# Context-Aware Detection of Manipulated Satellite Images

Matthew Chapman- supervisor Dr Deepayan Bhowmik

## Introduction and Motivation

The project aims to create an image manipulation detection system for satellite images, the development of a manipulation detection/prevention algorithm is important since maliciously altered images have previously been used to discredit celebrities/politicians, commit financial fraud and create misinformation to destabilise governments, create riots/ insurrections, with the spread of social media applications such as X (Twitter), Telegram and Facebook. In 2020 it was reported by the National Geospatial-Intelligence Agency that machine learning had been used to modify satellite images in order to create misinformation, thus creating a motivation to develop a system to detect maliciously altered satellite images [1]

### Ways Satellite images can be manipulated

Satellite images can be manipulated using several methods.

- GAN based synthesis methods
- Image removal and addition/Selective Editing
- Colour manipulation

In addition to these methods it is possible that the image could be made to fool neural networks by adding noise to it in order to create an adversarial attack and therefore cause a neural network system to misclassify the image

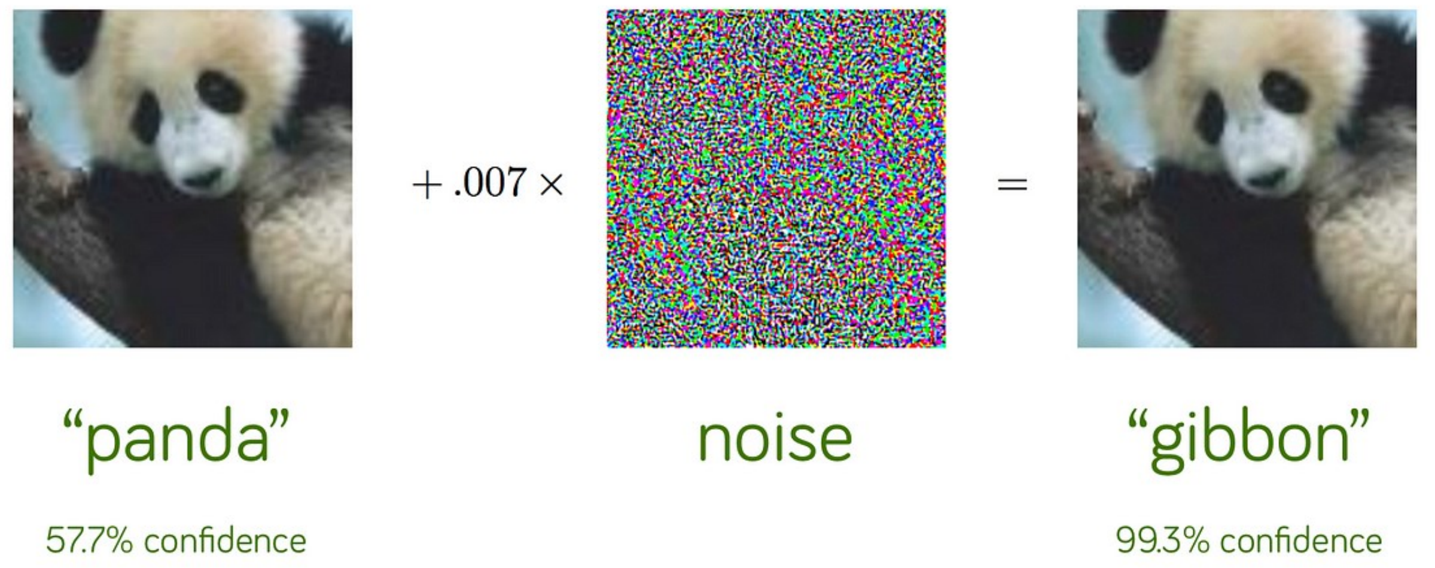

Figure 1: Brief Example depicting how an Adversarial Attack can happen. [2]

