Image Classification using EXIF Metadata

Juzlinda Mohd Ghazali[1], Shafinaz Mohammad Niyaz Khan[2], Lailatul Qadri Zakaria[3]

[1]*Computer Science Department, Faculty of Science and Information Technology, Selangor International Islamic University College (KUIS), Selangor, Malaysia*

[2]*IT & Multimedia Department, Faculty of Science and Information Technology, Selangor International Islamic University College (KUIS), Selangor, Malaysia*

[3]*Natural Language Processing, Faculty of Information Science and Technology, University Kebangsaan Malaysia, Selangor, Malaysia*

ABSTRACT

Semantic classification of unrestricted images is still an open problem regardless of all efforts done. Present methods have only mainly focused on features extracted from the image content (e.g. colour, texture, shape). Conversely, EXIF metadata recorded by the camera can be exploited to aid the classification process. Demonstrating scenery-object classification as an example, analysis of results has revealed different combinations of metadata features that contribute variedly in predicting scenery-object image classification when using different classifiers. The evaluation was done by using machine learning to which a dataset of 500 digital images, consisting of 250 random scenery images and 250 random object based images were trained and tested. For classification, Support Vector Machine (SVM) and K-Nearest Neighbour (KNN) classifiers were used. The research project achieved a result as high as 75.15% for the classification of scenery-object image using KNN.

Keywords : *EXIF metadata, image classification, KNN, machine learning, SVM.*

I. INTRODUCTION

Image classification is an effort to group images into semantically meaningful categories (e.g., scenery, indoor, sunset, tennis game) that is useful in content-based image retrieval and organisation application.

Even though there are many kinds of tags in EXIF metadata, only certain tags are used to determine certain image classifications [1], [2]. For example, in order to cluster photographs by events, timestamps have successfully been used for this kind of categorisation [3]. However, none of the prior research had harnessed the use of image capturing conditions of metadata (e.g., exposure time and flash) for classification purposes, and none was used specifically for scene classification.

This work has expanded the boundary of contribution in image classification by using new combination of selected EXIF features and by analysing them through selective classifiers or image classification algorithms

II. RESEARCH BACKGROUND

The demonstrated work is a result of theoretical and practical examination on EXIF metadata exploration of digital images. The project presents a new approach to semantic classification of arbitrary digital images. The study is not based on the image's content or low-level features such as colour or texture, but relies on the digital camera's technical parameters named EXIF metadata, that are embedded in the digital image captured by the digital camera [4].

In recent years of technological advancement, modern digital cameras have evolved at an extremely fast pace and are increasingly popular among the consumers to buy and take thousands of daily life photos and produce huge amount of image data around the world [3]–[6] by using various kinds of photo capturing devices that implant EXIF metadata in the captured digital images [7], [8].

Dealing with this huge accumulation of digital images, arises the need of automatic image organization through image classification in order to apply semantic meaning to these images. As for this work, it focuses on dealing with two types of classification, which are scenery and object classifications. [9]–[15] had conducted research on identifying different methods of scene classification problems, focusing on image content. [1], [2], [4], [6], [16] nonetheless embedded EXIF metadata. This work, particularly, has managed to identify some fascinating features of EXIF metadata that could assist in the scenery and object classification process.

In this work, most of the features taken into further examination were based from work done by other researchers; nonetheless a few more features had been taken into consideration based on their potential. Thus, a combination of seven features which are exposure time or shutter speed, f number (also known as F-stop) or aperture, International Standards Organisation (ISO), flash, focal length, image width and image height had been selected.

Various combinations of features were tested on a dataset comprising of 500 images (250 sceneries and

250 objects) using different classifiers in order to examine how these different combinations affect the success rate of the classification.

III. PROPOSED APPROACH

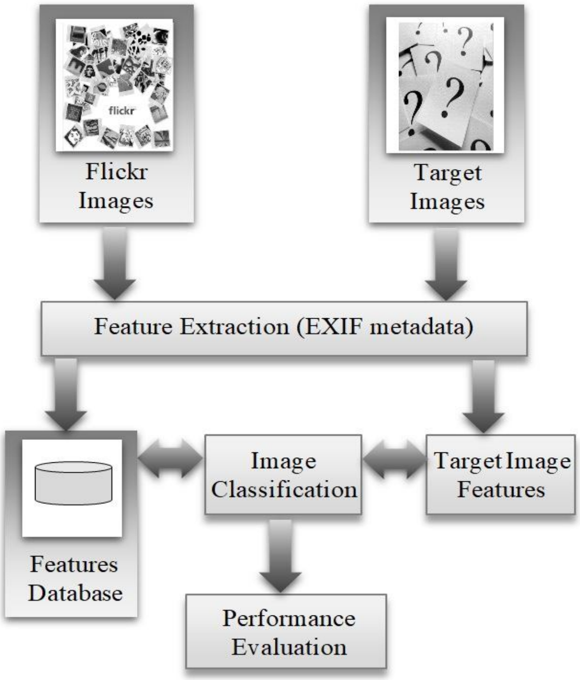


Fig. 1: Proposed approach to image classification

Fig. 1 illustrates the proposed approach to image classification. Social media sharing websites such as Flickr offers varied images that can be exploited. Concepts of image classification were chosen manually based on distinctive visual characteristics; confined within scenery and object images.

