

Tea Verification Using Triplet Loss Convolutional Network

Kun-Yi Chen¹, Chi-Yu Chang¹, Zhi-Ren Tsai¹, Chun-Ting Lee², Zon-Yin Shae^{1,*}

¹Department of Computer Science and Information Engineering, Asia University, Taichung, Taiwan

²Department of Psychology, Asia University, Taichung, Taiwan

Received 30 December 2020; received in revised form 17 August 2021; accepted 18 August 2021

DOI: <https://doi.org/10.46604/aiti.2021.6939>

Abstract

To solve tea image classification problems, this study focuses on triplet loss convolutional neural network to classify six high-mountain oolong tea classes. In the experiment, instead of using traditional deep learning training approach for local feature of tea images, an innovative image verification approach is proposed to learn the global feature of tea images by integrating the distributed tea leaves' features of all tea sub-images and using a majority voting mechanism to do classification. The results show that the proposed approach can work for small sample size dataset and have higher accuracy than normal transfer learning approach. The average accuracy of the proposed approach achieves 99.54%.

Keywords: convolutional neural network, tea image classification, tea image verification, triplet loss

1. Introduction

Tea is a vital export product for many countries to sustain their economic growth. Each year, there are at least six million tons of tea were produced globally [1]. In Taiwan, the price difference between high quality tea and inferior quality tea often reaches hundreds of US dollars. However, farmers who produce high quality tea often suffer from tea mixing issue, i.e., some middlemen mix the inferior quality tea with high quality tea and re-sell it as high quality tea products to earn extra margins. This can jeopardize the sustainability of local tea markets because of losing customers' trust.

To help tea farmers deal with this issue, tea industry conventionally applies biochemical techniques, such as DNA markers, mass spectrometry, gas-chromatography, or high performance liquid chromatography, to analyze tea samples and trace their original sources [2-3]. However, due to the fact that these techniques are time-consuming and have high costs, computer vision approaches were proposed in recent years as an alternative way to complement their disadvantages; the computer vision approaches often combine with signal processing algorithms or machine learning algorithms to extract tea image features [4-6].

This research aims to develop an image verification approach through majority voting mechanism, which can learn the global feature of tea images by integrating all tea sub-images' features, in order to help smallholder tea farmers address tea mixing issue. By using image cropping and majority voting, the proposed approach can work for small sample size dataset and have higher accuracy than traditional deep learning training approach.

The remainder of this study is arranged as follows. Section 2 explains how the initial experiment and two observations influence the proposed experiment design. Section 3 describes tea dataset details. Sections 4 and 5 show the proposed approach for tea image classification and its algorithms. Section 6 describes the whole image-preprocessing process, including the standardized procedure for photograph tea images and detailed image processing method. Sections 7 and 8 present the proposed model architecture and experiment results. Finally, section 9 concludes the study.

* Corresponding author. E-mail address: zshae1@asia.edu.tw

Tel.: +886-4-2332 3456 #1729

2. Initial Experiment and Observations

There are six high-mountain oolong tea classes in this study (for detailed dataset description, see section 3). The initial attempt to classify these tea classes is to use normal transfer learning training approach (Fig. 1), which has three parts: image pre-processing, image augmentation, and model training. In the experiment, four pre-trained models (pre-trained weight using ImageNet), i.e., Xception, VGG16, InceptionV3, and InceptionResNetV2 models [7-10], are used to classify tea classes. The overall transfer-learning model architecture with its parameters and neural layer information can be seen in Fig. 2. The first part in Fig. 1 is image pre-processing, where all tea images are resized to 299×299 pixel size with three color channels. The second part, data augmentation, is used to mitigate small data size issue. Data augmentation methods used in the experiment including randomly image rotation (0 ~ 30 degree), randomly horizontal flip, and randomly vertical flip. Through initial experiment, the result shows that the normal transfer learning training approach performs poorly for tea image classification task (Table 1). In Table 1, the first column displays several pre-trained deep learning models applied for tea classification. The third column shows the test accuracy of each trained model. The result shows that none of these models' accuracy is greater than 70%, which suggests that the normal transfer learning training approach is not suitable for the dataset in this study. Table 2 shows models' hyperparameters; the same hyperparameters are used for all models.

The low accuracy issue can be explained from two perspectives. First, the small data size used to train models can result in model overfitting. As illustrated in Fig. 3, the model overfitting tendency can be observed from the changes of validation loss, which shows that the training data is not enough. The reason of only using 360 tea images is that the application scenario of the study is to develop a tea image verification technique for smallholder tea farmers. Therefore, small data size is inevitable in this study and becomes a research problem to be solved via algorithm design. Second, there may be some distinct properties between traditional images and tea images such that the normal transfer learning approaches do not work. Here, traditional image means the images without distributed objects, such as tea leaves. To examine the rationality of the above-mentioned explanations for low accuracy, two observations are conducted.

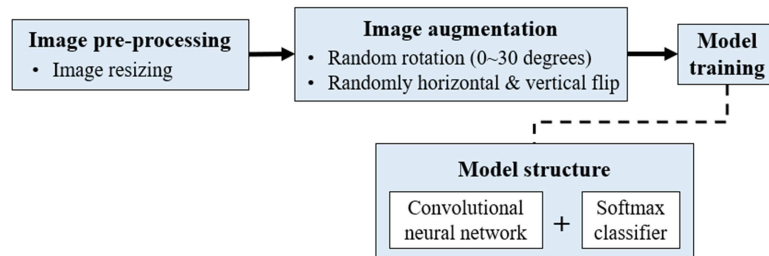


Fig. 1 Normal transfer learning approach in this study

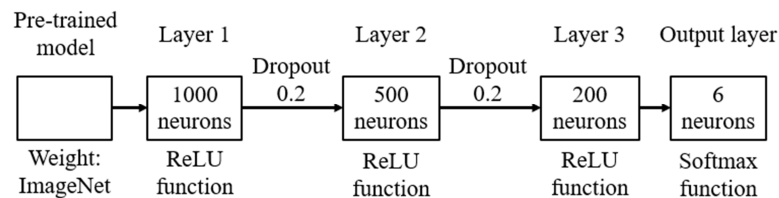


Fig. 2 Architecture of additional four dense layers

Table 1 Classification accuracy of training and testing data for oolong tea dataset in different pre-trained models using transfer learning

Pre-trained models	Accuracy of training data	Accuracy of testing data
Xception	73.75%	60.00%
VGG16	65.83%	69.99%
InceptionV3	25.00%	20.00%
InceptionResNetV2	18.75%	16.67%

Table 2 Parameters and hyperparameters of models

Transfer learning weight	Loss function	Optimization function	Additional dense layers	Epochs	Learning rate
ImageNet	Categorical cross-entropy	RMSprop	4	500	1e-5

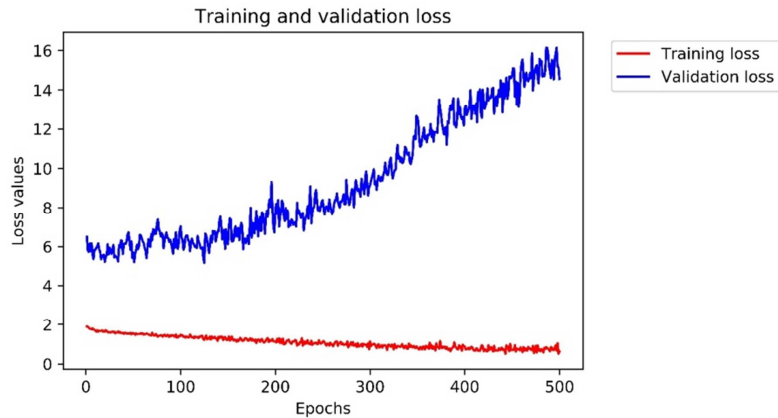


Fig. 3 Historical loss values in training process of InceptionV3 model using normal transfer learning approach

For the first observation, two algorithms, i.e., GradCAM++ and SmoothGrad [11-12], are used to visualize the pre-trained artificial intelligence (AI) model’s activation map and saliency map. In Fig. 4, there are three images presenting goldfish, dog, and tea. The goldfish and dog images are defined as traditional images because they do not include distributed objects as targets to be classified. The result of activation map and saliency map shows that: for traditional deep learning training approach, the model initially learns the low level local features from the image, such as edge and color, then integrates these low level local features into high level features, such as noses, eyes, and ears, and finally recognizes the image as an object, such as a dog’s face or a goldfish. However, unlike traditional images, tea image do not have local features, as shown in the activation map and saliency map. Tea image’s global feature is distributed into many objects, such as tea leaves and sticks. Therefore, these distributed objects’ feature need to be captured so that the image can be classified into the class it belongs to.

