



PERBANDINGAN CONVOLUTIONAL NEURAL NETWORK DAN VISION TRANSFORMER PADA KLASIFIKASI KUE TRADISIONAL INDONESIA

DOMAIN — ● ● ●

DS

NLP

CV



2022

Computer Vision

TEAM JUPYTER XXI & CORDOBA



**Halomoan Filipus
Simarmata**



**Nyayu Chika
Marselina**



**Athiya Shinta
Wulandari**



**Sukma
Imelda**



**Wella Novita
Andriani**

DAFTAR ISI



JUDUL



ALGORITMA



DOMAIN AI



EVALUASI MODEL



DATASET



BASIS AKHIR
PROJECT





DOMAIN ARTIFICIAL INTELLIGENCE



COMPUTER VISION





Dataset yang dipakai :

[https://www.kaggle.com/datasets/ilhamfp
31/kue-indonesia.](https://www.kaggle.com/datasets/ilhamfp/31/kue-indonesia)



TOTAL DATA

Dataset tersebut terdiri dari 1845 data citra dengan 8 kelas, yakni kue klepon, kue lumpur, kue kastengel, kue putri salju, kue serabi, kue dadar gulung, kue lapis, dan kue risoles.



Jenis Kue Tradisional	Kelas	
Kue Klepon	240 Data	
Kue Lumpur	238 Data	
Kue Kastengel	223 Data	
Kue Putri Salju	214 Data	
Kue Serabi	221 Data	KELAS
Kue Lapis	241 Data	2022
Kue Risoles	236 Data	
Kue Dadar gulung	232 Data	Computer Vision

Jumlah Data Kelas



ALGORITMA





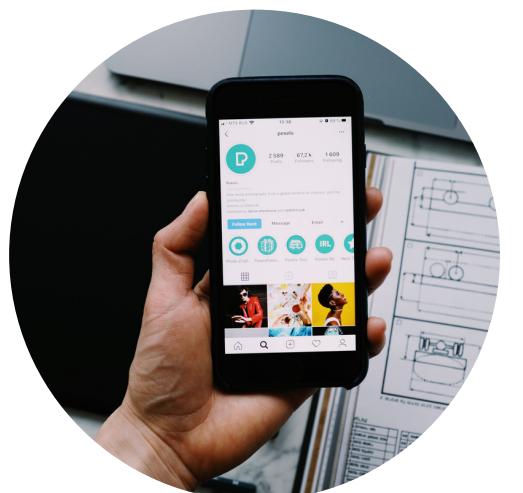
CONVOLUTIONAL NEURAL NETWORK



Convolutional Neural Network merupakan salah satu jenis algoritma Deep Learning yang dapat menerima input berupa gambar, menentukan aspek atau obyek apa saja dalam sebuah gambar yang bisa digunakan mesin untuk “belajar” mengenali gambar, dan membedakan antara satu gambar dengan yang lainnya.



VGG16



RESNET50

ARSITEKTUR
CNN

2022



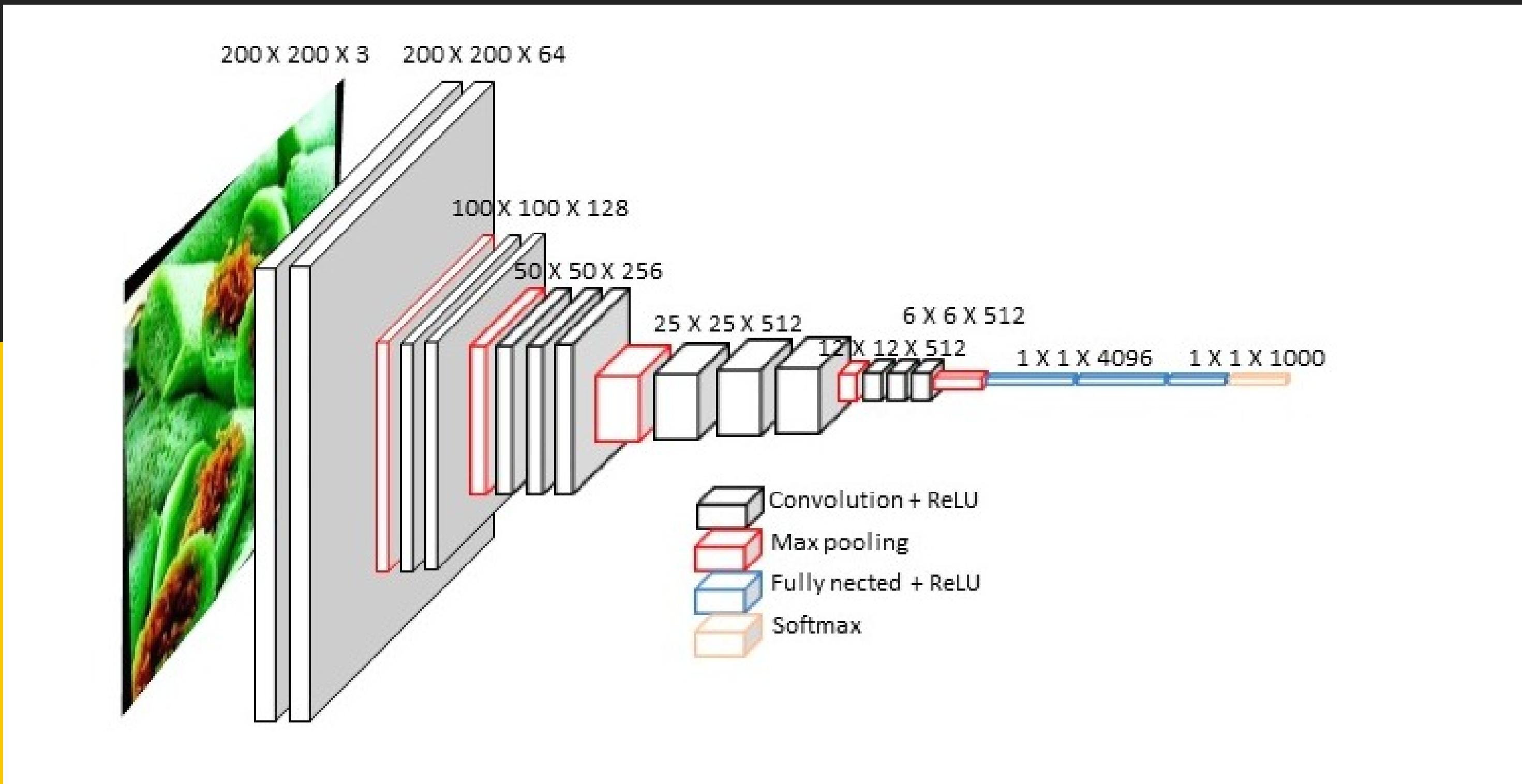
Computer Vision —

VGG16

VGG16 merupakan model CNN yang memanfaatkan convolutional layer dengan spesifikasi convolutional filter yang kecil (3×3). Dengan ukuran convolutional filter tersebut, kedalaman neural network dapat ditambah dengan lebih banyak lagi convolutional layer.



