# Report - Fundamentals of Deep Learning for Computer Vision

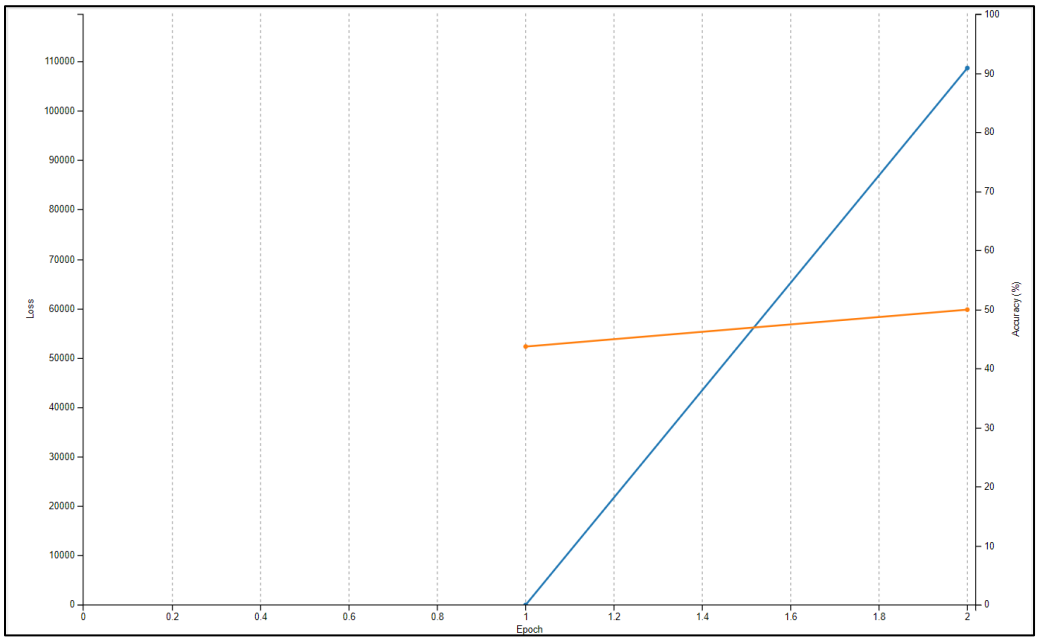
## Introduction

The Nvidia fundaments of deep learning for computer vision certification has a total of 5 tasks and 1 graded assessment. All the tasks are performed on the NVIDIA's DIGIT platform. It is a platform provided by NVIDIA which has multiple neural network models pre-designed for users to train.

## GPU TASK 1:

In this task, the objective is to load the dataset of images of beagles containing 8 labeled images of Louie and 8 labeled images of other dogs and classify whether the given image of the dog is Louie or not Louie. The dataset is very small for a deep neural network. AlexNet architecture is used to design the neural network for image classification. The experiment was performed with two sets of settings.

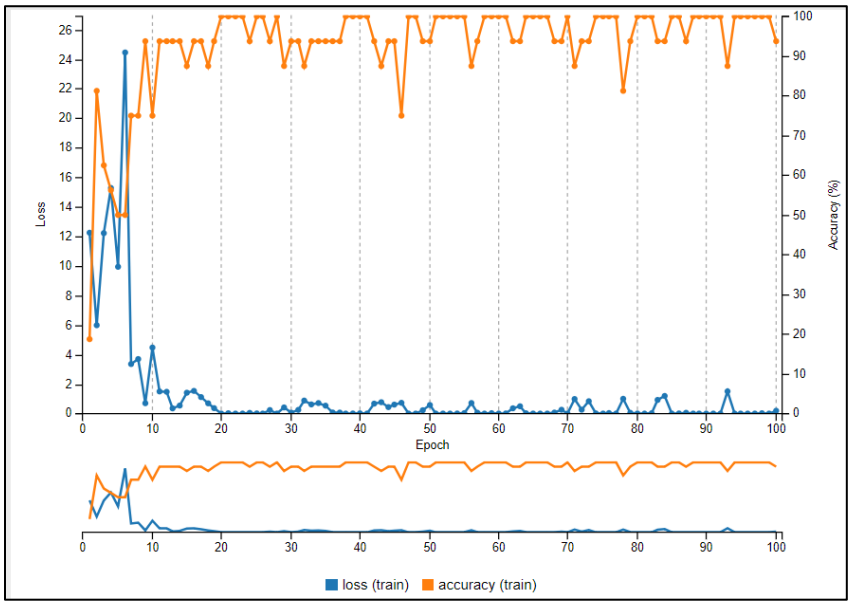
The first experiment was performed with AlexNet and less number of epochs and a faster learning rate. The model trained quickly but the accuracy for predicting Louie was very low(~48%).For the second experiment, the model was cloned and trained again with 100 epochs and slower learning rate. After training the prediction accuracy was 100%.




