A Transfer Learning Approach for Park Classification Using Images

Si Chen*
sichen12@illinois.edu
Department of Computer Science, University of Illinois at Urbana-Champaign
Urbana, Illinois, USA

Mario Lopez Gonzalez
mariol4@illinois.edu
Department of Computer Science, University of Illinois at Urbana-Champaign
Urbana, Illinois, USA

Hongpeng Guo
hg5@illinois.edu
Department of Computer Science, University of Illinois at Urbana-Champaign
Urbana, Illinois, USA

Klara Nahrstedt
klara@illinois.edu
Department of Computer Science, University of Illinois at Urbana-Champaign
Urbana, Illinois, USA

Brian Deal
deal@illinois.edu
Department of Landscape Architecture, University of Illinois at Urbana-Champaign
Champaign, Illinois, USA

Abstract

Deep convolutional neural network (CNNs) have proliferated in recent years as its great efficiency in robust feature extraction and information mining. This method has been widely applied in image classification as a state-of-the-art method in many fields, but less in the urban planning field, such as parks' image classification. Few studies have applied CNN-based or other computational models to classify parks, but still depend on traditionally manual classifications by local planners. To fill the gap, in this study, we propose a deep neural network framework based on transfer learning for park classification using geo-located images. The image dataset of urban parks in each domain are usually from local government or social media data, where the number and quality of images are limited. Another challenge in classifying park images is that the similarity between different parks' images are naturally high. To address these two challenges, we used transfer learning in the proposed framework to pretrain the CNN model to learn the generic feature extracting on a large and varied dataset before applying to the target dataset. Fine-tuning strategies and image augmentation strategies were applied in the proposed framework to better combine the target dataset knowledge with the pretrained CNN models and to decrease overfitting problems brought by small dataset. Since our final goal is to classify each park using a group of images, rather than the individual images, we designed a further step to aggregate the group of images' classification results. Experiment result on Seattle park images supported the effectiveness of the proposed park classification framework.

Keywords Image classification, Transfer learning, Small dataset, Park

1 Introduction

Classifying parks builds the basic principle for urban green space planning and design.The current park classifications used in urban planning field mostly depend on local planners manual labeling, using their empirical benchmarks, such as park size, vegetation, location and recreational activities. Such benchmarks vary among different cities, bringing bias and inefficiency. Some academic researchers have realized the gap, and attempted to apply data-based analysis in classifying parks or other places of interests (POIs). Ibes [11] used a multidimensional method to classify urban parks according to their physical, land cover, and built features, which however, still empirically selects features and weightings. Other studies tried to use big data, usually geo-located social media check-in or review data, to extract features for POIs [13] [8]. The geo-located images, however, were used in much fewer study in extracting POIs' features for classification, while they have potential to provide more robust and comprehensive information. For example, the landscape quality and vegetation features are usually missed in social media users' unprofessional descriptions, but provided in images.

For classifying images, deep convolutional neural networks (CNNs) is one of the most state-of-the-art method in various domains, having advanced development of computational capability. Deep CNNs typically can achieve high accuracy and efficiency in large dataset [21][22]. However,the park amount in each city domain usually only range from dozens to hundreds. And the images dataset for parks are mostly sourced from local governments' field studies or social media platforms. The number of open-access captured images from field studies or social media users for each parks are limited. So the qualified park image dataset in each city domain is generally not large, only having hundreds or thousands of images. Moreover, the quality of these images, according to the metrics of capture view, resolution and brightness are not guaranteed. The park images between different classifications also have high similarity. These challenges from park images may cause serious over-fitting problems when using deep CNNs for classification.

One popular approach to address small dataset issue in deep CNNs is transfer learning, which has a significant performance in keeping robustly high accuracy on small and low-varied training dataset[23][15]. In transfer learning, the deep network structure is pre-trained on a large image dataset to learn to extract generic features, and then applied on the small dataset in the target domain to further extract local features [15]. The transfer learning approach lowers the requirements of dataset size, variation, and hardware environment. These lower requirements increase models' adaptability and replicability across different city domains, since most urban models are difficult to be replicated because of being spatially sensitive and data hungry [25].Transfer learning, targeting for cross-domain tasks, can solve the obstacles in model replication. Few studies have applied transfer learning approach with deep CNNs in urban planning field for POIs' image classification, as urban planning is a relatively traditional and low-technique major.

