Sanjana. S
S7- CSE B

Classification of Real and Fake images using One-Class Variational Encoder

Deepfake videos are AI generated videos that look real but are actually fake. Using such videos and images it is easy for malicious abusers to create arbitrary fake news and fool and mislead the public. Generally Binary classification methods are used for classification of real and fake images. But it requires large dataset of real and fake images in advance. When new deepfake generation methods are introduced only little deepfake dataset are available for training the model. So the output of such models won't be accurate. So here one-class variational encoder is used as it needs only real images for training.

One class variational autoencoder consist of encoder and decoder. At Encoder side image is given as input and scaling is done using convolutional layers and applying batch normalization with ReLu activation. A distribution is returned by this encoder and latent space reparameterization is done to that distribution. Then by finding mean and variance z is calculated. This z is given as input to decoder and application of convdutional layers and batch normalization is done and image is reconstructed. RMSE value is calculated. If it is real image, its RMSE values will be low and if it is fake, RMSE values will be high.

Our project is also to detect Deepfakes. In our project we are using Recurrent Neural Network for Deepfake video detection. Like one-class variational encoder, ~~here~~ in our project we are using two-class encoder. CNN features are used both in our project and this literature survey paper.