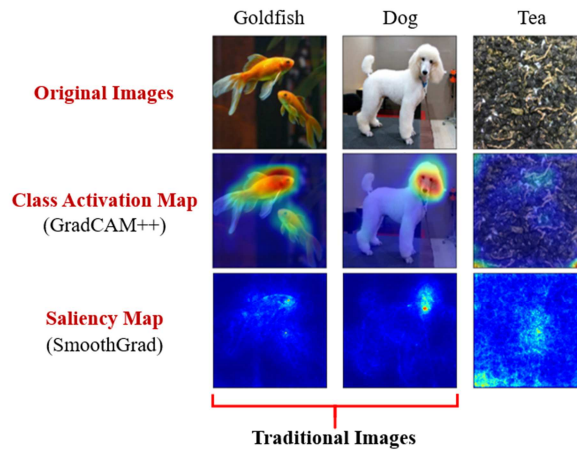


Fig. 4 Activation map and saliency map of GradCAM++ and SmoothGrad visualization algorithms

For the second observation, a tea image is cropped into many sub-images, and all sub-images look similar (Fig. 5). It can be seen that in a hypothetical feature space, the tea sub-images that belong to the same image should cluster together as their feature properties are statistically similar. To examine this idea, K-means are used to find the five most significant colors in the two tea classes (class 0 and class 3), and t-distributed stochastic neighbor embedding (t-SNE) algorithm is applied to plot these color feature vectors on a two-dimensional map (Fig. 6). Each dot in the figure represents a single sub-image. The red dots represent class 0’s cropped images and the blue dots represent class 3’s sub-images. The result suggests that the same class’s sub-images will cluster together in the feature space, and the two tea classes can be classified by using only tea color information.

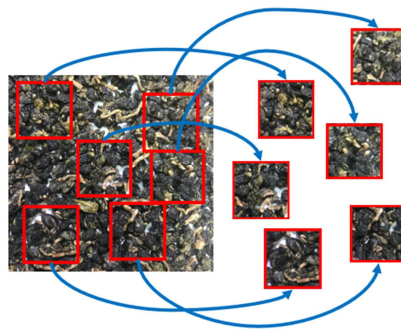


Fig. 5 Tea sub-images that belong to the same image

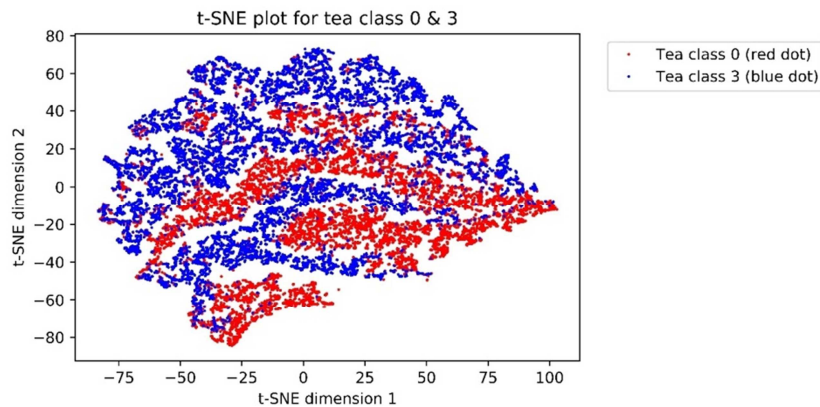


Fig. 6 t-SNE plot of two tea classes using the five most significant color information

The first observation using model visualization algorithms shows that the feature property between tea image and traditional image are different, i.e., distributed objects' feature property (tea images) versus non-distributed objects' feature property (traditional images). The second observation using K-means clustering algorithm and t-SNE algorithm finds that tea sub-images which belong to the same tea class have similar feature property, and suggests that: (1) small sample size problem may partially be solved by cropping every image into sub-images without distorting original tea image's property; (2) by leveraging tea image's unique characteristics, a new prediction method can be used to improve model's performance (see section 5.3.). These two observations support the idea that there exist some distinct properties between traditional images and tea images thus traditional deep learning training approach is not suitable to train tea images.

3. Dataset

There are six high-mountain oolong tea classes' samples collected in the study (Table 3 and Fig. 7). Each tea class has 60 images. 48 images are split for training set, 6 images are split for validation set, and 6 images are split for testing set. There are several reasons that make this dataset a challenging model training task for tea image classification comparing to the datasets of previous tea image classification studies [4-5, 13-16].

First, all tea classes belong to the same tea type (high-mountain oolong tea), meaning that all tea classes have the same tea product processing procedure. Second, among tea classes, only class 3 has different tea varieties. In addition, tea class 0, 1, and 2 belong to the same tea variety and they are planted in the same mountain; the only difference among them is their mountain adrets, meaning that they receive different sunlight degrees and angles while growing.

Finally, the small data sample size is challenging. It is known that deep learning models perform well when trained with large amount of data and perform poorly when the training data size is small, so why in this study the small sample size dataset is used to train model? The main reason is that this study's application scenario is for smallholder tea farmers whose tea products often go through middlemen before reach consumers, thus the amount of tea products each farmer can produce for every season is low. Therefore, it is reasonable to expect that the amount of image data obtained from original tea products is rather little.

Table 3 Information of six high-mountain oolong tea classes

Tea classes	Tea varieties	Places of origin	Note
Class 0	Chin-shin-oolong	Hehuan mountain, Nantou county	Adret: north
Class 1	Chin-shin-oolong	Hehuan mountain, Nantou county	Adret: east
Class 2	Chin-shin-oolong	Hehuan mountain, Nantou county	Adret: southwest
Class 3	Taiwan tea experiment station (TTES) No. 12	Zhushan township, Nantou county	-
Class 4	Chin-shin-oolong	Ali mountain, Chiayi county	-
Class 5	Chin-shin-oolong	Lugu township, Nantou county	-

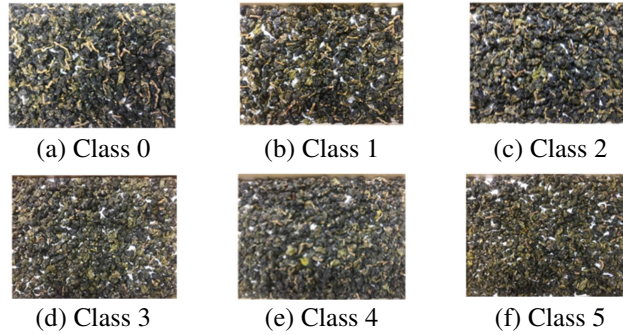


Fig. 7 Images of six high-mountain oolong tea classes

4. Experiment Design

Fig. 8 shows the overflow of the proposed approach to solve model convergence problem. This approach has five stages: image acquisition, image pre-processing, model training, and model evaluation. There are three innovative parts in the proposed approach. First, in image pre-processing stage, a tea image cropping process is added to get tea sub-images. Second, in model training stage, convolutional neural network (pre-trained InceptionV3 model) and XGBoost (extreme gradient boost tree) algorithm [3] are combined as a single model. Third, in model evaluation stage, instead of inputting sub-images directly to the trained model for prediction, a majority voting mechanism is added to evaluate model performance.