**GPU TASK 2:**

The second task is the image classification task where the objective is to classify dogs and cats. In the Louie example, our model was effective on data from training dataset, but not on out of sample data, rendering it is difficult in the real world.

In this task, we use a new dataset named Dogs vs Cats. This dataset contains images of both cats and dogs. We used of AlexNet which we used in the first dataset. Let's train the model to test the accuracy. In GPU task 2 the dataset contains 18750 labeled images of dogs and cats. The DIGITS is standardizing the images and again training them with AlexNet which is designed to



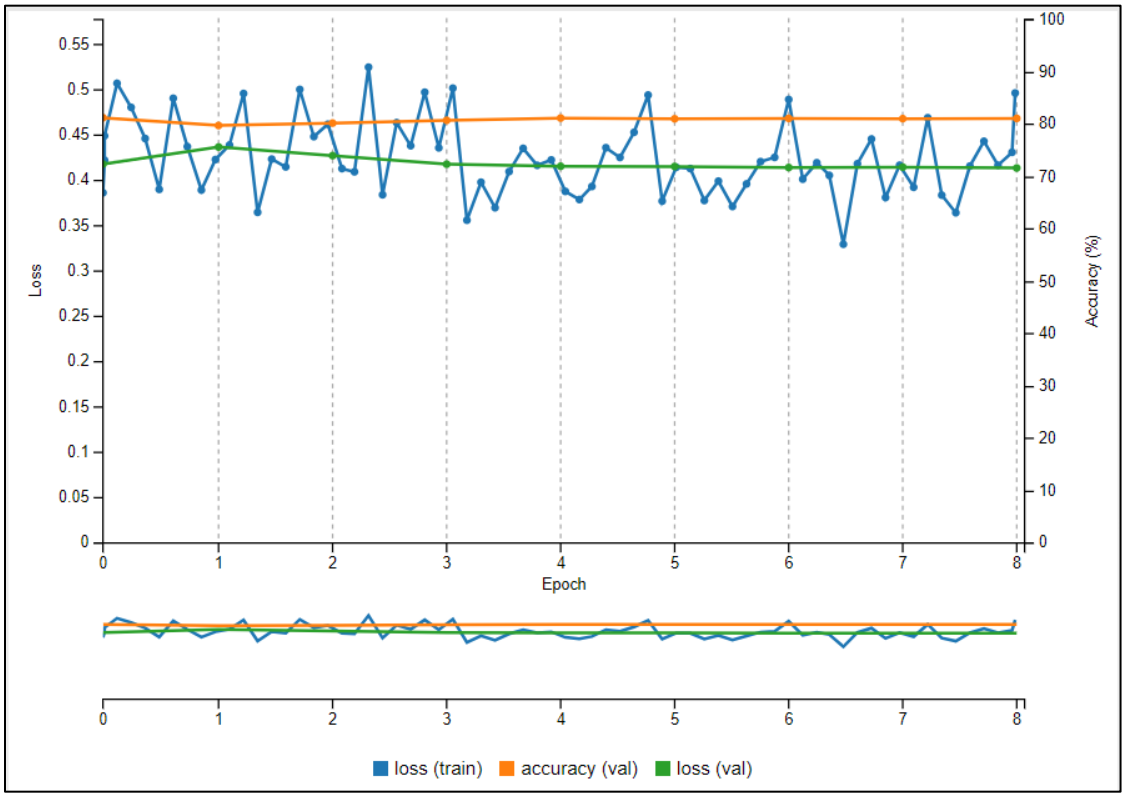

**GPU Task 3:**

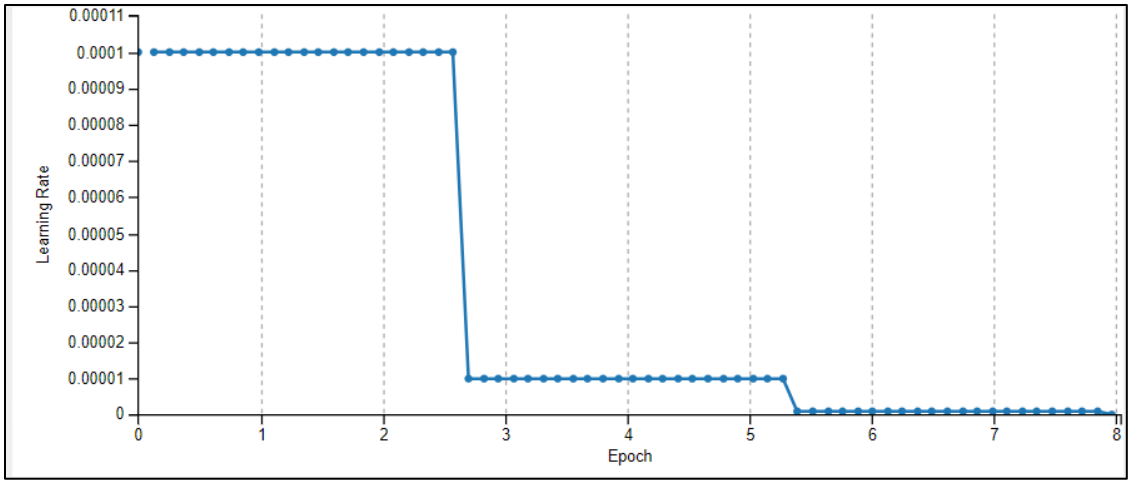
For the previous tasks we have trained a classification model on a smaller dataset and another model on larger dataset and with more number of epochs. The objective of this task is to deploy the model we trained previously. The model was used for deployment had weights and architecture of Dog vs cats model. We created a simulation to detect if a dog is present at the dog door or not.

