Department of Computer Science & Engineering

Rajshahi University of Engineering & Technology

Thesis Title:

Facial Expression Recognition Based on LBP and CNN: A Comparative Study Using SVM Classifier

Presented By:

Mir Mursalin Hossain
Roll no: 133047
Dept. of CSE, RUET

Supervised By:

Prof. Dr. Md. Rabiul Islam
Head,
Dept. of CSE, RUET

Outlines

1. Facial Expression Recognition
2. Steps of Facial Expression Recognition Model
3. LBP as Feature Extractor
4. CNN as Feature Extractor
5. Classifier: Support Vector Machine (SVM)
6. Literature Review
7. Work Flow
8. Dataset
9. Result & Comparative Analysis
10. Conclusion
11. Future Scopes
12. References

Facial Expression Recognition

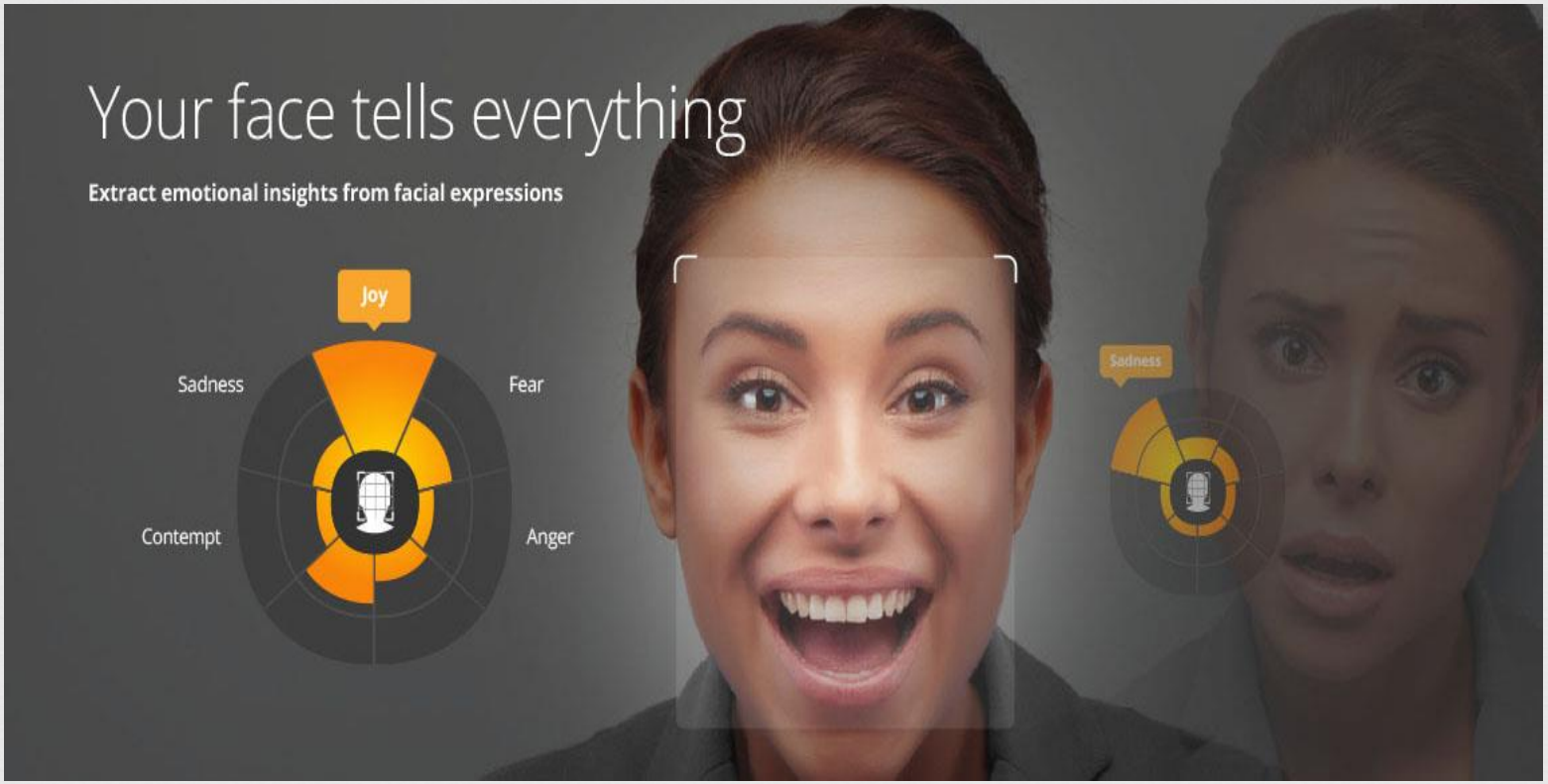


Figure 1: Emotion Analysis from Facial Expression [1]

Facial Expression Recognition (Cont.)

- Human emotion can be recognized through facial expression.
- Facial Expression Recognition (FER) system is a computer application for automatically identifying or verifying people's emotions reflected on their faces from a digital image or a video frame by comparing it with database.

