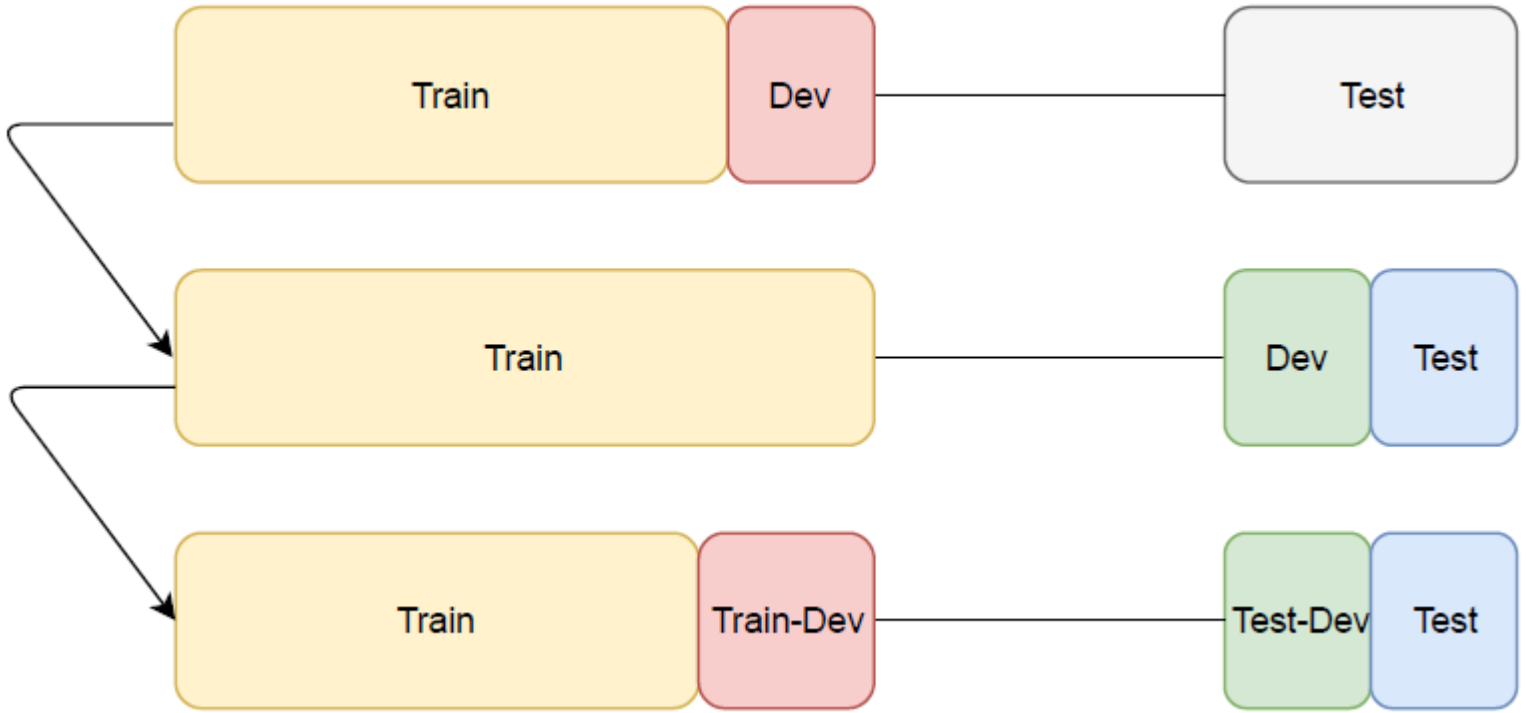
Lecture 4-5 Nuts and Bolts of building AI applications using Deep Learning

Source: Andrew Ng, Nuts and bolts of building AI applications using Deep Learning

Bias Variance Tradeoff

- How to split your data
  - If all your train and test data comes from same distributions, that's ok.
  - If large train data is from many places, and small test data is actual car data, split data can be tricky
  - Rule: dev and test to come from the same distribution

