Jyothi Engineering College
NAAC Accredited College with NBA Accredited Programmes*
Approved by AICTE & affiliated to APJ Abdul Kalam Technological University
A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR
JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777
NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.

Deep Fake Detection Using Machine Learning

Group no : 14

Department of CSE

Jyothi Engineering College

Thrissur

June 1, 2021

Group Members :

1. Karthik PC : JEC17CS061
2. M P Adithya Vijayan : JEC17CS069
3. Sanjana S : JEC17CS088
4. Thushara P : JEC17CS103

Guide : Ms. Aswathy Wilson

Vision of the Department

- Creating eminent and ethical leaders in the domain of Computational Sciences through quality professional education with a focus on holistic learning and excellence.

Mission of the Department

- To create technically competent and ethically conscious graduates in the field of Computer Science and Engineering by encouraging holistic learning and excellence.
- To prepare students for careers in Industry, Academia and the Government.
- To instill Entrepreneurial Orientation and research motivation among the students of the department.
- To emerge as a leader in education in the region by encouraging teaching, learning, industry and societal connect.

Contents

- Introduction
- Abstract
- Proposed System
- Requirements
- Literature Survey
- Design Architecture
- Modules
- Methodology
- UML
- Results
- Application
- Conclusions & Future Work
- References

Introduction

- Main objective of our project is to detect deep fake videos
- The amount of deep fakevideos are increasing rapidly
- Now there is need to detect whether a video is manipulated or not
- Our project aims to segregate deep fake videos and real videos
- Our system provides a method to detect these fake videos and thereby preventing the usage of these videos in creating political distress, blackmailing, fake terrorism events etc

Abstract

- Deep Fake videos are AI generated videos that look real but are actually fake
- Deep fake videos can have an adverse effect on a society
- Here we are using Xception model
- Images are transformed into various forms using albumentation which are used for classification
- Videos are classified into fake and real

Proposed System

- we are using Xception Network for Deepfake Detection
- As we are using more than one indicator for checking, it has more accuracy
- The system compare facial features
- According to the prediction value videos are classified as real and fake

Functional Requirements

- Detection of deepfake images
- Detection of deepfake images with varied resolution

