

BAB II

TINJAUAN PUSTAKA

2.1 Kajian Pustaka

Studi tentang CNN sedang dilakukan di berbagai bidang, termasuk kesehatan, pertanian, dan industri. Dalam penelitian ini, peneliti menggunakan 10 jurnal sebagai referensi untuk studi literatur. Penelitian yang dilakukan oleh Muhammad Mateen, dkk dalam penelitiannya dengan judul “*Fundus Image Classification Using VGG-19 Architecture with PCA and SVD*” [16]. Tujuan dari penelitian ini untuk mengklasifikasikan *diabetic retinopathy* (DR) dari citra fundus. Sumber data atau dataset yang digunakan diperoleh dari Kaggle dengan jumlah 35126 citra. Dalam penelitian ini CNN dengan model VGG-19 dan menggunakan analisis komponen utama (PCA) dan dekomposisi nilai tunggal (SVD) untuk fitur seleksinya dan FC menunjukkan akurasi klasifikasi 92,21%, 98,34%, 97,96%, dan 98,13% untuk FC7-PCA, FC7-SVD, FC8-PCA, dan FC8-SVD.

Pada penelitian Vimal K. Shrivastava dkk yaitu “*Rice Plant Disease Classification Using Transfer Learning Of Deep Convolution Neural Network.*” [17], menggunakan SVM untuk klasifikasi dan AlexNet untuk ekstraksi fitur. Dataset diambil menggunakan kamera *Canon Powershot SX530HS* dengan pengambilan gambar pada pukul 6.30-9.30 pagi dan 4.00-5.30 sore. Menggunakan 4 klasifikasi penyakit daun padi blast, blight, sheat blight, dan daun sehat. Dengan jumlah gambar sebanyak 619 gambar. Hasil dari akurasi klasifikasi tertinggi dari penelitian ini 80 persen hingga 20 persen partisi pelatihan-pengujian dengan jumlah 91,3 persen.

Studi dari Ardi Hidayat, dkk yaitu “*Detection Of Disease On Corn Plants Using Convolutional Neural Network Methods*” [18], menggunakan kumpulan data 3854 gambar dan tiga penyakit jagung yang berbeda: karat biasa, bercak

daun abu-abu, dan hawar daun utara. Studi ini menggunakan *deep network learning* metode CNN yang mencapai akurasi 99% untuk mendiagnosis penyakit jagung menggunakan.

Pada penelitian Endang Anggiratih dkk yang berjudul “Klasifikasi Penyakit Tanaman Padi Menggunakan Model *Deep Learning* Efficientnet B3 Dengan Transfer Learning” [19]. Bahan penelitian atau dataset yang digunakan dalam studi ini terdiri dari 857 citra dengan ukuran total 2848 x 4288. Gambar diambil dengan kamera digital NIKON D90. Ada dua klasifikasi, kelas bercak coklat dan kelas bakteri daun. Karena resolusi gambar yang berbeda, resolusi diubah menjadi ukuran 300 x 300 untuk arsitektur atau model Efficientnet B3 dan ukuran 224 x 224 untuk arsitektur atau model Mobilenet V3. Proses pengujian pada dataset uji dilakukan secara acak. Akurasi terbaik adalah untuk arsitektur Efficientnet dengan 79,53% pada 350 epoch, sedangkan arsitektur Mobilenet V3 memiliki akurasi 54,32% pada 300 epoch.

Penelitian Mohtar Khoirudin dkk, pada penelitiannya yang berjudul “Klasifikasi Penyakit Daun Padi Menggunakan Convolutional Neural Network” [20], menggunakan menggunakan *Confussion Matrix* untuk menghitung tingkat akurasi. Layer atau lapisan yang dipakai/digunakan dalam studi ini yaitu Conv, ReLU, dan 4-layer Maxpooling. Hasil akurasi pelatihan terbaik adalah akurasi 98% pada epoch 100, dan pengujian dengan matriks konfusi sebanyak 600 data menghasilkan nilai akurasi 98% dalam mengklasifikasi penyakit daun padi.

