DeepFace vs FER: A Comparative Analysis of Emotion Recognition Technologies

AMIR SULEIMENOV and GEORG IBER*, JKU, Austria
GEORG IBER, JKU, Austria

The practical work presented in this paper consists of having various pictures, including pictures of a drone (Parrot Anafi) at the institute, of groups of people. Those images will then be analyzed with the later introduced AI technologies DeepFace and FER. Both of them specialize in emotion recognition, which makes them really good to compare. Nowadays, AI takes a significant part in real-world applications. Our prototype aims to show the differences between two of those technologies, what the current state is, and how they differ in their functionality.

ACM Reference Format:
Amir Suleimenov, Georg Iber, and Georg Iber. 2023. DeepFace vs FER: A Comparative Analysis of Emotion Recognition Technologies. 1, 1 (June 2023), 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Emotions are very complex and really challenging to measure due to their subjective and context-dependent nature. As such, there is growing interest in AI systems capable of recognizing and interpreting human emotions, potentially enhancing human-machine interactions and personalizing services. AI emotion recognition is a subset of affective computing, identifies and classifies human emotions using inputs like facial expressions, voice, or physiological signals. It can be categorized as unimodal(single input) or multimodal(multiple inputs), explicit (direct expression of emotions) or implicit (indirect implication of emotions), and can operate at individual or group levels. The technology relies on machine learning, deep learning, computer vision, natural language processing and signal processing to analyze human emotions. Face recognition, a biometric technology identifying a person based on their facial features, is a crucial method in this domain.

1.1 Brief introduction to emotion recognition

Emotion recognition is important in various sectors

- It can analyze facial expressions to analyse customer reactions to products and to be used in marketing generally
- AI technology is able to assist in teaching, potentially by monitoring emotional states of students and teachers
- AI technology may be more benefitial in the future in the Healthcare system by diagnosing disorders affection facial expression

There are still some challenges with emotion recognition, for example with occlusion or pose variations. Most of the technologies use a three-step process involving face detection, expression detection, and expression classification. One

Authors' addresses: Amir Suleimenov, ; Georg Iber, , JKU, P.O. Box 1212, Linz, , Austria, 4040; Georg Iber, , JKU, Linz, Austria, 4020.

Manuscript submitted to ACM

of the best technologies right now is DeepFace, which is a highly accurate system, identifying human faces in images, employing a nine-layer neural network on them and is trained on millions of user-uploaded images from Facebook.

1.2 Importance of emotion recognition in AI

Emotion Recognition in AI is essential as it enhances understanding of human conditions, needs or reactions. It can be used in improving human-computer action and user experience. Key benefits include:

- AI can adjust its behavior to empathize with humans. Therefore we have better interactions.
- AI can offer personalized services if it understands emotional needs, which enhances the user experience
- AI can monitor emotional responses to situations, providing valuable feedback for more effective learning or gaming
- AI can detect negative emotions and prevent mental distress potentially

1.3 Architecture of the Prototype

Our prototype script uses OpenCV for image preprocessing and DeepFace and FER for emotion recognition. The process involves resizing and grayscaling transformation of some input image for easier face detection via the Haar Cascade classifier. Each detected face is then extracted into a smaller image and the location of it saved in a CSV file. The next step involves initializing another CSV file, that mirrors the original image dimensions and puts the analyzed emotion values from DeepFace and FER into. Therefore each face that got detected by OpenCV undergoes emtion analysis, and in case of our prototype, the happiness and the sad values get stored in a CSV File. As a last step, a small script generates a heatmap out of the CSV file, using the matplot and seaborn library in python.

2 DEEPFACE

DeepFace is a deep learning facial recognition technology created by Facebook. It can recognize human faces in digital images and therefore uses a 9-layer deep neural network. First introduced in 2014, it achieved a big accuracy rate of about 97.35 percent on the labeled faces in the wild (LFW) dataset. In the case of our prototype DeepFace provides us with more functionality than we need, because we are only interested in the emotion recognition part of it.

2.1 Advantages of DeepFace

DeepFace has a very high accuracy. It achieves near-human levels of face recognition and also emotion recognition. It can handle billions of photos, making it suitable for large-scale applications. And last but not least, DeepFace has a great impact on future technologies in this area by demonstrating its effectiveness for real-world applications. So DeepFace in general offers a high performance, robustness in terms of lighting, pose, and occlusion, and it is a state-of-the-art technology.

