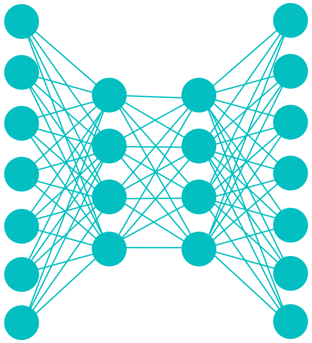
Lecture Notes for

Neural Networks and Machine Learning

Deep Transfer Learning and Demo

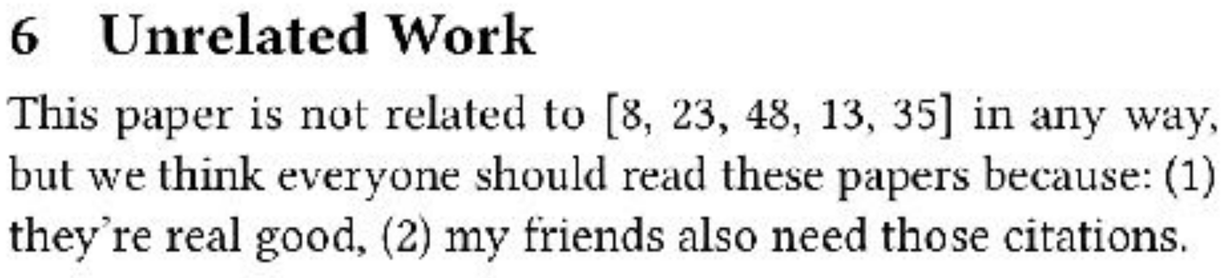


Logistics and Agenda

- Logistics
  - None!
- Agenda
  - Transfer Learning in Deep Learning
  - Transfer Learning Demo
- Next Time:
  - Transformers for Text and Vision
- Next Next Time
  - Self-Supervised Learning and Consistency Loss

Transfer Learning with Neural Networks

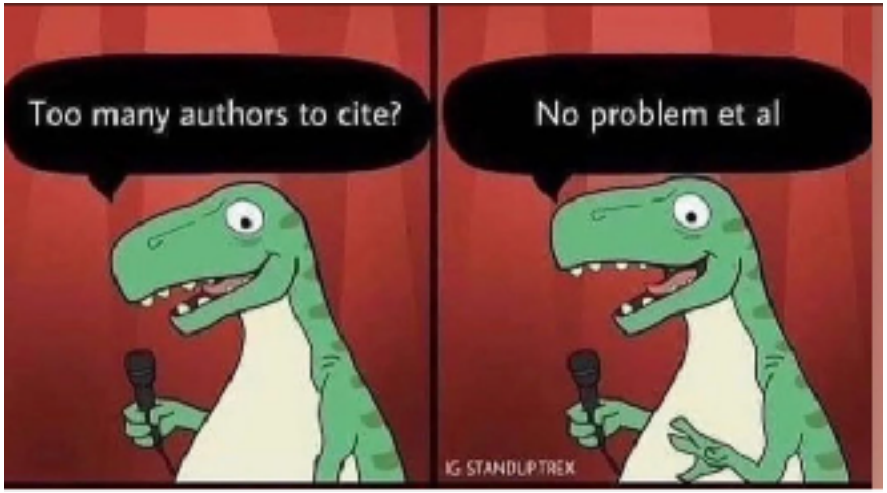
Found in a recent paper:



6 Unrelated Work

This paper is not related to [8, 23, 48, 13, 35] in any way, but we think everyone should read these papers because: (1) they're real good, (2) my friends also need those citations.