In this study, we use three different deep CNNs: VGG16 as a feature extractor (VGG16 FE), VGG16 as a feature extractor with image augmentation(VGG16 FE AUG), and VGG16 with fine-tuning and with image augmentation(VGG16 FT AUG). All of the three CNNs are pre-trained on ImageNet dataset[3] and designed to incorporate both low- and high-level information. Moreover, since each park has a group of images, to aggregate the classification results of image groups for each geo-located park, we proposed and tested three aggregation approaches, namely absolute voting, sum of probabilities, and sum of squared probabilities.

This study focuses on the application of transfer learning based deep CNNs in an interdisciplinary field of computer science and urban planning to provide advanced spatial analysis, in order to better support the future planning policies and strategies. The proposed park classification framework has potential to be used to classify all types of geo-locations using images and provide better analytic support for many other fields, such as geography and landscape architecture majors. The park image dataset of Seattle City, Washington is chosen for testing our proposed park classification framework. The VGG16 FT AUG model has the best performance, achieving the accuracy of 73.34%, which is satisfying performance considering the whole dataset only has 122 images for 55 parks. The contributions of this study can be summarized as follows.

1) To our knowledge, this is the first attempt to use transfer learning with deep CNNs for park classifications. This interdisciplinary study also provides a pioneer framework in applying advanced computational approaches for other POIs' feature extraction and classification.
2) We integrate and test the fine-tuning and image augmentation strategies with pre-trained deep CNNs on park images to better address the issues of small size and high similarity in park image data.
3) We design the summing approaches to aggregate the individual image classification results to each geo-located park.

2 Related Work

2.1 Deep CNNs

Deep convolutional neural networks (CNNs) have been popular in recent years as its great efficiency in robust feature extraction and information mining, especially on image data. A typical CNNs architecture is composed of convolutional layers, pooling layers and fully-connected layers.The three main CNN architectures in existing studies include AlexNet[14], VGG16[20] and VGG19. The AlexNet[14] aims to solve the object recognition problem, which acts as the first try to learn the network parameters in the object recognition task on large-scale database. As Figure 1 (a) shows, AlexNet has overall 26 layers, which can be divided into three repetitive sections and the last two SoftMax and classification layers. Each section is repeated several times to better adapt the specific data and extract robust features. The last section is consist of fully connected, relu and drop-out layer, which correspond to the non-linear activation unit and deal with the over-fitting problem during training.

To improve the CNN accuracy, Simonyan[20] proposed VGG16 CNN architecture for object detection, which went deeper in the network in order to extract more complicated and robust features as Figure 1 (b) shows. Compared with AlexNet, VGG16 architecture has more repeating times for deeper network, more replicative structure, but smaller size of receptive window for each convolutiona filter. The main novel part of VGG16 was to make a thorough evaluation of networks with increased based on small (3x3) convolution filters. The method achieved a significant improvement on the prior-art configurations through pushing the depth to 16-19 weight layers [20].VGG 19 was proposed to further deepen network for object detection task. It has similar replicative architecture with VGG16, but has several additional convolutional and relu layers in the repeated sections located at the network's middle part [12]. In this paper, we mainly extended VGG16 with image augmentation, transfer learning and fine-tuning strategies to achieve more robust image classification. [16].