Penelitian J. Raja dkk, dengan judul “*An Automated Early Detection of Glaucoma using Support Vector Machine Based Visual Geometry Group 19 (VGG-19) Convolutional Neural Network*” [21], menggunakan dataset sebanyak 175 dengan klasifikasi sebanyak 3 yaitu normal, *starting stage* dan *advanced stage* dari penyakit glukoma. Penelitian ini menggunakan VGG-19 untuk layernya dan SVM untuk klasifikasinya. Studi ini mencapai akurasi 92%.

Penelitian P. Naveen dan B. Diwan dengan judul “*Pre-trained VGG-16 with CNN Architecture to classify X-Rays images into Normal or Pneumonia*”

[22]. Pada penelitian ini menggunakan gambar x-ray pada dada yang menggunakan VGG-16 untuk mengklasifikasikan penyakit pneumonia atau normal. Dataset yang digunakan sebanyak 5216 gambar untuk *training set*, 624 gambar untuk *evaluation set*, dan 16 gambar untuk *validasi set*. Untuk menyaring data sampel agar seimbang digunakan data augmentasi. Penelitian dengan model ini memperoleh hasil 95.67 persen uji presisi dan 12.64 uji *loss* dan akurasi sebesar 50 persen.

Penelitian Regina Agustina, dkk dengan judul “Klasifikasi Kanker Kulit menggunakan Metode Convolutional Neural Network dengan Arsitektur VGG-16” [23]. Gambar dari kanker kulit diproses dengan arsitektur CNN model VGG-16. Kumpulan data dalam format gambar dari jaringan kanker berisi hingga 4000 jumlah citra atau gambar digunakan pada studi ini. Prosesnya dimulai dengan input citra, preprocessing, pelatihan model, dan pengujian sistem. Hasil terbaik diperoleh pada pengujian menggunakan pengoptimal hiperparameter SGD, rata-rata learning 0,001, epoch 50, batch size 32, tanpa preprocessing CLAHE dan Gaussian filtering. Akurasi, loss, recall, presisi yang diperoleh 99,70%, 0,0055, 0,9975, 0,9975, skor f1 0,9950.

Penelitian Yoze Rizki, dkk dengan judul “Klasifikasi Pola Kain Tenun Melayu Menggunakan Faster R-CNN” [24]. Studi ini menggunakan Faster R-CNN pada model VGG dengan K-Fold Cross Validation untuk dengan mengukur persentase presisi, akurasi dan *recall* yang diperoleh. Jumlah bahan penelitian atau data sampel yang digunakan adalah 100 gambar acak untuk setiap 5 konvolusi K-fold *cross-validation*. Bahan penelitian atau dataset tersebut dibagi 4:1 untuk data latih dan *testing data*. Selesai dilakukan persiapan data, preprocessing, dan penerapan arsitektur, dilakukan pengujian yang menghasilkan nilai rata-rata training loss sebesar 1,915 dari tahap awal sampai tahap terakhir dari data training berupa gambar tekstil Murray sawah. Studi ini memperoleh akurasi 82,14%, presisi 91,38%, dan tingkat recall 91,36%.

Penelitian selanjutnya dilakukan oleh Yosefina Finsensia Riti, dkk dengan judul “Klasifikasi Covid-19 Pada Citra CT Scans Paru-Paru Menggunakan Metode Convolution Neural Network” [25]. Dalam penelitian ini menggunakan metode CNN untuk mengklasifikasikan CT scan paru-paru dan matriks konfusi untuk mengukur keberhasilan metode CNN. Data sampel sebanyak 3216 citra yang menggunakan metode CNN untuk mengklasifikasikan kelas paru Covid-19 dan kelas paru normal. Akurasi didapatkan rata-rata 100% pada epoch yang diberikan.

