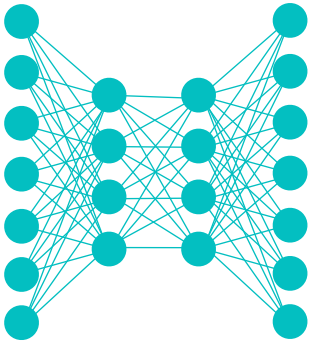
Lecture Notes for

Neural Networks and Machine Learning

Multi-Task Demo

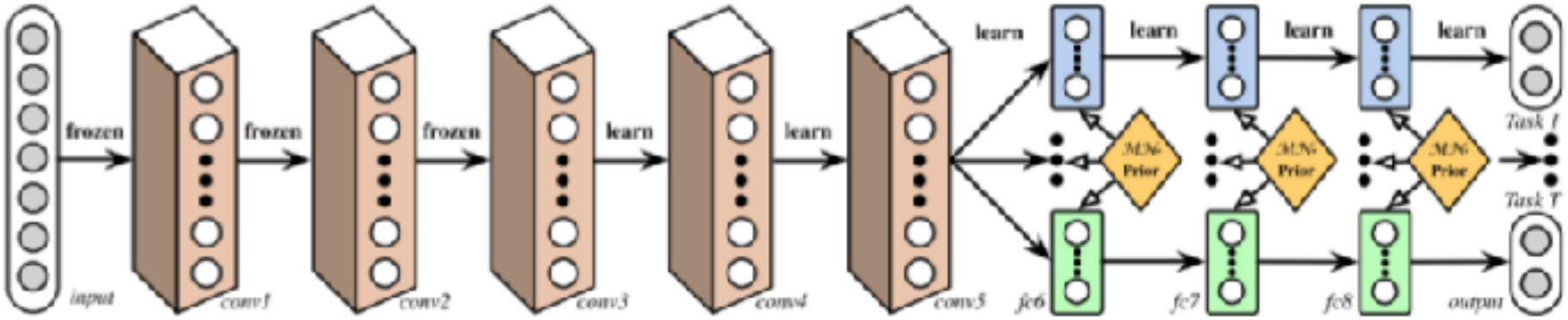


Logistics and Agenda

- Logistics
  - None!
- Agenda
  - Multi-Task Examples
  - Paper Presentation
  - Multi-Task Demos
  - Multi-Task Town Hall
- Next Time
  - Variational Auto-Encoders

Multi-Task Model Examples

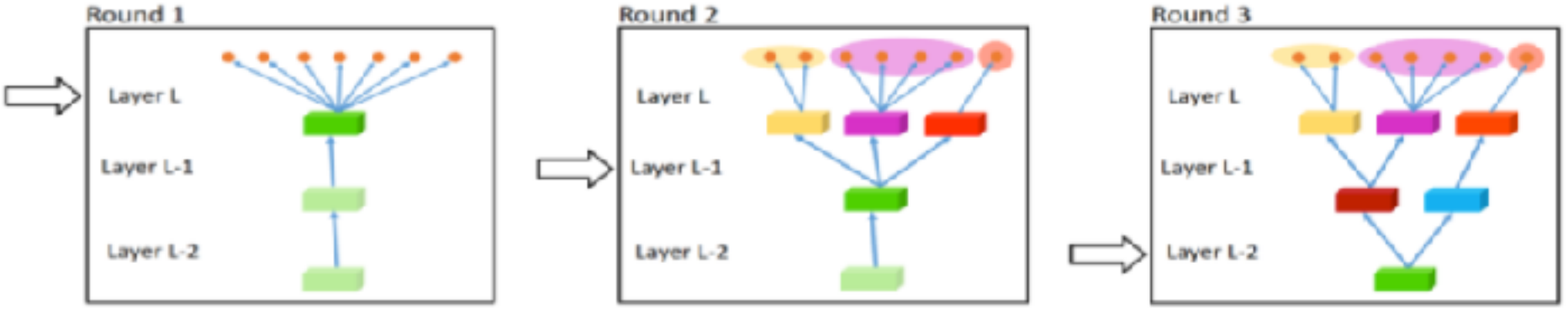
Multi-task: Deep Relationship Networks



- Start training traditionally
- Minimize Kroenecker Product between fully connected task specific layers
  - that is, make Covariance between layers close to identity
  - encourages feature maps in each task to be **less correlated** to feature maps of another task

