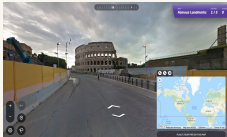
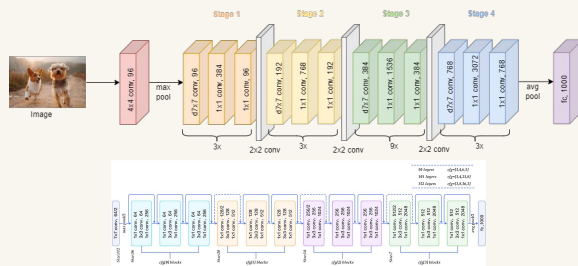


Motivation

The game GeoGuessr has players guess their location in Street View based on their surroundings. Is it possible to train a computer vision model to do the same?

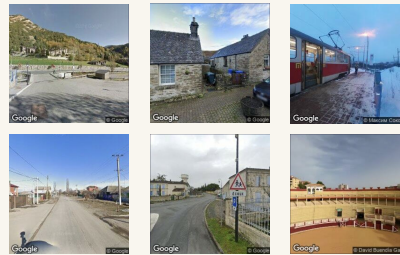


Architectures



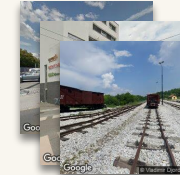
From top to bottom: ConvNext model and ResNet model used with BiT

Can you guess these countries?



Goal

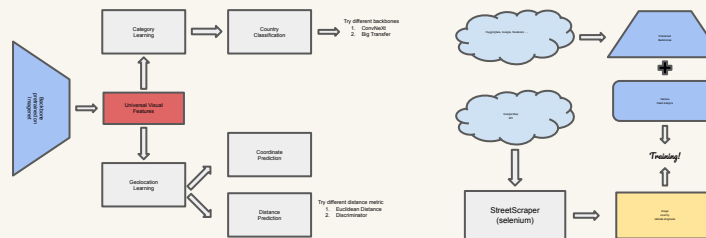
1. Finetune advanced CNNs to correctly classify country and predict coordinates/distance from Street View images above chance levels
2. Compare performance of ConvNeXt and BiT at these tasks



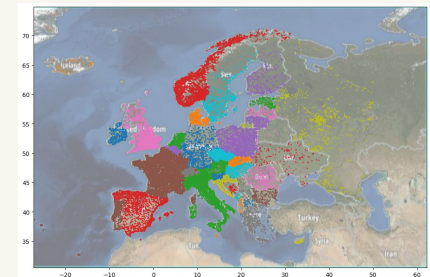
Coordinate

Country

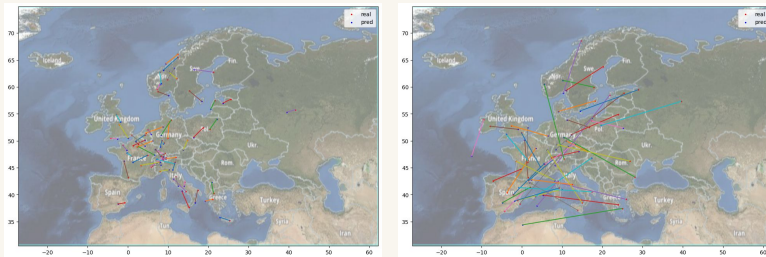
Methodology



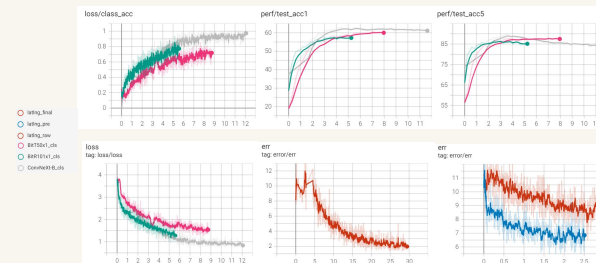
From left to right: high level overview of task, webscraping pipeline, data distribution by country



Results



From left to right: Positive and negative examples from ConvNext's lat-lng prediction.



Top from left to right: train acc1, test acc1, test acc5
Bottom from left to right: classification loss, coordinate error, distance prediction

Stage	Training		Testing	
Method	Acc@1	Acc@5	Acc@1	Acc@5
ConvNeXt-B	85.1%	97.9%	61.2%	86.6%
BiT-R50x1	63.4%	90.2%	59.0%	87.7%
BiT-R101x1	63.6%	91.0%	57.4%	86.3%

Method	error (degree)	loss	≤5 on Eval
Original from Imagenet	8.8	63.6	24%
Pretrained on Countries	6.9	41.8	44%
best model(Original)	6.35	3.55	54%
best model(Pretrained)	4.89	19.72	60%

Method	error (degree)	Acc@1	Acc@5
Euclidean	4.56	5%	20%
Discriminator	2.93	11%	36%

From top to bottom: classification accuracies of different backbones, errors in coordinate prediction, errors in distance prediction.

References

- Alexander Kolesnikov, Lucas Beyer, Xiaohua Zhai, Joan Puigcerver, Jessica Yung, Sylvain Gelly, and Neil Houlsby. Big transfer (bit): General visual representation learning. 2020.
- Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, and Saining Xie. A convnet for the 2020s. 2022.
- Eric Müller-Budack, Kader Pustularen, and Ralph Ewerth. Geolocation estimation of photos using a hierarchical model and scene classification. In Computer Vision – ECCV 2018, pages 575–592. Springer International Publishing, 2018.

Acknowledgements

We would like to thank Professor Tompkin, our mentor Alex Choi, and the entire staff of CSCI 1430.