Proceedings of the 53rd Hawaii International Conference on System Sciences | 2020

Understanding Customer Preferences Using Image Classification – A Case Study of an Online Travel Community

Ines Brusch
Brandenburg University of Technology Cottbus-Senftenberg
ines.brusch@b-tu.de

Abstract

Today, companies have a large amount of data at their disposal. In addition to classic data in text or table form, the number of images also increases enormously. This is particularly the case if the customer contact exists via the Internet, e.g., social networks, blogs or forums. If these images can be evaluated, they lead to a better understanding of the customer. Improved recommendations can be made and customer satisfaction can be increased. This paper shows by means of support vector machines (SVM), convolutional neural networks (CNN) and cluster analyses how it is possible for companies to evaluate image data on their own and thus to understand and classify the customer. The data of travel platform users serve as a case study. Advantages and disadvantages of, as well as prerequisites for SVMs and CNNs are pointed out and segmentation of the users on the basis of their images is made.

1. Introduction

Today, the Internet is available almost everywhere and offers an almost inexhaustible knowledge and information base. Starting from a pure information medium, the Internet has developed into an exchange platform in every respect. On the one hand, it is possible to buy almost any product or service via the Internet, on the other hand, one can exchange information about any interests and news. This information is of course also accessible to companies and can be evaluated using various algorithms.
Photography is another area that has increased in recent years. Within three years, the number of images increased from 660 billion to 1,200 billion in 2017 [36]. And the number of images that are posted online is also increasing. Today, photography has become part of our lives, including sharing images on social networking sites, instantly sharing images on smartphones, using images in blogs and forums, printing images, and more. In particular, the rise of the smartphone as the dominant camera and the expansion of the mobile Internet have driven this development. The images can also be evaluated by companies and provide important information about the activities, interests, and opinions of users and customers.
Since many images are taken and published online during holidays, in particular, a travel platform should serve as a use case for the analysis of images to identify user preferences. This use case was chosen because travel portals such as Holdaycheck, Tripadvisor or Travelfriends are very popular. Tripadvisor alone had 490 million unique users per month in 2018 [41]. Travelers not only inform themselves about travel destinations and insider tips before and during their holidays, they also exchange their experiences and aims. Of course, they also upload images that illustrate their activities and preferences.
In research on online travel communities, there are often questions about the reasons and motivations for the use [5, 45], positive and negative effects of word-of-mouth and its influence on customer loyalty [23, 32]. There are also some studies on the importance of images in travel communities. For example, some authors show that images or social media content can be used as a cost-efficient alternative to surveys to draw conclusions about user preferences [13, 19]. Further studies also deal with the contents of the images to illustrate activities and experiences [7, 16]. However, automatic recognition of the image content does not take place.
This is where this paper comes in and pursues the following research questions:

1. Is it possible for a company without special software to automatically recognize and categorize image content from everyday images?
2. In terms of the use case, are travel styles predictable based on images?

If these questions can be answered in the affirmative, further and improved recommendations can be made for users' next journeys. In this way, customer satisfaction can be further increased. This, in turn, leads to further customer recommendations, reuse

URI: https://hdl.handle.net/10125/63858
978-0-9981331-3-3

and additional usage, which leads to increased customer loyalty and higher profits. The question is also interesting from a market segmentation point of view. Companies have always tried to divide their customers into groups in order to ensure an optimal customer approach. However, this only works to a limited extent or is very time-consuming and cost-intensive. This paper shows an approach to automatically analyze the data provided by the customer and use it to improve customer communication.

First, a discussion of image analysis in general takes place. This is followed by the case study. The case study is divided into two parts. With the help of image analysis methods, such as support vector machines or neural networks, the first step is to categorize the images of the users of an online travel portal and to show which method is best suited for many different images and categories and can be easily implemented by companies. In a second step, the data of the real users and their images were collected. The captured images are automatically classified according to different image categories, e.g. food and beverages, historical sites or mountain panoramas, and segmented using cluster analysis. The clusters are then compared with the holiday styles specified by the user. A final conclusion and outlook follow the results of the investigation.

2. Classification of Images

The content-based evaluation of images has a long tradition. Only a few years ago it was standard to evaluate images via their low-level features, e.g. colors, textures or shapes. Especially color histograms enjoyed great popularity. A big advantage of the color histograms lies in the fact that they are both rotation- and translation-invariant and also robust against the scaling of images [12, 38]. Furthermore, color histograms can be created with little computational effort and require little memory capacity [12, 31]. Color histograms also have the advantage that different objects often generate characteristic color histograms [31]. A disadvantage, however, is that color histograms are very strongly influenced by changes in lighting conditions. That is because the color of the image also changes as the illumination changes. Two color histograms of the same object under different lighting conditions thus produce different histograms [28, 31]. However, there are numerous fields of application in which such an image analysis and classification of histograms seems reasonable and sufficient.

This classification can be done with support vector machines (SVM [43]), for example. Today SVM have strong theoretical foundations and a wide area of applications, e.g. medical science (e.g. [6, 30]), text categorization (e.g. [20, 40]) or image classification (e.g. [8, 48]). A good overview about the usage of SVM is also given in [27]. The starting point of SVM is a set of training objects, for which you know the class to which they belong. Each object is represented by a vector in the vector space. The task of the SVM is to fit a Hyperlayer into this space, which acts as a separating area and divides the training objects into two classes. The distance between the vectors which are closest to the layer is maximized. This wide, empty border will later ensure that even objects that do not correspond exactly to the training objects are classified as accurately as possible [8]. Especially for small training samples, SVMs were regarded as an efficient and stable method and as a positive by-product the high classification speed is to be mentioned.

