Overfitting

$\vec{y} \xleftarrow{f} \vec{x}$

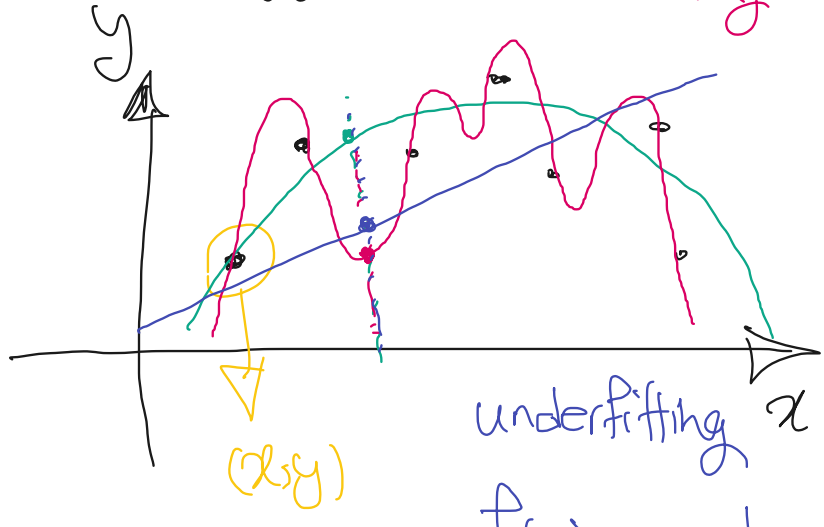
$\{\vec{x}, \hat{\vec{y}}\}_{i=1\cdots N}$

$\downarrow$

$\underset{w}{\operatorname{argmax}} L$

poly deg 6+

overfitting



$y = f(x) = ax + b$

$\{x_i, \hat{y}_i\}$

$\downarrow$

f

$\downarrow$

$x \to y$

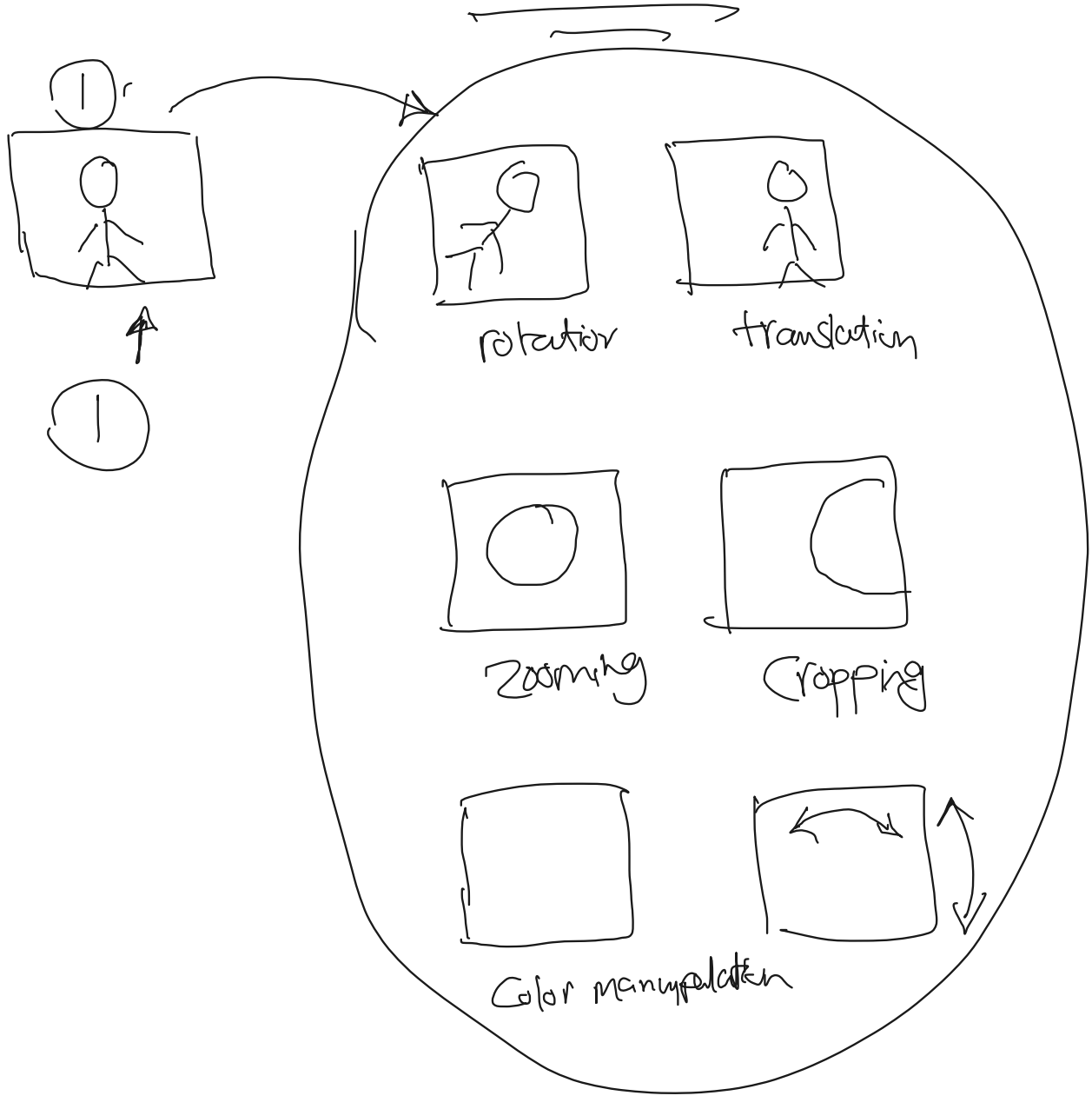
How to avoid overfitting:

1. avoid <u>too</u> complex models

(2) increase the # of samples

- collect/capture more data

- "data augmentation"

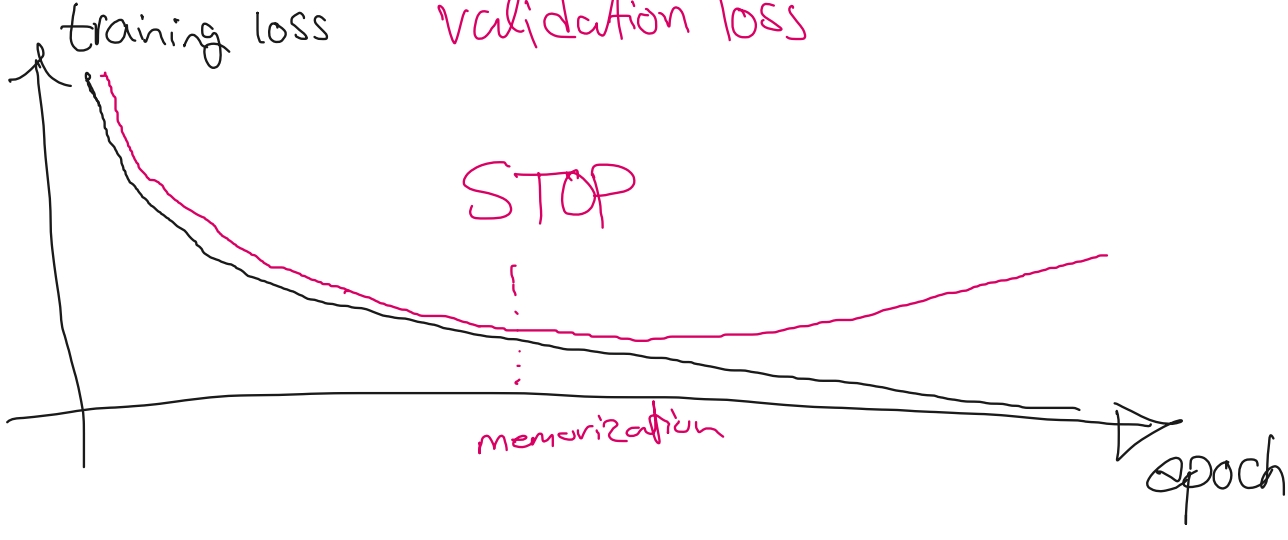


(3) adding "regularization term"

