dan grup Kompre Juni yang saling bantu-membantu dalam proses pengerjaan skripsi ini.
11. Teman-teman Ilmu Komputer angkatan 2018 yang telah memberikan pengalaman yang sangat menyenangkan bersama-sama.
12. Seluruh kakak tingkat dan adik tingkat Ilmu Komputer yang tidak bisa disebutkan satu persatu yang telah menjadi warna selama masa perkuliahan penulis.
13. Teman-teman Himakom dan BEM-F yang sudah mengajarkan banyak hal dalam berorganisasi, memberikan banyak pengalaman

Penulis Menyadari bahwa skripsi ini masih jauh dari kata sempurna, semoga skripsi ini membawa manfaat dan keberkahan bagi semua civitas Ilmu Komputer Universitas Lampung Aamiin Ya Rabbal Aalamiin.

Bandar Lampung, 20 Mei 2024

Muhammad Nur Ashiddiqi  
NPM. 1857051014

## DAFTAR ISI

	Halaman
<b>ABSTRAK</b> .....	iii
<b>MENGESAHKAN</b> .....	vi
<b>PERNYATAAN</b> .....	vii
<b>RIWAYAT HIDUP</b> .....	viii
<b>MOTTO</b> .....	ix
<b>PERSEMBAHAN</b> .....	x
<b>SANWACANA</b> .....	xi
<b>DAFTAR ISI</b> .....	xiii
<b>DAFTAR TABEL</b> .....	xvi
<b>DAFTAR GAMBAR</b> .....	xvii
<b>DAFTAR KODE PROGRAM</b> .....	xviii
<b>I. PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Batasan Masalah .....	3
1.4 Tujuan.....	4
1.5 Manfaat.....	4
<b>II. TINJAUAN PUSTAKA</b> .....	5
2.1 Penelitian Terdahulu.....	5
2.3 Sperma.....	8
2.4 Morfologi.....	8
2.5 Citra Digital .....	9

2.6	Pra-pemrosesan .....	10
2.7	<i>Machine Learning</i> (ML).....	11
2.8	<i>Deep Learning</i> .....	12
2.9	<i>Convolutional Neural Network</i> (CNN) .....	12
2.9.1	<i>EfficientNetB3</i> .....	13
2.9.2	<i>MobileNetV2</i> .....	14
2.9.3	<i>Feature Extraction</i> .....	14
2.9.4	<i>Classification Feature</i> .....	17
2.10	<i>Confusion matrix</i> .....	19
2.11	<i>Classification Report</i> .....	20
2.11.1	<i>Accuracy</i> .....	20
2.11.2	<i>Recall</i> .....	20
2.11.3	<i>Precision</i> .....	20
2.11.4	<i>F1-Score</i> .....	20
<b>III.</b>	<b>METODOLOGI PENELITIAN</b> .....	<b>21</b>
3.1	Tempat dan Waktu Penelitian .....	21
3.1.1	Tempat Penelitian.....	21
3.1.2	Waktu Penelitian .....	21
3.2	Data dan Alat.....	23
3.2.1	Data .....	23
3.2.2	Alat.....	23
3.3	Alur Kerja Penelitian .....	25
3.3.1	Pengumpulan Data .....	26
3.3.2	Pra-pemrosesan Data.....	26
3.3.3	Pemisahan Data.....	27
3.3.4	Augmentasi Data.....	27
3.3.5	Model <i>Convolutional Neural Network</i> (CNN).....	28
3.3.6	Evaluasi.....	28
<b>IV.</b>	<b>HASIL DAN PEMBAHASAN</b> .....	<b>29</b>
4.1	Implementasi Program .....	29
4.1.1	<i>Import Library</i> .....	29
4.1.2	Pra-pemrosesan .....	30
4.1.3	Pemisahan Data.....	32
4.1.4	Augmentasi Data.....	36
4.1.5	Arsitektur <i>Convolutional Neural Network</i> (CNN).....	38
4.1.6	Pelatihan Model .....	40
4.1.7	Pengujian Model .....	41
4.2	Evaluasi Hasil.....	42
4.3	Perbandingan Dengan Penelitian Sebelumnya.....	46

<b>V. SIMPULAN DAN SARAN .....</b>	<b>48</b>
5.1 Kesimpulan.....	48
5.2 Saran.....	49
<b>DAFTAR PUSTAKA.....</b>	<b>50</b>

## DAFTAR TABEL

Tabel	Halaman
1. Penelitian terkait.....	5
2. <i>Confusion matrix</i> (Han & Kamber, 2011).....	19
3. Alur Waktu Pengerjaan Penelitian.....	22
4. Jumlah Dataset.....	26
5. <i>Hyperparameter</i> .....	28
6. Hasil perbandingan klasifikasi tanpa augmentasi data.....	42
7. Hasil perbandingan klasifikasi dengan augmentasi data.....	43
8. Perbandingan penelitian sebelumnya.....	46

## DAFTAR GAMBAR

Gambar	Halaman
1. Struktur Sperma Manusia Normal .....	8
2. Morfologi Sperma Manusia. ....	9
3. Pra-pemrosesan data.....	11
4. Arsitektur <i>Convolutional Neural Network</i> .....	12
5. Arsitektur <i>EffecientNetB3</i> (Miftahuddin & Zaelani, 2022). ....	13
6. Arsitektur <i>MobileNetV2</i> (Hastomo, et al., 2021). ....	14
7. <i>Feature Extraction</i> . ....	15
8. <i>Convolutional Layer</i> . ....	15
9. <i>Pooling Layer</i> .....	16
10. Aktivasi <i>ReLU</i> .....	17
11. <i>Classification Feature</i> .....	17
12. <i>Sperm Morphology Image Data Set</i> .....	23
13. Alur Penelitian. ....	25
14. Pra-pemrosesan Data.....	27
15. Augmentasi Data. ....	28
16. Hasil Pra-pemrosesan data. ....	32
17. Grafik pemisahan data skenario pertama. ....	34
19. Grafik Pemisahan data skenario kedua. ....	36
20. Grafik perbandingan tanpa augmentasi data. ....	44
21. Grafik perbandingan dengan augmentasi data. ....	45
22. Grafik perbandingan penelitian sebelumnya.....	47

## DAFTAR KODE PROGRAM

Kode Program	Halaman
1. <i>Import library</i> .....	30
2. Pra-pemrosesan dan Import data. ....	31
3. Skenario pertama pemisahan data. ....	33
4. Skenario kedua pemisahan data. ....	35
5. Parameter Augmentasi Data.....	37
6. Model <i>Transfer Learning EfficientNetB3</i> .....	38
7. Model <i>Fine Tuning EfficientNetB3</i> . ....	39
8. Model <i>Transfer Learning MobileNetV2</i> .....	39
9. Model <i>Fine Tuning MobileNetV2</i> . ....	40
10. Pelatihan model tanpa augmentasi data. ....	40
11. Pelatihan model dengan augmentasi data latih. ....	40
12. Pengujian model dengan data testing.....	41

## I. PENDAHULUAN

### 1.1 Latar Belakang

Reproduksi merupakan kemampuan yang dimiliki makhluk hidup untuk menghasilkan keturunan baru (Sumiati, 2013). Bagi makhluk hidup tentu saja reproduksi ini sangat penting dalam kelangsungan hidup. Reproduksi pada manusia normal dilakukan dengan cara pembuahan dari sel gamet laki-laki dan sel gamet perempuan sehingga menghasilkan sel baru yang akan berkembang menjadi suatu zigot yang akan tumbuh sebagai cikal bakal manusia atau yang disebut juga dengan fertilisasi.

Sel spermatozoa (*Sperm Cell*) adalah sel gamet jantan umumnya diproduksi oleh laki-laki. Bentuk sel sperma pada umumnya berbentuk seperti kecebong, tersusun atas tiga bagian utama, yaitu kepala sperma (*sperm head*), bagian tengah sperma (*sperm midpiece*), dan ekor sperma (*sperm tail*) (Andrianto, 2017). Menurut WHO kelainan sperma terdapat pada bentuk kepala yaitu sperma cacat (*abnormal sperm*) dan sperma normal (*normal sperm*). Banyak kasus yang menyebabkan gagalnya pembuahan dalam fertilisasi yang disebabkan oleh kualitas sel sperma yang tidak dalam kondisi subur atau infertilitas. Banyak faktor-faktor penyebab turunnya produktivitas seorang pria seperti gaya hidup di masa kini, pola makanan, jenis makanan, dan kebiasaan sehari-hari yang tidak sehat membuat kualitas sperma menjadi buruk. Infertilitas adalah ketidakmampuan pasangan untuk menghasilkan anak setelah menikah dan melakukan senggama teratur tanpa memakai kontrasepsi

(Soebijanto, 2013). Infertilitas manusia dianggap sebagai penyakit serius pada sistem reproduksi yang mempengaruhi lebih dari 10% pasangan di seluruh dunia dan lebih dari 30% dari kasus yang dilaporkan terkait dengan pria (Iqbal, et al., 2020). Standar laboratorium untuk melihat kesuburan sperma masih dilakukan berdasarkan morfologi sperma (WHO, 2021).