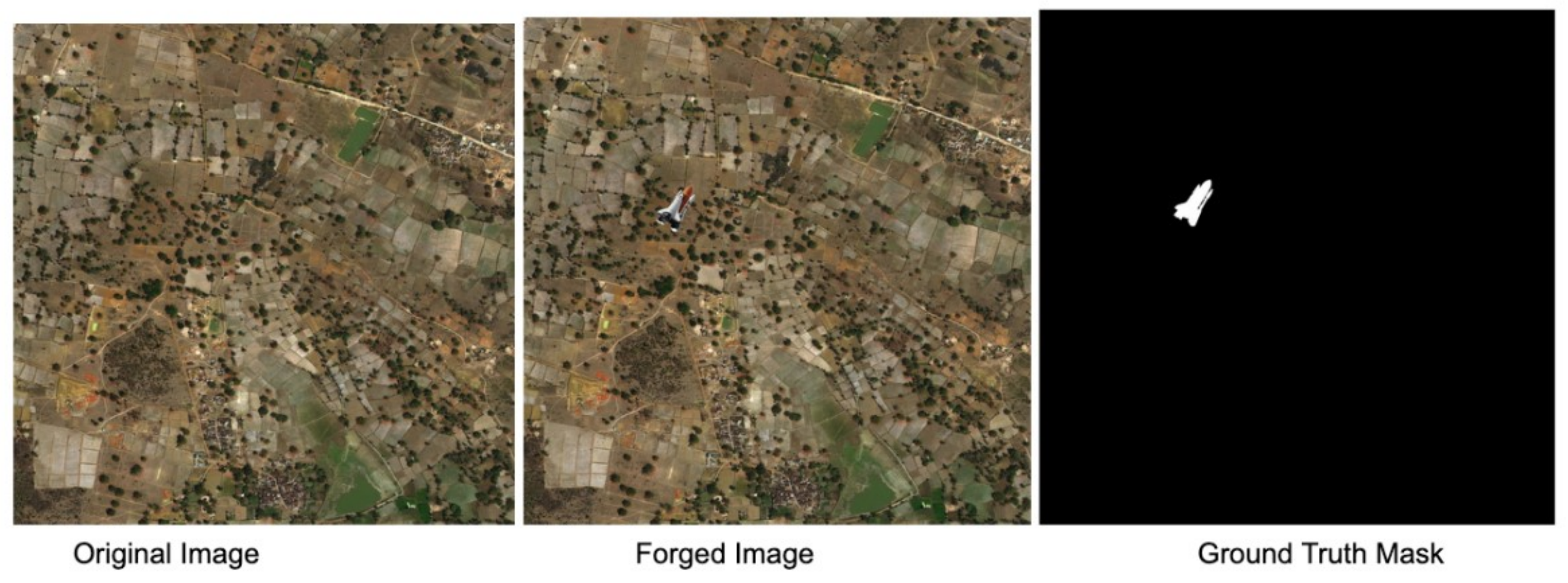

Figure 2: An example of a satellite image could be manipulated using Image removal and addition [3]

## Generative Adversarial Network (GAN) based methods

GANs are deep learning neural networks where two neural networks compete against each other to generate more realistic synthetic data from a training dataset as seen in figure 2. See figure 3 and 4 for satellite images produced using a StyleGAN neural network

