http://jecei.srttu.edu

SRTTU

Journal of Electrical and Computer Engineering Innovations

JECEI, Vol. 4, No. 1, 2016

Regular Paper

Fuzzy Neighbor Voting for Automatic Image Annotation

Vafa Maihami[1], and Farzin Yaghmaee[1,*]

[1]Department of Electrical and Computer Engineering, Semnan University, Semnan, Iran

*Corresponding Author's Information: f_yaghmaee@semnan.ac.ir

ARTICLE INFO

ARTICLE HISTORY:
Received 01 January 2016
Revised 12 June 2016
Accepted 13 June 2016

KEYWORDS:
Automatic image
Annotation
Fuzzy system
Image retrieval
Feature extraction

ABSTRACT

With quick development of digital images and the availability of imaging tools, massive amounts of images are created. Therefore, efficient management and suitable retrieval, especially by computers, is one of the most challenging fields in image processing. Automatic image annotation (AIA) or refers to attaching words, keywords or comments to an image or to a selected part of it. In this paper, we propose a novel image annotation algorithm based on neighbor voting which uses fuzzy system. The performance of the model depends on selecting the right neighbors and a fuzzy system with the right combination of features it offers. Experimental results on Corel5k and IAPR TC12 benchmark annotated datasets, demonstrate that using the proposed method leads to good performance.

1. Introduction

Nowadays, due to quick development in digital images and using different imaging tools such as mobile, camera etc., huge numbers of images are available and popularity of big visual data in many computer vision and multimedia based applications has increased. Therefore, automatic image retrieval is one of the interesting and challenging fields in the image processing. In the text-based image retrieval, the words must be associated with each single image. This means that we need to annotate images. Image annotation or tagging deals with attaching words, keywords or comments to an image or to a selected part of an image [1], [2]. The aim of image annotation is to assign the appropriate words in order to describe images as it is shown in Fig. 1 [41]. The use of images in human communication is hard and the process of digitization does not make image collections to be managed easily. Image annotation can be done by machine or humans but performing this action by humans is boring, costly, and associated with the error or ambiguity. In fact, image annotation is an automatic action which is performed by machines. Images have low level features (color, texture, shape...), while high level features are relative words to image. Image annotation can remove this gap between low level and high level features [1], [3], [4].

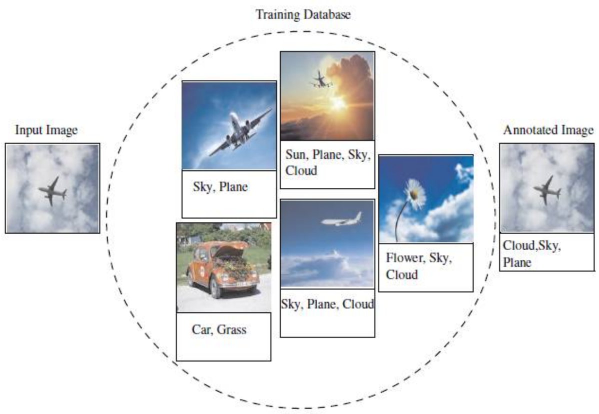

Figure 1: An example for image annotation [41]

Most of the existing methods of automatic image annotation use the visual content and they often rely heavily on supervised machine learning methods [5]–[10]. But in modern applications, especially real-time applications with a large and diverse visual content, a weak supervision which effectively and efficiently estimates tag/keywords relevance is needed. This paradigm is recently developed in references [11]–[13].

Generally, same tags/keywords are used to describe similar images. From the fact, a method called neighbor voting is proposed to annotate images [11]. Neighbor voting is a popular technique to estimate the relevance of image keywords/tags.

Neighbor voting comes with a simple yet effective mathematical model and comes with a low computational complexity. Also, this method has no limit on the number of images and the number of tags. These advantages have recently attracted substantial research attention [11], [13].

The effectiveness of neighbor voting depends on the neighbors used [11]. Identification of neighboring images is done based on low-level visual features, e.g., color, texture and local descriptors. Accordingly, to better identify the neighboring images, this paper propose an algorithm which uses fuzzy systems. Fuzzy logic for the first time was presented by Dr. Zadeh in 1965 [14]. After that fuzzy logic was used in different academic and industrial subjects [15]. Fuzzy systems are especially used in image processing and machine vision applications [15], [16]. We propose a novel neighbor voting algorithm for automatic image annotation based on fuzzy logic. As system input, given a query image annotated with keywords/tags, neighbor voting estimates the relevance of the keywords/tags with respect to the content of the query image by accumulating votes from annotated visual neighbors with the keywords/tags under consideration. In the fuzzy section, similarity scores are the fuzzy system inputs. System outputs are the best similarity score from combining features. Then, keywords/tags of *k* with most similar images are extracted and *w* keywords/tags with most repetitions in these *k* image are assigned to query image.

