Study Of Image Recognition Based Travel Destination Recommendation

A Liyanage
Department of Industrial Management
University of Kelaniya
Sri Lanka
liyanage_im15022@stu.kln.ac.lk

S Jayalal
Department of Industral Management
University of Kelaniya
Sri Lanka
shantha@kln.ac.lk

***Abstract*— Sri Lanka is a country with different types of tourist destinations. However, lack of popularity to many such attractive destinations is a major issue identified. Images of Sri Lankan destinations are published in Social Media platforms, often without the correct geo-location tagging. Therefore, most of the places are not popularized among international and local tourists. Key problem domain identified would be the inability to recognize specific location of such images. Due to high similarity and complex nature of images related to this field, effectiveness of existing methodologies has become questionable. Focus of this research would be a suitable image identification methodology-based model, which can identify an image based on their destination and to provide with advanced recommendations. Thereby, it would ensure promotion and focus towards attractive, yet widely unpopular destinations in Sri Lanka. However, in general this is not problem specific to Sri Lanka, but other countries like Sri Lanka as well. Therefore, in large, this model can be applicable to similar other countries as well, even though data and evaluations of this research are from Sri Lanka.**

Keywords— **Image recognition, Convolutional Neural Networks, Geo location, Destination identification.**

I. INTRODUCTION

A. *Tourism Industry of Sri Lanka*

Tourism is one of the most desirable industries in Sri Lanka due to the high income it provides. It is shown in the Fig. 1, Tourism Industry of Sri Lanka showed a drastic increase of foreign exchange it brought into the country with the end of 2010 up to 2018. Through the past few years tourism industry of Sri Lanka showed a drastic development with many local and foreign travellers exploring different parts of the country [1]. After 2019, there was an improvement of the number of tourists coming to the country, which has been again affected by the worldwide Pandemic situations [3]. Therefore, once the situation subsides, more focus has to be given to uplift tourism in the country. Further, it would be prominent to ensure arrivals of foreign travelers to Sri Lanka.

B. *Social Media Usage*

When an image is uploaded to most Social Media sites, it is possible to add the location of the destination. Most popular social media in Sri Lanka is Facebook, which is followed by Twitter, YouTube and Pinterest [4]. Most of these social media are image sharing platforms. They are used by travelers and photographers to publish their content, created in different areas of the world.

However, the current possibilities of identifying those images to use them in benefit of improving tourism to a certain place are low. There are few reasons for this; incorrect location tags, images without location tags or images having location tags without a specific location, but a general location etc. [5].

C. *Popularity and attractiveness of destinations*

There are many different places of travel in Sri Lanka. However, among many places to travel and explore in Sri Lanka, only very few places are popular among a majority of travellers. Being a country with a legacy of thousands of years, there are many places of historical and cultural importance as well [6]. As a justification for the identified research problem, an evaluation was done through different statistics available. The assumption is that some of the best and important places to travel are often not visited or known by a majority of the local and foreign travellers. Two main data sources were based for this evaluation. They are Google Map place ratings [7] and Sri Lanka Tourist Development Authority Annual Reports [1].

Google Maps contain details of different locations around the world. First as a justification of the problem domain, some destinations are analyzed with the reviews of the Google Map data, place ratings and number of ratings provided. These results can be used for a comprehensive analysis of popularity and attractiveness of different destinations [7]. A ranking of popularity was created for analysis purposes. For the context similarity, all chosen destinations are waterfalls in Sri Lanka. With the number of reviews received through Google Map place reviews, following ranking was created. Places with 1000 reviews are considered as having High Popularity. Places having between 400 to 1000 reviews are considered ad having Average Popularity. Places with less than 400 reviews are considered to have Low Popularity [7]. There is a Rating System for places in terms of Stars**.** This can be used for the identification of attractiveness/ importance of a particular destination. For this analysis, it is assumed that places with ratings above 4.0 are comparatively of high attractiveness. Table I contains results of analysis, through which it is evident that there are certain important places with lower popularity, which can be identified as justification of this research problem. There can be many reasons for certain places being unpopular. However, one of those reasons would be less popularity given to those images in Social media.

