

BAB II

DASAR TEORI

2.1 Kajian Pustaka

Penelitian ini menggunakan referensi dari penelitian sebelumnya yang relevan sebagai pedoman dalam penyusunan, bertujuan untuk memahami dan menemukan relevansi untuk menghindari kesamaan penelitian-penelitian sebelumnya dengan penelitian yang akan dilakukan.

Banyaknya manfaat dari pengembangan Deep Learning terutama bidang *computer vision* terutama untuk menyelesaikan permasalahan tentang kemampuan penglihatan. Hal tersebut terus berkembang dari riset-riset atau penelitian yang telah dilakukan sebelumnya di bidang *computer vision* oleh penelitian sebelumnya, baik mengenai Deep Learning menggunakan Convolutional Neural Network dan Deep Learning menggunakan Transfer learning sebagai metode *alternatif* dengan *dataset* rendah dan perolehan akurasi tinggi yang diterapkan oleh peneliti dalam berbagai macam kasus objek. Sebagaimana Penelitian sebelumnya dengan melakukan perbandingan-perbandingan dengan *pre-trained* model seperti yang dilakukan oleh Rifqi Rizqullah dan Muhammad Ichwan pada bidang medis untuk deteksi *Pneumonia* pada manusia [18], melakukan perbandingan menggunakan ResNet 50 dan ResNet 152 , dengan dataset citra *pneumonia* 1673 dan citra normal 1795, total citra 3468. Perolehan *accuracy* tertinggi pada *epochs* 150, ResNet 50 mendapatkan 88,7% dan ResNet 152 mendapatkan 89,3%.

Perbandingan metode Transfer learning VGG16, InceptionV3, ResNet50 dan Xception kemudian dilakukan *fine-tuning* untuk klasifikasi ras kucing dilakukan oleh Muhammad Afif dkk [13], penelitian ini membandingkan metode mana yang memiliki *accuracy* terbaik dari 4 *pre-trained* model dengan dataset yang diperoleh dari *Oxford-III* dengan 2393 citra dengan 12 class dengan data *train* 1068 citra, data *validation* 528 citra, dan data *testing* 797, hasilnya dari *fine tuning* tersebut

mendapatkan *accuracy* tiap model 87.73 %, 94.19 %, 97.94% 99.63%, dan testing tiap model 60.85%, 84.94%, 71.39%, dan 93.75%.

Perbandingan kualitas kayu kelapa menggunakan CNN Transfer Learning oleh Nurul Fathanah Mustamin, dkk [20], pengambilan dataset diambil melalui kamera canon 550D dengan jarak 1 meter di kalimatan selatan. Dengan 3 kategori yaitu kerapatan rendah, sedang, dan tinggi dengan total 174 citra, tidak diberi tahu berapa *split* data untuk *training* dan *validation* tetapi citra di *preprocessing* dengan ukuran 255 X 255 dengan bantuan software MATLAB kemudian didapatkan hasil GoogleNet *accuracy* 84.89%, ResNet 101 *accuracy* 78,41%, ResNet 50 *accuracy* 77,18%, ResNet 18 *accuracy* 72,94% dan AlexNet 65,84%.

Perbandingan penelitian untuk klasifikasi cacat kemasan kaleng dengan CNN oleh Rindi Kusumawardani dan Putu Dana Karningsih [21], *dataset* diambil sebanyak 600 citra diambil secara langsung dengan kondisi berbeda, ada 3 kategori *no defect*, *major defect*, dan *minor defect*, lalu *displit train* 80% dan *test* 20%, dengan size 224 X 224 yang mana pelatihan menggunakan arsitektur dari ShuffleNet, GoogLeNet, ResNet18, ResNet50, dan ResNet101, seluruh pengujian mendapatkan nilai *accuracy train* mencapai diatas 90%. Hasil terbaik didapatkan oleh ResNet 50 dan ResNet 101 dengan akurasi pengujian 95,56%.

