Machine Learning for Classifying Working State Images Recorded by Digital Tachograph System

Senlin Guan[1], Takeshi Shikanai[2] and Morikazu Nakamura[3]
[1]Faculty of Medicine, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8575, Japan
[2]Faculty of Agriculture, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan
[3]Faculty of Engineering, University of the Ryukyus, 1 Senbaru, Nishihara, Okinawa 903-0213, Japan
E-mail : [1]guansid@gmail.com, [2]tshikana@agr.u-ryukyu.ac.jp, [3]morikazu@ie.u-ryukyu.ac.jp

Abstract:

Machine learning is a powerful modelling and prediction tool for data analysis and decision-making in agriculture production, especially for the cases of dealing with large volume of data in diverse formats. In this paper, we present a case study of applying machine learning to classify the working states of harvesting sugarcane based on the time-series data, which is recorded by a digital tachograph system mounted on a small sugarcane harvester. The study aims at constructing a model and training it applicable to automatically learn from the time-series images and classify the images into different working states. Three machine learning models are implemented to evaluate the best accuracy for classification and the optimum parameters for the model. The result indicates that using machine learning is an effective way to distinguish the working states, and the average F1_score reaches 0.970 when recognizing the cutting state. The classification by the Support Vector Machine (SVM) model with Radial Basis Function (RBF) kernel gains higher accuracy than by that with linear kernel and by K-Nearest Neighbors (KNN).

1. Introduction

Machine learning as well as other artificial intelligence and deep learning is a powerful tool for learning from and making predictions on data [1, 5]. Because of having the outstanding capability for handling of large quantities of individual or contextual data in diverse formats, machine learning is widely used in many different fields, and also including in agricultural science. In order to improve agricultural productivity and economic growth, numerous Information and Communication Technologies (ICTs) have been applied in data collection and analysis for daily farming practices. The data of describing nature environment, crop status and farming activity is often recorded by using sensors, monitors, or other online/offline devices, and its possible formats involve discrete or continuous values, image, sound, movie, or their combinational forms. Making sense of the data is not easy, and accordingly some efficient techniques for data processing, analyzing and predicting like machine learning are required.

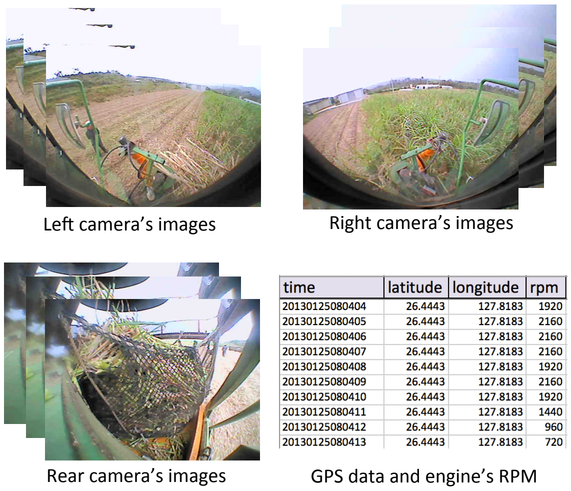
There exists many of agricultural real applications by using machine learning. McQueen et al. [2] gives an overall review of applying machine learning to agricultural data. A study using diverse machine learning techniques for fruits and vegetables is presented in [4]. Pietersma et al. [3] report an experimental performance analysis for applying machine learning on small data sets. More practical applications involve using machine learning to forecast the agricultural products such as shrimp and chicken export [8], and to classify field crop insects by multiple-kernel learning method [9]. Currently as a new and hot area of machine learning research, deep learning is also impacting the agricultural technique innovations. For example, Steen et al. [7] has successfully applied deep learning for obstacle detection in agricultural fields, and Zou et al. [10] dedicate an approach for feature selection based on deep learning for remote sensing scene classification.

In this study, we present a case study applying machine learning to working states classification in agricultural production. Shikanai and Guan [6] have used a digital tachograph system (DTS) for recording the working data for harvesting sugarcane, which is an important crop in Okinawa, Japan. The DTS records of the harvester consist of the GPS tracking data, the engine revolution speed and the working state images taken from multiple cameras. The proposed manual method in [6] is suitable for small volume of the data, however, for the large amount of the time-series data recorded by DTS, there requires an automatic model with the capabilities for calculating the working area and the work efficiency, learning from the datasets, and classifying and predicting the working data. This paper focus on constructing such models capable of automatic classification for the DTS data.

