1

$$\tilde{x} = \frac{x - \mu}{\sigma}$$

μ x σ x

2

2.1

1. 12 4
2. 2 4 4
3. Softmax 11

ReLU Softmax $\{\hat{y}_i\}_{i=0}^{i=11}$ $\hat{y}_i$ i i $prediction = \arg_i \max \hat{y}_i$ onehot

$$L = \sum_i y_i \log(\hat{y}_i)$$

y_i onehot $y_i = 1$
mini batch batch 256 1000 4×10^{-4}

3

3.1

10% 52% **??**

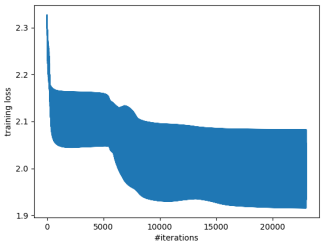

Figure 1: