

**IDENTIFIKASI SPERMA MANUSIA BERDASARKAN MORFOLOGI
DENGAN MENGGUNAKAN ALGORITMA YOLO**

(Skripsi)

Oleh

ARBI HIDAYATULLAH

1817051045



**JURUSAN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

**IDENTIFIKASI SPERMA MANUSIA BERDASARKAN MORFOLOGI
DENGAN MENGGUNAKAN ALGORITMA YOLO**

Oleh

ARBI HIDAYATULLAH

Skripsi

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

Pada

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



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG
2022**

ABSTRAK

IDENTIFIKASI SPERMA MANUSIA BERDASARKAN MORFOLOGI DENGAN MENGGUNAKAN ALGORITMA YOLO

Oleh

ARBI HIDAYATULLAH

Infertilitas merupakan masalah reproduksi yang sangat krusial yang dialami oleh pria maupun wanita. Infertilitas adalah ketidakmampuan untuk hamil dalam waktu satu tahun setelah hubungan seksual. Penelitian ini berfokus terhadap infertilitas pada pria. Banyak penyebab yang dapat menyebabkan infertilitas pada pria termasuk kualitas sperma. Saat ini identifikasi sperma manusia masih dilakukan secara manual dengan mengamati sperma dengan bantuan manusia melalui mikroskop, sehingga membutuhkan waktu dan biaya yang tinggi. Oleh karena itu, diperlukan teknologi tinggi untuk mengetahui kualitas sperma berupa teknologi *deep learning* berbasis video. Algoritma *deep learning* mendukung penelitian ini dalam mengidentifikasi sel sperma manusia. Sehingga *deep learning* dapat membantu mendeteksi video sperma secara otomatis dalam proses evaluasi sel sperma untuk menentukan infertilitas. Kami menggunakan teknologi pembelajaran mendalam untuk mengidentifikasi sperma menggunakan algoritma *You Only Look Once* versi 4 (YOLOv4). Tujuan dari penelitian ini adalah untuk menganalisis tingkat akurasi dari algoritma YOLOv4. Dataset yang digunakan bersumber dari dataset VISEM sebanyak 85 video. Hasil yang didapat adalah 90,31% AP (*Average Precision*) untuk objek sperma dan 68,19% AP (*Average Precision*) untuk objek non-sperma, kemudian untuk hasil pelatihan didapatkan model 79,58% mAP (*Mean Average Precision*). Penelitian kami menunjukkan hasil tentang identifikasi sperma manusia menggunakan YOLOv4. Hasil yang diperoleh model YOLOv4 dapat mengidentifikasi objek sperma dan non-sperma. Keluaran pada model YOLOv4 mampu mengidentifikasi objek pada data uji berupa video dan gambar.

Kata Kunci: *Deep learning*; klasifikasi; identifikasi; kepala sperma; sperma; *You Only Look Once* version 4;

Judul Skripsi : **IDENTIFIKASI SPERMA MANUSIA
BERDASARKAN MORFOLOGI DENGAN
MENGUNAKAN ALGORITMA YOLO**

Nama Mahasiswa : **Arbi Hidayatullah**

Nomor Pokok Mahasiswa : 1817051045

Program Studi : S1 Ilmu Komputer

Jurusan : Ilmu Komputer

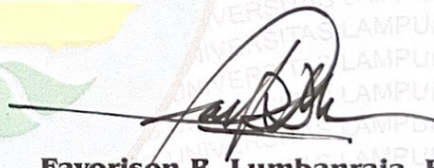
Fakultas : Matematika dan Ilmu Pengetahuan Alam

MENYETUJUI

1. Komisi Pembimbing

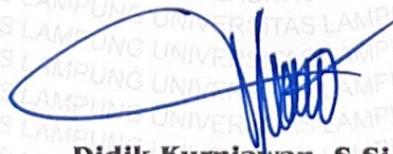


Aristoteles, S.Si., M.Si.
NIP 19810521 200604 1 002



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

2. Ketua Jurusan Ilmu Komputer




Didik Kurniawan, S.Si., M.T.
NIP 19800419 200501 1 004

MENGESAHKAN

1. Tim Penguji


Ketua : Aristoteles, S.Si., M.Si.



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



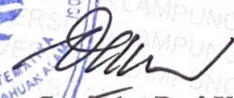
**Penguji II
Bukan Pembimbing : Rico Andrian, S.Si., M.Kom.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam



Dr. Eng. Suripto Dwi Yuwono, S.Si., M.T.
NIP 19740705 200003 1 001



Tanggal Lulus Ujian Skripsi : 1 Agustus 2022

PERNYATAAN

Saya yang bertanda tangan di bawah ini:

Nama : Arbi Hidayatullah

NPM : 1817051045

Dengan ini menyatakan bahwa skripsi saya yang berjudul **“Identifikasi Sperma Manusia Berdasarkan Morfologi Dengan Menggunakan Algoritma YOLO”** 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 di kemudian hari terbukti skripsi saya merupakan hasil penjiplakan atau dibuat orang lain, maka bersedia menerima sanksi berupa pencabutan gelar yang telah saya terima.

Bandar Lampung, 10 Agustus 2022



Arbi Hidayatullah

NPM. 1817051045

RIWAYAT HIDUP



Penulis dilahirkan di Bandar Lampung, pada tanggal 25 Maret 2000, sebagai anak pertama dari dua bersaudara.

Penulis menyelesaikan pendidikan formal di SDN 1 Rajabasa Raya dan selesai pada tahun 2012. Kemudian pendidikan menengah pertama di SMPN 22 Bandar Lampung yang diselesaikan pada tahun 2015, lalu melanjutkan ke pendidikan menengah atas di SMAN 13 Bandar Lampung yang diselesaikan pada tahun 2018.

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

1. Menjadi anggota Adapter Himpunan Mahasiswa Jurusan Ilmu Komputer pada periode 2018/2019.
2. Menjadi anggota Biro Kesekretariatan Himpunan Mahasiswa Jurusan Ilmu Komputer periode 2018/2019.
3. Menjadi Ketua Biro Kesekretariatan Himpunan Mahasiswa Jurusan Ilmu Komputer periode 2019/2020

4. Menjadi Asisten Dosen Jurusan Ilmu Komputer tahun 2019 hingga 2021.
5. Melaksanakan Kerja Praktek pada bulan Juli periode 2020/2021 di BP-KKN Universitas Lampung.
6. Melaksanakan Kuliah Kerja Nyata di Kelurahan Rajabasa Raya, Kecamatan Rajabasa, Kota Bandar Lampung, Lampung pada tahun 2020/2021 dengan program kerja pembuatan website Kelurahan Rajabasa Raya.
7. Penerima Beasiswa Peningkatan Prestasi Akademik (PPA) pada tahun 2019.

MOTO

1. “Sesungguhnya Allah tidak akan mengubah nasib suatu kaum sehingga mereka mengubah keadaan yang ada pada diri mereka sendiri.”

(QS. Ar-Ra’d:11)

2. “Hindarilah sifat malas dan bosan karena keduanya kunci keburukan. Sesungguhnya jika engkau malas, engkau tidak akan banyak melaksanakan kewajiban. Jika engkau bosan, engkau tidak akan tahan dalam menunaikan kewajiban.”

(Umar bin Khattab)

3. “Untuk mencapai sesuatu, harus diperjuangkan dulu. Seperti mengambil buah kelapa, dan tidak menunggu saja seperti jatuh durian yang telah masak.”

(Mohammad Natsir)

4. “Terus maju, jangan malu untuk memulai, jalanlah terus walau dirimu melakukan kesalahan, karena kesalahan yang kamu lakukan saat ini merupakan batu loncatan untuk sempurna di masa mendatang.”

(Anonim)

PERSEMBAHAN

Alhamdulillahillobbilamin

Puji dan syukur tercurahkan kepada Allah Subhanahu Wa Ta'alaah atas segala Rahmat dan Karunia-Nya sehingga saya dapat menyelesaikan skripsi ini. Shalawat serta salam selalu tercurahkan kepada Nabi Muhammad SAW.

Kupersembahkan karya ini kepada:

Kedua Orang Tuaku Tercinta

Yang senantiasa memberikan yang terbaik, dan melantunkan do'a yang selalu menyertaiku. Kuucapkan pula terima kasih sebesar-besarnya karena telah mendidik dan membesarkanku dengan cara yang dipenuhi kasih sayang, dukungan, dan pengorbanan yang belum bisa terbalaskan.

Seluruh Keluarga Besar Ilmu Komputer 2018

Yang selalu memberikan semangat dan dukungan.

Almamater Tercinta, Universitas Lampung dan Jurusan Ilmu Komputer

Tempat bernaung mengemban semua ilmu untuk menjadi bekal hidup.

SANWACANA

Puji syukur kehadirat Allah SWT atas berkah, rahmat dan hidayat-Nya, serta petunjuk dan pedoman dari Rasulullah Nabi Muhammad Sholallahu Alaihi Wasallam penulis dapat menyelesaikan skripsi yang berjudul “Identifikasi Sperma Manusia Berdasarkan Morfologi Dengan Menggunakan Algoritma YOLO” dengan baik dan lancar.

Terima kasih penulis ucapkan kepada semua pihak yang telah membantu dan berperan besar dalam menyusun skripsi ini, antara lain.

1. Kedua orang tua serta adik tercinta yang selalu memberi dukungan, do'a, semangat, motivasi, dan kasih sayang yang luar biasa tak terhingga. Semua yang telah kalian berikan tidak akan pernah mampu untukku balas. Semoga Allah SWT selalu memberikan kebahagiaan dan keberkahan dalam kehidupan kalian di dunia dan akhirat.
2. Bapak Aristoteles, S.Si., M.Si. sebagai pembimbing utama yang telah memberikan arahan, ide, motivasi, kritik serta saran kepada penulis sehingga dapat menyelesaikan skripsi ini dengan baik.
3. Bapak Favorisen R. Lumbanraja, Ph.D. sebagai pembimbing kedua dan pembimbing akademik yang juga selalu dapat memberikan waktu untuk

membimbing penulis dalam memberikan ide, kritik serta saran untuk dapat menyelesaikan skripsi ini.

4. Bapak Rico Andrian, S.Si., M.Kom. sebagai pembahas yang telah memberikan masukan yang bermanfaat dalam perbaikan skripsi ini.
5. Bapak Didik Kurniawan, S.Si., M.T. selaku ketua Jurusan Ilmu Komputer FMIPA Universitas Lampung.
6. Bapak Dr. Eng. Suropto Dwi Yuwono, S.Si., M.T. selaku Dekan FMIPA Universitas Lampung.
7. Bapak Dr. rer. nat. Akmal Junaidi, M.Sc. 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 telah memberikan ilmu dan pengalaman dalam hidup untuk menjadi lebih baik.
10. Teman-teman grup Rich yaitu Rika Ningtias Azhari, Annisa Nurwalikadani, Syela Septania, Abi Perdana Kusuma, Pandi Barep Arianza dan Yuan Ferdinan A K, lalu teman-teman grup heemahoora serta teman-teman seperjuangan kompre Muhammad Fajru Ramadhan dan Ridho Alrafi yang selalu memberikan semangat dan dukungan.
11. Keluarga Ilmu Komputer 2018 yang tidak bisa penulis sebut satu persatu. Keluarga kedua penulis, rekan kelompok, rekan diskusi, rekan bercanda, dan telah memberi arti dan warna serta pengalaman tak ternilai semasa duduk di bangku kuliah.

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 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, 10 Agustus 2022

Arbi Hidayatullah

NPM. 1817051045

DAFTAR ISI

	Halaman
ABSTRAK	iii
MENGESAHKAN	v
PERNYATAAN.....	vi
RIWAYAT HIDUP.....	vii
MOTO.....	ix
PERSEMBAHAN.....	x
SANWACANA.....	xi
DAFTAR ISI.....	xiv
DAFTAR TABEL.....	xvii
DAFTAR GAMBAR	xviii
DAFTAR PSEUDOCODE	xx
I. PENDAHULUAN	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan.....	4
1.5 Manfaat.....	4
II. TINJAUAN PUSTAKA	5
2.1 Penelitian Terdahulu.....	5
2.2 Sperma.....	9

2.3	Morfologi.....	9
2.4	Identifikasi.....	10
2.5	Citra Digital.....	10
2.6	Artificial Intelligence.....	11
2.7	Machine Learning.....	12
2.8	Deep Learning	13
2.9	<i>Convolutional Neural Network (CNN)</i>	14
2.10	You Only Look Once (YOLO).....	17
2.11	You Only Look Once V4 (YOLOv4).....	20
2.12	Pra proses Data.....	22
2.13	Confusion Matrix	22
2.14	IoU (<i>Intersection over Union</i>).....	25
2.15	mAP (<i>mean Average Precision</i>).....	25
2.16	<i>Non-Max Supression</i>	26
2.17	Konfigurasi <i>Hyperparameter</i>	26
III.	METODOLOGI PENELITIAN	29
3.1	Tempat dan Waktu Penelitian	29
3.2	Data dan Alat.....	31
3.3	Alur Kerja Penelitian.....	33
IV.	HASIL DAN PEMBAHASAN	37
4.1	Pra-Proses Data	37
4.1.1	Pengumpulan Data	37
4.1.2	<i>Resizing</i>	38
4.2	Anotasi Data.....	38
4.3	Pembagian Data.....	40
4.4	Konfigurasi <i>Hyperparameter</i>	41
4.4.1	<i>Backbone</i> Pada YOLOv4.....	42
4.4.2	<i>Neck</i> Pada YOLOv4.....	43
4.4.3	<i>Head</i> Pada YOLOv4	44
4.5	<i>Training Weights</i>	45
4.5.1	<i>Pre-trained Weights</i>	45
4.5.2	<i>Training Weights</i>	46
4.6	Implementasi Deteksi Model YOLOv4.....	46

