



Presentasi Skripsi

Pengenalan Emosi Manusia Menggunakan *Log-Gabor Convolutional Networks* Melalui Pendekatan *Facial Region Segmentation*

Naufan Rusyda Faikar 1607645

Pendahuluan



Manfaat pengenalan emosi

Latar Belakang Alinea 1

Pengenalan emosi manusia sangat bermanfaat di sektor-sektor terkait komputasi afektif.

Komputasi afektif adalah studi multidisipliner pada pengembangan sistem yang mampu mengenali, memahami dan menghasilkan emosi manusia.



Dalam *healthcare*

- Kapan seorang pasien
membutuhkan perhatian
tambahan? #sakit #sedih
- Bagaimana kepuasan
pasien terhadap pelayanan
dan fasilitas yang diberikan?

<https://sightcorp.com/blog/how-facial-recognition-is-used-in-healthcare/>



Dalam pendidikan di kelas nyata atau pembelajaran jarak jauh

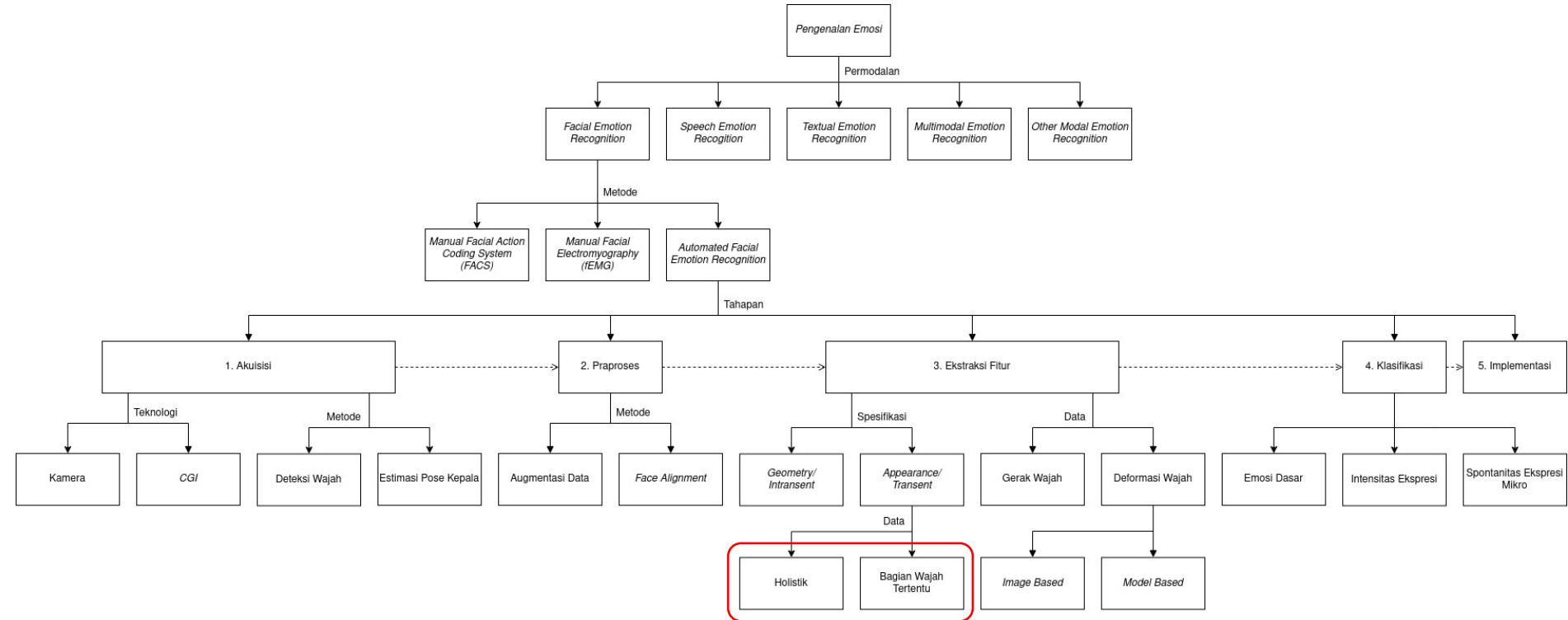
- Kapan seorang studen
membutuhkan penjelasan
tambahan? #bingung
- Bagaimana reaksi studen
terhadap pendekatan
belajar yang baru?

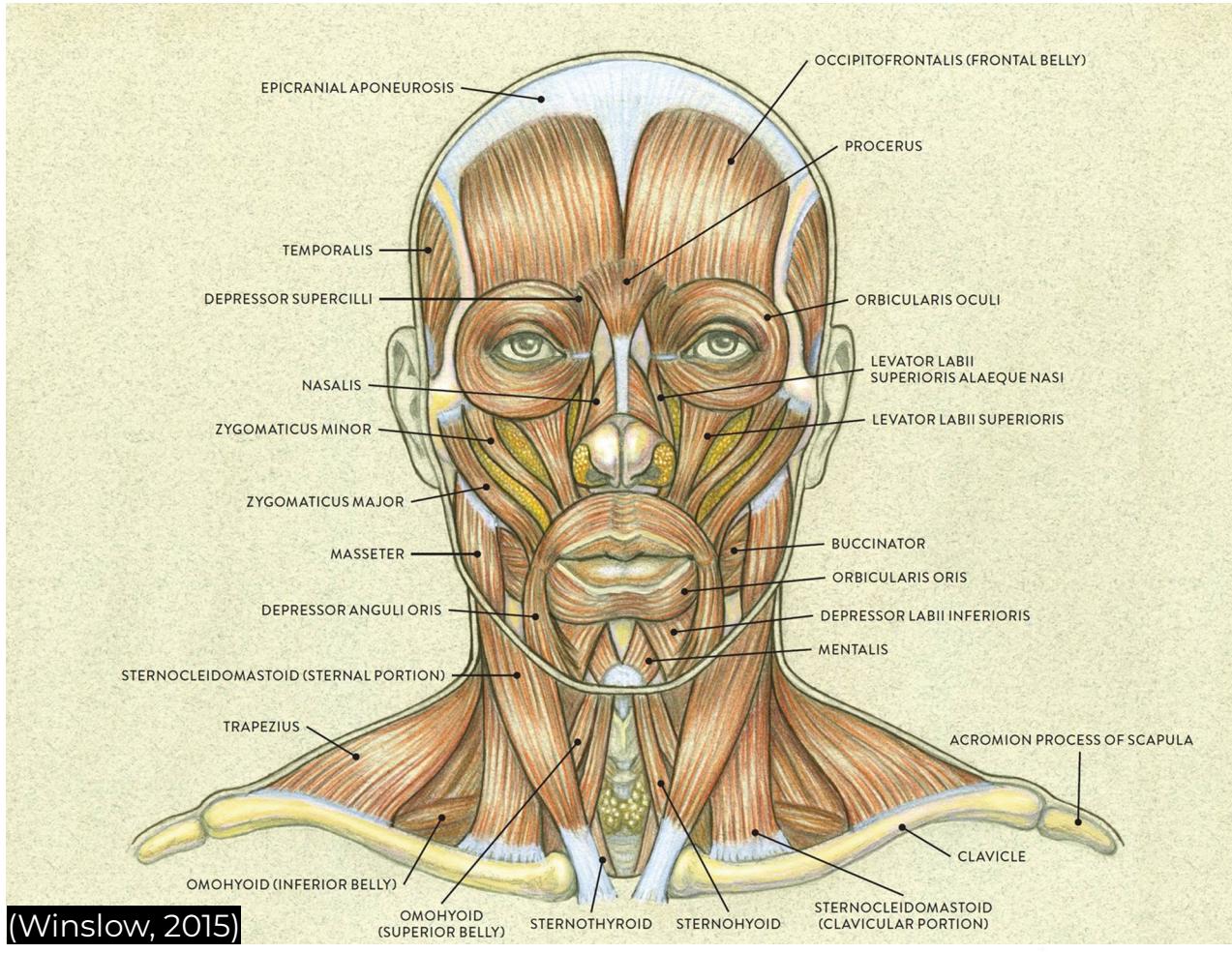


Cara-cara mengenali emosi

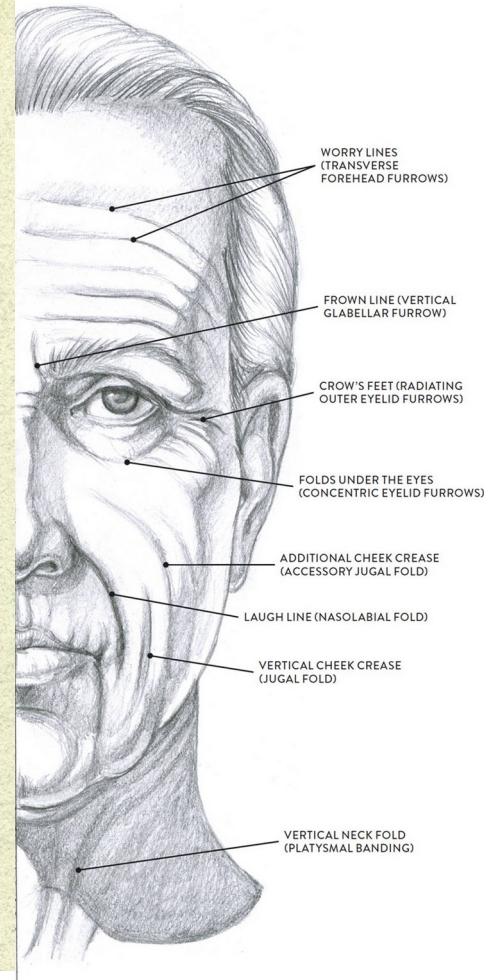
Latar Belakang Alinea 2

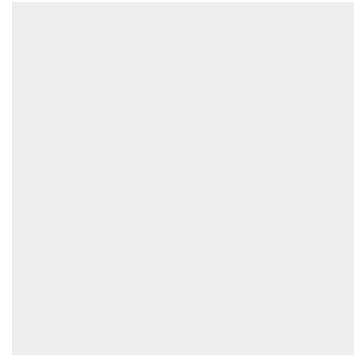
Pengenalan emosi melalui ekspresi wajah diklaim paling akurat*.





(Winslow, 2015)



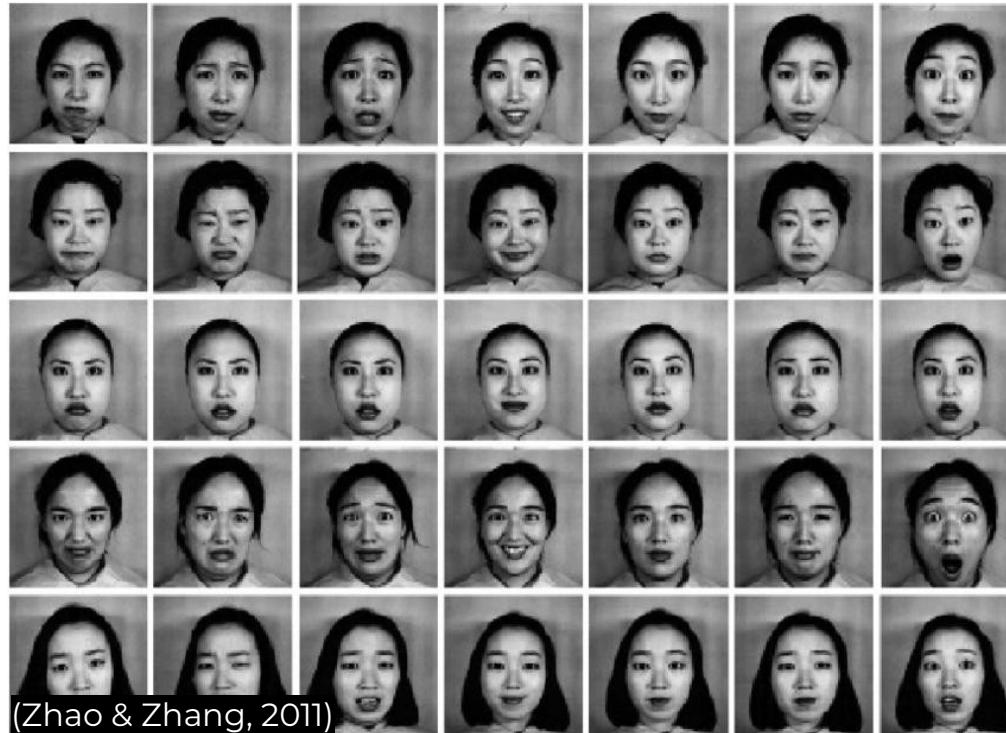


Masalah

Latar Belakang Alinea 4

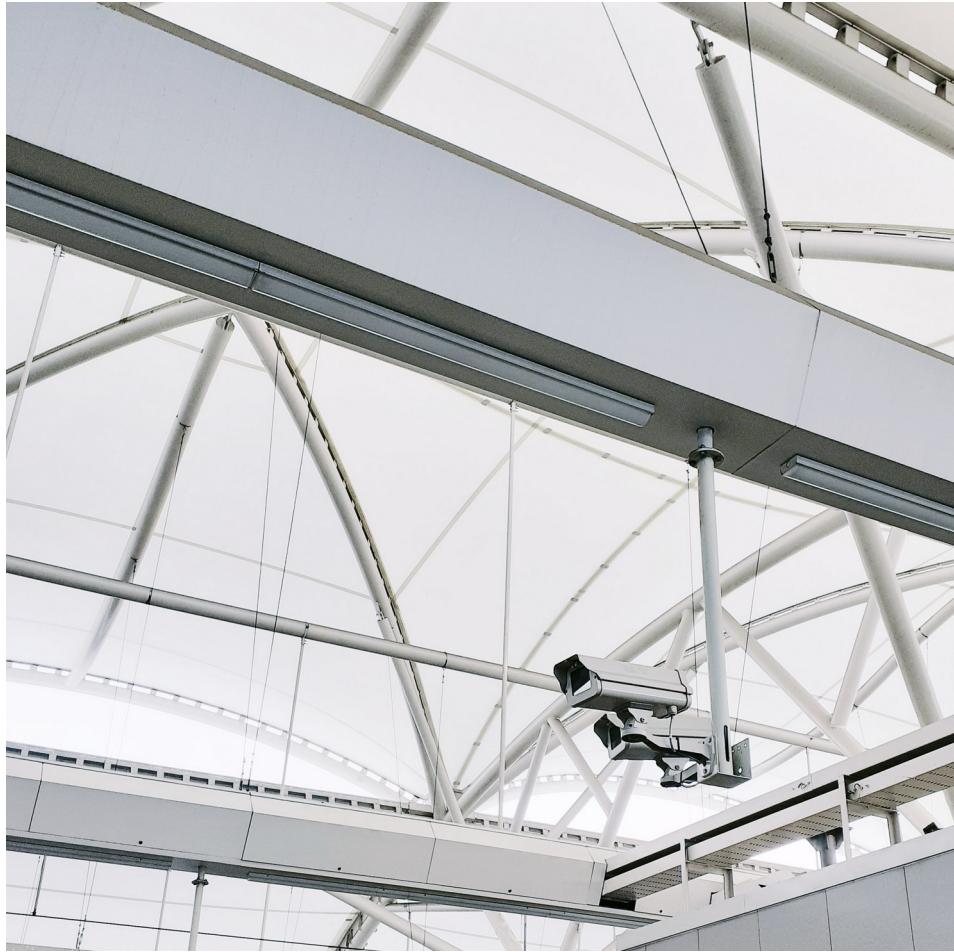
Keberhasilan pengenalan ekspresi wajah bergantung kepada kuantitas dan kualitas set data → sangat bervariasi.

