iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** VI **Month of publication:** June 2021

DOI: https://doi.org/10.22214/ijraset.2021.35516

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue VI Jun 2021- Available at www.ijraset.com

Hashtag Generation with Transfer Learning

M. Venu Gopalachari[1], Pavan Pawar[2], Nishtharth Kasireddy[3], S Rakesh [4]
[1]*Associate Professor,* [2,3]*UG Student,* [4]*Assistant Professor Department of IT, Chaitanya Bharathi Institute of Technology, Hyderabad*

Abstract*— In the recent past, Online Marketing applications have been a focus of research. Now-a-days we can observe a rapid growth of users on social media platforms which makes these platforms more suitable for Online Marketing. But still there are enormous challenges on the accuracy and authenticity of the content posted through social media. And if the social media business platforms are considered, majority of the users who try to add a market value to their own product face the problem of not getting enough attention from their target audience. Hashtags are super useful, not only do they help users find their desired product, they also help businesses reach their target audience. This paper focuses on providing trending hashtags for businesses to increase their reach in the market and also help users find their desired product. Firstly Data collection and preprocessing is done next few pre-trained model are selected for performing transfer learning, comparing results of various pre-trained models and choosing the best one and get the trending hashtags

***Keywords*— hashtag generation, deep learning, social media analytics, social networks, automatic recommendation**

I. INTRODUCTION

Now-a-days as most of the people are using social media, ordering clothes online, ordering food online etc., and on the other hand we can also see newly emerging businesses which strive to provide best quality products to its customers. But all these businesses fail to reach their target audience due to lack of knowledge about social media marketing[1, 7]. This is because social media is vast and also dynamic. So to solve this problem hashtags are at rescue. Hashtags are super useful, Not only do they help users find content, they also help recommendation systems to curate content to users, and that has intrinsic value in it. Better product recommendations means happier users and more happier businesses [10, 11]. There are many researches done and projects developed which generate a short caption on image provided but there are very fewer sources that talk about hashtags and their importance. As we know that social media is vast and tracking changes and finding the public interests is difficult and to solve this problem hashtags were introduced i.e., classifying the content on internet [8].

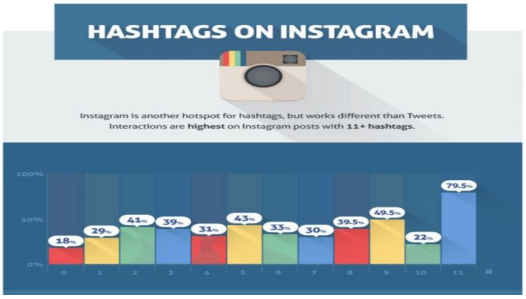


Figure 1 Like percentage vs No.of Hashtags(2020)

There are many researches done and projects developed which generate a short caption on image provided but there are very fewer sources that talk about hashtags and their importance [9]. As we know that social media is vast and tracking changes and finding the public interests is difficult and to solve this problem hashtags were introduced i.e., classifying the content on internet. Since hashtags increase the reach and engagement of content, they are a great way to increase traffic on Twitter, Instagram, and other social media platforms. The purpose of this work is to develop a safe and efficient trending hashtag generating application for social media marketers (business users) which generates trending and relevant hashtags for user content. The main problem related to this paper is identifying the image and giving hashtags which are on top of search results in Instagram. Instagram does not allow users to scrape its posts in large number so to tackle this, transfer learning and Instaloader API are used in the paper.

II. RELATED WORK

The analysis phase in [1] was focused on gathering information about the existing systems and analyzing the weaknesses and strengths of the respective systems which lead to the concept of developing the new system. Requirements of the new system were clearly understood during the analysis phase. The research team identified main users of the "Trending Hashtag Generator and Image Authenticator" are social media marketers, promoters and their target audience [5, 6]. After conducting a

©IJRASET: All Rights are Reserved

survey, the research team found that more than 70% of online marketers use hashtags to promote their content and majority of them were not satisfied with their audience reach even after using social media platform audience tools. 60% of online marketers and promoters were using existing systems to find matching hashtags for their content. Authors came into a conclusion that an efficient system which analyzes the content and generate "real time trending hashtags" would assist the online marketers and promoters to get the audience reach they expect.