Dilanjut penelitian dengan menggunakan CNN dan Transfer Learning untuk melakukan klasifikasi, seperti penelitian deteksi penyakit Covid-19 melalui *X-Ray* dengan Deep Residual Network oleh Yuli Sun Hariyani dkk [17], dataset diambil dari 2 sumber dari untuk Covid-19 diambil melalui GitHub seseorang dan *dataset* citra normal diambil melalui Kaggle, data pelatihan sebanyak 48 citra, validasi sebanyak 15 citra, dan data uji untuk Covid-19 sebanyak 42 citra dan untuk normal sebanyak 234 citra. Pelatihan dilakukan dengan Transfer Learning model ResNet 50 lalu diubah menjadi 3 model A dengan *block* 5C dari ResNet, model B terdiri dari *block* 5B dan 5C serta model C terdiri dari *block* 5A, 5B, dan 5C serta evaluasi menggunakan *confusion matrix*, dengan *epochs* 30 akurasi didapatkan dari ketiga model A, B, dan C yaitu 97%, 97%, dan 99%.

Selanjutnya penelitian yang dilakukan oleh Reynold Erwandi dan Suyanto melakukan klasifikasi kanker payudara dengan ResNet [14], dataset berasal dari 82

pasien *Lab Pathological Anatomy and Cytopathology (P&D)* di Brazil, diambil dari BreakHis dengan 7909 citra, , dilakukan *split* 64% pelatihan , 16% validasi dan 20% uji citra, kemudian arsitektur ResNet dimodifikasi dengan melakukan *freeze* kecuali *layer* terakhir, tidak disebutkan tipe ResNet apa yang digunakan dengan *learning rate* $10e-2$ serta *optimizer* SGD, evaluasi 2 kelas dan 8 kelas, dimana tingkat pembesaran dilakukan 4 tahap pelatihan yaitu 40x, 100x, 200x, dan 400x dengan *epochs* 450 mendapatkan hasil *accuracy* dengan model yang diusulkan pada 2 kelas didapatkan rata- rata 99,30% dan untuk 8 kelas mendapatkan rata- rata 94,60%.

Penelitian selanjutnya menggunakan CNN dan Transfer Learning MobileNet V2 yang dilakukan oleh M. Raihan Rafiiful Allaam dan Agung Toto Wibowo [22], melakukan klasifikasi 5 genus tanaman anggrek pengambilan *dataset* dilakukan secara langsung melalui lapangan dari rumah bunga Rizal lembang dengan 6000 citra (1200/genus) dan *scrapping website* sebanyak 185 citra, uji penelitian dilakukan dengan dengan evaluasi K-Fold =5 dengan pelatihan keategori 2 gambar *background bernoise* dan *background non-noise*, rata-rata hasil terbaik didapatkan bahwa dengan MobilenetV2, *Dropout* ,adam *learning rate* 0,0001 dengan kategori dengan *background noise* mendapatkan *accuracy* 94,78% dan kategori tanpa *background noise* MobilenetV2 *Dropout* adam *learning rate* 0,0001 mendapatkan *accuracy score* 95,62%.

Dilanjutkan Penelitian klasifikasi spesies bunga menggunakan CNN dan InceptionV3 oleh Santosh Giri [23], dataset didapatkan dari Oxford 17 bunga, dengan 17 kelas dan masing masing kelas terdapat 80 citra, kemudian dilakukan *resize* citra ukuran 299X299X3 , tahap pelatihan menggunakan 42 *layers* kedalaman dari *pre-trained* InceptionV3 hasil yang diperoleh adalah akurasi pelatihan 99% dan akurasi pengujian 86,4%.

