

EMOTION RECOGNITION AND DEPRESSION DETECTION USING DEEP LEARNING

Amey Chede, Dyaneshwar Jagadale , Saurabh Dayma , Dr.S.R.Ganorkar

1 Student, Electronics & Telecommunication, Sinhgad College Of Engineering Pune ,Maharashtra, India 2 Student, Electronics & Telecommunication, Sinhgad College Of Engineering Pune ,Maharashtra, India 3 Student, Electronics & Telecommunication, Sinhgad College Of Engineering Pune ,Maharashtra, India 4 HOD, Information Technology, Sinhgad College Of Engineering Pune, Maharashtra , India

ABSTRACT

Real-time analysis, labelling, and inference of cognitive affective states from a video capture of the face are made possible by emotion detection systems based on facial gesture. Since it is considered that when an emotion is experienced, facial expressions are temporarily activated, emotion detection can be accomplished by identifying the associated face expression. Depression is one of the six main emotions that are present and is quite important. The term "depression" refers to a mood condition. It might be characterized as despair, rage, or a sense of loss that interferes with daily tasks. Depression manifests itself differently for each person. Depression may sometimes result in deadly situations. To prevent any of them, depression must be identified as soon as possible and the victim must receive the proper care. The project's goal is to use real-time video to analyze a user's emotion. Convolutional neural networks [CNN] are used for this. If the feeling is determined to be depression, it has to be addressed as soon as possible. As the symptoms increase, a person's mental capacity deviates from normal, which results in a disorder. If depression is determined to be the emotion, a Chatbot pop-up that was created using the Tkinter library displays on the user's screen asking them to communicate their sentiments. This improves the user's mood while also assessing their level of depression and assisting them in overcoming it. A continuous evaluation is conducted to distinguish between sadness and depression if the user's feeling is determined to be sad.

Keyword : - Machine Learning, CNN, Deep Learning, Neural Network, ANN

1. INTRODUCTION

Emotions tend to rule our daily lives. An emotion is a complex psychological state that has three unique components, according to Don Hockenbury and Sandra E. Hockenbury's book "Discovering Psychology:" a subjective experience, a physiological reaction, and a behavioral or expressive response. Simply said, emotions encompass more than just feelings. It is the manner in which a person feels. It involves physical responses, such as the heart pounding with enthusiasm. Additionally, it contains vocalizations and facial gestures that convey emotion, such as when you utter "woah" in awe of something. Additional aspects of emotions include behaviors, such as shouting at someone when you're upset. A human can feel the following six fundamental emotions: happiness, sadness, fear, disgust, anger, and surprise.

Impaired emotion control is a symptom of depression. It is an aberrant emotional state that has a deep and long-lasting impact on our thoughts, perceptions, and behavior. A tough or stressful scenario, a loss, or a change in circumstances are not always the root causes of depression. In actuality, it frequently happens without any such stimuli. With 264 million sufferers globally, depression is a widespread disorder. Depression can cause major health conditions if it lasts for a long time and has a moderate to severe intensity. People frequently experience

suffering, which has an effect on their life. They do poorly in home, job, and school. Suicide can result from depression at its worst. A depressive episode can be classified as mild, moderate, or severe depending on how severe the symptoms are and how many there are.

To assist identify whether a person is sad or depressed, mental health practitioners utilize the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) criteria from the American Psychiatric Association. At least five of the following symptoms must persist. At least two weeks, and at least one of the symptoms must be decreased interest or pleasure or depression. The DSM-5 criteria include nine potential symptoms of depression:

1. trouble sleeping, or sleeping too much
2. trouble eating, or eating too much, coupled with weight gain or weight loss
3. irritability, restlessness, or agitation
4. extreme fatigue
5. unwarranted or exaggerated feelings of guilt or worthlessness
6. inability to concentrate or make decisions
7. suicidal thoughts or actions, or thinking a lot about death and dying

1.1 PROBLEM STATEMENT

Since it is considered that when an emotion is experienced, facial expressions are temporarily activated, emotion detection can be accomplished by identifying the associated face expression. Depression is one of the six main emotions that are present and is quite important. The term "depression" refers to a mood condition. It might be characterized as despair, rage, or a sense of loss that interferes with daily tasks. Depression manifests itself differently for each person. Depression may sometimes result in deadly situations.