Steps of Facial Expression Recognition Model

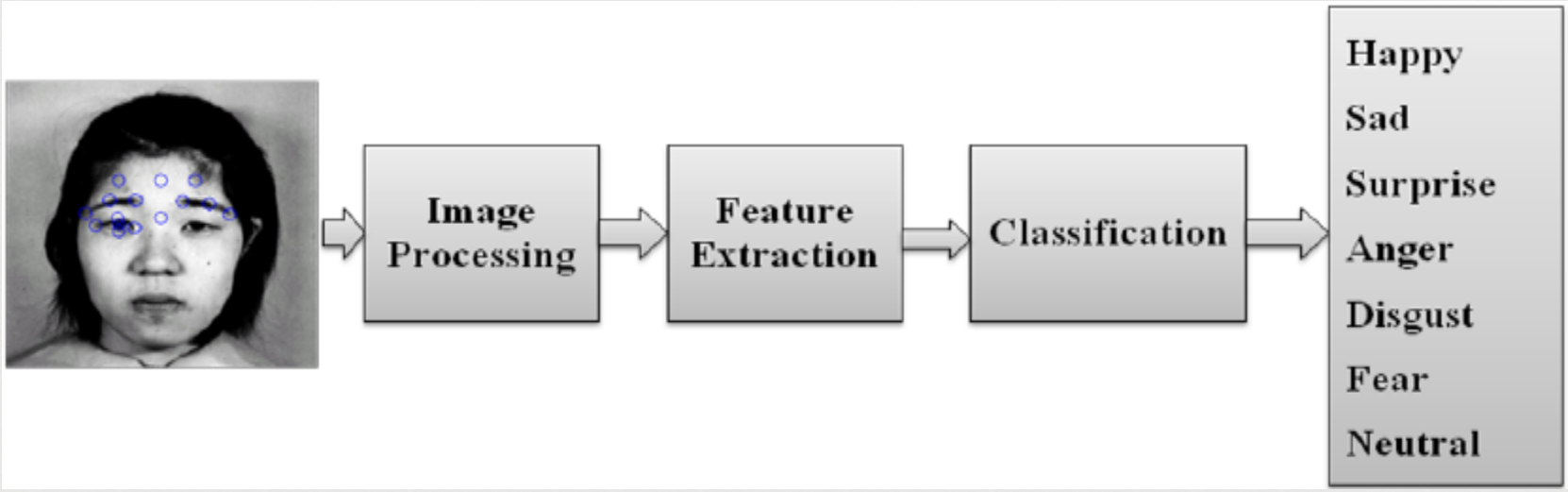


Figure 2: Basic Steps of Facial Expression Recognition Model [2]

LBP as Feature Extractor

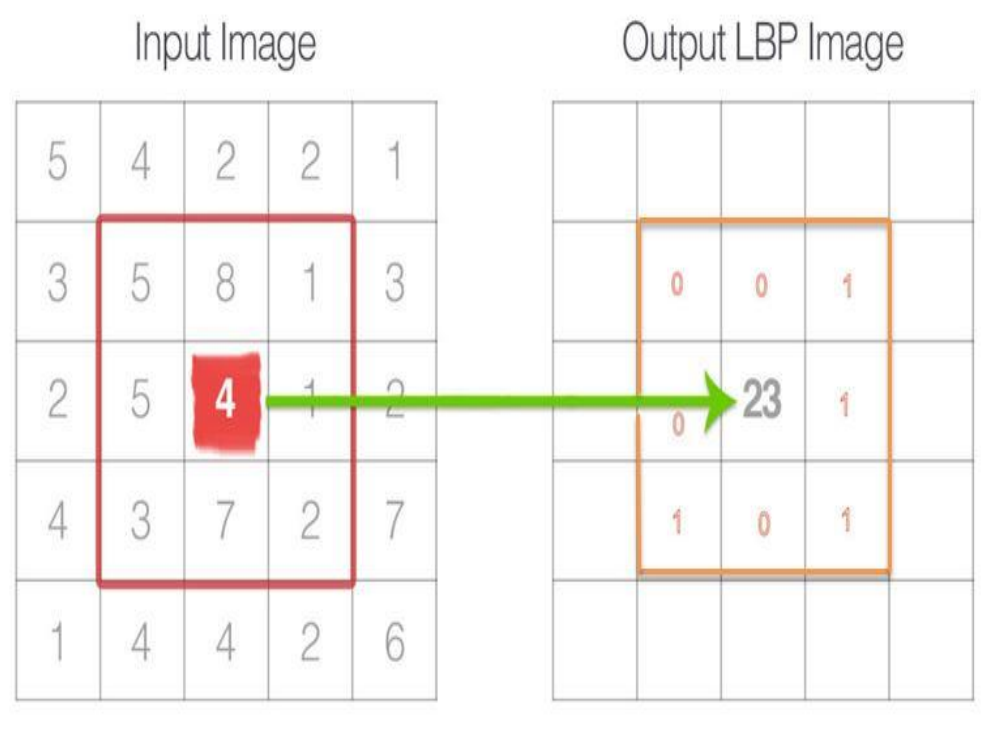


Figure 3: Calculation of LBP value from input image

- LBP features found from LBP image histogram are easy to compute.
- It can handle variable neighborhood sizes.
- It has tolerance against illumination changes. [3]

