

# **SKRIPSI**

## **Implementasi Convolutional Neural Networks menggunakan VGG-19 Net pada Image Depth Data Hand Posture**

Disusun dan diajukan oleh

**FATUR RAHMAN**

**H 131 16 518**



**PROGRAM STUDI ILMU KOMPUTER DEPARTEMEN MATEMATIKA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS HASANUDDIN**

**MAKASSAR**

**JANUARI 2020**

## LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

**Implementasi Convolutional Neural Networks menggunakan VGG-19 Net pada Image Depth Data Hand Posture**

adalah benar hasil karya saya sendiri, bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun.

Makassar, 31 Januari 2020



FATUR RAHMAN  
NIM. H 131 16 518

**Implementasi Convolutional Neural Networks menggunakan VGG-19 Net pada Image Depth Data Hand Posture**

Disusun dan diajukan oleh

**FATUR RAHMAN**

**H13116518**

Telah dipertahankan di hadapan Panitia Ujian yang dibentuk dalam rangka Penyelesaian Studi Program Sarjana pada Program Studi Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin dan dinyatakan telah memenuhi syarat kelulusan

Menyetujui,

Pembimbing Utama



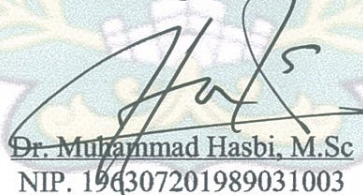
Dr. Hendra S.Si, M.Kom  
NIP. 197601022002121001

Pembimbing Pertama



Supri Bin Hj Amir, S.Si., M.Eng.  
NIP. 198805042019031012

Ketua Program Studi



Dr. Muzaymah Hasbi, M.Sc  
NIP. 196307201989031003



## HALAMAN PENGESAHAN

Skripsi ini diajukan oleh:

Nama : Fatur Rahman

NIM : H 131 16 518


Program Studi : Ilmu Komputer

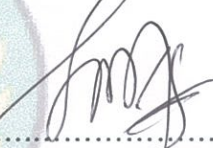
Judul Skripsi : Implementasi Convolutional Neural Networks menggunakan VGG-19 Net pada Image Depth Data Hand Posture

Telah berhasil dipertahankan di hadapan dewan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Sains pada Program Studi Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

### DEWAN PENGUJI

#### Tanda Tangan

1. Ketua : Dr. Hendra S.Si, M.Kom. (..........)

2. Sekretaris : Supri Bin Hj Amir, S.Si., M.Eng. (..........)

3. Anggota : Dr. Eng. Armin Lawi, S.Si., M.Eng. (..........)

4. Anggota : Dr. Loeky Haryanto, MS. M.Sc, M.Math. (.....)

Ditetapkan di : Makassar

Tanggal : 29 Januari 2020



## KATA PENGANTAR

Segala puji bagi Allah *Subhanahu Wa ta'ala*, Tuhan atas langit dan bumi beserta segala isinya. Karena, berkat nikmat dan karuniaNYA sehingga penulisan skripsi ini dapat terselesaikan. Shalawat serta salam semoga senantiasa tercurahkan kepada Baginda Rasulullah Muhammad *Shallallahu Alaihi Wasallam* dan kepada para keluarga serta sahabat beliau, yang senantiasa menjadi teladan yang baik.

Alhamdulillah, skripsi dengan judul “Implementasi Convolutional Neural Networks menggunakan VGG-19 Net pada Image Depth Data Hand Posture” yang disusun sebagai salah satu syarat akademik untuk meraih gelar Sarjana Sains pada Program Studi Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin ini dapat dirampungkan. Tentunya, dalam penulisan skripsi ini, penulis mampu melewati berbagai hambatan dan masalah berkat bantuan moril dan materiil, serta dorongan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan ucapan terima kasih yang tak terhingga kepada orang tua penulis, Ibunda **Siti Nuriyah** dan Ayahanda **Samuji**, sebagai tempat kembali setelah pergi, dan tempat terlelap dikala lelah, terima kasih atas kasih sayang, doa, dan nasihat yang tulus sebagai bekal kehidupan. Rasa terima kasih juga penulis tujukan kepada Kakanda **Siti Hanifah** , **Koiriyah**, **Kusniati**, sebagai kakak, motivator, dan rival dalam membanggakan kedua orang tua, terima kasih atas bimbingan yang penulis dapatkan selama ini.

Penghargaan dan ucapan terima kasih dengan penuh ketulusan juga penulis ucapkan kepada:

1. Bapak **Dr. Nurdin, S.Si. M.Si** , sebagai Ketua Departemen Matematika FMIPA Unhas. Penulis juga berterima kasih atas dedikasi dosen-dosen pengajar, serta staf Departemen atas ilmu dan bantuan yang bermanfaat.
2. Bapak **Dr. Hendra S.Si, M.Kom.** , sebagai dosen pembimbing utama sekaligus ketua tim penguji atas ilmu yang beliau berikan selama proses perkuliahan, dan

kesediaan beliau dalam membimbing, serta memotivasi penulis dalam penyusunan skripsi ini.

3. Bapak **Supri Bin Hj Amir, S.Si., M.Eng.** , sebagai dosen pembimbing pertama sekaligus sekretaris tim penguji atas ilmu yang beliau berikan selama proses perkuliahan, dan bimbingan serta segala bentuk bantuan yang telah beliau berikan dalam penyusunan skripsi ini.
4. Bapak **Dr. Eng. Armin Lawi, S.Si., M.Eng.** sebagai anggota tim penguji atas segala ilmu yang telah beliau berikan selama proses perkuliahan bentuk kritik dan masukan yang membangun selama proses penyusunan skripsi ini
5. Bapak **Dr. Loeky Haryanto, MS. M.Sc, M.Math.** , selaku dosen pembimbing akademik atas segala bentuk nasihat serta arahan beliau selama proses perkuliahan serta ilmu yang tak ternilai yang beliau berikan selama proses perkuliahan.
6. Sahabat-sahabatku **ILMU KOMPUTER 2016**, atas kebersamaan, kepedulian, suka-duka, canda tawa yang telah kita lewati selama ini. Semoga persahabatan kita yang telah terjalin tidak pernah usai.
7. Keluarga besar **A16ORITMA 2016** atas segala bentuk dukungan dan bantuan selama proses perkuliahan. Semoga kesuksesan selalu kita dapatkan dalam setiap langkah-langkah kita.
8. Keluarga besar **KM FMIPA Unhas** terkhusus kepada **MIPA 2016** atas persahabatan, kekerabatan, kerja sama, serta cerita-cerita lain yang telah kita ukir bersama.
9. Keluarga besar **Himatika FMIPA Unhas** terkhusus kepada **Pengurus BE Himatika FMIPA Unhas Periode 2018/2019** atas ilmu-ilmu organisatoris yang tidak didapatkan di bangku perkuliahan. Terima kasih pula atas kekeluargaan yang terjalin. Semoga kekeluargaan kita tetap terjalin hingga hari tua nanti.
10. Keluarga besar **KKN E-Commerce Luwu Utara** yang secara ikhlas dan tulus mengabdikan kepada masyarakat khususnya UKM UKM yang ada di Kabupaten Luwu Utara. Terkhusus kepada Bapak **Jafar Kadoeng** yang secara terbuka

