Snap and Snack: A Greedy Approach to Tuning a Food Classification Transfer Learning Task

Cristian Lomeli, Andrea Dahl, Kevin Tien

Overview

We propose and test a greedy approach to find a set of hyperparameters that outperform a baseline model, faster than other methods of hyper-parameter testing for a food image classification task.
We built an algorithm based on ResNet50 that, when presented an image of a prepared/cooked food, classifies the food as a specific dish.

Dataset

We use Muriz Serifovic's data scraper to obtain recipes and images from chefkoch.de, a German social media outlet for food. Our dataset contains 186 different food classes and to combat class imbalance, our dataset only includes classes with 100-125 images each.

Figure 1: Sample images of dishes from chefkoch.de

Methods - Architecture

We kept most of the architecture of ResNet50 untouched except for adding fully connected layers, dropout layers, and a softmax output layer. We modified the weights of Resnet50 by unfreezing some of the final convolutional blocks to allow improved fitting to our data-set. After allowing for trainable convolutional layers, the network was able to pick up on the higher level features of food items.

Methods - Tuning

After every 5 epochs, we alternated hyperparameters that we were focused on tuning, always picking the best performing option based on F1 score and continuing with it. We used an Adam Optimizer and categorical cross-entropy loss to train the model.

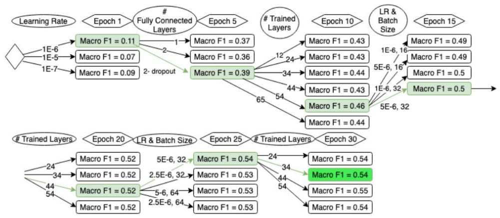

Figure 2: Flow chart showing the progression of hyperparameter tuning and selection throughout training epochs.

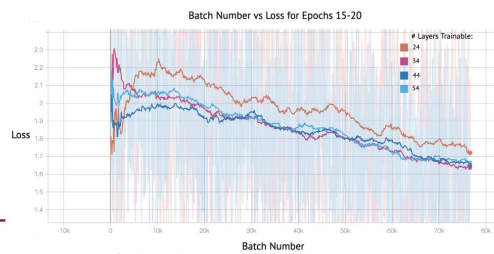

Figure 3: Defreezing the last 34 layers yielded lowest loss during training.

Figure 4: Class activation maps of last convolutional layer for several food items.

Results

We compared our model to the performance of 4 models initialized with random hyperparameter values also trained for 30 epochs. Our model slightly outperformed the other four models, suggesting our greedy approach to hyperparameter tuning was in fact effective.

Model and Hyper-Parameters	F1-Score
Model and Hyper-Parameters from our Greedy Approach	**.55**
lr=.00001, 2fc, 54 trainable, batch = 32	.54
lr=.00001, 1fc, 44 trainable, batch = 16	.54
lr=.000005, 2fc-dropout, 54 trainable, batch = 16	.54
lr=.00005, 2fc, 64 trainable, batch = 32	.53

Table 1: Greedy search approach outperforms random hyperparameter initializations.

Error Analysis

We found that the food items with the highest micro-F1 were items with unique textures: foods like Watermelon Shark(.96), Sushi(.86), and waffles(.84). The model performs poorly on foods like cakes (Lemon Cake(.23), Orange Cake(.00)), where discrepancies between flavors and styles might be much harder to distinguish and where the overall appearance of the different classes are extremely similar.

Figure 5: Sample images from best and worst performing classes

Future Work

Some logical next steps to improve our greedy algorithm and classification task involve tuning different hyperparameters, iterating over a different amount of epochs, and applying the model to other types of cuisine.
It would also be interesting to see how well the greedy algorithm generalizes to other image classification and deep learning tasks, as it can provide a computation and time effective method of training a model.

References

[1] Mezgec S., Korousic Seljak B. NutriNet: A Deep Learning Food and Drink Image Recognition System for Dietary Assessment. Nutrients. 2017;9:657. doi: 10.3390/nu9070657.

[2]He, Kaiming, et al. "Deep Residual Learning for Image Recognition." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, doi:10.1109/cvpr.2016.90.

[3] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna. "Rethinking the Inception Architecture for Computer Vision." IEEE Transactions on Pattern Analysis \& Machine Intelligence 7 arXiv:1512.00567 (2015)

[4] Kohavi, R., John, G.: Automatic Parameter Selection by Minimizing Estimated Error. In: Prieditis, A., Russell, S. (eds.) Proceedings of the Twelfth International Conference on Machine Learning, pp. 304–312. Morgan Kaufmann Publishers (1995)

[5] Smith, Samuel L., Pieter-Jan Kindermans, Chris Ying, and Quoc V. Le. "Don't decay the learning rate, increase the batch size." arXiv preprint arXiv:1711.00489 (2017).

Acknowledgments

We would like to thank our TA Sarah Najmark for her valuable guidance throughout the project.