4.7	Hasil <i>Training</i> Model YOLOv4	49
4.8	Pembahasan Hasil <i>Training</i> Model YOLOv4	53
4.9	Hasil Uji Model YOLOv4	59
4.10	Perbandingan Hasil Akurasi Dengan Penelitian Terdahulu	62
V.	PENUTUP	64
5.1	Simpulan.....	64
5.2	Saran.....	65
	DAFTAR PUSTAKA	67

DAFTAR TABEL

Tabel	Halaman
1. Penelitian terdahulu yang terkait dengan penelitian objek sperma menggunakan CNN (<i>Convolutional Neural Network</i>)	5
2. Alur Waktu Pengerjaan Penelitian.	30
3. <i>Hyperparameter</i> Pelatihan Model.	35
4. Hasil <i>average precision</i> pada <i>learning rate</i> 0.002	50
5. Hasil <i>average precision</i> pada <i>learning rate</i> 0.0002	50
6. Hasil <i>average precision</i> pada <i>learning rate</i> 0.00002	51
7. Hasil <i>precision</i> , <i>recall</i> , <i>F1-score</i> , dan <i>mAP</i> pada <i>learning rate</i> 0.002	51
8. Hasil <i>precision</i> , <i>recall</i> , <i>F1-score</i> , dan <i>mAP</i> pada <i>learning rate</i> 0.0002	52
9. Hasil <i>precision</i> , <i>recall</i> , <i>F1-score</i> , dan <i>mAP</i> pada <i>learning rate</i> 0.00002	52
10. Tabel perbandingan hasil <i>training</i> dengan tiga skenario pembagian data berdasarkan <i>learning rate</i>	54
11. Hasil evaluasi objek <i>test</i> pada model YOLOv4	59
12. Hasil perhitungan <i>precision</i> , <i>recall</i> , <i>AP</i> dan <i>mAP</i>	60
13. Perbandingan hasil <i>mAP</i> , <i>Precision</i> , <i>Recall</i> , <i>F1-score</i> dengan penelitian Deepsperm	62

DAFTAR GAMBAR

Gambar	Halaman
1. Morfologi sperma. (Sumber gambar http://www.proceptin.com /phc/sperm-cell.php .).....	9
2. Gambar morfologi sperma dan non sperma. (Sumber gambar https://datasets.simula.no/visem/ .)	10
3 Penerapan konsep <i>artificial Intelligence</i>	12
4. Ilustrasi dari model <i>deep learning</i> (Goodfellow, et al., 2016).	13
5. Arsitektur dari <i>Convolutional Neural Network</i> (CNN). (Sumber gambar https://kotakode.com/blogs/2707/Convolutional-Neural-Network-(CNN)). ...	14
6. Ilustrasi konvolusi matriks 6 x 6 (Rohim, et al., 2019).....	15
7. <i>Matrix pooling</i> . Ilustrasi di atas merupakan <i>matrix pooling</i> menggunakan 2 x 2 filter dan menggunakan 2 stride.	16
8. Tampilan representasi fungsi aktivasi ReLU (Gu, et al., 2017).....	17
9. Proses sistem deteksi YOLO (Redmon, et al., 2016).....	18
10. Model deteksi menggunakan bounding boxes (Redmon, et al., 2016).	19
11. Arsitektur YOLO (Redmon, et al., 2016).	19
12. Perbandingan YOLOv4 dengan detector objek yang lain (Bochkovskiy, et al., 2020).	20
13. Tahapan dari detector (Bochkovskiy, et al., 2020).	21
14. <i>Confusion matrix</i> (Rahma, et al., 2021).	23
15. Contoh tampilan deteksi IoU.	25
16. Dataset VISEM.	31
17. Alur kerja penelitian identifikasi sperma manusia menggunakan algoritma YOLO.....	33
18. Hasil <i>resizing</i> citra.....	34

19. Tampilan anotasi data.	34
20. Tampilan file hasil anotasi.	39
21. Penjelasan format file hasil anotasi.	39
22. Tampilan file obj.data.	39
23. Tampilan file obj.name.	40
24. Tampilan hasil anotasi.	40
25. Grafik nilai <i>precision</i> , <i>recall</i> dan <i>F1-score</i> pada <i>learning rate</i> 0.002.	55
26. Grafik nilai <i>precision</i> , <i>recall</i> dan <i>F1-sciore</i> pada <i>learning rate</i> 0.0002.	55
27. Grafik nilai <i>precision</i> , <i>recall</i> dan <i>F1-score</i> pada <i>learning rate</i> 0.00002.	56
28. Grafik perbandingan AP <i>sperm</i>	56
29. Grafik perbandingan AP non <i>sperm</i>	57
30. Grafik perbandingan hasil mAP.	57
31. Grafik data train hasil pembagian data <i>train</i> 70%, validasi 25%, dan <i>test</i> 5% dengan <i>learning rate</i> 0,0002.	59
32. Tampilan <i>testing</i> model YOLOv4.	61

DAFTAR PSEUDOCODE

Pseudocode	Halaman
1. Kode program ekstrak video menjadi <i>image</i>	38
2. Kode program <i>resizing</i> image.	38
3. Kode program pembagian data.	41
4. Kode program konfigurasi.	42
5. Kode program <i>backbone</i> YOLOv4.	43
6. Kode program SPP <i>neck</i> pada YOLOv4.	44
7. Kode program <i>head</i> pada YOLOv4	45
8. Kode program fungsi parser.	47
9. Kode program fungsi video_capture.	47
10. Kode program fungsi <i>inference</i>	48
11. Kode program fungsi <i>drawing</i>	48
12. Kode program main.	49

I. PENDAHULUAN

1.1 Latar Belakang

Reproduksi merupakan suatu proses biologis pada makhluk hidup untuk memperbanyak diri, agar dapat mempertahankan kelangsungan hidup suatu spesies. Proses reproduksi pada manusia bergantung pada penyatuan sel gamet pria dan wanita, yang masing masing memiliki setengah set kromosom, untuk membentuk individu baru dengan set kromosom yang lengkap. Organ reproduksi primer/gonad pada pria yakni sepasang testis dan pada wanita berupa sepasang ovarium. Pada kedua jenis kelamin gonad memiliki beberapa fungsi diantaranya pertama menghasilkan gamet atau gametogenesis, yaitu pada pria berupa sperma (spermatozoa) dan pada wanita berupa sel telur (ovum), serta mengeluarkan hormon seks yang pada pria berupa testosteron sedangkan pada wanita berupa estrogen dan progesteron (Sherwood, 2011).

Proses peleburan antara sel telur wanita dengan spermatozoa pria disebut fertilisasi. Proses ini merupakan cikal bakal terbentuknya suatu embrio manusia (Kumala, 2019). Dalam dua dekade terakhir ini, masalah reproduksi pada pria yang mendapatkan banyak perhatian adalah infertilitas. Menurut data dari badan kesehatan dunia WHO, tingkat kesuburan reproduksi telah turun drastis. Infertilitas didefinisikan sebagai ketidakmampuan untuk hamil dalam kurun waktu setelah setahun sering melakukan hubungan seksual tanpa menggunakan kontrasepsi (Oktarina, et al., 2014). Secara umum infertilitas pria menyumbang sebanyak 50% dari seluruh kasus infertilitas. Dari 50% tersebut sebanyak 6 – 27% merupakan kasus infertilitas pada pria namun tidak dapat dijelaskan penyebabnya (Syauqy, 2014).

Saat ini telah banyak penelitian yang memprediksi penyebab dari infertilitas. Metode yang banyak digunakan yakni dengan metode observasi dari data rekam medik pasien pada suatu rumah sakit (Dhyani, et al., 2020). Sejumlah penelitian menunjukkan bahwa faktor dari penyebabnya infertilitas diantaranya faktor usia, kebiasaan merokok, penggunaan ganja, heroin, gangguan hormone, dan gangguan imunologis (S.Ningsih & Farich, 2016). Hal-hal tersebut dapat meningkatkan resiko memiliki sperma yang abnormal sehingga menjadi infertilitas.

Berdasarkan penyebab dari infertilitas tersebut sebelum maraknya teknologi seperti saat ini, penelitian masih dilakukan secara evaluasi manual sampel sperma menggunakan mikroskop. Sehingga proses pengerjaannya banyak memakan waktu dan hasilnya tidak konsisten. Adanya perkembangan teknologi pada bidang komputer saat ini memengaruhi banyak penelitian untuk meningkatkan analisa infertilitas pada sperma pria.

Adapun penelitian yang terkait yakni menggunakan arsitektur CNN atau dikenal dengan *Convolutional Neural Network*. CNN dapat mengidentifikasi morfologi sperma secara akurat berdasarkan gambar sperma (Iqbal, et al., 2020). Hasil identifikasi yang akurat dipengaruhi oleh kualitas data gambar penuh atau dengan ukuran pixel yang besar, sehingga performa akurasi dan deteksi dapat lebih optimal (Nissen, et al., 2017).

Selain menganalisa berdasarkan morfologi dari sperma, CNN dapat mengidentifikasi sperma dengan motilitasnya. Untuk data yang akan diuji yakni berupa data video hasil rekaman dari sampel sperma. Memprediksi motilitas sperma secara otomatis ini menggunakan teknik yang termasuk ke dalam regresi linier sederhana. Proses pengerjaannya dilakukan dengan cepat dan mendapatkan hasil yang konsisten (Hicks, et al., 2019). Pengembangan dari CNN ialah R-CNN atau *Region Based Convolutional Neural Network* dapat digunakan dalam evaluasi motilitas kepala sperma pada data berupa video. Metode ini melakukan segmentasi kepala sperma, sedangkan untuk algoritma digunakan sebagai pelacakan koordinat pusat kepala sperma, sehingga memungkinkan untuk menghitung kecepatan motilitas dari kepala sperma tersebut (Valiuškaitė, et al., 2020).

Berdasarkan dari penelitian yang terkait belum ada suatu penggunaan metode yang melibatkan algoritma YOLO (You Look Only Once) untuk menganalisis dan identifikasi tentang sperma manusia. Algoritma YOLO (*You Look Only Once*) sendiri merupakan bagian dari algoritma machine learning yang berbasis CNN (*Convolutional Neural Network*) yang berfungsi sebagai *object detection* pada citra yang banyak (Redmon, et al., 2016). Algoritma YOLO ini merupakan metode yang lebih efisien dibandingkan dengan algoritma *detection object* pada *machine learning* lainnya (Bochkovskiy, et al., 2020). Maka dari itu, perlu dibangun suatu model menggunakan metode YOLO untuk mengidentifikasi morfologi sperma.

1.2 Rumusan Masalah

Berdasarkan latar belakang yang telah dijabarkan, permasalahan yang dapat diangkat adalah sebagai berikut :

1. Mengidentifikasi morfologi sperma manusia menggunakan algoritma YOLO.
2. Menguji hasil dari identifikasi morfologi sperma manusia pada algoritma YOLO berdasarkan data uji dan membandingkan hasil training berdasarkan skenario pembagian data.

1.3 Batasan Masalah

Agar penelitian menjadi lebih fokus, pembatasan masalah perlu dilakukan. Batasan pada penelitian ini adalah :

1. Penelitian ini menggunakan *dataset* Visem sperma manusia berupa dalam bentuk video berekstensi AVI (Audio Video Interleave).
2. Algoritma yang digunakan ialah algoritma YOLO.
3. Algoritma YOLO dibuat dengan menggunakan *platform* python.
4. Citra yang akan menjadi inputan adalah berupa video.

5. Objek yang dianotasi dari dataset hanya dua buah *class* yakni *sperm* dan *non sperm*.

1.4 Tujuan

Berdasarkan rumusan masalah di atas, didapatkan tujuan dari penelitian ini adalah sebagai berikut :

1. Melakukan proses identifikasi morfologi sperma manusia menggunakan algoritma YOLO.
2. Mengetahui hasil dari identifikasi algoritma YOLO pada studi kasus identifikasi sperma manusia berdasarkan morfologi dengan menggunakan algoritma YOLO.

1.5 Manfaat

Manfaat dari penelitian ini yaitu :

1. Membangun model identifikasi sperma dengan algoritma YOLO.
2. Membantu dalam pengembangan model ke depan untuk analisa sperma dapat lebih efisien terutama untuk penelitian infertilitas pada pria.
3. Sebagai referensi penggunaan metode YOLO dalam identifikasi morfologi sperma berupa video.