Artificial neural networks provide a further alternative for image classification. In 1943, artificial neural networks were used by McCulloch and Pitts. They used simple neural networks to generate Boolean functions AND, OR and NOR and their combinations. Their hypothetical nerve cells had only two possible outputs: on or off. Whether they became active depended on whether the inputs from other neurons exceeded a certain threshold value [25]. Even today, all artificial neural networks are based on this threshold logic - with a few variations. Artificial neural networks consist of artificial neurons that weight inputs and generate output via an activation function [47]. Introductions to the topic can be found in [1], [4], or [18]. One of the areas of application that particularly benefit from the innovations in the development of artificial neural networks is image recognition [15]. Basically, the classes of networks differ mainly in the different network topologies and connection types, like single-layer feedforward networks, multi-layer feedforward networks, backpropagation networks and networks with direct and indirect feedback and networks with lateral feedback and lattice structures.

In recent years, convolutional neural networks, in particular, have experienced a renaissance in image analysis. Convolutional Neural Networks (CNN [22]) are particularly suitable for image processing. By using GPUs, CNNs have a renaissance. Typical CNNs use 5 to 25 different layers for pattern recognition. CNNs extract localized features from input images and use filters to unfold these image fields. CNNs are state of the art in tasks of image classification [11, 21, 34] or object detection [17, 33].

The main advantage of CNN is that the learning of distributed representations allows the generalization to new combinations, which go beyond the characteristics learned during the training [3, 26]. Relevant features are automatically extracted from images and the task is completed automatically through the learning process. However, CNNs have also disadvantages. They require a lot of data (thousands of images) to train the model and produce high computational cost to process the data quickly (need of GPU). In the following, a case study will be carried out, which compares the different methods of image analysis and shows their possibilities.

3. Analyzing Customer Preferences

3.1. Image Categorization

Method. As a company, you have a multitude of different data on your consumers and interested persons available, including images. These can be determined e.g. in own databases, on the social media presence or by online queries. Since it is not efficient to evaluate the images manually, companies need methods that support them. Since many companies have neither the know-how nor the technology for highly specialized applications, the analyses should be applicable with minimal training and without additional technical equipment. In the present use case, no expensive special software and only a standard desktop PC (Intel Core i7 - 2600K CPU@ 3.40 GHz; 16.0 GB RAM) are used, so that the evaluation would be possible for every company. The evaluation is carried out using the free statistics software R.

This first step of the analysis serves to answer the first research question, which is as follows: Is it possible for a company without special software to automatically recognize and categorize image content from everyday images? For the case study, the data of 26 users of an online travel community and their 2,333 images were collected. Within their profile, users could select 19 travel styles to describe themselves. These were, for example, nature lover, beach goer, city explorer or a fan of peace and relaxation. Several styles can also be selected. Furthermore, they could rate hotels, places of interest and other points of interest, create a travel map, exchange ideas in the forum or post images of their travels. On average, the considered users had visited 65 cities all over the world and gave an average of 119 ratings on various sights, hotels, restaurants or other activities. On average, users uploaded 80.5 images (minimum 11 and maximum 333). Figure 1 shows two sample users and their information.

User name	*M*53** 	*o*nn**L 
Sociodemographic data	35-49 years old, man, Swiss	35-49 years old, woman, German
Travel information	32 cities visited, 64 reviews, 144 photos	154 cities visited, 14 reviews, 21 photos
Travel style (Top 4)	Nature lover (1), beach goer (2), city explorer (3), luxury traveler (4)	City explorer (1), nature lover (2), fan of peace and relaxation (3), beach goer (4)
Images		

Figure 1. Sample Users of the Travel Community

Since the taken images were as varied as the journeys of their photographers, it was first necessary to define main categories into which the images could be divided. A total of 18 categories were defined and a training dataset has been created. This training dataset includes images from the main categories of the SUN2012 database [46], a food image dataset [35] and five holiday-specific categories. In total, the training database contains 2,470 images and is needed to train an SVM and a CNN.

In order to answer the question of whether the images of tourists can be categorized automatically, four different methods were tested and compared:

1. An SVM using a combination of the low level features BIC (Border/ Interior Pixel Classification [37]), CEDD (Color and Edge Directivity Descriptor [9]) and FCTH (Fuzzy Color and Texture Histogram [10]). In [14], this combination turned out to be very promising for the analysis of holiday images.
2. A CNN consisting of 12 layers and 7.9 million trainable parameters.
3. The IMADAC software [2], which performs a cluster analysis using the Ward method based on various low-level features to categorize the images. As with SVMs, the features BIC, CEDD, and FCTH are used and weighted equally in this study.
4. Google's Inception.v3 [39], a CNN pre-trained in 1,000 categories with 1.2 million images

To evaluate the results of the different methods, there are different statistical quality criteria of classification, which calculated with the help of a confusion matrix. In total four results are possible: True positive (TP) = an image is part of a category and the test has notified this correctly; false negative (FN) = an image is part of a category and the test has not notified this; false positive (FP) = an image is not part of a category, but the test has notified it to the category; true negative (TN) = an image is not part of a category and the test has notified this correctly. Based on this matrix the following typical statistical quality criteria of classification can be used: Precision, Recall, Accuracy, and F-measure. The formulas are given in Figure 2 [29, 42].

