Flower Classification with Deep Learning

Jixiang Sun, Xinzhe Ni, Yiheng Dai

Artificial Intelligence, Shenzhen International Graduate School, Tsinghua University

Abstract

Deep learning methods have shown great potency on image classification tasks. However, few works focus on detailed classification of flowers. In this work, we construct **our own flower dataset**, using machine learning methods to perform feature visualization, and apply deep learning algorithms to make the classification.

Introduction

- Some deep learning algorithms [1] have achieved great performance on datasets like Imagenet [2]. However, few works have explored **detailed classification within one class**, for example, flowers.
- The accurate classification of flowers is meaningful since it can assist botanists with their research, and can also help normal people to identify unfamiliar flowers.
- Different species of flowers are visually alike, therefore, it is hard for human eyes to recognize them. In contrast, deep learning algorithms such as neural networks are able to capture the feature of flowers more precisely.

(a) Lilac (b) Amygdalus (c) Cerasus Serrulata (d) Bauhinia

(e) Chrysanthemoides (f) Cerasus dulosa (g) Orychophragmus (h) Peony

- We first use a flower dataset **collected by ourselves**.
- We then use **traditional machine learning methods** like t-Distributed Stochastic Neighbor Embedding (tSNE), Principal Component Analysis (PCA) and Isometric Mapping (ISOMAP) for visualization of the data.
- We finally respectively perform **unsupervised learning, supervised learning and deep learning methods** for classification.

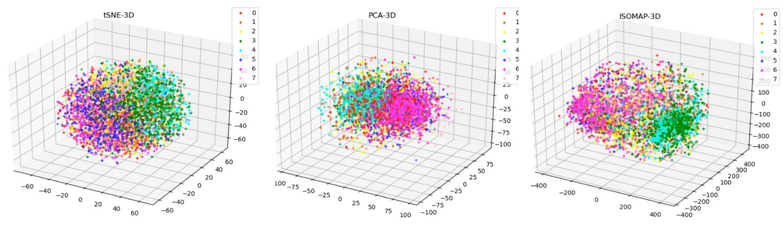
Methods

A. Dataset

We collect our flower dataset in the campus of **Tsinghua University**. The dataset consists of 8 species of flower: Lilac, Amygdalus, Cerasus Serrulata, Bauhinia, Chrysanthemoides, Cerasus Glandulosa, Orychophragmus and Peony.

B. Visualization

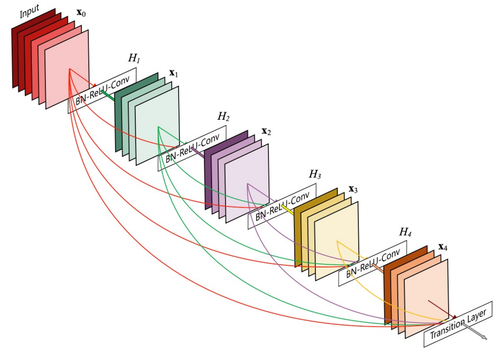
To better understand the difference of features between different species of flowers, we performed **tSNE, PCA and ISOMAP** operation on the dataset to obtain its feature visualization.



tSNE PCA ISOMAP

C. Classification

- We perform unsupervised learning algorithms like **k-means, hierarchical clustering and consensus clustering**.
- We perform supervised machine learning algorithms like Support Vector Machine (**SVM**).
- We perform deep learning methods like **DenseNet** [1] .



Dense block

Results

A. Unsupervised learning methods

We use normalized mutual information (NMI) to evaluate clustering results. It turns out that NMI scores are all low which means clustering methods are **not proper** for the classification.

	k-means	hierarchical clustering	consensus clustering
NMI score	0.1184	0.1510	0.1486/0.0923

B. Supervised learning and deep learning methods

- For supervised learning methods, we use SVM with different kernel including **RBF, linear, poly and sigmoid**.
- For deep learning methods, we use **DenseNet** with Adam optimizer, learning rate of 1e-4 and batchsize of 256. It turns out that deep learning method performs the best.

	SVM (RBF)	SVM (Linear)	SVM (Poly)	SVM (Sigmoid)	DenseNet
Accuracy	0.5375	0.4625	0.4813	0.4863	**0.9320**

Conclusions

- We collect a real world flower dataset of Tsinghua University by ourselves.
- We use traditional machine learning methods like tSNE, PCA and ISOMAP for visualization.
- We compare the classification performance of unsupervised learning, supervised learning and deep learning methods.

References

[1] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4700–4708, 2017.

[2] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li Fei- Fei. Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, pages 248–255. Ieee, 2009.