1. Advanced CNN architectures
   ○ Mostly covered by Lana: GoogLeNet-v1, ResNet
   ○ Beyond ResNets: GoogLeNet-v4, FractalNet, DenseNets
   ○ Attempts to explain ResNets: identity mappings in ResNets, ensembles of shallow networks, connections to RNNs and visual cortex
   ○ Other noteworthy architectures: Spatial Transformer, Wide And Deep, NoPooling, Highway Networks
   ○ Transfer Functions: ELU, PReLU, Leaky ReLU

2. Advanced training techniques
   ○ Gradient Descent Algos: Overview, another overview, Adagrad, RMSProp, Adam, Eve
   ○ Regularization: DropOut, Batch normalization
   ○ Network Initializations: Xavier/Glorot, Orthogonal, Data Dependent, Good Init

3. Network Compression and speeding up networks
   ○ Quantization, Pruning + Quantization + Encoding
   ○ XNORNet, Low Precision, Limited Precision
   ○ Low-rank expansions, CP-Decompositions, Non-linear Approx, Fastfood
   ○ Factorized/Separable Conv: Flattened Conv, Factored Conv
   ○ SqueezeNet, PerforatedCNN

4. Object Detection
   ○ Lana’s object detection slides from CS543
   ○ RCNN, SPPNet, Fast RCNN, Faster-RCNN, MultiBox, ION, YOLO, SSD, Feature pyramid networks
   ○ Speed/accuracy trade-offs for modern convolutional object detectors

5. Semantic segmentation, dense pixel labeling
   ○ ZoomOut, Hypercolumns
   ○ FCN, DeepLab, CRFasRNN, DeepMask, SharpMask
   ○ Dilated Conv
   ○ PixelNet
   ○ Optical Flow, Edge Detection
   ○ Fully connected deep structured networks
   ○ Stacked hourglass networks for human pose estimation

6. Similarity learning with CNNs
   ○ Siamese Networks, Triplet Loss, Overview
   ○ Applications: Patch Match, Visual Similarity, Overhead To Street, LIFT

7. Visualizing CNNs, adversarial examples
   ○ Visualizing and Understanding Convolutional Networks
○ Synthesizing the preferred inputs for neurons in neural networks via deep generator networks
○ Understanding Deep Image Representations by Inverting Them
○ Inverting Visual Representations with Convolutional Networks
○ Deep Inside Convolutional Networks
○ Deep networks are easily fooled
○ Intriguing properties of neural networks
○ Explaining and harnessing adversarial examples
○ Adversarial examples in the real world
○ Universal Adversarial Perturbations
○ Analysis of robustness to adversarial perturbations
○ Dense Associative Memory is Robust to Adversarial Inputs

8. Generative Adversarial Networks
○ Tutorial by Ian Goodfellow
○ GAN, Laplacian Pyramid of GANs, InfoGAN, EBGAN, Conditional GAN
○ Applications: Segmentation, Image-to-image translation, Video Generation, Text2Im
○ Advanced models: Coupled GANs, Stacked GANs

9. Variational Autoencoders
○ Original VAE paper
○ Tutorial by Carl Doersch
○ Conditional VAEs: Attribute2Im, Future Prediction
○ Plug n Play generative networks
○ Hybrid models: adversarial autoencoders, VAE+GAN with learned similarity
○ Other generative models (optional): Mixture Density Network, Gumbel Softmax

10. Other image generation methods
○ DRAW: A Recurrent Neural Network For Image Generation
○ Towards Conceptual Compression
○ Generative Image Modeling Using Spatial LSTMs
○ Pixel Recurrent Neural Networks
○ Conditional Image Generation with PixelCNN Decoders
○ Style transfer, transfer with perceptual losses