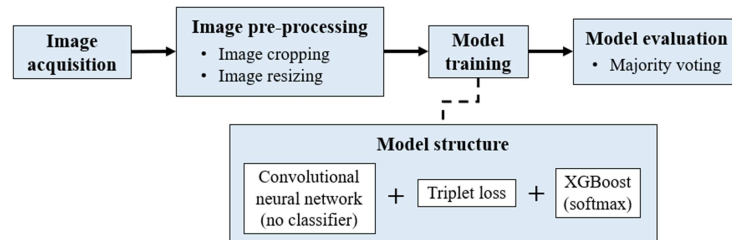


Fig. 8 The proposed approach for tea image classification

5. Algorithms

5.1. Triplet loss

Triplet loss is a metric learning [17]. It selects three tea images ($x^a_{(i)}$, $x^p_{(i)}$, and $x^n_{(i)}$) from any two tea classes, where $x^a_{(i)}$ is called anchor image, $x^p_{(i)}$ is called positive image, and $x^n_{(i)}$ is called negative image; $x^a_{(i)}$ and $x^p_{(i)}$ belong to the same tea class and $x^n_{(i)}$ belongs to a different tea class. The training goal is to minimize the L2 distance between anchor image and positive image ($D^{a,p}_{(i)} = \|x^a_{(i)} - x^p_{(i)}\|$) as they are same tea class images, and maximize the L2 distance between anchor image and negative image ($D^{a,n}_{(i)} = \|x^a_{(i)} - x^n_{(i)}\|$) as they are different tea categories images. In order to enlarge the learned distance between $D^{a,p}_{(i)}$ and $D^{a,n}_{(i)}$, a margin which is greater than 0 is also added:

$$\sum_i^n [D^{a,p}_{(i)} - D^{a,n}_{(i)} + margin]_+ \tag{1}$$

When using triplet loss in model training, one tea class is labeled as positive category and other five tea classes are labeled as negative category. Since there are six tea classes, six models can be trained, each time assigning one class as positive category and other five classes as negative category. For example, if tea class 0 is labeled as positive category, then tea class 1 ~ class 5 will be labeled as negative category.

5.2. Majority voting

After a test image is cropped into multiple tea sub-images and input to the trained model for evaluation, the majority voting mechanism will calculate how many percentage of tea sub-images are predicted as positive category and negative category, and the model's accuracy is obtained (Algorithm 1). For example, if a model is trained as class 0 being positive category and class 1 ~ class 5 being negative category; when a test image that belongs to class 0 is input, its sub-images predicted as class 0 will be counted as positive category, and its sub-images predicted as class 1 ~ class 5 will be counted as negative category. If the same tea sub-images' feature properties are statistically similar as Fig. 7 shows, most class 0's sub-images should be predicted as positive category in this example (Fig. 9).

Algorithm 1: majority voting

Abbreviations

PC: positive category

NC: negative category

PPOS: prediction percentage of sub-images

Input: tea images in the test set

Output: model's accuracy

```

1 true_positive = 0
2 false_positive = 0
3 true_negative = 0
4 false_negative = 0
5 for each_image in test_set:
6     if image's_label == PC:
7         if PPOS_in_PC > PPOS_in_NC:
8             true_positive += 1
9         else:
10            false_positive += 1
11    if image's_label == NC:
12        if PPOS_in_NC > PPOS_in_PC:
13            true_negative += 1
14        else:
15            false_negative += 1
16 model_accuracy =  $\frac{\text{true\_positive} + \text{true\_negative}}{\text{total\_test\_image\_number}} \times 100\%$ 

```

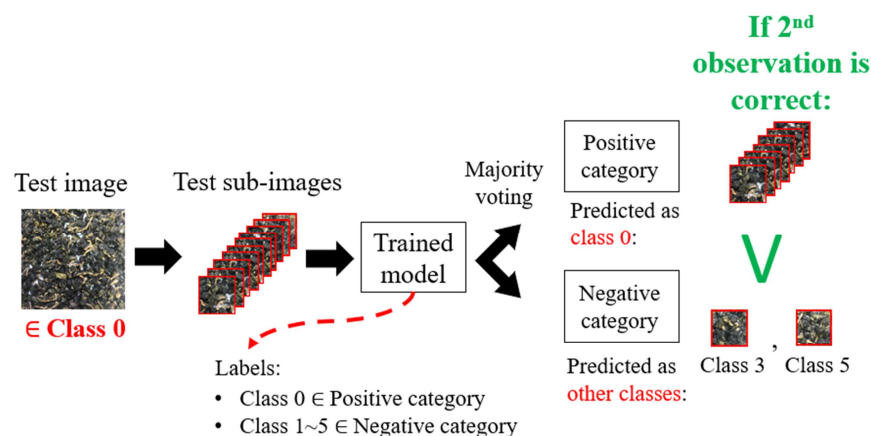


Fig. 9 Illustration of how majority voting works if the second observation is correct

6. Image-Preprocessing

6.1. Data acquisition

Due to tea image's sensitivity to brightness, shadow, and angle rotation changes, a standard operation procedure (SOP) is set to photograph the tea leaves. If tea image quality is low, then it is difficult to converge the model.

(1) SOP goal: Get a batch of clear and high-resolution tea photos with uniform brightness.

(2) Tools:

1. A paper box, having a size of 1/4 A4 paper area (Fig. 10).
2. A table lamp.
3. Six zipper storage bags.
4. A camera. In this study, iPhone 7 plus camera is used. The information about this camera are as follows:
 - A. Dual 12MP wide-angle and telephoto cameras.
 - B. Wide-angle: $f/1.8$ aperture; telephoto: $f/2.8$ aperture.
 - C. 2x optical zoom.
 - D. Autofocus with focus pixels.
 - E. Automatic exposure control.
 - F. Backside illumination sensor.
 - G. Hybrid IR filter.

(3) Procedure:

1. Preparation stage:
 - A. Pour tea leaves from the vacuum packaging into a zipper storage bag.
 - B. Put table lamp at paper box's left side.
2. Photograph stage:
 - A. Pour tea leaves from the zipper storage bag into paper box.
 - B. Shake the paper box to spread the tea leaves evenly on the paper.
 - C. Photograph the tea leaves using camera. Each image must have good focus, high resolution, and uniform brightness.
 - D. Repeat above steps (step A ~ C) until the desired image number is achieved.
 - E. Finally, check image quality. Replace unqualified image with new image.

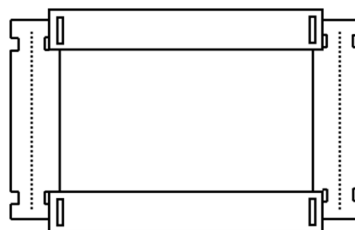


Fig. 10 Paper box template

6.2. Image pre-processing

There are three steps in tea image pre-processing stage:

Step 1: Assume an image's size is $X \times Y$, where X is the shorter side. (In this study, the image resolution is 3412×1920 pixels.) This image is cropped into $X \times X$ (Fig. 11).

Step 2: The second step is cropping the surrounding 100 pixels of image, in order to remove the paper box wall part in the image (Fig. 12) to get a new image (image size: $(X-200) \times (X-200)$) and resize the new image to 1536×1536 pixels (Fig. 13).

Step 3: The third step is cropping the square image into multiple 450×450 pixels sub-images where each sub-image is cropped for every 20 pixels spacing. The image cropping procedure serves two purposes: the first purpose is to examine the effectiveness of majority voting, which originates from the second observation; the second purpose is for data augmentation, which increases data sample size by cropping image into sub-images. Finally, all sub-images are resized into 224×224 pixels and converted into gray scale images (Figs. 14 and 15). These final gray scale sub-images will be input to the model for training.