## GPU TASK 4:

We have designed different models till task four. The models were designed and trained from scratch. But for the task 4 the objective is to use a pretrained model and we're further training it using alexnet. The model was trained for 16 hours previously

Model Training Summary:



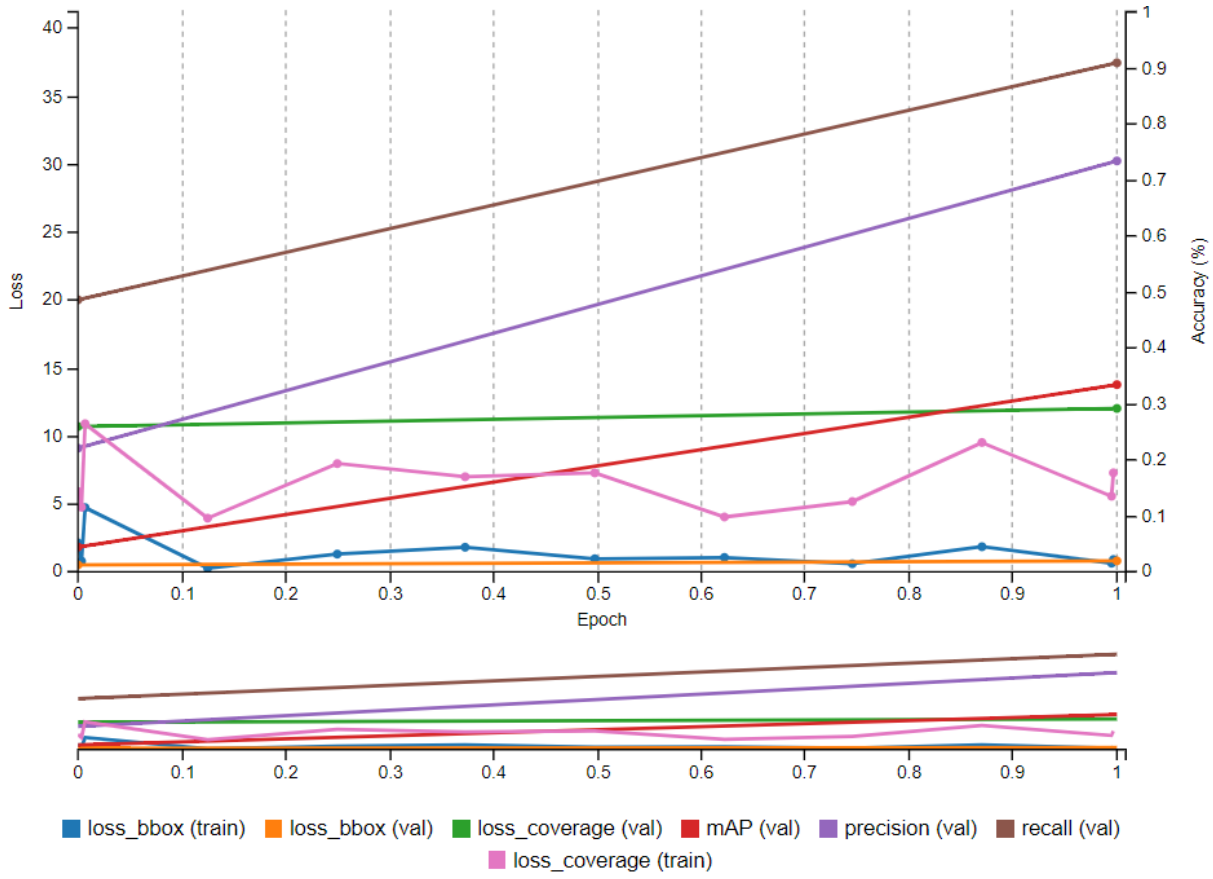



GPU TASK 5:

In this task, we used a 'sliding window' approach, where we took the small pieces of images which they called grid squares. Each grid square ran through an image classifier. The objective was to find Louie in the image. We combined deep learning with traditional computer vision and modified the internal of neural networks. We use a total of 3 approaches in this task to perform image detection. In the first approach, we use the traditional network. The second approach is rebuilding the existing neural network. The new network was built on the previous network. The new network was formed by changing the layer structure from previous AlexNet architecture. DetectNets are used in the third approach. When approach two and three are visualized we can say that the DetectNet is a Fully convolutional network which is configured to produce precise data representations. The bulk layers in DetectNet is identical to the GooGLeNet Network. The model was pretrained for 16 hours on NVIDIA Tesla K80. Since the model was well trained the DetectNets successfully detected the dog and drew a bounding box around it.