ARSITEKTUR VGG16



ResNet50



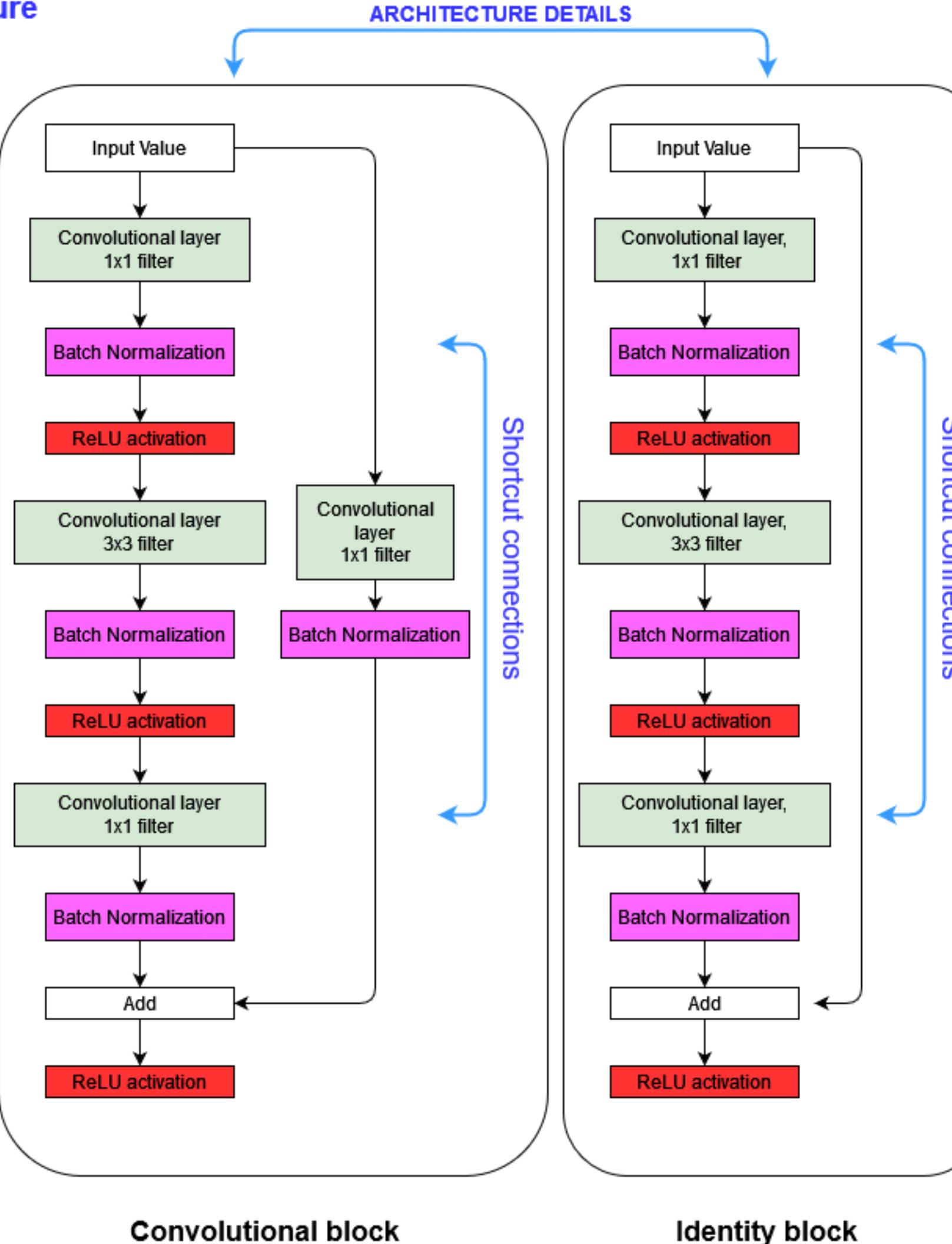
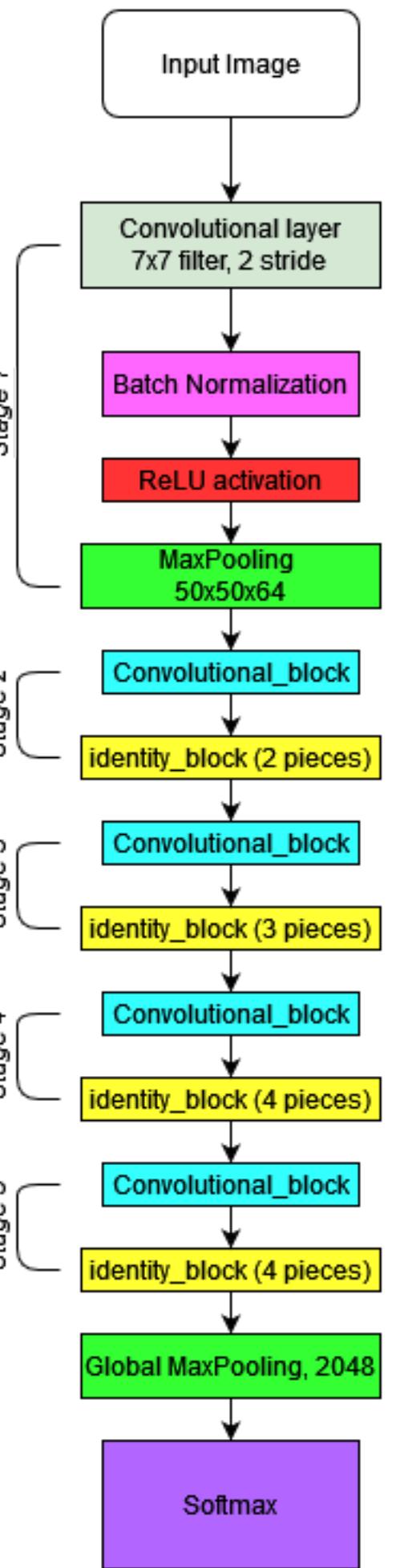
Model ResNet merupakan model yang menggunakan deep residual learning framework. Dengan menggunakan framework ini, setiap layer network memiliki referensi ke layer network sebelumnya, hal ini menjadikan proses optimasi menjadi lebih mudah daripada lapisan-lapisan network yang tidak memiliki keterhubungan. Karena proses optimasi lebih mudah, neural network yang dibentuk dapat memiliki jumlah layer yang banyak sampai dengan 34 layer dan akibatnya, akurasi meningkat dari neural network.



