WEB CRAWLER FOR CHILD SEXUAL ABUSE USING CONVUTIONAL NEURAL NETWORK

[1]Dr.Reshma Banu, [2]Harshitha N, [3]Harshitha V, [4]Tasmia Khan

[1]Assistant Professor   [2,3,4]Student

Department of Information Science and Engineering

[1,2,3,4]GSSS Institute of Engineering and Technology for Women, Karnataka, India.

Abstract: With the growing amount of pornography content over Internet and cases of Child Sex Abuse (CSA) possession and distribution, there is a rising demand for automatic detection of such content especially in certain environments such as educational or work places. The contribution of this project is we present a critical review of automatic pornography and CSA detection in images and videos. Second, we provide an empirical evaluation of five selected pornography detection approaches representing traditional skin detection based as well as more recent deep learning based methods. The evaluations are performed under common criteria using two publicly available pornographic databases. Finally, we assess these methods on a dataset of real-world CSA material provided by Spanish polices. This study observes that for pornography or CSA detection, the methods involving multiple features perform better than those using simple features like skin colour or single image descriptor. It is also found that deep learning based methods outperform all of the other methods and report current state-of-the-art.

Index Terms - – CSA, networks, image processing, web crawler, sever security.

I. INTRODUCTION

Sexual abuse toward children and adolescents is a stark reality worldwide. A common misperception about child sexual abuse (CSA) is that it is a rare event perpetrated against girls by male strangers in poor, inner-city areas. To the contrary, CSA is a much too common occurrence that results in harm to millions of children, boys and girls alike, in large and small communities, and across a range of cultures and socioeconomic backgrounds. These acts are perpetrated by many types of offenders, including men and women, strangers, trusted friends or family, and people of all sexual orientations, socioeconomic classes, and cultural backgrounds. And also they will be posting videos and pictures of these innocent children on social media in order to black mail their parents or for making money or for any other reason. Day by day these kinds of issues are reaching maximum to the worst. So in order to overcome this are developing an application.

Child Abuse!! It's the most common abuse that remains unnoticed. Be it online or through physical exploitation, the incidents of child abuse have increased exponentially. This year marks 16 years of the United Nations Convention on the Rights of Children (UNCRC), yet in the past two decades, the child exploitation is gaining momentum. The definition of child abuse is not confined to physical abuse, but it also includes online sexual abuse. Most often, it is the latter that remain unreported

In its 2018 report, the UN cited 18.4 million cases of suspected online abuse. The report also states that over the past decade, the Child Sexual Abuse Material (CSAM) has increased. An investigation by The Newyork Times revealed that big tech companies have reported exploitation of over 45 million images of children. Moreover, the Internet Watch Foundation found that 39% of CSAM online are the victims under 10 years of age and 43% displays the act of extreme sexual violence. The exploitation of child doesn't end here. Another report by the Internet Watch Foundation in the UK cites that online child abuse has surged by 50% during the COVID 19 lockdown. And despite the many strategies implemented to mitigate the incidents of child abuse, undoubtedly the international organizations have not reached its goals. .In order to prevent these kind of issues we are going to develop an application.

II. LITRTURE REVIEW

Before you begin to format your paper, first write and save the content as a separate text file. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper Do not number text heads—the template will do that for you.Finally, complete content and organizational editing before formatting. Please take note of the following items when proof reading spelling and grammar returns of the shares and estimated betas.[1] Pornography and child sexual abuse detection in image and video: A comparative evaluation. This paper describes certain aspects With the growing amount of pornography content over Internet and cases of Child Sex Abuse (CSA) material possession and distribution, there is a rising demand for automatic detection of such content especially in certain environments such as educational or work places. The contribution of this paper is three fold.[2] First, we present a critical review of automatic pornography and CSA detection in images and videos. Second, we provide an empirical evaluation of five selected pornography detection approaches representing traditional skin detection based as well as more recent deep learning based methods. The evaluations are performed under common criteria using two publicly available pornographic databases..[3]E Fidalgo describes to assess these methods on a dataset of real-world CSA material provided by Spanish Police Forces. This study observes that for pornography or CSA detection, the methods involving multiple features perform better than those using simple features like skin color or single image descriptor. It is also found that deep learning based methods outperform all of the other methods and report to the current state-of-the-art[4] Automatic detection of child pornography using color visual words. This paper addresses the computer detection of child sexual abuse (CSA) images, a challenge of growing importance in multimedia forensics and security. In contrast to previous solutions based on hash sums, file names, or the retrieval of visually similar images, we introduce a system which employs visual recognition techniques to automatically identify suspect material [5] Our approach is based on color-enhanced visual word features and a statistical classification using SVMs. The detector is adapted to CSA material in a training step. In collaboration with police partners, we have conducted a quantitative evaluation on several datasets (including real-world CSA material). Our results indicate that recognizing child pornography is a challenging problem (more difficult than the detection of regular porn). Yet, while skin detection - a popular approach in pornography detection - fails, our approach can achieve a prioritization of content (equal error $11-24\%$) to improve the efficiency of forensic investigations of child sexual abuse. Examples illustrate that the system employs color cues as key features for discriminating CSA content.
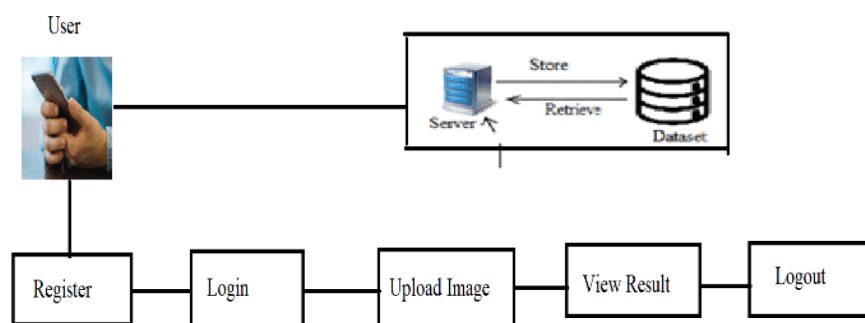
II METHODOLOGY

