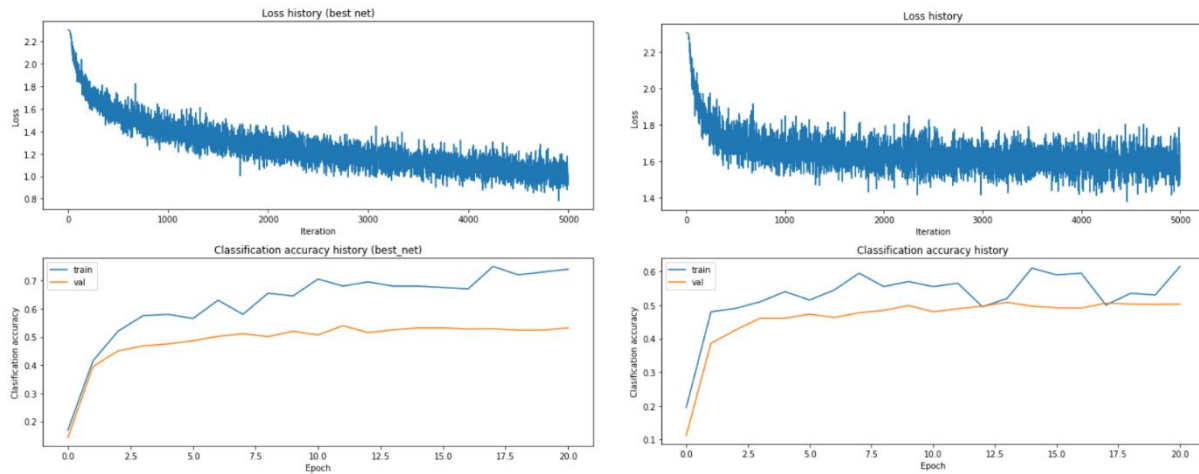


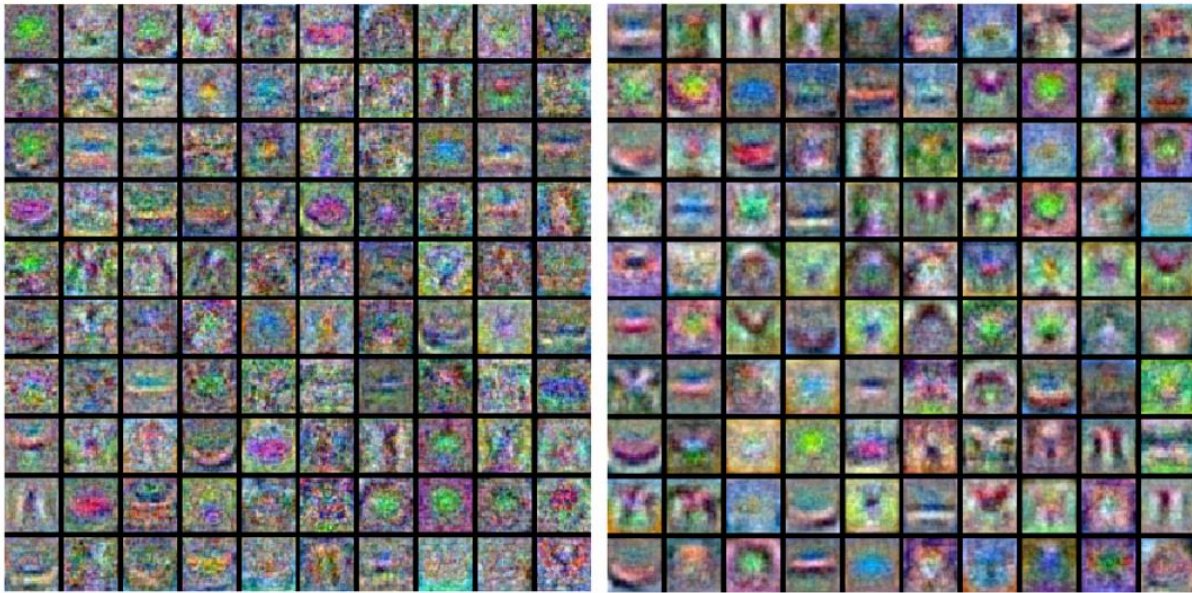
## Two layer net overfitting

In my first experiment, I achieved a classification accuracy of 54.5% on the validation set and 53% on the test set. The best hyperparameters for these results are: hidden\_size=100, learning\_rate=0.001, regularization\_strength=0.001. The plots with loss and classification accuracy during the training phase are represented in Figure 1 (left column).

