

BAB II

LANDASAN TEORI

2.1 Penelitian Terdahulu

Tabel 2.1 Tabel Penelitian Terdahulu

No	Nama Jurnal	Judul Artikel dan Author	Hasil Penelitian
1.	Journal of Plant Diseases and Protection[21]	<p>Judul Artikel: <i>Automated Tomato Leaf Disease Classification using Transfer Learning-based Deep Convolution Neural Network</i></p> <p>Authors: Rajasekaran Thangaraj, S. Anandamurugan, Vishnu Kumar Kaliappan</p> <p>Tahun: 2021 Volume: 128 Halaman: 73-86</p>	<p>Penelitian ini dilakukan untuk mendeteksi dan mengklasifikasikan penyakit pada daun tomat menggunakan algoritma <i>Convolution Neural Network</i> (CNN) dengan model <i>transfer learning Modified Xception</i>. Penelitian ini membandingkan 3 jenis <i>optimizer</i> yaitu <i>stochastic gradient descent</i>, <i>adaptive moment estimation</i>, dan RMSprop. Hasil yang diperoleh adalah nilai akurasi dari <i>adaptive moment estimation</i> lebih tinggi dari pada analisis lainnya dengan nilai 99.55%, dibandingkan dengan RMSprop 99.01% dan <i>stochastic gradient descent</i> 81.77%</p>
2.	Sustainability[22]	<p>Judul Artikel: <i>Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique</i></p> <p>Authors: Yonis Gulzar</p> <p>Tahun: 2023 Volume: 15 Issue: 3 Halaman: 1906</p>	<p>Penelitian ini dilakukan untuk membantu dalam mengklasifikasikan berbagai jenis buah sebanyak 40 kategori, dengan membandingkan berbagai macam model <i>transfer learning</i> diantaranya adalah AlexNet, VGG16, InceptionV3, ResNet, dan MobileNetV2. Masing-masing model yang digunakan memiliki perbedaan nilai evaluasi yang cukup berbeda berdasarkan akurasi, <i>precision</i>, <i>recall</i>, dan <i>F1 Score</i>. Didapatkan nilai evaluasi tertinggi yaitu model MobileNetV2.</p> <p>Selain itu penelitian ini melakukan <i>tuning model</i> MobileNetV2 dengan menerapkan <i>data augmentation</i>, <i>adaptive learning rate</i> sebesar 0.001,</p>

No	Nama Jurnal	Judul Artikel dan Author	Hasil Penelitian
			<i>model checkpointing</i> dan <i>dropout</i> yang menghasilkan nilai evaluasi lebih tinggi dibandingkan sebelum di <i>tuning</i> .
3.	IAES International Journal of Artificial Intelligence[23]	<p>Judul Artikel: <i>Lotus Species Classification using Transfer Learning Based on VGG16, ResNet152V2, and MobileNetV2</i></p> <p>Authors: Nachirat Rachburee, Wattana Punlumjeak</p> <p>Tahun: 2022 Volume: 11 Issue: 4 Halaman: 1344-1352</p>	<p>Penelitian ini berhasil mengklasifikasikan 5 jenis tanaman <i>lotus</i>. Penelitian ini menggunakan algoritma CNN dengan model VGG16, ResNet152V2, dan MobileNetV2. Dilakukan <i>fine-tuning</i> dengan menerapkan <i>dropout</i> untuk mengurangi <i>overfitting</i>. Dengan <i>optimizer adaptive moment estimation</i>, dan menambahkan <i>learning rate</i> sebesar 0,0001. Setelah dilakukan pengklasifikasian, didapatkan bahwa MobileNetV2 memiliki akurasi yang lebih tinggi dengan akurasi 99,5%, di lanjutkan dengan VGG16 98,1%.</p>
4.	ICIC Express Letters, Part B: Applications [24]	<p>Judul Artikel: <i>Classification of Skin Diseases and Disorders Using Convolutional Neural Network on a Mobile Application</i></p> <p>Authors: Nico Bernando Setiawan, Friska Natalia, Ferry Vincenttius Ferdinand, Sud Sudirman, Chang Seong Ko</p> <p>Tahun: 2021 Volume: 12 Issue: 8 Halaman: 715-721</p>	<p>Penelitian ini membahas penyakit kulit di Indonesia, yang memuat klasifikasi beberapa penyakit seperti <i>psoriasis</i>, kurap, eskim, dan kulit sehat dengan menggunakan 8000 macam gambar. Algoritma dalam penelitian ini diimplementasikan pada aplikasi android. Metode yang digunakan pada penelitian ini adalah metode <i>Deep Convolutional Neural Network</i> dengan menerapkan teknik pelatihan model <i>cross validation</i>. Penelitian ini membagi <i>dataset</i>-nya menjadi <i>validation</i> dan <i>training</i> dengan perbandingan 80:20. Setelah dilakukan pemodelan pertama, data <i>validation</i> memiliki akurasi sebesar 81.7% dengan <i>loss</i> 0.64 dan data <i>training</i> memiliki akurasi yang lebih besar sejumlah 93% dengan total <i>loss</i> yang lebih rendah sebesar 0.22. Selanjutnya, pada akhir percobaan dilakukan</p>