Masalah yang ditemukan pada penelitian sebelumnya proses *classification* menggunakan algoritma SVM dimana parameter utamanya yaitu kernel dan nantinya objek akan dipisah menjadi beberapa kelas yang memberikan tingkat *accuracy* sebesar 82.33% (Ilhan, et al., 2018). Algoritma SVM hanya menganggap setiap gambar yang diinput sebagai vector (Xinyan & Weigang, 2019). Penelitian ini memiliki hasil yang lumayan baik, namun tingkat keberhasilan masih dapat ditingkatkan lagi dengan cara menambahkan atau mengganti metode yang sudah diterapkan. Pengujian lain yang dilakukan dengan menggunakan algoritma *Convolutional Neural Network* (CNN) untuk pengklasifikasian sperma dengan membandingkan dataset yang berbeda. CNN secara signifikan mengurangi jumlah parameter yang perlu dilatih dibandingkan algoritma klasifikasi lainnya (Ghamisi, et al., 2016).

*Convolution Neural Network* adalah suatu jaringan saraf yang berisikan beberapa *layer* seperti *convolution layer* dan *fully connected layer*. *Convolutional Neural Network* memfokuskan parameter inputan dengan cara di latih (Albert, et al., 2020). CNN memiliki lapisan khusus yang bernama lapisan konvolusi. Saat sebuah citra dimasukkan maka akan menghasilkan sebuah pola dari beberapa bagian citra yang nantinya akan lebih mudah untuk diklasifikasikan. Kelebihan CNN yaitu mampu mengklasifikasi dan memiliki hasil paling signifikan dalam pengenalan objek. CNN dapat menangani citra secara langsung dan dapat mengeksplorasi setiap piksel dengan operasi konvolusi yang membuatnya lebih baik dalam mengklasifikasi gambar yang ada (Xinyan & Weigang, 2019).

Berdasarkan uraian diatas, judul penelitian ini adalah “**Klasifikasi Sperma Manusia Berdasarkan Morfologi Menggunakan *Deep Learning Convolutional Neural Network* (CNN)**”.

## **1.2 Rumusan Masalah**

Berdasarkan pemaparan latar belakang tersebut, maka rumusan masalah dalam penelitian ini adalah sebagai berikut:

1. Bagaimana cara mengklasifikasikan sperma bentuk normal, abnormal dan bukan sperma melalui citra gambar dengan menggunakan algoritma *Convolutional Neural Network* (CNN).
2. Membandingkan hasil klasifikasi antar arsitektur *Convolutional Neural Network* (CNN) yang digunakan pada penelitian ini, dalam mengklasifikasikan tiga kelas bentuk sperma cacat, sperma normal, dan bukan sperma.

## **1.3 Batasan Masalah**

Adapun batasan masalah dalam penelitian ini, yaitu :

1. Penelitian ini akan dilakukan menggunakan pemodelan dengan metode *Convolutional Neural Network* (CNN) untuk mengklasifikasi tiga kelas bentuk sperma yaitu sperma cacat, sperma normal dan bukan sperma.
2. Data yang digunakan berjumlah 3000 citra gambar sperma diperoleh dari *Sperm Morphology Image Data Set* (SMIDS) yang dapat diakses melalui link alamat <https://data.mendeley.com/datasets/6xvdhc9fyb/1>.
3. Hasil pengujian prediksi menggunakan model yang telah dibuat akan ditunjukkan melalui beberapa parameter, antara lain *accuracy*, *precision*, *recall*, dan *f1-score*.
4. Tidak dikembangkannya sistem pada penelitian ini, melainkan hanya berupa model klasifikasi.

## 1.4 Tujuan

Adapun tujuan dari penelitian ini sebagai berikut:

1. Mengevaluasi kinerja *Convolutional Neural Network* (CNN) dalam mengklasifikasikan bentuk sperma.
2. Mengetahui *hyperparameter* yang optimal pada *Convolutional Neural Network* (CNN) sehingga menghasilkan pemodelan dengan kinerja yang maksimal.

## 1.5 Manfaat

Dalam penyusunan penelitian ini memiliki manfaat sebagai berikut:

1. Memberikan kemudahan untuk ilmu kedokteran dalam mengklasifikasi bentuk sperma cacat, sperma normal, dan bukan sperma.
2. Dapat digunakan sebagai referensi bagi pihak yang membutuhkan informasi dalam penerapan model *Convolutional Neural Network* (CNN).

## II. TINJAUAN PUSTAKA

### 2.1 Penelitian Terdahulu

Penelitian ini tidak lepas dari penelitian sebelumnya yang telah dilakukan, sehingga penelitian yang akan dilakukan memiliki hubungan antara persamaan dan perbedaan objek yang diteliti. Ringkasan dari penelitian terdahulu dapat dilihat pada Tabel 1.

Tabel 1. Penelitian terkait klasifikasi sperma.

Penelitian	Data	Metode	Hasil
<i>Dual Tree Complex Wavelet Transform Based Sperm Abnormality Classification</i> (Ilhan, et al., 2018)	536 citra sperma masing-masing kelas yaitu citra normal sebanyak 179, citra abnormal sebanyak 109, citra non-sperm sebanyak 248.	<i>Dual Tree Complex Wavelet, SVM classification</i>	<i>Accuracy: 82.3%</i>
<i>The Effect of Nonlinear Wavelet Transform Based Denoising in Sperm Abnormality Classification</i> (Ilhan, et al., 2018)	536 citra sperma masing-masing kelas yaitu citra normal sebanyak 179, citra abnormal sebanyak 109, citra non-sperm sebanyak 248.	<i>Adaptive Denoising feature extraction by SURF descriptor SVM classification</i>	<i>Accuracy: 83.4%</i>

Penelitian	Data	Metode	Hasil
<i>A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison with conventional methods</i> (Ilhan, et al., 2020)	3000 citra sperma tumpang tindih dengan masing-masing kelas yaitu citra normal sebanyak 1021, citra abnormal sebanyak 1005, citra non-sperm sebanyak 974.	<i>Overlapping group shrinkage fuzzy-c means clustering data augmentation classification by MobileNet</i>	<i>Accuracy: 87%</i>
<i>Automated sperm morphology analysis approach using a directional masking technique</i> (Ilhan, et al., 2020)	3000 citra sperma tumpang tindih dengan masing-masing kelas yaitu citra normal sebanyak 1021, citra abnormal sebanyak 1005, citra non-sperm sebanyak 974.	<i>Adaptive Noising Overlapping Group Shrinkage image gradient directional masking feature extraction by MSER descriptor SVM classification (nonlinear kernel)</i>	<i>De- Accuracy: 85.7%</i>

### 2.1.1 Dual Tree Complex Wavelet Transform Based Sperm Abnormality Classification.

Penelitian ini dilakukan oleh Ilhan, et al, (2018). Pada penelitian ini fitur statistik berdasarkan *Dual Tree Complex Wavelet Transform (DTCWT)* yang berasal dari sperma normal, sperma cacat dan bukan sperma dimasukkan ke dalam pengklasifikasian *Support Vector Machine (SVM)* dengan tujuan untuk membedakan tiga kelas. Hasil yang diperoleh dengan menggunakan metode ini yaitu akurasi sebesar 82.33% (Ilhan, et al., 2018).

### 2.1.2 *The Effect of Non-linear wavelet Transform based De-Noising in Sperm Abnormality Classification.*

Penelitian ini dilakukan oleh Ilhan, et al, (2018). Dalam penelitian ini menggunakan fitur berbasis *descriptor* dalam klasifikasi sperma normal, sperma cacat, dan bukan sperma. Selain itu, peneliti menyelidiki efek dari dua teknik *De-Noising* dalam kinerja klasifikasi karena adanya *noise patch*. Hasil menunjukkan bahwa proses pengurangan *noise* adaptif berdasarkan wavelet secara dramatis meningkatkan kinerja hingga akurasi sebesar 86% yang mendukung pengklasifikasian kernel SVM polinomial (Ilhan, et al., 2018).

### 2.1.3 *A fully automated hybrid human sperm detection and classification system based on mobile-net and the performance comparison conventional methods.*

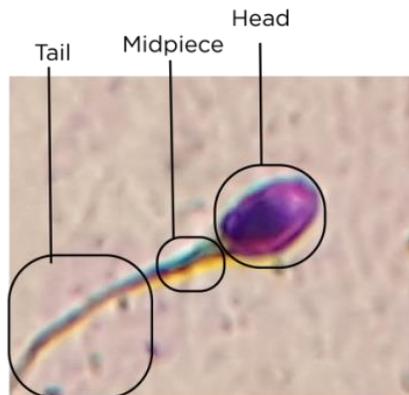
Penelitian ini dilakukan oleh Ilhan, et al, (2020). Dalam penelitian ini menggunakan arsitektur *MobileNet* untuk mengklasifikasi sperma cacat, sperma normal dan bukan sperma. Hasil yang didapatkan adalah pada arsitektur *MobileNet* mendapatkan akurasi sebesar 87% dengan lama waktu latih 1,9 Jam (Ilhan, et al., 2020).