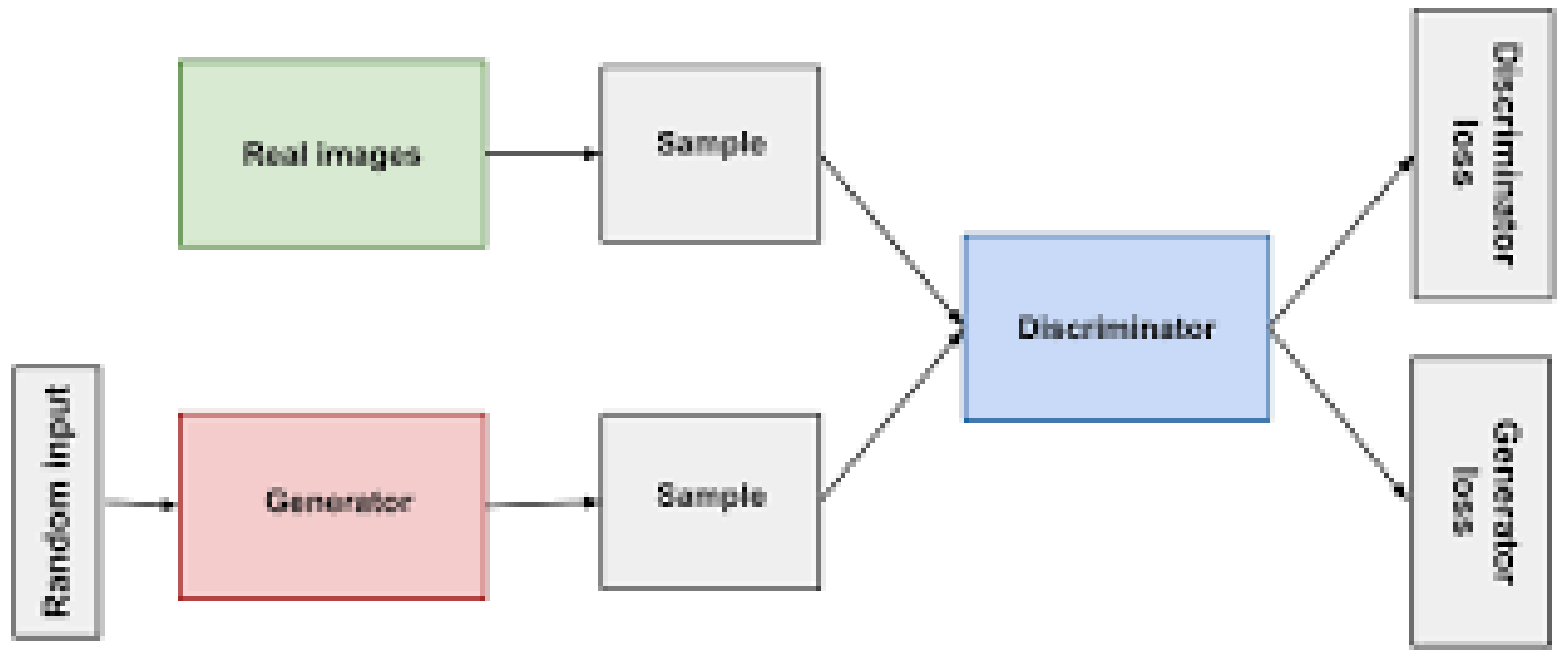

Figure 3

GAN model schematic [4]

