A Neural Network-based Approach for Apple Leaf Disease Detection in Smart Agriculture Application

Shengjie Gan*, Defeng Zhou, Yuan Cui, Jing Lv
Hunan Chemical Vocational Technology College
Zhuzhou 412004, Hunan, China

***Abstract*—Plant diseases significantly harm agriculture, which has an impact on nations' economies and levels of food security. Early plant disease detection is essential in smart agriculture. For the diagnosis of plant diseases, a number of methods, including imaging, have been used recently. Some of the existing methods for plant disease detection using imaging have limitations as firstly, high computational cost, some methods require complex image processing algorithms or manual design of features that can increase the time and resources needed for the detection. Secondly, low accuracy, most of the methods rely on simple classifiers or handcrafted features that may not capture the subtle differences between different diseases or healthy leaves. Thirdly, dependency on expert knowledge, some methods need human intervention or prior knowledge of the diseases and pests to perform the detection. These limitations are not suitable for the problem at hand because they can affect the efficiency of the detection system. In this study, three apple tree leaf diseases—apple black spot, Alternaria, and Minoz blight—are detected using a neural network (NN) and a digital image processing technique. The sample images are prepared, processed, and used to extract attributes using a digital image processing approach, and the NN is used to classify the diseases. An evaluation of the proposed system's performance in identifying illnesses in apple trees shows satisfactory accuracy and strong overall performance. Additionally, when compared to other techniques already in use, this strategy is more effective at diagnosing.**

Keywords*—*Smart agriculture; plant disease; apple leaf disease; image processing; neural network

I. INTRODUCTION

According to the statistics of FAO, America has managed to earn 1.1 billion dollars by producing 4.08 million tons of apples on average, in addition to meeting the domestic demand to be the largest exporter of this product worldwide [1-3]. Principally, the price and balanced market of agricultural products depend on the quality. Plant diseases can significantly reduce product quality and quantity, leading to a decline in the economies of nations that depend on exporting agricultural products [4]. If identified in the earliest stages of development, these disorders may, in some cases, be prevented and controlled. In this regard, some countries are looking for ways to diagnose plant diseases early [5].

Among the most common solutions in the past has been the visual diagnosis of the disease by experienced experts, which, in large farms, typically costs a lot and calls for ongoing plant pathologist monitoring [6]. However, many diseases don't have obvious symptoms in the first stages of the disease, and it is difficult to identify them with the naked eye. For this reason, it is necessary to investigate and provide a fast, automatic, low-cost, and precise tool for identifying plant diseases. In recent decades, the growth of technology has increased progress in various branches of science, industry, and agriculture; therefore, various researchers have sought to use new technologies for the early and timely diagnosis of various plant diseases [7]. The technologies used are mainly in two parts, the type of sample processing system and classification models. Among the different sample processing systems, we can mention the types of imaging systems, the smelling machine, and the tasting machine. The different types of statistical models, data mining, and artificial intelligence can also be mentioned from the different classification models [8]. Usually, sample processing systems are selected according to the type and symptoms of plant diseases [2].

The imaging system has attracted much attention recently due to its advantages, such as non-destructiveness, reduction of human resources costs, and high accuracy. Different researchers of this system diagnose palm leaf nutritional disease, classify minnow, black spot, and Alternaria (apple tree leaf) diseases, and diagnose tea leaf disease [9]. The diseases are internal powdery mildew, bacterial angular spot, ring spot, spot, gray rot, anthracnose, and powdery mildew [10]. Since no signs of them can be seen in the beginning and spread of the disease, more advanced systems than usual imaging should be used to extract the characteristics of the disease. Therefore, the researchers considered the hyperspectral imaging system because this system, in addition to the spatial features related to the disease, also extracts the spectral features, which are very useful in the early diagnosis of the disease. On the other hand, this system is costly, time-consuming, and requires trained people, which may only be possible for some farmers to use. The odor machine is another type of sample processing system that has advantages such as non-destructiveness, high reliability, and easy and fast use, and it is not effective in cases where there are no signs of discoloration of the disease in the samples [11].

Researchers have also developed more accurate classification models by studying the recent research done by different researchers to diagnose plant diseases, which can be advanced image processing systems [12]. In this regard, the classification model can be used in a study to diagnose and classify diseases (mosaic, leaf rust, and round spot) of apple tree leaves as well as diseases (brown spot, grey spot, and round spot) of corn.

