

BAB 2

TINJAUAN PUSTAKA

Didalam perancangan dan pembuatan Skripsi ini diperlukan beberapa teori-teori yang dapat menunjang dalam pembuatan Skripsi, maka penulis mengambil beberapa tinjauan pustaka yang dapat membantu menyelesaikan Skripsi tersebut. Salah satunya adalah dari referensial jurnal, dan teori yang dikutip dari beberapa artikel, buku, jurnal, penelitian terdahulu, catatan perkuliahan, dan internet. Berikut ini adalah beberapa penelitian yang berhubungan dengan rancangan sistem ini dan landasan teori yang dipakai sebagai rujukan dalam penelitian ini.

2.1. Kajian Penelitian Terkait

Pada penelitian pertama berupa Skripsi milik Maulana Muhammad Fathul Alim. Pada Skripsi ini, Para peneliti menggunakan berbagai model CNN, seperti VVG, ResNet, dan DenseNet, untuk membandingkan kinerja mereka dalam mendeteksi penyakit tanaman tomat menggunakan dataset PlantVillage, yang berisi 22.930 gambar daun tomat dalam 10 kategori. Hasil komparasi model CNN terbaik adalah *ResNet50* dengan 96% *accuracy*, 97% *precision*, 96% *recall*, 97% *f1-score*, dan 97,9% AUC. Dan *ResNet-50* mampu melakukan prediksi benar sebanyak 4402 data terhadap 4585 data yang telah di uji (Maulana Muhammad F.A., 2020).

Berikutnya penelitian yang kedua yaitu jurnal oleh Erlyna Nour Arrofiqoh dan Harintaka. Jurnal ini meneliti tentang bagaimana metode CNN mengklasifikasikan citra tanaman dimana lima kelas atau variabel yang digunakan adalah beras, bawang merah, kelapa, pisang, dan cabe. Dalam penelitian ini, nilai data verifikasi adalah 93%, dan akurasi data uji adalah 82% (Erlyna N.A., dkk. 2018).

Penelitian sejenis selanjutnya adalah sebuah skripsi yang disusun oleh Putri Navia Rena. Pada skripsi ini peneliti membuat sistem berdasarkan metode CNN Membantu orang biasa yang ingin belajar membaca lembaran musik. Keakuratan mengidentifikasi catatan input dalam sistem adalah 66,28% dengan menggunakan data sebanyak 2405 gambar dan 69% dengan menggunakan data *testing* baru sebanyak 400 gambar setelah diuji coba (Putri Navia R., 2019).

Yang keempat adalah sebuah thesis oleh Triano Nurhikmat. Dimana thesis ini juga meneliti dan membuat sistem Input Wayang Golek dan citra digital berbasis

CNN sebagai metode input. Kumpulan data yang digunakan adalah gambar yang dikumpulkan menggunakan proses crawling dari mesin pencari Google. Jumlah sampel yang digunakan adalah 240 data, dan jumlah cek sebanyak 60 data citra. Data terbagi menjadi 3 label: Cepot, Gatot Kaca, dan Semar. Nilai akurasi yang didapat adalah 95% pada pelatihan dan 90% pada pengujian (Triano Nurhikmat, 2018).

Penelitian serupa lainnya juga dimuat dalam jurnal Sarirotul Divineyah dan Agung Nilogiri. Dalam jurnal ini, peneliti membagi kumpulan data ke dalam 10 folder dengan 2000 gambar, yang dibagi secara acak menjadi gambar uji dan gambar pelatihan. Dan rata-rata tingkat akurasi hasil klasifikasi mencapai 85%. Pada saat yang sama, tingkat akurasi pengenalan mencapai 90%, yang diperoleh dari 40 gambar uji (Sarirotul Ilahiyah, dkk. 2018).

Selain itu penelitian sejenis terdapat pada jurnal oleh R. Mehindra Prasmatio, Basuki Rahmat, dan Intan Yuniar. Pada penelitian jurnal ini, Para peneliti menggunakan perpustakaan OpenCv untuk mengembangkan sistem CNN untuk pengenalan ikan real-time untuk pengenalan objek dan peralatan kamera. Para peneliti membagi setiap jenis ikan menjadi 900 gambar dari kumpulan data yang digunakan, 720 gambar digunakan untuk pelatihan dan 180 gambar digunakan untuk pengujian. Di sini, para peneliti mencapai tingkat akurasi 85,18% dalam 27 tes, mereka tidak dapat mengenalinya dalam 4 tes, dan mereka mampu mengenali gambar ikan sebanyak 23 kali (R. Mahindra Prasmatio, dkk. 2020).

