



An Improved Plant Identification System by Fuzzy c-means Bag of Visual Words Model and Sparse Coding

Soodabeh Safa¹, Fatimah Khalid²

¹ Center of Advanced Computing Technologies (CACT), Fakulti Teknologi Maklumat Dan Komunikasi (FTMK), Universiti Teknikal Malaysia (UTeM), Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia, soodabeh.safa@utem.edu.my

² Department of Multimedia, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, Malaysia fatimah.Kh@Upm.edu.my

ABSTRACT

Current work adopts the Fuzzy c-means Bag of Visual Words model and sparse coding for plant identification. Plant identification has become a significant research area in the botany field in recent years. SIFT features are distinctive invariant features based on scale-space because of the situation of its robust identical matching capabilities. Bag of visual words (BoVW) model and its variants are used effectively for the retrieval of images by many researchers. Classic bag of visual words algorithm is based on k-means clustering and every SIFT features belongs to one cluster and it leads to decreasing classification results. Data entities may belongs to further than one cluster in the fuzzy clustering (soft clustering), and a set of membership levels are allied with each group. This demonstrate the intensity of the correlation between that aspect of data and a specific cluster. In the classic Bag of visual words model, the Fuzzy c-means algorithm is replaced with K-means and the accuracy of SIFT matching is increased. Moreover, sparse coding has been commonly used in recent years for the purposes of retrieving and identifying images. The pure picture patch computes the atoms in an over-complete dictionary by adding them sparsely. Sparse representation prevents over-fitting in the classifier by eliminating redundancies and evaluating high-frequency patterns between feature vectors. Performance of proposed methods surpass the classic bag of words algorithm for plant identification tasks.

Key words : Plant identification, Fuzzy C-means algorithm, SIFT Descriptor, Bag of Visual Words algorithm, sparse coding.

1. INTRODUCTION

Identification and classifying different kinds of plants with taking advantage of technology ends up in large benefits in agriculture [1]. Plants identification refers to the comparison of unknown plants with a group of

recognized species to work out the actual of best-known species to determine the particular specie to which it belongs. For efficient agricultural biodiversity, detailed information on the identification and geographical distribution of plants is needed [2]. Image identification and matching are necessary to analysis in computer vision and image processing applications [3]. It is conjointly a necessary precondition for finding several sensible issues and have a distinctive perspective to contribute in serving to the globe to enhance the performance of image matching techniques, and have introduced a different range of algorithms [4]. Image matching algorithms, categorized into two types: i) global matching algorithms based on features and ii) local matching algorithms based on features. Local feature-based Image matching is more robust compared with these two matching algorithms. These have been used successfully in many real-world applications, such as object recognition, texture recognition, image retrieval, video data mining, panoramic design and object recognition. [5]. Local matching algorithms based on features require two stages: i) the detection of interest points, and ii) the description. The SIFT descriptor is presently the most commonly utilized local descriptor in computer vision applications, as the SIFT characteristics are extremely distinctive and invariant to adjustments in scale, rotation and illumination changes. SIFT descriptor yields the greatest results in object recognition and their applications and is grounded on key points extracted from the image. Alternative researchers tried to enhance the SIFT algorithmic rule. It's smart stability and unchangeableness likewise as detects local key points that contain a huge quantity of information. Owing to its distinctive benefits, it's become a preferred analysis topic. Several researchers perpetually push to enhance it. within the part of descriptor establishing, SIFT describes only local information and does not make use of global information [6], [7]. According to Mortensen [8], A global context SIFT descriptor (called GSIFT) was implemented which adds a global texture vector to the SIFT basis. In the key point detection process, SIFT uses

only the grayscale image details. Color images discarded by several color information. Abdel-Hakim [9] proposed CSIFT, which adds color invariance to the basis of SIFT and intends to overcome the disadvantage of SIFT for color images. Image classification tasks square measure the SIFT algorithm that combined with Bag-of- Visual-Words methodology. Intuitively, this Bag-of-Words algorithmic is a popular algorithmic for document classification. The document represented and classified by the set of words that they contain, and their frequencies. Hence, for the development of object recognition, this image Bag-of-Words study was conducted where image are represented by the set of features they contain. Recently, it has ' been used for the recognition of scenes [10].