Penelitian lainya dilakukan oleh Daniel Fernando S. S., dkk dengan judul “*Towards automated eye cancer classification via VGG and ResNet networks using transfer learning*” [26]. Penelitian ini menggunakan CNN untuk mendeteksi uveal melanoma (UM). Menurut penelitian ini, hasil yang didapatkan mengungguli berbagai algoritma komputasi canggih yang dipelajari untuk mendeteksi UM, khususnya peningkatan sensitivitas, presisi, dan akurasi, masing-masing mencapai 99%, 98% dan 99%. Selain itu, penelitian ini menggunakan dataset sebanyak 2048 gambar dan dua algoritma diterapkan untuk mengurangi bias dataset: algoritma augmentasi data menggunakan filter Gabor, dan algoritma untuk menghilangkan titik cahaya menggunakan pendekatan Navier–Stokes.

Tabel 2. 1 Studi Sebelumnya

No	Judul	Tahun	Metode	Hasil	Penelitian yang diusulkan
1.	<i>Fundus Image Classification Using VGG-19 Architecture with PCA and SVD</i> Muhammad Mateen, Junhao Wen, Nasrullah, Sun Song, dan Zhouping Huang	2018	CNN dan VGG-19	Penelitian ini menggunakan PCA dan SVD untuk fitur seleksi dengan <i>fully connected (FC) layers</i> menunjukkan akurasi klasifikasi masing-masing pada FC7-PCA, FC7-SVD, FC8-PCA, dan FC8-SVD sebesar 92,21%, 98,34%, 97,96%, dan 98,13%.	Penelitian ini menyarankan untuk menggunakan VGGnet dalam mengidentifikasi citra karena memiliki akurasi yang lebih tinggi dibanding dengan AlexNet. Dan Penggunaan VGG-19 sendiri digunakan karena hasil yang didapatkan lebih reliable.
2.	<i>Rice Plant Disease Classification Using Transfer Learning Of Deep Convolution Neural Network.</i> Vimal K. Shrivastava, Monoj K. Pradhan, Sonajharia Minz, Mahesh P. Thakur	2019	CNN, AlexNet dan SVM	Dalam penelitian ini menggunakan SVM untuk klasifikasi dan AlexNet untuk ekstraksi fitur. Sehingga menghasilkan akurasi klasifikasi sebenar 91.37% dari partisi <i>training-testing</i> sebesar 80%-20%.	Penelitian ini mengusulkan untuk menggunakan dataset dengan jumlah yang lebih banyak.
3.	<i>Detection Of Disease On Corn Plants Using Convolutional Neural Network Methods.</i> Ardi Hidayat, Ucu Darusalam, Irmawati	2019	CNN	Studi ini menghasilkan akurasi 99,6% dan loss 0,0321 dari epoch 100.	Penelitian ini menyarankan untuk menggunakan CNN karena memiliki akurasi sebesar 99,6%.

No	Judul	Tahun	Metode	Hasil	Penelitian yang diusulkan
4.	Klasifikasi Penyakit Tanaman Padi Menggunakan Model <i>Deep Learning Efficientnet B3 Dengan Transfer Learning</i> . Endang Anggiratih, Sri Siswanti, Saly Kurnia Octaviani, Arumsari	2021	Efficientnet B3 dan Mobilenet V3	Penelitian ini menggunakan Efficientnet B3 dan memperoleh akurasi sebesar 79,53%. Sedangkan untuk Mobilenet V3 memperoleh akurasi sebesar 54,32%	Penelitian ini menyarankan untuk menggunakan model lain yang memiliki akurasi lebih tinggi. Karena model/arsitektur yang digunakan pada studi ini memiliki akurasi yang rendah menurut peneliti.
5.	Klasifikasi Penyakit Daun Padi Menggunakan <i>Convolutional Neural Network</i> . Mohtar Khoiruddin, Apri Junaidi, Wahyu Andi Saputra	2021	CNN dan <i>Confusion Matrix</i>	Penelitian ini menghasilkan hasil akurasi pelatihan terbaik dengan akurasi 98% pada 100 epoch.	Penelitian ini menyarankan untuk menggunakan <i>transfer learning</i> agar mendapatkan hasil akurasi yang lebih baik.
6.	<i>An Automated Early Detection of Glaucoma using Support Vector Machine Based Visual Geometry Group 19 (VGG 19) Convolutional Neural Network</i> . J. Raja, P. Shanmugam, R. Pitchai	2021	CNN, VGG-19, SVM	Penelitian ini menghasilkan akurasi sebesar 92%	Penelitian ini mengusulkan untuk menggunakan VGG 19 enagn menggunakan SVM karena akurasi yang dihasilkan lebih tinggi.