Non functional Requirements

- The system should be scalable
- The system should be reliable

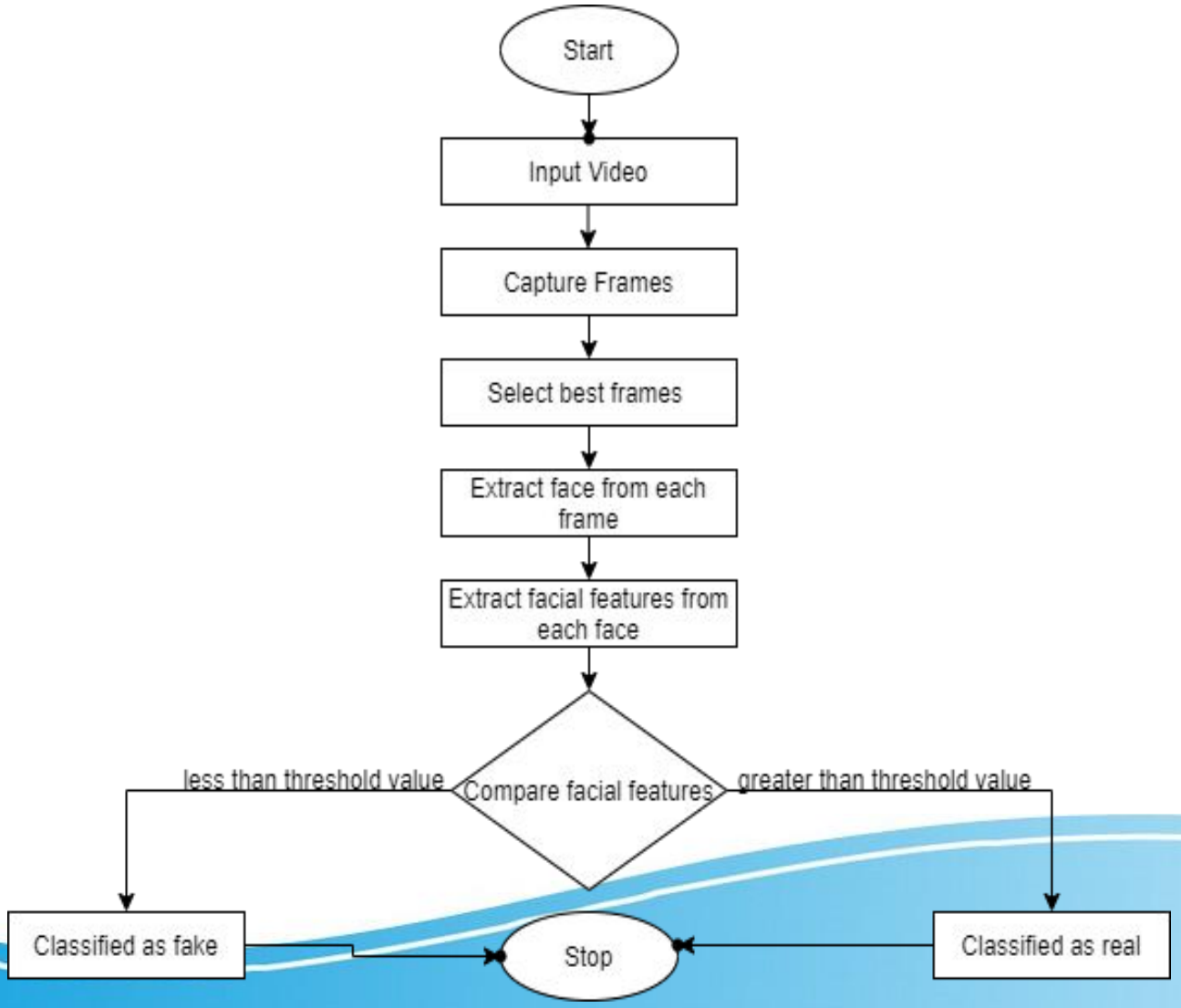
Literature Survey

- Deepfake Video Detection using Recurrent Neural Network
  - Uses a convolutional neural network to extract frame features.
  - Convolutional LSTM is used to predict whether an image is manipulated or not.
- Classification of Real and Fake Images Using One-Class Variational Encoder
  - It requires only real images for training so that data scarcity limitation can be solved
  - Eventhough it has 97% accuracy, better performance is only on NT and DFD dataset

- Deepfake Source Detection via Interpreting Residuals with Biological Signals
  - Uses PPG signal for the identification of images
  - Improves the accuracy of the network
- Classification of Deepfake Using Mouth Features
  - Two GAN Algorithms used one as encoder and other as decoder
  - CNN used for comparison
- Effective and Fast Deepfake Detection Method based on Haar Wavelet Transform
  - This method takes advantage of the fact that deep fake algorithms are only able to generate fake faces with specific size & resolution
  - This method is very complex and it has not been implemented yet

Jyothi Engineering College
NAAC Accredited College with NBA Accredited Programmes*
Approved by AICTE & affiliated to APJ Abdul Kalam Technological University
A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR
JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777
NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.

Design Architecture


Start
Input Video
Capture Frames
Select best frames
Extract face from each frame
Extract facial features from each face
Compare facial features
less than threshold value
greater than threshold value
Classified as fake
Stop
Classified as real

Modules

- Data acquisition module
- Image enhancement module
- Deepfake detection module

Data Acquisition Modules

- DeepFake dataset collection
  - Downloaded from face forensic++
- Classified the dataset into training and testing set

Image Enhancement Modules

- Captured frames from videos
- Frames are resized into required size
- Discarding unwanted frames
- Face extraction using blaze-face model

