Classification of Instagram photos: topic modelling vs transfer learning

Nicolas Tsapatsoulis
nicolas.tsapatsoulis@cut.ac.cy
Cyprus University of Technology
Limassol, Cyprus

ABSTRACT

The existence of pre-trained deep learning models for image classification, such as those trained on the well-known Resnet-50 architecture, allows for easy application of transfer learning to several domains including image retrieval. Recently, we proposed topic modelling for the retrieval of Instagram photos based on the associated hashtags. In this paper we compare content-based image classification, based on transfer learning, with the classification based on topic modelling of Instagram hashtags for a set of 24 different concepts. The comparison was performed on a set of 1944 Instagram photos, 81 per concept. Despite the excellent performance of the pre-trained deep learning models, it appears that text-based retrieval, as performed by the topic models of Instagram hashtags, stills perform better.

CCS CONCEPTS

• **Computing methodologies** → *Information extraction*; **Natural language processing**.

KEYWORDS

image classification, topic modelling, transfer learning, deep learning

ACM Reference Format:
Nicolas Tsapatsoulis. 2022. Classification of Instagram photos: topic modelling vs transfer learning. In *12th Hellenic Conference on Artificial Intelligence (SETN 2022), September 7–9, 2022, Corfu, Greece.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3549737.3549759

1 INTRODUCTION

In a recent study [27] we introduced topic modelling as a mean for effective retrieval of Instagram photos. The method proved semantically consistent with human interpretation [11, 12]. However, nowadays deep learning [4] is considered the state of the art approach for image classification [6]. So, the research question we examine in this study is whether image classification of Instagram photos, for which textual information in the form of hashtags also exist, is more effectively achieved on the basis of visual content via deep learning or metadata / tagging via topic modelling.

SETN 2022, September 7–9, 2022, Corfu, Greece

ACM ISBN 978-1-4503-9597-7/22/09.
https://doi.org/10.1145/3549737.3549759

Topic modeling, a very popular data analysis technique in the field of big data, is a form of text mining, employing unsupervised and supervised statistical machine learning techniques to identify patterns in a corpus or large amount of unstructured text. It processes huge collection of documents and group the words, or more generally tokens, they contain into clusters of words identifying topics by using word (tokens) similarity processes. In topic modelling each document is viewed as a mixture of various topics, identified of the whole document collection, where each topic is defined as a distribution over a vocabulary of terms. Thus, the relevance of each document with a specific topic can be quantified on the basis of prominence of that specific topic in the document.

While topic modelling of Instagram hashtags for the purpose of Automatic Image Annotation (AIA) is a new area of research (see Argyrou *et al.* [3]), several researchers applied topic model analysis on social media data. With the aid of topic modelling the main themes that pervade a large and otherwise unstructured collection of social media documents can be mined. Instagram photos are usually accompanied by hashtags [7] that the owners use to describe photos' content and, in several cases, their feelings and moments that relate with those photos. We have shown in a previous study [10] that Instagram hashtags along with the relevant images provide a rich source for creating training sets for AIA. At the same time the hashtag sets of those images can be utilised for image retrieval on the basis of topic modelling as explained in [27].

Transfer learning [22] based on deep neural network models [26] has been proved a very successful approach for image classification on a variety of cases [21, 23, 28], including classification of Instagram photos [30]. Transfer learning has been developed based on the observation that the requirement for the training and test data to be independent and identically distributed is very strict and, in several cases, properly trained (using massive training data) models can be effectively applied on similar application domains without the need of retraining or models' adjustment. Thus, transfer learning was proposed as a radical solution to the problem of developing learned classification models using insufficient training data.

The plethora of pre-trained deep learning models for image retrieval and classification made transfer learning the main benchmark method for every application in image classification and image retrieval. This is the main reason for conducting the current study: Our findings regarding the use of topic modelling of Instagram hashtags for the retrieval of Instagram photos [27] must be assessed on the basis of a state-of-the-art method for content-based image retrieval.

2 RELATED WORK

In this section we review relevant work related to the application of transfer learning in image classification, Instagram image classification using deep learning methods and topic model analysis of social media data.

2.1 Topic modelling of social media data

Manikonda *et al.* [18] concluded that on Twitter you can locate informational content while on Instagram the textual content is more personal and social in nature and does not always match the visual content of the associated images. To reach this conclusion the researchers performed textual analysis of tweets and Instagram comments and hashtags and visual analysis of Instagram photos posted on these two platforms from the same set of users. For their textual analysis they used Latent Topic Models with the aid of the Twitter-LDA API[1] which was developed for topic modeling of short text corpora to mine the latent topics [31]. The visual analysis targeted on image clustering using low level features (SURF features[2]).