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 terdahulu yang terkait dengan penelitian objek sperma menggunakan CNN (*Convolutional Neural Network*)

No	Judul	Metode	Data	Hasil
1	<i>Convolutional neural networks for segmentation and object detection of human semen</i> (Nissen, et al., 2017)	<i>Convolutional Neural Network (CNN)</i>	<i>Dataset : 765 grayscale images, 35 independent sperm sample</i>	presisi : 93,87% recall : 91,89%
2	<i>Machine Learning-Based Analysis of Sperm Videos and Participant Data for Male Fertility Prediction</i> (Hicks, et al., 2019)	<i>Linear Regression, Gaussian Process, SMOreg, Random Forests dan Convolutional Neural Network (CNN)</i>	<i>Dataset : VISEM-dataset 21.250 image</i>	<i>Deep learning : MAE = 11.091 Classical Machine Learning : MAE = 11.400</i>
3	<i>Deep Learning-Based Morphological Classification of Human Sperm Heads</i> (Iqbal, et al., 2020)	<i>Convolutional Neural Network (CNN)</i>	<i>dataset : SCIAN sejumlah 1132 dan 384 gambar kepala sperma, dataset :</i>	<i>SCIAN : Recall = 88% akurasi = 77% HuSHeM : recall = 95% akurasi = 96%</i>

				HuSHeM gambar sperma.	216 kepala	
4	<i>Deep Learning Based Evaluation of Spermatozoid Motility for Artificial Insemination</i> (Valiuškaitė, et al., 2020)	<i>Region Based Convolutional Neural Networks (R-CNN)</i>	Metode <i>DeepSperm Convolutional Neural Network (CNN)</i>	<i>Dataset</i> : VISEM-dataset 650 anotasi image	diambil dari 12 sperma banteng berbeda pada balai inseminasi buatan lembang Indonesia, dengan observasi video beresolusi 640 x 480 pixels dan 24 fps.	akurasi : 91.77% Mean Absolute Error (MAE) : 2.92
5	<i>Deep Sperm: A robust and real-time bull sperm-cell detection in densely populated semen videos</i> (Hidayatullah, et al., 2021)	<i>DeepSperm Convolutional Neural Network (CNN)</i>	Metode <i>DeepSperm Convolutional Neural Network (CNN)</i>	<i>Dataset</i> diambil dari 12 sperma banteng berbeda pada balai inseminasi buatan lembang Indonesia, dengan observasi video beresolusi 640 x 480 pixels dan 24 fps.	mAP = 94.11 F1 Skor = 0.93 kecepatan = 51.9 fps.	

Berdasarkan Tabel 1, terdapat beberapa penelitian terdahulu yang sudah dilakukan sebelumnya. Penelitian yang pertama dilakukan oleh Nissen, et al., pada tahun 2017, penelitian ini membandingkan satu set arsitektur dari *Convolutional Neural Network (CNN)* untuk melakukan segmentasi dan mendeteksi sel sperma manusia suatu gambar dari sampel sperma yang telah diambil. *Dataset* yang dibangun yakni 765 gambar dalam bentuk *grayscale* dari 35 sampel sperma yang berbeda. Gambar diakuisisi menggunakan *image cytometer* dengan 20 pembesaran optik dan resolusi 1920 x 1440 *pixels*. Intensitas gambar yang terkuantisasi yakni dari 14 bit sampai 8 bit gambar. Pada setiap gambar intensitas dinormalisasikan yang nilainya di antara nol dan satu. Gambar-gambar itu dianotasi oleh seorang ahli dan didaftarkan ke dalam dua *class* yakni *background* dan sel sperma. *Dataset* sampel ini dibagi menjadi 70% data *train* dan 30% data *test* berdasarkan pengambilan sampel bertingkat pada jumlah rata-rata sel sperma pada gambar penuh di setiap sampelnya. Arsitektur yang digunakan untuk perbandingannya ialah arsitektur 2-conv dan 3-conv yang telah dilatih dalam *dataset* dan ditest dengan gambar penuh dari *dataset*. Arsitektur yang telah dilatih untuk 200 *epoch* dengan *mini-batches* 256

patches. Hasil yang didapatkan dari pengujian yakni metode terbaik adalah 3-*conv-full-up-inc* yang memiliki presisi 93,87% dan recall 91,89% pada set pengujian dengan menggunakan threshold 150.

Pada penelitian terdahulu bagian kedua yakni oleh Hicks, et al., pada tahun 2019, dalam penelitian ini menjelaskan tentang metode untuk menganalisis otomatis suatu data sperma manusia. Sebelumnya dalam analisis sperma masih dilakukan secara evaluasi manual menggunakan mikroskop dan memakan waktu yang lama sehingga dilakukannya penelitian ini untuk menemukan metode yang lebih efisien. Peneliti menggunakan teknik yang termasuk regresi linier sederhana dan metode yang lebih canggih yakni *convolutional neural network* (CNN). Untuk *dataset* yang digunakan iyalah berasal dari VISEM-*dataset* dengan data berupa video sejumlah 85 video sampel dari partisipan yang berbeda. Data video tersebut diekstrak menjadi 250 *frame* per video, sehingga mendapatkan total sebesar 21.250 *frame* untuk semua video. Menggunakan pendekatan *classical machine learning* mendapatkan performa terbaik dengan MAE 10.800. Lalu menggunakan pendekatan Channel-wise Grayscale mendapatkan MAE sebesar 8.786. Hasil ini menunjukkan bahwa prediksi motilitas sperma berdasarkan deep learning menggunakan video prosesnya cepat dan konsisten.

Penelitian terdahulu bagian ketiga oleh Iqbal, et al., tahun 2020, penelitian ini membahas tentang klasifikasi morfologi dari kepala sperma menggunakan *convolutional neural network* (CNN) berdasarkan gambar sperma. *Dataset* yang digunakan bersumber dari 2 data yakni *dataset* SCIAN dan HuSHeM. Jumlah pada *dataset* SCIAN yakni 1132 dan 348 gambar kepala sperma, sedangkan untuk *dataset* HuSHeM hanya memiliki 216 gambar kepala sperma. Hasil pengujian penelitian ini mendapatkan nilai akurasi dan recall dari masing masing *dataset*. Untuk *dataset* SCIAN mendapatkan akurasi 77% dan recall 88%, sedangkan untuk *dataset* HuSHeM mendapatkan akurasi 96% dan recall 95%.

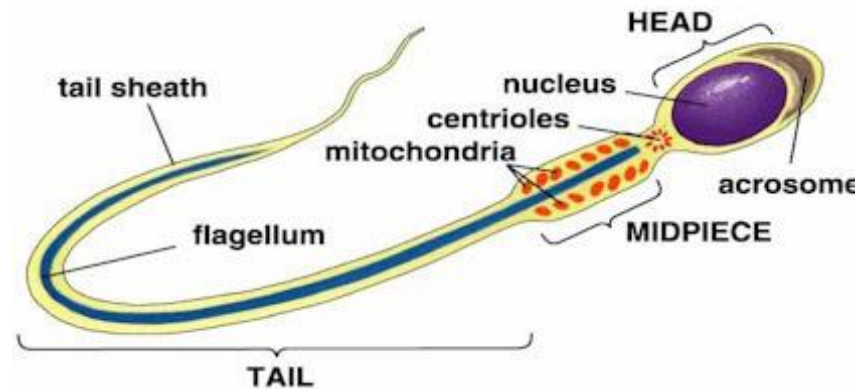
Selanjutnya penelitian terdahulu yang ke-empat yakni oleh Valiuškaitė, et al., pada tahun 2020, pembahasan penelitian ini hampir sama dengan penelitian Imran Iqbal et al., yang membahas tentang deteksi kepala sperma namun yang membedakan ialah metodenya. Metode yang digunakan masih dalam lingkup *deep learning* yakni menggunakan arsitektur *Region Based Convolutional Neural Network* (R-CNN). R-CNN ini digunakan untuk mengevaluasi motilitas kepala sperma pada video sperma manusia. Dataset yang digunakan yakni menggunakan VISEM-dataset. Akurasi yang didapatkan dalam penggunaan R-CNN ini berdasarkan dataset sebesar 91.77%, sedangkan untuk mean absolute error (MAE) prediksi vitalitasnya adalah 2.92. Namun, dalam penelitian terdapat beberapa kasus yakni gagalnya metode yang digunakan untuk mengenali kepala sperma dengan benar karena oklusi, benda yang serupa, atau kepala sperma yang terhubung ke batas bingkai video, yang membuat kepala sperma yaitu hanya tampak seperti titik hitam saja.

Penelitian terdahulu bagian kelima yaitu oleh Hidayatullah, et al., pada tahun 2021. Penelitian ini membahas tentang deteksi sel sperma pada banteng yakni membuat suatu metode dengan pendekatan *deep learning* yang bernama DeepSperm. Arsitektur yang diajukan pada penelitian ini menemukan satu lapisan deteksi yang khusus untuk deteksi objek kecil. Untuk menangani *overfitting* dan meningkatkan akurasi yakni dengan menetapkan resolusi jaringan *input* yang lebih besar, menggunakan *dropout layer* dalam saturasi dan *exposure*. Pengukuran akurasi pada metode ini menggunakan *Mean Average Precision* (mAP), *confusion matrix*, *precision*, *re-call*, dan *F1-score*. Lalu hasil dari DeepSperm akan dibandingkan dengan metode You Look Only Once (YOLO) v3 dan YOLOv4. Dalam hasil percobaannya mendapatkan 94.11 mAP pada dataset uji, *F1-score* 0.93, dan kecepatan pemrosesan 51.9 fps. Dibanding dengan YOLOv4 metode ini dapat 2,18 kali lebih cepat pada pengujiannya dan 2,9 kali lebih cepat pada pelatihan dengan kumpulan data kecil, sekaligus mencapai akurasi deteksi komparatif.

2.2 Sperma

Sperma atau spermatozoa, yaitu suatu sel yang tidak akan lagi mengalami pembelahan atau pertumbuhan. Sperma berasal dari gonosit yang menjadi spermatogonium, spermatosit primer dan sekunder yang berikutnya berubah menjadi spermatid dan akhirnya berubah menjadi spermatozoa atau disebut dengan sperma (Syauqy, 2014).

Sperma adalah sel memanjang yang dihasilkan oleh sistem reproduksi jantan pada manusia. Sperma memiliki bagian-bagian yakni terdiri dari kepala yang terdapat inti dan ekor yang digunakan sebagai motilitasnya (Agustinus, et al., 2018). Gambar 1 merupakan tampilan morfologi sperma.

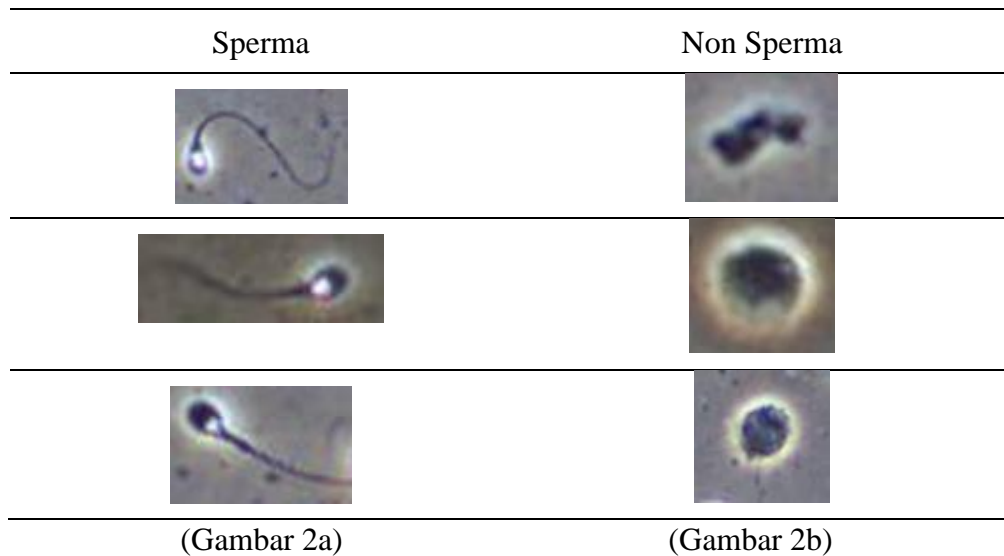


Gambar 1. Morfologi sperma. (Sumber gambar <http://www.proceptin.com/phc/sperm-cell.php>.)

2.3 Morfologi

Morfologi dalam arti bahasa merupakan cabang ilmu linguistik yang membahas atau mempelajari kompleksitas struktur kata dan pengaruh perubahan struktur kata terhadap kelas dan makna kata (Mustaqim, et al., 2018). Definisi morfologi di bidang biologi yakni merupakan suatu ciri yang mudah diamati pada tubuh suatu makhluk hidup. Setiap makhluk hidup pasti memiliki karakteristik morfologi yang sangat penting untuk mendeksi sifat khususnya. Pada umumnya morfologi pada makhluk hidup dipengaruhi oleh genetik dan juga dari faktor lingkungan hidup (Hadiyanti, et al., 2018). Gambar

2 merupakan morfologi sperma dan non sperma yang akan digunakan pada penelitian ini.



Gambar 2. Gambar morfologi sperma dan non sperma. (Sumber gambar [https://datasets.simula.no/visem/.](https://datasets.simula.no/visem/))

Gambar 2 merupakan gambar morfologi dari sperma dan non sperma yang bersumber dari dataset VISEM. Pada Gambar 2a adalah bentuk morfologi dari objek sperma yang terdiri dari dua bagian yakni kepala dan ekor. Pada Gambar 2b merupakan bentuk non sperma yang bentuk morfologinya tidak beraturan, seperti bulatan tidak beraturan yang tidak memiliki kepala dan ekor.

2.4 Identifikasi

Menurut Poerwadarminto (1976:369) dalam (Nalole, 2015) “identifikasi adalah penentuan atau penetapan identitas seseorang atau benda”. Menurut JP Chaplit pada (Profita, 2015) identifikasi merupakan proses suatu pengenalan, menempatkan suatu objek atau individu pada suatu kelas dengan karakteristik tertentu. Dari menurut pendapat ahli di atas dapat disimpulkan bahwa identifikasi merupakan pengenalan atau penentuan suatu objek atau individu yang dikelompokkan pada suatu karakteristik yang sama.