2. Experimental data and methods

The experimental data is recorded by the DTS mounted on a small sugarcane harvester. The DTS consists of a drive recorder, two front cameras and a rear camera. When the power of the harvester is on, the DTS starts recording the GPS clock and location data, the number of engine's Revolution Per Minute (RPM), and the working images toward the front-left, front-right and rear directions of the machine body. The front camera records a realtime working image per two seconds, while the rear camera shots per minute on a storage bag for storing the cut fragment canes and crashes. The GPS clock, together with location data, RPM value and working images are stored in the memory card of the DTS. Figure 1 shows the time-series samples data recorded by the DTS.

According to the RPM values and the working images, we can manually distinguish the activities of the harvester. When the harvester is stopping, the number of the RPM is zero, and the time-series images taken by the left or right camera are almost static. The number of RPM keeps in a lower range from zero to 1,000 when loading/unloading the storage bag, or idling for short-time preparation and maintenance, and in a higher range from 1,000 to 1,900 when backing or rotating the



time	latitude	longitude	rpm
20130125080404	26.4443	127.8183	1920
20130125080405	26.4443	127.8183	2160
20130125080406	26.4443	127.8183	2160
20130125080407	26.4443	127.8183	2160
20130125080408	26.4443	127.8183	1920
20130125080409	26.4443	127.8183	2160
20130125080410	26.4443	127.8183	1920
20130125080411	26.4443	127.8183	1440
20130125080412	26.4443	127.8183	960
20130125080413	26.4443	127.8183	720

Figure 1. The time-series data recorded by DTS

Figure 2. Four working states of the harvester

direction along with the moving lanes. While the harvester is cutting sugarcane, the engine turns to high speed of revolution over than 1,900 RPM, and simultaneously, the canes are attached on both the left and right cutting arms of the harvester body. Figure 2 shows the sample images of four working states recorded by the front-left camera.

The DTS data for experiment is recorded during the harvesting season from January to February, 2013. Except the days unavailable for harvesting due to weather or machine condition, there are 26 work days having valid DTS data. We consider using a major component of the DTS data, the time series images, to classify and predict the working states. Let m $(m = 1, ..., 26)$ be the number of a work day, the set of the images taken from the left, right and rear camera can be denoted as $SL_m, SR_m, SB_m (m = 1, ..., 26)$, respectively. Since the DTS using the same clock, a combined image in which the left and right camera's images are horizontally arranged is more meaningful than a single image from one side camera. Accordingly, we define an additional set SC_m for representing the combined images. According to the same prescriptive naming rule of the time series images, and the combined image SC_m can be created by matching the file names of the original images from the left and right cameras.

We considered three classifiers of the SVM with RBF kernel, the SVM with linear kernel, and the K-Nearest Neighbour (KNN) for distinguishing the working states. The brief steps applying the classifiers to the DTS data include (1) loading the original images and converting into representation images (2) splitting the representation images into a test set and a training set by k-fold cross-validation so that the validation set is no longer needed. (3) creating training vectors by using PCA for dimensionality reduction on the training set (4) creating a classifier for training the data, and (5) making prediction on the test set and evaluating the classification model quality.

3. Experimental results

The experiments are made by a Python program running on a MacBook Pro with Intel Core i5 and and 8GB RAM. Since we use the randomized singular value decomposition for dimensionality reduction in the PCA procedure, for avoiding contingent case all the computed result data is derived from the average values by five times of experiments. The dimensionality of features and the length of principle components of the PCA are defined as $128 \times 96 = 12,288$ and 150 in advance, respectively. Among all sets of original images, we are particularly interested in the set of the combined images containing both the left and right work states. Using this set, we plot some representative images transformed from the eigen values computed by the PCA in Figure 3. Each image is typical of a set of similar working states of the harvester.