7 Related Work



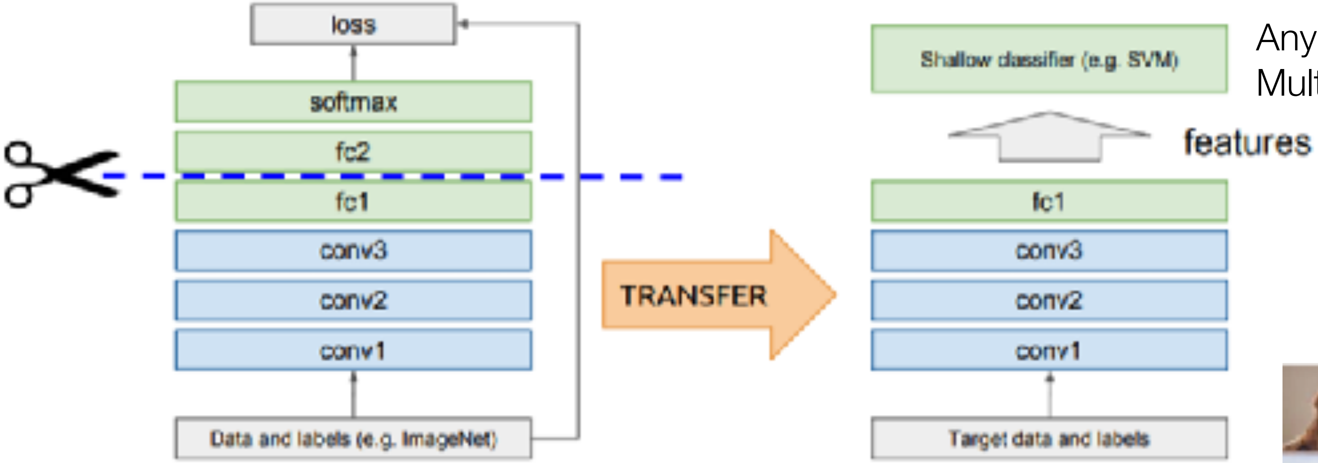
Deep Transfer Learning

- Almost always **Inductive Transfer**
  - (new task , same domain, or domain adaptation)
- Almost always **Feature Representation Transfer**
  - like image pre-training
- All other topics are mostly open research topics that maybe one of you will solve!

(Sun, B., Feng, J., & Saenko, K. (2016). Return of Frustratingly Easy Domain Adaptation)

Approaches with Deep Learning

- **Feature Extraction Transfer**
  - Most well known: use learned parameters from one task in another task in same domain
  - Most useful when labels for target domain are sparse



Any classifier: SVM, Random Forest, Multi-layer perceptron

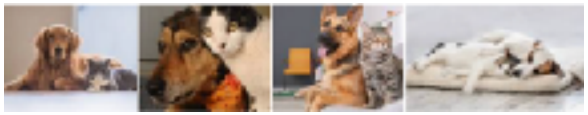


Image Net

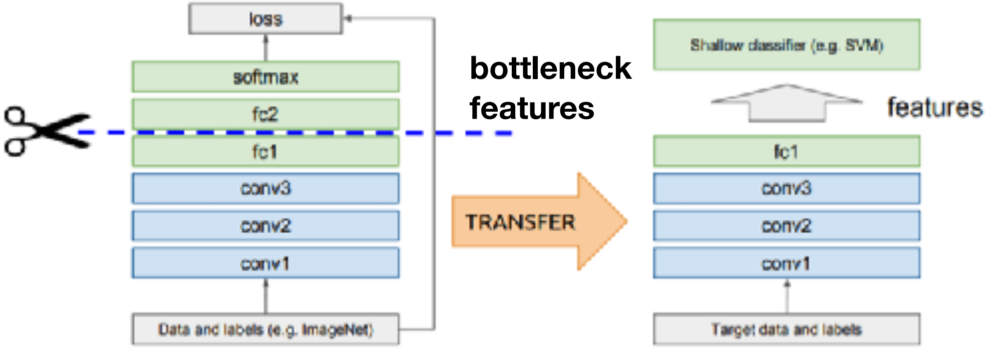
New domain: Dogs versus Cats

New domain: Gaze Classification

https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a

Defining the Bottleneck

- Frozen training layers before bottleneck:
  - Why waste computations?
  - Computing more than one forward pass on the same data—just save them out
  - *Unless* using augmentation



- In Keras, build multiple models with different entry points
  - **Input** to **Bottleneck**
  - **Bottleneck** to **Output**
  - **Input** to **Output**

```
inputs = Input(shape=(IMG_SIZE, IMG_SIZE, 3))
model_base = VGG(include_top=False,
            input_tensor=inputs, weights="imagenet")
```