The objective of this [2] is twofold. First, generate hashtags for an input image and second, use one of the suggested hashtags as seed text to produce a story as a caption for the image. To generate hashtags for the image, the attention-based framework introduced in [3] is used. The Encoder-Decoder framework uses a CNN-LSTM network to generate captions for the image. Originally, the framework was used to generate captions in the form of a sentence in natural language but this work leverages the model to produce hashtags instead of sentences. The model takes an image I as input and produces a list, X, of hashtags where $| X | \leq 1$. To generate story-like captions from the hashtags generated above, an character-level Recurrent (RNN) Neural Network is trained on a corpus of personal narratives. The RNN models the probability distribution of the characters in sequence given a sequence of previous characters. From the hashtags generated in the previous phase, a hashtag is chosen as seed text and using the character sequences of this seed text, new characters are generated in sequence. That is, the model is trained to generate narratives by adopting the writing style in the corpus using the hashtag [12, 13]. This [4] paper formulates a hashtag recommendation task as a multi-class classification problem. Networks handle input microblogs of varying length. Each dimension of the output layer represents the probability of a hashtag recommended. As discussed in the introduction section, we also follow the trigger words assumption, which has been successfully used in previous studies. Hence, the proposed model incorporates two channels: a local attention channel and global channel.

III.HASHTAG AUTOMATION

The major drawbacks observed in the existing system are Large Training sets Required, Slow Convergence rate, High Computational Complexity, Not trending Hashtags. The proposed system includes use of pretrained model RESNET50 (which is present in keras API) for classifying different branded cars and providing proper hashtags for them using Transfer Learning.

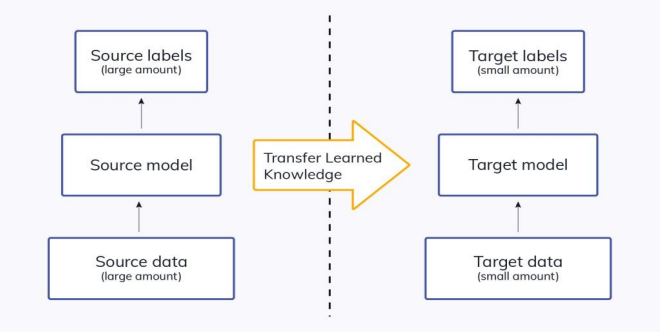


Figure 2 Transfer Learning Idea

The problem was to get trendy hashtags for set of flowers (five different categories), and there are pre-existing datasets of flowers in tensorflow datasets. But the problem with them was they were huge datasets nearly 1000 images for each category so computation for such big datasets was a heavy task and also there were fixed categories of flowers we could classify. So to solve above problems a custom dataset flowers of 3 categories has been created using Fatkun Chrome Extension as a tool to download images nearly 50 of each category. Then that images are preprocessed and converted into a numpy array which is of size 224x224 pixels(because the pretrained model used accepts images of that size) and the channel being RGB. Next few pre-trained model are selected for performing transfer learning and then comparing results of various pre-trained models and choosing the best one. And a transfer learning model is developed with number of class labels in the last layer of our neural network. The label generated is passed through a function called hashtag_generator() to provide respective hashtags. Instaloader is the API used to fetch TopSearchedHashtags on the instagram. In this step we basically do image rescaling,image Augumentation and also convert into numpy array(which are understood by neural networks). All the above mentioned functionalities are achieved by using ImageDataGenerator in Keras. Here we also put a validation split of 0.2 i.e., and at this step train,test and validation sets are created and are ready.

©IJRASET: All Rights are Reserved

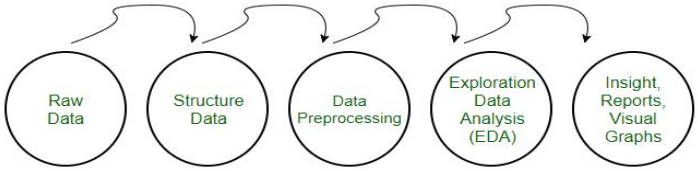


Figure 3 Data Preprocessing

Transfer learning is an approach where a model implemented for a particular task is reused as the initial point and to implement other tasks by modifying it which increases is the efficiency. Transfer learning is one of the popular approaches in deep learning and machine learning where already trained models are used as the starting point on computer vision and natural language processing tasks, or any image classification tasks. As classification tasks given the vast compute and time resources required to develop neural network models on these problems in order to reduce all the complexities transfer learning is a good solution.

In this model a pre-trained network called ResNet50(a Convolution Neural Network architecture) is considered which is trained with ImageNet, a dataset of over 14 million images in order to increase the number of images and added some more layers and remove some layer and replace it with other so it can classify skin lesions. The main idea behind this is that we take a model that was trained with a huge amount of data and use that for a new task where we don't have as much data or when we have less data this comes handy. By using this transfer learning as the model is pretrained with many images it works good not only with the user provided data set but also to the real time inputs.