Kemudian penelitian klasifikasi hama pada citra daun kopi dengan CNN Alexnet dilakukan oleh Dicki Irfansyah, Metty Mustikasari, dan Amat Suroso [24], Penelitian mengembangkan program MATLAB dengan arsitektur Alexnet sehingga program juga dapatk digunakan secara *offline*, dengan *dataset* diunduh melalui *website* RoCoLe dengan 3 kelas kategor kesehatan *health*, *rust* ,dan *red spider mite* yang masing masing berjumlah 100 per kategori, dibagi menjadi *train*

80% dan *test* 20%, hasil pelatihan mendapatkan *accuracy training* mencapai 80,56% dan *testing* 81,6%.

Penerapan Convolutional Neural Network dikembangkan dari segi arsitektur dan jumlah lapisan pada jaringan untuk pelatihan. Contohnya penelitian Arsitektur AlexNet [8]. Dataset ImageNet LSVRC-2010 kedalam 1000 kelas. Arsitektur yang dibuat oleh Alex Krizhevsky berhasil memperoleh *test error* terendah sebesar 17 %. Hasil tersebut membuktikan pengembangan arsitektur CNN dinilai bagus dengan jumlah data pelatihan yang sangat banyak

Adapun pengembangan Arsitektur bernama Residual Network pada kompetisi ILVRSC 2015 oleh he, dkk [19]. Penelitian menggunakan *dataset* ImageNet 2015 kedalam 1000 kelas memenangkan juara 1. Dibuat dari peningkatan VGG dengan menghindari kerumitan kedalaman dengan teknik Residual block serta memecahkan masalah *gradient* dengan tingkat *test error* terendah sebesar 3,57 %. Hasil tersebut jauh lebih baik dari AlexNet dalam pengembangan arsitektur Convolutional Neural Network (CNN) dari waktu ke waktu.

Lalu penelitian hanya menggunakan Convolutional Neural Network (CNN), Adapun penelitiannya dilakukan oleh Gusti Alfahmi Anwar dan Desti Riminarsih [25]. Penelitian ini melakukan klasifikasi 4 *genus panthera* yaitu harimau, jaguar, macan tutul, dan singa dengan total *dataset* 5600 citra diambil dari imagenet.org, lalu dibagi menjadi data pelatihan sebanyak 3840 citra, data validasi sebanyak 960 citra, dan data uji sebanyak 800 citra. Dilakuakn *resize* 128 X 128 dan normalisasi 0-1, dengan arsitektur CNN dengan 100 *epochs*, *Batch size* 64, waktu *training* 46 menit, mendapatkan nilai akurasi pelatihan 92,31% dan akurasi validasi 81,88%.

Melakukan klasifikasi batik riau menggunakan CNN oleh Hendry fonda, Yuda Irawan, Anita Febriani [26], penelitian ini menggunakan 168 gambar, tidak diberitahu sumber gambarnya tetapi 68 citra merupakan batik riau dan 100 citra bukan batik riau, selanjutnya dilakukan pelatihan sebanyak 100 *epochs* dengan hasil akurasi 65%, pengaruh tersebut didasari oleh warna dan macam-macam motif. Penelitian sebelumnya dipaparkan pada tabel 2.1 dibawah sebagai berikut:

Tabel 2.1 *Literature review* dari penelitian sebelumnya

No.	Penulis	Objek	Metode Penyelesaian Masalah	Metode Pengumpulan Data	Hasil	Kelemahan
1.	Rifqi rizqullah eka prasetyo, Muhammad ichwan [18].	Perbandingan metode ResNet 50 dan 152 untuk deteksi pneumonia pada manusia.	ResNet 50 dan ResNet 152.	Situs Kaggle Paul Money 2017.	Hasil akurasi tertinggi pada ResNet 50 dan 152, berada pada 88,7% dan 89,3%.	Penambahan <i>layers</i> sebelum output dapat menambah hasil optimal.
2.	Muhammad Afif Amanullah Fawwaz, Kurniawan Nur Ramadhani, S.T., M.T., Febryanti Sthevanie, S.T., M.T. [13]	Klasifikasi ras kucing	Perbandingan Transfer Learning + <i>fine tuning</i>	Oxford-IIIT	Menggunakan Bantuan Transfer Learning + <i>fine tuning</i> hasil testing VGG16 akurasi 60.85%, InceptionV3 akurasi 84.94%, ResNet 50 akurasi 71.39%, dan Xception akurasi 93.75%.	<i>Dataset</i> kurang pada setiap jenis kucing
3.	Nurul Fathanah Mustamin, Yuslena Sari, Husnul Khatimi [20]	Klasifikasi kualitas 3 jenis kayu kelapa	Perbandingan CNN <i>pre-trained</i>	Diambil oleh Hendriyane denga bantuan kamera Canon 550D	Iterasi pelatihan 10 hasil GoogleNet <i>accuracy</i> 84.89%, ResNet 101 <i>accuracy</i> 78,41%, ResNet 50 <i>accuracy</i> 77,18%, ResNet 18 <i>accuracy</i> 72,94% dan AlexNet <i>accuracy</i> 65,84%.	<i>Dataset</i> hanya 174 citra terlalu sedikit.