CNN as Feature Extractor

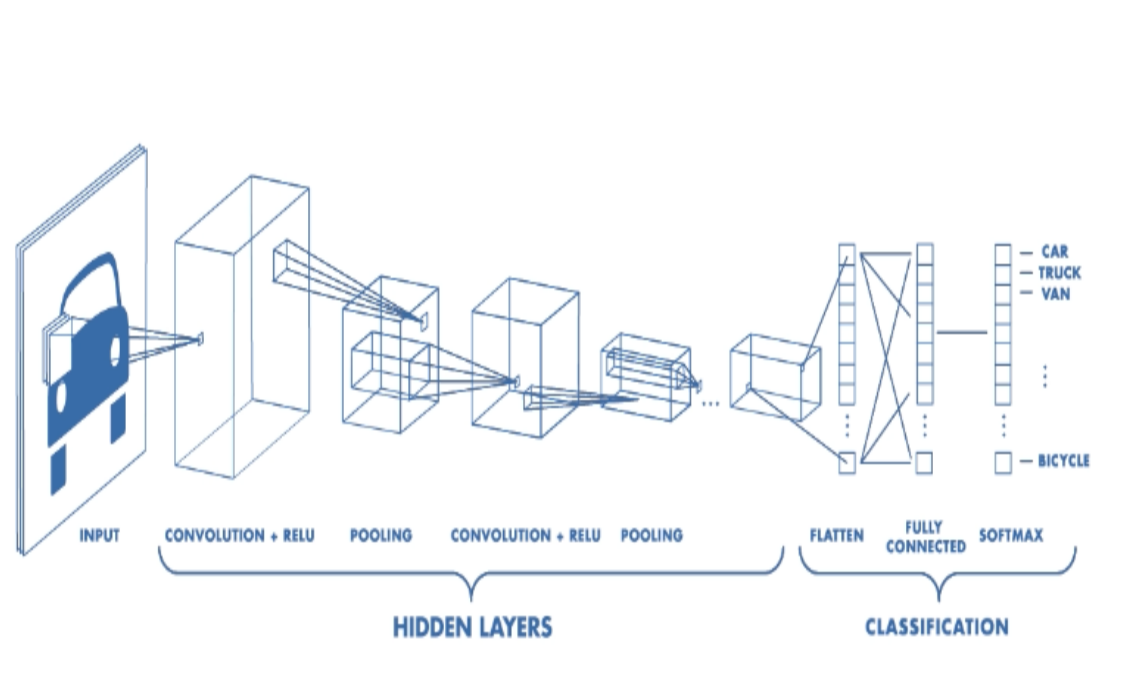


Figure 4: Parts of CNN with example network [4]

- CNN have two components:
  (1)Feature extraction part
  (2)Classification part
- The output of feature extraction part can be used for other classifier.

Classifier: Support Vector Machine (SVM)

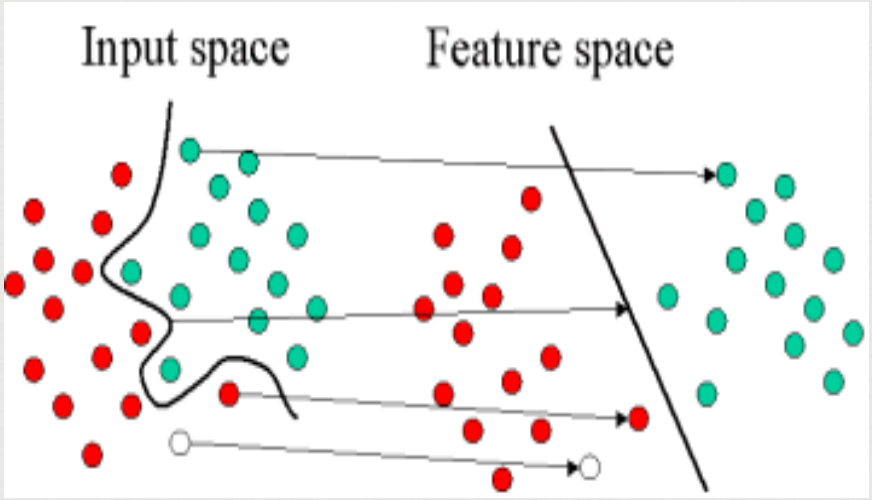


Figure 5: Classification using SVM [5]

- SVM is a supervised classification method for multi-class problem.
- It performs a mapping of data into a higher dimensional feature space and then tries to separate data.

Literature Review

- C. Shan, S. Gong, P. W. McOwan published an comprehensive study on facial expression recognition based on local binary patterns. [3]

- They found about 88.1% accuracy in CK+ dataset and 79.8% accuracy in JAFFE dataset for 7-class facial expression recognition for LBP based SVM (linear).

- They performed experiment on only posed facial expression dataset like CK+, JAFFE and MMI dataset.

Literature Review (Cont.)

- H. Medeiros, V. Pilla Jr, A. Zanellato, C. Bortolini used CNN for feature extraction and SVM for classification in facial expression recognition problem. [6]

- They achieved mean accuracy of 98.52% for 3-class facial expression recognition in CK+ dataset.

- They used pre-trained CNN of Alexnet [7] and only trained the CK+ posed dataset in SVM module.

Literature Review (Cont.)

Problems:

- No cross checking of test accuracy for different dataset.
- Different method used different way for image processing.
- Difference of required time and space for different method was not comparatively discussed.

