

Facial Emotion Image Classification based on Face orientation

By Zhansaya Sovetbek



Submitted to the Department of Data Science
in partial fulfillment of the requirements for the degree of
Master of Science in Data Science
at the NAZARBAYEV UNIVERSITY
April 2023

Thesis Supervisor: Martin Lukac, Associate Professor at
Nazarbayev University, PhD



Outline

- Introduction
- Datasets
- Related Works
- Methodology
- Results
- Conclusion



Introduction



Anger



Fear



Happy



Sad

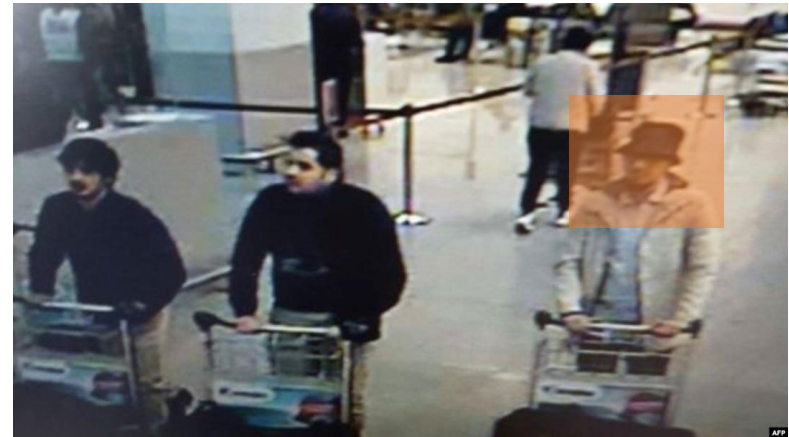
Emotion recognition, also known as affective computing, refers to the ability of machines to detect and interpret human emotions through various methods such as facial expressions, speech patterns, physiological signals, and behavioral cues.

Introduction

Motivation: Emotion Recognition could be used to determine, does a human need medical help, better understand people with emotional disabilities, and detect any suspicious movements in public places. Solving this problem requires advanced machine learning techniques, innovative algorithms, and a deep understanding of human emotions.

Problem Statement: The task of classifying facial emotions in images based on face orientation is a complex challenge in computer vision, machine learning, and deep learning. It involves accurately recognizing the emotions portrayed in an image while considering the angle and position of the face, as such variations can affect the accuracy of recognition.

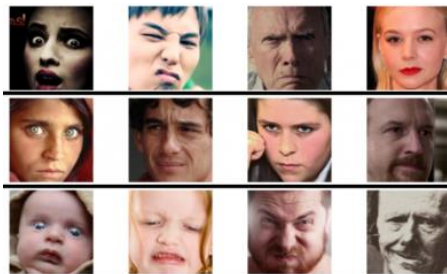
Solution: to explore state-of-the-art recognition works and investigate the impact of both facial orientation and emotional category on the accuracy of recognition.



Datasets

AffectNet

- It is a new face expression database that was created by gathering and annotating facial photos
- Overall: 420299 images
- In research, I extracted 291,651 images to work with



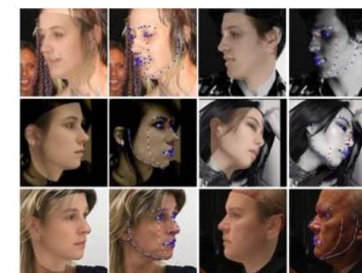
Pointing'04

- head-position image database, a collection of images, was created to help with head pose estimate studies in the field of computer vision.
- 15000 images



AFLW2000-3D

- a dataset of images that have been annotated with image-level 68-point 3D facial landmarks.
- 2000 images



Imbalanced Dataset

Direction	Anger	Contempt	Disgust	Fear	Happy	Neutral	Sad	Surprise
Forward	1504	1334	1454	2052	1661	1823	1554	1769
Left	411	310	371	361	347	397	391	412
Right	367	251	337	281	300	380	387	376
Up	788	399	750	1058	553	706	972	1005
Down	1194	1323	709	507	1356	913	896	663
Overall	4264	3617	3621	4259	4217	4219	4200	4225

Table 3.1: Raw Imbalanced Affectnet distribution

Direction	Anger	Contempt	Disgust	Fear	Happy	Neutral	Sad	Surprise
Forward	35%	37%	40%	48%	39%	43%	37%	42%
Left	10%	9%	11%	8%	8%	10%	10%	10%
Right	9%	7%	10%	7%	7%	9%	9%	9%
Up	18%	11%	20%	25%	13%	17%	23%	24%
Down	28%	36%	19%	12%	33%	21%	21%	15%
Overall	4264	3617	3621	4259	4217	4219	4200	4225

Table 3.2: Percentage Distribution of Raw Imbalanced dataset

Neutral	75374
Happy	134915
Sad	25959
Surprise	14590
Fear	6878
Disgust	4303
Anger	25382
Contempt	4250
None	33588
Uncertain	12145
Non-Face	82915
Total	420299

AffectNet

Balanced Dataset For
Baseline Model
Overall: 5600 images
Train: 3200 images
Val: 1600 images
Test: 800 images

Balanced Dataset For
Proposed Model
Overall: 11240 images
Train: 8840 images
Val: 1200 images
Test: 1200 images