https://arxiv.org/pdf/1506.02117.pdf

Multi-task: Adaptive Feature Sharing



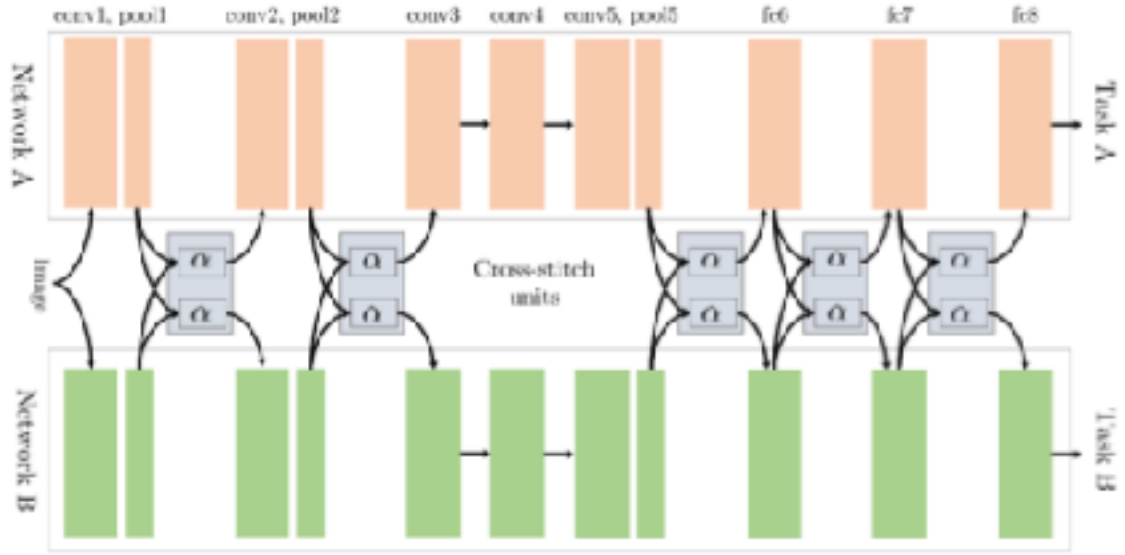
- Train
- Repe
  - D
  - Cut weights and fine tune network
  - Decrement current layer index

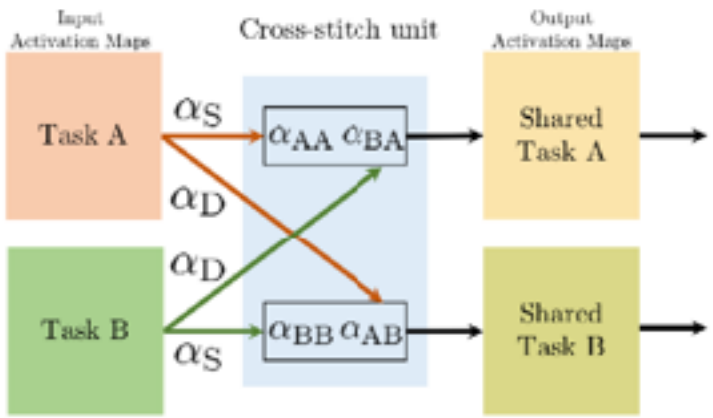
$$A^\star, \omega^\star(l) = \underset{A \in \mathbb{R}^{d \times d'}, |\omega| = d'}{\arg\min} ||W^{p,l} - A W^{p,l}_{\omega:}||_F, \qquad (2)$$

where $W^{p,l}_{\omega:}$ is a truncated weight matrix that only keeps the rows indexed by the set ω. This problem is NP-hard, however, there exist approaches based on convex relaxation

http://openaccess.thecvf.com/content_cvpr_2017/papers/Lu_Fully-Adaptive_Feature_Sharing_CVPR_2017_paper.pdf