Work Flow

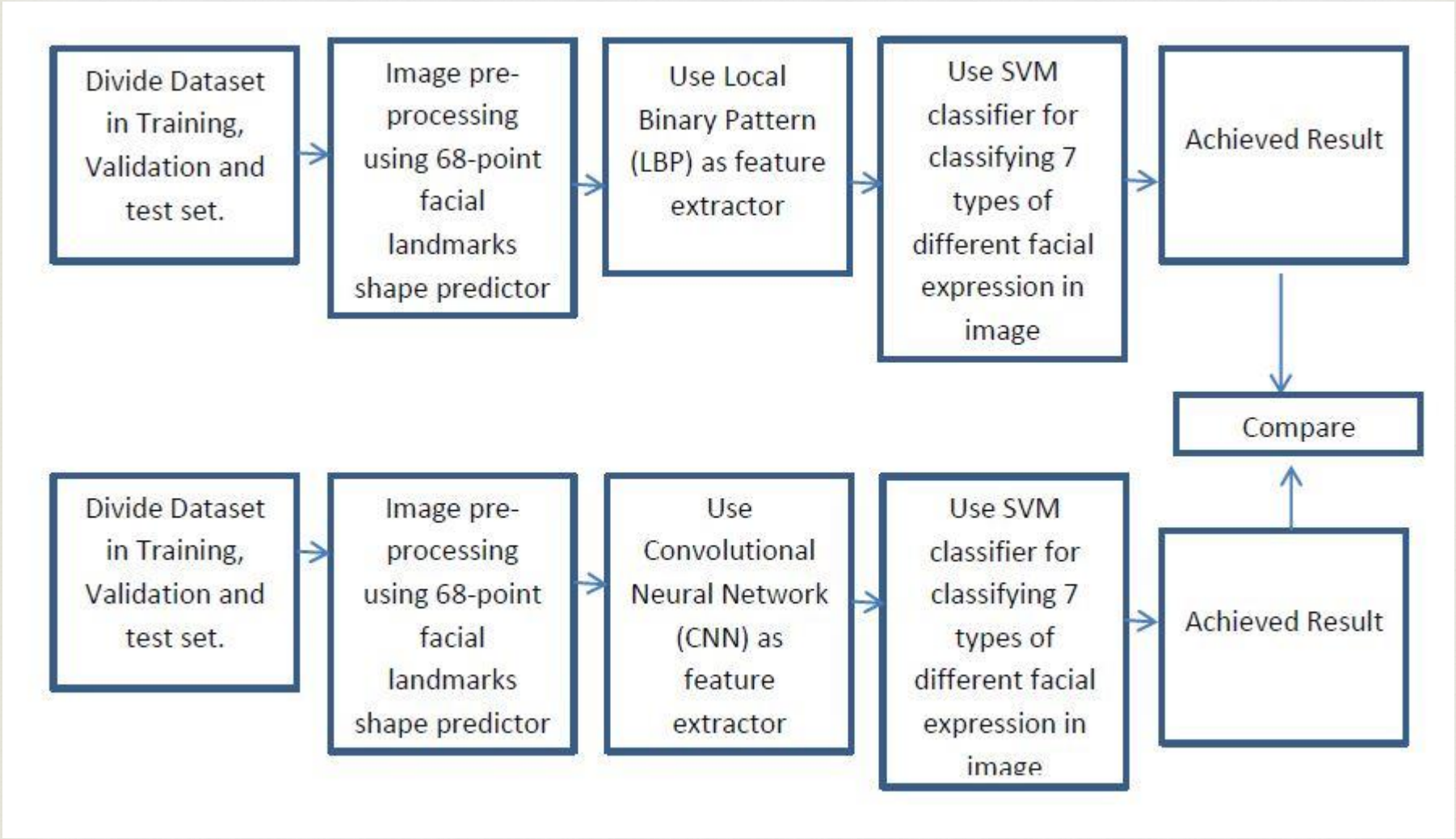


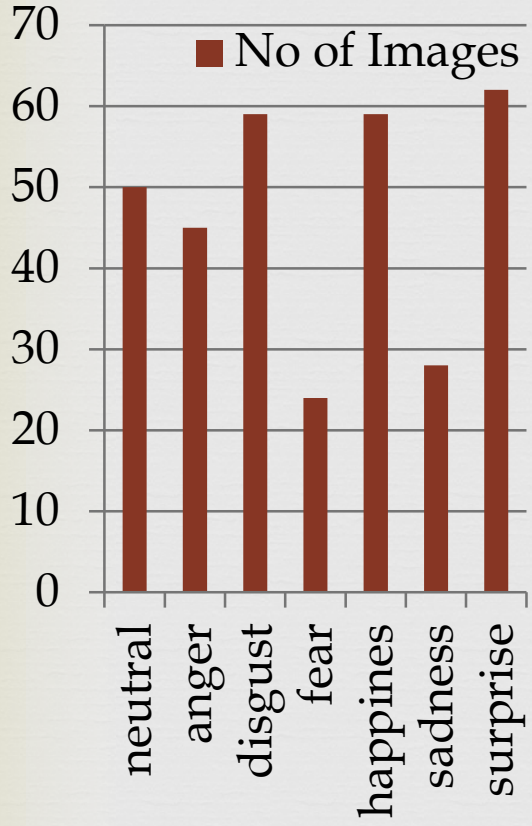
Figure 6: Proposed Work Flow Diagram

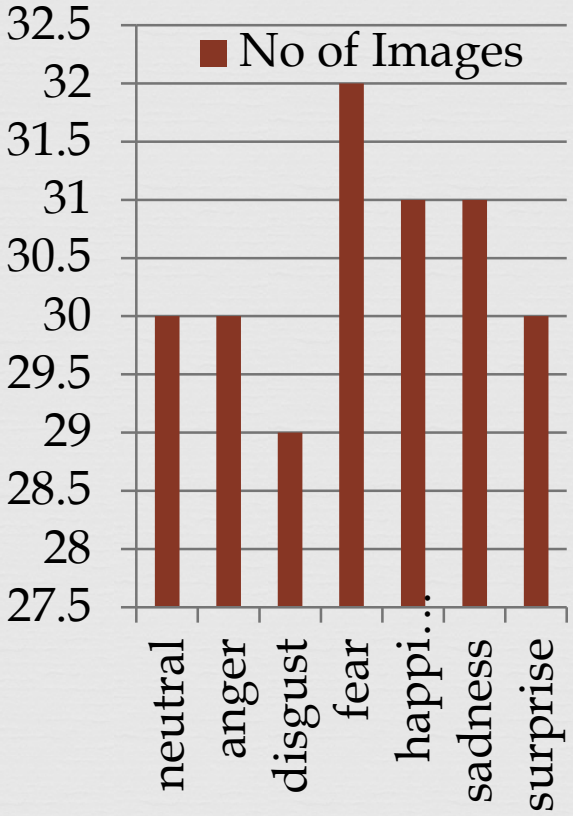
Dataset

Name	Resolution	Type	No of Subjects	No of Training Image	No of Testing Image
CK+	640*490	Posed	123	327	0
JAFFE	256*256	Posed	10	213	0
KDEF	562*762	Posed	70	490	0
FER2013	48*48	Non-Posed	Unknown	28709	3589

Table 1: Dataset Details for 7-class Facial Expression Recognition

Dataset (Cont.)