No	Nama Jurnal	Judul Artikel dan Author	Hasil Penelitian
			pengubahan <i>k-fold</i> menjadi 20 sehingga <i>training</i> dan <i>validation</i> menjadi 95:5, yang menghasilkan nilai akurasi dari <i>validation</i> sebesar 90% dan loss sebesar 0.39
5.	International Journal of Online and Biomedical Engineering (iJOE)[26]	<p>Judul Artikel: <i>Brain Tumor Classification Deep Learning Model Using Neural Networks</i></p> <p>Authors: Gisella Luisa Elena Maquen Niño, Ariana Ayelen Sandoval-Juarez, Robinson Andres Veliz-La Rosa, Gilberto Carrión-Barco, Ivan Adrianzén-Olano, Hugo Vega-Huerta, Percy De-La-Cruz-VdV</p> <p>Tahun: 2023 Volume: 19 Issue: 9 Halaman: 81-92</p>	Penelitian ini melakukan pengklasifikasian <i>brain tumor</i> menggunakan algoritma CNN, dengan model ResNet50. Penelitian ini menggunakan <i>metodology Cross Industry Standard Process for Data Mining (CRISP-DM)</i> sebagai model proses <i>data mining</i> yang mencakup keseluruhan penelitian.
6.	Applied Sciences (Switzerland) [17]	<p>Judul Artikel: <i>RecipeIS—Recipe Recommendation System Based on Recognition of Food Ingredients</i></p> <p>Authors: Miguel Simões Rodrigues, Filipe Fidalgo, dan Ângela Oliveira</p> <p>Tahun: 2023 Volume: 13 Issue: 13</p>	Penelitian ini melakukan pembentukan sistem rekomendasi menu makanan berdasarkan citra bahan makanan. Digunakan sebanyak 36 kategori bahan makanan untuk penelitian ini. Sistem ini mengimplementasikan algoritma CNN dengan model ResNet-50 yang berhasil mendapatkan akurasi sebesar 96%.
7.	Mathematical Modelling of Engineering Problems [27]	<p>Judul Artikel: <i>Development of a Web and Mobile Applications-Based Cassava Disease Classification Interface Using Convolutional Neural Network</i></p> <p>Authors: Kennedy Okokpujie, Imhade P. Okokpujie, Odumuyiwa Ayomikun, Abidemi Orimogunje, Adebayo Ogundipe</p>	Penelitian ini melakukan pengidentifikasian pada penyakit singkong, dengan membuat sistem berbasis <i>website</i> dan aplikasi <i>mobile</i> . Penelitian ini melakukan perbandingan model MobileNetV2, VGG16, dan ResNet50 dengan menggunakan 2756 data gambar yang terbagi menjadi enam kategori. Datanya dibagi menjadi tiga, <i>training</i> sebesar 70%, <i>testing</i> 20% dan <i>validasi</i> 10%. Penelitian ini juga

No	Nama Jurnal	Judul Artikel dan Author	Hasil Penelitian
		Tahun: 2023 Volume: 10 Issue: 1 Halaman: 119-128	menerapkan <i>transfer learning</i> untuk pembentukan modelnya. Hasil evaluasi modelnya didapati bahwa MobileNetV2 memiliki nilai tertinggi dengan nilai akurasi pelatihan sebesar 98%, dilanjutkan dengan VGG16 dengan nilai akurasi pelatihan sebesar 92%, dan ResNet50 dengan nilai akurasi sebesar 75%.
8.	Frontiers in Nanotechnology [28]	Judul Artikel: SKCV: <i>Stratified K-fold Cross-Validation on ML Classifiers for Predicting Cervical Cancer</i> Authors: Sashikanta Prusty, Srikanta Patnaik, Sujit Kumar Dash Tahun: 2022 Volume: 4	Penelitian ini melakukan klasifikasi untuk memprediksi penyakit <i>cervical cancer</i> . Pada penelitian ini menerapkan <i>cross validation</i> dengan menggunakan <i>stratified cross validation</i> untuk melakukan pengolahan modelnya. Digunakan nilai <i>n_split</i> sebesar 5 untuk membagi proporsi data menjadi pelatihan dan uji dengan distribusi yang merata tiap kelasnya. Penggunaan <i>stratified classification</i> ini bekerja dengan lebih baik untuk menghasilkan evaluasi yang lebih akurat.
9.	Journal of The Institution of Engineers (India): Series B [29]	Judul Artikel: <i>Competitive Deep Learning Methods for COVID-19 Detection using X-ray Images</i> Authors: H. Swapnarekha, Himansu Sekhar Behera, Debanik Roy, Sunanda Das, Janmenjoy Nayak Tahun: 2021 Volume: 102 Halaman: 1177-1190	Penelitian ini melakukan pendeteksian COVID-19 menggunakan data gambar <i>X-Ray</i> sebanyak 624. Data yang digunakan pada penelitian ini melalui tahapan augmentasi yang menerapkan <i>rotation, fill mode, horizontal dan vertical shift range, shear dan zoom range, horizontal flip</i> . Penelitian ini melakukan pengujian pada dua model MobileNetV2 dan VGG16, serta melakukan pengujian <i>optimizer</i> Adam dan RMSprop. Didapati bahwa model MobileNetV2 berhasil mengungguli model VGG16 dan <i>optimizer</i> terbaik menggunakan RMSprop. Dihasilkan akurasi sebesar 93%, <i>precision</i> 92%, <i>recall</i> 94%, dan <i>f1-score</i> 93%.
10.	Telecommunication, Computing, Electronics	Judul Artikel: <i>Vision: A Web Service for Face</i>	Penelitian ini dilakukan untuk membentuk sistem pengenalan

No	Nama Jurnal	Judul Artikel dan Author	Hasil Penelitian
	and Control (TELKOMNIKA)[30]	<i>Recognition using Convolutional Network</i> Authors: Akino Archilles, Arya Wicaksana Tahun: 2020 Volume: 18 Issue: 3 Halaman: 1389-1396	wajah dengan menggunakan algoritma CNN yang berbasis <i>website</i> . Dalam mengatasi potensi kesalahan pengidentifikasian, penelitian ini menerapkan mekanisme <i>thresholding</i> dengan nilai ambang batas 0.9 dan 0.8. Jika <i>confidence level</i> berada di bawah 0.8 maka sistem tidak dapat mengenali.