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 *continue* 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

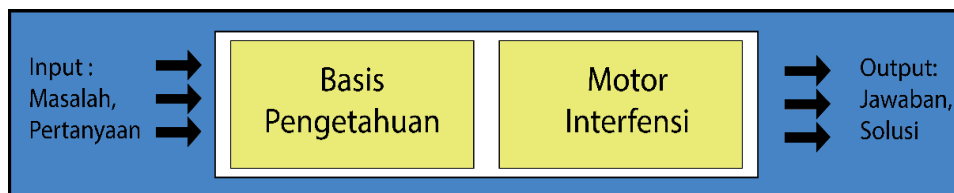
G : nilai skala keabuan

2.6 Artificial Intelligence

Artificial Intelligence atau kecerdasan buatan adalah suatu ilmu yang mempelajari bagaimana komputer dapat digunakan untuk mensimulasikan dan memperluas fungsi kinerja seperti otak manusia. Sehingga komputer dapat memiliki kemampuan layaknya manusia seperti belajar, penyimpanan,

menyelesaikan masalah, memori, pengetahuan dan penambahan bahasa alami manusia (Halim & Prasetyo, 2018).

Pembuatan suatu aplikasi kecerdasan buatan terdapat 2 bagian utama yang sangat dibutuhkan, diantaranya adalah Basis Pengetahuan (*Knowledge Base*) dan Motor Inferensi (*Inference Engine*). Pada basis pengetahuan yakni bersifat fakta-fakta, teori, pemikiran dan hubungan satu dengan yang lainnya. Sedangkan untuk Motor Inferensi merupakan kemampuan untuk menarik kesimpulan berdasarkan pengetahuan dan pengalaman (Dahria, 2008). Penerapan konsep *artificial intelligence* pada komputer dijelaskan pada Gambar 3.



Gambar 3 Penerapan konsep *artificial Intelligence*.

2.7 Machine Learning

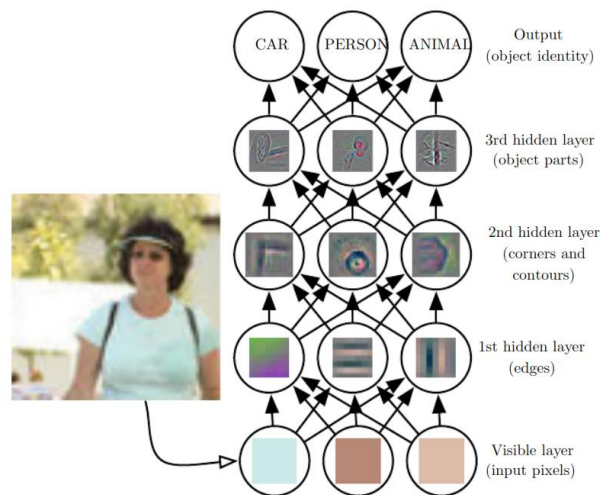
Machine learning atau pembelajaran mesin adalah suatu pendekatan dalam *artificial intelligence* yang sedang marak digunakan untuk menirukan perilaku manusia guna menyelesaikan suatu masalah atau melakukan automasi. Setidaknya terdapat dua aplikasi utama dalam *Machine Learning* yakni klasifikasi dan prediksi. *Machine Learning* memiliki ciri khas diantaranya adalah proses pelatihan, pembelajaran atau *training* (Ahmad, 2017).

Terdapat beberapa tipe algoritma yang digunakan pada *Machine Learning* diantaranya adalah *Supervised Learning*, *Unsupervised Learning*, *Reinforcement Learning*, dan *Evolutionary Learning*. Tipe yang sering digunakan banyak orang dalam *Machine Learning* ialah *Supervised Learning*. *Supervised Learning* adalah suatu training set dari banyak contoh dengan respon (target) yang benar disediakan, berdasarkan training set ini, algoritma

digeneralisasi untuk merespon dengan benar semua input yang tepat (Marsland, 2015).

2.8 Deep Learning

Deep learning adalah suatu cabang ilmu dari *machine learning* yang terinspirasi dari kortek manusia dengan mengimplementasikan jaringan syaraf buatan yang banyak memiliki *hidden layer* (Santoso & Ariyanto, 2018). Menurut Abu Ahmad *deep learning* adalah teknik yang terdapat pada *neural network* dengan menggunakan teknik tertentu yakni seperti *Restricted Boltzmann Machine* (RBM) untuk meningkatkan kecepatan proses pembelajaran pada *neural network* yang menggunakan *layer* yang banyak atau lebih dari 7 *layer* (Ahmad, 2017). Beberapa jenis *deep learning* yakni *Deep Auto Encoder* (DAE), *Deep Belief Nets* (DBN), *Convolutional Neural Network* (CNN), dan lain lain. Ilustrasi dari proses *deep learning* dari input hingga mendapatkan *output* dapat dilihat pada Gambar 4.



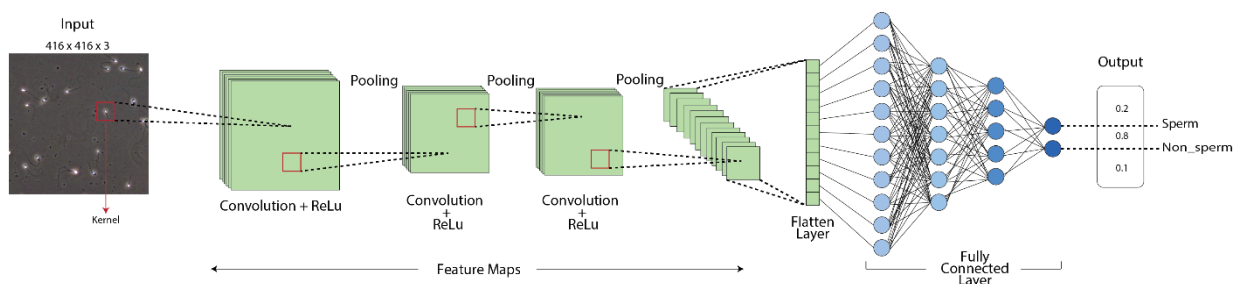
Gambar 4. Ilustrasi dari model *deep learning* (Goodfellow, et al., 2016).

Sulit bagi komputer untuk memahami arti data input sensorik mentah seperti gambar diatas yang dipresentasikan dengan sekumpulan *pixel*. *Deep learning* mampu menyelesaikan kesulitan tersebut dengan memecah pemetaan rumit yang diinginkan menjadi serangkaian pemetaan sederhana bersarang. Dari gambar diatas terdapat 5 proses *deep learning* dalam menyelesaikan

permasalahan tersebut. Pertama *layer* terlihat yang masih berupa *pixel* warna dan kemudian akan diekstrak menjadi bentuk yang abstrak *layer* ini disebut dengan *hidden layer*. Dalam *hidden layer* ini tahap pertama mampu mengidentifikasi garis atau tepi, *hidden layer* kedua mampu mengidentifikasi pencarian sudut dengan mudah serta kontur diperpanjang atau kumpulan tepi, berikutnya *hidden layer* ketiga sudah mampu mendeteksi seluruh bagian dari objek tertentu. Akhirnya, deskripsi gambar dalam hal bagian-bagian objek yang telah dikelola dapat digunakan untuk mengenali hasil objek yang ada dalam gambar (Goodfellow, et al., 2016).

2.9 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) merupakan salah satu cabang algoritma *deep learning*, digunakan pada *computer vision* untuk menyelesaikan suatu kasus atau masalah tertentu seperti mengklasifikasi dan mendeteksi suatu objek pada gambar atau video. CNN termasuk kedalam variasi *Multi Layer Perceptron* (MLP) yang terinspirasi dari jaringan syaraf pada manusia. Ciri pada CNN yakni memiliki susunan neuron 3D (tinggi, lebar, dan kedalaman). Lebar dan tinggi merupakan suatu ukuran pada *layer* sedangkan untuk kedalaman yakni menunjukkan pada jumlah *layer* (Rahim, et al., 2020). Gambar 5 merupakan ilustrasi arsitektur dari *Convolutional Neural Network* (CNN).



Gambar 5. Arsitektur dari *Convolutional Neural Network* (CNN). (Sumber gambar [https://kotakode.com/blogs/2707/Convolutional-Neural-Network-\(CNN\)\)](https://kotakode.com/blogs/2707/Convolutional-Neural-Network-(CNN)))).

2.9.1 Konvolusi

Konvolusi atau convolution merupakan suatu proses memanipulasi citra dengan menggunakan eksternal *mask* atau subwindows untuk menghasilkan suatu citra baru. Definisi konvolusi secara matematis ialah jumlah total dari perkalian antara elemen yang bersesuaian (memiliki koordinat yang sama) dalam dua matriks atau dua vektor (Rohim, et al., 2019). Contoh gambaran konvolusi diilustrasikan pada Gambar 6.

3	0	1	2	7	4
1	5	8	9	3	1
2	7	2	5	1	3
0	1	3	1	7	8
4	2	1	6	2	8
2	4	5	2	3	9

6 x 6

1	0	-1
1	0	-1
1	0	-1

3 x 3

*

-5	-4	0	8
-10	-2	-2	3
0	-2	-4	-7
-3	-2	-3	-16

4 x 4

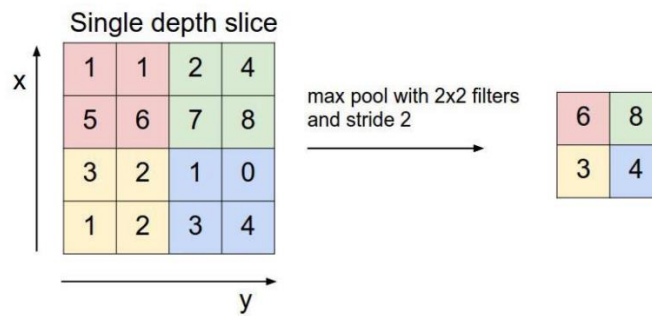
=

Gambar 6. Ilustrasi konvolusi matriks 6 x 6 (Rohim, et al., 2019).

Gambar 6 menampilkan proses untuk konvolusi pada CNN, langkah pertama dapat di lihat pada matriks input 6 × 6 yang akan dioperasikan dengan filter 3 x 3. Proses pengoperasian ini nilai matriks input akan dikalikan dengan nilai matriks filter mulai dari bagian yang telah di hijaukan dan seterusnya dengan langkah/stride 1. Sehingga nilai yang dihasilkan pada setiap perkalian dengan filter akan mengisi nilai matriks baru. Berdasarkan contoh pada Gambar 6 hasil output matriks yang dihasilkan ialah matriks 4 x 4 disebabkan oleh penggunaan langkah/stride 1.

2.9.2 Pooling

Pooling berfungsi untuk mereduksi ukuran spasial dari citra dan jumlah parameter dalam jaringan (Rohim, et al., 2019). *Pooling* juga dapat mempercepat komputasi dan mengontrol terjadinya *overfitting* (Zufar & Setiyono, 2016). Pada Gambar 7 merupakan ilustrasi dari proses *pooling*.



Gambar 7. *Matrix pooling*. Ilustrasi di atas merupakan *matrix pooling* menggunakan 2 x 2 filter dan menggunakan 2 *stride*.

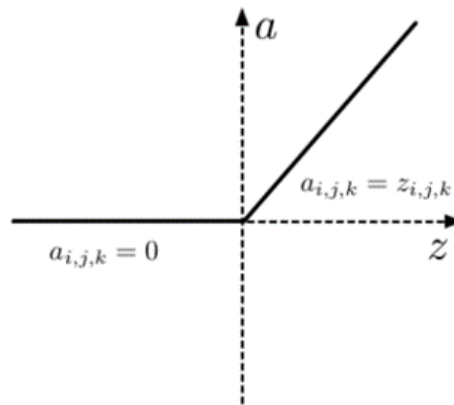
Pada Gambar 7 menampilkan contoh dari pooling yakni jenis max pooling. Max pooling ini mengambil nilai terbesar dari bagian tersebut, pada contoh Gambar 7 menggunakan filter 2 x 2 sehingga max pooling akan mengambil nilai terbesar dari setiap bagian. Dapat di lihat pada bagian warna merah muda terdapat nilai baris dan kolom diantaranya 1, 1, 5, dan 6. Sehingga dari bagian warna merah muda tersebut diambil angka yang terbesar yakni 6 untuk dijadikan nilai pada matriks baru.

2.9.3 ReLU

Rectified Linier Unit (ReLU) yakni salah satu fungsi aktivasi *non-saturated* yang paling populer. Fungsi aktivasi ini didefinisikan pada Persamaan 3 :

$$a_{i,j,k} = \max(z_{i,j,k}, 0) \dots\dots\dots(3)$$

Pada Persamaan 2, $z_{i,j,k}$ adalah input dari fungsi aktivasi di lokasi (i, j) pada *channel k-th*. ReLU adalah potongan fungsi linier yang memangkas bagian negatif menjadi nol dan mempertahankan bagian positif pada hasil konvolusi. Representasi ReLU dapat dilihat pada Gambar 8.



Gambar 8. Tampilan representasi fungsi aktivasi ReLU (Gu, et al., 2017)

Banyak penelitian telah menunjukkan bahwa ReLU bekerja lebih baik dari pada fungsi aktivasi sigmoid dan tanh secara empiris (Gu, et al., 2017).

2.9.4 Fully Connected Layer

Fully Connected Layer adalah suatu lapisan yang terhubung secara penuh atau keseluruhan, lapisan neuron ini terhubung langsung dengan neuron lainnya dengan dua lapisan yang berdekatan tanpa terhubung dengan lapisan apapun (Artyani, 2019). *Fully Connected Layer* merupakan suatu *layer* yang dipakai dalam melakukan transformasi pada suatu dimensi data sehingga dapat diklasifikasikan secara linier (Alwanda, et al., 2020).

2.10 You Only Look Once (YOLO)

You Only Look Once atau YOLO merupakan suatu pendekatan terbaru untuk deteksi objek. YOLO berbeda dengan penelitian sebelumnya yang mendeteksi suatu objek menggunakan kembali pengklasifikasinya, melainkan YOLO meringkai deteksi objek sebagai masalah regresi dengan *bounding boxes* (kotak pembatas) yang terpisah secara spasial dan probabilitas kelas yang terkait. Proses gambar di dalam YOLO ini sederhana dan mudah dapat dilihat pada Gambar 9.