### 2.1.4 *Automated sperm morphology analysis approach using a directional masking technique.*

Penelitian ini dilakukan oleh Ilhan, et al, (2020). Dalam penelitian ini menurut hasil tingkat akurasi 85,7% yang telah dicapai dengan *algoritma* SVM berbasis kernel *non-linear* dalam klasifikasi fitur MSER yang diekstraksi dari pra-pemrosesan kaskade yang diterapkan pada gambar SMIDS, meningkatkan tingkat akurasi sebesar 5% dengan gambar aslinya (Ilhan, et al., 2020).

### 2.3 Sperma

Sperma adalah sel yang sangat khusus dan padat yang tidak mengalami pembelahan atau pertumbuhan setelah terbentuk. Sel ini berasal dari genosit, yang kemudian berdiferensiasi menjadi spermatogonium, dan akhirnya menjadi sperma matang (Syauqy, 2016). Struktur sperma terdiri dari tiga bagian utama, yaitu kepala (*head*), bagian tengah (*midpiece*), dan ekor (*tail*) seperti yang ditunjukkan pada Gambar 1.



Gambar 1. Struktur Sperma Manusia Normal (Ilhan, et al., 2020)

Dalam sperma, kepala merupakan bagian terpenting karena mengandung materi genetic yang ditransfer dari ayah ke anak saat pembuahan terjadi. Bagian tengah mengandung mitokondria yang menyediakan energi yang diperlukan untuk pergerakan sperma. Sementara itu, ekor sperma berfungsi sebagai struktur penggerak yang memungkinkan sperma bergerak menuju sel telur (Syauqy, 2016).

### 2.4 Morfologi

Morfologi merupakan cabang ilmu linguistik yang membahas atau mempelajari kompleksitas struktur kata maupun struktur bentuk dan pengaruh perubahan struktur kata terhadap kelas kata atau makna kata (Mustaqim, et al., 2018). Pengertian morfologi dalam bidang biologi adalah suatu karakteristik yang mudah diamati pada suatu makhluk hidup. Setiap makhluk hidup memiliki ciri morfologi yang dapat diamati untuk mendeteksi sifat khusus (Hadiyanti, et al., 2018).



Gambar 2 Morfologi Sperma Manusia (Ilhan, et al., 2020).

Gambar 2 merupakan gambar morfologi sperma, abnormal\_sperm jenis sperma dengan kelainan struktur atau bentuk yang tidak normal dapat menghambat pergerakan dan pembuahan, non-sperm merupakan elemen-elemen lain dalam sampel semen selain sel sperma seperti sel darah merah, sel epitel atau debris sel, yang dapat menunjukkan masalah Kesehatan atau infeksi jika berlebihan, normal\_sperm merupakan sperma dengan bentuk dan struktur sehat yang memungkinkan pergerakan lancar menuju sel telur dan proses pembuahan yang baik.

## 2.5 Citra Digital

Citra Digital tersusun dari dua suku kata yakni citra berarti suatu gambar visual yang diam (foto) atau bisa juga dengan gambar visual yang bergerak (video). Sedangkan untuk arti dari kata digital ialah pengolahan gambar atau video yang dilakukan menggunakan komputer secara digital (Ratna, 2020). Citra dilihat dari sudut pandang sistematis merupakan fungsi kontinu dari intensitas cahaya pada suatu bidang dua dimensi  $f(x,y)$ ,  $x$  dan  $y$  merupakan koordinat spasial dan amplitudo  $f$  pada koordinat  $(x,y)$  yang disebut intensitas atau derajat keabuan citra pada titik tersebut (Putri, 2016).

Bidang dari  $f(x,y)$  pada suatu citra merupakan suatu matriks dua dimensi yang terdiri dari  $M$  kolom dan  $N$  baris. Baris dan kolom tersebut menghasilkan suatu perpotongan yang disebut *pixel* atau suatu elemen paling kecil pada suatu citra

(Kusumanto & Tompunu, 2011). Berikut adalah Persamaan 1 dari bentuk matriks citra digital berukuran  $M \times N$ .

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \cdots & f(0, N-1) \\ f(1,0) & f(1,1) & \cdots & f(1, N-2) \\ \vdots & \vdots & \vdots & \vdots \\ f(M-1,0) & f(M-1,1) & \cdots & f(M-1, N-1) \end{bmatrix} \dots\dots\dots (1)$$

Suatu citra  $f(x, y)$  dalam fungsi sistematis dapat dituliskan pada Persamaan 2 (Kusumanto & Tompunu, 2011) :

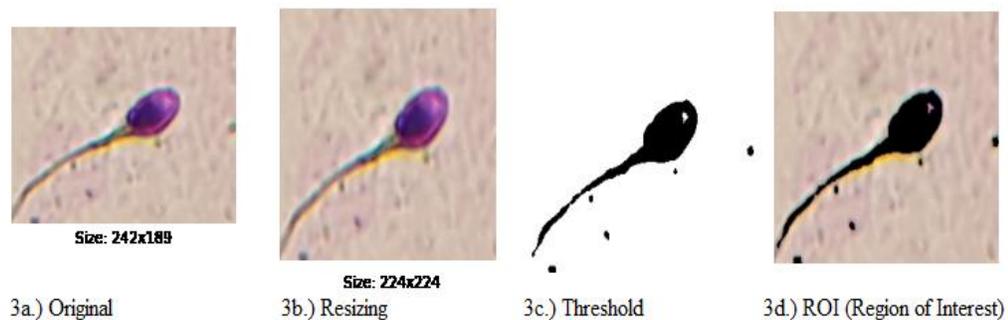
$$\begin{aligned} 0 &\leq x \leq M - 1 \\ 0 &\leq y \leq N - 1 \quad \dots\dots\dots (2) \\ 0 &\leq f(x, y) \leq G - 1 \end{aligned}$$

Dimana :

- $M$  : jumlah *pixel* baris pada array citra
- $N$  : jumlah *pixel* kolom pada array citra
- $G$  : nilai skala keabuan

## 2.6 Pra-pemrosesan

Pra-pemrosesan merupakan teknik yang digunakan pada data untuk menghapus atau mengurangi *noise*, *missing value*, dan *inconsistent data*. *Missing value* menyebabkan data yang dimiliki menjadi tidak akurat karena terdapat informasi atau nilai yang hilang pada data tersebut. *Missing value* menjadi masalah yang sering terjadi pada sebuah data karena proses pengambilan data yang kurang maksimal. *Noise* pada data berupa data yang dapat mengurangi informasi dari citra. Salah satu penyebab adanya *noise* adalah kesalahan pemberian label pada data. *Data inconsistency* terjadi saat data yang ada memiliki format yang berbeda-beda, nama yang berbeda-beda, dan lain sebagainya (Nurvinda, 2021). beberapa contoh metode dalam pra-pemrosesan data dapat dilihat pada Gambar 3.



Gambar 3. Pra-pemrosesan data.

Gambar 3 merupakan tahapan pra-pemrosesan data. Pada 3a merupakan gambar asli belum memiliki perubahan apapun yang akan digunakan sebagai dasar melakukan proses lebih lanjut. Pada 3b *resizing* atau perubahan ukuran gambar digunakan untuk menyesuaikan ukuran gambar agar sesuai dengan ukuran atau format tertentu. Pada 3c *threshold* atau ambang merupakan proses untuk mengkonversikan gambar dari citra *grayscale* menjadi biner. Pada 3d *roi* atau *region of interest* merupakan area tertentu dalam gambar yang dipilih untuk dianalisis atau diproses lebih lanjut.

## 2.7 Machine Learning (ML)

*Machine learning* atau pembelajaran mesin merupakan pendekatan dalam *artificial intelligence* yang banyak digunakan untuk menirukan perilaku manusia untuk penyelesaian masalah atau melakukan otomatisasi (Ahmad, 2017). *Machine Learning* menurut (Marsland, 2015) adalah ilmu dari *Artificial Intelligence* yang akan membuat komputer dapat memiliki suatu kemampuan dengan menggunakan sebuah algoritma yang digunakan komputer untuk belajar dan melakukan tugas tanpa harus adanya instruksi dari pengguna.

