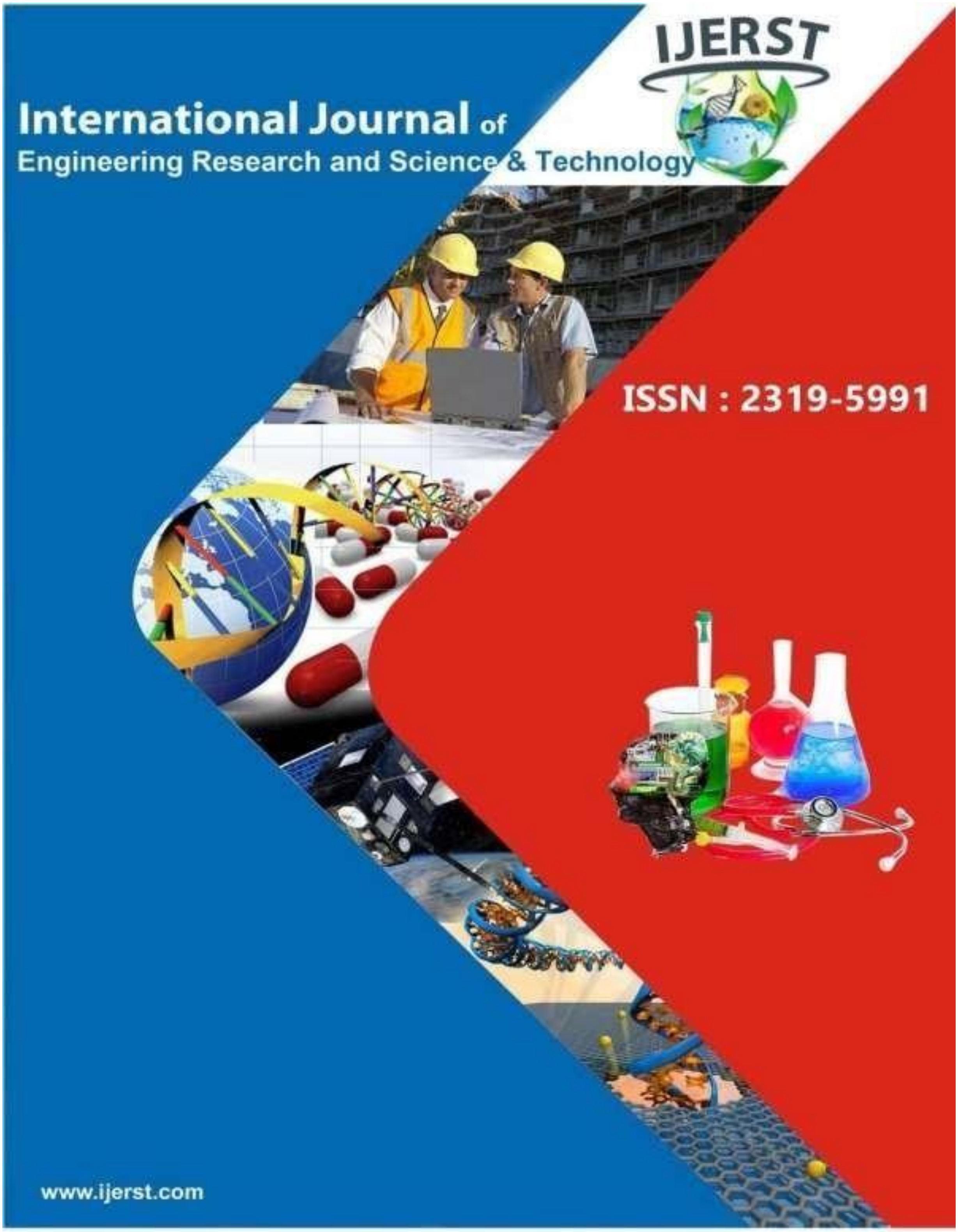


IJERST

International Journal of Engineering Research and Science & Technology

ISSN : 2319-5991

www.ijerst.com

Email: editor@ijerst.com or editor.ijerst@gmail.com

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

DETECTION OF DEEP FAKE VIDEOS USING LONG DISTANCE ATTENTION

[1] Mr.K.LAKSHMANA REDDY, [2] PULIDINDI HARITHA

[1](Associate Professor), Dept of MCA, S.V.K.P & Dr K.S. Raju Arts & Science College(A) Penugonda, W.G.District, Andhra Pradesh, klreddyin@gmail.com
[2]PG scholar, S.V.K.P & Dr K.S.Raju Arts & Science College(A) Penugonda, W.G.District,Andhra Pradesh, pulidindiharitha9@gmail.com

ABSTRACT

The exponential growth of videos on YouTube has attracted billions of viewers among which the majority belongs to a young demographic. Malicious uploaders also find this platform as an opportunity to spread upsetting visual content, such as using animated cartoon videos to share inappropriate content with children. Therefore, an automatic real-time video content filtering mechanism is highly suggested to be integrated into social media platforms. In this study, a novel deep learning-based architecture is proposed for the detection and classification of inappropriate content in videos. For this, the proposed framework employs an ImageNet pre-trained convolutional neural network (CNN) model known as EfficientNet-B7 to extract video descriptors, which are then fed to bidirectional long short-term memory (BiLSTM) network to learn effective video representations and perform multiclass video classification. An attention mechanism is also integrated after BiLSTM to apply attention probability distribution in the network. These models are evaluated on a manually annotated dataset of 111,156 cartoon clips collected from YouTube videos. Experimental results demonstrated that EfficientNet-BiLSTM (accuracy = 95.66%) performs better than attention mechanismbased EfficientNet-BiLSTM (accuracy = 95.30%) framework. Secondly, the traditional machine learning classifiers perform relatively poor than deep learning classifiers. Overall, the architecture of EfficientNet and BiLSTM with 128 hidden units yielded state-of-the-art performance (f1 score = 0.9267). Furthermore, the performance comparison against existing state-of-the-art approaches verified that BiLSTM on top of CNN captures better contextual information

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

of video descriptors in network architecture, and hence achieved better results in child inappropriate video content detection and classification.

INTRODUCTION

In an attempt to provide a safe online platform, laws like the children’s online privacy protection act (COPPA) imposes certain requirements on websites to adopt safety mechanisms for children under the age of 13. YouTube has also included a ‘‘safety mode’’ option to filter out unsafe content. Apart from that, YouTube developed the YouTube Kids application to allow parental control over videos that are approved as safe for a certain age group of children . Regardless of YouTube’s efforts in controlling the unsafe content phenomena, disturbing videos still appear even in YouTube Kids [20] due to difficulty in identifying such content. An explanation for this may be that the rate at which videos are uploaded every minute makes YouTube vulto unwanted content. Besides, the decision-making algorithms of YouTube rely heavily on the metadata of video (i.e., video title, video description, view count, rating, tags, comments, and community flags). Hence, filtering videos based on the metadata and community flagging is not sufficient to assure the safety of children. Many cases exist on YouTube where safe video titles and thumbnails are used for disturbing content to trick children and their parents. The sparse inclusion of child inappropriate content in videos is another common technique followed by malicious uploaders displays an example among such cases where video title and video clips are safe for children (as shown in but included inappropriate scenes in this video. The concerning thing about this example, including many similar cases, is that these videos have millions of views with more likes than dislikes, and have been available for years. Many other cases (as shown in Fig. 1(d)) also identified where videos or the YouTube channel is not popular, yet contains child unsafe content especially in the form of animated cartoons. It is evident from examples that this problem persists irrespective of channel or video popularity. Furthermore, YouTube has disabled the dislike feature of videos which resulted in viewers being incapable of getting the indirect video content.

