Image classification by Distortion-Free Graph Embedding and KNN-Random forest

Askhat Temir, Kamalkhan Artykbayev, M. Fatih Demirci
Department of Computer Science, Nazarbayev University
{askhat.temir, kamalkhan.artykbayev, muhammed.demirci}@nu.edu.kz

NAZARBAYEV UNIVERSITY

Abstract

Image classification algorithms play an important role in various computer vision problems such as object tracking, image labeling, and object segmentation. A number of methodologies have been proposed to tackle this problem. One of the possible approaches employed extensively in the literature is to represent an image as a graph based on its hand-crafted features. However, recent advancements in deep neural networks have shown their ability to learn more discriminative and representative features. Therefore, the deep features have become considerable alternatives of hand-crafted ones. In this paper, we propose a novel framework based on distortion-free graph embedding using deep features and KNN-Random forest. Our method outperforms the state-of-the-art graph embedding-based image classification approach for the task of image classification. Particularly, the proposed framework obtains 97.5% top - 1 image classification accuracy for the ImageNet dataset for 5 classes and 93.3% for 10 classes and STL-10 top - 1 90.79% accuracy.

Introduction

Image classification based on graph embedding has been actively studied in many frameworks. Traditionally, these approaches represent images using hand-crafted features and construct graphs where vertices show features and edges encode relations between the features. The approaches then perform graph embedding into some geometric space such that similar graphs are located nearby while dissimilar graphs are placed further away. Since the successful use of AlexNet for image classification, deep neural networks (DNNs) such as VGG, Inception, ResNet, DenseNet have become the dominating approach for this task. Moreover, DNNs have shown their ability to learn more representative and discriminative features for image classification. Consequently, graph embedding approaches utilize deep features as opposed to hand-crafted features for graph construction.

In this paper, we propose a novel image classification using a distortion-free graph embedding with deep features. Specifically, after extracting deep features from images, we construct a complete graph such that its vertices represent features and its edges show the distances between the corresponding features. We then perform a distortion-free graph embedding under ℓ_∞ to represent the input graph as a set of points in the geometric space. Finally, we use KNN-Random Forest to perform the image classification. Experimental evaluation of the proposed method including a comparison with the previous graph embedding framework demonstrates its effectiveness.

Proposed method

Our framework consists of 6 main steps. 1) Fine-tuning of a pre-trained deep learning model for the dataset used in the proposed framework, 2) Extracting deep features for each image and them in descending order by their stability value computed in the deep learning model, 3) Creating a complete graph such that each node represents a feature and the weight of each edge reflects the absolute difference between the corresponding feature values, 4) Embedding each graph into a geometric space under ℓ_∞ without distortion, 5) Applying the hybrid algorithm KNN-Random forest for image classification, and 6) Generating the classification output.

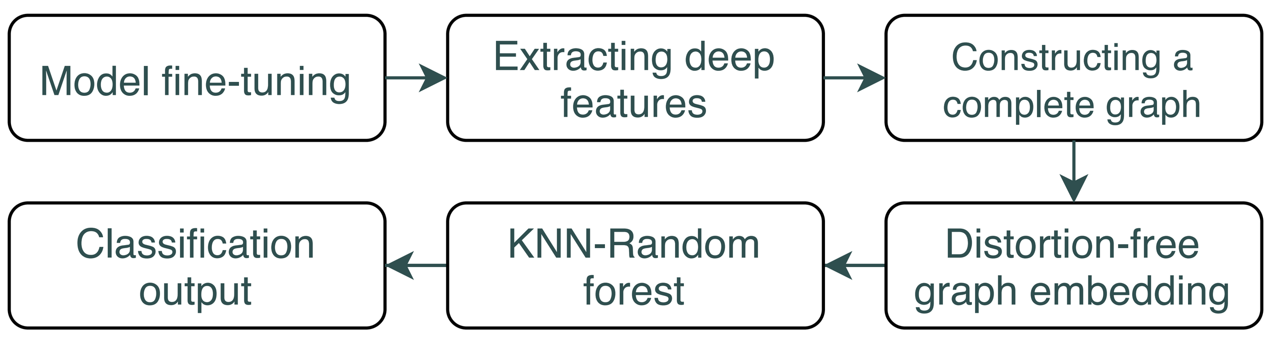

Figure 1: Overview of the proposed method

Distortion-free graph embedding

Formally, a distortion-free graph embedding for a graph $G = (V, E)$ is a mapping $f : v_i \to y_i \in \mathbb{R}^d$, $\forall i \in [n]$ such that $d = |V|$ and the function f preserves the second-order proximity based on a graph G.

