May-June 2020
ISSN: 0193-4120 Page No. 9698 – 9701

Children Inappropriate Content Detection Using Mask R- Cnn

Nithesh M- [1.] UG, Student, Sri Krishna College of Engineering and Technology

Naveen R [- 2.] UG, Student, Sri Krishna College of Engineering and Technology

Naveen V [- 3.] UG,Student ,Sri Krishna College of Engineering and Technology

Article Info
Volume 83
Page Number: 9698 – 9701
Publication Issue:
May - June 2020

Article History
Article Received*: 19 November 2019*
Revised*: 27 January 2020*
Accepted*: 24 February 2020*
Publication*: 18 May 2020*

Abstract:
The amount of inappropriate content on the Internet grows daily. Much of such content is unconstrained and freely-available for all users, requiring parents to adapt to parental control strategies for protecting their children. This is an attempt to mask or hide the data which are not applicable for the children. There are two major classes of data to be masked which have completely different features from one another (i) Adult content can be efficiently flagged as offensive based on the tradition and interests of the community who views it. The social-media applications such as tiktok and instagram allow users to post such content which is neither tagged as offensive content nor fall under age restriction which acts directly or indirectly as an threat to the adolescent aged community.Based on the above scenario it is evident and clear that we are compromising our culture and tradition and are moving towards the western culture. Certain content can be most appropriate to the western people and culture however it is not apt for the Indian culture.It could be more precisely said that the filter criteria algorithms do not understand the highly prideful tradition of the Indian and hence don't filter efficiently based on our ideologies and practices.(ii) horror content Some movies like "SAW" contain some violent scenes which display blood and organs to the audience which are not appropriate for the children because on a high level such movies inculcates violent behaviour on the children and it could make the audience faint on sight of blood due to a medical condition called Vasovagal syncope. This as a combination rises and stands as a new and unsolved problem statement.This project detects the space in a frame which has inappropriate content for children and masks it by using a M R-CNN model.

Keywords*: Image Processing, Mask R-CNN, Convolutional Neural Networks, Deep Learning, Machine Learning.*

I. INTRODUCTION

The aim is to restrict the easy access of underage youngsters to potentially harmful material. children are exposed to a lot of sources such as social media ,torrent websites and other no reliable sources where they have access to content that can disturb their mentality and well-being. We have categorised the offensive content into two major categories:-(i) body horror and (ii) body exposure. The first category is a direct and known problem since that is understood and accepted as inappropriate content for children.

When it comes to building a product one thing has to be kept in mind i.e the beliefs , traditions and cultures changes from region to region, in fact it changes from religion to religion in some region like India so, the product cannot sometime be acceptable in all the regions. This is the exact scenario that is taking place in our Region right now. The problem is that the threshold that is kept to classify the content as inappropriate to children are not appropriate for our region and also found that filtering based on this criteria gives results that contradicts the culture.

May-June 2020
ISSN: 0193-4120 Page No. 9698 – 9701

Thus we see this problem as a subset of the nudity content detection as a computer-vision problem. Though there are a lot of advancement in computer vision and especially in detecting complete body exposed content, this cannot detect the partial nudity content because the features completely differ in both the problems .though it logically locks like a subset of nudity detection it is not .Hence, it was required to create a dataset with the features relating to the problem . One further step in image processing was using CNN for the problem. We have used one of the advanced algorithms in CNN which is called "MASK-RCNN" which uses multiple neural networks for different tasks required to complete this process. We have used a MASK R-CNN model with RESNET 101 backbone from "matterport" and the weights from training coco dataset to perform transfer learning on the dataset that we created.

II. OBJECTIVE

- TO OBSTRUCT THE DISPLAY OF HORROR CONTENT.
- TO OBSTRUCT THE DISPLAY OF BODY EXPOSED CONTENT TO THE CHILDREN.
- TO SUPPORT THE BETTERMENT OF AUTOMATIC PARENT CONTROL.