After loading the original images and converting to the representation images, the classifier splits the data into a training dataset with 75% size and a test dataset with 25% size of the images set. And then, the classifier transforms the training dataset by the PCA procedure, fits the classifier and makes the prediction on the test dataset. The experiment results are shown in Table 1. In the table, every SC_m corresponds a set of images taken in a workday. The columns with labels of "SC-2, SC-2L, SC-2K" list the result data computed by the classifiers of the SVM with RBF kernel, the SVM with linear kernel and the KNN, for two states classification, respectively. Similarly, those with the labels of "SC-4, SC-4L, SC-4K" contain the result data for four states classification. Here, two states classification indicates classifying the images into a cutting state or a not-cutting state, and four states classification indicates classifying the images into a state of stop, idle, moving or cutting. In the case of two states classification, the average values of F1_score for the three classifiers reach 0.970, 0.960, 0.955, and in four states classification, the comparatively lower values of 0.937, 0.921, 0.915 are obtained.

In addition to the SC_m, the sets SL_m and SR_m use the same parameters for running the classifiers. Using these parameters, we can easily compare the performance between the two classification and the four states classification, and between the three classifiers. Figure 4 (a) shows the two paral-

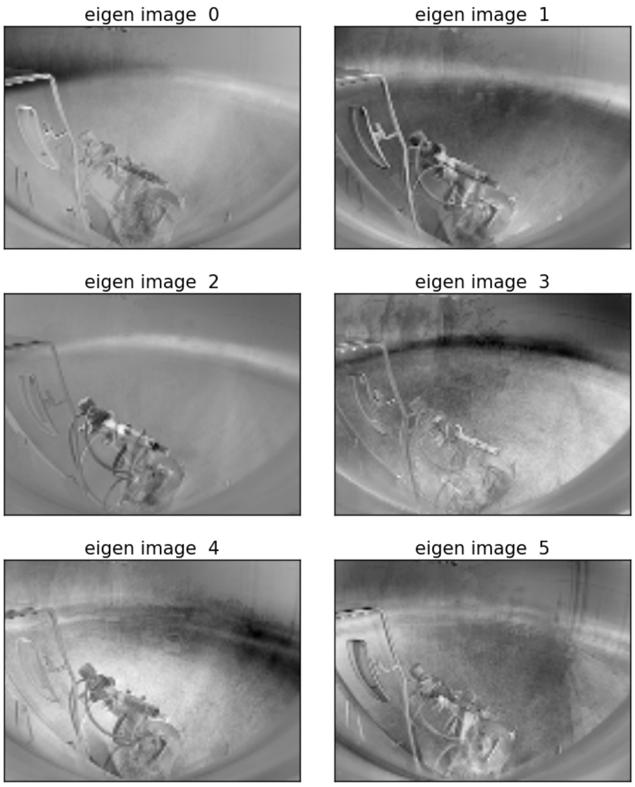


Figure 3. Transformed images from principal components

lel curves with the markers of the average values of F1_score for the two states classification and the four states classification. It is very distinct that the two states classification gains the higher performance than the four states classification. It is considered as that when the harvester is cutting the sugarcane, the canes attaching to the part of the machine body facilitates distinguishing the cutting state. Contrastively for the four states classification, the working image of stop, idle and moving does not have notably separable features, and thus the average F1_score is slightly low. In Figure 4 (b,c), the classifier of the SVM with RBF kernel "SC-2, SC-4" exhibit the highest performance among the threes classifiers. This has been demonstrated in all the experiments running on the original images from the left camera and the right camera, and their combined images. The figure only shows the result on the two and four states classification for the combined images.

In order to ascertain the factors causing the prediction error, we list 10 prediction error sample images taken by the left camera on the workday of January 25, 2013 (Figure 5). There are 44 prediction error images for predicting the 3,187 images in the test set. The labels under the image name conform the naming rule of Prediction (True). Referring to the RPM number and the other neighbour time series images, we find that the images of No. 6 and 7 border the time of starting cutting operation, and the image of No. 8 borders the finish time of cutting. The RPM number of the first image is 1,680, but of its former and latter series images are mostly over 1,920, and thus we consider the work state is exactly in cutting because