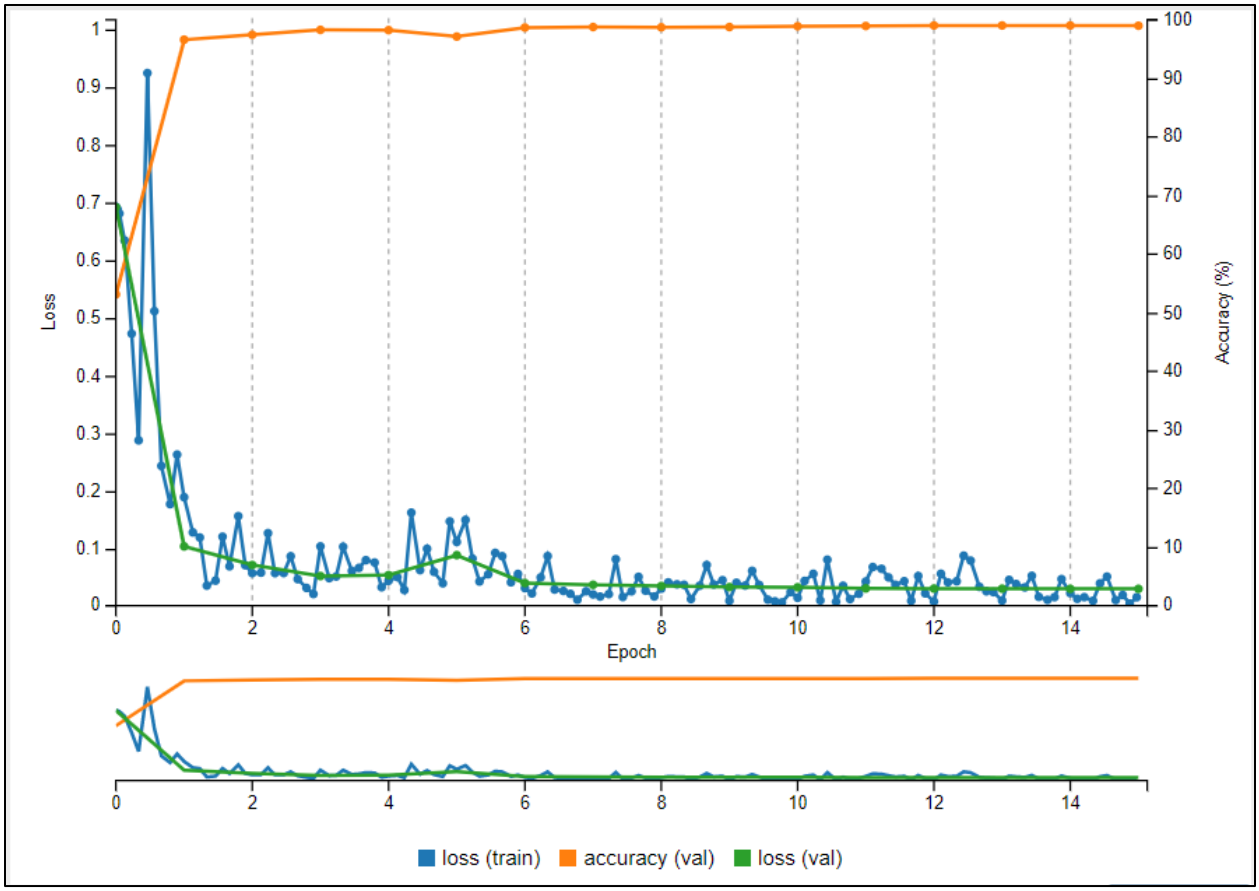
Output of Object Detection-

Model Training Summary




Assessment:

The assessment is the final Computer Vision certification program. The aim of the task is to classify the images of whales. The images from the dataset is given in whale and no-whale format. The first task is to load the dataset of whales. For the second task we are training the models with AlexNet. When the model is trained we're updating the weights of model. After updating weights and architecture we're testing the trained model on the given image to classify the image.




The output of Assessment

```
whale
```

```
not whale
```

The objective of the assessment was achieved to classify the image into whale or no-whale.