Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 65 (2015) 691 – 700

Procedia Computer Science

International Conference on Communication, Management and Information Technology (ICCMIT 2015)

An Assistive Object Recognition System for Enhancing Seniors Quality of Life

Esraa Elhariri[a,e,*], Nashwa El-Bendary[b,e], Aboul Ella Hassanien[c,d,e], and Vaclav Snasel[f]

[a] Faculty of Computers & Information, Fayoum University, Fayoum, 63511 ,Egypt
[b] Arab Academy for Science, Technology, and Maritime Transport, Dokki, Giza,12311, Egypt
[c] Faculty of Computers & Information, Cairo University, 5 Ahmed Zewal St., Orman, Giza, Cairo, Egypt
[d] Faculty of Computers & Information, Beni Suef University, Beni Suef, Egypt
[e] Scientific Research Group in Egypt(SRGE), http://www.egyptscience.net
[f] Faculty of Electrical Engineering and Computer Science, VÅB-TU of Ostrava, Czech Republic

Abstract

This paper presents an indoor object recognition system based on the histogram of oriented gradient and Machine Learning (ML) algorithms; such as Support Vector Machines (SVMs), Random Forests (RF) and Linear Discriminant Analysis (LDA) algorithms, for classifying different indoor objects to improve quality of elderly people's life. The proposed approach consists of three phases; namely segmentation, feature extraction, and classification phases. Datasets used for these experiments, are totally consisted of 347 images with different eight indoor objects used for both training and testing datasets. Training dataset is divided into eight classes representing the different eight indoor objects. Experimental results showed that RF classification algorithm outperformed both SVM and LDA algorithms, where RF achieved 80.12 %, SVM and LDA achieved 77.81% and 78.76% respectively.

Peer-review under responsibility of Universal Society for Applied Research

Keywords: features extraction; Linear Discriminant Analysis (LDA); object recognition; support vector machines (SVMs); random forests (RF).

Peer-review under responsibility of Universal Society for Applied Research
doi:10.1016/j.procs.2015.09.013

* Corresponding author. Esraa Elhariri
E-mail address: eng.esraa.elhariri@gmail.com

1. Introduction

In 2010, Elderly who aged 65 or older were 524 million approximately of the world's population and the number of elderly rapidly grow. So, by 2050, it is expected to be nearly triple to about 1.5 billion world's population [1]. As a result of the growth in the elderly population, the attention to their complex needs will be increased. The elderly suffer from different problems and diseases such as, dementia, memory difficulty and low vision. Most of the elderly are not able to manage daily life by themselves due to these problems.

For Dementia, it isn't a specific disease. It is a degenerative condition and has no known cure. Symptoms of dementia disease are memory loss, cognitive weakness, communication difficulty, mood disturbances, executive dysfunction, aphasia, agnosia, apraxia, visuospatial impairment, behavioral and activity, these symptoms get worse over time an affect people's daily life and activities [2–4]. Also, for visually impaired, low vision people (by birth or by an accident or due to old age), they suffer from poor vision, sometimes they distinguish differ between objects [5]. Elderly people also suffer from memory difficulty; as they know how to use objects, but, they don't remember their names.

These situations are depressing for the individual and irritating for their loved ones. Individuals, who suffer from these problems, are worrying about their life as well as privacy as they can't act independently. Although there is no cure for these problems, they can still have a good quality of life, if the right long-term care plan was provided in a place. Since the cost of caring for elderly is very steep [4], so an indoor object detection and recognition system for elderly, visually impaired and low vision people, is required to support these people to act independently. Object detection and recognition is one of the most challenging problems in computer vision. Utilizing computer vision with machine learning techniques has recently been widely used in developing real world object recognition system based on the use of local features [6].

This paper is a part of the framework for designing an intelligent assistive system for indoor object detection and recognition for elderly people. The proposed system depends on computer vision and machine learning techniques. The dataset used in this paper, is downloaded from UMass Lowell University website. The rest of this paper is organized as follows. Section 2 introduces related research work. Section 3 describes the different phases of the proposed system. Section 4 discusses the tested image dataset and presented the obtained experimental results. Finally, section 5 presents conclusions and future work.