Dataset was increased 2 times



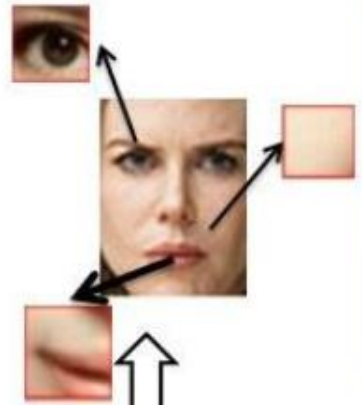
Related Works

The Facial Emotion Detection Problems



Large intra-class
variances

1



Facial Emotion
Feature
Extraction

2

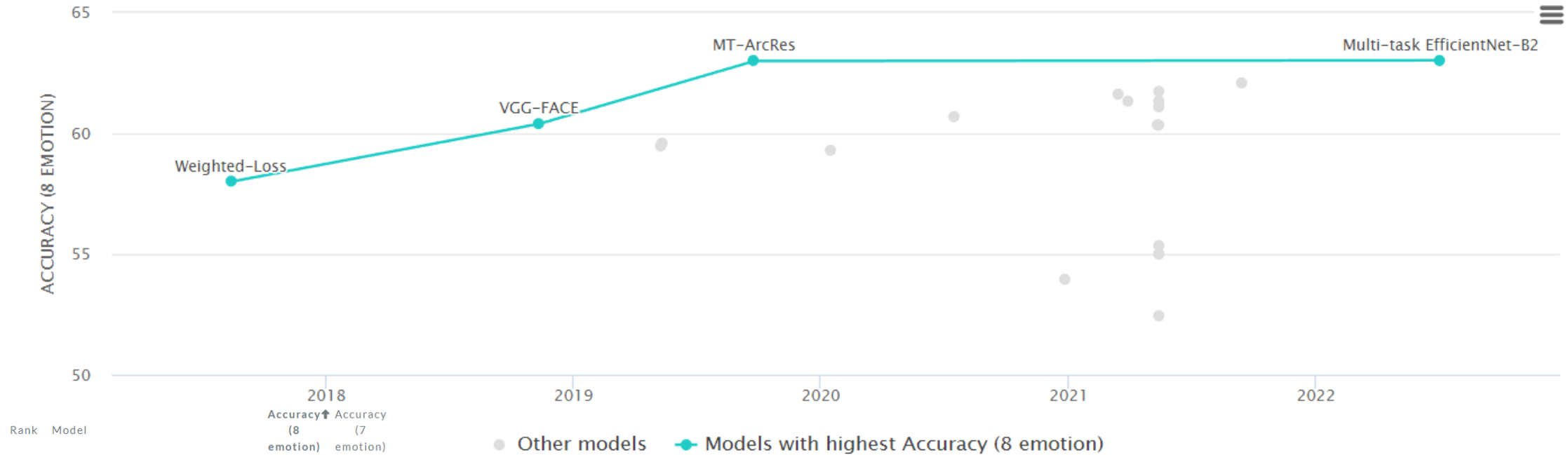


Image
segmentation

