Large Field/Close-Up Image Classification: From Simple to Very Complex Features

Quyet Tien Le, Patricia Ladret, Huu-Tuan Nguyen, Alice Caplier

Large field / close-up image classification: From simple to very complex features

Quyet-Tien LE[1,2], Patricia LADRET[1], Huu-Tuan NGUYEN[2], and Alice CAPLIER[1]

[1] GIPSA Lab, Grenoble Alpes University, Saint Martin d'Heres Cedex FRANCE
{quyet-tien.le, patricia.ladret, alice.caplier}
@gipsa-lab.grenoble-inp.fr
[2] Vietnam Maritime University, Hai Phong, VIETNAM
{tienlqcnt, huu-tuan.nguyen}
@vimaru.edu.vn

Abstract. In this paper, the main contribution is to explore three different types of features including Exchangeable Image File (EXIF) features, handcrafted features and learned features in order to address the problem of large field / close up images classification with a Support Vector Machine (SVM) classifier. The impacts of every feature set on classification performances and computational complexities are investigated and compared to each other. Results prove that learned features are of course very efficient but with a computational cost that might be unreasonable. On the contrary, it appears that it is worthy to consider EXIF features when available because they represent a very good compromise between accuracy and computational cost.

Keywords: image classification · large field image · close-up image · handcrafted features · exif features · learned features · feature evaluation · feature selection · support vector machine · transfer learning.

1 Introduction

Image classification has been studied for many years and the main idea is to use image features that are computed from image data either by hand [2, 24] or via a learning algorithm [6, 7] to separate images into different categories. The focused problem in this paper is large field / close-up image classification (image samples can be seen in Fig. 1). This classification can be used in many applications. By analyzing tourists' photos, it is possible to provide tourists with valuable information about places including beautiful panorama scenes (mountains, rivers, castles,. . .) or local views (food, local specialty, exhibit,. . .) [24]. This classification also helps understanding what attract viewers' attention to improve aesthetic quality assessment [11].

Until now, there are few researches about this topic. In [26], Wang et al. propose a method using color coherence vector and color moments to classify close-up and non close-up images. In another study, Zhuang et al. [27] divide an image into 256 parts. The number of edge points is counted in each part to build a 256 bin histogram. The 256 bin values and standard deviation of those values are the key features to classify close-up and distance view images. In [24], Tong et al. use features representing the distributions of high frequencies in the first classification stage and the spatial size and the conceptal size are used in the second one to classify distance / close-up view images. All features used in those classification methods are handcrafted features. The role of Exchangeable Image File (EXIF) features and learned features for that task is still a question.

Handcrafted features have been widely used for image classification [16]. Nowadays, deep learning approaches are the must for object classification [19]. At the same time, EXIF data has not been widely used for image classification. EXIF data are metadata (data information of data) and tags revealing photo information such as picture-taking time, picture-taking conditions [23]. Surprisingly, EXIF features have been only occasionnaly used in researches. In [9], Huang et al. use the manufacturer, camera model, date and time stamp and some other EXIF parameters as watermark information to protect image copyright. In [15], aperture, exposure value, ISO and picture-taking time are exploited to enhance region of interest detection. In [3, 4], to classify in-door and out-door images, Boutell et al. present a method integrating image content and EXIF data consisting of exposure time, flash use and focal length.

In this study, the problem of large field / close-up image classification with Support Vector Machine (SVM) is considered. The performances of classification based on EXIF features, handcrafted features or learned features are compared in terms of accuracy and computational complexity. Handcrafted based feature method is considered as the reference for that study. In order to evaluate the influences of the different feature types fairly, SVM is chosen because of its simplicity. If complex classifiers are used, the accuracy of the classifications could be affected not only by input features but also by the suitability between the model structure and input features.

Experiments in this paper have been focused on 2 datasets [1]. First of all, the Flickr dataset including 800 large field images and 800 close-up images is used for experiments of large field / close-up image classification because for this database EXIF data are available which is not always the case. Another different database, the CUHKPQ dataset with 600 large field and 600 close-up images [22] without EXIF data is used only for the feature selection process among the handcrafted and learned features.

The paper is organized as follows. In section 2, handcrafted features and learned features are defined and selected. EXIF features are described in section 3. Section 4 presents experiments and results. Conclusions are drawn in the last part.

