AI-Generated Image Detection

Dr. Li, Aryan Singh — College of Engineering and Computer Science, University of Michigan-Dearborn

SURE EXP+ Undergraduate Research

INTRODUCTION

- Rapid rise of AI-generated images creates challenges in distinguishing real from artificial content.
- Growing concerns over AI-generated art's impact on artists' careers.
- GenImageDetector introduces a framework using advanced detection methods to identify AI-generated images.

PURPOSE

- Develop a reliable method to differentiate AI-generated images from real ones.
- Utilize neural networks, metadata analysis, forensic analysis, and error level analysis.
- Ensure adaptability and longevity of the framework as AI technologies evolve.

CONCLUSION

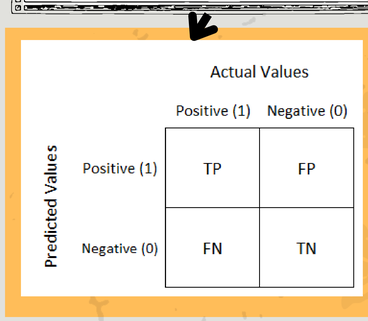
- Through rigorous testing we are able to determine traditional DNN detection models are not sufficient enough to accurately determine an AI-generated image
- ResNet50 provided a more accurate calculation of AI-generated images from CNNSpot

METHODOLOGY

- Framework Design: Combines multiple detection techniques for comprehensive image classification with GenImage models: CNNSpot and ResNet50.
- Benchmark Datasets: Utilized popular datasets for training and evaluating AI image detection algorithms.
- Performance Evaluation: Rigorous testing using benchmark datasets, focusing on accuracy, reliability, and adaptability, with Confusion Matrices measuring detection performance.

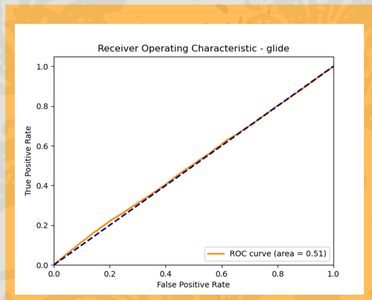
CONFUSION MATRIX

A confusion matrix is a table that compares predicted values to actual values in a dataset to evaluate the performance of a classification model. It's also known as an error matrix

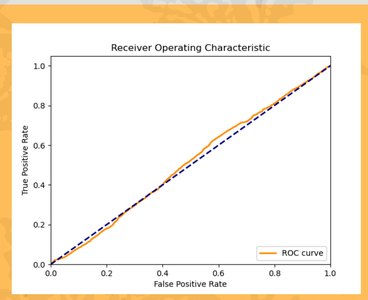


RESULTS

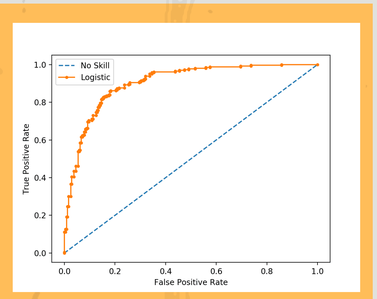
- Below are an example of results from CNNSpot (traditional method) and ResNet50 (modern method)



CNNSpot tested on glide dataset



ResNet50 tested on glide dataset



ResNet50 tested on all datasets

AI DETECTION MODEL

The process of training and evaluating an AI model can be daunting. There is an example of the workflow of this project below.

```
├── imagenet_ai
│   ├── train
│   │   ├── ai
│   │   ├── nature
│   ├── val
│   │   ├── ai
│   │   ├── nature
```

Organization of dataset

```
# Python 3.x
pip3 install torch
```

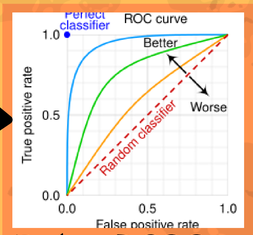
Install proper packages

```
Model resnet50 created, param count:25557032
Data processing configuration for current model + dataset:
        input_size: (3, 224, 224)
        interpolation: bicubic
        mean: (0.485, 0.456, 0.406)
        std: (0.229, 0.224, 0.225)
        crop_pct: 0.95
AMP not enabled. Training in float32.
Scheduled epochs: 60
```

train model for 50 iterations

```
Validating in float32. AMP not enabled.
Loading pretrained weights from url (https://github.com/rwig
Model resnet50 created, param count: 25557032
Data processing configuration for current model + dataset:
        input_size: (3, 224, 224)
        interpolation: bicubic
        mean: (0.485, 0.456, 0.406)
        std: (0.229, 0.224, 0.225)
        crop_pct: 0.95
```

Validate model accuracy on all datasets



Evaluate ROC Curve from confusion matrices

PRO SKILLS

- Communication: Maintaining constant communication to ensure proper encapsulation of the mentors vision. Any problems that would arise within the code could be resolved quickly.
- Technology: Usage of Python coding language and AI related python libraries such as Pytorch. Learning curve on what it means to train, test, and evaluate industry standard AI tools. Utilized university hardware through ssh server.

NEXT STEPS

- More organization when it comes to examining the codebase of different models
- Increased accessibility across all operating systems to quickstart model training and testing
- Possible mixture of best performed models to produce a new model

9.09.2024