Github page

BMVC 2021

Semi-Supervised Learning with Taxonomic Labels

Jong-Chyi Su   Subhransu Maji

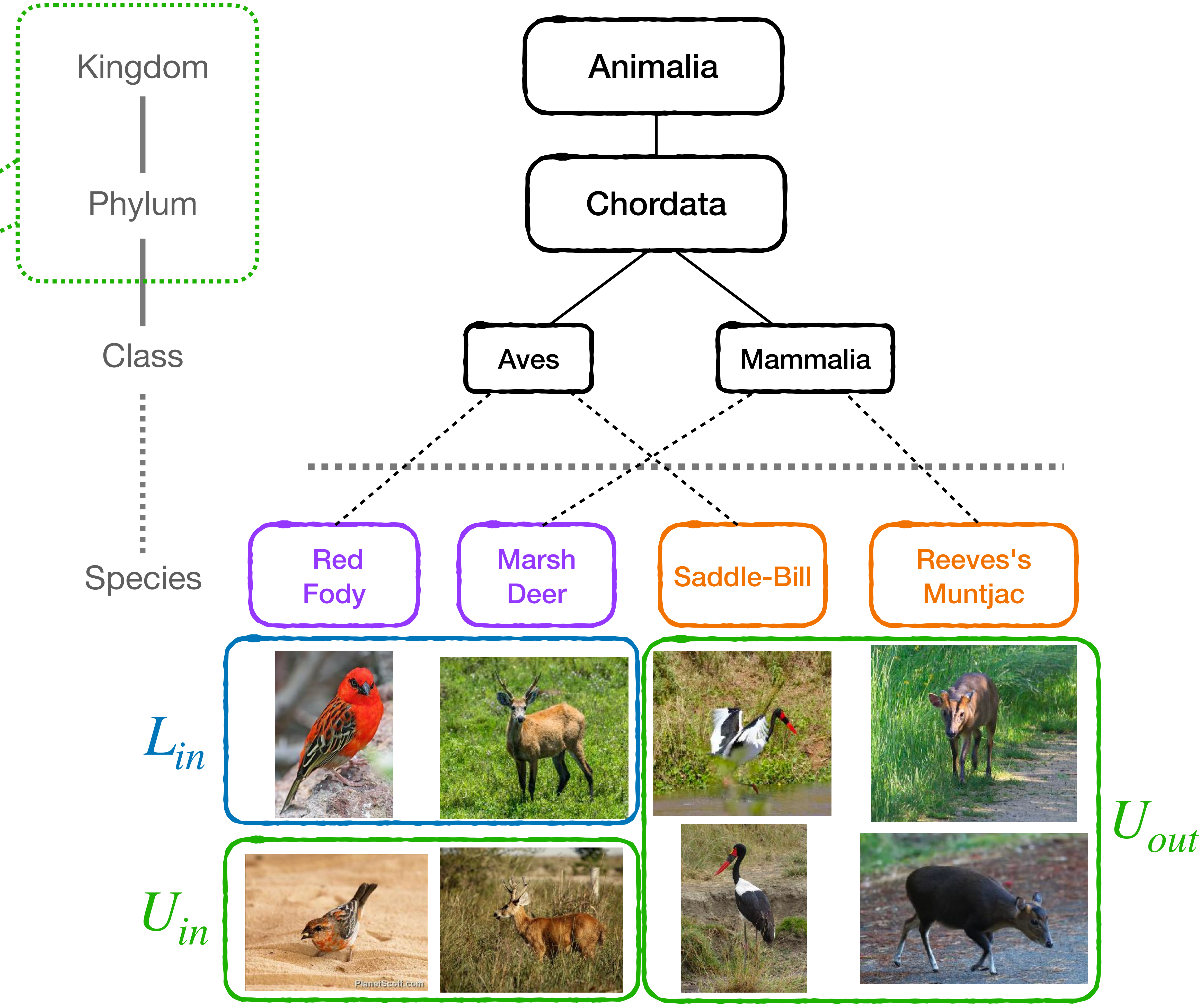
UMassAmherst


Motivation

- Acquiring annotations is time-consuming
  - especially in fine-grained domains
- Can we leverage coarse taxonomic labels to improve fine-grained classification?
- Can we incorporate them in a semi-supervised setting?

