Subfigure and Multi-Label Classification using a Fine-Tuned Convolutional Neural Network

Ashnil Kumar[1], David Lyndon[1], Jinman Kim[1], and Dagan Feng[1,2]

[1] School of Information Technologies, University of Sydney, Australia
[2] Med-X Research Institute, Shanghai Jiao Tong University, China
{ashnil.kumar}@sydney.edu.au

Abstract. This paper describes the submission of the BMET group to the Subfigure Classification and Multi-Label Classification tasks of the ImageCLEF 2016 medical subtrack. Our method creates a new optimised feature extractor by using medical images to fine-tune a CNN that has been pre-trained on general image data. Our classification method shows promising result in both the the subfigure classification and multi-label classification subtasks.

Key words: convolutional neural network, fine-tuning, subfigure classification, multi-label classification

1 Introduction

This paper describes the submission of the BMET group to two of the the ImageCLEF 2016 Medical Tasks: Subfigure Classification and Multi-Label Classification [1, 2]. A primary challenge of these tasks is to automatically extract relevant representations (content and semantics) from the image data that allow easy differentiation of different modalities [3]. Previous attempts combined a vast range of image derived features that were sampled both globally over the whole image and locally over several different sub-patches [4, 5]. These features were designed by humans to represent some characteristic of the underlying image data, e.g., textures, colours, binary patterns.

Convolutional neural networks (CNNs) were used to optimise feature extraction for the ImageCLEF Medical Tasks in 2015 [6]. In our prior work, we designed a new CNN for both modality classification [7] and x-ray body region identification [8]. Choi [9] used generic features learned by a CNN from a large, well-labelled natural image dataset. However, the size of the challenge dataset limited the ability to learn the best image features.

In this paper, we describe a method for modality classification that uses a smaller medical image dataset to fine-tune (optimise) a CNN that was pre-trained on a large natural image dataset. This method allows us to adapt or adjust the generic features learned from natural images to be more specific for the medical imaging modalities in the ImageCLEF datasets. We apply our method to the Subfigure Classification and Multi-Label Classification tasks.

2 Materials

We used the Subfigure Classification training dataset (6776 images, 30 classes) to fine-tune our CNN. The test datasets consisted of 4166 images for Subfigure Classification and 1084 images for Multi-Label Classification. A full description of the datasets can be found in the ImageCLEF 2016 overview papers [1, 2].

3 Methods

We used the well-established AlexNet architecture [10] pre-trained (initialised) on the ImageNet natural image dataset (1000 classes, > 1 million samples) [11]. We fine-tuned the initial AlexNet filter weights (derived from natural images) for 100 epochs using back-propagation so that they were more appropriate for the 30 classes in the dataset. Dropout was used to avoid overfitting.

We increased the robustness of our algorithm to translation and orientation using a 24-fold data augmentation scheme. We generated 6 crops (original, top left, top right, bottom left, bottom right, centre) and 4 reflections (no flip, x axis, y-axis and both axes) of each training sample; 90% of the augmented dataset was used for fine-tuning and 10% for validation.

The fine-tuned CNN produced a 4096-dimensional feature vector for each input image. To improve efficiency, we reduced the dimensionality using Principle Component Analysis (PCA) [12] to select the principle components that explained 99% of the variation in the data (1453 dimensions).

We trained a multi-class support vector machine (SVM) using the PCA-reduced features extracted from all 24 augmented variations of the training dataset. During classification, we generated the feature vectors for each test image and its 5 crops, and used the SVM to obtain the posterior probability and per-class score that each crop depicted a particular modality. When using per-class SVM scores, we linearly scaled them to the range $[0, 1]$ to reduce the impact of very large outlier scores. We investigated several different schemes to determine the class of an input image, as described in our runs.

We implemented our method in MATLAB, using the MatConvNet library [13] for our implementation of CNN fine-tuning. For our experiments we used the pre-trained AlexNet provided as a part of MatConvNet.

4 Runs

We submitted 5 runs to the Subfigure Classification (SC) task and 2 runs to the Multi-Label Classification (ML) task.

SC1 Mean SVM posterior probabilities of all crops.
SC2 Mean of the per-class SVM score, which were scaled across all crops.
SC3 Mean of the per-class SVM scores, scaled separately for each crop.
SC4 Majority class across all crops. The per-class scores were not scaled.
SC5 Maximum SVM posterior probability across all crops.
ML1 For each crop, the label was the modality with the highest SVM score.
ML2 For each crop, the label was the modality with the highest probability.