The objective of our work is to design a novel algorithm for image annotation based on neighbor voting and with the help of fuzzy systems. The rest of this paper is organized as follows: in Section II, the fuzzy logic will be explained briefly, and then in section III the designed system is completely presented in details. Experimental results of the proposed method is explained in Section IV, and finally conclusion is presented in Section V.

2. BACKGROUND

This section reviews the most important conception which are influenced in automatic image annotation.

A. Related Works

The overall image annotation approaches usually can be done either manually or automatically. The manual approach has high accuracy. Manual annotation process is expensive, time- consuming and fully depends on personality and knowledge about the subject. In automatic image annotation (AIA), machines are used instead of humans. Accuracy in this approach is lower than humans, but speed is very high and its cost is low, too. Image annotation in automatic/semi-automatic approach is divided into text-based methods and image-based methods. In Text-based methods, image annotation is based on some texts which are near to the image (like images caption in a web page). This approach cannot use the image content and is usually divided in two categories: ontology- based [17] and field- based [18].

On the other hand, Image-based methods use low level features such as color, texture, and shape to obtain annotation. Image-based methods are usually classified in two approaches: Global feature-based methods (features extraction and annotation by using single feature vector per image), Regional feature-based methods} (segmentation process and feature extraction applied for every segment, annotation using a set of feature vectors per image - single feature vector per each segment).

The approaches of AIA can be classified into three following models [2], [19], [22]:

Classification-based approaches: In this approach, various classifiers are used to divide the image into different categories. Based on these categories, each category is determined by appropriate keywords (labels). References [20], [21] are two examples of this category.

Probabilistic-based approaches: In this approach, the basis is visual feature extraction and image Low-level features (color, texture, shape, etc.) from the training data and then, checking the probability distribution between the visual features of images and high-level semantic features. Finally, results from probability distribution are used for annotating a of query (new) images. References [6], [9] are two examples of this category.

Nearest-neighbor-based approaches: This method, solves image annotation as a retrieval problem. Nearest neighbors are determined by the average of several distances such as fuzzy measures [23], histogram intersection [24] and Euclidean distance [25]. Keywords are then transferred from neighbors to the given image.

A comparative study of the aforementioned schemes has been done recently in [1]. In this paper, the focus is mainly on the two major aspects of automatic image annotation, feature extraction and semantic learning/annotation [1]. A recent survey shows that tags propagated from visual neighbors are as effective as tags predicted by some heavily trained models for describing image content [43].

We review baseline neighbor voting in the next section. Good studies have been done on improving the method of neighbor voting. The authors of [26] proposed novel automatic image tagging method aims at automatically discovering more complete tags associated with information importance for test images. First, near-duplicate clusters are discovered. Then, the candidate tag set obtained clusters is expanded by considering the implicit multi-tag associations mined from all the clusters' documents. A visual relevance score is also computed for each candidate tag to the test image to reduce noisy tags. In [13] in order to improve the effectiveness of the tag relevance learning technique, visually weighted neighbor voting has been proposed. The authors enhanced the effectiveness of neighbor voting by assigning a weight to each vote that is proportional to the visual similarity between the seed image and the neighbor casting the vote. The authors of [37] proposed an automatic image annotation model fuzzy and decision tree based. This approach firstly obtains association rules which represent the correlations between image features and high-level semantic concepts of training images. Then by adding the decision tree, the unnecessary rules is reduced. In [39] nearest neighbors by the average of several distances [called joint equal contribution (JEC)] are determined from different visual features. Johnson et al [42] uses image metadata to generate neighborhoods of related images, and then uses a deep neural network to blend visual information from the image and its neighbors. The authors of [36] proposed an automatic image annotation model based on the optimization of classes' scores using particle swarm optimization (PSO). In addition, random forest classifier and normalized cuts algorithm have been applied for automatic image classification, annotation, and clustering.

B. Baseline neighbor voting

In this section, are described the basic ideas behind baseline neighbor voting [11].