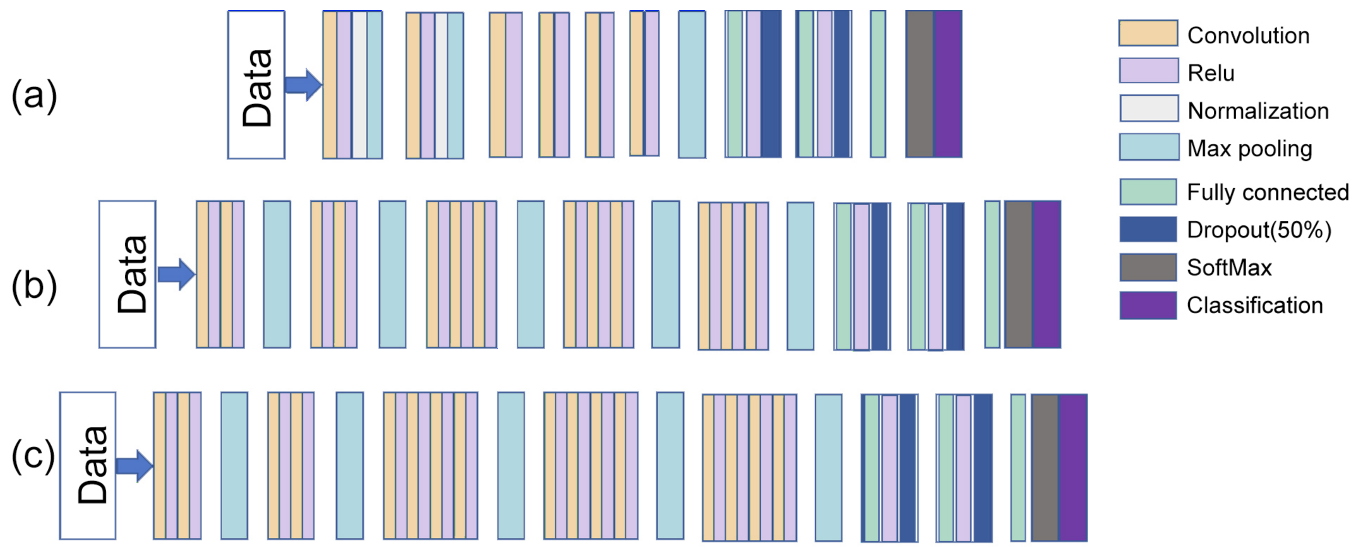

Figure 1. Three existing CNN architectures[16]. (a) AlexNet[14] (b) VGG16[20] (c) VGG19[12]

2.2 Transfer learning

Transfer learning targets to learn from related tasks and handle the different distribution issue across various domains. Since Ando and Zhang [1] explored to transfer knowledge from supervised learning to unsupervised learning and proposed that there were common structure of hypothesis spaces shared among the multiple tasks, a number of studies have developed multiple transfer learning methods to achieve cross-domain models [26][6]. For example, Tian, Tao and Rui[24] mapped source domain and target domain into a new space to reduce the bias caused by cross-domain correspondence to implement transfer learning. Their focus was to resort the sparse coding method to look for a proper mapping space.

In more recent years, a number of advanced transfer learning techniques [17][5][18] were proposed for processing text data from different languages or the image data from different domains. They can be summarized into two main categories: (1) to re-weight instances[2] and (2) to find common features [5][10]. The transfer learning approach used in Yosinski et al.[26]'s study was a popular one of the second category, which transferred the pre-learned knowledge from a large image dataset - ImageNet, to the target domain's image dataset, demonstrating the transferability of image features. In his study, fine-tuning strategies were also used in deep CNNs to better learn the features in the targeted datasets. The transfer learning used in this study can be seen as an extension based on Yosinski et al.[26]' study.

2.3 Image augmentation

Image augmentation technique is typically used to expand the dataset to decrease the over-fitting problems during training on small dataset. The main image augmentation approaches include flipping images, adding noise and applying affine transformations (translation, zoom, flips, shear, mirror and color perturbation)[19][7]. Many studies[27][9] have proved the efficiency of image augmentation on conbatting the over-fitting problems. For example, Yosinski et al.[26] augmented the MNIST dataset (a popular dataset containing handwritten digit images), through simply setting with randomly (by at most 5%) translated images,and found the testing error greatly decreased from 28% to 20%.

Other recent studies explored more image augmentation approaches. Goodfellow et al.[9] proposed an Generative Adversarial Networks (GANs) to regenerate graphs. The GANs was composed of two competing neural network models in order to drive the generated artificial data to be indistinguishable from real graphs. The experiments of GANs on MNIST showed good performance on re-generating high-quality artifical digit images. DeVries and Taylor[4] proposed a relatively simple, but highly applicable domain-agnostic approach for data augmentation. The proposed approach departed from existing data points and applied simple transformations, such as noise adding or extrapolating between them, which focused on performing the transformation in a learned feature space, rather than in the input space. These approaches have described multiple ways to implement image augmentation, and demonstrated the efficiency of image augmentation (even simple approach) in improving the model performance. In this study, we applied and tested the image flipping and distortion strategies with deep CNNs in the proposed park classification framework.

3 Methodology

Our proposed framework firstly classifies individual images, and then aggregates the classification results of image groups for each park to decide final park classification. For classifying the images, we designed and tested three deep CNNs based on the classical VGG16 model [20] and applied a transfer learning approach to all three CNNs.