Terakhir, penelitian terkait ada pada jurnal milik Ari Ashari Jaelani, Dr. Fiky Yosef Supratman, S.T, M.T, dan Nur Ibrahim, S.T,M.T. Penelitian pada jurnal ini membahas bagaimana menerapkan metode CNN untuk mengklasifikasikan tipe klon gambung berdasarkan foto daun. Dimana peneliti menggunakan 1966 dataset yang dibagi menjadi dua jenis dataset, yang pertama memiliki jumlah dataset sebanyak 1136 foto dan 830 foto. Dari hasil pengujian yang didapatkan, diperoleh nilai akurasi sebesar 60% dan hasil pengujian aplikasi sebesar 25% (A. Ashari Jaelani, dkk. 2020).

2.2. Landasan Teori

a. Melon

Melon atau disebut *Cucumis melo* mulai dikenal di Indonesia pada era 1980-an, saat tanaman anggota keluarga *Cucurbitaceae* ini di tanam di Kalianda, Lampung dan Cisarua, Bogor. Teknologi budidaya melon yang masih asing bagi para petani di perkenalkan oleh pada ahli dari Taiwan. Meskipun benih melon yang ditanam saat itu didatangkan dari berbagai negara, tetapi yang mendominasi adalah benih asal Pulau Formosa. Wajar jika sampai sekarang melon seakan identik dengan Taiwan. Daerah penanaman melon saat ini berkisar dari Jawa Barat, Jawa Tengah hingga Jawa Timur.

Keragaman melon sungguh luar biasa, ini terjadi karena keluarga *Cucumis melo* berasal dari spesies yang sama, sehingga biasa dikawinsilangkan dan menghasilkan varietas baru. Aneka warna dan bentuk melon tersebar di seluruh dunia: Afrika, India, Jepang, Spanyol, Eropa, amerika, Asia, dan China. Keberagaman melon itu bisa di kelompokkan dalam beberapa grup besar yaitu *reticulatus* (melon berjaring, daging buahnya yang hijau atau oranye dan beraroma), *flexuosus* (bentuknya persis mentimun, tetapi lebih panjang), *conomon* (melon yang umumnya berbentuk silindris, warnanya hijau, putih atau belang dan berdaging putih), *cantaloupe* (daging buahnya berwarna oranye, meskipun ada pula yang putih dan beraroma), *inodorus* (melon tanpa jaring. Daging buah tidak beraroma, dan biasanya berwarna hijau atau putih meskipun ada juga yang oranye), *dudaim* (melon berbentuk bulat atau oval, berukuran kecil dengan daging buah putih kekuningan dan warna kulit buah hijau, kuning atau strip hijau), dan *momordica* (bentuknya oval atau silindris dengan kulit yang halus). Tiga dari grup tersebut bernilai ekonomis tinggi, seperti *cantaloupe*, *inodorus*, dan *momordica* (Daryono, Budi Setiadi., dan Sigit Dwi Maryanto. 2017).

Ada dua kelompok besar buah melon di Indonesia yaitu melon berjaring dan tanpa jaring. Melon berjaring adalah melon yang pertama kali di kenal di Indonesia. Kulit buahnya tebal, keras, dan tahan lama jika disimpan. Ada dua jenis melon berjaring. Pertama musk melon *Cucumis melo var. reticulatus*. buah

berbentuk bulat, berkulit keras dengan daging buah hijau. Contoh yang paling terkenal dari musk melon ialah sky rocket. Tipe kedua dari melon berjaring ialah cantaloupe *Cucumis melo var. cantelupensis*. nama *cantaloupe* diambil dari nama desa italia, *Cantalup*, tempat pertama kali varietas ini di tanam. Ukuran *cantaloupe* relative lebih besar dengan warna daging kuning kemerahan dan beraroma harum. Daging buah agak tipis. Varietas *new century* contoh *cantaloupe* yang pernah beredar sejak era 1980-an. Dan jenis ini adalah *honeydew* (The Best Melon-Jakarta Trubus, 2011).



