

MACHINE LEARNING METHODS AND MODELS

FACIAL EMOTION RECOGNITION

Pierre Emmanuel Gerard Ankouane

University of Bradford -23042929

Table of Contents

[INTRODUCTION 2](#_Toc194638048)

[Dataset Selection and Preprocessing 3](#_Toc194638049)

[Machine Learning Methods implementation 4](#_Toc194638050)

[ML System Design and Development 5](#_Toc194638051)

[Result and Discussion 8](#_Toc194638052)

[Conclusion and Future Work 9](#_Toc194638053)

[References 10](#_Toc194638054)

[Appendices 11](#_Toc194638055)

**INTRODUCTION**  
  
This project demonstrates the use of machine learning techniques including both supervised and unsupervised learning methods. The aim is to identify human emotions from facial expression using Convolutional neural networks (CNN) for classification and K means clustering with Principal component analysis (PCA) for emotion grouping. With this advancement of artificial intelligence, automatic recognition has improved multiple sectors such as healthcare, surveillance and education fields building a system capable of analysing comparing, detecting and classifying different Facial emotions Combining deep learning knowledge for accurate recognition and clustering for better structure of data. The core goal of developing a unified application capable of detecting and grouping emotion expression using two such as TensorFlow, keras, open CV, and Sckit- learn to implement and evaluate the models. Supervised learning models on this project has been trained on FER2013 and CK48+ For this project performing a validation accuracy of 56% over 10 epochs. The set of images structured into different emotions label in to train and test directories to allow easier manipulation with open CV an efficiency in training using PIXELS grayscale images 48X48 Classify into happy, sad, angry, discussed, here, surprised to permit real time detection with webcam based on emotion.

# **Dataset Selection and Preprocessing**

In this project, open-source Facial expression data set from Kaggle (FER2013 & CK48+) containing 48x48 pixels images categorized in seven emotions: Angry, discussed, Fear, sad, surprised, and neutral. This image is for the being converted into image folder format Was used to simplify preprocessing. Including main directories: train/ and test/ each containing subfolders for each emotion, to easily support both training CNN model I'm clustering the images. Preprocessing steps included:  
  
Greyscale normalisation changing the picture value scale to [0, 1] range.  
Image resizing to 48x48 for consistency   
Data augmentation in CNN training to increase viability  
Flattening images and PCA for image to be reduced to 2D using PC a unsupervised clustering to ensure a uniform input for both supervised and unsupervised pipelines.

Images where visually inspected for low quality, denoising process using open CV filters (Gaussian blur and median filters) 2 this this smoothness of sharp pixel noise. Pixel intensity values were killed this step was essential to stabilise training and help neural network coverage faster.

**Machine Learning Methods implementation**  
  
 Supervised machine learning implementation  
 This model CNN implementation in tensorflow and keras including convolutional layers, Max pooling dropout for regularisation and a dense layer connected to be trained to classify images into one of this seven emotion categories. As key aspect of this supervise model, input size of 48x48 grayscale images was used with the activation of functions ReLU for hidden layers and softMax for output with loss function categorical cross entropy which led to a promising accuracy in real time detection when integrated with open CV and haar cascade for face detection and trained model.  
  
**Key configurations:**

**Loss function:** Categorical Crossentropy

**Optimizer:** Adam

**Activation:** ReLU (hidden), Softmax (output)

**Output classes:** 7 emotions  
  
 Unsupervised machine learning implementation  
   
To the contrast of the supervised model, this implemented strategy using k-means clustering to flatten and normalise the images using principal component analysis (PCA) to reduce dimensionality to two components of visualisation. Following some key steps such as image loading and flattening from data set, clustering into 7 clusters coma scatter plot visualisation showing group emotion based on pick said intensity patterns. Although unsupervised learning cannot cluster directly, visualisation revealed this grouping that roughly correspond to the different emotions categories.  
  
