Open-Set Classification with Binary Classifiers

A typical task for classification is to identify, which of a concrete set of classes a certain object belongs to. The classification of objects from images has a long history. Before the deep learning era, the main approach was to use a set of binary classifiers such as Support Vector Machines (SVMs). These classifiers were trained in a one-vs-all fashion, so that the samples for one class counted as positives, and the remaining classes contributed negatives. One such classifier was trained for each of the known classes. Finally, a test sample is predicted to belong to the class that has the highest classification score of its corresponding classifier.

When deep learning arose, these sets of binary classifiers got forgotten quickly. Instead, for a given sample, the network directly outputs a logit score for each class, which are finally transformed into probabilities using SoftMax. During training, Categorical Cross-Entropy loss is employed to train the network to output probability of 1 for the correct class, and 0 for all other classes. One particular issue (or design choice) of SoftMax is that the sum over all probabilities of all K classes is 1: $\sum_k p_k = 1$. This works well when the test sample belongs to one of the known classes, in which case the correct class should get probability 1, and the remaining classes get probability 0. However, when presenting a sample from an unknown class, in theory all classes should get probability of 0, but SoftMax prevents this from happening. Particularly, it happens rather often that such samples get assigned one of the known classes with high probability [Dhamija et al., 2018, Palechor et al., 2023].

One partial solution to this problem is to train the network to provide the lowest possible SoftMax score, which is $\frac{1}{K}$ [Dhamija et al., 2018]. However, this still requires probabilities to sum up to 1, and often one class has a high probability score, especially in difficult conditions [Palechor et al., 2023]. Another solution is to add an output class for unknown, which can either be trained using samples of unknown classes [Palechor et al., 2023] or estimated based on deep feature similarities [Bendale and Boult, 2016]. However, how to select effective unknown samples, or how to model deep feature similarities is difficult.

The solution that should be explored in this Master thesis is to incorporate a set of binary one-vs-all classifiers into deep network training. Particularly, the SoftMax activation, which requires exactly one output to be active, should be replaced with a set of binary classifiers, which allows several or even no outputs to be active at the same time. Since these binary classifiers will have highly unbalanced training data (K-1 out of K times, the training sample is negative), different loss functions that can deal with highly unbalanced data should be investigated, such as Focal Loss [Lin et al., 2017] or balancing weights [Rudd et al., 2016]. Also, hard negative mining strategies can be employed to reduce the number of negative samples.

Another way of including binary classifiers into open-set classification is to randomly split the training set into two classes multiple times, and train a separate binary classifier for each split [Vareto and Schwartz, 2020, Vareto et al., 2023]. For a sample of a known class, the probability for each of the random partition containing this class should be high, while for unknown samples, probabilities will be distributed more uniformly across splits. While we have recently shown this to work for face recognition [Vareto et al., 2023], a task with very restricted inputs, its extensibility to large-scale image classification should be investigated.

Experiments will be conducted on our newly-developed open-set evaluation framework [Palechor et al., 2023] on ImageNet, which allows testing on various difficulties of open-set problems. Possibly, combinations of the newly-trained networks with existing modeling techniques can be investigated [Bisgin et al., 2023]. Since these protocols also allow training on some unknown samples, ideas on how to incorporate these samples into the training process shall be developed.

Requirements:

- Successful participation in my Deep Learning course.
- A reasonable understanding of deep neural networks and how they learn.
- Programming experience in python and a deep learning framework, optimally, pytorch.

References

[Bendale and Boult, 2016] Bendale, A. and Boult, T. E. (2016). Towards open set deep networks. In *Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE.

[Bisgin et al., 2023] Bisgin, H., Palechor, A., Suter, M., and Günther, M. (2023). Large-scale evaluation of open-set image classification techniques. *Journal of Machine Learning Research (JMLR)*. **under submission**.

[Dhamija et al., 2018] Dhamija, A. R., Günther, M., and Boult, T. E. (2018). Reducing network agnostophobia. In *Advances in Neural Information Processing Systems (NeurIPS)*.

[Lin et al., 2017] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollar, P. (2017). Focal loss for dense object detection. In *International Conference on Computer Vision (ICCV)*.

[Palechor et al., 2023] Palechor, A., Bhoumik, A., and Günther, M. (2023). Large-scale open-set classification protocols for ImageNet. In *Winter Conference on Applications of Computer Vision (WACV)*. CVF/IEEE.

[Rudd et al., 2016] Rudd, E. M., Günther, M., and Boult, T. E. (2016). MOON: A mixed objective optimization network for the recognition of facial attributes. In *European Conference on Computer Vision (ECCV)*. Springer.

[Vareto et al., 2023] Vareto, R. H., Günther, M., and Schwartz, W. R. (2023). Toward open-set face recognition with neural ensemble, maximal entropy loss and feature augmentation. In *Conference on Graphics, Patterns and Images (SIBGRAPI)*. **to appear**.

[Vareto and Schwartz, 2020] Vareto, R. H. and Schwartz, W. R. (2020). Unconstrained face identification using ensembles trained on clustered data. In *International Joint Conference on Biometrics*. to appear.