Gambar 2.1. Buah Melon Sky Rocket

Pada Gambar 2.1 yaitu buah melon jenis sky rocket adalah varietas yang bertahan lama dan sudah dikenal sejak 1980-an. Bentuk buahnya bulat berkulit hijau dengan kadar kemanisan 9° briks. Tekstur daging agak juicy dan lembut.

b. Citra Digital

Disiplin yang berhubungan dengan teknologi pemrosesan gambar (termasuk foto dan video) disebut pemrosesan gambar digital. Proses ini dilakukan dengan mengubah citra menjadi citra diskrit melalui proses scanning. Secara khusus, digitalisasi dicapai dengan mengubah gambar, teks, dan suara dalam data digital menjadi data elektronik untuk tujuan tertentu. Dalam konteks ini, pengolahan citra digital dipahami sebagai data dua dimensi. Dalam hal ini, kata bilangan berarti pengolahan citra atau citra dilakukan oleh komputer.

(Sutoyo, dkk 2009). Dalam citra digital, terdapat deretan bilangan real atau bilangan kompleks yang diwakili oleh urutan bit tertentu.

Pemrosesan citra digital menggunakan komputer dilakukan dengan membandingkan citra dua dimensi pada grid dengan elemen piksel (elemen terkecil dari citra). Setiap piksel memiliki nomor, dan setiap nomor dalam piksel mewakili saluran warna. Dimana dalam proses tersebut akan mengurutkan angka-angka dalam piksel untuk pengolahan tertentu. Matriks pada citra digital terdiri dari M kolom dan N baris, yang memiliki dua parameter yaitu intensitas warna dan koordinator. Nilai yang diterapkan pada koordinator (x, y) adalah f(x, y), yaitu besar intensitas atau warna dari piksel dititik itu. Oleh karena itu, citra dapat dituliskan kedalam sebuah matriks pada rumus (2.1):

$$f(x, y) = \begin{bmatrix} f(0,0) & f(0,1) & \dots & f(0, M - 1) \\ f(1,0) & f(1,1) & \dots & f(1, M - 1) \\ \vdots & \vdots & \vdots & \vdots \\ f(N - 1,0) & f(N - 1,1) & \dots & f(N - 1, M - 1) \end{bmatrix} \quad (2.1)$$

Berdasarkan rumus (2.1), suatu citra f(x, y) dapat dituliskan kedalam fungsi matematis seperti rumus (2.2) berikut ini:

$$\begin{aligned} 0 \leq x \leq M - 1 \\ 0 \leq y \leq N - 1 \\ 0 \leq f(x, y) \leq G - 1 \end{aligned} \quad (2.2)$$

Dimana

M = Jumlah piksel baris pada array citra

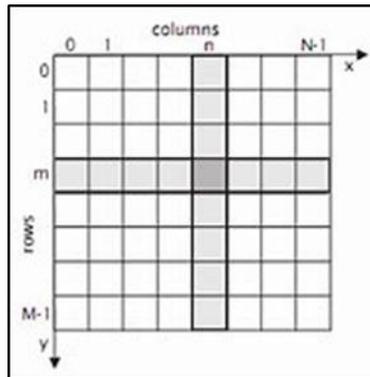
N = Jumlah piksel kolom pada array citra

G = Nilai skala keabuan (grayscale)

Nilai M, N, dan G biasanya pangkat 2, seperti terlihat pada persamaan berikut (2.3):

$$M = 2^m / N = 2^n / G = 2^k \quad (2.3)$$

Nilai m, n, dan k adalah bilangan positif. Interval (0, G) disebut (skala abu-abu). Nilai G tergantung pada proses pemindaian. Biasanya 0 (nol) abu-abu mewakili intensitas hitam dan 1 (satu) intensitas putih. Untuk gambar 8-bit, nilai G adalah $2^8 = 256$ warna (tingkat abu-abu). Seperti Gambar 2.2 sebagai berikut:



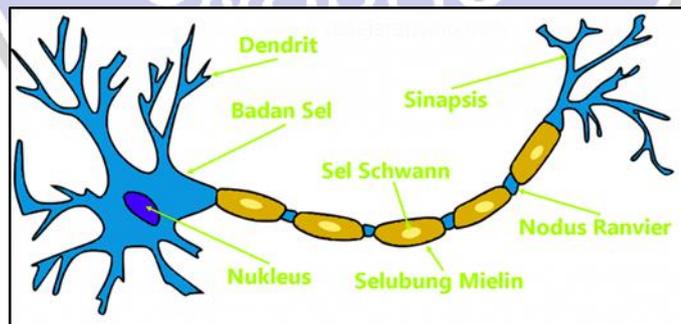
Gambar 2.2. Representasi Citra Digital 2d
(Bernd, 2000)