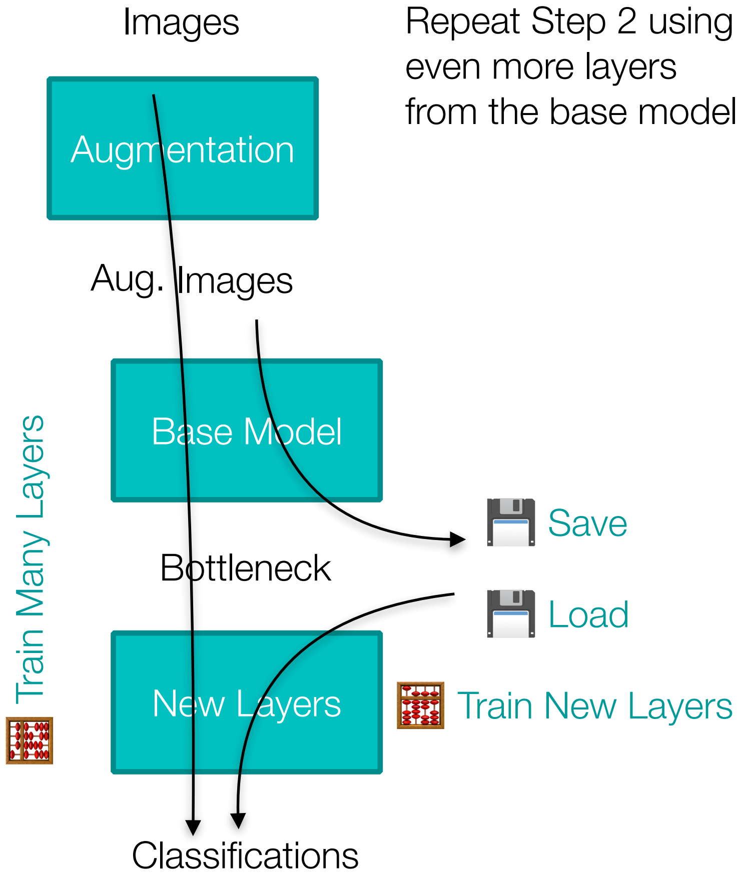
```
bottleneck = Input(shape=model_base.output.shape)
outputs = Dense(NUM_CLASSES)(bottleneck)
model_top = Model(bottleneck, outputs)
```

```
model_total = Model(inputs, outputs)
```

https://keras.io/examples/vision/image_classification_efficientnet_fine_tuning/

Freezing and Fine-tuning

- Step 1, Freeze base model:
  - No update during back-propagation
  - Only update layers after the bottleneck
  - *Optional: Augment a set of training data*
  - Send training dataset through base model
    - Save out bottleneck features
  - Train bottleneck features in new task
    - Typically 5-10 epochs is sufficient, easy to overfit
    - Larger training step size is okay
- Step 2, Fine-tune, unfreeze a few layers in base model:
  - *Optional: Setup images to use some type of augmentation*
  - Attach newly trained model to pre-trained model
  - Train to your hearts content, use smaller training step size



Step 3: Rinse and Repeat Step 2 using even more layers from the base model

Bottlenecking on a GPU

Dogs versus Cats

eclarson Eric Larson

justinledford Justin Ledford

Member of 8000net

Updated for tf==2.12 in the Main Repository:
02 Transfer Learning.ipynb

Original Example: https://github.com/8000net/Transfer-Learning-Dolphins-and-Sharks

Justin Ledford ·

Another Great Example:
https://keras.io/examples/vision/image_classification_efficientnet_fine_tuning/

Popular Transfer Learning Models

- **Vision**:
  - Conv Architectures:
    - VGG, Inception, ResNet, Xception, EfficientNet
    - ViT: Huge ViT
- **Audio**:
  - WaveNet, VGGish
- **Text**:
  - Word Embedding
    - Glove, Word2Vec, ConceptNet
  - Sentence Embedding
    - Universal Sentence Encoders (Google)
    - BERT (Google)

- **From SMU PhD Students in my Research Lab:**
  - VGG for transferring to gaze classification
  - VGG for swapped face detection
  - Domain adaptation for speaker authentication
  - YOLO/DarkNet for surgical instrument detection
  - GLOVE for similar instructions in a maintenance manual
  - BERT for student vocabulary acquisition

Lecture Notes for

Neural Networks and Machine Learning

Transfer Learning

Next Time:
Transformers

Reading: None