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

LITERATURE SURVEY

Content detection and classification of YouTube videos have garnered significant academic and industry interest due to the vast and diverse nature of the platform's content. The goal is to develop systems that can automatically analyze, categorize, and manage video content efficiently. This literature survey reviews key research areas, methodologies, challenges, and advancements in this field. Object Detection and Recognition: Techniques such as Convolutional Neural Networks (CNNs) and Region-based CNNs (R-CNN) have been employed to detect and recognize objects within video frames .Scene Understanding: Research has focused on identifying and classifying different scenes in videos, leveraging techniques like semantic segmentation and scene graph generation .Speech Recognition: Automatic Speech Recognition (ASR) systems convert spoken language into text, which can then be analyzed for content classification .Audio Event Detection: Methods to identify non-speech audio events (e.g., music, background noises) using techniques like spectrogram analysis and neural networks**.**Text Analysis: Extracting and analyzing textual content from video descriptions, comments, and transcriptions using NLP techniques such as sentiment analysis, topic modeling, and named entity recognition (NER) Content Summarization: Summarizing video content using NLP to generate concise descriptions and metadata .Multi-Modal Learning:Combining Visual and Audio Features: Integrating visual and audio data to improve classification accuracy using multi-modal neural networks .Cross-Modal Retrieval: Enhancing content retrieval by leveraging correlations between different data modalities .Supervised Learning: Training models on labeled datasets to classify content into predefined categories. Techniques include Support Vector Machines (SVM), Random Forests, and deep learning models like CNNs and Recurrent Neural Networks (RNNs) .Unsupervised Learning: Employing clustering algorithms (e.g., K-means, DBSCAN) to group similar content without prior labels .Transfer Learning: Utilizing pre-trained models (e.g., VGG, Re Net) and fine-tuning them for specific tasks .Visual Features: Extracting features from video frames using methods like Histogram of Oriented Gradients (HOG), Scale-Invariant Feature Transform (SIFT), and deep feature extractors from CNNs . AI-Powered Moderation: Leveraging AI models to automatically detect and flag content that violates platform policies .The field of content detection and classification for YouTube videos is rapidly evolving, driven by advancements in machine learning, deep learning, and multi-modal analysis. While significant progress has been made, challenges remain in scalability, accuracy, and ethical considerations. Ongoing research and innovation are essential to address these challenges and enhance the capabilities of automated content moderation systems.

EXISTING SYSTEM

In the past few years, the performance of general image classification tasks has been significantly improved. From the amazing start of Alex net in image net, the

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

method based on Deep learning almost dominate the Image net competition. However, for fine-grained object recognition there are still great challenges. The main reason is that the two objects are almost the same from the global and apparent point of visual. Therefore, how to recognize the subtle differences in some key parts is a central theme for fine-grained recognition.

Earlier works leverage human-annotated bounding box of key parts and achieve good results. But the disadvantage is that it needs expensive manual annotation, and the location of manual annotation is not always the best distinguishing area which completely depends on the cognitive level of the annotator Since the key step of fine-grained classification is focusing on more discriminative local areas [42], many weakly supervised learning methods [23], [40], [43] have been proposed. Most of them use kinds of convolutional attention mechanisms to find the pivotal parts for detection. Fu et al. [43] use a recurrent attention convolutional neural network (RA-CNN) to learn discriminative region attention. Hu et al. [44] propose a channel-wise attention method to model interdependencies between channels. In [40], a multi-attention convolutional neural network is adopted and more fine-grained features can be learned. Huetal [23] propose a weakly supervised data augmentation network using attention cropping and attention dropping.

Deep fake detection and fine-grained classification are similar, that attempt to classify very similar things. Thus we learn from the experience in this field and leverage the attention maps generated with long range information to make the networks focus on pivotal

PROPOSED SYSTEM

The experience of the fine-grained classification field is introduced, and a novel long distance attention mechanism is proposed which can generate guidance by assembling global information.

- It confirms that the attention mechanism with a longer attention span is more effective for assembling global information and highlighting local regions. And in the process of generating attention maps, the non-convolution module is also feasible.
- A spatial-temporal model is proposed to capture the defects in the spatial domain and time domain, according to the characteristics of deep fake videos, the model adopts the long distance attention as the main mechanism to construct a multi-level semantic

guidance. The experimental results show that it achieves the state-of-the-art performance.

Disadvantages

- The spatial attention model is not designed to capture the artifacts that existed in the spatial domain with a single frame.
- The system not implemented Effectiveness of spatial-temporal model which leads the system less effective.