menerima kami di kediamannya. Semoga Bapak diberikan kesehatan dan rezeki yang berkecukupan.

11. Seluruh pihak yang yang tidak dapat disebutkan satu per satu atas segala bentuk kontribusi, partisipasi, serta motivasi yang diberikan kepada penulis selama ini. Semoga apa yang kita berikan, dilipatgandakan oleh Tuhan Yang Maha Kaya.

Penulis menyadari bahwa masih banyak kekurangan dalam tugas akhir ini, untuk itu dengan segala kerendahan hati penulis memohon maaf. Akhir kata, semoga tulisan ini memberika manfaat untuk pembaca.

Makassar, 28 Januari 2020



FATUR RAHMAN  
H13116518

**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR  
UNTUK KEPENTINGAN AKADEMIS**

---

Sebagai civitas akademik Universitas Hasanuddin, saya yang bertanda tangan di bawah ini: \_

Nama : Fatur Rahman  
NIM : H 131 16 518  
Program Studi : Ilmu Komputer  
Departemen : Matematika  
Fakultas : Matematika dan Ilmu Pengetahuan Alam  
Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul:


**“Implementasi Convolutional Neural Networks menggunakan VGG-19 Net  
pada Image Depth Data Hand Posture”**

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar, pada tanggal 28 Januari 2020

Yang menyatakan

  
(Fatur Rahman)



## ABSTRAK

Penelitian ini dilakukan untuk mengetahui implementasi *Convolutional Neural Network* pada *Image Depth Data Hand Posture*. *Hand Posture* digunakan untuk berinteraksi dimana bentuk tangan dan gerakan dimanfaatkan untuk berkomunikasi. *Hand Posture* sendiri sulit untuk di Klasifikasi karena beragam objek tangan manusia merupakan objek artikulasi yang kompleks. Pengklasifikasian Gambar dengan *Convolutional Neural Network* sekarang ini banyak menarik perhatian , pendekatan pembelajaran untuk pengklasifikasian gambar dari objek yang sama namun dari subjek yang berbeda sangat terbatas. Model yang biasa digunakan untuk Klasifikasi citra menggunakan *Convolutional Neural Networks* sangat beragam, penentuan arsitektur dalam pembuatan model merupakan penentu hasil Klasifikasi pada citra. Pada penelitian ini Arsitektur yang digunakan dalam pembuatan model adalah Arsitektur *VGG-19 Net*. Berdasarkan hasil penelitian pada *Image Depth Data Hand Posture*, persentase tingkat akurasi pengklasifikasian menggunakan model *Convolutional Neural Network* dengan Arsitektur *VGG-19 Net* pada setiap subjek masing masing adalah 0.9976, 1.0, 0.9984, 1.0 dan 0.9992 dan *AUC* masing masing sebesar 1.0 , 1.0 , 0.9992 , 1.0 , 1.0.

**Kata Kunci:** Convolutional Neural Netrowks, *VGG-19 Net*, Klasifikasi, Hand Posture.

## ABSTRACT

This research was conducted to determine the implementation of *Convolutional Neural Network* in Image Depth Data Hand Posture. Hand Posture used to interact where the hand and movement are utilized to communicate. Hand Posture itself is difficult to classify because the various objects of human hands are complex articulation objects. Classifying images with the *Convolutional Neural Network* is now attracting a lot of attention, a learning approach for classifying images of the same object but from different subjects is very limited. The models commonly used for image classification using *Convolutional Neural Networks* are very diverse, the determination of architecture in model making is a determining result of classification on imagery. In this research the architecture used in model making is the *VGG-19* Net architecture. Based on the research results in Image Depth Data Hand Posture, the percentage of classifying accuracy rate using *Convolutional Neural Network* model with *VGG-19* Net architecture on each subject are 0.9976, 1.0, 0.9984, 1.0 and 0.9992 and *AUC* respectively of 1.0, 1.0, 0.9992, 1.0, 1.0.

**keywords:** *Convolutional Neural Networks, VGG-19 Net, Classification, Hand Posture*

## DAFTAR ISI

BAB I	PENDAHULUAN.....	1
1.1	Latar Belakang.....	1
1.2	Rumusan Masalah .....	3
1.3	Batasan Masalah.....	3
1.4	Tujuan Penelitian.....	3
1.5	Manfaat Penelitian.....	4
BAB II	TINJAUAN PUSTAKA .....	5
2.1	Deep Learning .....	5
2.2	Neural Networks.....	5
2.3	Convolutional Neural Networks.....	7
2.3.1	Lapisan Konvolusi .....	8
2.3.2	Fungsi Aktivasi ReLU.....	10
2.3.3	Pooling .....	11
2.3.4	Fully Connected Layer.....	12
2.4	Softmax <i>Loss</i> Function .....	13
2.5	Image Depth Data.....	14
2.6	Confusion Matrix.....	14
2.7	Arsitektur <i>VGG Net</i> .....	15
2.8	Google Colaboratory .....	17
BAB III	METODE PENELITIAN .....	19