 All these represented with a great and suitable graphical user interface design and function ability using Tkinter to provide a user friendly application to be able to demonstrate its capability which help launch supervised that's real camera webcam emotion detection also launched on supervised display PCA clustered Facial image  
 comparing models side by side to provide output of comparison from supervised and unsupervised approach.  
  
Key configurations:

**Dimensionality reduction:** PCA (n=2)

**Clustering:** KMeans (k=7)

**Evaluation:** Cluster visualization, manual emotion matching  
  
**Tools and Frameworks**

• TensorFlow & Keras: Used for building and training the CNN model.

• OpenCV: For face detection, image manipulation, and real-time webcam integration.

• Matplotlib/Seaborn: For data visualization (accuracy plots, confusion matrices).

• PyTorch: Optionally explored for experimentation and future scalability.

**ML System Design and Development**  
  
 Machine learning system design and development used for this project that support every time emotion recognition using the camera an unsupervised emotion clustering includes:

**Data Preprocessing and Augmentation**

This is the initial set of the pre-processing of the data set, widely uses it's original CSV format with pixel values stored as a string flowing processes such as parsing CSV into numpy arrays and we shipping the images, normalisation scaling pixels values [0,255] to [ 0, 1], Label encoding to convert into integers labelled, splitting into training (80%) and validation (20%) sets To prevent overfitting and several image augmentation techniques webbing used including horizontal flipping, zooming and shifting To diversify and help the model generalise the real world variability.

**Model Architectural overview**  
  
Input Layer: Webcam (real-time) or dataset images

Preprocessing: Grayscale conversion, resizing, normalization

Convolutional neural network selected with its strong spetial features extraction to help the model compile with loss function, optimiser, metrics on model checkpointing for optimal performance

**Real time deployment**

As final solution using open CV real time interface on webcam frames, expression Marsha to compare from uploaded images with user current facial expression, use for user interface to select modes and test were implemented to detect emotion live, and analyse clustering results through visual plot.

**Supervised Pathway:**

Face detection using Haar Cascades (OpenCV)

Emotion classification via trained CNN model (emotion\_recognition\_model.h5)

Display emotion predictions in real-time via OpenCV window

**Unsupervised Pathway:**

Load dataset from folders

Convert images to 48x48 grayscale and flatten

Use PCA to reduce dimensionality

Apply KMeans clustering to organize images into 7 clusters

Plot the PCA-reduced data points with colours representing different emotions

**Graphical User Interface (GUI):**

A Tkinter-based interface allows the user to:

Launch Supervised Mode (real-time webcam classification)

Launch Unsupervised Mode (clustering visualization)

**System Tools Used:**

TensorFlow/Keras for CNN architecture

Scikit-learn for PCA and KMeans

OpenCV for camera access and face detection

Matplotlib for cluster visualization

Tkinter for GUI controls

This design ensures scalability, testability, and visual interpretability while maintaining user-friendly usability for non-technical users, Launch Comparison View (evaluate results from both pipelines)  
A diagram of a pipeline

AI-generated content may be incorrect.  
**Figure 1:** Machine Learning Pipeline.

**Result and Discussion**

The results from both supervised and unsupervised provided valuable insights for each approach. The CNN model reaches an accuracy of approximately 83% on both FER2013 & CK48+ which demonstrate the model high precision and recall in recognising emotion such as happy, and surprise due to the distinct facial patterns. Emotions like disgust and fear mostly turned out to be misclassified reflecting the inherent similarity between the expressions. Applying the clustering K means + PCA, result into scattered plot displayed visually separable clusters. (k=7) where each cluster was manually associated to an emotion based on the observation of representative image within that cluster. This jewel model setup made the project more complete and visually interpretable providing a much more deeper understanding and more interactive user experience.

