# Multi-Task Self-Training for Learning General Representations

(Self-supervised learning에 초점이 맞춰진 논문)

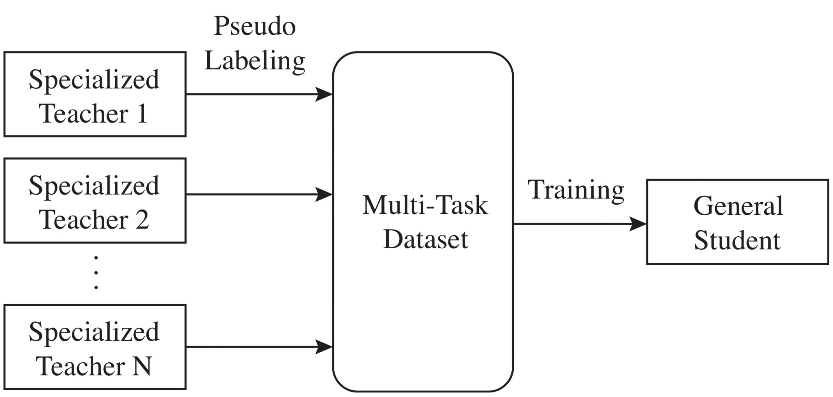
3 Steps:

① Train specialized teachers independently on labeled datasets

② Use the specialized teachers to label an unlabeled dataset to create a multi-task pseudo labeled dataset.

③ The dataset, which now contains pseudo labels from teacher models trained on different datasets/tasks, is then used to train a student model with multi-task learning

Despite the wide adoption of transfer learning from supervised training, the features may not necessarily be useful for downstream tasks.



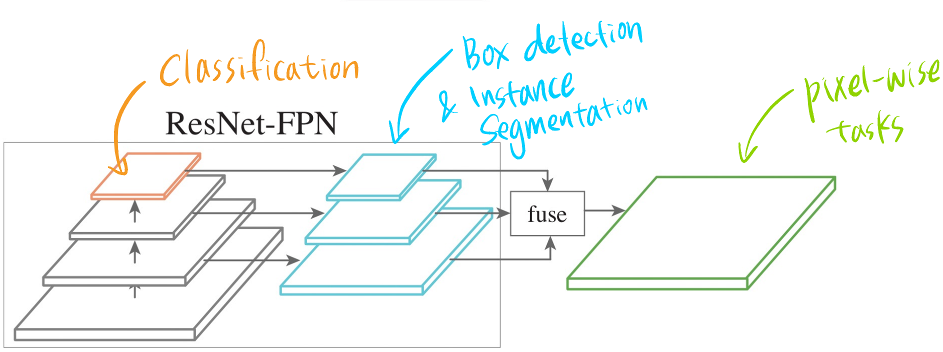




Figure 3. **The ResNet-FPN backbone architecture for multi-task learning.** **Orange:** the top-level features for classification. **Cyan:** multi-scale features for box detection and instance segmentation. **Green:** the high resolution features for pixel-wise tasks (*e.g.*, segmentation, depth, and surface normal estimation.)

<Method>

## #1. Specialized Teacher Models

- Classification
- detection
- segmentation
- depth estimation

→ transfer the knowledge in specialized teacher models to unlabeled or partially labelled datasets by pseudo labeling.

## #2. Multi-task Student Model

Loss = weighted sum of losses from all tasks

$\mathcal{L} = \sum_i w_i \mathcal{L}_i$

$w_i$ : loss contribution for the task $i$

<Experiments>

- Self-supervised and supervised pre-training on ImageNet does not learn features that generalize nearly as well to tasks other than image classification
- As we continue to add pseudo labels from different tasks/datasets our representations improve in quality.