JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

DETECTING FACIAL EXPRESSIONS OF EMOTIONS USING MACHINE LEARNING

[1] Gudiwaka Vijayalakshmi,[2] Pampana Nandini,
[1] Assistant professor, [2] MCA 2nd year,
[1]Computer Science and Engineering, [2]Master of Computer Applications,
Sanketika Vidya Parishad Engineering College, Visakhapatnam, India

ABSTRACT
CNNs have achieved significant progress in several domains, notably in computer vision. Their application in analysing images and distinguishing facial expressions [1] with exceptional accuracy has been particularly noteworthy. A framework based on convolutional neural networks [2] has been developed to identify human emotions [3] in real-time, leveraging dynamic facial expressions. The framework encompasses three stages: initial face detection [4] utilizing Haar Cascades [5], subsequent normalization, and emotion recognition [6] through a CNN [7] trained on the FER 2013 dataset [8], comprising seven distinct emotions. Experimental outcomes unequivocally demonstrate the feasibility [9] of facial emotion recognition, offering educators the opportunity to customize their presentations [10] based on their students' emotional states.

Keywords: facial expressions, human emotions, haar cascades, feasibility, recognition, emotional states

I.INTRODUCTION

Facial expressions are a powerful means of communicating [11] emotions, and detecting them accurately [12] can provide valuable insight [13] into a person's emotional state. The ability to detect facial expressions of emotion has numerous applications, from improving human-computer [14] interaction to enhancing medical diagnoses [15] and improving mental health care. With advances in computer vision [16] and machine learning [17], facial expression recognition has become a promising field of research, and several techniques have been developed for detecting and analysing facial expressions of emotion. This technology can recognize and interpret facial expressions, providing valuable information about a person's emotional state and helping us to better understand and communicate with each other.

II.SYSTEM ANALYSIS

System analysis for detecting facial expressions of emotions using machine learning involves analysing the requirements, design, and implementation of the system. Requirements analysis [18] involves identifying the needs and expectations of the users, stakeholders [19], and the system itself. In this case, the requirements may include high accuracy [20] in recognizing facial expressions, real-time processing, the ability to detect multiple emotions, and robustness to variations in lighting and facial features. The design phase involves creating a high-level architecture [21] of the system, selecting appropriate machine learning algorithms [22], and determining the data requirements. Facial expression recognition typically involves the use of convolutional neural networks (CNNs), which are designed to recognize patterns in images. The input data may include facial images, facial landmarks, or a combination of both. Implementation involves building and testing [23] the system. This may involve training the CNN model using a large dataset of facial expressions, such as the popular Facial Expression Recognition Challenge (FERC) dataset, and fine-tuning the model to improve accuracy. The system may also include pre-processing steps, such as face detection and alignment, to ensure that the input images are properly prepared for analysis. The performance of the system can be evaluated using metrics such as accuracy, precision, recall, and F1-score [24]. The system may also be tested on a variety of datasets to evaluate its robustness [25] and generalization ability. Overall, the system analysis for detecting facial expressions of emotions using machine learning involves a comprehensive understanding of the requirements, design, and implementation of the system, as well as ongoing evaluation and refinement to ensure high accuracy and robustness.

III.EXISTING SYSTEM

Convolutional neural networks (CNNs) are a popular type of deep learning algorithm that has been used for emotion detection in a variety of applications. EmoReact is an open-source dataset and benchmark for emotion recognition from facial expressions that uses a CNN-based architecture. The system uses a combination of convolutional, pooling, and fully connected layers to learn features from facial images and predict emotions. EmoReact has achieved state-of-the-art performance on benchmark datasets, demonstrating the effectiveness of CNNs for emotion detection.

