computer sciences and mathematics forum

MDPI

Proceeding Paper

Combine Transfer Deep Learning with Classical Machine Learning Models for Multi-View Image Analysis †

Zheng Xu [1,*] and Cong Wu [2]

[1] Department of Mathematics and Statistics, Wright State University, OH 45435, USA
[2] Department of Computer Science and Engineering, Wright State University, OH 45435, USA
* Correspondence: zheng.xu@wright.edu; Tel.: 1-937-775-2103

† Presented at the 1st International Online Conference on Mathematics and Applications; Available online: https://iocma2023.sciforum.net/.

Abstract: Deep learning has become widely used in image analysis. Transfer learning can make use of information from other data sets for the analysis of this data set. When there is a small number of images at hand, transfer learning using pre-trained models with coefficients already estimated from other data sets is recommended. This is in contrast to deep learning with most model parameters re-estimated. Because deep transfer learning uses pre-trained models with weight parameters in the lower layers fixed, deep learning can be viewed as a two-stage approach: (1) feature extraction from lower neural network layer, and (2) estimate a neural network using the extracted features as input. Since deep transfer learning is a feature extraction, we can extend the two-stage approach in a more general two-stage framework: (1) feature extraction using multiple methods, (2) machine learning methods taken extracted features as input. We evaluate the performance of methods with different Stage 1 approach and Stage 2 based on a multi-view plant imaging data set in predicting the phenotype leaf numbers based on images. This paper contains a study to conduct evaluation of different two-stage machine learning methods for multi-view image data in plant image phenotyping.

Keywords: deep learning; transfer learning; feature extraction; random forest; plant image phenotyping; multiview images

MSC: 68T45; 62P10

check for updates

Citation: Xu, Z.; Wu, C. Combine Transfer Deep Learning with Classical Machine Learning Models for Multi-View Image Analysis. *Comput. Sci. Math. Forum* **2023**, *1*, 0. https://doi.org/

Academic Editor: Firstname Lastname

Published: 1 December 2022

1. Introduction

Deep learning has become widely used in image analysis. A typical deep learning model can have millions of parameters. For example, VGG16, VGG19 and ResNet50 respectively has 138.4 million, 143.7 million, and 25.6 million parameters [1–3]. Because deep learning models have millions of parameters, large datasets have to be used to train the model to estimate model parameters. Researchers have spent a lot of resources (time and money) to collect and annotate large datasets for deep learning model purpose.

However, in many applications, it is not necessary to fully re-train the model based on a large annotated dataset. Fully re-estimate the deep learning model can incur a lot of cost, and the strategy of fully re-estimate the deep learning model is not feasible due to budget and resource limit. In addition, fully re-estimate the model with a small data set may not have good performance. For a small dataset, a simple model is often preferred than a complex model[4]. For example, if only limited observations are available for the regression, a linear regression or polynomial regression (with degree less than 4) is often preferred than non-parametric regression models[4]. The other way is to use pre-trained parameters in a complex model such as deep learning neural network models are also recommended [4]. Researchers develop the transfer learning approach to solve the issue.

Transfer learning allow researchers to make use of pretrained model based on other dataset to analyze their own problem using their small or medium-size data set [5]. One typical method of transfer learning for image analysis is a two-stage approach. In Stage 1, the lower neural network layers in the deep learning model (for example, in VGG16 model, the 16 convolutional neural network layers) are used with pretrained weights from another data set, which is a standard big data set such as ImageNet [6] to transform the input images into features. In Stage 2, the extracted features and the ground truth response (y) are fed into a neural network with all neural network parameters estimable [7]. In this way, satisfactory performance can be obtained even with a small data set.

We extend the two-stage approach in general. In Stage 1, a feature-extraction method is used to extract features from input images. The feature extraction methods can be principal component analysis or pre-trained deep neural network models with parameters fixed. In Stage 2, it is a supervised learning problem (regression or classification) with extracted features in Stage 1 as input and response variable (continuous or categorical y) as output. The method in Stage 2 can be a neural network or random forest. We intend to evaluate the performance of our general two-stage approach with different Stage 1 methods and different Stage 2 methods.

