Deepfake Detection Using Deep Learning

ASHOK V

MDL20CSIP02

M.Tech Computer Science and Engineering
with Specialisation in Image Processing

Department of Computer Engineering

Government Model Engineering College Thrikkakara

February 3, 2022

Overview

1. Introduction
2. Literature Survey
3. Literature Review
4. Problem Statement
5. Proposed Solution
6. Work Schedule
7. Hardware and Software Requirements
8. References

Introduction

- Massive expansion of machine learning and artificial intelligence, any manipulation in images and videos has become possible in recent years.
- ML and AI powered deep fake tools have contributed to the creation of fake news and distortion of politicians and celebrities.
- Deep learning is an important method in image processing area.
- Deep Fakes detection using histogram oriented gradients and support vector machine classifier is exist.
- It gave good accuracy on more accurate fakes, but still have problems with low resolution video fakes.

Literature Survey

Sl.No	Paper Details	Methodology	Advantages	Disadvantages
1	Passive Copy-Move Forgery Detection in Videos	Copy- Move forgery is a special kind of video forgery technique, in which copy-move of the frame region is performed in intra frame or inter frames.	Accuracy is high, Not a complex task	Images created using AI can not be detected. Real time detection is not possible
2	BusterNet: Detecting Copy-Move Image Forgery with Source/Target Localization	BusterNet is a pure, end-to-end trainable, deep neural network solution. It features a two-branch architecture followed by a fusion module.	Consist of CNN architecture for each domain (face, iris, and fingerprint)	Only works with biometric images

Literature Survey

Sl.No	Paper Details	Methodology	Advantages	Disadvantages
3	DeepFake Detection on Publicly Available Datasets using Modified AlexNet	Used a modified AlexNet constructed of an arrangement of 6 layers.	Used 3 publicly available datasets for training. Got a consistent accuracy on these three datasets	Accuracy is 87.49%
4	Deep Representations for Iris, Face, and Fingerprint Spoofing Detection	CNN for detecting the fake iris in biometric images.	First CMFD algorithm with discernibility to localize source/target regions. And it outperforms state-of-the-art copy-move detection algorithms by a large margin on the two publicly available datasets, CASIA and CoMoFoD	Real time detection is not possible

Literature Survey

Sl.No	Paper Details	Methodology	Advantages	Disadvantages
5	A Liveness Detection Method for Face Recognition Based on Optical Flow Field	The method is based on the assumption that a 3D face generates a special 2-D motion which is higher at central face parts (e.g. nose) compared to the outer face regions (e.g. ears)	It recognize spoofing attacks based on the correlation between optical flows in foreground and background regions	Works only with videos. And Low accuracy on deepfake dataset
6	Exposing Deep Fakes Using Inconsistent Head Poses	This method is based on the observations that Deep Fakes are created by splicing synthesized face region into the original image, and in doing so, introducing errors that can be revealed when 3D head poses are estimated from the face images.	Using features based on this method, an SVM classifier is evaluated using a set of real face images and Deep Fakes.	Low accuracy

Literature Survey

Sl.No	Paper Details	Methodology	Advantages	Disadvantages
7	Detection of Deep Network Generated Images Using Disparities in Color Components	This method analyze the disparities in color components between real scene images and deep fake images. Existing deep networks generate images in RGB color space and have no explicit constrains on color correlations; therefore, deep fake images have more obvious differences from real images in other color spaces, such as HSV and YCbCr, especially in the chrominance components.	It proposes a feature set to capture color image statistics for detecting deep fake images.	It can not detect deep fakes in videos and need high quality images for detection.
8	DeepVision: Deepfakes detection using human eye blinking pattern	It detect Deepfakes generated through the generative adversarial network (GANs) model via an algorithm called DeepVision. It analyze a significant change in the pattern of blinking.	Deepfakes can be determined through integrity verification by tracking significant changes in the eye blinking patterns	The integrity verification may not be applicable to people with mental illnesses or problems in nerve conduction pathways.