IV.PROPOSED SYSTEM

A Convolutional Neural Network (CNN) is a type of deep artificial neural network used for identifying patterns in images without requiring extensive pre-processing. In contrast to traditional methods, CNNs learn filters automatically to extract relevant features. The primary building block of a CNN is a neuron, and these neurons are connected in layers such that the output of one-layer feeds into the input of the next. The first layer in a CNN is typically a Convolutional Layer, which extracts features from the input image. This layer performs convolution, a mathematical function that multiplies two matrices, one representing the image and the other representing a filter or kernel. This operation preserves the structural relationships between pixels and extracts features from the image. After the Convolutional Layer, a Pooling Layer is often added to reduce the spatial size of the feature maps, making the network more robust to variations in the input image. There are three types of pooling layers: Max Pooling, Average Pooling, and Global Pooling. Max Pooling selects the maximum element from a region of the feature map covered by the filter, while Average Pooling computes the average of the elements in the region. Global Pooling reduces each channel in the feature map to a single value, resulting in a 1x1xnc feature map. This can be accomplished using either global max pooling or global average pooling.

V.SPECIFICATION

A. HARDWARE REQUIREMENTS (Minimum Requirement)

1.RAM:4GB+RAM
2.PROCESSOR: i3 5th Gen 2.2 Ghz

B. SOFTWARE REQUIREMENTS

1.Domain: Python
2.Version: Python IDLE (3.8.0)
3.Code Editors: PyCharm
4.Frameworks and Dependencies: Tensor flow, Keras, Open CV
5.Operating System: Windows 10

C. CODE EDITORS

PyCharm

PyCharm is an integrated development environment (IDE) used in computer programming, specifically for the Python language. It is developed by the Czech company Jet Brains (formerly known as IntelliJ). It provides code analysis, a graphical debugger, an integrated unit tester, integration with version control systems (VCSes), and supports web development with Django as well as data science with Anaconda

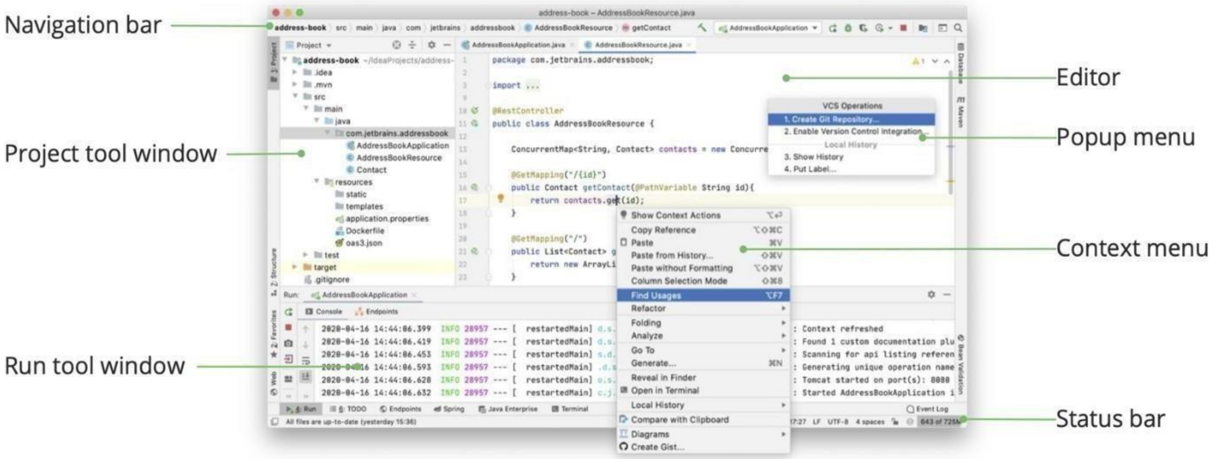

Figure1: PyCharm screen

Coding assistance and analysis, with code completion, syntax and error highlighting, linter integration, and quick fixes. Project and code navigation: specialized project views, file structure views and quick jumping between files, classes, methods and usages. Python refactoring: includes rename, extract method, introduce variable, introduce constant, pull up, push down and others. Integrated Python debugger. Integrated unit testing, with line-by-line code coverage. Google App Engine Python development. Version control integration: unified user interface for Mercurial, Git, Subversion, Perforce and CVS with change lists and merge. Support for scientific tools like matplotlib, NumPy and SciPy [professional edition only]. PyCharm provide an API so that developers can write their own plugins to extend PyCharm features. Several plugins from other JetBrains IDE also work with PyCharm. There are more than 1000 plugins which are compatible with PyCharm