2. MATERIAL AND METHODS

2.1 Data Set

The dataset Pl@ntLeaves used in this study within the CLEF image [11] contained 11572 images: 6630 scans, 2216 photographs, and 2726 scan-like photos (Figure 1). The three image varieties illustrated with an equivalent species *Celtis australis* L. the foremost necessary motivation for selecting supported this dataset of pictures contains vast variety of image rotation, scale, noise and luminance These types of plant-leaf images poses a big obstacle for researchers (plant species) to identification community to produce an appropriate automatically determine plant species technique [12].

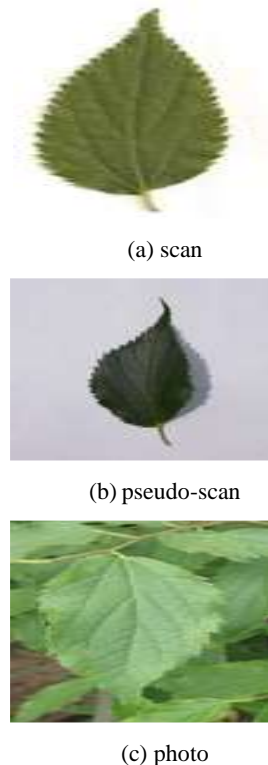


Figure 1: Three image types illustrated with the same species *Celtis australis* L, scan (a), pseudo-scan (b), and photo categories (c).

2.2 Key point reduction using dense grid SIFT

Throughout this method, key points (or points of interest) in an image are placed on a scale-invariant basis and thus the descriptors calculated at these extreme point units of measurement relative to those in another image. A high amount of matching key points indicates the similarity between two images and matching is long and recognition phases. The cluster technique is used to omitting the quantity of key points by omitting similar points and choose additional distinct Key points. At now, addressing a productive implementation methodology is required (Figure 2).



Figure 2: A gridded leaf image in the dataset.

2.3 Combining Bag of visual words with Fuzzy c-means clustering

Fuzzy c-means (FCM) is a methodology of clustering that permits one piece of information to belong to 2 or a lot of clusters and permits a homogenized grouping of categories obviously [13].

This technique developed as a common pattern recognition [14]. Step-down was supported by [15] the subsequent objective functions:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C u_{ij}^m \|x_i - c_j\|^2 \quad (3)$$

$$1 \leq m < \infty$$

Where, m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j , x_i is i th of d -dimensional measured data, c_j is the d -dimension center of the cluster, and $\|*\|$ is any norm expressing the similarity between any measured data and the center. Fig