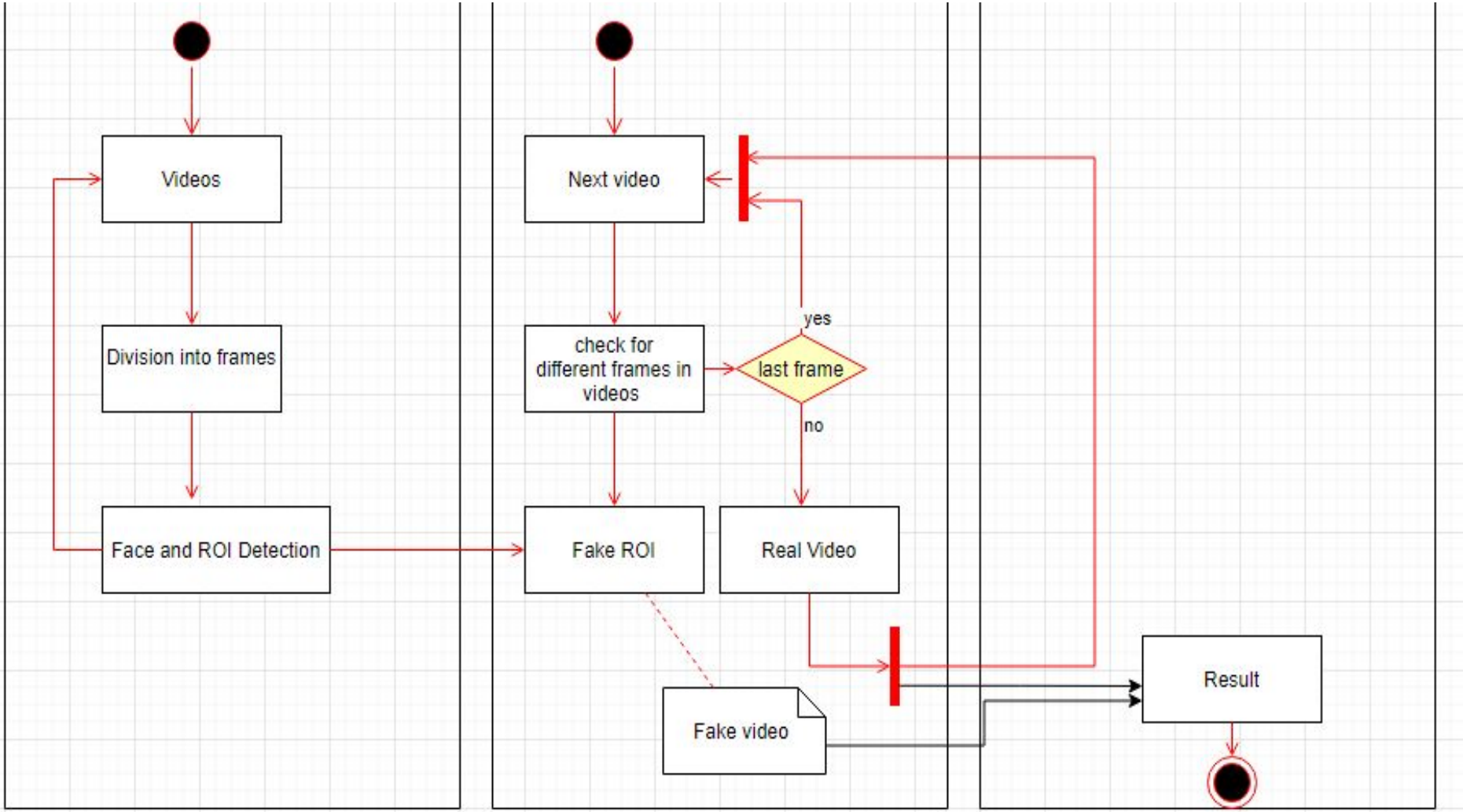
Deepfake Detection Modules

- Data loader function used to load training data
- Facial features are extracted
- Xception binary classifier is used to train
- Images are classified as real or fake

Methodology

- Input dataset is given and are converted to frames
- Face is extracted from real and fake frames
- Extracted face is transformed using albumentation
- Facial features are extracted
- Xception model is trained using these extracted features
- Model is trained for specified number of epochs
- Required test sample is passed
- Extraction of face and facial features from test sample
- Extracted features are compared with Xception model
- Videos are classified into fake and real