Tabel 2.2 *Literature review* selanjutnya

No.	Penulis	Objek	Metode Penyelesaian Masalah	Metode Pengumpulan Data	Hasil	Kelemahan
4.	Rindi Kusumawardani dan Putu Dana Karningsih [21]	Klasifikasi cacat kaleng 3 kategori	Metode Transfer Learning	Diambil secara langsung dengan kondisi berbeda sebanyak 600 gambar,	Hasil terbaik didapatkan oleh ResNet 50 dan ResNet 101 dengan akurasi pengujian 95,56%.	Fine-tuning dapat diterapkan untuk mendapatkan acc lebih.
5.	Yuli Sun Hariyani, Sugondo Hadiyoso, dan Thomhert Suprpto Siadari [17]	Klasifikasi Covid-19 pada <i>X-Ray</i>	Residual Network	Github dan Kaggle	Model A, B, dan C pada iterasi 30 mendapatkan akurasi 97%, 97%, dan 99%	Data pelatihan dan validasi sedikit
6.	Reynold Erwandi, Suyanto [14]	Klasifikasi kanker payudara dengan kategori 2 kelas dan 8 kelas	ResNet yang diusulkan (modifikasi)	Diambil melalui 82 pasien di <i>Lab Pathological Anatomy and Cytopathology (P&D)</i> di Brazil total 7909 citra	<i>epochs</i> 450 mendapatkan hasil <i>accuracy</i> dengan model yang diusulkan pada 2 kelas didapatkan rata-rata 99,30% dan untuk 8 kelas mendapatkan rata-rata 94,60%.	Modifikasi yang perlu ditingkatkan untuk klasifikasi pelatihan 8 kelas

Tabel 2.3 *Literature review* selanjutnya

No.	Penulis	Objek	Metode Penyelesaian Masalah	Metode Pengumpulan Data	Hasil	Kelemahan
7.	M.Raihan Rafiiful Allaam, Agung Toto Wibowo [22].	Klasifikasi 5 genus tanaman Anggrek.	CNN dan MobileNetV2.	Dataset lapangan rumah rizal lembang dan <i>scrapping web</i> total jumlah 6000 citra.	Hasil validasi terbaik pada 95,62%.	Pelatihan menggunakan arsitektur Transfer Learning lain dapat menambah optimalisasi nilai akurasi.
8.	Santosh Giri [23]	Klasifikasi 17 spesies bunga	CNN dan Inception v3	Oxford 17 <i>flower dataset</i>	Mendapatkan akurasi pelatihan 99% dan akurasi testing 86,4%	<i>Dataset</i> per kategori hanya 80 (bergantung pada dataset oxford)
9.	Dicki Irfansyah, Metty Mustikasari, Amat Suroso [24]	Klasifikasi hama pada daun tanaman kopi	CNN dan AlexNet	RoCoLe: <i>A robusta coffee leaf images dataset</i>	Akurasi saat <i>training</i> 69.44-80.56%, dan uji 40 data testing 81,6%	Perlu adanya perubahan arsitektur CNN agar akurasi dapat di naikkan.
10.	Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton [8]	Klasifikasi 1000 kategori berbeda	CNN	ImageNet LSVRC-2010 kontes	<i>top-5 testting</i> dengan <i>error</i> 15.3%,	Komputasi berat di sarankan memiliki GPU untuk pelatihan.