VI.ALGORITHM

In this case we are using CNN algorithm

Convolutional Neural Networks (CNNs) have revolutionized various applications, becoming the most renowned deep learning architecture. Their widespread acceptance and success have fuelled the recent surge in interest in deep learning. CNNs emerged with Alex Net in 2012 and have since evolved significantly, progressing from eight layers to 152 layers in ResNet within just three years. CNNs have become the go-to model for image-related tasks, surpassing competing approaches with remarkable accuracy. Moreover, they find utility in recommender systems, natural language processing, and more. The key advantage of CNNs over their predecessors lies in their ability to automatically learn essential features without human intervention. For instance, when presented with numerous images of cats and dogs, CNNs autonomously learn distinctive characteristics for each group. In addition to their effectiveness, CNNs are computationally efficient. They employ unique convolution and pooling techniques, along with parameter sharing. Consequently, CNN models can now operate on diverse devices, making them universally accessible. In summary, CNNs appear almost magical. They offer highly efficient and eco-friendly models that achieve superhuman precision through automated feature extraction (even surpassing humans in image classification). This post aims to uncover the secrets and methodologies behind this remarkable approach.

CNN Algorithm

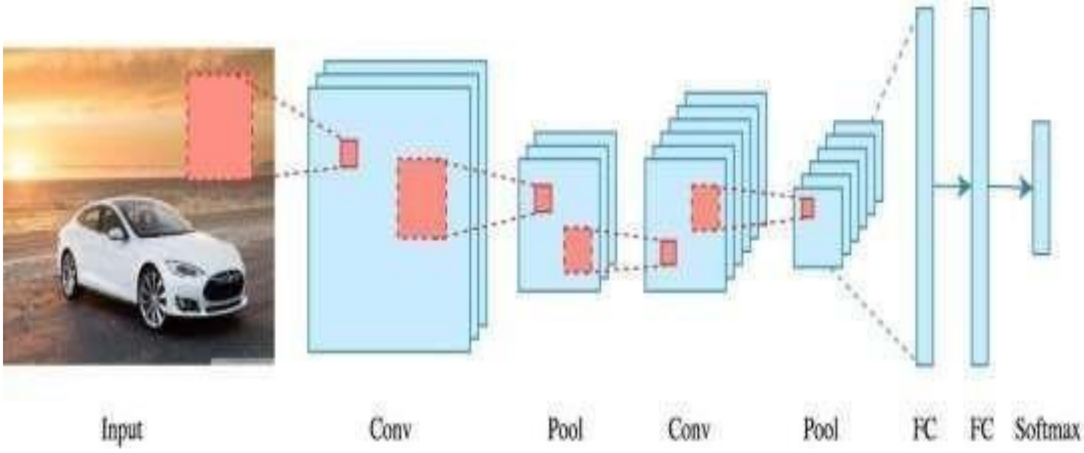

Figure2: CNN algorithm

The CNN, or Convolutional Neural Network [18], is a type of deep learning model designed to classify images by extracting key features from them. Compared to other classification algorithms, CNNs require minimal pre-processing of the input image

It is made up of four layers.

a) Convolution Layer
b) ReLu Layer
c) Pooling Layer
d) Fully Connected Layer

Figure3: pixels structure

This image is composed of three-color channels: Red, Green, and Blue. Images can be represented in various colour spaces, such as Grayscale, RGB, HSV, CMYK, and more. As image dimensions increase, such as in the case of 8k resolution (7680x4320), the CNN algorithm plays a crucial role in compressing the image into a more manageable format for processing while retaining essential features for accurate predictions.