Liu and Jansson [16] tried to identify city events from Instagram data. They created a dataset with posts, comments and hashtags from publicly accessible Instagram accounts in the Helsinki metropolitan region. Then, they applied an LDA based topic modelling method to the set of relevant posts in order to discover clusters of targeted events. Instagram hashtags were kept during their analysis but only as a part of the container post / message. The authors concluded that it is necessary to remove frequent non-topical terms, such as compliments, excitements or other positive tone and sentiments in order to bring up more novel topics. They examined, also, the importance of hashtags' presence in the Instagram posts and drew the conclusion that keeping hashtags in the analysis brings additional value into the mined topics.

2.2 Classification of Instagram photos via deep learning

Xie *et al.* [30] trained an EfficientNet model on labeled ImageNet images and used it as a teacher (training data) to generate pseudo labels on 300M unlabeled Instagram images. They trained, then, a larger EfficientNet as a student model on the combination of labeled and pseudo labeled Intsagram images. They iterated this process by putting back the student as the teacher, adn so on. During the generation of the pseudo labels, the teacher was not noised so that the pseudo labels are as accurate as possible. However, during the learning of the student, the authors injected noise such as dropout, stochastic depth and data augmentation via RandAugment to the student so that the student generalizes better than the teacher. The authors claim a 88.4% top-1 accuracy on ImageNet, which is 2.0% better than the state-of-the-art model that requires 3.5B weakly labeled Instagram images. On robustness test sets, it improves ImageNet-A top-1 accuracy from 61.0% to 83.7%, and reduces ImageNet-C mean corruption error from 45.7 to 28.3, and reduces ImageNet-P mean flip rate from 27.8 to 12.2.

Argyris *et al.* [2] collected more than 45000 images and social media usage behaviors over 26 months. They then applied deep-learning algorithms to automatically classify each image and used social media analytics to disclose hidden associations between visual elements and brand engagement. Their hypothesis testing results provide empirical support for VCSI (Visual-Congruence-induced Social Influence), advancing theories into the rapidly growing fields of multimodal content and Influencer marketing.

Mittal *et al.* [20] inspect some prominent user interaction properties and photo properties to understand users' engagements towards posts on Instagram. They used user interaction properties such as hashtags and photo post time and user's posted photo features or image contexts such as image filters. They applied these user interaction and photo properties analysis task on eight major cities' Instagram posts and further classified the posts of those cities in five categories using Non-negative Matrix Factorization and the Latent Dirichlet Allocation algorithm. The four prime influencing analyses were computed to get the ecology of the users on Instagram photo posts, namely: Time based analysis (TBA), Image Filter analysis (IFA), Image Hashtags analysis (IHA) and Image categorization analysis (ICA). According to the authors, this multivariate feature based Instagram analysis could help users to gain insight of popular content and make their respective content popular so as to reach out to a maximum number of people.

2.3 Transfer learning for image classification

Abdullah and Hasan [1] in their study to classify 200 images in five categories, namely: Binoculars, Planes, Faces of people, Watches, and Motorbikes, used the AlexNet Model, a pre-trained CNN, and they concluded that with the help of that pre-trained model improved accuracy of classification is achieved. Chaib *et al.* [5] used two pre-trained Convolutional Neural Networks (VGG-Net and CaffeNet) for the classification of Very High-Resolution (VHR) satellite images. They compare their method with other state-of-the-art methods, and they concluded that their transfer-learning based methodology outperforms the other state-of-the-art methods. Shima [24] also applied transfer learning with the aid of Alexa-Net pre-trained models. They investigated object classification, using the STL-10 database, for ten classes, and they achieved an average of 84.38% test-set accuracy.

Nasiri *et al.* [21] examined automatic identification of grapevine cultivar by leaf image using pre-trained ImageNet models, in a transfer learning context. They reported an overall accuracy of 99.11% on six grapevine cultivars. Taheri-Garavand *et al.* [25] proposed a transfer-learning based methodology for chickpea variety identification and discrimination. Four commercial chickpea varieties (Adel, Arman, Azad, and Saral) were used in their experiments. They used pre-trained ImageNet models to fine-tune their models and they reached an average classification accuracy of over 94%.

Fiallos *et al.* [8] collected 7382 pictures associated with the hashtag *#allyouneedisecuador* which was created during a campaign entitled *"All you need is Ecuador"* in an effort to strengthen tourism in Ecuador. They calculated the similarity of topics mined from user descriptions (hashtags and post text) and topics mined from visual

[1] https://github.com/minghui/Twitter-LDA
[2] https://en.wikipedia.org/wiki/Speeded_up_robust_features

analysis of the photos, called visual descriptions. Visual descriptions were extracted with the aid of Microsoft Cognitive Services[3]. The visual descriptions produced 962 terms (reduced to 838 after the preprocessing stage) while the user descriptions produced 21972 (reduced to 18810 terms after pre-processing). Topic modelling was applied to both description sets separately by combining TF-IDF with either the Non-Negative Matrix Factorization algorithm or the K-Means clustering algorithm. The authors discovered low similarity between the topics mined from the user descriptions and the visual descriptions and attributed this deviation to the fact users usually refer to situations or opinions regarding the photos while visual analysis produces tags more related with the actual content of the corresponding images.

