"Detecting Deep Fake Images using Convolutional Neural Networks"

NAGA SAHITHI VUNNAM , SHIVA REDDY DUBBAKA , VARSHITH GANDHE, KARTHIKEYA ARRA
UMKC School Of Science and Engineering

Acknowledgments

GUIDE: PROF. **SYED JAWAD SHAH**
- **CAM**
- **UDICA**
- **REGNIER INSTITUTE for E&I**
- **T-Mobile**

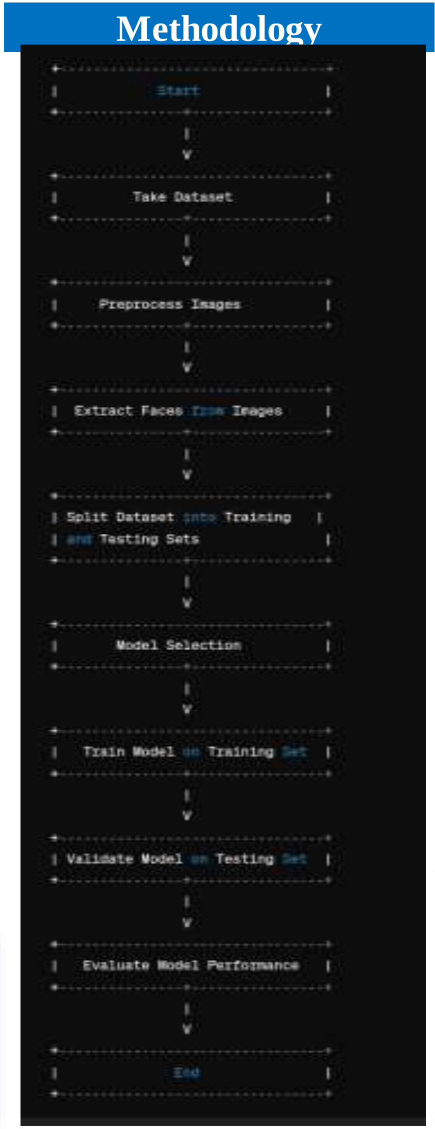
Introduction

The rapid advancement of deep fake technology poses a significant challenge in distinguishing between authentic and manipulated images. With the proliferation of deep fake content across various online platforms, there is a pressing need for robust detection methods to combat the spread of misinformation and safeguard digital authenticity. This project aims to address this critical issue by developing a deep learning model capable of accurately identifying deep fake images, thereby mitigating the potential societal and security implications associated with their deceptive use.

Objectives

➢Develop a deep learning model to accurately detect deep fake images.
➢Investigate the effectiveness of convolutional neural networks (CNNs) in distinguishing between real and fake images.

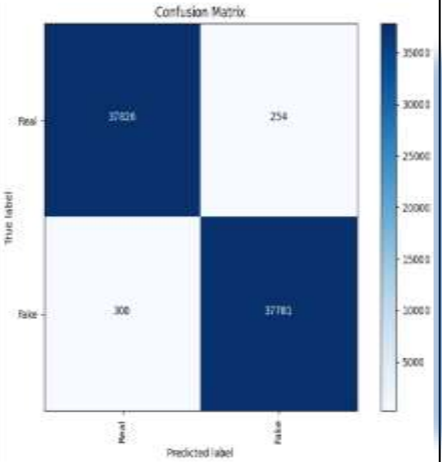
Methodology



```
+-------------------------------------+
|              Start                  |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
|          Take Dataset               |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
|      Preprocess Images              |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
|  Extract Faces from Images          |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
| Split Dataset into Training        |
| and Testing Sets                    |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
|        Model Selection              |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
|  Train Model on Training Set        |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
| Validate Model on Testing Set       |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
|  Evaluate Model Performance         |
+------------------+------------------+
                   |
                   v
+-------------------------------------+
|               End                   |
+-------------------------------------+
```

Results

Visual examination of the confusion matrix revealed a higher number of correctly classified real images compared to fake images, indicating the model's effectiveness in distinguishing between the two classes.



Discussions

Accurate detection of deep fake images is crucial for maintaining the integrity and trustworthiness of digital media in various domains, including journalism, politics, and entertainment.
The high accuracy and precision of the developed model demonstrate its potential for mitigating the harmful effects of deep fake technology by enabling automated detection and removal of manipulated content.

Conclusion

The project successfully developed and evaluated a deep learning model for detecting deep fake images with high accuracy and precision.
The model's performance metrics demonstrate its effectiveness in distinguishing between real and fake images, thereby addressing the pressing need for reliable deep fake detection methods..

Future work

Future research endeavors could focus on enhancing the model's robustness to adversarial attacks and novel deep fake generation techniques.
Additionally, collaborative efforts with industry stakeholders and policymakers are essential for implementing effective deep fake detection mechanisms at scale and enforcing regulatory measures to curb the proliferation of deceptive content online.

References

•Manjil Karki. (2020). Deepfake and Real Images. Kaggle. Retrieved from https://www.kaggle.com/datasets/manjilkarki/deepfake-and-real-images

•Siwei Lyu, Ming-Ching Chang, and Luisa Verdoliva. (2021). DeepFake Detection Dataset (DFDD). Zenodo. DOI: 10.5281/zenodo.5528418