$$P = \frac{TP}{TP + FP} \qquad R = \frac{TP}{TP + FN}$$

$$A = \frac{TP + TN}{TP + TN + FP + FN} \qquad F_{\propto} = \frac{P * R}{(1 - \propto)P + \propto R}$$

Figure 2. Quality Criteria of Classification

Principally accuracy is a good measure to evaluate the proportion of correctly classified items. However, accuracy conflates the performance on relevant images (TP) with the performance of irrelevant images (TN). Because of this, the F-measure is also used. The F-measure is the weighted harmonic mean of P and R. A high alpha value causes high importance of precision and a low causes a high importance of recall. In this study, an alpha value of 0.5 is assumed. This is the best compromise between P and R. With this weighting it can be assumed that a maximum F-measure between 0.4 and 0.9 can be achieved.

Results. As described in the previous section, the user images have now been categorized using the methods SVM, CNN, IMADAC, and Inception.v3. At first, the 10 most frequent image categories of the users were analyzed. These categories accounted for 95% of the images. Table 1 shows the results of the Accuracy (A) and F measure (F) calculations as well as the number of training images and test images used in every category. No training data is required for the IMADAC and Inception.v3 methods. For Inception.v3, the image database used as the basis for learning is already online. IMADAC clusters the test images based on their low-level features, no prior learning is required. In addition, the analysis is carried out for only three categories. This serves to show the strengths and weaknesses of the individual analysis methods. Furthermore, it sometimes makes sense for companies to identify only a few main categories in order to bundle capacities. The process times are also given for all calculations. These serve to weigh up the costs and benefits of the individual steps.

Table 1. Results of Content-Based Image Analysis of 3 and 10 Holiday Categories Using Accuracy (A) and F-measure (F)

	Mountains & desert	House (indoor) & Hotel	Houses (outdoor) & gardens	Water & snow	Animals	Historical places	Shopping	Cultural	Selfies & persons	Food & drinks
Train	371	201	285	457	40	454	170	239	132	121
Test	108	329	395	220	87	221	124	199	102	213
SVM 3 categories (Proc. time: 1.32 sec + 1.02 min)										
F	0.57	0.86	0.70							
A	0.84	0.88	0.76							
SVM 10 categories (Proc. time: 13.77 sec + 2.45 min)										
F	0.31	0.60	0.12	0.48	-	0.43	0.23	0.34	0.11	0,53
A	0.92	0.87	0.81	0.83	0.96	0.82	0.89	0.85	0.91	0,92
CNN 3 categories (Processing time: 26.38 min)										
F	0.53	0.77	0.68							
A	0.83	0.82	0.73							
CNN 10 categories (Proc. time: 74.38 min)										
F	0.25	0.32	0.22	0.37	0.02	0.23	0.14	0.15	0.06	0,03
A	0.88	0.77	0.73	0.83	0.95	0.83	0.86	0.81	0.89	0,89
IMADAC 3 categories (Proc. time: 11 sec + 51 sec)										
F	0.45	0.67	0.54							
A	0.79	0.73	0.63							
IMADAC 10 categories (Proc. time: 32 sec + 1.50 min)										
F	0.20	0.26	0.23	0.38	0.19	0.24	0.25	0.23	0.09	0,34
A	0.92	0.73	0.76	0.88	0.90	0.84	0.86	0.90	0.84	0,86
Inception.v3 3 categories (Proc. time: 4.78 min)										
F	0.43	0.76	0.53							
A	0.81	0.81	0.72							
Inception.v3 10 categories (Proc. time: 4.78 min)										
F	0.64	0.79	0.58	0.71	0.79	0.65	0.64	0.60	0.36	0.87
A	0.96	0.93	0.85	0.94	0.98	0.92	0.95	0.93	0.95	0.97

Looking at the 3 categories, it can be seen that a simple SVM achieves the best results. Both Accuracy and F-measure are the best in all three categories. Then follows CNN, Inception.v3 and IMADAC. In addition, the SVM impresses with its speed, the calculation of the categories is completed within 1.32

seconds. 1.02 minutes are needed to extract the image features. However, the feature extraction only needs to be done once. The saved features can then be used for further calculations. The CNN programmed in R has clear weaknesses here, even if the results are promising. Here the process takes over 26 minutes and only 857 training images. IMADAC also convinces with a fast process, but shows only mediocre results. Inception.v3 cannot demonstrate its strength in a few main categories. It is trained to recognize individual objects, which cannot be assigned to the 3 main categories without manual intervention. Here a manual rework would be necessary.

It can be seen that Inception.v3 achieves the best results across all categories. Especially in the category food & drinks Inception.v3 achieved excellent results. 88% of all images in this category were recognized correctly. If you draw the pictures that do not belong to the category, 99.9% of the pictures are categorized correctly in relation to the category food and drinks. Inception.v3 also achieved very good results in the category of animals. Here SVM and CNN had the most problems and did not recognize the animals. The problem lies mainly in the small training data set. This problem does not exist with Inception.v3. Inception.v3 is trained on 1.2 million images, which led to these very good results. The CNN programmed here took 74.38 minutes to train the categories at 60 iterations. Confirming various publications, it can be stated that CNNs require an enormously large training data set and GPU-based calculations to be promising. In such small training data sets as used in this paper, SVMs show better results than CNN. Thus, Table 1 shows that the SVM performs better in all categories. In addition, it convinces again by a very fast process time. Moreover, if there is also a limited number of objectives (e.g. the comparison of interior and exterior images), SVMs achieve satisfactory results. It would then make sense to optimize these with regard to the different categories by adjusting and enlarging the low-level features. IMADAC also convinces with its process time and shows better results than the programmed CNN. However, the inclusion of further low-level features to improve the results should be examined.