Tabel 2.4 *Literature review* selanjutnya

No.	Penulis	Objek	Metode Penyelesaian Masalah	Metode Pengumpulan Data	Hasil	Kelemahan
11.	Kaiming He, Xiangyu Zhang, Shaoqing Ren, dan Jian Sun [19]	Klasifikasi 1000 kategori berbeda	CNN	ImageNet LSVRC-2015 kontes	<i>top-5 testing</i> dengan <i>error</i> 3,57%,	Komputasi berat di sarankan memiliki GPU untuk pelatihan.
12.	Gusti Alfahmi Anwar dan Desti Rimirasih [25]	Klasifikasi 4 genus panthera	CNN	Imagenet.org	nilai akurasi pelatihan 92,31% dan akurasi validasi 81,88%	Tidak menggunakan Transfer Learning.
13.	Hendry fonda, Yuda Irawan, Anita Febriani [26]	Klasifikasi Batik Riau	CNN	Diambil <i>dataset</i> 168 gambar batik Riau	Hasil akurasi batik Riau 65% dan bukan batik Riau 65%.	<i>Dataset</i> kurang banyak, tidak augmentasi data.

Referensi rujukan utama:

- a. Klasifikasi Kanker Payudara Menggunakan Residual Neural Network [14].
- b. Perbandingan Metode Deep Residual Network 50 dan Deep Residual Network 152 untuk Deteksi Penyakit *Pneumonia* pada manusia [18].

2.2 Dasar Teori

2.2.1 Taksonomi Hewan

Karakteristik suatu hewan memiliki ciri ciri yang berbeda antar lainnya, didasarkan pada beberapa pengamatan, tingkah laku, dan cara kerja tubuh hewan tersebut. Hal tersebut dilakukan cara untuk melakukan pengelompokan yang disebut taksonomi hewan yang didasarkan oleh persamaan ciri ataupun perbedaan cirinya, morfologi, anatomi, manfaat, habitat, dan cara hidup hewan. Tujuan utama taksonomi adalah mempermudah mengenali, membandingkan, mengidentifikasi, dan mempelajari makhluk hidup [27].

2.2.2 Deep Learning

Deep Learning adalah bagian dari Machine learning yang mencoba meniru proses mempelajari tingkah laku manusia. Metode pembelajaran mesin ini adalah pembelajaran *supervised*, pembelajaran *semi-supervised*. Dalam sistem pembelajaran mesin, model mempelajari perhitungan dan menganalisis data untuk membuat keputusan, tahap *feature extraction* dan *classification* berada pada satu tahap [13].

2.2.3 Data Preprocessing

Data Preprocessing adalah fase penting sebelum melatih atau menguji model. Tahap ini memiliki proses untuk pengumpulan 2 grup yaitu *data train* dan *data validation*, serta pengubahan ukuran gambar. *Resizing* digunakan untuk mengubah ukuran gambar dengan tujuan menyesuaikan untuk pelatihan atau pengujian. *Preprocessing* sendiri bertujuan untuk

mengoptimalkan akurasi meningkatkan kualitas gambar, sehingga menyederhanakan, dan mempercepat kinerja sistem komputasi [28].

2.2.4 Data Augmentation

Augmentasi data adalah teknik manipulasi yang tidak menghilangkan ekstraksi informasi tentang data dalam sebuah dataset menjadi berbeda, untuk mereplikasi kumpulan data dengan cara efektif sehingga model mencapai akurasi optimal serta menghindari *overfitting* dan *underfitting* [9], [10].

