Segmentation as Auxiliary Information for Reducing Spurious Correlations

Zev Nicolai-Scanio and Catherine Yeo
Harvard CS 282r | Spring 2022

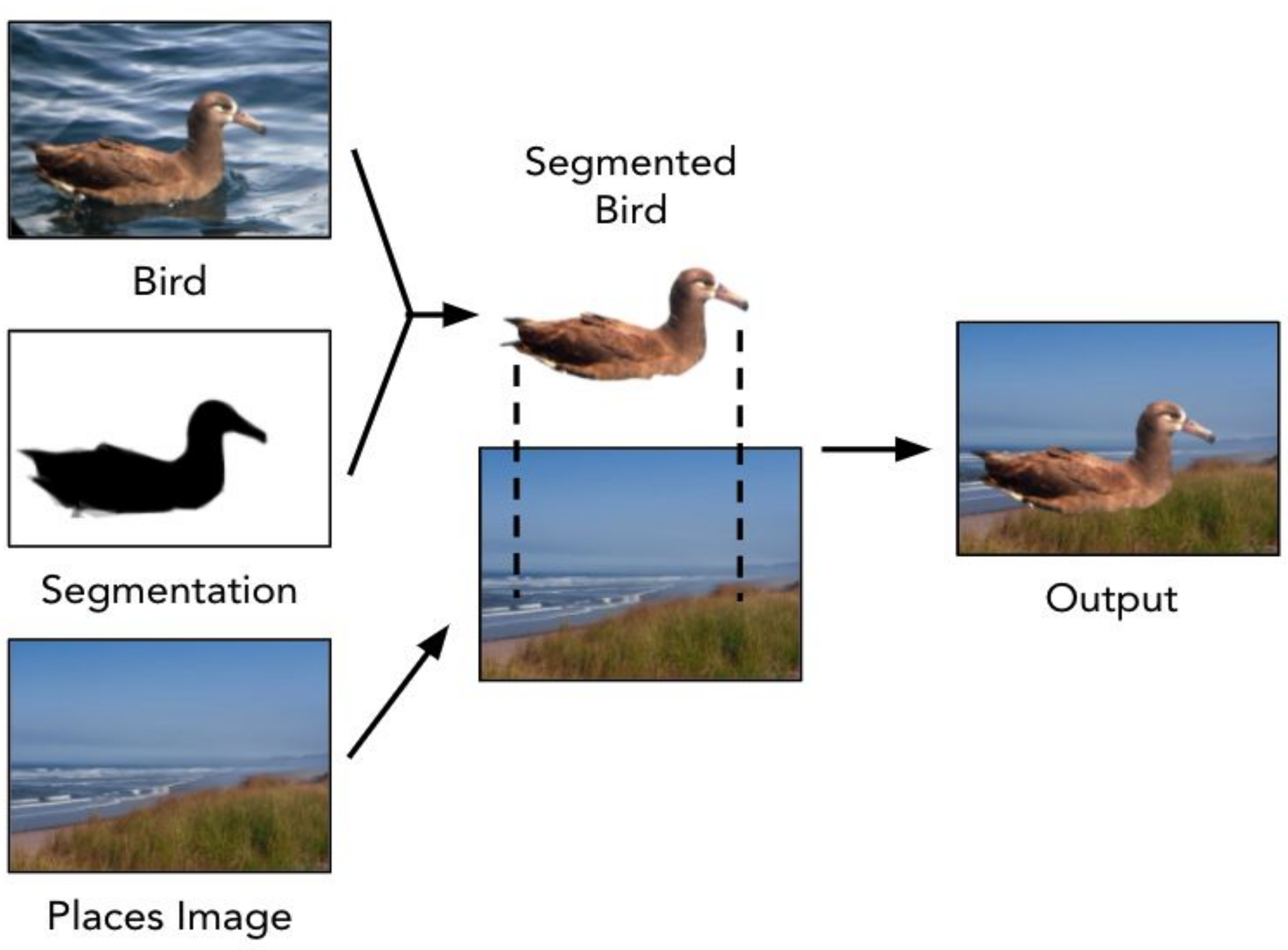
Harvard John A. Paulson School of Engineering and Applied Sciences

Background and Problem

- This project explores novel methods for training models to prevent heavy reliance on spurious correlations.
- We plan to test two hypotheses for reducing spurious correlation, focusing on image classification problems:
  - Including auxiliary structural information —specifically image segmentation information — with the input image can help the model leverage this to learn to avoid the spurious correlations.
  - By requiring the model to perform an additional task during training that requires structural understanding of the image, the model will learn internal representations that are less susceptible to spurious correlations.
- Issues with models learning spurious correlations that may be present in the training data can have serious detrimental effects to models' abilities to generalize. Our work would help contribute to better performance and more trustworthy models without significantly hindering inference-time performance.

Datasets and Data Pipeline

- The modified Waterbirds dataset is a synthetic dataset produced from the combination of the Birds-200-2011 dataset and Places dataset.
- There are 2 types of birds (water and land), and 2 types of backgrounds (water and land).
- The confounding percentage is the correlation between the bird type and background type.



