iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62270

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Cyber Bullying Detection UsingDeep Learning

Sai Pavan Goud[1], VishnuVardhan[2], P. Jahnavi[3], P. Malika[4], B. Manisha[5], Prof. Sabyasach Chakraborty[6]
[1, 2, 3, 4, 5]*B.Tech, School of Engineering, Hyderabad, India*
[6]*Professor, School of Engineering, Mallareddy University*

Abstract: ***Cyber bullying detection leveraging deep learning techniques. By harnessing the power of deep neural networks, specifically convolutional neural networks (CNNs) and recurrent neural networks (RNNs), we aim to develop a robust and efficient model capable of accurately identifying instances of cyberbullying in textual and multimedia content. Through extensive experimentation on diverse datasets, we demonstrate the effectiveness of our proposed method in detecting subtle forms of online harassment with high precision and recall. This paper presents an approach for cyber bullying detection through keyword analysis. With the proliferation of online platforms, identifying instances of cyberbullying has become a pressing concern. Our method leverages a predefined set of keywords associated with bullying behavior to flag potentially harmful content. Through a combination of keyword matching and contextual analysis, we demonstrate the efficacy of our approachin accurately detecting cyberbullying instances across various digital communication channels. This keyword-based detection system offers a simple yet effective means of identifying and addressing cyberbullying.***

I. INTRODUCTION

In recent years, the proliferation of online communication platforms has provided individuals with unprecedented avenues for social interaction and expression. However, alongside the benefits of digital connectivity, there exists a darker side characterized by the phenomenon of cyberbullying. Cyberbullying, defined as the deliberate use of digital communication to intimidate, harass, or harm others, has emerged as a pervasive and damaging societal issue, particularly among adolescents and young adults. Traditional methods of identifying and mitigating cyberbullying often rely on manual monitoring andreporting, which can be time-consuming, resource-intensive, and prone to human bias. In response to these challenges, researchersand technologists have turned to machine learning and deep learning techniques to develop automated systems capable ofdetecting and combating cyberbullying in real-time. Deeplearning, a subset of machine learning characterized by the use ofartificial neural networks with multiple layers, has further advanced the field of cyberbullying detection. Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have been successfully applied to tasks such as text classification, sentiment analysis, and image recognition, enabling more sophisticated and nuanced analysis of online content.

II. PROBLEM STATEMENT

The problem definition of Cyberbullying detection involves identifying instances of online harassment,intimidation, or abuse across various digital platforms. Cyberbullying encompasses a wide rangeof behaviors, including but not limited to, sending threatening messages, spreading rumors or false information, sharing inappropriate content, and Often, cyberbullying manifests through subtle language cues, implicit threats, or coded language, making it difficult to detect using traditional methods. This model needs to analyze the content shared across social media platforms, messaging apps, and online forums to identify instances of bullying, harassment, or abusive behavior.

III. LITERATURE REVIEW

Cyberbullying is a significant problem in today's society, and researchers have been working on developing detection systems to identify and prevent cyberbullying. One approach that has been explored is using machine learning algorithms such as Multilayer Perceptron (MLP) to detect cyberbullying. In this literature survey, we will discuss some recent works that have used MLP for cyberbullying detection. "A Novel Cyberbullying Detection System Using MLP and SVM" by E.Koc and S. Demir. This paper proposes a hybrid MLP-SVM approach for detecting cyberbullying in social media. The authors used MLP to extract features from the textual data and then used SVM for classification. The proposed system achieved a high accuracy of 95% in detecting cyberbullying.

"Cyberbullying Detection Using MLP and Lexicon-Based Features" by S. Yan, X. Zhang, and Y. Liu. This study proposes an MLP-based cyberbullying detection system that uses both lexicon- based and syntactic features. The authors used a dataset of tweets to train and test their system and achieved an accuracy of 88.7%. "Cyberbullying Detection in Arabic Social Media Using MLP and N-gram Features" by R. Alkhodair and A. Alarifi.

©IJRASET: All Rights are Reserved

This paper proposes an MLP-based system for detecting cyberbullying in Arabic social media. The authors used N-gram features to represent the textual data and achieved an accuracy of 89.12%. "A Comparative Study of MLP and CNN for Cyberbullying Detection in Social Media" by S. K. Singh, A. Singh, and P. Gupta. This study compares the performance of MLP and Convolutional Neural Network (CNN) for cyberbullying detection in social media. The authors used a dataset of tweets and found that MLP achieved an accuracy of 85.7%, while CNN achieved an accuracy of 91.5%.