This study proposes a method to identify and categorize three apple tree leaf diseases using digital image processing and a neural network (apple black spot, Alternaria, and Minoz blight). The proposed method aims to overcome these limitations by using a neural network and a digital image processing technique that can extract features and classify diseases with high accuracy and low complexity.

The main contributions of this study are as follows:

1) The study, which focuses on the disease of apple tree leaves, provides a digital image processing technique specifically made for identifying plant diseases. The preparation, processing, and extraction of pertinent properties from sample images made possible by this technique enable precise and effective disease identification.

2) This study investigated a neural network (NN) to classify diseases and trained it using information gathered from digital image processing technology. This strategy shows how NNs are effective at correctly classifying and diagnosing diseases of apple trees, advancing automated and trustworthy plant disease diagnosis systems.

The rest of this paper structure as follows, Section II reviews the related works. Section III discuss about proposed method. Section IV presents results and discussion. Finally, this paper concludes in Section V.

II. Related Works

In study [13], a method presented for detecting apple leaf disease. This approach is according to technologies for pattern recognition and image processing. The input RGB (Blue, Green, and Red) picture was first given a color transformation structure, and the RGB model was then transformed into the Hue, Saturation, and Intensity (HSI), YUV, and grey models. Following removing behind-the-scenes, the sickness spot image was segmented using a region-growing algorithm (RGA) and a predefined threshold value. Each spot picture had thirty-eight color, texture, and form classification characteristics extracted. Combining (GA) with a genetic algorithm allows for correlation-based feature selection of the most critical factors, enhancing the apple leaf disease detection accuracy while minimizing the feature space's complexity (CFS).

An automated disease leaf recognition approach is proposed for identifying and grading leaf diseases using machine vision and digital image processing [14, 15]. The proposed system is broken up into two phases. The plant is recognized in the first stage according to the characteristics of its leaves. This phase involves pre-processing leaf image data and feature extraction, followed by training and classification using an artificial neural network to identify leaf features [16]. In the second phase, the disease that affects the leaf is categorized, which entails segmenting the defective area using k-means, feature extraction from the wrong area, and disease classification using an artificial neural network (ANN). The degree of disease present in the leaf is then considered when assigning a disease grade.

A prediction model for plant leaf disease detection and classification utilizing computer vision and machine learning approaches was proposed in [17]. Pre-processing, segmentation, and extraction of properties, including shape, color, texture, vein, and so on, are done on the raw picture of a leaf. Several machine-learning classifiers are used to categorize the leaf image. The experimental outcomes are assessed and contrasted with those of K-Nearest Neighbor, Random Forest, Support Vector Machine, and Artificial Neural Network.

For precisely identifying plant leaf disease, the Boosted support vector machine-based Arithmetic optimization algorithm (BSVM-AOA) has been presented [18]. In this instance, the greyscale co-occurrence matrix is employed for feature extraction, and the vector value active contour model is used for picture segmentation. Furthermore, performance indicators, such as f-rating, recall, accuracy, and specificity, are used to gauge how well the proposed technique performs. The proposed method is compared to the various existing processes in a comparative analysis. Their findings revealed that the BSVM-AOA technique had a 98.6%.

III. Proposed Method

In this study, to diagnose apple tree leaf diseases, a method consisting of an image processing method and a neural network model has been developed. Fig. 1 depicts the proposed approach.

A. Pre-processing

1) Dataset collection: This research investigated three apple tree leaf diseases: Apple black spot, Alternaria disease, and minnow pest. Some sample images are collections inspired by [19]. Six hundred forty different samples of apple tree leaves, including data augmentation described in the next section, the dataset contains 320 leaves infected with Alternaria disease, 184 infected with apple black spot disease, and 136 infected with Minoz pest.

2) Data augmentation: Preparing images of leaf samples to prepare pictures of leaves and produce sample images using data augmentation. In this data augmentation, filters such as noising, blurring, contrast, and brightness are added to the sample images to extend the dataset [19]. Fig. 2 displays examples of the dataset's photos.

3) Image conversion: This step is image conversion RGB channel to L*a*b* (Lab*) channel. The Lab* is a color space that is based on the opponent color model of human vision, where red and green form an opponent pair and blue and yellow form an opponent pair. It expresses color as three values: L* for perceptual lightness and a* and b* for the four unique colors of human vision: red, green, blue and yellow.

