DEEPFAKE: UNVEILING DIGITAL DECEPTIONS

GROUP 5 MEMBERS

- OLIVE MULOMA
- ABIGAIL MWENDWA
- HAWKINS MURITHI
- HARRY ATULAH

TABLE OF CONTENTS

01 OVERVIEW

02 BUSINESS UNDERSTANDING

03 DATA UNDERSTANDING

04 MODELLING

05 EVALUATION

06 RECOMMENDATION

01

OVERVIEW

- **Goal** : Create a deepfake detection system.
- **Functions** : Detect and verify fake images.
- **Process** :
  1. Collect data
  2. Build model and Evaluate accuracy
  3. Test and deploy

02

BUSINESS UNDERSTANDING

STAKEHOLDERS

- **Media and Content Verification Teams**: Gain better tools for authenticating images.
- **General Public**: Benefit from more trustworthy visual media.
- **Content Moderation**: Use detection models to ensure image authenticity.

BUSINESS PROBLEM

Threat : Advanced deepfake tech creates convincing fake media that misleads and threatens security.

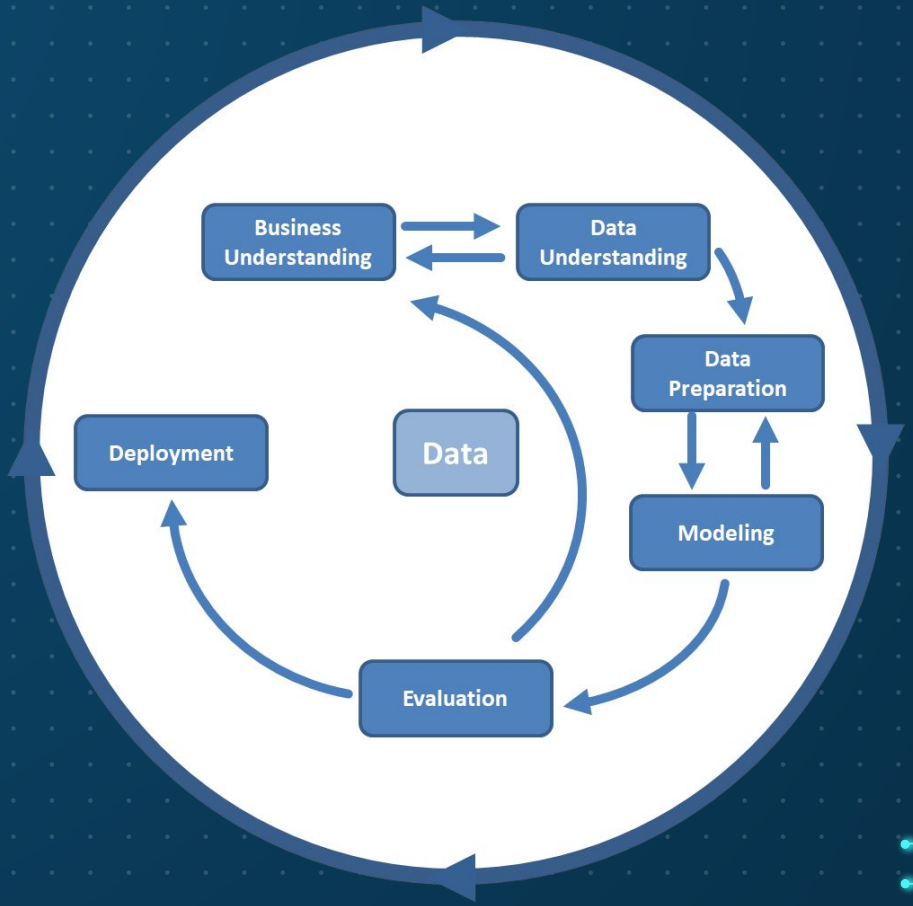
Malicious Use : Deepfakes distort reality and damage reputations.

Goal : Develop a solution to accurately identify and distinguish fake media.