Results

	Training Accuracy	Validation Accuracy
Baseline, 95% confounder	0.9283	0.8249
Baseline, 50% confounder	0.8717	0.9208
Gaussian noise, 95% confounder	0.9155	0.8265
Grayscale, 95% confounder	0.9195	0.8473

- Finetuned a ResNet-50 on the baseline and augmented datasets at the listed confounder correlation.
- We then run the model on held out data with uncorrelated backgrounds to get the validation accuracy.

Experimental Setup

1) Data Augmentation

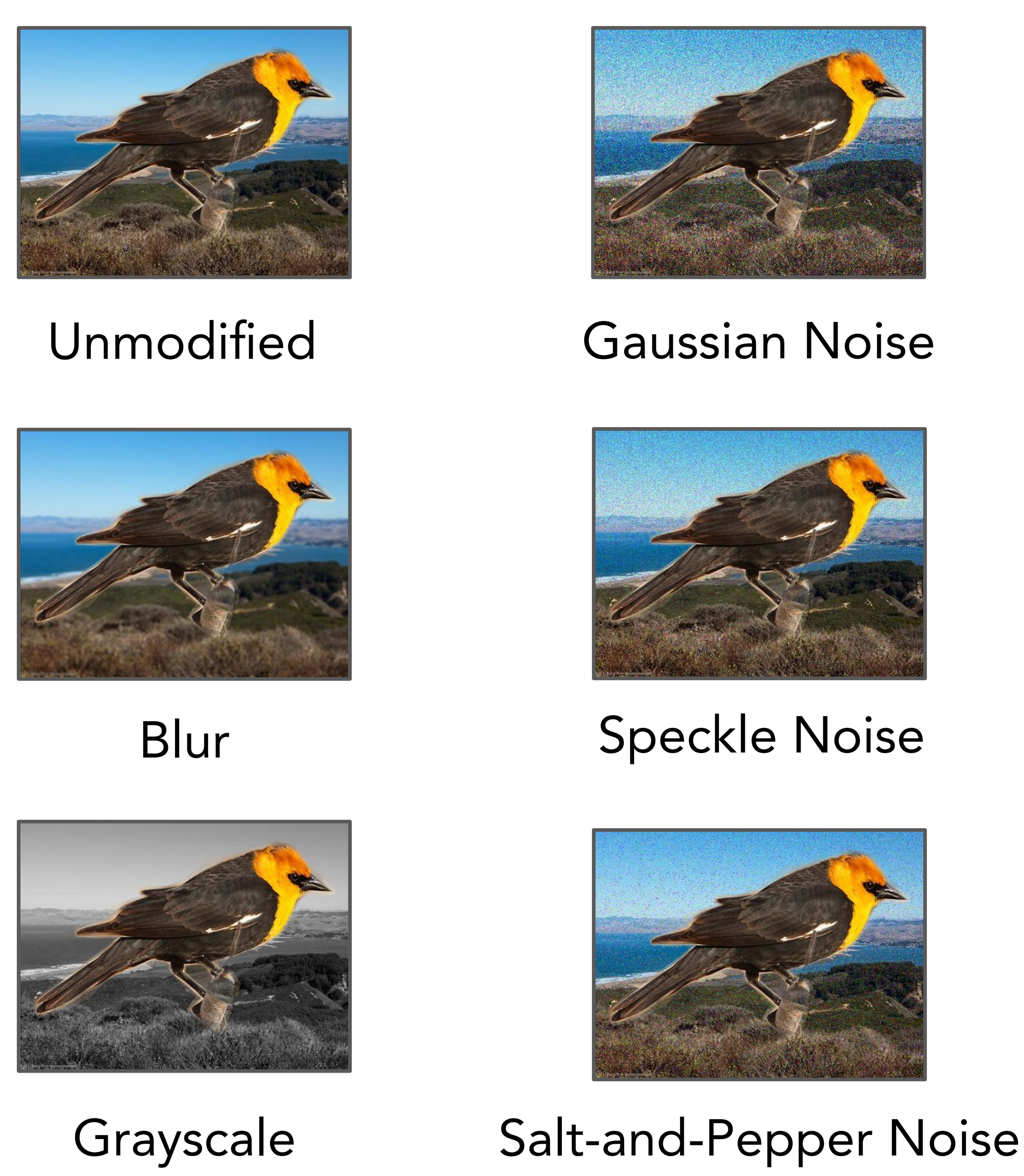
- Use image segmentation to inform data augmentation.
- Perform augmentations on the background places images, including: adding Gaussian noise, converting to grayscale, adding blur, etc.

2) Segmentation as Input

- Add segmentation mask of bird as an input channel to the model.
- Idea: mask will serve as information on what part of image model should care about

3) Segmentation as Output

- Consider segmentation as an output.
- Build a multi-headed model using ResNet-50 backbone in DeepLabv3 for both classification and segmentation tasks.
- Idea: segmentation task will result in internal representations that are more robust to the confounder



References

[1] Makar, M., Packer, B., Moldovan, D., Blalock, D., Halpern, Y., and D'Amour, A. (2022). Causally Motivated Shortcut Removal Using Auxiliary Labels.
[2] Sagawa, S., Koh, P. W., Hashimoto, T.B., and Liang, P. (2021). Distributionally Robust Neural Networks for Group Shifts: On the Importance of Regularization for Worst-Case Generalization.