



A pair of random images from the training set is fed into a CNN.

**Augmentation Network**

Augmented and target image are compared to generate a loss (or none)



**Classification Network**

Second network generates a classification loss

$$\begin{bmatrix} 2 \\ 5 \\ 1 \\ 8 \end{bmatrix}$$