OBJECTIVES

- **Detection** : Build a precise model to identify fake media.
- **Performance** : Ensure consistent results with diverse data.
- **Usability** : Create an easy-to-use verification interface.

CRISP-DM – Project framework



DATASET

Dataset Title : CIFAKE - Real and AI-Generated Synthetic Images

Size : 1.34 GB

Total Images : 120,000

Categories : Real and AI-generated synthetic images

Labels :

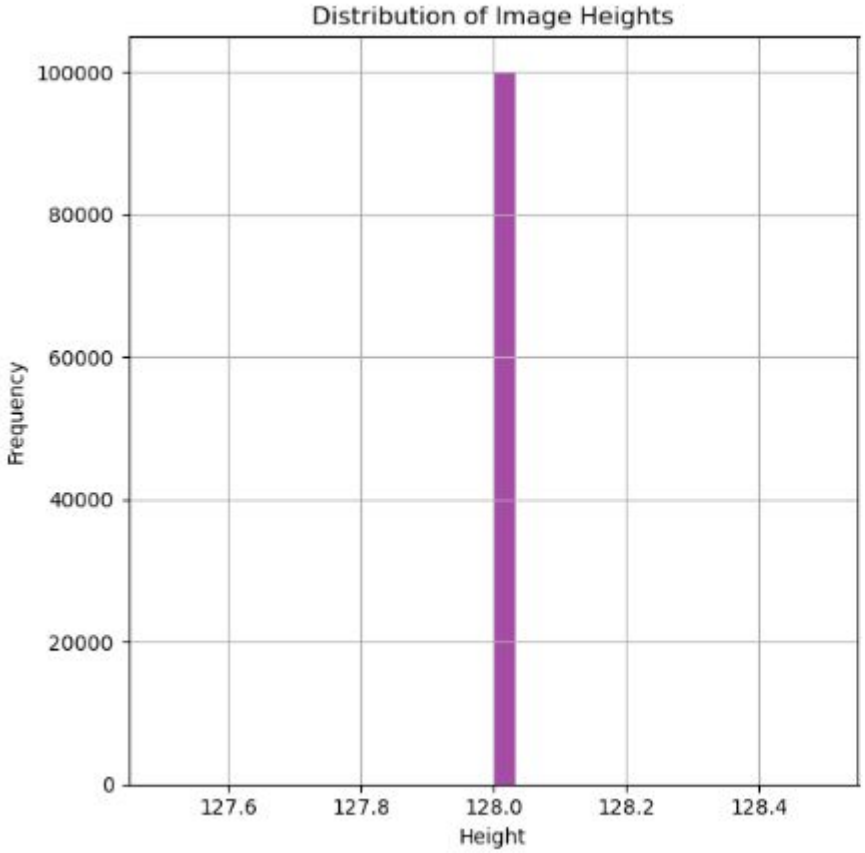
- 0: Real images
- 1: AI-generated images

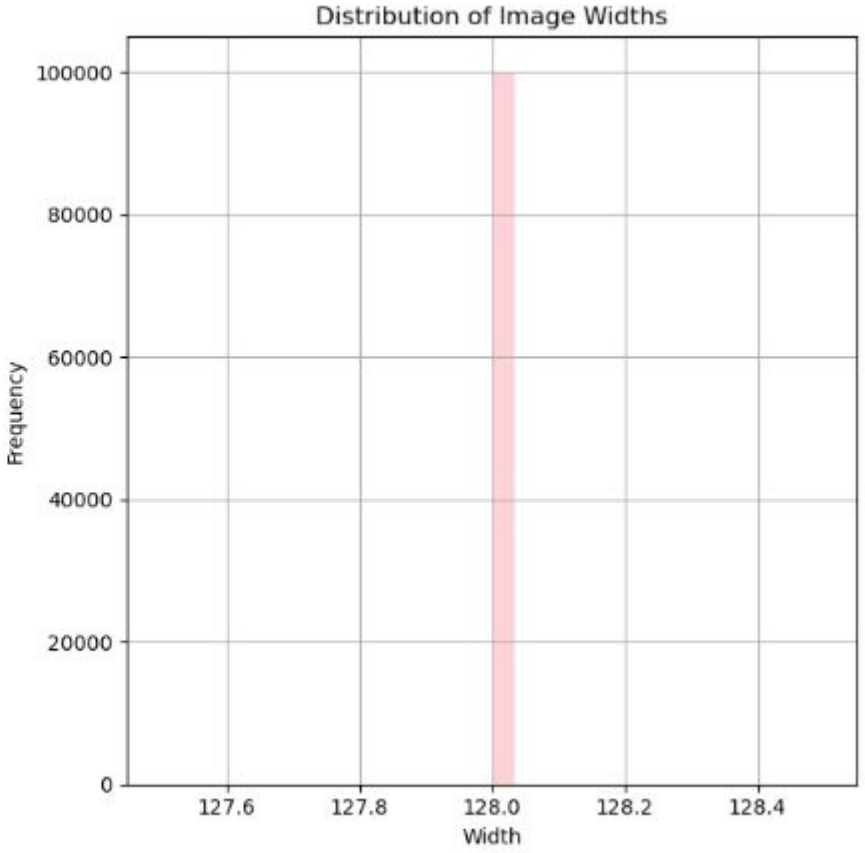
DATA ANALYSIS

- **Data Analysis : Examined 60,000 real and 60,000 AI-generated images to understand patterns and ensure quality.**