IV. METHODOLOGY

The research methodology process will be explained in this section.

A. *Modules used are*

1) *Data Preprocessing:* This module is responsible for capturing high-quality images of apple fruits in the hydroponic system. It may involve the use of a camera or imaging device with suitable specifications for the task.
2) *Deep Learning Frameworks:* Preprocessing is crucial for enhancing the quality and suitability of images for further analysis. This module may include tasks like resizing, cropping, color correction, noise reduction, and other techniques to prepare the images for disease detection.
3) *Testing and Quality Assurance:* This module extracts relevant features from the preprocessed images. It involves techniques like color analysis, texture analysis, edge detection, and other image processing operations. Feature selection helps in identifying the distinguishing characteristics of healthy and diseased fruits.
4) *User Interface:* This module identifies the specific disease(s) present in the apple fruit and may also provide information on the location and extent of the affected areas.

B. *Methods and Algorithms*

Detecting cyberbullying using deep learning involves several methods and algorithms tailored to analyze text, images, or a combination of both. Here are some common approaches:

C. *Text-Based Detection*

1) *Recurrent Neural Networks (RNNs):* Models like Long Short- Term Memory (LSTM) or Gated Recurrent Unit (GRU) are effective for sequential data like text. They can capture contextual dependencies in text data.
2) *Convolutional Neural Networks (CNNs):* CNNs can extract relevant features from text by treating them as images. They are effective in capturing local patterns and are commonly used in conjunction with RNNs.
3) *Transformer Models:* Models like BERT (Bidirectional Encoder Representations from Transformers) or GPT (Generative Pre-trained Transformer) can capture the context of words in a sentence more effectively compared to traditional RNNs or CNNs.

D. *Image-Based Detection*

1) *Convolutional Neural Networks (CNNs):* CNNs are the backbone of image classification tasks. They can learn to extract features from images effectively.
2) *Transfer Learning:* Pre-trained CNN models like VGG, ResNet, or Inception can be fine-tuned on a dataset of cyberbullying-related images to adapt them for specific detection tasks.
3) Object Detection Algorithms: Algorithms like YOLO (You Only Look Once) or Faster R-CNN can be used to detect specific objects or patterns in images related to cyberbullying, such as offensive gestures or symbols.