1.) Pengolahan Citra

Pengolahan citra adalah proses pengolahan citra dengan komputer untuk memperoleh kualitas yang lebih baik. Dalam hal ini, editing gambar bertujuan untuk memudahkan orang dalam meningkatkan kualitas gambar dan agar lebih mudah dipahami.

c. Artificial Neural Network

Artificial Neural Network (ANN) Ini adalah model komputasi paralel yang mensimulasikan fungsi sistem jaringan saraf biologis otak manusia. Otak manusia terdiri dari miliaran neuron yang saling berhubungan. Hubungan ini disebut *Synapses*. Komponen saraf terdiri dari nukleus yang memproses informasi, akson, dan setidaknya satu dendrit. Informasi yang masuk diterima oleh dendrit. Selain itu, melalui pemrosesan informasi, dendrit juga mengandung akson sebagai keluaran dari suatu pemrosesan informasi (Adrian Rosebrock, 2017).



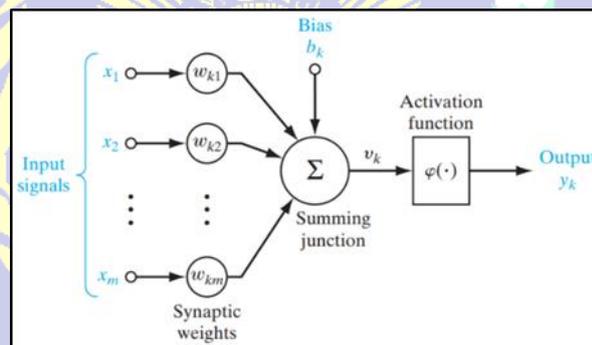
Gambar 2.3. Jaringan Syaraf Manusia

Sumber : www.eruptingmind.com

Fungsi sistem saraf pada Gambar 2.3 di atas harus dimulai dengan sinyal input melalui dendrit ke badan sel. Kemudian sinyal tersebut diproses di dalam badan sel sesuai dengan fungsi tertentu (proses penjumlahan). Ketika sinyal yang diproses melebihi ambang batas tertentu, sinyal menciptakan neuron untuk melanjutkan sinyal. Jika di bawah ambang batas, sinyal diblokir (inhibited), dan kemudian sinyal yang akan diputar dikirim ke axon, dan akhirnya ke neuron lain melalui synapse (Purnomo, Mauridhi Hery, dan Agus Kurniawan. 2006).

1.) Komponen Neural Network

Neural Network terdiri dari beberapa jenis yang berbeda, tetapi hampir semua komponennya sama. Mirip dengan jaringan saraf di otak manusia, Neural Network juga terdiri dari beberapa unit neuron yang saling berhubungan. Pada setiap neuron ini berkomunikasi dengan neuron lain untuk mengubah informasi yang diterimanya. Hubungan ini sering disebut bobot (*Weight*). Informasi disimpan dalam nilai tertentu dengan bobot tertentu. Ini adalah struktur neuron dalam jaringan saraf (Neural Network).



Gambar 2.4. Struktur Neural Network

Sumber : www.innoarchitech.com

Berdasarkan Gambar 2.4 diatas menunjukkan struktur *neural network*.

Adapun komponen bangunan tersebut adalah sebagai berikut:

- a) *Input* terdiri dari variable independen ($X_1, X_2, X_3, \dots, X_m$) yang merupakan sebuah sinyal yang masuk ke sel syaraf.
- b) Bobot (*Weight*) terdiri dari beberapa bobot ($w_{k1}, w_{k2}, w_{k3}, \dots, w_{km}$) yang berhubungan dengan masing-masing node.
- c) *Threshod* merupakan nilai ambang batas internal dari node. Besar nilai ini mempengaruhi aktivitas dari *output* node y .

d) *Activation Function* (fungsi aktifasi) merupakan operasi matematika yang dikenal pada sinyal *output* y .

Fungsi dari struktur jaringan saraf yang disebutkan di atas tidak berbeda secara signifikan dari jaringan saraf manusia. Informasi (data record) dikirim dengan bobot kedatangan tertentu. Masukan tersebut kemudian diproses oleh fungsi ekstensi yang menjumlahkan nilai semua bobot yang masuk. Fungsi aktivasi masing-masing neuron membandingkan hasil penjumlahan ini dengan ambang (*threshold*) tertentu. Jika sinyal input melebihi ambang batas tertentu, neuron akan memicu. Jika tidak, neuron tidak akan memicu. Ketika sebuah neuron diaktifkan, ia akan mengirimkan outputnya ke semua neuron yang terhubung dengannya berdasarkan bobot outputnya, dan seterusnya (Hermawan, 2006).