Table 1: Subfigure Classification

Run	Type	Correctness (%)
SC1	visual	77.55
SC2	visual	77.53
SC3	visual	77.50
SC4	visual	77.26
SC5	visual	76.38

Table 2: Multi-Label Classification

Run	Hamming Loss	F-Measure
ML1	0.0131	0.295
ML2	0.0135	0.320

5 Results and Discussion

Table 1 shows the results of our Subfigure Classification runs. The best outcome came from averaging the posterior probabilities calculated from classifying each crop. Table 2 shows the results of our Multi-Label Classification runs. The low Hamming Loss indicates that our runs had very few incorrectly predicted labels.

6 Conclusions

We presented a method for subfigure modality classification and multi-label classification that used fine-tuned CNNs as a feature extractor. We expect improved results through the use of deeper CNNs such as Deep Residual Networks [14].

Acknowledgments. We gratefully acknowledge the support of NVIDIA Corporation through the donation of the Tesla K40 GPU used for this research.

This work was supported in part by ARC grants. This work was also supported by the Faculty of Engineering and Information Technologies, The University of Sydney, under the Faculty Research Cluster program.

References

1. Villegas, M., Müller, H., García Seco de Herrera, A., Schaer, R., Bromuri, S., Gilbert, A., Piras, L., Wang, J., Yan, F., Ramisa, A., Dellandrea, E., Gaizauskas, R., Mikolajczyk, K., Puigcerver, J., Toselli, A.H., Sánchez, J.A., Vidal, E.: General Overview of ImageCLEF at the CLEF 2016 Labs. Lecture Notes in Computer Science. (2016)
2. García Seco de Herrera, A., Schaer, R., Bromuri, S., Müller, H.: Overview of the ImageCLEF 2016 Medical Task. In: CLEF2016 Working Notes. (2016)

3. Kumar, A., Kim, J., Cai, W., Feng, D.: Content-based medical image retrieval: a survey of applications to multidimensional and multimodality data. Journal of Digital Imaging **26**(6) (2013) 1025–1039
4. Pelka, O., Friedrich, C.M.: FHDO biomedical computer science group at medical classication task of ImageCLEF 2015. In: CLEF 2015 Working Notes. Volume 1391. (2015)
5. Abedini, M., Cao, L., Codella, N., Connell, J.H., Garnavi, R., Geva, A., Merler, M., Nguyen, Q.B., Pankanti, S.U., Smith, J.R., Sun, X., Tzadok, A.: IBM research at ImageCLEF 2013 medical tasks. In: CLEF 2013 Working Notes. Volume 1179. (2013)
6. Villegas, M., Müller, H., Gilbert, A., Piras, L., Wang, J., Mikolajczyk, K., de Herrera, A.G.S., Bromuri, S., Amin, M.A., Mohammed, M.K., Acar, B., Uskudarli, S., Marvasti, N.B., Aldana, J.F., del Mar Roldán García, M.: General Overview of ImageCLEF at the CLEF 2015 Labs. Lecture Notes in Computer Science. (2015)
7. Lyndon, D., Kumar, A., Kim, J., Leong, P.H.W., Feng, D.: Convolutional neural networks for medical classification. In: CLEF 2015 Working Notes. Volume 1391. (2015)
8. Lyndon, D., Kumar, A., Kim, J., Leong, P.H.W., Feng, D.: Convolutional neural networks for medical clustering. In: CLEF 2015 Working Notes. Volume 1391. (2015)
9. Choi, S.: X-ray Image Body Part Clustering using Deep Convolutional Neural Network: SNUMedinfo at ImageCLEF 2015 Medical Clustering Task. In: CLEF 2015 Working Notes. (2015)
10. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In Pereira, F., Burges, C., Bottou, L., Weinberger, K., eds.: Advances in Neural Information Processing Systems 25. Curran Associates, Inc. (2012) 1097–1105
11. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L.: Imagenet large scale visual recognition challenge. International Journal of Computer Vision **115**(3) (2015) 211–252
12. Jolliffe, I.: Principal component analysis. Wiley Online Library (2002)
13. Vedaldi, A., Lenc, K.: MatConvNet – Convolutional Neural Networks for MATLAB. In: Proceeding of the ACM Int. Conf. on Multimedia. (2015) 689–692
14. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. arXiv preprint arXiv:1512.03385 (2015)