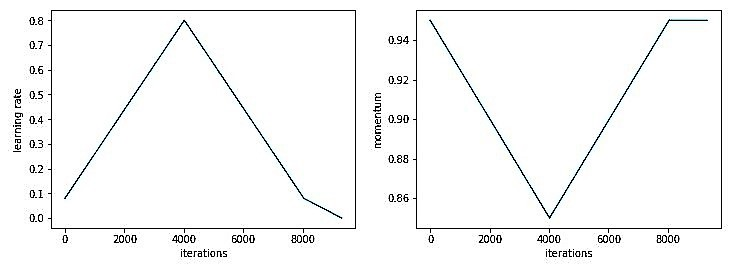
1)what is the concept of cyclical momentum ?

Ans : Cyclic learning rates and cyclic momentum which usually goes hand-in-hand is a learning rate scheduling technique for faster training of a network and a finer understanding of the optimal learning rate. Cyclic learning rates have an effect on the model training process known somewhat fancifully as "superconvergence".To apply cyclic learning rate and cyclic momentum to a run, begin by specifying a minimum and maximum learning rate and a minimum and maximum momentum. Over the course of a training run, the learning rate will be inversely scaled from its minimum to its maximum value and then back again, while the inverse will occur with the momentum. At the very end of training the learning rate will be reduced even further, an order of magnitude or two below the minimum learning rate, in order to squeeze out the last bit of convergence.



The maximum should be the value picked with a learning rate finder procedure, and the minimum value can be ten times lower.

One-cycle training

This cyclic learning rate policy is meant to be applied over one entire learning cycle: e.g. one epoch. Fast.AI calls this the one cycle training. After each cycle, you are supposed to re-apply the learning rate finder to find new good values, and then fit another cycle, until no more training occurs.Momentum is counterproductive when the learning rate is very high, which is why momentum is annealed in the opposite of the way in which the learning rate is annealed in the optimizer.Cyclic learning rate scheduling thus combines discovery of the maximum practical learning rate (given a certain batch size) with learning rate annealing. You can use cyclic learning to discover the maximum practical learning rate (too high and it will diverge, too low and it will not stabilize at the top), and then switch to adaptive learning rate with that learning rate.

2)what callback keeps track of hyperparameter value(along with other data)during training ?

Ans :A callback is a set of functions to be applied at given stages of the training procedure. You can use callbacks to get a view on internal states and statistics of the model during training.you want to automate some tasks after every training/epoch that help you have controls over the training process. This includes stopping training when you reach a certain accuracy/loss score, saving your model as a checkpoint after each successful epoch, adjusting the learning rates over time, and more.hyperparameter is a parameter whose value is used to control the learning process. By contrast, the values of other parameters (typically node weights) are derived via training.

3)In the color dim plot,what does one column pixels represent ?

Ans : An image is divided into pixels (Short for Picture Element ), each pixel represents the smallest single point on the screen. A pixel can only be one color at a time, and when a picture is saved, the color of each individual pixel must be stored in binary .digital color image pixel is just numbers representing a RGB data value (Red, Green, Blue). Each pixel's color sample has three numerical RGB components (Red, Green, Blue) to represent the color of that tiny pixel area. These three RGB components are three 8-bit numbers for each pixel.PIXEL DIMENSIONS are the horizontal and vertical measurements of an image expressed in pixels. The pixel dimensions may be determined by multiplying both the width and the height by the dpi.

5)does a batch normalization layer have any trainable parameters ?

Ans : Batch normalization can be implemented during training by calculating the mean and standard deviation of each input variable to a layer per mini-batch and using these statistics to perform the standardization.These parameters are used for re-scaling (γ) and shifting(β) of the vector containing values from the previous operations. These two are learnable parameters, during the training neural network ensures the optimal values of γ and β are used. That will enable the accurate normalization of each batch.

Two learnable parameters called beta and gamma.Two non-learnable parameters (Mean Moving Average and Variance Moving Average) are saved as part of the ‘state’ of the Batch Norm layer.

6)In batch normalization during Preparation,what statistics are used to normalize.?what about during the validation process ?

Ans : mean ,mode & standard deviation are used to normalize the data.the mean and standard deviation of inputs for the layer can be set as mean values observed over the training dataset.

we will call Inference mode, the mean and standard deviation are based on previous estimations, meaning the running average. This is usually used during validation and inference.

7)why do batch normalization layer help models generalize better ?

Ans : Batch normalization solves a major problem called internal covariate shift. It helps by making the data flowing between intermediate layers of the neural network look, this means you can use a higher learning rate. It has a regularizing effect which means you can often remove dropout.

8)explain between max pooling & Average pooling number eight.

Ans : Max pooling: The maximum pixel value of the batch is selected.

Max pooling selects the brighter pixels from the image. It is useful when the background of the image is dark and we are interested in only the lighter pixels of the image. For example: in MNIST dataset, the digits are represented in white color and the background is black. So, max pooling is used. Similarly, min pooling is used in the other way round.

Average pooling: The average value of all the pixels in the batch is selected.

Average pooling method smooths out the image and hence the sharp features may not be identified when this pooling method is used.

9)what is the purpose of the pooling layer ?

Ans : Pooling Layer The main purpose of pooling layer is to progressively reduce the spatial size of the input image, so that number of computations in the network are reduced. Pooling performs downsampling by reducing the size and sends only the important data to next layers in CNN.The purpose of pooling is mainly to down-sample the feature maps and to learn larger-scale image features that are invariant to small local transformations (e.g., translation, scaling, and rotation).

10)why do we end up with completely Connected layers ?

Ans : The output from the convolutional layers represents high-level features in the data. While that output could be flattened and connected to the output layer, adding a fully-connected layer is a cheap way of learning non-linear combinations of these features.

11)what do you mean by parameters ?

Ans : Batch Norm layer also has parameters of its own: Two learnable parameters called beta and gamma. Two non-learnable parameters (Mean Moving Average and Variance Moving Average) are saved as part of the 'state' of the Batch Norm layer.

12)what formulas are used to measure these parameters ?

Ans : Camera parameters include intrinsics, extrinsics, and distortion coefficients. To estimate the camera parameters, you need to have 3-D world points and their corresponding 2-D image points. You can get these correspondences using multiple images of a calibration pattern, such as a checkerboard.