Pada lapisan neuron, susunan neuron-neuron dikumpulkan pada lapisan-lapisan neuron, kemudian neuron-neuron pada lapisan ini dihubungkan bolak-balik selain lapisan input dan output. Informasi yang diperoleh pada tahap awal masuk berpindah dari satu level ke level lainnya, dari level masuk ke level keluar. Lapisan ini biasanya disebut lapisan tersembunyi (*hidden layer*). Umumnya, setiap neuron pada lapisan yang sama memiliki status yang sama. Oleh karena itu, setiap neuron di setiap lapisan memiliki fungsi aktifasi yang sama. Hubungan antara lapisan dan neuron harus selalu terhubung. Faktor terpenting yang menentukan perilaku neuron adalah mode bobot dan fungsi aktivasinya

d. Deep Learning

Teknologi deep learning adalah teknologi pembelajaran mesin, yang mengajarkan komputer kemampuan serupa dengan kemampuan alami manusia yaitu belajar dari pengalaman mereka sendiri. Karena hasilnya yang bagus, metode deep learning baru-baru ini mendapat banyak perhatian penelitian.

Pembelajaran simulasi komputer untuk menggunakan pembelajaran mendalam untuk mengklasifikasikan gambar, teks, atau suara secara langsung. Penggunaan pembelajaran mendalam ini mencapai tingkat akurasi yang tinggi, dan dalam beberapa kasus, kinerjanya melebihi kinerja pada skala manusia.

Model ini menggunakan sejumlah besar data yang dikombinasikan dengan arsitektur jaringan saraf multilayer untuk pelatihan.

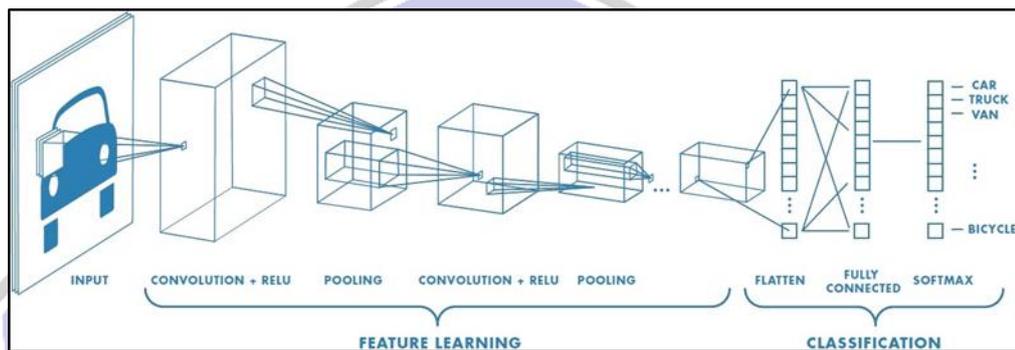
Meskipun istilah deep learning telah muncul di jaringan saraf sejak tahun 2006, istilah tersebut tidak muncul kembali hingga beberapa tahun yang lalu. Ini dilakukan karena dua alasan: Pertama, metode pembelajaran mendalam membutuhkan banyak data, sehingga keterampilan pembelajaran mendalam presisi tinggi terlihat nyata. Alasan Kedua, Pembelajaran mendalam dengan kumpulan data besar membutuhkan komputer berperforma tinggi untuk meningkatkan kecepatan pemrosesan. Alasan kedua membuat metode deep learning sangat populer di kalangan peneliti, karena semakin banyak perangkat komputasi, seperti GPU berperforma tinggi dan komputasi awan, yang membantu perkembangan metode ini dengan cepat (Hermawan, 2006).