The previous discussion reveals that topic modelling on Instagram hashtags has not been thoroughly investigated yet for image indexing and retrieval. Instagram hashtags were, typically, used only as a part of the container post / message while in many of the reported works dealing with Instagram photo classification were totally ignored. In addition, the work of Fiallos *et al.* [8] indirectly suggests that visual descriptions, largely used in content-based image retrieval, lack semantic meaning and textual descriptions, including hashtags serve complimentary to that.

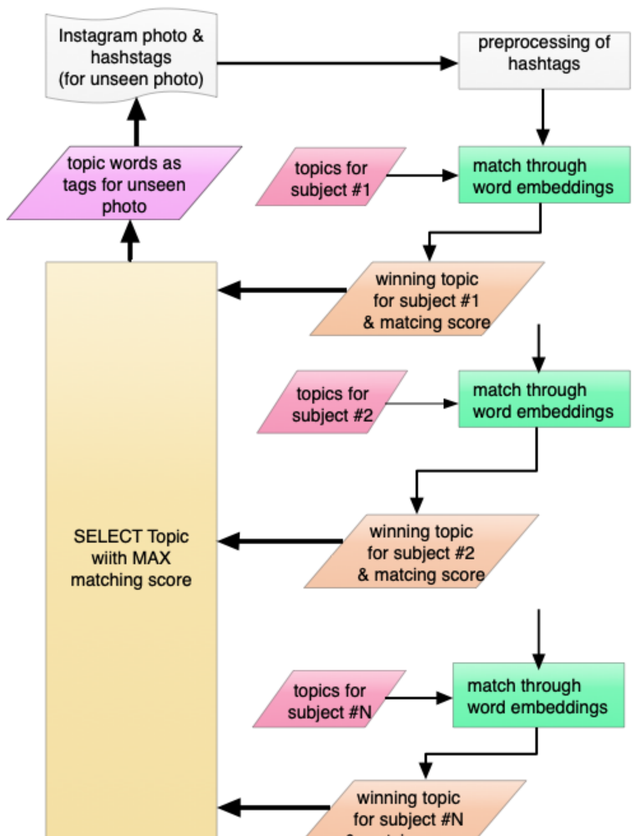

Figure 1: The topic modelling tagging process

[3]https://azure.microsoft.com/en-us/services/cognitive-services/directory/vision/

3 METHODOLOGY

Topic modelling of Instagram hashtags and the use of developed topic models for Instagram photo retrieval is explained in detail in [27]. Below we provide a short recap so as for the reader to better understand the method as well as the complexity compared to transfer learning.

Figure 2 shows an example of how the comparison of the two methods is done: The visual content of the Instagram post (photo to the left) is classified via the ResNet-50 model. In case the winning class is one of the *hat* related classes (i.e., *sombrero*, *cowboy hat* - see Table 1) the transfer learning based classification is successful. The hashtags of the Instagram post (see the right part of Figure 2) are classified via the trained topic models. Again, if the winning class is *hat* topic modelling based classification is successful.

3.1 Topic modelling of Instagram hashtags

The architecture of the proposed technique for image retrieval of Instagram images based on topic modelling of Instagram hashtags is shown in Figure 1. First, topics models are created from a collection of Instagram hashtags of photos belonging to the same concept (i.e., queried by the same hashtags, say *#bear*). This results in a set of topics $\mathcal{S}^q = \{\mathcal{T}_1^q, \mathcal{T}_2^q, ..., \mathcal{T}_t^q, ...\mathcal{T}_{q_k}^q\}$ for the q-th concept. Second, the matching score of the hashtags $\mathcal{H}_I$ of an unseen Instagram image I, preprocessed in the same way as the training instances, with each one of the topics $\mathcal{T}_t^q$ of the q-th subject is computed. Both $\mathcal{H}_I$ and $\mathcal{T}_t^q$ are sets of words while the latter includes also the importance of each word expressed through its relative frequency in the topic. The matching score $R(\mathcal{H}_I, \mathcal{T}_t^q)$ between these two sets can be computed as a weighted sum of the pair similarities of their word embeddings [19] used for the Glove project[4], as shown in eq. 1. Those word embeddings were pre-trained on external sources such Google News and Wikipedia.

$$R(\mathcal{H}_I, \mathcal{T}_t^q) = \frac{1}{|\mathcal{H}_I| \cdot |\mathcal{T}_t^q|} \sum_{h_I \in \mathcal{H}_I} \sum_{w_{t\xi}^q \in \mathcal{T}_t^q} \alpha[w_{t\xi}^q] \cdot cc(\vec{h}_I, \vec{w}_{t\xi}^q) \quad (1)$$

where $|\mathcal{A}|$ denotes the cardinality of set $\mathcal{A}$, $\vec{h}_I$ and $\vec{w}_{t\xi}^q$ are the word embeddings of hashtag h_I and topic word $w_{t\xi}^q$ respectively, $\alpha[w_{t\xi}^q]$ is the authority score, computed with the aid of the HITS (Hyperlink-Induced Topic Search) algorithm, of word $w_{t\xi}^q$ and $cc(.,.)$ is the similarity measure used with the word embeddings of Gensim models[5].