Table 1. Average F1_scores

WorkDate	Dataset (m)	Samples (N)	SC-2	SC-2L	SC-2K	SC-4	SC-4L	SC-4K
2013_01_24	SC01	12,306	0.966	0.959	0.945	0.931	0.919	0.915
2013_01_25	SC02	12,243	0.983	0.976	0.972	0.963	0.951	0.944
2013_01_26	SC03	12,312	0.963	0.951	0.949	0.942	0.920	0.927
2013_01_27	SC04	12,819	0.955	0.943	0.945	0.923	0.904	0.902
2013_01_28	SC05	10,791	0.956	0.946	0.939	0.917	0.909	0.897
2013_01_29	SC06	12,125	0.965	0.953	0.952	0.934	0.913	0.912
2013_01_30	SC07	4,982	0.963	0.959	0.946	0.938	0.934	0.914
2013_01_31	SC08	11,886	0.970	0.956	0.958	0.950	0.926	0.932
2013_02_01	SC09	11,167	0.969	0.957	0.949	0.928	0.915	0.908
2013_02_02	SC10	12,667	0.973	0.957	0.959	0.944	0.925	0.926
2013_02_03	SC11	10,024	0.982	0.972	0.969	0.937	0.928	0.917
2013_02_04	SC12	11,498	0.971	0.963	0.962	0.940	0.913	0.920
2013_02_05	SC13	8,775	0.965	0.944	0.945	0.922	0.907	0.900
2013_02_06	SC14	13,494	0.960	0.943	0.938	0.929	0.907	0.894
2013_02_07	SC15	12,109	0.986	0.980	0.967	0.953	0.940	0.926
2013_02_08	SC16	12,375	0.965	0.943	0.942	0.929	0.906	0.895
2013_02_09	SC17	12,804	0.987	0.976	0.973	0.962	0.948	0.944
2013_02_10	SC18	4,133	0.981	0.980	0.971	0.955	0.954	0.926
2013_02_13	SC19	8,648	0.976	0.968	0.963	0.923	0.892	0.908
2013_02_14	SC20	13,159	0.986	0.977	0.976	0.932	0.902	0.923
2013_02_15	SC21	5,608	0.956	0.957	0.926	0.919	0.915	0.871
2013_02_16	SC22	7,563	0.955	0.945	0.938	0.926	0.914	0.905
2013_02_20	SC23	13,671	0.971	0.960	0.959	0.939	0.919	0.912
2013_02_21	SC24	13,137	0.981	0.966	0.965	0.948	0.928	0.928
2013_02_22	SC25	4,102	0.982	0.981	0.976	0.965	0.957	0.942
2013_02_26	SC26	13,065	0.951	0.946	0.942	0.922	0.898	0.908
	Max	13,671	0.987	0.981	0.976	0.965	0.957	0.944
	Min	4,102	0.951	0.943	0.926	0.917	0.892	0.871
	Average	10,672	0.970	0.960	0.955	0.937	0.921	0.915
	Sum	277,463						

of the possibility of lowly recording the RPM value or a short machine body adjustment sometimes. The true state of the next four images of No. 2, 3, 4, 5 is in not-cutting, and of the remaining two images of No. 8 and 10 is exactly in cutting.

4. Discussions

In this study, we have implemented three machine learning models for automatically classifying the working states by the DTS data. For each model, 780 times of experiments were made for training and predicting the dataset converted from the original images in 26 workdays. The result data was derived from the average values by five times of experiments, and therefore was reliable for quantitative evaluation. The optimal trade-off parameters $C = 500, \gamma = 0.01$ for the SVM with RBF kernel were adequate values for avoiding the overfitting and underfitting.

In the experiments, every dataset was derived from only one workday data. The average time for modelling computation was 34.8 s after loading the initial data when using the SVM with RBF kernel. With respect to the possible maximum input data size of the machine learning model, we experimented on the datasets derived from multiple workdays. Up to the dataset size of 8 workdays containing 143,595 images,