A.Convolution Layer

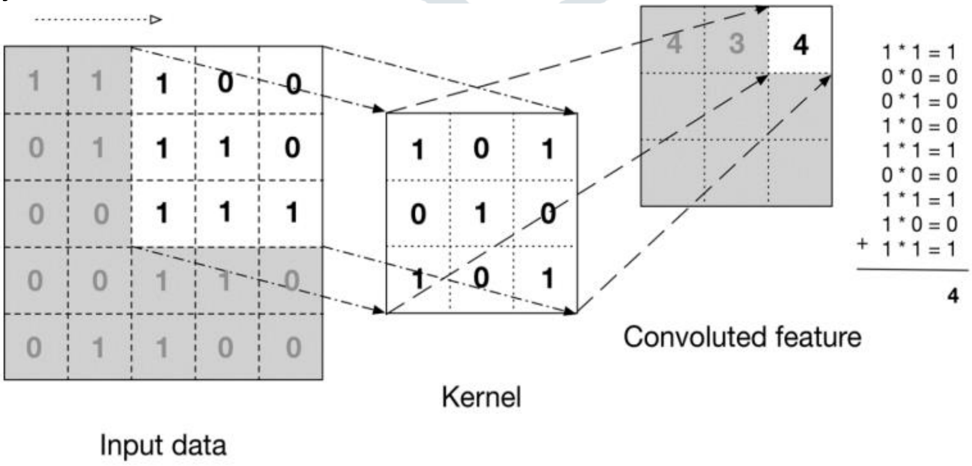

Figure4: convolution layer

In CNNs, grayscale images are processed by applying filters or kernels to extract features. These filters, often 3x3 or 5x5 in size, convolve with corresponding input arrays to generate filtered outputs, resulting in feature maps with reduced dimensions.

B.ReLu Layer

ReLU is a function that transforms negative pixel values in a layer to 0, making images more non-linear as they naturally are. It helps counteract linearity introduced during convolution, removing dark parts and keeping only positive values (grey and white colours). The rectified version of an image sets all negative pixel values to 0, distinguishing it from the non-rectified version.

C.Pooling Layer

The pooling layer reduces input size, improving computation efficiency and enhancing the spatial variance handling of CNNs. It also aids in preventing overfitting by summarizing discovered features and producing down sampled feature maps.

D.Fully Connected Layer

The primary objective of a convolutional neural network (CNN) is to extract informative features from input data for accurate image identification. The interconnected neurons in the network recognize patterns, transmit signals to the output layer, and assign classes based on weights, while the fully connected layer captures specific features and communicates them to the output classes in trained images.

VII.IMPLEMENTATION

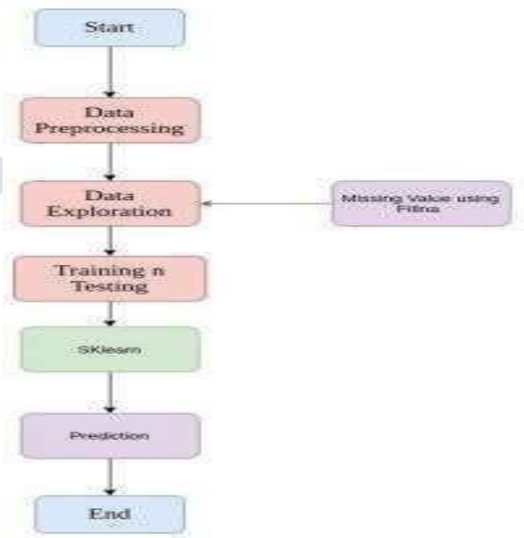

Figure5: Implementation

Data Collection: Data is collected from the Kaggle dataset, consisting of 35,500 grayscale 48x48 pixel photos of faces, categorized by facial expressions (happiness, neutral, sadness, anger, surprise, disgust, fear).

Data Pre-Processing: Data pre-processing techniques, including data augmentation such as cropping, padding, and horizontal flipping, are applied to enhance performance and increase data diversity. Pre-processed images are obtained from the Kaggle dataset for faster and more accurate processing.

Train And Test Modelling: The data is split into training and test sets, where the training set is used for model training, and the test set evaluates model performance. Data augmentation is employed to expand the training set, providing diverse images. Optionally, data augmentation can also be applied to the test set to observe the model's behaviour on augmented images.

Modelling: A CNN model is built and saved for further use. Four different models, including three neural networks implemented using Keras and TensorFlow, and a decision tree implemented using Sci-Kit Learn in Python 3.6, were evaluated. The training algorithms specific to each model are described alongside their model descriptions. The table displays the final accuracy comparison among the models.