Nomenclature

ML	Machine Learning
SVMs	Support Vector Machines
RFs	Random Forests
HOGs	Histogram of Gradients
LDA	Linear Discriminant Analysis
PCA	Principal Component Analysis
SIFT	Scale Invariant Feature Transform

2. Related Work

As previously stated in the introduction section, object recognition is one of the most challenging problems in computer vision. In the last few years, there has been substantial work in the computer vision field, which tackling the problem of object recognition. This section gives a brief survey of recent work on object recognition.

In [7], an object categorization approach based on a combination of efficient image representation and support vector machines (SVMs) has been presented. Image representation based on local descriptor such as (SIFT, JET and image patch). Firstly, they compute different local descriptor feature vectors, then they test their approach using different combinations of SVM kernel and local descriptors. Finally, they compare the performance and SIFT descriptor was found to be the best one.

While in [8], an image recognition method for Smartphone based on local features and SVM has been proposed. The local features used in this research are HOG and Color patch descriptor. Firstly, they extract both HOG and Color patch, and they reduce their dimensionality using PCA. Then, they are Fisher Vector represented. Finally, they applied SVM as a classification algorithm. This system has achieved 79.2% classification rate with the top five candidates.

Also in [9], a face recognition approach based on SIFT features has been proposed. The approach works as follows: SIFT features are extracted from all faces images. Then, SIFT features are extracted from a given test face image. Finally, the extracted features from the test face are compared against the features from each face. Face image with the largest number of matching points, which is greater than or equal to specific threshold, is the nearest face to test one. This approach was compared with two different approaches Eigenfaces and Fisherfaces. The obtained results show the excellence of this approach specially using smaller training sets.

In [10], authors proposed a robust object recognition system which, based on a new set of features. The used features are a set of scale and translation invariant C2 features. Firstly, they compute C2 features, and then they apply linear SVM as a classifier algorithm. This system exhibits excellent performance on a variety of image datasets.

This paper presents an intelligent assistive system for indoor objects recognition based on the histogram of oriented gradient in addition to SVMs, RF and LDA Machine Learning algorithms, to improve seniors quality of life. The objects dataset used in this paper has been downloaded from UMass Lowell University website.

3. The Proposed Indoor Object Recognition System

In particular, the proposed framework is capable of recognizing different indoor objects. The proposed system exploits HOG descriptor to identify different indoor objects in a multi-class scenario. The steps of it are shown at Fig. 1.

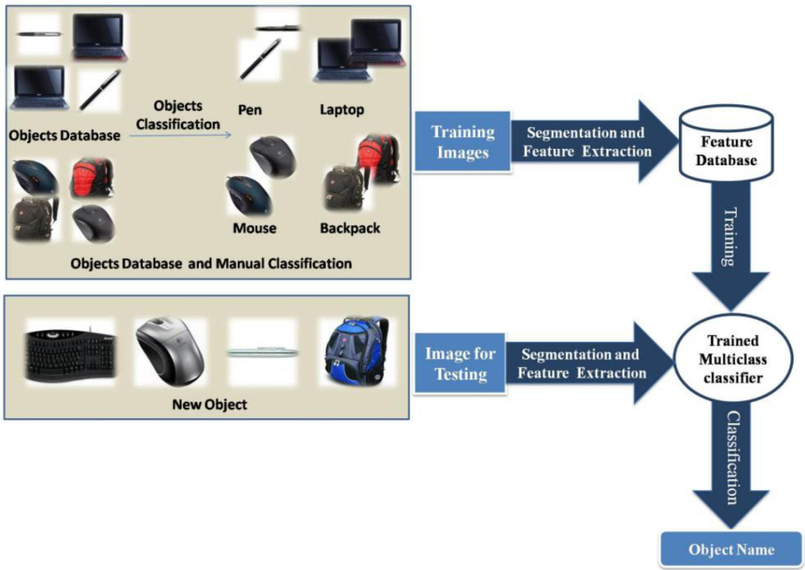

Fig. 1. Architecture of the proposed recognition system.

Segmentation, feature extraction, training and classification are the major tasks to be performed. For the indoor objects recognition problem, a good segmentation is required to extract a region of interest to achieve more accurate results. An image segmentation based on edge detection and morphological operations is used to detect only object part (region of interest).

3.1. Segmentation

Image segmentation is an appropriate and efficacious method for detecting foreground objects in images with stable background. There are many methods for image segmentation. In our proposed system, that used images objects have sufficient contrast from the background, so objects can be easily detected in an image. Here, an image segmentation approach based on edge detection and basic morphological operations is used to detect our objects. Fig.2. shows some segmentation examples.