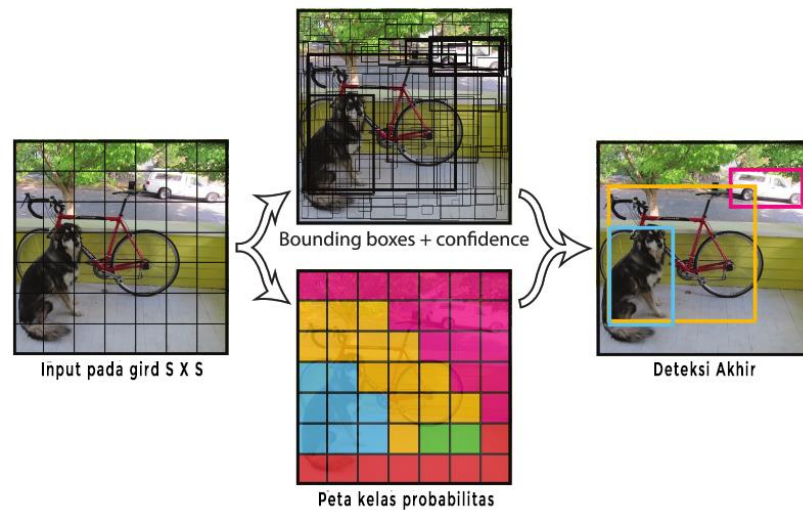


Gambar 9. Proses sistem deteksi YOLO (Redmon, et al., 2016).

Tahapan awal dalam proses gambar dalam YOLO yakni pertama mengubah ukuran gambar input menjadi 448×448 , selanjutnya menjalankan jaringan konvolusi tunggal pada gambar, dan yang terakhir yakni membatasi deteksi yang dihasilkan dengan *model's confidence*. YOLO membagi gambar input menjadi *grid* $S \times S$. Jika pusat dari suatu objek jatuh ke dalam sel *grid*, maka sel *grid* tersebut yang akan bertanggung jawab untuk mendeteksi objek tersebut. Setiap sel *grid* memprediksi B *bounding boxes* dan *confidence scores*. *Confidence score* mempresentasikan seberapa yakin model bahwa *bounding boxes* itu berisi suatu objek dan seberapa akurat objek tersebut diprediksi. Umumnya *confidence* didefinisikan sebagai $Pr(Object) * IOU_{pred}^{truth}$, jika tidak ada objek pada sel tersebut maka nilai *confidence* adalah nol. Setiap sel *grid* juga memprediksi probabilitas kelas bersyarat C atau $Pr(Class_i|Object)$. Objek yang diprediksi hanya satu set probabilitas kelas per sel *grid*, terlepas dari jumlah *boxes* B . Dalam pengujian mengalikan probabilitas kelas bersyarat dan prediksi *individual box confidence*,

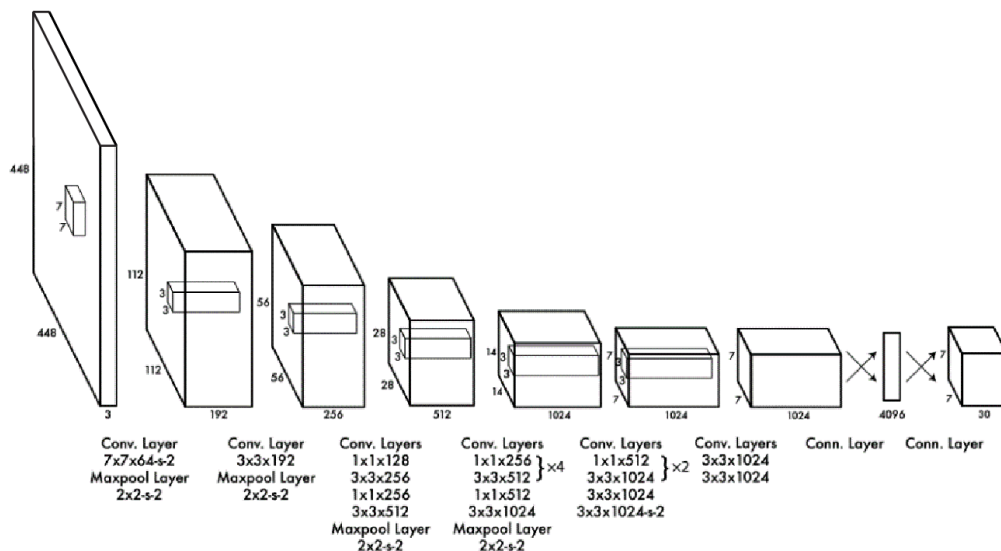
$$Pr(Class_i|Object) * Pr(Object) * IOU_{pred}^{truth} = Pr(Class_i) * IOU_{pred}^{truth} \quad (4)$$

Persamaan 4 merupakan skor *confidence* kelas khusus untuk setiap kotak. Skor ini mengkodekan probabilitas kelas itu muncul didalam kotak dan sebar baik kotak yang diprediksi cocok dengan objek. Gambar 10 merupakan gambaran dari prediksi *bounding boxes* B yang dikodekan sebagai tensor $S \times S \times (B * 5 + C)$ (Redmon, et al., 2016).



Gambar 10. Model deteksi menggunakan *bouding boxes* (Redmon, et al., 2016).

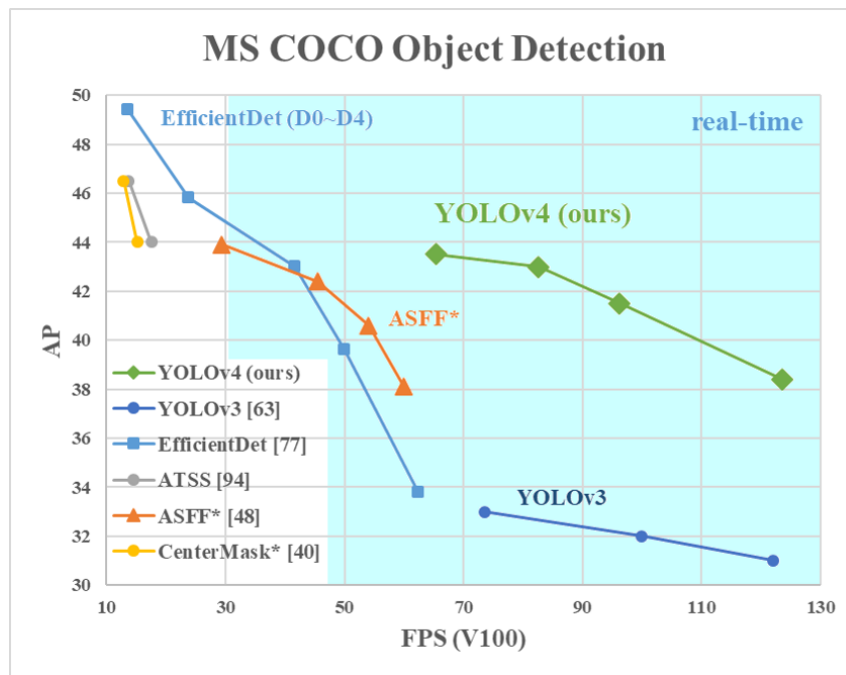
YOLO diimplementasikan sebagai *convolutional neural network*. Arsitektur ini terinspirasi oleh model GoogleNet untuk klasifikasi gambar. Jaringan pada YOLO ini memiliki 24 lapisan konvolusi diikuti oleh 2 *fully connected layers*. Secara sederhana menggunakan 1×1 reduksi *layer* yang diikuti oleh 3×3 *convolutional layers*. Prediksi dari output akhir jaringan YOLO ini adalah tensor $7 \times 7 \times 30$. Untuk gambaran arsitektur jaringannya dapat di lihat pada Gambar 11 (Redmon, et al., 2016).



Gambar 11. Arsitektur YOLO (Redmon, et al., 2016).

2.11 You Only Look Once V4 (YOLOv4)

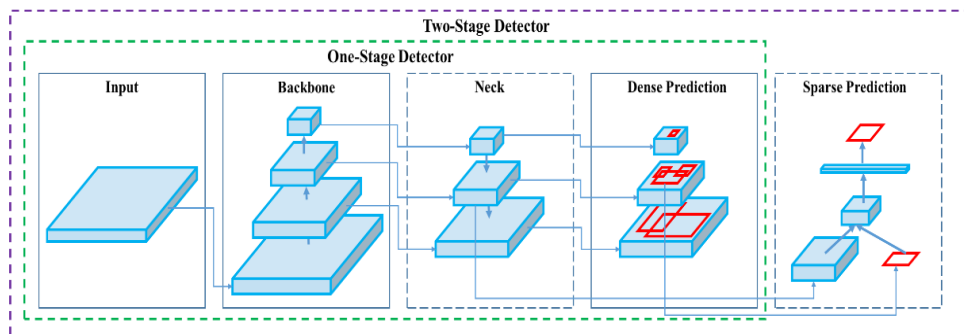
YOLOv4 dapat meningkatkan akurasi pendeteksi object secara *real-time*, memungkinkan penggunaannya mendapatkan sistem rekomendasi yang menghasilkan petunjuk serta dapat manajemen proses yang berdiri sendiri dan pengurangan input dari manusia. YOLOv4 mengatasi masalah yang ada dengan membuat CNN yang beroperasi secara *real-time* pada GPU konvensional, dan untuk training hanya memerlukan satu GPU konvensional. Tujuan dari YOLOv4 ini merancang kecepatan operasi dari detector objek dalam memproduksi sistem dan optimasi untuk komputasi paralel. YOLOv4 ini berharap objek yang dirancang dapat dengan mudah dilatih dan digunakan. Siapapun yang menggunakan GPU konvensional untuk melatih dan menguji dapat mencapai hasil deteksi objek *real-time*, kualitas tinggi, dan meyakinkan (Bochkovski, et al., 2020). Berdasarkan hasil YOLOv4 dapat dilihat pada Gambar 12.



Gambar 12. Perbandingan YOLOv4 dengan *detector* objek yang lain (Bochkovski, et al., 2020).

Berdasarkan Gambar 12 YOLOv4 berjalan dua kali lebih cepat daripada EfficientDet dengan kinerja yang sebanding. YOLOv4 memperoleh nilai AP sebesar 43.5% pada dataset COCO dengan kecepatan *real-time* 65 FPS dapat

di lihat pada garis grafik berwarna hijau. YOLOv4 meningkatkan AP dan FPS pada Tesla V100 melebihi YOLOv3 masing masing sebesar 10% dan 12%. Detector modern biasanya terdiri dari dua bagian terdiri dari *backbone* dan *head*, untuk *backbone* yang telah dilatih sebelumnya dengan ImageNet. *Head* yang digunakan untuk memprediksi *class* dan *bounding boxes* dari objek. Untuk detektor yang berjalan pada platform GPU, *backbone* yang digunakan dapat berupa VGG, ResNet, ResNetXt, atau DenseNet. Untuk detektor yang berjalan pada platform CPU, *backbone* yang digunakan berupa SqueezeNet, MobileNet, atau ShuffleNet. Pengembangan detektor objek dalam beberapa tahun terakhir sering menyisipkan beberapa *layer* diantara *backbone* dan *head*, *layer* ini biasanya digunakan untuk mengumpulkan *feature maps* dari tahapan yang berbeda. *Layer* tersebut dapat kita sebut dengan *neck object detector*. Pada umumnya *neck* terdiri dari beberapa *bottom-up path* dan *topdown path*. Jaringan yang dilengkapi oleh mekanisme ini antara lain *Feature Pyramid Network* (FPN), *Path Aggregation Network* (PAN), BiFPN, dan NAS-FPN (Bochkovskiy, et al., 2020). Berikut merupakan tampilan tahapan dari detektor pada Gambar 13.



Gambar 13. Tahapan dari detektor (Bochkovskiy, et al., 2020).

Dalam penelitian ini diperkenalkan beberapa istilah yakni *Bag of Freebies* (BoF) dan *Bag of Specials* (BoS). Definisi dari *Bag of Freebies* (BoF) adalah peneliti dapat melakukan suatu optimasi untuk menghasilkan akurasi yang lebih baik serta tidak meningkatkan biaya inferensi dengan menggunakan metode *training*. BoF yang digunakan untuk *backbone* adalah *DropBlockRegularization*, *Class Label Smoothing*, dan *CutMix* and *Mosaic Data Augmentation*. BoF untuk *backbone* ini berguna meningkatkan variabilitas pada input gambar agar model yang dibangun memiliki kualitas

yang lebih tinggi untuk gambar yang diperoleh dari *environment* yang berbeda. Definisi untuk *Bag of Special* (BoS) merupakan serangkaian modul plugin dan metode *post-processing* yang hanya meningkatkan biaya inferensi dalam jumlah kecil tetapi dapat secara signifikan meningkatkan akurasi pada deteksi objek. BoS yang digunakan pada *backbone* antara lain adalah *Mish Activation*, *Cross-Stage Partial Connections* (CSP), *Multi input Weight Residual Connection* (MiWRC) (Bochkovskiy, et al., 2020).

2.12 Pra proses Data

Pra proses merupakan suatu fase utama yang dapat menentukan kualitas untuk proses berikutnya. Tujuan dari fase pra proses ini ialah untuk memperoleh data siap olah untuk diproses oleh data mining dari data awal yang masih berupa data tekstual (Trihanto, 2017). Pra proses dapat dilakukan seperti mengubah ukuran (*resize*) dataset yang berupa sebuah citra. Tujuan dari *resize* yakni untuk mempermudah dan mempercepat citra ketika diproses (Amatullah, et al., 2021).