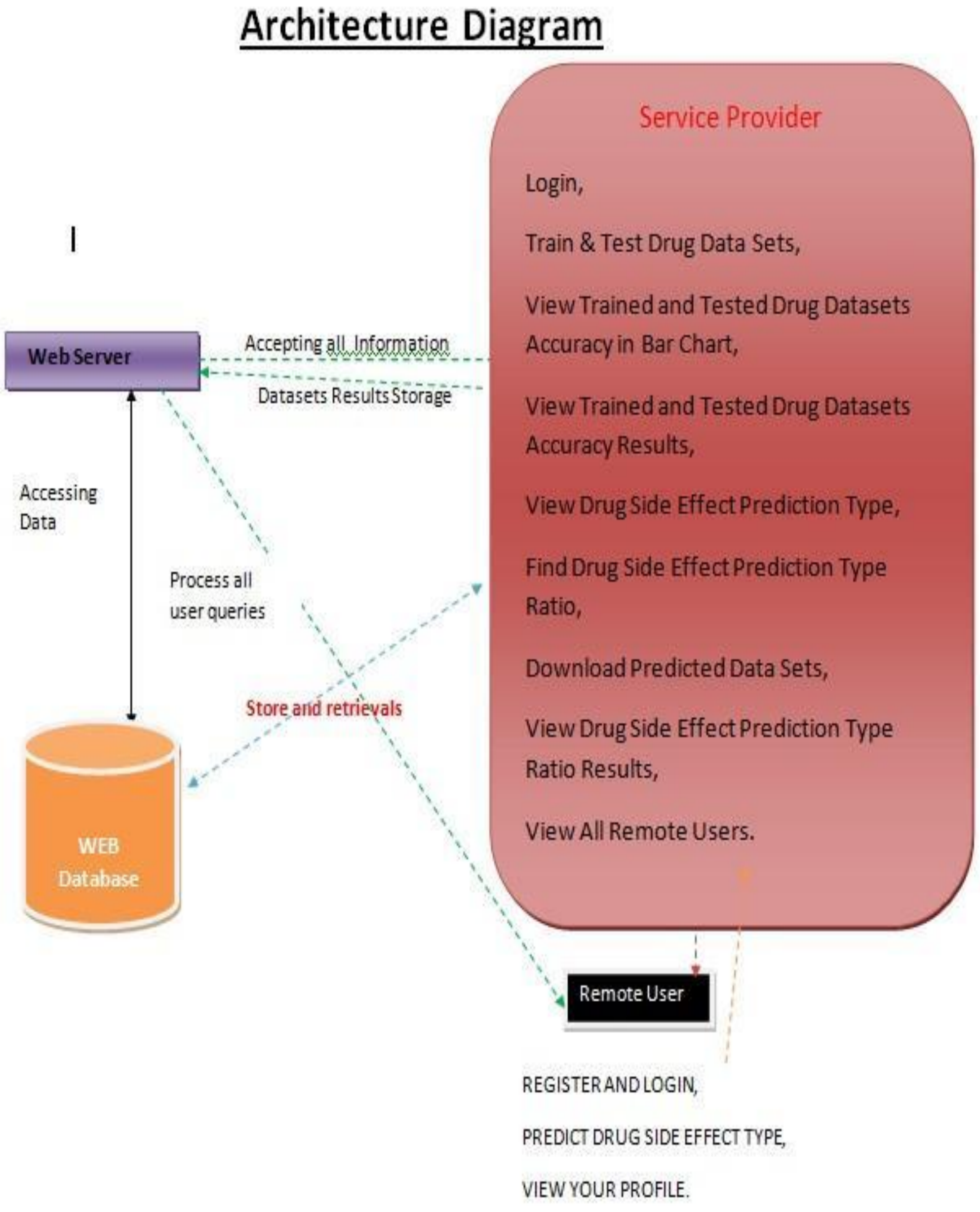
Advantages

- In the proposed system, the motivation to use long distance attention is given first and then the proposed model is described briefly. As aforementioned, there is no precise global constraint in the deepfake generation model, which always introduces disharmony between local regions in the face forgery from a global perspective.

- In addition to the artifacts that exist in each forgery frame itself, there are also inconsistencies (e.g., unsmooth lip movement) between frame sequences because the deepfake videos are generated frame by frame. To capture these defects, a spatial-temporal model is proposed, which has two components for capturing spatial and temporal defects respectively. Each component has a novel long distance attention mechanism which can be used to assembling the global information to highlight local regions.

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

ARCHITECTURE


Architecture Diagram
Service Provider
Login,
Train & Test Drug Data Sets,
View Trained and Tested Drug Datasets Accuracy in Bar Chart,
View Trained and Tested Drug Datasets Accuracy Results,
View Drug Side Effect Prediction Type,
Find Drug Side Effect Prediction Type Ratio,
Download Predicted Data Sets,
View Drug Side Effect Prediction Type Ratio Results,
View All Remote Users.
Web Server
Accepting all Information
Datasets Results Storage
Accessing Data
Process all user queries
Store and retrievals
WEB Database
Remote User
REGISTER AND LOGIN,
PREDICT DRUG SIDE EFFECT TYPE,
VIEW YOUR PROFILE.

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

MODULES

SERVICE PROVIDER

In this module, the Service Provider has to login by using valid user name and password. After login successful he can do some operations such as Login, Train & Test YouTube Content Data Sets, View Trained and Tested Accuracy in Bar Chart , View Trained and Tested Accuracy Results ,View Prediction Of YouTube Content Data Sets, View YouTube Content Data Sets Ratio , Download Predicted Data Sets ,View YouTube Content Data Sets Ratio Results ,View All Remote Users.

VIEW AND AUTHORIZE USERS

In this module, the admin can view the list of users who all registered. In this, the admin can view the user’s details such as, user name, email, address and admin authorizes the users.

REMOTE USER

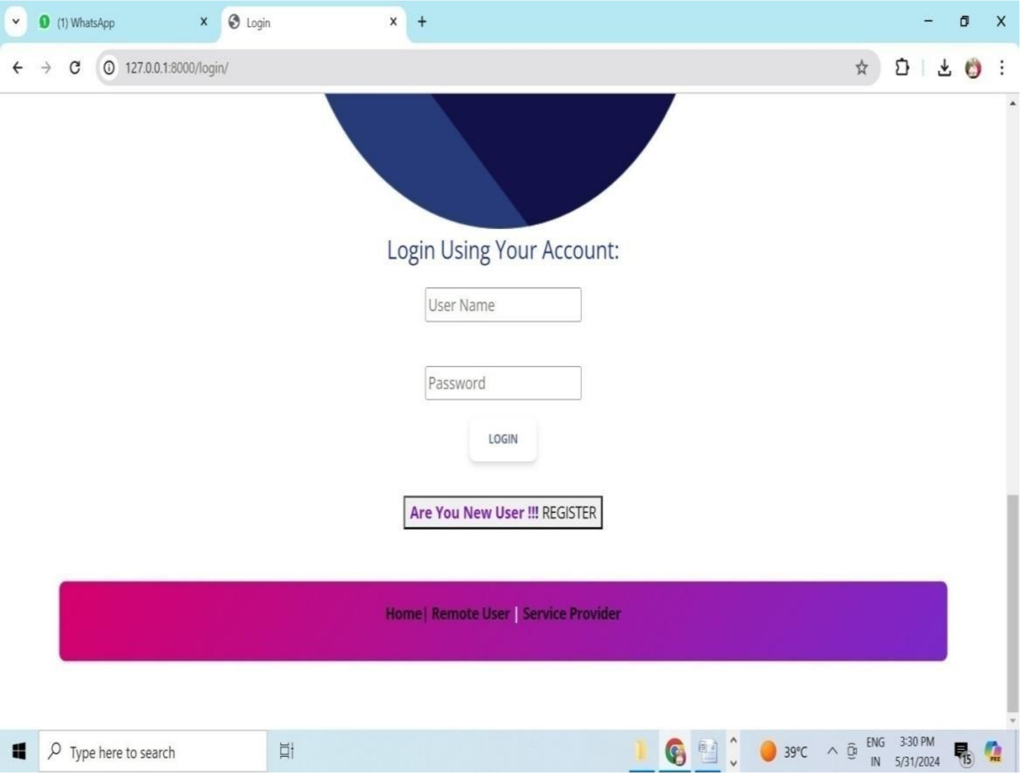
In this module, there are n numbers of users are present. User should register before doing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user will do some operations like REGISTER AND LOGIN, PREDICT YOUTTUBE CONTENT TYPE, VIEW YOUR PROFILE

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue 2, 2024

OUTPUT SCREENS

USER LOGIN PAGE



Login Using Your Account:

User Name

Password

LOGIN

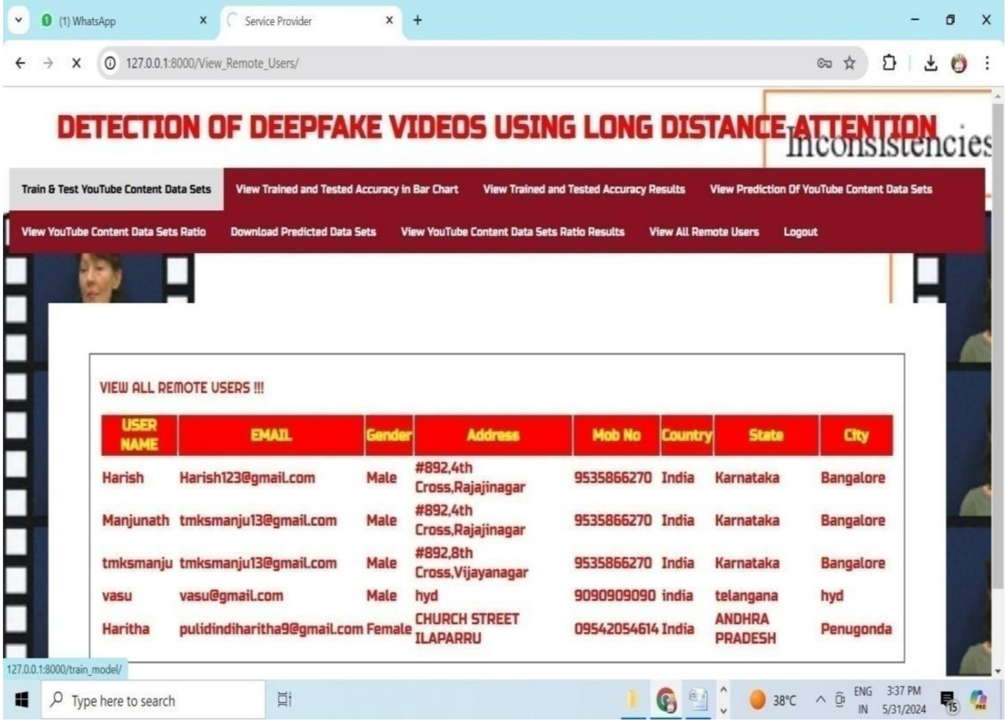
Are You New User !!! REGISTER

Home| Remote User | Service Provider

VIEW ALL REMOTE USERS

ISSN 2319-5991 www.ijerst.com

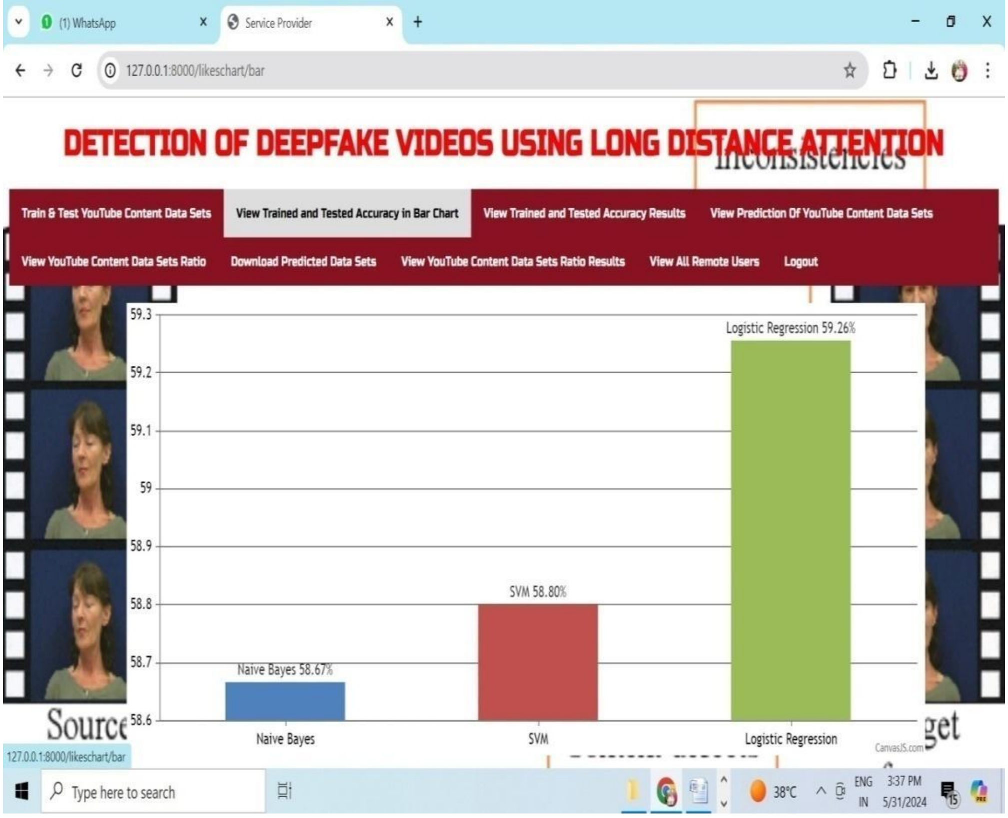
Vol. 17, Issue 2, 2024


(1) WhatsApp
Service Provider
127.0.0.1:8000/View_Remote_Users/
DETECTION OF DEEPFAKE VIDEOS USING LONG DISTANCE ATTENTION
Inconsistencies
Train & Test YouTube Content Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of YouTube Content Data Sets
View YouTube Content Data Sets Ratio
Download Predicted Data Sets
View YouTube Content Data Sets Ratio Results
View All Remote Users
Logout
127.0.0.1:8000/train_model/
Type here to search
38°C
ENG IN
3:37 PM 5/31/2024

VIEW ALL REMOTE USERS !!!

USER NAME	EMAIL	Gender	Address	Mob No	Country	State	City
Harish	Harish123@gmail.com	Male	#892,4th Cross,Rajajinagar	9535866270	India	Karnataka	Bangalore
Manjunath	tmksmanju13@gmail.com	Male	#892,4th Cross,Rajajinagar	9535866270	India	Karnataka	Bangalore
tmksmanju	tmksmanju13@gmail.com	Male	#892,8th Cross,Vijayanagar	9535866270	India	Karnataka	Bangalore
vasu	vasu@gmail.com	Male	hyd	9090909090	india	telangana	hyd
Haritha	pulidindiharitha9@gmail.com	Female	CHURCH STREET ILAPARRU	09542054614	India	ANDHRA PRADESH	Penugonda