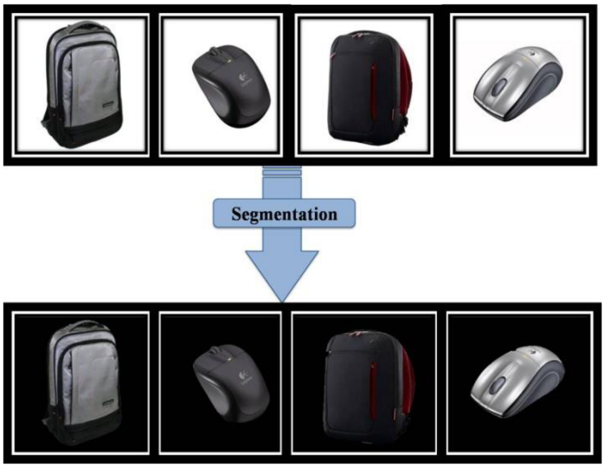

Fig. 2. Sample images before and after applying segmentation step.

The segmentation approach works as shown at Algorithm (1):

Algorithm 1 Segmentation Algorithm

1: Convert image $Img_{original}$ from RGB color space into grayscale G_{img}

2: Compute object edges via Apply Sobel edge detection filter with a defined threshold T on G_{img}.

3: Apply image dilation using linear structuring elements(vertical one followed by horizontal one) SE

4: Perform holes filling to fill holes in the interior of the object

5: Smooth the object by eroding the image twice with a diamond structuring element SE, Now we have object binary mask $Object_{mask}$

6: Multiply $Img_{original}$ by $Object_{mask}$

7: Output is segmented Object Image

3.2. *Feature Extraction*

In this paper, HOGs are used as local descriptors. They are commonly used in computer vision and image processing for the purpose of object detection. The technique counts occurrences of gradient orientation in localized portions of an image. This method is similar to other features such as edge orientation histograms, scale invariant feature transform (SIFT), SURF descriptors, and shape contexts. The difference is that is computed on a dense grid of uniformly spaced cells and it is very simple and fast to extract [11, 12].

HOG computation algorithm can be performed by using Algorithm (2) [12]:

Algorithm 2 HOG Algorithm

1: Compute gradient values by applying the following process:
- first the grayscale image is filtered to obtain x and y derivatives of pixels using those masks:

$$D_y = \begin{vmatrix} 1 \\ 0 \\ -1 \end{vmatrix}, \quad D_x = \begin{vmatrix} -1 & 0 & 1 \end{vmatrix}$$

- compute magnitude and orientation of the gradient using the following equations:

$$|G| = \sqrt{{I_x}^2 + {I_y}^2} \quad (1)$$

$$\theta = \arctan \frac{I_y}{I_x} \quad (2)$$

2: Divide the image into 2×2 block with total of 4 blocks.
3: Extract HOG regarding eight orientations for each block and perform block normalization.
4: Combine blocks HOGs into 1D feature vector with length 32.
5: Normalize the final HOG vector.

3.3. Classification

The final step in object recognition using Histogram of Oriented Gradient descriptors is to feed the descriptors into some recognition system based on supervised learning. For classification phase, the proposed approach applied three different algorithms for classification of Indoor Objects SVM, RF and LDA. In this section, we will discuss the basics of SVM, RF and LDA.

1) Support Vector Machines (SVMs): One of the machine learning (ML) algorithms that are used for classification and regression of high dimensional datasets with excellent results is Support Vector Machines (SVMs) [13–15]. For classification problem solving, SVM is trying to find an optimal separating hyperplane between classes.

It depends on support vectors, these are training cases which are placed on the edge of class descriptor, any other cases are discarded [14–16]. SVM algorithm seeks to maximize the margin around a hyperplane that separates a positive class from a negative class [13–15].