3.1 Pretrained CNNs for image classification

In this section, we adopt three pre-trainned deep CNN architectures in classifying individual images, including VGG16 as a feature extractor, VGG16 as a feature extractor with image augmentation strategy, and VGG16 with fine-tuning and image augmentation strategies. We pre-trained the three deep CNNs on a nature image dataset (ImageNet) before training on the target park image dataset. ImageNet is a large open-access image dataset containing over 1 million natural images with 1000 categories, which has been used for pre-training in many studies [3]. Through pre-training, the deep CNNs could learn to extract distinct generic features. After pre-training, the deep CNNs can be applied to extract discriminative features from park images with no need to train from scratch on target tasks.

The pre-training, as a typical transfer learning approach, enables learning feature extraction from ImageNet and transferring to target park images. Two benefits can be obtained from applying transfer learning approach for our study. Firstly, transfer learning can significantly improve deep CNNs' performance on small dataset, because it helps to reduce overfitting through pre-training process[16]. Secondly, transfer learning lowers the requirement of training dataset size and variation for target task, because the model's ability to extract general features has been pre-trained. The training time on target task can also be reduced because only the last layers of the deep CNNs models need to be trained in new domains. So the proposed park classification framework can have better applicability and replicability when applied in different city domains since the lower requirement for new data and hardware environment. The three adopted deep CNNs are specifically described as follows:

VGG16 as a feature extractor (VGG16 FE). VGG16 FE adds a pre-trained process on VGG16[20], which keeps the same network structure with VGG16 as shown in 2. The fully connected layers, combined with Dense layers and Dropout layers, enables regularization and reduces over-fitting problem. In VGG16 FE, we replace VGG16's original top layers with our own fully connected layers, to use the pre-trained model as a feature extractor.

VGG16 as a feature extractor with image augmentation (VGG16 FE AUG). This approach also keeps the same structure with VGG16 FE as Figure 2 shows, but contain an extra image augmentation strategies. We flip and distort the existing images in the target domain to generate more images for training. The image flipping strategies we use include rotating, zooming in and out and shifting the image, as displayed in Figure 3. Image augmentation can efficiently

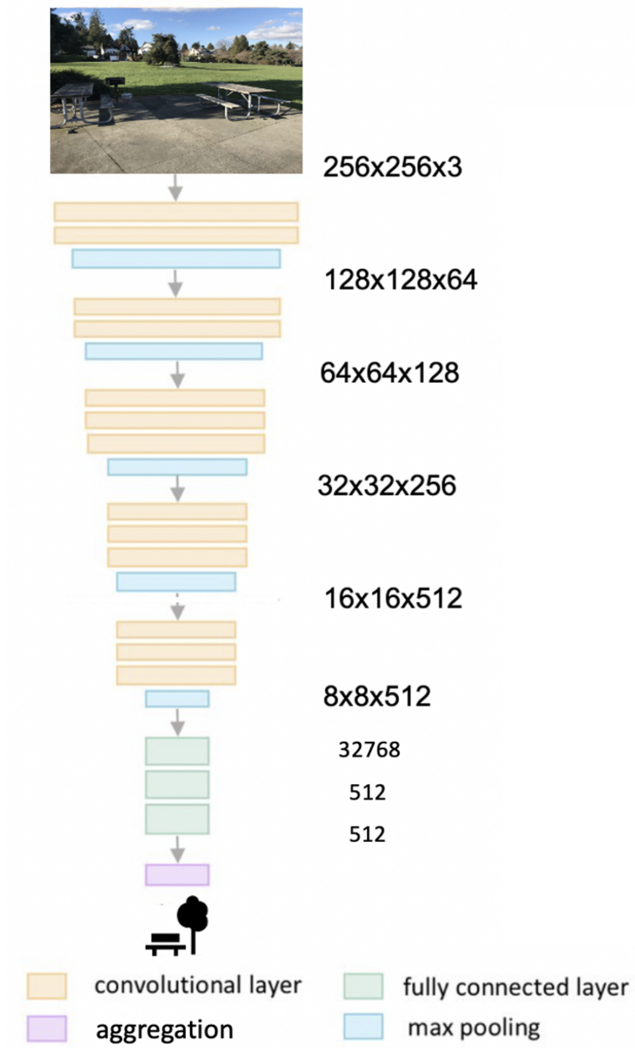

Figure 2. VGG16 FE structure

decrease the overfitting problem, which usually exists during training on small dataset [26]. It can also help the CNN models avoid learning irrelevant patterns, thus boosting overall performance.