Set data wajah frontal & nonfrontal (1/2)

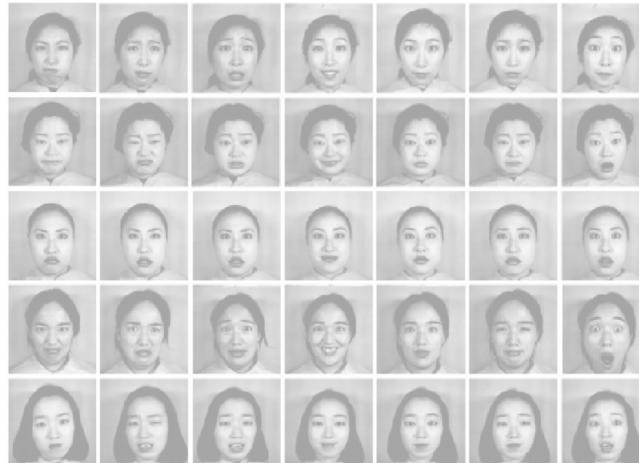


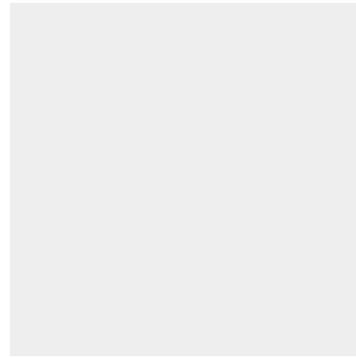
(Zhao & Zhang, 2011)





Set data wajah frontal & nonfrontal (2/2)





Latar Belakang Alinea 6

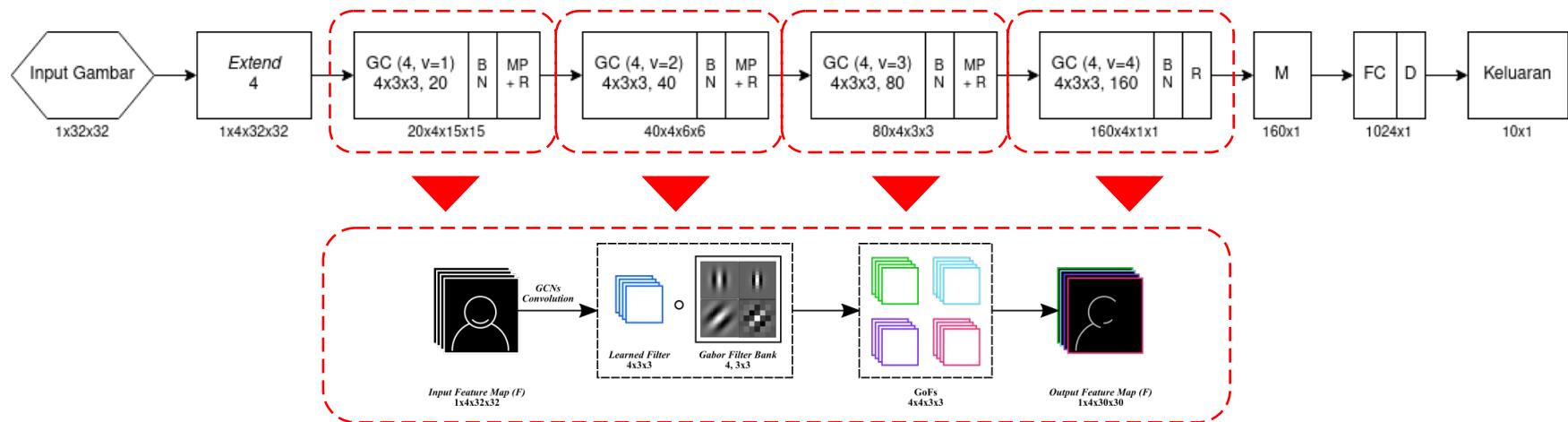
Penelitian terhadap set data wajah nonfrontal belum mencapai kepuasan.

Akurasi 100% untuk CK+: Zhou, J., & Wang, T. (2019). Facial Feature Recognition based on ASNMF Method.

Akurasi 75,42% untuk FER-2013: Georgescu, M. I., Ionescu, R. T., & Popescu, M. (2019). Local learning with deep and handcrafted features for facial expression recognition. IEEE Access, 7, 64827–64836.

Gabor convolutional networks

- Filter Gabor + CNN? Bukan, sudah banyak!
- CNN + filter Gabor → Gabor CNN (Luan dkk., 2018)



Facial region segmentation

Segmentasi 2/3 bagian wajah: kedua mata, hidung & mulut?

Ya, namun untuk set data wajah nonfrontal!

Filter log-Gabor

Filter Gabor → filter log-Gabor (Field, 1987)

- Untuk penelitian pengenalan ekspresi wajah
- Modifikasi Gabor CNN → log-Gabor CNN

Tujuan penelitian (1/2)

Latar Belakang Alinea 8

1- Rancangan pemodelan pengenalan emosi menggunakan *Gabor Convolutional Networks* (GCNs) melalui pendekatan *facial region segmentation* (FRS).

2- Analisis kinerja model menggunakan GCNs dengan dan tanpa FRS.

Tujuan penelitian (2/2)

Latar Belakang Alinea 8

3- Analisis kinerja model sebelum dan sesudah modifikasi instruksi pada GCNs melalui penggantian filter Gabor menjadi log-Gabor.

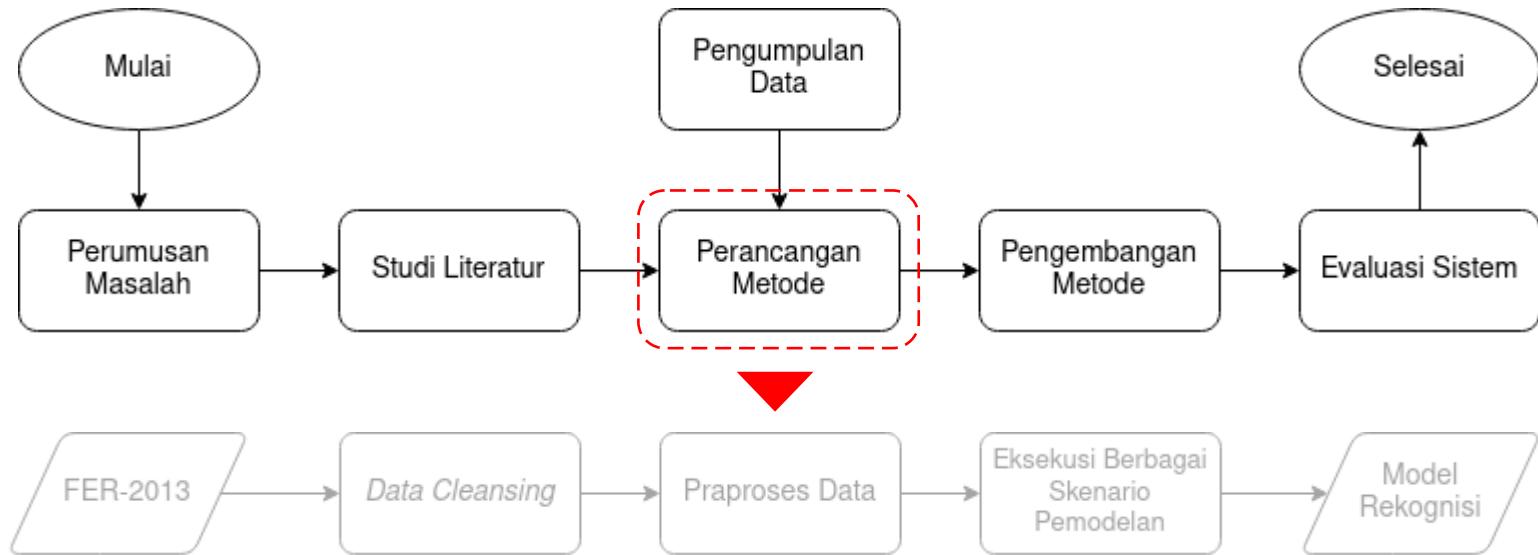
Batasan penelitian

- Diasumsikan bahwa ekspresi wajah adalah representasi paling akurat dari emosi manusia.
- Diasumsikan bahwa ekspresi wajah adalah universal untuk tujuh kelas emosi, yaitu *angry, disgust, fear, happy, sad, surprise, dan neutral.*

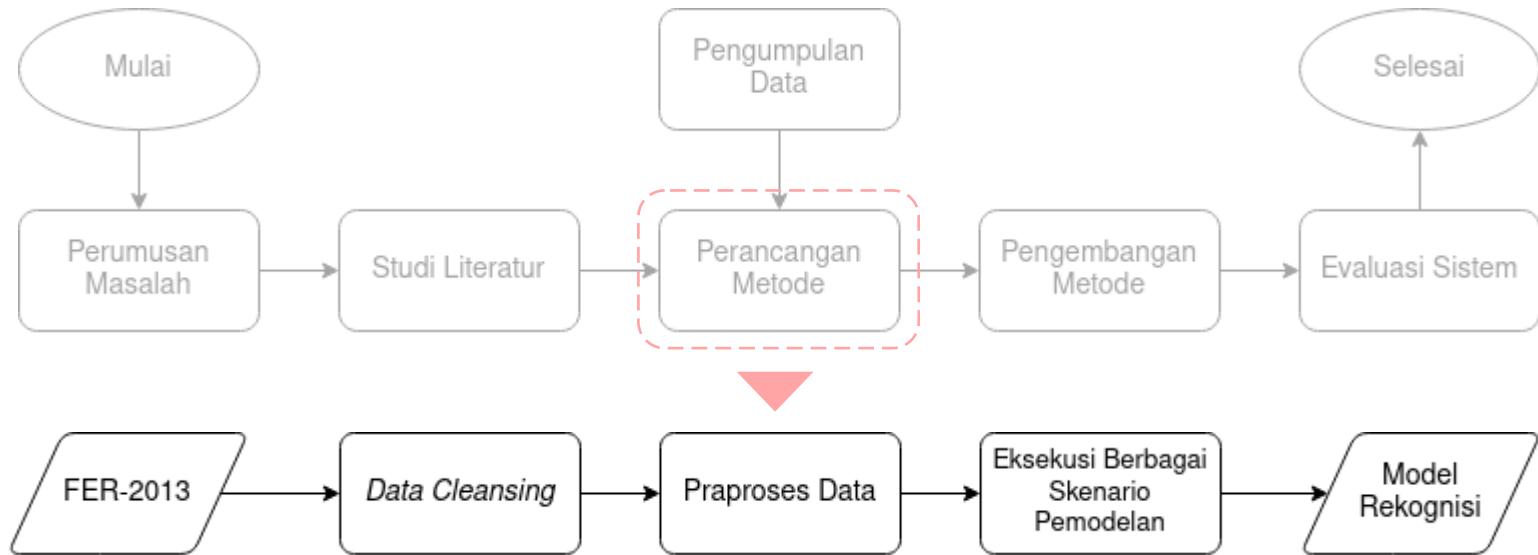
Manfaat penelitian

- Memberikan pengetahuan dan pengalaman pertama bagi saya di penelitian pengenalan emosi otomatis melalui ekspresi wajah.
- Memberikan kontribusi kepada masyarakat di penelitian pengenalan ekspresi wajah.
- Memberikan dorongan kepada industri dalam pemanfaatan pengenalan ekspresi wajah sebagai salah satu usaha mencapai tujuan bisnis mereka.

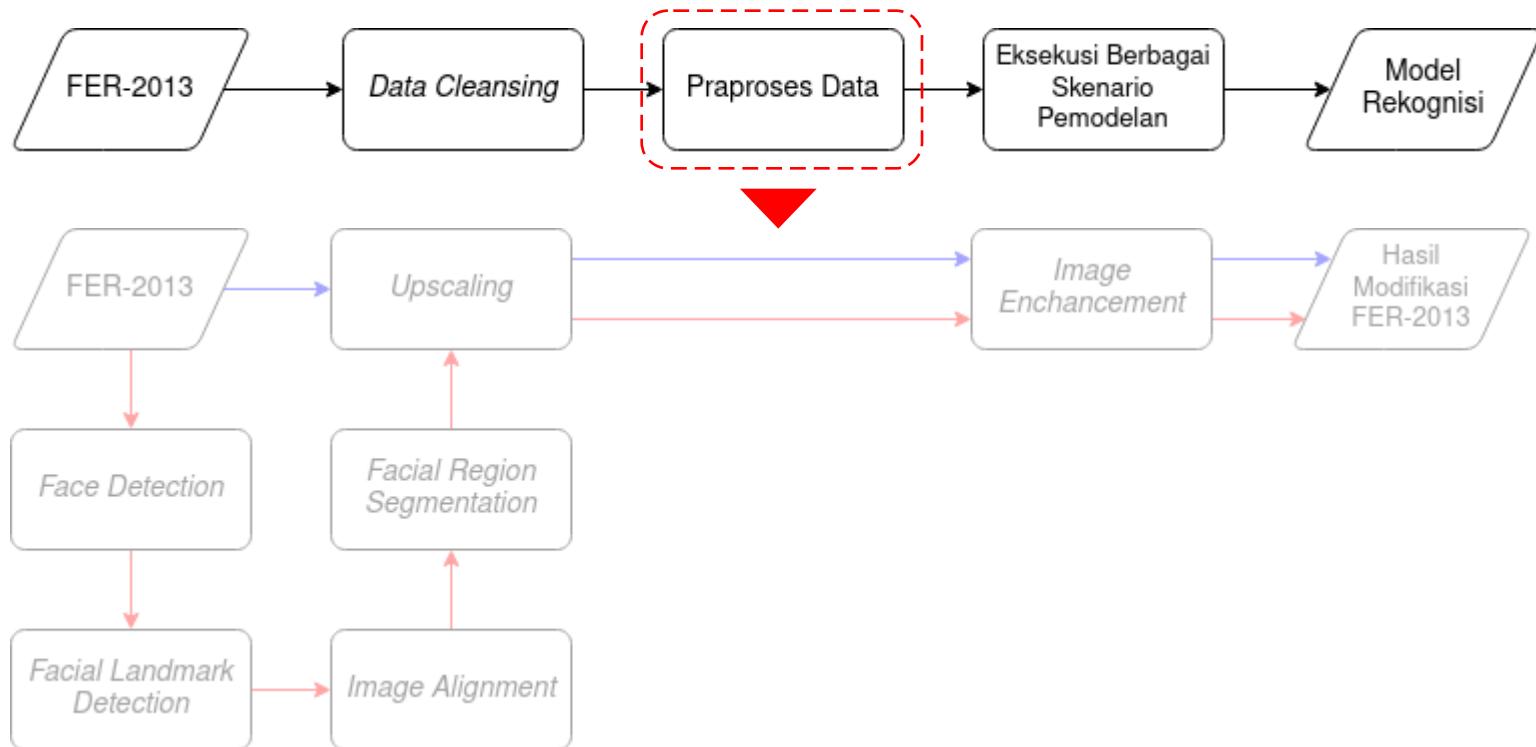
Desain Eksperimen (1/3)