Berdasarkan tabel 2.1 penelitian terdahulu, menunjukkan berbagai penelitian terdahulu mengenai klasifikasi dan pengolahan gambar menggunakan algoritma *Convolutional Neural Network* (CNN) dengan berbagai model *transfer learning*, studi kasus, dan implementasi yang berbeda. Pada penelitian ini akan menggunakan algoritma CNN untuk mengenali jenis bahan makanan dengan tujuan untuk memberikan rekomendasi terkait menu MP-ASI [17], [24]. Penggunaan metode *transfer learning* akan diterapkan pada penelitian ini, untuk memaksimalkan model dan meminimalkan komputasi yang berlebih serta waktu pada proses pengembangan model, sama seperti yang diterapkan pada penelitian [21], [22], [23]. Penerapan MobileNetV2 dan VGG16 dilakukan berdasarkan hasil penelitian terdahulu yang membandingkan berbagai model, dan menghasilkan kedua model tersebut sebagai model terbaik [23], [27]. Dalam penelitian tersebut, MobileNetV2 dan VGG16 terbukti sebagai model yang paling unggul dalam evaluasi metrik. Oleh karena itu, kedua model tersebut dipilih untuk digunakan dalam penelitian ini. Penelitian ini akan membandingkan model MobileNetV2 dan VGG16 sebagaimana dilakukan pada penelitian [22], [23], [27], [29]. Penerapan MobileNetV2 dan VGG16 dilakukan didasari hasil penelitian terdahulu yang Penelitian ini akan menerapkan teknik *data mining* sebagai kerangka kerja dari penelitian ini, dengan menggunakan CRISP-DM [25] perapan *splitting* data *training* dan *validation* data sebesar 80% : 20% sebagaimana dilakukan pada percobaan pertama di penelitian mengacu pada penelitian [21], [24]. Data yang diuji akan dilakukan augmentasi untuk memberikan gambaran yang lebih luas untuk melatih model, seperti yang dilakukan pada penelitian [29]. Penelitian ini akan menerapkan metode

optimizer menggunakan *adaptive moment estimation* seperti pada penelitian [21], [22], serta akan menerapkan *fine tuning* pada keseluruhan model agar menghasilkan evaluasi terbaik [22], [23]. Diterapkan juga pelatihan model menggunakan *cross validation* seperti pada penelitian [24], [28]. Model terbaik akan dilihat berdasarkan evaluasi yang dilakukan berdasarkan metrik evaluasi akurasi, presisi, *recall*, dan *f1-score*, sama halnya yang diterapkan pada penelitian [22], [29]. Output dari penelitian ini adalah mengimplementasikan model terbaik dalam bentuk aplikasi Android seperti yang dilakukan pada penelitian [24], serta fitur rekomendasi menu MP-ASI seperti pada penelitian [17]. Agar mengatasi potensi kesalahan mengidentifikasi dan mengklasifikasikan bahan makanan, diterapkan mekanisme *thresholding* untuk menghindari kesalahan dalam memberikan rekomendasi makanan, karena disebabkan salahnya pengidentifikasian bahan makanan [30]. Berdasarkan penelitian terdahulu yang sudah dilakukan sebelumnya yang dapat dilihat pada Tabel 2.1, terdapat beberapa perbedaan penelitian. Pada Penelitian ini menerapkan teknik *data mining Cross Industry Standard Process for Data Mining (CRISP-DM)* dengan permasalahan yang berbeda, untuk memberikan rekomendasi MP-ASI dalam mencegah *stunting* [17], [25]. Penelitian akan menerapkan mekanisme *thresholding* untuk membantu menghindari kesalahan rekomendasi akibat salah dalam pendeteksian dan pengklasifikasian karena terlalu rendahnya *confidence level* [30]. Pelatihan model pada penelitian ini akan diterapkan sebanyak dua menggunakan *stratified cross validation* dan tanpa *stratified cross validation*.

2.2 Tinjauan Teori

2.2.1 Gizi Seimbang Bayi dan Balita

Gizi menjadi salah satu faktor penentu untuk mencapai tumbuh kembang yang optimal pada bayi dan balita [31]. Ketika gizi awal diberikan mengalami kekurangan akan berakibat gagal tumbuh pada anak sehingga berakibat anak lebih pendek dari normal. Pengaruh lainnya adalah kondisi kognitif anak akan terganggu, anak akan rentan memiliki penyakit, dan yang paling fatal akan berakibat kematian. Masa bayi dan balita merupakan masa periode emas pertumbuhan untuk mempercepat pemulihan, perkembangan,

dan mengurangi intensitas penyakit yang menginfeksi [32]. Kerusakan yang terjadi pada periode tersebut akan berakibat fatal karena bersifat tidak dapat pulih.

Gizi seimbang pada makanan yang harus dikonsumsi sehari-hari mencakup 2 kategori, makronutrien dan mikronutrien [33]. Masing-masing kategori ini terdiri dari [32], [34]:

a) Makronutrien

Makronutrien merupakan zat yang terkandung untuk memberikan energi kepada tubuh yang dibutuhkan dalam jumlah besar dengan satuan gram. Makronutrien terdiri atas:

1) Karbohidrat

Karbohidrat merupakan sumber penghasil energi utama yang diperlukan oleh tubuh. 1 gram karbohidrat akan menghasilkan 4 kkal. Karbohidrat berperan dalam mengatur metabolisme yang ada pada tubuh. Pemberian karbohidrat yang berlebih akan berdampak pada penyakit tidak menular seperti obesitas.

2) Protein

Protein merupakan komponen utama sistem komunikasi antar sel, yang berperan sebagai penyusun enzim, hormon, dan pengangkut zat gizi.

3) Lemak

Lemak merupakan senyawa organik yang terbentuk menyerupai karbohidrat dengan kandungan oksigen yang lebih sedikit dibandingkan karbohidrat, namun menghasilkan kalori lebih banyak. Lemak berperan sebagai pelindung organ dalam tubuh, memproteksi dari cedera.

b) Mikronutrien

Mikronutrien merupakan zat yang terkandung untuk menambah nilai fungsi pada zat gizi makro agar berfungsi dengan baik. Zat gizi

mikro dibutuhkan dalam jumlah kecil dengan satuan miligram. Mikronutrien terdiri atas:

1) Vitamin

Vitamin merupakan zat gizi yang penting yang memberikan dampak besar pada kesehatan. Vitamin akan membantu balita terhindar dari penyakit kronis pada saat dewasa. Vitamin dibutuhkan dalam jumlah kecil pada tubuh namun memiliki peranan penting pada pertumbuhan, menjaga kesehatan, dan metabolisme. Mengacu pada Kemenkes dibutuhkan vitamin A, vitamin B kompleks, vitamin C, dan vitamin D untuk membantu tumbuh kembang bayi [31].

2) Mineral

Mineral memiliki peranan penting dari berbagai enzim sebagai proses metabolisme. Mineral juga berperan memelihara fungsi pada tubuh. Mengacu pada Kemenkes Mineral yang dibutuhkan pada bayi adalah kalsium, zat besi, yodium, zink dan fosfor [31].