3.1	Waktu dan Tempat .....	19
3.2	Tahapan Penelitian .....	19
3.2.1	Tahap Pra Penelitian(Persiapan) .....	19
3.2.2	Penelitian.....	19
3.3	Deskripsi Data .....	20
3.4	Alur Penelitian.....	21
BAB IV	HASIL DAN PEMBAHASAN .....	23
4.1	Pembuatan Label .....	23
4.2	Pembagian Dataset untuk Training Data dan Data Uji .....	24
4.3	Arsitektur Jaringan VGG-19 Net.....	25
4.4	Proses Training dan Validation .....	29
4.5	Akurasi .....	29
4.5.1	Akurasi dan <i>Loss</i> Subjek 1 .....	29
4.5.2	Akurasi dan <i>Loss</i> Subjek 2 .....	30
4.5.3	Akurasi dan <i>Loss</i> Subjek 3 .....	32
4.5.4	Akurasi dan <i>Loss</i> Subjek 4.....	33
4.5.5	Akurasi dan <i>Loss</i> Subjek 5 .....	34
4.6	Evaluasi Model.....	36
BAB V	KESIMPULAN DAN SARAN .....	37
5.1	Kesimpulan.....	37
5.2	Saran .....	37
Lampiran	.....	41
Lampiran 1	: Confusion Matrix Subjek 1.....	41

Lampiran 2 : Confusion Matrix Subjek 2.....	42
Lampiran 3 : Confusion Matrix Subjek 3.....	43
Lampiran 4 : Confusion Matrix Subjek 4.....	43
Lampiran 5 : Confusion Matrix Subjek 5.....	44
Lampiran 6 : Presisi , Recall , F1 setiap kelas Subjek 1.....	45
Lampiran 7 : Presisi , Recall , F1 setiap kelas Subjek 2.....	45
Lampiran 8 : Presisi , Recall , F1 setiap kelas Subjek 3.....	47
Lampiran 8 : Presisi , Recall , F1 setiap kelas Subjek 4.....	48
Lampiran 9 : Presisi , Recall , F1 setiap kelas Subjek 5.....	49
Lampiran 10 : Source Code.....	70

## DAFTAR GAMBAR

Gambar 1 Ilustrasi Neural Networks.....	6
Gambar 2 Ilustrasi Neuron .....	6
Gambar 3 Gambaran Umum Proses CNN .....	7
Gambar 4 Ilustrasi penerapan CNN pada sebuah Image .....	8
Gambar 5 ReLU .....	10
Gambar 6 Contoh Fungsi Aktivasi ReLU.....	11
Gambar 7 Contoh Operasi Max Pooling.....	11
Gambar 8 Downsampling .....	12
Gambar 9 Proses Flatten .....	12
Gambar 10 Arsitektur CNN .....	13
Gambar 11 Contoh penggambaran kedalaman data.....	14
Gambar 12 Contoh arsitektur vgg-net.....	16
Gambar 13 Arsitektur VGG-19 Net.....	16
Gambar 14 Perbedaan setiap Arsitektur VGG Net .....	17
Gambar 15 Jumlah Parameter(perjuta) Arsitektur VGG .....	17
Gambar 16 Contoh Data Hand posture .....	21
Gambar 17 Contoh Bobot pada Model .....	28
Gambar 18 Training Akurasi dan Validasi Akurasi Subjek 1.....	30
Gambar 19 Training Loss dan Validasi Loss Subjek 1 .....	30
Gambar 20 Training Akurasi dan Validasi Akurasi Subjek 2.....	31
Gambar 21 Training Loss dan Validasi Loss Subjek 2.....	32
Gambar 22 Training Akurasi dan Loss Subjek 3 .....	33
Gambar 23 Validasi Akurasi dan Validasi Loss Subjek 3 .....	33
Gambar 24 Training Akurasi dan Validasi Akurasi Subjek 4.....	34
Gambar 25 Training Loss dan Validasi Loss Subjek 4.....	34
Gambar 26 Training Akurasi dan Validasi Akurasi Subjek 5.....	35
Gambar 27 Training <i>loss</i> dan Validasi <i>Loss</i> Subjek 5 .....	35

## DAFTAR TABEL

Tabel 1 Confusion Matrix .....	15
Tabel 2 Label Kelas .....	23
Tabel 3 Shape Data .....	25
Tabel 4. Hasil Akhir Akurasi dan Loss .....	36

# BAB I

## PENDAHULUAN

### 1.1 Latar Belakang

Di era teknologi sekarang, perkembangan yang begitu pesat terjadi disemua bidang ilmu pengetahuan termasuk perkembangan teknologi, Perkembangan teknologi berdampak penyebaran informasi begitu mudah dan bisa di akses dimana saja. Informasi berupa visual merupakan informasi yang kini banyak digunakan oleh generasi milenial, terbukti sebanyak 62 % dari generasi milenial menginginkan pencarian Informasi berupa Informasi *Visual* (ViSenze, 2018), ini tidak mengeherankan karena Informasi *Visual* adalah Informasi yang mudah dipahami, penelitian membuktikan 90 % dari Informasi yang dikirimkan ke otak manusia adalah Informasi *Visual* (Anne, 2018).

Informasi *Visual* biasanya berbentuk gambar atau citra, perkembangan pertumbuhan data citra dapat membantu manusia mengenali objek sehingga dapat menyajikan sebuah Informasi secara *Visual*. Namun data citra dapat menjadi sebuah data yang tidak menyajikan sebuah Informasi ketika citra tersebut tidak diketahui objeknya sehingga dalam suatu citra biasanya tidak hanya berupa hanya satu objek namun ada beberapa objek.

Banyak media berita sekarang menggunakan *Hand posture* untuk menyampaikan berita untuk membantu penderita disabilitas tuna rungu. *Hand posture* adalah Bahasa *Isyarat* yang direpresentasikan oleh bentuk jari. Beragam pola jari dapat mewakili huruf yang ada sehingga dapat membantu seseorang untuk menyampaikan Informasi. *Hand Posture* biasa digunakan untuk membantu seseorang yang tidak dapat mendengar (*tuna rungu*) atau tidak dapat berbicara (*tuna wicara*) agar dapat berkomunikasi, menurut WHO pada tahun 2005, terdapat 278 juta orang menderita gangguan pendengaran (Rahayu, Jimmy, & Dannes., 2015). *Hand Posture* memuat bentuk yang mewakili setiap huruf dan angka yang ada, sebanyak 27 huruf abjad dan 10 angka dasar



memiliki bentuk tangan masing-masing. Sehingga banyak bentuk tangan tersebut perlu dilakukan pengenalan objek dari setiap bentuk tersebut sehingga dari bentuk tangan tersebut kita dapat mengetahui maksud Informasi yang diberikan oleh *Hand Posture* tersebut. Sehingga perlu dilakukan penelitian mengenai pengenalan obyek *Hand Posture* dalam hal ini penelitian klasifikasi data citra *Hand Posture*.