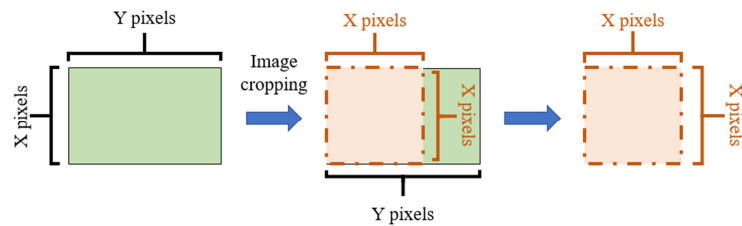


Fig. 11 Step 1 in image pre-processing stage



Fig. 12 The paper box wall part in the tea image

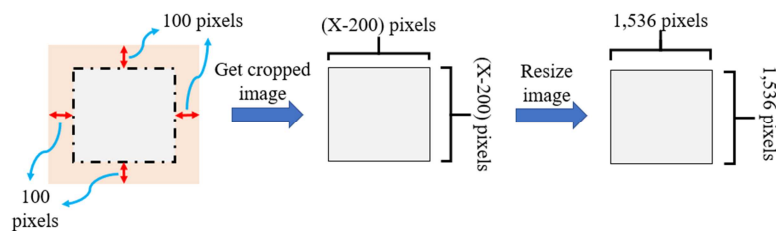


Fig. 13 Step 2 in image pre-processing stage

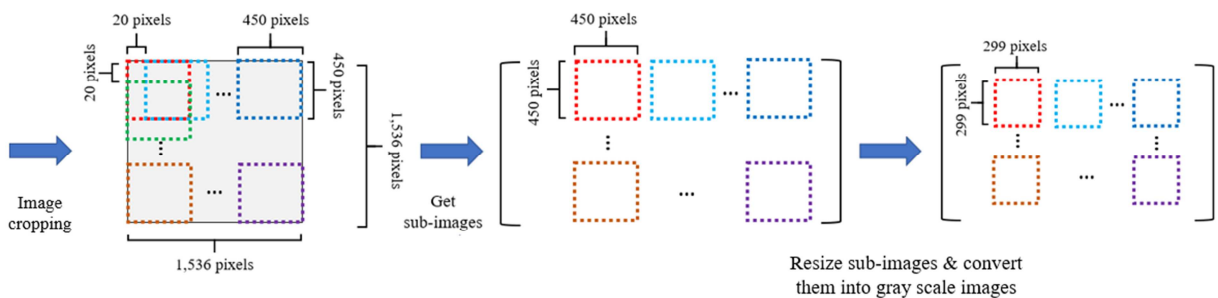


Fig. 14 Step 3 in image pre-processing stage

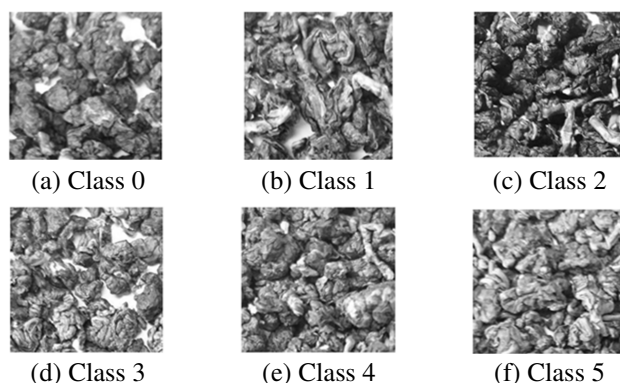


Fig. 15 Oolong tea classes gray scale sub-images

7. Experiment

7.1. Machine learning model

The whole model is composed of two sub-models: the first sub-model is a pre-trained InceptionV3 model (weight: ImageNet) plus three additional layers, and the second sub-model is XGBoost model. Fig. 16 shows overall model architecture and layer shape information. Table 4 and Table 5 show the two sub-models' parameters and hyperparameters. In training process, n 224×224 pixels gray scale training sub-images are input to the pre-trained InceptionV3 model conducting transfer learning. After the first trained sub-model is obtained, these n sub-images are input to the first trained sub-model for prediction, and n 128-dimensional embedding vectors are obtained. Finally, these vectors are input to train the XGBoost model.

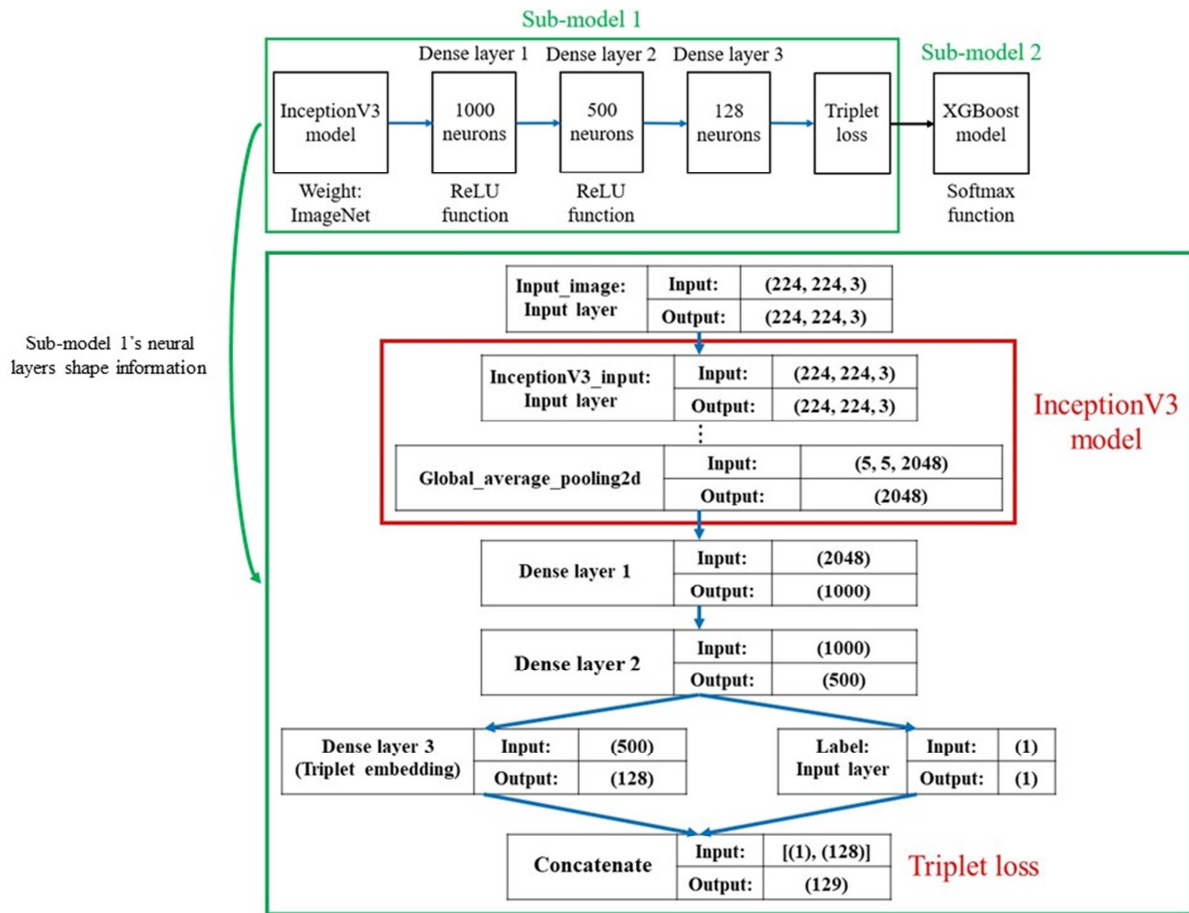


Fig. 16 Model architecture and neural layer shape

Table 4 Parameters and hyperparameters of InceptionV3

Model	Optimization function	Epochs	Learning rate	Loss function	Triplet loss threshold value
InceptionV3	Adaptive moment estimation (ADAM)	5000	1e-6	Triplet loss	0.001

Table 5 Parameters and hyperparameters of XGBoost model

Model	Class number	Tree depth	Epochs	Learning rate	Tree number of boosting iterations
XGBoost	6	20	500	0.1	20

7.2. Hardware and software

Programs are run on NVIDIA DGX-1 machine, which has 8 GPUs (Tesla V100-SXM2) and 80 CPUs (20-Core Intel® Xeon® E5-2698). For software, Python v3.6.7 is used with modules Tensorflow (1.15.0), XGBoost (v0.90), IPython (v7.5.0), Matplotlib (v3.1.0), Scikit-learn (v0.19.1), Pandas (v0.24.2), and Numpy (v1.16.2).