11. 3D + graphics
○ Deep Convolutional Inverse Graphics Network
○ DeepStereo: Learning to Predict New Views from the World’s Imagery
○ PoseNet: A Convolutional Network for Real-Time 6-DOF Camera Relocalization
○ Perspective Transformer Nets
○ Unsupervised Learning of 3D Structure from Images
○ Single Image 3D Interpreter Network
12. Self-supervised learning
   - 3D face from image
   - Context as Supervisory Signal
   - Unsupervised Learning of Visual Representations using Videos
   - Slow and steady feature analysis
   - Learning Visual Features from Large Weakly Supervised Data
   - Split-brain Autoencoders
   - Context Encoders
   - Colorful Image Colorization, Automatic Colorization
   - Ambient Sound Provides Supervision for Visual Learning
   - Unsupervised learning through video prediction

13. Deep reinforcement learning: Q Learning
   - Deep RL tutorial
   - Q Learning Tutorial
   - Playing Atari with DQN
   - Deep Reinforcement Learning with Double Q-learning
   - Learning to Play in a Day with DQNs
   - Application to recognition (optional): Object Localization

14. DRL II: Policy Gradients, planning
   - Blog: http://karpathy.github.io/2016/05/31/rl/
   - REINFORCE - historic paper
   - AlphaGo
   - PGQ: Combining policy gradient and Q-learning
   - Reinforcement Learning with Unsupervised Auxiliary Tasks
   - Value iteration networks
   - The Predictron: End-To-End Learning and Planning
   - Applications to recognition (optional): Recurrent Models of Visual Attention, Action Detection

15. Deep learning for manipulation, navigation (both RL and self-supervised)
   - End-to-end training of deep visuomotor policies
   - Learning to poke by poking
   - Learning hand-eye coordination with large-scale data collection
   - The curious robot
   - Supersizing self-supervision
   - Learning to Navigate in Complex Environments
   - Real single-image flight without a single real image
16. Recurrent architectures: LSTM, GRU, RNN
   ○ Mostly covered by Arun:
     ■ LSTM: A Search Space Odyssey
     ■ An Empirical Exploration of Recurrent Network Architectures
   ○ Visualizing and understanding recurrent networks
   ○ Identity RNN
   ○ Unitary Evolution Recurrent Neural Networks
   ○ Recurrent Dropout without Memory Loss
   ○ Recurrent Batch Normalization
   ○ Architectural Complexity Measures of RNNs
   ○ Collection of useful papers and applications:
     https://github.com/kjw0612/awesome-rnn#applications
   ○ Application of recurrent models to recognition: Feedback networks

17. Image captioning with recurrent models, attention
   ○ Show And Tell, Follow-up
   ○ NeuralTalk
   ○ From Captions to Visual Concepts and Back
   ○ Attributes for Captioning
   ○ Show, Attend, and Tell
   ○ Attention Correctness in Neural Image Captioning

18. Image-text embeddings, grounding
   ○ Datasets: Flickr30k Entities, Visual Genome
   ○ Deep structure-preserving embeddings
   ○ Order embeddings
   ○ Grounding by Reconstruction
   ○ DenseCap

19. Visual Question Answering
   ○ The VQA dataset
   ○ Simple Baseline for Visual Question Answering
   ○ Revisiting Visual Question Answering Baselines
   ○ Where To Look: Focus Regions for Visual Question Answering
   ○ Hierarchical Question-Image Co-Attention for Visual Question Answering
   ○ Multimodal Compact Bilinear Pooling (original CBP)
   ○ Neural module networks
   ○ CLEVR

20. Deep learning for NLP
   ○ Word Embeddings: Word2Vec, Glove, Doc2Vec, Skip Thought
   - **Semantic Parsing**
   - **Language Modeling with Gated Convolutional Networks**
   - **Language Modeling with Outrageously Large Neural Networks**

22. Deep learning for audio
   - **Recognition:** LSTM for Acoustic Modeling, RNN, Deep Speech, Raw Waveform
   - **Generation:** WaveNet

23. Architectures with memory
   - **Neural Turing Machines**
   - **Memory Networks**
   - **End-to-End Memory Networks**
   - **Pointer Networks**
   - **Differentiable Neural Computers**

24. Meta-algorithms
   - **Learning to learn by gradient descent by gradient descent**
   - **Neural architecture search with reinforcement learning**
   - **Designing Neural Network Architectures using Reinforcement Learning**
   - **HyperNetworks**
   - **Learning to learn for global optimization of black box functions**