III. TECHNOLOGY

- Python
- NVIDIA Tesla P100 (Google Collab)
- Tensorflow
- Keras

IV. DESIGN OF PROPOSED SYSTEM

A. *Existing System*

Current parental control devices depend on human intervention, and hence there is the need of computational approaches for automatically detecting and blocking the content that is in any way inappropriate for the children. Where traditional methods of site blocking based on black listing the websites are not applicable since there are flaws in the filtering criteria itself and thus methodologies analyzing the content of the files present in the websites have been proposed. A number of solutions exist for "NOT SAFE FOR CHILDREN" images detection and fewer for video. The most typical means of identifying pornographic material are strongly based on skin detection but as said earlier these methods didn't work properly for specific conditions on adult content, however body horror detection was a direct problem.

B. Proposed System Design

The algorithm used for content detection is mask RCNN, which is built on Faster RCNN. The function of Mask RCNN is it outputs a masked layer over the detected object by the number of classes provided. The method of detecting the outline of the object by pixels is Instance Segmentation. The process is done for two groups of detection, the detection of body parts and detection of horror scenes. Body parts include four classes, upper, middle, lower and back. It is a dual step process where the images are scanned at first and the bounding boxes masks are created at the time. There are various levels of MRCNN to provide masking. The Backbone levels gather the higher traits of the images. The next level includes RPN that holds the image objects. It glances over the image and generates traits optimally. It slides over the image at a minimal time frame. RPN gathers the foreground and background data and hence it provides ROI. It is discarded when the background class is generated. Based on the ROI align methods, a map is sampled and given for bilinear interpolation. The mask network is formed at the last level. It is an add on to fast RCNN. It holds the positive areas of ROI. The final masks are scaled to values that improve accuracy. The created masks are given by the pixels of the image. The results are improved by using a huge dataset with different aspects of the subject. By training the module, the accuracy rate can be improved. At a high level, the proposed system consists of the following four modules:

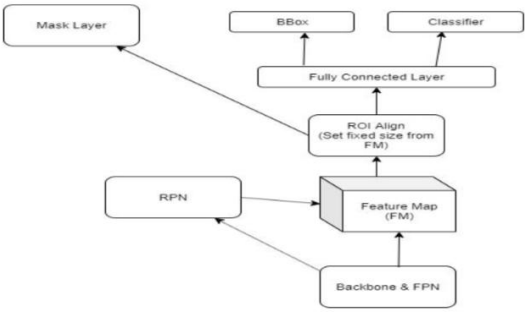
May-June 2020
ISSN: 0193-4120 Page No. 9698 – 9701

Backbone - level 1:
The first level in this system collects the required traits from the images. The images are given as tensors. The traits of high level are extracted by the inner layers. The loaded image gets converted into tensors The FPN augments to traits to detect the object size. Here the traits are augmented and are improvised by standard Convnet.
Region Proposal Network (RPN) Structure - level 2: The second layer includes the RPN structure. It acts like a scanner that slides over the image in a very short time span and identifies the regions that are holding the objects. The anchors present in the image act like boxes enveloping the image. The anchors depend on the image size. The bigger the image more anchors are generated. Hence it takes time for the sliding process. There are thousands of anchors generated for the region. The output generated from sliding through the images are of two types, the foreground and background. To fit the anchors within the regions, they are refined and duplicate anchors are removed, so that the object is mapped properly. The change in dimension is calculated by delta.
ROI Classifier & Bounding Box Regressor - level 3: The ROIs are produced in this layer. The foreground and background data are generated from this. Once the ROI is produced, the BG is discarded. Bounding box refinement is carried to tune the boxes that are holding the objects. The ROI s are aligned according to the object mapped. The ROI pooling process is carried out to remove different types of bounding boxes produced. The dimensions of the image used are standardized using the ROI pool. The feature map is taken for sampling the points that map the objects. The principle is like cropping the image and resizing it. Here the sampling process is carried out. The tensor is truncated for the sampling.
Segmentation Masks - level 4 :
The Mask R-CNN utilizes an add-on to the Faster RCNN, called the mask network. The mask network is used as by the Mask RCNN that is better than the faster RCNN. The mask branch consists of a convolutional network. The positive areas from the box are taken by the classifier. The masks are produced based on the positive areas picked by the ROI classifier. They contain more details than binary masks by using float and called as soft masks. During inferring, the final masks are given by scaling up the predicted masks to the dimension of the ROI bounding box. By training the boxes, the final masks are used to render the regions.