Given an image dataset φ, baseline neighbor voting estimates the relevance of a tag *w* with respect to the content of an image *I* as the difference between 'the number of images annotated with *w* in a set of *k* neighbors of *I* retrieved from φ by means of visual search and the number of images annotated with w in a set of *k* neighbors of *I* retrieved from φ by means of random sampling. Summary of the procedure for learning tag relevance by neighbor voting is shown in Algorithm 1 [11].

$$Prior(w,k) \approx k.\frac{L_w}{\varphi} \tag{1}$$

where *k* is the number of visual neighbors, L_w the number of images labeled with *w*, and φ is the size of the image dataset.

Algorithm 1 Baseline neighbor voting algorithm

```
Input: A user-tagged image I.
Output: tagRelevence(w,I,k) , i.e., the tag relevance value
  of each w tag in I.
Find k nearest visual neighbors of I from the collection
  with the unique-user constraint, i.e., a user has at most
  one image in the neighbor set.
for tag w in tags of I do
  tagRelevence(w,I,k)=0
end for
for image J in the neighbor set of I do
  for tag w in (tags of J ∩ tags of I) do
    tagRelevence(w,I,k)=tagRelevence(w,I,k)+ 1
  end for
end for
tagRelevence(w,I,k)=tagRelevence(w,I,k)- Prior(w,k)
tagRelevence(w,I,k)=max(tagRelevence(w,I,k),1)
```

C. Fuzzy system

In classical set theory, membership in a set is defined as either non-complete (=0) or complete (=1). In fuzzy set theory, membership in a set ranges from non-complete (=0) to complete (=1) [14]. A fuzzy logic system consists of three main parts: fuzzier, rules and inference engine, and defuzzier. These components and the general architecture of a fuzzy logic system is shown in Fig. 2.

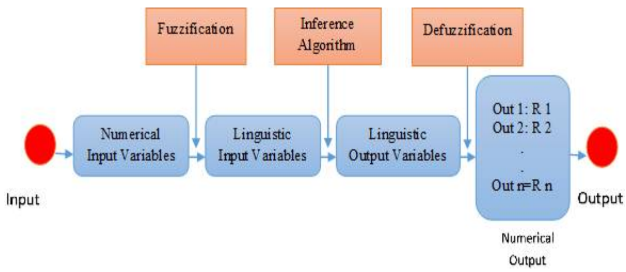

Figure 2: A Fuzzy Logic System [40].

In the process of fuzzy logic, firstly, a crisp set of input data is gathered and converted to a fuzzy set using fuzzy linguistic variables, fuzzy linguistic terms and membership functions. This step is known as fuzzification. Afterwards, an inference is made based on a set of rules. The linguistic values assigned to the input variables in the previous step fire some rules and output variables are calculated. A fuzzy rule is a simple IF-THEN rule with a condition and a conclusion. After the inference step, the overall result is a fuzzy value. Lastly, the resulting fuzzy output is mapped to a crisp output using the membership functions, in the defuzzfication step [40].

Fuzzy theory from the time of presenting up to now has been known as a method that can act with the vague and unclear data well. This theory has the ability of working on the plans that their data are not defined clearly. This matter has many uses in the management controlling systems that act on vague data. An expert fuzzy system provides relations between input spaces nonlinear data and the output spaces by means of classical outputs and inputs. This relation is done by some system IF-THEN rules. In a universe of discourse U, a fuzzy subset A of U is defined by a membership function μ A(x) which maps each element x in U to a real number in the interval [0 1]. The function value μ A(x) denotes the grade of membership of x in A. The larger the value of A(x), the stronger the grade of membership for x in U is.

For demonstrating membership function we can use many methods. We have used triangle method and trapezoid method here because of their simplicity and clarity. The fuzzy systems are mostly demonstrated by these membership functions. Triangle membership is defined:

$$A(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } x \epsilon (a,b) \\ \frac{c-x}{c-b} & \text{if } x \epsilon (b,c) \\ 1 & \text{if } x > c \end{cases} \quad (2)$$

And trapezoid membership is defined:

$$A(x) = \begin{cases} 0 & \text{if } x \leq a \\ \frac{x-a}{b-a} & \text{if } x \epsilon (a,b) \\ 1 & \text{if } x \epsilon (b,c) \\ \frac{d-x}{d-c} & \text{if } x \epsilon (c,d) \end{cases} \quad (3)$$

Mamdani method has been used for fuzzy inference system in the proposed system. This system was presented by Mamdani for the first fuzzy control in a system that was a steam engine in London University. He used some rules for expressing human experiences in his system. Most of the today's fuzzy control systems are made of a set of defined rules and according to Mamdeni's method. Centroid method has also been used to defuzzification of system. The three main parts of a fuzzy system are fuzzification, rules and defuzzification module. The input data are in fuzzy format. Then, according to the defined rules in the system for the inputs necessary conclusions are done. At the next phase, the necessary wanted fuzzy outputs are produced. Finally, the data is changed to an acceptable and classic form by industrial machines. Mainly, a fuzzy system is to achieve a set of local input-output relationships that describe a process. Many applications are based on fuzzy system due to their ability to process imprecise data and ambiguous concepts [15], [16], [27].