UML

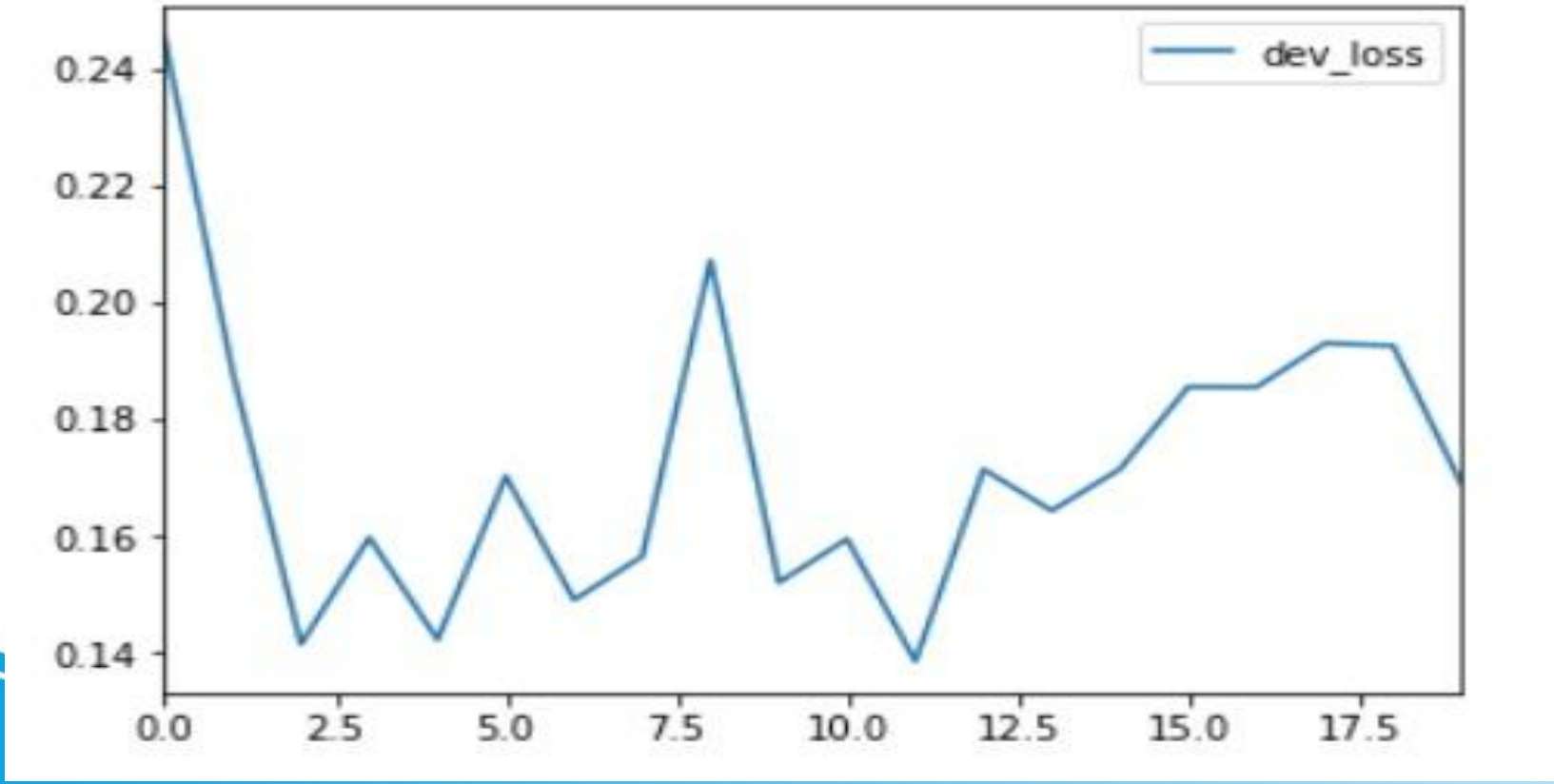

Videos
Division into frames
Face and ROI Detection
Next video
check for different frames in videos
yes
last frame
no
Fake ROI
Real Video
Fake video
Result

