CAN AI ACCURATELY DISTINGUISH REAL FROM AI IMAGES ?

جامعة الملك سعود
King Saud University
كلية علوم الحاسب والمعلومات
قسم تقنية المعلومات

Introduction

Artificial Intelligence (AI) has achieved major breakthroughs in image generation.
This project investigates whether AI can truly create images indistinguishable from real photographs by training a model to classify real versus AI-generated images, revealing AI's current capabilities and limitations.

Objectives

- **Can** a model correctly tell real and AI images apart without knowing the category?
- **What** features (like edges, colors) help in classification?
- **How** do image edits (like resizing or brightness changes) affect results?
- **What** mistakes does AI make during classification?
- **How** well do humans perform compared to AI?

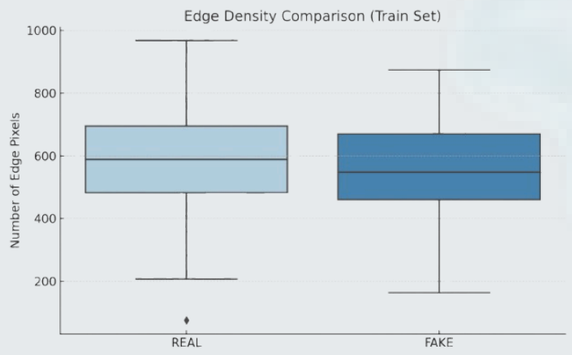
Data Collection

- **Dataset**: CIFAKE (from Kaggle)
- **Content**:
  - 60,000 real images (from CIFAR-10 categories).
  - 60,000 AI-generated images (via Stable Diffusion model).
- **Preparation**:

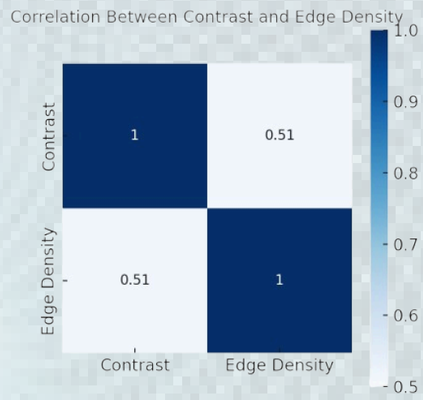
Organized into training and testing sets, ensuring diversity in lighting, backgrounds, and styles.

Data Analysis

- **Key Findings from EDA:**
  - Real images have natural edges and texture variations.
  - AI images often show overly smooth or sharp features.
  - Color distribution showed minimal impact.
- **Important Observations:**
  - Blurring images significantly reduced model accuracy.
  - Preprocessing (e.g., flipping, brightness changes) confused the model in some cases.



Examples of real images with low edge density.

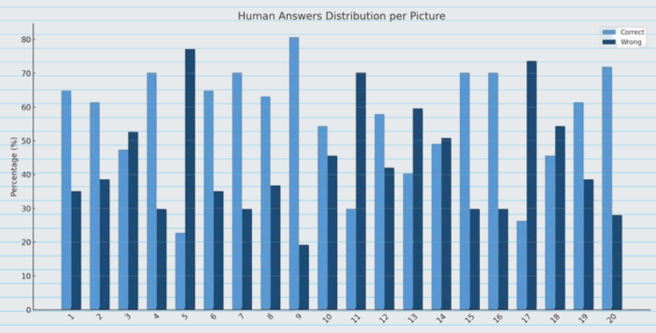


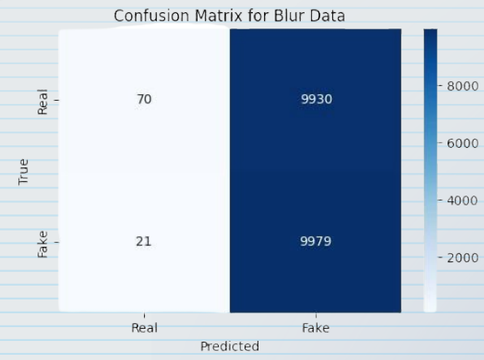
Moderate correlation between contrast and edge density.

Models and Findings

- Models Used: Pre-trained Vision Transformer (ViT), Lightweight CNN, Deeper CNN.
- Best Accuracy: 98% (Pre-trained ViT).
- Insights:
- Fine details (edges, texture) are critical.
- Color alone is not enough.
- Transformations (blur, rotation) reduce accuracy.
- **Human vs. AI:**

AI outperformed humans in classification: **98.41% vs. 56.14%.**





Accuracy dropped after blurring, confirming the importance of fine details.

Conclusions

- **AI** models effectively distinguish real from AI-generated images.
- **Fine** details are crucial for accurate classification.
- **Preprocessing** impacts model reliability.
- **Real-world** image variations still challenge AI performance.
- **Humans** achieved only 56.14% accuracy, confirming AI's superiority.