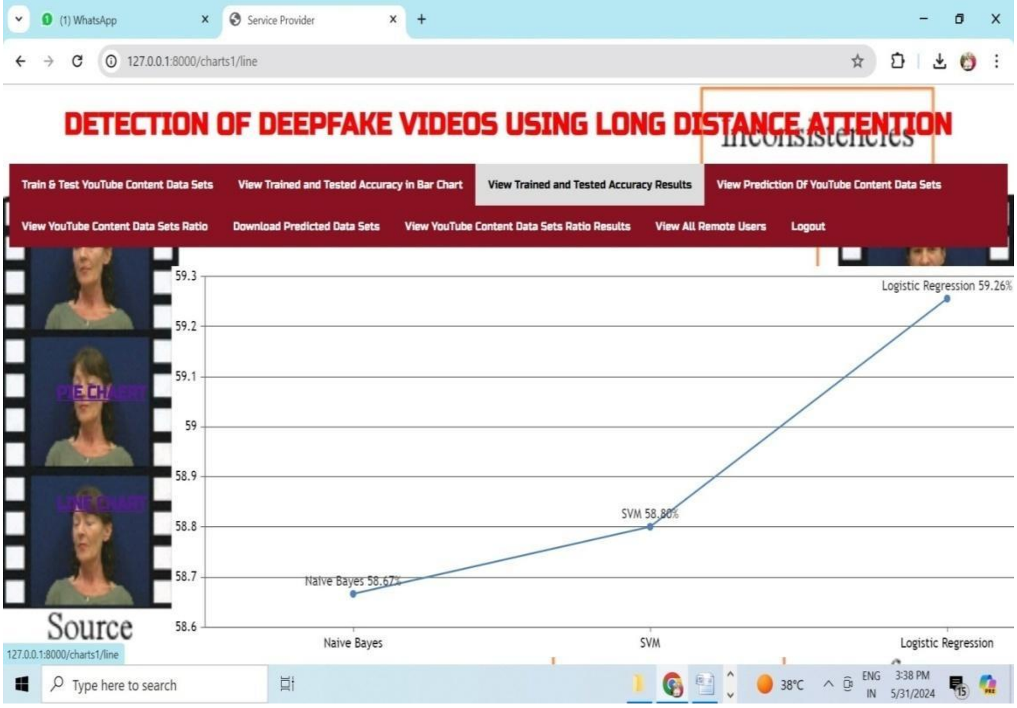
BAR CHART


(1) WhatsApp
Service Provider
127.0.0.1:8000/likeschart/bar
DETECTION OF DEEPFAKE VIDEOS USING LONG DISTANCE ATTENTION
Inconsistencies
Train & Test YouTube Content Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of YouTube Content Data Sets
View YouTube Content Data Sets Ratio
Download Predicted Data Sets
View YouTube Content Data Sets Ratio Results
View All Remote Users
Logout
59.3
59.2
59.1
59
58.9
58.8
58.7
58.6
Logistic Regression 59.26%
SVM 58.80%
Naive Bayes 58.67%
Naive Bayes
SVM
Logistic Regression
CanvasJS.com
Source
get
127.0.0.1:8000/likeschart/bar
Type here to search
38°C
ENG IN
3:37 PM 5/31/2024

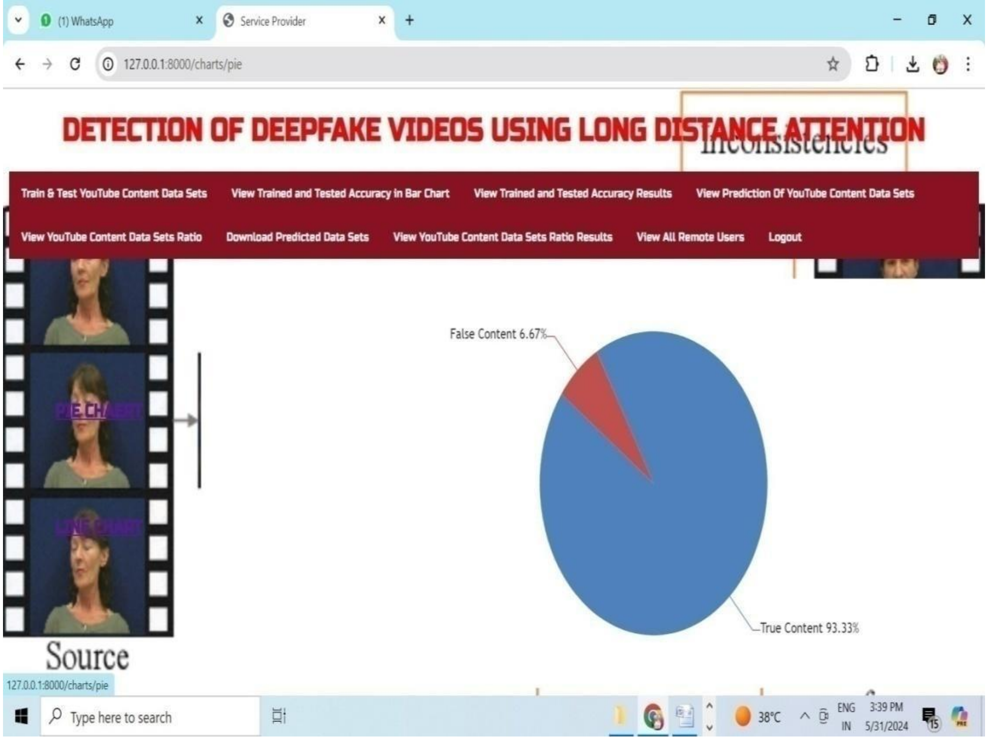
ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue 2, 2024

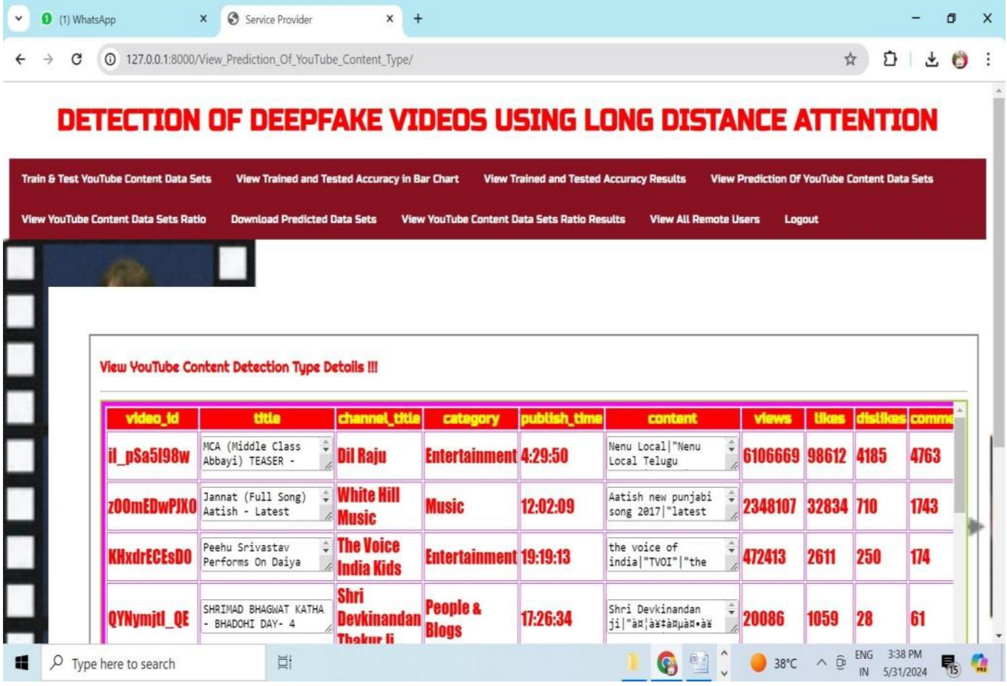
LINE CHART


(1) WhatsApp
Service Provider
127.0.0.1:8000/charts1/line
DETECTION OF DEEPFAKE VIDEOS USING LONG DISTANCE ATTENTION
Train & Test YouTube Content Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of YouTube Content Data Sets
View YouTube Content Data Sets Ratio
Download Predicted Data Sets
View YouTube Content Data Sets Ratio Results
View All Remote Users
Logout
59.3
59.2
59.1
59
58.9
58.8
58.7
58.6
Logistic Regression 59.26%
SVM 58.80%
Naive Bayes 58.67%
Naive Bayes
SVM
Logistic Regression
Source
Type here to search
38°C
ENG IN
3:38 PM 5/31/2024