Pengenalan obyek dalam citra merupakan hal yang menantang dalam bidang *computer vision*. Bagaimana komputer dapat mengenali objek merupakan suatu hal yang dapat membantu manusia pada saat ini. Dataset *Hand Posture* mempunyai karakteristik yang membedakan terhadap dataset lainnya, seperti gambar yang mencakup berbagai macam tangan dengan menggunakan kondisi pencahayaan yang berbeda.

*Convolution Neural Networks* merupakan metode yang sedang *trend* pada saat ini untuk mengklasifikasikan gambar, dimana telah dilakukan penelitian terkait menggunakan CNN terbukti memiliki akurasi yang tinggi, seperti yang telah dilakukan Simonyan dan Zisserman yang menggunakan arsitektur *VGG-Net* memenangkan kompetisi *Large Scale Visual Recognition Challenge 2014*(ILSVRC2014) dan mendapatkan akurasi 92.7 % (Simonyan & Zisserman, 2015), Dennis dan Bogdan dalam penelitian *Hand posture recognition using Convolutional Neural Networks* menggunakan arsitektur *LeNet* menghasilkan akurasi 97 % (Fernandes & Kwolek, 2017), Adiguna dan Yustinus dalam penelitian *CNN Based Posture Free Hand Detection* menggunakan arsitektur *CUDA-Convnet2structure* menghasilkan akurasi 93.9 % (Adiguna & Soelistio, 2018). Maka dari itu algoritma *Convolutional Neural Networks* merupakan teknik pembelajaran mesin yang berhasil mengklasifikasikan berbagai objek.

Di dalam penelitian ini metode *CNN* dengan arsitektur *VGG19-Net* yang digunakan untuk mengklasifikasikan objek. Difokuskan pada pembuatan model dalam melakukan ekstraksi fitur dan klasifikasi objek penelitian. Sehingga dilakukan penelitian mengenai

implementasi *Convolutional Neural Networks* Dengan Arsitektur *VGG19-Net* pada *Image Depth Data Hand Posture*.

## **1.2 Rumusan Masalah**

Berdasarkan latar belakang tersebut,berikut diberikan rumusan masalah pada penelitian ini,yaitu :

1. Bagaimana implementasi *Convolutional Neural Networks* dengan Arsitektur *VGG-19 Net* pada *Image Depth Data Hand Posture*?
2. Bagaimana tingkat akurasi yang didapatkan dari hasil *Convolutional Neural Networks* dengan Arsitektur *VGG-19 Net* pada *Image Depth Data Hand Posture* ?

## **1.3 Batasan Masalah**

Permasalahan yang dibahas pada penelitian ini memiliki ruang lingkup yang luas, sehingga diberikan batasan-batasan sebagai berikut:

1. Dataset yang digunakan adalah dataset *Hand Posture* .
2. Dataset memiliki 31 kelas.
3. Penelitian ini tidak membahas tentang penyakit yang diderita.
4. Data latih di ambil dari *Fingerspeeling Recognition* dari halaman *github byeongkeun-kang*.

## **1.4 Tujuan Penelitian**

Berdasarkan rumusan masalah, maka tujuan dari penelitian ini adalah:

1. Mengetahui implementasi *Convolutional Neural Networks* dengan Arsitektur *VGG-19 Net* pada *Image Depth Data Hand Posture*.
2. Mengetahui tingkat akurasi yang didapatkan dari *Convolutional Neural Networks* dengan Arsitektur *VGG-19 Net* pada *Image Depth Data Hand Posture*.

3. Mengetahui model dan jumlah filter yang efektif yang digunakan dalam *Convolutional Neural Networks* dengan Arsitektur *VGG-19 Net* pada *Image Depth Data Hand Posture*.

### **1.5 Manfaat Penelitian**

Hasil penelitian ini diharapkan dapat bermanfaat:

1. Sebagai rujukan untuk mengatasi dataset dengan kedalaman gambar yang sangat sering dijumpai.
2. Menjadi sumber informasi mengenai performa *Convolutional Neural Networks*.
3. Menjadi sumber Informasi penggunaan *Convolutional Neural Networks* dengan Arsitektur *VGG-19 Net* pada *Image Depth Data Hand Posture*.

## **BAB II**

### **TINJAUAN PUSTAKA**

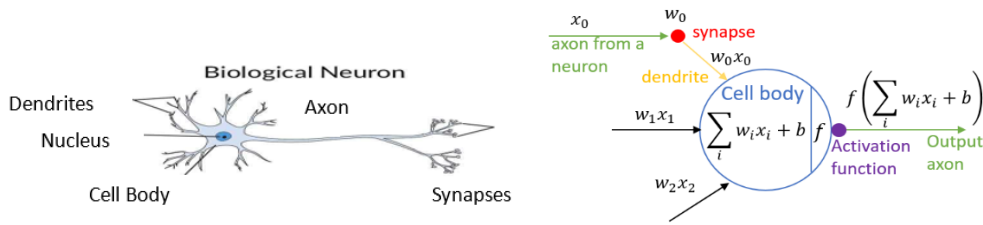
#### **1.6 Deep Learning**

*Deep Learning* adalah bagian dari *machine learning*. *Deep Learning* banyak digunakan untuk kebutuhan analisis data yang mengkombinasikannya dengan kecerdasan buatan. Aplikasi dari penggunaan *deep learning* banyak seperti klasifikasi gambar, mobil tanpa pengemudi, klasifikasi berita dan banyak digunakan oleh perusahaan besar, seperti Google, Microsoft, Facebook, IBM, Baidu, Apple, Adobe, Netflix, NVIDIA dan NEC (Alpaydin E., 2009).

Teknik *Deep Learning* memberikan model yang bagus untuk algoritma *supervised learning*. Dengan menambahkan data latih maka model pembelajaran tersebut dapat mewakili kelas yang lebih baik. Dalam pembuatan algoritma *deep learning*, *artificial intelligence* dapat menguatkan pembelajaran pada data yang dilatih. Interaksi tersebut dapat memberikan umpan balik antara sistem pembelajaran dan pengalaman interaksi yang berguna untuk meningkatkan kinerja dalam tugas yang sedang dipelajari (Hindawi, 2017).