Our proposed two-stage machine learning strategy is a general approach in that researchers can select appropriate Stage 1 method and Stage 2 method according to the research problem, objective, and data. In our previous work [8], we propose machine learning methods to predict continuous phenotypes and binary phenotypes based on plant images. Our methods belong to the general-framework of our proposed two-stage approach. In Stage 1, we adopt principal components analysis (PCA) to extract features, i.e. PCs. In Stage 2, we use a range of machine learning methods (Random Forest, Partial Least Squares and LASSO) to predict the plant phenotypes including leaf number of a plant, based on plant images. Our proposed methods work for plant image phenotyping [8].

Another example of method that belongs to our proposed two-stage method is the standard deep transfer learning[9]. In Stage 1, the lower layers of neural network models with pre-trained fixed weights are used to extract features. In Stage 2, these features were fed into the upper layers of neural networks. Thus, a deep learning with the weights of lower layers fixed also belongs to our proposed general framework [9].

Image-based plant phenotyping, i.e. plant image phenotyping, refers to a rapidly emerging research area concerned with quantitative measurement of the structural and functional properties of plants based on plant images[8]. Image-based plant phenotyping facilitates the extraction of traits noninvasively by analyzing a large numbers of plants in a relatively short period of time. Plant image phenotyping has the advantage of low cost, high throughput, and being a non-destructive measurement [10]. Based on plant image phenotyping, agricultural and biological researchers can track the growth dynamics of plants, identify the time of critical events (such as plant flowering), morphological changes (such as leaf number, plant size, position of each leaf) so that they can better analyze the problem such as how different factors (fertilizer usage amount, temperature, moisture) influence plants[8]. In this article, We illustrate our method by evaluating the performance of our general two-stage framework with different Stage 1 and Stage 2 methods. We evaluate how these methods works for plant image phenotyping, especially in detect the number of leaves of plants by analyzing RGB images.

The remaining of the paper is as follows: Section 2 specified methods and data. Section 3 shows results. Section 4 makes discussions. Section 5 draws conclusions.

2. Materials and Methods

The proposed method is a general two-stage approach. In Stage 1, a feature-extraction method is used to extract features from input images. We adopt principal component analysis in this paper. In our ongoing project, we are evaluating the performance of other feature extraction methods especially the use of pre-trained deep neural network models with parameters predetermined and fixed. As transfer learning, the values of neural network parameters are pre-trained using a large dataset, such as ImageNet. ImageNet is a large image dataset organized according to the WordNet hierarchy. Each meaningful concept in WordNet is described by multiple words or word phrases [6]. ImageNet includes 80,000 nouns with each noun illustrated on average 1000 images, so that satisfy the researchers' critical need for more data to enable more general machine learning methods [6]. The use of ImageNet to pre-train neural network parameters to obtain parameter values and use of fixed imageNet weight in deep learning have shown advantage in the literature [11].

In Stage 2, it is a supervised learning problem (regression or classification) with extracted features in Stage 1 as input and response variable (continuous or categorical y) as output. We adopted Partial Least Square (PLS) and LASSO as regression methods, and Partial Least Square-Discriminant Analysis (PLS-DA) and LASSO as classification methods. LASSO often show good prediction performance for high-dimensional data with the use of $L1$ penalty [4]. When model interpretation is preferred instead of model prediction. LASSO method is often used to identify predictors impacting the response variable, assuming sparse signals [4]. With recent development in explanatory machine learning and artificial intelligent, researchers try to develop models with good interpretation, instead of a black-box machine learning model. When explanatory is preferred, LASSO methods and decision trees are often used due to their good interpretability [4]. A range of methods on visual interpretability for deep learning have been developed in literature [12]. In our ongoing project, we are working on evaluating the performance of random forest and neural networks as our Stage 2 methods.

The dataset used in our study is University of Nebraska Lincoln (UNL) Component Plant Phenotyping Dataset (CPPD) [13]. UNL-CPPD data set consists of images of 13 maize plants for two side views (0 degree and 90 degree). Plants were imaged once per day from 2 days to 28 days after planting using RGB camera of UNL Lemnatec Scanalyzer 3D high-throughput phenotyping facility.

