

Summary, a practical guide to deep network design

© 2019 Philipp Krähenbühl and Chao-Yuan Wu

Loss functions

Continuous output?

N

Y

Classification

Regression

2 classes

>2 classes

BCEWithLogitsLoss

CrossEntropyLoss

L1Loss / MSELoss

Activation functions

- Try ReLU
- If ReLU fails, try:
 - Leaky ReLU / PReLU

SGD

- Select batch size
- Tune learning rate
- momentum = 0.9