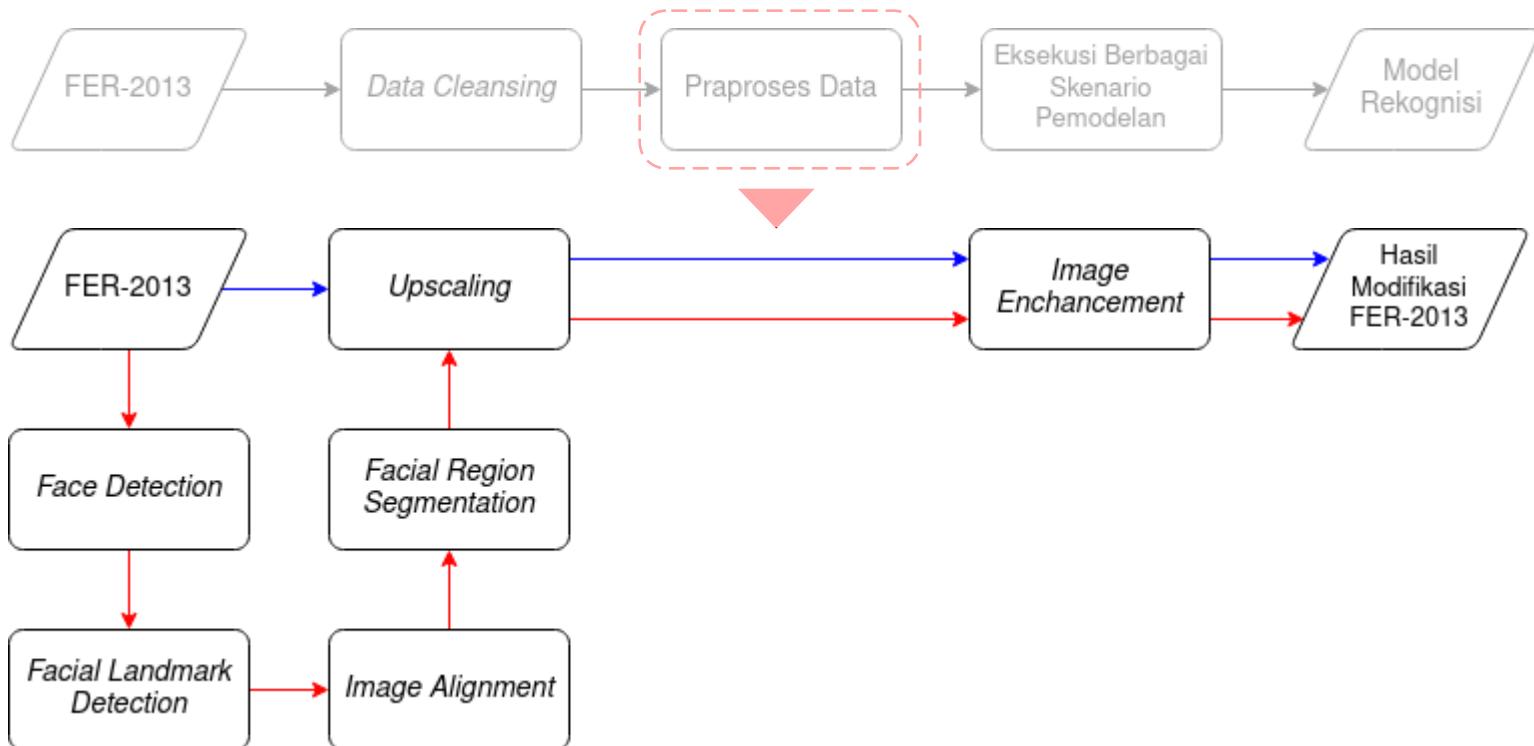
Desain Eksperimen (1/3)



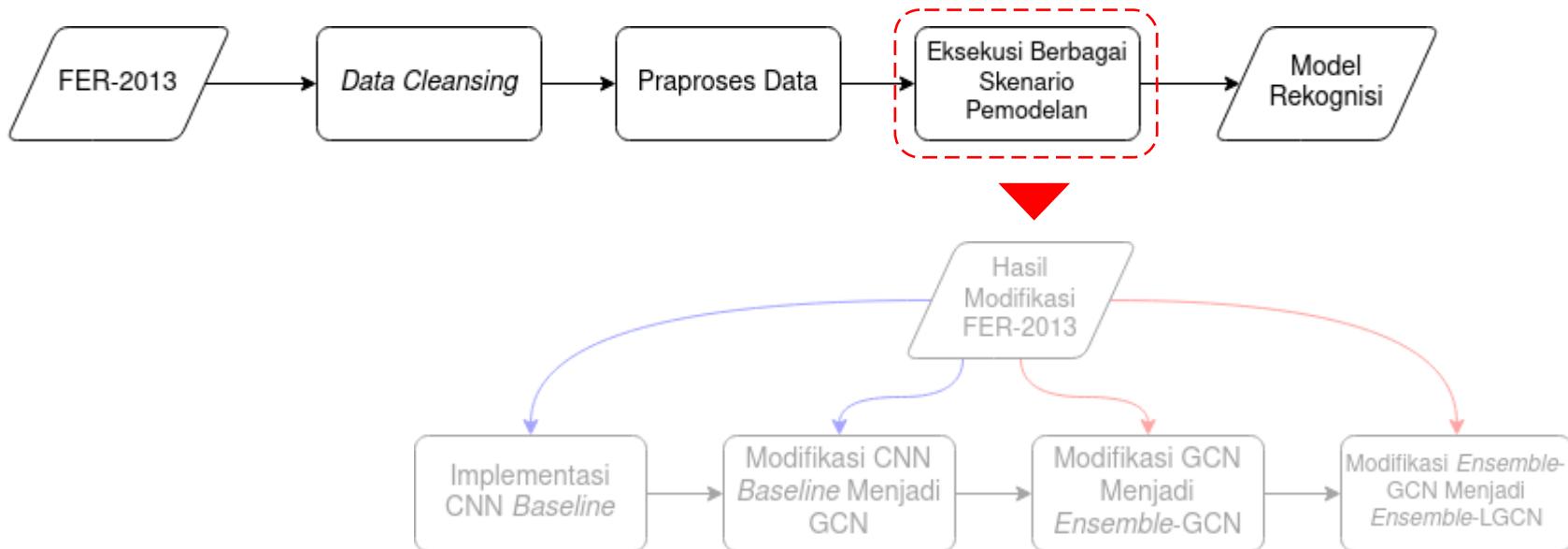
Desain Eksperimen (2/3)



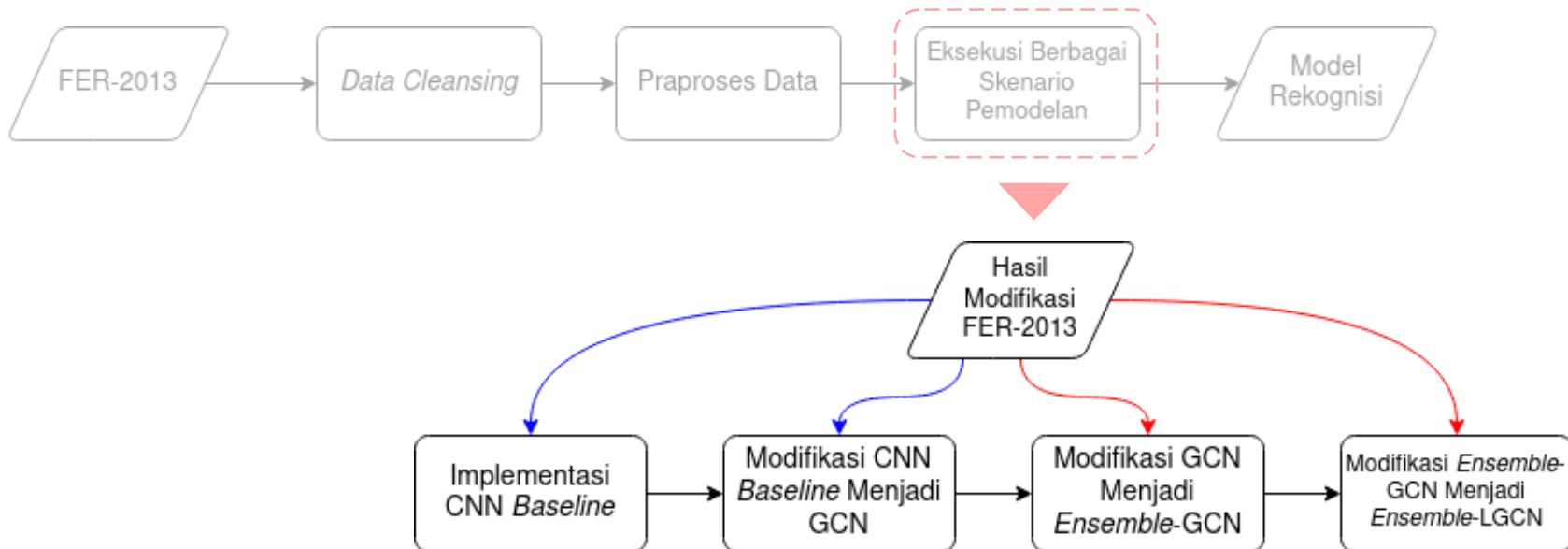
Desain Eksperimen (2/3)



Desain Eksperimen (3/3)



Desain Eksperimen (3/3)



Persiapan



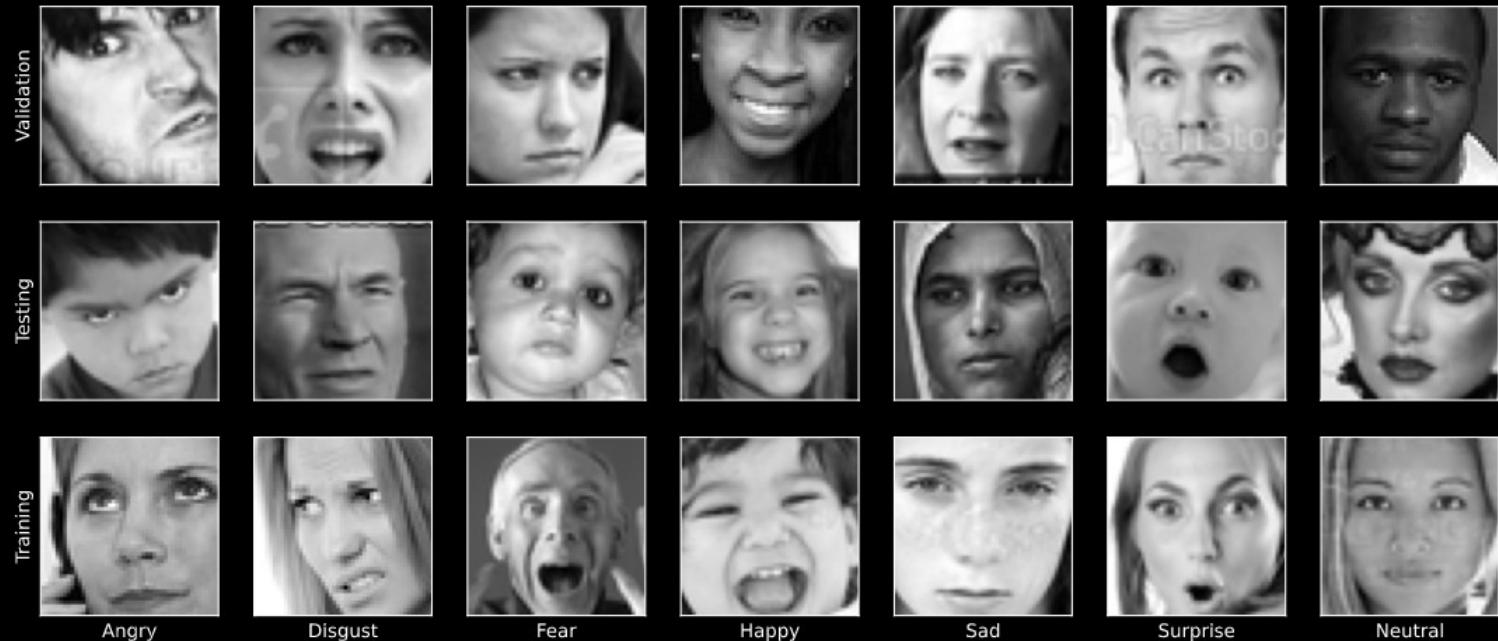
Eksplorasi set data (FER-2013)

Diambil dari <https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data>

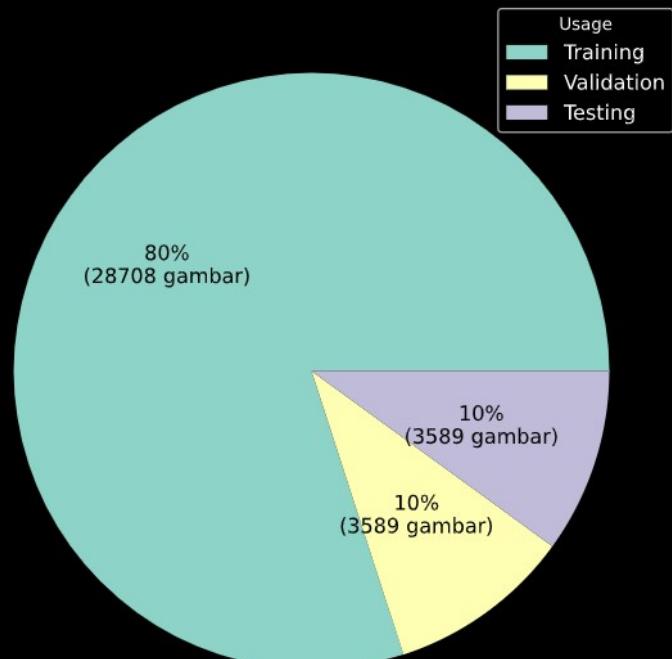
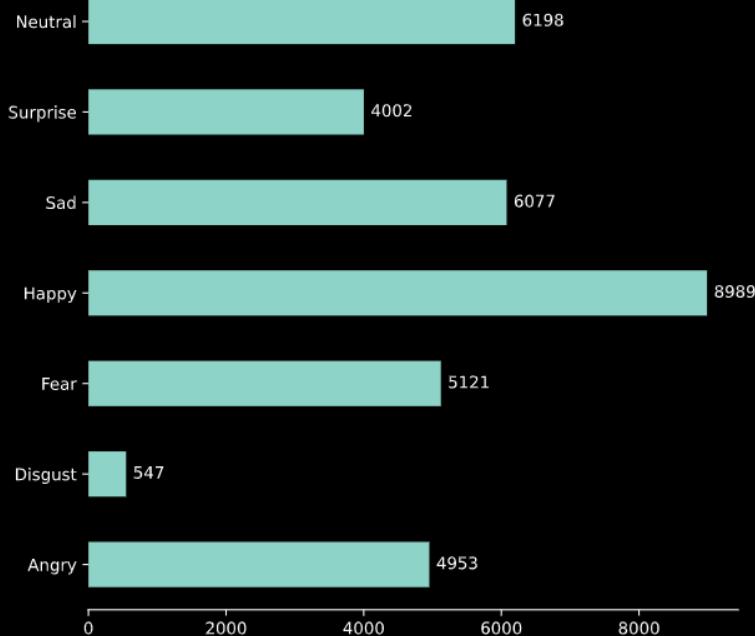
Terdiri dari:

- 35.887 gambar wajah manusia (dalam 48x48 piksel)
- 7 jenis label emosi: *Angry, disgust, fear, happy, sad, surprise* dan *neutral*
- Rasio data *training:validation:testing* adalah 80:10:10

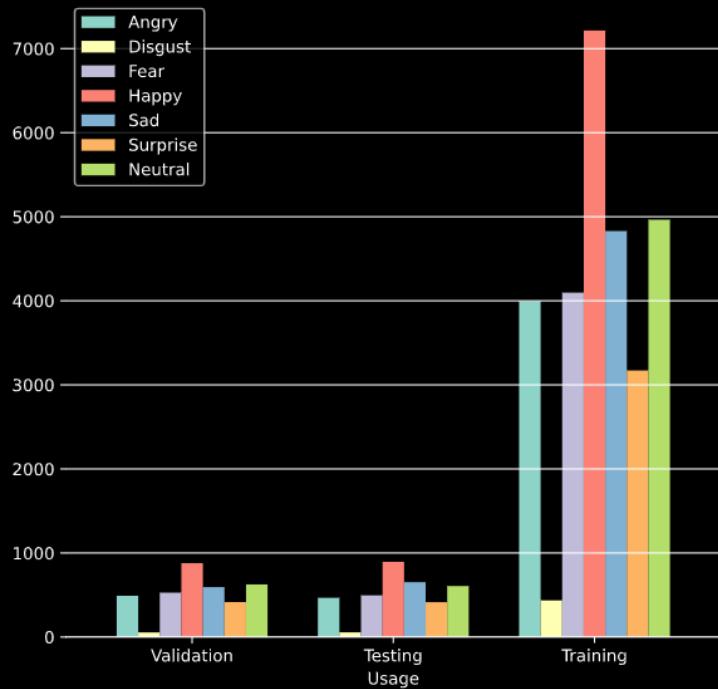
Pratinjau set data



Distribusi set data (1/2)



Distribusi set data (2/2)

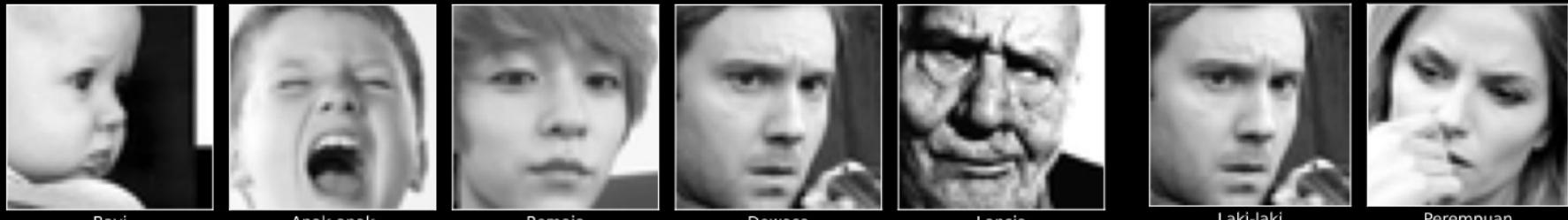


Perhatian!

Penelitian ini dan semua penelitian rujukan memakai distribusi ini, namun penelitian *baseline* menggunakan rasio 80:20 untuk data *training:testing*.

Variasi set data (1/3)

Variasi usia



Variasi gender



Variasi jarak kamera



Variasi rotasi

Variasi set data (2/3)

Variasi kualitas



Normal

Overbright

Underbright

Blurry

Blocky

B&W

Variasi tipe subjek



Foto manusia

Lukisan

Kartun 2D

Kartun 3D

Foto patung

Variasi arah iluminasi



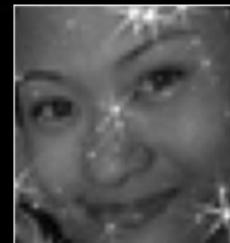
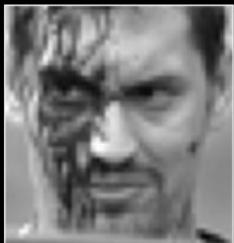
Dari arah kanan

Dari arah kiri

Dari arah atas

Variasi set data (3/3)

Variasi gangguan



Sun, Y., Shan, C., Tan, T., Tong, T., Wang, W., & Pourtaherian, A. (2019). Detecting discomfort in infants through facial expressions. *Physiological measurement*, 40(11), 115006.

Benitez-Garcia, G., Nakamura, T., & Kaneko, M. (2018). Multicultural facial expression recognition based on differences of western-caucasian and east-asian facial expressions of emotions. *IEICE TRANSACTIONS on Information and Systems*, 101(5), 1317–1324.

Pembersihan data rusak

Ditemukan 85
data gambar
rusak melalui
pengecekan
manual satu-s

Termasuk data tidak relevan seperti gambar multiwajah.



Spesifikasi komputer & sistem (1/2)

Spesifikasi komputer:

- CPU: Intel i5 8250u (8) @ 1.60GHz-3.40GHz
- GPU: Nvidia GTX 1050M 4GB
- RAM: 16GB 2400MHz

Spesifikasi sistem:

- OS: Fedora 32 Workstation
- Kernel: Linux 5.7.8-200.fc32.x86_64

Spesifikasi komputer & sistem (2/2)

- CUDA driver 10.2
- Nvidia driver 440.100
- GCC 10.1.1 20200507 (Red Hat 10.1.1-1)
- Python 3.8.3
- Pustaka-pustaka: gcn 0.1, imgaug 0.4.0, numpy 1.18.5, opencv-python 4.2.0.34, scikit-learn 0.23.1, torch 1.5.0, matplotlib 3.2.1 dll.

Konfigurasi *training* (1/2)

```
""" Fixed random generator for reproducibility """
seed = 123

import os
os.environ['PYTHONHASHSEED'] = str(seed)

import random
random.seed(seed)

import numpy as np
np.random.seed(seed)

import torch
torch.manual_seed(seed)

torch.backends.cudnn.benchmark = True # must be set to False
torch.backends.cudnn.deterministic = False # must be set to True
```

Konfigurasi *training* (2/2)

```
criterion = torch.nn.CrossEntropyLoss(reduction='sum').cuda()
optimizer = torch.optim.Adam(model.parameters(), lr=.001,
                             betas=(.9, .999), eps=1e-7,
                             weight_decay=.01)
scheduler = torch.optim.lr_scheduler.ReduceLROnPlateau(optimizer,
                                                       factor=np.sqrt(.1),
                                                       patience=10,
                                                       min_lr=.5e-6)
```

```
from imgaug import augmenters as iaa

affine = iaa.Affine(rotate=(-5, 5), scale=(.9, 1.1), shear=(-5, 5),
                    translate_percent={'x': (-.1, .1), 'y': (-.1, .1)},
                    random_state=seed)
hflip = iaa.Fliplr(.5, random_state=seed)
cutout = iaa.Cutout(size=.3, squared=False, random_state=seed)
```

Hasil & pembahasan (I/III)



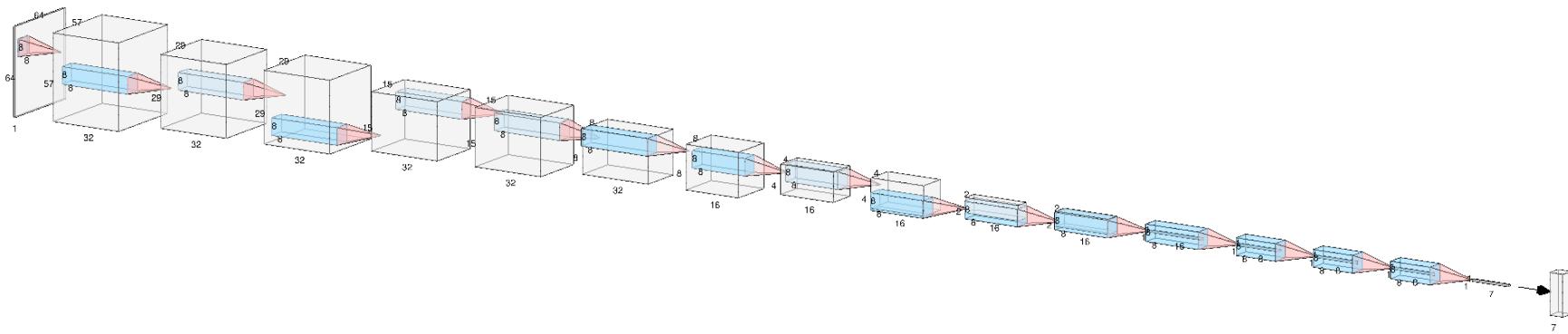
Eksperimen I-A (1/7)

Percobaan ke-	1	2	3
Akurasi tes (%)	<u>53,79</u>	<u>53,99</u>	<u>54,24</u>

Implementasi arsitektur
jaringan *baseline* tanpa
data augmentasi.

Eksperimen I-A (2/7)

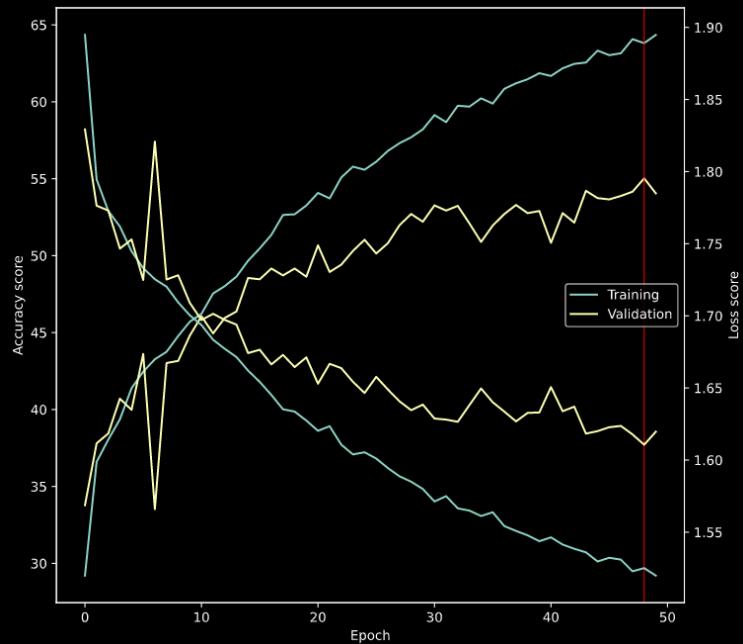
Arsitektur CNN *baseline*



The image above was created using <http://alexlenail.me/NN-SVG/AlexNet.html>

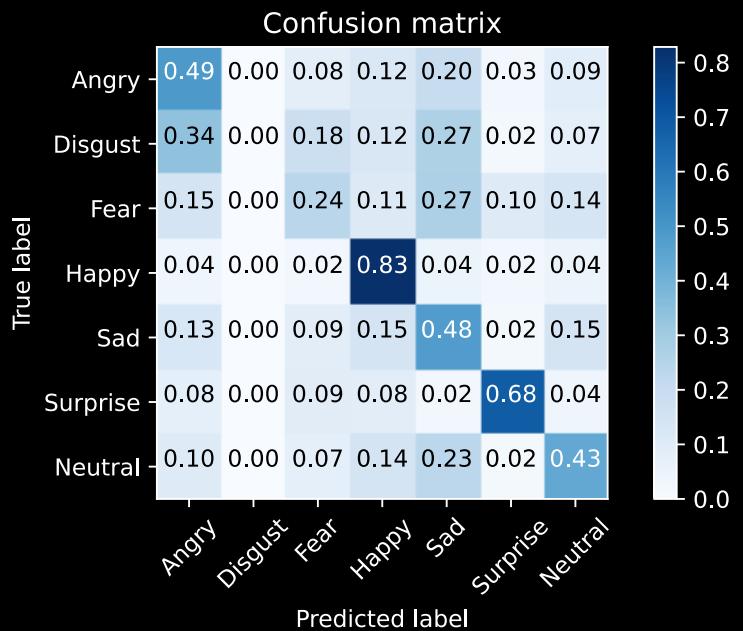
Arsitektur di atas adalah “model2” oleh Agrawal, A., & Mittal, N. (2020). Using cnn for facial expression recognition: a study of the effects of kernel size and number of filters on accuracy. The Visual Computer, 36(2), 405–412.

Eksperimen I-A (3/7)



Percobaan ke-	<u>3/3</u>
Waktu <i>training</i> (jam)	<u>0,31</u>
Optimal di epoch ke-	<u>48</u>
Akurasi tes (%)	<u>54,24</u>

Eksperimen I-A (4/7)

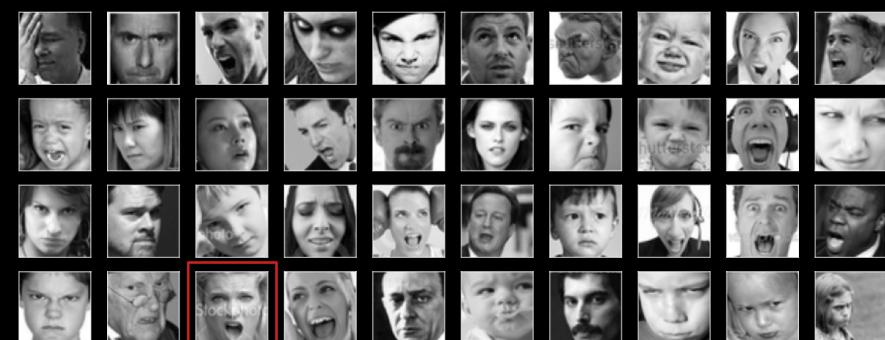


1- Data tes berlabel “*disgust*”
tidak ada yang berhasil
dikenali dengan benar.

