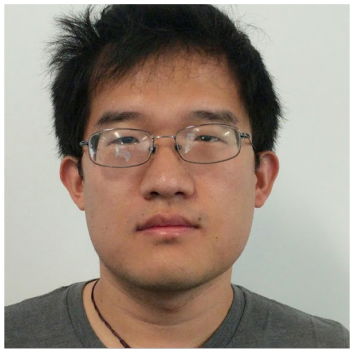
Improving Generalizability of Self-supervised Speech Models

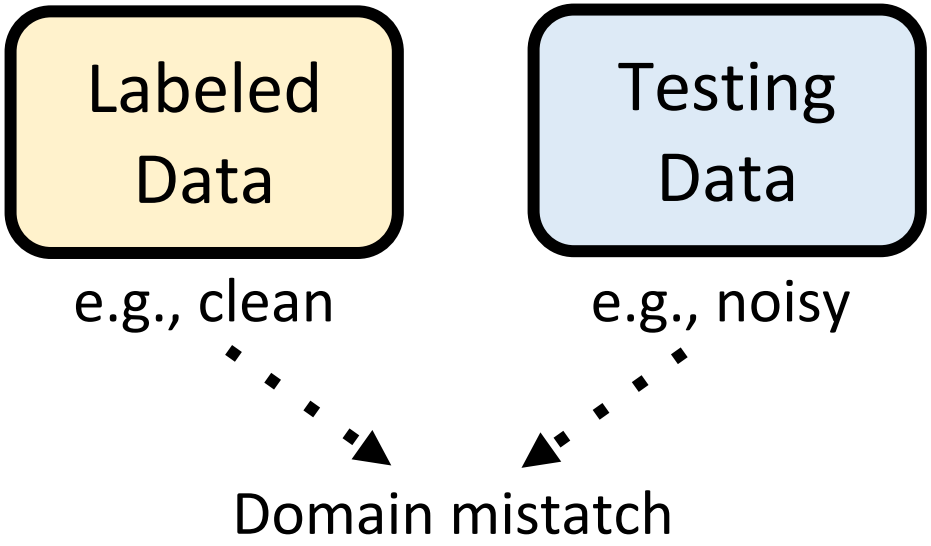
Kuan-Po Huang



Yu Zhang

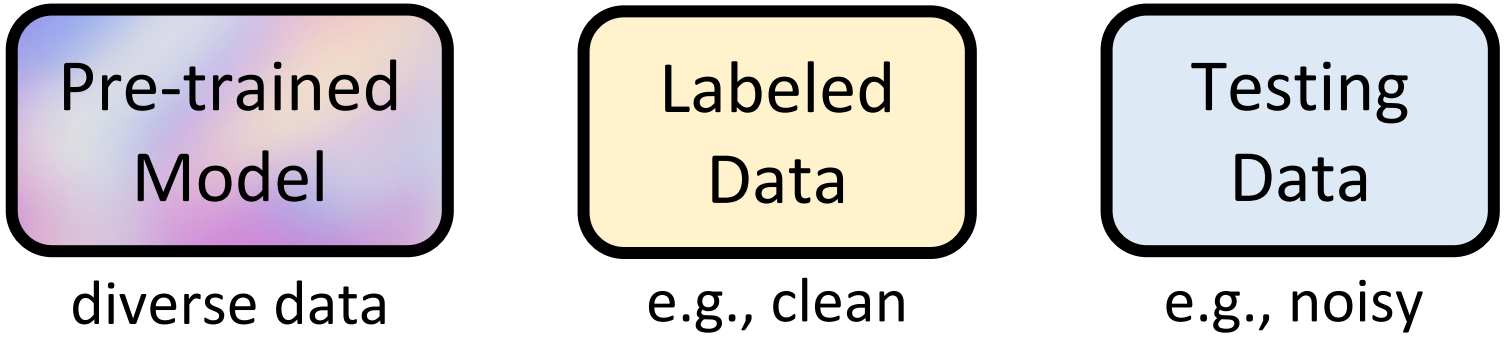
Hung-yi Lee

Motivation



Different domains: recording conditions, speaking styles (read vs. spontaneous), accents/dialects, languages