Arsitektur ResNet-50



ResNet-50 Architecture



VISION TRANSFORMER

Kelebihan dari ViT adalah memiliki ukuran model yang lebih efisien dibandingkan konvolusi konvensional dan dapat memprediksi beresolusi tinggi.



HYPERPARAMETER ARSITEKTUR VIT

Hyperparameter	Arsitektur A	Arsitektur B
learning_rate	0,001	0,0001
weight_decay	0,0001	0,00001
batch_size	32	64
image_size	72	144
patch_size	6	12
projection_dim	64	128
num_heads	4	8



EVALUASI MODEL

Confusion matrix adalah tabel yang menyatakan klasifikasi jumlah data uji yang benar dan jumlah data uji yang salah.

$$Akurasi = \frac{TP+TN}{TP+TN+FP+FN} \times 100\%$$

$$Presisi = \frac{TP}{TP+FN} \times 100\%$$

$$Recall = \frac{TP}{TP+FN} \times 100\%$$

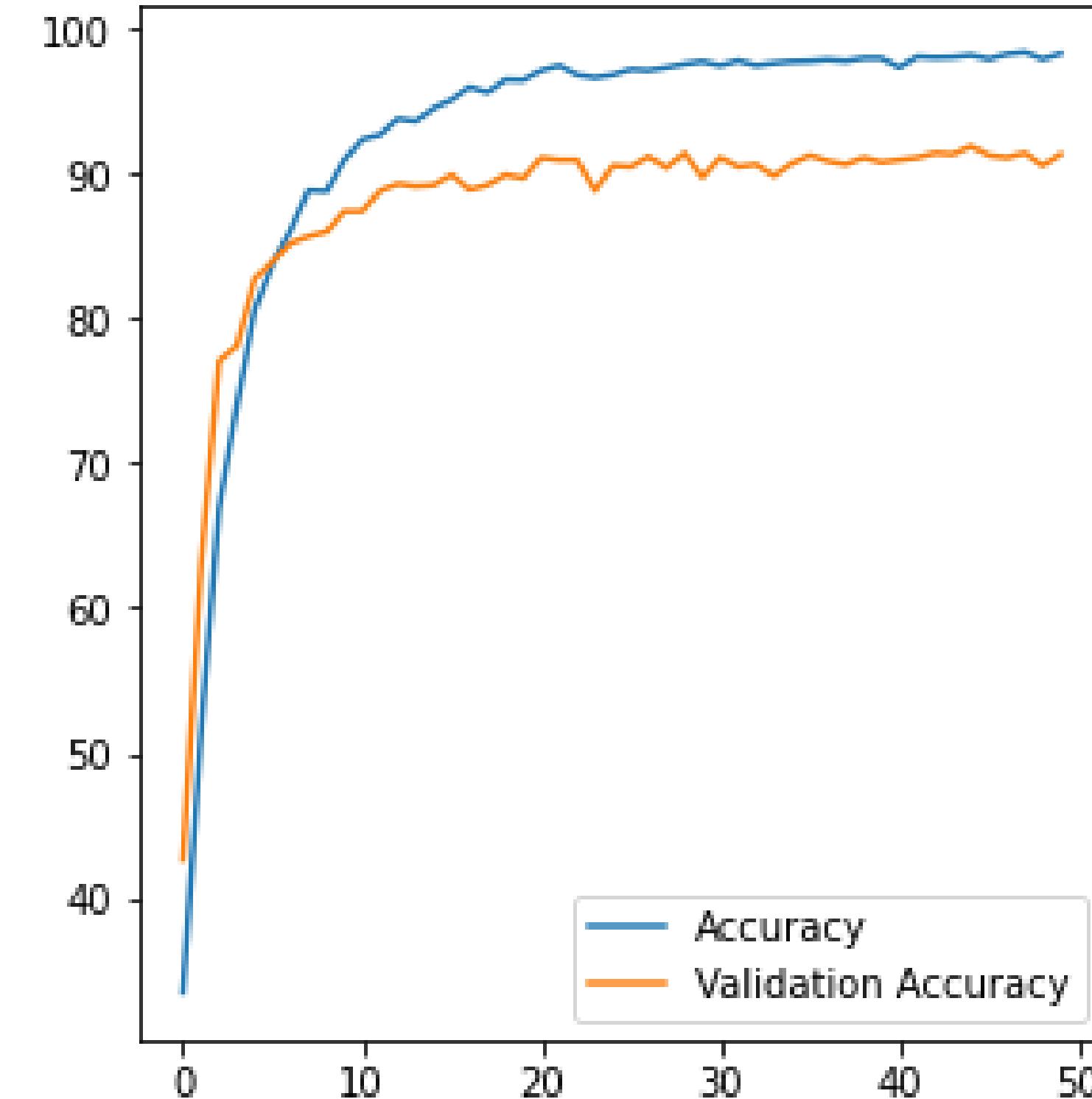
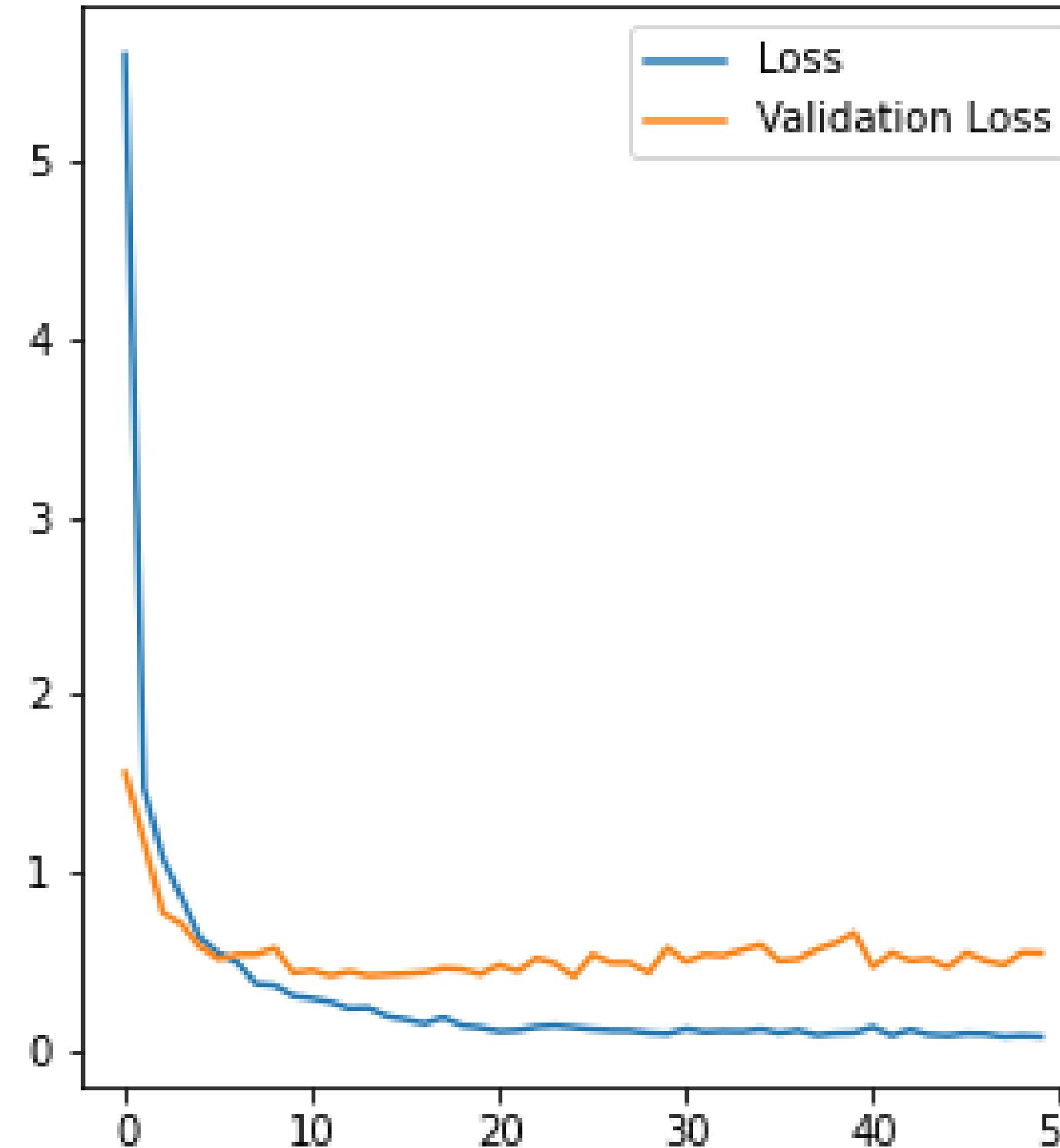
$$F1 - Score = 2 \times \frac{Presisi \times Recall}{Presisi + Recall} \times 100\%$$