Features database was made available through careful selection and extraction from Flickr image metadata. These features were used as a training set to classify target images accordingly. Subsequently, the performance of image classification was measured using Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) classifiers. SVM and KNN were chosen as they are among the most popular and simplest of machine learning algorithms [6] and [17].

A. Support Vector Machine (SVM)

Support Vector Machine (SVM) is a very popular method in machine learning theory. SVM maximises the margin around the separating hyperplane, which is known as large margin classifiers. The decision function is fully specified by a subset of training samples, which are the support vectors. Solving SVM is a quadratic programming problem. It is seen by many as the most successful current text classification method; it alsoprovides very similar performance for other discriminative methods.

B. K-Nearest Neighbour (KNN)

K-Nearest Neighbour (KNN) algorithm is one of the simplest methods in the field of machine learning. The idea of this classifier is that the training samples of the same category are grouped together in a multidimensional feature space. If most of the K nearest neighbours of a test sample belong to a single category, this test sample can also be included in this category. K can be determined by the user in the classification phase; usually it is an odd number which will not give any ties

IV. RESULTS AND DISCUSSIONS

Throughout this section, the metadata features are addressed as shown in Table 1. There were no missing metadata values during the testing of the datasets with the usage of PyCharm software. Various accuracy results were obtained from two to seven features tested by using SVM and KNN classifiers. Some combinations of EXIF metadata features worked better in a given classifier than other combinations.

Table 1: EXIF metadata feature label

Label	EXIF Metadata Feature
F1	Exposure Time/Shutter Speed
F2	F Number/Aperture
F3	ISO
F4	Flash
F5	Focal Length
F6	Image Width
F7	Image Height

Table 2 shows the recorded accuracy values of more than 70% using KNN classifier for various metadata feature combinations. The combination of image width (F6) and image height (F7) shows the best accuracy of 75.15%. Looking at a slight reduction in accuracy of 72.73%, it is noticeable that the same features F6 and/or F7 appeared in the combination of features used.

Table 2: Accuracy using various metadata feature combinations with KNN

Feature Types	Accuracy (KNN)
F2 + F5	70.91%
F2 + F6	72.73%
F2 + F7	70.30%
F3 + F7	72.73%
F4 + F7	70.30%
F5 + F6	72.73%
F6 + F7	75.15%
F2 + F5 + F6	70.91%
F2 + F5 + F7	70.91%
F2 + F6 + F7	72.73%
F5 + F6 + F7	72.73%
F2 + F5 + F6 + F7	70.91%

Table 3 also shows accuracy values more than 70% but this time using SVM classifier for various metadata feature combinations. The best recorded accuracy for SVM is 72.73%. There are seven different combinations having this accuracy score. All of these combinations included F number (F2) and flash presence (F4). Notably, F6 and/or F7 were also included in some of these top scorers. Thus, it is shown that F number, flash presence, image width, and height are metadata features that mainly contribute to the accuracy of image scenery-object classification.