Given a training dataset with n samples $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$, where x_i is a feature vector in a v-dimensional feature space and with labels $x_i \in -1, 1$ belonging to either of two linearly separable classes C_1 and C_2. Geometrically, the SVM modeling algorithm finds an optimal hyperplane with the maximal margin to separate two classes, which requires to solve the optimization problem, as shown in equations (3) and (4).

$$\text{maximize} \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{n} \alpha_i \alpha_j y_i y_j . K(x_i, x_j) \quad (3)$$

$$\text{Subject} - \text{to}: \sum_{i=1}^{n} \alpha_i y_i, 0 \leq \alpha_i \leq C \quad (4)$$

Where, α_i is the weight assigned to the training sample x_i . If $\alpha_i > 0$, x_i is called a support vector. C is a regulation parameter used to trade-off the training accuracy and the model complexity so that a superior generalization capability can be achieved. K is a kernel function, which is used to measure the similarity between two samples.

2) Random Forests (RF): One of the best known classification and regression techniques, which has the ability to classify large datasets with great accuracy is Random Forests (RF) algorithm. It generates an ensemble of decision trees. The main principle of ensemble methods is to group weak learners together to build a strong learner [17], [18]. The input is entered at the top and as it traverses down the tree, the original data is sampled in random, but with replacement into smaller and smaller sets. Sample class is determined using random forests trees, which are of an arbitrary number. To perform RF algorithm, Use Algorithm (3) [17]:

Algorithm 3 Random Forest Algorithm

1: Draw N_{tree} bootstrap samples from the original data.
2: For each of the bootstrap samples, grow an un-pruned classification or regression tree.
3: At each internal node, rather than choosing the best split among all predictors, randomly select m_{try} of the M predictors and determine the best split using only those predictors.
4: Save tree as is, alongside those built thus far (Do not perform cost complexity pruning).
5: Predict new data by aggregating the predictions of the N_{tree} trees.

The predictions of the Random Forests are taken to be the majority votes of the predictions of all trees for classification and for regression are taken to be the average of the predictions of the all trees as shown in equation (5) [17], [18]:

$$S = \frac{1}{K} \sum_{k=1}^{K} K^{th} \quad (5)$$

Where S is a random forests prediction, K^{th} is a tree response, and k is the index runs over the individual trees in the forest.

3) Linear Discriminant Analysis (LDA): Linear Discriminant Analysis (LDA) is a commonly used technique for data classification and dimensionality reduction. Linear Discriminant Analysis easily handles the case where the within class frequencies are unequal and their performances has been examined on randomly generated test data. it's basic idea is to find a linear transformation that best discriminate among classes, then classification can be performed in transformed space based on some metrics(Euclidean distance) [19].

Given data $y_i, x_i, i =_1^n$, where y_i 1, 2, ...,K is the class label, k is the number of classes and x_i is a vector of features or predictors, we seek to find the best direction in the predictor space in which the classes are separated as much as possible.

Mathematically, LDA implementation is carried out via scatter matrix analysis. For all samples of all classes, we define two measures [19], [20]:

• Within-class scatter matrix. It is defined by:

$$S_w = \sum_{j=1}^{K} \sum_{i=1}^{N_j} \left(x_i^j - \mu_j\right)\left(x_i^j - \mu_j\right)^T \quad (6)$$

Where x_j^i is the i^{th} sample of class j , μ_j is the mean of class j , K is the number of classes, and N_j is the number of samples in class j ; and

• Between-class scatter matrix. It defined by:

$$S_b = \sum_{j=1}^{K} (\mu_j - \mu)(\mu_j - \mu)^T \quad (7)$$

Where μ represents the mean of all classes. This method maximizes the ratio of between-class measure to the within-class measure in any particular data set thereby guaranteeing maximal separability. The maximization of $\frac{\delta|S_b|}{\delta|S_w|}$.

For SVM and LDA, The inputs are training dataset feature vectors and their corresponding classes, whereas the outputs are the names of objects in the testing dataset. For RF, The inputs are a number of trees, training dataset feature vectors and their corresponding classes, whereas the outputs are the names of objects in the testing dataset. Since SVM is a binary class classification method and our problem is an N-class classification problem, so in this research the SVM algorithm is applied to Multi-class problem [21], [22]. We used one-against-one approach to do that.

SVM was trained and tested using (Linear and Polynomial with order = 3 kernel functions) and cross-validation.

4. Experimental Results

Simulation experiments in this paper used a dataset of total 347 images for different indoor objects for both training and testing datasets with 10-fold cross-validation. Samples of some indoor objects are shown at Fig. 3.

Fig. 3. Samples of different indoor objects.

Training dataset is divided into eight classes representing the different classes of indoor objects. The classes are Mug, Pen, Monitor, Mouse, Keyboard, Backpack, Laptop, and Mobile. Some samples of both training and testing datasets Fig. 4.