2022

Computer Vision

ARSITEKTUR VGG16

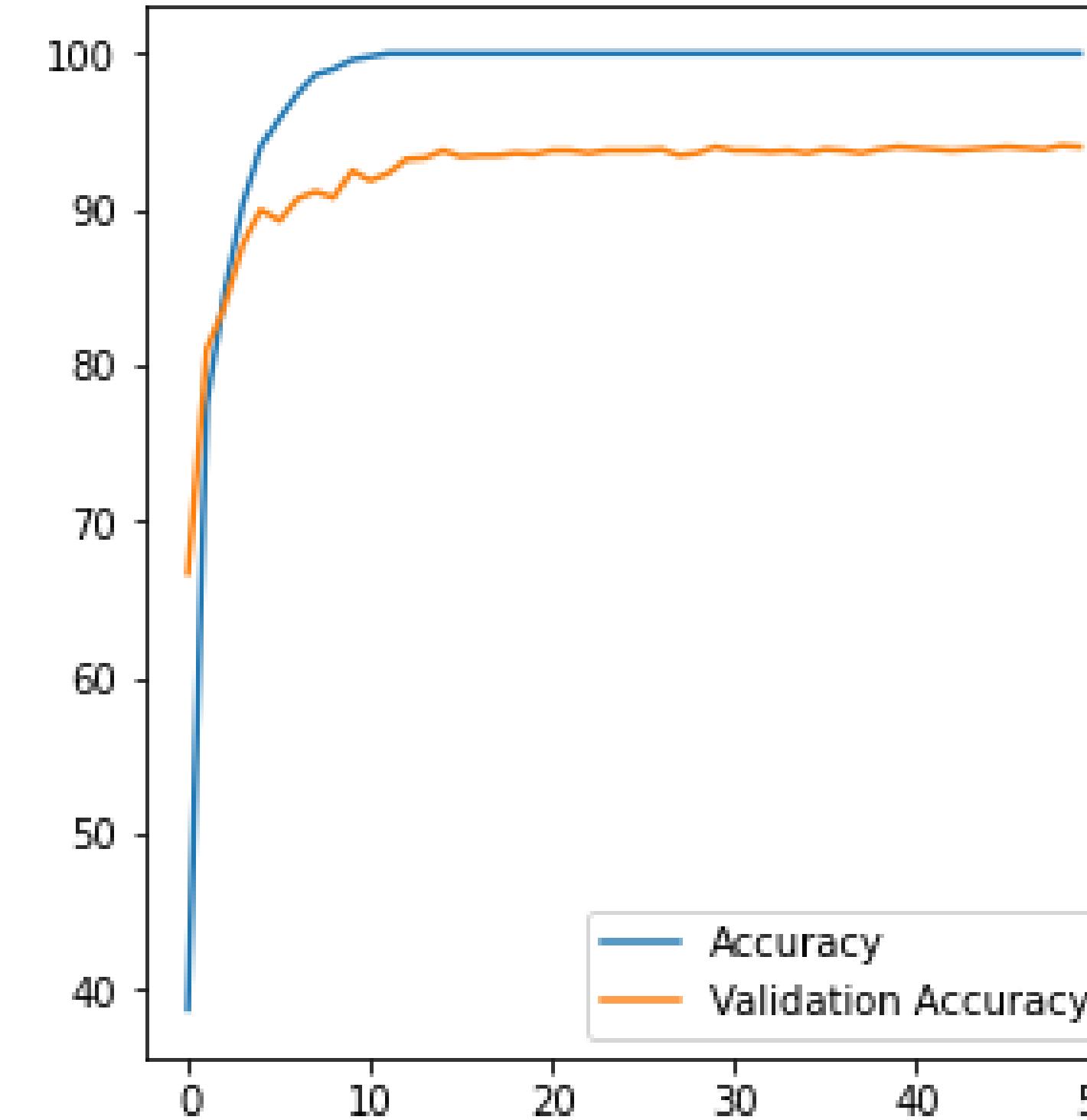
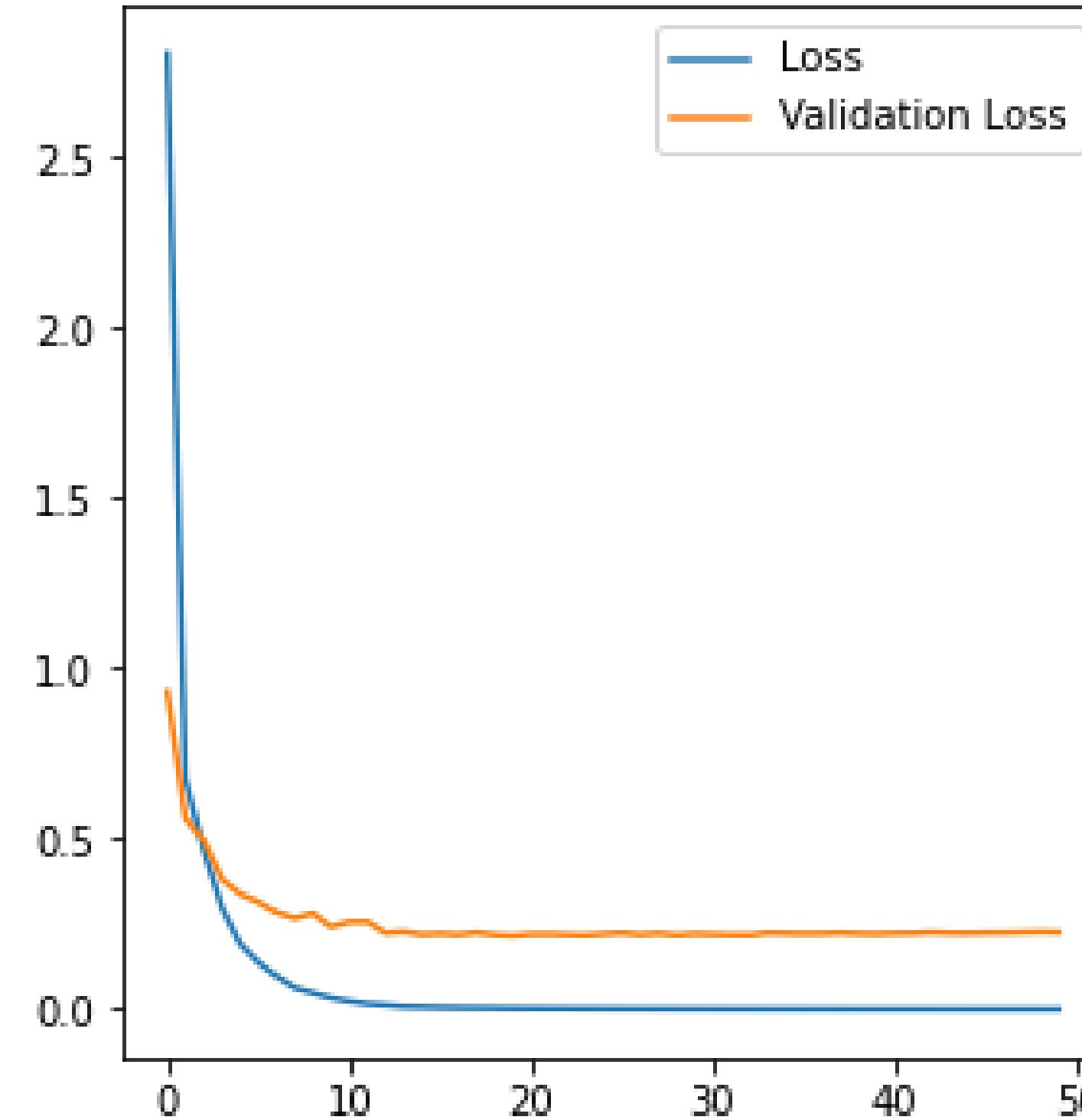