No	Judul	Tahun	Metode	Hasil	Penelitian yang diusulkan
7.	<i>Pre-trained VGG-16 with CNN Architecture to classify X-Rays images into Normal or Pneumonia</i> P Naveen, B Diwan	2021	CNN dan VGG-16	Peneliti memperoleh akurasi sebanyak 50%. Dataset yang digunakan sebanyak 5216 gambar untuk training set, 624 gambar untuk evaluation set, dan 16 gambar untuk validasi set.	Penelitian ini menyarankan untuk menggunakan model lain yang memiliki akurasi lebih tinggi dibanding dengan model yang digunakan peneliti.
8.	Klasifikasi Kanker Kulit menggunakan Metode Convolutional Neural Network dengan Arsitektur VGG-16. Regita Agustina, Rita Magdalena, Nor Kumalasari Caecar Pratiwi	2021	CNN dan VGG16	Akurasi yang dicapai dalam penelitian ini adalah 99,70%, loss 0,0055, presisi 0,9975, recall 0,9975, skor f1 0,9950.	Studi ini menyarankan untuk menggunakan metode CNN model VGG-16 untuk mengklasifikasikan gambar karena memiliki akurasi yang tinggi.
9.	Klasifikasi Pola Kain Tenun Melayu Menggunakan Faster R-CNN. Yoze Rizki, Reny Medikawati Taufiq, Dinia Putri, Harun Mukhtar	2021	Faster R-CNN	Penelitian ini memilih metode deteksi objek Faster R-CNN melalui validasi K-Fold Cross Validation dengan nilai k=5, didapatkan 82.14%, 91.38% dan 91.36% untuk akurasi, presisi dan recallnya.	Penelitian ini menyarankan untuk menggunakan Faster R-CNN karena unggul dari penelitian sebelumnya dengan <i>study case</i> sama.

No	Judul	Tahun	Metode	Hasil	Penelitian yang diusulkan
10.	Klasifikasi Covid-19 Pada Citra CT Scans Paru-Paru Menggunakan Metode Convolution Neural Network. Yosefina Finsensia Riti, Stephanus Surijadiarma Tandjung	2022	CNN dan <i>Confusion Matrix</i>	Penelitian ini mendapatkan akurasi keseluruhan rata-rata sebesar 100% setiap epoch.	Penelitian ini menyarankan untuk menggunakan CNN untuk mengklasifikasikan data dari dua dimensi dengan gambar yang berbeda atau gambar yang sama.
11.	<i>Towards automated eye cancer classification via VGG and ResNet networks using transfer learning.</i> Daniel Fernando Santos-Bustos, Binh Minh Nguyen, Helbert Eduardo Espitia	2022	CNN, VGG dan ResNet	Penelitian ini mendapatkan hasil sensitivitas, presisi, dan akurasi, masing-masing mencapai 99%, 98% dan 99%.	Penelitian ini menyarankan untuk menggunakan momen invarian geometris Hu dapat digunakan sebagai metode eksperimental untuk melakukan konfigurasi dan pra-inisialisasi CNN, dan menggunakan arsitektur yang berbeda dalam skema ini.