3

Related Works

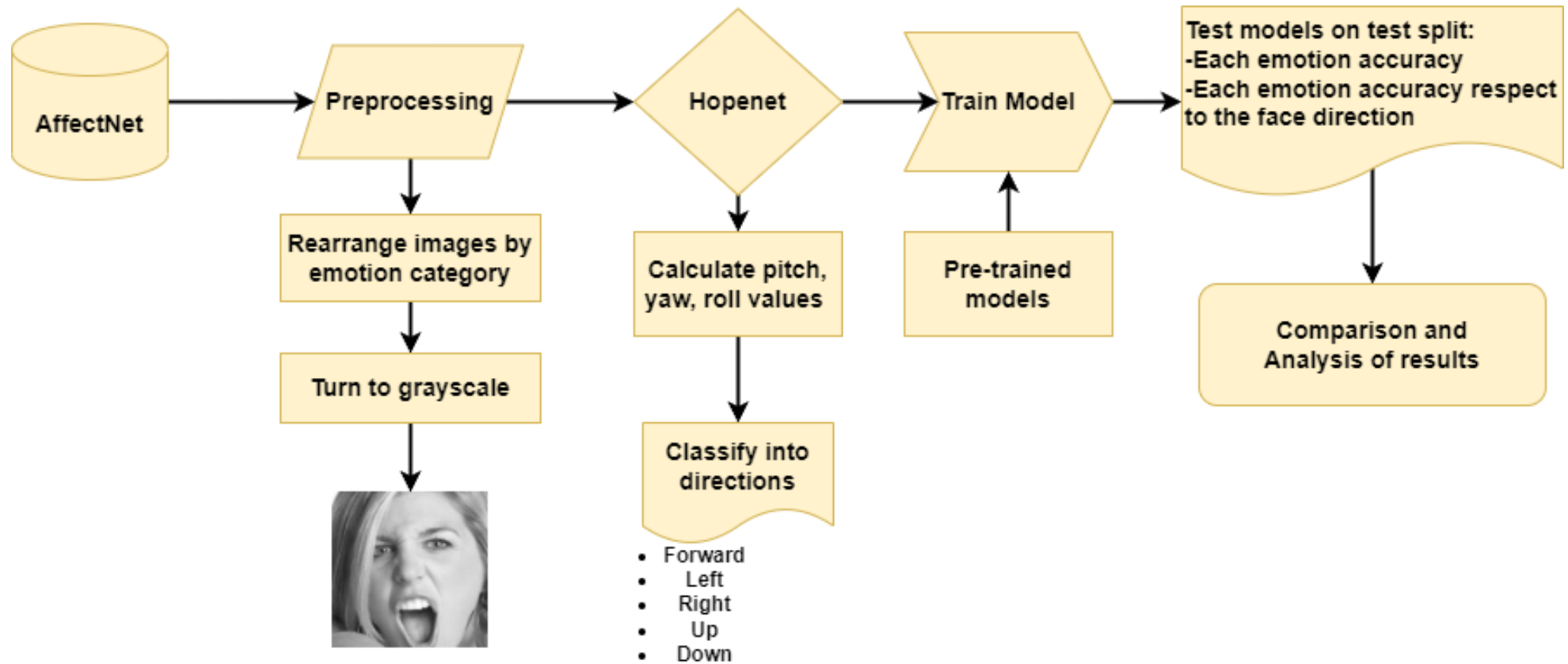
Models on AffectNet



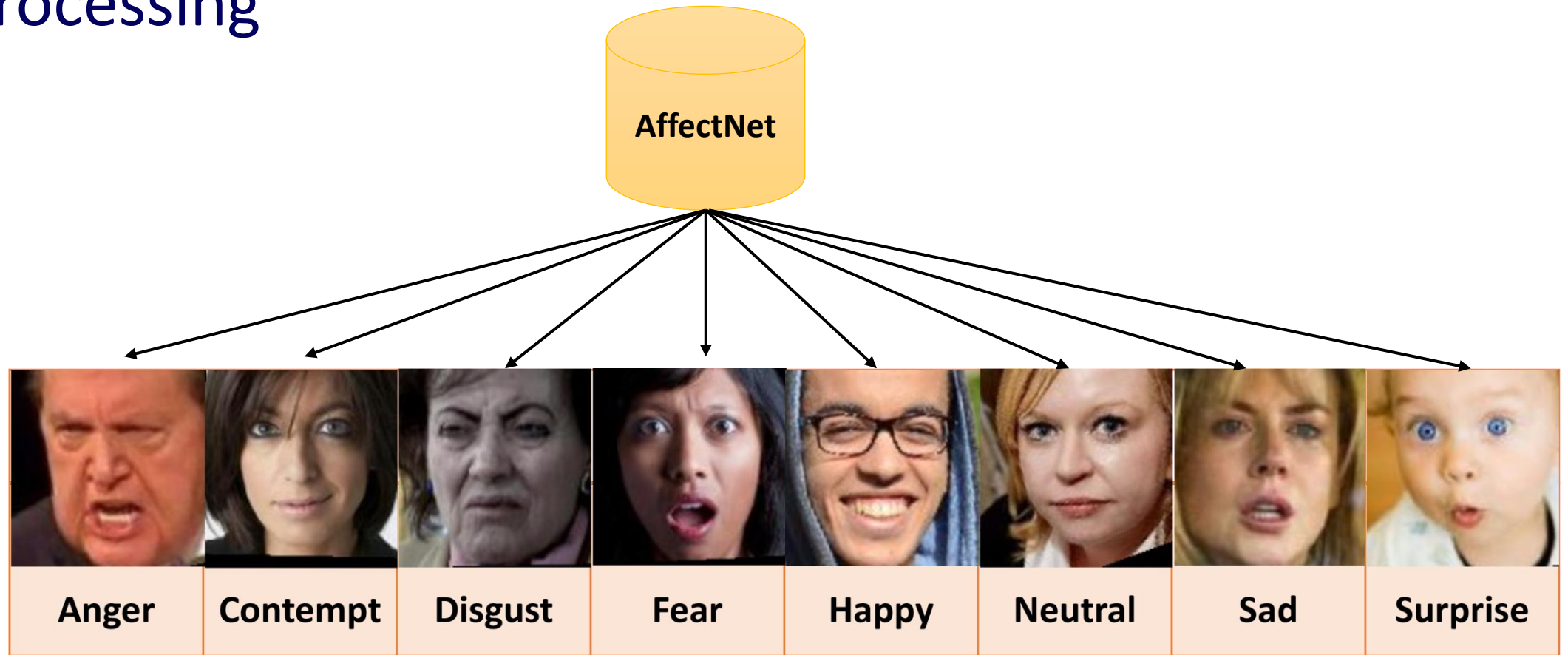
Rank	Model	Accuracy (8 emotion)	Accuracy (7 emotion)
1	Multi-task EfficientNet-B2	63.03	66.29
2	MT-ArcRes	63	
3	DAN	62.09	65.69