In conclusion, it can be said that pre-trained convolutional neuronal networks are a very good and easily implementable way to categorize a wide variety of images available in companies. However, developing one's own CNNs requires a very large training database and a very high level of computational effort, which only makes sense for specialized applications. Classic SVMs are a good alternative for smaller image collections with few categories. When expanding the training database, it is conceivable that both SVM and CNN will achieve better results. There may also be distortions between the training data set and the test data set. For example, the training data set for animals could consist of many domestic animals and few wild animals, but the vacation pictures from the test data set could consist of a large number of wild animals. Adjustments would be useful here. In addition, it would be possible to integrate further features into the analysis in order to improve the results. In order to optimize CNN, further tests would be useful regarding the number of layers and optimization settings.

At the end of the first test, Table 2 shows the results of Inception.v3 across all 18 image categories. Even the high number of nonspecialized categories shows very good results. It should be mentioned once again that the aim of the analysis was not to detect objects but to classify images from the everyday context.

Table 2. Accuracy (A) and F-measure (F) of 18 Holiday Categories Using Inception.v3

	Number of images in the dataset	Inception.v3	
		F	A
Mountains & Desert	108	0,58	0,96
House (indoor) & Hotel	329	0,78	0,93
Houses (outdoor) & gardens	395	0,53	0,84
Water & snow	220	0,66	0,93
Animals	87	0,72	0,97
Historical places	221	0,63	0,92
Shopping	124	0,58	0,94
Cultural	199	0,58	0,94
Selfies & Person	102	0,35	0,96
Food & drinks	213	0,84	0,97
Forest & field	55	0,32	0,98
Flowers	41	0,22	0,98
Transport	27	0,38	0,98
Parks	74	0,17	0,97
Commercial markets	0	-	1,00
Sports & leisure	40	-	0,98
Night	6	0,13	0,99
Signs & writings	67	0,16	0,97

The first research question can therefore be answered as follows:

- Yes, it is possible for a company to automatically categorize and recognize everyday images based on their content without special software.
- However, there is no universal method that allows a perfect automated analysis of such complex problems, where characteristic objects as well as scenes have to be recognized.

However, it is conceivable that with a sufficiently large training database and high computing power a CNN could be trained for this question. Another possibility is to use methods for multi-label analysis (e.g. [44]), because images usually have several labels that are relevant for the context. This could extract larger semantic information, which would then have to be assigned to the different categories. Overall, however, Inception.v3 showed the best results, so that this neural network was used as the basis for the next step of the investigation.

3.2. User classification

Method. The second step now serves to clarify the second research question, which is as follows: Are travel styles predictable based on the images? To this end, the user database was initially expanded to 80 users, their data and travel styles recorded, and their 6,919 images analyzed. Inception.v3 was used for categorization. This procedure turned out to be the best in the first step and can also be carried out in a relatively short time using a standard computer. Afterward the category shares of the images per user were computed and standardized by means of z-transformation. In this way, better comparability between different image quantities can be guaranteed. In total, users of the platform were able to choose between 19 travel styles. In the following analyses, the focus is on the eight most frequently chosen travel styles. They were selected by at least 25 of the 80 users.

Results. T-tests (Table 3) were first used to check whether there were significant differences between the members and non-members of a travel style. However, there were few significant differences, some of them questionable. For example, the t-test showed significant differences between nature lovers and non-nature lovers in the category signs & writings (p=0.03). However, the image category generally contained very few images, so that significant differences between few and very few images are questionable. However, some results were understandable. For example, history lovers take significantly more images of historical places than non-history lovers (p=0.01). Beach goers take significantly more images of houses and hotels (indoor) than non-beach goers (p=0.02). This can be explained by the fact that beach goers are often less active on holiday and therefore have fewer opportunities to take images of places of interest or excursions. Overall, it is not possible to distinguish the members of a travel style from the non-members of a travel style by the differences of the individual image categories. The travel styles gourmet and fan of peace and relaxation did not show any significant difference in the image categories of their members and non-members.

Table 3. T-Tests Between Members and Non-members of a Travel Style (Note: t test with statistical significance at a level of *: p<0.05, **: p<0.01, ***: p<0.001)