Dari tabel 2.1 dapat disimpulkan dari sebelas tinjauan pustaka tersebut menggunakan berbagai macam metode untuk mengklasifikasikan *image*. Beberapa yang menggunakan CNN dengan arsitektur yang berbeda-beda. Akurasi yang didapatkan dengan menggunakan CNN rata-rata sebesar lebih dari 90%. Sehingga pada penelitian ini penulis ingin mengklasifikasikan penyakit pada daun padi dengan menggunakan metode CNN model VGG-19. Alasan memilih VGG-19 adalah karena VGGNet lebih dalam dan arsitektur yang lebih reliabel untuk teknologi ImageNet [16]. Selain itu penulis juga ingin mengetahui berapa besar akurasi yang didapatkan dengan menggunakan metode CNN model VGG-19.

2.2 Dasar Teori

Klasifikasi penyakit daun padi menjadi topik dalam penelitian ini, sehingga dalam merealisasikan penelitian dibutuhkan dasar teori dari berbagai bidang keilmuan, berikut dasar teori yang dapat membantu penelitian ini.

2.2.1 Tanaman Padi

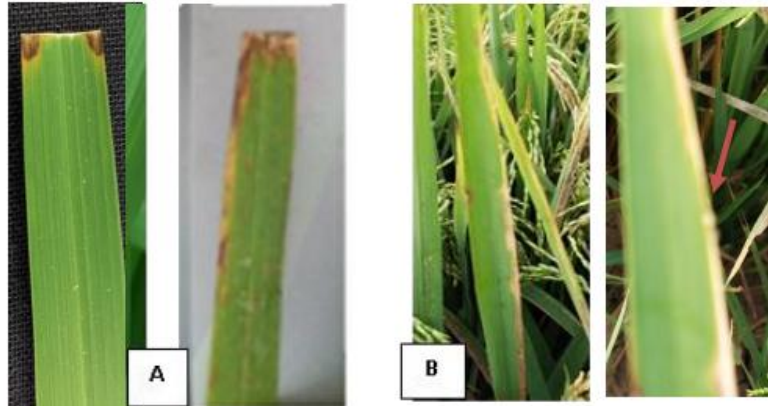
Tanaman padi yang memiliki nama ilmiah *Oryza sativa* L. termasuk tumbuhan/tanaman spermatophyta bersifat hidrofit [27]. Beras yang diproduksi dari tanaman padi merupakan bahan pangan pokok terpenting, yang merupakan bahan baku strategis dan menempati posisi yang sangat penting [28].

2.2.2 Penyakit Pada Daun Padi

Berikut beberapa *leaf rice disease* (penyakit daun padi) yang akan diklasifikasikan pada penelitian ini, yaitu:

- a. Penyakit hawar daun bakteri dengan bahasa ingris *bacterial leaf blight* ini dikarenakan adanya bakteri dengan nama ilmiah *Xanthomonas campestris* pv *Oryzae* dan bisa menyusutkan nilai panen hingga 30%. Penyakit ini sering terjadi saat penghujan basah maupun saat musim

kering. Hal ini sangat umum di sawah yang terus-menerus becek dengan N tinggi pupuk (>250 kg/ha urea) [10].



Gambar 2. 1 Penyakit hawar (blight) [29]

- b. Rice blast* atau penyakit blas, penyakit utama padi dikarenakan oleh cendawan *Pyricularia grisea*. *Rice blast* pertama kali terjadi di lahan sawah dataran tinggi, namun sejak tahun 1980-an ditemukan di lahan sawah beririgasi dan telah menyebar ke lahan sawah dataran rendah [30].



Gambar 2. 2 Penyakit blast [31]

- c. Rice tungro virus* atau tungro padi, dikarenakan oleh kontaminasi dua virus yang berbeda, virus RTBV dan virus RTSV. Ciri kontaminasi tungro yang menonjol yaitu adanya warna daun yang berubah dan pertumbuhan terhambat/kerdil. Warna pada daun tanaman yang

terkontaminasi bervariasi dari agak kuning hingga ungu, warna pada daun mulai berubah dari atas dan memanjang sampai bawah [10].



