# Generative Adversarial Networks for Improving Imbalanced Classification Performance

Snehal Bhatia, Master of Science in Computer Science (Future Networked Systems)

University of Dublin, Trinity College, 2019

Supervisor: Prof. Rozenn Dahyot

The advent of deep learning has resulted in extremely powerful classification models which have led to significant progress in many domains of computer vision such as object recognition and image classification. However, most existing classifiers assume the underlying training dataset to be evenly distributed. In real-life datasets, it is often observed that some of the classes have far less quantity of data samples than the others. An example of this could be a dataset for animal detection collected from a wildlife sanctuary where very few animals belonging to 'endangered species' class were present. Due to the underrepresentation of the minority classes, the data samples belonging to them often pass as noise or outliers, or are ignored, ultimately resulting in their misclassification. More often than not, the correct detection of minority class instances is of the utmost importance. Therefore, we study the detrimental effects of class imbalance on the classification performance of the Hybrid CNN-SVM architecture, and aim to introduce measures to overcome this issue. We artificially introduce class imbalance in two benchmark datasets, FMNIST and CIFAR-10, and observe that the classification Accuracy and F1-Score both drop by an average of 7%, when compared to the performance on the original dataset. To combat this problem, we use data augmentation strategies to re-balance the datasets. We first explore the traditional data augmentation practices of applying geometric and photo-metric transformations on the existing images of minority classes, such as image rotation, scaling, zooming, blurring, whitening, shearing, etc. Then, we propose the use of a modified architecture of Generative Adversarial Networks (GAN), called Wasserstein GAN with Gradient Penalty (WGAN-GP) to generate new data samples. Since their introduction in 2014, GANs have shown immense potential in mimicking data distribution and generating realistic images using low amounts of training data. Training the classifier on the datasets augmented using WGAN-GP, we observe an average increase of 4% in the classification Accuracy and F1-Score for FMNIST dataset, and this is even more for the more complex CIFAR-10 dataset, where a 6.2% improvement is achieved. This is significantly better than the improvement achieved using the baseline method of dataset augmentation using image transformations, and it has proven to be a more promising solution for real-world datasets which are becoming increasingly complex and diverse.