8. Experiment Results

Table 6 shows the experiment result using the proposed approach. The test accuracy of all models is greater than 97%, which is better than normal transfer learning training approach. Fig. 17 shows the majority voting result of class 0 ~ class 5 oolong tea test images using “class 0 as positive category” model, which shows that the test sub-images that belong to class 0 get higher prediction percentage in positive category and the test sub-images that belong to class 1 ~ class 5 get the higher prediction percentage in negative category. In each sub-plot, the X-axis is test image sequence and Y-axis is sub-images prediction percentage of positive category and negative category. Positive category label is presented as red dots and negative category label is presented as blue dots. Fig. 17 suggests that the second observation is correct, meaning that the tea sub-images that belong to the same image have similar property, thus they should cluster together on the feature space. This implies that if an image belongs to positive category, most of its sub-images should be predicted as positive category. In the appendix, Figs. 18-22 show the majority voting results of class 0 ~ class 5 oolong tea test images using class 1 ~ class 5 as positive category models. Table 7 shows the result comparison between the proposed work and the previous work. Table 8 shows other model’s performance using class 0 as positive category.

Table 6 Classification accuracy of testing data using the proposed approach and other publications’ results

Pre-trained InceptionV3 models	Accuracy of testing data
Class 0 as positive category	100.00%
Class 1 as positive category	97.22%
Class 2 as positive category	100.00%
Class 3 as positive category	100.00%
Class 4 as positive category	100.00%
Class 5 as positive category	100.00%

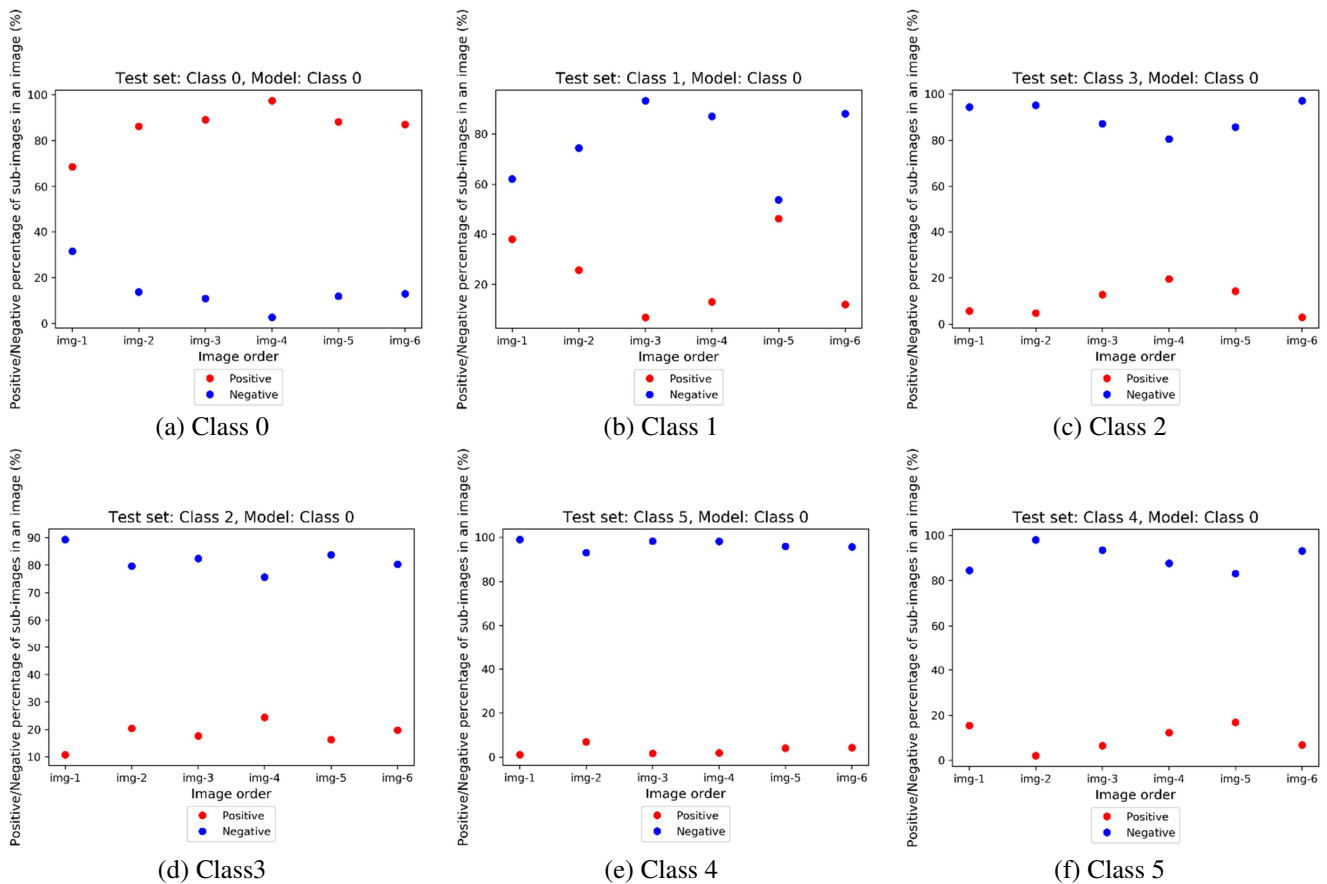


Fig. 17 The sub-images prediction percentage results of class 0 ~ class 5 oolong tea test images using “class 0 as positive category” model

Table 7 The comparison between the proposed work and the previous work

Ref.	Models	Tea types	Total sample number	Average accuracy
This work	Pre-trained InceptionV3 model + XGBoost + majority voting	6 high-mountain oolong tea classes (Among them, 5 classes belong to the same tea variety.)	360 images	99.54%
[4]	12-layer convolutional neural network	3 tea classes including green tea, oolong tea, and black tea	900 images	98.33%
[5]	Histogram equalization + gray-level co-occurrence matrix + support vector machine (SVM)	10 tea classes from different brands and origins, including 1 white tea, 2 different kinds of flowering tea, 2 different kinds of green tea, 1 oolong tea, and 4 different varieties of black tea	3000 images	94.64%
[15]	Principal component analysis + linear discriminant analysis	5 green tea classes from different varieties, brands, and origins	120 images	98.33%
[16]	Fuzzy SVM + winner-take-all method	3 tea classes including green tea, oolong tea, and black tea	300 images	97.77%

Table 8 Classification accuracy of testing data using the proposed approach

Pre-trained models	Accuracy of testing data (class 0 as positive category)
VGG16	100.00%
Xception	100.00%
InceptionResNetV2	100.00%

9. Conclusions

This study proposed an image verification approach to improve deep learning model's performance. The result outperforms normal transfer learning training approach in the challenging tea dataset. Future studies can explore feature engineering methods to extract tea image's feature properties and complement the lack of interpretability of deep learning model. In addition, the proposed approach may have broader application to other crop types, such as rice, coffee bean, wheat seed, etc., which seem to have similar distributed objects' feature with tea images.

