**Preprocessing**

Videos are preprocessed to have a fixed number of samples and resolution, there is also the option to use 3 color channels (RGB) or have it in gray scale.

**Model Training**

After preprocessing, the sampled and normalized images are fed into a CNN network to extract features from the images, finally these features are fed to RNN layers to process the sequential relations, and at the end there is a fully connected layer followed by one neuron in the output layer that outputs the probability of deception in the input video.

In this approach, feature extraction is done in the CNN layers.