3. PROPOSED FUZZY SYSTEM FOR IMAGE ANNOTATION

The main components of the proposed image annotation fuzzy system is shown in Fig. 2.

First, the feature extraction component extracts low-level features (color, texture and etc) from input image. In this paper, we use four features including: color moment and color correlogram to color features and, wavelet and texture moment to textture features.

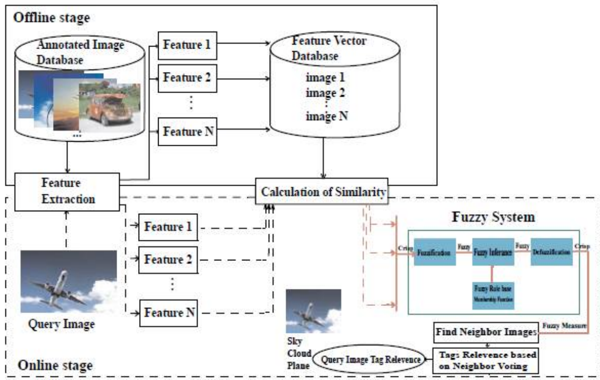

Figure 3: The model of the proposed method

Next, any features are obtained and compared with image dataset based on its features and the best similarity between those using Euclidean distance (Eq. 4) is obtained.

$$ED(M^k, M^t) = \sqrt{\sum_{n=1}^{N} ED(M_i^k, M_i^t)^2} \quad (4)$$

where M^k and M^t are image query and image database respectively and *i* is a feature range. Closer distance represents the higher similarity between images. Before it is given to fuzzy input, we separately normalize the features into a length unit between 0 and 1([0-1]). In order to normalization we use equation 5 where *V(i)* is input value, and Max(u) and Min(u) are maximum value and minimum value u vector, respectively.

$$\frac{V(i) - \text{Min } (u)}{Max(u) - Min(u)} \quad (5)$$

These similarities are input crisps to fuzzy system. Output fuzzy is a number which shows the best similarity between image dataset. Finally, find *k* nearest neighbor images with max similarity. Calculate repeat keywords from *k* image with max similarity and write *N* keywords with max repeat (Image annotation results).

For doing the necessary fuzzy processing on the data collected by the system features, first fuzzification process was done on the inputs. In order to do so three linguistic variants consisting of low, normal and high and for output parameter variants consisting of very low, low, normal, high and very high have been used for all of the system input parameters. Expression language for system parameters is shown in Table 1.

TABLE 1
EXPRESSION LANGUAGE FOR SYSTEM PARAMETERS

Parameter	Type	Generate with Fuzzy
Feature1	Input	Low-Normal-High
Feature2	Input	Low-Normal-High
Feature3	Input	Low-Normal-High
Feature4	Input	Low-Normal-High
Similarity	Output	Very Low-Low-Normal-High-Very High

The presented system was completely carried out on Matlab fuzzy toolbox. In Fig. 3, system membership function of one of the inputs has been presented. Other inputs and outputs are similar. Fuzzy system will be concluded according to the dependence of every input data on the basis of the rules that defined for the system. It makes the similarity to be adjusted on its basis.

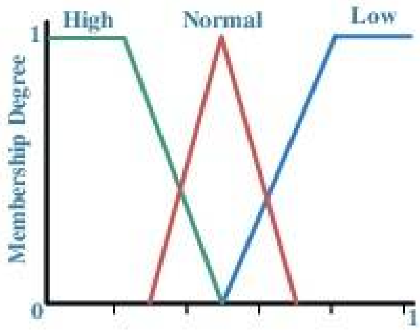

Figure 4: An example of membership function.

In the designed system, 81 rules are considered according to the conditions of the images and also its most limits. Table 2 shows some of these rules that have been performed in the environment. For example, according to the first rule, if all of the inputs have low surface, similarity becomes very low.

TABLE 2
DESIGNED RULES FOR SYSTEM

Rule	Feature1	Feature2	Feature3	Feature4	Similarity
1	Low	Low	Low	Low	Very Low
2	Low	Low	Low	Normal	Low
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
81	High	High	High	High	Very High

4. EXPERIMENTAL RESULTS

Our experimental results are explained in this section.

