Kermit -> Random Forest 1fps classification

Feature 1: Hu Moments

Feature 2: Haralick Texture

Feature 3: Color Histogram

(Audio) Feature 4: Librosa feature extraction <https://www.kaggle.com/maghili/urban-sound-classification-using-randomforest>

Ms. Piggy -> DNN with Keras:

(Audio) Librosa feature extraction <https://www.kaggle.com/maghili/urban-sound-classification-using-randomforest>

Use one of those: VGG, a resnet, an autoencoder

<https://elitedatascience.com/keras-tutorial-deep-learning-in-python>