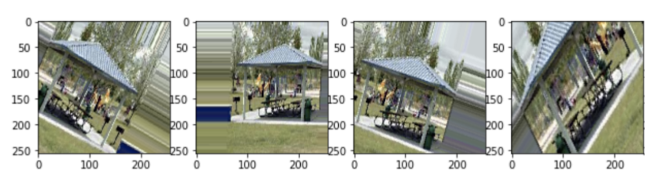

Figure 3. Example of Image Augmentation

VGG16 as a feature extractor with fine-tuning and with image augmentation (VGG16 FT AUG). We develop this model based on the VGG16 FE and VGG16 FE AUG. Except the pre-training and image augmentation, the fine-tuning strategy is added on some convolutional layers in VGG16 FT AUG. So we not only train the fully connected layers, but train several convolutional blocks as shown in Figure 4 on the target park image dataset. Through fine-tuning strategy, the coarse features that are detected at lower layers remain the same, but the last layers that extract more specific-domain features are re-trained on park images. That means the local features in target task can be better learned. However, if re-training too many convolutional blocks, the benefits of transfer learning would be decreased and more over-fitting problems would occur. After testing for the balance between local features learning and transfer learning benefits on park images, we found re-trained the last two convolutional blocks has the best performance.

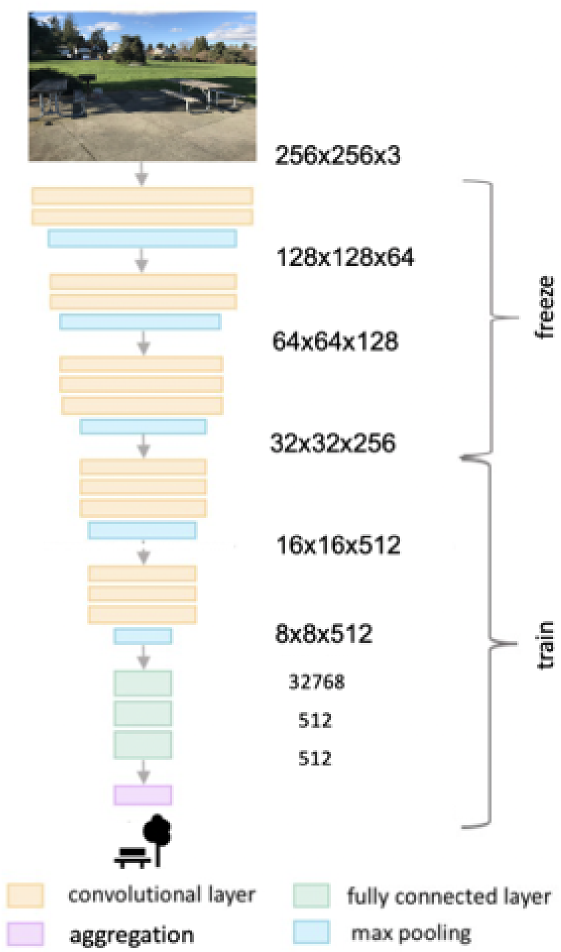

Figure 4. VGG16 FT AUG architecture

3.2 Aggregating image classification results to parks

Our final goal is to classify the geo-located park. Each park generally has more than one images. However, the deep CNNs' classifications are based on individual images. So we need to aggregate the group of individual image classification results to each park. We adopt and test three aggregation approaches, which are described as follows.

1) Absolute Voting, which selects the most frequently appeared classification label within the image group of each park, and assign that label as park classification. It is a simple and direct choosing method, which is expected to have worst performance. We adopt it as a baseline method for comparison purpose.
2) Sum of Probabilities, which firstly sums the probabilities of individual image classifications within the image group of each park and then select the classification label having the highest sum as park classification. This method considers the confidence of individual image classification, which represents their weight in the final classification decision for park.
3) Sum of Squared Probabilities, which uses the same summing method with Sum of Probability. The difference is that the probabilities of the individual image classifications are squared before they are added up within each image group. Through squaring, the high-confidence classifications for individual images can occupy a bigger weight in the final park classification. This method better considers the situations that some park classifications have key features, which are easy to detect and specific to that classification. For example, the detected residential building in a park image can narrow the park classification down to Community Park or Neighborhood Park. Such key features can provide higher confidence for image classification, as well as the final park classification.