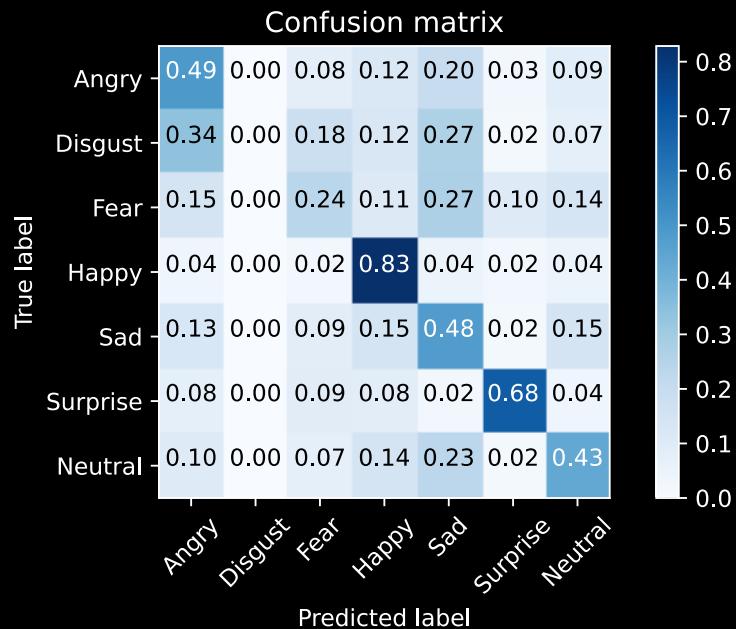
- Apakah karena data *training*-nya karena terlalu sedikit?
- Apakah karena emosi “*disgust*” sulit dibedakan dari “*angry*”?



disgust* versus *angry



Eksperimen I-A (5/7)

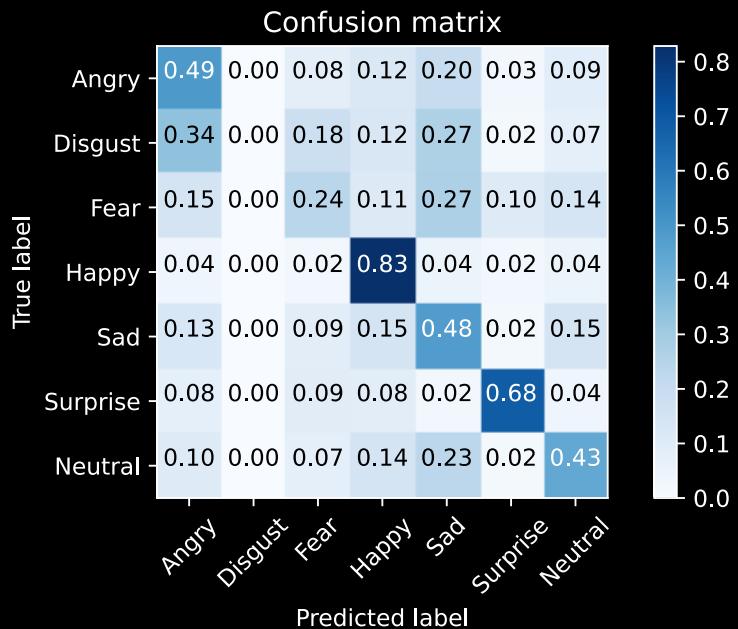


2- Data tes berlabel “fear”

sangat sedikit yang berhasil
dikenali dengan benar.

Apakah karena emosi “fear” sulit
dibedakan dari emosi yang lain,
terutama dengan emosi “sad”?

Eksperimen I-A (6/7)



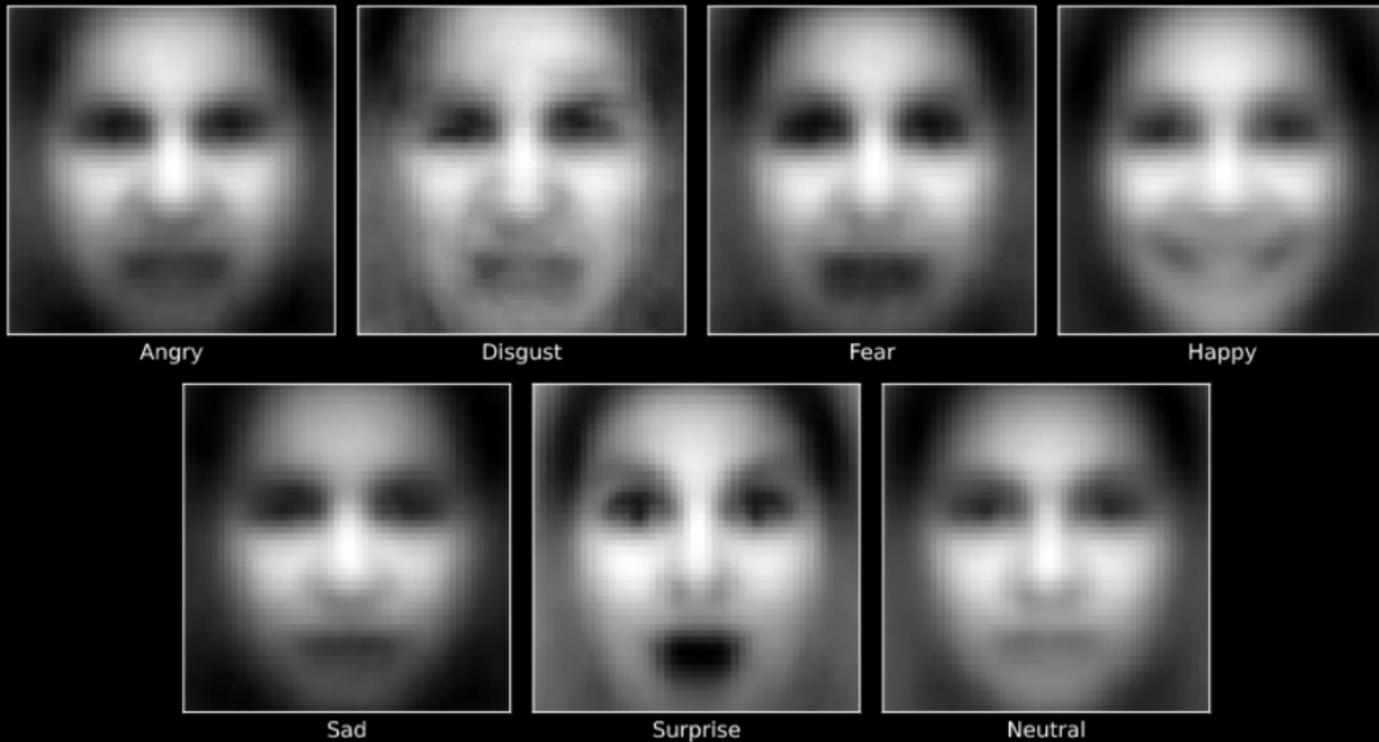
3- Data tes berlabel “sad”

masih tidak ada separuhnya

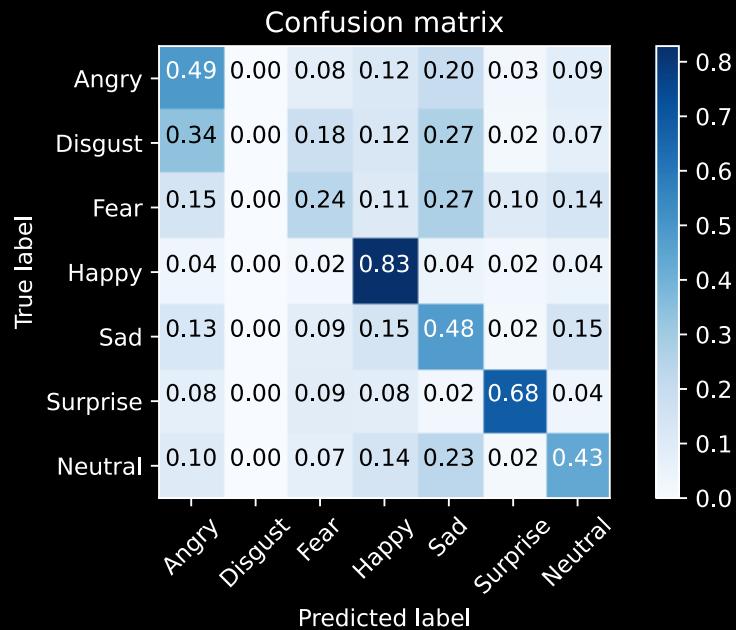
yang berhasil dikenali dengan
benar.

Apakah karena emosi “sad” sulit
dibedakan dari emosi yang lain,
terutama dengan emosi “neutral”?

Rerata gambar wajah per label emosi



Eksperimen I-A (7/7)



- Apakah hanya karena arsitektur jaringan terlalu sederhana untuk dapat belajar?

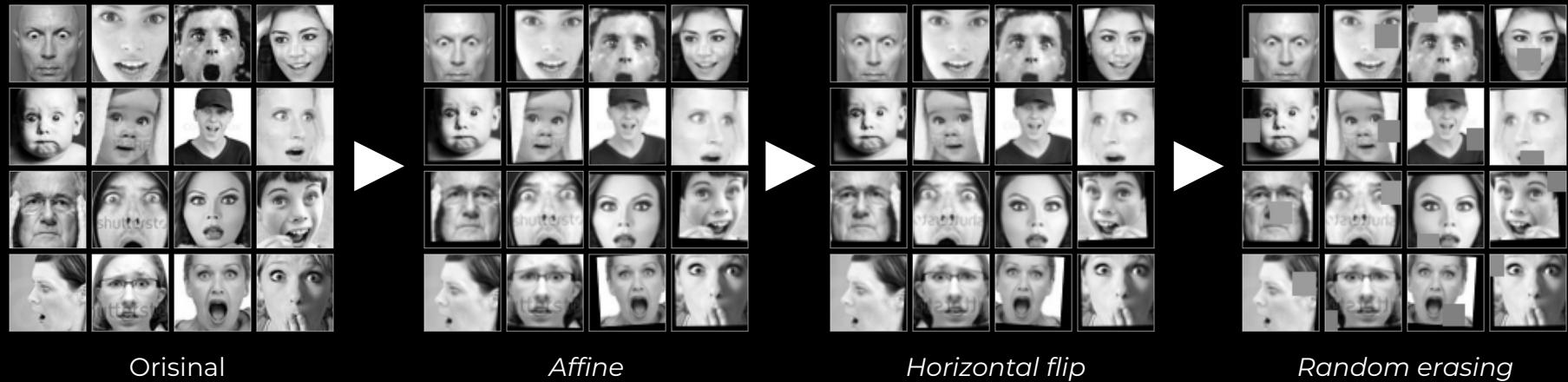
- Ataukah karena terdapat kesalahan dalam pelabelan data?

Eksperimen I-B (1/3)

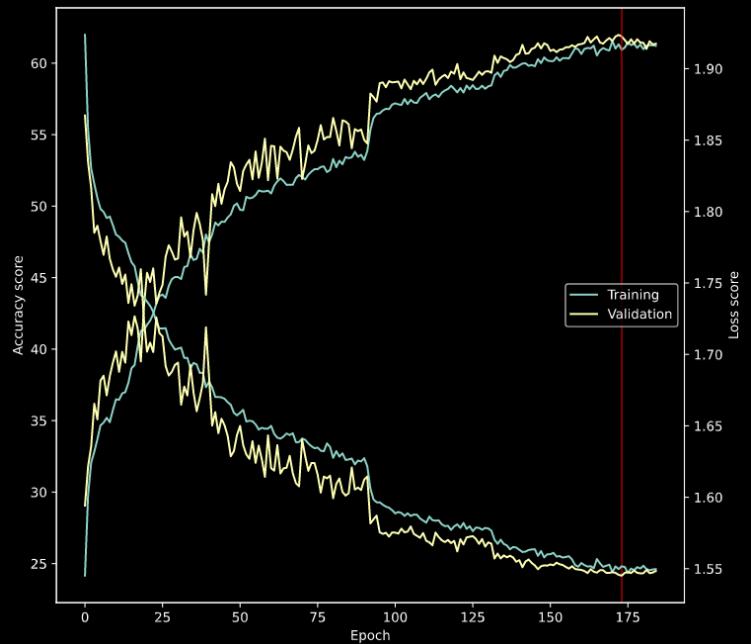
Percobaan ke-	1	2	3
Akurasi tes (%)	<u>59,71</u>	<u>60,27</u>	<u>59,21</u>

Implementasi arsitektur
jaringan *baseline* dengan
data augmentasi.

Augmentasi data

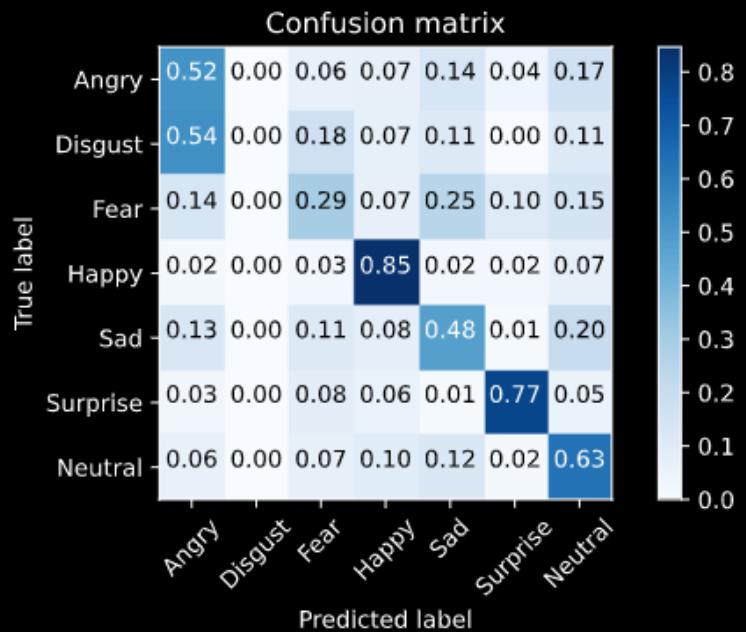


Eksperimen I-B (2/3)



Percobaan ke-	<u>2/3</u>
Waktu <i>training</i> (jam)	<u>1,64</u>
Optimal di epoch ke-	<u>174</u>
Akurasi tes (%)	<u>60,27</u>

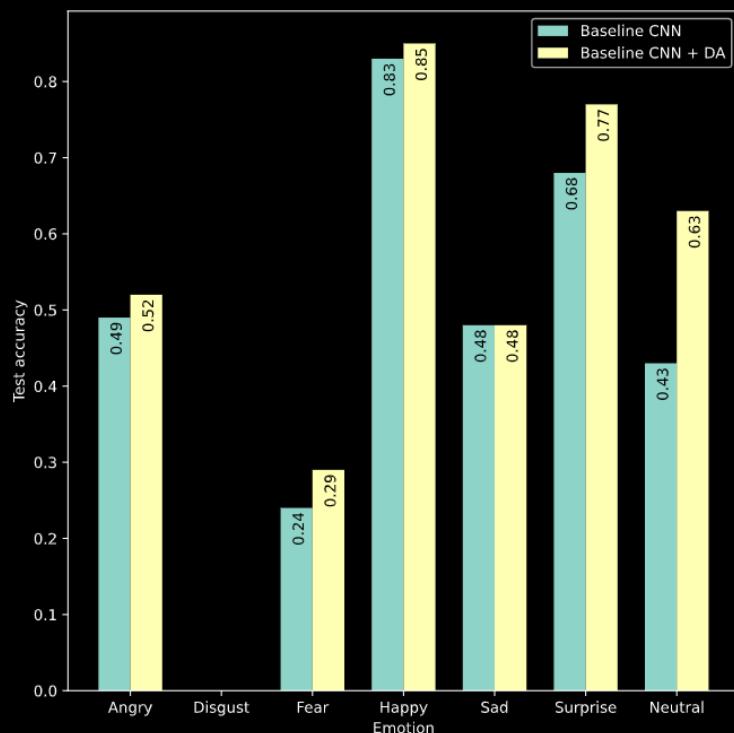
Eksperimen I-B (3/3)



Data tes berlabel “*disgust*”
masih belum ada yang berhasil
dikenali dengan benar.