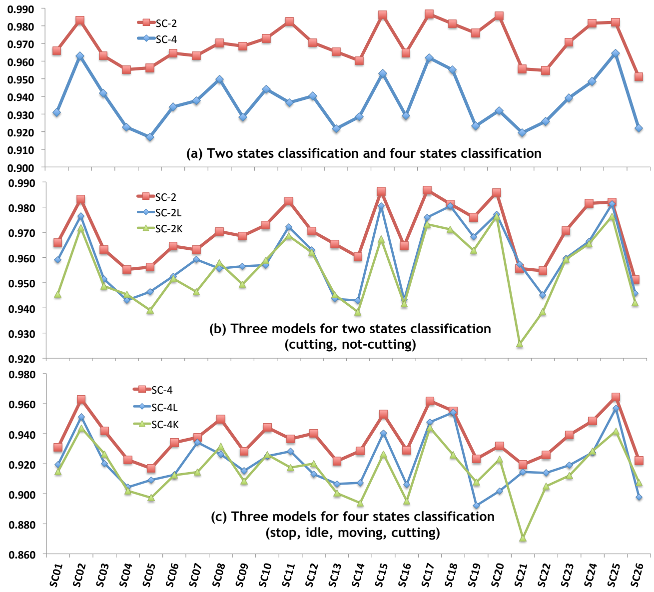


Figure 4. Comparison results of the average F1_score

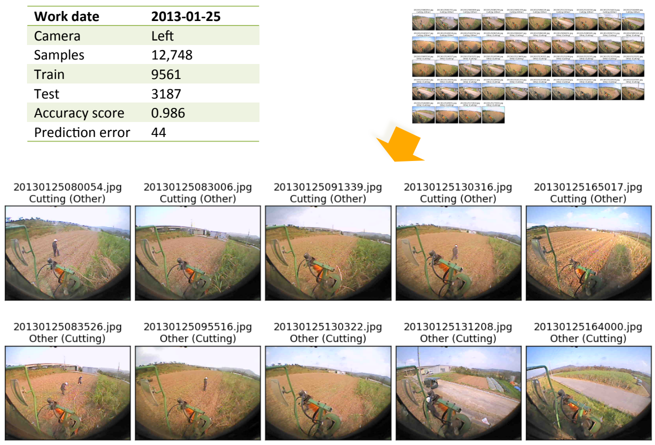


Figure 5. Prediction error images

the model gained the average F1_score 0.965 in 3.9 h. For more larger input dataset size, the model exited the computing process by raising a memory error. The strategies such reducing the typical features of the image in advance and managing the computer memory usage are required for modelling the datasets in all of workdays.

There exist many contextual information between the neighbour time series images or other data. In the experiments, a proportion of prediction errors increased when the harvester machine changed its working state between cutting and not-cutting, and the others prediction error images arranged nearly the middle of a series images of a stable working state. Referring to the contextual information is helpful for reducing these prediction errors. The extension of this study will focus on the solution for improving the modelling accuracy and robustness, modelling the large size of input data, and discovering the potential information.

5. Conclusions

The analysis and utilization for a large size of time series data gathered in the agricultural production require an effective tool. In this study, we apply the machine learning techniques to classify and predict the work states by using the time series data recorded by the DTS. Three classifiers are implemented in the experiments for quantitative comparison. The SVM with RBF kernel shows the best prediction accuracy in both two types of classification for all datasets, and its average F1_score reaches 0.970. The high recongization rate for the cutting state enables us to make an automatic tool to compute the working efficiency instead of the manual manipulation, and to elucidate the relationships between the work efficiency and the condition of the farmland and the crop.

References

[1] Bishop, C. M., 2006. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA.

[2] McQueen, R. J., Garner, S. R., Nevill-Manning, C. G., Witten, I. H., 1995. Applying machine learning to agricultural data. Computers and Electronics in Agriculture 12 (4), 275–293.

[3] Pietersma, D., Lacroix, R., Lefebvre, D., Wade, K. M., 2003. Performance analysis for machine-learning experiments using small data sets. Computers and Electronics in Agriculture 38 (1), 1–17.

[4] Rocha, A., Hauagge, D. C., Wainer, J., Goldenstein, S., 2010. Automatic fruit and vegetable classification from images. Computers and Electronics in Agriculture 70 (1), 96–104.

[5] Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural Networks 61, 85–117.

[6] Shikanai, T., Guan, S., 2015. Recording operation of sugar cane harvester and analysis of working efficiency by digital tachograph system. Agricultural Information Research 24 (4), 101–111.

[7] Steen, K. A., Christiansen, P., Karstoft, H., Jørgensen, R. N., 2016. Using deep learning to challenge safety standard for highly autonomous machines in agriculture. Journal of Imaging 2 (1), 6.

[8] Sujjaviriyasup, T., Pitiruek, K., 2013. Agricultural product forecasting using machine learning approach. Int. Journal of Math. Analysis 7, 1869–1875.

[9] Xie, C., Zhang, J., Li, R., Li, J., Hong, P., Xia, J., Chen, P., 2015. Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning. Computers and Electronics in Agriculture 119, 123—132.

[10] Zou, Q., Ni, L., Zhang, T., Wang, Q., 2015. Deep learning based feature selection for remote sensing scene classification. IEEE Geoscience and Remote Sensing Letters 12 (11), 2321–2325.