Acknowledgments

This research is partially supported by the Ministry of Science and Technology through both the Center of Precision Medicine Research and Artificial Intelligence Research Center of Asia University Taiwan under the grants SKI MOST 110-2321-B-468-001, MOST 110-2321-B-468-001, and 108-2511-H-468-001-MY3.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] V. Voora, S. Bermúdez, and C. Larrea, "Global Market Report: Tea," <https://www.iisd.org/publications/global-market-report-tea>, December 16, 2019.
- [2] W. Pongsuwan, E. Fukusaki, T. Bamba, T. Yonetani, T. Yamahara, and A. Kobayashi, "Prediction of Japanese Green Tea Ranking by Gas Chromatography/Mass Spectrometry-Based Hydrophilic Metabolite Fingerprinting," *Journal of Agricultural and Food Chemistry*, vol. 55, no. 2, pp. 231-236, January 2007.
- [3] W. Shao, C. Powell, and M. N. Clifford, "The Analysis by HPLC of Green, Black and Pu'er Teas Produced in Yunnan," *Journal of the Science of Food and Agriculture*, vol. 69, no. 4, pp. 535-540, December 1995.
- [4] Y. D. Zhang, K. Muhammad, and C. Tang, "Twelve-Layer Deep Convolutional Neural Network with Stochastic Pooling for Tea Category Classification on GPU Platform," *Multimedia Tools and Applications*, vol. 77, no. 17, pp. 22821-22839, September 2018.
- [5] Y. Chen, "Identification of Tea Leaf Based on Histogram Equalization, Gray-Level Co-Occurrence Matrix and Support Vector Machine Algorithm," *International Conference on Multimedia Technology and Enhanced Learning*, April 2020, pp. 3-16.
- [6] Y. Zhang, X. Yang, C. Cattani, R. V. Rao, S. Wang, and P. Phillips, "Tea Category Identification Using a Novel Fractional Fourier Entropy and Jaya Algorithm," *Entropy*, vol. 18, no. 3, 77, March 2016.

- [7] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," IEEE Conference on Computer Vision and Pattern Recognition, July 2017, pp. 1251-1258.
- [8] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," <https://arxiv.org/pdf/1409.1556.pdf>, April 10, 2015.
- [9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," IEEE Conference on Computer Vision and Pattern Recognition, June 2016, pp. 2818-2826.
- [10] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning," <https://arxiv.org/pdf/1602.07261.pdf>, August 23, 2016.
- [11] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," IEEE International Conference on Computer Vision, October 2017, pp. 618-626.
- [12] D. Smilkov, N. Thorat, B. Kim, F. Viégas, and M. Wattenberg, "SmoothGrad: Removing Noise by Adding Noise," <https://arxiv.org/pdf/1706.03825.pdf>, June 12, 2017.
- [13] L. Zhang, *Advances in Forest Management under Global Change*, London: IntechOpen, 2020.
- [14] X. Li, P. Nie, Z. J. Qiu, and Y. He, "Using Wavelet Transform and Multi-Class Least Square Support Vector Machine in Multi-Spectral Imaging Classification of Chinese Famous Tea," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11149-11159, September 2011.
- [15] Q. Chen, J. Zhao, and J. Cai, "Identification of Tea Varieties Using Computer Vision," *Transactions of the Asabe*, vol. 51, no. 2, pp. 623-628, March-April 2008.
- [16] S. Wang, X. Yang, Y. Zhang, P. Phillips, J. Yang, and T. F. Yuan, "Identification of Green, Oolong and Black Teas in China via Wavelet Packet Entropy and Fuzzy Support Vector Machine," *Entropy*, vol. 17, no. 10, pp. 6663-6682, October 2015.
- [17] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," IEEE Conference on Computer Vision and Pattern Recognition, June 2015, pp. 815-823.



Copyright© by the authors. Licensee TAETI, Taiwan. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-NC) license (<https://creativecommons.org/licenses/by-nc/4.0/>).

Appendix

The appendix shows the sub-images' prediction percentage results of class 0 ~ class 5 oolong tea test images using class 1 ~ class 5 as positive category models (Figs. 18-22).

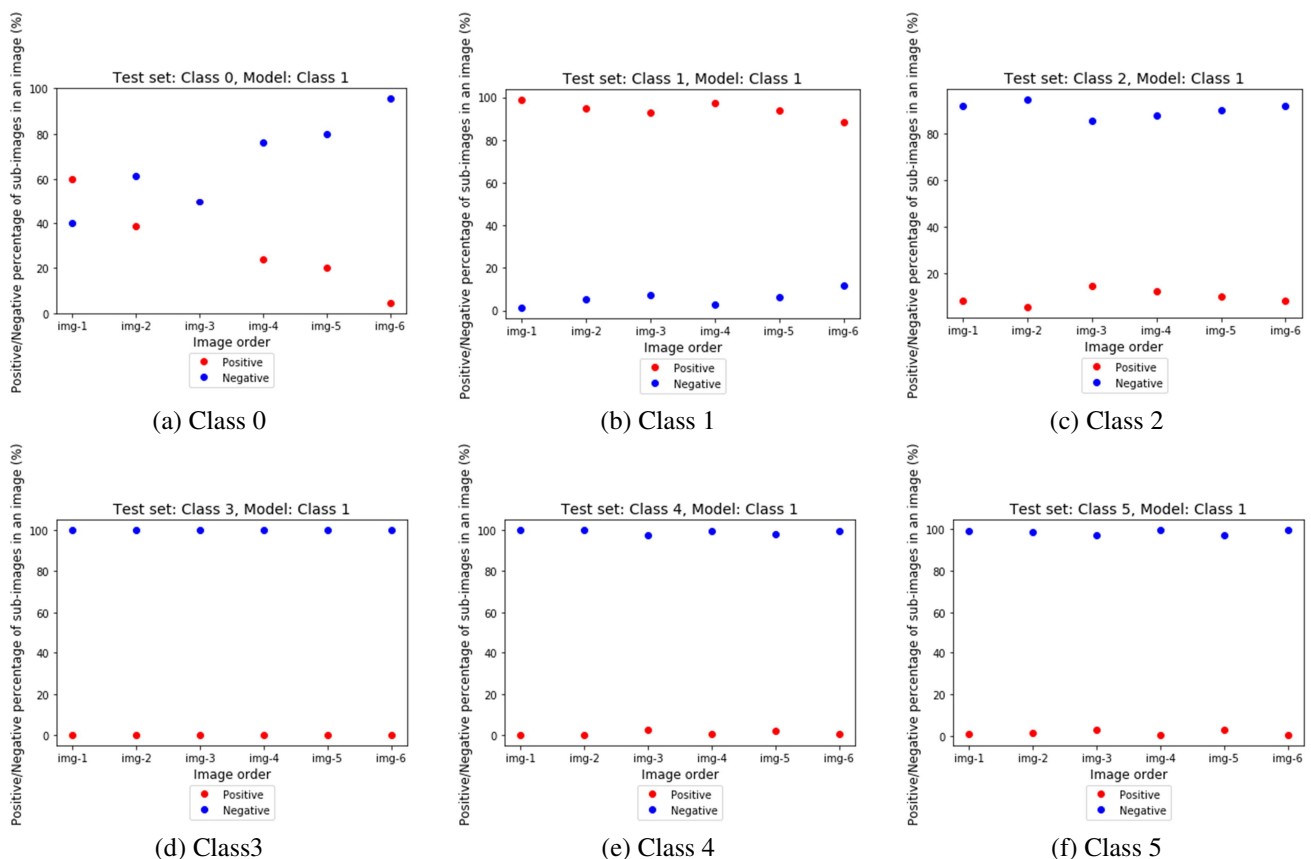


Fig. 18 The sub-images prediction percentage results of class 0 ~ class 5 oolong tea test images using "class 1 as positive category" model