Methodology

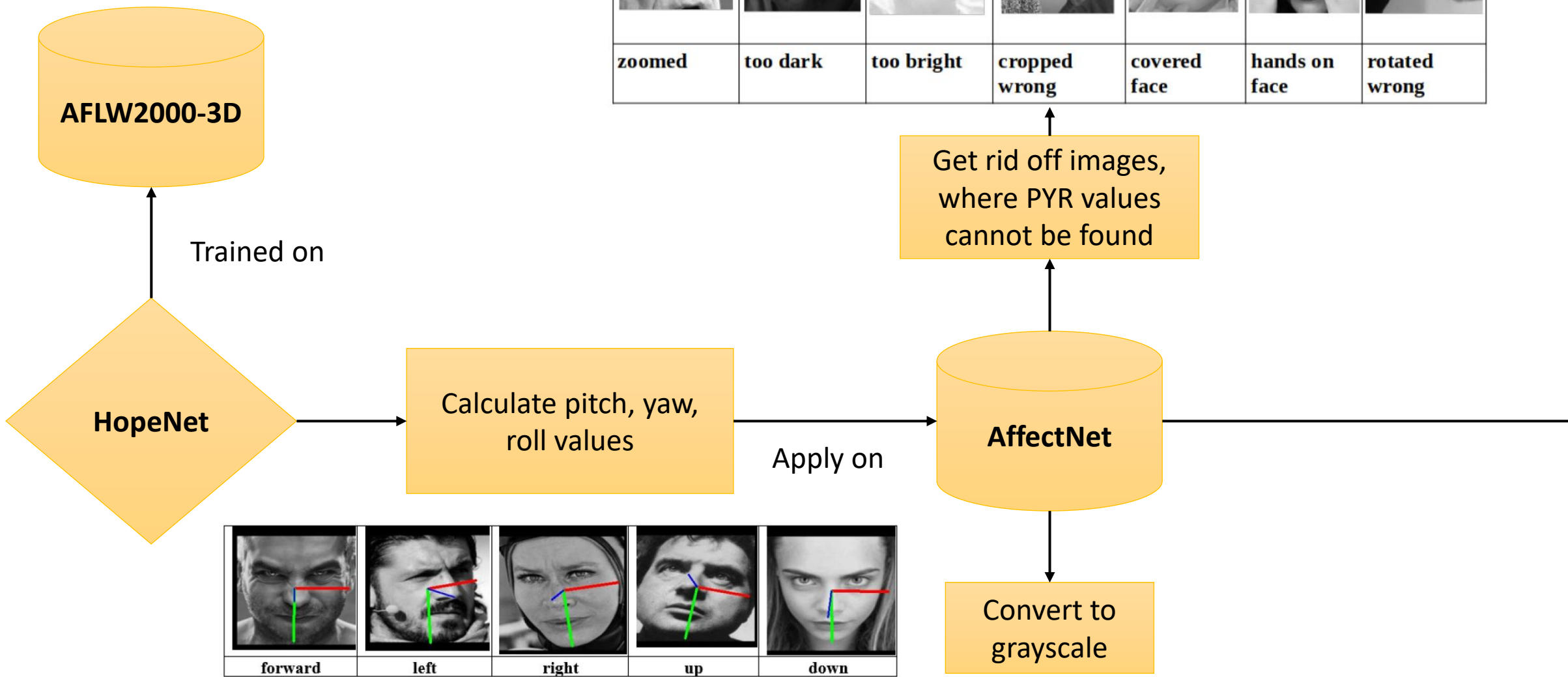
Thesis Workflow










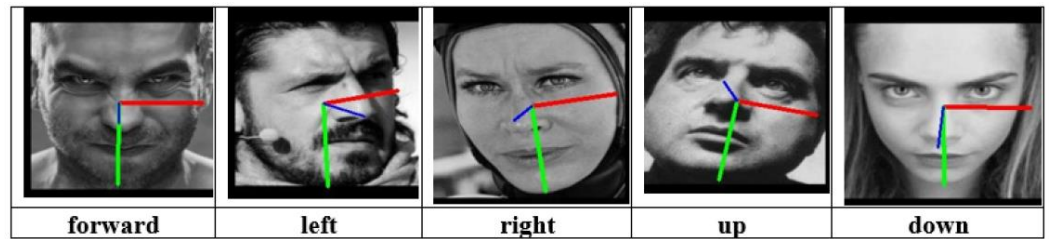
Preprocessing



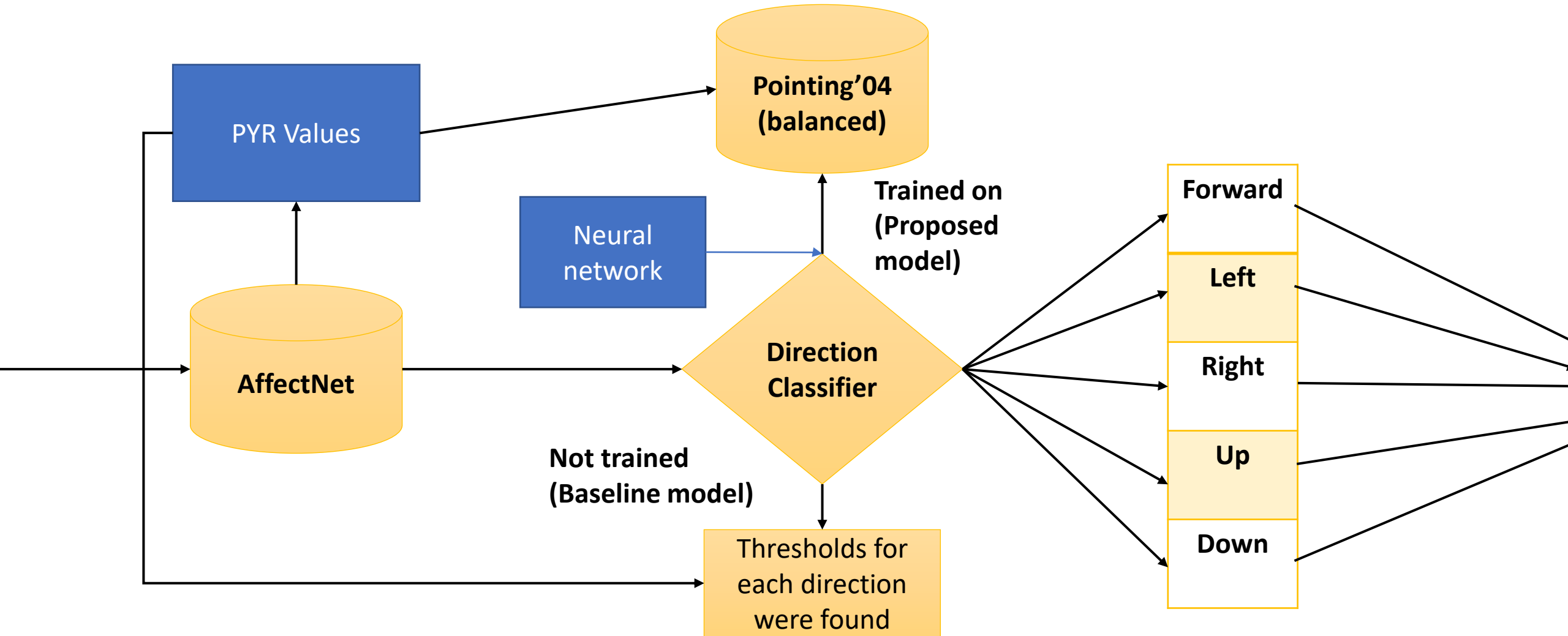
Preprocessing



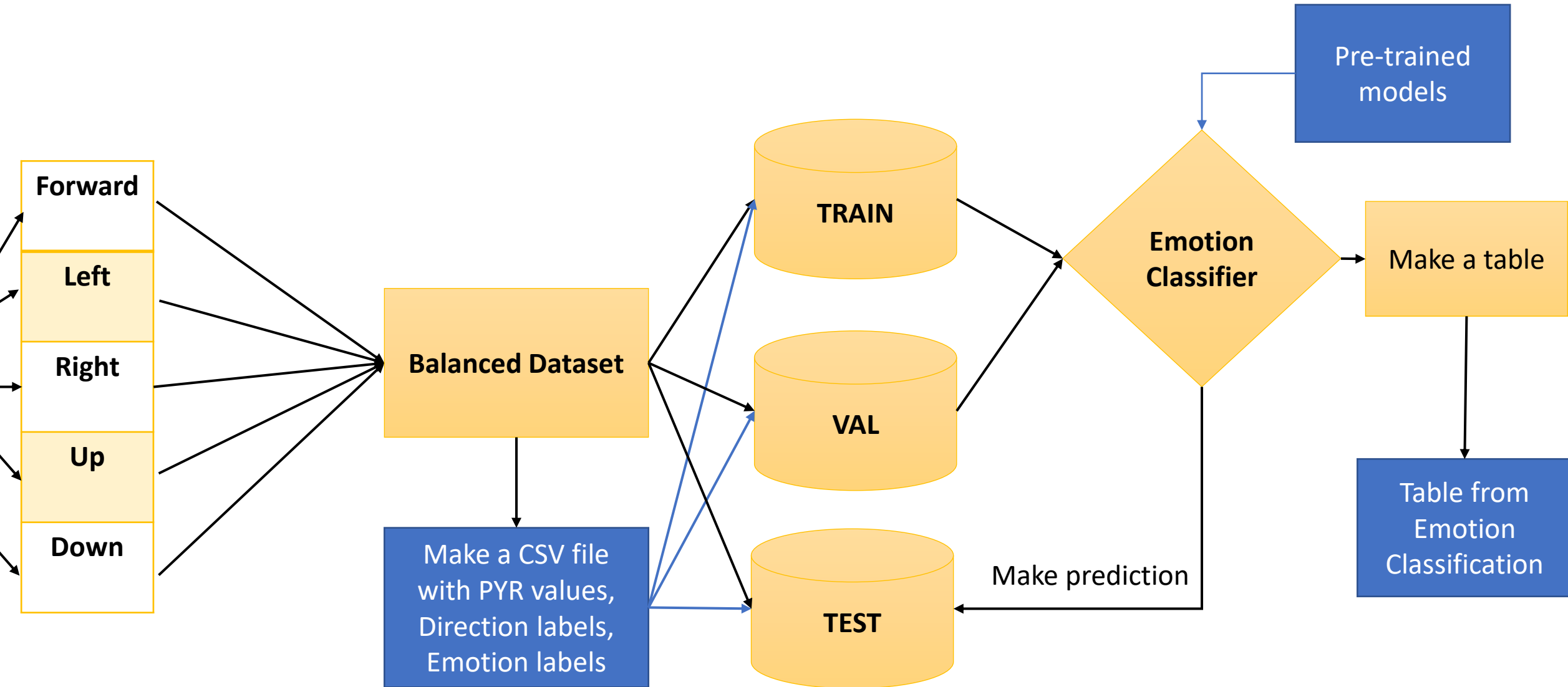
						
zoomed	too dark	too bright	cropped wrong	covered face	hands on face	rotated wrong



Direction Classification



Emotion Classification



Direction Classification models

Model	Accuracy
Neural Network	85%
Googlenet	84%
MobileNetV3-small	83%
ResNet-18	82%

Table 3.3: Overall Direction Classification by each model

- 0-forward
- 1-left
- 2-right
- 3-up
- 4-down

	A	B	C	D	E	F
1	Unnamed: 0	Filename	Pitch	Yaw	Roll	Direction
2	0	266849.jpg	-9.3948	-4.8767	1.6852	0
3	1	70015.jpg	5.1716	-13.1784	9.0178	3
4	2	163428.jpg	-1.4012	4.5605	1.1525	0
5	3	238887.jpg	-14.5345	-3.0493	-0.1093	4
6	4	405896.jpg	1.8919	-4.3922	-6.7957	0
7	5	53269.jpg	-12.8206	-2.9103	-0.9192	4
8	6	288337.jpg	-8.3703	-7.2591	-5.1563	0
9	7	134314.jpg	-4.2796	-14.2695	2.6693	1
10	8	20279.jpg	4.9284	-9.6153	-1.7466	0
11	9	306105.jpg	-1.7223	-4.1033	-4.5332	0
12	10	19804.jpg	-7.1609	-4.0537	-2.464	0
13	11	249494.jpg	2.5132	2.1206	0.855	0
14	12	403241.jpg	-11.1235	16.0069	2.6036	4
15	13	114570.jpg	-2.6276	-7.5053	7.0635	1
16	14	184964.jpg	-11.1139	-4.4037	-6.2698	4
17	15	40039.jpg	-14.6325	-7.7016	-2.5477	4
18	16	231956.jpg	2.5908	-0.984	-1.2828	0

Predict directions

370	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/195402.jpg	0.712	-5.8267	1.9993	0