2022

Computer Vision

ARSITEKTUR RESNET50

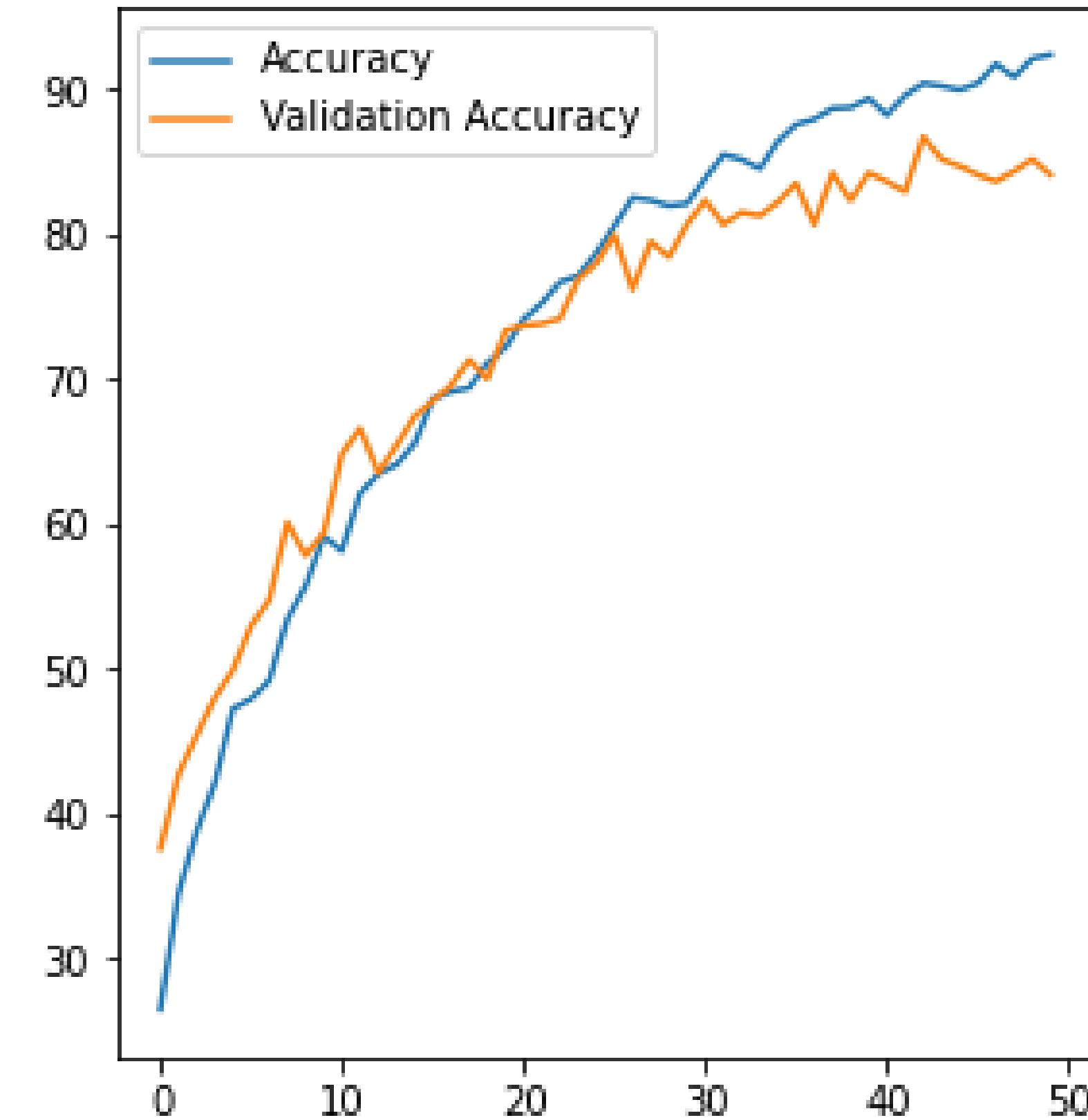
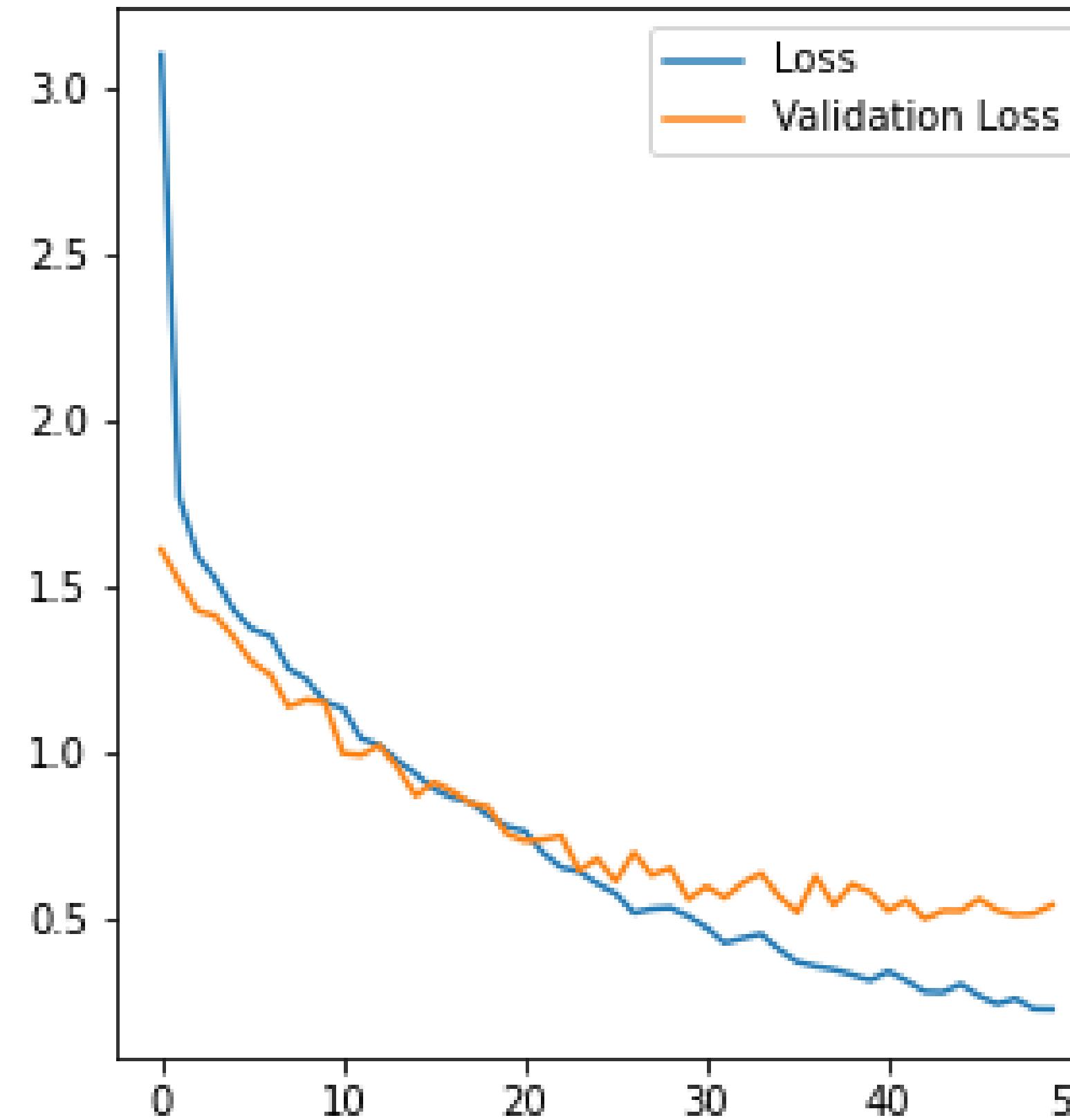




