The literature has proposed various ways to transfer knowledge between models: i) instance-based, ii) adversarial network-based, iii) knowledge distillation-based, and iv) model parameter transfer-based. Instance-based methods augment target dataset with reweighted source dataset and learn the target task. Therefore, they need to be retrained again with source data, increasing training time. Adversarial-based techniques particularly address a sub-field of transfer learning (TL), domain adaptation. In domain adaptation problem, source and target task share the same label space, but input distribution is different. However, there can be situation in TL when target and source task doesn’t necessarily share the same label space, which adversarial-based techniques are not designed to address. Distillation-based techniques employ a teacher-student paradigm to transfer knowledge from teacher to student. Finally, parameter transfer-based techniques transfer learned parameters from the source model, and freeze or fine-tune them for target task as necessary. This approach is commonly employed in practice as it needs less training time and require less instrumentation, as evident from the support from popular DL libraries for this paradigm. In terms of selective transfer, other approaches proposed ways to transfer knowledge selectively. However, to our knowledge, no techniques have been proposed for selective parameter transfer. In our paper, we propose a decomposition-based technique to fill that gap.

The notion of selective or partial transfer exists in other approaches of TL. For example, among adversarial network-based approaches, Zhangjie et al. [1] proposes a partial transfer of knowledge. They selectively transfer knowledge from a subset of source classes instead of all. However, it is applicable when target labels are a subset of source labels. In traditional TL setup, it is not necessarily true always. For example, source task may have learnt a general taxonomy, while target task may learn a fine-grained taxonomy. Instance-based techniques too proposed to select a closest subset of source data to jointly train with target dataset [2,3]. However, these instance-selection based approaches would tend to overfit when target dataset is very small, also needing higher training time. Knowledge distillation-based techniques too proposed to transfer knowledge selectively [5]. They evaluate the effect of knowledge transfer via every possible pairs of features (links), and choosing to transfer via closest matched links. However, this is very expensive method as it involves inner-loop architecture for evaluating all pairs and therefore, may not apply in practical scenarios [4]. In contrast to these techniques, we decompose source features in terms of their relevance to target task, and transfer most relevant ones. It retains the benefit of traditional parameter transfer, i.e., shorter training time, applicability to scenario where labels spaces are different etc. In opposed to traditional parameter transfer, it offers following benefits: i) alleviate negative transfer problem when domain distance increases, ii) reduce the size of the target model, thereby, reducing training and inference time.

[1] Cao, Zhangjie, et al. "Partial transfer learning with selective adversarial networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

[2] Ge, Weifeng, and Yizhou Yu. "Borrowing treasures from the wealthy: Deep transfer learning through selective joint fine-tuning." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

[3] Qu, Chen, et al. "Learning to selectively transfer: Reinforced transfer learning for deep text matching." Proceedings of the twelfth ACM international conference on web search and data mining. 2019.

[4] Ji, Mingi, Byeongho Heo, and Sungrae Park. "Show, attend and distill: Knowledge distillation via attention-based feature matching." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 35. No. 9. 2021.

[5] Jang, Yunhun, et al. "Learning what and where to transfer." *International Conference on Machine Learning*. PMLR, 2019.