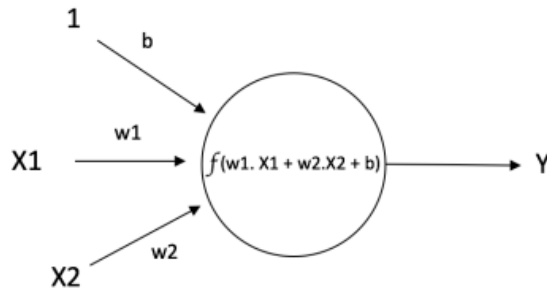
#### **1.7 Neural Networks**

*Neural Networks* atau jaringan saraf merupakan sebuah model dari jaringan saraf otak manusia yang ditiru oleh banyak peneliti untuk diadopsi cara kerjanya di berbagai bidang kajian seperti biologi, fisika, ilmu komputer, dll (Muhammad, 2018). Di bidang ilmu komputer terdapat istilah *Artificial Neural Networks* (ANN). Konsep ANN diadopsi dari jaringan saraf sehingga ANN biasa disebut jaringan saraf tiruan. Model yang dibuat oleh ANN dapat untuk beradaptasi, belajar dan mengelompokkan berbasis pemrosesan paralel.



Gambar 1 Ilustrasi Neural Networks

Adapun beberapa hal yang mendasar pada komputasi *neural Networks* adalah neuron, atau biasa disebut juga dengan node. Node dapat menerima input dari node lainnya atau dari sumber eksternal kemudian dihitung untuk mendapatkan sebuah output. Setiap input memiliki bobot( $w$ ) tersendiri, yang mana bobot ini diberikan dengan dasar hubungannya dengan input lainnya. Node tersebut menggunakan fungsi ke dalam input yang telah diberi bobot. Contoh Ilustrasi neuron dapat dilihat di gambar 2.

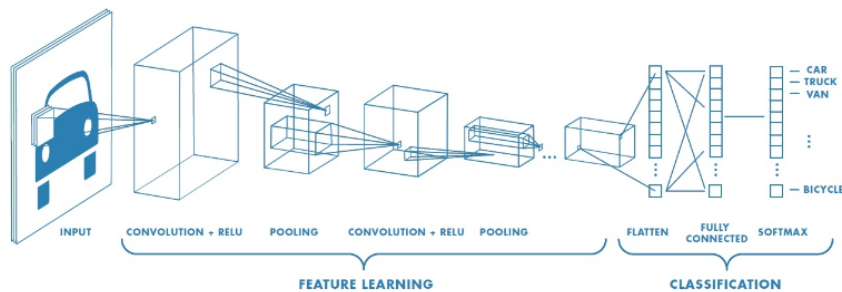


Gambar 2 Ilustrasi Neuron

Sebuah neuron input berupa  $x_1$  dan  $x_2$  dengan bobotnya yaitu  $w_1$  dan  $w_2$ . Selain dari input dan bobot juga terdapat inputan lain berupa 1 dengan bobot  $b$  (bias). Fungsi utama dari bias adalah untuk menyediakan setiap node dengan sebuah nilai constant yang dapat dilatih (sebagai tambahan selain dari input normal yang diterima). Output pada neuron tersebut yaitu  $Y$  yang merupakan hasil perhitungan dari sebuah fungsi non linear  $f$  yang disebut dengan fungsi aktivasi. Tujuan dari sebuah fungsi aktivasi adalah untuk menunjukkan non-linear terhadap output dari neuron tersebut (Muhammad, 2018).

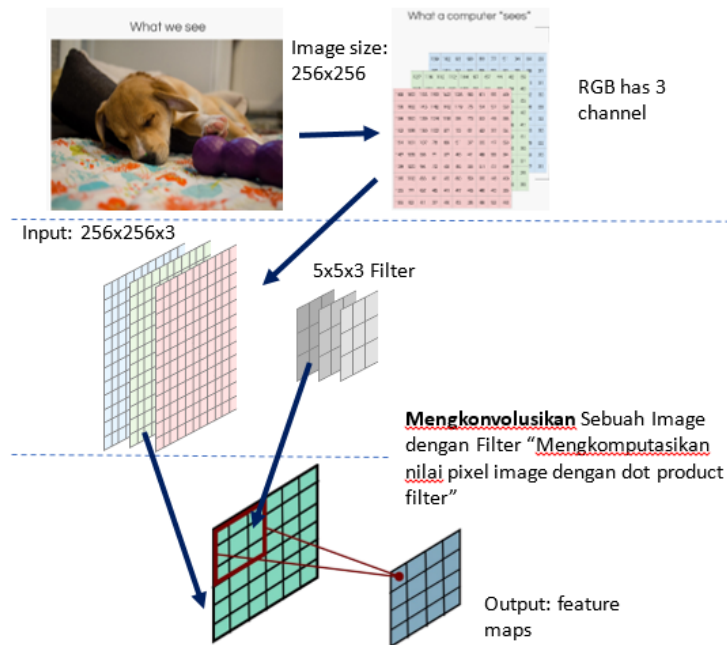
## 1.8 Convolutional Neural Networks

*Convolutional Neural Networks (CNN)* adalah salah satu metode machine learning dari pengembangan *Multi Layer Perceptron (MLP)* yang didesain untuk mengolah data dua dimensi. CNN termasuk dalam jenis *Deep Neural Networks* karena dalamnya tingkat jaringan dan banyak diimplementasikan dalam data citra (Bishop M., 2016). CNN dapat menggunakan operasi konvolusi sebagai dasar dari algoritma nya. Nama konvolusi sendiri merupakan operasi aljabar linear yang mengalikan matriks dari filter pada citra yang akan di proses, sehingga proses ini disebut lapisan konvolusi dan merupakan salah satu jenis layer yang dimiliki oleh CNN. *Convolutional Neural Networks* input yang berupa kumpulan pixel yang berupa array. Gambaran umum *Convolutional Neural Networks* dapat dilihat di gambar 3 :



Gambar 3 Gambaran Umum Proses CNN

Lapisan lapisan pada gambar merupakan lapisan yang paling umum digunakan dalam arsitektur *Convolutional Neural Networks*. Lapisan lapisan ini melakukan operasi yang mengubah data berdasarkan lapisan yang telah dilalui dengan maksud mempelajari fitur-fitur khusus pada data. Contoh Ilustrasi CNN pada sebuah image dapat dilihat pada gambar 4.