2. LITERATURE SURVEY

Aliaa A computer vision system for automatic face emotion identification is presented by A. A. Youssif and Wesam A. A. Asker in their paper (AFER). AFER consists of three main phases, the first of which is the identification of faces in scenes. The facial features that are displaying a facial expression are extracted in the second stage, and the facial displays that are visible on the face are classified in the third. The open source code library (OpenCV) is used for face detection, and it uses a face detection method based on Viola & Jones characteristics. The face image is then divided into three areas, including the mouth, nose, two eyes, and two brows, before the Facial Features Extraction process is carried out. Second, facial characteristic points (FCPs) are extracted from each component of the face, including the lips, nose, eyes, and brows. A feature vector made up of geometric features and appearance characteristics, which display a pattern for different facial expression classes, is created by applying the feature extraction procedure to a face picture. In order to recognise face expressions, the feature vector is then sent as an input to the radial basis function artificial neural network. With identification rates between 90% and 99% in a person-dependent dataset and between 83% and 100% in a person-independent dataset, the results demonstrate that the AFER system properly categorizes facial expressions. [1]

The paper on emotion recognition using convolutional neural networks was suggested by the authors Enrique Correa, Arnoud Jonker, Michael Ozo, and Rob Stolck. A few hundred high-resolution photographs to tens of thousands of lesser images are used in this technique. The size of the training dataset from FEREC needs to be raised from 9000 photos to 20000 images in order to improve the accuracy of emotions recognised with those of other techniques, including SVM and LVQ. 90% of users are satisfied, 80% are neutral, and 77% are startled. [2]

3. METHODOLOGY

Proposed block diagram of Emotion Detection and Depression Detection is shown in Fig 3.1

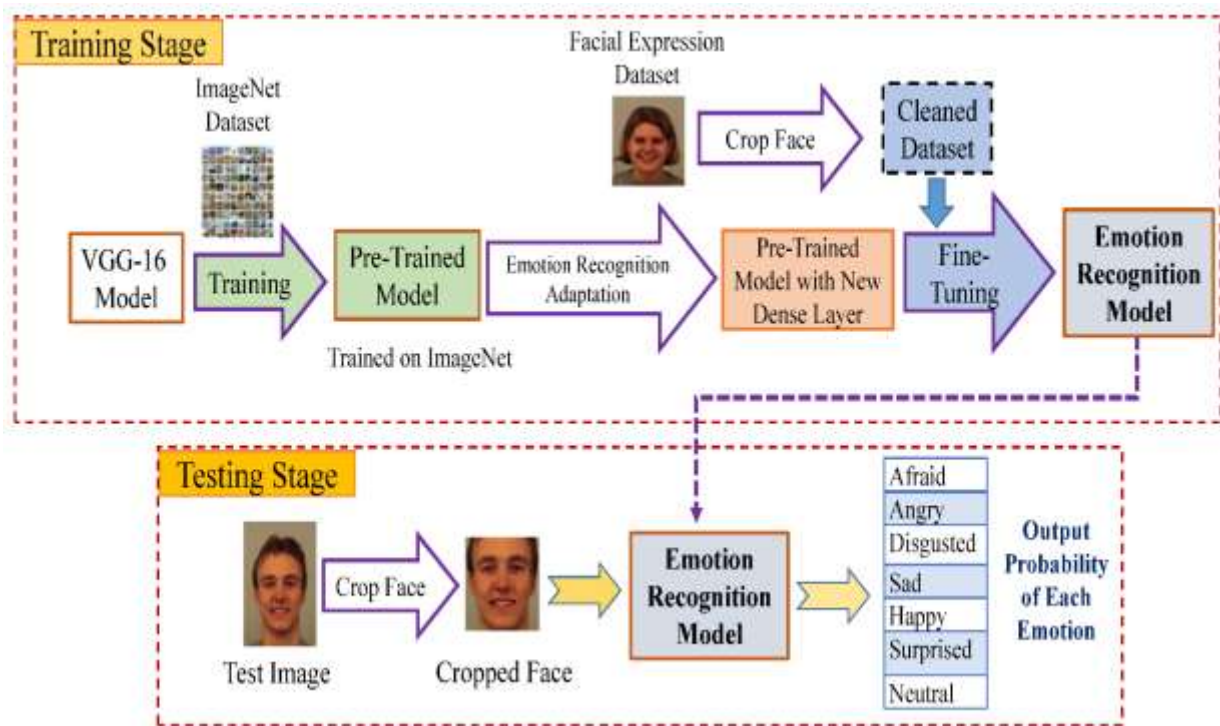


Fig 3.1 Block Diagram of Emotion Recognition and Depression Detection using Deep Learning

