International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 04 | Apr 2022 www.irjet.net p-ISSN: 2395-0072

Fake News Analyzer

Nitin Sharma

Student, Dept. Electronics & Telecommunications, K. J. Somaiya Institute Of Engineering & Technology, Sion, Mumbai, India

Ritik Sharma

Student, Dept. Electronics & Telecommunications, K. J. Somaiya Institute Of Engineering & Technology, Sion, Mumbai, India

Sharan Shetty

Student, Dept. Electronics & Telecommunications, K. J. Somaiya Institute Of Engineering & Technology, Sion, Mumbai, India

Prof. Harshwardhan Ahire

Assistant Professor, Dept. Electronics & Telecommunications, K. J. Somaiya Institute Of Engineering & Technology, Sion, Mumbai, India

---

Abstract - This paper presents a method to detect misinformation in news content and face tempering in videos. The system is broken down into three segments. 1. A simple UI's designed for the users to insert alleged text, paragraph, headline or statement to analyses and process it.2.Text□based misinformation detection; the system would run a text classifier once a sentence, statement, or paragraph has been inserted into the system in text format. Datasets like LIAR (with over 12.8k manually labeled short statements were collected) and different datasets from Kaggle were used for training the model. Sentimental analysis was executed after cleaning, lemmatizing, and splitting the text data. The text than is processed on the basis of probability of truth or false it propagates. 3. Deep fake detection; Deep-Fake and Face2Face were used to build a model that could automatically and efficiently detect forged faces in a video. It follows a deep learning approach to detect whether the given video is subjected to forgery or not. The whole system is vertical integration of analyzing different aspects that contribute to misinformation and fake data (in the form of images, videos, and texts). The system would contribute to fighting fake news.

1. Introduction

With the advancements in the field of technology and the increase in the number of people using those technologies, there tends to have certain repercussions which come along with all the pleasures of inventions. The major concern for this generation is to educate the citizens about the consumption of content. What we see, read, and listen to on the pool of the internet impacts the way we see things around us and also our perspectives. These very contents are designed to deliberately mislead masses about certain subjects and to induce a very specific view about them. These contents could be false, intentionally plotted, misreported, polarized, persuasive information, and citizen journalism with improper facts. Hence, a system/ application is needed that would identify and break down any alleged information on the basis of detailed analysis. This paper tries to present a case for an ML-based application that would use a simple UI for users to insert or upload their query text, paragraph, headline, statement or alleged video to process it for fake content analysis and detect its authenticity. This would deal with Deep fakes and misinformation in the text format. The system is divided into two different segments that would focus on two different mediums of spreading fake news i.e. Video, and text. Videos are subjected to deep fakes and wrongful captions which then are forwarded and reposted so many times that it becomes difficult to verify their true source. We worked on a deep fake and an object detection model that would help us to understand whether a video is face forged or not. And notify its true source. We have used a microscopic level of analysis that would help us in detecting face forgery which is hard to identify from

a naked human eye. The model is based on a well-performing network of image classification. Text based classification is divided into 6 different classes that would define if the given statement is Original (True), True (True), Mostly True (True), Half True (True), Barely True (False), False (False) and Pants Fire (False). Based on these classifications, the system concludes whether a given statement, paragraph or a headline is trustworthy or not.

1.1 Deep-Fake

The concept of deep-fake refers to images, audios or videos that are fakes, that is they depict events that never occurred. Unlike methods of manipulating media in the past like Photoshop, these deep fakes are created by deep neural networks to be nearly indistinguishable from their real counterparts the advances in the field of deep fakes are equal parts impressive and alarming. On the upside the fidelity with which we can alter media will certainly lead to some world class memes but in the wrong hands this technology can be used to spread misinformation and undermine public trust. Almost like a sci fi type of identity theft, where you can get anyone to say anything and its vice versa. This means that as we get better at generating deep fakes, we must also get better at identifying those deep fakes Deep Fakes can be generated by using auto encoders. The auto encoders work as follows, when the data is processed such as image data, data gets compressed by an encoder. The purpose of this compression is to suppress the effect of noise in the data and to reduce convolutional complexity; conversely the original image can be restored at least approximately by passing the compressed version of the image through a decoder. Now suppose we want to create a deep fake that blends Van Gogh's 'Starry night' and Da Vinci's 'Mona Lisa'. To do so when we train the auto encoders for different data sets. We allow the encoders to share weights while keeping the decoders separate, that way an image of the Mona Lisa can be compressed