[1] The databases are available at "http://www.gipsa-lab.fr/∼quyettien.le/projets_en.html" and "http://www.mediafire.com/file/58e8jui7547mam2/LargeFieldCloseupImageDatabase.zip/file"

Fig. 1. The first, second, third and fourth rows (separated by the red lines) present the best close-up, large field image classifications (images being classified correctly and having the biggest distances to the hyper-plane) and the worst large field and close-up image classifications (images being classified incorrectly and having the biggest distances to the hyper-plane) based on the EXIF, handcrafted and learned features respectively. A: Aperture, F: Focal length. E: Exposure time, I: Illumination measure.

2 Handcrafted and learned features

2.1 Handcrafted features

The main goal of this part is to select among usual features computed from image data for large field / close-up image classification the most relevant ones. Thus, a large handcrafted feature set is first built from common handcrafted features appearing in different researches [1, 5, 12, 17, 25]. Features consisting of color histogram, sharpness, hue, saturation, brightness, color saliency and contrast are computed for different regions including the whole image, regions of interest(regions attracting viewers' eyes), background regions and regions split by symmetry rules, landscape rule, rule of thirds (See Fig. 2) to define 2,003 handcrafted features.

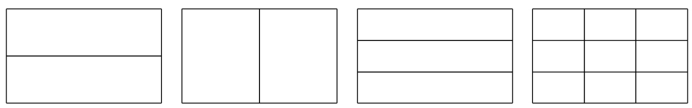

Fig. 2. The two first left images present regions split by symmetry rules. The others show regions split by landscape rule and rule of third respectively.

2.2 Learned features

Beside being handcrafted from images, features can also be learned from images by employing deep learning [14]. VGG16 [20] is a well-known deep convolutional neural network. It includes 3 main parts including convolutional layers, fully connected layers and a prediction layer. If the prediction layer is removed, that model can be considered as a feature extractor. From images of size 244 $\times$ 244, 4,096 features are learned by this model. Although those features have been learned for the task of classifying objects in images, they can be applied for different tasks [18] such as image quality assessment [10, 21]. In this study, the VGG16 model without the prediction layer pre-trained on the ImageNet dataset for the task of classifying objects in images is considered to compute the learned features for the large field / close-up image classification on the corresponding database.

2.3 Feature selection

Because those handcrafted and learned features have been primarily designed for a different task, the impact of each feature for the purpose of large field / close-up image classification is estimated by using the relief method [13]. CUHKPQ dataset is split into a training set and a testing set. Each set includes 300 large field (L) and 300 close-up (C) images. All features of each image in the training set are calculated and normalized to range the [0,... 1]. The relevance of a given feature f for the classification is calculated as:

$$r(f) = dif(f, L, C) - dif(f, L, L) - dif(f, C, C) \tag{1}$$

$$dif(f,X,Y) = \frac{\sum_{i=1}^{\|X\|}\sum_{j=1}^{\|Y\|}(d(f,X_i,Y_j))}{\| X \| \times \| Y \|} \quad (2)$$

where $r(f)$ is a combination of the interclass and intraclass differences. The most relevant features are associated to the highest $r(f)$ values. $\| X \|$ is the number of images in set X. X_i is the i^{th} image in the set X and $d(f,x,y)$ is the absolute difference between f values of the 2 images x and y.

After calculating and normalizing the relevance values to the range [0,. . . 1], a threshold needs to be determined to select the most relevant features for the classification. To do it, the features of the testing images are calculated. An algorithm inspired of the binary search algorithm is proposed to compute the threshold as:

Input:
- Feature set: $F = \{f_0, f_1, \ldots f_m\}$
- Relief set of F: $R = \{r_0, r_1, \ldots r_m\}$
- Training set S_1 and testing set S_2

Output: Threshold T

Algorithm:
- Set 2 thresholds: $T_1 = 0$, $T_2 = 1$
- For i=1 to k (in this study k is set to 50)
  + For j=1 to 2
    Use T_j as a threshold $F_T = \{f_x | r_x \geq T_j\})$
    Use F_T, S_1, S_2 to train and test an SVM classifier.
    Gain testing accuracy A_j
    End of loop j
  + If $A_1 < A_2$
    $T_1 = T_1 + \frac{T_2 - T_1}{5}$
  + If $A_1 > A_2$
    $T_2 = T_2 - \frac{T_2 - T_1}{5}$
  + If $A_1 = A_2$
    $T_1 = T_1 + \frac{T_2 - T_1}{5}$
    $T_2 = T_2 - \frac{T_2 - T_1}{5}$
  End of loop i
- Return threshold $T = \frac{T_1 + T_2}{2}$

Applying the feature selection algorithm on the handcrafted and learned feature sets, 21 handcrafted features and 925 learned features are selected from the 2,003 handcrafted features and the 4,096 VGG16 features respectively.