In the third step, the best matching, to the set of hashtags $\mathcal{H}_I$, topic $\mathcal{T}_{opt}^q(\mathcal{H}_I)$ of the q-th concept is selected with aid of eq. 2:

$$\mathcal{T}_{opt}^q(\mathcal{H}_I) = argmax\{R(\mathcal{H}_I, \mathcal{T}_t^q)\} \quad (2)$$

Finally, the best matching topic $\mathcal{T}_{opt}(\mathcal{H}_I)$ for set $\mathcal{H}_I$, across all concepts, denotes the tags that will be assigned to Instagram image I (in case of AIA purposes) or the winning category (in the case of image classification), and is given by eq. 3:

$$\mathcal{T}_{opt}(\mathcal{H}_I) = argmax\{\mathcal{T}_{opt}^q(\mathcal{H}_I)\} \quad (3)$$

[4]https://nlp.stanford.edu/projects/glove/
[5]https://radimrehurek.com/gensim/models/word2vec.html

h

Figure 2: An example of Instagram image along with its hashtags. The photo is classified via transfer learning while the hashtags are classified based on topic modelling (see also the Python Code in Appendix)

3.2 Transfer learning

Residual Neural Networks (ResNets) developed by He *et al.* [13] are state of the art deep learning architectures. ResNets deal with the vanishing gradient and the degradation problem that appear during the ultra-deep CNNs train. They include stacked residual units (building blocks) containing skip connections to link the input and output of each unit. CNNs with residual units were proved to perform better than plain counterparts. ResNet-50, the specific architecture adopted, consists of 50 layers, providing residual connection between them.

The Resnet-50 pre-trained model could be ideal for Instagram image classification in the context of transfer learning. In this context it has been used in a variety of studies dealing with image classification [14, 15, 17, 23, 29] with excellent results. The Resnet-50 models classify images into one of 1000 image classes corresponding to a variety of concepts. As shown in Table 1 there is no 1-to-1 matching between the concepts identified by the 24 Instagram query hashtags and the Resnet-50 image classes. While for some concepts, such as *hamster*, *lion*, *microphone* and *zebra*, there is an exact match this is not the cases for the majority of the remaining Instagram concepts. The Instagram concept *dog*, for instance, corresponds to more than 40 Resnet-50 image classes. Nevertheless, whenever transfer learning based classification results in a winning class related to the corresponding Instagram concept it is considered as a successful classification.

3.3 Corpus

To examine the proposed methodology we constructed a training dataset of 1944 Instagram images (see Table 1) along with their hashtags by querying with 24 different concepts / hashtags (i.e., #*bear*, #*cat*, etc.) using the *Instaloader* Python library[6] as it can be seen in Appendix A. From the results we selected manually the 81 most relevant, in terms of visual content, images for each concept. The great majority (if not all) of those images were uploaded to Instagram by different users. The hashtags surrounding these images were also automatically crawled and stored in a 24 different files one for each concept. Each file contains 81 columns with each column corresponding to the hashtags of an individual Instagram image within that concept.

A total of 33067 hashtags were collected across all concepts, corresponding to an average of 17 hashtags per image. As expected from the trend of Instagram users the same hashtags may appear in more than one subject (see [9] for an extended discussion on this issue).

[6] https://instaloader.github.io/

4 RESULTS & DISCUSSION

Table 1 summarises the results of the current study. It is shown there the top-1 and top-5 retrieval accuracy per concept (or aggregated concepts in the case of transfer learning based classification) for the compared methods along with the corresponding matching scores. Oviosulsy the scale of matching scores of the two compared methods varies and this is the reason we have included in the results the corresponding standard deviation of matching scores. The matching scores seem to be correlated with the top-1 accuracy, especially in the case of transfer learning.

On average topic modelling based classification outperforms transfer learning based classification in both top-1 and top-5 accuracy (83.5% vs 71.7% for top-1 and 97.5% vs 87.4% for top-5) . This result indicates that text-based image retrieval stills perform better than content-based image retrieval even in the era of deep learning.

In a more quantitative basis we observe that the performance of Resnet-50 models varies greatly among the various concepts. For instance, the *chair*, *piano*, and *hat* related classes achieve less than 40% top-1 accuracy while *dog*, *elephant* and *zebra* related classes achieve a top-1 accuracy higher than 98.5%. The visual appearence of concepts such us *chair*, *piano*, and *hat* in Instagram photos (and in general) is vague and in the majority of cases these objects appear with other other objects in the same photo. Figure 2 shows such an example: The hat goes with a human in a scene showing also sea and sky. Chairs also appear in many cases with tables and pianos with humans (pianists). This means that deep learning models are still based on low level visual characteristics rather than semantic information.