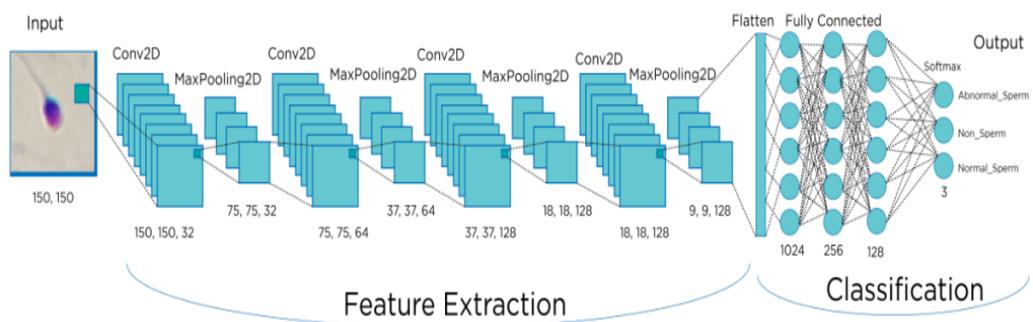
## 2.8 Deep Learning

*Deep Learning* merupakan metode pembelajaran yang memanfaatkan *Artificial Neural Network* (ANN) yang memiliki *multi layer* (berlapis-lapis). ANN dirancang mirip otak manusia, dimana neuron-neuron terkoneksi satu sama lain sehingga membentuk sebuah jaringan yang sangat rumit (Nugroho, et al., 2020).

*Deep Learning* atau *deep structured learning* atau *deep neural* merupakan metode pembelajaran yang memanfaatkan *multiple non-linier transformation*, *deep learning* pada umumnya dipandang sebagai gabungan *artificial intelligence* dan *machine learning* (Primartha, 2018). Salah satu algoritma yang termasuk dalam kategori *Deep Learning* yaitu *Convolutional Neural Network* (CNN).

## 2.9 Convolutional Neural Network (CNN)

*Convolutional Neural Network* (CNN) adalah suatu algoritma *deep learning* yang dikembangkan dari *multi layer perceptron* (MLP) yang dirancang untuk mengelola data dua dimensi (2D). CNN tidak jauh berbeda dengan arsitektur *neural network*, yaitu memiliki *weight*, *bias*, dan *activation function* (Hanin et al., 2021). Arsitektur CNN sendiri memiliki struktur bagian yaitu *feature extraction* dan *classification*. Arsitektur CNN dapat dilihat pada Gambar 4. Pada penelitian ini, ada dua arsitektur CNN yang akan digunakan, yaitu *MobileNetV2* dan *EfficientNetB3*.

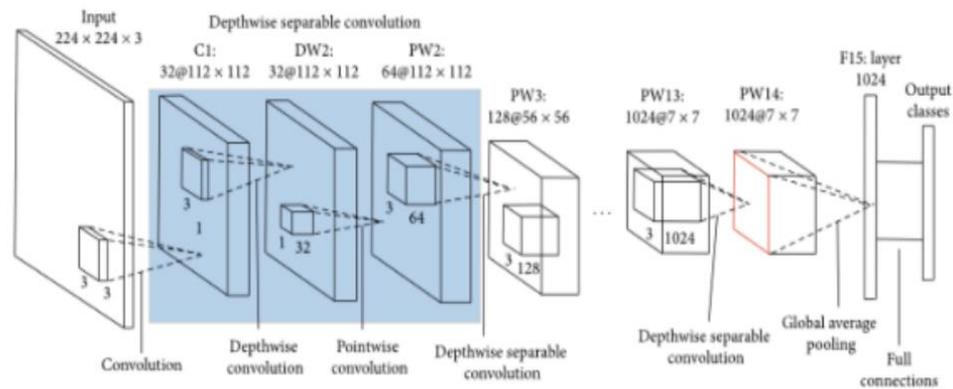


Gambar 4. Arsitektur *Convolutional Neural Network* (Pangestu, et al., 2020).



### 2.9.2 *MobileNetV2*

Salah satu arsitektur dalam metodologi jaringan saraf konvolusi, yang juga dikenal sebagai CNN, yang dapat digunakan untuk mengatasi kebutuhan komputasi yang berat atau membutuhkan sumber daya komputasi yang tinggi adalah *MobileNetV2*. Perbedaan utama antara *MobileNetV2* dan arsitektur CNN konvensional terletak pada penggunaan lapisan konvolusi dengan filter ketebalan yang sesuai dengan ketebalan gambar input. *MobileNetV2* membagi proses konvolusi menjadi dua tahap, yaitu konvolusi *depthwise* dan konvolusi *pointwise* (Miftahuddin & Zaelani, 2022).



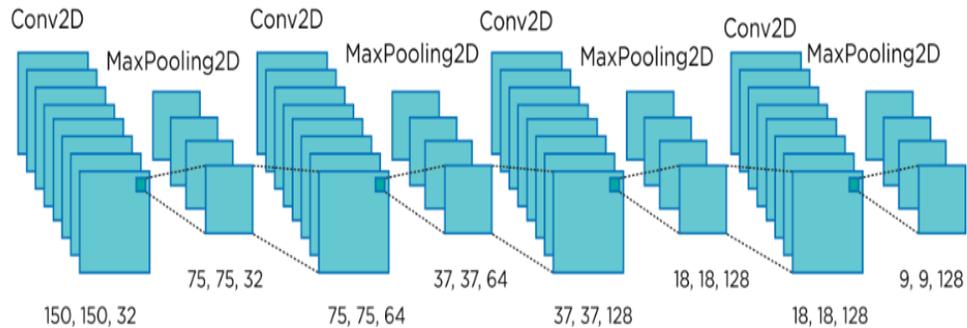
Gambar 6. Arsitektur *MobileNetV2* (Hastomo, et al., 2021).

Seperti Gambar 6 Dalam penggunaan model *MobileNetV2*, langkah pertama adalah memasukkan citra sebagai input. Setelah itu, *MobileNetV2* menggunakan layer 2D. *MobileNetV2* memiliki fitur lapisan *bottleneck*. Kedua *blok* ini memiliki perbedaan dalam proses yang disebut sebagai *bypass* atau residu. Setelah proses tersebut selesai, dilakukan *Global Average Pooling* (GAP) pada lapisan tertentu.

### 2.9.3 *Feature Extraction*

*Feature extraction* merupakan teknik pengambilan ciri dari suatu bentuk untuk mendapatkan nilai yang akan proses selanjutnya (Religia, 2019). Pada algoritma CNN *Feature extraction* terdapat dua lapisan utama yaitu

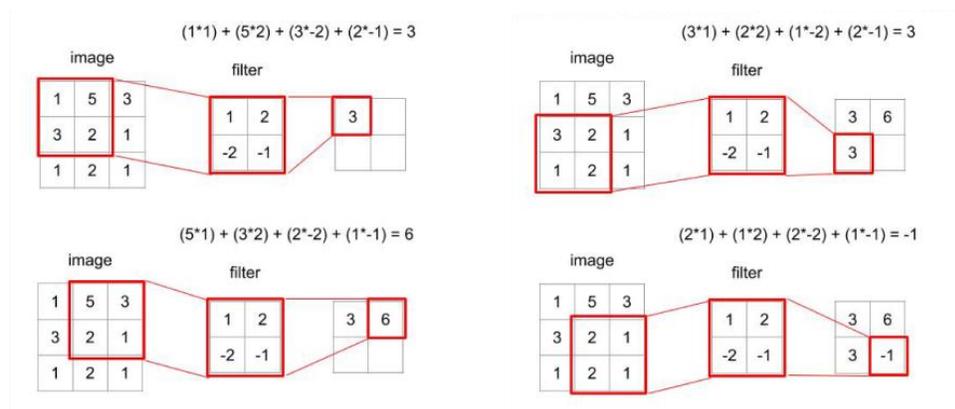
*convolutional layer* dan *pooling layer*. Terdapat fungsi aktivasi yang digunakan pada penelitian ini yaitu *ReLU*. Jumlah lapisan untuk menghasilkan akurasi yang baik dengan melakukan banyak percobaan seperti pada Gambar 7.



Gambar 7. *Feature Extraction* (Pangestu, et al., 2020).