3.1 Software Requirements

3.1.1 Arduino IDE

The open-source software known as the Arduino IDE is used to create and upload code to Arduino boards. For different operating systems, including Windows, Mac OS X, and Linux, the IDE programme is appropriate. The programming languages C and C++ are supported. Integrated Development Environment is referred to in this sentence.

Sketching is a common term for writing a programme or piece of code in the Arduino IDE. To upload the sketch created in the Arduino IDE software, we must link the Genuino and Arduino board with the IDE. The drawing has the ".ino" file extension.

The Arduino IDE will appear as:



Fig. 3.2 Arduino IDE

3.1.2 Python

The Raspberry Pi uses Python, an interpreted high level programming language, to manage the status and voltage level of the pins as well as to run machine learning (ML)-based algorithms on input photos to find and classify animals. Python is a strong language since it has many different libraries and frameworks that enable speedier execution of quotations with less lines of code. Python is useful in this project since it serves as an embedded language for GPIO setup, which controls the voltage levels and raspberry pi pin statuses. The backend framework is also in charge of teaching this CV to recognize the animals in the image. Weights are applied to train the cv2 so that it will correctly encircle the animal in the output image and display a rectangle as the boundary.

3.1.3 OpenCV

A common infrastructure of computer vision applications was created with the help of OpenCV, an open-source computer vision machine learning software library, to hasten the adoption of perception in commercial goods. It features 2500 optimized algorithms, including a vast array of traditional and cutting-edge CVs. The majority of these are used to recognize objects, categories human behaviors, detect vehicles, extract 3D models, extract stereo camera images, and many other things. It is also utilized in an easy-to-use and efficient way to improve up scaling, downscaling interpolation, and other image processing applications. It allows multi-paradigm programming for higher level project synthesis and offers interfaces for C++, Java, and MATLAB. To detect the region of interest in a picture and extract data if an animal is discovered, cv2 is trained using weights in this research. Animal presence may be recognized and identified in many instances because to the robust support libraries of CV.

3.1.4 CNN

A deep learning neural network designed for analyzing organized arrays of data, such as images, is known as a convolutional neural network, or CNN. In the input image, design elements like lines, gradients, circles, or even eyes and faces are extremely successfully picked up by CNN. Convolutional neural networks are extremely effective for computer vision because of this feature. CNN does not require any preprocessing and may be used

immediately to an underdone image. A feed-forward neural network with up to 20 layers is a convolutional neural network. The convolutional layer, a specific sort of layer, is what gives convolutional neural networks their power.