RGB to Python programming language and the Open CV library are employed for image processing and system implementation [20]. Image processing steps are composed of removing the background, removing leaf tails, removing unwanted regions such as noises, and the RGB channel to the L*a*b* channel conversion of the picture due to the proximity of this channel to the human visual system and removing the L component to eliminate the effect of brightness.

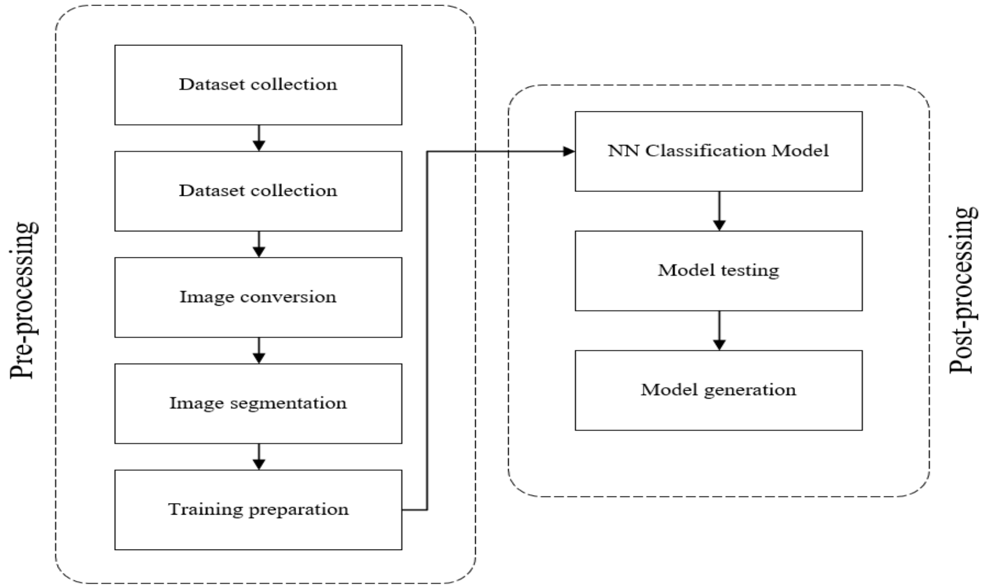

Fig. 1. The proposed method.

Fig. 2. Sample images in the dataset.

4) *Images segmentation:* Segmentation was done based on the area by extracting the contaminated area using the k-means technique. Image clustering commands were applied to only *a and *b components using the k-means method, and the images were divided into healthy and infected leaf areas. After this stage, because the disease spot had a smaller surface than the leaf surface, this area was selected. The remaining image processing and feature extraction operations were used on the picture of the diseased spot.

5) *Feature extraction:* In this stage, a collection of characteristics is extracted to describe the contaminated regions. It is according to color, shape, and textural attributes used to describe the regions. Since the elements of color and texture are different for each apple leaf disease, only these features are used for classification. The wavelet and co-occurrence matrix from the grey level is used for texture features. Hence, a maximum element including 14 color features including intensity, angle, four statistical features belonging to the *a as well as *b components of the *L*a*b and R space from the RGB space, and 24 wavelet features, two statistical elements for four coefficients in three Wavelet level, 32 features consisting of the co-occurrence matrix, eight statistical characteristics in four directions of 0, 45, 90 and 145 degrees were extracted for each type of disease.

6) *Training preparation:* In this section, the structure of different sets of neural networks is generated. This network structure divides the dataset into training, verification, likewise testing sets. In this study, training data accounted for 70% of the total data, 10% was used for validation data, and 20% was considered for the testing set.

B. *Post-Processing*

1) *Proposed NN classification model:* Choosing a suitable architecture for the neural network significantly impacts its classification and diagnosis performance. In this study, a multi-layer perceptron (MLP) neural network has been used,

where the first layer has 33 input nodes, and the last layer contains one output node.

It is also essential to choose the correct number of hidden layers and neurons for each layer. According to extensive experimental and results comparison, the structure of the MLP involves 70 nodes for the input layer and one node for the output layer. Furthermore, one and two hidden layers were used, and the number of 3 to 5 neurons was also considered for each of these layers. For the neurons of the network's hidden layers, the hyperbolic and sigmoid tangent transfer function was selected. For the last layer, it was decided to use the linear transfer function.