	Experience like a native	Fan of peace and relaxation	Gourmets	History lovers	Lovers of art and architecture	Nature lover	City explorer	Beach goers
Mountains & Desert	.49	.19	.21	.01 *	.36	.57	.51	.39
House (indoor) & Hotel	.03 *	.43	.68	.04 *	.50	.39	.15	.02 *
Houses (outdoor) & gardens	.64	.71	.12	.60	.10	.54	.79	.42
Water & snow	.57	.05	.30	.89	.00 ***	.04 *	.22	.38
Animals	.24	.82	.28	.86	.67	.45	.14	.86
Historical places	.85	.57	.42	.01 *	.68	.49	.64	.20
Shopping	.14	.38	.36	.80	.86	.99	.27	.30
Cultural	.08	.75	.60	.13	.39	.18	.58	.84
Selfies & Person	.42	.58	.69	.46	.58	.85	.38	.01 *
Food & drinks	.91	.22	.05	.00 ***	.13	.48	.91	.91
Forest & field	.82	.15	.87	.89	.79	.26	.41	.29
Flowers	.51	.46	.20	.58	.61	.76	.75	.57
Transport	.62	.35	.57	.30	.21	.88	.04 *	.49
Parks	.54	.33	.26	.39	.51	.82	.80	.10
Sports & leisure	.93	.17	.89	.10	.79	.45	.90	.21
Signs & writings	.70	.29	.40	.72	.15	.03 *	.47	.01 *

Because T-tests did not yield satisfactory results, a hierarchical cluster analysis is performed as follows. The squared euclidean distance was used as distance measure and the Ward method was used as the segmentation algorithm. The elbow criterion was used to determine the optimal number of clusters, this was five clusters. Figure 3 displays the distribution in form of a dendrogram. It can be seen that a very large cluster was created, as well as a medium and three smaller clusters. A comparison with Figure 1 shows that although both users specify three identical travel styles (nature lover, beach goer, city explorer), they also specify a different one (luxury traveler vs. fan of peace and relaxation) and a different order of styles. This suggests that the two users have some different interests. As the dendrogram shows, they are also segmented into different clusters with regard to their images.

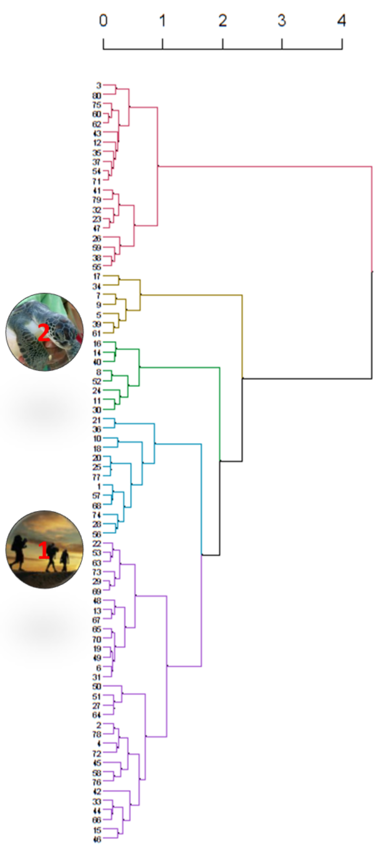

Figure 3. Dendrogram of the Hierarchical Cluster Analysis Including Sample Users

Table 4 shows the image distributions within the clusters and the distribution of holiday styles. This ratio was calculated with the following formula:

$$\frac{\frac{no.travel\ style\ users\ (i)\ in\ cluster/4}{no.users\ in\ cluster}}{\sum_{i=1}^{8}\frac{no.travel\ style\ users\ (i)\ in\ cluster/4}{no.users\ in\ cluster}} * \frac{5}{no.travel\ style\ users\ (i)}$$

Table 4. Characteristics of the Clusters

	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5
Number of users	13	32	20	7	8
Characteristics (mean of z-transformation values)					
Mountains & Desert	-0.301	0.456	-0.313	-0.610	-0.016
House (indoor) & Hotel	-0.377	-0.048	-0.476	2.541	-0.229
Houses (outdoor) & gardens	-0.336	0.538	-0.480	-0.651	0.163
Water & snow	-0.425	-0.030	-0.487	-0.448	2.420
Animals	-0.087	0.378	-0.247	-0.653	-0.181
Historical places	1.428	-0.079	-0.272	-0.618	-0.784
Shopping	-0.322	0.067	-0.055	0.148	0.263
Cultural	1.609	-0.339	-0.141	-0.402	-0.557
Selfies & Person	0.180	0.125	-0.106	-0.460	-0.125
Food & drinks	-0.725	-0.482	1.443	-0.017	-0.489
Forest & field	-0.047	0.279	0.011	-0.569	-0.569
Flowers	-0.362	0.248	0.009	-0.306	-0.159
Transport	-0.289	0.420	-0.083	-0.483	-0.578
Parks	-0.275	-0.023	0.008	-0.404	0.873
Sports & leisure	0.289	0.047	-0.021	-0.241	-0.392
Signs & writings	0.511	-0.277	0.036	0.607	-0.342
Travel style					
Experience like a native	0.010	0.029	0.023	0.040	0.000
Fan of peace and relaxation	0.013	0.034	0.012	0.000	0.043
Gourmet	0.019	0.018	0.028	0.037	0.013
History lover	0.040	0.022	0.017	0.000	0.023
Lovers of art and architecture	0.024	0.023	0.031	0.012	0.000
Nature lover	0.024	0.021	0.020	0.016	0.034
City explorer	0.026	0.021	0.023	0.015	0.026
Beach goer	0.021	0.015	0.019	0.050	0.035

In cluster 1, most users are history lovers, city explorers, and lovers of art and architecture. An outstanding number of images of historical places as well as cultural objects characterizes the cluster. There is also a below-average number of images of food & drinks, and water & snow. Thus, it can be seen that the images of cluster 1 fit very well to the most common travel styles of this cluster.