Salah satu teknik deep learning yang digunakan oleh deep neural network adalah Convolutional Neural Networks (CNN). Nama tersebut berasal dari operasi metode yang paling penting, yaitu operasi konvolusi, yang kemudian digabungkan dengan metode jaringan saraf. CNN ini merupakan model algoritma deep neural network terbaru yang dirancang untuk mengolah data berupa data dua dimensi seperti gambar. Beberapa algoritma pemrosesan gambar lainnya, algoritma dari CNN ini, menghilangkan kebutuhan untuk ekstraksi fitur manual, sehingga ekstraksi fitur otomatis ini memungkinkan model pembelajaran mendalam untuk melakukan tugas visi komputer dengan sangat akurat, seperti klasifikasi objek.

e. Convolutional Neural Network

Convolutional Neural Network (CNN) Ini adalah pengembangan dari *multi layer perceptron* (MLP) untuk memproses data gambar 2D. Karena kedalaman jaringannya yang tinggi, CNN adalah jenis jaringan saraf yang dalam dan sering diterapkan pada data gambar. Pada prinsipnya, klasifikasi citra dapat digunakan dengan MLP, tetapi metode MLP tidak cocok untuk digunakan karena tidak menyimpan informasi spasial dalam data citra, dan mengasumsikan bahwa setiap piksel adalah fitur independen, sehingga menghasilkan hasil yang buruk.

Penelitian asli di balik penemuan CNN pertama kali dilakukan oleh Hubel dan Wiesel (Hubel & Wiesel, T, 1968) yaitu tentang *visual cortex* pada indera penglihatan kucing. Secara teknis, CNN adalah arsitektur multi-langkah yang dapat dilatih. *Input* dan *output* dari setiap level terdiri dari beberapa *matriks*, biasanya disebut *feature map*. Setiap tahap terdiri dari tiga lapisan, yaitu lapisan *convolutional*, lapisan fungsi aktivasi dan lapisan pengelompokan (*pooling layer*). Gambar 2.5 adalah jaringan arsitektur *Convolutional Neural Network*, Sebagai Berikut:



Gambar 2.5. Arsitektur Convolutional Neural Network

Bersasarkan Gambar 2.5 diatas, Tahap pertama dalam arsitektur CNN adalah tahap convolution. Tahap ini dilakukan dengan ukuran kernel tertentu. Perhitungan jumlah kernel yang digunakan tergantung dari jumlah fungsi yang dibuat. Kemudian nonaktifkan fungsi tersebut, umumnya menggunakan ReLU (Rectifier Linear Unit) untuk mengaktifkan fungsi tersebut, kemudian melewati proses pengelompokan (*pooling*) setelah menyelesaikan proses aktivasi fungsi. Proses ini akan diulang beberapa kali hingga peta fitur yang cukup untuk menavigasi ke jaringan saraf yang terhubung penuh diperoleh, dan kelas output dari jaringan saraf yang terhubung penuh ditemukan (Hui Li *et al.*, 2018).

1.) Convolution Layer

Lapisan convolutional adalah bagian dari adegan dalam arsitektur CNN. Pada tahap ini dilakukan konvolusi pada output dari layer sebelumnya. Lapisan ini merupakan proses utama yang membentuk fondasi dari arsitektur jaringan CNN. Konvolusi adalah istilah matematika yang mengulangi penerapan satu fungsi ke output dari fungsi lain. Operasi konvolusi adalah operasi dua fungsi parameter bernilai nyata. Operasi ini

menerapkan fungsi *output* sebagai fungsi yang dipetakan ke gambar *input*. Input dan output ini dapat dianggap sebagai dua parameter yang valid (Sri Kusumadewi, 2004). Operasi konvolusi dapat dituliskan pada rumus (2.4) sebagai berikut:

$$s(t) = (x * t)(t) = \sum_a^{\infty} = -\infty x(a) * w(t - a) \quad (2.4)$$

Keterangan:

- S(t) = Fungsi hasil operasi konvolusi
- X = Input
- W = Bobot (kernel)

Fungsi S(t) memberikan hasil tunggal dalam bentuk peta fitur. Parameter pertama adalah input x, dan parameter kedua adalah w sebagai kernel atau filter. Jika kita memperlakukan input sebagai gambar, kita dapat mengatakan bahwa t adalah piksel dan menggantinya dengan i dan j. oleh sebab itu, operasi pada konvolusi ke input dengan lebih dari satu dimensi dapat dituliskan pada (2.5 & 2.6) sebagai berikut:

$$s(ij) = (K * I)(ij) = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} I(i - m, j - n) K(m, n) \quad (2.5)$$

$$s(ij) = (K * I)(ij) = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} I(i + m, j + n) K(m, n) \quad (2.6)$$