Gambar 4 Ilustrasi penerapan CNN pada sebuah Image

Ada beberapa lapisan yang dilalui sebuah image pada arsitektur *convolutional neural networks*, lapisan yang paling umum digunakan seperti lapisan konvolusi, lapisan fungsi aktivasi, lapisan *pooling* dan *fully connected layer*.

### 1.8.1 Lapisan Konvolusi

Lapisan Konvolusi menempatkan input yang berupa gambar melalui serangkaian filter konvolusional, yang masing-masing mengaktifkan fitur tertentu dari gambar (Mathworks, n.d.). Operasi Konvolusi biasanya dinotasikan dengan Persamaan 1 :

$$C(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n). \quad (1)$$

Dimana :

$C(i, j)$  = Matriks hasil konvolusi

$I$  = Template Gambar

$K$  = Kernel

- m = ukuran baris template
- n = ukuran kolom template
- i = ukuran baris kernel
- j = ukuran kolom kernel
- \* = notasi untuk konvolusi

Proses ini dilakukan dengan menggeser filter di gambar. Setiap filter memiliki bobot dan nilai bias yang sama di seluruh gambar selama proses ini. Ini disebut pembagian bobot dan memberikan kemampuan untuk mewakili fitur yang sama pada seluruh gambar.

Pada lapisan konvolusi, dapat di modifikasi sesuai kebutuhan peneliti. Pada saat input masuk ke lapisan konvolusi input memiliki beberapa parameter yaitu Width(W), Height(H) dan D (Depth) serta yang menjadi inti berupa nilai dari setiap pixel input tersebut. Untuk mengatur sebuah ukuran output atau feature map dari lapisan konvolusi, parameter yang mempengaruhi output tersebut biasanya dituliskan dengan persamaan 2 :

$$FM = \frac{W-F+2P}{S} + 1 \quad (2)$$

Dimana :

- FM = Feature Map
- W = Volume Size
- F = Ukuran Filter
- P = Zero Padding
- S = Stride

Nilai 1 diambil dari nilai Bias.

Parameter diatas harus diatur untuk memenuhi output yang di inginkan, contoh :



Sebuah input yang berukuran 5 x 5 x 3 menggunakan zero padding 2 dikonvolusikan dengan Filter 3 x 3 dan menggunakan Stride 2 maka output dari layer tersebut adalah

$$\text{Output Layer} = (W - F + 2P) / s + 1 \text{ maka } (5 - 3 + 2(2)) / 2 + 1 = 4$$

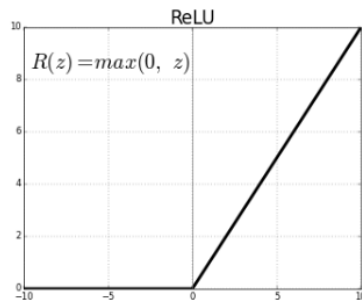
Jika kasus diganti dengan Kasus yang sama namun Stride yang digunakan adalah 3 dan padding hanya 1 maka filter ini tidak dapat di konvolusikan ke input. Perhatikan, sesuai rumus maka  $\text{Output Layer} = (5 - 3 + 2(1)) / 3 + 1$  dimana  $(5 - 3 + 2) / 3$  hasilnya bukan merupakan sebuah bilangan integer.

### 1.8.2 Fungsi Aktivasi ReLU

Fungsi Aktivasi merupakan layer yang memungkinkan pelatihan lebih cepat dan lebih efektif dengan memetakan nilai berdasarkan fungsi aktivasi yang digunakan. Fungsi aktivasi yang sering digunakan adalah fungsi aktivasi *ReLU* atau *Rectified linear unit* yaitu fungsi aktivasi yang memetakan nilai negative menjadi nol dan mempertahankan nilai positif. Fungsi *ReLU* didefinisikan dengan fungsi :

$$f(y) = \begin{cases} y, & \text{jika } y > 0 \\ 0, & \text{jika } y < 0 \end{cases} \quad (3)$$

Inputan  $y$  adalah inputan dari aktivasi nonlinear  $f$  pada semua fitur input , sehingga persamaan (3) sama dengan  $f(y) = \max\{0, y\}$ .



Gambar 5 ReLU

Contoh dari proses aktivasi *ReLU* dapat dilihat di gambar 6.

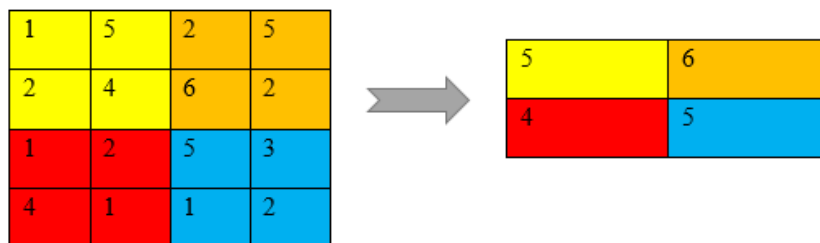
9	2	-5	2
9	-8	5	-3
-1	2	4	5
1	5	-4	6

9	2	0	2
9	0	5	0
0	2	4	5
1	5	0	6

Gambar 6 Contoh Fungsi Aktivasi ReLU

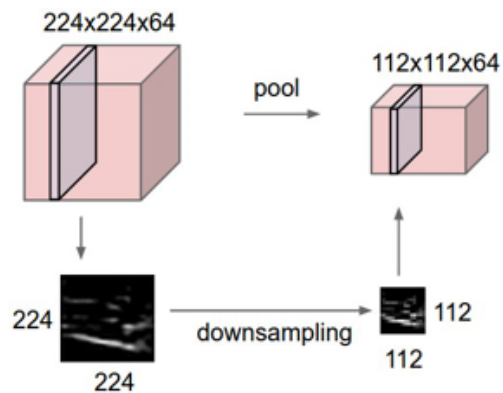
### 1.8.3 Pooling

*Pooling* merupakan layer untuk menyederhanakan output yang sebelumnya telah di convolusikan. Prinsipnya menggunakan sebuah filter dengan ukuran dan stride tertentu yang akan bergeser pada seluruh area feature map. *Pooling* yang biasa digunakan adalah *Max Pooling* dan *Average Pooling*, namun dalam penelitian ini akan menggunakan operasi *Max Pooling*. Contoh operasi max pooling dapat dilihat di gambar 7.