Literature Survey

Sl.No	Paper Details	Methodology	Advantages	Disadvantages
9	Exposing Region Splicing Forgeries with Blind Local Noise Estimation	Region splicing is a simple and common digital image tampering operation, where a chosen region from one image is composited into another image with the aim to modify the original image's content	It describe a method to expose region splicing by revealing inconsistencies in local noise levels, based on the fact that images of different origins may have different noise characteristics introduced by the sensors or post-processing steps.	It can not detect AI generated fake images
10	Detection of face spoofing using visual dynamics	Dynamic mode decomposition,Local binary pattern and support vector machine	Fast response and leading to better detection in video attacks	Vulnerable to image quality

Literature Survey

Sl.No	Paper Details	Methodology	Advantages	Disadvantages
11	Computationally efficient face spoofing detection with motion magnification	Eulerian motion maginification, LBP and Histogram of oriented optical flow	Efficient for video attacks and it can detect deepfake in real time	Less feasible to detect print attacks and 3D spoof models are undetectable
12	Face anti-spoofing countermeasure: Efficient 2DMaterials Classification	Stokes degree of linear polarization (SDOLP) with polarizer	High accuracy in both image and video deep fake. And fast response	Very expensive and external hardware is required
13	Integration of image quality and motion cues for face anti-spoofing: A neural network approach	Shearlet based image quality analysis and optical flow map	Better performance and fast response	Difficult to implement and requires high performance computer to run

Literature Survey

Sl.No	Paper Details	Methodology	Advantages	Disadvantages
14	Face2Face Manipulation Detection Based on Histogram of Oriented Gradients	This paper uses Histogram Oriented Gradients for feature extraction and Support Vector Machine for classification.	Real time detection is possible. It is very simple to implement.	The drawback of HOG is that its computation speed is slow while detecting an object and accuracy is not highly reliable compared to the current neural networks
15	Xception: Deep Learning with Depthwise Separable Convolution	Xception is a deep convolutional neural network architecture that involves Depthwise Separable Convolutions.	XCeption is an efficient architecture and give better performance than inception v3,VGG,ResNet	It is expensive to train, but are pretty good improvements compared to Inception. Transfer learning brings part of the solution when it comes to adapting such algorithms for a specific task.

Literature Review

- Reviewed some milestone deep learning methods like Histogram Oriented Gradients, InceptionNet, XceptionNet,etc
- Using deep learning it is possible to learn robust and high level feature representation of an image
- Detecting deep fakes in an image / video is possible with deep learning
- Among different deep learning methods, XceptionNet is showing high accuracy and immediate response.

	Top-1 accuracy	Top-5 accuracy
VGG-16	0.715	0.901
ResNet-152	0.770	0.933
Inception V3	0.782	0.941
Xception	**0.790**	**0.945**

Figure: Comparison of different architectures

Problem Statement

A system that automatically and accurately detect deepfakes / faceswaps in images or videos in real time.