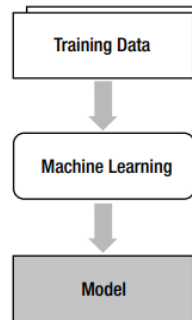
Gambar 2. 3 Penyakit tungro [31]

2.2.3 *Artificial Intelligence/AI*

Artificial intelligence/AI merupakan salah satu cabang dari ilmu komputer. Cabang ilmu ini mempelajari tentang bagaimana memberikan komputer dengan kualitas atau kecerdasan seperti manusia. Kecerdasan dihasilkan dan dipasang ke dalam mesin (komputer) untuk menyelesaikan tugas yang dapat dilakukan manusia. [32].

2.2.4 *Machine Learning*

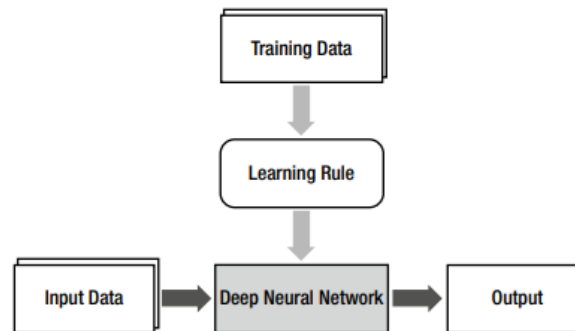
Teknik yang meningkatkan kinerja sistem dengan belajar dari pengalaman melalui metode komputasi disebut *Machine learning*. Dalam sistem komputer, pengalaman ada dalam bentuk data, dan tugas utama *machine learning* adalah mengembangkan *learning algorithms* yang membangun model dari data [32]. *Machine learning* termasuk bagian dari *AI* [33]. Phil Kim dalam bukunya menjelaskan bahwa “*learning*” ini digunakan karena prosesnya seperti sedang melatih data dalam memecahkan masalah untuk menemukan sebuah model. Sehingga, data yang digunakan dalam *modeling process* disebut dengan data “*training*”[33].



Gambar 2. 4 Proses dalam *machine learning* [33]

2.2.5 *Deep Network Learning*

Deep network learning adalah cabang pembelajaran *machine/mesin* yang terdiri atas algoritma pemodelan abstrak tingkat tinggi untuk data dengan seperangkat fungsi transfer nonlinier yang diatur dalam lapisan dan kedalaman [34]. *Deep network learning* secara sederhana dapat didefinisikan sebagai sebuah teknik *Machine Learning* yang menggunakan JSC[33]. *Transfer learning* adalah teknik yang menggunakan model yang telah dilatih sebelumnya (pretrained model) untuk mengkategorikan kumpulan data baru, menghilangkan persyaratan untuk melakukan data pelatihan dari awal [36].

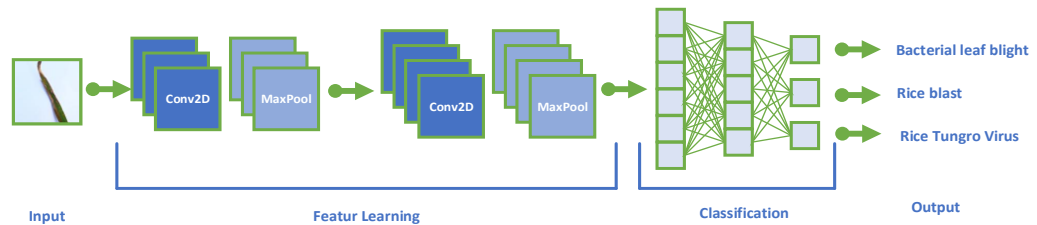


Gambar 2. 5 Konsep JSC dan hubungannya dengan *machine learning* [33]

2.2.6 CNN

CNN merupakan sebuah *deep network learning* khusus untuk pengenalan gambar [33]. CNN utamanya terdiri dari 5 layer, yaitu : *input*