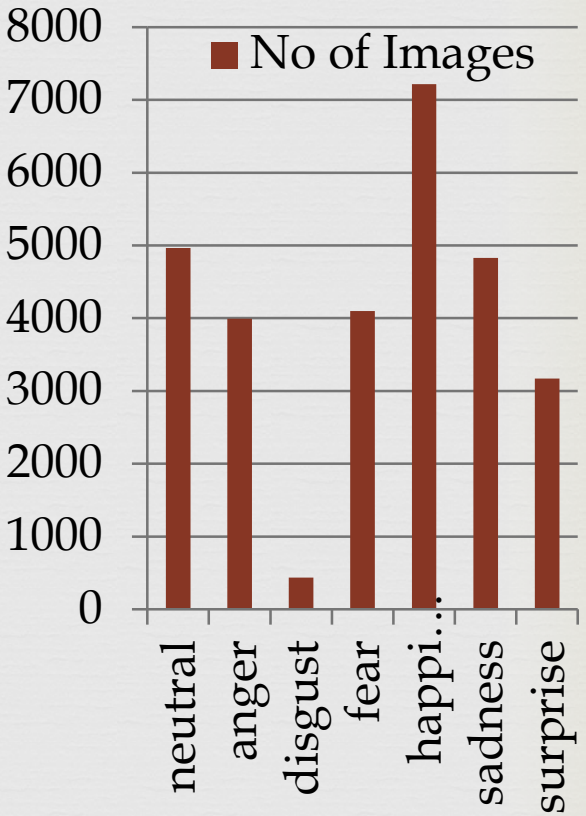


Chart 1: No of Training Images per Expression Label In CK+, JAFFE and FER2013 Dataset Respectively

Dataset (Cont.)