2.13 Confusion Matrix

Confusion Matrix adalah suatu pengukuran kinerja atau performa dalam menyelesaikan masalah klasifikasi *machine learning*, dimana hasil *output* dapat berupa dua kelas atau lebih. *Confusion matrix* merupakan suatu alat analisis prediktif yang menampilkan dan membandingkan nilai sebenarnya dengan nilai hasil prediksi model. Prediksi model yang dapat digunakan untuk mendapatkan hasil matriks evaluasi yakni Akurasi, *Precision*, *Recall* dan *F1 Score* (Rahma, et al., 2021). Berikut dapat dilihat pada Gambar 14 merupakan 4 kombinasi berbeda dari nilai prediksi dan nilai sebenarnya.

		Nilai Sebenarnya	
		1 (<i>Positive</i>)	0 (<i>Negative</i>)
Nilai Prediksi	1 (<i>Positive</i>)	TP (<i>True Positive</i>)	FP (<i>False Negative</i>)
	0 (<i>Negative</i>)	FN (<i>False Negative</i>)	TN (<i>True Negative</i>)

Gambar 14. *Confusion matrix* (Rahma, et al., 2021).

TP (*True Positive*) : Jumlah data yang bernilai positif dan diprediksi benar sebagai positif.

FP (*False Negative*) : Jumlah data yang bernilai negatif tetapi diprediksi sebagai positif.

FN (*False Negative*) : Jumlah data yang bernilai positif tetapi diprediksi sebagai negatif.

TN (*True Negative*) : Jumlah data yang bernilai negatif dengan prediksi benar sebagai negatif.

2.13.1 Average Precision (AP)

Average Precision (AP) adalah standar *de facto* untuk evaluasi kinerja dalam deteksi objek. Deteksi objek gambar statis yang populer, deteksi objek video, dan penelitian deteksi objek video *online* terutama melaporkan hasil AP dan *mean-AP*, AP tampaknya menjadi satu-satunya kriteria yang digunakan untuk membandingkan metode deteksi objek (Oksuz, et al., 2018). Dalam perhitungan AP dapat menggunakan *N-Point Interpolation*, untuk nilai *recall* dilambangkan dengan $R_r(n)$ dan sedangkan untuk nilai *precision* dilambangkan dengan $Pr_{interp}(R_r n)$. Untuk persamaan dari $R_r(n)$ dapat dilihat pada persamaan 5.

$$R_r(n) = \frac{N-n}{N-1}, \quad n = 1, 2, \dots, N. \dots\dots\dots (5)$$

Sehingga untuk persamaan AP menjadi :

$$AP = \frac{1}{N} \sum_{n=1}^N Pr_{interp}(R_r(n)). \dots\dots\dots (6)$$

Persamaan 6 merupakan persamaan untuk mencari nilai AP menggunakan *N-Point Interpolation*. Biasanya nilai N yang dipakai pada mencari hasil AP yakni $N = 11$ (Padilla, et al., 2021).

2.13.2 Precision

Precision adalah suatu rasio dari jumlah objek dengan benar atau *true positif* dibandingkan dengan seluruh data yang diprediksi positif. Nilai *precision* akan semakin rendah jika nilai dari *False Positive* (FP) semakin besar dan begitupun sebaliknya (Kusuma, et al., 2021). Persamaan presisi dapat dilihat pada Persamaan (7).

$$Precision = \frac{TP}{TP+FP} \times 100\% \dots\dots\dots(7)$$

2.13.3 Recall

Recall merupakan rasio dari jumlah objek terdeteksi dengan benar atau *True Positive* dibandingkan dengan seluruh data yang positif, *recall* yang memiliki nilai yang tinggi menunjukkan bahwa sistem dapat mengklasifikasi kelas objek dengan benar (Kusuma, et al., 2021). Persamaan recall dapat dilihat pada Persamaan (8).

$$Recall = \frac{TP}{TP+FN} \times 100\% \dots\dots\dots(8)$$

2.13.4 F1 Score

F1 Score adalah suatu perbandingan rata-rata dari nilai presisi dan *recall*. *F1 Score* memiliki nilai tertinggi sebesar 1 dan terendah sebesar 0. Nilai *F1 Score* ini semakin mendekati 1 menunjukkan bahwa kinerja sistem telah baik (Kusuma, et al., 2021). Persamaan *F1 Score* dapat dilihat pada Persamaan (9).

$$F1\ Score = 2 \times \frac{Recall \times Presisi}{Recall + Presisi} \dots\dots\dots(9)$$

2.14 IoU (*Intersection over Union*)

Intersection over union digunakan untuk menghitung suatu luas area yang berpotongan kemudian membaginya dengan luas area gabungan antara 2 *bounding box*. Pada umumnya nilai IoU ini dipakai untuk menentukan seberapa baik performa dari *bounding box* yang diprediksi dengan luas objek sesungguhnya pada suatu citra (Fandisyah, et al., 2021). Persamaan IoU dapat di lihat pada Persamaan (10).

$$IoU = \frac{\text{Area Irisan}}{\text{Area Gabungan}} \dots\dots\dots(10)$$



Gambar 15. Contoh tampilan deteksi IoU.

Gambar 15 merupakan tampilan ilustrasi dari deteksi IoU pada objek kucing, dapat dilihat terdapat dua *bounding box*. *Bounding box* merah adalah *ground truth* dan *bounding box* berwarna biru adalah *bounding box* prediksi. Sehingga jika ingin mendapat nilai IoU dari objek kucing tersebut dapat diimplementasikan Persamaan 10.

2.15 mAP (*mean Average Precision*)

Mean average precision atau mAP adalah suatu matriks yang mengevaluasi kinerja dari model deteksi objek, dan mengukur seberapa bagus performansi dari *weights file* hasil *training data* (Kusuma, et al., 2021). Nilai mAP didapatkan dari setiap nilai presisi item relevan yang dihasilkan dan menggunakan nilai 0 untuk item relevan yang tidak dihasilkan oleh sistem (Fandisyah, et al., 2021). Persamaan mAP dapat di lihat pada Persamaan (11).

$$mAP = \frac{1}{c} \sum_{i=1}^c AP_i \dots\dots\dots(11)$$

2.16 Non-Max Supression

Non-max supression digunakan untuk menyeleksi *bounding box* yang muncul *overload* atau berlebihan pada objek yang sama dengan membandingkan nilai *confidence* masing-masing *bounding box*, hanya nilai *confidence* yang paling tinggi dan yang akan dipertahankan (Nazilly, et al., 2020). *Non Max Suppression* membuat seleksi berdasarkan IoU dengan mengurangi *bounding box* yang muncul secara berlebihan (Fandisyah, et al., 2021).

2.17 Konfigurasi Hyperparameter

Hyperparameter merupakan konstanta yang dapat disesuaikan dengan kebutuhan untuk mendapatkan model yang hasil kinerja yang lebih optimal (Sutama, et al., 2020). *Hyperparameter* dapat juga dijadikan sebagai kontrol sebuah algoritma dengan melakukan penyetelan agar mendapatkan prediksi yang lebih baik (Septiani, et al., 2020). Adapun untuk macam-macam *hyperparameter* sebagai berikut.

2.17.1 Max Batches

Max Batches merupakan banyaknya iterasi pada proses *training* data. Semakin besar nilai dari *max batch* yang dipakai, maka sistem akan semakin banyak mempelajari data *training*. Jumlah dari suatu *training* data tidak boleh lebih dari jumlah *max batch*, dikarenakan nilai *max batches* perlu disesuaikan dengan jumlah kelas dari objek yang di deteksi (Kusuma, et al., 2021). Persamaan 12 berikut ini merupakan persamaan untuk mencari nilai *max batch* berdasarkan jumlah *class*.

$$\text{Max batches} = \text{number of class} \times 2000 \dots\dots\dots(12)$$

2.17.2 Subdivision

Subdivisions yaitu membagi nilai dari *batch* menjadi lebih kecil lagi yang disebut dengan *mini-batch*. Prosoes *subdivisions* ini bertujuan untuk mempercepat proses *training* dan juga meningkatkan akurasi dengan bantuan GPU. Sebagai contoh jika menggunakan nilai *batch*

sebesar 64 dan dibagi dengan 8 *subdivisions*, maka akan menghasilkan nilai 8 yang berarti dilakukannya proses *training* untuk 8 gambar tiap *mini-batch* nya. Proses tersebut terjadi selama 8 kali hingga proses *training* pada satu *batch* tersebut selesai. Selanjutnya, sistem memproses *batch* selanjutnya yang juga memiliki nilai 64 (Kusuma, et al., 2021).

2.17.3 Steps

Steps adalah *epoch* pada saat dimana *learning rate* mengecil sesuai dengan *scale factor* atau bisa juga dengan persentasi dari *max batches*. *Steps* juga sebagai pemeriksa dari jumlah iterasi yang di mana *scale* akan diterapkan (Hidayatullah, 2021).

2.17.4 Scale

Scale atau disebut juga dengan *scale factor* digunakan untuk pengecil *learning rate*. *Scale* pula sebagai koefisien dari *learning rate* untuk menentukan bagaimana *learning rate* akan diubah selama peningkatan jumlah iterasi pada proses *training* (Hidayatullah, 2021).

2.17.5 Momentum

Momentum merupakan perubahan bobot berdasarkan dari arah gradien pola terakhir dan pola sebelumnya. Ukuran nilai dari momentum ialah antara 0 sampai 1. Jika nilai momentum = 0, maka perubahan bobot hanya akan dipengaruhi oleh gradiennya. Sedangkan jika nilai momentum = 1, maka perubahan bobot akan sama dengan perubahan bobot sebelumnya (Faza, 2018).

2.17.6 Decay

Decay merupakan suatu *hyperparameter* untuk menghilangkan ketidakseimbangan dalam *dataset*. Nilai default yang digunakan pada *decay* ini ialah 0.0005 yang hasil bobotnya sudah baik (Hidayatullah, 2021).

2.17.7 *Learning rate*

Learning rate adalah penentuan jumlah bobot yang diperbaharui selama proses *back propagation*. *Learning rate* juga menentukan laju iterasi sehingga dapat mencapai fungsi kerugian yang minimum. Semakin cepat *training* maka semakin tinggi pula *learning rate*-nya. Namun *learning rate* yang memiliki nilai terlalu tinggi dapat menyebabkan nilai *loss function* turun-naik tidak menentu, sehingga diperlukan beberapa kali percobaan untuk mendapatkan nilai *learning rate* yang optimal (Kusuma, et al., 2021). *Learning rate* semakin kecil proses *training* akan lebih lama dan memungkinkan model sistem untuk melewati banyak nilai *weight* (bobot) yang penting dalam proses *training* (Hermawan, et al., 2021).

2.17.8 *Filters*

Filters dalam parameter ini yakni menyatakan untuk banyaknya jumlah dari filter yang digunakan. penentuan filter yang akan digunakan yakni menyesuaikan dengan jumlah *class* (Hidayatullah, 2021). Untuk perhitungan filter dapat dilihat pada Persamaan 13.

$$filters = (jumlah\ kelas + 5) \times 3 \dots\dots\dots(13)$$

2.17.9 *Classes*

Classes adalah parameter yang menyatakan jumlah objek *class* yang akan di *train* atau dilatih. Objek *class* yang digunakan haruslah sesuai dengan objek *class* pada data yang telah dianotasi (Hidayatullah, 2021).

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 bulan September 2021 di semester tujuh ganjil hingga penyelesaian pada bulan Maret 2022. Pengerjaan dibagi menjadi tiga tahap, yaitu :

a. Tahap Penelitian Awal

Tahap ini merupakan tahap pengumpulan data, pemahaman studi literatur dan penyusunan *draft* usul untuk bab 1 – 3. Proses ini memerlukan waktu pengerjaan selama kurang lebih 2 bulan.

b. Tahap Penelitian Lanjutan

Pada tahap ini mulai mengerjakan penelitian dengan menjalankan program untuk praproses dataset, anotasi dataset, Merancang model. Proses ini memerlukan waktu pengerjaan selama ± 3 bulan.

c. Tahap Evaluasi

Pada tahap evaluasi dilakukan untuk mengetahui tingkat keakuratan kinerja model dari uji dataset serta penulisan *draft* hasil untuk bab 4 – 5 dan penyampaian hasil penelitian melalui seminar hasil.

Alur waktu pengerjaan dapat dilihat pada Tabel 2.

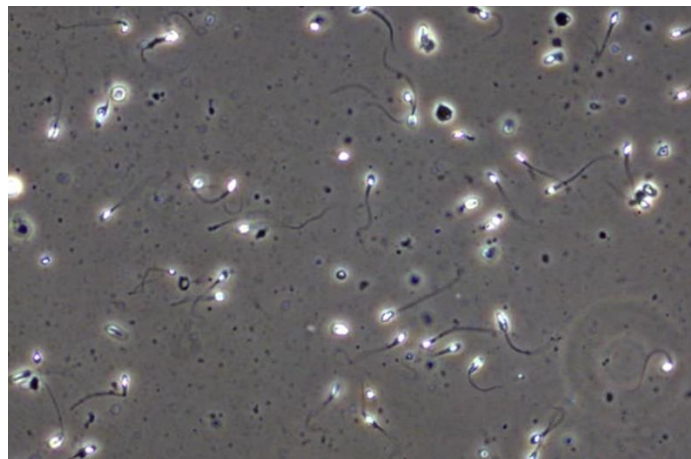
Tabel 2. Alur Waktu Pengerjaan Penelitian.