In cluster 2 most users are from the categories experience like a native, fan of peace and relaxation, lovers of art and architecture, and nature lover. An outstanding number of images of houses & gardens, as well as mountains & desert, characterizes the cluster. There is also a below-average number of images of food & drinks, as well as cultural objects. Here the architecture lovers of contemporary buildings, as well as people seeking peace and relaxation in the mountains, can be found.

Cluster 3 is characterized primarily by users of travel styles experience like a native, gourmet, lovers of art and architecture and city explorer. Their images are characterized by an outstanding number of images of food & drinks, as well as forest & field. This cluster also has the highest proportion of images showing transportation and the second highest proportion of cultural images. Images of houses & gardens, as well as water & snow, are very rare. Here you will find the gourmets who report on their food and drinks, as well as the lovers of art and architecture who are more interested in works of art.

Cluster 4 unites the beach goers, experience like a native and gourmets. The images of the cluster are characterized by the themes house and hotel, as well as shopping. Houses and gardens, as well as animals are very rare on images of this cluster. All users of the cluster call themselves gourmets or beach goers. Especially with beach goers, it seems plausible that they have a high number of interior shots of the hotel. They will more often relax on the beach than take images of activities and sights. This category of gourmets suggests that they enjoy their food rather than photograph it. Shopping also includes grocery stores and markets. Therefore, it could be that these gourmets prepare their own food.

Finally, cluster 5 mainly includes users from the categories fan of peace and relaxation, beach goer and nature lover. Their images show water & snow as well as parks most frequently. Both are perfectly suited to the three categories. Transportation and historical places are rarely motifs of the images. These users love to be at the seaside, whether as nature lovers, beach goers or to find peace and relaxation.

Overall, it can be seen that the three most common travel styles in the clusters are described very well by the images of their users. Table 5 shows the correspondence between the clusters and the self-assessment of the users, i.e. the hit rate of the segmentation. If only the first choice of users is considered, 58.8% of the self-selected number one is already assigned. However, the users do not need to classify the travel styles according to their preferences, it is also possible to do this alphabetically or according to spontaneous associations. In addition, most travelers are not limited to one travel style; for example, travelers take a city trip in spring and a beach holiday in summer. Besides, overlaps in travel styles should not be ignored. It is often the case that people who make many city trips are also interested in architecture. Therefore, it makes sense to consider the other three chosen travel styles as well. If the second choice is added, the segmentation result rises to 76.3% and with the third choice to 88.8%. One of the user's first four preferences is met in 92.5% of all cases. This is a good result, only 7.5% of the users are assigned to a cluster whose focus they do not correspond to. This raises the question of whether users have misjudged their travel styles or whether a certain degree of variation in the images is permissible. Finally, users do not upload all their holiday images but make a pre-selection that matches their assessments of activities, accommodation or destinations.

Table 5. Hits and Hit Rates of the User Segmentation Within the Clusters

User selected travel style	Clu-ster 1	Clu-ster 2	Clu-ster 3	Clu-ster 4	Clu-ster 5	Hit	Hit Rate in %
	History & City & Art	Peace & Native & Art & Nature	Art & Gourmet & Native & City	Beach &Native & Gourmet	Peace & Beach & Nature		
First choice	8	14	17	5	3	44	58.8
Second choice	9	23	17	6	6	62	76.3
Third choice	11	28	18	7	7	72	88.8
Fourth choice	12	29	18	7	8	75	92.5

In clusters 4 and 5, a total hit rate of 100% was achieved. In the worst case, the hit rate was 90.6% in cluster 2

Overall, the second research question can be answered as follows:

- Travel styles are predictable on the basis of user images
- Cluster analysis offers in this case a very good possibility for user segmentation

The analysis of contents of images can be considered as a further and good possibility to understand users of travel platforms and to make them targeted offers regarding new destinations.

4. Conclusion and Outlook

The present work deals with the question of whether it is possible for a company to automatically classify a multitude of different images and to predict user preferences based on them. A travel platform serves as an application example, where users can comment on their travels, exchange information and show images of their journeys. If the prediction of customer preferences is successful via image analysis, for example, the customer satisfaction of the platform can be increased, which leads to increased customer loyalty. The question is also interesting from a market segmentation point of view. If existing data can be evaluated automatically and customer segments can be formed, an improved customer approach is possible. This saves time and money compared to the past.

To answer the research question, several analytical methods for image analysis (SVM, CNN, Image Classification Software IMADAC, Inception.v3) were first tested. If only a few image classes are existent, methods like SVM can be used. However, if there are many different image types, deep learning methods like CNN have enormous advantages. The study showed that Inception.v3, in particular, achieved very good results and also recognizes a large number of images in different categories. Thus the first research question can be answered with yes. It is possible for a company without special software to automatically recognize and categorize image content from everyday images. In addition to Inception v3, there are now numerous other pre-trained CNNs that can be used for image analysis. Here a company can integrate context-dependently tailored CNNs. If, for example, a company is interested in the age and gender distribution at events, CNNs such as those implemented in [24] can be used.

In a second step, the images of several real-world users were automatically classified and evaluated using t-tests and hierarchical cluster analysis to predict travel styles and find user segments. T-tests alone were not sufficient to assign users to their travel styles. With the help of cluster analysis, at least one travel style of the user could even be correctly determined in 92.5% of the cases.

