ABSTRACT

With the advancement of artificial intelligence (AI) technology, digital media manipulation has reached alarming levels, especially with the rise of deepfakes. Deepfakes, created using techniques such as Generative Adversarial Networks (GANs), can produce fake videos that are highly realistic and difficult to distinguish from genuine footage. This threat becomes serious in a security context, especially in CCTV systems which are often used for surveillance and criminal investigations.

This research aims to develop a deepfake detection method on CCTV footage using a deep learning GANs approach. The dataset used consists of original images taken from CCTV footage and deepfake images created manually using the Roop tool. The GANs model that was built was tested and compared with comparison methods such as Convolutional Neural Networks (CNN) using the VGG16, ResNet50, and InceptionV3 models.

Test results show that GANs have an accuracy of 61% and an F1-Score of 69% for deepfake data, while the VGG16 CNN model achieves an accuracy of 63% and an F1-Score of 44% for deepfake data, but is superior in classifying real data with a Recall value of 97 % and F1-Score 72%. This shows that although GANs are better at detecting deepfakes, the VGG16 model CNN is more effective at identifying genuine images.

Keywords: CCTV, Deepfake, Generative Adversarial Networks, Manipulation, Machine Learning.