3.2 Flow Chart

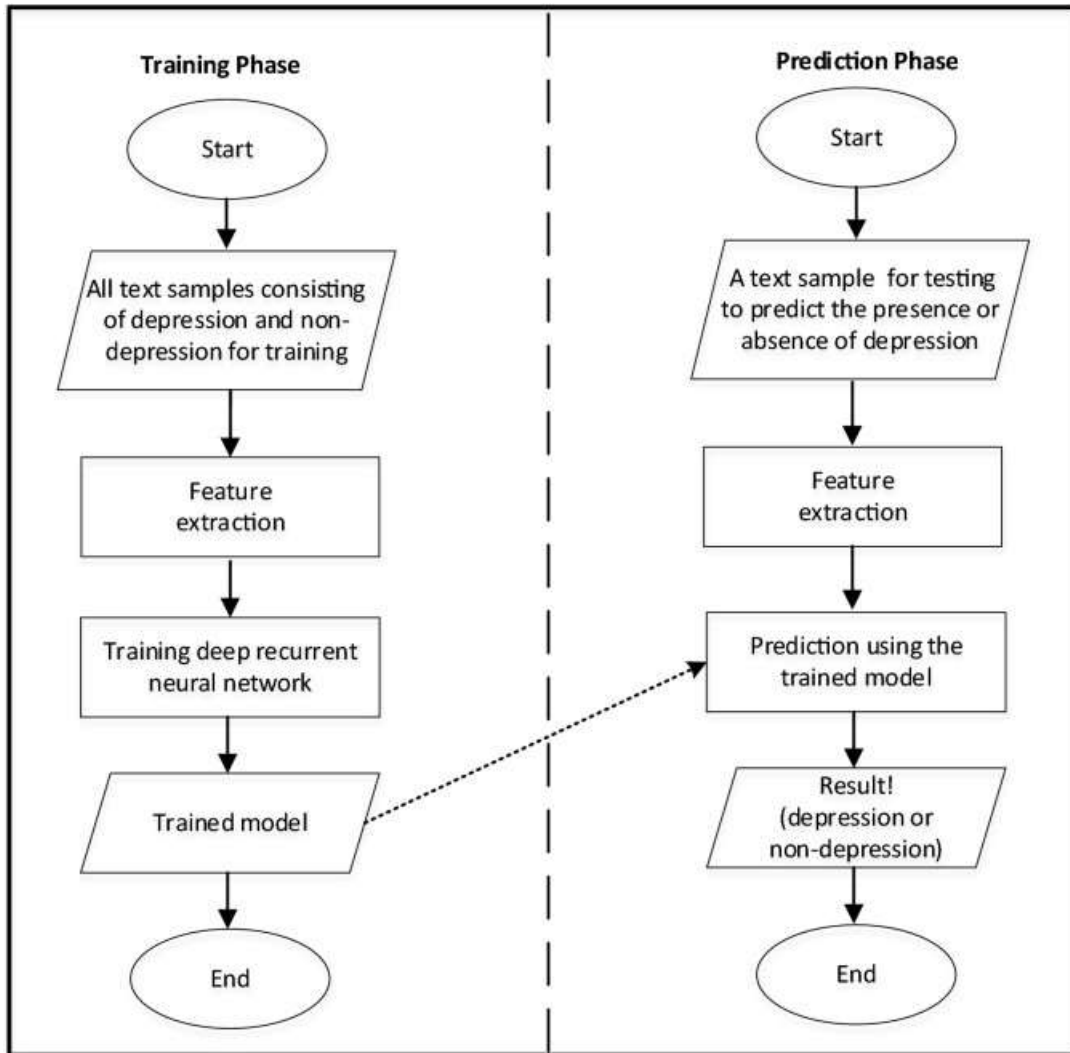


Fig. 3.3 shows flowchart of Emotion Recognition and Depression Detection using Deep Learning

4. RESULT AND DISCUSSION

4.1 Convolutional Neural Network

A ConvNet's architecture was influenced by how the Visual Cortex is organised and is similar to the connection network of neurons in the human brain. The i -th layer of a convolutional neural network is made up of the convolutional layer and the pooling layer together. The number of these layers may be raised to capture even more low-level information, but doing so will need more computer power depending on how complex the pictures are.

4.1.1 Facial Expression dataset

In the literature, there are several open-access facial expression datasets. We utilised the Kaggle dataset for facial expression, which contains 48x48 pixel grayscale portraits of people. There are 28,709 samples with seven different emotions in the training set (happy, sad, surprised, fearful, angry, disgusted, and neutral).

4.1.2 Image Preprocessing

The Haar Cascade library was used to extract the facial circumference from the images. These discovered rectangular face expressions were then captured and cut. The photographs were also made into grayscale versions and fed into neural networks. This procedure was carried out to prevent the neural networks from having extra density

1.DataSets

Datasets in image processing are collections of images with corresponding labels or annotations that are used for training and testing deep learning models. These datasets are crucial for developing and evaluating machine learning algorithms for computer vision tasks, such as image classification, object detection, semantic segmentation, and image generation.

A good dataset should have a sufficient number of images with diverse visual features and variations to ensure that the trained models can generalize well to unseen data. The images should be accurately labeled or annotated to provide ground truth information for the learning algorithms. In addition, the dataset should have a good balance of classes or categories to avoid bias towards specific types of objects or scenes.

There are various publicly available datasets for image processing in deep learning, ranging from small-scale datasets like MNIST and CIFAR-10 to large-scale datasets like ImageNet and COCO. These datasets have different characteristics in terms of size, complexity, and task-specific annotations, and they are often used as benchmarks for evaluating the performance of deep learning models.