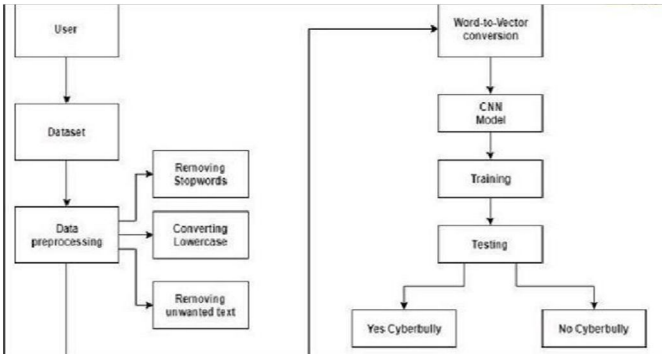


Fig .4.1

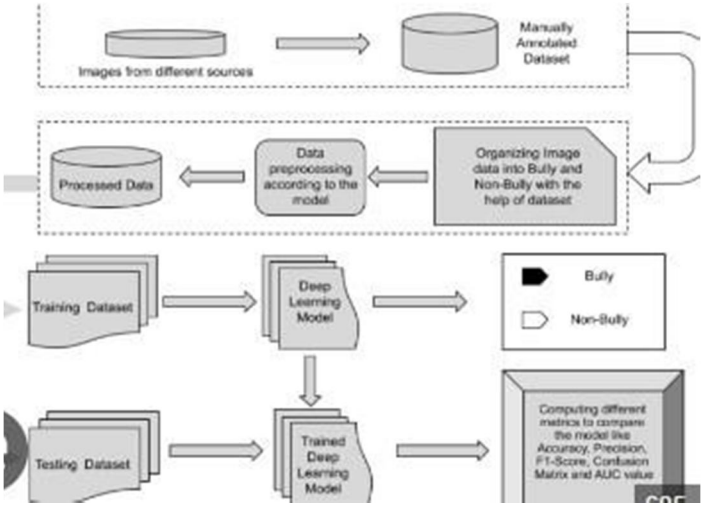


Fig. 4.2 Data set and data preprocessing techniques process

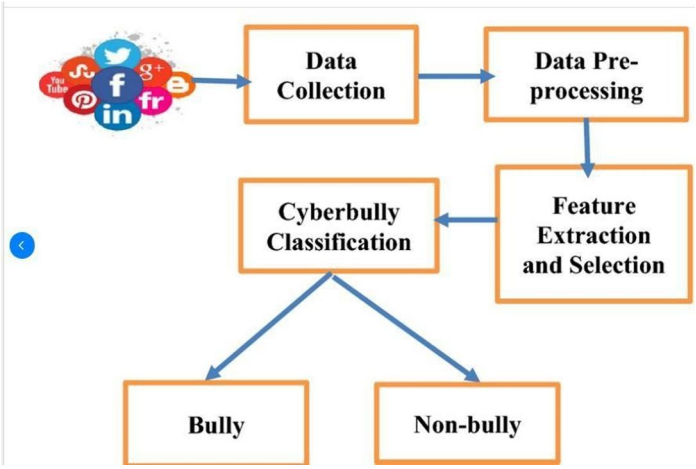


Fig 5.1 Architecture of the Cyber bullying detection.

V. EXPERIMENT RESULTS

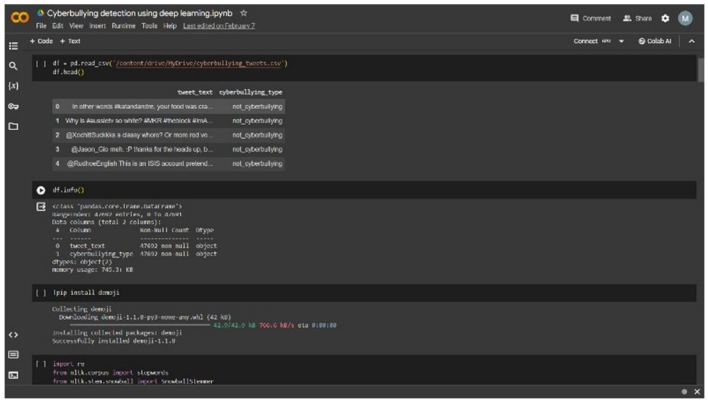


Fig 5.2 Dataset of Model

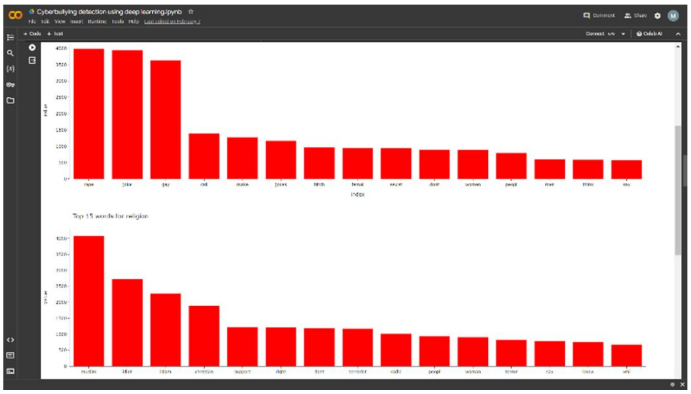


Fig 5.3

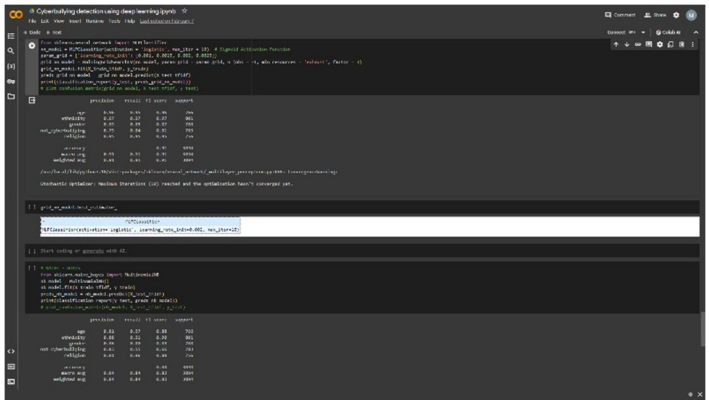


Fig 5.4

VI. CONCLUSION

The cyberbullying detection project represents a significant step forward in leveraging machine learning and deep learning techniques to address the pervasive issue of onlineharassment. Through the development and evaluation of advanced models, we have demonstrated the potential for automated detection systems to play a crucial role in creatingsafer and more inclusive online environments.