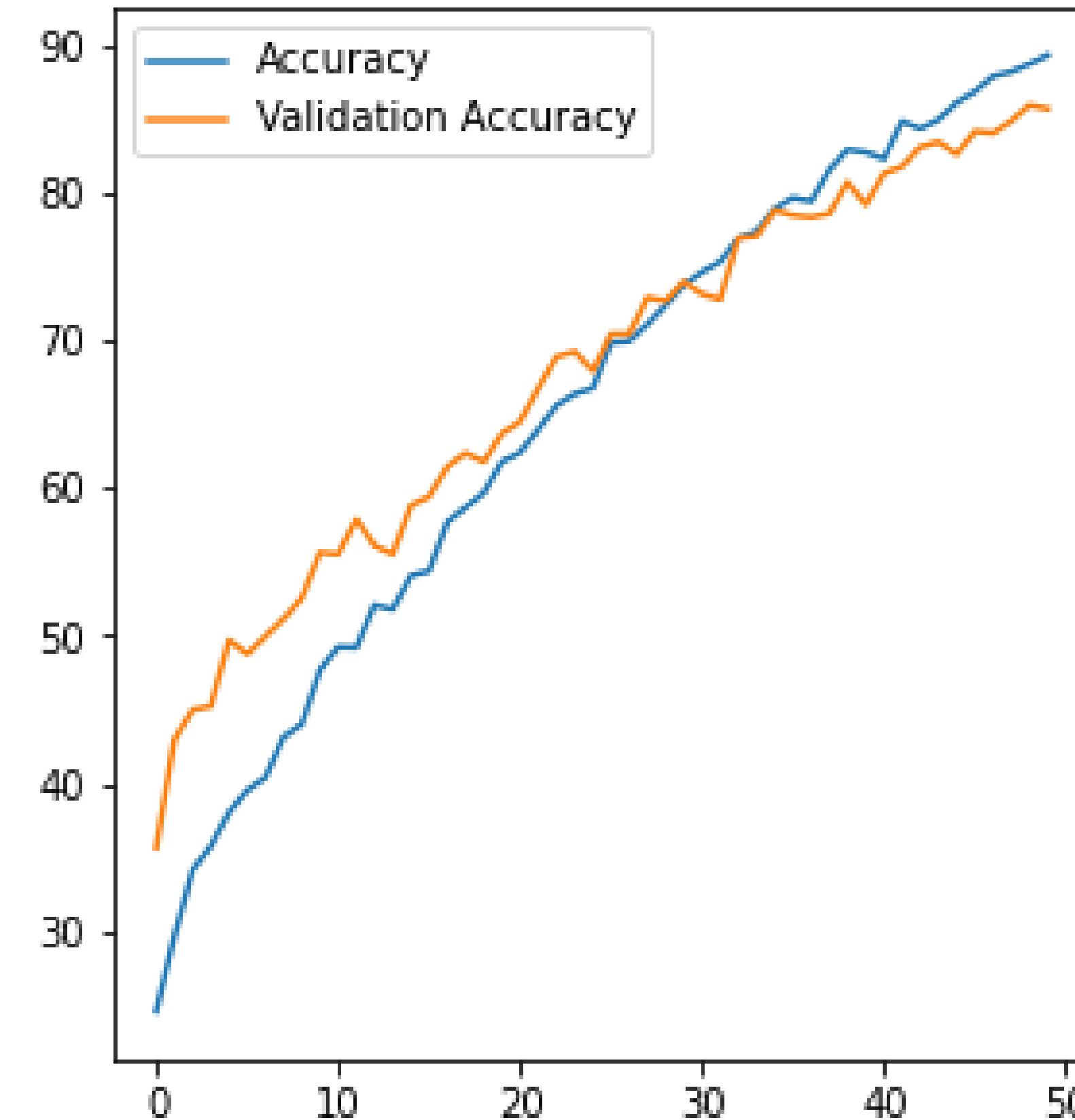
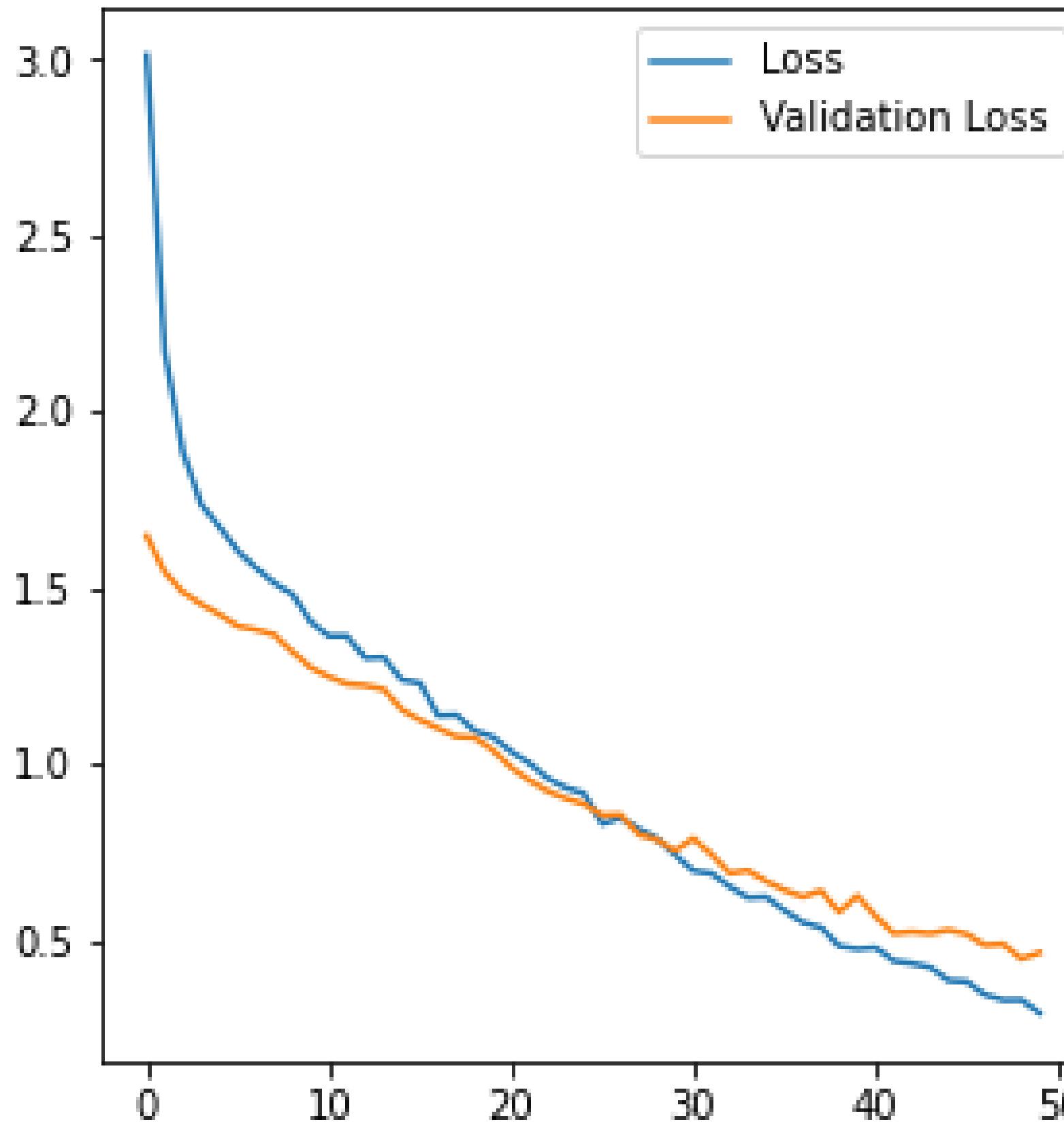
ARSITEKTUR VIT A