PIE CHART


(1) WhatsApp
Service Provider
127.0.0.1:8000/charts/pie
DETECTION OF DEEPFAKE VIDEOS USING LONG DISTANCE ATTENTION
Train & Test YouTube Content Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of YouTube Content Data Sets
View YouTube Content Data Sets Ratio
Download Predicted Data Sets
View YouTube Content Data Sets Ratio Results
View All Remote Users
Logout
False Content 6.67%
True Content 93.33%
Source
Type here to search
38°C
ENG IN
3:39 PM 5/31/2024

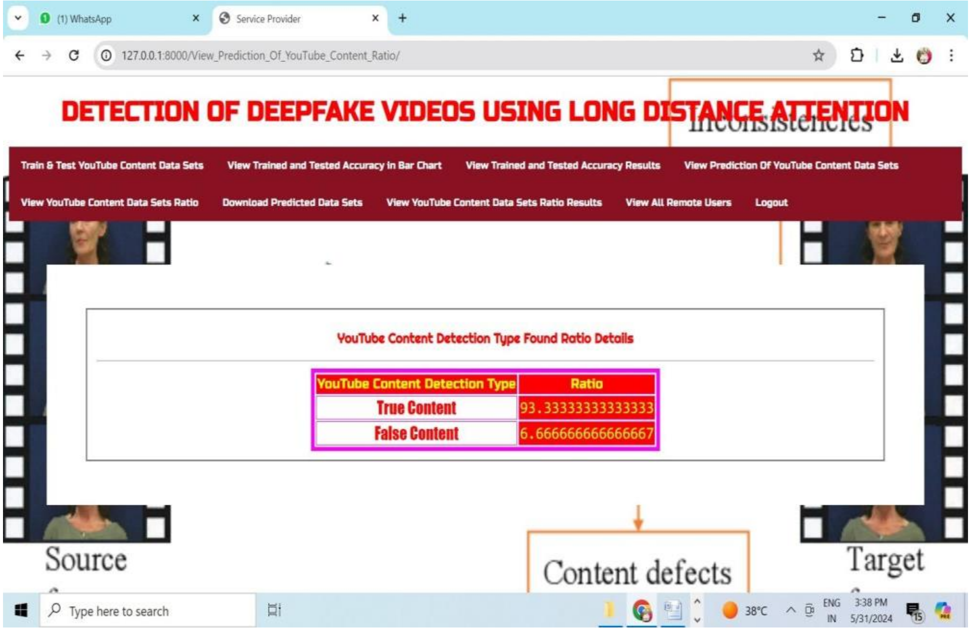
DETAILS


(1) WhatsApp
Service Provider
127.0.0.1:8000/View_Prediction_Of_YouTube_Content_Type/
DETECTION OF DEEPFAKE VIDEOS USING LONG DISTANCE ATTENTION
Train & Test YouTube Content Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of YouTube Content Data Sets
View YouTube Content Data Sets Ratio
Download Predicted Data Sets
View YouTube Content Data Sets Ratio Results
View All Remote Users
Logout
Type here to search
38°C
ENG
IN
3:38 PM
5/31/2024

View YouTube Content Detection Type Details !!!

video_id	title	channel_title	category	publish_time	content	views	likes	dislikes	comme
il_pSa5I98w	MCA (Middle Class Abbayi) TEASER -	Dil Raju	Entertainment	4:29:50	Nenu Local|"Nenu Local Telugu	6106669	98612	4185	4763
z00mEDwPJK0	Jannat (Full Song) Aatish - Latest	White Hill Music	Music	12:02:09	Aatish new punjabi song 2017|"latest	2348107	32834	710	1743
KHxdrECEsD0	Peehu Srivastav Performs On Daiya	The Voice India Kids	Entertainment	19:19:13	the voice of india|"TVOI"|"the	472413	2611	250	174
QYNymjtl_QE	SHRIMAD BHAGWAT KATHA - BHADOHI DAY- 4	Shri Devkinandan Thakur Ji	People & Blogs	17:26:34	Shri Devkinandan ji|"à¤¦à¥‡à¤µà¤•à¥	20086	1059	28	61

RATIO


(1) WhatsApp
Service Provider
127.0.0.1:8000/View_Prediction_Of_YouTube_Content_Ratio/
DETECTION OF DEEPFAKE VIDEOS USING LONG DISTANCE ATTENTION
inconsistencies
Train & Test YouTube Content Data Sets
View Trained and Tested Accuracy in Bar Chart
View Trained and Tested Accuracy Results
View Prediction Of YouTube Content Data Sets
View YouTube Content Data Sets Ratio
Download Predicted Data Sets
View YouTube Content Data Sets Ratio Results
View All Remote Users
Logout
Source
Content defects
Target
Type here to search
38°C
ENG
IN
3:38 PM
5/31/2024

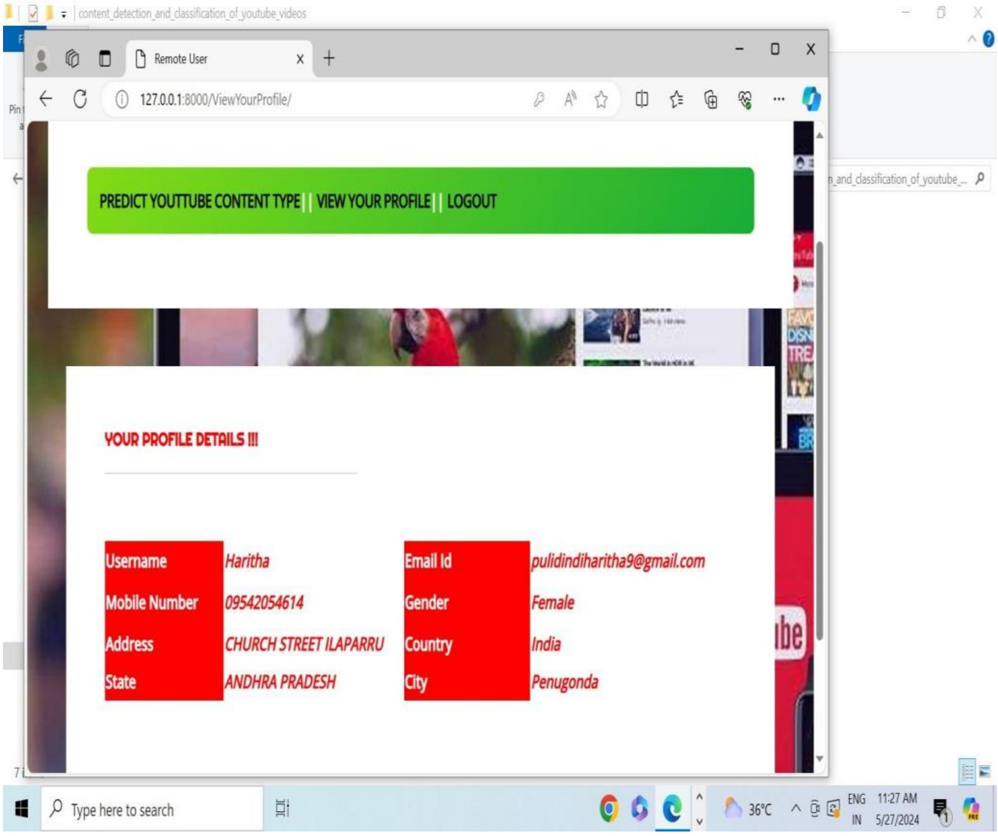
YouTube Content Detection Type Found Ratio Details

YouTube Content Detection Type	Ratio
True Content	93.33333333333333
False Content	6.666666666666667

ISSN 2319-5991 www.ijerst.com

Vol. 17, Issue 2, 2024

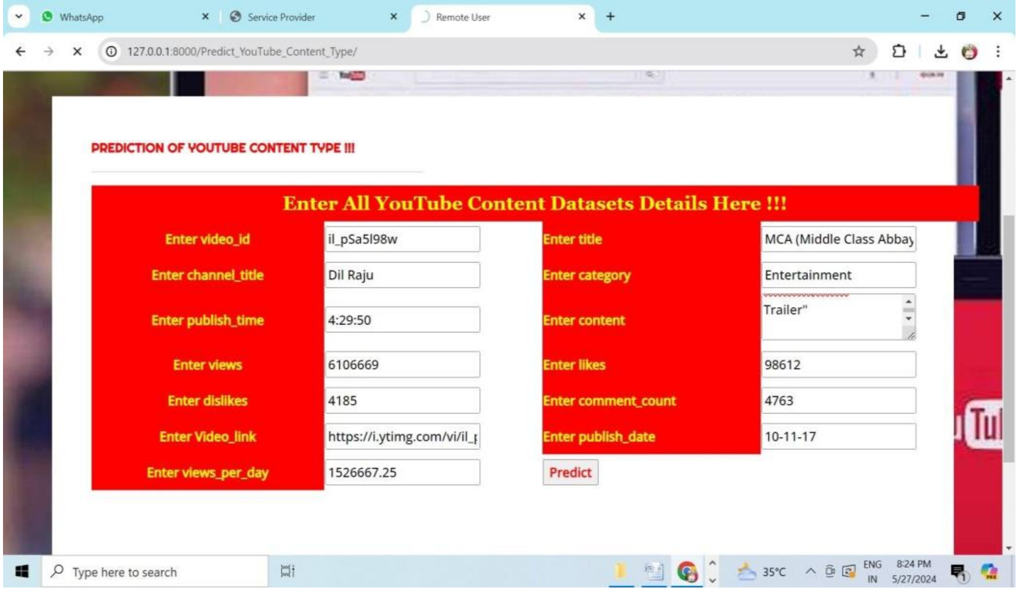
REMOTE USER


Remote User
127.0.0.1:8000/ViewYourProfile/
PREDICT YOUTTUBE CONTENT TYPE || VIEW YOUR PROFILE || LOGOUT
Type here to search
36°C
ENG IN
11:27 AM 5/27/2024

YOUR PROFILE DETAILS !!!

Username	Haritha	Email Id	pulidindiharitha9@gmail.com
Mobile Number	09542054614	Gender	Female
Address	CHURCH STREET ILAPARRU	Country	India
State	ANDHRA PRADESH	City	Penugonda

PREDICT CONTENT TYPE



PREDICTION OF YOUTUBE CONTENT TYPE !!!

Enter All YouTube Content Datasets Details Here !!!

Enter video_id	il_pSa5l98w	Enter title	MCA (Middle Class Abbay
Enter channel_title	Dil Raju	Enter category	Entertainment
Enter publish_time	4:29:50	Enter content	Trailer"
Enter views	6106669	Enter likes	98612
Enter dislikes	4185	Enter comment_count	4763
Enter Video_link	https://i.ytimg.com/vi/il_	Enter publish_date	10-11-17
Enter views_per_day	1526667.25		

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

CONCLUSION

In this paper, a novel deep learning-based framework is proposed for child inappropriate video content detection and classification. Transfer learning using EfficientNet-B7 architecture is employed to extract the features of videos. The extracted video features are processed through the BiLSTM network, where the model learns the effective video representations and performs multiclass video classification. All evaluation experiments are performed by using a manually annotated cartoon video dataset of 111,156 video clips collected from YouTube. The evaluation results indicated that proposed framework of Efficient Net-BiLSTM (with hidden units D 128) exhibits higher performance (accuracyD95.66%) than other experimented models including Efficient Net-FC, Efficient Net-SVM, Efficient Net-KNN, Efficient Net-Random Forest, and Efficient Net-BiLSTM with attention mechanism-based models (with hidden units D 64, 128, 256, and 512). Moreover, the performance comparison with existing state-of-the-art models also demonstrated that our BiLSTM-based framework surpassed other existing models and methods by achieving the highest recall score of 92.22%. The advantages of the proposed deep learning-based children inappropriate video content detection system are as follows:

1) It works by considering the real-time conditions by processing the video with a speed of 22 fps using Ef_cientNet-B7 and BiLSTM-based deep learning framework, which helps in filtering the live-captured videos.

2) It can assist any video sharing platform to either remove the video containing unsafe clips or blur/hide any portion with unsettling frames.

3) It may also help in the development of parental control solutions on the Internet through plugins or browser extensions where child unsafe content can be filtered automatically.

Furthermore, our methodology to detect inappropriate children content from YouTube is independent of YouTube video metadata which can easily be altered by malicious up loaders to deceive the audiences. In the future, we intend to combine the temporal stream using optical low frames with the spatial stream of the RGB frames to further improve the model performance by better understanding the global representations of videos. We also aim to increase the classification labels to target the different types of inappropriate children content of YouTube videos.

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

REFERENCES

[1] L. Ceci. *YouTube Usage Penetration in the United States 2020, by Age Group*. Accessed: Nov. 1, 2021. [Online]. Available: https://www.statista.com/statistics/296227/us-youtube-reach-age-gender/

[2] P. Covington, J. Adams, and E. Sargin, ``Deep neural networks for YouTube recommendations," in *Proc. 10th ACM Conf. Recommender Syst.*, Sep. 2016, pp. 191_198, doi: 10.1145/2959100.2959190.