Motivation

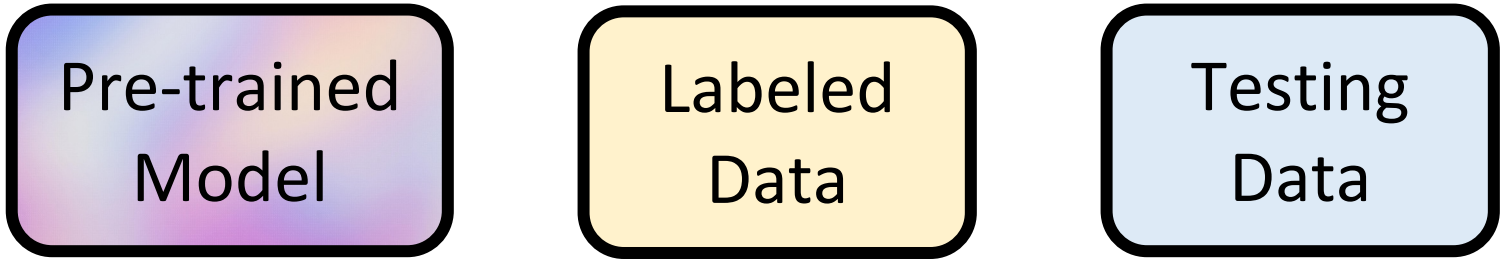


Does pre-training improve model generalizability during fine-tuning?

- In **NLP,** one can fine-tune on English labeled data and test on Chinese.
- How about speech?

Experiments from
Kuan-Po Huang

Step 1: Investigation

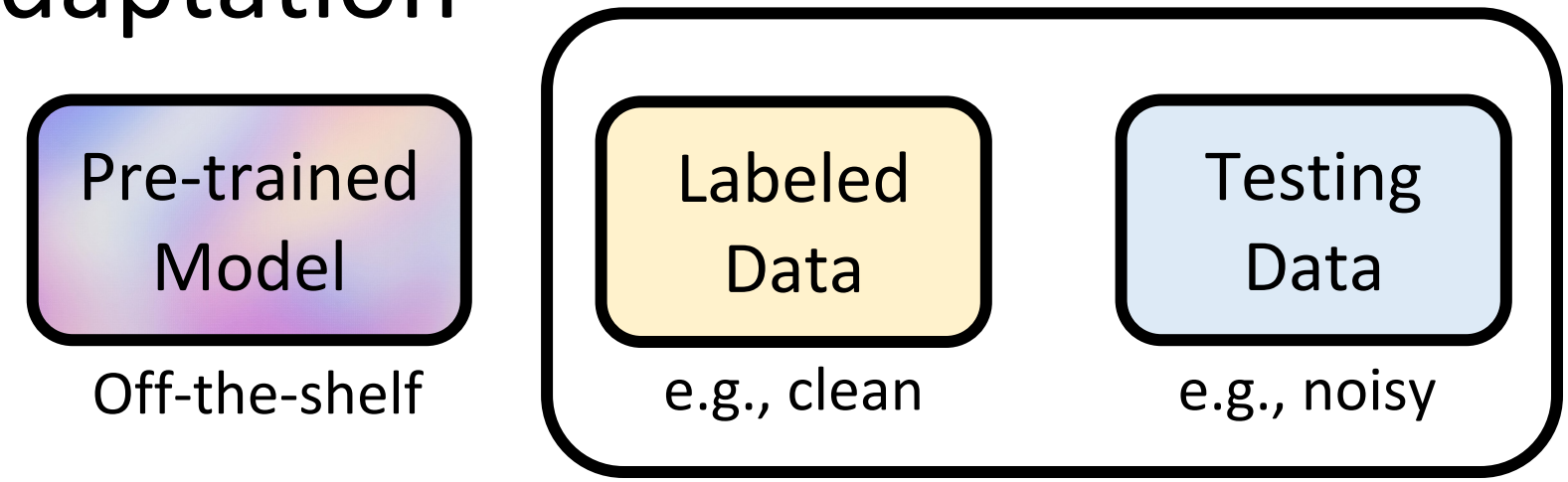


diverse data | e.g., clean | e.g., noisy

Does pre-training improve model generalizability during fine-tuning?