Table 3: Accuracy using various metadata feature combinations with SVM

Feature Types	Accuracy (SVM)
F1 + F2	70.91%
F2 + F4	72.73%
F1 + F2 + F3	70.91%
F1 + F2 + F4	70.91%
F1 + F2 + F5	70.91%
F1 + F2 + F6	70.91%
F1 + F2 + F7	70.91%
F2 + F4 + F5	72.73%
F2 + F4 + F6	72.73%
F2 + F4 + F7	72.73%
F1 + F2 + F3 + F4	70.91%
F1 + F2 + F3 + F5	70.91%
F1 + F2 + F3 + F6	70.91%
F1 + F2 + F3 + F7	70.91%
F1 + F2 + F4 + F5	70.91%
F1 + F2 + F4 + F6	70.91%
F1 + F2 + F4 + F7	70.91%
F1 + F2 + F5 + F6	70.91%
F1 + F2 + F5 + F7	70.91%
F1 + F2 + F6 + F7	70.91%
F2 + F4 + F5 + F6	72.73%
F2 + F4 + F5 + F7	72.73%
F2 + F4 + F6 + F7	72.73%
F1 + F2 + F3 + F4 + F5	70.91%
F1 + F2 + F3 + F4 + F6	70.91%
F1 + F2 + F3 + F4 + F7	70.91%
F1 + F2 + F3 + F5 + F6	70.91%
F1 + F2 + F3 + F5 + F7	70.91%
F1 + F2 + F3 + F6 + F7	70.91%
F1 + F2 + F4 + F5 + F6	70.91%
F1 + F2 + F4 + F5 + F7	70.91%
F1 + F2 + F4 + F6 + F7	70.91%
F1 + F2 + F5 + F6 + F7	70.91%
F2 + F4 + F5 + F6 + F7	72.73%
F1 + F2 + F3 + F4 + F5 + F6	70.91%
F1 + F2 + F3 + F4 + F5 + F7	70.91%
F1 + F2 + F3 + F4 + F6 + F7	70.91%
F1 + F2 + F3 + F5 + F6 + F7	70.91%
F1 + F2 + F4 + F5 + F6 + F7	70.91%
F1 + F2 + F3 + F4 + F5 + F6 + F7	70.91%

Fig. 2 illustrates the image classification performance, as a whole, comparing between SVM and KNN classifiers by the number of feature combinations. For SVM, there is a consistent score from two to five feature combinations of 72.73% and slightly decreases for six and seven feature combinations, with 70.91%. Nevertheless, for KNN, the scores vary throughout the numbers of feature combination. The best score is with two feature combination, at 75.15%. Then, it continuously decreases at 72.73% for three-feature combination, 70.91% for four-feature combination, and 68.48% for five-feature combination. It increases at 70.9% for six-feature combination and falls back at 66.67% for seven-feature combination.

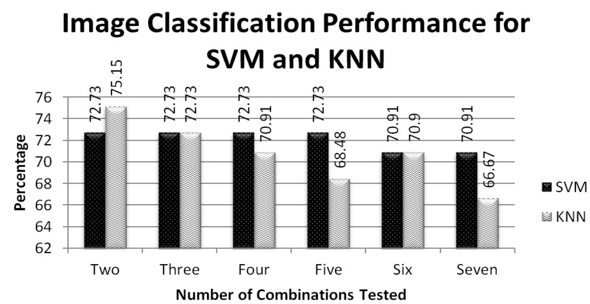


Fig. 2: Image Classification Performance for SVM and KNN

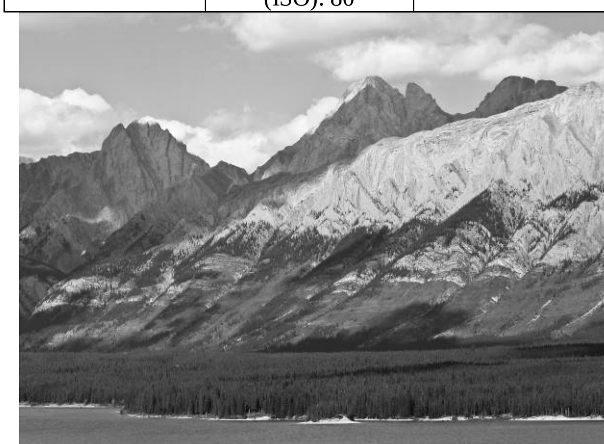
The different scenarios for scenery and object class of the seven features are demonstrated in Table 4. The comparative values between both classes are shown for all seven features.

Table 4: EXIF metadata comparison to identify classes of scenery and object-based image

Category	Scenery	Object
Exposure Time / Shutter Speed	Slow shutter speed (example: 2 to 1/2 second).	Fast shutter speed (example: 1/1000 to 1/4000 second).
F Number / Aperture	High f number (small opening area of the camera lens).	Low f number (big opening area of the camera lens).
International Standards Organisation (ISO)	High ISO speed.	Low ISO speed.
Flash	Not being used.	Usually being used.
Focal Length	Short focal length (wide angle of lens view).	Long focal length (narrow angle of lens view).
Image Width	In normal cases, width value is higher than its height value.	In normal cases, width value is lower than its height value.
Image Height	In normal cases, height value is lower than its width value.	In normal cases, height value is higher than its width value.

Fig. 3 displays three sample images with features that contains EXIF metadata which carry the opposite class values. Scenery images (the first two images) have the values of exposure time or shutter speed of 1/640 and 1/500, International Standards Organisation (ISO) of 80 for both, and focal length of 20.9 and 27.9 respectively. Whereas object image (the third image) has 1/60, 280, and 18.0 for exposure time or shutter speed, International Standards Organisation (ISO), and focal length respectively. This shows that all three features' values of exposure time or shutter speed, International Standards Organisation (ISO), and focal length for the three sample images portrayed the values that are of the other class. These are conditions where images were classified wrongly when using these features.

