# TRIBHUVAN UNIVERSITY
# INSTITUTE OF ENGINEERING

**Khwopa College Of Engineering**
Libali, Bhaktapur
**Department of Computer Engineering**


A PROPOSAL ON

## DEEP FAKE DETECTION

*Submitted in partial fulfillment of the requirements for the degree*

**BACHELOR OF COMPUTER ENGINEERING**

Submitted by

| | |
|---|---|
| Manish Pyakurel | KCE/077/BCT/020 |
| Rupak Neupane | KCE/077/BCT/028 |
| Sarjyant Shrestha | KCE/077/BCT/033 |
| Srijan Gyawali | KCE/077/BCT/036 |

**Under the Supervision of**
Er.Dinesh Gothe
Department Of Computer Engineering

**Khwopa College Of Engineering**
Libali, Bhaktapur
2023-00

Introduction

# 1 Background Introduction

In recent years, the landscape of digital image manipulation has undergone a transformative shift with the emergence of deep fake techniques. This innovative approach, rooted in deep learning methodologies, has gained significant traction as a means of fabricating images by seamlessly replacing facial features from one individual with those of another. Coined as "deep fakes" by a Reddit user in 2017, these manipulations often leverage advanced adversarial models, such as Generative Adversarial Networks (GANs). Notably, this technology has been controversially utilized to superimpose celebrity faces onto explicit content, raising concerns related to fake pornography, misinformation, financial fraud, and hoaxes. Despite the ethical challenges associated with deep fakes, it is essential to acknowledge the positive applications within fields such as virtual reality, film editing, and production. The core working principles behind deep fakes involve intricate processes of merging, replacing, combining, and superimposing images. Leveraging deep learning and machine learning techniques, these manipulations give rise to convincingly altered digital images and videos, demonstrating both the potential benefits and ethical considerations associated with this rapidly advancing technology.

# 2 Problem Statement

In the rapidly evolving landscape of computer and automation technologies, the realm of possibilities continues to expand. Artificial Intelligence (AI) stands as a pivotal force, driving unprecedented advancements in areas such as predictive analytics, weather forecasting, automation, and the creation of sophisticated entities like deep fakes, which encompass AI-generated videos, audios, and images. While these technological strides are undeniably transformative, the misuse and exploitation of such capabilities pose significant concerns.

In recent times, there's been a surge in the creation of deep fakes, where the faces of celebrities or ordinary people are manipulated using just a single image and advanced deep learning algorithms. This issue is becoming more significant, as it circulates potentially harmful and illegal images of the victims to the public.

The rise of these deceptive practices not only threatens individual privacy but also has broader implications for public trust and safety. As deep fakes become increasingly convincing, the potential for malicious use, misinformation, and damage to reputations grows. It is crucial to address this issue head-on by developing sophisticated detection mechanisms to safeguard against the harmful consequences of manipulated images. This proposal seeks to contribute to the ongoing efforts in mitigating the risks associated with deep fakes, reinforcing the integrity of visual content in the age of advanced AI technologies.

# 3 Objective

The main aim of this project is:

- To identify manipulated digital media content, particularly facial features and images.
- To implement cutting-edge deep learning and machine learning techniques.

Literature Review

Deep fakes, which involve the unauthorized swapping of face images, are frequently carried out without the knowledge or consent of individuals, including celebrities and politicians. Notably, historical instances, such as the facial image swapping in a photograph of Abraham Lincoln (Badale et al., 2018), underscore the longstanding nature of this challenge. Addressing these concerns, Yang, Li & Lyu (2019) proposed a model leveraging head pose inconsistency to detect deep fakes, enabling the creation of synthetic faces for various individuals while preserving the original facial expressions. Jagdale & Shah (2019) introduced the NA-VSR algorithm for super resolution, involving video conversion into frames, median filtering to remove noise, and bicubic interpolation for pixel density enhancement. Additionally, Yadav & Salmani (2019) elucidated the working principles of deep fake techniques, emphasizing the high precision value in face image swapping. Generative Adversarial Neural Networks (GANs) play a pivotal role in deep fake generation, comprising a generator and a discriminator. The generator synthesizes fake images from a given dataset, while the discriminator evaluates the authenticity of the generated images. The inherent risks of deep fakes, including character defamation, potential harm to individuals, and the dissemination of fake news in society, highlight the importance of addressing these challenges. Existing approaches encounter issues such as inefficiency in detecting deep fake images, high error rates, prolonged computation times, and data access inaccuracies. This work, FF-LBPH-DBN, focuses on minimizing computational complexity while efficiently applying various metrological parameters. Table 1 presents a survey-based overview of existing approaches for detecting fake images (Vivek et al., 2018).

Methodology

# 4 SOFTWARE DEVELOPMENT APPROACH

Prototyping model is a type of software develpoment model. It is an iterative approach where a basic prototype is constructed to gain a better understanding of the project. This prototype is typically incomplete or lacking many components. The model is then refined based on feedback and system is reconstructed iteratively until desired conditions are met.

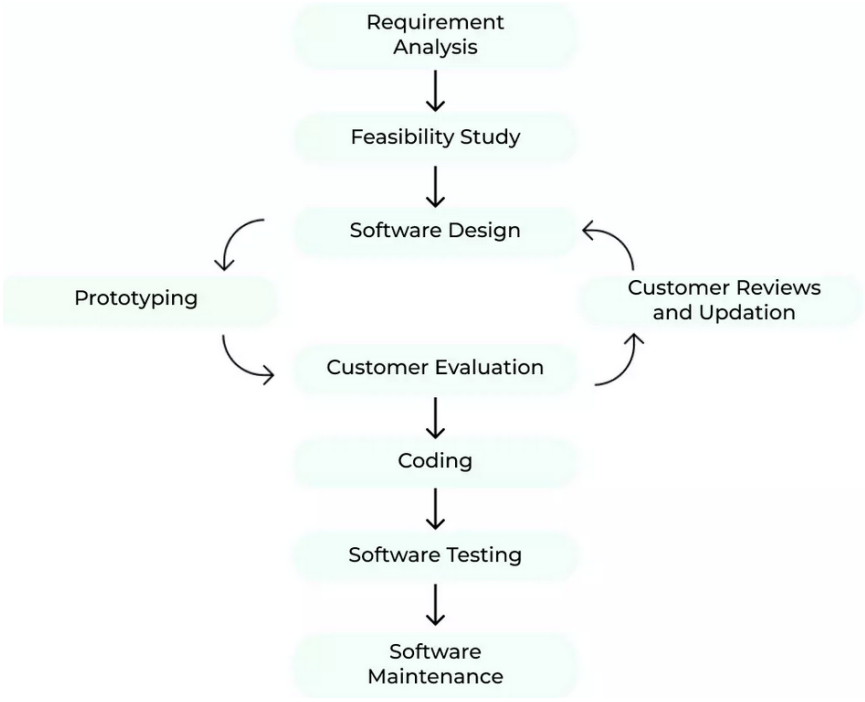


Figure 1: Prototype Model for Software Development

# 5 DATA COLLECTION

We have found many datasets on the internet from popular platforms like kaggle, github. For this project we will be using the datasets provided by ondyari/ FaceForensics `https://github.com/ondyari/FaceForensics`.
For POS tagging, the data from NELRALEC [10] The following table shows the amount of data collected from different sources and their usage in our project;