2.2.2 *Stunting*

Stunting merupakan gangguan pertumbuhan tinggi badan yang tidak sesuai dengan umur seusianya yang terjadi dalam jangka waktu yang panjang yang disebabkan oleh ketidakcukupan konsumsi gizi. Kondisi ini terjadi akibat dari kekurangan gizi yang kronis, yang terjadi dalam masa kandungan dan pada awal setelah kelahiran yang disebut dengan 1000 Hari Pertama Kehidupan (HPK) yang dampaknya akan terlihat ketika anak berusia 2 tahun. Dampak lain *stunting* akan berdampak pada perkembangan kognitif anak dengan tingkat produktivitas diperkirakan hanya setengah dari potensinya. Beragam faktor yang menyebabkan *stunting* dapat mempengaruhi status gizi anak baik secara langsung maupun tidak langsung. Faktor langsung mencakup asupan makanan, penyakit menular, dan penilaian pribadi anak seperti jenis kelamin dan berat badan saat lahir. Sementara itu, faktor tidak

langsung meliputi pemberian ASI yang tidak eksklusif, layanan kesehatan, dan karakteristik keluarga [32].

2.2.3 Bahan Makanan

Bahan makanan merupakan elemen yang diperoleh dari sektor perkebunan, pertanian, peternakan, dan teknologi makanan yang mengandung semua unsur gizi yang penting, serta dikonsumsi secara rutin untuk memenuhi kebutuhan nutrisi dan pertumbuhan tubuh [35]. Bahan makanan terbagi menjadi beberapa jenis berdasarkan sumber dan karakteristiknya, diantaranya [35], [36]:

a) Bahan makanan nabati

Bahan makanan nabati merupakan jenis bahan makanan yang berasal dari tumbuh-tumbuhan yang dapat dikonsumsi dan memiliki nilai gizi. Bahan makanan nabati meliputi sayur-sayuran, buah-buahan, kacang-kacangan, umbi-umbian, sereal, dan produk olahannya.

b) Bahan makanan hewani

Bahan makanan hewani merupakan bahan makanan yang berasal dari hewan, yang mencakup daging, unggas, telur, dan susu. Bahan makanan hewani ini merupakan sumber protein yang baik bagi tumbuh kembang.

c) Lemak dan minyak

Lemak merupakan senyawa yang terbentuk dari asam dan gliserol, macam-macam bahan lemak berupa mentega, keju, susu. Lemak juga terbagi menjadi lemak nabati dan lemak hewani. Lemak memiliki peranan penting sebagai sumber energi bagi tubuh. Minyak merupakan bahan cair yang terbentuk dari lemak nabati maupun hewani.

d) Ikan dan hasil laut

Ikan dan hasil laut merupakan binatang yang memiliki ekosistem atau hidup berada di air tawar maupun air laut. Ikan juga merupakan bahan makanan yang memiliki sumber protein.

e) *Food additive*

Food additive merupakan bahan yang sengaja ditambahkan pada makanan dengan takaran yang sedikit untuk memperbaiki citarasa, bentuk, warna, tekstur, dan memperpanjang waktu simpan. *Food additive* ini bekerja untuk mempertahankan ataupun memperbaiki mutu makanan.

2.2.4 Makanan Pendamping ASI(MP-ASI)

Air Susu Ibu (ASI) adalah cairan yang penting dan diperlukan bayi, yang diberikan secara eksklusif sampai dengan usia 6 bulan. ASI digunakan untuk menggantikan makanan dan minuman selama 6 bulan tersebut demi mendukung tumbuh kembang bayi tanpa ada tambahan. Ketika bayi sudah mencapai umur 6 bulan, penting bagi orang tua untuk memberikan Makanan Pendamping ASI (MP-ASI) hingga berusia 24 bulan atau lebih. MP-ASI digunakan untuk memperoleh tambahan asupan untuk memenuhi gizi dan energi bagi bayi karena ASI yang diberikan tidak cukup memenuhi gizi. Perlu diperhatikan pemberian MP-ASI yang terlalu dini akan berdampak pada bayi yang menyebabkan penyakit pencernaan, diare, alergi makanan, dan perubahan selera makanan[37].

Terdapat kriteria dalam pemberian makanan pada bayi berdasarkan umur, yang dapat dilihat pada Tabel 2.2 [38], [39].

Tabel 2.2 Tabel Rekomendasi MP-ASI pada Bayi

Usia Bayi	Tekstur	Frekuensi	Jumlah
6-8 Bulan	Bubur kental, makanan lumat	2-3 kali/hari pada menu utama. 1-2 kali/hari pada makanan selingan	Mulai dengan 2-3 sdm tiap kali makan, tingkatkan

			bertahap hingga 1/2 mangkok (125 ml)
9-11 Bulan	Dicincang halus, serta makanan yang dapat digenggam bayi	3-4 kali/hari pada menu utama. 1-2 kali/hari pada makanan selingan	1/2 - 3/4 mangkok ukuran 250 ml (125 – 200 ml)
12-23 Bulan	Makanan keluarga	3-4 kali/hari pada menu utama. 1-2 kali/hari pada menu selingan	3/4 - 1 mangkok ukuran 250 ml

Menurut *World Health Organization* (WHO) dan Ikatan Dokter Anak Indonesia (IDAI) merekomendasikan pemberian MP-ASI saat bayi dengan rentang usia 6-8 bulan komposisinya adalah 70% ASI dan 30% MP-ASI, sedangkan rentang usia 9-11 bulan perbandingan komposisinya adalah 50% untuk masing-masing MP-ASI dan ASI. Pada usia 12-23 bulan komposisinya adalah 70% MP-ASI dan 30% ASI [40].