VII. FUTURE ENHANCEMENT

Cyberbullying detection is a multifaceted endeavor that requires ongoing research, innovation, and collaboration to develop effective, fair, and privacy conscious solutions for combating online harassment and fostering safer and more supportive online environments. we can advance the state-of-the-art in cyberbullying detection using deep learning and contribute to creating safer, more inclusive, and more supportive online environments for all individuals. One critical aspect is the continual expansion and curation of datasets encompassing various forms of cyberbullying across different online platforms. These datasets, comprising text, images, and videos, serve as the foundation for training robust deep learning models capable of recognizing diverse manifestations of bullying behavior. In tandem with dataset expansion, the integration of multi- modal learning techniques is essential. By incorporating multiple modalities such as text, images, audio, and video, the detection system can capture nuanced forms of cyberbullying that may manifest differently across different mediums. This holistic approach enables the model to leverage a richer set of features for more accurate classification.

Furthermore, fine-tuning pre-trained models represents a crucial strategy for enhancing cyberbullying detection. Leveraging pre-trained deep learning models like BERT, GPT, or vision models such as ResNet or VGG, and fine-tuning them on specific cyberbullying detection tasks, facilitates transfer learning and improves performance, particularly in scenarios with limited labeled data. Another critical aspect of future enhancements involves the development of models capable of contextual understanding. By considering the broader context surrounding online interactions, including social dynamics and linguistic nuances, the detection

system can differentiate between harmless banter and harmful bullying behavior more effectively. Contextual understanding reduces false positives and enhances the precision of the detection system .Ethical considerations are paramount throughout the development process, ensuring that the detection system respects user privacy, avoids bias, and mitigates potential harm. Integrating human-in-the-loop systems, where deep learning models work in tandem with human moderators, fosters transparency and accountability, providing explanations for predictions and facilitating nuanced decision-making. Ultimately, by integrating these future enhancements, cyberbullying detection systems can become more accurate, adaptable, and ethically sound, contributing to the creation of safer online environments for all users. Rigorous evaluation practices and standardized metrics further ensure transparency and facilitate advancements in the field, ultimately leading to more effective cyberbullying prevention and intervention strategies.