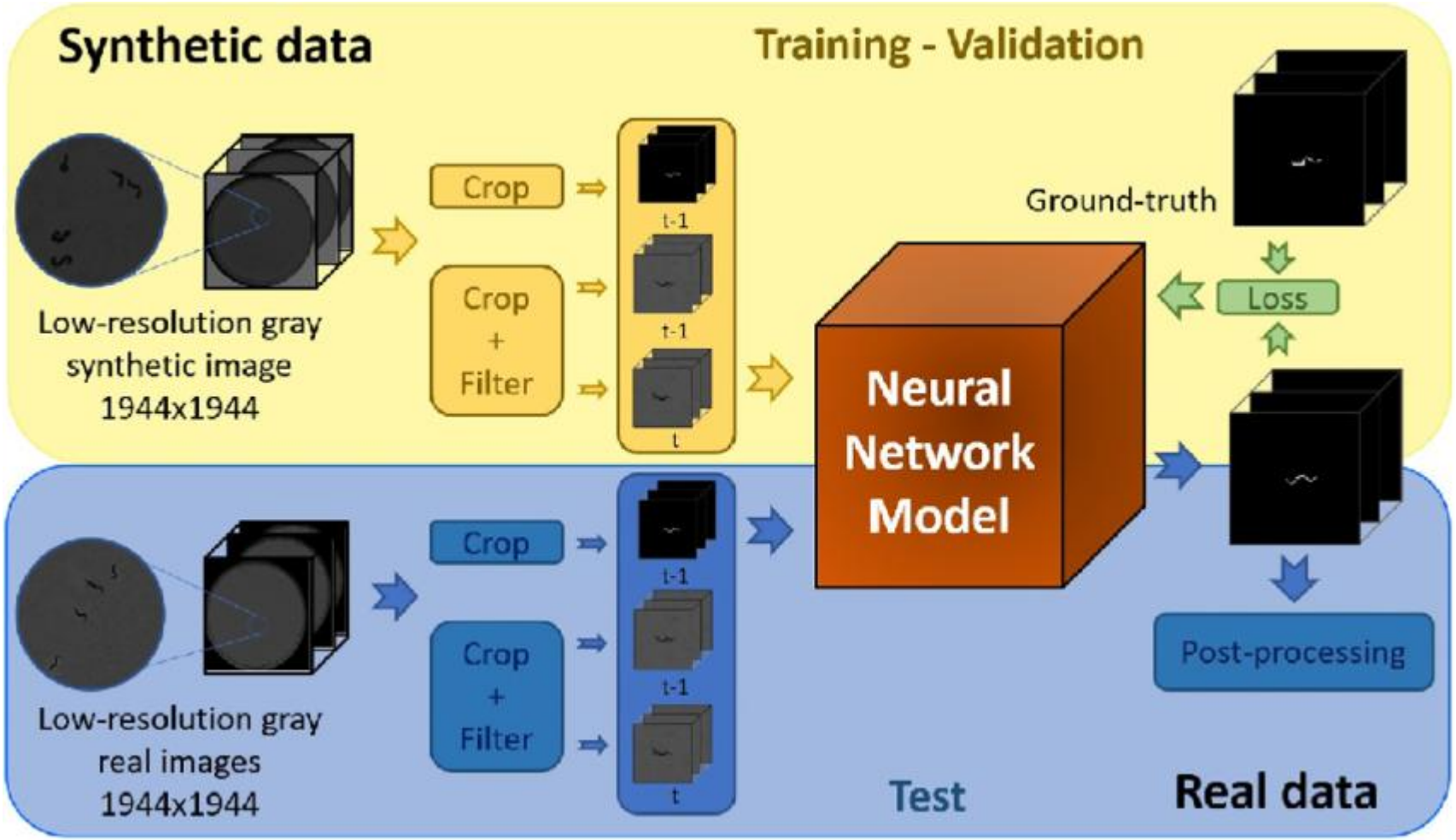


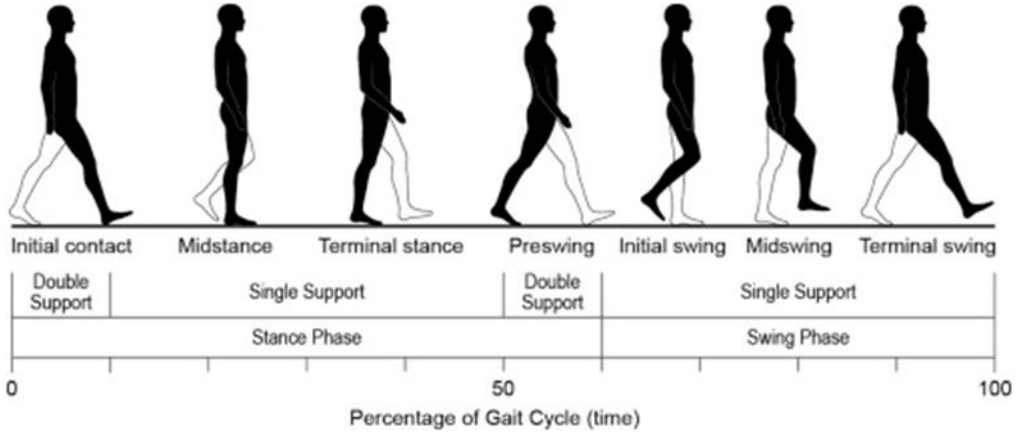
Validation 只從 大的training data來
