

# Project Explanation Due Oct-4

October 9, 2024

## 1 Dataset Overview

- **Dataset Name:** CASIA1 Ground Truth
- **Data Composition:** The dataset is composed of two main classes: real images and fake images. Fake images are altered using methods such as image splicing (cutting and pasting parts from other images) and simple post-processing techniques like color adjustments and blurring.
- **Format:** Images are provided in standard formats (jpg, tif), and ground truth data is available, labeling which images are tampered with and which are not. This ground truth is essential for training machine learning models.
- **Size:** The dataset includes a substantial number of images. The images come from diverse settings, ensuring variability in image content which is crucial for training a robust model.

## 2 Problem Description

We are addressing fake image detection in news media. With the rise of misinformation, there has been an increasing challenge in identifying manipulated or fake images used in news outlets and social media platforms to mislead the public. These fake images can sway public opinion, contribute to the spread of false information, and cause societal harm.

The goal is to develop a machine learning model that can automatically detect these fake images with a high degree of accuracy. This model will need to be trained on labeled data (real vs. fake) to learn the subtle differences between authentic and tampered images, potentially using features like pixel inconsistencies, unnatural edges, or lighting mismatches that are indicative of tampering.

## 3 Stakeholders of the Project

- **News Agencies:** Media outlets that are concerned with the integrity of the images they publish.
- **Social Media Platforms:** Companies like Twitter, Facebook, and Instagram, where fake images can spread rapidly.
- **Government and Law Enforcement Agencies:** In matters of public safety or legal cases where image manipulation could influence outcomes.
- **Advertising companies:** Fake image content might harm brand credibility.
- **Journalists and Reporters:** Professionals who are committed to ensure credibility of their reports and need tools to verify the news.
- **Fact-Checking Organizations:** Websites like Snopes or Politifact which verify the authenticity of information.
- **Social Media Users**

## 4 Company Key Performance Indicators (KPIs)

- **Detection Accuracy:** The percentage of fake images accurately detected by the model.
- **False Positive Rate (FPR):** The rate at which real images are incorrectly classified as fake.
- **False Negative Rate (FNR):** The rate at which fake images are incorrectly classified as real. **Processing Speed:** The time it takes for the model to classify an image, which is crucial for real-time applications.
- **Scalability:** The model's ability to handle large volumes of data without significant degradation in performance.
- **User Adoption:** The extent to which news agencies and platforms integrate and use the tool for detecting fake images.
- **Reduction in Fake Image Spread:** Measuring the reduction in the dissemination of fake images after the tool's implementation.
- **Model Robustness:** The model's ability to detect various forms of tampering, including new or sophisticated manipulation techniques.