The distortion free graph embedding under ℓ_∞ consists of several steps. Let $G = (V, E)$ be an input graph and let $V = \{v_0, v_1, v_2, v_3, v_4\}$ be its set of nodes. An embedding for a node in this graph is the set $\Omega = \{d_0, d_1, d_2, d_3, d_4\}$ where d_i is the shortest distance to the corresponding node in the graph. For instance, the vector representation for the v_0 in Fig. 2 would be $\{0.0, 2.0, 3.5, 2.0, 1.0\}$ where each element in the set is the shortest distance to a corresponding node. In the same way, we compute and get the embedding for the v_3 which gives $\{2.0, 2.5, 4.0, 0.0, 1.5\}$. In order to find an embedding for the entire graph, we find the coordinates for every node.

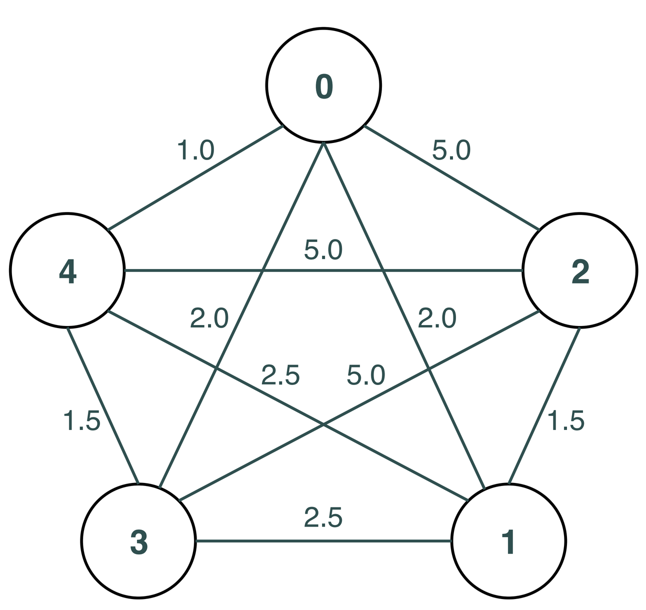

Figure 2: A sample connected graph with edge weights

In the proposed framework, we simply use the values as computed in the deep learning model. This ensures the stability of our ordering process. In addition, since there is always the same number of features obtained for an input image, we do not deal with the problem of equalizing features unlike the previous work. In particular, after obtaining ~150 deep features from an image, we sort them in descending order based on their values and construct a fully connected graph. The distance between nodes is computed as the absolute difference between the features. This distance is also known as Chebyshev or chessboard distance. This equals the limit of the L_p metric

$$\lim_{p\to\infty}\left(\sum_{i=1}^{n}|x_i - y_i|^p\right)^{1/p} \qquad (1)$$

which is also known as the ℓ_∞ metric.

Experiments

ImageNet is an image dataset, which was organized according to the "WordNet" hierarchy [2]. Each class is represented with 1000-1500 images per category. To increase the number of images in our dataset, we have used data augmentation, which artificially expands the size of a training and testing datasets by creating modified versions of all images in the dataset. This technique improves the ability to generalize for a model by feeding new variations of images. Particularly, we use a rotation of 20 degrees and the horizontal flip for each image, creating additional 5 variations per image in the dataset. Figure 3 shows this data augmentation where the first image in each row is the original image whereas the rest presents the transformed images for the classes "violet", "tarantula", and "cheetah" respectively.

In order to evaluate the classification framework based on the distortion-free graph embedding using different datasets, an experiment using STL-10 [1] has been conducted. STL-10 consists of 5000 training images with 10 predefined folds which are similar to CIFAR10 with the size of 96x96. Furthermore, there are 100,000 unlabeled images in the dataset. However, unlabeled images are not utilized in this experiment and purely based on the labeled 5000 images part.