Gambar 7 Contoh Operasi Max Pooling

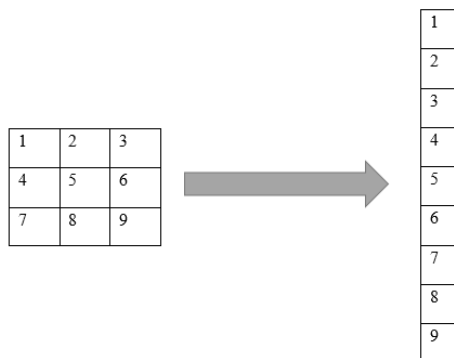
Tujuan dari *pooling layer* adalah mengurangi dimensi dari *feature map*, sehingga mempercepat komputasi karena parameter yang harus di update semakin sedikit dan dapat mengatasi *overfitting*, pooling layer juga mengakibatkan proses downsampling terjadi, contoh proses downsampling dapat dilihat seperti digambar 8.



Gambar 8 Downsampling

#### 1.8.4 Fully Connected Layer

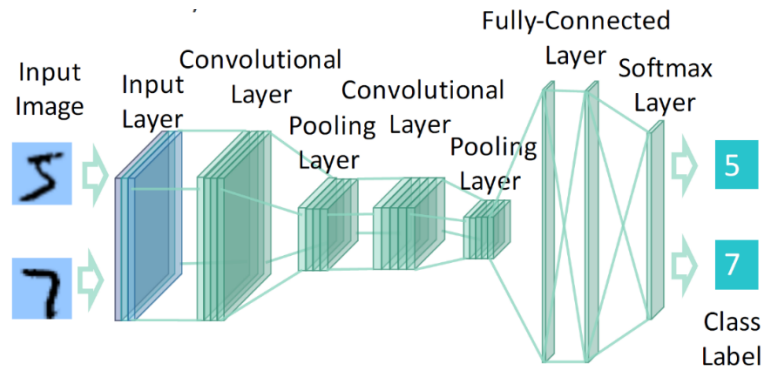
*Fully Connected Layer* menerima input sebuah vektor dari layer sebelumnya, sehingga pada saat feature map hasil yang berbentuk multidimensional *array* dilakukan *flatten* sehingga input menjadi sebuah *vector*. Proses *flatten* dapat dilihat pada gambar 9.



Gambar 9 Proses Flatten

*Fully-Connected Layer* (FC Layer) adalah layer yang biasa digunakan dalam penerapan *multilayer perceptron*(MLP), memiliki *hidden layer*, *activation function*, *output layer* dan *loss function/activation function*. *FC layer* bertujuan untuk melakukan transformasi pada dimensi data agar dapat diklasifikasikan secara linier (Basha, Dubai,

Pulabaigari, & Mukherjee, 2019). Secara umum Full Connected Layer terletak sebelum layer penentuan kelas Klasifikasi , hal ini dapat dilihat di arsitektur CNN di gambar 10.



Gambar 10 Arsitektur CNN

Lapisan *fully connected* menghubungkan semua *neuron* dari lapisan sebelumnya dengan neuron di lapisan selanjutnya. Setiap aktivitas pada layer sebelumnya dihubungkan sehingga terdapat algoritma *backpropagation* yang bekerja untuk memperbarui bobot berdasarkan error yang telah didapat sebelumnya. Kunci keberhasilan CNN terletak pada hubungan neuron yang mengurangi jumlah bobot dan membuat jaringan lebih mudah untuk di optimalkan.

### 1.9 Softmax Loss Function

Softmax Layer adalah layer yang digunakan untuk menentukan kelas prediksi dari proses *training* suatu *Networks*. Output dari layer ini sebenarnya digunakan untuk menentukan *cross entropy loss*. Persamaan *Cross Entropy loss* di definisikan dengan persamaan 4.

$$H(p, q) = - \sum_x p(x) \log[q(x)] \quad (4)$$

Dimana :

$H(p, q)$  = *Cross Entropy Loss*

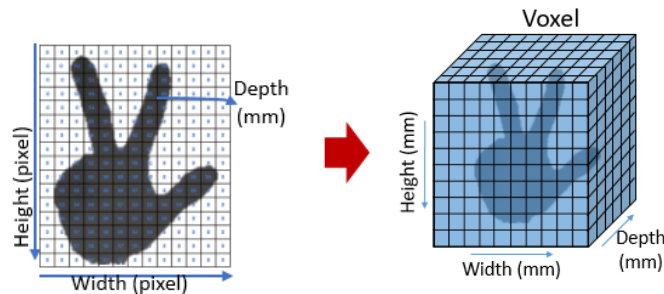
$p(x)$  = Label kelas sebenarnya

$q(x)$  = Output prediksi dari Networksnya

Dimana fungsi utama *softmax* berfungsi untuk mengubah vektor ke bentuk probabilitas output yang bernilai antara 0 dan 1. Setelah nilai *softmax* didapatkan akan dihitung *cross-entropy loss* dengan menggunakan persamaan (4), sehingga dari persamaan tersebut dapat dihasilkan nilai error yang diperoleh yang dapat digunakan dalam proses *back-propagation*.

### 1.10 Image Depth Data

*Image Depth Data* adalah suatu data set yang memiliki kedalaman bit yang mengacu pada Informasi warna pada suatu gambar yang berbeda beda tiap gambarnya. Semakin tinggi kedalaman bit suatu gambar, semakin banyak warna yang dapat di simpan. Gambar paling sederhana adalah gambar 1 bit , yang hanya dapat menampilkan dua warna, yaitu hitam dan putih. Gambar 8 bit dapat menyimpan 256 warna sedangkan gambar 24 bit dapat menampilkan lebih dari 16 juta warna. Gambaran mengenai kedalaman suatu gambar dapat dilihat di gambar 11.