371	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/271244.jpg	-14.1771	-23.6118	-12.817	1
372	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/102291.jpg	-6.2657	-6.8625	4.0954	1
373	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/357219.jpg	1.0945	-14.0032	-3.2365	1
374	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/56704.jpg	-20.8301	6.0202	4.7611	4
375	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/235025.jpg	5.1416	-6.3639	6.3499	0
376	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/412851.jpg	1.6603	7.2143	-10.4206	3
377	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/348412.jpg	-6.2151	-10.2315	4.9976	3
378	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/235285.jpg	-11.0819	3.8733	-2.1519	0
379	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/39516.jpg	-1.1546	7.8054	9.3022	3
380	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/80838.jpg	-13.7984	22.726	13.2071	4
381	/mnt/nas4/Zhansaya_affectnet/train_set/balanced/train/Anger/255915.jpg	-11.3505	34.5335	4.0163	2

Baseline model vs Proposed model Emotion Classification

Baseline Model

Pre-trained model	Accuracy	F1 score
MobileNetV3-small	46%	0.4623
Googlenet	43%	0.4285
ResNet-18	17%	0.1395
VGG-16	13%	0.0554
Alexnet	12%	0.0278
AdaBoost	12%	0.0277

Proposed Model

Pre-trained model	Accuracy	F1 score
MobileNetV3-small	52%	0.5177
Googlenet	52%	0.5160
ResNet-18	49%	0.4920
AdaBoost	16%	0.1658
VGG-16	12%	0.1277
Alexnet	12%	0.1276