SUPERB task	Clean	Musan noise	Gaussian noise	Reverberation
Intent Classification	0.9834	0.9056	0.8571	0.9634
Keyword Spotting	0.9630	0.9111	0.7683	0.9464
……				

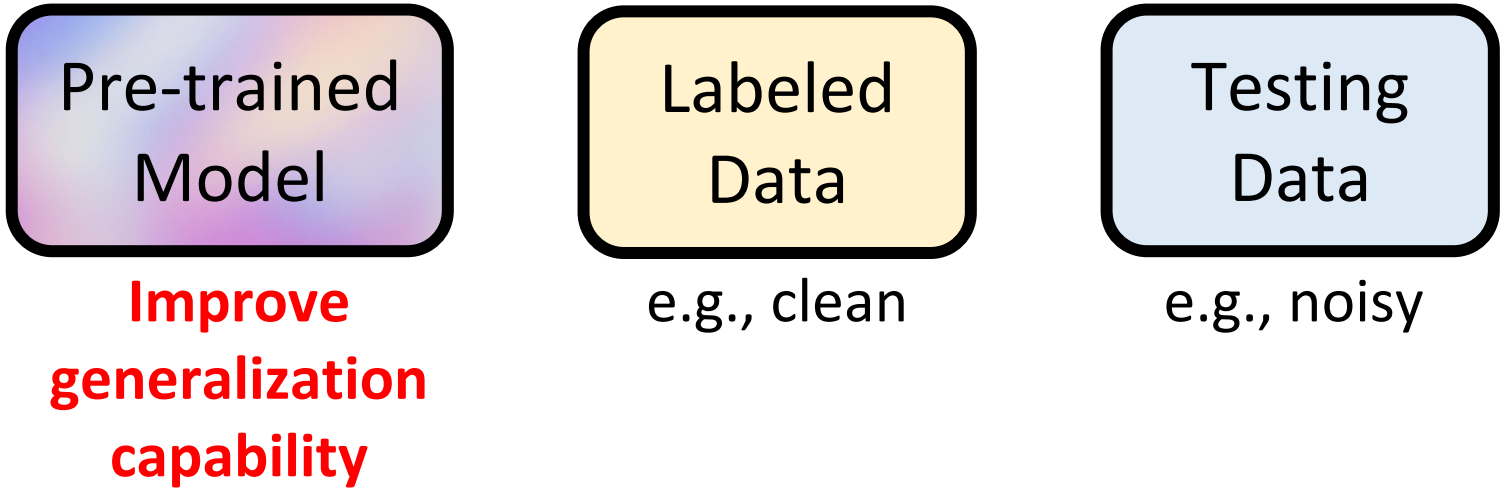
Step 2: Typical Domain Adaptation



Dealing with domain mismatch problem at the fine-tuning stage:

- Domain adversarial training
- Pseudo-labeling
- Testing time training (TTT)

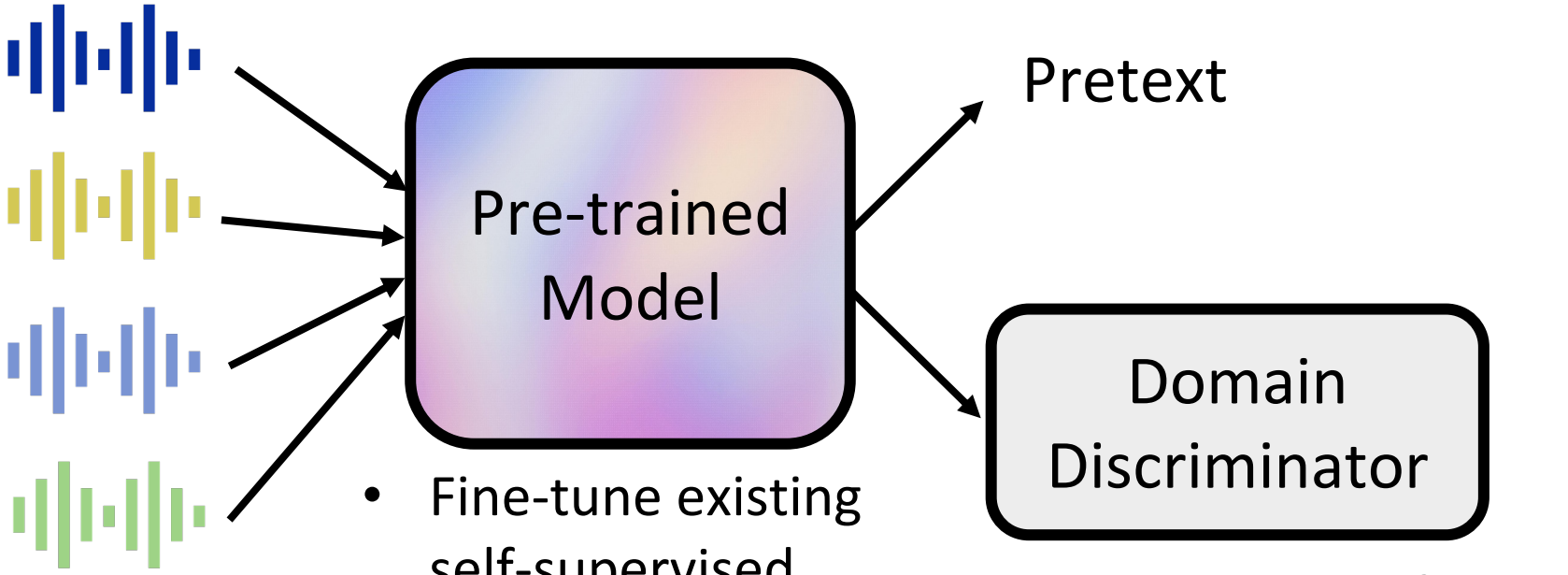
Step 3: Robust Pre-trained Models by Existing Approaches