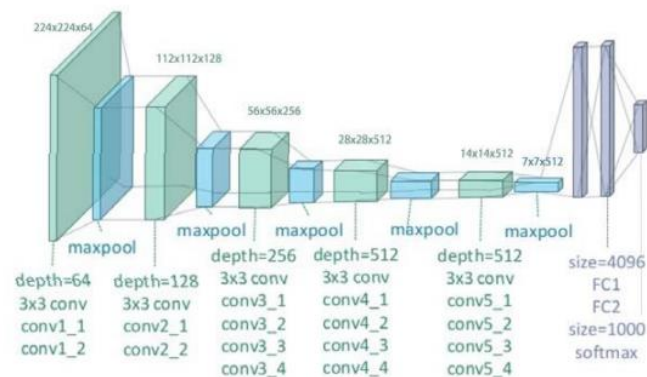
layer, layer konvolusi/Conv, layer *pool*, layer *all-layer* dan layer *softmax* [35]. Arsitektur CNN dapat dibagi menjadi 2 bagian. Pertama layer untuk fitur ekstraksi terdiri dari layer konvolusi dan *pooling layer*. Kedua *fully connected layer* dimana terjadinya proses *deep network learning* [36].



Gambar 2. 6 Arsitektur CNN [20]

2.2.7 VGG-19

VGG atau VGGNet (*visual geometry group network*) merupakan *deep network learning* dengan operasi berlapis/multilayer yang didasarkan pada model CNN dan diaplikasikan ke kumpulan data citra/gambar [16]. VGG-19 adalah model sederhana karena lapisan konvolusi atas menggunakan lapisan konvolusi 3x3 untuk meningkatkan kedalaman jaringan. VGG-19 menggunakan lapisan *max-pooling* untuk mengurangi kebisingan [16]. VGG-19 memiliki 16 layer konvolusi dan 3 lapisan *fully connected layer*. Lapisan pada arsitektur VGG memiliki struktur yang berurutan. Citra yang diinput akan melalui *convolution layer* yang berukuran 3x3 dan *max pooling layer* secara berulang. Kemudian citra akan diklasifikasikan pada *Fully Connected Layer*.



Gambar 2. 7 VGG-19 [37]

Berikut penjelasan tiap lapisannya:

- Lapisan Input: lapisan input CNN terdiri dari kumpulan data. Data input akan direpresentasikan sebagai matriks 3x3[29].
- Lapisan Konvolusi: lapisan yang menggunakan filter untuk belajar dari bagian data input yang lebih kecil untuk mendapatkan fitur dari gambar [29].

$$\begin{bmatrix} 1 & 1 & 1 & 3 \\ 4 & 6 & 4 & 8 \\ 30 & 0 & 1 & 5 \\ 0 & 2 & 2 & 4 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 7 & 5 & 9 \\ 4 & 7 & 9 \\ 32 & 2 & 5 \end{bmatrix}$$

Gambar 2. 8 Operasi Matriks pada Lapisan Konvolusi[33]

- Pooling Layer: lapisan ini digunakan untuk mengecilkan dimensi gambar, menurunkan daya pemrosesan yang diperlukan untuk lapisan berikutnya[29].
- Max pooling: piksel dengan nilai maksimum sebagai input dipilih dan ditransfer ke output saat mengurai input. Ini adalah pendekatan yang paling banyak digunakan dibandingkan dengan pooling rata-rata.
- Flatten: Pada tahap ini hasil dari pooling diratakan ke FC [29].

- f. Fully Connected Layer (Dense): ini adalah salah satu lapisan terakhir CNN, dan dapat mengenali fitur yang terhubung secara signifikan dengan kelas keluaran. Hasilnya adalah vektor satu dimensi yang dibuat dengan meratakan hasil pooling layer[29].
- g. Lapisan SoftMax: ini adalah lapisan terakhir jaringan yang membantu dalam mengklasifikasikan gambar masukan individual dari kumpulan data ke dalam beberapa kelas bergantung pada properti yang dipelajari dari jaringan [29].
- h. Lapisan Output: lapisan keluaran menampung hasil klasifikasi akhir[29].