### 2.9.3.1 Convolutional Layer

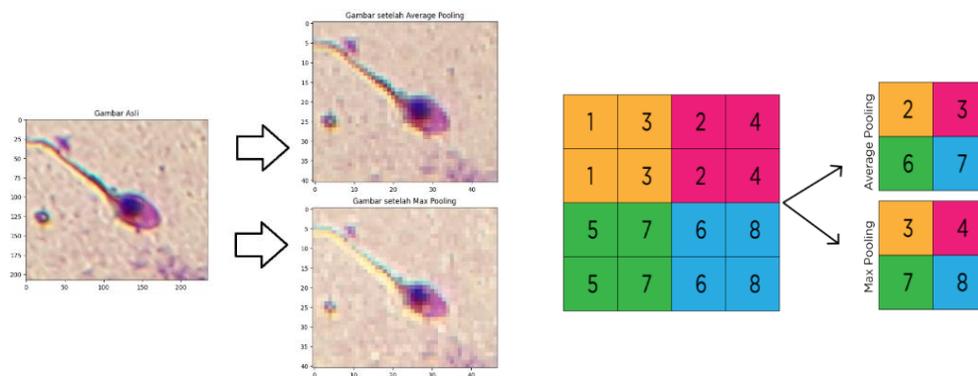
*Convolutional Layer* adalah layer pertama yang akan memproses citra yang dimasukkan, sehingga menjadi layer ini sangat penting pada arsitektur *convolutional neural network* (CNN) (Pangestu, et al., 2020). Pada lapisan ini terjadi operasi konvolusi dan filterasi untuk mendapatkan ekstraksi fitur dari data atau citra yang di masukkan. Operasi ini dinamakan *feature map* atau *activation map* seperti yang ditunjukkan pada Gambar 8.



Gambar 8. *Convolutional Layer*.

### 2.9.3.2 Pooling Layer

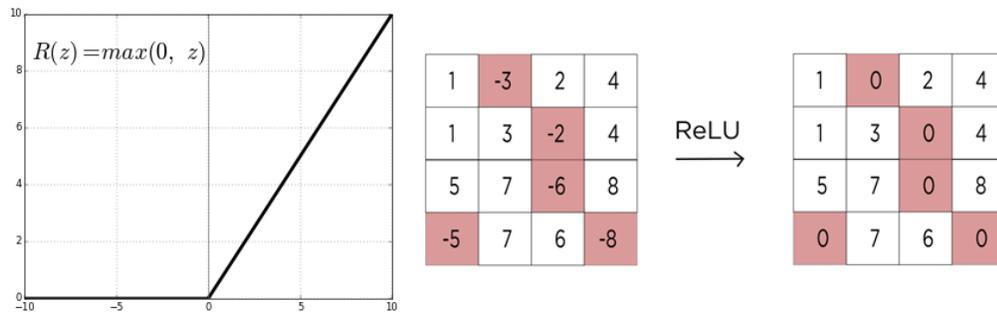
*Pooling layer* adalah lapisan yang mengolah output dari *convolutional layer* dengan tujuan untuk mereduksi ukuran sebuah data citra (Putra, et al., 2016). *Pooling layer* dapat meningkatkan invariansi posisi dari fitur. Sebagian besar CNN metode yang sering digunakan adalah *max pooling* dan *average pooling*. *Max pooling* membagi output dari *convolutional layer* menjadi beberapa grid kecil lalu mengambil nilai tertinggi dari setiap *grid* sedangkan *average pooling* mengambil nilai rata-rata dari setiap *grid*, keduanya akan menghasilkan matriks baru seperti yang ditunjukkan pada Gambar 9. *Grid* yang berwarna kuning, merah, hijau dan biru merupakan kelompok *grid* yang akan dipilih nilainya sehingga hasil proses tersebut dapat dilihat pada tampilan *grid* sebelah kanan.



Gambar 9. Pooling Layer.

### 2.9.3.3 ReLU (Rectified Linear Units)

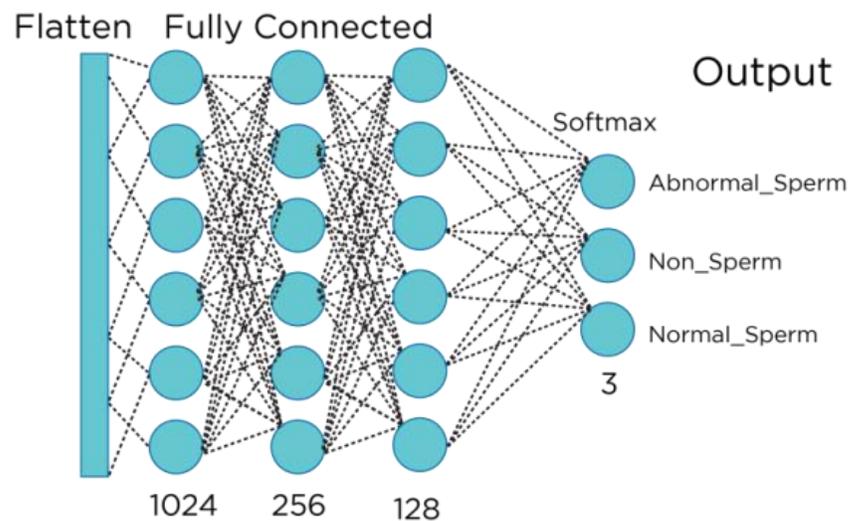
*Rectified linear units (ReLU)* merupakan fungsi sebagai pengubah atau menghilangkan nilai negatif pada citra. Perubahan tersebut bertujuan untuk meningkatkan kualitas citra untuk meminimalisir kesalahan (Pangestu, et al., 2020). Cara kerja fungsi aktivasi ReLU ini yaitu dengan mengganti nilai negatif pada citra atau *feature maps* dengan nilai 0 seperti pada Gambar 10.



Gambar 10. Aktivasi ReLU.

#### 2.9.4 Classification Feature

*Classification feature* atau klasifikasi fitur merupakan proses untuk menyatakan suatu objek kedalam sebuah kategori yang sudah didefinisikan sebelumnya (Pangestu, et al., 2020). Dalam penelitian ini membahas tentang mengambil ciri dari suatu *feature* dengan dataset berisi kumpulan citra sperma yang akan digunakan untuk klasifikasi sperma cacat, sperma normal dan bukan sperma. Pada tahap ini terdapat satu lapisan yaitu *fully connected layer* dan fungsi aktivasi yang digunakan pada penelitian ini yaitu *softmax* seperti pada Gambar 11.



Gambar 11. Classification Feature.

#### 2.9.4.1 Flatten

*Flatten* merupakan proses membentuk ulang fitur menjadi sebuah vector agar bisa digunakan sebagai *input* dari *fully connected layer*. Pada umumnya *flatten* akan merubah hasil *feature map* dari *feature extraction* berbentuk dua dimensi array menjadi satu dimensi array (Pangestu, et al., 2020).

#### 2.9.4.2 Fully Connected Layer

*Fully connected layer* merupakan lapisan yang menerapkan MLP dengan tujuan melakukan transformasi pada dimensidata agar data dapat diklasifikasi secara linear. Pada lapisan ini neuron-neuron saling terhubung dengan semua neuron pada lapisan sebelumnya. Lapisan ini akan menerima input dari lapisan aktivasi sebelumnya untuk menentukan fitur yang paling berkorelasi dengan kelas tertentu berdasarkan *training dataset* (Pangestu, et al., 2020).

#### 2.9.4.3 Softmax

*Softmax* merupakan fungsi yang mengambil inputan vektor dari bilangan real  $k$ , dan normalisasi menjadi probabilitas terdistribusi yang terdiri dari probabilitas  $k$ . Sebelum menerapkan *softmax*, beberapa komponen vektor bisa negatif atau lebih besar dari satu dan mungkin tidak berjumlah 1, tetapi setelah menerapkan *softmax* setiap komponen akan berada dalam interval (0-1) dan komponen akan bertambah hingga 1, sehingga dapat diartikan sebagai probabilitas, komponen *input* yang lebih besar akan sesuai dengan probabilitas yang lebih besar. *Softmax* sering digunakan dalam neural network, fungsi ini digunakan untuk menentukan *output* yang sesuai (Pangestu, et al., 2020). Persamaan 3 yang digunakan pada fungsi aktivasi *softmax* adalah sebagai berikut.

$$\rho(x) = \frac{e^x}{\sum_{k=1}^k e^x} \dots\dots\dots(3)$$

### 2.10 Confusion matrix

*Confusion matrix* merupakan sebuah cara yang sangat berguna untuk menganalisis seberapa baik *classifier* mengenali sebuah data termasuk kedalam kategori (Anreaja, et al., 2022 ). Ada beberapa istilah pada *confusion matrix*, istilah tersebut dapat dilihat pada Tabel 2.

Tabel 2. *Confusion matrix* (Anreaja, et al., 2022 ).

<i>Actual Class</i>	<i>Predicted Class</i>		<b>Total</b>
	<i>Yes</i>	<i>No</i>	
<i>Yes</i>	<i>TP</i>	<i>FN</i>	<i>P</i>
<i>No</i>	<i>FP</i>	<i>TN</i>	<i>N</i>
<b>Total</b>	<i>P'</i>	<i>N'</i>	<i>P + N</i>

Adapun istilah yang ada pada tabel 3 tersebut, sebagai berikut (Rahayu, et al., 2021).

- TP (True Positive)*, yaitu jumlah data dengan nilai sebenarnya positif dan nilai prediksi positif.
- FP (False Positive)*, yaitu jumlah data dengan nilai sebenarnya negatif dan nilai prediksi positif.
- FN (False Negative)*, yaitu jumlah data dengan nilai sebenarnya positif dan nilai prediksi negatif.
- TN (True Negative)*, yaitu jumlah data dengan nilai sebenarnya negatif dan nilai prediksi negatif.