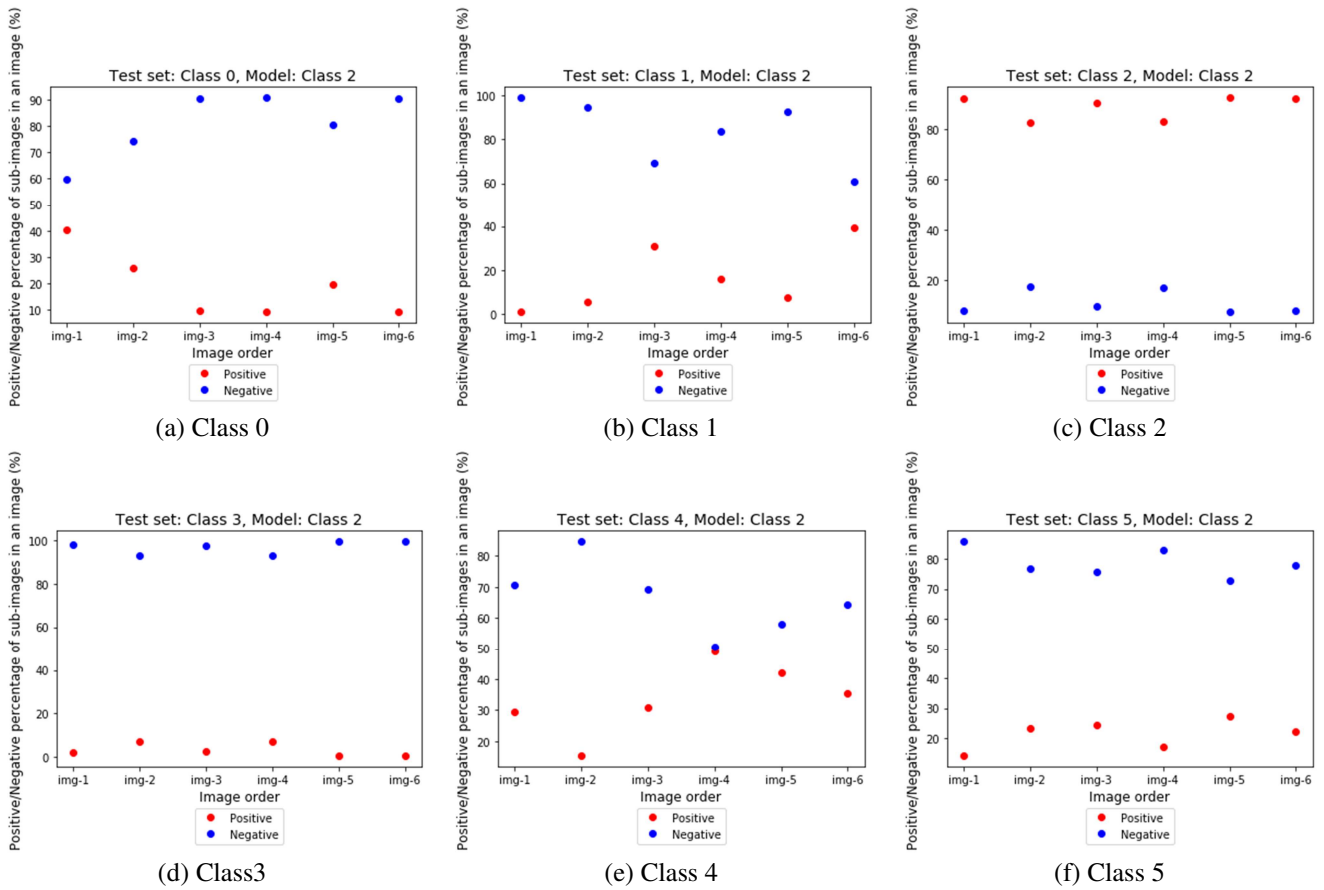


Fig. 19 The sub-images prediction percentage results of class 0 ~ class 5 oolong tea test images using “class 2 as positive category” model

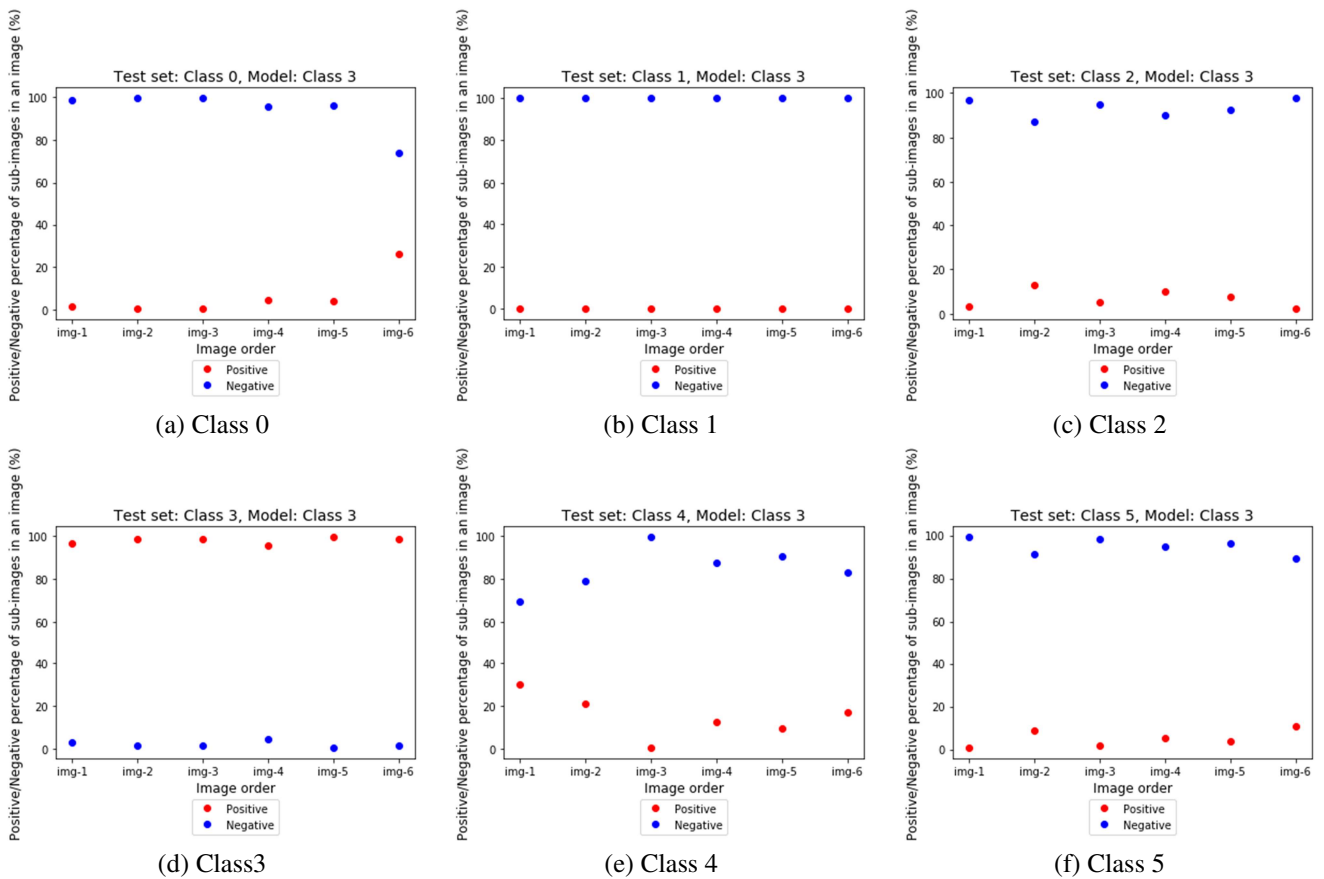


Fig. 20 The sub-images prediction percentage results of class 0 ~ class 5 oolong tea test images using “class 3 as positive category” model