A. Dataset

Corel5k. One of the benchmarks for evolutionary image annotation is Corel5k image dataset. This set contains 5000 images in 50 categories. In each category, there are 100 almost similar images, such as cars, sunset, animals, flowers and so on. The images are stored in JPEG format with size 192x128. The set split into 4500 training and 500 test examples. All images are also annotated from a dictionary of 374 keywords, with each image annotated by 1 to 5 keywords. Distribution number of images for keywords between 1 to 5 are 25, 341, 1416, 2708 and 18, respectively [28].

IAPR TC12: This set consists of 19627 images. Appropriate for testing the scalability of annotation algorithms. The set splits into 17665 training and 1962 test examples. All images are also annotated from a dictionary of 291 keywords, with each image annotated by 1 to 23 keywords and average 5.7 for each image.

To evaluate the performance annotation results we used three measures: average Precision (P), average Recall (R) and F-score (F+). The annotation precision for a test image is defined as the number of keywords that are assigned correctly, divided by the total number of keywords predicted to have the image. The annotation recall is defined as the number of keywords that are assigned correctly, divided by the number of manual keywords image. The averaged precision and recall over all test images as well as the number of individual images with positive recall are used for evaluation.

$$Precision = \frac{Number\ of\ keywords\ that\ assigned\ correctly}{Total\ number\ of\ keywords\ predicted} \tag{6}$$

$$Recall = \frac{Number\ of\ keywords\ that\ assigned\ correctly}{Number\ of\ manual\ keywords\ image} \tag{7}$$

Another measure is F-score, which is obtained from a relation between precision and recall. F-score are defined as:

$$F-score = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{8}$$

We adopt a combined 124-D global feature as a tradeoff between effectiveness and efficiency. The feature is calculated as follows. For each image, we extract 64-D color correlgoram [29], 14-D color texture moment [30], 40-D Wavelet [31] and 6-D RGB color moment. Then, we separately normalize the three features into unit length and use the Euclidean distance as a dissimilarity measurement.

B. Results

We consider presented algorithm on Corel5k and IAPR TC12 datasets. Number of images for nearest neighbors similarity and number of words to any test image are assumed to be 50 and 4, respectively.

We compare the proposed fuzzy neighbor voting algorithm to the following six state-of-the-art algorithms for automatic image annotation:

Cross-Media Relevance Model (CMRM)[32]: an automatic approach to annotate and retrieve images based on a training set of images. Regions in an image can be described using a small vocabulary of blobs. Blobs are generated from image features using clustering.

Continuous Relevance Model (CRM)[33]: A new statistical generative model for learning the semantics of images.

NonParametric Density Estimation (NPDE)[34]: A simple framework for automated image annotation based on nonparametric density estimation.

Mutli-Instance Learning (MIL)[35]: A model based on mutli-instance learning algorithm and Gaussian mixture model.

Random Forest with Clustering-Particle Swarm Optimization (RFC-PSO)[36]: The proposed approach is based on the optimization of classes' scores using particle swarm optimization. In addition, random forest classifier and normalized cuts algorithm have been applied for automatic image classification, annotation, and clustering.

Baseline neighbor voting [11]: Explained in section 2.2.

Table 3 and Fig. 4 compare the proposed models in this article and previous traditional annotation models such as CMRM, CRM, NPDE, MIL, RFC-PSO and baseline neighbor voting. Table 3 and Fig. 5 show that the proposed fuzzy method has the highest average precision, average recall and F-score AIA models on Corel5k.

TABLE 3
PERFORMANCE OF VARIOUS ANNOTATION MODELS ON COREL5K VS. PROPOSED METHOD

Models	Recall	Precision	F-Score
CMRM[30]	0.09	0.10	0.9
CRM[31]	0.19	0.16	0.17
NPDE[32]	0.21	0.18	0.19
MIL[33]	0.22	0.20	0.21
RFC-PSO[34]	0.22	0.26	0.24
Voting[11]	0.29	0.25	0.27
Fuzzy	0.32	0.27	0.29

Compared to baseline neighbor voting, fuzzy neighbor voting allows improving the effectiveness of image tag recommendation with 10 percent (from 0.29 to 0.32) to recall and with 8 percent (from 0.25 to 0.27) precision. We can observe similar results in F-score measure with 11 percent improvement (from 0.268 to 0.293).