Gambar 11 Contoh penggambaran kedalaman data

### 1.11 Confusion Matrix

Mengukur kinerja atau performa suatu model adalah parameter baik atau bagusnya sebuah model. Model diperoleh dari pelatihan data yang telah *Confusion Matrix* adalah sebuah tabel yang memiliki Informasi mengenai hasil klasifikasi yang dikerjakan. *Confusion matrix* merupakan metode untuk menganalisis seberapa baik kinerja sebuah

model yang telah dibuat dalam mengidentifikasi data dari kelas yang berbeda-beda (Muhammad, 2018). Contoh tabel *Confusion Matrix* dapat dilihat di tabel 1:

Tabel 1 *Confusion Matrix*

	<i>Class 1 Predicted</i>	<i>Class 2 Predicted</i>
<i>Class Actual 1</i>	<i>True Positif</i>	<i>False Negative</i>
<i>Class Actual 2</i>	<i>False Positif</i>	<i>True Negative</i>

Class 1 menyatakan kelas *positive* sedangkan class 2 menyatakan kelas yang *negative*. Hasil dari tabel *confusion matrix* dapat menghitung akurasi model yang telah dibuat dengan persamaan berikut.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Dimana:

$TP = True\ Positif$

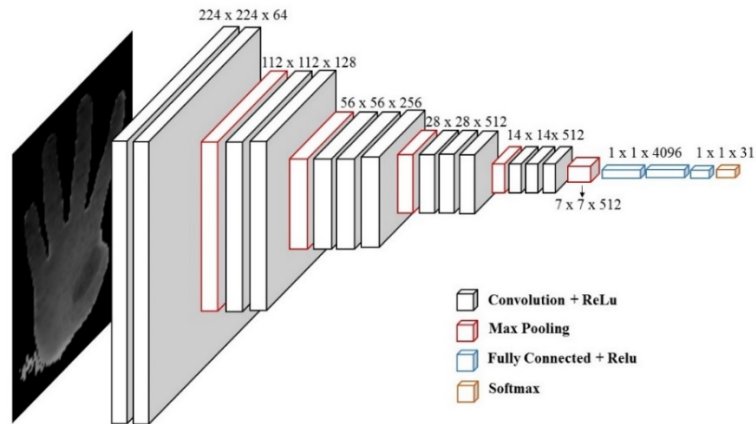
$TN = True\ Negative$

$FP = False\ Positive$

$FN = False\ Negative$

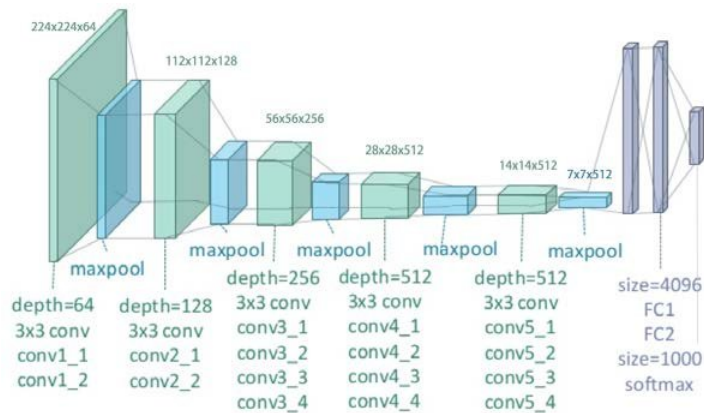
### 1.12 **Arsitektur VGG Net**

Arsitektur jaringan Visual Geometry Group (*VGG Networks*) diperkenalkan oleh *Simonyan* dan *Zisserman* dalam papernya (Simonyan & Zisserman, 2015). Arsitektur *VGG Net* memiliki kemampuan kinerja akurasi yang tinggi. Arsitektur yang dibuat oleh Visual Geometry Group ini memiliki 6 type Arsitektur. Arsitektur tersebut memiliki lapisan konvolusi dan *pooling* yang berulang kali . contoh arsitektur *VGG Net* dapat dilihat di gambar 12.



Gambar 12 Contoh arsitektur vgg-net

Pada gambar 12 merupakan tampak arsitektur VGG-16 , namun pada penelitian ini menggunakan arsitektur VGG-19 Net , dimana Arsitektur VGG-19 Net mirip dengan arsitektur VGG-16 Net, perbedaannya layer yang digunakan oleh arsitektur VGG-19 Net lebih banyak dibanding VGG-16 Net, dimana layer yang digunakan pada arsitektur VGG-19 Net berjumlah 19 layer, yaitu 16 layer konvolusi dan 3 *Fully Connected Layer* sedangkan VGG-16 Net hanya 13 layer konvolusi dan 3 *Fully Connected Layer*. Adapun Arsitektur VGG-19 Net dapat dilihat pada gambar 13.



Gambar 13 Arsitektur VGG-19 Net

Layer pada arsitektur VGG merupakan layer yang berurutan dimana data yang akan ditrain melewati layer konvolusi secara berurutan, sedangkan setiap 1 blok layer

konvolusi berisi lapisan fungsi aktivasi *ReLU* dan pooling *Max Pooling* . kernel yang digunakan pada saat konvolusi berukuran 3 x 3 dan 2 x 2 untuk *pooling* dengan stride 2. Adapun perbedaan jumlah layer pada setiap arsitektur VGG Net dapat dilihat pada gambar 14 :

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 <b>LRN</b>	conv3-64 <b>conv3-64</b>	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 <b>conv3-128</b>	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 <b>conv1-256</b>	conv3-256 conv3-256 <b>conv3-256</b>	conv3-256 conv3-256 conv3-256 <b>conv3-256</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 <b>conv1-512</b>	conv3-512 conv3-512 <b>conv3-512</b>	conv3-512 conv3-512 conv3-512 <b>conv3-512</b>
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Gambar 14 Perbedaan setiap Arsitektur VGG Net

Pada Gambar 14 terdapat 6 jenis arsitektur VGG Net, setiap jenis arsitektur tersebut memiliki jumlah parameter yang berbeda beda, dimana arsitektur VGG-19 Net merupakan arsitektur yang paling banyak memiliki parameter yaitu sekitar 149 juta, adapun jumlah parameter setiap arsitektur VGG Net dapat dilihat pada gambar 15 :

Network	A,A-LRN	B	C	D	E
Number of parameters	133	133	134	138	144

Gambar 15 Jumlah Parameter(perjuta) Arsitektur VGG

### 1.13 Google Colaboratory

*Google Colaboratory* atau *Google Colab* adalah sebuah *tools* yang dibuat oleh *google* untuk pengolahan data atau machine learning untuk memberikan fasilitas



komputasi yang begitu cepat. Google menyediakan layanan GPU gratis sebagai backend komputasi yang dapat digunakan selama 12 jam pada setiap session. *Google Colab* dibuat berdasarkan *Environment Jupyter* sehingga mirip dengan *Jupyter Notebook*.