However, it must be mentioned that the sample is relatively small with 6,919 images and a larger validation is necessary. Usually companies have user databases of thousands of users and therefore also of countless images. A validation on this scale would be very interesting.

Furthermore, subsequent studies should deal with further algorithms for image analysis. Would generic algorithms or other deep learning methods be better suited? Alternatives to hierarchical cluster analysis, such as k-means or fuzzy clusterwise regression, could also be addressed here.

In addition, users have chosen and limited their travel styles themselves, which may include a bias, e.g. that not all their preferences are covered by the styles or that some preferences were intentionally concealed. Here, subsequent studies should examine what further possibilities there would be for determining a travel style. In addition, users do not upload all holiday images, but only the part that matches their assessments of places and activities. A certain degree of freedom must also be granted here.

It is also interesting to include additional user information, e.g. reviews of the users. The sample user 1 writes e.g. about a lonely camping site. Does this really speak for a luxury traveler? He also gives restaurant reviews, so that he may also be found in the group of gourmets. The destinations themselves can also provide information about the traveler and should be included in the analysis. Here one has to consider how much information is necessary to describe the user in the best possible way.

Future research could also include the labels of images and additional text descriptions, which would make it easier, for example, to recognize a person's attitude. If an image in the ice hockey stadium is for example described with the text “What a great evening” or with “I had to go to the game with my friend”.

The investigation of the influence of images on the customer's choice would also be an interesting research aspect. For example, does it influence a Hawaii traveler's choice of destinations to see a lot of pictures of the Hawaii Volcanoes National Park before the trip begins?

Overall, this paper shows the possibilities of image analysis in the economic context and expands the classical business method canon to a promising method of information technology.

10. References

[1] Aleksander, I. and H. Morton, An introduction to neural computing, Chapman & Hall, London, 1990.
[2] Baier, D., I. Daniel, S. Frost, and R. Naundorf, "Image data analysis and classification in marketing", Advances in Data Analysis and Classification, 6(3), 2012, pp. 253–276.

[3] Bengio, Y., "Learning Deep Architectures for AI", Foundations and Trends® in Machine Learning, 2(1), 2009, pp. 1–127.
[4] Bishop, C.M., Neural networks for pattern recognition, Oxford University Press, 1995.
[5] Bronner, F. and R. de Hoog, "Vacationers and eWOM: Who Posts, and Why, Where, and What?", Journal of Travel Research, 50(1), 2011, pp. 15–26.
[6] Brown, M.P.S., W.N. Grundy, D. Lin, N. Cristianini, C.W. Sugnet, T.S. Furey, M. Ares, and D. Haussler, "Knowledge-based analysis of microarray gene expression data by using support vector machines", Proceedings of the National Academy of Sciences, 97(1), 2000, pp. 262–267.
[7] Cederholm, E.A., "The use of photo-elicitation in tourism research–framing the backpacker experience", Scandinavian Journal of Hospitality and Tourism, 4(3), 2004, pp. 225–241.
[8] Chapelle, O., P. Haffner, and V.N. Vapnik, "Support vector machines for histogram-based image classification", IEEE transactions on neural networks, 10(5), 1999, pp. 1055–1064.
[9] Chatzichristofis, S.A. and Y. Boutalis, "CEDD: Color and Edge Directivity Descriptor: A Compact Descriptor for Image Indexing and Retrieval", in Computer Vision Systems, A. Gasteratos, M. Vincze, and J. Tsotsos, Editors. 2008. Springer Berlin Heidelberg.
[10] Chatzichristofis, S.A. and Y.S. Boutalis, "FCTH: Fuzzy Color and Texture Histogram - A Low Level Feature for Accurate Image Retrieval", in Image Analysis for Multimedia Interactive Services, 2008. WIAMIS '08. Ninth International Workshop on : Image Analysis for Multimedia Interactive Services, 2008. WIAMIS '08. Ninth International Workshop on : Image Analysis for Multimedia Interactive Services, 2008. WIAMIS '08. Ninth International Workshop on. 2008.
[11] Chen, L.-C., G. Papandreou, I. Kokkinos, K. Murphy, and A.L. Yuille, "DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs", IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), 2018, pp. 834–848.
[12] Choraś, R.S., "Image Feature Extraction Techniques and Their Applications for CBIR and Biometrics Systems", International Journal of Biology and Biomedical Engineering, 1(1), 2007, pp. 6–16.
[13] Daniel, I. and D. Baier, "Lifestyle Segmentation Based on Contents of Uploaded Images Versus Ratings of Items", in Algorithms from and for Nature and Life, B. Lausen, D. van den Poel, and A. Ultsch, Editors. 2013. Springer: Heidelberg.
[14] Daniel, I. and D. Baier, "Towards Lifestyle Segmentation via Uploaded Images from Surveys and Social Networks", in ICIS 2015 Proceedings, Fort Worth. 2015.
[15] Egmont-Petersen, M., D. de Ridder, and H. Handels, "Image processing with neural networks—a review", Pattern Recognition, 35(10), 2002, pp. 2279–2301.
[16] Garrod, B., "Understanding the relationship between tourism destination imagery and tourist photography", Journal of Travel Research, 47(3), 2009, pp. 346–358.
[17] Girshick, R., J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014.
[18] Gurney, K., An introduction to neural networks, CRC press, 2014.
[19] Hausmann, A., T. Toivonen, R. Slotow, H. Tenkanen, A. Moilanen, V. Heikinheimo, and E. Di Minin, "Social media data can be used to understand tourists' preferences for nature-based experiences in protected areas", Conservation Letters, 11(1), 2018, e12343.
[20] Joachims, T., "Text categorization with Support Vector Machines: Learning with many relevant features", in Machine learning: ECML-98: Proceedings, C. Nédellec and C. Rouveirol, Editors. 1998. Springer: Berlin [etc.].
[21] Krizhevsky, A., I. Sutskever, and G.E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", Advances in Neural Information Processing Systems. 2012.
[22] LeCun, Y., B. Boser, J.S. Denker, D. Henderson, R.E. Howard, W. Hubbard, and L.D. Jackel, "Backpropagation applied to handwritten zip code recognition", Neural computation, 1(4), 1989, pp. 541–551.
[23] Lee, K.-H. and S.S. Hyun, "A model of value-creating practices, trusting beliefs, and online tourist community behaviors: Risk aversion as a moderating variable", International Journal of Contemporary Hospitality Management, 28(9), 2016, pp. 1868–1894.
[24] Levi, G. and T. Hassner, "Age and gender classification using convolutional neural networks", in The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. 2015.
[25] McCulloch, W.S. and W. Pitts, "A logical calculus of the ideas immanent in nervous activity", The bulletin of mathematical biophysics, 5(4), 1943, pp. 115–133.
[26] Montúfar, G.F. and J. Morton, "When Does a Mixture of Products Contain a Product of Mixtures?", SIAM Journal on Discrete Mathematics, 29(1), 2015, pp. 321–347.
[27] Mountrakis, G., J. Im, and C. Ogole, "Support vector machines in remote sensing: A review", ISPRS Journal of Photogrammetry and Remote Sensing, 66(3), 2011, pp. 247–259.
[28] Muselet, D. and L. Macaire, "Combining color and spatial information for object recognition across illumination changes", Pattern Recognition Letters, 28(10), 2007, pp. 1176–1185.
[29] Olson, D.L. and D. Delen, Advanced Data Mining Techniques, Springer Berlin Heidelberg, 2008.
[30] Orru, G., W. Pettersson-Yeo, A.F. Marquand, G. Sartori, and A. Mechelli, "Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review", Neuroscience & Biobehavioral Reviews, 36(4), 2012, pp. 1140–1152.
[31] Pass, G., R. Zabih, and J. Miller, "Comparing images using color coherence vectors", in Proceedings of the fourth ACM international conference on Multimedia. 1996. ACM: Boston, Massachusetts, United States.
[32] Sanchez-Franco, M.J. and F.J. Rondan-Cataluña, "Virtual travel communities and customer loyalty:

Customer purchase involvement and web site design", Electronic commerce research and applications, 9(2), 2010, pp. 171–182.
[33] Shaoqing, R., H. Kaiming, R. Girshick, Z. Xiangyu, and S. Jian, "Object Detection Networks on Convolutional Feature Maps", IEEE Transactions on Pattern Analysis and Machine Intelligence, 39(7), 2017, pp. 1476–1481.
[34] Simonyan, K. and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", 2015/04/10.
[35] Singla, A., L. Yuan, and T. Ebrahimi, "Food/non-food image classification and food categorization using pre-trained googlenet model", in Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management, Amsterdam, October 2016. 2016. ACM.
[36] https://www.statista.com/chart/10913/number-of-photos-taken-worldwide, accessed 4-25-2019.
[37] Stehling, R.O., M.A. Nascimento, and A.X. Falcão, "A compact and efficient image retrieval approach based on border/interior pixel classification", in Proc. of the Eleventh International Conference on Information and Knowledge Management, ACM, Editor. 2002: New York.
[38] Swain, M.J. and D.H. Ballard, "Color indexing", International Journal of Computer Vision, 7(1), 1991, pp. 11–32.
[39] Szegedy, C., V. Vanhoucke, S. Ioffe, and J. Shlens, "Rethinking the inception architecture for computer vision", in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[40] Tong, S. and D. Koller, "Support vector machine active learning with applications to text classification", Journal of machine learning research, 2(Nov), 2001, pp. 45–66.
[41] TripAdvisor, "About TripAdvisor", 2018.
[42] van Rysbergen, C.J., Information retrieval, 2nd edn., Butteerworth, London, 1979.
[43] Vapnik, V.N. and A. Cervonenkis, CA (1974). Teoriya Raspoznavaniya Obrazov: Statisticheskie Problemy Obucheniya (Theory of Pattern Recognition: Statistical Problems of Learning), 1974.
[44] Wang, J., Y. Yang, J. Mao, Z. Huang, C. Huang, and W. Xu, "CNN-RNN: A Unified Framework for Multi-Label Image Classification", in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.
[45] Wang, Y. and D.R. Fesenmaier, "Modeling participation in an online travel community", Journal of Travel Research, 42(3), 2004, pp. 261–270.
[46] Xiao, J., K.A. Ehinger, J. Hays, A. Torralba, and A. Oliva, "Sun database: Exploring a large collection of scene categories", International Journal of Computer Vision, 119(1), 2016, pp. 3–22.
[47] Yegnanarayana, B., Artificial neural networks, PHI Learning Pvt. Ltd, 2009.
[48] Zhang, D., M.M. Islam, and G. Lu, "A review on automatic image annotation techniques", Pattern Recognition, 45(1), 2012, pp. 346–362.