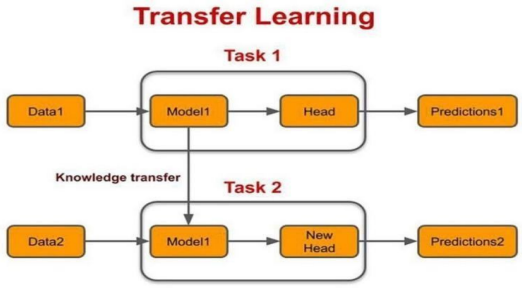


Figure 4 Transfer Learning

IV. RESULTS

To get good accuracies same above procedures are carried out with three different pre-trained models ResNet50, MobileNetV2, VGG16 and following are the result comparisons(accuracy and error rates). ResNet-50 is a convolutional neural network that is 50 layers deep. You can load a pretrained version of the network trained on more than a million images from the ImageNet database . The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224. VGG16 is a convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford in the paper. The model achieves 92.7% top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes. MobileNet is a type of convolutional neural network designed for mobile and embedded vision applications. They are based on a streamlined architecture that uses depthwise separable convolutions to build lightweight deep neural networks that can have low latency for mobile and embedded devices.

TABLE I
COMPARISON OF MODELS ACCURACY

Models	Training Accuracy(%)	Test Accuracy(%)	Number of Epochs
ResNet50	84.29	75.00	20
VGG16	96.30	86.25	20
MobileNetV2	97.40	96.25	20