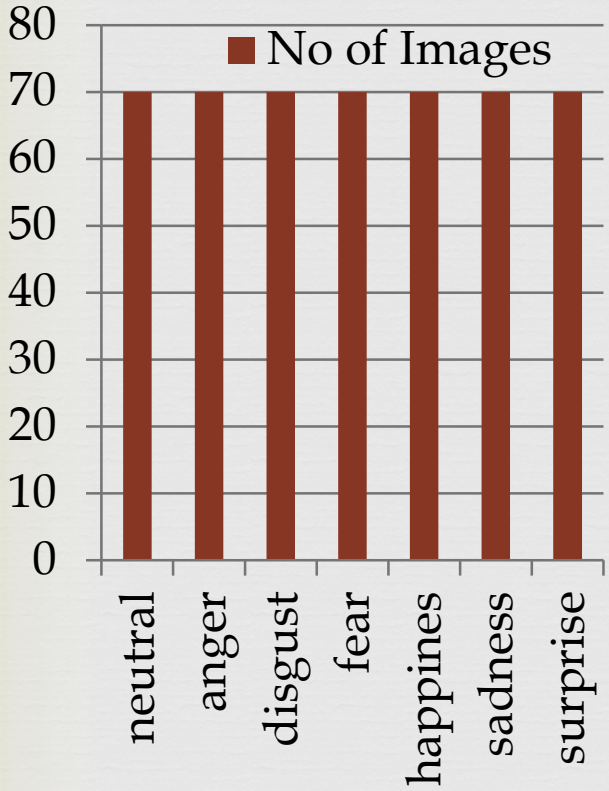


Chart 2: No of Training Images per Expression Label In KDEF Dataset

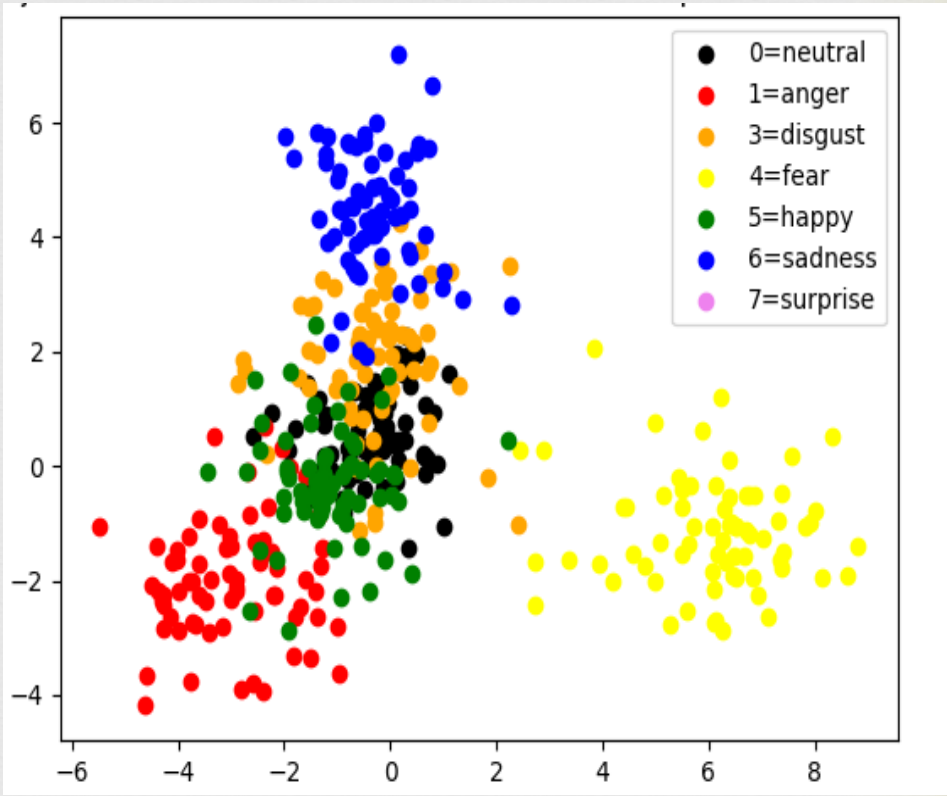


Figure 7: Projected 2D plot of KDEF Dataset

Dataset (Cont.)

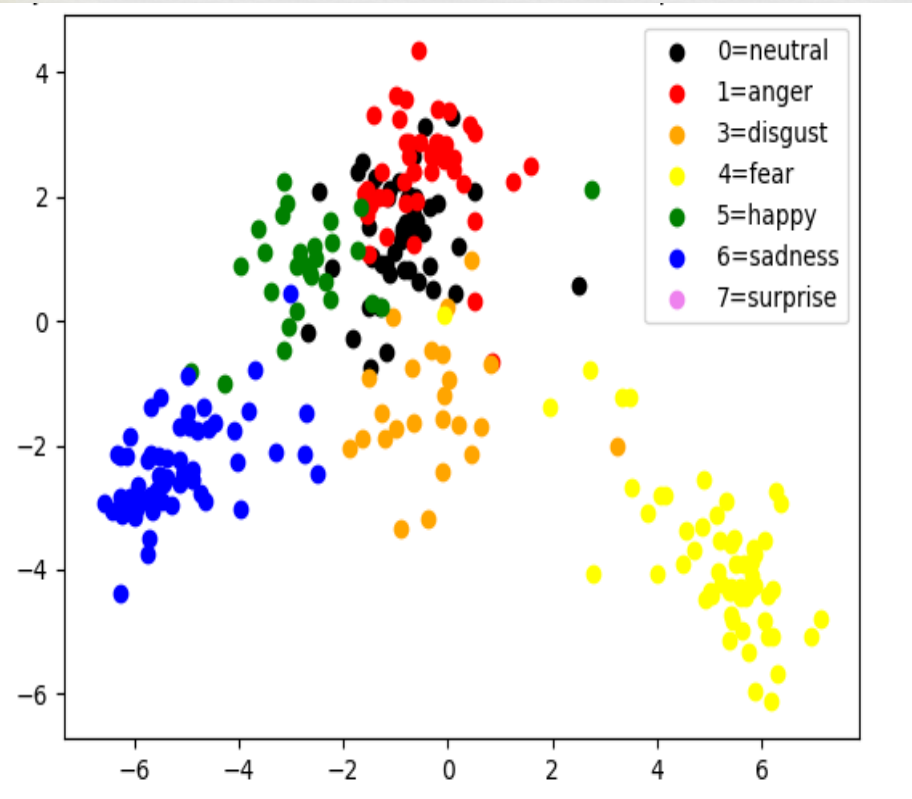


Figure 8: Projected 2D Plot of CK+ Dataset

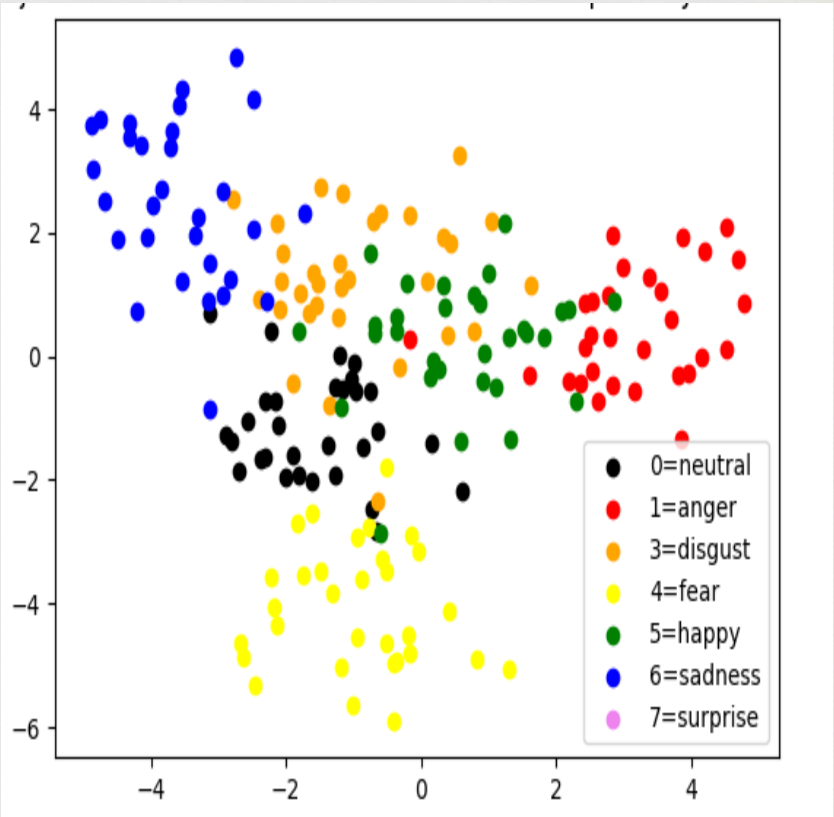
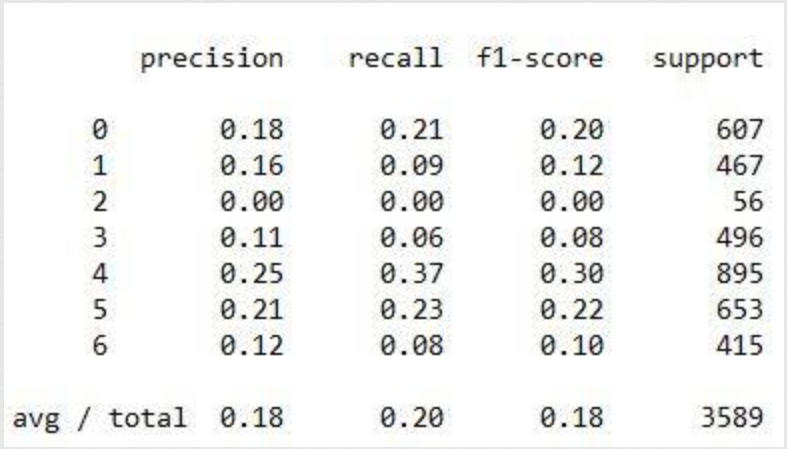