Tahapan	Kegiatan	2021								2022															
		November				Desember				Januari				Februari				Maret				April			
		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
Penelitian Awal	Penentuan Tema dan Pengumpulan Studi Literatur	■	■																						
	Pengumpulan Dataset	■	■	■	■																				
	Penyusunan Draft Proposal (Bab I - III)					■	■	■	■	■	■	■	■												
Penelitian lanjutan	Praproses Dataset					■	■																		
	Anotasi Data					■	■	■	■	■	■	■	■	■	■	■	■	■	■						
	Merancang Model									■	■	■	■	■	■	■	■	■	■						
Evaluasi	Pengujian Dataset													■	■	■	■	■	■						
	Penyusunan draft hasil (Bab IV-V)																			■	■	■	■	■	■

3.2 Data dan Alat

3.2.1 Data

Data yang digunakan bersumber dari Simula Open Dataset dengan beralamat <https://datasets.simula.no/visem/>. Dataset VISEM ini merupakan kumpulan data *multi-modal* yang berisi sumber data seperti video, data analisis biologis, dan data peserta, namun dalam penggunaan untuk penelitian ini hanya menggunakan data yang berupa video yang dijadikan dataset penelitian. *The VISEM dataset* berisikan 85 video rekaman data anonim dari 85 peserta pendonor yang berbeda. Peserta dalam dataset ini adalah laki-laki berusia 18 tahun keatas. Dataset ini dikumpulkan sejak tahun 2008 hingga 2013 dan dalam pengumpulan dataset ini diakhiri pada bulan Desember 2017. Video dengan berekstensi AVI (Audio Video Interleave) serta resolusi setiap video berukuran 640 x 480 *pixel* dengan 50 *fps frame rate*. Ukuran seluruh data video ini cukup besar yakni 35 gigabyte, dengan setiap video berdurasi antara dua hingga tujuh menit. Tampilan dari VISEM video dapat dilihat pada Gambar 16.



Gambar 16. Dataset VISEM.

3.2.2 Alat

3.2.2.1 Perangkat Keras (*Hardware*)

Perangkat keras yang digunakan pada penelitian ini yaitu ;

- a. *Processor* : Intel Core i7-9750H CPU 2.60GHz ~2.6 GHz

- b. *Installed RAM* : 8 GB
- c. *VGA* : NVIDIA GeForce GTX 1650 4GB

3.2.2.2 Perangkat Lunak (*Software*)

Perangkat lunak yang digunakan dalam penelitian ini yaitu :

- a. Sistem Operasi : Windows 10 *Home Single 64-bit*
- b. Python 3.7.7

Python merupakan bahasa pemrograman tinggi yang mampu melakukan eksekusi sejumlah instruksi secara interpretatif. Python dirilis pertama kali pada tahun 1991 oleh Guido Van Rossum.

- c. Darknet

Darknet merupakan *framework* pada *neural network* yang bersifat *open source*. *Framework* darknet ini dibuat menggunakan bahasa C dan CUDA. Kelebihan darknet ini ialah cepat, mudah dalam penginstallan dan mendukung untuk komputasi pada CPU dan GPU.

- d. OpenCV

OpenCV merupakan singkatan dari *Open Source Computer Vision Library*, sebuah *library* bersifat *open source* yang dikembangkan oleh intel. OpenCV bertujuan untuk focus dalam menyederhanakan *programming* terkait citra digital.

- e. CMake

CMake merupakan *build tools* yang dapat digunakan untuk *built*, *testing*, dan *packaging*. CMake sering digunakan untuk membantu proses pengembangan aplikasi yang dibangun dengan C++.

- f. Visual Studio

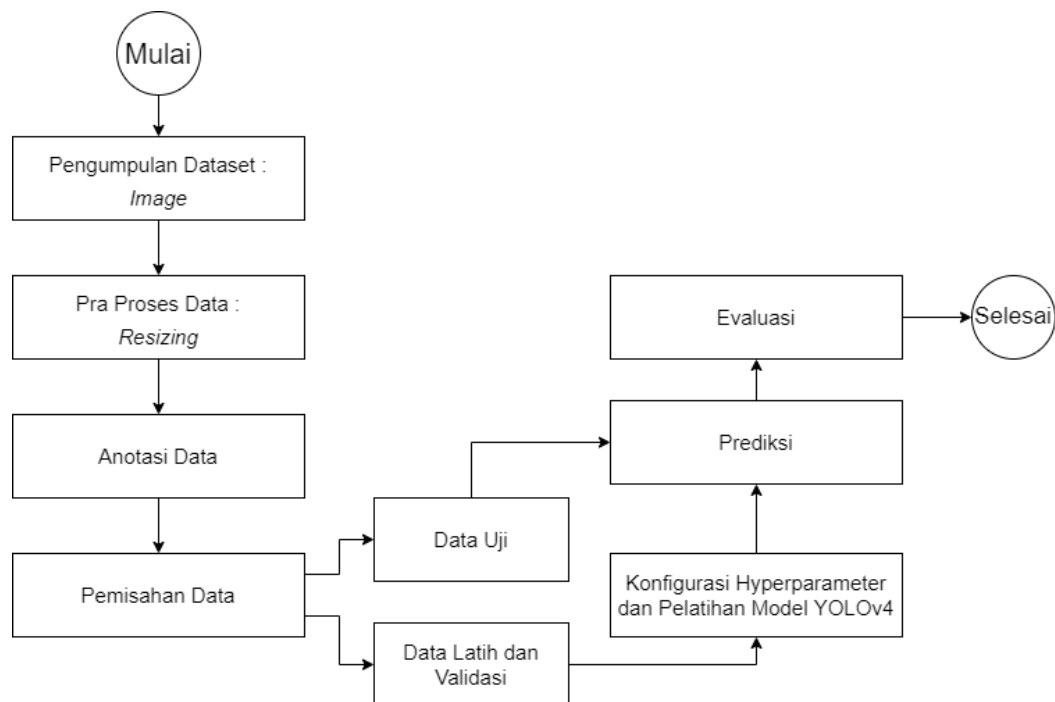
Visual studio merupakan sebuah *software* yang digunakan untuk melakukan pengembangan aplikasi, seperti aplikasi bisnis, personal, dalam bentuk aplikasi *console*, aplikasi windows maupun aplikasi web.

g. YOLOv4

YOLOv4 merupakan pengembangan dari YOLO yang menjadi *state of the art* dari model deteksi objek berbasis *deep learning*, serta dapat mendeteksi objek dengan akurat secara real time.

3.3 Alur Kerja Penelitian

Alur kerja penelitian ini memiliki beberapa tahapan yang dapat dilihat pada Gambar 17.



Gambar 17. Alur kerja penelitian identifikasi sperma manusia menggunakan algoritma YOLO.

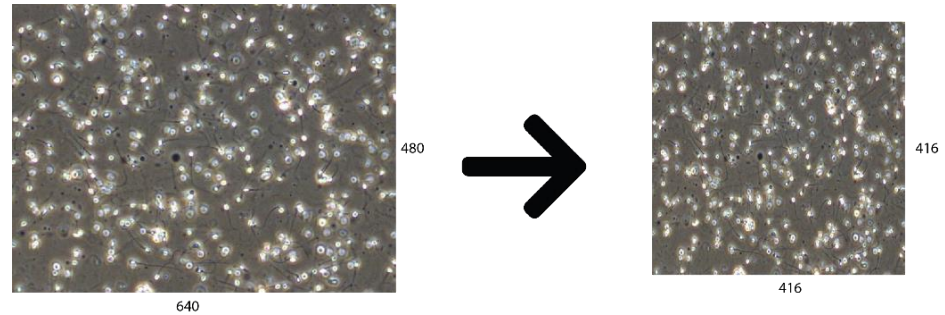
Berikut merupakan penjelasan dari alur kerja penelitian ini.

3.3.1 Pengumpulan *Dataset*

Pengumpulan *dataset* merupakan pengumpulan citra sperma yang merupakan suatu data sekunder yang didapatkan dari *Simula Open Dataset*. *Dataset* ini diolah dengan cara mengekstrak seluruh data yang masih berupa video menjadi suatu *image*. Berdasarkan 85 video rekaman dapat dihasilkan sebanyak 1330 *image*. Komposisi dari ekstraksi video yang dilakukan yakni dengan menjadikan setiap 30 detik video dibuat menjadi 2 *frame image*.

3.3.2 Pra proses *Dataset*

Pra pemrosesan data dilakukan setelah melakukan ekstraksi video menjadi image. Dikarenakan ukuran video 640 x 480 *pixel*, sehingga harus dilakukan *resizing image* untuk menyamakan dengan ukuran citra yang sesuai dengan literatur acuan dengan ukuran 416 x 416 (Redmon, et al., 2016). Ilustrasi untuk *resizing* dapat dilihat pada Gambar 18.



Gambar 18. Hasil *resizing* citra.

3.3.3 Anotasi Data

image yang akan dilatih. *Class* yang dibuat pada anotasi ini adalah *sperm* dan *non sperm*. Masing-masing *frame* diberikan anotasi berupa *bounding boxes* pada bagian morfologi kepala sperma dan suatu benda yang bukan sperma. Untuk anotasi ini menggunakan *Yolo_mark* yang bersumber dari github https://github.com/AlexeyAB/Yolo_mark.git. Berikut tampilan dari anotasi data dapat dilihat pada Gambar 19.



Gambar 19. Tampilan anotasi data.

3.3.4 Pembagian Data

Pada tahap ini dilakukan dengan membagi keseluruhan data untuk proses *train*, *validation*, dan *test*. Data yang digunakan untuk *train* yaitu sebanyak 80%, untuk data *validation* sebanyak 10%, dan data *test* sebanyak 10% dari total seluruh data.

3.3.5 Konfigurasi *Hyperparameter* dan Pelatihan Model YOLOv4

Hyperparameter dan pelatihan yang digunakan dalam model yang dirancang ini dapat dilihat pada Tabel 3.

Tabel 3. *Hyperparameter* Pelatihan Model

<i>Hyperparameter</i>	<i>Default</i>	Deepsperm	YOLOv4 Sperm
<i>Batch</i>	64	64	64
<i>Subdivision</i>	64	16	64
<i>Width dan height</i>	608 x 608	640 x 640	416 x 416
<i>Max_batch</i>	500500	4000	6000
<i>Steps</i>	400000	3600	4800, 5400
<i>Scale</i>	.1, .1	.1, .1	.1, .1
<i>Momentum</i>	0.949	0.9	0.949
<i>Decay</i>	0.0005	0.0005	0.0005
<i>Learning rate</i>	0.00261	0.001	0.001
<i>Filters</i>	255	18	21
<i>Classes</i>	80	1	2

Batch merupakan jumlah sampel data yang akan diproses dalam satu iterasi. *Subdivision* yaitu jumlah dari *minibatch* dalam satu *batch*. Dalam hal ini *subdivision* yang digunakan sebesar 64, dikarenakan model akan di train menggunakan GPU (*Graphics Processing Unit*). *Width* dan *height* merupakan dari resolusi layer input. *Width* dan *height* yang digunakan yakni menyesuaikan dengan ukuran *default* dari YOLOv4. *Max_batch* merupakan jumlah *epoch* pelatihan. Dalam perhitungan *max_batch* harus disesuaikan dengan jumlah kelas objek. Sehingga perhitungan untuk menentukan *max_batch* yakni jumlah

kelas x 2000. Namun apabila dari perhitungan tersebut nilai *max_batch* kurang dari 6000 maka *max_batch* nya ditetapkan sebesar 6000.

Steps adalah *epoch* pada saat *learning rate* mengecil sesuai dengan *scale factor*. *Step* memiliki nilai yang didapatkan berdasarkan 80% dan 90% dari nilai *max_batch*. Sehingga didapatkan untuk nilai *steps* sebesar 4800 dan 5400. *Scale* digunakan untuk *scale factor* pengecil *learning rate*. Nilai *scale* ditentukan berdasarkan dari nilai *step*, dikarenakan nilai *step* lebih dari 600 maka pada *scale* menggunakan nilai 1. *Momentum* merupakan *hyperparameter* yang digunakan untuk membantu mengetahui langkah selanjutnya dengan pengetahuan langkah sebelumnya pada proses *train*. *Decay* merupakan *hyperparameter* untuk menghilangkan ketidakseimbangan dalam *dataset*. *Learning rate* adalah *hyperparameter* yang menentukan seberapa cepat jaringan memperbarui *weights* pada saat proses *train*. *Filters* dalam parameter ini menyatakan sebagai jumlah dari *filter* atau kumpulan *filter* yang digunakan dengan menyesuaikan jumlah *class*. Untuk perhitungan nilai *filters* ini yakni $(\text{jumlah kelas} + 5) \times 3$ sehingga nilai *filters*-nya adalah 21 sebab memiliki 2 kelas objek. *Classes* adalah jumlah objek *class* yang akan di *train* atau dilatih.

3.3.6 Evaluasi

Proses evaluasi dilakukan untuk mengetahui tingkat keakuratan suatu kinerja model. Keakuratan kinerja model dapat diuji dengan menghitung nilai IoU antara prediksi kotak pembatas dengan kotak pembatas yang sebenarnya. Selanjutnya nilai IoU tersebut akan digunakan pada *confusion matrix* untuk menghitung *accuracy*, *recall*, *precision*, dan mAP.

V. PENUTUP

5.1 Simpulan

Berdasarkan penelitian dan pembahasan yang sudah dilakukan mengenai identifikasi sperma manusia berdasarkan morfologi menggunakan algoritma YOLO dapat diambil beberapa kesimpulan, antara lain.