©IJRASET: All Rights are Reserved

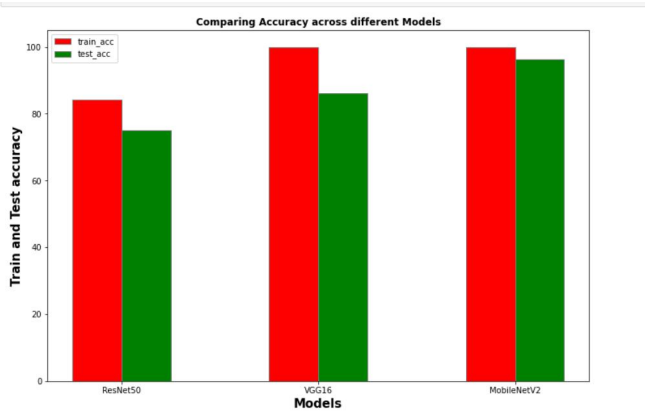


Figure 5 Barplot for comparing models

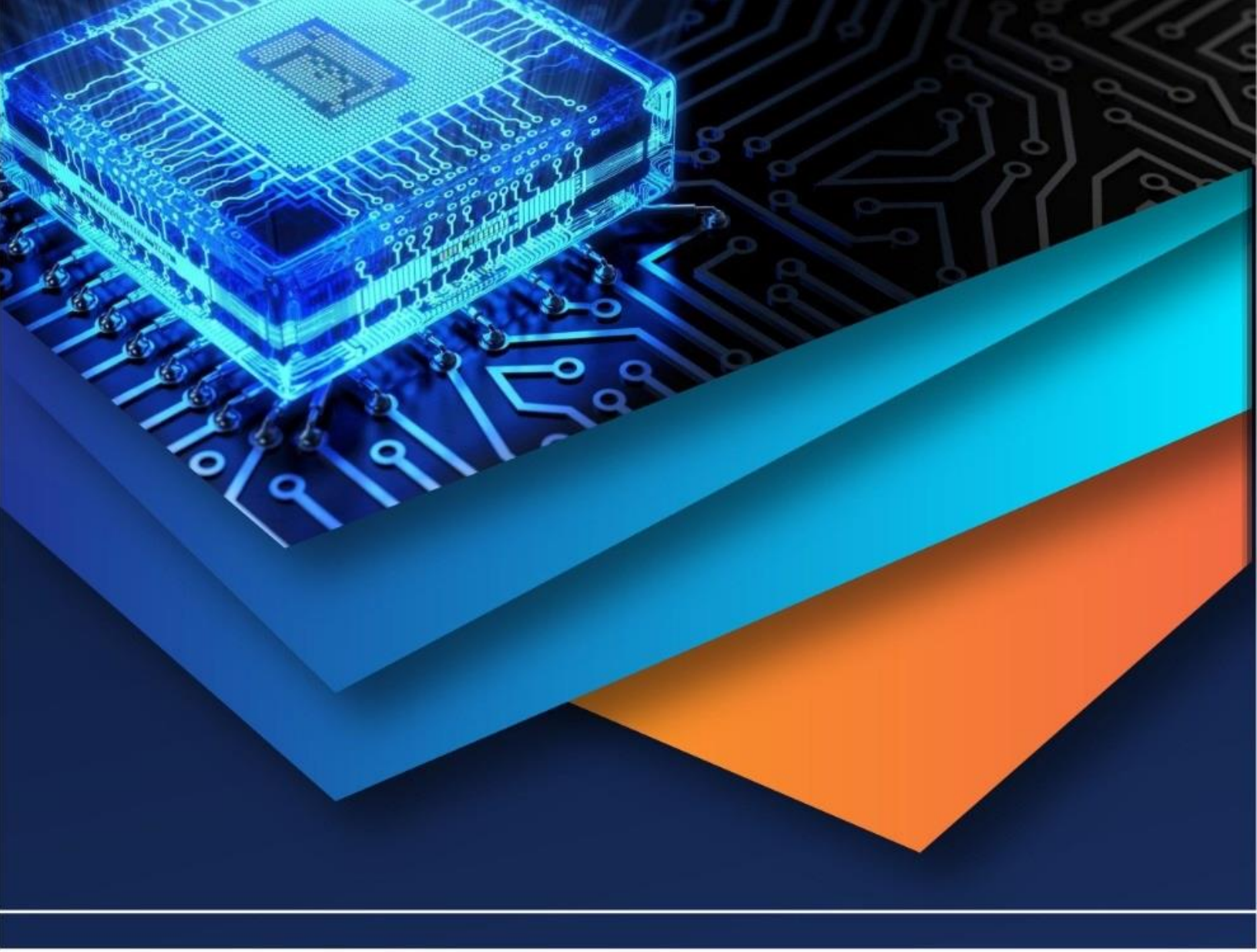
V. CONCLUSIONS

This paper provides trending hashtags for businesses to increase their reach in the market and also help users find their desired product. We have applied transfer learning model on three different pre-trained models and model with best accuracy was MobileNetV2 and it gave around 96 percent accuracy. And after getting labels from model they are fed to Instaloader API which suggests us top searched Hashtags on Instagram and all these are wrapped around flask server and results are displayed on a web application with simple UI. The project can be deployed using Heroku to gain a Real-time experience.

REFERENCES

[1] Kavinga Yapa Abeywardana, Ginige A.R., Herath N, Somarathne H.P, Thennakoon T.M.N.S, “Hashtag Generator and Content Authenticator”, IEEE, 2018.
[2] Shivam Gaur, “Generation of a short narrative caption for an image using the suggested hashtag”, IEEE, 2019.
[3] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhutdinov, R.Zemel and Y. Bengio, “Show, Attend and Tell: Neural Image Caption Generation with Visual Attention,” arXiv.org, 2016.
[4] Yuyun Gong, Qi Zhang, “ Hashtag Recommendation using Attention-based Convolutional Neural Network”, Springer, 2017.
[5] M. Park, H. Li and J. Kim, ”HARRISON: A Benchmark on HAshtag Recommendation for Real-world Images in Social Networks,” arXiv.org, 2016.
[6] J. Xu, X. Ren, Y. Zhang, Q. Zeng, X. Cai and S. Xu, ”A Skeleton-Based Model for Promoting Coherence Among Sentences in Narrative Story Generation,” arXiv.org, 2018.
[7] M. S. Lukin, K. Bowden, C. Barackman and A. M. Walker, ”PersonaBank: A Corpus of Personal Narratives and Their Story Intention Graphs,” arXiv.org, 2017.
[8] X. Wang, W. Chen and W. Yuan-Fang, ”No Metrics Are Perfect: Adversarial Reward Learning for Visual Storytelling,” arXiv.org, 2018.
[9] E. Denton, J. Weston, M. Paluri, L. Bourdev and R. Fergus, ”User Conditional Hashtag Prediction for Images,” in ACM SIGKDD International Conference on knowledge discovery and data mining, 2015.
[10] C. Park, B. Kim and G. Kim, ”Attend to You: Personalized Image Captioning with Context Sequence Memory Networks,” arXiv.org, 2017.
[11] A. Veit, M. Nickel and S. Belongie, ”Separating Self-Expression and Visual Content in Hashtag Supervision,” arXiv.org, 2017.
[12] L. Wang, X. Chu, W. Zhang, Y. Wei, W. Sun and C. Wu, ”Social Image Captioning: Exploring Visual Attention and User Attention,” Sensors, vol. 18, no. 2, 2018.
[13] R. Kiros, Y. Zhu, R. Salakhutdinov, S. R. Zemel, A. Torralba, R. Urtasun and S. Fidler, ”Skip-Thought Vectors,” arXiv.org, 2015.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)