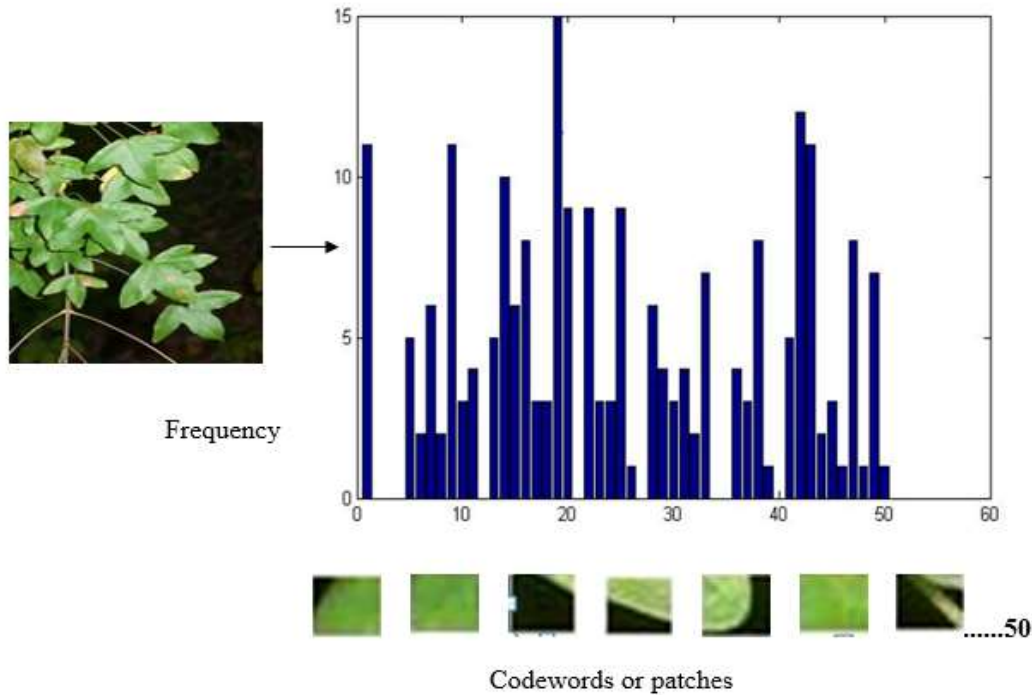


Figure 3: Example of Image and its Histogram with 50 k-means centers (code words).

According to Alitappeh, Saravi, and Mahmoudi [12] and Sukthakar [16], Fuzzy partitioning is achieved by an iterative optimisation of the target function, which is previously mentioned, with the update of membership u_{ij} and the cluster centers c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left[\frac{\|x - c_j\|}{\|x - c_k\|} \right]^{\frac{2}{m-1}}} \quad (4)$$

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m} \quad (5)$$

This iteration will stop when

$$\max_{ij} \{ |u_{ij}^{(k+1)} - u_{ij}^{(k)}| \} < \varepsilon$$

and where ε is a termination criterion between 0 and 1, whereas k is the iteration steps. This procedure converges to a local minimum or a saddle point of J_m .

The algorithm is composed of the following steps:

1. Initialize $U=[u_{ij}]$ matrix, $U^{(0)}$
2. At k -step: calculate the centers vectors $C^{(k)}=[c_j]$ with $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}$$

3. Update $U^{(k)}$, $U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left[\frac{\|x - c_j\|}{\|x - c_k\|} \right]^{\frac{2}{m-1}}}$$

4. If $\|U^{(k+1)} - U^{(k)}\| < \varepsilon$ then STOP; otherwise return to step 2

In fuzzy clustering (soft clustering), data features will belong to quite more than one cluster, and related to every feature could be a set of membership levels. These indicate the strength of the association between that data feature and a specific cluster. Fuzzy clustering could be a method of distribution these membership levels and so

using them to assign data features to at least one or additional clusters [13], [17]. Fuzzy C-means algorithm is replaced with K-means in the classic Bag of words algorithm and improves the accuracy of SIFT matching (figur4). The Bag of words algorithm is showed below:

- Extract local SIFT method vectors from any image collection;
- Place all vectors of these local features into one set. At this stage individuals don't even need to store the image from local vector that was extracted;
- Use a clustering algorithm (Fuzzy c-means) over a collection of local feature vectors to locate the centroid coordinates and assign each centroid a specific id. This set of centroid will be the vocabulary.

- The global function vector is a histogram that records how many times each centroid in each image actually occurred. Find the nearest centroid for local function vector to compute the histogram [14], [16].

The pseudo-code of Bag of visual words with Fuzzy c-means:

```

In the BoVW (calculate dictionary)
Dictionary = k means (centers, sift_ all,
options);
Replace k means with [fuzzy c-means].
    
```

The concept of fuzzy c-means based on the BoVW is depicted in Figure 4.