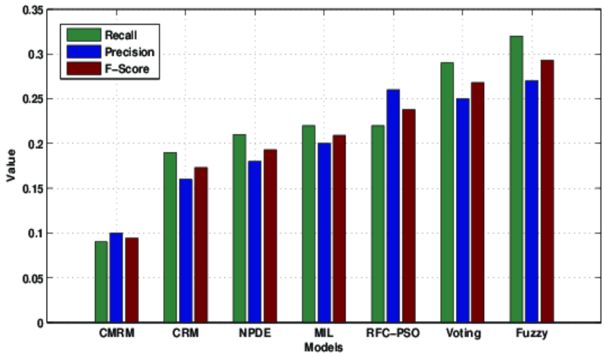

Figure 5: Performance of various annotation models on Corel5k vs. Proposed method.

Next, we present experimental results obtained on the IAPR TC12 dataset. Table 4 and Fig. 6 compare the current state-of-the-art for this dataset vs. our method. It outperforms other annotation algorithms.

TABLE 4
PERFORMANCE OF VARIOUS ANNOTATION MODELS ON IAPR TC12 VS. PROPOSED METHOD

Model	Recall	Precision	F-Score
MBRM[38]	0.23	0.24	0.24
JEC[39]	0.19	0.29	0.23
Voting[11]	0.25	0.27	0.26
Fuzzy	0.27	0.30	0.28

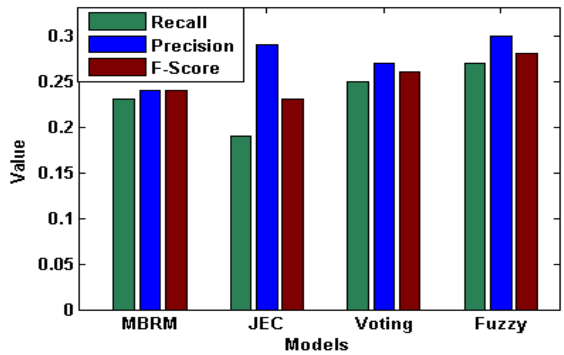

Figure 6: Performance of various annotation models on IAPR TC12 vs. Proposed method

Compared to the baseline neighbor voting, fuzzy neighbor voting allows improving the effectiveness of image tag recommendation with 8 percent (from 0.25 to 0.27) to recall and with 11 percent (from 0.27 to 0.30) precision. We can observe similar results in F-score measure with 9 percent improvement (from 0.259 to 0.284).

5. CONCLUSION

Automatic image annotation process makes metadata (text or keywords) for a digital image based on its visual content. Image annotation can fill the gap between low level features (color, texture, shape, etc.)

extraction and high level semantic features (text and keywords about the image).

This paper presented an algorithm for automatic image annotation using novel neighbor voting. For this purpose, a fuzzy model to combine features was used. Experimental results showed that the proposed method compared with other methods has the highest accuracy, recall and F-score AIA models. Also, fuzzy neighbor voting improved performance baseline neighbor voting to 10 percent of precision, 8 percent of recall and 11 percent of F-score.

We can identify a number of directions for future research. First, use for other features including texture, shape, edge or combination of these as fuzzy input. One way is use for the fuzzy multi-level systems that each level has one feature or combination of features. Second, classification of retrieved images before obtaining the final tags. Clustering algorithms can be used. Third, sensitivity analysis on important model parameters especially number of images for nearest neighbor's similarity and number of words to any test image.

REFERENCES

[1] D. Zhang , Md. Monirul Islam, and Guo jun Lu, "A review on automatic image annotation techniques", *Pattern Recognition* , vol. 45, pp. 346-362, 2012.

[2] F. Wang, "A survey on automatic image annotation and trends of the new age", *Procedia Engineering*, vol. 23, pp. 434-438, 2011.

[3] R. Datta, D. Joshi, J. Li, and J. Wang, "Image retrieval: Ideas, influences, and trends of the new age", *ACM Comput. Surveys (CSUR)*, vol. 40, no. 2, pp. 5, 2008.

[4] Y. Liu, D. Zhang, G. Lu, and W. Ma, "survey of content-based image retrieval with high-level semantics", *Pattern Recognition*, vol. 40, no. 1, pp. 262-282, 2007.

[5] Yuan. Ying, F. Wu, J. Shao, and Y. Zhuang, "Image annotation by semi-supervised cross-domain learning with group sparsity", *Journal of Visual Communication and Image Representation*, vol. 24, no. 2, pp. 95-102, 2013.

