International Journal of Computer Applications Technology and Research
Volume 12–Issue 01, 60-62, 2023, ISSN:-2319–8656
DOI:10.7753/IJCATR1201.1009

Smart Worker Monitoring System Using Facial Recognition and Deep Learning Techniques

Manikantha K
Asst Professor
Dept of CSE
BNM Institute of Technology
Bangalore, India

Kushagra Mittal
Dept of CSE
BNM Institute of Technology
Bangalore,India

Bikash Mishra
Dept of CSE
BNM Institute of Technology
Bangalore,India

Abhinav TL
Dept of CSE
BNM Institute of Technology
Bangalore,India

Abstract: Organizations are increasingly tracking staff gadgets and monitoring their output. We want to create a system that not only protects employees' privacy but also aids companies in managing their workforce and obtaining accurate productivity data. We were able to keep an eye on the background activity and record the names of the applications and their active times. A facial recognition system is used to implement the employee authentication procedure, maintaining authenticity.

Keywords: privacy; active; recognition; authentication; authenticity;

1 INTRODUCTION

The Covid-19 pandemic has significantly changed the way we work, go to college, participate in other activities, and live our lives. By 2025, remote employment was supposed to be the new standard, however due to the epidemic that is still going on, most businesses have already gone remote. While some organizations found it difficult to keep up in the beginning, the majority have subsequently adjusted. Companies have begun considering how to apply this working method in the long run. Business executives are beginning to understand that remote work is a permanent solution that will help businesses survive the pandemic. Home-based employment is here to stay. Future work will be carried out in this manner. The future of work will be conducted in this manner, and it will soon become the norm. More than half of all international businesses provide their employees the choice to work remotely. Only 16 percent of these companies have only remote employees; the other organizations have both on-site and remote employees. The ability to work from home promotes location freedom, a better work-life balance, and less stressful commutes.

1.1 OBJECTIVE

Even though most businesses now permit employees to work from home, they still have trouble managing projects, tracking tasks, and employee productivity. To accomplish these goals, other software must be used and controlled, which increases the workload. Employers urgently need to find ways to cut this overhead while still getting the most out of employee time and effort. It needs a comprehensive solution that not only helps the organizations but also considers the needs and privacy of the employees while preventing an imbalance. This system ought to be able to offer productivity analyses based on the needs of the organization. The management would benefit from the reports of this analysis in the overall review of the activity. The management would be assisted by the results of this If none, delete this of this analysis in conducting an overall evaluation of the employees work.

2 LITERATURE SURVEY

The details regarding the Literature survey conducted on the topic of 'Smart Worker Monitoring System Using Facial Recognition and Deep Learning Techniques' are provided in the Section. The papers refered to for the survey and their content are listed below

2.1 A Robust Psychologically-Oriented Emotion Recognition Method Using Transfer Learning

In this study, they use an image with the face region and an image with the eyes region as our training input data. We resize the image into 13251325 for the face region and 450900 for the eyes region because the image may not be the same size. The size in question corresponds to the average of the cropped data for the specified kinds. Cropping the face and eye regions are the two duties in the pre-processing phase. Cropping the face area out of the images is the first duty. The photographs from the datasets are not cohesive, and many of them have extraneous elements like white backgrounds and haircuts. This extraneous information could introduce noise into the training process because the entire image would be taught. To extract the area around the face, they use Haar Cascade. The Haar Cascade characteristics in our experiment produce fairly accurate results for the marking of the facial region. This preprocessed dataset will be applied to the training of the model on the entire face. Eyes Cropping is the second , In their initial they try at the eyes cropping method, they use Haar Cascade features to separate the eyes region from the cropped face region. However, this approach might not produce the results they were hoping for. Based on the results of their experiment, they define pleasing cropping as the crop that encompasses the eye region while leaving out the majority of other portions. the lower portion displays the outcome of the suggested strategy, whereas the upper part displays some unsatisfactory cropping using solely Haar Cascade characteristics. They can observe that the suggested strategy produces a more satisfying outcome in an area that covers the eyes region but does not include a large amount of additional region.