Performance variation among the concepts is also observed, in a lower level though, in the topic modelling based classification. We see, for instance, that for some concepts such as *hedgehog* and *table* the top-1 accuracy is less than 30%. The *hedgehog* concept is vaquely described through hashtags while the semantic description of the *table* concept is similar to that of the *chair* concept. This is reflected in the top-5 accuracy of the two concepts which reaches 100% in both cases.

5 CONCLUSION

In this paper we compare text based image retrieval, implemented via topic modelling of Instagram hashtags as suggested by Tsapatsoulis [27], with contet-based image retrieval realized via tranfer learning of the well-known Resnet-50 models. The comparison was

Table 1: Classification performance for 81 test cases per concept

	Transfer Learning					Topic Modelling			
Subject	related concepts	top 1%	top 5%	Mean Score	St. Dev.	top 1%	top 5%	Mean Score	St. Dev.
bear	brown bear, American black bear, ice bear, sloth bear	93.8	98.8	0.6798	0.2467	100	100	0.2987	0.1055
cat	tabby, Angora, Siamese cat, Persian cat Madagascar cat, Egyptian cat, tiger cat	82.7	93.8	0.4396	0.2757	98.8	100	0.2665	0.1028
chair	rocking chair, folding chair, barber chair park bench, cradle, studio couch	37.0	65.4	0.1603	0.2019	93.8	100	0.1758	0.0754
dog	various (more than 40 concepts)	98.8	100	0.7543	0.2034	96.3	100	0.2130	0.1079
elephant	African elephant, Indian elephant, tusker	98.8	100	0.6716	0.1822	79.0	100	0.2012	0.0831
fish	barracouta, coho, gar, sturgeon, tench, anemone fish goldfish, starfish, hammerhead, puffer. goldfish, great white shark, etc.	64.6	82.9	0.3799	0.3176	69.5	97.6	0.0637	0.0271
guitar	electric guitar, acoustic guitar	76.5	90.1	0.5374	0.3202	100	100	0.2334	0.0771
hamster	hamster	80.2	95.1	0.6255	0.3231	90.1	100	0.2381	0.0971
handbag	purse, backpack, mailbag	67.9	81.5	0.3096	0.2476	90.1	100	0.1875	0.0684
hat	sombrero, cowboy hat	38.3	55.6	0.2224	0.2882	92.6	100	0.1527	0.0541
hedgehog	porcupine, echidna	80.2	88.9	0.6802	0.3451	25.9	45.7	0.0501	0.0655
horse	sorrel	46.9	70.4	0.2705	0.2902	98.8	100	0.1850	0.0646
laptop	notebook, laptop, hand-held computer	61.0	90.2	0.3169	0.2308	93.9	100	0.1120	0.0542
lion	lion	87.7	91.4	0.7229	0.2796	66.7	100	0.1421	0.0536
mic	microphone	67.9	93.8	0.5757	0.3726	81.5	98.8	0.0846	0.0300
monkey	macaque, baboon, capuchin, chimpanzee, titi howler monkey, guenon, patas, langur, orangutan spider monkey, gorilla	90.1	97.5	0.5995	0.2625	97.5	100	0.2357	0.0925
parrot	macaw, lorikeet	49.4	72.8	0.3274	0.3599	86.4	98.8	0.1799	0.0596
piano	grand piano, accordion	37.0	87.7	0.2730	0.3112	79.0	100	0.1115	0.0420
rabbit	angora, wood rabbit, hare	75.3	91.4	0.4005	0.2624	100	100	0.3213	0.1179
sunglasses	sunglasses, sunglass	69.1	88.9	0.3434	0.2355	98.8	100	0.2319	0.1158
table	dining table, desk, pool table	56.8	84.0	0.3649	0.3210	29.6	100	0.1464	0.0459
turtle	loggerhead, leatherback turtle, box turtle, terrapin chiton, mud turtle	91.4	96.3	0.5158	0.2227	55.6	98.8	0.1153	0.0414
watch	stopwatch, analog clock, digital watch	69.1	80.2	0.3682	0.2579	92.6	100	0.1778	0.0640
zebra	zebra	100	100	0.8954	0.0885	86.4	100	0.2062	0.0756
Average		71.7	87.4	0.4764	0.2686	83.5	97.5	0.1804	0.0717

made on a set of 1944 Instagram posts (images and associated hashtags) collected through 24 different main hashtags (concepts). The main conclusion of the current study is that topic modelling based classification outperforms transfer learning based classification in both top-1 and top-5 accuracy. Top-5 accuracy is by far better for the topic modelling approach but for a much fewer number categories (24 in topic modelling based classification versus 1000 image classes in Resnet-50) despite that in the transfer learning case many similar concept models do exist.

