Image Classification: AI-Generated and Real Images

Sanya Madan	Parisha Agrawal	Brinda Muralie
2021561	2021270	2021140
CSAI'25	CSAI'25	ECE'25
sanya21561@iiitd.ac.in	parisha21270@iiitd.ac.in	brinda21140@iiitd.ac.in

Abstract

This project aims to develop robust classification methods for distinguishing between real and AI-generated images. Using datasets such as CIFAKE which includes both types of images, extracting discriminative features, and evaluating various models, the project seeks to address the challenge of identifying increasingly realistic AI-generated images, with potential applications in image regulation both online and in real world.

1. Motivation

Differentiating between real and AI-generated images has become a complex task for humans due to the remarkable progress in AI image generation. Misleading images that convincingly imitate reality can cause significant harm like spreading fake news, damaging reputations, and manipulating public opinion through media. The judicial system also often relies on visual evidence. Identification and mitigation of fake images is crucial to ensure the authenticity of visual content.

2. Related Work

[1] CIFAKE: Advancing AI-Generated Image Recognition : Introduces dataset, enhancing AI-generated image recognition to 92.98%.

[2] GenImage: Advancing AI-Generated Fake Image Detection : Empowers development of fake image detection methods with a comprehensive dataset.

[3] Enhancing Detection of Diffusion-Generated Images : Proposes multiLID method for accurate detection of deep diffusion model-generated images.

3. Final Outcome:

By using AI's capability to detect AI, we aim to bridge the gap between human limitations and AI's enhanced perception, leading to more accurate image classification. This project aims to help combat the issue of validating authenticity of images and design a reliable image classification system by vectorizing images using our model that can detect patterns and anomalies in synthetic images, which can further be used by regulators and platforms.

4. Timeline

Weeks	Tasks
Week 2-4	Researching data sets, Data visualization and preprocessing (FDA, LDA, PCA, outlier detection). Also analyse their use and relevance.
Week 4-6	Apply different ML models (Linear, logistic regression, Naive bayes, Decision tree, Random forest, SVM)
Week 6-8	Hyper-parameter tuning, Analysis of outcomes
Week 10-12	Further optimisation, use Neural Networks (multilayer perceptron and CNN)
Week 13	Writing Report

5. Individual Tasks

Name	Tasks
Parisha	Data Visualization, Data Preprocessing (PCA), Grid Search, Applying model (Decision tree, Random forest, CNN), Hyperparameter tuning, Report Writing
Sanya	Data Visualization, Data Preprocessing (FDA, LDA), Grid Search, Applying model (Naive Bayes, SVM), Hyperparameter tuning, Report Writing
Brinda	Data Visualization, Data Preprocessing, Grid Search, Applying model (Regression, KNN, K-Means, MLP), Hyperparameter tuning, Report Writing

Note: This is the tentative task distribution based on our understanding of the models until now, but we will mostly be collaborating and working together on all the models.