3 EXIF Features

3.1 EXIF feature definition

In photography, camera tunnings are stored by digital cameras as EXIF data. 4 EXIF parameters and a combination of some of them are analyzed in this study.

Aperture Aperture refers to the size of lens opening for light when a picture is captured. This parameter is stored as a *f-stops* such as $f/1.4$, $f/2$, $f/2.8$,... in which $f\text{-}stops = \frac{f}{D}$ where f is the focal length and D is diameter of the entrance in a camera. A smaller *stops* value represents a wider aperture. The Depth Of Field (DOF) and brightness of pictures are affected by the aperture value.

Focal length Focal length exhibits the distance from the middle of the lens to the digital sensor and it also decides the angle of view in the picture. This parameter is measured in millimeters. A long focal length makes a narrow view and a wide scene is captured when using a short focal length.

Exposure time Exposure time represents the total time for light falling on the sensor of the camera during shooting. It is measured in seconds. In weak light conditions or to create some special effects, photographers use long exposure time. A short exposure time is regularly used when capturing moving objects like taking sport photos.

ISO ISO describes the sensitivity level of the sensor in a camera. ISO parameter is measured with numbers such as 100, 200, 400,... The lower ISO value represents the less sensitive mode of the sensor. The brightness of a photo decreases with the decrease of ISO.

Illumination measure Illumination measure refers to the light falling on a surface [8]. This feature is calculated as:

$$I_{measure} = \log_{10}\left(\frac{aperture^2}{exposure\ time}\right) + \log_{10}\left(\frac{250}{ISO}\right) \tag{3}$$

3.2 EXIF feature selection

In this subsection, the influence of EXIF features on large field / close-up image classification is investigated. At the first step, EXIF values of 400 large field and 400 close-up photos extracted from Flick dataset (The training set in the next experiments of large field / close-up image classification) are displayed Fig. 3. It appears that the differences of EXIF parameters between close-up and large field images are significant in aperture, focal length, illumination measure and to a smaller extent in exposure time.

Unsurprisingly the aperture data is very significant to distinguish between close up and large field images. Actually, a high aperture value is regularly chosen to highlight the objects by low DOF effect. In the other hand, because large field scenes are far from the camera, a small aperture value is set for capturing a large field photo to gain a deep DOF.

Focal length is the second discriminating feature. A large field scene is wide so photographers often use a short focal length to get the whole scene. In contrast, to focus on close-up objects, a longer focal length is regularly used to take close-up pictures.

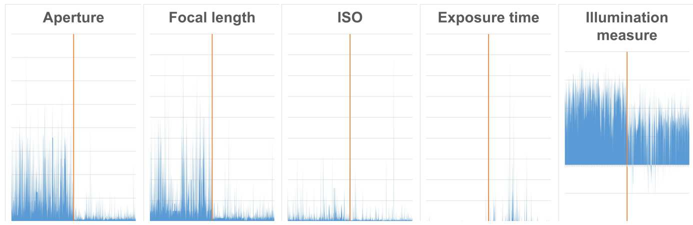

Fig. 3. EXIF values of 400 close-up images (the left side) and 400 large field images (the righ side).

Illumination measure and exposure time are also going to be considered for large field / close-up image classification. On the contrary, ISO feature is not relevant enough.

4 Experiments

4.1 Dataset and setup

Large field / close-up image classification is performed separately with EXIF, handcrafted and learned features. An SVM classifier is trained and tested to evaluate the classification performances for each feature set.

The experiments are performed on 1600 images including 800 large field and 800 close-up images collected and categorized from Flickr website by our team. Half of the large field and close-up images are selected randomly to train the classifiers while the others are used to test. Each SVM classifier is applied with linear kernel, $C = 0.5, g = 1 \times 10^{-5}, e = 1.192 \times 10^{-7}$. Evaluation criteria in this study include overall Accuracy (A), Balanced accuracy (B), Precision (P) and Recall (R) where TP, TN, FP, FN are true positive, true negative, false positive and false negative respectively and they are expressed as a number of images (cf. Table 1).