2) Model testing: After proposing the NN model, it is required to test the classification using testing data. In this research, to assess the accuracy of the proposed model, two assessment metrics are used Correlation Coefficient (CC) and Mean Square Error (MSE).

$$CC = \frac{\sum_{i=1}^{N}(X_i - \underline{X})(Y_i - \underline{Y})}{\sqrt{(X_i - \underline{X})(Y_i - \underline{Y})}} \tag{1}$$

$$MSE = \sqrt{\frac{1}{N}\sum_{N=1}^{N}(X_n - Y_n)^2} \tag{2}$$

In the above Equations, N is the number of data, Y_i is the predicted value, X_i is the measured value, and $\underline{X}$, $\underline{Y}$ respectively represent the average measured and predicted values. In the performance evaluation of the proposed NN model, the input data enters the network through the first layer. After passing through the different layers of the network, the network's output is obtained. Now, having the network output and the measured output value, the error value (RMSE) is calculated.

3) Model generation: After ensuring the effectiveness of the proposed system in classification, it must be utilized as a tool for diagnosing apple tree leaf disease.

IV. Results and Discussion

This section provides the experimental result, performance evaluation and discussion of the proposed method in this study. In the next sub-section, the discussion is also presented to discuss in detail the results and clarify the efficiency of the method.

A. Experimental Results and Performance Evaluation

The first step to collecting the proposed model results is to prepare the data presented in full in the previous section. The next step is choosing the neural network architecture. To achieve the best neural network model, various network architectures were implemented and evaluated. As mentioned above, the MLP architecture was used as a high-efficiency architecture for neural network training. Fig. 3 shows experimental results using the generated model. In Fig. 3, leaf disease and non-disease are shown in red and green color boxes.

Table I evaluates the ability of the neural network trained by the NN model algorithm to diagnose apple tree diseases. Moreover, experimental results for the proposed method and Support Vector Regression (SVR) model were collected for apple tree disease detection, and their results were evaluated and compared. Table II indicates the result of the SVR model.

Six hundred forty data samples are available in the dataset to test the models. Among the available data, 320 of these leaves were infected with Alternaria disease, 184 were infected with apple black spot disease, and 136 were infected with Minoz blight. Table III shows how each model could correctly diagnose the number of disease samples. For example, the neural network model has been able to accurately analyze 316 samples of 320 samples related to Alternaria disease and wrongly diagnose four samples as black spot disease.

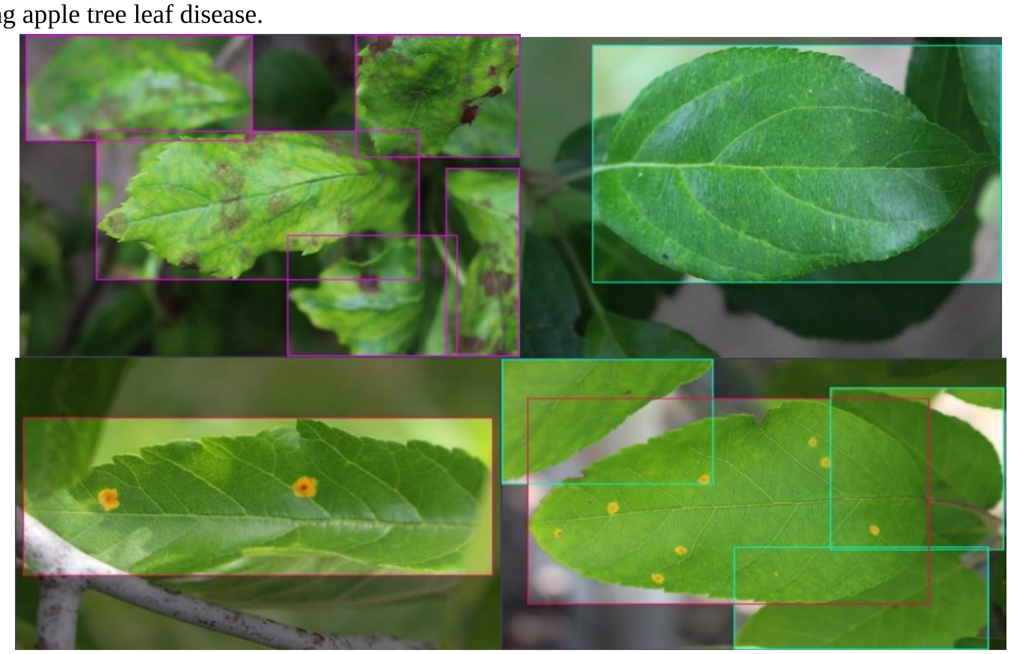

Fig. 3. Experimental results (includes leaf disease and non-disease).

