TITLE SELECTION REPORT

Degree	B.Tech
Department	Computer Science and Engineering
Year	IV
Section	A
Team Members	
Shiyam Saran S (122012012830)	
Rahman N (122012012822)	
Sathesh Kumar B (122012012879)	

Submitted to Guide,
Dr.R.Poonguzhali
Head of the Deparment(CSE)

Project Title 1:

AI Model to Detect Deepfake Images

1.Introduction:

In today's digital world, it's becoming incredibly difficult to tell what's real and what's fake. With the rise of AI tools, creating fake but realistic images what we call deepfake is getting easier every day. This is a huge problem. It's not just about misinformation; it's a real security threat that can be used for malicious purposes, like creating fake evidence. Our project is all about fighting back against this. We want to build a AI model that can reliably spot these fakes, helping to keep our digital media honest and authentic.

2.Area/Domain:

Our project falls under the domains of **Artificial Intelligence (AI)**, **Computer Vision**, and **Web Application Development**. It specifically focuses on deep learning for image classification and building a user interface for multimedia forensics.

3.Abstract:

We are proposing a project to build a reliable deepfake detection system, complete with a user-friendly website. The idea is simple: a user uploads an image to our site, and our AI model, which uses a powerful technology called a **Convolutional Neural Network (CNN)**, will quickly analyze it and tell them if it's real or fake. This isn't about building a model from scratch. We'll be using a smart technique called **"transfer learning"** to fine-tune a pre-trained model, which is much more efficient. Our research suggests that AI-generated images have tiny, hidden flaws that are invisible to us but obvious to a well-trained model. We'll leverage these imperfections to provide an accurate and trustworthy result. Ultimately, our goal is to put a powerful tool in the hands of ordinary people to help them navigate an increasingly deceptive digital landscape.

5.Objectives:

- ➢ To create a robust, automated system that can reliably detect and identify GAN-generated images.

- ➢ To design and build a simple and intuitive web interface that allows any user to upload an image for analysis.

- ➢ To integrate our deep learning model seamlessly into the website to process the uploaded image and provide a classification result either real or fake.

- ➢ To ensure our model is trained on a diverse dataset, so it can detect fakes from all sorts of AI generators, not just the ones we trained it on.

- ➢ To present the final result in a clear, easy-to-understand way for the user.

6.Methodology with Diagram:

- ➢ **Web Interface (UI) Development:** We'll build a clean and straightforward website. It will have a clear area for a user to upload their image. While our model is working its magic, a loading bar or animation will let the user know what's happening. Finally, a dedicated section will clearly show the final verdict: "Real Image" or "Fake Image."

- ➢ **Backend Model Integration:** When a user uploads an image, our server will take over. First, it will prepare the image for our model by resizing it and normalizing its data. Then, this processed image will be fed directly into our powerful deep learning model.

- ➢ **Model Architecture and Training:**

  - ▪ **Model Choice:** We've learned that a specific type of CNN called EfficientNetB4 is incredibly effective at this task, even achieving 100% accuracy on a test dataset . We'll use this as our primary model.

  - ▪ **Transfer Learning:** Instead of training from scratch, we'll use a model that has already learned to recognize patterns from millions of images on a dataset like ImageNet. We'll simply retrain its final layer to specialize in detecting our two categories: "real" and "fake".

  - ▪ **Dataset:** We'll use a unique dataset called Real or Synthetic Images (RSI). It includes images from different AI models and three different types of generation tasks (image-to-image, sketch-to-image, and text-to-image). This diversity is key to ensuring our model can handle fakes from anywhere.
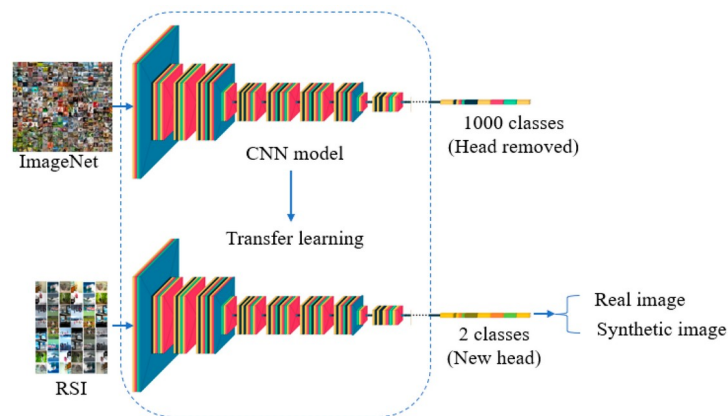
- ➢ **Result Delivery:** As soon as our model makes a decision, the result will be sent back to the website. The user will see a clear, bold message, like This image is **"REAL"** or This image is **"FAKE!"**, possibly with a confidence score to back it up.

7.Expected Outcomes:

- ➢ A functional web application that is easy for anyone to use.
- ➢ A highly accurate AI model for deepfake detection, with strong performance even on fakes it hasn't seen before.
- ➢ A practical tool that can serve as a first line of defense against manipulated digital media and help users verify the authenticity of images they encounter online.