	A	B	C	D	E	F	G
1	Unnamed: 0	Filename	Pitch	Yaw	Roll	Direction	Emotion
2	0	266849.jpg	-9.3948	-4.8767	1.6852	0	anger
3	1	70015.jpg	5.1716	-13.1784	9.0178	3	happy
4	2	163428.jpg	-1.4012	4.5605	1.1525	0	happy
5	3	238887.jpg	-14.5345	-3.0493	-0.1093	4	surprise
6	4	405896.jpg	1.8919	-4.3922	-6.7957	0	sad
7	5	53269.jpg	-12.8206	-2.9103	-0.9192	4	sad
8	6	288337.jpg	-8.3703	-7.2591	-5.1563	0	neutral
9	7	134314.jpg	-4.2796	-14.2695	2.6693	1	disgust
10	8	20279.jpg	4.9284	-9.6153	-1.7466	0	contempt
11	9	306105.jpg	-1.7223	-4.1033	-4.5332	0	disgust
12	10	19804.jpg	-7.1609	-4.0537	-2.464	0	fear
13	11	249494.jpg	2.5132	2.1206	0.855	0	surprise
14	12	403241.jpg	-11.1235	16.0069	2.6036	4	sad
15	13	114570.jpg	-2.6276	-7.5053	7.0635	1	happy
16	14	184964.jpg	-11.1139	-4.4037	-6.2698	4	disgust
17	15	40039.jpg	-14.6325	-7.7016	-2.5477	4	happy
18	16	231956.jpg	2.5908	-0.984	-1.2828	0	contempt

Baseline Model Accuracy and Loss Diagrams

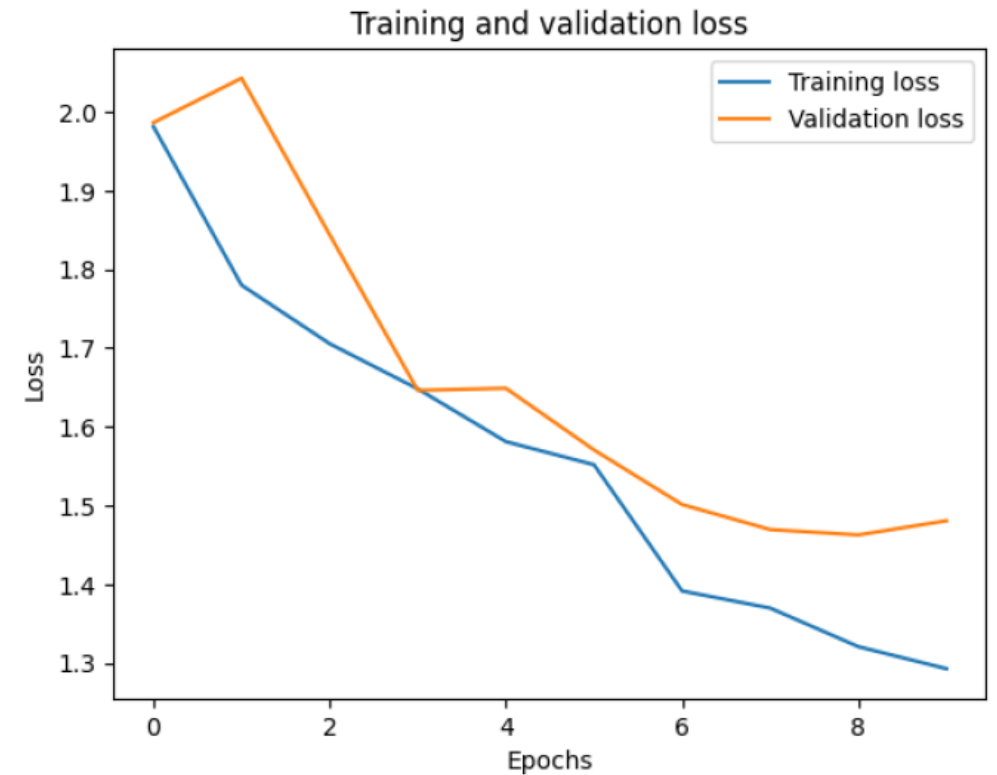
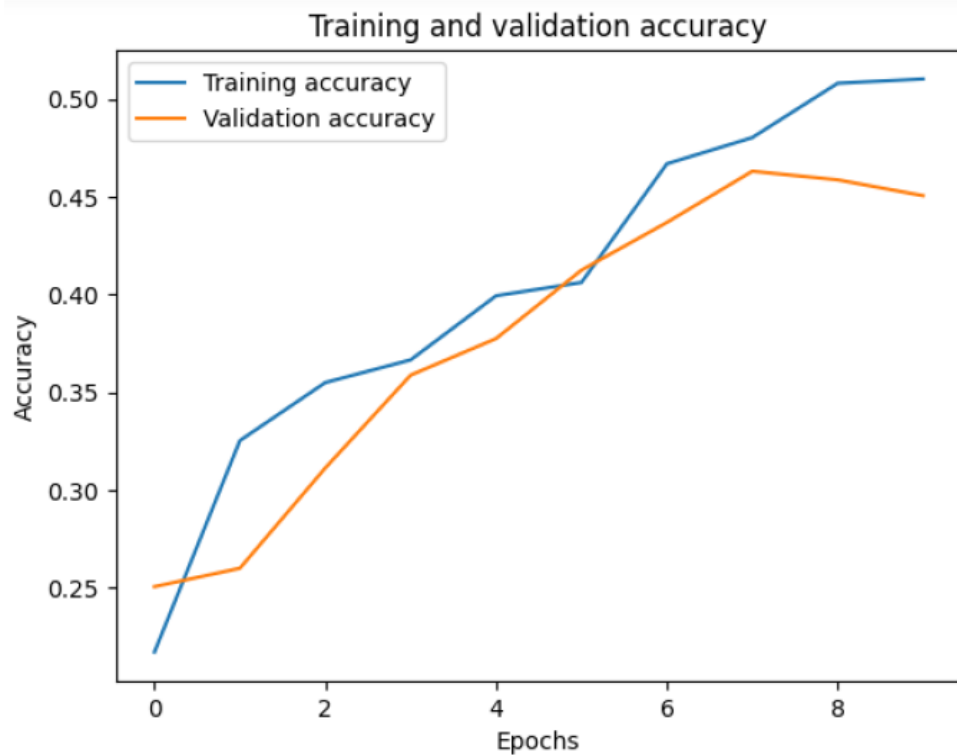