Possible mismatch

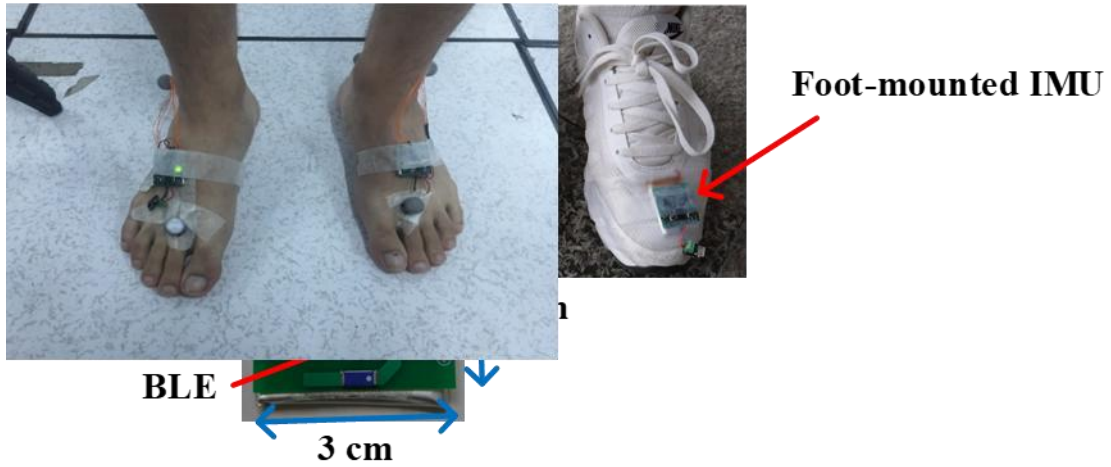
Validation 只從 小的testing data來
Possible mismatch