VIII.REAL TIME TEASTING

Run the python app code.py. Immediately after running the app, the camera will be switched on automatically to capture the live video. A rectangular box willbe formed around the maximum area of the face to detect the emotion of the detected face from the camera. And the emotion of the face will be

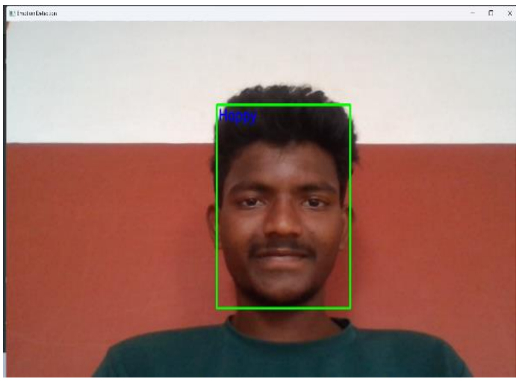

Figure6: happy face expression from live camara

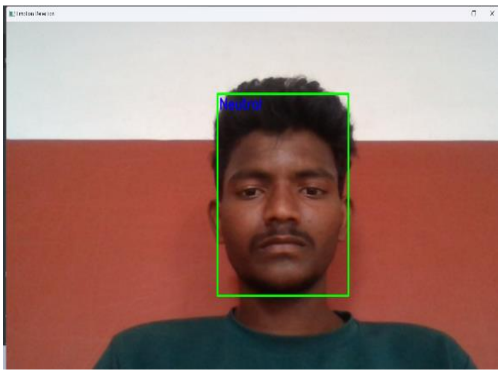

Figure7: neutral face expression from live camara

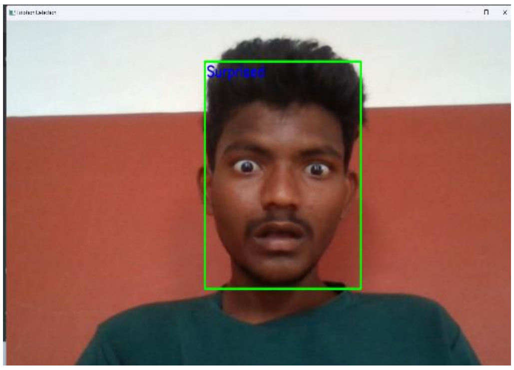

Figure8: surprised face expression from live camara

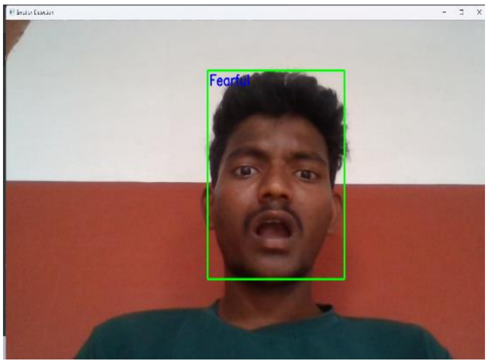

Figure9: fearful face expression from live camara

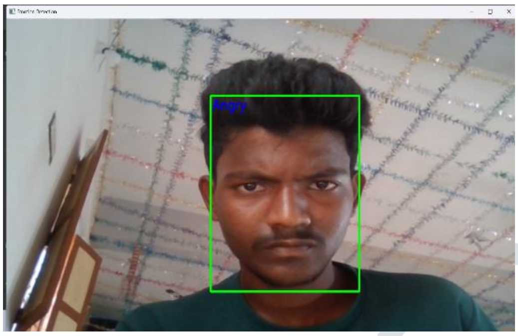

Figure10: angry face expression from live camara

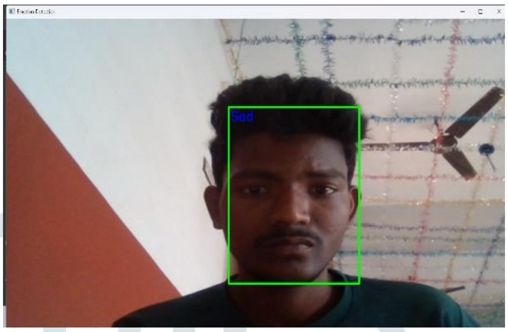

Figure11: sad face expression from live camara

The screenshot showcases the output of your Python code, which has successfully detected and captured facial expressions from a video.