Nilai dari TP dan TN memberikan informasi ketika *classifier* benar dalam melakukan klasifikasi data, sedangkan nilai FP dan FN memberikan informasi ketika *classifier* salah dalam melakukan klasifikasi data.

## 2.11 Classification Report

*Classification report* adalah metrik evaluasi untuk mengukur keefektifan dari suatu model klasifikasi, pada umumnya metrik ini digunakan untuk mengevaluasi performa model dalam tugas pengklasifikasian menggunakan metrik seperti *accuracy*, *recall*, *precision*, dan *f1-score* (Anreaja, et al., 2022).

### 2.11.1 Accuracy

*Accuracy* adalah menghitung nilai probabilitas berdasarkan nilai yang benar dari label kelas. Akurasi memiliki Persamaan 4, sebagai berikut (Rahayu, et al., 2021).

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \dots\dots\dots (4)$$

### 2.11.2 Recall

*Recall* atau *sensitivity* adalah keakuratan dari nilai positif yang ada. *Recall* memiliki Persamaan 5, sebagai berikut (Rahayu, et al., 2021).

$$recall = \frac{TP}{TP+FN} \times 100\% \dots\dots\dots (5)$$

### 2.11.3 Precision

*Precision* adalah nilai kebenaran dari prediksi yang dilakukan oleh *classifier* dengan label kelas yang sudah ada. *Precision* memiliki Persamaan 6, sebagai berikut (Rahayu, et al., 2021).

$$precision = \frac{TP}{TP+FP} \times 100\% \dots\dots\dots (6)$$

### 2.11.4 F1-Score

*F1-score* adalah nilai rata-rata harmonik dari *recall* dan *precision* (Rahayu, et al., 2021). *F1-score* dinyatakan dengan Persamaan 7, sebagai berikut.

$$F1\ Score = 2 \times \frac{Recall \times Presisi}{Recall + Presisi} \dots\dots\dots (7)$$

### III. METODOLOGI PENELITIAN

#### 3.1 Tempat dan Waktu Penelitian

##### 3.1.1 Tempat Penelitian

Penelitian dilaksanakan di Jurusan Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Lampung.

##### 3.1.2 Waktu Penelitian

Penelitian dilakukan pada Januari 2023 s.d. Juli 2023. Pengerjaan dibagi menjadi tiga tahapan, yaitu:

a. Pengumpulan Data

Data citra sperma manusia ini dikumpulkan oleh (Ilhan, et al., 2020) dan dipublikasikan pada tahun 2020 yang kemudian dapat diakses melalui link <https://data.mendeley.com/datasets/6xvdhc9fyb/1>.

b. Pra-pemrosesan Data

Dilakukan pra-pemrosesan data berupa *resizing*, *threshold*, *region of interest* dan augmentasi data pada data latih.

c. Pemodelan Data

Dibuat model *convolutional neural network* (CNN) dengan arsitektur *MobileNetV2* Dan *EfficientNetB3*. Masing-masing arsitektur dibagi dua skenario berdasarkan pembagian data dan *hyperparameter fine tuning*. Alur waktu pengerjaan penelitian dapat dilihat pada Tabel 3.

Tabel 3. Alur Waktu Pengerjaan Penelitian.

Kegiatan	2023																															
	Januari				Februari				Maret				April				Mei				Juni				Juli				Agustus			
	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Pemahaman dan Pengumpulan Data	■																															
Pra-pemrosesan Data									■																							
Pemodelan Data, Prediksi CNN, dan Evaluasi Kinerja													■				■				■				■							
a. Arsitektur MobileNet V2													■																			
• 60% <i>training</i> , 20% <i>validation</i> , 20% <i>testing</i>													■				■															
• 70% <i>training</i> , 10% <i>Validation</i> , 20% <i>testing</i>													■				■															
b. Arsitektur EfficientNet B3													■																			
• 60% <i>training</i> , 20% <i>validation</i> , 20% <i>testing</i>													■				■															
• 70% <i>training</i> , 10% <i>Validation</i> , 20% <i>testing</i>													■				■															

## 3.2 Data dan Alat

### 3.2.1 Data

Data yang digunakan pada penelitian ini berjumlah 3000 citra Citra sperma yang merupakan citra gambar RGB dengan ekstensi bmp (*bitmap*) yang memiliki variasi ukuran piksel dikumpulkan oleh (Ilhan, et al., 2020) dan dipublikasikan pada tahun 2020 dengan judul *A Fully Automated Hybrid Human Sperm Detection and Classification System Based on Mobile-net and the Performance Comparison with Conventional Methods*. Data ini bersumber dari alamat <https://data.mendeley.com/datasets/6xvdhc9fyb/1>. Data tersebut dibagi menjadi tiga kelas yaitu 1005 citra sperma normal (*Normal\_Sperm*), 974 citra bukan sperm (*Non-Sperm*), dan 1021 citra sperma abnormal (*Abnormal\_Sperm*). Data dapat dilihat pada Gambar 12.



Gambar 12. *Sperm Morphology Image Data Set*.

### 3.2.2 Alat

#### 3.2.2.1 Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan pada penelitian ini yaitu:

- a. *Processor*: AMD ® Ryzen 5 3400G
- b. *Installed RAM*: 16 GB.
- c. *Storage*: SSD M.2 NVMe 500 GB
- d. *Video Graphics Array (VGA)*: TESLA K80, 24GB of GDDR5 VRAM Memory, 4992 Nvidia CUDA cores

### 3.2.2.2 Perangkat Lunak (*Software*)

Perangkat lunak yang digunakan pada penelitian ini yaitu:

- a. Sistem Operasi: Ubuntu 20.04.3 LTS.
- b. *Tools*:
  1. Python 3.9.16 dan Jupyter Notebook 6.3.0

2. *Library*:

- A. *Library* Open-CV 4.6.4

OpenCV (*Open Source Computer Vision*) merupakan *library* Python yang digunakan untuk pengolahan citra gambar dan video untuk mendapatkan informasi yang ada didalamnya (Lazaro, et al., 2017).

- B. *Library* Numpy 1.22.3

Numpy (*Numerical Python*) adalah sebuah *library* python yang digunakan untuk kebutuhan *scientific* dan matematis (Retnoningsih & Pramudita, 2020).

- C. *Library* Pandas 1.5.3

Pandas merupakan *library* python yang biasa digunakan untuk menganalisis data berbagai format data seperti file .txt, .csv, atau lainnya (Retnoningsih & Pramudita, 2020).

- D. *Library* Tensorflow-gpu 2.4.1

Tensorflow adalah *Library* pada python yang merupakan APIs atau *tools* untuk *artificial intelligence* (AI) dan *machine learning* (ML).

- E. *Library* Keras-gpu 2.4.3

Keras merupakan *library* yang memiliki fitur untuk mempermudah pengembangan lebih dalam tentang *Deep Learning* (Santoso & Ariyanto, 2018).

- F. *Library* Scikit-learn 1.0.2

Scikit-learn merupakan *library* di Python yang biasa digunakan untuk membantu dalam melakukan *data splitting*, *data processing*, dan *deep learning* (Retnoningsih & Pramudita, 2020).

### G. Library Matplotlib 3.7.1

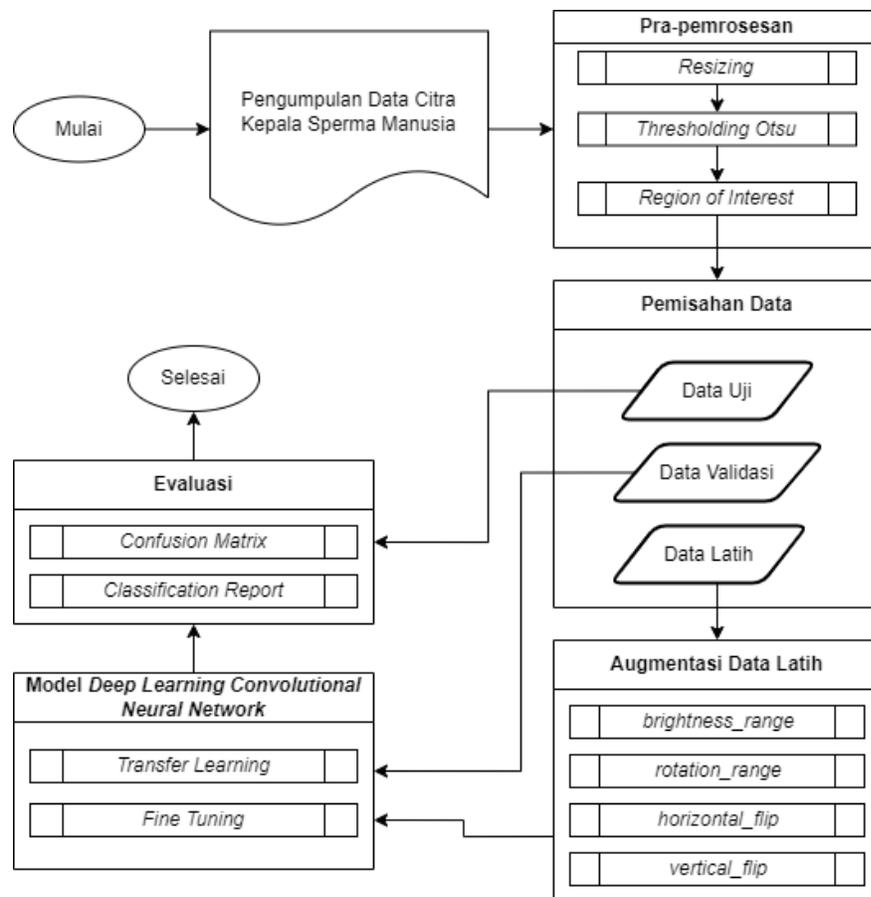
Matplotlib adalah *library* Python yang digunakan untuk memvisualisasikan data dalam bentuk plot atau grafik dua dimensi (2D) maupun tiga dimensi (3D).