Multi-task: Cross Stitch Networks

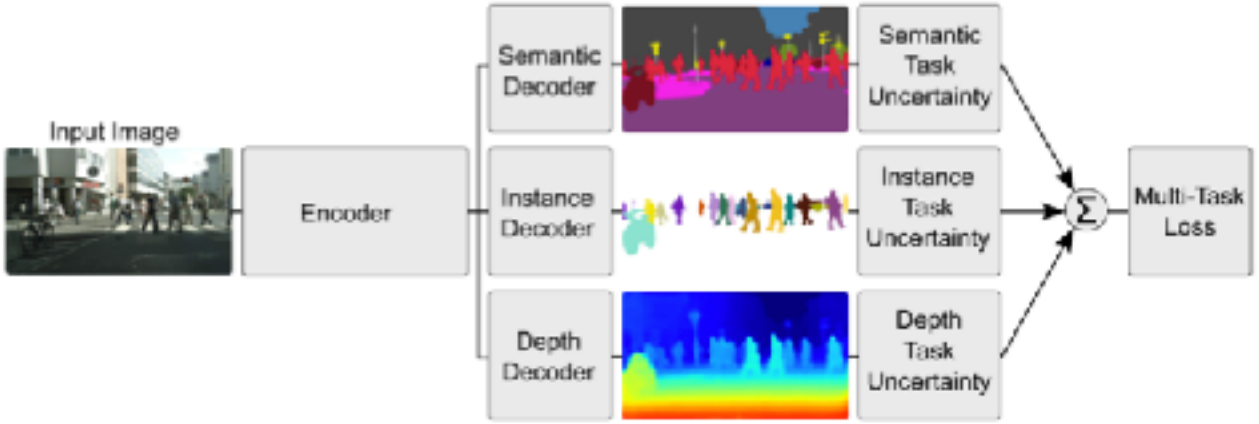




- Only works for simultaneous multi-label problems
  - like semantic segmentation and surface normal segmentation (clustering similarly facing objects)
- Take a learned weighted sum of the activations
- Works a little better than single task, but no worse

https://arxiv.org/pdf/1604.03539.pdf

Multi-task: Uncertainty Weighting



- Use variance of each loss function from each task to normalize
  - call it homoscedastic without sound reasoning because that feels better than “normalized variance”
  - talk about homoscedasticity for no reason
- Write an entire paper in a “mathy” way to make it seem like more of a contribution
- Profit because you are Oxford/Cambridge and reviewers give you a pass

https://arxiv.org/pdf/1705.07115.pdf

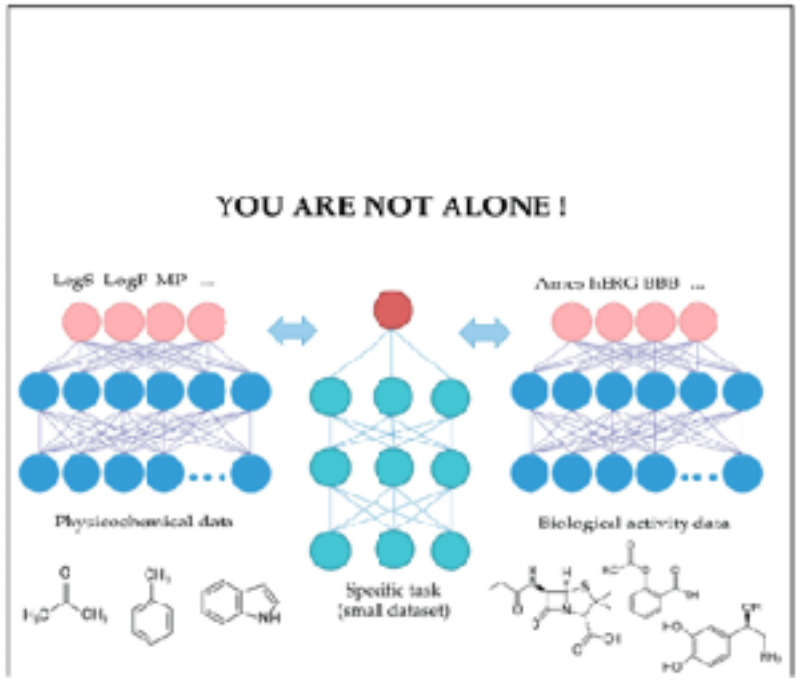
Current Multi-task Research

- Incredibly diverse sets of solutions
- Mostly not evaluated on similar datasets
- Reasoning given is mostly ad-hoc…
- Theory is wildly under developed
  - because the problem is incredibly difficult
- Neural architecture search is an option…