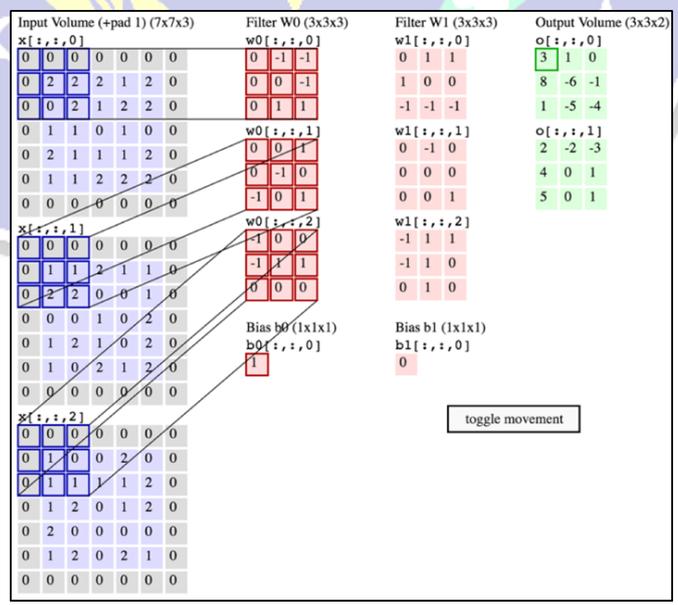
Berdasarkan dua persamaan di atas, ini adalah perhitungan berdasarkan konvolusi, di mana i dan j adalah piksel pada gambar. Perhitungan kumulatif dan terjadi ketika K adalah kernel, maka s adalah input, dan kernel adalah input reversibel. Atau, pertimbangkan operasi konvolusi sebagai perkalian matriks antara gambar input dan kernel, dan gunakan produk skalar untuk menghitung output. Selain itu, definisi output dapat didefinisikan untuk setiap lapisan *hyperparameter*. *Hyperparameter* yang digunakan dalam persamaan di bawah ini digunakan untuk menghitung jumlah *neuron* yang ditembakkan pada *output* (Sri Kusumadewi, 2004). Perhatikan persamaan (2.7) berikut:

$$\frac{(W - f + 2P)}{(S + 1)} \quad (2.7)$$

Keterangan :

- W = Ukuran volume gambar
- F = Ukuran filter
- P = Nilai padding yang digunakan
- S = Ukuran pergeseran

Berdasarkan persamaan di atas, ukuran volume (W), filter (F), langkah aplikasi (S), dan jumlah padding nol yang digunakan (P) dapat digunakan sebagai hyperparameter untuk menghitung ukuran spasial volume keluaran. Stridenya adalah nilai yang digunakan untuk mengimbangi filter pada gambar *input*, dan *padding* nol adalah nilai yang digunakan untuk mendapatkan angka nol di tepi gambar. Di bawah ini adalah operasi *convolutional layer* yang terdiri dari *neuron-neuron* yang disusun membentuk filter dengan panjang dan tinggi (dalam piksel). Misalnya, lapisan pertama dalam lapisan ekstraksi fitur biasanya *conv*. Lapisan 3x3x3. Panjang 3 piksel, tinggi 3 piksel dan ketebalan/kuantitas 3 blok, tergantung pada saluran gambar. Ketiga filter ini dipindahkan ke semua bagian gambar. Setiap shift melakukan operasi "titik" antara input dan nilai filter untuk menghasilkan output, atau biasanya dalam bentuk peta aktivasi atau peta fitur (*feature map*). Perhatikan ilustrasi Gambar 2.6 berikut:

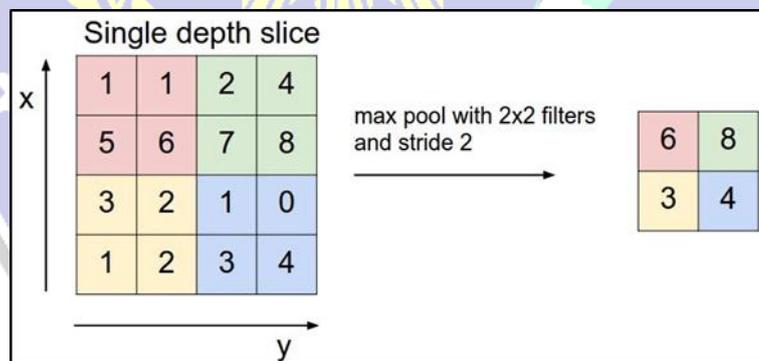


Gambar 2.6. Proses Convolution Layer

Sumber: CNN For Visual Recognition

2.) Pooling Layer

Pooling adalah penggunaan operasi gabungan untuk mengurangi ukuran *array*. Tingkat pengelompokan biasanya setelah *conv*. Pada dasarnya, tingkat pengelompokan terdiri dari filter dengan ukuran dan panjang langkah tertentu, yang secara bergantian bergerak melintasi seluruh area objek grafik (*feature map*). Pada tingkat pengelompokan, ada dua kelompok yang umum digunakan, yaitu pengelompokan tengah dan pengelompokan terbesar. Nilai asumsi pada kelompok tengah adalah nilai rata-rata, dan nilai asumsi pada kelompok maksimum (*max-pooling*) adalah nilai maksimum. Dalam arsitektur model CNN, pengelompokan lapisan penyisipan di antara lapisan konvolusi yang berurutan dapat secara bertahap mengurangi ukuran volume *output* di peta fungsi, sehingga mengurangi parameter dan perhitungan dalam jaringan untuk mengontrol *over-fitting*. Lapisan pengelompokan menggunakan dan mengurangi ukuran tumpukan peta fitur apa pun. Cara umum untuk mengelompokkan lapisan adalah dengan menggunakan filter 2x2, yang diterapkan dalam dua langkah dan bekerja pada setiap segmen input. Berikut ini Gambar 2.7 adalah contoh gambar operasi *maxpooling*:



Gambar 2.7. Operasi Max-Pooling

Sumber: CNN For Visual Recognition (n.d.)

Pada Gambar 2.7 di atas menunjukkan proses pengelompokan maksimum. Hasil dari proses grouping adalah array yang lebih kecil dari citra aslinya. Pengelompokan level yang disebutkan di atas bekerja secara

bergantian dengan setiap segmen kedalaman volume input (Adrian Rosebrock, 2017).

Seperti yang ditunjukkan pada Gambar 2.7, operasi pengelompokan maksimum menggunakan filter 2x2. Ukuran input proses adalah 4x4. Untuk masing-masing dari 4 angka dalam operasi *input*, ambil nilai maksimum, lalu buat ukuran *output* 2x2 baru.

3.) Fully Connected Layer

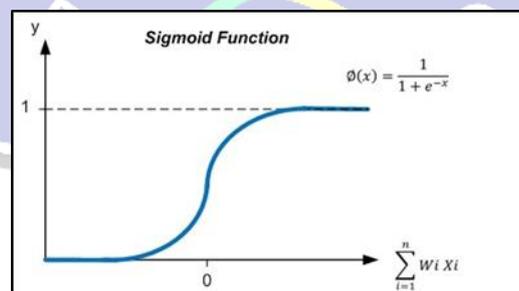
Fully-Connected Layer adalah lapisan di mana semua neuron yang diaktifkan di lapisan atas terhubung ke neuron di lapisan berikutnya seperti jaringan saraf normal. Lapisan ini terutama digunakan untuk MLP (*Multilayer Perceptron*), yang tujuannya adalah untuk mengubah ukuran data sehingga data dapat diklasifikasikan secara linier.

Seperti pada Gambar 2.5 Perbedaan antara fully connected layer dan convolutional layer biasa adalah bahwa neuron pada convolutional layer hanya terhubung ke area input tertentu, sedangkan pada fully connected layer memiliki neuron yang terhubung penuh (Adrian Rosebrock, 2009).

4.) Aktivasi Sigmoid

Fungsi aktivasi sigma (*sigmoid*) adalah fungsi nonlinier. Input dari fungsi pemicu ini adalah bilangan real, dan rentang output pemicu adalah 0 hingga 1. Berikut adalah fungsi (2.8) yang diberikan:

$$A = \frac{1}{1 + e^{-x}} \quad (2.8)$$



Gambar 2.8. Fungsi Aktivasi Sigmoid

(Yopi Andry Lesnussa & Lexy Janzen Sinay, 2017)

Jika masukan dari sebuah node pada jaringan syaraf memiliki nilai *negative* maka hasil yang diperoleh adalah 0, maka apabila inputannya

mempunyai nilai positif maka *outputnya* bernilai 1. Kekurangan dari fitur ini adalah jika aktivasi neuron menyebabkan nilai berada pada kisaran 0 atau 1, maka sigmoid dapat mematikan gradien, dimana gradien hampir 0. Maka keluaran sigmoid tidak akan berpusat pada nol (*zero-centered*) (Yopi Andry Lesnussa & Lexy Janzen Sinay, 2017).