2.2.5 Transfer Learning

Transfer Learning merupakan pembelajaran mendalam dari CNN yang dilatih sebelumnya untuk dapat dimodelkan kembali dengan masalah baru [11]. Transfer Learning adalah teknologi yang sangat penting dan efektif dalam *training*. Tujuan dari Transfer Learning adalah tidak membuang informasi yang berguna dari data sebelumnya dan menerapkan pengetahuan yang dipelajari sebelumnya [12]. Dalam Transfer Learning, mesin menggunakan pengetahuan dari masalah sebelumnya yang memiliki hubungan yang sama agar dapat meningkatkan generalisasi untuk hal baru daripada harus dilakukan *learning* dari awal [13].

2.2.6 Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNNs) adalah jenis jaringan saraf yang tiruan. CNN adalah jaringan saraf yang umum digunakan sebagai pemrosesan data gambar. Convolutional Neural Network (*CNN*) pertama kali diusulkan di tahun 1960-an[6] adalah metode Deep Learning dalam klasifikasi suatu objek gambar dengan nilai akurasi tinggi [4]. CNN bekerja menggunakan kernel untuk mengekstrak fitur. Keuntungan CNN pada input yang menggunakan operasi konvolusi adalah membantu mengurangi jumlah parameter dengan tujuan mengurangi beban komputasi. Lalu kinerja CNN sangat baik. banyaknya lapisan yang dapat digunakan untuk menyaring setiap proses *node* bertujuannya untuk membuat arsitektur pembelajaran yang lebih baik, pada setiap iterasi dalam pelatihan [13].

Beberapa *layer* yang banyak digunakan terdapat 4 tahapan yaitu *Convolutional layer*, *Pooling layer*, *BatchNormalization layer*, *Dropout layer*, dan *Flatten* [1].

2.2.7 Tensorflow

Tensorflow adalah kerangka kerja Deep Learning gratis dan *open source* serta perpustakaan ilmu data yang dikembangkan oleh ilmuwan di tim Google. Tensorflow dapat digunakan di berbagai bidang. *Framework API* tensorflow yang merupakan alat yang dapat digunakan untuk menyederhanakan proses pembuatan, pelatihan, dan penerapan model untuk pemecahan masalah menggunakan Deep Learning [29].

2.2.8 Open-CV

Open-CV merupakan sebuah *library* pemrosesan gambar yang berlisensi BSD dan gratis, biasa digunakan untuk *computer vision* sebagai *tools* akademis maupu komersial [30].

2.2.9 Arsitektur Residual Network

Residual Network adalah salah satu struktur CNN yang mengurangi beban komputasi dibandingkan VGGNet. Diperkenalkan oleh He, didasarkan oleh masalah saat jaringan Deep Learning semakin dalam dan *error* akan naik lebih tinggi, untuk mengatasinya menggunakan metode Skip connection atau Identity mapping yaitu terobosan rute alternatif [17]. Kelebihan arsitektur ResNet membuat nilai akurasi tidak menurun meskipun dengan desain maju dan menghindari permasalahan *gradient* [14]. Untuk ResNet152 memiliki *block* dengan *convolution layer* 3x3, *Batch normalization* 1x1, dan *ReLU activation* 1x1, tahapan arsitektur awal lapisan conv1 memiliki *convolution layer* 7x7 dengan keluaran dimensi 112x112, selanjutnya lapisan conv2 memiliki *maxpooling* 3x3 dan proses *block* yang dilakukan 3x dengan output dimensi 56x56. kemudian lapisan conv3 memiliki proses *block* yang dilakukan 8x dengan keluaran dimensi 28x28. kemudian lapisan conv4

memiliki *block* yang dilakukan 36x dengan keluaran dimensi 14x14. kemudian lapisan conv5 memiliki *block* yang dilakukan 3x dengan keluaran dimensi 7x7. kemudian *average pooling*, *fully connected layer*, dan lapisan terakhir sesuai kelas dengan *softmax activation* dengan keluaran dimensi 1x1 pada arsitektur [18].