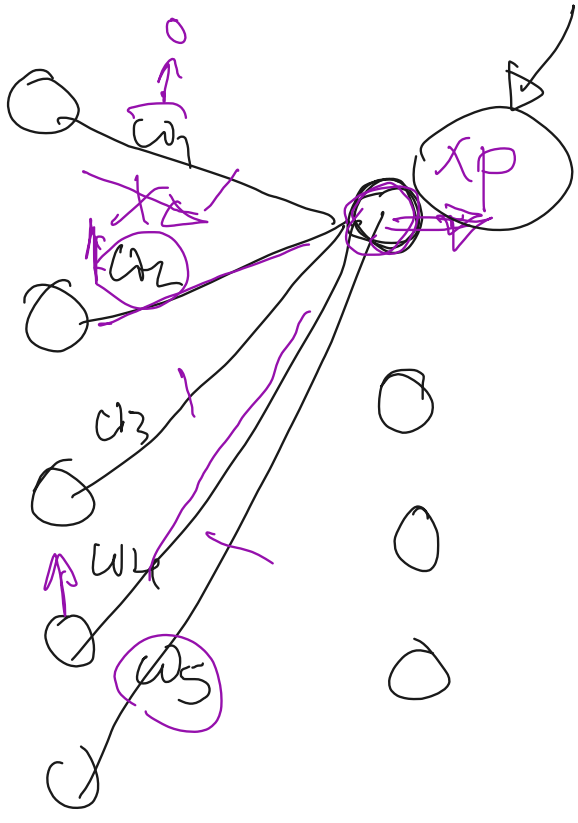
$$L = \sum_{i=1}^{N} \|\vec{y} - \hat{\vec{y}}\|_2 + \lambda \|\vec{w}\|_2$$

ignore bias

(4) early stopping



(5) Dropout



dropout

Training

no-dropout

Testing

normalization

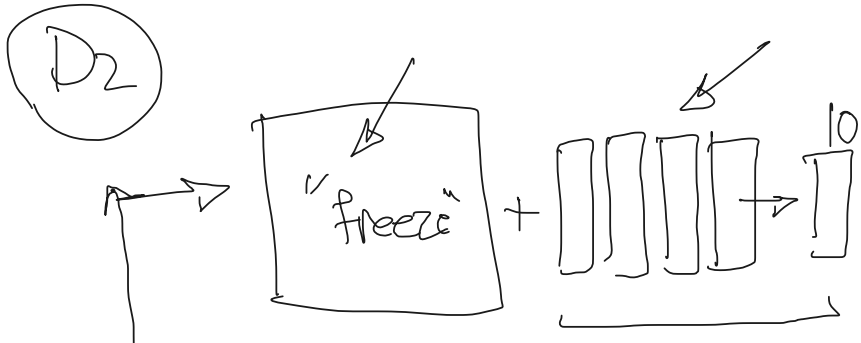
all layers except input $p = 0.5$

input $p \geq 1$

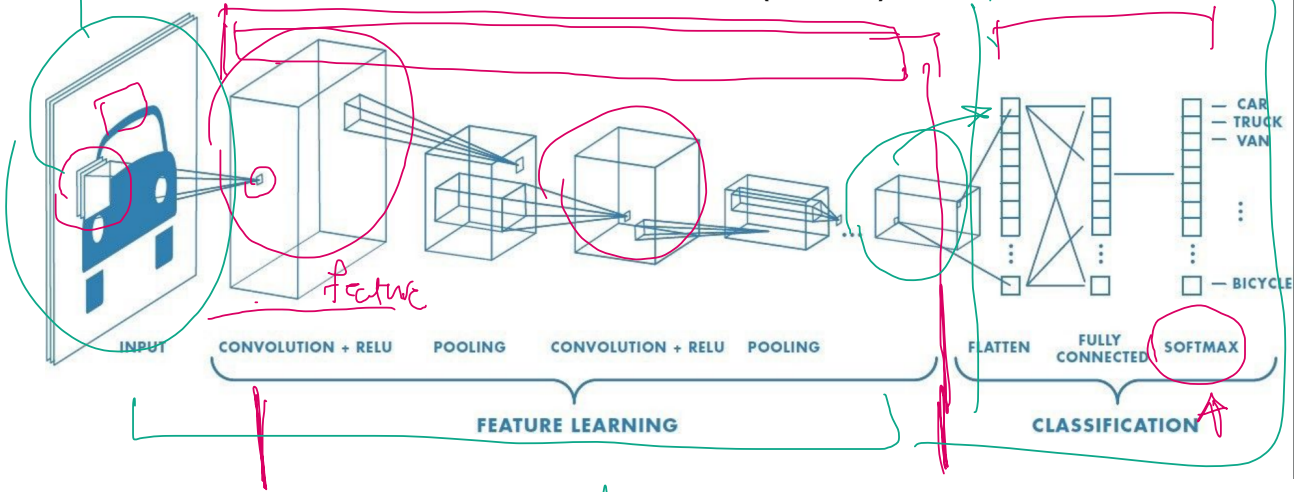
Transfer Learning

Model A { D_1, T_1 → Segmen

Model B { D_2, T_2 → classif



Convolutional Neural Network (CNN)



general

specific

imagenet

1000 ← VGG

AlexNet

← ResNet

representation learning