[3] M. M. Neumann and C. Herodotou, ``Evaluating YouTube videos for young children," *Educ. Inf. Technol.*, vol. 25, no. 5, pp. 4459_4475, Sep. 2020, doi: 10.1007/s10639-020-10183-7.

[4] J. Marsh, L. Law, J. Lahmar, D. Yamada-Rice, B. Parry, and F. Scott, *Social Media, Television and Children*. Shef_eld, U.K.: Univ. Shef_eld, 2019. [Online]. Available: https://www.stac-study.org/downloads/ STAC_Full_Report.pdf

[5] L. Ceci. *YouTube_Statistics & Facts*. Accessed: Sep. 01, 2021. [Online]. Available: https://www.statista.com/topics/2019/youtube/

[6] M. M. Neumann and C. Herodotou, ``Young children and YouTube: A global phenomenon," *Childhood Educ.*, vol. 96, no. 4, pp. 72_77, Jul. 2020, doi: 10.1080/00094056.2020.1796459.

[7] S. Livingstone, L. Haddon, A. Görzig, and K. Ólafsson, *Risks and Safety on the Internet: The Perspective of European Children: Full Findings and*

ISSN 2319-5991 www.ijerst.com
Vol. 17, Issue 2, 2024

Policy Implications From the EU Kids Online Survey of 9-16 Year Olds and Their Parents in 25 Countries. London, U.K.: EU Kids Online, 2011. [Online]. Available: http://eprints.lse.ac.U.K./id/eprint/33731

[8] B. J. Bushman and L. R. Huesmann, ``Short-term and long-term effects of violent media on aggression in children and adults," *Arch. Pediatrics Adolescent Med.*, vol. 160, no. 4, pp. 348_352, 2006, doi: 10.1001/archpedi. 160.4.348.

[9] S. Maheshwari. (2017). *On YouTube Kids, Startling Videos Slip Past Filters*. The New York Times. [Online]. Available: https://www.nytimes.com/ 2017/11/04/business/media/youtube-kids-paw-patrol.html

[10] C. Hou, X. Wu, and G. Wang, ``End-to-end bloody video recognition by audio-visual feature fusion," in *Proc. Chin. Conf. Pattern Recognit. Comput. Vis. (PRCV)*, 2018, pp. 501_510, doi: 10.1007/978-3-030-03398-9_43.

[11] A. Ali and N. Senan, ``Violence video classi_cation performance using deep neural networks," in *Proc. Int. Conf. Soft Comput. Data Mining*, 2018, pp. 225_233, doi: 10.1007/978-3-319-72550-5_22.

[12] H.-E. Lee, T. Ermakova, V. Ververis, and B. Fabian, ``Detecting child sexual abuse material: A comprehensive survey," *Forensic Sci. Int., Digit. Invest.*, vol. 34, Sep. 2020, Art. no. 301022, doi: 10.1016/j.fsidi. 2020.301022.

[13] R. Brandom. (2017). *Inside Elsagate, The Conspiracy Fueled War on*

Creepy YouTube Kids Videos. [Online]. Available: https://www.theverge. com/2017/12/8/16751206/elsagate-youtube-kids-creepy-conspiracytheory

[14] Reddit. *What is ElsaGate?* Accessed: Dec. 14, 2020. [Online]. Available: https://www.reddit.com/r/ElsaGate/comments/6o6baf/

[15] B. Burroughs, ``YouTube kids: The app economy and mobile parenting,'' *Soc. media*C *Soc.*, vol. 3, May 2017, Art. no. 2056305117707189, doi: 10.1177/2056305117707189.

[16] H. Wilson, ``YouTube is unsafe for children: YouTube's safeguards and the current legal framework are inadequate to protect children from disturbing content,'' *Seattle J. Technol., Environ. Innov. Law*, vol. 10, no. 1, p. 8, 2020. [Online]. Available: https://digitalcommons. law.seattleu.edu/sjteil/vol10/iss1/8

[17] S. Alshamrani, A. Abusnaina, M. Abuhamad, D. Nyang, and D. Mohaisen, ``Hate, obscenity, and insults: Measuring the exposure of children to inappropriate comments in YouTube,'' in *Proc. Companion Proc. Web Conf.*, Apr. 2021, pp. 508_515, doi: 10.1145/3442442. 3452314.

[18] N. Elias and I. Sulkin, ``YouTube viewers in diapers: An exploration of factors associated with amount of toddlers' online viewing,'' *Cyberpsychol., J. Psychosoc. Res. Cyberspace*, vol. 11, no. 3, p. 2, Nov. 2017, doi: 10.5817/cp2017-3-2.

[19] D. Craig and S. Cunningham, ``Toy unboxing: Living in a (n unregulated)

material world," *Media Int. Aust.*, vol. 163, no. 1, pp. 77_86, May 2017, doi: 10.1177/1329878X17693700.

[20] K. Papadamou, A. Papasavva, S. Zannettou, J. Blackburn, N. Kourtellis, I. Leontiadis, G. Stringhini, and M. Sirivianos, ``Disturbed YouTube for kids: Characterizing and detecting inappropriate videos targeting young children," in *Proc. Int. AAAI Conf. Web Soc. Media*, 2020, pp. 522_533. [Online]. Available: https://ojs. aaai.org/index.php/ICWSM/article/view/7320/7174

[21] R. Kaushal, S. Saha, P. Bajaj, and P. Kumaraguru, ``KidsTube: Detection, characterization and analysis of child unsafe content & promoters on YouTube," in *Proc. 14th Annu. Conf. Privacy, Secur. Trust (PST)*, Dec. 2016, pp. 157_164, doi: 10.1109/pst.2016. 7906950.

[22] R. Tahir, F. Ahmed, H. Saeed, S. Ali, F. Zaffar, and C. Wilson, ``Bringing the kid back into YouTube kids: Detecting inappropriate content on video streaming platforms," in *Proc. IEEE/ACM Int. Conf. Adv. Soc. Netw. Anal. Mining*, Aug. 2019, pp. 464_469, doi: 10.1145/3341161. 3342913. [23] A. Ulges, C. Schulze, D. Borth, and A. Stahl, ``Pornography detection in video bene_ts (a lot) from a multi-modal approach," in *Proc. ACM Int. Workshop Audio Multimedia Methods Large-Scale Video Anal.*, 2012, pp. 21_26, doi: 10.1145/2390214.2390222.

[24] C. Caetano, S. Avila, S. Guimaraes, and A. D. A. Araújo, ``Pornography

detection using BossaNova video descriptor," in *Proc. 22nd Eur. Signal Process. Conf.*, 2014, pp. 1681_1685. [Online]. Available: https://ieeexplore.ieee.org/document/6952616

[25] L. Duan, G. Cui, W. Gao, and H. Zhang, ``Adult image detection method base-on skin color model and support vector machine," in *Proc. Asian Conf. Comput. Vis.*, 2002, pp. 797_800. [Online]. Available: http://aprs.dictaconference.org/accv2002/accv2002_proceedings/Duan797.pdf