V. APPLICATION.

- Automatic censoring of adult content on pop up ads
- Can be used in safe browsing
- Can be used in social-networking sites to ensure that the content is appropriate

A. *Figures*



A.

Mask R CNN Model

VI. IMPLEMENTATION OF PROPOSED SYSTEM

A. *Our Approach*

We have adopted a Mask R-CNN model developed by "matterport" for instance segmentation. new code is written that initiates, utilizes the model and applies the splash of colours which is used to mask the inappropriate content on the image. A new dataset is created which holds train-test split in the ratio 3:1.The dataset has images of body horror and body exposure. the images of body exposure are selected based on two main criteria:-
(i)if the image shows four parts of the body such as the stomach region, thighs, back and the breast

May-June 2020
ISSN: 0193-4120 Page No. 9698 – 9701

(ii)if the above mentioned parts of the body are visible in such a manner that it contradicts the day to day clothing habits and culture of the people in our region.

these images are labelled into five main categories as below:-

This implementation mainly follows the Mask RCNN paper and has the four main modules as said in the proposed system. At first the dataset is loaded to the model and then the model is trained. the knowledge gained in the training process can be saved as weights. To make detections on an image, the image and the saved weights is fed to the same model .The image is resized to 1024x1024 pixels to increase the batch size so that more images can be processed by the model. The resized image is then fed to the region proposal network (RPN) which runs a lightweight binary classifier on a lot of boxes (anchors) over the image and returns object/no-object scores. then the proposals from the RPN is given to the classifier heads to make classifications and generate bounding boxes. The detections (refined bounding boxes and class IDs) from the bounding box regressors are given to the mask head and it generates segmentation masks for every instance. Thus we get the areas where inappropriate content are present on the image and those pixels in the image are masked. Each image has N sampled ROIs and a mini-batch has 2 images per GPU. Wetrainon 8 GPUs (so effective mini batch size is 16) for 30 epochs, with a learning rate of 0.02. We use a weight decay of 0.0001 and momentum of 0.9.the loss was 0.2110 at the end of 30th epoch.

VII.**CONCLUSION**

From the implementation process we conclude that the contents found inappropriate for children both in terms of body exposure and body horror are detected and the regions of interest are masked using the MASK R-CNN algorithm and are applied a splash of colours for presentation. The CID dataset is created to cover up the adult and horror explicit images. The algorithm works efficiently providing about 60-70% accuracy in masking the contents. We can either use this model with minimal changes for simply detecting images with inappropriate content and make decisions based on the classification or else this project can directly be used to mask the content on social media such as tic-toc or instagram which is the main objective of the project.

References

1. J. Cruz, M. Dimaala, L. Francisco, E. Franco, A. Bandala, E. Dadios, “Object recognition and detection by shape and color pattern recognition using ANN”, 2013 IEEE 2013 International Conference of Information and Communication Technology, 2013 IEEE.
2. A. Caglayan, A. Can, “3D Convolutional Object Recognition using Volumetric Representations of Depth Data”, 2017 Fifteenth IAPR International Conference on Machine Vision Applications, MVA.
3. “Mask R -CNN” by Kaiming He , Georgia Gkioxari, Piotr Doll´ar ,Ross Girshick Facebook AI Research (FAIR)
4. C. Lee, K. Won oh, H. Kim, “Comparison of faster R-CNN models for object detection”, 2016 16th International Conference on Control, Automation and Systems, 16–19, 2016 in HICO
5. A. Nguyen, D. Kanoulas, G. Caldwell, and N. Tsagarakis, “Detecting Object Affordances with Convolutional Neural Networks”, 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), October 9-14, 2016
6. F. Pourghahestani, E. Rashedi, “Object detection in images using artificial neural network and improved binary gravitational search algorithm”, 2015 4th IEEE CFIS.
7. X. Zhou, W. Gong, W. Fu, F. Du, “Application of deep learning in object detection”, 2017 IEEE ICIS, May 24-26,2017.