[6] J. Liu, M. Li, Q. Liu, H. Lu, and S. Ma, "Image annotation via graph learning, *Pattern Recognition"*, vol. 42, no. 2, pp. 218-228, Feb. 2009.

[7] Ch. Huang, F. Meng, W. Luo, and Sh. Zhu, "Bird breed classification and annotation using saliency based graphical model", *Journal of Visual Communication and Image Representation*, vol. 25, no. 6, pp. 1299-1307, 2014.

[8] D. Zhang, , M. Islam, and G. Lu, "Structural image retrieval using automatic image annotation and region based inverted file", *Journal of Visual Communication and Image Representation*, vol. 24, no. 7, pp. 1087-1098, 2013.

[9] Ja-Hwung Su, et al. "Effective semantic annotation by image-to-concept distribution model", *Multimedia, IEEE Transactions on*, vol. 13.3, pp. 530-538, 2011.

[10] Y. Yang, Z. Huang, Y. Yang, J. Liu, H. Tao Shen, and J. Luo, "Local image tagging via graph regularized joint group sparsity", *Pattern Recognition*, vol. 46, no. 5, pp. 1358-1368, 2012.

[11] X. Li, C.G.M. Snoek, and M. Worring, "Learning Social Tag Relevance by Neighbor Voting", *IEEE Trans. Multimedia*, vol. 11, no. 7, pp. 1310-1322, Nov. 2009.

[12] L. Wu, R. Jin, and A. K. Jain, "Tag Completion for Image Retrieval", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 716-727, March 2013.

[13] S. Lee, W. De Neve, Y. Man Ro, "Visually weighted neighbor voting for image tag relevance learning", *Multimed Tools Appl*, April, 2013. DOI 10.1007/s11042-013-1439-3.

[14] L. H. Zadeh, Fuzzy sets, *Information and control*, 1965.

[15] Ross, J Timothy, "Fuzzy logic with engineering applications", *John Wiley & Sons*, 2009.

[16] J. A. Sanz, , M. Galar, A. Jurio, A. Brugos, M. Pagola, and H. Bustince, "Medical diagnosis of cardiovascular diseases using an interval-valued fuzzy rule-based classification system", *Applied Soft Computing*, 2013.

[17] S. Dasiopoulou, C. Doulaverakis, V. Mezaris, I. Kompatsiaris, M.G. Strintzis, "An Ontology-Based Framework for Semantic Image Analysis and Retrieval", *Semantic-Based Visual Information Retrieval, Yu-Jin ZHANG (Eds), Idea Group Inc.*, 2007.

[18] Zh. Hua, X. Wang, Q. Liu, H. Lu, "Semantic knowledge extraction and annotation for web images", *Proceedings of the 13th annual ACM international conference on Multimedia*, Hilton, Singapore, November 06-11, 2005.

[19] M. Han, X. Zhu, W. Yao, "Remote sensing image classification based on neural network ensemble algorithm", *Neurocomputing*, vol. 78 (1), pp. 133-138, 2012.

[20] Y. Han, F. Wu, Q. Tian, Y. Zhuang "Image annotation by input-output structural grouping sparsity", *IEEE Transactions on Image Processing* (99), 2012.

[21] Z. Chen, Zh. Chi, H. Fu, D. Feng, "Multi-instance multi-label image classification: A neural approach", *Neurocomputing*, vol. 99 , pp. 298-306, 2013.

[22] Sh. Zhang, J. Huang, H. Li, and D. N. Metaxas, "Automatic image annotation and retrieval using group sparsity", *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 42, no. 3, pp. 838-849, 2012.

[23] T. Chaira, and A. K. Ray, "Fuzzy measures for colour image retrieval", *Fuzzy Sets and Systems* , pp. 545-560 , 2005.

[24] F. Long, H. Zhang, and D.D. Feng, "Fundamentals of content-based image retrieval", *in: Multimedia Information Retrieval and Management: Technological Fundamentals and Applications, Springer, 2003.*

[25] S. Jeong, C.S. Won, R.M. Gray, Image retrieval using colour histograms generated by Gauss mixture vector quantization, *Comput. Vision Image Underst.* vol. 94 (1–3), pp. 44-66, 2004.

[26] Y. Yang, Z. Huang, H. T. Shen, Zhou, X., "Mining multi-tag association for image tagging", *World Wide Web* vol. 14(2), 133-156., 2011.

[27] P. Villar, A. Fernandez, R. A. Carrasco, and F. Herrera, "Feature selection and granularity learning in genetic fuzzy rule-based classification systems for highly imbalanced data-sets.", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, vol. 20 (03),369-397, 2012.