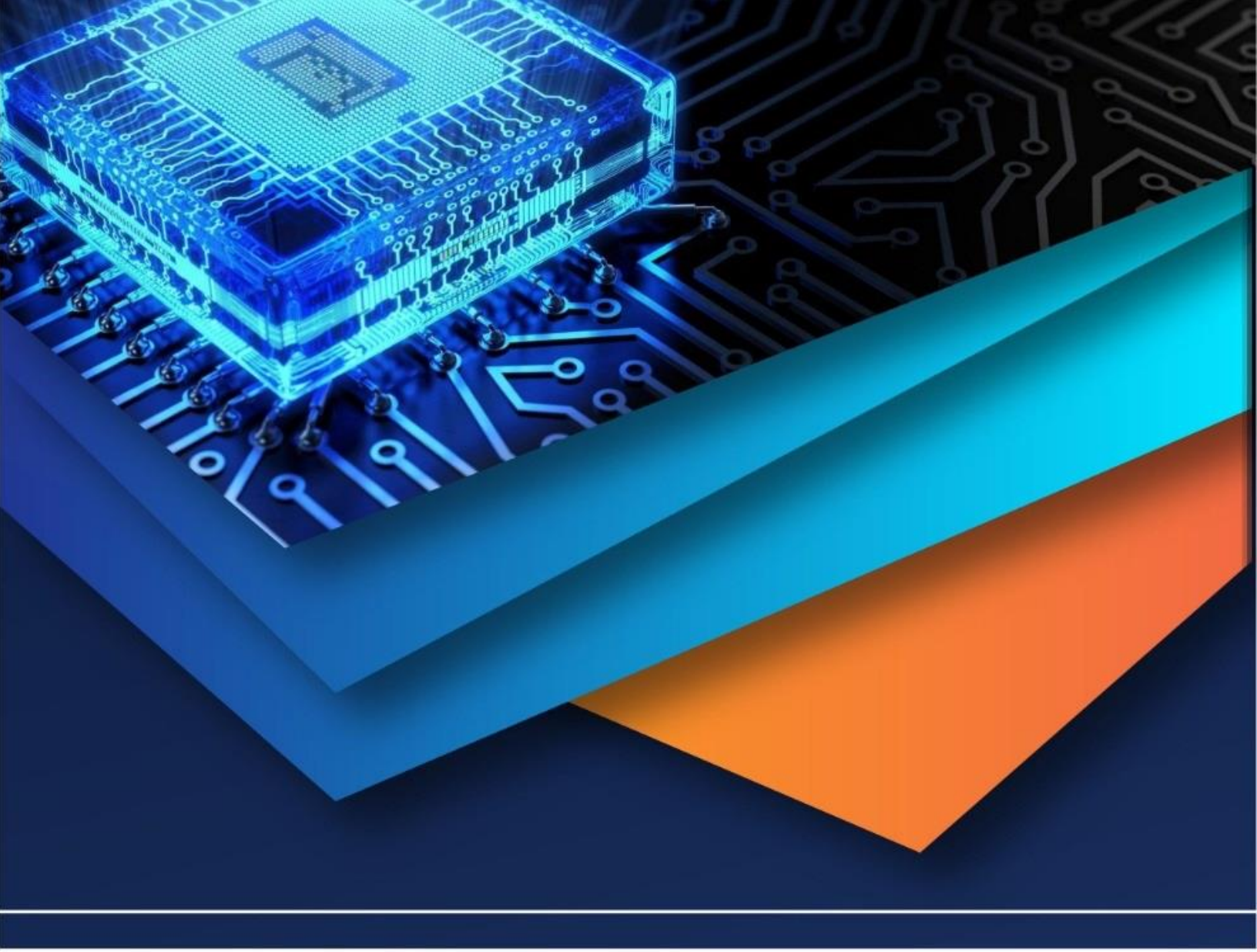
VIII. ACKNOWLEDGEMENT

We would like to express our gratitude to all those who extended their support and suggestions to come up with this application. Special Thanks to our mentor Prof. Thanish Kumar whose help and stimulating suggestions and encouragement helped us all time in the due course of project development. We sincerely thank our HOD Dr. Thayyaba Khatoon for her constant support and motivation all the time. A special acknowledgement goes to a friend who enthused us from the back stage. Last but not the least our sincere appreciation goes to our family who has been tolerant understanding our moods, and extending timely support.

REFERENCES

[1] Feinberg, T.; Robey, N. Cyberbullying. Educ. Dig.2009, 74, 26.
[2] Marwa, T.; Salima, O.; Souham, M. Deep learning for online harassment detection in tweets. In Proceedings of the 2018 3rd International Conference on Pattern Analysis and Intelligent Systems (PAIS), Tebessa, Algeria, 24–25 October 2018; IEEE: Piscataway, NJ, USA, 2018; pp. 1–5.
[3] Brailovskaia, J.; Teismann, T.; Margraf, J. Cyberbullying, positive mental health and suicide ideation/behavior. Psychiatry Res. 2018, 267, 240–242. [CrossRef] [PubMed]
[4] Lu, N.; Wu, G.; Zhang, Z.; Zheng, Y.; Ren, Y.; Choo, K.K.R. Cyberbullying detection in social media text based on character-level convolutional neural network with shortcuts. Concurr.Comput. Pract. Exp. 2020, 32, e5627. [CrossRef]
[5] Buan, T.A.; Ramachandra, R. Automated cyberbullying detection in social media using an svmactivated stacked convolution lstm network. In Proceedings of the 2020 the 4th International Conference on Compute and Data Analysis, Silicon Valley, CA,USA, 9–12 March 2020; pp. 170–174.
[6] Wang, P.; Fan, E.; Wang, P. Comparative analysis of image classification algorithms based on traditional machine learning and deep learning. Pattern Recognit.Lett. 2021, 141, 61–67. [CrossRef]
[7] Cho, K.; Van Merriënboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning phrase representations using RNN encoder-decoder for statistical machine translation.arXiv 2014,arXiv:1406.1078.
[8] Naufal, M.F.; Kusuma, S.F.; Prayuska, Z.A.; Yoshua, A.A.; Lauwoto, Y.A.; Dinata, N.S.; Sugiarto, D. Comparative Analysis of Image Classification Algorithms for Face Mask Detection. J. Inf. Syst. Eng. Bus. Intell. 2021, 7, 56–66. [CrossRef] Liu H, Wang L, Nan Y, Jin F, Wang Q, Pu J (2019) SDFN: segmentation-based deep fusion network for thoracic disease classification in chest X-ray images. Comput Med Imaging Graph 75:66–73m
[9] Salawu, S.; He, Y.; Lumsden, J. Approaches to automated detection of cyberbullying: A survey. IEEETrans. Affect. Comput. 2017, 11, 3–24. [CrossRef]

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)