The proposed approach has been implemented considering the One-against-One multi-class SVM system, LDA and RF using 10-fold cross validation and a total of 347 images of different indoor objects for both of training and testing datasets. The used features for classification are HOGs. Moreover, SVM algorithm was employed with different kernel functions that are: Linear kernel and Polynomial kernel for object classification.

Fig. 4. Examples of training and testing samples.

Figure 5 shows classification accuracy for object recognition obtained via applying each kernel function of SVM, LDA and RF.

The accuracy is computed using equation (8):

$$Accuracy = \frac{Number\ of\ correctly\ classified\ images}{Total\ number\ of\ testing\ images} * 100 \quad (8)$$

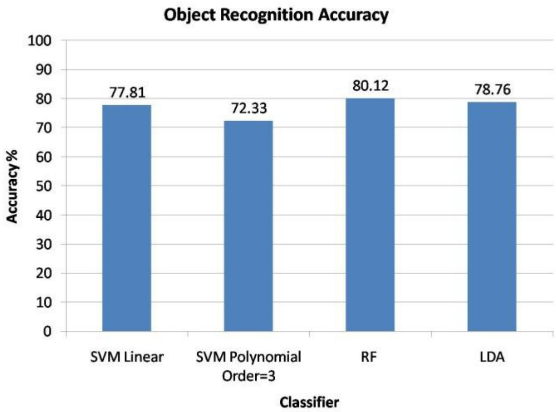

Fig. 5. Object recognition System Accuracy.

In this paper, we tested SVM, RF and LDA classifiers for the problem of object recognition. The results showed that RF classifier is better than SVM and LDA and achieved good accuracy 80.12 %.

This is because of the working method of RF described above. RF achieves high accuracy with huge amount of dataset which has high number of variables. RF has one particular strength point, it is really doing feature selection, because it will build trees on good features by chance, and favor these over trees that were built on noise features.

5. Conclusions and Future Work

In this paper, an object recognition system based on Histogram of Oriented Gradients along with Support Vector Machines (SVMs), Random Forests (RF) and Linear Discriminant Analysis (LDA) classifiers has been presented. The proposed classification approach was implemented by applying a segmentation approach based on edge detection and morphological operations. Then, feature extraction was applied to each pre-processed image, HOG for each image is obtained as a feature vector. Finally, SVM, RF and LDA models are developed for object recognition. The proposed approach has been implemented via applying One-against-One multi-class SVM system, RF and LDA system using 10-fold cross-validation.

Based on the obtained the experimental results, RF outperformed both SVM and LDA classification algorithms. Paper's weakness is the selection of objects in the dataset collection, look far from the everyday life of elderly. So, more related objects to elderly will be included in our future work.

For future work, other features as well as other machine learning classifiers could be applied in order to achieve better recognition accuracy. Also, other objects, which elderly faced it daily will be included, as the currently available datasets didn't include them .In addition to other datasets; such as faces, could be used in order to extend the applicability of the proposed system.