Malah makin banyak dikenali
sebagai emosi “*angry*”.

Eksperimen I-A & I-B



↑6,03% terhadap eksperimen sebelumnya

	Baseline CNN	Baseline CNN + DA
Acc (%)	<u>54,24</u>	<u>60,27</u>

Acc — akurasi test

Kesimpulan

Terbukti bahwa penting untuk melakukan data augmentasi terhadap set data FER-2013.

Eksperimen I-C (1/4)

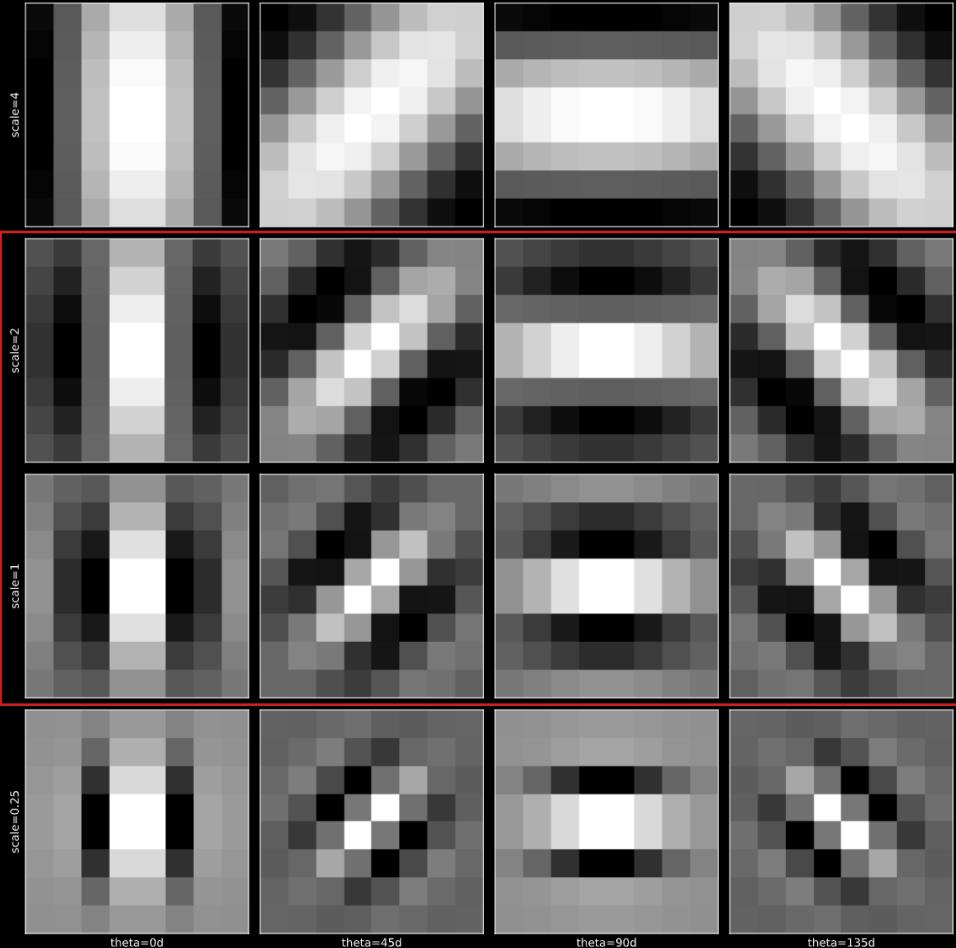
Percobaan ke-	1	2	3
Akurasi tes (%)	<u>63,51</u>	<u>62,14</u>	<u>61,89</u>

Implementasi *Gabor convolutional network* pada arsitektur jaringan *baseline*.

Eksperimen I-C (2/4)

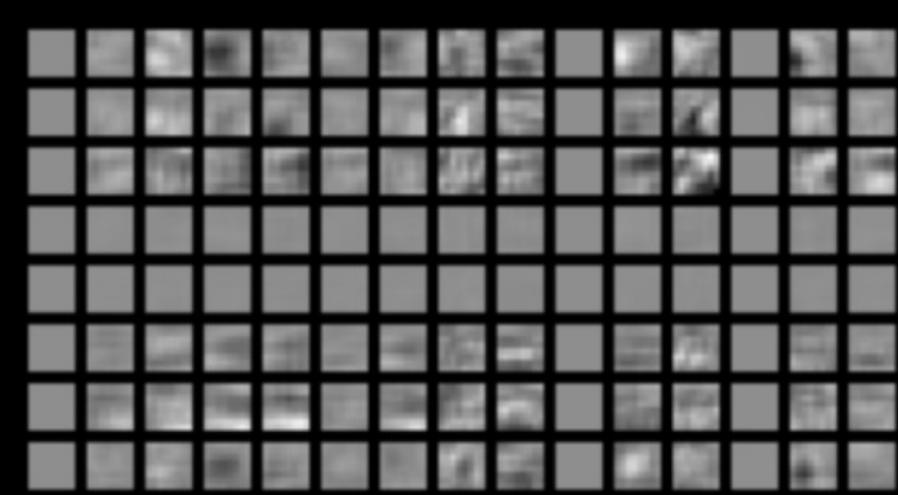
Arsitektur jaringan *baseline* yang ditingkatkan

- Mengubah setiap *convolutional layer* menjadi Gabor *convolutional layer*
- Penambahan sebuah *max layer* dan *fully connected layer* di paling akhir



Konfigurasi Gabor *convolutional layer*

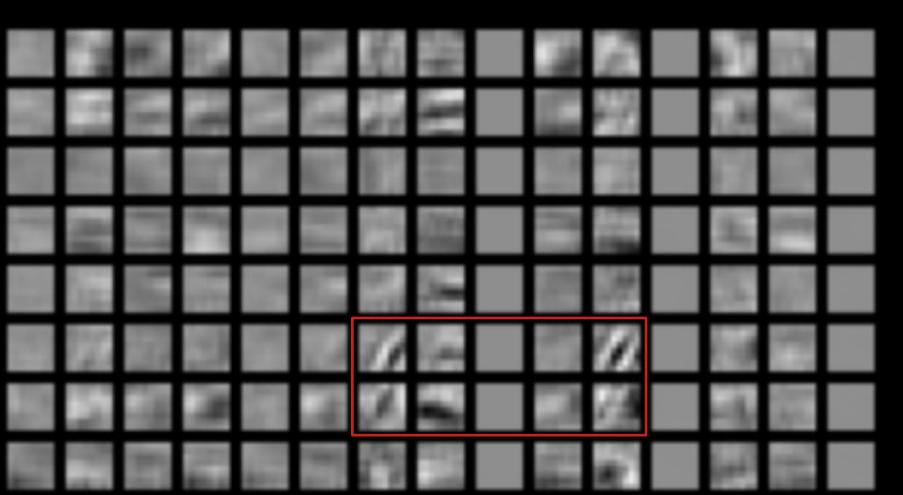
<u>kernel_size</u>	8
M	4
<u>scale (GConv<1-4>)</u>	<u>2</u>
<u>scale (GConv<5-8>)</u>	<u>2</u>
<u>scale (GConv<9-12>)</u>	<u>1</u>
<u>scale (GConv<13-16>)</u>	<u>1</u>



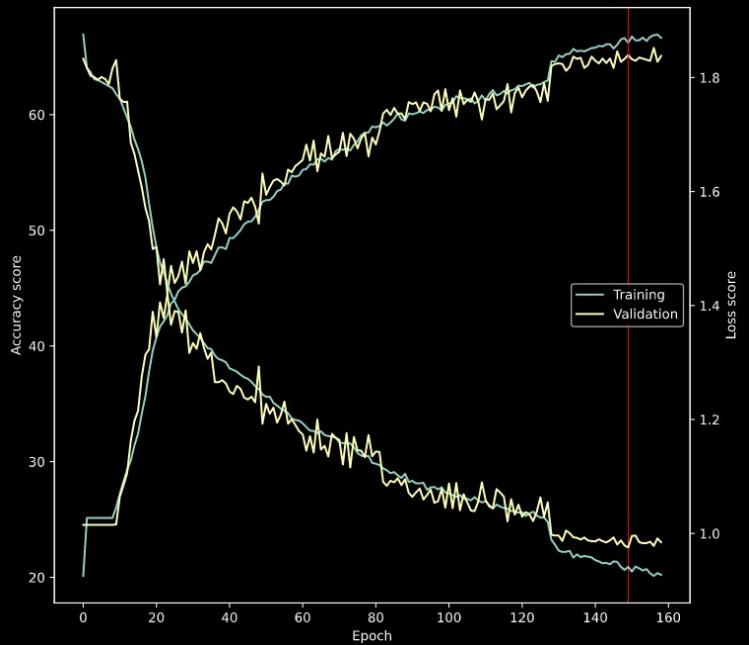
Kernel
L8

ver
sus

kernel
L9

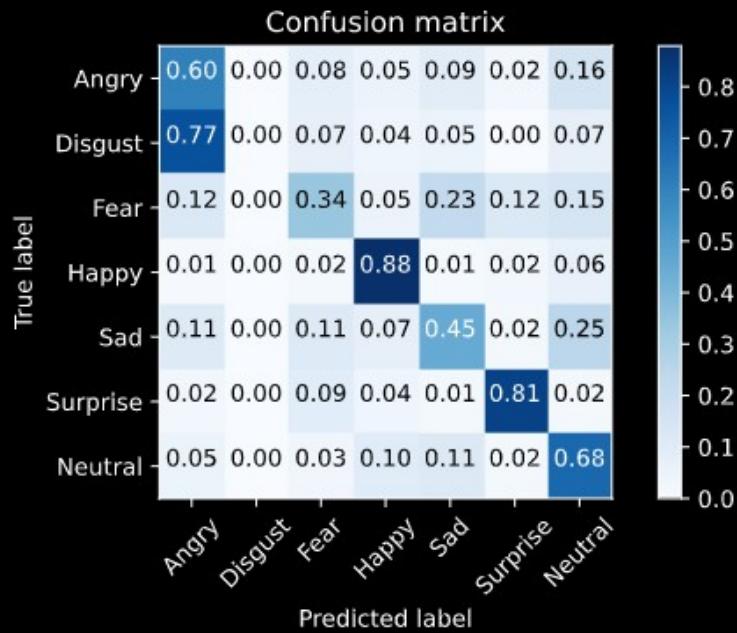


Eksperimen I-C (3/4)



Percobaan ke-	<u>1/3</u>
Waktu <i>training</i> (jam)	<u>7,95</u>
Optimal di epoch ke-	<u>150</u>
Akurasi tes (%)	<u>63,51</u>

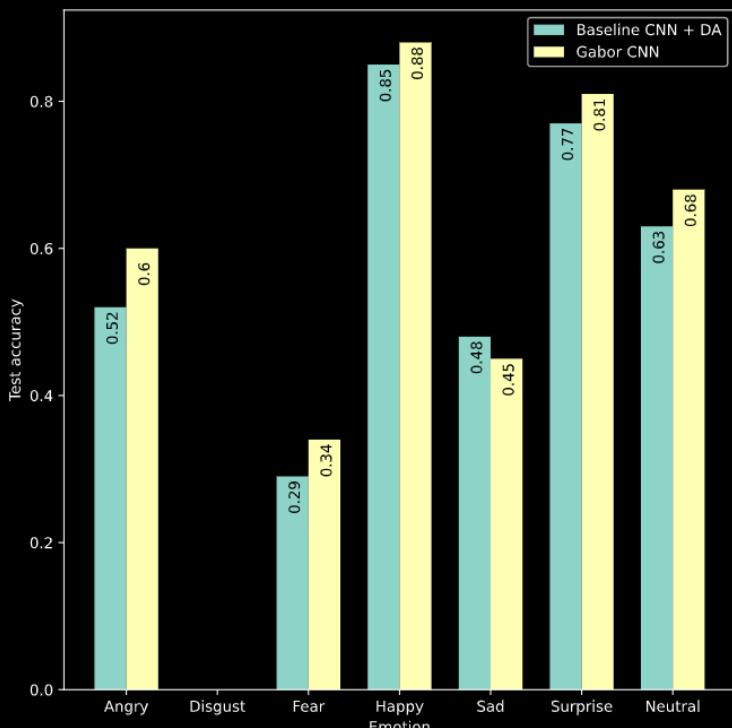
Eksperimen I-C (4/4)



Data tes berlabel “*disgust*”
masih belum ada yang berhasil
dikenali dengan benar.

Malah makin banyak lagi dikenali
sebagai emosi “*angry*”.

Eksperimen I-B & I-C



↑3,24% terhadap eksperimen sebelumnya

	Baseline CNN + DA	Gabor CNN
Acc (%)	<u>60,27</u>	63,51

Acc — akurasi test

Kesimpulan

Terbukti bahwa penggunaan arsitektur *Gabor convolutional network* dapat meningkatkan performa model.

Hasil & pembahasan (II/III)



Eksperimen II-A

Nama library	Loss (%)	Waktu (s)	Marks
face_alignment* (<u>face_detector=sfd</u>)	<u>3,10</u>	<u>2786</u>	<u>68</u>
RetinaFace** (<u>net=mobilenet0.25</u>)	<u>42,23</u>	<u>334</u>	<u>5</u>
RetinaFace** (<u>net=resnet50</u>)	<u>7,05</u>	<u>476</u>	

Loss — persentase banyak data yang tidak terdeteksi wajah untuk threshold=0,5; Waktu — waktu total pendeksi *facial landmark* seluruh data gambar menggunakan GPU; Marks — banyak *facial landmarks* per wajah

Implementasi pustaka
untuk mendeksi *facial landmarks*.