Figure 9: Projected 2D plot of JAFFE Dataset

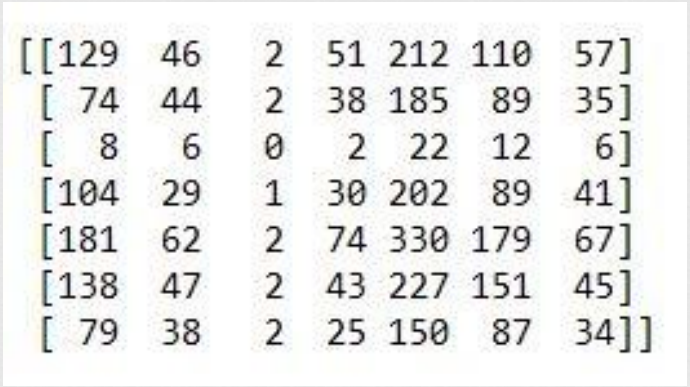
LBP Based SVM for FER2013 Dataset

Kernel	Time	Dataset	Accuracy
linear	148.4 sec	FER2013 PublicTest Dataset	40.3%
		FER2013 PrivateTest Dataset	18.1%
rbf	463.0 sec	FER2013 PublicTest Dataset	47.7%
		FER2013 PrivateTest Dataset	20.0%

Table 2: Testing Accuracy of LBP Based SVM for FER2013 Training Dataset [8]



	precision	recall	f1-score	support
0	0.18	0.21	0.20	607
1	0.16	0.09	0.12	467
2	0.00	0.00	0.00	56
3	0.11	0.06	0.08	496
4	0.25	0.37	0.30	895
5	0.21	0.23	0.22	653
6	0.12	0.08	0.10	415
avg / total	0.18	0.20	0.18	3589



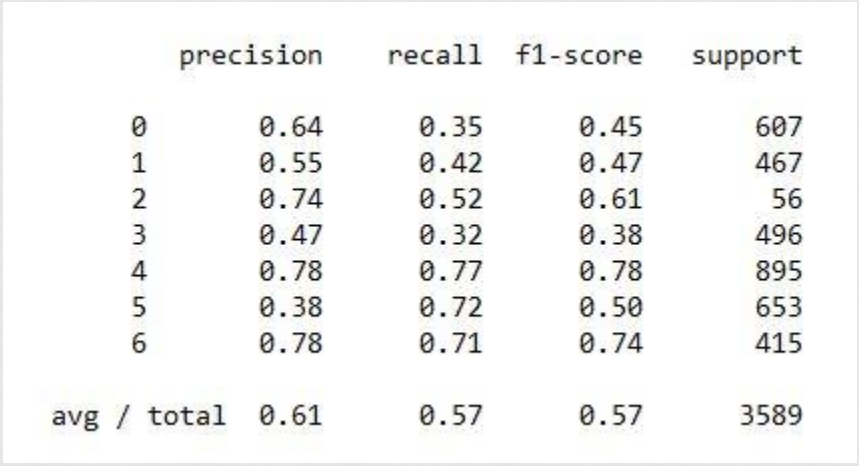
```
[[129  46   2  51 212 110  57]
 [ 74  44   2  38 185  89  35]
 [  8   6   0   2  22  12   6]
 [104  29   1  30 202  89  41]
 [181  62   2  74 330 179  67]
 [138  47   2  43 227 151  45]
 [ 79  38   2  25 150  87  34]]
```

Figure 10: Classification Report & Confusion Matrix for 'rbf' Kernel in FER2013 PrivateTest Dataset

CNN Based SVM for FER2013 Dataset

Kernel	Time	Dataset	Accuracy
linear	47.2 sec	FER2013 PublicTest Dataset	57.8%
		FER2013 PrivateTest Dataset	57.3%
rbf	914.4 sec	FER2013 PublicTest Dataset	31.9%
		FER2013 PrivateTest Dataset	25.0%