Exposure Time / Shutter Speed: 1/640	International Standards Organisation (ISO): 80	Focal Length: 20.9



Exposure Time / Shutter Speed: 1/500	International Standards Organisation (ISO): 80	Focal Length: 27.9

Exposure Time / Shutter Speed: 1/60	International Standards Organisation (ISO): 280	Focal Length: 18.0

Fig. 3: Image samples which were wrongly classified

V. CONCLUSION

This work has shown that F number, flash presence, image width, and height are metadata features that mainly contribute to the accuracy of image scenery-object classification. KNN scored the best image classification performance using two metadata features with 75.15% accuracy. Using EXIF metadata as features for semantic classification has proven to be promising, with the additional advantage of avoiding computational complexity. Interesting directions include exploring other scene classes and experimenting on other classifiers

ACKNOWLEDGEMENT

We would like to thank KUIS for the research grant given to conduct this research.

REFERENCES

[1] Boutell, M., & Luo, J. (2004a, June 27). "*Bayesian fusion of camera metadata cues in semantic scene classification*". In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on (Vol. 2: 623-630). IEEE.

[2] Boutell, M., & Luo, J. (2004b, August 23). "*Photo classification by integrating image content and camera metadata*". In Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on (Vol. 4: 901-904). IEEE.

[3] Platt, J. C. (2000, June 16). "*AutoAlbum: Clustering digital photographs using probabilistic model merging*". Incbaivl: 96-101. IEEE.

[4] Hingorani, M. (2016, September). "*Into The Technicalities of Photography: An EXIF Data Exploration*". University of California, Santa Barbara, California, United States of America.

[5] Keim, D., Andrienko, G., Fekete, J. D., Gorg, C., Kohlhammer, J., & Melançon, G. (2008, August). "*Visual analytics: Definition, process, and challenges*". Lecture notes in computer science, 4950: 154-175.

[6] Guo, F., Bossard, L., & Gool, L. V. (2011, August 18). "*Indoor outdoor image classification*". Bachelor Thesis.

[7] "*Japan Electronics and Information Technology Industries Association*" (JEITA), CP-3451 (2002, April). Exchangeable image file format for digital still cameras: EXIF version 2.2.

Technical Standardization Committee on AV & IT Storage Systems and Equipment.

[8] Alvarez, P. (2004). "*Using extended file information (EXIF) file headers in digital evidence analysis***"**. International Journal of Digital Evidence, 2(3): 1-5.

[9] Boutell, M., Luo, J., & Gray, R. T. (2003, July 6). "*Sunset scene classification using simulated image recomposition***"**. In Multimedia and Expo, 2003. ICME'03. Proceedings. 2003 International Conference on (Vol. 1: 37-40). IEEE.

[10] Carson, C., Belongie, S., Greenspan, H., & Malik, J. (2002 August). "*Blobworld: Image segmentation using expectation-maximization and its application to image querying***"**. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(8): 1026-1041.

[11] Oliva, A., & Torralba, A. (2001 May 1). "*Modeling the shape of the scene: A holistic representation of the spatial envelope***"**. International journal of computer vision, 42(3): 145-175.

[12] Paek, S., & Chang, S. F. (2000). "*A knowledge engineering approach for image classification based on probabilistic reasoning systems. In Multimedia and Expo***"**, 2000. ICME 2000. 2000 IEEE International Conference on (Vol. 2: 1133-1136). IEEE.

[13] Song, Y., & Zhang, A. (2003). "*Analyzing scenery images by monotonic tree***"**. Multimedia Systems, 8(6): 495-525.

[14] Szummer, M., & Picard, R. W. (1998, January 3). *"Indoor-outdoor image classification. In Content-Based Access of Image and Video Database*," 1998. Proceedings, 1998 IEEE International Workshop on (pp. 42-51). IEEE.

[15] Vailaya, A., Figueiredo, M., Jain, A., & Zhang, H. J. (1999, July). "*Content-based hierarchical classification of vacation images***"**. In Multimedia Computing and Systems, 1999. IEEE International Conference on (Vol. 1: 518-533). IEEE.

[16] J. Ghazali, S. A. Noah, and L. Zakaria. (2013). "*Classification of images for automatic textual annotation: A review of techniques***"**. Journal of Applied Sciences, vol. 13, no. 6. pp. 760–767, 2013.

[17] Ghazali J M, Ahmad S N, Elias K A, Khan S M N, Noah S A M, Zakaria L Q. (2018). "*Annotating images by semantic representation using the open knowledge base***"**. J. Fundam. Appl. Sci., 2018, 10(3S), 691-702