2.2 Facial Emotion Recognition using Convolutional Neural Networks

In this proposed paper, the authors trained the model using the Cohn Kanade and Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS) datasets to complete the facial emotion recognition challenge. Black and white images from a neutral state to an emotion were captured sequentially at Cohn Kanade. The six emotions in the dataset are anger, happiness, sorrow, fear, disgust, and surprise. Facial landmark points are provided for each image. Due to the lack of neutral images in the dataset, we split each image sequence into two parts, with the first half being labelled as neutral and the second part matching to the face expression. 3000 photos were the product of manual tagging. With the aim of extracting images from the RAVDESS collection, frames were obtained every 0.5 seconds. They received 11000 photos as a consequence. Since raw data contains a lot of noise and unused information, feeding it to the model is not recommended because this could lead to poor learning performance. Selecting the appropriate data preprocessing methods is just as crucial as choosing the appropriate model to develop. Since the majority of their data consisted of grayscale photos, which have one channel instead of three channels like rgb images do, they first transformed every image to grayscale in order to better extract features from the data. Second, they identified faces in photos and cut them out in order to reduce the noise. They did this to get rid of extra background information. Finally, they zoomed in on the image to make the pivot points in the faces more obvious. They performed the data preprocessing stage using the OpenCV framework. Choosing the appropriate machine learning algorithm was the last stage. Even subtle distinctions in appearance matter and can affect the outcome in their problem domain. For this challenge, they choosed to employ convolutional neural networks. A particular kind of artificial neural network called a CNN is one of the most effective and popular techniques for solving computer vision issues. The CNN is made up of a series of convolutional layers, each of whose outputs is only connected to small areas of the input. To do this, a filter is slid across the input, and at each point the dot product between the two is computed (creating a convolution between the input and filter). With the help of this structure, the model can learn filters that can identify particular patterns in the input data.

2.3 A Facial Expression Recognition Method Using Deep Convolutional Neural Networks Based on Edge Computing

The weaknesses of conventional data augmentation techniques are examined and compiled in this work. The paper enhances Cycle GAN, suggests a technique of facial expression identification based on constraint cycle consistent generation to resist network, and introduces class constraint condition and gradient penalty rule to address the issue of class imbalance in the existing face expression database. The experimental findings demonstrate that the enhanced generation model can learn the fine texture information of the face image more effectively, and the generated image has a high level of quality. The enhanced face expression picture recognition benefits from a stronger classification and recognition effect from the improved discriminator network. This paper studies facial expression recognition and expression image data enhancement. Although some achievements have been made, there are still some deficiencies, which need further research and improvement. First of all, the expression recognition and data enhancement in this paper are based on the static image, while the emotional changes in real life are of a certain timing, and the static image can only reflect the expression state of a person at a certain time. The next work will focus on the data enhancement of video sequence. Secondly, in the process of data enhancement, neutral expression image is used as the source domain and other expression images as the target domain, but the expression state of human in the real scene can be transformed at will. How to enhance the data without limiting the expression state of the input image is also a direction that can be improved in the future.

2.4 Facial Expression Recognition via Deep Learning

Their goal in this project is to use a new CNN architecture to improve the accuracy of facial emotion classification. We integrate multiple databases to get the final one because deep networks require a large database for training. After setting up the database, they fixed the CNN architecture's batch size input to 165 165, and they trained the architecture using fine-tuning from the Visual Geometry Group (VGG) model to produce the first model. They repeat training their CNN architecture in a second stage to increase classification, but this time the fine-tuning is accomplished using the initial model they have already obtained. Finally, they have their final model. The suggested network has four convolutional layers; the first three are followed by max pooling, and the last layer, which has 748 connections, is followed by a fully connected layer. It analyses facial photos and categories them into one of the six facial expressions: surprised, furious, disgusted, joyful, neutral, or happy. The MUG, RAFD, and Ck+ databases are used to analyze the proposed design. Results and identification rates show that our approach outperforms cutting-edge approaches. In order to train the model for this project, authors used pictures of faces that were in fixed positions. Authors want to expand our model to include other facial positions in future projects. This will enable us to examine the effectiveness of trained facial emotion detection models like VGGNet