The RGB images were converted to grayscale images and resized to 224×224, which is the size of input images for deep learning models including VGG16, VGG19 and ResNet50 [1–3]. In this paper, each grayscale image was converted into a numerical matrix of 224 rows and 224 columns, which was vectorized/reshaped to a column vector of length $224^2 = 50176$.The data were centered and scaled before extraction of principal components. Principal components were extracted from the centered and scaled vectors representing the images. The extracted principal components were then fed into Stage-2 machine learning methods (any appropriate supervised learning method can be used) to make prediction.

The phenotype leaf number refers to the number of leafs in a plant image. It is an integer and we treat it as a continuous phenotype. Then the binary variable "leafy" was created as leafy= 1 if leaf number is more than the median leaf number. We applied regression methods to predict the phenotype " plant leaf numbers" and apply classification methods for the binary phenotype "leafy".

Five-fold cross validation (CV) was used to evaluate the performance. In the regression problem, the performance evaluation metrics are Mean Square Error (MSE), Root Mean Square Error (RMSE) and Mean Absolute Deviation (MAD) specified as

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2; \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}; \quad (2)$$

$$MAD = \frac{1}{n}\sum_{i=1}^{n}|Y_i - \hat{Y}_i|, \quad (3)$$

where Y_i is the true response value and $\hat{Y}_i$ is the predicted response value for observation i. In the classification problem, the performance evaluation metric is accuracy, which is the number of correct classifications divided by the total number of classifications.

3. Results

3.1. Performance of regression methods

We reported in Table 1 the performance of regression methods for continuous trait leaf number. We found the performance of PLS method and LASSO method are nearly the same with the performance of LASSO method slightly better than the performance of PLS in the regression problem. Although in literature, researchers often report the superior performance of LASSO over PLS, we need to point out that we are not sure which way (LASSO or PLS) is better for plant image phenotyping since only one dataset is studied. We are working on evaluate the methods in other dataset in our ongoing project. Given current results, we recommend adopting our proposed general framework, and try a range of Stage 1 methods and Stage 2 methods to decide which specific method should be used in Stage 1, and which to use in Stage 2. In our ongoing project, we will provide more thorough results based on multiple datasets.

Table 1. Performance of Machine Learning Methods for Continuous Traits

Method	Criteria	Performance Score
PLS	MSE	2.960
PLS	RMSE	1.720
PLS	MAD	1.024
LASSO	MSE	2.959
LASSO	RMSE	1.719
LASSO	MAD	1.020

3.2. Performance of classification methods

We reported in Table 2 the performance of classification methods for binary trait "leafy". We found the performance of LASSO method is slightly better than the performance of PLS-DA in the classification problem. In practice, although LASSO method often show superior performance than least square regression, ridge regression and partial least squares in real application [4], which method is preferred still depends on specific applications so that LASSO method and partial least squares are likely to be the best performing method for other data sets. We note that the performance of methods are evaluated in only one data set (UNL-CPPD multi-view images). In our ongoing projects, we are evaluating our methods in other datasets. We point out that evaluation of methods based on multiple datasets can provide more evidence of the methods, and the results based on one data-set is of limited scope. In our ongoing project, we are providing more through results based on multiple data-sets. The purpose of this article is to propose and communicate our methods with experts and more through analysis will be provided in our ongoing project.

Table 2. Performance of Machine Learning Methods for Binary Traits.

Method	Criteria	Performance Score
PLS-DA	Accuracy	0.890
LASSO	Accuracy	0.895

4. Discussion

Our proposed methods is a general two-stage framework allowing the choice of Stage 1 method and Stage 2 method. When Stage 1 method and Stage 2 method are based on the same neural network model, it reduces to the deep transfer learning model. The current report is principal component as Stage 1 method, partial least square and LASSO as Stage 2 methods based on UNL COPD dataset. We note that current report is of limited scope, and more methods and more datasets are needed make more thorough analysis and report. In our ongoing project, we are working to evaluate the performance of two-stage approach using different Stage 1 methods (deep neural network and principal component) and

Stage 2 methods (partial least squared, LASSO and random forest), and compare their performance with the deep transfer learning in the literature based on multiple datasets.