2.3 Kerangka Kerja, Algoritma, Teknik *Processing* dan Metode Evaluasi

2.3.1 CRISP-DM

Cross Industry Standard Process Model for Data Mining (CRISP-DM) merupakan suatu kerangka kerja guna berorientasi pada pemrosesan masalah bisnis yang diimplementasikan pada *data mining* dengan tujuan menghasilkan suatu gambaran umum [41]. CRISP-DM itu sendiri terdiri dari 6 tahapan diantaranya [42], [43]:

a) *Business Understanding*

Business understanding merupakan tahapan penentuan tujuan dan kebutuhan yang ingin dicapai dari sudut pandang dan permasalahan bisnis, kemudian diolah dan diselesaikan dengan proses *data mining*.

b) *Data Understanding*

Data Understanding merupakan fase mempersiapkan pemahaman data yang digunakan, dengan melakukan pengumpulan, pengidentifikasian, dan mengevaluasi data yang akan digunakan berdasarkan atributnya.

c) *Data Preparation*

Data Preparation merupakan fase dalam melakukan pemilihan data, membersihkan data seperti menghilangkan *missing values* dan melakukan transformasi data sehingga data dapat dilakukan pemodelan.

d) *Modeling*

Modelling merupakan tahapan pemodelan data dengan membangun model algoritma sesuai dengan data dengan tujuan menghasilkan pola yang maksimal.

e) *Evaluation*

Evaluation merupakan tahapan dalam mengevaluasi seluruh rangkaian dan model yang telah dibuat agar tetap sejalan dengan tujuan yang ada pada *business understanding*, serta memastikan standar yang dibangun dalam pembentukan pemodelan algoritma.

f) *Deployment*

Deployment adalah tahapan penerapan model yang telah dibuat dengan melakukan penyebaran informasi dengan tujuan memberikan gambaran kebijakan yang harus dilakukan yang dapat termuat pada laporan ataupun analisa.

2.3.2 *Machine Learning*

Machine learning merupakan pembelajaran atau penelitian yang berfokus pada ilmu komputer dan teknik yang menerapkan sistem kecerdasan buatan yang memungkinkan pembuatan pola yang bermanfaat. Penggunaan *machine learning* bertujuan untuk mempermudah pekerjaan dengan jalan keluar suatu hal yang rumit dengan membentuk suatu algoritma untuk menghasilkan pola tertentu yang sulit ditemukan oleh manusia [44].

a) Metode *Machine Learning*

Terdapat tiga jenis metode pada *machine learning* yang sering digunakan untuk melakukan pendeteksian dan pengklasifikasian pola, metode tersebut adalah:

1) *Supervised Machine Learning*

Supervised machine learning merupakan teknik yang digunakan untuk pendekatan secara universal untuk hubungan non-linear antara data *input* dan *output*. Metode ini meminimalkan kesalahan antar *output* yang dihasilkan. Metode *supervised learning* yang paling sering adalah analisis *classification*. Beberapa algoritma *supervised learning* yaitu *Random Forest*, *Convolutional Neural Network* (CNN), dan lainnya [45].

2) *Unsupervised Machine Learning*

Unsupervised machine learning merupakan teknik yang digunakan untuk menemukan kesamaan atau pola antara nilai dengan asumsi bahwa beberapa data dalam kumpulan data akan memiliki tingkat kesamaan yang lebih tinggi daripada yang lain. Metode *unsupervised learning* yang paling sering adalah analisis *clustering*, yang digunakan dalam analisis data untuk mencari pola atau pengelompokan tersembunyi dalam data. Beberapa algoritma *unsupervised learning* yaitu *K-Means*, *DBSCAN*, dan lainnya [46].

3) *Reinforcement Machine Learning*

Reinforcement machine learning merupakan teknik yang digunakan untuk menemukan strategi perilaku yang paling optimal untuk *trial* dan *error*. Terdapat dua tipe *reinforcement learning* yaitu *positive* dan *negative* [47].

2.3.3 Citra Digital

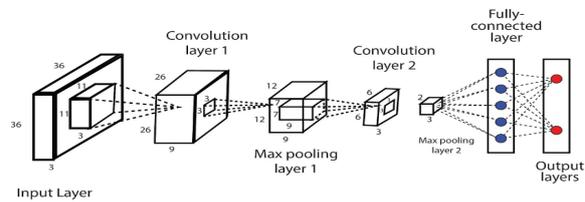
Citra merupakan suatu gambaran yang menggambarkan kemiripan pada objek tersebut. Citra digital ini merepresentasikan suatu visual berupa foto dan video, yang dikelola oleh komputer. Citra digital berbentuk biner dengan dua dimensi diskrit $f(x, y)$ yang akan membentuk *pixel*. Terdapat beberapa pembeda pada citra yaitu citra berwarna, citra *grayscale* dan citra biner [48].

2.3.4 Deep Learning

Deep learning merupakan sub-bagian dari *machine learning* yang bekerja dengan merepresentasikan dan mempelajari berbagai tingkatan secara hirarki yang berbeda, sehingga memungkinkan pembangunan konsep yang kompleks menjadi yang lebih sederhana dan bermakna. Deep Learning memiliki lapisan dengan jumlah lapisan sampai ratusan lapis yang berurutan. Lapisan tersebut dapat secara otomatis mempelajari data *training*. *Deep learning* ini melakukan pembelajaran untuk mendeteksi kata, mengidentifikasi gambar, dan membuat prediksi [49]. Terdapat tiga langkah utama dalam melakukan evaluasi pada model *deep learning* yaitu pemrosesan data, pemodelan dataset, dan validasi data [50].

2.3.5 Convolutional Neural Network

Convolutional Neural Network (CNN) merupakan suatu alat yang memungkinkan untuk melakukan pemecahan masalah pada suatu pola tertentu. Penggunaan CNN terkenal dengan lapisan *input* dan *output*, serta lapisan yang tersembunyi. Lapisan tersembunyi ini adalah *convolutional layer*, *pooling layer*, dan *full connected layer*. *Convolutional Neural Network* merupakan penerapan dari *Artificial Neural Network* yang memiliki model terbaik untuk melakukan penyelesaian pada permasalahan suatu pengenalan pada objek atau citra [51].