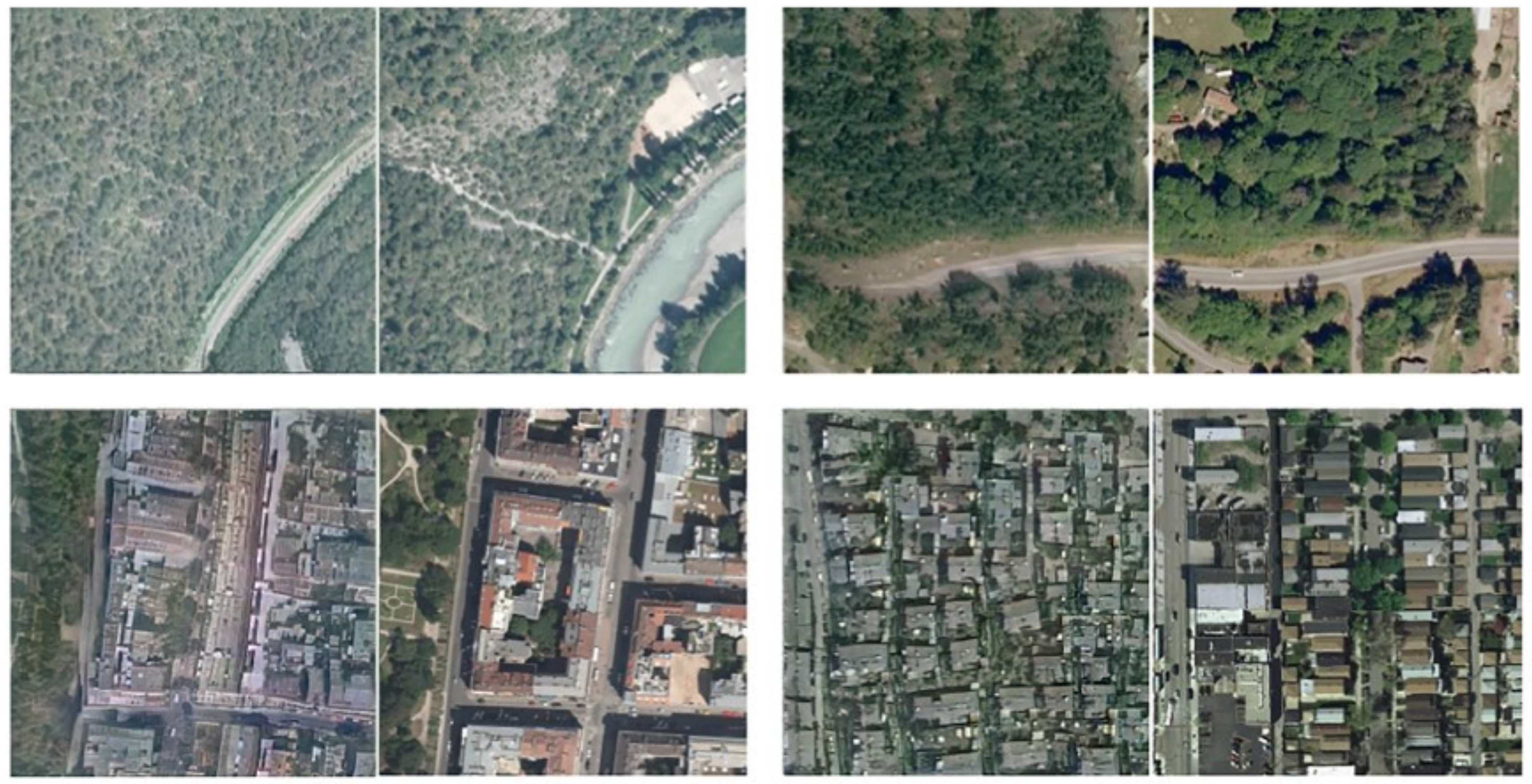

Figure 4 GAN Generated satellite images (Real: right), (associated latent-space image: left)[5]

## Image manipulation detection methods

There are several ways manipulated images can be detected these are as follows

- Image fingerprinting
- Statistical Analysis
- Computational Neural Networks (CNN)

### Bibliography

[1]Tucker, “The Newest AI-Enabled Weapon: ‘Deep-Faking’ Photos of the Earth,” Defense One, 2019. Available: https://www.defenseone.com/technology/2019/03/next-phase-ai-deep-faking-whole-world-and-china-ahead/155944/

[2]Na, “Adversarial Attacks in Machine Learning and How to Defend Against Them,” Medium, Dec. 19, 2019. Available: https://towardsdatascience.com/adversarial-attacks-in-machine-learning-and-how-to-defend-against-them-a2beed95f49c

[3]F. F. Niloy, “Dataset,” dash-lab.github.io, Jun. 20, 2023. Available: https://dash-lab.github.io/Dataset/. [Accessed: Mar. 24, 2024]

[4]Google, “Overview of GAN Structure | Generative Adversarial Networks,” Google Developers, 2019. Available: https://developers.google.com/machine-learning/gan/gan_structure

[5]M. Yates, G. Hart, R. Houghton, M. Torres Torres, and M. Pound, “Evaluation of synthetic aerial imagery using unconditional generative adversarial networks,” ISPRS Journal of Photogrammetry and Remote Sensing, vol. 190, pp. 231–251, Aug. 2022, doi: https://doi.org/10.1016/j.isprsjprs.2022.06.010