### H. Library Seaborn 0.12.2

Seaborn adalah *library* python yang digunakan untuk membuat visualisasi data yang menarik, *library* ini berfungsi sebagai antarmuka yang lebih tinggi untuk *library* matplotlib.

## 3.3 Alur Kerja Penelitian

Alur penelitian ini memiliki beberapa tahap yang dapat dilihat pada Gambar 13.



Gambar 13. Alur Penelitian.

Berikut ini penjelasan dari setiap tahapan pada Gambar 13.

### 3.3.1 Pengumpulan Data

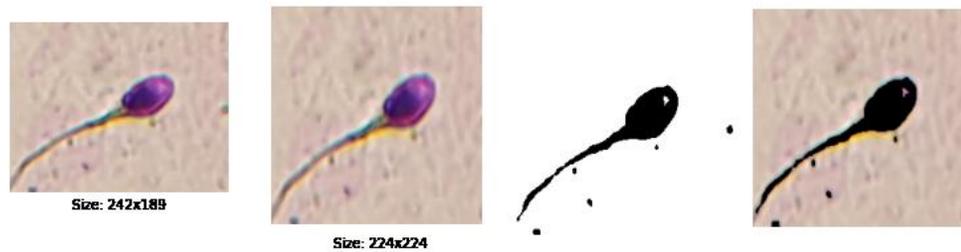
Pengumpulan data merupakan data sekunder dan didapatkan dari *Mendeley Data*. Data yang dikumpulkan merupakan citra sperma dengan format bmp. SMIDS didalam folder terdapat 3 kelas jenis sperma dengan total citra 3000. Jumlah masing-masing folder dataset dapat dilihat pada Tabel 4.

Tabel 4. Jumlah Dataset.

Jenis Data	Jumlah Data
<i>Abnormal_Sperm</i>	1005
<i>Non-Sperm</i>	974
<i>Normal Sperm</i>	1021
Total Data	3000

### 3.3.2 Pra-pemrosesan Data

Pada tahap pra-pemrosesan dilakukan dengan tiga langkah. Pertama menggunakan metode *resizing* untuk mengubah ukuran data *input* menjadi ukuran sesuai *input* pada model sehingga mempermudah proses selanjutnya. Kedua menggunakan *threshold* untuk merubah citra ke dalam citra biner dengan membatasi nilai intensitas piksel tertentu, sehingga mempermudah untuk mengidentifikasi objek dan mengurangi kompleksitas data. Terakhir dilakukan identifikasi dan isolasi area yang relevan menggunakan *Region of Interest* untuk memfokuskan analisis hanya dalam bagian tertentu dalam citra. Contoh penerapan pra-pemrosesan data dapat dilihat pada Gambar 14.



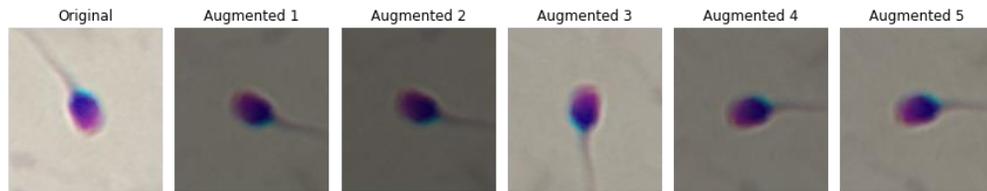
Gambar 14. Pra-pemrosesan Data.

### 3.3.3 Pemisahan Data

Pemisahan data akan dilakukan dengan *random splitting* Menggunakan fungsi *train\_test\_split* dari *library scikit-learn*. Fungsi tersebut akan memilih data secara acak sesuai dengan persentase pemisahan data. Dua skenario pembagian data yang akan dilakukan. Skenario pertama 60% *training*, 20% *validation* dan 20% *testing*. Skenario kedua 70% *training*, 10% *validation* dan 20% *testing*.

### 3.3.4 Augmentasi Data

Augmentasi data dilakukan untuk memperbanyak data *training* agar terhindar dari *overfitting* dan menghasilkan model yang lebih baik. Augmentasi data dilakukan dengan menggunakan *ImageDataGenerator* yang terdapat pada *library Keras*. Beberapa parameter yang akan digunakan seperti *brightness\_range* untuk mengatur rentang penyesuaian kecerahan pada gambar. Nilai yang akan digunakan untuk menentukan kecerahan di antara 0.2 hingga 0.8, di mana nilai di bawah 1.0 akan membuat gambar menjadi lebih gelap, dan nilai di atas 1.0 akan membuat gambar menjadi lebih terang. Kemudian *rotation\_range* mengatur rotasi acak pada gambar. Kemudian setelah gambar di *rotasi horizontal\_flip* mengatur untuk melakukan *flipping* secara *horizontal* pada gambar secara acak. Ini berarti gambar dapat dibalik secara *horizontal* secara acak, menghasilkan versi cermin dari gambar tersebut. Kemudian yang terakhir *vertical\_flip* untuk melakukan *flipping* secara vertikal pada gambar secara acak. Contoh penerapan augmentasi data dapat dilihat pada Gambar 15.



Gambar 15. Augmentasi Data.

### 3.3.5 Model *Convolutional Neural Network* (CNN)

Pada penelitian ini akan dibuat model *Convolutional Neural Network* (CNN). Dengan dua jenis arsitektur, yaitu arsitektur *MobileNetV2* dan *EfficientNetB3*. Dimensi *input* yang digunakan adalah  $224 \times 224$  piksel. Dalam pelatihan model menggunakan dua percobaan di setiap skenario. Percobaan pertama tanpa *augmentasi* dan percobaan kedua menggunakan *augmentasi* dengan tujuan mendapatkan model yang optimal. Dua skenario pelatihan pada Tabel 5.

Tabel 5. *Hyperparameter*.

Nama Parameter	Skenario 1	Skenario 2
<i>batch_size</i>	32	32
<i>epoch</i>	25	50
<i>optimizer</i>	SGD	SGD
<i>learning_rate</i>	0,001	0.0001
<i>momentum</i>	-	0.9
<i>fine_tuned_layer</i>	-	-7

### 3.3.6 Evaluasi

Pada tahap ini, dilakukan evaluasi untuk mengetahui kinerja model *convolutional neural network* (CNN) dalam mengklasifikasi sperma menggunakan *confusion matrix* dan *classification report* pengujian berasal dari data latih dan data uji yang menghasilkan nilai *confusion matrix* dan *classification report* dari data *testing* yang akan mengeluarkan hasil seperti *accuracy*, *recall*, *precision*, dan *f1 score*.

## V. SIMPULAN DAN SARAN

### 5.1 Kesimpulan

Berdasarkan penelitian dan pembahasan yang sudah dilakukan mengenai klasifikasi sperma manusia berdasarkan morfologi menggunakan *deep learning Convolutional Neural Network* (CNN) dapat diambil kesimpulan, antara lain.

1. Dalam proses pengujian model terdapat dua skenario pemisahan data, *learning rate*, dan *epoch*, serta dilakukan *fine tuning* pada model. Pada percobaan pertama, yakni tanpa augmentasi data, terdapat nilai terbaik pada model *EfficientNetB3* dengan pemisahan data 70% untuk pelatihan, 10% untuk validasi, dan 20% untuk pengujian, yang menghasilkan rata-rata akurasi sebesar 83%, presisi (*precision*) 83.03%, *recall* 83.31%, dan *f1-score* 83.13%. pada percobaan kedua dengan augmentasi data, terdapat nilai terbaik pada model *MobileNetV2* dengan pemisahan data 70% untuk pelatihan, 10% untuk validasi, dan 20% untuk pengujian, yang mencapai rata-rata akurasi sebesar 83.33%, presisi (*precision*) 84.01%, *recall* 83.54% dan *f1-score* 83.52%.
2. Dengan menerapkan teknik pelatihan kedua model dapat mencapai peningkatan yang signifikan dalam nilai *accuracy*, *presisi*, *recall*, dan nilai *f1-score* pada model *transfer learning*.