Validation 從 大的training data
和小的testing data各拿一部分

A Bad Example



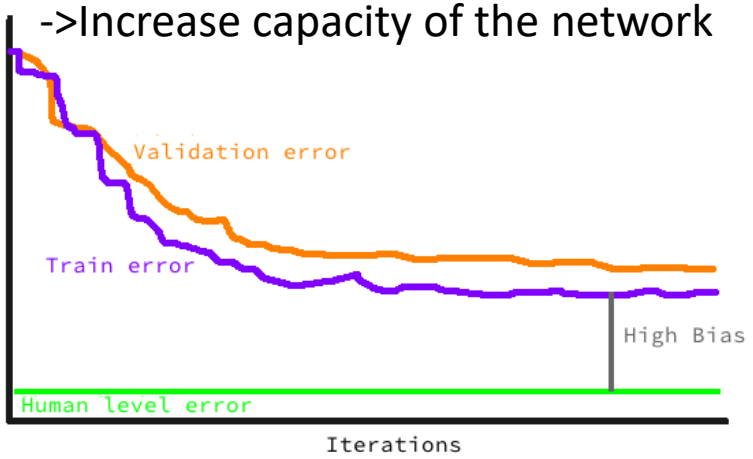




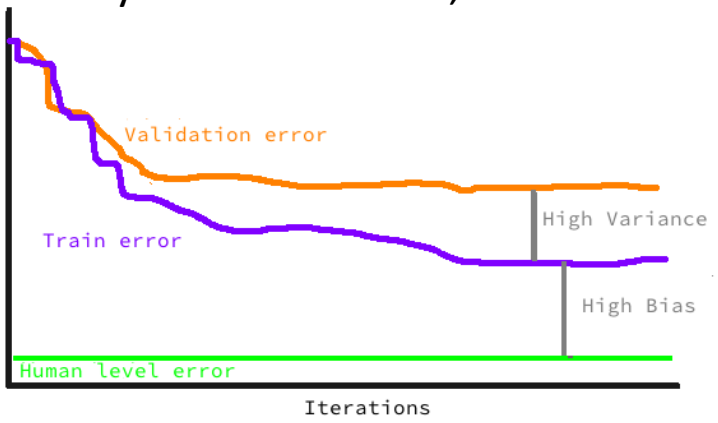
Test on P1	Sampling rate	Result	w/o	w/ 2.9%	w/ 3.4%	w/ 4.3%	Improved
S.L.	500Hz	Percentage error(%)	10.7	4.6	3.4	2.5	8.2
		Error(cm)	8.83	3.53	2.72	1.89	6.94
S.L.	50Hz	Percentage error(%)	15.5	6.7	5.4	4.5	11.0
		Error(cm)	14.7	5.83	4.75	3.82	10.88

Bias Variance Tradeoff

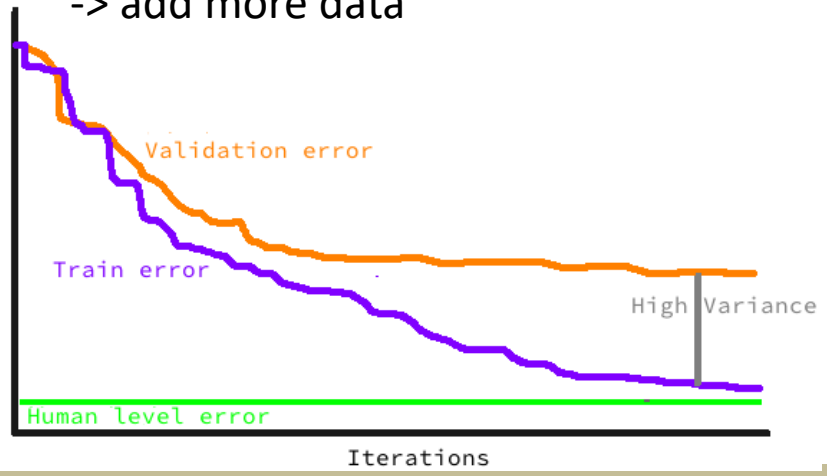
high bias in the model — Under fitting
->Increase capacity of the network