Proposed Solution

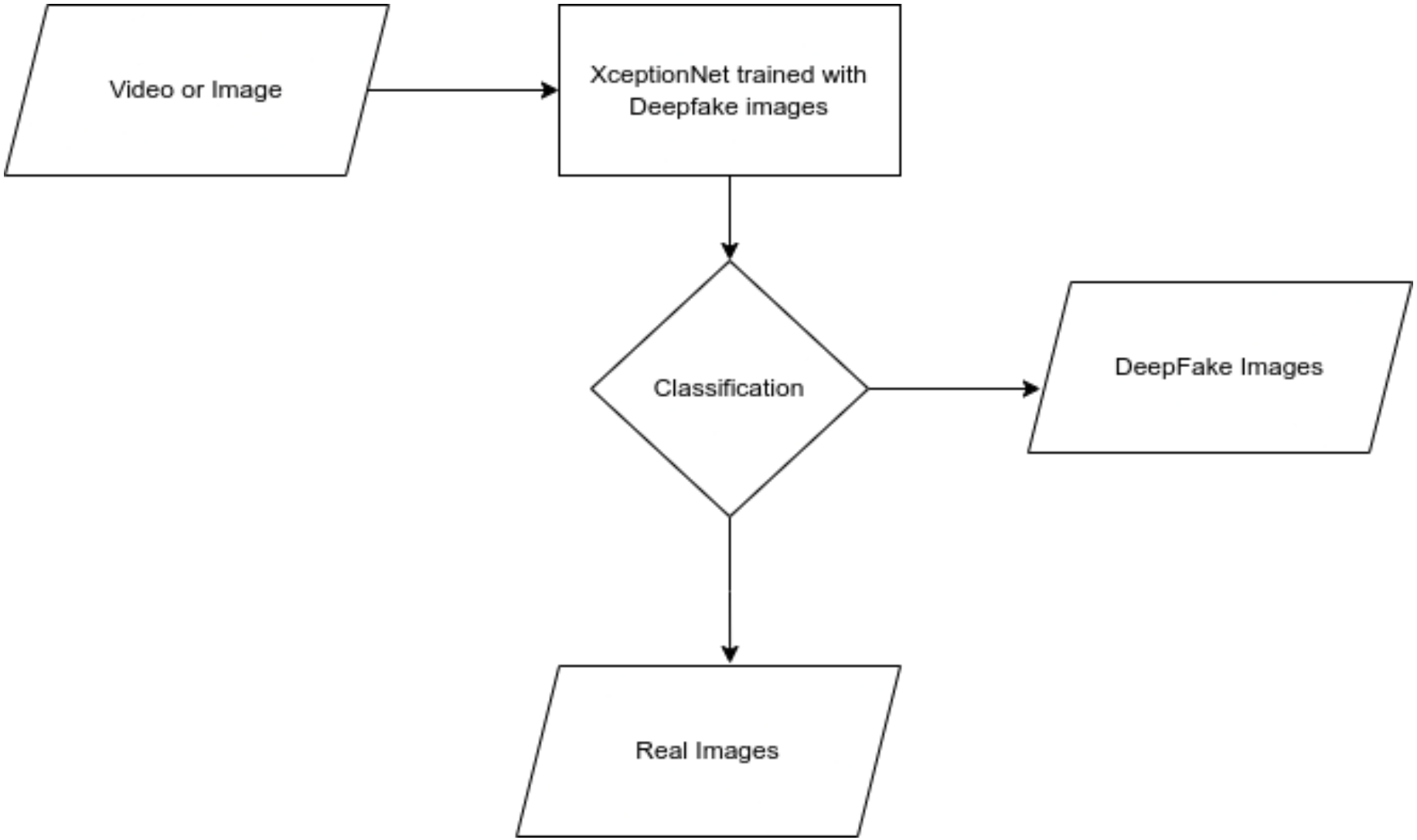

Figure: Proposed System

Work Schedule

Project Design	Feb 4^{th} Week
Initial Phase	April 4^{th} Week
Prototype	June 1^{st} Week
Final Implementation	June 4^{th} Week

Table: Tentative Schedule

Hardware and Software Requirements

OS	Windows 8, 10, Linux
Software	Anaconda Python distribution, OpenCV, Numpy, and Deep Learning libraries (Tensorflow, Keras)
CPU	Intel core i5 7^{th} generation processor or higher or an AMD equivalent processor
RAM	8GB
GPU	4GB
Disk Storage	50 GB of free disk space

Table: Hardware and Software Requirements

References

Pandey, Ramesh Chand, Sanjay Kumar Singh, and K. K. Shukla (2014)

Passive copy-move forgery detection in videos

International conference on computer and communication technology (ICCCT). IEEE.

Wu, Yue, Wael Abd-Almageed, and Prem Natarajan (2018)

Busternet: Detecting copy-move image forgery with source/target localization

Proceedings of the European Conference on Computer Vision (ECCV).

Xie, Daniel (2020)

DeepFake Detection on Publicly Available Datasets using Modified AlexNet

2020 IEEE Symposium Series on Computational Intelligence (SSCI).

Menotti, David (2015)

Deep representations for iris, face, and fingerprint spoofing detection

IEEE Transactions on Information Forensics and Security 10.4.

References

Bao, Wei (2009)

A liveness detection method for face recognition based on optical flow field

2009 International Conference on Image Analysis and Signal Processing. IEEE.

Yang, Xin, Yuezun Li, and Siwei Lyu (2019)

Exposing deep fakes using inconsistent head poses

ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE.

Li, H (2018)

Detection of deep network generated images using disparities in color components. arXiv 2018

arXiv preprint arXiv:1808.07276.

Jung, Tackhyun, Sangwon Kim, and Keecheon Kim (2020)

DeepVision: deepfakes detection using human eye blinking pattern

IEEE Access 8 (2020): 83144-83154.

References

Lyu, Siwei, Xunyu Pan, and Xing Zhang (2014)

Exposing region splicing forgeries with blind local noise estimation

International journal of computer vision 110.2 (2014): 202-221.

YFerrara, Pasquale, et al (2012)

Image forgery localization via fine-grained analysis of CFA artifacts

IEEE Transactions on Information Forensics and Security 7.5 (2012): 1566-1577.

Megahed, Amr, and Qi Han (2020)

Face2Face Manipulation Detection Based on Histogram of Oriented Gradients.

2020 IEEE 19th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom).

Jung, Tackhyun, Sangwon Kim, and Keecheon Kim (2020)

DeepVision: deepfakes detection using human eye blinking pattern

IEEE Access 8 (2020): 83144-83154.

Thank You !