Regarding Stage 1 methods to extract features from images, two widely-used methods are (1) principal component analysis and (2) pre-trained deep learning. Both methods work for plant image phenotyping (image regression, classification and segmentation) as shown in literature including two of our previous studies [8,14].

Although our two-stage method is a general framework, the most widely used one is deep transfer learning which already show great success in literature. We want to explore the possibilities of other models by using different Stage 1 and Stage 2 methods. In terms of prediction performance, we expect deep transfer learning may achieve the best prediction performance whereas it still deserves to compare different methods. The objective of this article is to compare different methods belonging to our two-stage general framework for better prediction and interpretation so that researchers can have better understanding and more tools when they want to develop novel machine learning methods.

5. Conclusions

We have proposed methods to extend the two-stage deep transfer learning models in the literature. Our general two-stage approach can include different Stage 1 methods and different Stage 2 methods. We evaluated the performance of our general two-stage approach with principal component analysis as our Stage 1 method and partial least square (PLS), partial least square - discriminant analysis (PLS-DA) and LASSO as our Stage 2 methods based on UNL-CPPD plant phenotyping data set.

Data Availability Statement: All data used in this paper are publicly available.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

CPPD	Component Plant Phenotyping Dataset
CV	Cross Validation
LASSO	Least Absolute Shrinkage and Selection Operator
PC	Principal Component
PLS	Partial Least Squares
PLS-DA	Partial Least Squares - Discriminant Analysis
RF	Random Forest
VGG	Visual Geometry Group from Oxford

References

1. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* **2014**.
2. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
3. Shafiq, M.; Gu, Z. Deep residual learning for image recognition: a survey. *Applied Sciences* **2022**, *12*, 8972.
4. Hastie, T.; Tibshirani, R.; Friedman, J.H.; Friedman, J.H. *The elements of statistical learning: data mining, inference, and prediction*; Vol. 2, Springer, 2009.
5. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep learning*; MIT press, 2016.
6. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. 2009 IEEE conference on computer vision and pattern recognition. Ieee, 2009, pp. 248–255.
7. Weiss, K.; Khoshgoftaar, T.M.; Wang, D. A survey of transfer learning. *Journal of Big data* **2016**, *3*, 1–40.
8. Xu, Z.; Wu, C. Machine Learning and Statistical Approaches for Plant Phenotyping. In *Intelligent Image Analysis for Plant Phenotyping*; CRC Press, 2020; pp. 195–220.
9. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A survey on deep transfer learning. Artificial Neural Networks and Machine Learning–ICANN 2018: 27th International Conference on Artificial Neural Networks, Rhodes, Greece, October 4-7, 2018, Proceedings, Part III 27. Springer, 2018, pp. 270–279.

10. Das Choudhury, S.; Samal, A.; Awada, T. Leveraging image analysis for high-throughput plant phenotyping. *Frontiers in plant science* **2019**, *10*, 508.
11. Shermin, T.; Teng, S.W.; Murshed, M.; Lu, G.; Sohel, F.; Paul, M. Enhanced transfer learning with imagenet trained classification layer. Image and Video Technology: 9th Pacific-Rim Symposium, PSIVT 2019, Sydney, NSW, Australia, November 18–22, 2019, Proceedings 9. Springer, 2019, pp. 142–155.
12. Zhang, Q.s.; Zhu, S.C. Visual interpretability for deep learning: a survey. *Frontiers of Information Technology & Electronic Engineering* **2018**, *19*, 27–39.
13. Das Choudhury, S.; Bashyam, S.; Qiu, Y.; Samal, A.; Awada, T. Holistic and component plant phenotyping using temporal image sequence. *Plant methods* **2018**, *14*, 1–21.
14. Miao, C.; Xu, Z.; Rodene, E.; Yang, J.; Schnable, J.C.; others. Semantic segmentation of sorghum using hyperspectral data identifies genetic associations. *Plant Phenomics* **2020**, *2020*.