**Problem & Motivation**

**Problem:** Building a convolutional neural network for image classification requires a large amount of labeled images to train. However, manually labeling images is a tedious and time-consuming process.

**Semi-Supervised Training:** Instead of using only labeled images, semi-supervised training takes advantage of unlabeled images to train an image classifier. We use a large set of unlabeled images and a small set of labeled images to train the CNN.

**Approach**

**Step 1:** Train an auto-encoder to compress and decompress unlabeled images in order to learn image features

**Step 2:** Use learned features and a small amount of labeled images to train a CNN to classify images

---

**Experiment Setup:**

- Trained auto-encoder on 50,000 images from the cifar10 dataset
  - Compresses 3 channel, 32x32 images to 3 channel, 16x16 images
  - Auto-encoder structure: Input - (256)3c - (128)3c - 2p - (64)3c - (33)2c - 2up - (128)3c - (256)3c - (33)2c - Output
  - Input and output are 32x32 RGB images. (128)3c denotes a convolutional layer with 128 feature maps and a filter of size 3x3, 2p means 2x2 max-pooling, and 2up denotes a 2x2 upsampling layer
  - Training used the L2 loss between the original image $x$ and the reconstructed image $\hat{x}$

**Experiment Results:**

- Trained 3 CNNs on the cifar10 dataset
  - CNN Structure: Input - (256)3c - (128)3c - 2p - (64)3c - (33)2c - 128fc - 10fc
  - Input is a 32x32 RGB image, 128fc denote a fully connected layer with 128 nodes, and the output is the image’s class

---

**Experiments**

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Basic</th>
<th>Semi-supervised, trained with 10,000 images</th>
<th>Semi-supervised, trained with 50,000 images</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>44.7%</td>
<td>54.71%</td>
<td>55.79%</td>
</tr>
</tbody>
</table>

**Stacked What-Where Auto-encoder**

- A stacked what-where auto-encoder (SWWAE) replaces the max-pooling and upsampling layers used in a basic auto-encoder with what-where pooling/unpooling
  - SWWAEs have been shown to greatly improve semi-supervised training [1]

**SWWAE Results**

- We modified a VGG16 network and created a deeper auto-encoder
  - We trained a basic auto-encoder and a SWWAE with this deeper structure

---

**References**