In summary, datasets play a crucial role in deep learning for image processing, and choosing the right dataset is important for developing and evaluating effective machine learning algorithms for computer vision tasks.

2.Model Training

Training consists in learning a relation between data and attributes from a fraction of the training dataset, and testing consists in testing predictions of this relation on another part of the dataset (since you know the prediction, you can compare the output of the relation and the real attributes).

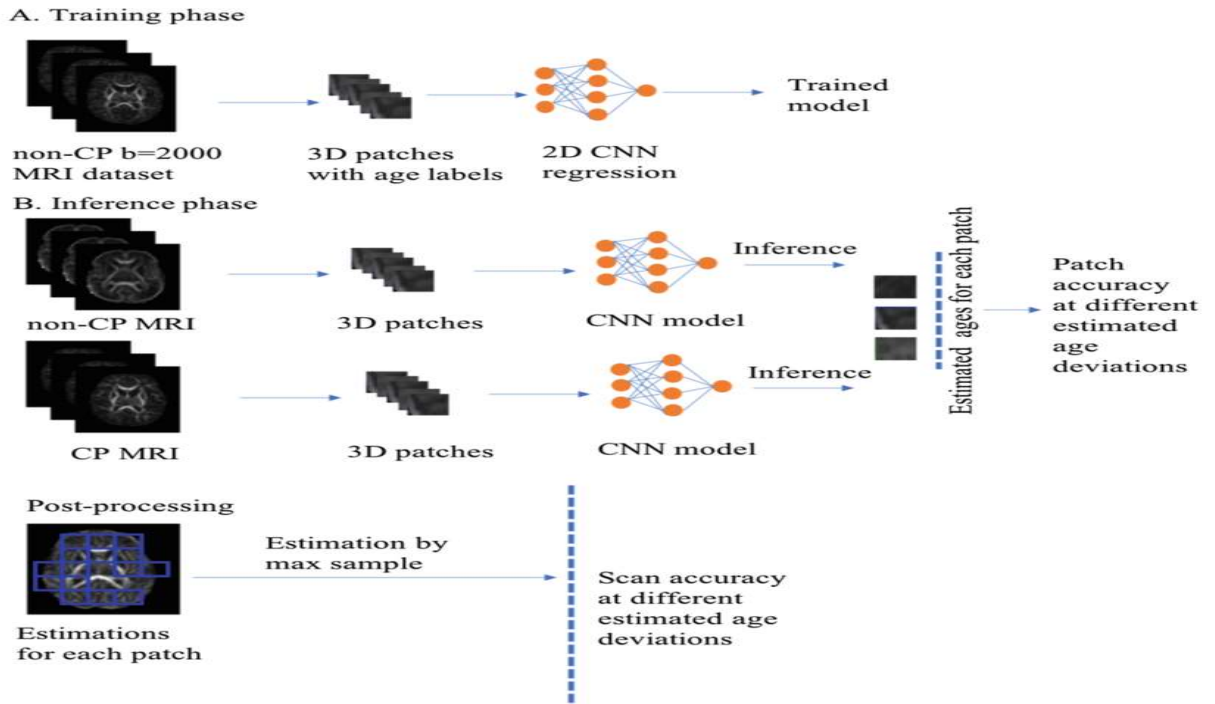


Fig 4.1 End to end training, inference and post-processing phases with a deep learning CNN regression network

4.1.3 Result

Using the Kaggle database, which contains seven emotions, we trained our Convolutional Neural Network model (happiness, anger, sadness, disgust, neutral, fear and surprise) The grayscale pictures of the discovered faces were then scaled to 4848 pixels and utilised as inputs to the CNN model. The emotion identification accuracy rate that we attained was 64%.