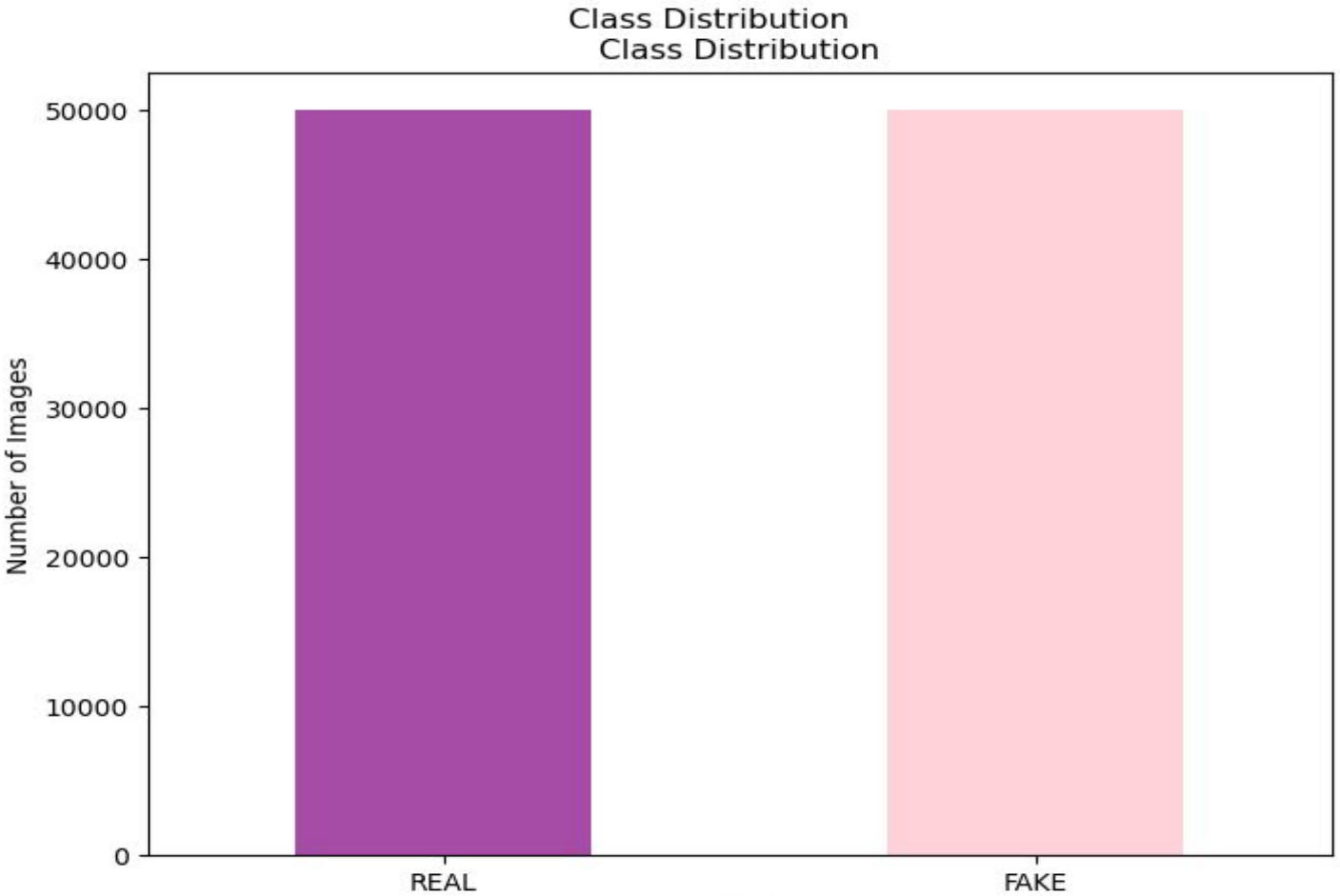
- **Data Cleaning : Used visual checks to guide preprocessing and feature extraction.**