[28] P. Duygulu, K. Barnard, J. De Freitas, and D. Forsyth, "Object recognition as machine translation:learning a lexicon for a fixed image vocabulary", *Proceedings of European Conferenceon Computer Vision(ECCV)*, vol. 2353, pp. 97-112., 2002.

[29] J. Huang, S. Kuamr, M. Mitra, W.-J. Zhu, R. Zabih, "Image indexing using colour correlogram", in: Proceedings of the *CVPR97*, pp. 762-765., 1997.

[30] H. Yu, M. Li, H. Zhang, and J. Feng, "Color texture moment for content- based image retrieval", in Proc. ICIP, pp. 929–932, 2002.

[31] B. S. Manjunath, and W. Y. Ma, "Texture features for browsing and retrieval of image data, Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol 18, no. 8, pp. 837-842., 1996.

[32] J. Jeon, V. Lavrenko, R. Manmatha, "Automatic image annotation and retrieval using cross-media relevance models", In: *26th annual international ACM SIGIR conference on research and development in information retrieval*. ACM, Toronto, 28 July-1 August 2003, pp 119-126.

[33] V. Lavrenko, R. Manmatha, J. Jeon, "A model for learning the semantics of pictures", In: *16th conference on advances in neural information processing systems* (NIPS 16), Vancouver. MIT Press, Canada,8-13 December 2003.

[34] A. Yavlinsky, E. Schofield, and S. Ruger, "Automated image annotation using global features and robust nonparametric density estimation", *in Proc. ACM Int. Conf.* Image Video Retrieval,pp. 507-517, 2005.

[35] S. Zhu, X. Tan, "A novel automatic image annotation method based on multi-instance learning", *Procedia Eng*, vol. 15:3439-3444, 2011.

[36] N. El-Bendary , T. h. Kim , A. Hassanien , M. Sami, "Automatic image annotation approach based on optimization of classes scores", *Computing*, 96(5), pp. 381-402, 2014.

[37] Li, Zhixin, L. Li, K. Yan, and C. Zhang, "Automatic image annotation based on fuzzy association rule and decision tree." InProceedings of the 7th International Conference on Internet Multimedia Computing and Service, p. 12. ACM, 2015.

[38] S.L. Feng, R. Manmatha, and V. Lavrenko, "Multiple bernoulli relevance models for image and video annotation", In Computer Vision and Pattern Recognition. *CVPR 2004*. Proceedings of the 2004 IEEE Computer Society Conference on (vol. 2, pp. II-1002). IEEE. 2004.

[39] A. Makadia, V. Pavlovic, and S. Kumar, "A new baseline for image annotation", In Computer VisionECCV 2008 (pp. 316-329). *Springer Berlin Heidelberg*. 2008.

[40] D. Arias-Aranda, J. L. Castro, M. Navarro, J. M. Sánchez, and J. M. Zurita, "A fuzzy expert system for business management", *Expert Systems with Applications* 37, no. 12 (2010): 7570-7580.

[41] V. Maihami, F. Yaghmaee, "Color Features and Color Spaces Applications to the Automatic Image Annotation", Book chapter in *Emerging Technologies in Intelligent Applications for Image and Video Processing*. 2016 Jan 7:378.

[42] J. Johnson, L. Ballan, and L. Fei-Fei, "Love thy neighbors: Image annotation by exploiting image metadata". In Proceedings of *the IEEE International Conference on Computer Vision* (pp. 4624-4632), 2015.

[43] X. Li, T. Uricchio, L. Ballan, M. Bertini, CG. Snoek, A. Del Bimbo, "Socializing the semantic gap: A comparative survey on image tag assignment, refinement and retrieval", *ACM Computing Surveys*. arXiv preprint arXiv:1503.08248. 2016.

BIOGRAPHIES

Vafa Maihami received the M.Sc. degree in computer engineering from the University of Kurdistan, Sanandaj, Iran, in 2012. He is currently working toward the Ph.D. degree in computer engineering at Semnan University, Semnan, Iran, working on Artificial Intelligence algorithms. His research interests include image Processing, machine learning, machine vision, and wireless sensor networks.

Farzin Yaghmaee received his Ph.D. in 2010 and M.Sc. in 2002 both in Artificial Intelligence from Sharif University of Technology, Iran and received B.Sc. from AmirKabir University of Technology. He is now a faculty member of Electrical and Computer Engineering Department of Semnan University. His research interests are: image and video processing, text mining and Persian language processing tools.