

The amelioration of Deep Learning algorithms was a milestone in the history of Computer Vision, because for the first time in the history, a computer program was able to see better than human being. The Convolutional Neural Networks are known for giving better accuracy than humans for task like Image Classification and Object Detection. However, the problem with these networks is their need for data. The more data is given to a CNN, the better it will perform. Although, there are real world problems where there's impossible to get handful of data, to exemplify, If a small company wants to install a face detecting system for their employees, then the dataset would be very small for the training of the model. In such cases, it is vital to use some other techniques along with CNN to achieve good accuracy. In this research, the adverse effect of dataset size on the accuracy of the CNN classifier is studied and the aim is to find out the techniques which work best with the dataset. The datasets used here are MNIST, FMNIST and DogsVsCats (from Kaggle) are used. To verify the effects of fewer training example on the classifier, the MNIST and FMNIST dataset are used where their 10%, 50% and 100% subsets were taken. The accuracy increased with the rise in size for FMNIST, but not for MNIST, because of its trivial nature. Once this is done, two helper techniques were used to improve the classification accuracy that are: Data Augmentation and Transfer Learning. Both techniques are implemented on FMNIST and DogsVsCats. For Data Augmentation, Geometric transformation and Elastic distortion is used and an overall improvement of 2.92% on FMNIST and 16.67% on DogsVsCats was recorded. Furthermore, for Transfer Learning, MobileNetV2 and VGG19 models were used along with the CNN classifier. The accuracy for FMNIST dropped by 6.12% while DogsVsCats illustrated a rise of 24.04%. The ups and downs in the accuracy on different datasets demonstrated that the nature of dataset also plays important role in conjunction with its size. If the dataset is trivial and uncomplicated like MNIST or FMNIST, then simpler and straightforward algorithms tend to work better while the fancy algorithms result into overfitting. Complicated datasets call for advanced techniques, that's why DogsVsCats gives best results with Transfer Learning.