4 Evaluation

4.1 Dataset Description

We evaluated the proposed park classification framework on Seattle park image dataset, which is obtained from the open-access Seattle GeoData Portal. The dataset is consist of 122 images and 55 geo-located parks. Each park has 2 4 images, captured from different views by the urban Splanning department of the Seattle City. Since these images have been sampled by local planners in order to contain information (e.g. facility, activity, building) presenting park identities, the whole dataset was used for evaluation. These images have different resolutions, brightness, sizes and focused items. The 55 parks are classified in 4 categories manually by the local planners: Recreational Park, Natural Park, Community Park, Neighborhood Park shown in Figure 5. The ground-truth classification labels were assigned empirically based on the planners' observations on the park size, canopy rate, location (eg. beside neighborhood, or beside city boundary), activities (eg. boat, swimming, picnic), landscape style (eg.human-built green spaces, wetlands, or natural preservation), vegetation, ecosystem and other features that can be detected by field studies or images manually. The image similarity between different classifications appears to be high from human observations. The image similarity is especially high between Community Park and Neighborhood Park, both of which are close to residential areas and serve for daily activities. The small size, low variation between different classification, and uneven quality in the Seattle park data represents the typical problems in most cities' park image data, which form big challenges for image classification. However, such dataset can better guide us to meliorate the proposed park classification framework and provide a robust evaluation result.

Figure 5. Typical images for each park classification

4.2 Experimental Results and Analysis

The 122 images in the dataset are randomly divided into 70 images for training, 26 images for validation and 26 images for testing. We firstly train the three proposed deep CNNs using the loss functions, and compare their performance in terms of the over-fitting and accuracy on individual image classifications. Then, we test the three aggregation methods on park classification by randomly spliting the 55 parks into 38 for training and 17 for testing. All the three deep CNNs are trained fro 30 epochs using the cross entropy loss function, and compared in terms of loss and accuracy. The hardware we use for testing the models is a MacBook Pro laptop, with 3.5 GHz Intel Core i7 double core and 16 GB 2133 MHz.

VGG16 FE cost approximately 30 minutes for 30 epochs' training. Its accuracy and loss during training are displayed in Figure 6. As expected, due to the small size of the training park dataset, even after pre-training on the ImageNet dataset, VGG16 FE still suffered high over-fitting problem.

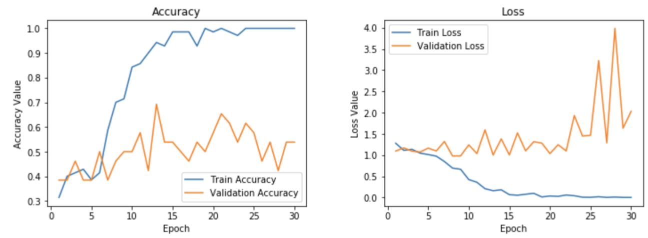

Figure 6. Accuracy and loss for VGG16 as a feature extractor (VGG16 FE)

VGG16 FE AUG cost about 70 minutes for 30 epochs' training and the evaluation results are shown in Figure 7. The over-fitting problem decrease a lot through the image augmentation strategies, as the training and validating accuracy and loss has much smaller difference compared with VGG16 FE. The overall accuracy of the model in classifying individual images is also improved.

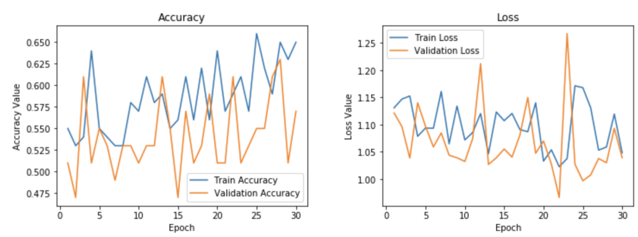

Figure 7. Accuracy and loss for VGG16 as a feature extractor with image augmentation (VGG16 FE AUG)

VGG FT AUG takes about 2 hours for training and evaluation results are in Figure 8. Even suffering some extent of over-fitting problem compared with VGG FE AUG, the fine-tuning strategies appiled in VGG FE AUG model significantly improves the accuracy in both stability and precision value than the VGG FE and VGG FE AUG. The balancing between the VGG FT AUG and VGG FE AUG is the balancing between over-fitting and accuracy. It is important to test which convolutional layer should be chosen for re-training on new domain's park images to minimize over-fitting and maximize accuracy. After testing several scenarios, we found re-training the last two convolutional layers achieves best overall performance.