Dataset

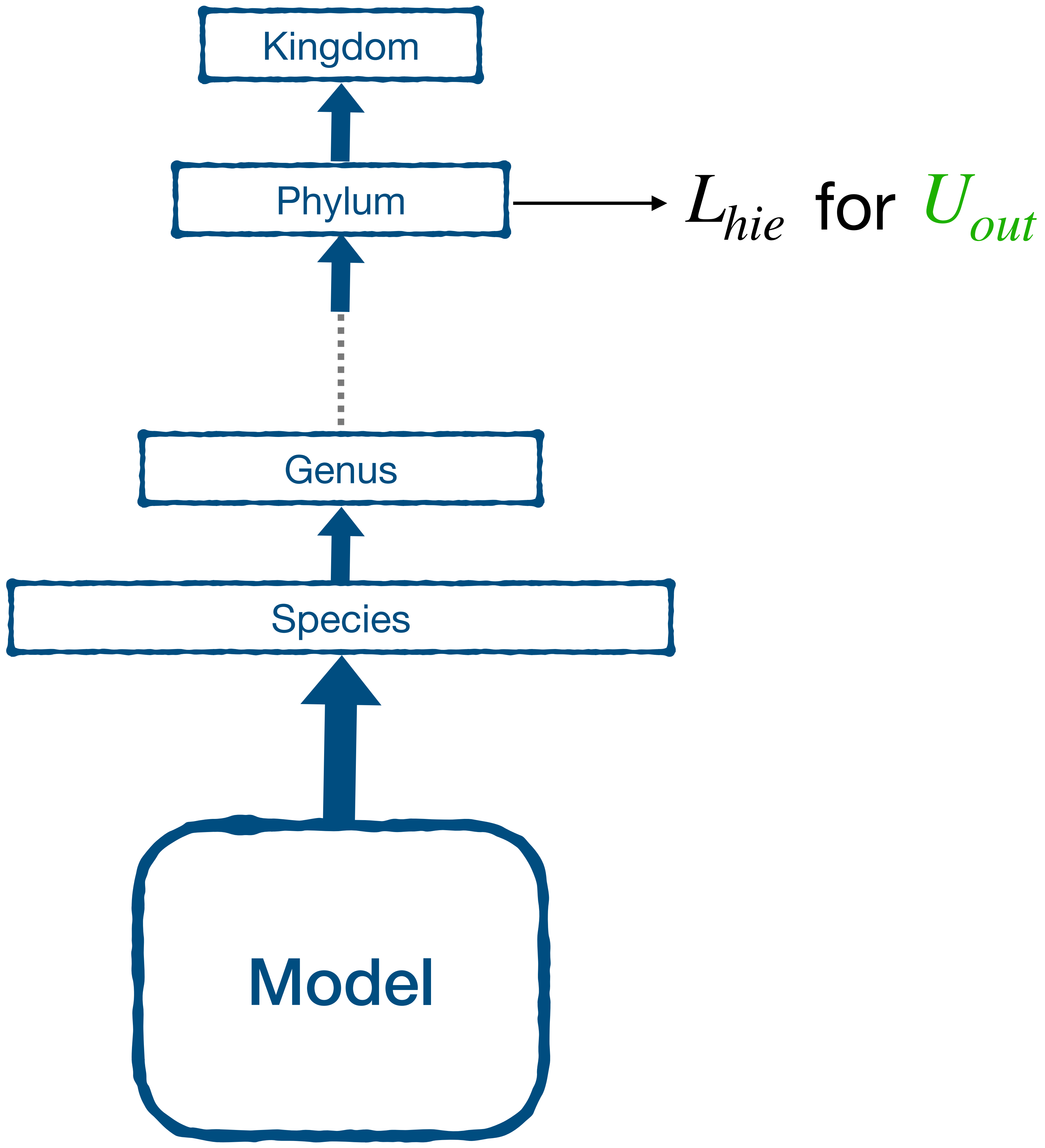
- Semi-iNat [1]
- 810 in-class species
  - labeled data L_{in}
  - coarsely labeled data U_{in}
- 1629 out-of-class species
  - coarsely labeled data U_{out}
- Test on in-class species



[1] Su and Maji, **The Semi-Supervised iNaturalist Challenge at the FGVC8 Workshop**, *arXiv:2106.01364*