Fig1:system architecture

As depicted in the fig1, In order to overcome the problems of the existing system we are introducing a methodology where the user first needs to register to our application before he is using it, once after registering then he will be logging and he will be uploading the images and once after uploading our application will detect whether the image conatins any contents related to child abuse if so

then he will not be allowed to upload the image. So once after uploading the user will view the result of upload like whether it is approved or not then he will be logging out of the application.

Proposed System

The proposed system uses AI which is a intrusion detection image that sexually abused and remove that picture from the page. Developing an application by using artificial intelligence in order to overcome the problem facing in current days. When a person uploads any videos on social media our application will check whether the video is healthy, suppose if not that means if the video contains any such kind of stuffs then it will be recognized by this application and it will be deleting right then . So that the person will not be able to upload these kinds of videos. These will be saving lot many innocent children.

System Modules:

- User: The user is the one who will register to the app and user is the one who needs to upload the images on the social media platform .
- Register: The user module is the one where he must register to the app and then he can be able to upload the image if he is not registered then cannot be able to upload the image.
- login: Once after registering he needs to login to our application again for the safety purpose he has to login with his user id and the password which he had been created while registering to the app.
- Post image: user will be posting some image
- View result: Once after posting our application will detect whether the Image contains anything's related to child abuse, if yes our application will not allow that person to post it. It will display a message that it cannot be uploaded.
- Log out: Finally user wills logout of our application.

CONCLUSION AND FUTURE SCOPE

The project helps to a review some well-known approaches for pornography and CSA detection along with presenting their categorization. The review indicates that as time and technology progressed, more robust and effective computer vision algorithms have been explored for offensive content detection to apprehend offenders. Further, through the review and empirical study conducted by us, it is found that the deep learning based methods report state-of-the-art accuracies and significantly outperforms all other methods based on skin detection models or image descriptor based pipelines for pornography and CSA detection both. The results obtained are quite impressive and for porn image and video classification both, the best average accuracies reported are around 96.5%. For CSA detection the best accuracy is around 87.56% without performing any fine-tuning.

REFERENCES

[1]M. Short, L. Black, A. Smith, C. Wetterneck, D. Wells, A review of internet pornography use research: methodology and content from the past 10 years, Cyber. Behav. Soc. Netw. 15 (1) (2012) 13–23.

[2] M. Fleck, D. Forsyth, C. Bregler, Finding naked people, in: Proceedings of the European Conference on Computer Vision (ECCV), Vol. 1065, 1996

[3] R. Ap-apid. An Algorithm for Nudity Detection. In Proceedings of the 5th Philippine Computing Science Congress (PCSC), 2005.

[4] H. Rowley, Y. Jing, S. Baluja, Large scale image-based adult-content filtering, in: VISAPP, 2006, pp. 290–296.

[5] H. Zuo, W. Hu, O. Wu, Patch-based skin color detection and its application to pornography image filtering, in: Proceedings of the International Conference on World Wide Web (WWW), 2010, pp. 1227–1228.

[6] M. Basilio, Explicit image detection using YCbCr space color model as skin detection, CEA 5th Proc. American Int. Conf. Applied Mathematics: on Computer Engineering and Applications (ACM , US , 2011). [7] T. Deselaers, L. Pimenidis, H. Ney, Bag-of-visual-words models for adult image classification and filtering, in: ICPR, 2008, pp. 1–4.

[8] A. Lopes, S. Avila, A. Peixoto, R. S. Oliveira, M. Coelho, A. Araújo, A bag-of-features approach based on HueSIFT descriptor for nude detection, in: Proceedings of EUSIPCO, 2009, pp. 1552–1556.

[9] S. Avila, N. Thome, M. Cord, E. Valle, A. Araújo, BOSSA: extended BoW formalism for image classification, in: 18th IEEE International Conference on Image Processing (ICIP), 2909-2912, 2011.

[10] C.M. Steel, The Mask-SIFT cascading classifier for pornography detection, in: Proceedings of the World Congress on Internet Security (WorldCIS), 2012DOI:10.14722/ndss.2016.23047 Conference: Network and Distributed System Security Symposium.

[11] Avila, S., Thome, N., Cord, M., Valle, E., and Araujo, A., Pooling in image representation: The visual codeword point of view. Computer Vision and Image Understanding, 117(5), pp. 453 – 465 (2013).

[12] C. C. Yan, Y. Liu, H. Xie, Z. Liao, J. Yin, Extracting salient region for pornographic image detection, Journal of Visual Comm. &Image Representation (JVCIR) 2014

[13] C. Caetano, S. Avila, S. Guimaraes, and A. de A. Araujo, "Representing local binary descriptors with Bossa Nova for visual recognition," in ACM SAC, 2014, pp. 49–54.

[14] Geng Zhen, Zhuo Li, Zhang Jing, Li Xiaoguang, "A comparative study of local feature extraction algorithms for web pornographic image recognition", The 2015 International Conference on Progress in Informatics and Computing (PIC-2015), December 18-20, 2015.

[15] C. Caetano, S. Avila, S. Guimarães, A. Araújo, Pornography detection using bossanova video descriptor, in: Proceedings of EUSIPCO, 2014.