4.2 Results and discussions

Handcrafted features Table 2 shows the results of the classification based on the 21 handcrafted features. The reference classification rate using those handcrafted features is 0.873.

In order to prove the efficiency of the selected handcrafted features, the classification based on those features is compared with the classifications based on other handcrafted features including Wang's [26], Zhuang's [27] features. The results are given in Table. 2. Despite of using more features, the classifications with Wang's (105 features) and Zhuang's (257 features) feature sets have lower accuracy at 0.774 and 0.854 respectively.

Table 1. Overview of evaluation criteria.

Evaluation criteria	**Formula**
Overall accuracy	$A = \frac{TP+TN}{TP+FP+TN+FN}$
Balanced accuracy	$B = 0.5 \times \frac{TP}{TP+FP} + 0.5 \times \frac{TN}{TN+FN}$
Precision	$P = \frac{TP}{TP+FP}$
Recall	$R = \frac{TP}{TP+FN}$

Table 2. Results for large field / close-up image classification based on the 21 selected handcrafted features compared with the classification based on other handcrafted features.

		Prediction	
		Close-up image	Large field image
Ground truth	Close-up image	TP = 349	FN = 51
	Large field image	FP = 51	TN = 349
Overall accuracy	**0.873**	Balanced accuracy	0.873
Precision	0.873	Recall	0.873
Overall accuracy of classifications based on other handcrafted feature sets			
Wang's features (105 features)			**0.774**
Zhuang's features (257 features)			**0.854**

Table 3. Results for large field / close-up image classification using the 4 selected EXIF features where TP, TN, FP FN are expressed as a number of images.

		Prediction	
		Close-up image	Large field image
Ground truth	Close-up image	TP = 348	FN = 52
	Large field image	FP = 46	TN = 354
Overall accuracy	0.878	Balanced accuracy	0.878
Precision	0.883	Recall	0.870

Table 4. Results for large field / close-up image classification based on the 925 selected VGG16 features.

		Prediction	
		Close-up image	Large field image
Ground truth	Close-up image	TP = 392	FN = 8
	Large field image	FP = 1	TN = 399
Overall accuracy	0.989	Balanced accuracy	0.989
Precision	0.997	Recall	0.980

EXIF features The results of the classification based on the 4 selected EXIF features are presented in Table 3. Using a very small number of simple features, the classification accuracy 0.878 is impressive.

Learned features The results of classification with the 925 best features learned from VGG16 are shown in Table 4. It is obvious, the classification with learned features has the highest overall accuracy (0.989) but the number of features is also the biggest (925 features) among the studied feature sets.

Comparisons To start with, it appears that EXIF features are quite powerful for large filed /close-up image classification since the accuracy (0.878) is obtained with only 4 EXIF features. With handcrafted features, the number of features is higher (21 versus 4) while the classification accuracy is almost the same (0.873). Secondly, the classification with learned features has the highest accuracy (0.989). However the number of selected features is also the biggest (925 learned features against 21 handcrafted features and 4 EXIF features). In

Table 5. Results for large field / close-up image classifications based on the 4 EXIF features, 21 handcrafted features, top 21 and top 4 most relevant learned features.

Feature set	**TP**	**FP**	**TN**	**FN**	**A**	**B**	**P**	**R**
EXIF features	348	46	354	52	0.878	0.878	0.883	0.870
4 learned features	389	9	391	11	0.975	0.975	0.977	0.973
Handcrafted features	349	51	349	51	0.873	0.873	0.873	0.873
21 learned features	391	6	394	9	0.981	0.981	0.985	0.978

order to compare the role of those features, the classifications using the top 21 and top 4 most relevant learned features are performed and the results are shown in Table 5. It appears that the learned features are very efficient for large field / close-up image classification since with the same number of features as handcrafted features (21 features) the accuracy of the classification based on the 21 most relevant learned features is higher than that of the handcrafted features (0.981 versus 0.873). Similarly, with only 4 features as EXIF features, the accuracy of the classification based on the 4 best learned features is 0.975, a very high accuracy while the classification accuracy with EXIF features is 0.878.