Then, we tested the three aggregation methods, Absolute Voting, Sum of Probabilities and Sum of Squared Probabilities, in summing the group of image classifications for classifying parks. The results are shown in Table 1, 2 and 3 respectively. Under all three aggregation methods, VGG16 FT AUG keeps

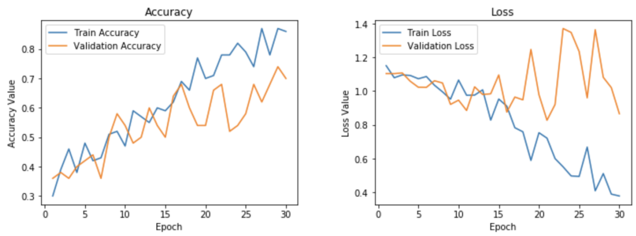

Figure 8. Accuracy and loss for VGG16 as a feature extractor with fine-tuning and with image augmentation (VGG16 FT AUG)

Metric	VGG16 FE	VGG16 FE AUG	VGG16 FT AUG
Precision	49.66%	56.12%	66.66%
Recall	42.12%	46.46%	62.16%

Table 1. Precision and recall of the model with absolute voting method.

Metric	VGG16 FE	VGG16 FE AUG	VGG16 FT AUG
Precision	49.66%	56.12%	69.87%
Recall	44.44%	50%	66.66%

Table 2. Precision and recall of the model with sum of probabilities aggregation method.

the best performance in classifying parks. It further confirms our choice to re-train two last convolutional layers for VGG FT AUG.

In terms of aggregation methods, the sum of squares has overall best precision and recall for all three deep CNNs models in final park classifications. It proves that the individual image's high-confidence classifications should occupy more weights in summing image prediction probabilities within image group for park classification.

The VGG16 FT AUG with the aggregation method of Sum of Squared Probabilities brings the best park classification accuracy of 73.34%, and the best recall of 69.86%. The over 70% accuracy is a satisfying result, considering the dataset only has 122 images with uneven quality, and the image similarity between the different park classifications are quite high. Compared with the best 51.11% precision using VGG16 FE and 58.12% using VGG16 FE AUG, the VGG16 FT AUG has a jumply improved accuracy, around 15%. Such result demonstrate the significant benefits and efficiency of applying fine-tuning strategies in improving pre-trained deep CNNs' performance on small park image dataset.

5 Conclusion

In this study, we propose a park classification framework using geo-located image data. To address the challenges of small dataset, high image similarity between different park

Metric	VGG16 FE	VGG16 FE AUG	VGG16 FT AUG
Precision	51.11%	58.12%	73.34%
Recall	44.44%	50%	69.87%

Table 3. Precision and recall of the model with sum of squares aggregation method.

classifications, and uneven image quality in most cities' park image data, we apply a transfer learning approach for all three designed deep CNNs. Through pre-training process, the transfer learning approach enables to alleviate the requirement for dataset size and variation for park images in target city, and at the same time provides robust and powerful classification. VGG FT AUG has the best performance classifying in the experiment, which proves the efficiency of adopting the image augmentation and fine-tuning strategies. We also compare the three aggregation methods to summing group of image classifications. The Sum of Squared Probabilities method has best performance, which highlights the weightings of high-confidence individual image classification.

Our proposed park classification framework can provide strong and robust classification for small, low-quality and low-variation image dataset. The hardware requirement is also low for running the framework in multiple tasks since the training working load is significant decreased by transfer learning approach. These benefits allows it to be generally applicable to most domains and running environments. This framework can potentially be applied to all other types of POIs' classifications, such as stores and restaurants. In urban planning, this tool is very helpful to train the model in a main city, and easily replicates in new urban settings.

In future work, we plan to evaluate the framework on other cities' park image data, and other POIs' images to test the universality of the framework. A data sampling strategy will be designed to filter out the less informative images (e.g. park images taken in the woods) to increase the robustness of the framework testing. Moreover, we will explore to future improve the framework with more complicated fine-tuning strategies,such as assigning different learning rate to layers, and other image augmentation strategies, such as generative adversarial networks (GANs). In addition, we plan to try unsupervised or weakly supervised classifications to more objectively classify parks, rather than only relying on the manually created classifications.