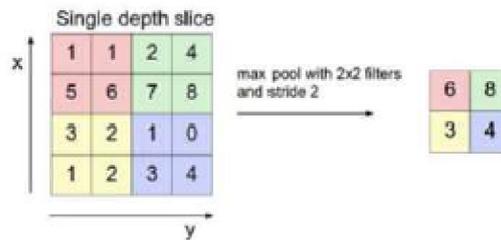
Gambar 2.1 Arsitektur Convolutional Neural Network [52]

a) *Convolutional Layer*

Convolutional layer merupakan lapisan yang mempelajari, mendeteksi dan menangkap inputan gambar. Selain itu, blok ini merupakan inti dari jaringan saraf pada *convolutional* yang bertindak sebagai filter untuk mengekstrak pengidentifikasian objek yang bermakna pada gambar. Filter ini mencakup lebar, tinggi, tipe konvolusi, dan nilai rata-rata pada elemen filter. *Convolutional layer* melakukan pemrosesan gambar dengan melakukan pergeseran pada filter, yang akan melakukan pemecahan gambar menjadi potongan kecil (*pixel*). *Rectified Linear Unit (ReLU)* akan bertugas untuk menyatukan dan memperbaiki kembali *pixel* yang bersifat *nonlinear* [53].

b) *Pooling Layer*

Pooling layer merupakan lapisan yang mengurangi dimensi dari *feature map* dengan tujuan untuk melakukan *downsampling* sehingga dapat mempercepat komputasi. Terdapat tiga jenis *pooling* yaitu *max pooling*, *min pooling*, dan *average pooling*. *Max pooling* adalah nilai terbesar pada matriks yang akan memilih nilai terbesar sesuai dengan bentuk filter, contoh *max pooling* terdapat pada Gambar 2.1. *Min pooling* merupakan nilai yang terkecil pada matriks, yang memiliki nilai filter terkecil. *Average pooling* merupakan filter yang akan dirata-rata terlebih dahulu [54].



Gambar 2.2 Max Pooling [55]

c) *Full Connected Layer*

Full connected layer merupakan lapisan akhir yang mengumpulkan keseluruhan hasil layer sebelumnya. Dapat dilihat pada Gambar 2.2 *full connected layer* berfungsi untuk melakukan proses klasifikasi akhir yang akan dilakukan komputasi skor untuk tiap kelas dengan probabilitas tertentu berdasarkan hasil *training*. Hal tersebut dapat digunakan untuk menggabungkan setiap fitur yang telah dilakukan proses agar menjadi vektor untuk diklasifikasi [56].

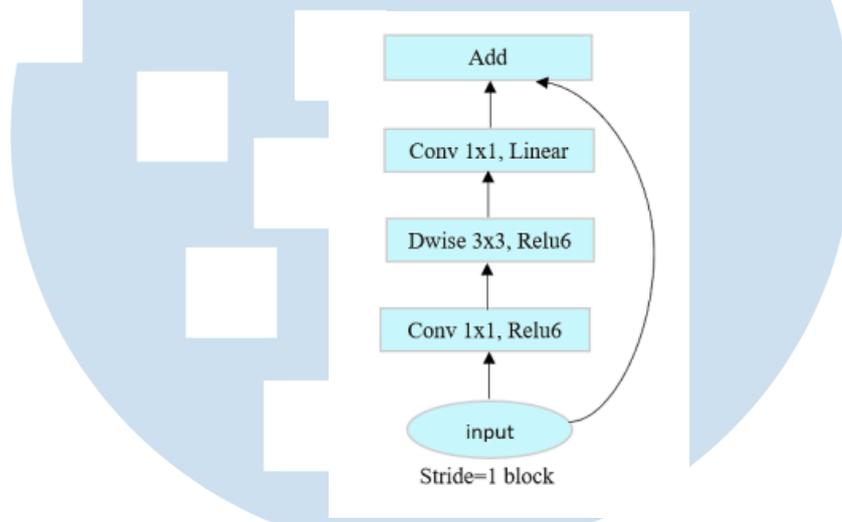
2.3.6 *Transfer Learning*

Transfer Learning merupakan salah satu metode *machine learning* yang menggunakan *pre-trained* model dengan tujuan untuk mengulangi pekerjaan yang telah dilakukan sebelumnya. Pembelajaran ini berkaitan dengan kecepatan sistem dalam menyesuaikan situasi, lingkungan, dan tugas baru. Pembelajaran ini memberikan kemudahan sistem untuk memecahkan masalah meskipun dengan ketersediaan data yang minim. Pembelajaran ini juga memungkinkan penerapan ulang karena pengetahuan yang dilakukan sebelumnya [57].

a) MobileNetV2

MobileNet merupakan suatu arsitektur dari *Convolutional Neural Network* yang digunakan dalam *image processing*. MobileNet itu sendiri terdiri dari beberapa versi, yang salah satunya adalah versi kedua yaitu MobileNetV2. MobileNetV2 memiliki keunggulan pada modelnya yang ringan dari versi sebelumnya, selain itu MobileNetV2 dirancang untuk perangkat *mobile*. Arsitektur dari MobileNetV2 ini

menerapkan *inverted bottleneck blocks* dan *residual connections* yang dipergunakan untuk menggantikan *convolutions* yang besar sehingga dapat mengurangi jumlah kinerja sistem dan memori yang dibutuhkan namun dengan mempertahankan akurasi yang sama bisa dilihat pada Gambar 2.3. Input gambar pada model ini direkomendasikan menggunakan ukuran 224 x 224 *pixels* dengan 3 RGB *channels* [23].



Gambar 2.3 *Convolutional Blocks* MobileNetV2[23]

b) VGG16

VGG16 merupakan suatu arsitektur dari *Convolutional Neural Network* untuk melakukan pengolahan gambar, yang terdiri dari beberapa lapisan konvolusional yang aktif melalui Rectified Linear Unit (ReLU) [58].