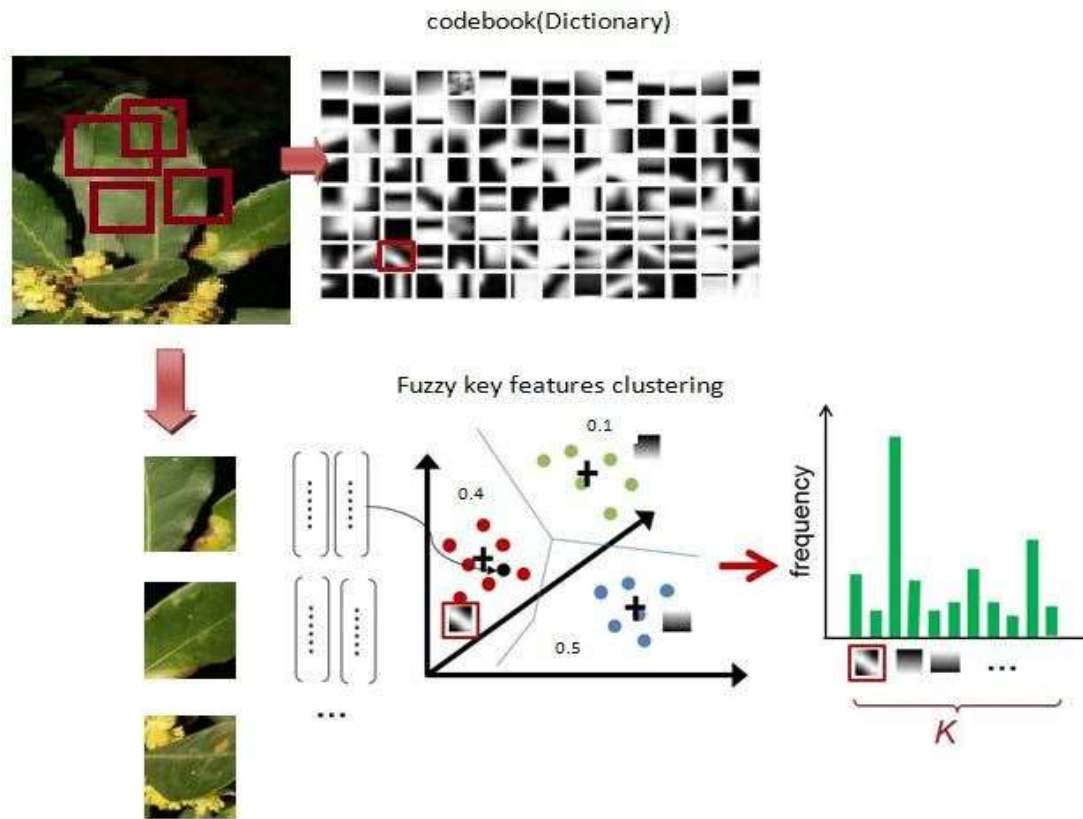


Figure 4: Fuzzy c-means based Bag of visual words model

2.4 Sparse Representation of Features

Sparsity is outlined as having few non-zero components or having few elements not near to zero. The prerequisite that coefficients α_i be sparse means given an input vector, many of the coefficients got to be as off from zero as potential. Within the approaches supported structured sparsity regularization, L1 (Norm one) and L2 (Norm 2) are used for structured dictionary learning. A matrix norm or vector norm returns the

magnitude of that matrix or vector parts. By the elimination of redundancies and therefore the analysis of high-frequency patterns between feature vectors, the dictionary learning methods for sparse representation prevent over-fitting in the classifier. There are three main methods used in dictionaries learning: employed in dictionaries learning: Figure 5 shows the block diagram of proposed sparse based image identification.