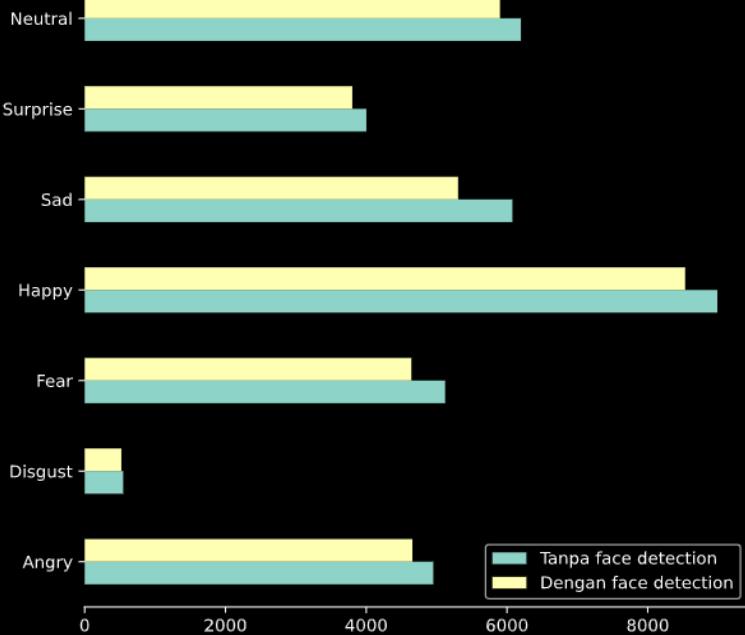
* <https://github.com/ladrianb/face-alignment>

** https://github.com/biubug6/Pytorch_Retinaface

Keluaran face_alignment

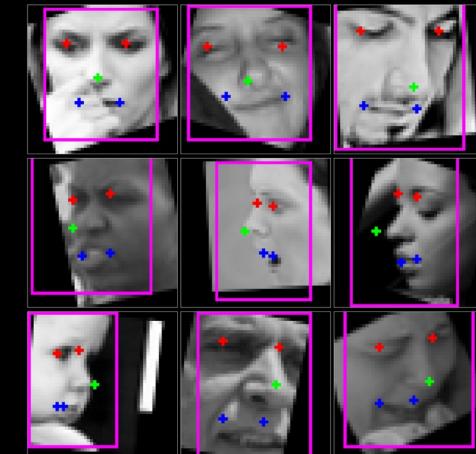
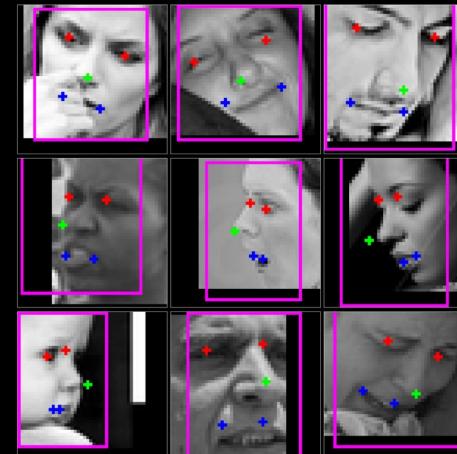
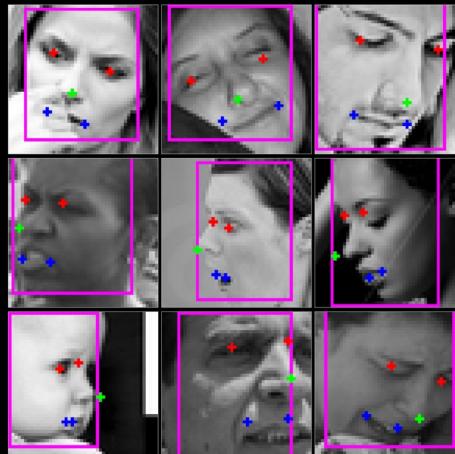


Keluaran RetinaFaces

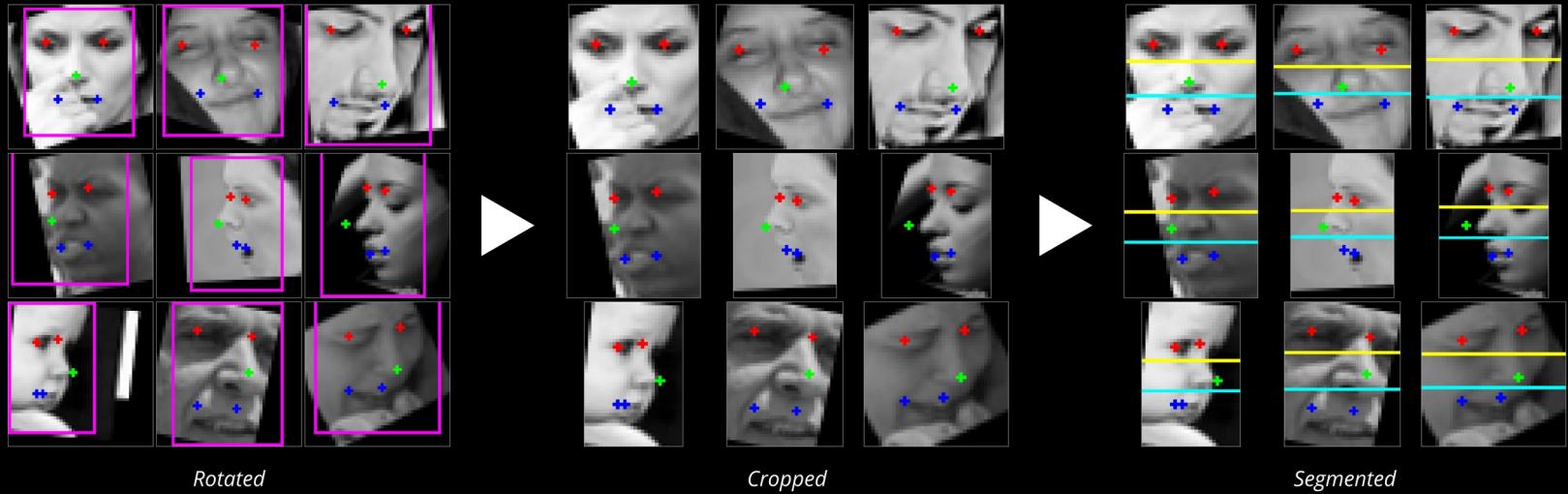


$(0,5 < \text{visualization_threshold} < 0,6)$

Facial region segmentation (1/2)



Facial region segmentation (2/2)



Eksperimen II-B (1/2)

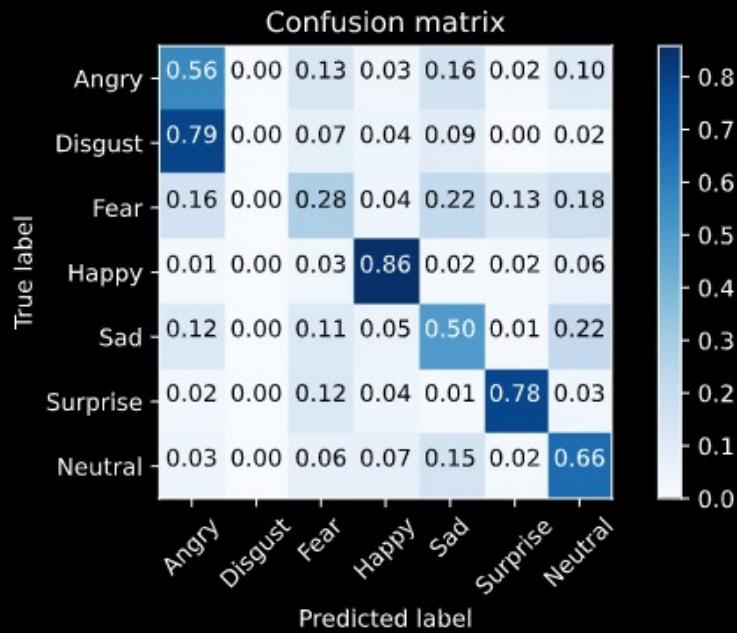
Fitur	Waktu training (jam)	Epoch	Akurasi tes (%) pada perhitungan		
			Single	Simple average	Weighted average
EN + NM	<u>4,17</u>	<u>124</u>	<u>58,09</u>	<u>62,20</u>	<u>62,20</u>
	<u>4,70</u>	<u>141</u>	<u>58,06</u>		
E + N + M	<u>3,44</u>	<u>153</u>	<u>17,41</u>	<u>44,10</u>	<u>53,22</u>
	<u>4,05</u>	<u>181</u>	<u>25,45</u>		
	<u>2,85</u>	<u>129</u>	<u>53,88</u>		
E + N	Ibid.			<u>17,74</u>	<u>17,62</u>
N + M				<u>52,77</u>	<u>53,61</u>
E + M				<u>44,91</u>	<u>53,28</u>
EN+NM (concat.)	<u>7,34</u>	<u>103</u>	<u>59,71</u>	-	
E+N+M (concat.)	<u>8,36</u>	<u>131</u>	<u>54,12</u>		
ENM	<u>6,80</u>	<u>131</u>	<u>62,08</u>		

E — Eyes atau daerah mata; N — Nose atau daerah hidung; M — Mouth atau daerah mulut;

Weighted average — rerata yang dihitung berdasarkan akurasi tes tiap-tiap model

Implementasi *Gabor convolutional network*
pada set data hasil *facial region segmentation*.

Eksperimen II-B (2/2)



Data tes berlabel “*disgust*”
masih belum ada yang berhasil
dikenali dengan benar.

Malah makin banyak lagi dikenali
sebagai emosi “*angry*”.

Hasil & pembahasan (III/III)



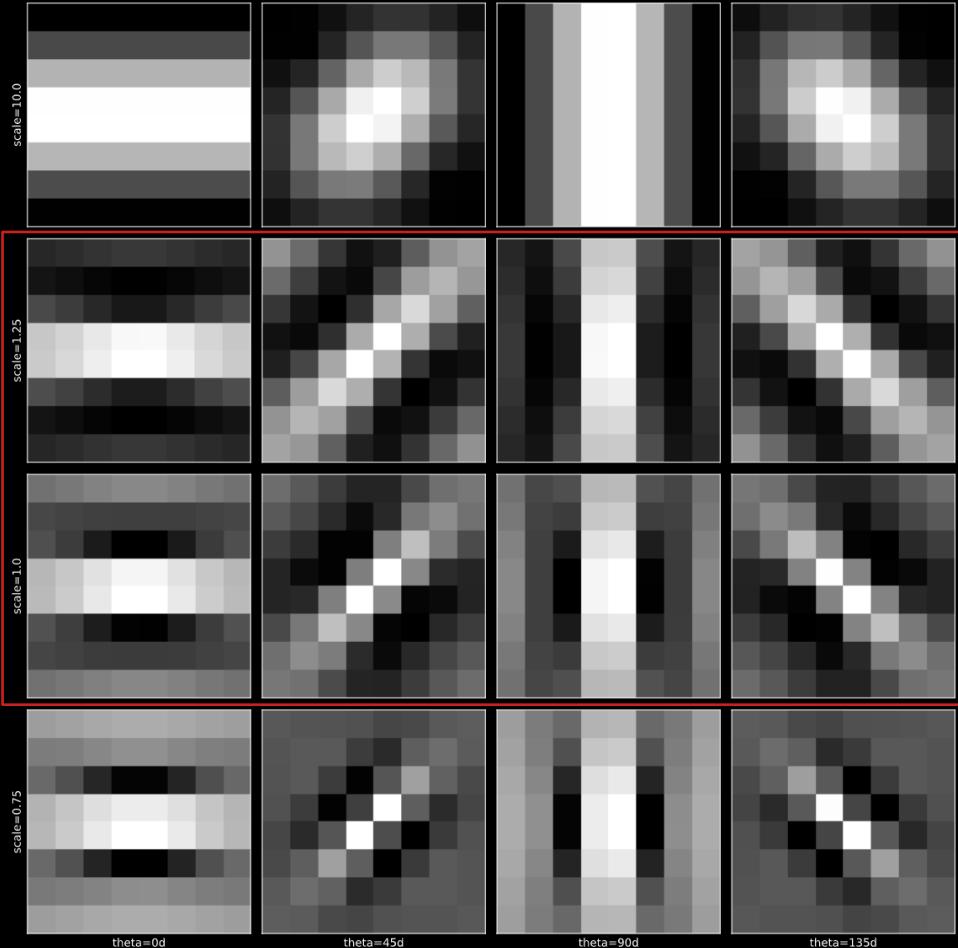
Eksperimen III-A (1/3)

Percobaan ke-	1	2	3
Akurasi tes (%)	<u>65,13</u>	???	???