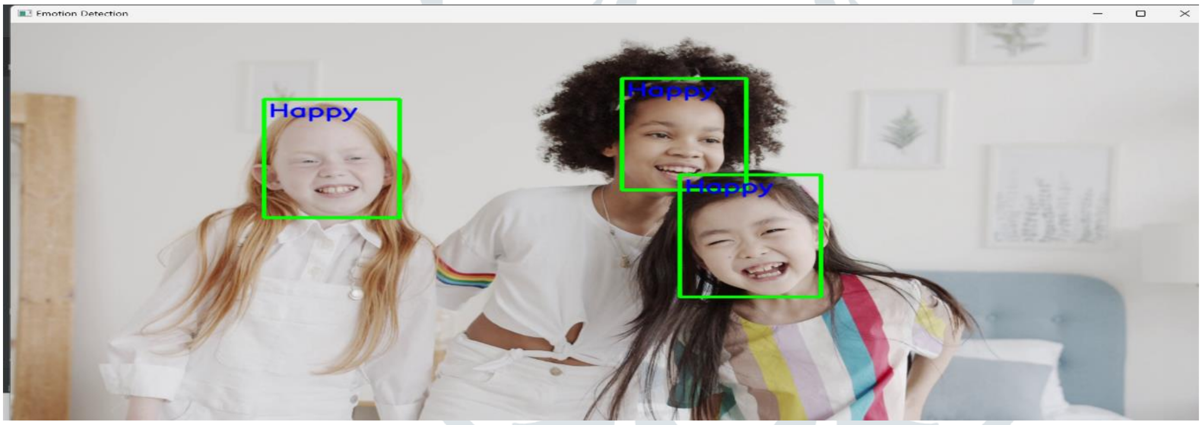

Figure12: facial expressions from recorded video

IX.CONCLUSION

In this project, we have introduced a novel methodology for facial expression identification using a CNN model that effectively extracts facial features. In contrast to existing methods, our proposed approach has the ability to automatically learn pattern features and mitigate the limitations caused by unnatural design elements. Our method, known as FERC (Facial Expression Recognition and Classification), leverages the strengths of CNN and supervised learning, made feasible by the availability of large datasets. One key advantage of the FERC algorithm is its capability to handle different facial orientations, including angles less than 30°, thanks to the utilization of a unique 24-digit-long EV feature matrix. Moreover, the inclusion of background removal significantly enhances the accuracy of emotion detection. The FERC algorithm can serve as a fundamental building block for various emotion-based applications such as lie detectors and mood-based learning systems for students. The proposed approach directly utilizes pixel values from sample training images, allowing for independent learning and the discovery of more hidden facial feature expressions. The training process incorporates suitable weight initialization, which plays a crucial role in weight adjustment. Through extensive experimental investigations, we have demonstrated that our proposed strategy can significantly improve the detection rate of facial expressions, even in challenging background conditions. Compared to existing approaches, our CNN model exhibits faster convergence and higher recognition rates, particularly in complex backdrop scenarios.

X.FUTURE SCOPE

The future scope of this work involves enhancing the accuracy and efficiency of emotion recognition systems by utilizing RGB datasets in uncontrolled conditions, incorporating deep neural networks as emotion classifiers, and considering compound emotions and micro expressions. The integration of multi-modal behavioural systems like body movements, facial expressions, and voice can contribute to the development of robust automatic recognition systems. Improving system security can be achieved through the utilization of cloud storage resources and cancellable biometrics. Additionally, further research is needed to analyse and improve the accuracy of the model's predictions, especially for challenging emotions such as disgust. The real-time capability and high accuracy of the model make it adaptable for various applications, including healthcare, marketing, and the video game industry.