Figure 4-3: Baseline model Accuracy Diagram for MobileNetV3-small

Figure 4-2: Baseline model Loss Diagram for MobileNetV3-small

Proposed Model Accuracy and Loss Diagrams

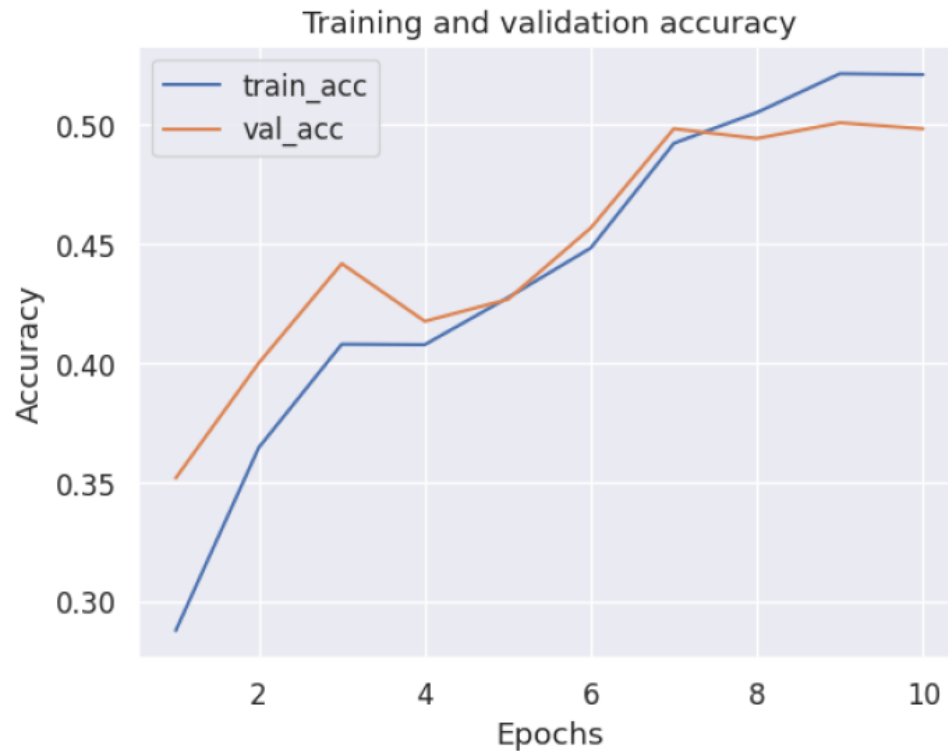
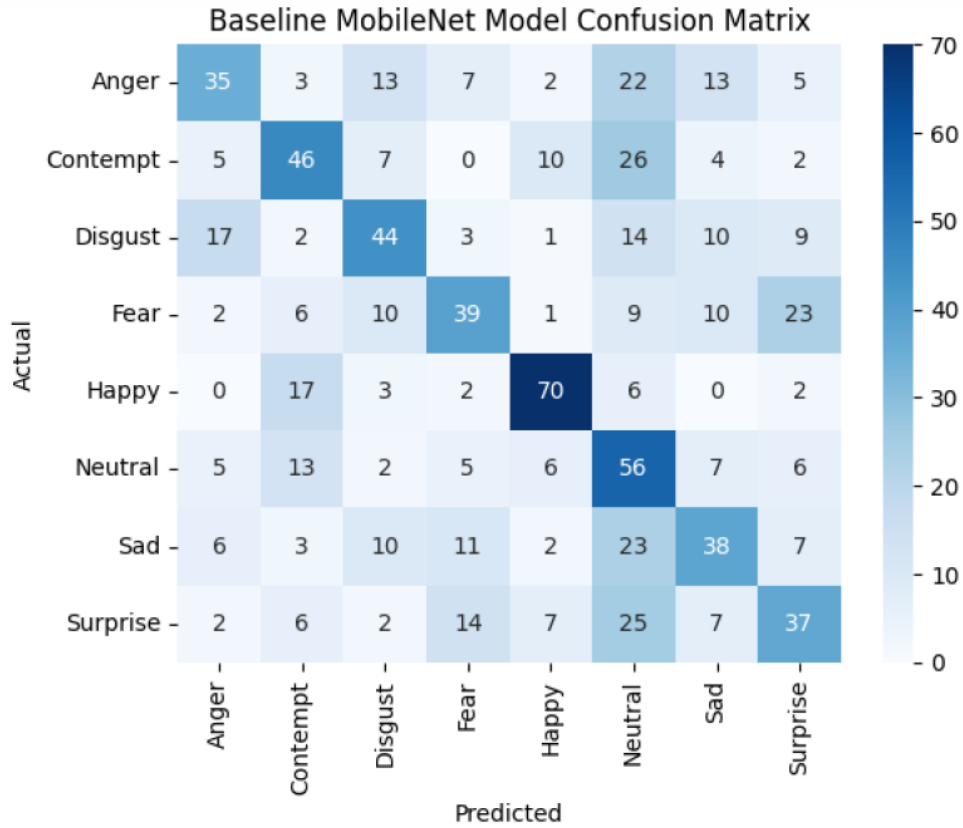


Figure 4-5: Proposed model Accuracy Diagram for MobileNetV3-small Figure 4-4: Proposed model Accuracy Diagram for MobileNetV3-small