Native idea (baseline): add more data from different domains for pre-training

Step 3: Robust Pre-trained Models by Existing Approaches

- Domain adversarial training during pre-training



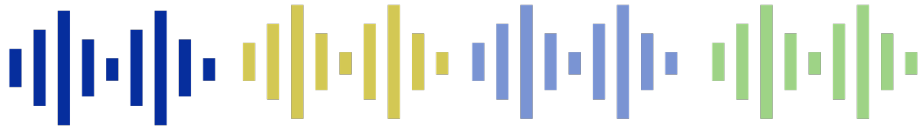
- Fine-tune existing self-supervised models
- Add adapters, and only train them

Representations from different domains have same distribution.

Step 3: Robust Pre-trained Models by Existing Approaches

https://arxiv.org/abs/1710.03463

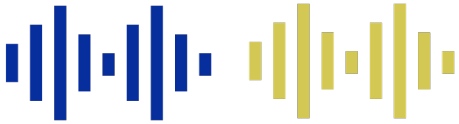
- Domain Generalization



Pre-training data (from different domains)

How to pre-train a model that can generalize to unseen new domains?

Training Domain



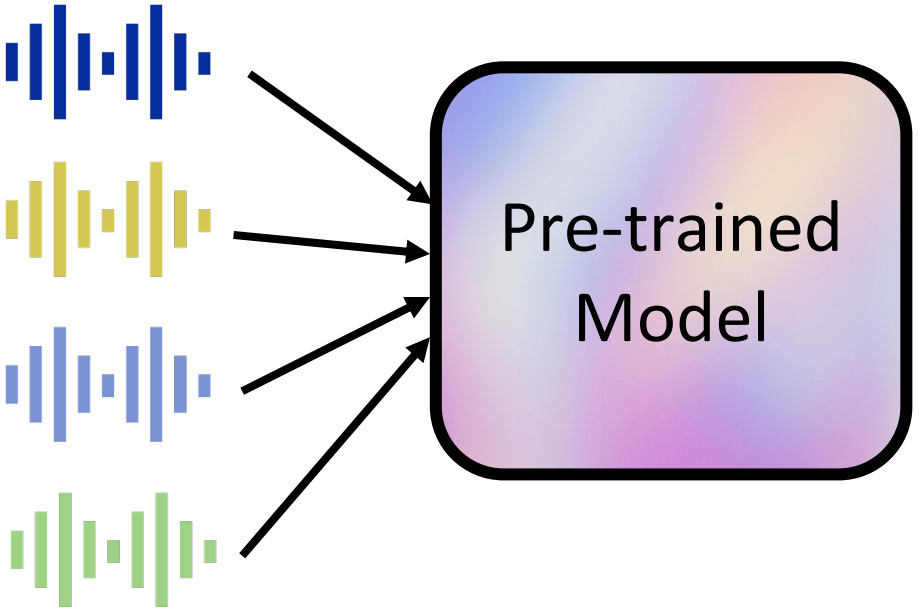
(Pseudo) Testing Domain

Training Target

- Pre-training **on training domains**
- Good pretext performance on **(pseudo) testing domains**

Step 4: Robust Pre-trained Models by Novel Ideas

- Any novel ideas ……



Having some domain specific parameters

Remove the parameters

Domain-agnostic model

(random thoughts, not carefully considered)

Concluding Remarks & Timeline

- March / April
  - Step 1: Generalizability of self-supervised models
  - Step 2: Focus on typical adaptation approach for fine-tuning
- May
  - Step 3: Using existing approaching to make self-supervised model generalize better
- June / July
  - Worst case: writing paper for the finding
  - Step 4: Implement new ideas

Think new idea