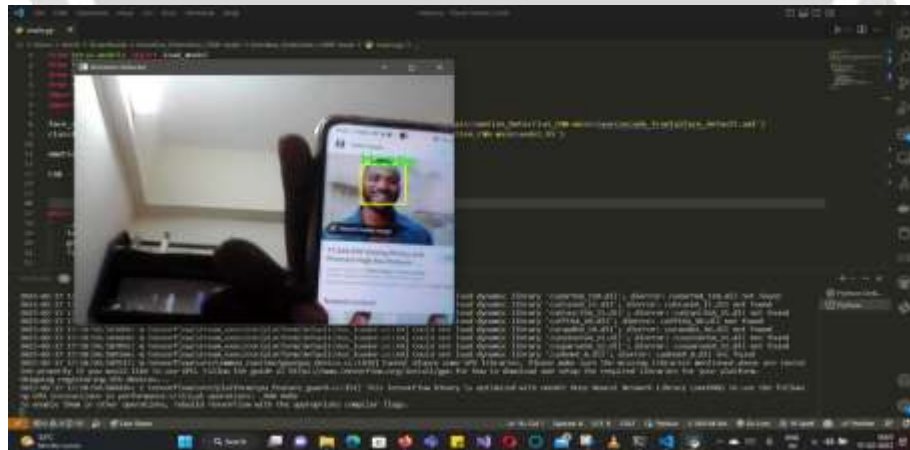


Fig4.2 emotion recognition

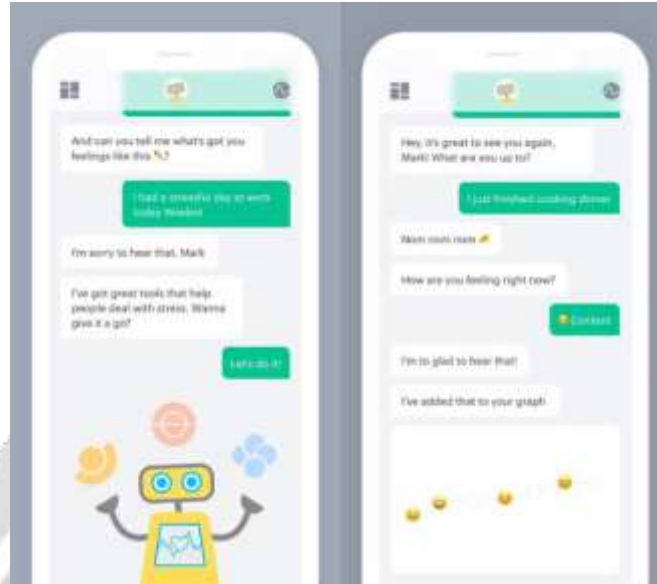


Fig 4.3 Chabot implementation

5. CONCLUSIONS

Numerous uses for an automated facial expression recognition system may be found in applications for human-computer interaction and psychological research. Because they can disclose a person's affective state, cognitive activity, personality, intention, and psychological condition, the system plays a communication function in interpersonal relationships. The system consists of three modules: face detection using HaarCascade; emotion recognition using CNN's Keras, which primarily focuses on identifying emotions that may indicate depression in a person. In the last module, a chatbot that can identify depression is employed to assist further distinguish between sorrow and depression.

6. REFERENCES

- [1] [1] Enrique Correa, Arnoud Jonker, Michael Ozo and Rob Stolk proposed their paper of emotion recognition using Convolutional Neural Network.
- [2] Emotion Detection using Viola Jones Algorithm- Kartika Candra Kirana, Slamet Wibawanto and, Heru Wahyu Herwant
- [3] A Survey Paper on Chatbots, Aafiya Shaikh, Dipti More, Ruchika Puttoo, Sayli Shrivastav, Swati Shinde Rehman, A.; Saba, T.
- [4] Performance analysis of character segmentation approach for cursive script recognition on benchmark database. Digit. Signal Processing 2011, 21, 486–490.
- [5] Dr. D.Venkataraman, Namboodri Sandhya Parameswaran, “Extraction of Facial Features for Depression Detection among Students”, International Journal of Pure and Applied Mathematics, 2018.
- [6] Ma Xiaoxi Lin Weisi, Huang Dongyan, Dong MinGhui, Haizhou Li, “Facial Emotion Recognition”, IEEE, 2017.

[7] Facial expression recognition based on Local Binary Patterns: A comprehensive study by Caifeng Shana, Shaogang Gongb, Peter W. McOwanb Nicu Sebe, Michael S,Lew ,Ira Cohen, Ashutosh Garg, Thomas S Huang.

[8] “Emotion Recognition using Cauchy Naive Bayes Classifier”,IEEE,2002.

[9] Automatic Facial Expression Recognition System Based on Geometric and Appearance Features- Aliaa A. A. Youssif, Wesam A. A. Asker