Fig 1 shows the top 9 best classifications (images being classified correctly and having the biggest distances to the hyper-plane of the SVM classifiers) and the top 9 worst classifications (images being classified incorrectly and having the biggest distances to the hyper-plane) of each category. It appears that the feature sets are acting totally differently since there are no overlapping images between those results. The best classified close-up images using EXIF features are mostly low DOF images because of wide aperture values. Almost of the best close-up images (7 of 9) have high apertures, high illumination measures and long exposure time ($A \geq 10$ and $I \geq 4.0$ and $E \geq \frac{1}{250}$) while no image of the best or worst large field photos and only one of the worst close-up images satisfies this condition. Additionally, 6 of the 9 best large field images have small focal lengths, short exposure time and illumination measures ranging from 2.75 to 3.418 ($F \leq 50$ *and* $E \leq \frac{1}{250}$ *and* $2.75 \leq I \leq 3.418$) while no image of the best close-up photos and only one of the worst large field images have EXIF data in those ranges. With handcrafted features the best classified close up images almost have blank background because some features are handcrafted to estimate the number of background details of close-up images (those features can not be used to classify blank background or blur background) so the classifier focuses on blank background. VGG16 being pre-trained on the ImageNet dataset for purpose of classifying objects in images, the extracted features have been designed to recognize objects very well. It explains why the top classified close-up images using those features are images with fish, bird, chicken, insect. Additionally, learned features seem to focus on the high frequency details in foreground of close-up images. In contrast, the differences between the best large field image classifications and the differences between the worst classifications are not clear. Last but not least, an experiment has been conducted on a PC equipped an Intel

Table 6. Feature computational time and classification time. In this table, the total classification time is the sum of the feature computational time and the time of SVM classification based on the computed features.

Feature set	**Feature computational** time (ms)	**SVM classification time (ms)**	**Total classification time (ms)**	**Overall accuracy**
4 EXIF features	1	3	4	0.878
21 handcrafted features	23,994	4	23,998	0.873
21 learned features	347,186	3	347,189	0.981
925 learned features	347,186	55	347,241	0.989

Core i7-2670QM CPU 2.40 GHz and 11.9 GB memory to evaluate the feature computational time and classification time of EXIF, handcrafted and learned features. The feature computational time and classification time with 800 images are shown in Table 6. It is clear that EXIF features are the simplest ones when only one EXIF feature (illumination measure) needs to be computed and its feature computational time is only 1 ms. In contrast, the feature computational time of learned features is over 14 times of the handcrafted features (347,186 ms versus 23,994 ms). Additionally, it is impossible to compute a part of learned features. Thus, the feature computational time for the 21, 925 or 4,096 learned features is the same. Although the time of SVM classification based on the computed feature sets is almost the same (3-4 ms) except the one of the 925 learned features (55 ms), the differences in the total classification time (the sum of the feature computational time and the time of SVM classification based on computed features) between those feature sets are significant. It points out that the classification based on EXIF features is very fast (only 4 ms). The classification based on handcrafted features is slower (24 seconds) while the classification with learned features is very slow (approximately 347 seconds) but the accuracy is not increasing in the same proportions so the question is: is the additional computational cost worthy regarding the gain in accuracy? The answer might depend on the considered application.

5 CONCLUSION

In this paper, large field / close-up image classification task is studied and 3 types of features are evaluated in terms of classification accuracy, complexity, running time. It appears that learned features are very powerful for that task, the accuracy of the classification reach to 0.989 with 925 features and 0.975 with only 4 features learned from VGG16 although they are complex and it is not easy to understand them. EXIF features are quite efficient for large filed / close-up image classification since it is possible to obtain the same and quite good classification score by using 4 very simple EXIF features than by using 21 complex handcrafted features. EXIF features are simple, efficient but they are not always available.

References

1. Aydin, T.O., Smolic, A., Gross, M.: Automated aesthetic analysis of photographic images. IEEE Transactions on Visualization and Computer Graphics **21**(1), 31–42 (2015)
2. Bosch, A., Zisserman, A., Munoz, X.: Image classification using random forests and ferns. In: 2007 IEEE 11th International Conference on Computer Vision. pp. 1–8 (Oct 2007)
3. Boutell, M., Luo, J.: Photo classification by integrating image content and camera metadata. Proceedings - International Conference on Pattern Recognition **4**(January), 901–904 (2004)
4. Boutell, M., Luo, J.: Beyond pixels: Exploiting camera metadata for photo classification. Pattern Recognition **38**(6), 935 – 946 (2005), image Understanding for Photographs
5. Datta, R., Joshi, D., Li, J., Ze Wang, J.: Studying aesthetics in photographic images using a computational approach. Proceedings of the 9th European Conference on Computer Vision, (Graz) **3**, 288–301 (05 2006)
6. Guo, T., Dong, J., Li, H., Gao, Y.: Simple convolutional neural network on image classification. In: 2017 IEEE 2nd International Conference on Big Data Analysis. pp. 721–724 (March 2017)
7. He, S., Xu, C., Guo, T., Xu, C., Tao, D.: Reinforced multi-label image classification by exploring curriculum. In: The Thirty-Second AAAI Conference on Artificial Intelligence (2018)
8. Hiscocks, P.D., Syscomp, P.E.: Measuring luminance with a digital camera. In: Syscomp Electronic Design Limited 2011 (2011), http://www.atecorp.com/ATECorp/media/pdfs/data-sheets/Tektronix-J16_Application.pdf