Another important finding is that topic models of similar subjects (for instance for the concepts *chair* and *table*) are also similar making the retrieval process explainable to the end user while deep learning models are still based on low level visual characteristics. Thus, the two methods should be seen as complimentary as also suggested by Fiallos *et al.* [8]. We are planning to investigate this in a future study.

Not all concepts are modelled, either visually or semantically, with the same success. In the content-based image retrieval, coexistence of concepts in the same image or small objects are the main reasons for ineffective modelling while vagueness in semantic description of concepts such as *hedgehog* or similarity of semantic descriptions, as in the case of *chair* and *table*, are the main reasons for poor topic modelling via Instagram hashtags.

REFERENCES

[1] Abdullah and Mohammad S. Hasan. 2017. An application of pre-trained CNN for image classification. In *Proceedings of the 20th International Conference of Computer and Information Technology (ICCIT)*. IEEE, 1–6. https://doi.org/10.1109/ICCITECHN.2017.8281779

[2] Young Anna Argyris, Zuhui Wang, Yongsuk Kim, and Zhaozheng Yin. 2020. The effects of visual congruence on increasing consumers' brand engagement: An empirical investigation of influencer marketing on instagram using deep-learning algorithms for automatic image classification. *Computers in Human Behavior* 112 (2020), 106443. https://doi.org/10.1016/j.chb.2020.106443

[3] Argyris Argyrou, Stamatios Giannoulakis, and Nicolas Tsapatsoulis. 2018. Topic modelling on Instagram hashtags: An alternative way to Automatic Image Annotation?. In *Proceedings of the 13th International Workshop on Semantic and Social Media Adaptation and Personalization* (Zaragoza, Spain) *(SMAP)*. IEEE, 61–67.

[4] Jianfang Cao, Aidi Zhao, and Zibang Zhang. 2020. Automatic image annotation method based on a convolutional neural network with threshold optimization. *PLoS ONE* 15 (2020). https://doi.org/10.1371/journal.pone.0238956

[5] Souleyman Chaib, Hongxun Yao, Yanfeng Gu, and Moussa Amrani. 2017. Deep feature extraction and combination for remote sensing image classification based

on pre-trained CNN models. In *Proceedings of the 9th International Conference on Digital Image Processing (ICDIP 2017)*, Charles M. Falco and Xudong Jiang (Eds.), Vol. 10420. SPIE, 712–716. https://doi.org/10.1117/12.2281755

[6] Qimin Cheng, Qian Zhang, Peng Fu, Conghuan Tu, and Sen Li. 2018. A survey and analysis on automatic image annotation. *Pattern Recognition* 79 (2018), 242–259. https://doi.org/10.1016/j.patcog.2018.02.017

[7] Alice R. Daer, Rebecca Hoffman, and Seth Goodman. 2014. Rhetorical Functions of Hashtag Forms Across Social Media Applications. In *Proceedings of the 32nd ACM International Conference on The Design of Communication CD-ROM* (Colorado Springs, CO, USA) *(SIGDOC'14)*. Association for Computing Machinery, Article 16. https://doi.org/10.1145/2666216.2666231

[8] Angel Fiallos, Karina Jimenes, Carlos Fiallos, and Stalin Figueroa. 2018. Detecting Topics and Locations on Instagram Photos. In *Proceedings of the 2018 International Conference on eDemocracy - eGovernment (ICEDEG)*. IEEE, 246–250. https://doi.org/10.1109/ICEDEG.2018.8372314

[9] Stamatios Giannoulakis and Nicolas Tsapatsoulis. 2016. Defining and identifying stophashtags in instagram. In *INNS Conference on Big Data* (Thessaloniki, Greece). INNS, 304–313.

[10] Stamatios Giannoulakis and Nicolas Tsapatsoulis. 2016. Evaluating the descriptive power of Instagram hashtags. *Journal of Innovation in Digital Ecosystems* 3, 2 (2016), 114–129. https://doi.org/10.1016/j.jides.2016.10.001

[11] Stamatios Giannoulakis and Nicolas Tsapatsoulis. 2021. Topic Identification via Human Interpretation of Word Clouds: The Case of Instagram Hashtags. In *Artificial Intelligence Applications and Innovations*, Ilias Maglogiannis, John Macintyre, and Lazaros Iliadis (Eds.). Springer International Publishing, Cham, 283–294. https://doi.org/10.1007/978-3-030-79150-6_23

[12] Stamatios Giannoulakis and Nicolas Tsapatsoulis. 2022. Topic Identification of Instagram Hashtag Sets for Image Tagging: An Empirical Assessment. In *Metadata and Semantic Research*, Emmanouel Garoufallou, María-Antonia Ovalle-Perandones, and Andreas Vlachidis (Eds.). Springer International Publishing, Cham, 159–170. https://doi.org/10.1007/978-3-030-98876-0_14

[13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2016. Deep Residual Learning for Image Recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 770–778. https://doi.org/10.1109/CVPR.2016.90

[14] Heechul Jung, Min-Kook Choi, Jihun Jung, Jin-Hee Lee, Soon Kwon, and Woo Young Jung. 2017. ResNet-based vehicle classification and localization in traffic surveillance systems. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition Workshops*. IEEE, 61–67.