Paper Presentation: Multi-task with Chemical Fingerprints

A Survey of Multi-task Learning Methods in Chemoinformatics

Sergey Sosnin,[a] Mariia Vashurina,[a] Michael Withnall,[c] Pavel Karpov,[b] Maxim Fedorov,[a,d] and Igor V. Tetko*[a]



Multi-Task Learning in Keras with Multi-Label Data

Fashion week, colors and dresses

"finish demo"

Follow Along: https://www.pyimagesearch.com/2018/06/04/keras-multiple-outputs-and-multiple-losses/

Multi-Task Learning

School Data, Computer Surveys

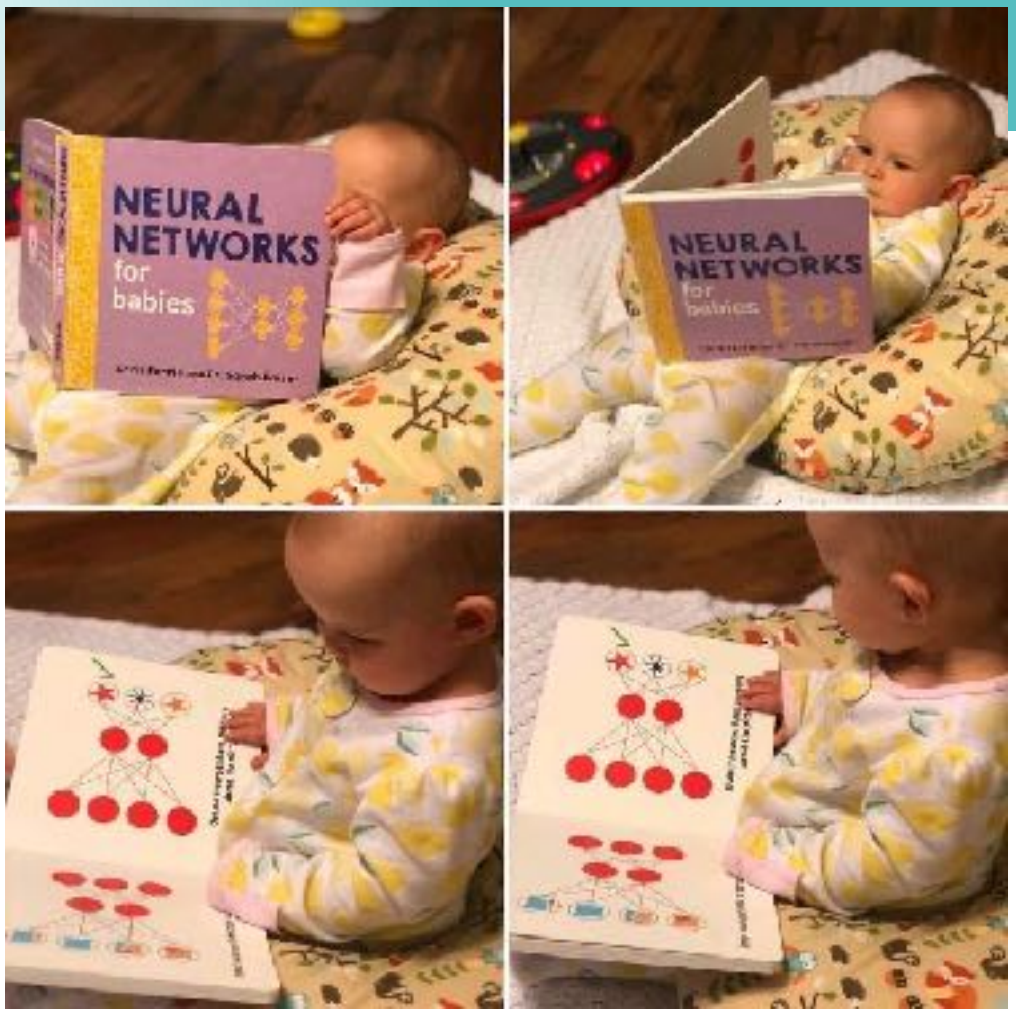
“finish demo”

Traian Pop

Luke Wood

```
Follow Along: LectureNotesMaster/
05 LectureMultiTask.ipynb
```


Lab Three Town Hall



Multi-Task Networks
Multi-Modal Networks

Lecture Notes for

Neural Networks and Machine Learning

Demo Multi-Task

Next Time:
GANs

Reading: Chollet 8.1-8.5