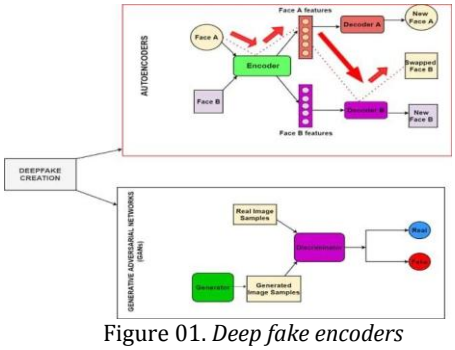

Figure 01. *Deep fake encoders*

according to a general logic. It considers things like the illumination position and the expression of her face but when it gets restored this will be done according to the logic specific to the starry night, which has the effect of overlaying Van Gogh's distinctive style onto Davinci's masterpiece.

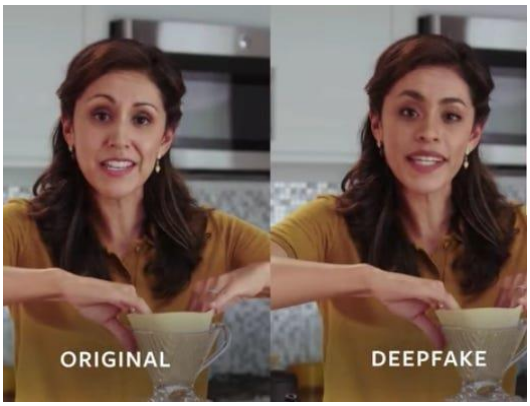

Figure 02. *Examples of Deep-Fake images*

1.2 News Headline Analysis

Text messages, news headlines from various micro blogging and news media outlets are forwarded through their URL links to multiple internet users via different social media platforms and networks such as WhatsApp, Facebook, Telegram, Instagram, etc. Most of the times these contents produced by several independent and unverified sources are responsible to spread lies and misinformation among the users. Their main aim is to write an eye-catching headline that could garner internet user's attention and better SEO operations give them better share ratio and hence giving them a viral status that could fire up the entire internet with false and maligned information. Most of the time many users read the headline which could be constructed in such a way that it creates controversial waves so that it helps the publishers to gain more visitors and sometimes many different political and independent institutions spread these very headliners through WhatsApp and several other chatting applications making it difficult to verify the source of the material. And hence it gets difficult on the users end to verify the legitimacy of the material. These headlines are then spread through user's tweets, texts, and chats which are in normal text format. And once this information is condensed into a simple text statement it gets complicated to determine the source and the construction of the sentences that was originally crafted

Figure 03. *Examples of Deep-Fake videos*

2. Proposed Method

We employed a web application to scan, take text input, and/or upload media (image/video) which then could be analyzed according to their respective analyzing systems. As the consumption of fake news is mainly through social media apps, the system treats two input systems in two separate ways; text classification and video manipulation detection.

We used existing machine learning models and trained it using our data collected from random videos and headlines found on the internet. We didn't use the existing deep fake or acclaimed fake news headlines for this data set. We built a custom pipeline for unorganized and real- world data to raise the uncertainty of the model and study the model's operations to different live examples.

For deep fake videos, we compiled a raw data from YouTube, Reddit, news network's archives and several other meme pages to build a real- life practical data set for the machine learning model to process. The random videos add to the uncertainty and put the model into test for real life examples. The GUI helps the user to add any video directly into the system and we build a custom pipeline that breaks those moving images into multiple single frames and feeds it into the Meso- 4 model. After it processes the video file, it gives a probability score of dependency ratio for the data. The nature of the file is projected through percentage score that lets you decide whether the alleged video file is trustworthy or not.

For Headline text analysis, we have worked with different NLP models and tool kits to give us a stable system that could take input in a text format and analyses it to project whether the given statement is how much trustworthy. We have used LIAR dataset to enhance our system by using existing pre-trained models and constantly adding newer data points to keep updating the system. It lets you identify mis- quoted statements, misinformation and intestinally fabricated statements by classifying them into various trust levels.