References

[1] Ando, R. K., and Zhang, T. A framework for learning predictive structures from multiple tasks and unlabeled data. *Journal of Machine Learning Research 6*, Nov (2005), 1817–1853.

[2] Dai, W., Yang, Q., Xue, G.-R., and Yu, Y. Boosting for transfer learning. In *Proceedings of the 24th international conference on Machine learning* (2007), pp. 193–200.

[3] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. Imagenet:

A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (2009), Ieee, pp. 248–255.

[4] DeVries, T., and Taylor, G. W. Improved regularization of convolutional neural networks with cutout. *arXiv preprint arXiv:1708.04552* (2017).

[5] Ding, Z., Shao, M., and Fu, Y. Latent low-rank transfer subspace learning for missing modality recognition. In *Proceedings of the Twenty-Eighth AAAI Conference on Artificial Intelligence* (2014), pp. 1192–1198.

[6] Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., and Darrell, T. Decaf: A deep convolutional activation feature for generic visual recognition. In *International conference on machine learning* (2014), pp. 647–655.

[7] Frid-Adar, M., Diamant, I., Klang, E., Amitai, M., Goldberger, J., and Greenspan, H. Gan-based synthetic medical image augmentation for increased cnn performance in liver lesion classification. *Neurocomputing 321* (2018), 321–331.

[8] Gong, Y., Lin, Y., and Duan, Z. Exploring the spatiotemporal structure of dynamic urban space using metro smart card records. *Computers, Environment and Urban Systems 64* (2017), 169–183.

[9] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. Generative adversarial nets. In *Advances in neural information processing systems* (2014), pp. 2672–2680.

[10] Hu, J., Lu, J., and Tan, Y.-P. Deep transfer metric learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 325–333.

[11] Ibes, D. C. A multi-dimensional classification and equity analysis of an urban park system: A novel methodology and case study application. *Landscape and Urban Planning 137* (2015), 122–137.

[12] Kandasamy, K., Neiswanger, W., Schneider, J., Poczos, B., and Xing, E. P. Neural architecture search with bayesian optimisation and optimal transport. *Advances in neural information processing systems 31* (2018), 2016–2025.

[13] Kovacs-Györi, A., Ristea, A., Kolcsar, R., Resch, B., Crivellari, A., and Blaschke, T. Beyond spatial proximity—classifying parks and their visitors in london based on spatiotemporal and sentiment analysis of twitter data. *ISPRS International Journal of Geo-Information 7*, 9 (2018), 378.

[14] Krizhevsky, A., Sutskever, I., and Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Communications of the ACM 60*, 6 (2017), 84–90.

[15] Nguyen, L. D., Lin, D., Lin, Z., and Cao, J. Deep cnns for microscopic image classification by exploiting transfer learning and feature concatenation. In *2018 IEEE International Symposium on Circuits and Systems (ISCAS)* (2018), IEEE, pp. 1–5.

[16] Shaha, M., and Pawar, M. Transfer learning for image classification. In *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)* (2018), IEEE, pp. 656–660.

[17] Shao, L., Zhu, F., and Li, X. Transfer learning for visual categorization: A survey. *IEEE transactions on neural networks and learning systems 26*, 5 (2014), 1019–1034.

[18] Shekhar, S., Patel, V. M., Nguyen, H. V., and Chellappa, R. Generalized domain-adaptive dictionaries. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2013), pp. 361–368.

[19] Shorten, C., and Khoshgoftaar, T. M. A survey on image data augmentation for deep learning. *Journal of Big Data 6*, 1 (2019), 60.

[20] Simonyan, K., and Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

[21] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2015), pp. 1–9.

[22] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016), pp. 2818–2826.

[23] Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., and Liu, C. A survey on deep transfer learning. In *International conference on artificial neural networks* (2018), Springer, pp. 270–279.

[24] Tian, X., Tao, D., and Rui, Y. Sparse transfer learning for interactive video search reranking. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 8*, 3 (2012), 1–19.

[25] Wagner, H. H., and Fortin, M.-J. Spatial analysis of landscapes: concepts and statistics. *Ecology 86*, 8 (2005), 1975–1987.

[26] Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. How transferable are features in deep neural networks? In *Advances in neural information processing systems* (2014), pp. 3320–3328.

[27] Yosinski, J., Clune, J., Nguyen, A., Fuchs, T., and Lipson, H. Understanding neural networks through deep visualization. *arXiv preprint arXiv:1506.06579* (2015).