XI.REFERENCES

[1]AN Article Reference of facial expressions
https://link.springer.com/article/10.1134/S1054661814010106
[2] AN Article Reference of neural networks
https://www.sciencedirect.com/science/article/abs/pii/S1751157719303979
[3] AN Article Reference of human emotions
https://link.springer.com/article/10.1007/s42761-022-00125-6
[4] AN Article Reference of face detection
https://link.springer.com/article/10.1007/s10462-018-9650-2
[5] AN Article Reference of Haar cascades
https://iopscience.iop.org/article/10.1088/1757-899X/1074/1/012033/meta
[6] AN article Reference of emotion recognition
https://www.frontiersin.org/articles/10.3389/fnins.2020.622759/full
[7] AN article Reference of CNN
https://link.springer.com/article/10.1007/s00202-020-01099-z
[8] AN Web Reference of FER 2013 dataset
https://www.mdpi.com/1424-8220/20/4/1087
[9] AN Web Reference of feasibility
https://onlinelibrary.wiley.com/doi/full/10.1111/hir.12108
[10] AN Web Reference of presentations
https://www.mdpi.com/2076-3417/10/20/7325
[11] A Book Reference of communicating
https://books.google.co.in/books?hl=en&lr=&id=7WWWZLaiv70C&oi=fnd&pg=PR7&dq=An+Article+Reference+of+communicating&ots=hWn9ypqYqp&sig=g0ixDv-Ai0EixoKFbPXqNH2_TkQ&redir_esc=y#v=onepage&q=An%20Article%20Reference%20of%20communicating&f=false
[12] AN article Reference of accurately
https://www.nature.com/articles/s41587-019-0074-6
[13] AN Web Reference of valuable insight
https://informationr.net/ir/6-3/paper105.html
[14] AN Article Reference of human-computer
https://www.sciencedirect.com/science/article/pii/S1110016820300168
[15] AN Article Reference of medical diagnoses
https://academic.oup.com/jamia/article-abstract/1/1/8/849986
[16] AN Book Reference of computer vision
https://dl.acm.org/doi/abs/10.5555/580035
[17] AN Article Reference of machine learning
https://link.springer.com/article/10.1007/s11192-018-2944-y
[18] AN Book Reference of Requirements analysis
https://dl.acm.org/doi/abs/10.5555/78225
[19] AN Web Reference of stakeholders
https://www.tandfonline.com/doi/full/10.1080/17453674.2021.1918389
[20] AN Article Reference of high accuracy
https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1005151
[21] AN Article Reference of high-level architecture
https://ieeexplore.ieee.org/abstract/document/9303396/?casa_token=AJrH6Wc7CA8AAAAA:zoxMgv-HEobHMd69JXZuNaxIi0V8xOcSjCDNnfkbzJVIBCC91EqL5GT_HYlEYmI5WJqbboFj3DI
[22] AN Article Reference of machine learning algorithms
https://www.ceeol.com/search/article-detail?id=1030033
[23] AN Book Reference of testing
https://books.google.co.in/books?hl=en&lr=&id=sU92osXzjAsC&oi=fnd&pg=PA1&dq=An+Article+Reference+of+testing&ots=Hpc_hyQeSQ&sig=F4izCCyJTQvdq8ViEr4pKWIq9CA&redir_esc=y#v=onepage&q=An%20Article%20Reference%20of%20testing&f=false
[24] AN Article Reference of F1-score
https://link.springer.com/article/10.1186/s12864-019-6413-7
[25] A Web Reference of robustness
https://www.tandfonline.com/doi/full/10.1080/21655979.2020.1870074

XII.BIBLOGRAPHY

Gudiwaka Vijaya Lakshmi working as an Assistant professor in the Department of Computer Science and Engineering, Sanketika Vidya Parishad Engineering College, Visakhapatnam Andhra Pradesh. with 2 years of experience in Computer science and engineering (CSE), accredited by NAAC. With her area of interests in java, python, .Net, HTML.

Pampana Nandini is studying her 2nd year, Master of Computer Applications in Sanketika Vidya Parishad Engineering College, affiliated to Andhra University, accredited by NAAC. With her interest in python, web development and as a part of academic project, detecting facial expression of emotions using machine learning. As a result of a desire to comprehend. In completion of his MCA.