Distribution of Image Sizes

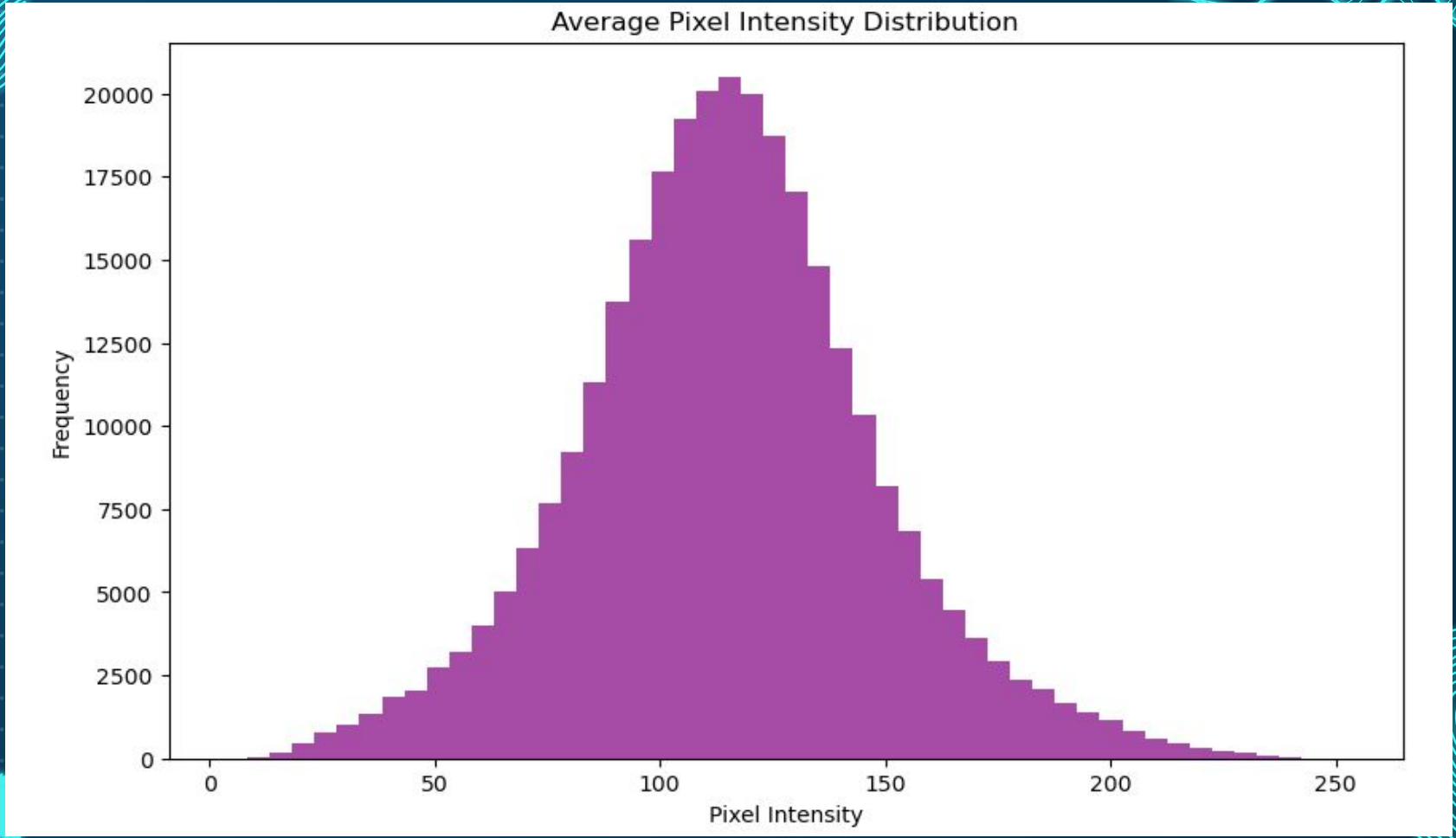




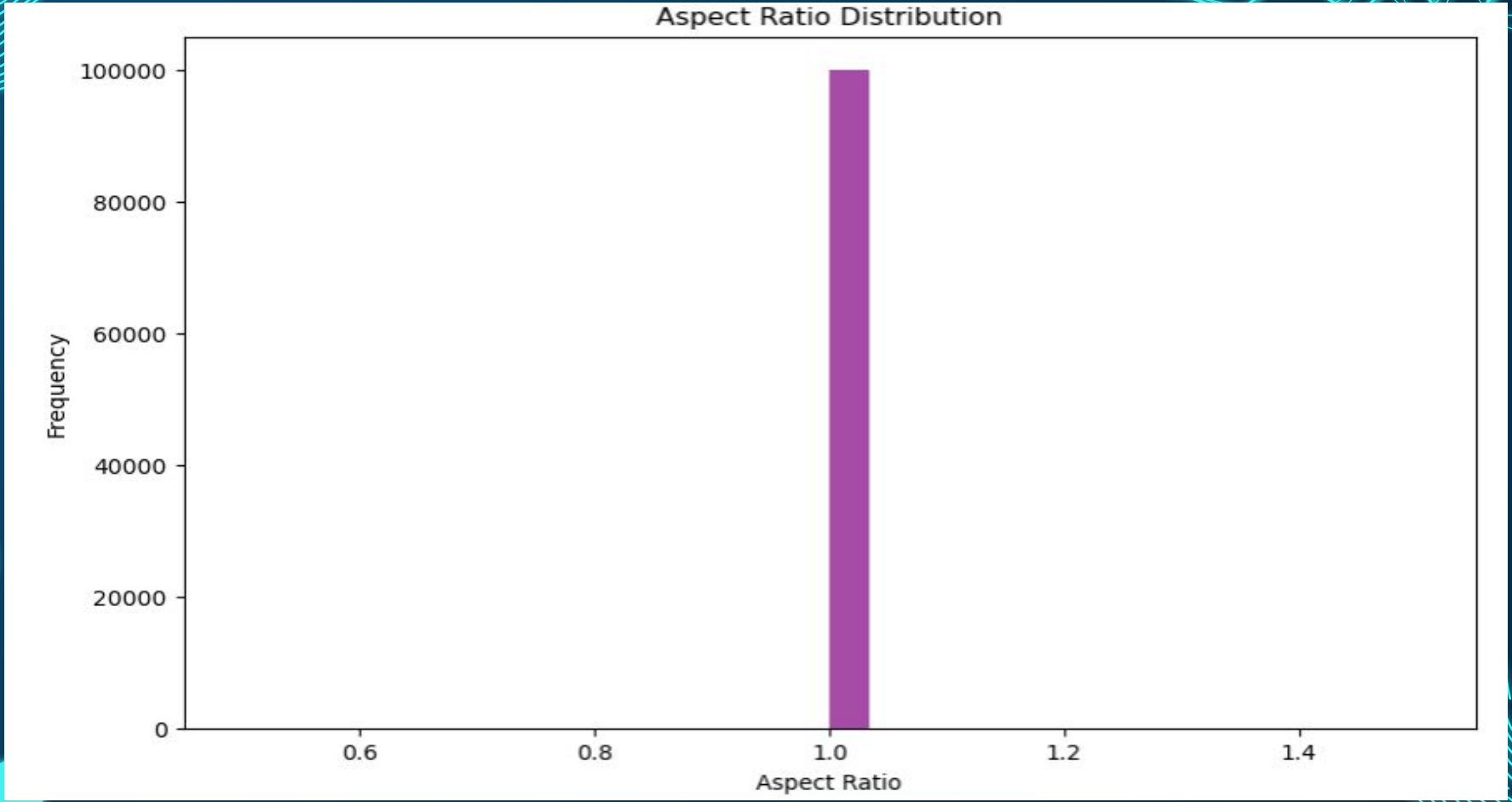
Class distribution



Average Pixel Intensity Distribution



Aspect Ratio Distribution



04

MODELLING

- **CNN Development** : Built and improved a model to detect deepfakes by identifying whether images are real or fake.

- **Architecture** : Uses layers to pick out details in images and sort them, with methods for effective training and strong results.