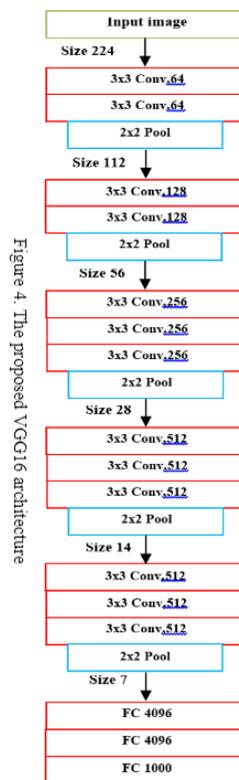


Figure 4. The proposed VGG16 architecture

Gambar 2.4 VGG16 Arsitektur [59]

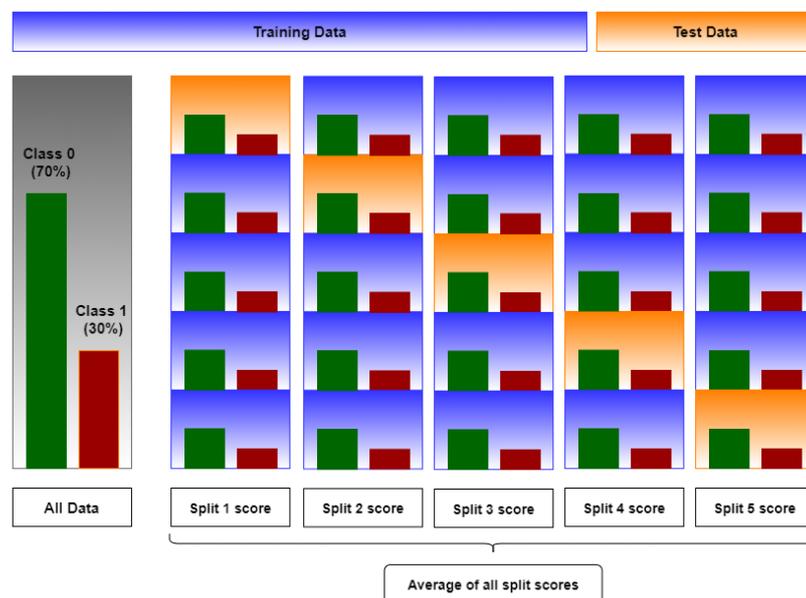
Lapisan konvolusional ini berukuran 3 x 3, *spatial pooling* dijalankan dengan 5 *max pooling layers*. *Max pooling* dilakukan dengan ukuran 2 x 2 *pixels*. Input model ini menggunakan gambar dengan ukuran 224 x 224 *pixels* dengan 3 RGB *channels*. Arsitektur VGG16 ini memiliki 16 lapisan yang terbagi menjadi 13 lapisan adalah *convolutional layers*, 2 *full connected layers*, dan 1 *softmax classifier* yang dapat dilihat pada Gambar 2.4[59].

2.3.7 Teknik *Processing*

a) *Cross Validation*

Cross validation merupakan teknik yang digunakan pada *machine learning* untuk mengevaluasi kinerja suatu model[60]. *Cross validation* melibatkan pembagian *dataset* menjadi beberapa lipatan serta melatih model pada lipatan tersebut. Proses ini diulang untuk setiap lipatan dan

hasilnya digabungkan untuk memberikan estimasi keseluruhan kinerja model. Pembagian *dataset* tersebut terbagi menjadi dua yaitu data yang digunakan untuk pelatihan dan data yang digunakan untuk memvalidasi model. Evaluasi dari *cross validation* ini dipengaruhi seberapa baik partisi data yang dibuat. Salah satu jenis *cross validation* adalah *Stratified K-Fold Cross Validation*. *Stratified K-Fold Cross Validation* merupakan teknik yang digunakan untuk menilai kinerja model dengan tujuan menjaga keseimbangan distribusi kelas di tiap foldnya, terutama pada data yang tidak seimbang atau distribusi kelas yang tidak sama. Dalam hal ini, *dataset* dibagi menjadi *K-Fold*, dan setiap lipatan memiliki rasio yang sama dari variabel target yang ada dalam *dataset*, terlihat pada Gambar 2.5. Hal tersebut, dilakukan untuk memastikan bahwa model dilatih dan diuji pada sampel yang mewakili dari setiap kelas, memberikan estimasi yang lebih akurat tentang kinerja model [28].



Gambar 2.5 Implementasi *Stratified K-Fold* [61]

2.3.8 Evaluasi

a) *Confusion Matrix*

Confusion matrix merupakan alat evaluasi yang digunakan untuk menganalisis kinerja prediktif dari suatu model klasifikasi. *Confusion matrix* memberikan ringkasan tentang seberapa baik model tersebut dapat memprediksi kelas-kelas yang benar dan yang salah dalam dataset. *Confusion matrix* terdiri dari empat entri utama: *True Positive* (TP), *True Negative* (TN), *False Positive* (FP), dan *False Negative* (FN). Dapat dilihat pada Tabel 2.3 mewakili kombinasi antara kelas yang diprediksi dengan benar dan yang salah, serta kelas yang seharusnya diprediksi dengan benar[60].

Tabel 2.3 *Confusion Matrix*

		Prediksi (<i>Predicted</i>)	
		<i>Negative</i>	<i>Positive</i>
Sebenarnya (<i>Actual</i>)	<i>Negative</i>	TN	FP
	<i>Positive</i>	FN	TP

b) Akurasi

Akurasi merupakan suatu nilai yang dihasilkan melalui hasil bagi keseluruhan data yang diprediksi benar, serta penggunaan akurasi dapat dievaluasi untuk keberhasilan model lainnya. Akurasi dihitung dengan melakukan pembagian jumlah prediksi yang benar dengan total jumlah prediksinya.[62]

$$\text{Akurasi} = \frac{(TP + TN)}{(TP + FP + TN + FN)}$$

Rumus 2.1 Rumus Mencari Nilai Akurasi[63]

Keterangan Rumus:

TP (True Positive): Jumlah sampel positif yang diprediksi dengan benar

TN (True Negative): Jumlah sampel negatif yang diprediksi dengan benar.

FP (*False Positive*): Jumlah sampel negatif yang salah diprediksi sebagai positif.

FN (*False Negative*): Jumlah sampel positif yang salah diprediksi sebagai negatif.

c) *Precision*

Presisi merupakan suatu nilai ketepatan informasi yang dikeluarkan atau dihasilkan oleh prediksi sistem atau model untuk menunjukkan secara jelas dan benar mengenai informasi data kelas positif [64]. Presisi berkaitan dengan seberapa banyak prediksi positif yang benar.

$$Precision = \frac{(TP)}{(TP + FP)}$$

Rumus 2.2 Rumus Mencari Nilai *Precision* [63]

Keterangan Rumus:

TP (*True Positive*): Jumlah sampel positif yang diprediksi dengan benar

TN (*True Negative*): Jumlah sampel negatif yang diprediksi dengan benar.