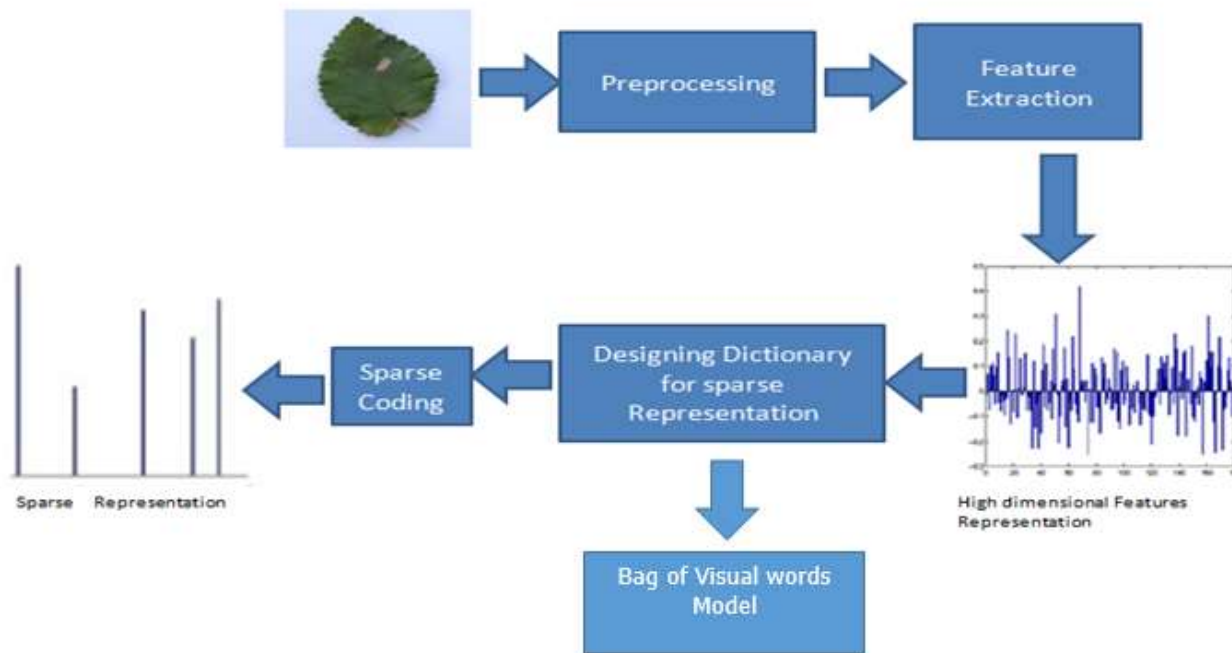


Figure 5: The block diagram of proposed sparse based image identification system

1) Probabilistic such as Method of Optimal Directions (MOD), Maximum a Posteriori (MAP), and Lewicki. 2) Clustering-based such as K-means, K-Singular Value Decomposition (K-SVD).3) Parametric methods. K-SVD is used during this work [18]. Sparse coding utilized in this work emphasizes on the patch info illustration and is completely different from the SIFT-based descriptor. SIFT computed supported the gradient, although sparse coding is approximated by the mix of various atoms that square measure during a lower mathematical space. Figure six shows the diagram of the planned sparse based plant identification. To noise removing and broken half correction of pictures, some image process steps before feature extraction is required.

3. RESULTS AND DISCUSSION

The dataset used in this study, was based on PI@ntLeaves from image CLEF [11], consists of at least 126 tree species. It contains 11572 pictures subdivided into three different kinds of pictures: scans (57%), scan-like photos (24%), and free natural photos (19%). Here is the classification result with a Random Forest classifier that is a comparatively new classifier

from the combination of Breiman’s “bagging” concept and also the random choice of features to construct a group of decision trees with controlled variation [13]. The proposed FCM and sparse method work better for scan and scan-like images rather than natural photo images, because of complex background in the natural photos, especially for needle-shaped leaves and the experiment shows impressive improvement in leaf image classification. However, more study on using specific classifiers for every image type need to recommend and, possibly, the FCM technique might a novel improvement in leaf image classification.

Table 1: Accuracy according to the image type for the classic bag of words and fuzzy c-means bag of words.

Leaf Image Type	Scan	Scan-like	Photo
Classic Bag of visual Words with 50 k-means centers (code words)	45%	40%	41%
Fuzzy C-means Bag of visual Words	72.6 %	67.9%	62.2%
Sparse Bag of visual Words	74%	69%	67.5%

According to Table 1, the proposed FCM Bag of Visual words and Sparse Bag of visual Words methods work better for scan and scan-like images rather than natural photo images, especially for needle shape leaves as will be seen in Figures 6 and 7.



Figure 6: Well identified images (scan)



Figure 7: Images with a lower rate of identification (scan like)

4. CONCLUSION

In this study, Fuzzy c-means algorithm with Bag of visual words model is combined subjected to plant leave images species classification. Moreover, sparse coding is applied for the SIFT descriptor, and instead of key points, sparse atoms considered as image feature using by Bag of Visual words models. Results suggested that proposed methods outperform the classic Bag of visual words.

5. FUTURE SCOPE

Future work ought to explore the possibility of making new models of plant identification with alternative options and descriptors, like D-Nets and U-Nets. what is more exploitation alternative strategies of distributed illustration and wordbook learning strategies besides combining some ancient classification algorithms with evolutionary algorithms like intelligent watermarking (IW) and PSO are compared to the present work for plant identification problems.

ACKNOWLEDGEMENT

The authors would like to thank the Fakulti Teknologi Maklumat dan Komunikasi (FTMK), Universiti Teknikal Malaysia Melaka (UTeM), Centre of Advanced Computing Technologies (CACT) for supporting this research.