9. Huang, H., Chen, Y., Chen, S.: Copyright protection for images with exif metadata. In: 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing. pp. 239–242 (Aug 2008)
10. Jin, B., Segovia, M.V.O., Susstrunk, S.: Image aesthetic predictors based on weighted CNNs. In: Proceedings - International Conference on Image Processing, ICIP. vol. 2016-Augus, pp. 2291–2295 (2016)
11. Kao, Y., Huang, K., Maybank, S.: Hierarchical aesthetic quality assessment using deep convolutional neural networks. Signal Processing: Image Communication **47**, 500 – 510 (2016)
12. Ke, Y., Tang, X., Jing, F.: The design of high-level features for photo quality assessment. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition **1**, 419–426 (2006)
13. Kira, K., Rendell, L.A.: A practical approach to feature selection. In: Proceedings of the Ninth International Workshop on Machine Learning. pp. 249–256. ML92, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1992)
14. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (May 2017)
15. Li, Z., Fan, J.: Exploit camera metadata for enhancing interesting region detection and photo retrieval. Multimedia Tools Appl. **46**(2-3), 207–233 (Jan 2010)
16. Lu, D., Weng, Q.: A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing **28**(5), 823–870 (2007)
17. Luo, Y., Tang, X.: Photo and Video Quality Evaluation: Focusing on the subject. Proceedings of the 10th European Conference on Computer Vision **8**(08), 386–399 (2008)
18. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering **22**(10), 1345–1359 (Oct 2010)
19. Rawat, W., Wang, Z.: Deep convolutional neural networks for image classification: A comprehensive review. Neural Computation **29**(9), 2352–2449 (2017)
20. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings (2015)
21. Talebi, H., Milanfar, P.: Nima: Neural image assessment. IEEE Transactions on Image Processing **27**, 3998–4011 (2018)
22. Tang, X., Luo, W., Wang, X.: Content-based photo quality assessment. IEEE Transactions on Multimedia **15**(8), 1930–1943 (2013)
23. Technical Standardization Committee on AV & IT Storage Systems and Equipment: Exchangeable image file format for digital still cameras: Exif Version 2.2. Tech. Rep. JEITA CP-3451 (April 2002)
24. Tong, S., Loh, Y.P., Liang, X., Kumada, T.: Visual attention inspired distant view and close-up view classification. In: 2016 IEEE International Conference on Image Processing (ICIP). pp. 2787–2791 (Sept 2016)
25. Vailaya, a., Figueiredo, M., Jain, a., Zhang, H.J.Z.H.J.: Content-based hierarchical classification of vacation images. Proceedings IEEE International Conference on Multimedia Computing and Systems **1**(2), 518–523 (1999)
26. Wang, Y., Hu, B.G.: Hierarchical Image Classification Using Support Vector Machines. Asian Conference on Computer Vision (January), 23–25 (2002)
27. Zhuang, C., Ma, Q., Liang, X., Yoshikawa, M.: Anaba: An obscure sightseeing spots discovering system. 2014 IEEE International Conference on Multimedia and Expo (ICME) pp. 1–6 (July 2014)