[15] Riaz Ullah Khan, Xiaosong Zhang, Rajesh Kumar, and Emelia Opoku Aboagye. 2018. Evaluating the Performance of ResNet Model Based on Image Recognition. In *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence (ICCAI 2018)*. Association for Computing Machinery, New York, NY, USA, 86–90. https://doi.org/10.1145/3194452.3194461

[16] Shuhua Liu and Patrick Jansson. 2017. City event detection from social media with neural embeddings and topic model visualization. In *2017 IEEE International Conference on Big Data (Big Data)*. 4111–4116. https://doi.org/10.1109/BigData.2017.8258430

[17] Arpana Mahajan and Sanjay Chaudhary. 2019. Categorical Image Classification Based On Representational Deep Network (RESNET). In *Proceedings of the 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 327–330. https://doi.org/10.1109/ICECA.2019.8822133

[18] Lydia Manikonda, Venkata Vamsikrishna Meduri, and Subbarao Kambhampati. 2016. Tweeting the Mind and Instagramming the Heart: Exploring Differentiated Content Sharing on Social Media. In *Proceedings of the 10th International AAAI Conference on Web and Social Media (AAAI)*. PKP|PS, 639–642.

[19] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed Representations of Words and Phrases and Their Compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2 (NIPS'13)*. Curran Associates Inc., Red Hook, NY, USA, 3111–3119.

[20] Vatsala Mittal, Aastha Kaul, Santoshi Sen Gupta, and Anuja Arora. 2017. Multivariate Features Based Instagram Post Analysis to Enrich User Experience. *Procedia Computer Science* 122 (2017), 138–145. https://doi.org/10.1016/j.procs.2017.11.352

[21] Amin Nasiri, Amin Taheri-Garavand, Dimitrios Fanourakis, Yu-Dong Zhang, and Nikolaos Nikoloudakis. 2021. Automated Grapevine Cultivar Identification via Leaf Imaging and Deep Convolutional Neural Networks: A Proof-of-Concept Study Employing Primary Iranian Varieties. *Plants* 10, 8 (2021). https://doi.org/10.3390/plants10081628

[22] Sinno Jialin Pan and Qiang Yang. 2010. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* 22, 10 (2010), 1345–1359. https://doi.org/10.1109/TKDE.2009.191

[23] A. Sai Bharadwaj Reddy and D. Sujitha Juliet. 2019. Transfer Learning with ResNet-50 for Malaria Cell-Image Classification. In *2019 International Conference on Communication and Signal Processing (ICCSP)*. 945–949. https://doi.org/10.1109/ICCSP.2019.8697909

[24] Yoshihiro Shima. 2018. Image Augmentation for Object Image Classification Based On Combination of Pre-Trained CNN and SVM. *Journal of Physics: Conference Series* 1004 (apr 2018), 012001. https://doi.org/10.1088/1742-6596/1004/1/012001

[25] Amin Taheri-Garavand, Amin Nasiri, Dimitrios Fanourakis, Soodabeh Fatahi, Mahmoud Omid, and Nikolaos Nikoloudakis. 2021. Automated In Situ Seed Variety Identification via Deep Learning: A Case Study in Chickpea. *Plants* 10, 7 (2021). https://doi.org/10.3390/plants10071406

[26] Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, and Chunfang Liu. 2018. A Survey on Deep Transfer Learning. In *Artificial Neural Networks and Machine Learning – ICANN 2018*, Věra Kůrková, Yannis Manolopoulos, Barbara Hammer, Lazaros Iliadis, and Ilias Maglogiannis (Eds.). Springer International Publishing, Cham, 270–279.

[27] Nicolas Tsapatsoulis. 2020. Image retrieval via topic modelling of Instagram hashtags. In *Proceedings of the 15th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*. IEEE, 1–8. https://doi.org/10.1109/SMAP49528.2020.9248465

[28] Nicolas Tsapatsoulis and Katerina Diakoumopoulou. 2021. Face Verification in Practice: The Case of Greek Artist Leonidas Arniotis. In *Computer Analysis of Images and Patterns*, Nicolas Tsapatsoulis, Andreas Panayides, Theo Theocharides, Andreas Lanitis, Constantinos Pattichis, and Mario Vento (Eds.). Springer International Publishing, Cham, 131–139.