2022

Computer Vision



ARSITEKTUR VIT B



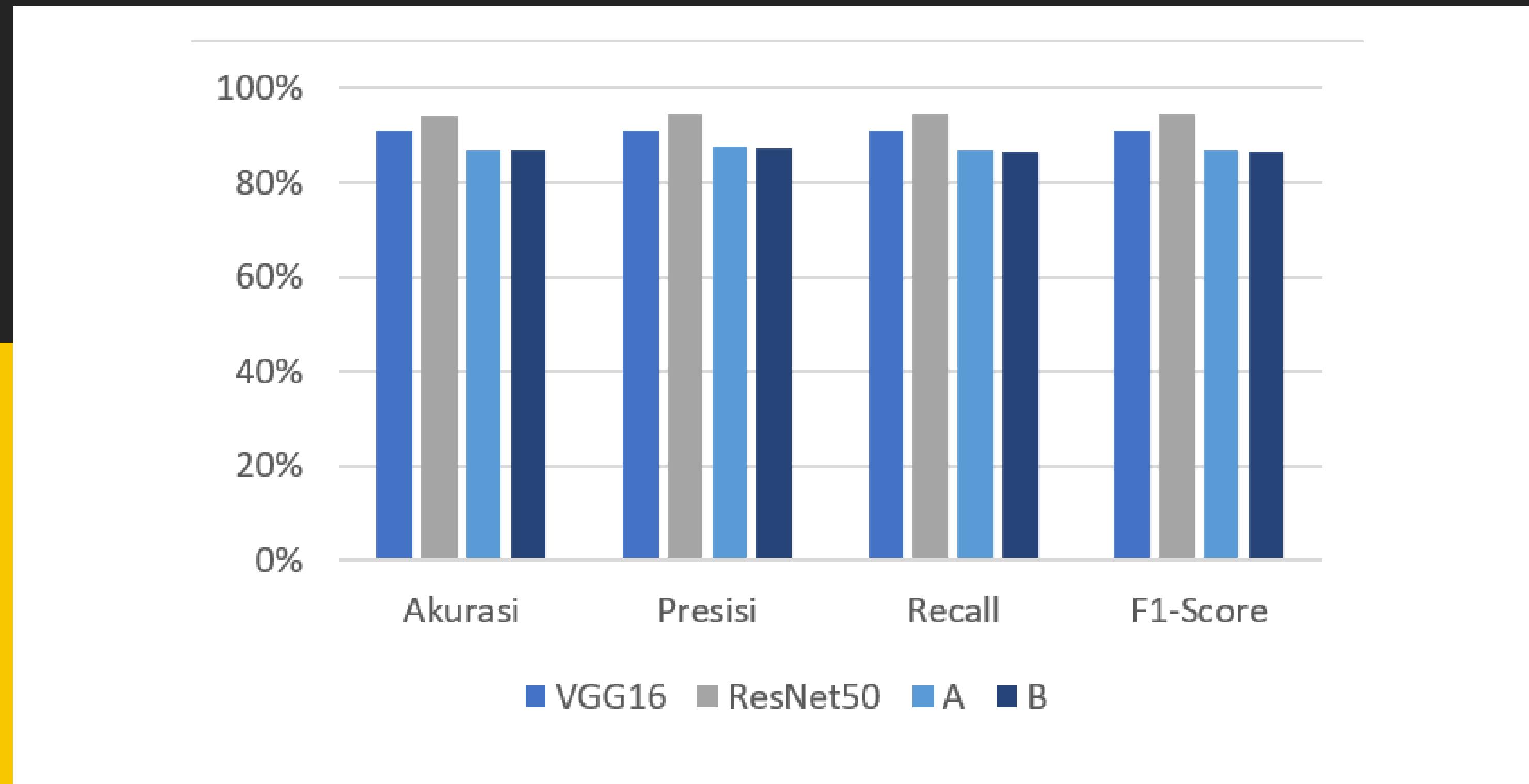
2022

Computer Vision

HASIL MODEL

Algoritma	Arsitektur	Akurasi	Presisi	Recall	F1-Score
CNN	VGG16	91%	91,13%	91,13%	90,88%
	ResNet50	94%	94,38%	94,5%	94,38%
ViT	A	87%	87,63%	86,75%	86,75%
	B	87%	87,13%	86,5%	86,63%

HASIL MODEL





2022

Computer Vision

PENGERJAAN PROJECT AKHIR YANG TIM KAMI PILIH ADALAH PAPER BASED DALAM BENTUK JURNAL, BERIKUT LINK GITHUB KAMI LAMPIRKAN:

[HTTPS://GITHUB.COM/MOANFS/JUCOR
_KUE.ID.GIT](https://github.com/moanfs/jucor_kue.id.git)

BASIS PROJECT AKHIR





Thank You



2022

Computer Vision