Baseline Model Results

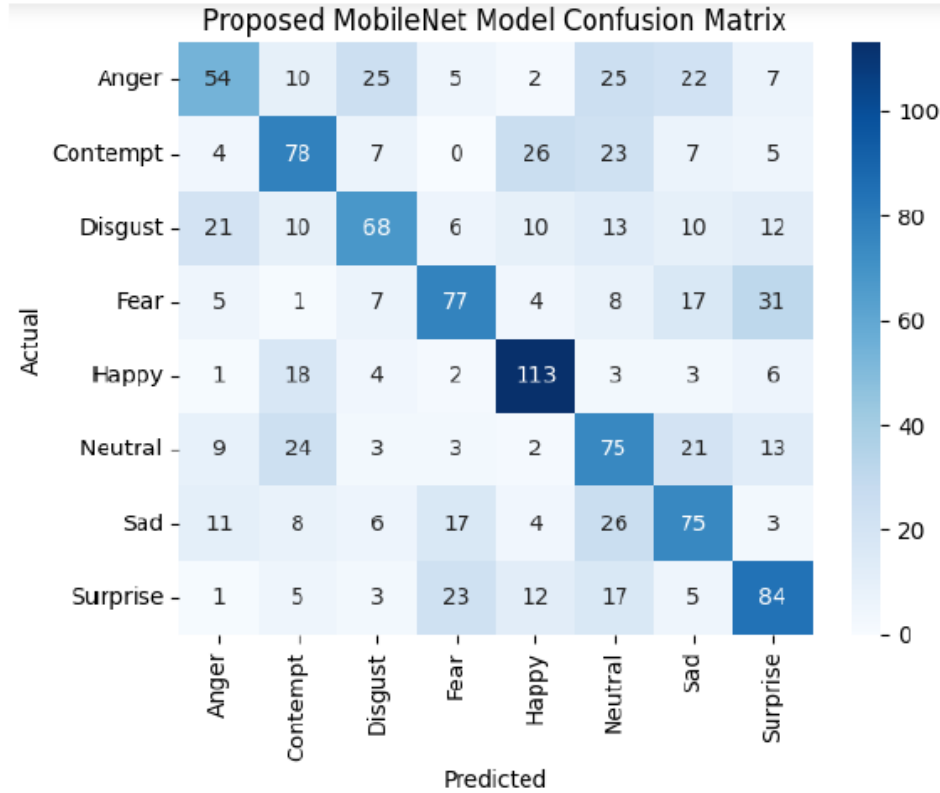


Direction	Anger	Contempt	Disgust	Fear	Happy	Neutral	Sad	Surprise	mean	var	std
Forward (20)	8	6	9	5	17	10	9	6	8.75	12.4375	3.526
Left(20)	5	13	9	9	13	13	6	8	9.5	9.0	3.0
Right(20)	10	13	9	7	12	11	9	9	10.0	3.25	1.802
Up(20)	5	7	8	11	14	10	8	6	8.625	7.4843	2.735
Down(20)	7	7	9	7	14	12	6	8	8.75	6.9375	2.633
mean	7.0	9.2	8.8	7.8	14.0	11.2	7.6	7.4			
variance	3.6	9.76	0.159	4.16	2.8	1.36	1.839	1.44			
st deviation	1.897	3.124	0.399	2.039	1.673	1.166	1.356	1.2			
Emotion accuracy	35/100	46/100	44/100	39/100	70/100	56/100	38/100	37/100			

Table 4.2: Direction by Emotion true labeled table

Figure 4-1: Baseline model Confusion matrix for MobileNetV3-small

Proposed Model Results



Direction	Anger	Contempt	Disgust	Fear	Happy	Neutral	Sad	Surprise	mean	var	std
Forward											
(30)											
Left(30)	14	19	12	13	24	11	14	20	15.87	18.35	4.28
Right(30)	12	15	14	17	16	18	17	18	15.87	3.85	1.96
Up(30)	8	16	16	18	24	16	14	15	15.87	17.10	4.13
Down(30)	11	15	14	14	25	14	12	14	14.87	16.10	4.01
mean	9	13	12	15	24	16	18	17	15.5	17.75	4.21
variance	10.8	15.6	13.6	15.4	22.6	15.0	15.0	16.8			
st deviation	4.56	3.84	2.23	3.44	11.039	5.6	4.8	4.56			
st deviation	2.13	1.95	1.49	1.85	3.32	2.36	2.19	2.13			
Emotion accuracy	54/150	78/150	68/150	77/150	113/150	75/150	75/150	84/150			

Table 4.6: Direction by Emotion true labeled table for Proposed Model

Figure 4-6: Proposed model Accuracy Diagram for MobileNetV3-small

Baseline Model vs Proposed Model

Baseline Model

Direction	Anger	Contempt	Disgust	Fear	Happy	Neutral	Sad	Surprise
Forward	23%	14%	20,4%	13%	25%	18%	24%	16%
Left	14%	28%	20,4%	23%	18%	23%	15%	22%
Right	28%	28%	20,4%	18%	17%	20%	24%	24%
Up	14%	15%	18,4%	28%	20%	18%	21%	16%
Down	20%	15%	20,4%	18%	20%	21%	16%	22%
Emotion accuracy	35%	46%	44%	39%	70%	56%	38%	37%

Table 4.3: The highest percentage of direction per each emotion(by column)

Proposed Model

Direction	Anger	Contempt	Disgust	Fear	Happy	Neutral	Sad	Surprise
Forward	26%	24%	18%	17%	21%	15%	19%	24%
Left	22%	19%	20%	22%	14%	24%	23%	21%
Right	15%	21%	24%	23%	21%	21%	19%	18%
Up	20%	19%	20%	18%	23%	19%	15%	17%
Down	17%	17%	18%	20%	21%	21%	24%	20%
Emotion accuracy	36%	52%	45%	51%	75%	50%	50%	56%

Table 4.7: The highest percentage of direction per each emotion(by column)

Visualization of Proposed Model Results

Anger- Forward	Contempt- Forward	Surprise- Forward	Neutral- Left	Disgust- Right	Fear- Right	Happy- Up	Sad- Down
							

Conclusion

- There were two approaches: Baseline Model and the Proposed Model
- All training, validation, and testing were preformed on balanced datasets.
- With the increase in the size of the dataset, the accuracy of pre-trained models also increased.
- Even though the percentage distribution of directions is almost the same, we can see that some directions suppress others, and could predict emotion.
- Future work: to increase the dataset size and try other machine learning or deep learning models for direction and emotion classification

Thank you for your attention!