TABLE I. PERFORMANCE EVALUATION FOR THE PROPOSED MODEL

Class	The proposed NN model						
	0	1	2	3	4	Total	Performance
1	0	316	4	0	0	320	0.98
2	0	4	179	1	0	184	0.97
3	0	1	5	130	0	136	0.95
Total	0	321	188	131	0	640	-

TABLE II. PERFORMANCE EVALUATION FOR THE SVR MODEL

Class	SVR model						
	0	1	2	3	4	Total	Performance
1	2	313	3	2	0	320	0.97
2	1	2	174	7	0	184	0.95
3	3	0	5	128	0	136	0.94
Total	6	315	182	137	0	640	-

TABLE III. RESULTS OF EVALUATION METRICS

Model	Training data		Testing **data**	
	RMSE	CC	RMSE	CC
ANN	0.091	0.978	0.098	0.976
SVR	0.12	0.978	0.30	0.963

Table III demonstrates that the proposed model's CC index for the test data is equivalent to 0.976, which is higher than the CC values obtained from other models. In addition, the RMSE index of the proposed model for the test data is equal to 0.098, which is lower than the values of other models. As a result, it can be said that the proposed model performs superior to other models in diagnosing the type of apple tree leaf disease.

B. *Discussion*

This section presents a discussion and in more detail of the proposed methodology in this research. As discussed earlier, this research emphasizes evaluating and comparing the proposed approach for identifying apple tree illnesses, which is based on a neural network (NN) model algorithm. The proposed method's outcomes are contrasted with the Support Vector Regression (SVR) model.

The evaluation of the neural network trained using the NN model technique for identifying diseases in apple trees is shown in Table I. We gathered and compared the experimental findings for the SVR model and the proposed method. The dataset utilized to test the models included 640 data samples, as shown in Table II. 320 of these samples contained Alternaria disease, 184 contained apple black spot disease, and 136 contained Minoz blight infections.

The effectiveness of each model in diagnosing the various illness samples is detailed in Table III. For instance, the neural network model correctly identified Alternaria disease in 316 out of 320 samples but mistakenly identified black spot disease in four samples. The proposed model outperformed other models in terms of the CC index, which is shown in Table III as having a value of 0.976 for the test data. The proposed model's RMSE index for the test data is also 0.098, which is lower than the results produced by previous models.

Finally, based on these findings, we observed that the model performs better than previous models in identifying the specific apple tree leaf disease type. The proposed method outperforms previous models in detecting and classifying apple tree illnesses, according to the high CC index and low RMSE index.

V. CONCLUSION

The paper presents a novel method for detecting three apple tree leaf diseases: apple black spot, Alternaria, and Minoz blight. The method consists of two main components: digital image processing and neural network (NN) techniques. The digital image processing component is responsible for preparing, processing, and extracting features from the leaf images. The neural network component is responsible for classifying the diseases based on the features. The paper describes the details of each component and evaluates the performance of the method using a dataset of 300 leaf images. The paper also compares the method with other existing methods and shows that the proposed method achieves higher accuracy and performance in identifying the apple tree leaf diseases. The paper concludes by suggesting some future directions for improving the method, such as extending it to a deep learning-based model and exploring various convolutional neural networks (CNN) for better feature extraction and classification.

ACKNOWLEDGMENT

This research was supported by the Natural Science Foundation of Hunan Province, Systematical Development and Application of Intelligent Greenhouse based on the Internet of Things (2022JJ50090).

REFERENCES

[1] A. A. Bharate and M. S. Shirdhonkar, "A review on plant disease detection using image processing," 2017: IEEE, pp. 103-109.

[2] V. Binnar and S. Sharma, "Plant Leaf Diseases Detection Using Deep Learning Algorithms," 2023: Springer, pp. 217-228.