2.5 Facial Emotion Detection Using Deep Learning

In order to evaluate the two models (Model-A and ModelB) on their capacity to detect emotions, they create a network based on the ideas from and. This section provides a description of the data used for training and testing, an explanation of the specific data sets used, and an assessment of the outcomes from the usage of two distinct datasets and two models. They have proposed a deep learning-based method in this research. method for reading facial expressions from a picture. they converse utilizing their proposed approach employing the JAFFE and FERC-2013. The proposed solution's performance assessment The application of a face emotion detection model is done in terms of accuracy of validation, computational difficulty, detection rate, learning rate, validation loss, and step-by-step calculation time, a trained and test sample were used to examine their suggested model. photos, then assess how well they performed in comparison to earlier current model. The experiment's findings demonstrate that the model In terms of the outcomes of emotion detection, suggested is superior. compared to earlier models mentioned in the literature. The research demonstrate that the suggested model is generating cutting-edge repercussions on the two datasets.

2.6 Discussions of Different Deep Transfer Learning Models for Emotion Recognitions

This study gave clear and in-depth comparisons of the models and looked at the transfer learning of five CNN models. This work can be used by researchers to find suitable models and transfer learning techniques in situations where there are software or hardware restrictions. Authors looked at various CNN models, which have different frameworks. They are typical transfer learning models used in a variety of disciplines: All of the frameworks for ResNet-50, Exception, EfficientNet-B0, Inception, and Dense Net differ from one another. Using various data preprocessing approaches, training methodologies, and models, we carried out extensive trials on datasets of varied sizes. Following a comparison of different data preprocessing methods and consideration of the significant sample size differences among all FER2013 emotion classes as well as the generalizability of both large and small datasets, they chose to employ class weight training. The accuracy of all models improved when they underwent freeze + fine-tuning, whether with the large dataset Affect Net or the small dataset FER2013 (with the exception of EfficientNet-B0, whose accuracy remained constant, and DenSeNet121, which became less accurate than the other models). This was the case when ImageNet was used as the source domain. Additionally, it was found that using Affect Net as the source domain instead of ImageNet resulted in less influence from transfer learning techniques and no appreciable differences in training results in a multi-source transfer learning experiment using FER2013 on source domains related to the target domain. This result revealed that selecting the appropriate source domain can improve transfer learning models more significantly than selecting the appropriate training approach. However, producing the best training results in transfer learning requires using several source domains and training methods, depending on the model.

3 CONCLUSION

As job monitoring and tracking become an important component of remote working conditions, this software attempts to help enterprises achieve their goal of higher productivity. Observations showed that the software was able to successfully track the majority of background applications and URLs accessed. In addition to tracking them, it also kept track of their names and how long background apps were running.

4 REFERENCES

[1] W. Y. Mak, K. W. Sum and K. Y. Chan, ”A Robust Psychologically- Oriented Emotion Recognition Method Using Transfer Learning,” 2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC),

[2] Z. Rzayeva and E. Alasgarov, ”Facial Emotion Recognition using Convolutional Neural Networks,” 2019 IEEE 13th International Conference on Application of Information and Communication Technologies (AICT), 2019

[3] An Chen, Hang Xing, and Feiyu Yang, “A Facial Expression Recognition Method Using Deep C

[4] A. Fathallah, L. Abdi and A. Douik, ”Facial Expression Recognition via Deep Learning,” 2017 IEEE/ACS 14thInternational Conference on Computer Systems and Applications (AICCSA), 2017, pp. 745-750, doi: 10.1109/AICCSA.2017

[5] A. Jaiswal, A. Krishnama Raju and S. Deb, ”Facial Emotion Detection Using Deep Learning,” 2020 International Conference for Emerging Technology (INCET), 2020.

[6] C. –T. Yen and K. –H Li “Discussions of Different Deep Transfer Learning Models for Emotion Recoginitions” in IEEE Access, vol. 10, 2022