high bias and variance in the model
rethink your architecture, and add more data

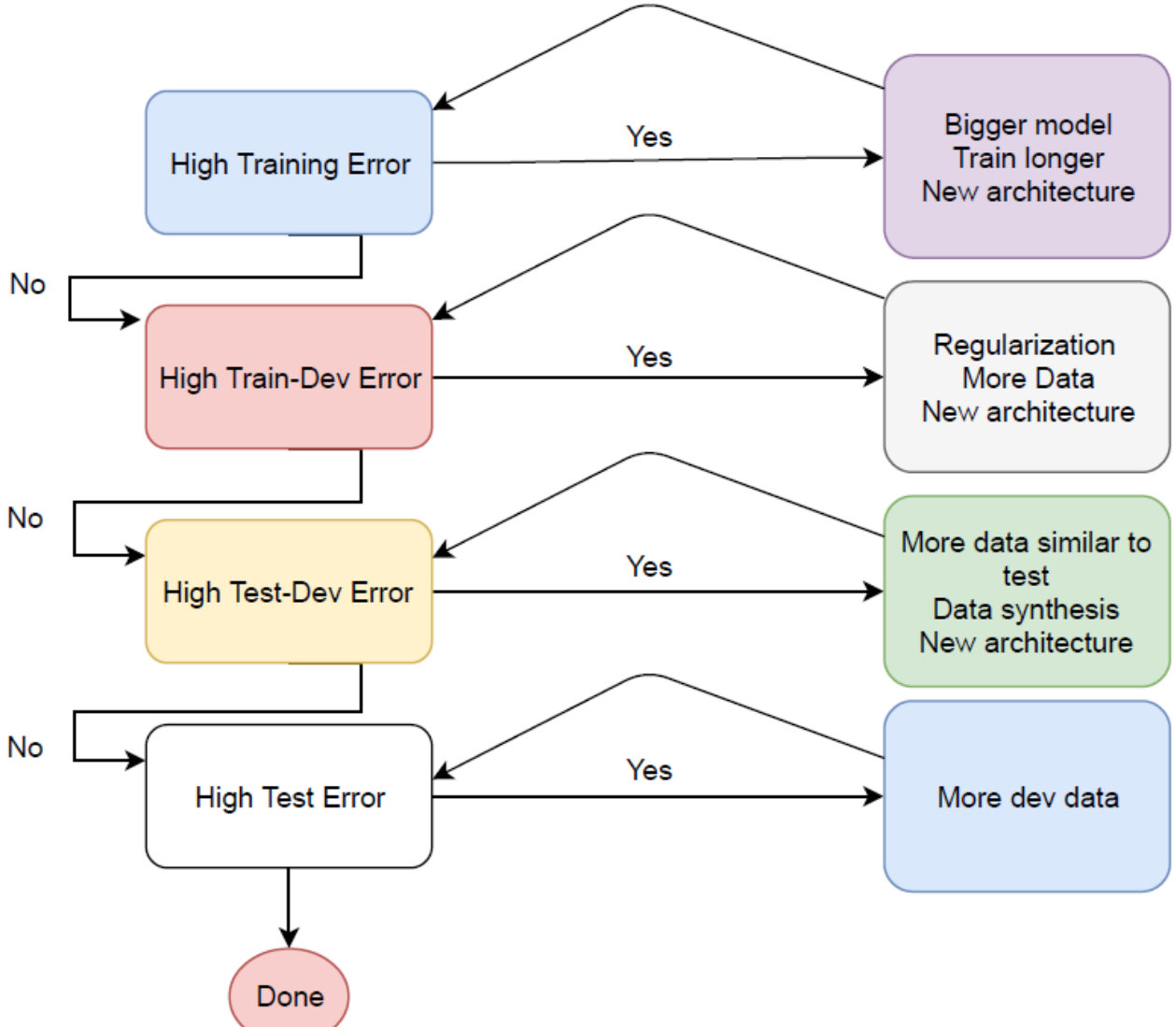


high variance in the model — over fitting
->regularization to limit model freedom
-> add more data



Flowchart for working with a model

- You can always take one of two simpler options: more data or bigger model



Human Level Performance

- DL models tend to plateau once they have reached or surpassed human-level accuracy
- Reasons for the plateau
  - a noisy subset of the data makes further improvement futile
  - But human are good at these tasks so trying to make progress beyond that suffers from diminishing returns
- how should one pick the criteria for human-level accuracy
  - E.g. typical human: 5%, general doctor: 1%, specialized doctor: 0.8%, group of specialized doctors: 0.5%
  - always the best accuracy possible