[3] Z. Iqbal, M. A. Khan, M. Sharif, J. H. Shah, M. H. ur Rehman, and K. Javed, "An automated detection and classification of citrus plant diseases using image processing techniques: A review," Computers and electronics in agriculture, vol. 153, pp. 12-32, 2018.

[4] P. Kartikeyan and G. Shrivastava, "Review on emerging trends in detection of plant diseases using image processing with machine learning," International Journal of Computer Applications, vol. 975, no. 8887, 2021.

[5] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease recognition–A review," Information processing in agriculture, vol. 8, no. 1, pp. 27-51, 2021.

[6] X. Chao, G. Sun, H. Zhao, M. Li, and D. He, "Identification of apple tree leaf diseases based on deep learning models," Symmetry, vol. 12, no. 7, p. 1065, 2020.

[7] A. Rastogi, R. Arora, and S. Sharma, "Leaf disease detection and grading using computer vision technology & fuzzy logic," in 2015 2nd international conference on signal processing and integrated networks (SPIN), 2015: IEEE, pp. 500-505.

[8] H. Azgomi, F. R. Haredasht, and M. R. S. Motlagh, "Diagnosis of some apple fruit diseases by using image processing and artificial neural network," Food Control, vol. 145, p. 109484, 2023.

[9] E. Omrani, B. Khoshnevisan, S. Shamshirband, H. Saboohi, N. B. Anuar, and M. H. N. M. Nasir, "Potential of radial basis function-based support vector regression for apple disease detection," Measurement, vol. 55, pp. 512-519, 2014.

[10] S. Zhang, X. Wu, Z. You, and L. Zhang, "Leaf image based cucumber disease recognition using sparse representation classification," Computers and electronics in agriculture, vol. 134, pp. 135-141, 2017.

[11] A. Loutfi, S. Coradeschi, G. K. Mani, P. Shankar, and J. B. B. Rayappan, "Electronic noses for food quality: A review," Journal of Food Engineering, vol. 144, pp. 103-111, 2015.

[12] A. Tomar, G. Gupta, W. Salehi, C. H. Vanipriya, N. Kumar, and B. Sharma, "A Review on Leaf-Based Plant Disease Detection Systems Using Machine Learning," Recent Innovations in Computing: Proceedings of ICRIC 2021, Volume 1, pp. 297-303, 2022.

[13] Z. Chuanlei, Z. Shanwen, Y. Jucheng, S. Yancui, and C. Jia, "Apple leaf disease identification using genetic algorithm and correlation based feature selection method," International Journal of Agricultural and Biological Engineering, vol. 10, no. 2, pp. 74-83, 2017.

[14] A. Rastogi, R. Arora, and S. Sharma, "Leaf disease detection and grading using computer vision technology & fuzzy logic," 2015: IEEE, pp. 500-505.

[15] K. R. Gavhale and U. Gawande, "An overview of the research on plant leaves disease detection using image processing techniques," Iosr journal of computer engineering (iosr-jce), vol. 16, no. 1, pp. 10-16, 2014.

[16] S. S. Harakannanavar, J. M. Rudagi, V. I. Puranikmath, A. Siddiqua, and R. Pramodhini, "Plant leaf disease detection using computer vision and machine learning algorithms," Global Transitions Proceedings, vol. 3, no. 1, pp. 305-310, 2022.

[17] N. Ganatra and A. Patel, "A multiclass plant leaf disease detection using image processing and machine learning techniques," International Journal on Emerging Technologies, vol. 11, no. 2, pp. 1082-1086, 2020.

[18] M. Prabu and B. J. Chelliah, "An intelligent approach using boosted support vector machine based arithmetic optimization algorithm for accurate detection of plant leaf disease," Pattern Analysis and Applications, vol. 26, no. 1, pp. 367-379, 2023.

[19] Z. Ghasemi Varjani, S. S. Mohtasebi, H. Ghasemi, and E. Omrani, "Developing a new hybrid system for detection of apple tree leaves diseases," Iranian Journal of Biosystems Engineering, vol. 49, no. 2, pp. 215-225, 2018.

[20] M. Ang, E. Sundararajan, K. Ng, A. Aghamohammadi, and T. Lim, "Investigation of Threading Building Blocks Framework on Real Time Visual Object Tracking Algorithm," Applied Mechanics and Materials, vol. 666, pp. 240-244, 2014.