2.2.8 *Epoch*

Epoch adalah proses pelatihan dalam jaringan saraf sampai loop kembali ke tahap awal ketika seluruh dataset melewati proses ini [38]. Atau dapat diartikan serangkaian perhitungan dari awal hingga akhir, diumpankan kembali ke proses awal sebagai satu putaran atau satu epoch [39].

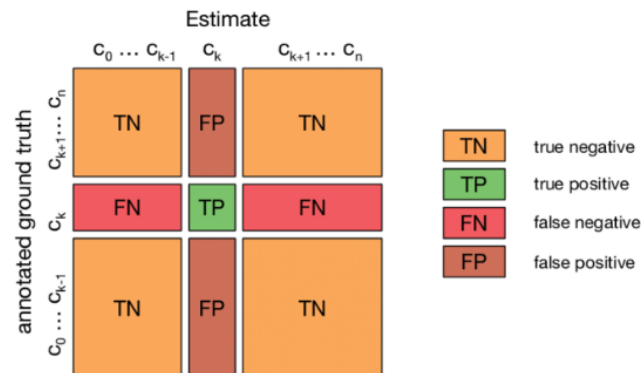
2.2.9 **Augmentasi Data**

Skema yang memungkinkan peneliti/praktisi untuk secara signifikan meningkatkan variasi data yang tersaji untuk model training/pelatihan tanpa benar-benar mengumpulkan/mengambil data baru [40]. Tergantung pada kebutuhan sistem, ada berbagai cara untuk memperluas data, beberapa di antaranya umum untuk pemrosesan gambar, seperti pemotongan gambar, pengubahan ukuran gambar, pencerminan gambar, rotasi gambar, dan pembesaran gambar yang digunakan untuk tujuan tertentu [41].

2.2.10 **Confusion Matrix**

Machine learning menggunakan *confusion matrix* sebagai *visual assessment tool*. Dalam confusion matrix, baris sesuai dengan hasil kelas

yang sebenarnya dan kolom dengan hasil kelas yang diprediksi. Ini mencantumkan setiap skenario yang mungkin muncul dari masalah klasifikasi [42]. Detail tabel prediksi dari *confusion matrix* terdapat pada gambar 2.9.



Gambar 2. 9 Confusion Matrix [43]

Dari gambar 2.9 diperoleh rumus untuk *accuracy*, *precision*, *recall* dan *f1-score* sebagai berikut [23].

Persentase data dari kumpulan data lengkap yang diprediksi secara akurat adalah akurasi. [24].

$$Accuracy = \frac{True\ hawar + True\ blas + True\ Tungro}{Jumlah\ sampel}$$

Beberapa data yang relevan dari semua hasil tes yang sebenarnya diterapkan untuk mendapatkan presisi [24].

$$Precision\ hawar = \frac{True\ hawar}{True\ hawar + False\ hawar}$$

$$Precision\ blas = \frac{True\ blas}{True\ blas + False\ blas}$$

$$Precision\ tungro = \frac{True\ tungro}{True\ tungro + False\ tungro}$$

Recall dihitung dengan membagi total *true positive* dan *false negative* dengan jumlah *true positive* [24].

$$\text{Recall hawar} = \frac{\text{True hawar}}{\text{True hawar} + \text{False hawar}}$$

$$\text{Recall blas} = \frac{\text{True blas}}{\text{True blas} + \text{False blas}}$$

$$\text{Recall tungro} = \frac{\text{True tungro}}{\text{True tungro} + \text{False tungro}}$$

Rata-rata presisi dan *recall* yang diselaraskan adalah definisi *F1-score*.

$$F1 - \text{score hawar} = 2 * \frac{\text{precision hawar} * \text{recall hawar}}{\text{precision hawar} + \text{recall hawar}}$$

$$F1 - \text{score blas} = 2 * \frac{\text{precision blas} * \text{recall blas}}{\text{precision blas} + \text{recall blas}}$$

$$F1 - \text{score tungro} = 2 * \frac{\text{precision tungro} * \text{recall tungro}}{\text{precision tungro} + \text{recall tungro}}$$