1. Momentum:

The ~~Dropout~~ Momentum adds an average of the lasts steps. This should increase the speed if the model always takes the same step. It also helps to overcome local minimum by adding an average of the last steps. The added average "pushes" it over small slopes

Learningrate:

The learning rate determines how large a step is. If it is too high, it skips the optimum. If it is to small, it takes longer to reach the optimum.

Learning rate decay:

The above-mentioned problems should be avoided by decreasing the learning rate. The learningrate should be high at the beginning and the longer the model is trained, the lower the learning rate should be to prevent overshooting the optimum. This is achieved by using a learning rate decay.

2. Dropout:

With dropout regularization, some neurons are randomly set to zero during the forward pass. This helps the model to learn more generalizations and makes it less sensitiv to certain neurons, making the model less overfitting and better able to adapt to new, unseen data.

3. Filtersize:

= Output_size = (Input_size - Filter_size)/Stride + 1

= 5	= (10 - Filter_size)/2 + 1	1-1
= 4	= (10 - Filter_size)/2	1-2
= 3	= 10 - Filter_size	1-10
= -2	= - Filter_size	1-(-1)
= 2	= Filter_size	