[29] Weibin Wang, Dong Liang, Qingqing Chen, Yutaro Iwamoto, Xian-Hua Han, Qiaowei Zhang, Hongjie Hu, Lanfen Lin, and Yen-Wei Chen. 2020. *Medical Image Classification Using Deep Learning*. Springer International Publishing, Cham, 33–51. https://doi.org/10.1007/978-3-030-32606-7_3

[30] Qizhe Xie, Minh-Thang Luong, Eduard Hovy, and Quoc V Le. 2020. Self-training with noisy student improves imagenet classification. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. IEEE, 10687–10698.

[31] W. X. Zhao, J. Jiang, J. Weng, J. He, E.-P. Lim, H. Yan, and X. Li. 2011. Comparing twitter and traditional media using topic models. In *Proceedings of the 33rd European Conference on Advances in Information Retrieval*. 338–349.

A PYTHON CODE

```
#------------------- Crawling Instagram Data ------------------------------
>>> import instaloader
>>> L = instaloader.Instaloader()
>>> hashtag = instaloader.Hashtag.from_name(L.context, 'bear')
>>> posts=hashtag.get_posts()
>>> k=1; hash = []
L.download_hashtag('bear', max_count=100)
>>> for post in posts:
...     if k>100:
...             break
...     else:
...             hash+=[post.caption_hashtags]
...             k+=1
...
>>> hash[0]
['sunday', 'barbudo', 'domingo', 'beardedmen', 'beard', 'bearded', 'barba',
       'bear', 'beardstyle',
'pose', 'beardlove', 'instabeard', 'beardstagram', 'hairy', 'fitness', 'fit',
       'model', 'muscle']

#-------------------- Image Classification ------------------------------------
import mxnet as mx
import gluoncv as cv
# Gets the trained topics and a csv file with the hashtags for each image
# (see for instance the zebra.csv file) and returns the top1 and top5
       percentages
# as well as the mean matching score and the corresponding standard deviation
 def TopicScores(dir, concept, topics):
# Example Call: TopicScores(dir, 'turtle', topics)
   concept = concept.lower()
   filepath = concept+'/'+concept+'.csv'
   filename = dir+filepath
   E = topic_hashtag_csvRead(filename)
   L = len(E['Hashtags'])
   for i in np.arange(L):
      key = E['\ufeffkey'][i].lower()
      H = tknzr.tokenize(E['Hashtags'][i])
      H = [k for k in H if len(k)>1]
      TM[key] = {}
      TM[key] = TopicMatch(topics, H)
   keys = [key for key in TM.keys() if key[:3]==concept[:3]]
   scores=[sorted(TM[key].items(), key=operator.itemgetter(1),
          reverse=True)[0][0] for key in keys]
   f = FreqDist(scores)
   top1 = f[concept.lower()]/f.N()
   LM = []
```

```
   for key in keys:
      LM+=sorted(TM[key].items(), key=operator.itemgetter(1), reverse=True)[:5]
   scores = [key[0] for key in LM]
   f = FreqDist(scores)
   top5 = f[concept]/L
   scores=[TM[key][concept] for key in keys]
   return top1, top5, np.mean(scores), np.std(scores)

# Given the topic models and a hashtag set H find the match between each topic
      and the
# hashtag set
def TopicMatch(topics, H):
   scores = {}
   for key in topics.keys():
      S = []
      for subkey in topics[key]['b'].keys():
         S+=[matchingScore(topics[key]['b'][subkey], H, stoph, 0.01)[0]]
      scores[key]=max(S)
   return scores

def ImageClassificationScores(dir, concept, net, synonyms):
# Example Call: ImageClassificationScores(dir, 'watch', net, ['stopwatch',
      'analog clock', 'digital watch'])
   concept = concept.lower()
   dir = dir+concept+'/'
   filekeys = [f for f in os.listdir(dir) if f[-3:]=='jpg']
   L = len(filekeys)
   for key in filekeys:
      filename = dir+key
      img = mx.image.imread(filename)
      transformed_img = cv.data.transforms.presets.imagenet.transform_eval(img)
      pred = net(transformed_img)
      prob = mx.nd.softmax(pred)[0].asnumpy()
      ind = mx.nd.topk(pred, k=10)[0].astype('int').asnumpy().tolist()
      key=key[:-4].lower()
      IC[key]={}
      IC[key]['class'] = [net.classes[ind[i]] for i in np.arange(10)]
      probs = [prob[net.classes.index(t)] for t in synonyms]
      IC[key]['prob'] = max(probs)
   scores = [IC[key[:-4].lower()]['class'][0] for key in filekeys]
   f = FreqDist(scores)
   T = 0; scores = []
   for t in synonyms:
      T += f[t]
      scores += [key for key in filekeys if t in
            IC[key[:-4].lower()]['class'][:5]]
   top1 = T/L; top5 = len(set(scores))/L
   scores = [IC[key[:-4].lower()]['prob'] for key in filekeys]
   return top1, top5, np.mean(scores), np.std(scores)
```