REFERENCES

- [1] N. Nasharuddin, N. Yusoff, and S. A.-I. J. of A. T. in, “**An Multi-feature Vegetable Recognition using Machine Learning Approach on Leaf Images System.**”
- [2] J. Wäldchen and P. Mäder, “**Plant Species Identification Using Computer Vision Techniques: A Systematic Literature Review,**” *Arch. Comput. Methods Eng.*, 2018, doi: 10.1007/s11831-016-9206-z.
- [3] J. Wu, Z. Cui, V. S. Sheng, P. Zhao, D. Su, and S. Gong, “**A comparative study of SIFT and its variants,**” *Meas. Sci. Rev.*, 2013, doi: 10.2478/msr-2013-0021.
- [4] K. Abratkiewicz, D. Gromek, P. Samczynski, J. Markiewicz, and W. Ostrowski, “**The concept of applying a SIFT algorithm and orthophotomaps in SAR-based augmented integrity navigation systems,**” 2019, doi: 10.23919/IRS.2019.8768136.
- [5] R. Bandara, L. Ranathunga, and N. A. Abdullah, “Ravimal Bandara, Lochandaka Ranathunga, Nor Aniza Abdullah (2019) **Nature Inspired Dimensional Reduction Technique for Fast and Invariant Visual Feature Extraction,**” *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, no. 3, pp. 696–706, Jul. 2019, doi: 10.30534/ijatcse/2019/57832019.
- [6] A. Raza, T. Nawaz, H. Dawood, H. D.-M. T. and, and undefined 2019, “**Square texton histogram features for image retrieval,**” *Springer*, Accessed: Aug. 28, 2020. [Online]. Available: <https://link.springer.com/content/pdf/10.1007/s11042-018-5795-x.pdf>.
- [7] F. Guo, J. Yang, Y. Chen, B. Y.-2018 3rd International, and undefined 2018, “**Research on image detection and matching based on SIFT features,**” *ieeexplore.ieee.org*, Accessed: Aug. 28, 2020. [Online]. Available: <https://ieeexplore.ieee.org/abstract/document/8376448/>.
- [8] E. N. Mortensen, H. Deng, and L. Shapiro, “**A SIFT descriptor with global context,**” 2005, doi: 10.1109/CVPR.2005.45.
- [9] A. E. Abdel-Hakim and A. A. Farag, “**CSIFT: A SIFT descriptor with color invariant characteristics,**” 2006, doi: 10.1109/CVPR.2006.95.
- [10] T. Botterill, S. Mills, and R. Green, “**Speeded-up bag-of-words algorithm for robot localisation through scene recognition,**” 2008, doi: 10.1109/IVCNZ.2008.4762067.
- [11] H. Goëau *et al.*, “**The ImageCLEF plant identification task 2013,**” 2013, doi: 10.1145/2509896.2509902.
- [12] R. J. Alitappeh, K. J. Saravi, and F. Mahmoudi, “**Key point reduction in SIFT descriptor used by subtractive clustering,**” 2012, doi: 10.1109/ISSPA.2012.6310683.
- [13] K. S. Chuang, H. L. Tzeng, S. Chen, J. Wu, and T. J. Chen, “**Fuzzy c-means clustering with spatial**

- information for image segmentation,”** *Comput. Med. Imaging Graph.*, 2006, doi: 10.1016/j.compmedimag.2005.10.001.
- [14] L. Ledwich and S. Williams, “**Reduced SIFT features for image retrieval and indoor localisation,**” *Aust. Conf. Robot. Autom.*, 2004.
- [15] S. Lazebnik, C. Schmid, and J. Ponce, “**A sparse texture representation using local affine regions,**” *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2005, doi: 10.1109/TPAMI.2005.151.
- [16] Y. Ke and R. Sukthankar, “**PCA-SIFT: A more distinctive representation for local image descriptors,**” 2004, doi: 10.1109/cvpr.2004.1315206.
- [17] A. Hassan, W. Shah, A. Husein, M. Talib, and A. Mohammed, “**Clustering Approach in Wireless Sensor Networks based on K-means: limitations and recommendations 120.**”
- [18] J. Liu, W. Liu, S. Ma, M. Wang, L. Li, and G. Chen, “**Image-set based face recognition using K-SVD dictionary learning,**” *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 5, pp. 1051–1064, May 2019, doi: 10.1007/s13042-017-0782-5.