IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

KEEPING CHILDREN SAFE ONLINE WITH LIMITED RESOURCES

[1]Nayana.S, [1] Devika Ayyappan, [1]Shabana.S, [2]Resmi B.S

[1]UG Scholar, Department of Computer science and Engineering,

[2]Asst Prof, Department of Computer Science and Engineering, UKF College of Engineering and Technology,

Kerala, India

ABSTRACT: ***Nowadays increase in internet use and facilitating access to social media platforms has helped the predatory to establish online relationships with children which has boosted to increase in online solicitation. The main of our system is to detect child predators based on chat, comments, and posts on social media accounts and sent predator records to cyber cell admin and user of the PAN12 data set is done for the next classification purpose.***

Keywords: Machine learning, Training Module, Neural Network, Data Set, Casper.

1. INTRODUCTION

This is a widely studied field, both from a psychological standpoint of how the children's well-being is affected by these types of dangers [4], and from a cyber-security perspective, which focuses on detecting and preventing the occurrence of unwanted content. Individual systems exist that are used to detect unwanted images [3], videos [8], texts [2], and audios [6]. We apply the algorithm on a set of captured browsing sessions that contain both appropriate and inappropriate materials. We measure success by our ability to detect inappropriate parts of the manually labelled data-set. Our main contributions are our modular system architecture, as well as our manually labelled data-set, which was used for both training, validation and testing or our image detection algorithm from screenshots

2. RELATED WORKS

A lot of child predator detection systems exists in the fields of audio chat, gamming and in various other online entertainment platforms. During playing games or audio chatting, a child predator system exits that can detect online sexual harassment and prevents children from abuse or harassment by sexual predators. As children are actively using social media, it requires a child predator detection system.

3. PROPOSED SYSTEM

3.1 THE AUDIO MODULE

All audio input and output components, such as the microphone and speakers should be accessible and controllable. The audio module does not have any built-in classification or analysis capabilities, as it is deferred to the text processing module. [5] is an ideal option for this transcription task, because it can be integrated at the OS level, making the audio signals completely available to our solution. Kaldi is appropriate for the children's protection context mainly because it is flexible in controllingall the parts of the speech-to-text conversation and could easily adapt to different noisy environments by integrating different acoustic modeling scripts at the operating system level.

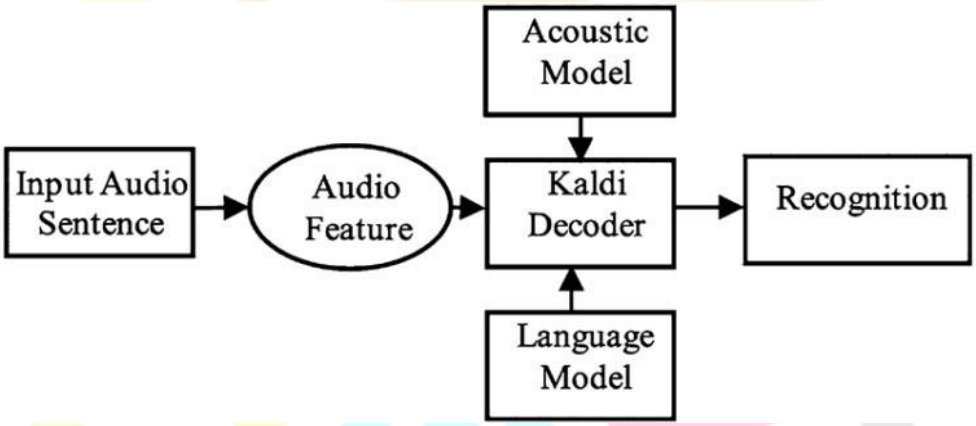

Fig 3.1.1: The design of the transcription system

3.2 THE IMAGE PROCESSING MODULE

The main goal of the image processing module is to detect image/video content on the user's screen from the screenshots recorded by the developed HCI Tracker software and for the classification of the identified content as appropriate or inappropriate for kids. One screenshot is saved every 3 seconds by the software during the sessions. Each of the saved screenshots was manually annotated by using the Yolo Mark [1], annotating tool [7]. This resulted in the screen segmentation dataset that contains 5967 image areas and 4052 textual areas respectively. The image areas were annotated as containing pornography, nudity, or as being neutral, but this

information was ignored for the purpose of training a network that will be able to isolate image areas. Larger image databases existfor objectionable content detection.