Table 3: Testing Accuracy of CNN Based SVM for FER2013 Training Dataset [8]



	precision	recall	f1-score	support
0	0.64	0.35	0.45	607
1	0.55	0.42	0.47	467
2	0.74	0.52	0.61	56
3	0.47	0.32	0.38	496
4	0.78	0.77	0.78	895
5	0.38	0.72	0.50	653
6	0.78	0.71	0.74	415
avg / total	0.61	0.57	0.57	3589

```
[[214  33   1  40  67 240  12]
 [ 21 196   5  47  31 150  17]
 [  0  11  29   1   3  12   0]
 [ 21  36   0 159  24 226  30]
 [ 43  19   2  20 692  99  20]
 [ 29  55   2  44  45 472   6]
 [  8   9   0  24  25  53 296]]
```

Figure 11: Classification Report & Confusion Matrix for 'linear' Kernel in FER2013 PrivateTest Dataset

LBP Based SVM for KDEF Dataset

Kernel	Time	Dataset	Accuracy
linear	1.4 sec	JAFFE Dataset	13.1%
		CK+Dataset	11.0%
rbf	0.2 sec	JAFFE Dataset	15.0%
		CK+Dataset	14.7%

Table 4: Testing Accuracy of LBP Based SVM for KDEF Training Dataset [8]

	precision	recall	f1-score	support
0	0.15	0.14	0.14	50
1	0.17	0.20	0.18	45
2	0.25	0.15	0.19	59
3	0.07	0.12	0.09	24
4	0.26	0.17	0.20	59
5	0.08	0.21	0.12	28
6	0.12	0.06	0.08	62
avg / total	0.17	0.15	0.15	327

```
[[ 7  7  5  8  7 13  3]
 [ 3  9  5  8  4 10  6]
 [12  7  9  8  4 12  7]
 [ 3  2  2  3  2  8  4]
 [ 7 10  6  6 10 11  9]
 [ 4  4  3  6  4  6  1]
 [11 14  6  7  8 12  4]]
```

Figure 12: Classification Report & Confusion Matrix for 'rbf' Kernel in CK+ Dataset

CNN Based SVM for KDEF Dataset

Kernel	Time	Dataset	Accuracy
linear	0.1 sec	JAFFE Dataset	37.1%
		CK+Dataset	64.2%
rbf	0.3 sec	JAFFE Dataset	14.6%
		CK+Dataset	8.6%

Table 5: Testing Accuracy of CNN Based SVM for KDEF Training Dataset [8]

	precision	recall	f1-score	support
0	0.53	0.48	0.51	50
1	0.64	0.47	0.54	45
2	0.90	0.61	0.73	59
3	0.27	0.62	0.38	24
4	0.92	0.93	0.92	59
5	0.42	0.79	0.55	28
6	0.90	0.60	0.72	62
avg / total	0.72	0.64	0.66	327

```
[[24  2  0 12  2  7  3]
 [ 5 21  4  7  1  7  0]
 [10  7 36  3  1  2  0]
 [ 2  1  0 15  1  4  1]
 [ 1  0  0  2 55  1  0]
 [ 1  2  0  3  0 22  0]
 [ 2  0  0 14  0  9 37]]
```

Figure 13: Classification Report & Confusion Matrix for 'linear' Kernel in CK+ Dataset

Result & Comparative Analysis

- It is seen that LBP based SVM works better when 'rbf' kernel is used and CNN based SVM works better when 'linear' kernel is used.
- CNN based SVM gives comparatively better result than LBP based SVM for facial expression recognition.
- Training in SVM 'rbf' kernel takes comparatively longer time than training in SVM 'linear' kernel.
- Performance drop happens when testing is done in different dataset (Ex: KDEF is trained and CK+ and JAFFE is tested) instead of testing in same dataset (Ex: FER2013 Training Dataset is trained and FER2013 PublicTest & PrivateTest Dataset is tested).

Conclusion

- The thesis work highlighted implementing LBP and CNN as feature extractor for SVM classifier of 7-class of Facial Expression Recognition System.

- Applying same experiment procedure for every method and evaluating for different dataset gives us comparative overview about performance of SVM classifier and possibilities of LBP and CNN as feature extractor for facial expression recognition.

Future Scopes

- Recognizing facial expression of non-frontal face.
- Recognize facial expression through video stream in real time.
- Finding more accurate features and classification method for facial expression recognition.
- Evaluating performance of different type dataset for training and testing.

References

- [1] brandspurng.com, "*LIGHTSPEED INTRODUCES EMOTION ANALYTICS*", https://brandspurng.com/lightspeed-introduces-emotion-analytics/
- [2] Taqdir and J. Kaur, "Facial Expression Recognition with PCA And LDA", *IJCSIT*, 2014
- [3] C. Shan, S. Gong and P. W. McOwan, "Facial expression recognition based on local binary patterns: A comprehensive study", *Elsevier*, 2008
- [4] mathworks.com, "*Introduction to Deep Learning*", https://www.mathworks.com/videos/introduction-to-deep-learning-what-are-convolutional-neural-networks--1489512765771.html

References (Cont.)

- [5] statsoft.com, "*Support Vector Machines (SVM) Introductory Overview*", http://www.statsoft.com/textbook/support-vector-machines
- [6] H. Medeiros, V. Pilla Jr, A. Zanellato and C. Bortolini , "Facial expression classification using convolutional neural network and support vector machine", *Semanticscholar*, 2016.
- [7] A. Krizhevsky, I. Sutskever and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", *NIPS*, 2012.
- [8] A. Mursalin, "*Fer-landmarks-cnn-lbp-svm*", https://www.kaggle.com/ankur133047/kernels

Thank You