**Flipkart Object Localization Challenge 2019**

**Team: Tensors**

**Sandeep Pattanayak | Swagat Das**

**Indian Institute of Technology, Kharagpur**

We started off segregating the images into 2 categories: images which had the ground truth labels available and boxes which did not have then. We applied Sobel Filter and Canny Filter on the 14000 images, whose bounding box coordinates were given. Because we were using Keras for our purpose, and it does not contain support for IoU metric, we manually defined the metric. We then trained a number of models, that had approaches ranging from simple stacked convolutional layer models to Models having skip connections to models inspired from Fractal Architectures. Due to less number of images available to us, usage of deep models was not possible, and we stuck to shallower architectures. We tried transfer learning using the two set of images, but could not get good enough results.

After experimenting with a lot of architectures, we finalised 2 models. The first Model was a 3-layer convolutional layer model, followed by a Fully Connected layer. This model was trained on first 128 by 128 pixels Sobel images, 256 by 256 pixels Sobel images and finally on 128 by 128 pixeled Canny images. We got 3 predictions using this model. The other model was a Skip Connection model, that had mild residual connections, apart from having stacked Convolutions and FC layers. Predictions on this model was done using 128 by 128 pixeled Sobel images. Finally we ensembled predictions in a slightly complex manner, ensembling the first 3 predictions first using a specific set of weights, and then ensembled the resulting predictions with predictions obtained from the skip connection model. This prediction recorded an IoU of 0.826 on the unseen test dataset.