The gap between validation and training accuracy (Figure1, left column) suggests that this model overfitted the train data. On the right column can be seen the classification accuracy history for a larger regularization strength (reg=1). A larger regularization term helped to avoid overfitting, but the accuracy decreased a little (validation accuracy: 51.4%, test accuracy: 50.9%).

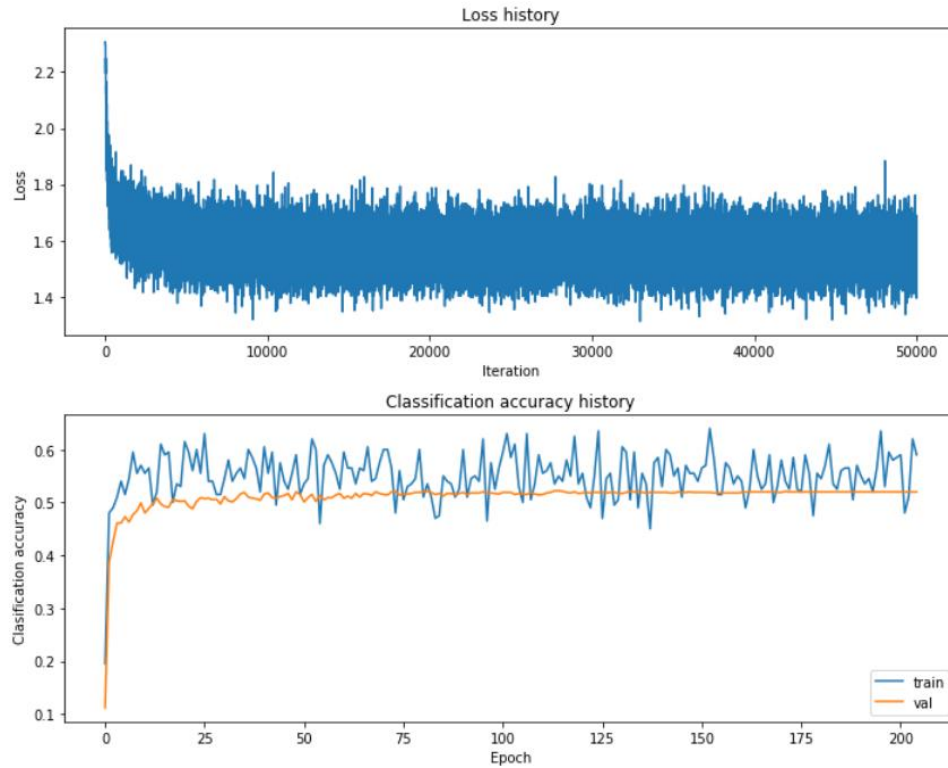


**Figure 1.** Loss and classification history. Left column: (hs=100, lr=0.001, reg=0.001), right column: (hs=100, lr=0.001, reg=1).



**Figure 2.** The weights for above models. Left column: (hs=100, lr=0.001, reg=0.001), right column: (hs=100, lr=0.001, reg=1). The left column, that corresponds to the model that overfitted the train data, is more noisy.

In the previous examples, I used for training 5 000 iterations with a batch size of 200 images. For the following experiment, I increased the number of iterations to 50 000 and retrained the model with a large regularization strength ( $\text{reg}=1$ ). The classification accuracy increased (validation accuracy: 52%, test accuracy: 53.5%), while the gap between the train and validation accuracy remained very small. The plots and weights for this model are represented in Figure 3 and Figure 4.



**Figure 3.** Loss and classification history for model: ( $\text{hs}=100$ ,  $\text{lr}=0.001$ ,  $\text{reg}=1$ ,  $\text{num\_iters}=50\,000$ ).



**Figure 4.** The weights for model: ( $\text{hs}=100$ ,  $\text{lr}=0.001$ ,  $\text{reg}=1$ ,  $\text{num\_iters}=50\,000$ ).

The quantitative results of my experiments are presented in the Table 1.

Hyperparameters of the models	Train accuracy	Validation accuracy	Test accuracy	Overfitting
(hs=100, lr=0.001, reg=0.001, num_iters=5 000).	65.2%	54.5%	53.0%	+
(hs=100, lr=0.001, reg=1, num_iters=5 000).	52.7%	51.4%	50.9%	-
(hs=100, lr=0.001, reg=1, num_iters=50 000).	55.2%	52.0%	53.5%	-

**Table 1.** Results of my experiments.