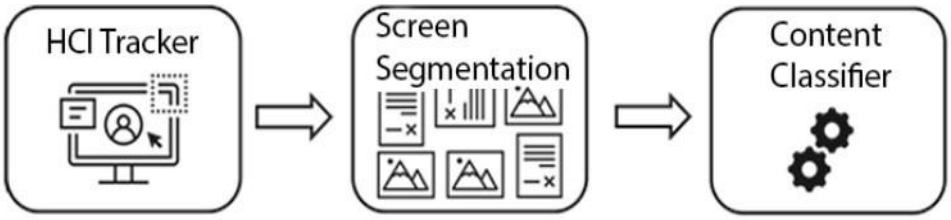

Fig 3.2.1: The design of the image processing module

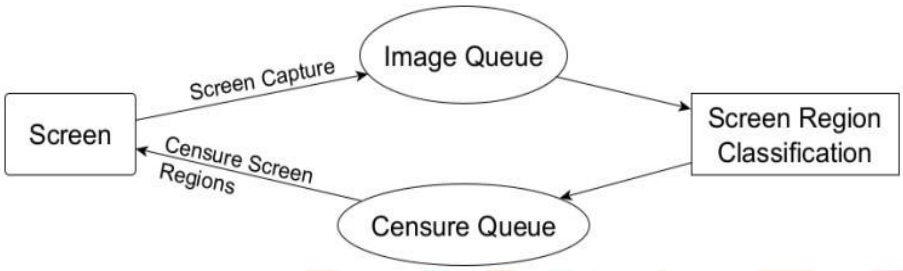

Fig 3.2.2: The design of the image processing module

3.3 THE TEXT PROCESSING MODULE

The main aim of this module is to implement multilingual OCR and cyberbullying detection algorithms. The input to these algorithms is the frames which are basically screenshots from the user's screen. These algorithms should be capable of detecting and recognizing cyberbullying from recorded frames/images/screenshots that contain text written in an arbitrary language. The text detection and recognition algorithms are capable of detecting the region with text content that is present on the screen and recognizing the text within text regions.

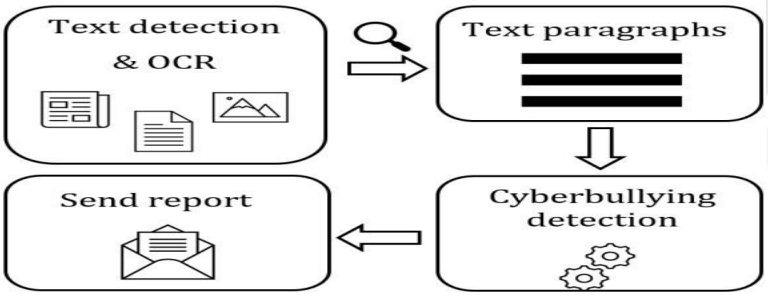

Fig 3.3.1 The workflow of the text module

3.4 CASPER SYSTEM ARCHITECTURE

This work focuses on extracting text via the speech recognition and OCR modules and also with the visual content by first identifying the visually salient non text regions and

analyzing them via the Non Text visual content and content analysis modules, respectively. The A.I. module integrates the outputs of the Image and text analysis, as seen in the figure. Alternatively individual modules, such as the content analysis module, can independently reach their own categorizations based on the content they see, and take independent actions, such as blocking certain content, or notifying the children's parents.

