The Future of Radio
Learning Efficient Signal Processing Systems

Tim O’Shea, oshea@vt.edu
Founder/Chief Scientist, DEEPSIG Inc
Research Faculty, Virginia Tech
Deep Learning Trends

• Large Neural Networks are Disrupting Signal Processing
  • Bigger change than most people realize
    • Feature Learning
    • End-to-end learning
    • Widely applicable to many domains

• Feature engineering is becoming irrelevant
  • Expert transforms unnecessary to achieve state of the art performance
  • Engineered features created barriers to learning anyway
  • Engineered algorithms created work in optimizing disparate algorithms
  • Just embrace everything as a dense multiply accumulate with some arbitrary set of weights

Hierarchy of trained representations

Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]
Deep Learning Trends

• Things that aren’t that exciting anymore
  • Computer Vision (Object Recognition)
    • Self driving cars (Tesla, Comma.ai, etc)
  • Voice Recognition (Siri / Google Assistant)

• It’s clear deep learning can (has) destroy the state of the art in these fields
  • Custom silicon can drastically bring down power costs of these networks
    • Apple “Bionic” processor, Google TPU, etc
  • These ships are already sailing wholesale
What is still brewing ...

• Probably every signal processing algorithm on earth could be reconsidered/improved
  • Especially learned representations ...

• Simple construct of the autoencoder
  • Learns entirely new representations of information
  • Based on reconstruction loss or other loss functions

\[
\phi : \mathcal{X} \rightarrow \mathcal{F} \quad \psi : \mathcal{F} \rightarrow \mathcal{X}
\]

\[
\phi, \psi = \arg \min_{\phi, \psi} \| X - (\psi \circ \phi) X \|^2
\]

Minimize encoder and decoder loss

Only for relevant distribution of \( X \)!
What's interesting ... and coming next...

- **Image Compression Schemes**
  - Context aware compression: both 0.08 bits/pixel
  - New JPEG Standards

- **Encryption**
  - Simply a min/max reconstruction optimization

- **Video Compression**
  - Netflix / MIT adoption now
Better Estimators under impaired channels

- Re-consider estimators and representations
- In the context of actual distribution information
- Better estimation under impairments!
- Especially good for short-time windows
Efficient Approximate Decoders

- Learning decoders for ‘near optimal’ error correction codes! (Partitioning to scale)
  - FEC decoding / detection currently the #1 power consuming operation in radio baseband devices
  - Work from Cammerer, Gruber, Hoydis, ten Brink (University of Stuttgart / Nokia-Bell Labs)
  - Shows near-optimal polar code decoding performance with partitioned neural networks at lower complexity than successive cancellation or Belief Propagation!

- Potentially a major advance in error correction
  - Learn approximate decoders on code-word sets
  - Low latency one-shot decoding at higher efficiency

Reproducing at one point either exactly or approximately a message selected at another point”

- C. E. Shannon, “A mathematical theory of communication,” 1948

- All communication systems need to do is optimize for reconstruction loss
  - Everything else is a secondary sub-task
  - Let's not get hung up on minutia

- This actually works really well
  - Matches coded modulation baselines immediately
Radio Methods for Saliency

• How can we reduce the search space?
  • Leverage things we know about propagation physics?

• Introduce domain aware attention mechanisms in the right way –
  • Decomposition of receiver
    • Learned estimation modules (Attention model)
    • Expert transformation modules to match physical world propagation models/effects
    • Learned demapping/representation modules
  • Joint learning of encoder/modulator, synchronizer, decoder/demodulator, and over the air representation
    • Lower complexity learning problem
      • Converges faster, less overfitting
      • Only imparts propagation medium expert knowledge (things we can’t change)
      • Learn everything else end-to-end
What if we have to share the channel?

- Can easily extend this method to multi-access channel
  - Learn a better solution than orthogonality
  - Same basic principals
  - Comes up with immediately interpretable results

(a) (1,1) (b) (2,2) (bits, symbols)
Synthesizing Complex Multi-user PHYs

- Comparison with existing methods
  - Compare the multi-user access channel
  - Orthogonal (Time-slicing (TS)) vs learned method
  - Learns new never before seen PHY scheme
  - Infinite number of possible waveforms!
  - In this case pseudo-orthogonal superposition code
Synthesizing Complex Multi-Antenna PHYs

- Extend this technique to multi-antenna
  - Same basic principals
  - Complex MIMO channel effects
- Incorporate CSI feedback
- Entirely new MIMO scheme
Synthesizing Complex Multi-Antenna PHYs

- Can learn incredibly complex joint solutions
  - Soft joint-modulation-coding schemes
  - Outperform current baselines (zero forcing MIMO)
- Enormous potential for distributed wireless
  - MIMO system performance
  - Secrecy and privacy

Complex MIMO PHY Learning
Non-standard MIMO QAM Modes

Learned 2x2 Constellations
1. Transmitted
2. Diag Rx
3. Uniform Rx

Overall Best 2x2 Scheme Performance

Bit Error Rate (BER)

Signal to Noise Ratio (dB)
Thanks! Questions?

ML Driven RF Systems are coming FAST
Come and work with us / talk to us!

Next generation radio sensing and communications systems

Applied Research
Mature Capabilities
info@deepsig.io

Unrestricted Fundamental
Enabling Research
oshea@vt.edu