REFERENCES

1. World Health Organization and National Institute on aging, National Institutes of Health and U.S. Department of Health and Human Services, "Global health and aging",2011,*World Health Organization*.
2. Carrie B Peterson, Neeli R Prasad, and Ramjee Prasad, "Framework for Dementia Quality of Life Assessment with Assistive Technology Intervention*", in The proceeding of the 7th conference on Biomedical Engineering (BioMED 2010)* , Innsbruck, Austria, February 17 - 19 ,pp. 129–134,2010, IEEE.
3. Sposaro Frank, Justin Danielson, and Gary Tyson. "iWander: An Android Application for Dementia Patients", *2010 Annual International Conference of the IEEE on Engineering in Medicine and Biology Society (EMBC)*, Buenos Aires, Argentina, September 1-4, pp. 3875-3878, 2010,IEEE.
4. Prince Martin, Malenn Guerchet, and Matthew Prina, "World Alzheimer Report 2013: Journey of Caring: An Analysis of Long-Term Care for Dementia". *Alzheimer's Disease International*, 2013.
5. Y. Jayanta Singh and Shalu Gupta, "Speedy Object Detection Based on Shape", *In The International Journal of Multimedia & Its Applications (IJMA)*, vol. 5, No.3, pp. 15–23, June 2013, arxiv.org.
6. Jaruwan Toontham and Chaiyapon Thongchaisuratkrul, "The Comparison of Object Recognition and Identification by Using Image Processing Base on the Neural Network, the Hough transform and the Harris Corner Detection", *in International Journal of Machine Learning and Computing*, vol. 1, No. 4, pp. 400–404, October 2011.
7. Jan Eichhorn and Olivier Chapelle, "Object categorization with svm: kernels for local features". *In Advances in Neural Information Processing Systems (NIPS) Journal*, July 2004, Max Planck Institute for Biological Cybernetics.
8. [8 Yoshiyuki Kawano and Keiji Yanai, "Rapid Mobile Object Recognition Using Fisher Vector*", In 2nd IAPR Asian Conference on Pattern Recognition (ACPR)*, Okinawa, Japan, November 5-8, pp. 476-480, 2013, IEEE.
9. ALY, Mohamed, "Face recognition using SIFT features*", In CNS/Bi/EE report*, vol. 186, 2006.
10. [10] Serre, T., Wolf, L. and Poggio, T., "Object recognition with features inspired by visual cortex". *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR2005)*, San Diego, CA, USA, June 20-26, vol. 2, pp. 994-1000,2005, IEEE.
11. Dalal, Navneet and Bill Triggs, "Histograms of oriented gradients for human detection", *In Computer Society Conference on Computer Vision and Pattern Recognition(CVPR 2005)*, San Diego, CA, USA, June 20-26, pp. 886–893, vol. 1, 2005, IEEE.
12. Kawano, Yoshiyuki, and Keiji Yanai, "ILSVRC on a Smartphone", *In IPSJ Transactions on Computer Vision and Applications*, vol. 6, pp. 83-87, 2014, Information Processing Society of Japan.
13. Q. Wu and D.-X. Zhou, "Analysis of support vector machine classification, " *Journal of Computational Analysis and Application*., vol. 8, no. 2, pp. 99-119, 2006.
14. H. M. Zawbaa, N. El-Bendary, A. E. Hassanien, and A. Abraham, "SVM-based Soccer Video Summarization System, " *in Proc. The Third IEEE World Congress on Nature and Biologically Inspired Computing (NaBIC2011)*, Salamanca, Spain, October 19-21 , pp. 7–11, 2011, IEEE.
15. H. M. Zawbaa, N. El-Bendary, A. E. Hassanien, and T. H. Kim, "Machine Learning-Based Soccer Video Summarization System, " *in Proc. Multimedia, Computer Graphics and Broadcasting FGIT-MulGraB (2)*, Jeju Island, Korea, vol. 263, pp. 19–28, 2011, Springer.
16. S.R. Suralkar and A. H. Karode, and P. W. Pawade, "Texture Image Classification Using Support Vector Machine," *International Journal of Computer Applications in Technology*, vol. 3, no. 1, pp. 71–75, 2012.

17. Vrushali Y Kulkarni and Pradeep K Sinha, "Efficient Learning of Random Forest Classifier using Disjoint Partitioning Approach", *in Proc. of the World Congress on Engineering*, London, U.K, July 3-5, vol. 2, 2013.
18. Anna Bosch, Andrew Zisserman and Xavier Munoz, "Image Classification using Random Forests and Ferns," *in IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, Brazil, October 14-20, pp. 1–8, 2007, IEEE.
19. Li, Tao, Shenghuo Zhu, and Mitsunori Ogihara, "Using discriminant analysis for multi-class classification: an experimental investigation", *Knowledge and information systems*, vol. 10, no. 4, pp. 453-472, 2006, Springer.
20. Fisher, Ronald A, "The use of multiple measurements in taxonomic problems", *Annals of eugenics* , vol.7, no.2, pp. 179-188, 1936, Wiley Online Library.
21. Y. Liu and Y. F. Zheng, "One-against-all multi-class SVM classification using reliability measures*," in Proc. IEEE International Joint Conference on Neural Networks (IJCNN'05)*, Montreal, Quebec, Canada, vol. 2, pp. 849–854, 2005, IEEE.
22. G. Anthony, H. Gregg, and M. Tshilidzi, "Image Classification Using SVMs: One-against-One Vs One-against-All, *" in Proc. of the 28th Asian Conference on Remote Sensing*, Kuala Lumpur, Malaysia, November 12-16, 2007, Asian Association on Remote Sensing (AARS).