Results

```
Epoch 1/20, LR: 0.001000, Loss: 0.2207: 100%|██████████| 1013/1013 [07:57<00:00,  2.12it/s]
Dev loss: 0.2451, Acc: 0.897430, Kaggle: 0.281889
Saving best model...
Epoch 2/20, LR: 0.001000, Loss: 0.1623: 100%|██████████| 1013/1013 [07:57<00:00,  2.12it/s]
Dev loss: 0.1880, Acc: 0.917976, Kaggle: 0.244787
Saving best model...
Epoch 3/20, LR: 0.001000, Loss: 0.1416: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Dev loss: 0.1416, Acc: 0.940979, Kaggle: 0.206655
Saving best model...
Epoch 4/20, LR: 0.001000, Loss: 0.1232: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Dev loss: 0.1597, Acc: 0.946218, Kaggle: 0.204922
Epoch 5/20, LR: 0.001000, Loss: 0.1127: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Dev loss: 0.1423, Acc: 0.943189, Kaggle: 0.206867
Epoch 6/20, LR: 0.001000, Loss: 0.1038: 100%|██████████| 1013/1013 [07:57<00:00,  2.12it/s]
Dev loss: 0.1703, Acc: 0.938114, Kaggle: 0.217404
Epoch 7/20, LR: 0.001000, Loss: 0.0997: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Dev loss: 0.1491, Acc: 0.942453, Kaggle: 0.206499
Epoch 8/20, LR: 0.001000, Loss: 0.0894: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Dev loss: 0.1565, Acc: 0.940324, Kaggle: 0.211782
Epoch 9/20, LR: 0.001000, Loss: 0.0867: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Epoch     8: reducing learning rate of group 0 to 7.0000e-04.
Dev loss: 0.2072, Acc: 0.936640, Kaggle: 0.224125
Epoch 10/20, LR: 0.000700, Loss: 0.0785: 100%|██████████| 1013/1013 [07:59<00:00,  2.11it/s]
Dev loss: 0.1521, Acc: 0.949329, Kaggle: 0.198995
Epoch 11/20, LR: 0.000700, Loss: 0.0673: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Dev loss: 0.1595, Acc: 0.947937, Kaggle: 0.202024
Epoch 12/20, LR: 0.000700, Loss: 0.0652: 100%|██████████| 1013/1013 [07:58<00:00,  2.12it/s]
Dev loss: 0.1386, Acc: 0.950311, Kaggle: 0.197535
Saving best model...
Epoch 13/20, LR: 0.000700, Loss: 0.0615: 100%|██████████| 1013/1013 [07:59<00:00,  2.11it/s]
Dev loss: 0.1714, Acc: 0.944499, Kaggle: 0.207326
```

```
Epoch 14/20, LR: 0.000700, Loss: 0.0607: 100%|██████████| 1013/1013 [08:00<00:00,  2.11it/s]
Dev loss: 0.1644, Acc: 0.947610, Kaggle: 0.201195
Epoch 15/20, LR: 0.000700, Loss: 0.0572: 100%|██████████| 1013/1013 [08:00<00:00,  2.11it/s]
Dev loss: 0.1715, Acc: 0.947610, Kaggle: 0.204213
Epoch 16/20, LR: 0.000700, Loss: 0.0575: 100%|██████████| 1013/1013 [08:00<00:00,  2.11it/s]
Dev loss: 0.1855, Acc: 0.942043, Kaggle: 0.212465
Epoch 17/20, LR: 0.000700, Loss: 0.0568: 100%|██████████| 1013/1013 [07:59<00:00,  2.11it/s]
Dev loss: 0.1855, Acc: 0.947446, Kaggle: 0.206291
Epoch 18/20, LR: 0.000700, Loss: 0.0553: 100%|██████████| 1013/1013 [07:59<00:00,  2.11it/s]
Epoch    17: reducing learning rate of group 0 to 4.9000e-04.
Dev loss: 0.1930, Acc: 0.944172, Kaggle: 0.211803
Epoch 19/20, LR: 0.000490, Loss: 0.0475: 100%|██████████| 1013/1013 [07:59<00:00,  2.11it/s]
Dev loss: 0.1926, Acc: 0.943435, Kaggle: 0.212842
Epoch 20/20, LR: 0.000490, Loss: 0.0478: 100%|██████████| 1013/1013 [07:59<00:00,  2.11it/s]
Dev loss: 0.1689, Acc: 0.947937, Kaggle: 0.202401
```

Training loss graph



```
submission_df_xception.head()
```

	filename	label	result
0	aassnaulhq.mp4	0.971394	FAKE
1	aayfryxljh.mp4	0.009612	REAL
2	acazlolrpz.mp4	0.871641	FAKE
3	adohdulfwb.mp4	0.005082	REAL
4	ahjnxtiamx.mp4	0.748403	FAKE

Applications of Proposed System

- Decrease the spread of fake videos so that malicious abuser could not create fake news and mislead public
- Fake videos cannot be used for political distress and blackmailing
- Can be used in cyber crime detection centres
- Protection against fake celebrity pornographic videos
- Videos are classified as real and fake

Conclusion and Future work

- In this we extract features from face and compare with the pretrained model to predict whether it is fake or not.
- Our system provides a method to detect these fake videos and thereby preventing the usage of these videos in creating political distress, blackmailing, fake terrorism events, etc.
- Efficiency can be improved in future
- We will add UI for better usability

References

- .H. Khalid and S. S. Woo, **"OC-FakeDect: Classifying Deepfakes Using One-class Variational Autoencoder,"** 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 2020, pp. 2794-2803, doi: 10.1109/CVPRW50498.2020.00336.
- .R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "**Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization**," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74.
- D. Afchar, V. Nozick, J. Yamagishi and I. Echizen, "**MesoNet: a Compact Facial Video Forgery Detection Network**," 2018 IEEE International Workshop on Information Forensics and Security (WIFS), Hong Kong, Hong Kong, 2018, pp. 1-7, doi: 10.1109/WIFS.2018.8630761.
- Z. Boulkenafet, J. Komulainen and A. Hadid, "**Face Spoofing Detection Using Colour Texture Analysis**," in IEEE Transactions on Information Forensics and Security, vol. 11, no. 8, pp. 1818-1830, Aug. 2016, doi: 10.1109/TIFS.2016.2555286.
-] D. Guera and E. J. Delp, "**Deepfake Video Detection Using Recurrent Neural Networks**," 2018 15th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2018. [6] X. Yang, Y. Li, and S. Lyu, "Exposing Deep Fakes Using Inconsistent

Thank You