05

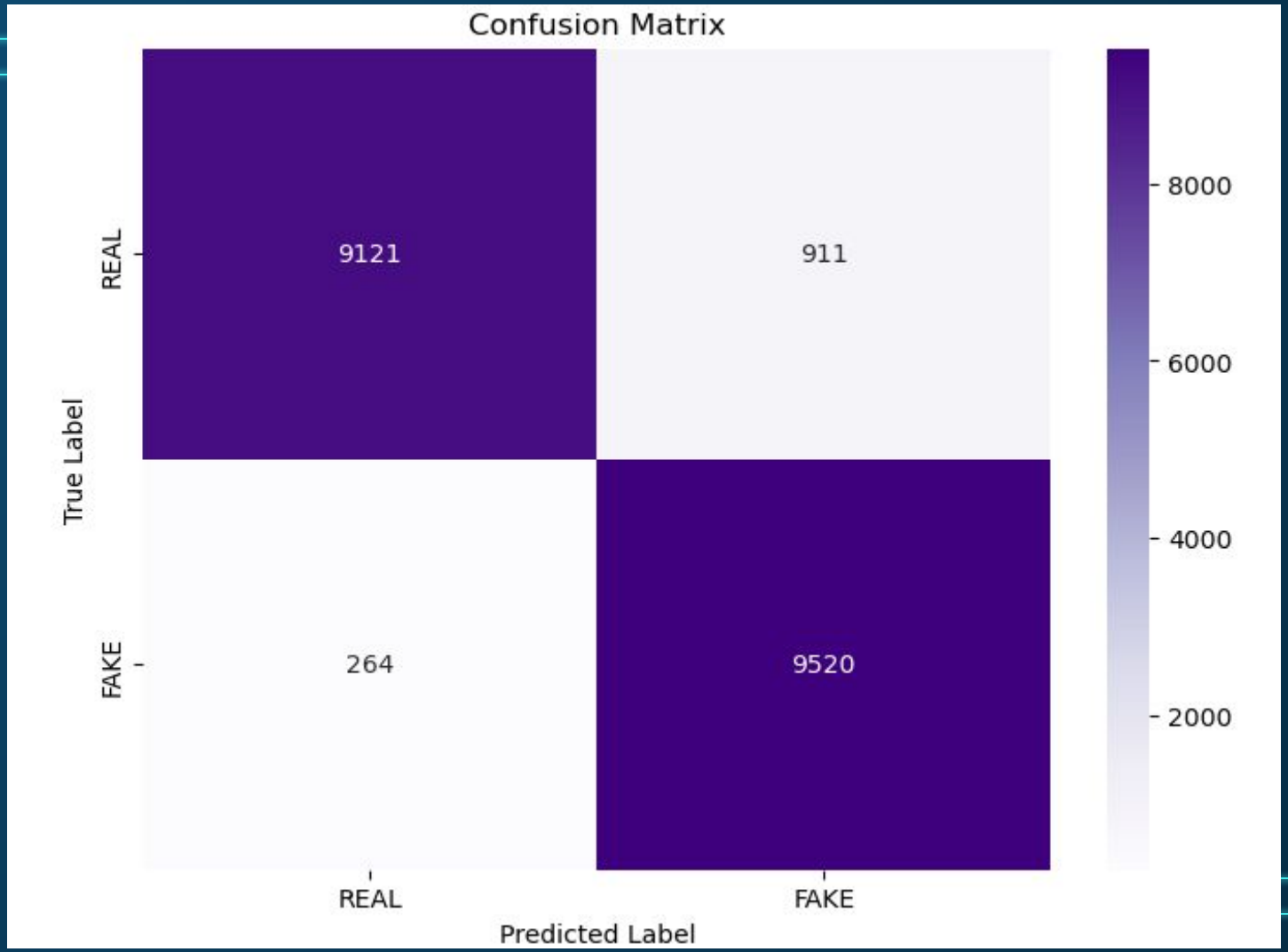
EVALUATION

- **Accuracy** : The model correctly identified 94% of images as real or fake.

- **Precision and Recall** : Both were high, with a 94% score for detecting real and fake images.

- **F1 Score** : The model was slightly better at spotting fake images (97%) than real ones (91%).

- **Robustness Testing** : The model was tested on different sets of images to ensure it performed well in various situations.


Confusion Matrix
9121
911
264
9520
REAL
FAKE
True Label
REAL
FAKE
Predicted Label
8000
6000
4000
2000

05

USER INTERFACE & DEPLOYMENT

- **Web Application** : A simple website where users upload photos to check if they are real or fake.

- **Deployment** : The site handles many files at once and adjusts as more people use it.

- **Access** : Available online at "Deepfake-Detector" for easy media verification.

Deepfake Detection

Deepfake Detection

Upload a photo to check if it's real or fake.

Choose File No file chosen

Predict

© 2024 Deepfake Detection

15°C
Partly cloudy
Search
23:44
19/08

127.0.0.1:5000

Deepfake Detection

Deepfake Detection

Deepfake Detection

Upload a photo to check if it's real or fake.

Choose File WhatsApp Image 2024-08-16 at 12.27.16 PM.jpeg

Predict

Prediction: Fake

© 2024 Deepfake Detection

07

CHALLENGES

- **Hosting** : Used DigitalOcean for large model.
- **Hardware** : Used a more powerful laptop.
- **Dataset** : Added own images for accuracy.
- **Costs** : DigitalOcean for cost-effective hosting.
- **Interface** : Created a simple design.

06

RECOMMENDATIONS

- **Verify Media** : Establish pre-publication checks for authenticity.
- **Boost Collaboration** : Enhance teamwork between creators and verification teams.
- **Use Detection Tools** : Integrate deepfake detection in moderation systems.

FUTURE WORKS

- **Include Videos** : Add video analysis capability.
- **Expand Dataset** : Use a more diverse dataset for better accuracy.
- **Advanced Techniques** : Explore newer models for improved performance.
- **User Feedback** : Integrate feedback to enhance the system.

~Why did the deepfake go to therapy?
Because it had an identity crisis!~

THANK YOU!