Implementasi *log-Gabor*

convolutional network

pada set data tanpa *facial
region segmentation.*

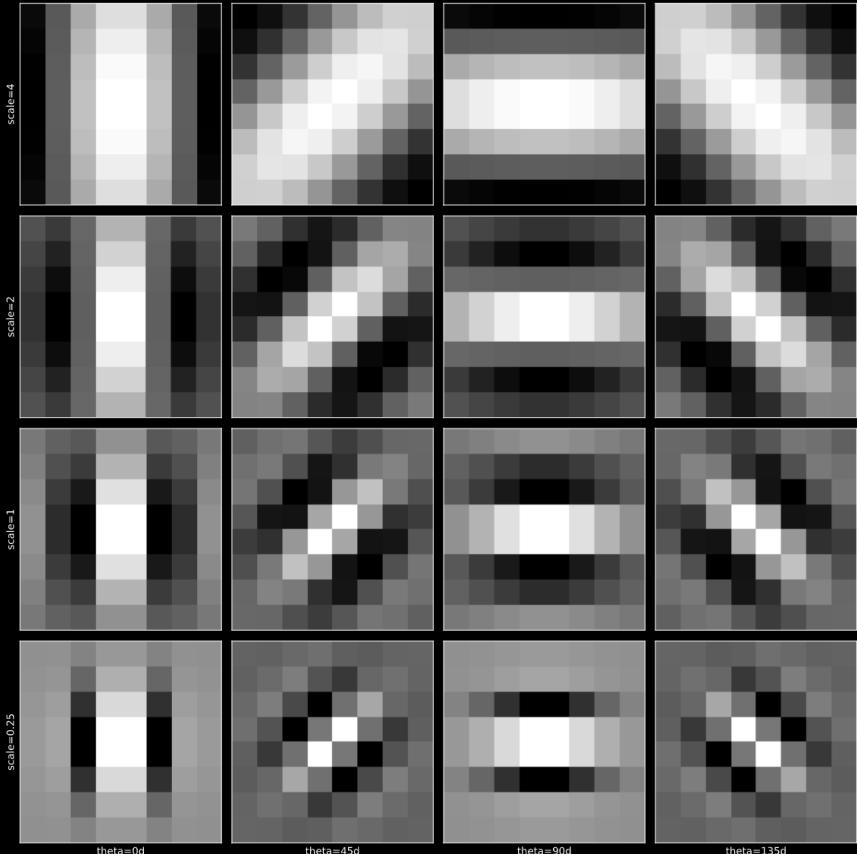
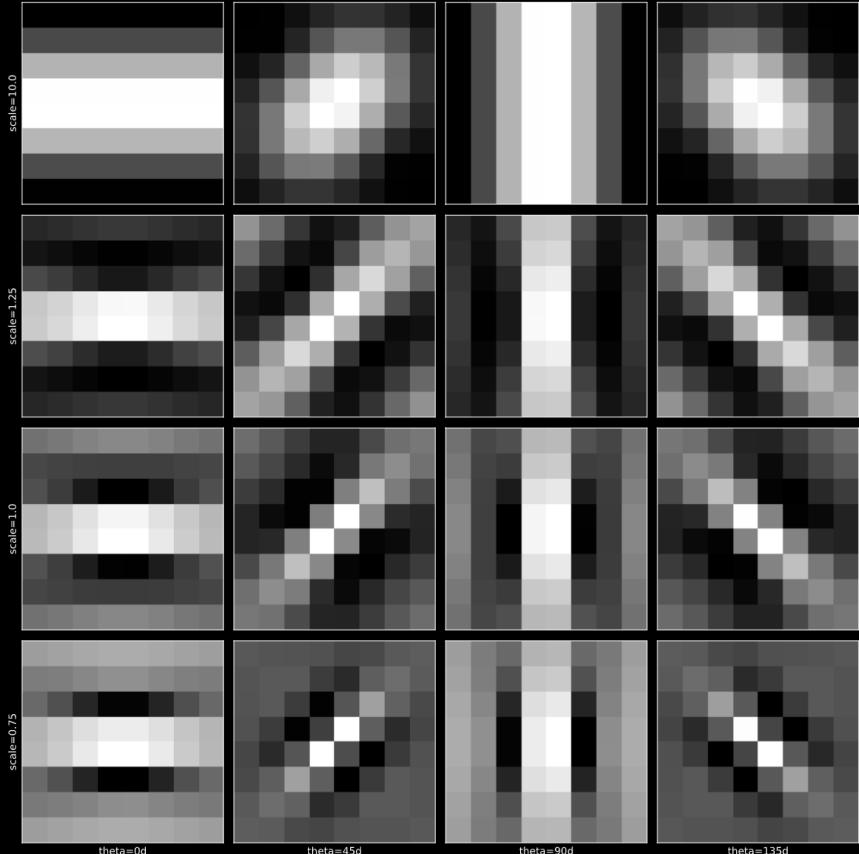


Konfigurasi log-Gabor *convolutional layer*

<u>kernel_size</u>	8
M	4
<u>scale (LGConv<1-4>)</u>	1,25
<u>scale (LGConv<5-8>)</u>	1,25
<u>scale (LGConv<9-12>)</u>	1,0
<u>scale (LGConv<13-16>)</u>	1,0

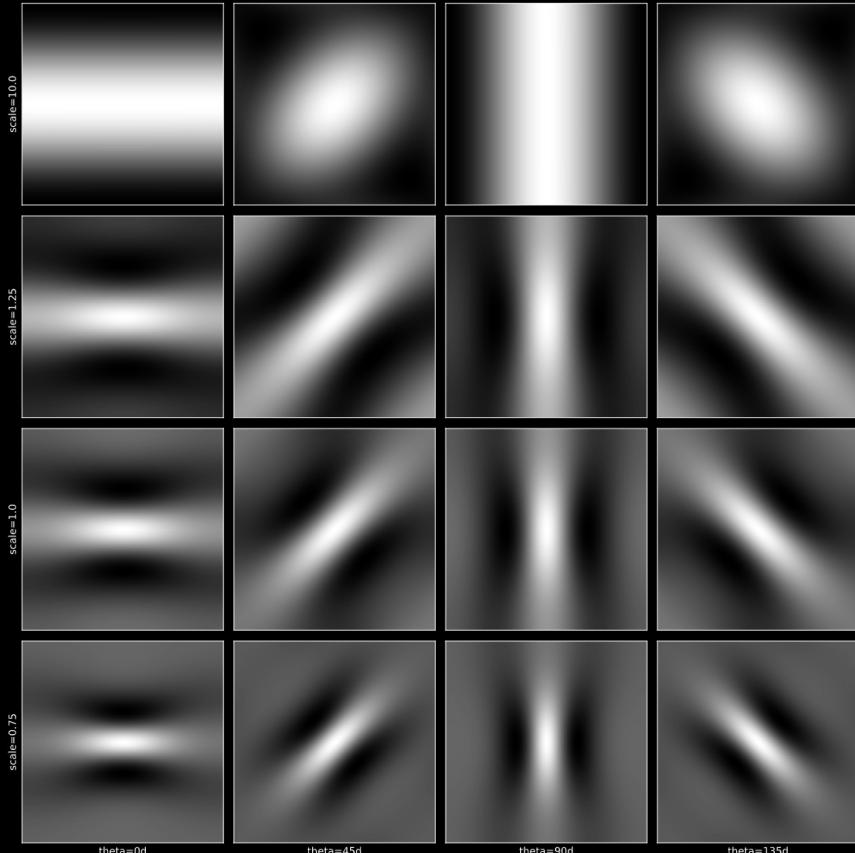
scale — sf_0=0,0075/scale

Log-Gabor versus Gabor (1/2)

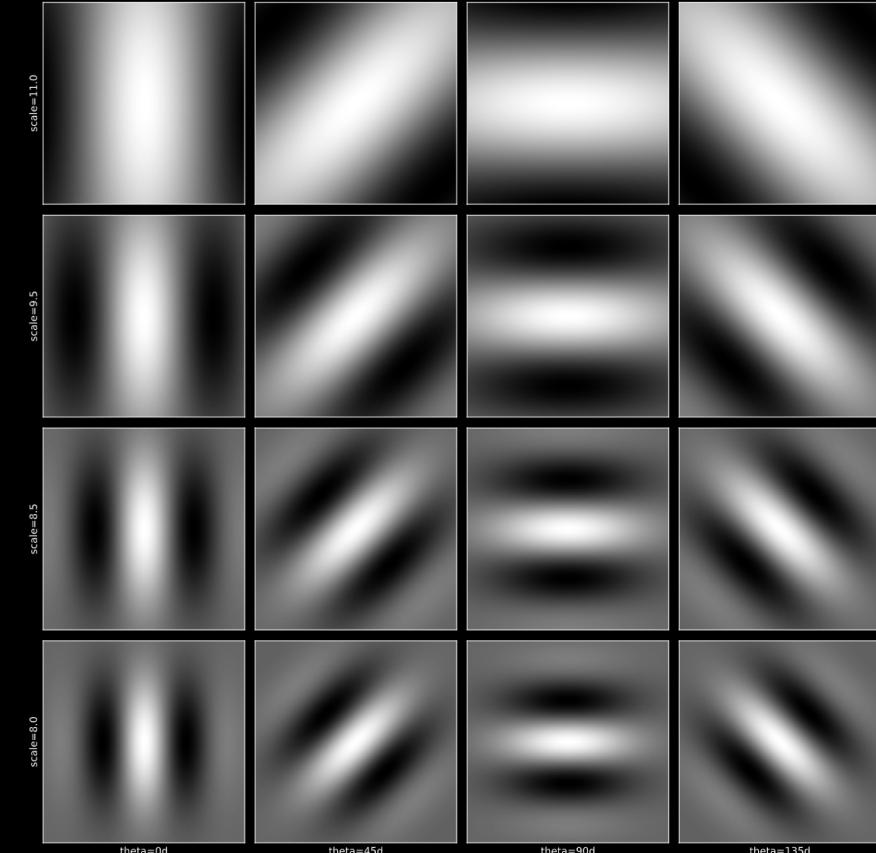


8 x 8

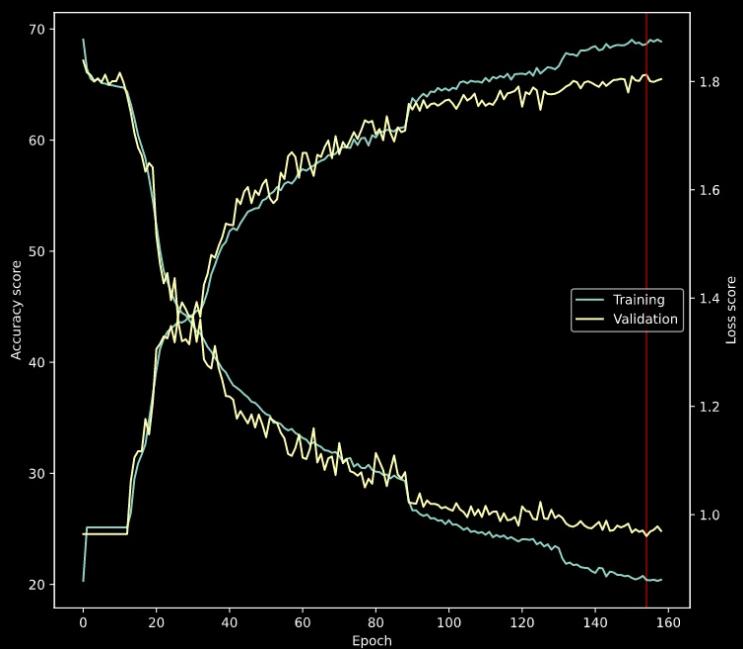
Log-Gabor versus Gabor (2/2)



100 x 100

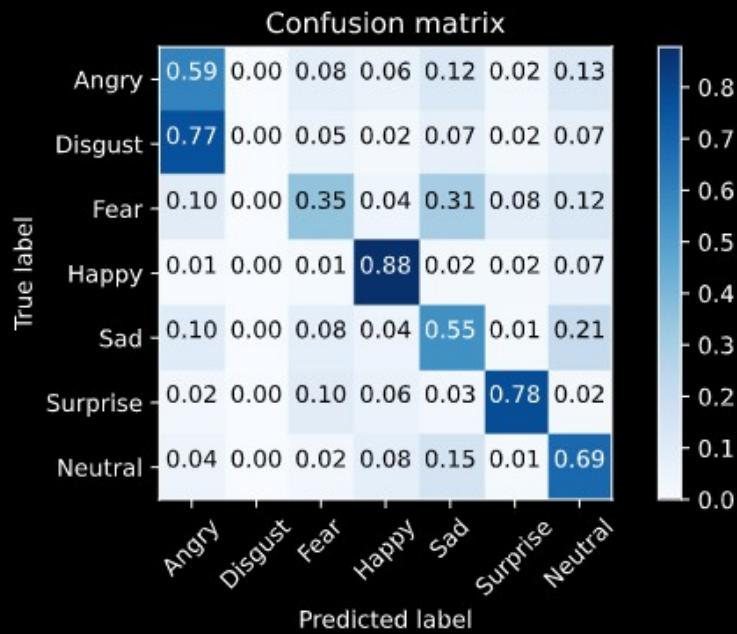


Eksperimen III-A (2/3)



Percobaan ke-	<u>1/x</u>
Waktu <i>training</i> (jam)	<u>8,38</u>
Optimal di epoch ke-	<u>155</u>
Akurasi tes (%)	<u>65,13</u>

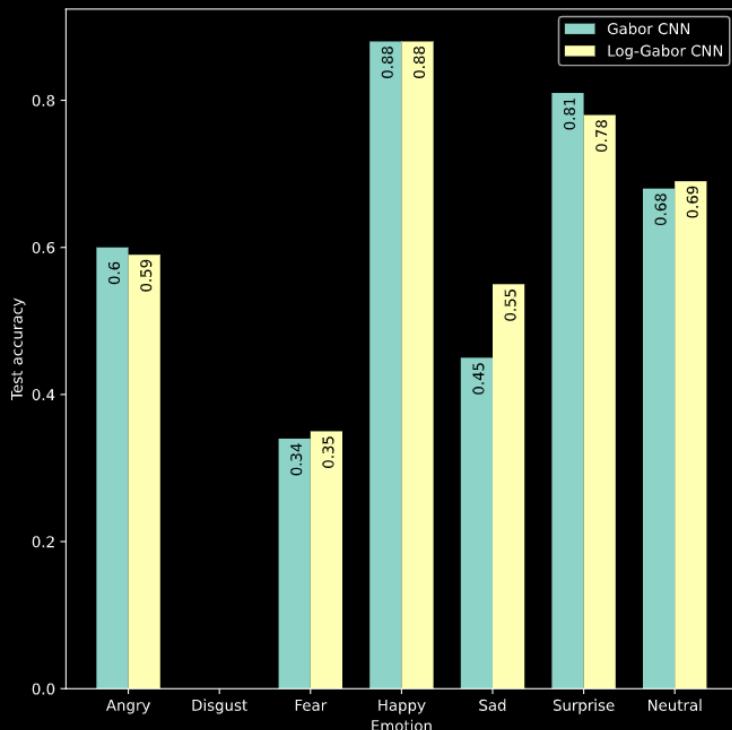
Eksperimen III-A (3/3)



Data tes berlabel “*disgust*”
masih belum ada yang berhasil
dikenali dengan benar.

Tidak ada perubahan yang cukup
berarti.

Eksperimen I-C & III-A



↑1,62% terhadap eksperimen sebelumnya

	Gabor CNN	Log-Gabor CNN
Acc (%)	<u>63,51</u>	<u>65,13</u>

Acc — akurasi test

Kesimpulan

Terbukti bahwa pemanfaatan filter Log-Gabor terhadap filter Gabor dapat meningkatkan performa model.

Lihat di #revisi...



Kesimpulan & saran



Kesimpulan

- 1- Pemodelan pengenalan emosi menggunakan GCNs melalui pendekatan FRS berhasil dirancang dan dieksekusi
- 2- Terbukti bahwa GCN dengan/tanpa FRS dapat meningkatkan performa model *baseline*
- 3- Terbukti bahwa filter log-Gabor dapat meningkatkan performa GCNs

Saran (1/2)

- Beberapa peluang peningkatan metode: (1) penggunaan teknik *facial landmark detection* yang lebih andal untuk set data wajah nonfrontal, (2) penggunaan teknik *image enhancement* yang andal terutama untuk data gambar wajah yang *underbright* atau *overbright*, (3) pemanfaatan teknik *upsampling* terutama untuk data berlabel *disgust*, dan (4) penggunaan teknik *network architecture search*.

Saran (2/2)

- Perbaikan metode pengenalan emosi menggunakan teknik FRS, terutama untuk area wajah bagian atas (dari mata hingga hidung), yang relevan diterapkan pada masa pandemi COVID-19. Misalnya mengadopsi teknik *face frontalization* sebelum FRS.

Thanks!

From Unsplash