2.2 Limitations of DeepFace

While not bothering for our prototype, the fact that it needs massive amounts of labeled data to be trained effectively shouldn't be ignored. This might limit the applicability when the data is not readily available. DeepFace is also ver complex and resource-intensive. Training a technology like that requires a lot of computational resources and expertise. In the case of our prototype, using DeepFace driven by python was very accesible and easy to setup.

3 FER TECHNOLOGY

Facial Expression Recognition (FER) is a field of study in computer science and machine learning. It focuses on detection human emotions from facial expressions in photos or videos. Not like DeepFace, it's focus is to classify facial expressions. And like DeepFace does as well, FER categorizes the found emotions(happy, sad, angry, surprise, disgust, fear and neutral). To summarize, DeepFace as a technology is very complex and needs large amounts of labeled training data to achieve its high performance.

3.1 Advantages of FER

FER specializes in understanding non-verbal cues, which can be a valuable tool for enhancing human-computer interactions, while DeepFace for example is mostly used internally by FaceBook. FER also works more effectively on smaller datasets, since it's using a wider range of techniques.

3.2 Limitations of FER

The first limitation would be the performance. Depending on which technique you use, e.g deep learning-based or traditional machine learning-based, the performance may vary. Also FER struggles with variations in lighting, pose or facial occlusions.

4 METHODOLOGY

This study compares the DeepFace and FER emotion recognition through a structured process. I tests both technologies on images categorized as simple, intermediate and complex, based on the subjective difficulty of emotion recognition. After applying both AI technologies, the results are visually represented with heatmaps and Tables.

- The 'simple' category is an image, where the faces are easily discernable. Every subject is oriented towards the camera and lighting is bright with minimal shadows. This represents an ideal, straightforward environment for FER and DeepFace.
- The 'intermediate' category feature subjects with varying head angles, some not directly facing the camera. This adds an additional layer of complexity to the task of emotion detection.
- Lastly, the 'complex' category encompasses images with more challengine lighting conditions, multiple subjects and a variation in head orientations. It tests the robustness of both DeepFace and FER in less than optimal conditions. By this methodology with such categories, we aim to make sure that comparison of the two technologies DeepFace and FER is done properly, thereby highlighting their limitations and capabilities.

5 RESULTS

For evaluating our prototype more, multiple photos of people have been used to improve the accuracy and also the source code itself. For the purpose of this paper we chose to present the images, that had the biggest differences in terms of their analysis. We will present from 'simple' to 'complex'. Lets take a look at the first image.

5.1 Simple Category

Our first image contains a group of four people. Next to the group photos are all the cut out images that OpenCV was able to detect. As described in the architecture section, OpenCV was used to detect the faces inside the image, altough DeepFace would be able to do it as well. But to make it a fair comparison, we had both DeepFace and FER analyze on

Fig. 1. Simple category group image

the same input from OpenCV. The first thing that you can see, is that OpenCV only detects 3 of the 4 faces in the image. As for the emotion analysis of DeepFace and FER, it doesn't really matter to us, since OpenCV is not a topic of the study. We still would subjectively categorize this image as 'simple', since the lighting conditions are good, and everyone is looking rather straight into the camera.

As for the results we take a look at the table that shows the different emotion values of each face, both DeepFace and FER.

Face ID	Emotion	DeepFace	FER
Face1	Angry	0.00046	0.0
	Disgust	0.0	0.0
	Fear	0.00006	0.0
	Happy	99.42	0.99
	Sad	0.01	0.0
	Surprise	0.00054	0.0
	Neutral	0.57	0.0
Face2	Angry	0.01	0.5
	Disgust	0.00083	0.0
	Fear	1.46	0.09
	Happy	98.31	0.15
	Sad	0.16	0.25
	Surprise	0.02	0.01
	Neutral	0.03	0.0
Face3	Angry	0.00004	0.01
	Disgust	0.0	0.0
	Fear	0.008	0.0
	Happy	99.49	0.95
	Sad	0.01	0.03
	Surprise	0.00073	0.0
	Neutral	0.49	0.0

