SAI KUNDAN SUDDAPALLI

Beyond the Veil: Piercing the Facade of Deepfakes

ABSTRACT :

Deepfakes, synthetic media manipulated to depict nonexistent events, pose a significant challenge. This work investigates a novel deep learning architecture, codenamed Argus (after the multi-eyed giant symbolizing vigilance), for deepfake detection. Argus leverages the strengths of a pre-trained VGG16 model for feature extraction and a Vision Transformer (ViT) for classification. This combination captures both spatial features and long-range dependencies in images, aiming for robust deep fake identification.

The Karggel Faces dataset, known for its deep fake content, serves as the training and evaluation ground. To assess Argus's generalizability and robustness, we compare its performance against three adversarial attacks: Patch Attack, DeepFool, and Carlini-Wagner Attack (C&W). Furthermore, we explore the potential of cutting-edge Vision Large Language Models (Vision LLMs) like ChatGPT Vision ,Gemini Vision Pro,ViViT (Visual ViT), ALIGN (Aligner), and BEiT (Beijing Tiny Image Transformer). These models are trained on massive image datasets and may possess capabilities complementary to deep learning architectures. We investigate how integrating these Vision LLMs into the analysis pipeline, potentially implemented using TensorFlow, might enhance deep face detection.

The findings of this research will contribute to the development of more robust deepfake detection systems. We will shed light on the potential of Vision LLMs in this domain and explore how deep learning and LLMs can be combined for more comprehensive deepfake analysis.

Keywords: Deepfakes, Deep Learning, VGG16, Vision Transformer, Adversarial Attacks, Patch Attack, DeepFool, Carlini-Wagner Attack, Vision Large Language Models, ViViT, ALIGN, BEiT, Deep Face Detection, TensorFlow.