2.1 Deep- Fake Detection

Meso-4 is a convolutional neural network with four convolutional blocks followed by one fully connected hidden layer. We first import different libraries necessary for the operations and then we created a dictionary to store all the image dimensions and the no, of color channels. The height and width is set to 256 since 3 channels were used to process the image. The system uses classifier class to load weights and make predictions. A meso-4 class was created to store classifier class. 3 dimensions of the image data was passed as the input layer which then proceeds to create 4 convolutional blocks. Convolutional layers always include convolution layer and max pooling layer and a batch normalization layer. Conv2d here represents the convolutional layer. Filters were used to assign variable parameters of the image data. CNN proceeds to higher order feature representations, from lines to corners to shapes to faces. Meso-4 has 4 blocks in its convolutional layer with one hidden layer and one output layer for the predictions. Binary class mode was established to classify images into true or false value; the method is called as binary classification task manager. We have built a custom preprocessor to divide the moving images into separate isolated independent images.

Figure 04. *Results of real image detection*

Figure 05. *Results of Deep-Fake images*

2.2 News Headline Analysis

We built a simple UI for web application that could help the user insert i.e. simply copy paste the text into a system and then it would initiate the process of identifying the source and nature of the content uploaded. We used LIAR dataset which was developed by William Yang Wang. We built few classifiers to detect the text pasted. After going through text extraction, we fed the extracted data into different classifiers. Naïve-Bayes, Logical Regression, Linear SVM, Stochastic Gradient different other classifiers were used to build a trustworthy classifying pipeline. Grid Search CV methods were used to perform parameter tuning. Models which suited the conditions better were selected to identify fake news. 50 additional features were extracted from the term frequency tfidf vectorizer to classify the most important words in class

Figure 06. *Results of news headline analysis (a)*

Figure 07. *Results of news headline analysis (b)*

Acknowledgement

A successful work of dissertation is the result of inspiration, support, guidance, motivation, and co- operation of facilities during study. It gives a great pleasure to acknowledge our deep sense of gratitude to present our project titled; "Fake News Analyzer". We would like to give sincere thanks to our Principal Dr.Suresh Ukarande and H.O.D. Dr. Jayashree Khanapuri for giving me the opportunity to present this seminar. I am also thankful to my respected guide Prof. Harshwardhan Ahire for this wholehearted support and affectionate encouragement. His dynamism, vision, and sincerity have deeply inspired us. He has taught us the methodology to present the research work as clearly as possible.

REFERENCES

[1] The psychology of fake news, Gordon Pennycook, David G Rand, Trends in cognitive sciences 25 (5), 388-402, 2021

[2] Fake news, Thorsten Quandt, Lena Frischlich, Svenja Boberg, Tim Schatto-Eckrodt

[3] The science of fake news. David MJ Lazer, Matthew A Baum, Yochai Benkler, Adam J Berinsky, Kelly M Greenhill, Filippo Menczer, Miriam J Metzger, Brendan Nyhan, Gordon Pennycook, David Rothschild, Michael Schudson, Steven A Sloman, Cass R Sunstein, Emily A Thorson, Duncan J Watts, Jonathan L Zittrain

[4] Science 359 (6380), 1094-1096, 2018F. Chollet. Xception: Deep learning with depth wise separable convolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1800–1807,2017. 5

[5] F. Chollet et al. Keras. http s://keras.io, 2015. 5

[6] W. Shi, F. Jiang, and D. Zhao. Single image superresolution with dilated convolution based multi-scale information learning inception module. arXiv preprint arXiv:1707.07128, 2017.S. Fan, R. Wang, T.-T. Ng, C. Y.-C. Tan, J. S. Herberg,
H. Farid. A Survey Of Image Forgery Detection. *IEEE Sig- nal Processing Magazine*, 26(2):26–25, 2009. 1

[7] P. Garrido, L. Valgaerts, O. Rehmsen, T. Thormahlen, P. Perez, and C. Theobalt. Automatic face reenactment. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4217–4224, 2014.

[8] MesoNet: a Compact Facial Video Forgery Detection Network, https://github.com/ondyari/FaceForensics

[9] Detecting fake news article Jun Lin, Computer Science and Engineering, Louisiana State University, Baton Rouge, LA, USA, Glenna Tremblay-Taylor, Computer Science, Keene State College, Keene, NH, USA, Guanyi Mou, Computer Science, Worcester Polytechnic Institute, Worcester, MA, USA, Di You, Computer Science, Worcester Polytechnic Institute, Worcester, MA, USA, Kyumin Lee, Computer Science, Worcester Polytechnic Institute, Worcester, MA, USA