Fig. 2. Table containing the emotion detection values

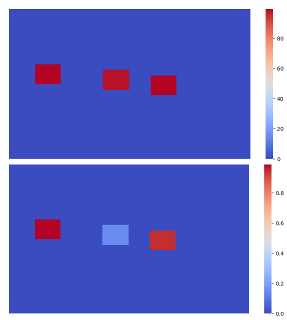


Fig. 3. Heatmap containing the happy values of DeepFace and FER

- It appears that DeepFace is more sensitive than FER in detecting subtle emotions. DeepFace gives a value to every emotion, no matter how small, while FER normalizes it to 0, if the value is too small.
- Both models agree on the primary emotion for all faces, which is 'Happy'. The percentage assigned by DeepFace is very high compared to FER
- For the secondary emotions both Technologies identify emotions like 'Sad', 'Fear', 'Neutral', 'Angry' and 'Surprise' as well. But FER tends to give less importance to secondary emotions, since a lot of them are assigned to zero.
- DeepFace appears to detect the 'Neutral' emotion much better than FER.
- The biggest Discrepancy appears for face 2. Notably, FER detected a significant amount of anger (50 percent) while DeepFace detected only a tiny amount of anger (0.01 percent). This could suggest differences in how the models perceive or interpret certain facial features.

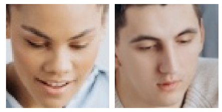


Fig. 4. Intermediate category group image

5.2 Intermediate Category

Unfortunately OpenCV is not really reliable in this case, so it could only cut out 2 of the 4 faces shown in the picture, because of the head orientation. But we did cut the other 2 faces manually to check their emotion values, since it's not about OpenCV functionality, but about DeepFace and FER. But lets stay with these images for now. The respective emotion value table looks like this.

Face ID	Emotion	DeepFace	FER
Face1	Angry	2.26	0.72
	Disgust	0.00	0.01
	Fear	32.88	0.01
	Happy	60.55	0.18
	Sad	0.36	0.02
	Surprise	3.13	0.03
	Neutral	0.45	0.02
Face2	Angry	39.84	0.02
	Disgust	0.03	0.00
	Fear	3.06	0.01
	Happy	0.01	0.05
	Sad	8.29	0.03
	Surprise	0.00	0.00
	Neutral	48.77	0.9

Fig. 5. Table containing the emotion detection values

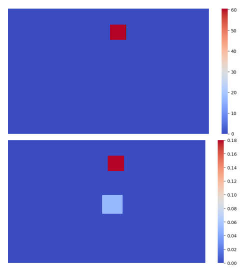


Fig. 6. Heatmap containing the happy values of DeepFace and FER

- DeepFace and FER disagree on the primary emotion for both faces. For Face1, DeepFace indicates 'Happy' while FER suggests 'Angry'. For Face2, DeepFace suggests 'Angry' as the primary emotion, while FER suggests 'Neutral'
- Similar to the previous comparison, DeepFace is more sensitive with subtle emotions
- There are significant differences between the two models in several categories. For instance, in Face1, DeepFace shows a high percentage of 'Fear' while FER indicates a very low percentage. The same for Face2, where DeepFace indicates a high 'Angry' percentage compared to FER

These differences show the variability that can exist between different emotion recognition models, which may be due to different training datasets, algorithms, or the ways in which the facial expressions are interpreted.

5.3 Complex Category

In the last picture, OpenCV did a very good job, detecting 6 of the 7 faces in the picture. Most probably because of the higher resolution compared to the other pictures.

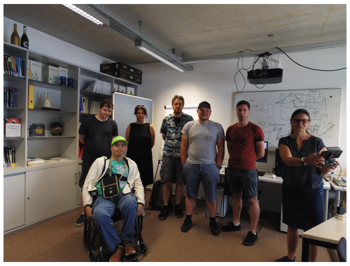
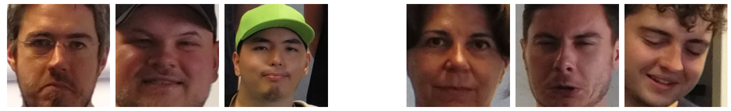