References

1. Aydin, T.O., Smolic, A., Gross, M.: Automated aesthetic analysis of photographic images. IEEE Transactions on Visualization and Computer Graphics **21**(1), 31–42 (2015)
2. Bosch, A., Zisserman, A., Munoz, X.: Image classification using random forests and ferns. In: 2007 IEEE 11th International Conference on Computer Vision. pp. 1–8 (Oct 2007)
3. Boutell, M., Luo, J.: Photo classification by integrating image content and camera metadata. Proceedings - International Conference on Pattern Recognition **4**(January), 901–904 (2004)
4. Boutell, M., Luo, J.: Beyond pixels: Exploiting camera metadata for photo classification. Pattern Recognition **38**(6), 935 – 946 (2005), image Understanding for Photographs
5. Datta, R., Joshi, D., Li, J., Ze Wang, J.: Studying aesthetics in photographic images using a computational approach. Proceedings of the 9th European Conference on Computer Vision, (Graz) **3**, 288–301 (05 2006)
6. Guo, T., Dong, J., Li, H., Gao, Y.: Simple convolutional neural network on image classification. In: 2017 IEEE 2nd International Conference on Big Data Analysis. pp. 721–724 (March 2017)
7. He, S., Xu, C., Guo, T., Xu, C., Tao, D.: Reinforced multi-label image classification by exploring curriculum. In: The Thirty-Second AAAI Conference on Artificial Intelligence (2018)
8. Hiscocks, P.D., Syscomp, P.E.: Measuring luminance with a digital camera. In: Syscomp Electronic Design Limited 2011 (2011), http://www.atecorp.com/ATECorp/media/pdfs/data-sheets/Tektronix-J16_Application.pdf
9. Huang, H., Chen, Y., Chen, S.: Copyright protection for images with exif metadata. In: 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing. pp. 239–242 (Aug 2008)
10. Jin, B., Segovia, M.V.O., Susstrunk, S.: Image aesthetic predictors based on weighted CNNs. In: Proceedings - International Conference on Image Processing, ICIP. vol. 2016-Augus, pp. 2291–2295 (2016)
11. Kao, Y., Huang, K., Maybank, S.: Hierarchical aesthetic quality assessment using deep convolutional neural networks. Signal Processing: Image Communication **47**, 500 – 510 (2016)
12. Ke, Y., Tang, X., Jing, F.: The design of high-level features for photo quality assessment. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition **1**, 419–426 (2006)
13. Kira, K., Rendell, L.A.: A practical approach to feature selection. In: Proceedings of the Ninth International Workshop on Machine Learning. pp. 249–256. ML92, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (1992)
14. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Commun. ACM **60**(6), 84–90 (May 2017)
15. Li, Z., Fan, J.: Exploit camera metadata for enhancing interesting region detection and photo retrieval. Multimedia Tools Appl. **46**(2-3), 207–233 (Jan 2010)
16. Lu, D., Weng, Q.: A survey of image classification methods and techniques for improving classification performance. International Journal of Remote Sensing **28**(5), 823–870 (2007)

17. Luo, Y., Tang, X.: Photo and Video Quality Evaluation: Focusing on the subject. Proceedings of the 10th European Conference on Computer Vision **8**(08), 386–399 (2008)
18. Pan, S.J., Yang, Q.: A survey on transfer learning. IEEE Transactions on Knowledge and Data Engineering **22**(10), 1345–1359 (Oct 2010)
19. Rawat, W., Wang, Z.: Deep convolutional neural networks for image classification: A comprehensive review. Neural Computation **29**(9), 2352–2449 (2017)
20. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings (2015)
21. Talebi, H., Milanfar, P.: Nima: Neural image assessment. IEEE Transactions on Image Processing **27**, 3998–4011 (2018)
22. Tang, X., Luo, W., Wang, X.: Content-based photo quality assessment. IEEE Transactions on Multimedia **15**(8), 1930–1943 (2013)
23. Technical Standardization Committee on AV & IT Storage Systems and Equipment: Exchangeable image file format for digital still cameras: Exif Version 2.2. Tech. Rep. JEITA CP-3451 (April 2002)
24. Tong, S., Loh, Y.P., Liang, X., Kumada, T.: Visual attention inspired distant view and close-up view classification. In: 2016 IEEE International Conference on Image Processing (ICIP). pp. 2787–2791 (Sept 2016)
25. Vailaya, a., Figueiredo, M., Jain, a., Zhang, H.J.Z.H.J.: Content-based hierarchical classification of vacation images. Proceedings IEEE International Conference on Multimedia Computing and Systems **1**(2), 518–523 (1999)
26. Wang, Y., Hu, B.G.: Hierarchical Image Classification Using Support Vector Machines. Asian Conference on Computer Vision (January), 23–25 (2002)
27. Zhuang, C., Ma, Q., Liang, X., Yoshikawa, M.: Anaba: An obscure sightseeing spots discovering system. 2014 IEEE International Conference on Multimedia and Expo (ICME) pp. 1–6 (July 2014)