## 5.2 Saran

Saran pada penelitian ini, antara lain.

1. Penelitian lebih lanjut dapat diarahkan untuk mengembangkan model atau metode yang lebih canggih dengan memanfaatkan arsitektur *deep learning* terbaru dengan *hyperparameter tuning*. Penggunaan model yang lebih kompleks dapat meningkatkan performa dan akurasi klasifikasi.
2. Melakukan *augmentasi dataset* untuk menyeimbangkan data perkelas ataupun memperbanyak ukuran *dataset*. *Dataset* yang lebih besar akan membantu mengatasi masalah *overfitting* dan membuat model lebih baik dalam menghadapi data baru.
3. Selain klasifikasi, model *deep learning CNN* juga dapat dieksplorasi untuk tugas-tugas lain yang relevan dalam analisis sperma, seperti prediksi potensi fertilisasi dan penilaian morfologi lainnya.

## DAFTAR PUSTAKA

- Ahmad, A., 2017. Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *Jurnal Teknologi Indonesia*, pp. 1-5.
- Albert, Gunadi, K. & Setyati, E., 2020. Deteksi Helm pada Pengguna Sepeda Motor dengan Metode Convolutional Neural Network. *Jurnal Infra*, 8(1), pp. 295-301.
- Andrianto, H., 2017. *Biologi Sel dan Molekuler*. Yogyakarta: Deepublish.
- Anreja, L. J., Harefa, N. N., Prima Negara, J. G., Pribyantara, .V. N. H., & Prasetyo, A. B., 2022 . Naive Bayes and Support Vector Machine Algorithm for Sentiment Analysis Open Sea Mobile Application Users in Indonesia. *JISA (Jurnal Informatika dan Sains)*, 5(1), pp. 62-68.
- Ghamisi, P., Yushi, C. & Xiao Xiang, Z., 2016. A Self-Improving Convolution Neural Network for the Classification of Hyperspectral Data. *IEEE Geoscience and Remote Sensing Letters*, 13(10), p. 1537–1541.
- Hadiyanti, N., Supriyadi & Pardono, 2018. Keragaman Beberapa Tumbuhan Ciplukan (*Physalis spp.*) di Lereng Gunung Kelud, Jawa Timur. *Jurnal Ilmu-Ilmu Hayati*, 17(2), pp. 134-139.
- Hanin, M. A., Patmasari, R., & Nur, R. Y. 2021. Sistem Klasifikasi Penyakit Kulit Menggunakan Convolutional Neural Network (CNN) Skin Disease Classification System Using Convolutional Neural Network ( CNN ). *E-Proceeding of Engineering*, 8(1), pp. 273–281.

- Hastomo, W., Sugiyanto & Sudjiran, 2021. Convolution Neural Network Arsitektur Mobilenet-v2 Untuk Mendeteksi Tumor Otak. Seminar Nasional Teknologi Informasi dan Komunikasi STI&K (SeNTIK), 5(1), pp. 17-21.
- Ilhan, H. O., Serbes, G. & Aydin, N., 2018. Dual Tree Complex Wavelet Transform Based Sperm Abnormality Classification. *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, pp. 577-580.
- Ilhan, H. O., Serbes, G. & Aydin, . N., 2020. Automated Sperm Morphology Analysis Approach Using a Directional Masking Technique. *Computers in Biology and Medicine*, 122(1), pp. 1-24.
- Ilhan, H. O., Sigirci, I. O., Serbes, G. & Aydin, N., 2018. The Effect of Nonlinear Wavelet Transform Based De-noising in Sperm Abnormality Classification. *2018 3rd International Conference on Computer Science and Engineering (UBMK)*, pp. 658-661.
- Ilhan, H. O., Sigirci, I. O., Serbes, G. & Aydin, N., 2020. A Fully Automated Hybrid Human Sperm Detection and Classification System Based on Mobile-net and The Performance Comparison with Conventional Methods. *Medical & Biological Engineering & Computing*, 58, pp. 1047-1068.
- Iqbal, I., Mustafa, G. & Ma, J., 2020. Deep Learning-Based Morphological Classification of Human Sperm Heads. *Diagnostics*, 10(5), p. 1-21.
- Kusumanto, R. & Tomponu, A. N., 2011. Pengolahan Citra Digital Untuk Mendeteksi Obyek Menggunakan Pengolahan Warna Model Normalisasi RGB. Seminar Nasional Teknologi Informasi & Komunikasi Terapan 2011 (Semantik 2011), pp. 1-7.
- Lazaro, A., Buliali, J. L. & Amaliah, B., 2017. Deteksi Jenis Kendaraan di Jalan Menggunakan OpenCV. *Jurnal Teknik ITS*, 6(2), pp. 293-299.

- Marsland, S., 2015. *Machine Learning & Pattern Recognition Series*. Second Edition ed. New York: Taylor & Francis Group, LLC.
- Miftahuddin, Y. & Zaelani, F., 2022. Perbandingan Metode Efficientnet-b3 Dan Mobilenet-v2 Untuk Identifikasi Jenis Buah-buahan Menggunakan Fitur Daun. *Jurnal Ilmiah Teknologi Informasi Terapan*, 9(1), pp. 1-11.
- Mustaqim, N., Hanye, P. & Seli, S., 2018. Morfologi Bahasa Dayak Pompakng. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa*, 7(4), pp. 3-10.
- Nugroho, P. A., Fenriana, I. & Arijanto, R., 2020. Implementasi Deep Learning Menggunakan CNN Pada Ekspresi Manusia. *Jurnal Algor*, 2(1), pp. 12-21.
- Nurvinda, G., 2021. Langkah Awal dalam Pemrosesan Data: Data Preprocessing dalam Data Mining. DQLab. <https://dqlab.id/pentingnya-preprocessing-dalam-pengolahan-data-statistik>. Diakses pada 22 maret 2023
- Pangestu, R. A., Rahmat, B. & Anggraeny, F. T., 2020. Implementasi Algoritma CNN Untuk Klasifikasi Citra Lahan dan Perhitungan Luas. *Jurnal Informatika dan Sistem Informasi (JIFoSI)*, 1(1), pp. 166 - 174.
- Primartha, R., 2018. *Belajar Machine Learning : Teori Dan Praktik*. Bandung: Informatika.
- Putra, I. W. S. E., Wijaya, A. Y. & Soelaiman, R., 2016. Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101. *Jurnal Teknik ITS*, 5(1), pp. 65-69.
- Putri, A. R., 2016. Pengolahan Citra Dengan Menggunakan Webcam Pada Kendaraan Bergerak Di Jalan Raya. *JUPI (Jurnal Ilmiah Pendidikan Informatika)*, 1(1), pp. 1-6.
- Ratna, S., 2020. Pengolahan Citra Digital dan Histogram Dengan Python dan Text Editor Pycharm. *Technologia*, 11(3), pp. 181-186.

- Religia, Y., 2019. Feature Extraction Untuk Klasifikasi Pengenalan Wajah Menggunakan Support Vector Machine dan K-Nearest Neighbor. *Pelita Teknologi: Jurnal Ilmiah Informatika, Arsitektur dan Lingkungan*, 14(2), pp. 85-92.
- Retnoningsih, E. & Pramudita, R., 2020. Mengenal Machine Learning Dengan Teknik Supervised dan Unsupervised Learning Menggunakan Python. *Bina Insani ICT Journal*, 7(2), pp. 156-165.
- Santoso, A. & Ariyanto, G., 2018. Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah. *Jurnal Teknik Elektro*, 18(1), pp. 15-21.
- Soebijanto, S., 2013. Prediksi Keberhasilan Kehamilan Teknik Fertilisasi in Vitro pada Berbagai Umur Istri. *Medica Hospitalia*, 2(1) , pp. 1-5.
- Sumiati, 2013. Sistem Reproduksi Manusia. *Jurnal Biologi*, 2(2), pp. 1-13.
- Syauqy, A., 2016. Evaluasi Kromatin Sperma Sebagai Indikator Kualitas Sperma. *Jambi Medical Journal* , 2(1), pp. 87-97.
- WHO, 2021. WHO Laboratory Manual for The Examination and Processing of Human Semen. 6th pendant. Switzerland: World Health Organization.
- Xinyan, F. & Weigang, Z., 2019. Research on SAR Image Target Recognition Based on Convolutional Neural Network. *Journal of Physics: Conference Series*, 1213(4), pp.1-8.