FP (*False Positive*): Jumlah sampel negatif yang salah diprediksi sebagai positif.

d) *Recall*

Recall merupakan suatu nilai yang menampilkan rasio tingkat keberhasilan suatu informasi data yang diprediksi positif dan benar positif dari data kelas positif secara keseluruhan [64]. *Recall* memberikan gambaran seberapa baik model dapat menangkap semua contoh positif.

$$Recall = \frac{(TP)}{(TP + FN)}$$

Rumus 2.3 Rumus Mencari Nilai *Recall* [63]

Keterangan Rumus:

TP (True Positive): Jumlah sampel positif yang diprediksi dengan benar

TN (True Negative): Jumlah sampel negatif yang diprediksi dengan benar.

FN (False Negative): Jumlah sampel positif yang salah diprediksi sebagai negatif.

e) F1-Score

F1-Score merupakan *harmonic mean* antara nilai presisi dan recall. Nilai ini membantu dalam mengevaluasi kinerja model klasifikasi, terutuk pada kelas target yang tidak seimbang. Skor F1 memiliki rentang nilai dari 0 hingga 1, di mana 1 menunjukkan klasifikasi yang sempurna dan 0 menunjukkan klasifikasi yang sangat buruk [63].

$$F1 - Score = 2 \times \frac{(Precision \times Recall)}{(Precision + Recall)}$$

Rumus 2.4 Rumus Mencari Nilai F1-Score [63]

2.3.9 Sistem Rekomendasi

Sistem rekomendasi adalah suatu sistem yang menyajikan *item* maupun saran yang bermanfaat bagi pengguna. Sederhananya, sistem ini membantu pengguna menemukan hal-hal yang mungkin mereka sukai atau butuhkan, berdasarkan data dan pola perilaku mereka [65], [66]. Terdapat beberapa jenis sistem rekomendasi yang umum digunakan diantaranya [66]:

a) *Collaborative Filtering*

Collaborative filtering merupakan metode rekomendasi yang membentuk basis data berdasarkan preferensi pengguna dengan tujuan memberikan rekomendasi yang sesuai dengan minat dan ketertarikan pengguna berdasarkan prefrensinya.

b) *Content-Based Filtering*

Content-based filtering merupakan metode rekomendasi yang merekomendasikan berdasarkan kemiripan atribut yang disukai

oleh pengguna, jadi atribut tersebut menjadi penentu untuk menghasilkan *item* atau objek yang inginkan pengguna.

c) *Hybrid System*

Hybrid system merupakan metode yang digunakan untuk membentuk rekomendasi dengan menggunakan lebih dari satu metode, hal ini dilakukan untuk memberikan rekomendasi yang lebih akurat dan sesuai dengan yang diinginkan pengguna.

2.4 Alat Penelitian

2.4.1 Python

Python adalah bahasa pemrograman tingkat tinggi yang sering digunakan banyak orang, yang dibuat oleh *Guido van Rossum* pada tahun 1980-an. Penggunaan kode Python dapat digunakan untuk menulis semua hal, dari skrip kecil hingga program besar. Python adalah bahasa pemrograman yang bersifat *open source* [67].

Python mendukung pemrograman berorientasi objek dan pemrograman fungsional. Python merupakan pemrograman yang dinamis dengan memiliki sintaks yang jelas dan ekspresif dengan manajemen memori yang otomatis. Python dapat digunakan untuk segala jenis pengembangan perangkat lunak *multiplatform*. Python telah didistribusikan dengan beberapa lisensi dan versi yang berbeda. Python dapat digunakan dalam kepentingan komersial karena bersifat *open source* [68].

2.4.2 Visual Studio Code

Visual Studio Code (VS Code) merupakan kode editor sederhana yang didukung dengan operasi pengembangan seperti *debugging*, *running task*, dan *version control*. Visual Studio Code ini merupakan *tools* yang digunakan oleh *developer* untuk membantu siklus pembuatan kode, pembangunan sistem, dan *debug* yang cepat dan kompleks yang disebut *Integrated Development Environment* (IDE). VS Code ini dapat diakses melalui banyak sistem operasi seperti Linux, macOS, Windows, serta VS Code ini gratis untuk digunakan. VS Code dapat menerapkan banyak *environment* menyesuaikan dengan

kebutuhan, namun perlu dilakukan instalasi terlebih dahulu, serta ketersediaan banyak bahasa pemrograman untuk membentuk sistem [69].

2.4.3 Firebase

Firebase merupakan teknologi yang milik Google yang dikenal sebagai sistem *Backend as a Service* (BaaS) yang dapat mendukung penerapan data *real time* [70]. Firebase pertama kali diperkenalkan pada tahun 2011 oleh Andrew Lee dan James Tamplin, namun pada tahun 2014 diakuisisi oleh Google. Tidak hanya menyediakan layanan *database*, Firebase juga memiliki layanan lainnya yang diperuntukkan dalam mempermudah *developer*, seperti layanan Firebase Analytics, Firebase Cloud, Firebase Authentication, dan layanan lainnya. Firebase dapat diimplementasikan dalam aplikasi *mobile* maupun *website* sesuai dengan kebutuhan *developer* itu sendiri [71].

2.4.4 Android Studio

Android Studio merupakan *Integrated Development Environment* (IDE) resmi dalam melakukan pengembangan Android yang diperkenalkan oleh Google dan didasarkan pada IntelliJ IDEA dari JetBrains. IDE ini memberikan penawaran berbagai fitur untuk membantu melakukan pengembangan aplikasi android dengan kualitas tinggi. Android Studio mendukung berbagai macam bahasa pemrograman seperti Java, Kotlin, serta bahasa pemrograman lain dengan menggunakan plugin yang tersedia. Fitur yang disediakan Android lainnya adalah membantu pengembang mengidentifikasi masalah yang terjadi pada aplikasi melalui *log*, serta memudahkan pengembangan dalam membangun, menguji dan menerapkan aplikasi secara langsung [72].

U N I V E R S I T A S
M U L T I M E D I A
N U S A N T A R A