Proceedings of Machine Learning Research 191:1–3, 2022 ECMLPKDD Workshop on Meta-Knowledge Transfer

Experiments in Cross-domain Few-shot Learning for Image Classification

Hongyu Wang HW168@STUDENTS.WAIKATO.AC.NZ
Huon Fraser HF71@STUDENTS.WAIKATO.AC.NZ
University of Waikato, New Zealand
Henry Gouk HENRY.GOUK@ED.AC.UK
University of Edinburgh, Scotland
Eibe Frank EIBE@WAIKATO.AC.NZ
Bernhard Pfahringer BERNHARD@WAIKATO.AC.NZ
Michael Mayo MMAYO@WAIKATO.AC.NZ
Geoff Holmes GEOFF@WAIKATO.AC.NZ
University of Waikato, New Zealand

Editors: P. Brazdil, J. N. van Rijn, H. Gouk and F. Mohr

Abstract

We summarise experiments (Wang et al., 2022) evaluating cross-domain few-shot learning (CDFSL) with feature extractors trained on ImageNet. The work explores the transfer performance of extracted features on five target domains with different degrees of similarity to ImageNet. These experiments compare robust classifiers and normalisation methods, consider multi-instance learning algorithms, and evaluate the effect of using features extracted by different ResNet backbones at various levels of their convolutional hierarchies. The cosine similarity classifier and 1-vs-rest logistic regression with ℓ_2 regularisation are the top-performing robust classifiers in the evaluation, and ℓ_2 normalisation improves performance on all five target domains when using LDA as the robust classifier. The results also show that feature extractors with the highest capacity do not always achieve the best CDFSL performance. Lastly, simple multi-instance learning methods are shown to improve classifier accuracy.

Keywords: Cross-Domain Few-Shot Learning, Pretrained Feature Extractors, Normalisation

1. Introduction

Modern computer vision models are usually deep convolutional neural networks with high model capacity. While capable of achieving good image classification performance, they generally require a large number of labelled training images to be effective. Training such models on a small labelled image set from scratch usually leads to subpar performance due to their capacity to overfit. Transfer learning can be used to combat overfitting in such scenarios, i.e., a model pretrained on a pre-existing large labelled dataset (source domain) is refitted to the small dataset (target domain). Cross-domain few-shot learning (CDFSL) encompasses transfer learning tasks where target domain instances are especially scarce (usually in the tens or lower hundreds), and source and target domain distributions are explicitly different. This extended abstract summarises work presented in (Wang et al., 2022), where we evaluate methods and techniques used in a common practical CDFSL

Table 1: CDFSL accuracy with different methods and techniques

Method	CropDisease	EuroSAT	Food101	ISIC	ChestX
No normalisation	93.74±0.46	86.74±0.49	77.90±0.71	42.18±0.56	24.89±0.41
ℓ_2 normalisation	**93.95±0.45**	**87.30±0.48**	**79.06±0.69**	**42.48±0.56**	**25.11±0.40**
ResNet101-conv4	**95.55±0.36**	**90.69±0.41**	69.93±0.73	**45.33±0.56**	**25.98±0.42**
ResNet152-conv4	95.47±0.38	90.24±0.41	70.92±0.73	45.15±0.56	25.51±0.42
ResNet101-conv5	92.13±0.49	86.57±0.50	78.57±0.71	42.70±0.56	24.63±0.40
ResNet152-conv5	92.51±0.48	85.54±0.54	**80.01±0.70**	42.34±0.57	25.22±0.43
1-vs-rest logistic ℓ^2	**93.79±0.46**	87.62±0.47	**79.42±0.69**	**43.28±0.56**	**25.30±0.42**
Cosine similarity	93.77±0.47	**87.67±0.47**	79.34±0.69	42.98±0.56	25.22±0.42
Baseline mono-instance	92.32±0.47	86.00±0.50	**81.26±0.64**	42.98±0.57	25.29±0.44
Simple multi-instance	**93.94±0.41**	**87.26±0.51**	80.56±0.66	**44.07±0.50**	**25.81±0.42**

scenario: a robust classifier is trained on target domain feature vectors extracted by a model, pretrained on source domain data, that is fixed after pretraining to minimise overfitting. We measure performance of different classifiers, normalisation methods, and feature extractor configurations, as well as multi-instance learning.

2. Methods and Experiments

We use the CDFSL benchmark by Guo et al. (2020), with ImageNet as its source domain, and four target domains: CropDisease, EuroSAT, ISIC, and ChestX. We include Food101 (Bossard et al., 2014) as an additional target domain. We evaluate the following robust classifiers: logistic regression, linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), random projection ensembles with LDA, linear SVMs, Naïve Bayes, k-nearest neighbours, random forests, nearest shrunken centroid, and cosine similarity classifier. The effect of normalising feature vectors with ℓ_p-norm is also evaluated. A set of ResNet (He et al., 2015) models of various sizes (18, 34, 50, 101, and 152) are used as feature extractors, and experiments are conducted to determine the optimal feature extractor and level of convolution for each target domain. We also evaluate multi-instance learning methods in this CDFSL context using weakly augmented target domain instances.

3. Results and Conclusions

We find that ℓ_2 normalisation improves LDA performance on all five target domains, as shown by the first group of rows in Table 1, where features are extracted using the highest-capacity model available. We also find that features extracted by a ResNet101 model at its penultimate convolutional block lead to better performance on four of the five target domains (except Food101) than features extracted by the highest-capacity model, i.e., a ResNet152 model at its final convolutional block, as shown by the second group of rows in Table 1, where a cosine similarity classifier is applied to the features.

We also find that the cosine similarity classifier and 1-vs-rest logistic regression with ℓ_2 regularisation are consistently top performers amongst the algorithms evaluated, indicating that the old and established ℓ_2-regularised logistic regression is a viable alternative to the newly popular cosine similarity classifier in CDFSL. The third group of rows in Table 1 show results for these two methods. Our multi-instance experiments show that simple multi-instance learning methods applied with weakly-augmented bags of instances improve accuracy for most of the target domains, as shown by the last group of rows in Table 1. Detailed comparisons of the algorithms are available in (Wang et al., 2022).

References

Lukas Bossard, Matthieu Guillaumin, and Luc Van Gool. Food-101 – mining discriminative components with random forests. In *ECCV*, pages 446–461, 2014.

Yunhui Guo, Noel C. Codella, Leonid Karlinsky, James V. Codella, John R. Smith, Kate Saenko, Tajana Rosing, and Rogerio Feris. A broader study of cross-domain few-shot learning. In *ECCV*, 2020.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CVPR*, pages 770–778, 2015.

Hongyu Wang, Henry Gouk, Huon Fraser, Eibe Frank, Bernhard Pfahringer, Michael Mayo, and Geoffrey Holmes. Experiments in cross-domain few-shot learning for image classification. *Journal of the Royal Society of New Zealand*, pages 1–23, 2022.