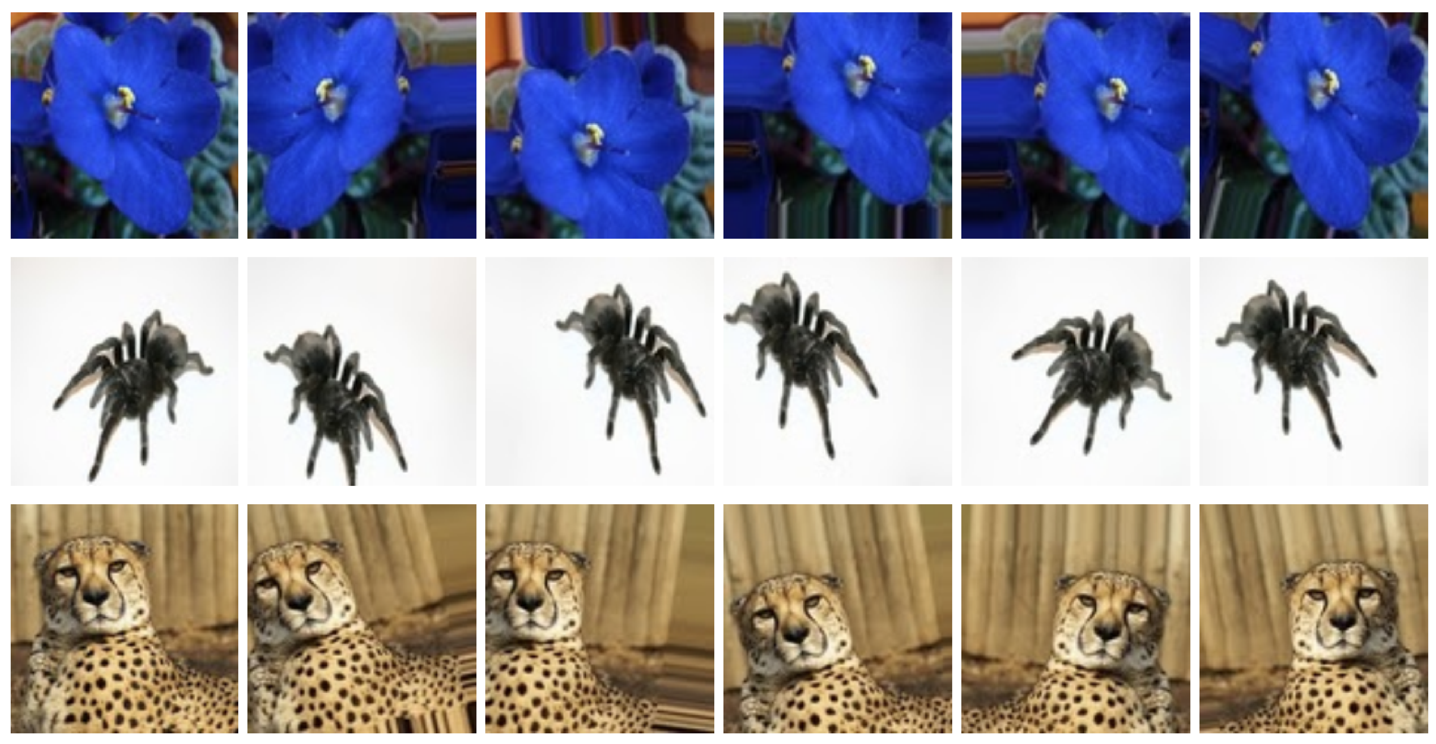

Figure 3: Augmentation examples

Conclusions

- Proposed a novel image classification approach based on distortion-free graph embedding with deep learning and KNN-random forest.
- Our method outperforms the state-of-the-art graph embedding-based image classification approach for the task of image classification.
- The proposed framework obtains 97.5% top - 1 image classification accuracy for the ImageNet dataset for 5 classes and 93.3% for 10 classes and STL-10 top - 1 90.79% accuracy.

Forthcoming Research

Graph embedding techniques have been employed by several different frameworks for several problems, such as image classification, feature correspondence, and image indexing. In this paper, we have proposed an image classification framework based on distortion-free graph embedding with deep features. Although such distortion-free graph embedding has been proposed with handcrafted features before, the way that we apply this embedding using deep features overcoming some problems faced by the alternative technique is novel. We have shown the effectiveness of the proposed framework using STL-10 and the subset of an ImageNet. However, our future goal is to perform a more comprehensive evaluation in a larger dataset and compare it with more alternative methods.

References

[1] Adam Coates, Andrew Ng, and Honglak Lee. An analysis of single-layer networks in unsupervised feature learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 215–223, 2011.

[2] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.