4. RESULT ANALYSIS

We are visiting to discuss the results in terms of the classification accuracy for various classification methods. Online child grooming is defined as a process to approach, persuade, and interact with a baby, the victim, in gender by using the internet as a medium. Perpetrators approach the victim to create not only sexual but also emotional relationships. The massive proliferation of social media has opened possibilities for perpetrators to conduct the crime of online childgrooming on an exceeding large scale. To reveal this kind of crime, the investigator usually relies on the conversation texts where the grooming patterns are carefully analyzed. With the vast amountof conservation text data, the tactic becomes extremely difficult and recures a significant amount of sometimes. The manual approach to investigating grooming patterns is additionally error-prone. Besides, the grooming process typically takes a pair of months on the common. Finally, we proposea straightforward classification method, which needs very low computational cost and makes it suitable for implementation within electronic mobile devices. The proposed method is to classify the conversation on the premise of this number of grooming characteristics. This method is suggested by observing the actual fact that the number of grooming characteristics is markedly different.

5. CONCLUSION

An automatic system to detect online child grooming has an important role in analyzing the vast amount of conversation texts. For this reason, many studies have been performed using various pattern detection schemes. In the current work, seventeen characteristics of grooming conversation are identified and utilized for classification. As each and everyone, even children is using the internet nowadays and getting harassed by predators so in order to stop these predators it is very important to detect and punish them. The main aim of the groomer is to build a relationship with a child in order to gain access to that child. When grooming takes place, it is common that an adult groomer is pretending to be a child with common hobbies or interests to build a relationship with the child. In this project, we detect child predators for child safety. And send predator reports to cyber admin for action.

6. ACKNOWLEDGEMENT

We would wish to thank everyone who contributed to the successful completion of this project. We could like to express our gratitude to our project guide Ms. Resmi B.S Assistant Professor in

computer science and engineering who gave valuable suggestions for our project. We express our deep-felt gratitude to beloved HOD Dr.Ramani.K, heads of the department for providing necessary information regarding the project and also here support in completing it. We also thank our project coordinator Mr. Jithin Jacob, Assistant Professor who gave expert supervision encouragement, and constructive criticism amidst his busy schedule throughout the project. we are also grateful to all the authors of the books and papers which have been referred to publish this paper.

REFERENCE

[1] N. AlDahoul, H. Karim, M. Abdullah, M. Fauzi, A. Wazir, S. Mansor, and J. See, ‘‘Transfer detection of YOLO to focus CNN’s attention on nude regions for adult content detection,’’ *Symmetry*, vol. 13, no. 1, p. 26,2020.

[2] C. Van Hee, G. Jacobs, C. Emmery, B. Desmet, E. Lefever, B. Verhoeven, G. De Pauw, W. Daelemans, and V. Hoste, ‘‘Automatic detection of cyberbullying in a social media text,’’ *PLoS ONE*, vol. 13, no. 10, Oct. 2018, Art. no. e0203794, DOI: 10.1371/journal.pone.0203794.

[3] J. Xin, W. Yuhui, and T. Xiaoyang, ‘‘Pornographic image recognition via weighted multiple instance learning,’’ 2019, *arXiv:1902.03771*. [Online].Available: http://arxiv.org/abs/1902.03771

[4] R. Slonje and P. K. Smith, ‘‘Cyberbullying: Another main type of bullying?’’ *Scandin. J. Psychol.*, vol. 49, no. 2, pp. 147–154, Apr. 2008.

[5] J. Guglani and A. N. Mishra, ``DNN based continuous speech recognition system of Punjabi language on kaldi toolkit," *Int. J. Speech Technol.*, vol. 24, no. 1, pp. 41_45, Mar. 2021.

[6] Y. Liu, Y. Yang, H. Xie, and S. Tang, ``Fusing audio vocabulary with visual features for pornographic video detection," *Future Gener. Comput. Syst.*, vol. 31, pp. 69_76, Feb. 2014.

[7] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, ``You only look once: Uni_ed, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 779_788.

[8] J.Wehrmann, G. S. Simões, R. C. Barros, andV. F. Cavalcante, ``Adult content detection in videos with convolutional and recurrent neural networks," *Neurocomputing*, vol. 272, pp. 432_438, Jan. 2018.