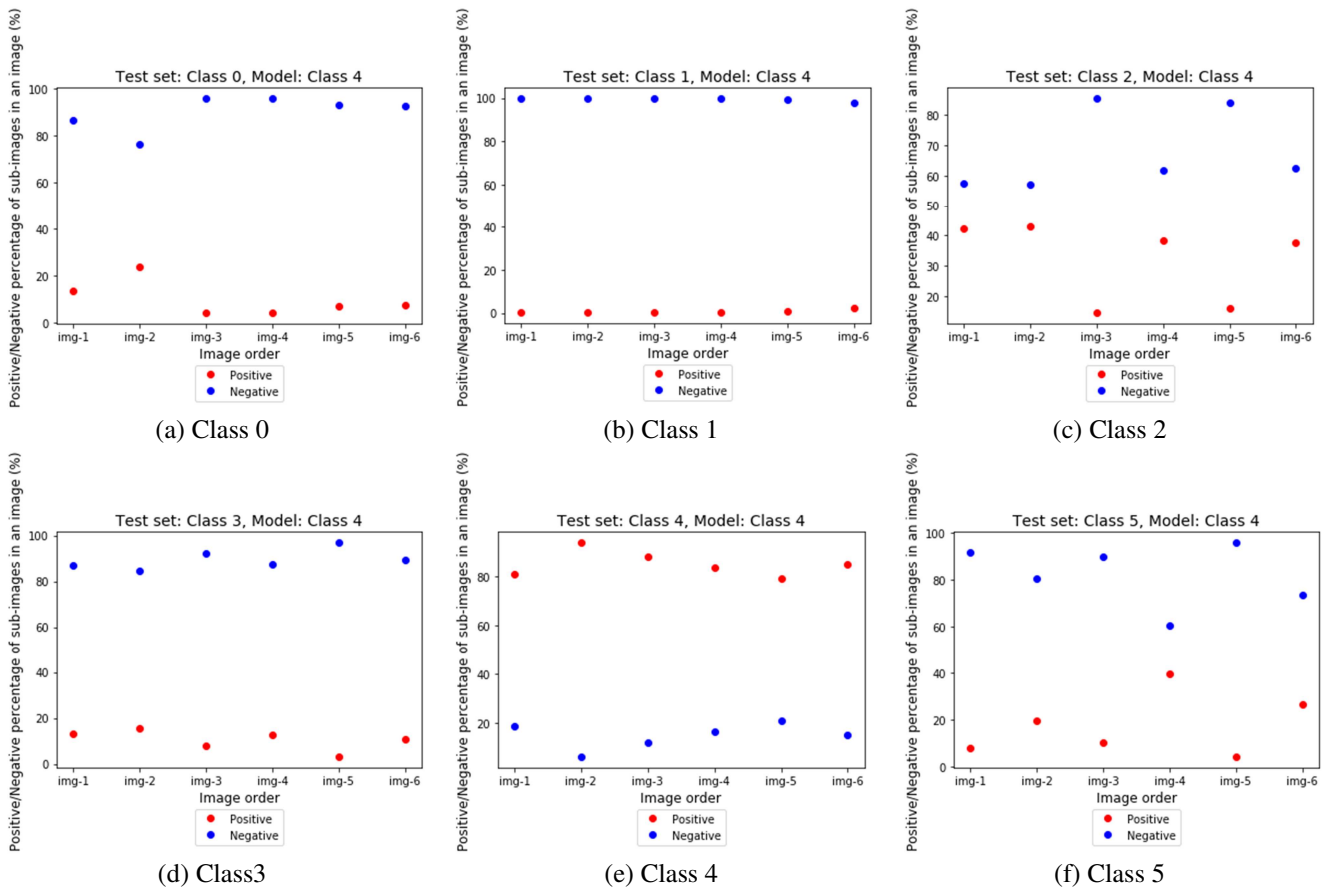


Fig. 21 The sub-images prediction percentage results of class 0 ~ class 5 oolong tea test images using “class 4 as positive category” model

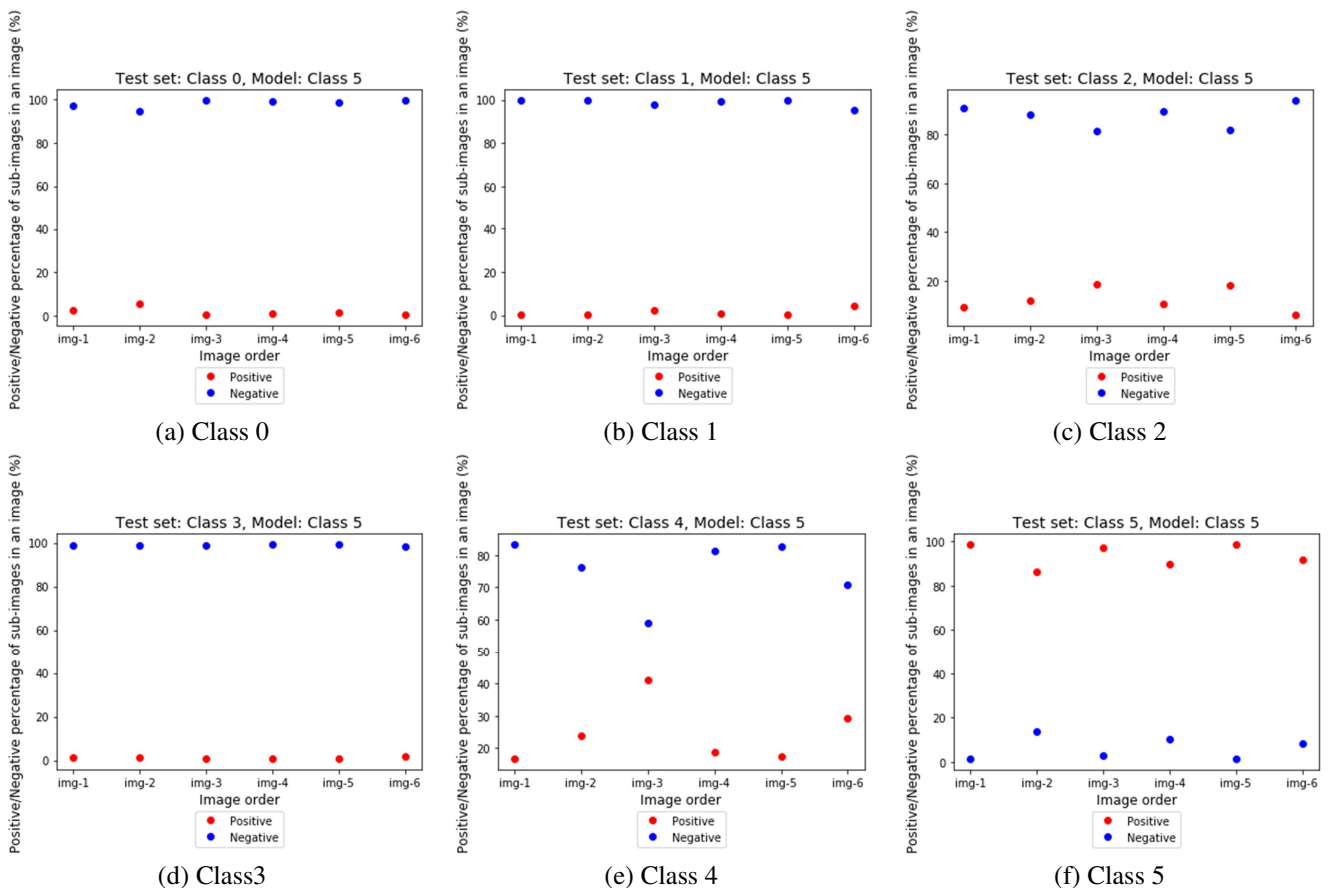


Fig. 22 The sub-images prediction percentage results of class 0 ~ class 5 oolong tea test images using “class 5 as positive category” model