**A graph of a graph with blue and orange lines

AI-generated content may be incorrect.  
Figure 2:** CNN Model Accuracy Curve – Training vs Validation

This graph of CNN model accuracy surges a learning relevant feature without overfitting. From the graph above, the epochs shows that the training set is having a large amount of intra class variation making it harder to fit. Relatively small number of an epoch so the model hasn't overfitted yet. More conventional layers could be tested do too and upward consistent trajectory of both curves indicating healthy convergence however to improve the result, hyperparameters tuning could be a solution such as changing learning rate to yield better generalisation, using a large all pre augmented data set could boost the performance especially for classes like disgust. The CNN model was trained over 10 epochs using grayscale official emotion images of size 48X48 from data set to classify Fisher expression into seven emotion that is happy, disgust, angry, fear, sad and neutral, surprise. This performance during the training track by accuracy metrics on both training and validation data set.

F1-score: varied by emotion (best: happy, weakest: disgust)

**Validation included:**

Manual webcam testing

Cross-checking with known emotion samples

Inspecting cluster purity (KMeans)

**A graph with many colored dots

AI-generated content may be incorrect.**

**Figure 3:** Facial Emotion Clustering using PCA and K-Means – Each color represents a different emotion class.

**Conclusion and Future Work**  
  
 Several challenges were encountered during the development of this project such as  
 database format which was in CSV format end required to be converted into image version so that it can easily be manipulated also, model overfitting was faced with regularisation techniques like dropout used to prevent CNN over fitting. Webcam bees realtor recognition required permission for it to be able to execute normally and obey the ethics of AI, image errors which fail to load, dimension issues importation and integration of the data set into the right path document to be able to easily read by the computer, solved by converting images from CSV to JPEG and PNG using AI, moving image folder into the right workspace folder, resizing the images to the right format to ease readability by the system and ensure proper execution.  
  
Improvements and Future work  
  
Important improvement on the long run use to improve current project capacities and extend its capabilities will include factors such as   
  
**Emotion History Tracking :** Which will provide the users with recent emotions over time to observe their emotion trends   
Comparison dashboard : Which will display predictions from both models side-by-side with evaluation metrics at real-time.  
  
**Model tuning :**To enhance the accuracy exploring advanced architectures like ResNet   
  
**Larger dataset :** Providing more diverse datasets to trail the model for a better improved generalization.  
 **Emotional intensity :**Predicting not just the type but the emotion intensity (e.g. Slightly angry vs very angry )

# **References**

1. Mollahosseini, A., Hasani, B. and Mahoor, M.H., 2016. AffectNet: A database for facial expression, valence, and arousal computing in the wild. IEEE Transactions on Affective Computing.
2. Kaggle. FER2013 Dataset. [https://www.kaggle.com/datasets/msambare/fer2013]
3. OpenCV. Haarcascade Documentation. [https://docs.opencv.org/3.4/db/d28/tutorial\_cascade\_classifier.html]
4. Kotsiantis, S. B., Zaharakis, I., & Pintelas, P. (2007).Supervised MachineLearning: A Review of Classification Techniques. Emerging Artificial Intelligence Applications in Computer Engineering, 160, 3–24. Useful to justify the choice of supervised learning models. <https://doi.org/10.3233/978-1-58603-818-3-3>
5. Tan, P. N., Steinbach, M., & Kumar, V. (2018).Introduction to Data Mining (2nd ed.). Pearson.
6. Ekman, P. (1999). Basic Emotions. In T. Dalgleish & M. Power (Eds.), Handbook of Cognition and Emotion (pp. 45–60). Wiley.

# **Appendices**

* Cluster plot visualization screenshots
* Link to GitHub project: [[https://github.com/ByEmG/F.E-version-2-]](https://github.com/ByEmG/F.E-version-2-)
* Link to video demo: [<https://drive.google.com/file/d/1CZ6x6N809iPbfUb_zXE2p-MF4UUj4y0g/view?usp=sharing>]

A screen shot of a computer program

AI-generated content may be incorrect.

This CNN model was trained on 48x48 grayscale images using FER2013 data. The architecture includes two convolutional layers, followed by flattening and dense layers for classification.

A screenshot of a computer screen

AI-generated content may be incorrect.

This code applies Principal Component Analysis (PCA) to reduce image data to two dimensions, allowing visualization of the data structure. KMeans clustering then identifies emotion-based groupings, and each cluster is assigned an emotion label.