1. YOLO dapat mengidentifikasi sperma dan non sperma berdasarkan morfologi, yang dimana model telah mampu membuat *bounding box* setiap objek yang ada pada *frame* video dataset VISEM.
2. Dalam proses *training* menggunakan tiga skenario pembagian data dan tiga jenis *learning rate*. Skenario pertama yakni data *train* 80%, validasi 10%, dan *testing* 10%, skenario kedua data *train* 70%, validasi 25%, dan *testing* 5%, serta skenario ketiga data *train* 60%, validasi 20%, dan *testing* 20%. Semua skenario telah dilakukan uji *training* dengan tiga jenis *learning rate* yakni 0.002, 0.0002, dan 0.00002.
 - a. Hasil uji *training* menggunakan *learning rate* 0.002 pada skenario pertama dengan pembagian data *train* 80%, validasi 10%, dan *testing* 10% menghasilkan mAP 78%, *precision* 0.61, *recall* 0.65, dan F1-score 0.63. Pada skenario kedua dengan pembagian data *train* 70%, validasi 25%, dan *testing* 5% menghasilkan mAP 77.83%, *precision* 0.75, *recall* 0.87, F1-score 0.81. Lalu pada skenario ketiga dengan pembagian data *train* 60%, validasi 20%, dan *testing* 20% menghasilkan mAP 76.88%, *precision* 0.7, *recall* 0.88, F1-score 0.78.
 - b. Hasil uji *training* menggunakan *learning rate* 0.0002 pada skenario pertama dengan pembagian data *train* 80%, validasi 10%, dan

testing 10% menghasilkan mAP 78.83%, *precision* 0.77, *recall* 0.84, dan F1-score 0.8. Pada skenario kedua dengan pembagian data *train* 70%, validasi 25%, dan *testing* 5% menghasilkan mAP 79.58%, *precision* 0.73, *recall* 0.87, F1-score 0.8. Lalu pada skenario ketiga dengan pembagian data *train* 60%, validasi 20%, dan *testing* 20% menghasilkan mAP 78.64%, *precision* 0.77, *recall* 0.82, F1-score 0.8.

- c. Hasil uji *training* menggunakan *learning rate* 0.00002 pada skenario pertama dengan pembagian data *train* 80%, validasi 10%, dan *testing* 10% menghasilkan mAP 74.74%, *precision* 0.74, *recall* 0.82, dan F1-score 0.77. Pada skenario kedua dengan pembagian data *train* 70%, validasi 25%, dan *testing* 5% menghasilkan mAP 76.11%, *precision* 0.73, *recall* 0.83, F1-score 0.77. Lalu pada skenario ketiga dengan pembagian data *train* 60%, validasi 20%, dan *testing* 20% menghasilkan mAP 75.76%, *precision* 0.76, *recall* 0.8, F1-score 0.78.

Hasil akurasi terbaik yakni terdapat pada percobaan dengan nilai *learning rate* 0.0002 yang nilai akurasinya sebesar 79.58% mAP, *precision* 0.73 pada pembagian data *train* 70%, validasi 25% dan *test* 5%. Setiap proses percobaan untuk *training* menggunakan 6000 iterasi untuk membuat data latihnya. Pengujian dalam penelitian ini menggunakan video yang hasilnya semua objek dapat dideteksi dengan baik dan telah berlabel dengan *bounding box*. Dalam penelitian ini terdapat kasus dimana model tidak mampu mendeteksi secara optimal dikarenakan data video yang digunakan terdapat objek yang blur dan objek sperma yang terpotong oleh *frame*.

5.2 Saran

Saran pada penelitian ini, antara lain.

1. Memperbanyak dataset baru dengan kualitas gambar yang baik dan jelas serta memerlukan jumlah data class yang telah dianotasi yang seimbang.

2. Pada penelitian selanjutnya dapat memodifikasi model YOLO sehingga model YOLO dapat menghitung jumlah objek yang sedang diidentifikasi.
3. Dapat ditambahkannya jumlah class objek baru seperti class objek sperma normal dan abnormal.

DAFTAR PUSTAKA

- Agustinus, I'tishom, R. & Pramesti, D. M., 2018. *BIOLOGI REPRODUKSI PRIA*. Surabaya: Airlangga University.
- Ahmad, A., 2017. Mengenal Artificial Intelligence, Machine Learning, Neural Network, dan Deep Learning. *Jurnal Teknologi Indonesia*, p. 4.
- Alwanda, M. R., Ramadhan, R. P. K. & Alamsyah, D., 2020. Implementasi Metode Convolutional Neural Network Menggunakan Arsitektur LeNet-5 untuk Pengenalan Doodle. *Jurnal Algoritme*, Volume 1, pp. 45-56.
- Amatullah, L., Ein, I. & Santoni, M. M., 2021. Identifikasi Penyakit Daun Kentang Berdasarkan Fitur Tekstur dan Warna Dengan Menggunakan Metode K-Nearest Neighbor. *Seminar Nasional Mahasiswa Ilmu Komputer dan Aplikasinya (SENAMIKA)*, pp. 783-786.
- Artyani, I., 2019. Simulasi Metode Convolutional Neural Network dan Long Short-Term Memory untuk Generate Image Captioning Pada Gambar Lalu Lintas Kendaraan Berbahasa Indonesia. *Journal Teknik Informatika UINJKT*, p. 27.
- Bochkovskiy, A., Wang, C.-Y. & Liao, H.-Y. M., 2020. YOLOv4: Optimal Speed and Accuracy of Object Detection. *Cornell University*, pp. 1-17.
- Dahria, M., 2008. Kecerdasan Buatan (Artificial Intellegence). *Jurnal SAINTIKOM*, Volume 5, p. 185.
- Dhyani, I. A. D., Kurniawan, Y. & Negara, M. O., 2020. Hubungan Antara Faktor-Faktor Penyebab Infertilitas Terhadap Tingkat Keberhasilan IVF-ICSI di

- RSIA Puri Bunda Denpasar Pada Tahun 2017. *JURNAL MEDIKA UDAYANA*, pp. 2-5.
- Fandisyah, A. F., Iriawan, N. & Winahju, W. S., 2021. Deteksi Kapal di Laut Indonesia Menggunakan YOLOv3. *JURNAL SAINS DAN SENI ITS*, Volume 10, pp. D26-D30.
- Faza, S., 2018. *PENINGKATAN KINERJA DALAM PENGKLASIFIKASIAN MENGGUNAKAN DEEP LEARNING*. Medan: Universitas Sumatera Utara.
- Goodfellow, I., Bengio, Y. & Courville, A., 2016. *Deep Learning*. London: The MIT Press Cambridge, Massachusetts.
- Gu, J. et al., 2017. Recent Advances in Convolutional Neural Networks. *ELSEVIER*, p. 8.
- 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.
- Halim, C. & Prasetyo, H., 2018. Penerapan Artificial Intelligence Dalam Computer Aided Instructure(CAI). *Jurnal Sistem Cerdas*, Volume 1, p. 1.
- Hermawan, M. I., Tritoasmoro, I. I. & Ibrahim, N., 2021. PENGATURAN LAMPU LALU LINTAS BERDASARKAN KEPADATAN KENDARAAN MENGGUNAKAN METODE YOLO. *e-Proceeding of Engineering*, 8(1), pp. 198-205.
- Hicks, t. A. et al., 2019. Machine Learning-Based Analysis of Sperm Videos and Participant Data for Male Fertility Prediction. *Springer Nature*, pp. 1-5.
- Hidayatullah, P., 2021. *Baku Sakti Deep Learning Computer Vision Menggunakan YOLO untuk Pemula*. Cimahi: Stunning Vision AI Academy.
- Hidayatullah, P. et al., 2021. DeepSperm : A robust and real-time bull sperm-cell detection in densely populated semen videos. *ELSEVIER*, pp. 1-12.
- Iqbal, I., Mustafa, G. & Ma, J., 2020. Deep Learning-Based Morphological Classification of Human Sperm Heads. *Diagnostics*, pp. 2-5.

- Kumala, D. R., 2019. Pengaruh Model Pembelajaran Kooperatif Tipe Team Assisted Individualization (TAI) Terhadap Penguasaan Konsep Pada Materi Sistem Reproduksi Kelas XI Di SMA Negeri 1 Bandar Lampung. *repository radenintan*, pp. 61-63.
- Kusumanto, R. & Tompunu, 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. 1-2.
- Kusuma, T. A. A. H., Usman, K. & Saidah, S., 2021. PEOPLE COUNTING FOR PUBLIC TRANSPORTATIONS USING YOU ONLY LOOK ONCE METHOD. *Jurnal Teknik Informatika (JUTIF)*, Volume 2, pp. 60-64.
- Marsland, S., 2015. *Machine Learning & Pattern Recognition Series*. Second Edition ed. New York: Taylor & Francis Group, LLC.
- Mustaqim, N., Hanye, P. & Seli, S., 2018. Morfologi Bahasa Dayak Pompakng. *Jurnal Pendidikan dan Pembelajaran Khatulistiwa*, pp. 3-10.
- Nalole, M., 2015. Meningkatkan Kemampuan Mengidentifikasi Sifat-Sifat Bangun Datar Melalui Model Pembelajaran Kooperatif Tipe Teams Games Turnament (Tgt) Pada Siswa Kelas VI SDN 5 Pantungo Kabupaten Gorontalo. *UNG*, p. 18.
- Nazilly, M. L., Rahmat, B. & Puspaningrum, E. Y., 2020. IMPLEMENTASI ALGORITMA YOLO (YOU ONLY LOOK ONCE) UNTUK DETEKSI API. *Jurnal Informatika dan Sistem Informasi (JIFoSI)*, 1(1), pp. 81-91.
- Nissen, M. S. et al., 2017. Convolutional neural networks for segmentation and object detection of human semen. *Cornell University*, pp. 1-6.
- Oksuz, K., Cam, B. C., Akbas, E. & Kalkan, S., 2018. Localization Recall Precision (LRP): A New Performance Metric for Object Detection. *ECCV 2018*, pp. 504 - 519.

- Oktarina, A., Abadi, A. & Bachsin, R., 2014. Faktor-faktor yang Memengaruhi Infertilitas pada Wanita di Klinik Fertilitas Endokrinologi Reproduksi. *Majalah Kedokteran Sriwijaya*, Volume 46.
- Padilla, R. et al., 2021. A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. *Electronics*.
- Profita, R. P., 2015. Identifikasi Motif Menonton Tayangan Program Televisi “Laptop Si Unyil” TRANS 7 Pada Siswa SDN 010 Kec. Samarinda Utara Kel. Sungai Pinang Dalam Samarinda. *eJournal Ilmu Komunikasi*, p. 4.
- Putri, A. R., 2016. Pengolahan Citra Dengan Menggunakan Web Cam Pada Kendaraan Bergerak Di Jalan Raya. *JUPI (Jurnal Ilmiah Pendidikan Informatika)*, Volume 1, p. 2.
- Rahim, A., Kusriani & Luthfi, E. T., 2020. CONVOLUTIONAL NEURAL NETWORK UNTUK KLASIFIKASI PENGGUNAAN MASKER. *Jurnal Teknologi Informasi dan Komunikasi*, Volume 10, p. 110.
- Rahma, L., Syaputra, H., Mirza, A. & Purnamasari, S. D., 2021. Objek Deteksi Makanan Khas Palembang Menggunakan Algoritma YOLO (You Only Look Once). *Jurnal Nasional Ilmu Komputer*, pp. 214-217.
- Ratna, S., 2020. Pengolahan Citra Digital dan Histogram Dengan Python dan Text Editor Pycharm. *Technologia*, Volume 11, p. 2.
- Redmon, J., Divvala, S., Girshick, R. & Farhadi, A., 2016. You Only Look Once: Unified, Real-Time Object Detection. *IEEE Xplore*, pp. 1-9.
- Rohim, A., Sari, Y. A. & Tibyani, 2019. Convolution Neural Network (CNN) Untuk Pengklasifikasian Citra Makanan Tradisional. *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, Volume 3, p. 7038.
- S.Ningsih, Y. J. & Farich, A., 2016. Determinan Kejadian Infertilitas Pria di Kabupaten Tulang Bawang. *Jurnal Kesehatan*, pp. 8-5.
- Santoso, A. & Ariyanto, G., 2018. Implementasi Deep Learning Berbasis Keras Untuk Pengenalan Wajah. *Jurnal Teknik Elektro*, Volume 18, p. 15.

- Septiani, A., Jondri & Rizal, A., 2020. Klasifikasi Suara Paru Normal dan Abnormal dengan Menggunakan Discrete Wavelet Transform dan Support Vector Machine. *Jurnal Tugas Akhir Fakultas Informatika*, 8(1), pp. 731-742.
- Sherwood, L., 2011. *Fisiologi manusia : dari sel ke sistem*. 6 ed. Jakarta: Buku Kedokteran EGC.
- Sutama, V. A., Wibowo, S. A. & Rahmania, R., 2020. Investigasi Pengaruh Step Training pada Metode Single Shot Multibox Detector untuk Marker dalam Teknologi Augmented Reality. *JURNAL ILMIAH FIFO*, XII(1), pp. 1-11.
- Syauqy, A., 2014. Evaluasi Kromatin Sperma Sebagai Indikator Kualitas Sperma. *Jambi Medical Journal*, Volume 2, pp. 1-4.
- Trihanto, W. B., 2017. Information Retrieval System Untuk Menentukan Tren Judul Jurnal Berbahasa Indonesia Menggunakan Metode TF-IDF Dan Naïve Bayes Classifier. *UNIVERSITAS NEGERI SEMARANG*, pp. 13-14.
- Valiuškaitė, V. et al., 2020. Deep Learning Based Evaluation of Spermatozoid Motility for Artificial Insemination. *Sensors*, pp. 2-8.
- Zufar, M. & Setiyono, B., 2016. Convolutional Neural Networks untuk Pengenalan Wajah Secara Real-Time. *JURNAL SAINS DAN SENI ITS*, Volume 5, p. 74.