Fig. 7. Intermediate category group image



Face ID	Emotion	DeepFace	FER	Face ID	Emotion	DeepFace	FER
Face1	Angry	3.11	0.07	Face4	Angry	0.19	0.03
	Disgust	0.00	0.0		Disgust	0.00	0.0
	Fear	1.10	0.22		Fear	22.28	0.08
	Happy	0.05	0.0		Happy	67.66	0.22
	Sad	18.52	0.31		Sad	9.33	0.14
	Surprise	0.00	0.01		Surprise	0.54	0.02
	Neutral	77.22	0.38		Neutral	0.00	0.52
Face2	Angry	0.17	0.02	Face5	Angry	0.31	0.01
	Disgust	0.01	0.0		Disgust	0.00	0.0
	Fear	0.25	0.0		Fear	7.46	0.02
	Happy	89.75	0.16		Happy	0.00	0.0
	Sad	8.04	0.02		Sad	92.19	0.77
	Surprise	0.05	0.0		Surprise	0.00	0.0
	Neutral	1.73	0.8		Neutral	0.05	0.19
Face3	Angry	1.03	0.08	Face6	Angry	0.00	0.01
	Disgust	0.05	0.0		Disgust	0.00	0.0
	Fear	2.26	0.0		Fear	0.00	0.0
	Happy	68.13	0.61		Happy	98.88	0.95
	Sad	5.86	0.07		Sad	0.08	0.01
	Surprise	0.38	0.02		Surprise	0.00	0.0
	Neutral	22.28	0.22		Neutral	1.03	0.02

Fig. 8. Table containing the emotion detection values

- DeepFace and FER agree on the dominant emotion again. For the top left picture (Face1), DeepFace gives a sadness score of 18 percent while FER gives it 31 percent
- For Face2 (middle left) both agree on the dominant emotion happiness with a high percentage
- In general, both models broadly agree on the dominant emotion of each face but the results can be different in the secondary emotion. As well as in the other difficulty categories, the results are pretty similar. FER just simplifies a few values by normalizing it to zero.

6 CONCLUSION

6.1 Scale of Scoring

DeepFace operates on a slightly different scale than FER. While FER goes from 0 to 1, DeepFace gives percentage values from 0 to 100.

6.2 Consistency in Dominant Emotions

Both technologies agree on the dominant emotion for most of the given faces event if there are some exceptions.

6.3 Variance in Results

In some cases, the difference is significant. These differences highlight the subjectivity in emotion recognition tasks, and also that different training data can lead to different results.

6.4 Summarization

Both DeepFace and FER have their strengths and can demonstrate how powerful they are. Depending on what needs to be done in real-world applications, DeepFace has a more detailed and nuanced result, while FER provides a more straightforward and binary interpretation.

7 RESOURCES

Rawal, N., Stock-Homburg, R.M. Facial Emotion Expressions in Human–Robot Interaction: A Survey. Int J of Soc Robotics 14, 1583–1604 (2022). https://doi.org/10.1007/s12369-022-00867-0 Facial Emotion Expressions in Human–Robot Interaction: A Survey Niyati Rawal1 · Ruth Maria Stock-Homburg Evaluating Distances in Tactile Human-Drone Interaction Marc Lieser Ulrich Schwanecke Jorg Berdux M. Lieser, U. Schwanecke, and J. Berdux, "Evaluating Distances in Tactile Human-Drone Interaction," in 2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR), 2016, pp. 92-97, doi: 10.1109/SSRR.2016.7784268. A. Rosebrock, "OpenCV Haar Cascades," PyImageSearch, April 12, 2021. [Online]. Available: https://pyimagesearch.com/2021/04/12/opencv-haar-cascades/. [Accessed: April 19, 2023]. J. Saragih, S. Lucey, and J. Cohn, "Deformable model fitting by regularized landmark mean-shift," in 2011 International Conference on Computer Vision, 2011, pp. 593-600, doi: 10.1109/ICCV.2011.6126271. "DeepFace: Closing the Gap to Human-Level Performance in Face Verification" [Online]. Available: https://research.facebook.com/publications/deepface-closing-the-gap-to-human-level-performance-in-face-verification/. [Accessed: June 8, 2023]. "Menpo: A Comprehensive Suite for Interactively Developing, Evaluating, and Deploying Deformable Models," Menpo.org. [Online]. Available: https://www.menpo.org/. [Accessed: April 19, 2023].

"Cascade Classifier"[Online] Available: https://docs.opencv.org/3.4/db/d28/tutorialcascadeclassifier.html. [Accessed: May 12, 2023].