Method

- Add cross-entropy loss on coarsely-labeled data
- First predicts on the finest level (*species*)
- Marginalize the probabilities
- Combine with semi-supervised losses

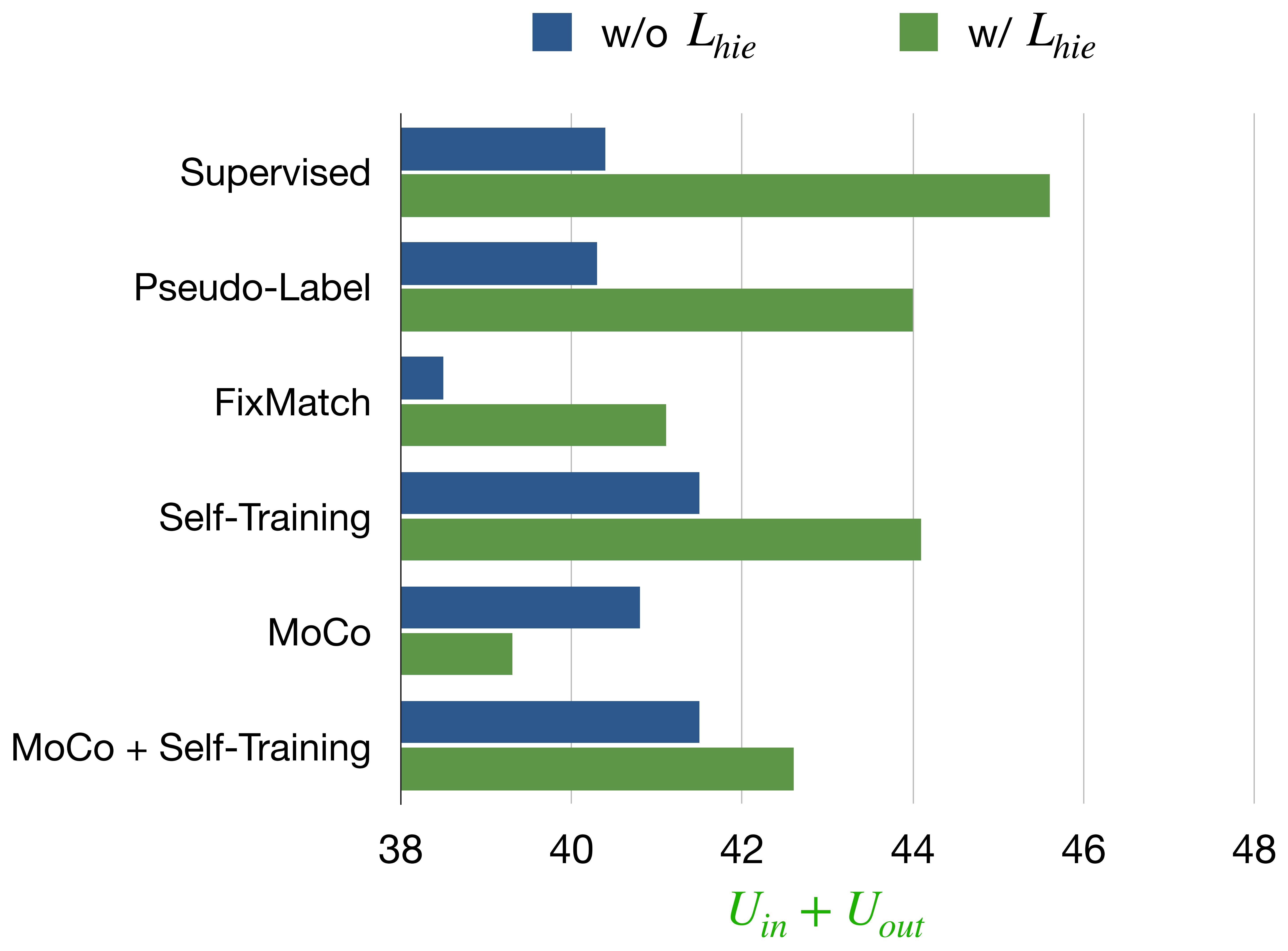


Adding hierarchical loss improves Semi-SL

- Having U_{in} only
- Improves all Semi-SL methods
- FixMatch performs the best

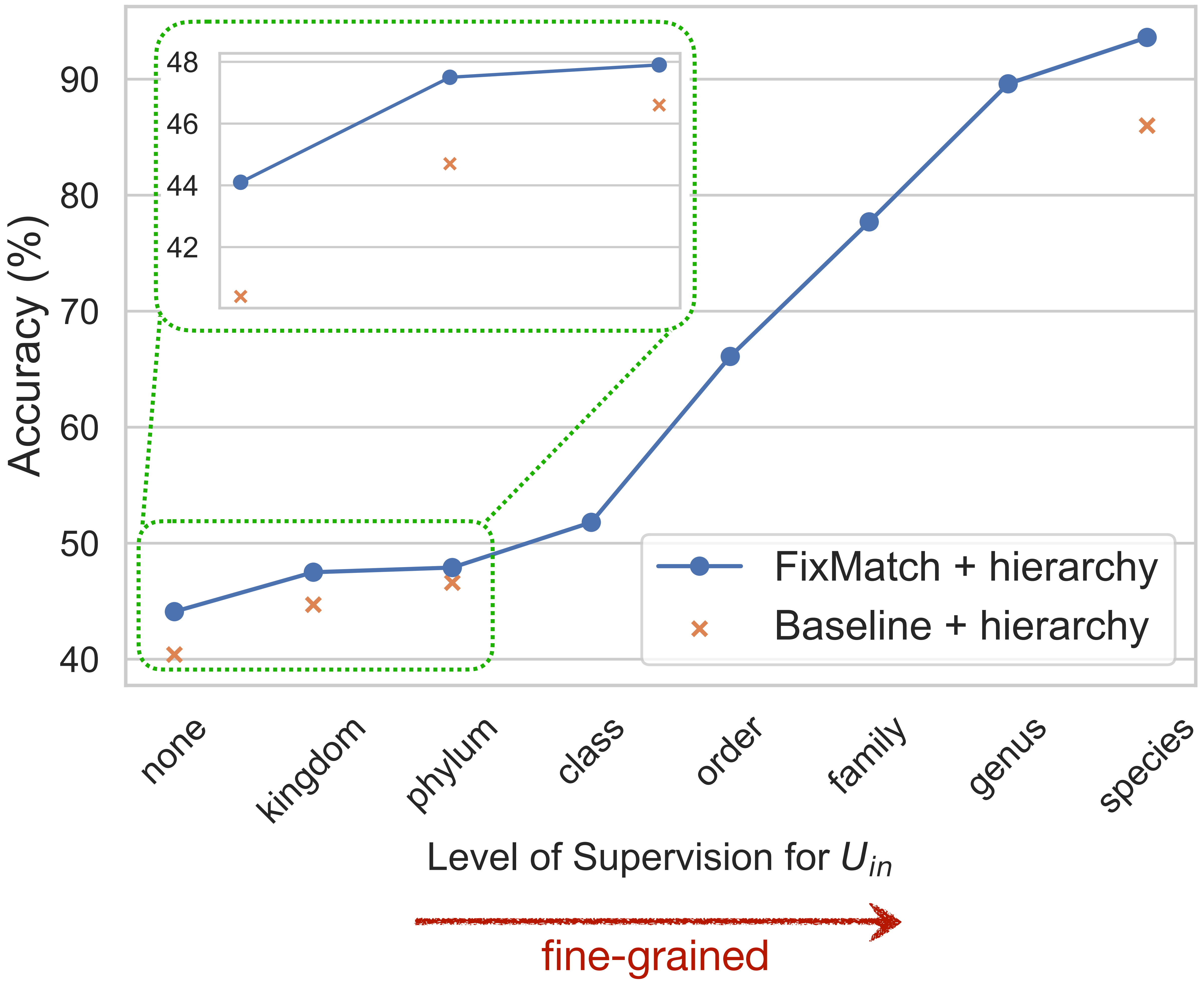
Semi-SL methods are *not* robust to domain shift

- Having $U_{in} + U_{out}$
- Less improvements
- Supervised + hierarchical loss performs the best



Different levels of hierarchical supervision

- Having U_{in} only
- Having more fine-grained labels improves the performance
- Semi-SL still gives improvements



Conclusion

- Coarse labels can improve fine-grained classification
- Semi-supervised learning also improves the performance
- However, Semi-SL is not robust to domain shift