II. LITERATURE REVIEW

A. *Social Media related Review*

In a research on Twitter data of the text-based geo location, it has been identified that text-based geolocation accuracy

TABLE I. EVALUATION OF POPULARITY OF DESTINATIONS

Destination	Reviews	Popularity	Rating
Bambarakanda Falls	1515	High	4.8
Aberdeen Waterfall	795	Average	4.8
Galboda Falls	145	Low	4.7
Dumbara ella	57	Low	4.9

depends on population demographics as well, such as age factor [8]. Around 4.3 percent of the images in Flickr are geo-tagged with the latitude and longitude values. Therefore, it is approximately 180 million geo-tagged images, which when considered as a percentage of total images uploaded is very low [9]. Through [10] it is evident that accuracy of geotag information is highly dependent on the popularity of the location. This research is done based on image data in Flickr.

B. Image Recognition Technology and Application Review

Though there are many images that are available online there are only few technologies that are using these images for image recognition, object identification or categorizing for the use of those images in a beneficial manner. However, there have been some research into this arena, which has led to development some tools. These tools are generally known as Reverse Image Search Applications or web browser extensions etc. Some of the applications that are available for reverse image searching include Google Reverse Image Search, Tin Eye Search Engine, Bing Image Match etc. [22]. Out of these the most popular and the used one is Google Reverse Image search. To create multileveled equations, it may be necessary to treat the equation as a graphic and insert it into the text after your paper is styled.

Through one research [11] authors have analyzed that Reverse image search can be used to search either data related to the query image or the images related to that image or exact images. Research on a specific domain related images, to analyze the accuracy of reverse image search for images of that domain could be highlighted [12]. This particular research queried Google's ability to accurately identify similar images with matched diagnoses. Google Reverse Image Search permits users to upload an image and search for similar images and content.

This research analyses the accuracy of the image recognition done by Google reverse image search on their domain. They identify that there are error-prone results generated by the generic algorithm. Therefore, in building a more accurate and domain specific system, it is of wide importance to rationalize the initial hypothesis that the existing algorithm is error some or sometimes maybe less accurate. Therefore, this particular research is important as it shows that there are issues of accuracy [12] .

C. Algorithmic Review

There are some researches done about the existing algorithms that could be used for the image and object recognition. Key focus of this review would be towards Convolutional Neural Networks (CNNs) and the researches done on analysis of different architectures and relative comparison of accuracy and performance.

Research on analysis of convolutional neural networks for image classification compares and contrasts the performances of VGG Nets, GoogLeNet and ResNet50. The final analysis of this research shows that GoogLeNet and ResNet50 are able to recognize objects with better accuracy compared to Alex Net. Further, it shows that performance of trained Convolutional Neural Networks is varying based on the categories of objects used in training [13].

Then there is this particular research which identifies and introduces the Siamese Neural Network as important and effective for the image classification and identification [14]. This research is focused on exploring a method for learning this Neural Network architecture as it uses a natural ranking method based on similarity of its given inputs. According to the author, once a network is tuned, it is possible to generalize the power of prediction. Its usefulness would not limit to new data, but also to new classes of data from new unknown distributions of data. This research is important since it discusses the CNNs for image recognition, based on Siamese Neural Network to identify similar images in a data set. Therefore, it could be understood that Siamese Network is a novel area with higher performance and accuracy.

When Siamese Neural Networks are considered, research on Snake classification by Abeysinghe et. al. is one of the most recent researches which has highlighted the key importance of Twin Network/ Siamese Network in 'One-shot image classification'. They identify this network architecture as a yet virgin area of research in image recognition, where it could be implemented in many challenging classification functions in the future [15].

There are several researches that have identified VGG Nets including VGG16, VGG19 as important architectures for the image recognition and classification. Another medical based research [16], indicates imaging modality by VGG Nets and ResNet50, through analyzing their behavior related to breast cancer images. Here they have identified through experimental results that VGG16 shows the highest performance with 92.6% accuracy.

The original paper on development of AlexNet for ImageNet, contains the basic ideas of AlexNet and its underlying concepts [21]. Even though, at that time of the development of CNNs it was a very accurate and high performing architecture, later onwards there were some limitations in the architecture. Removal of any one of Convolution Neural layers in AlexNet would reduce its performance challengingly [17].

Therefore, VGG16, VGG19 and AlexNet are architectures, standout in identification of images from challenging data sets [17]. Further, Siamese Networks standout in providing a Similarity Score related classification for the image data sets with considerable validity.

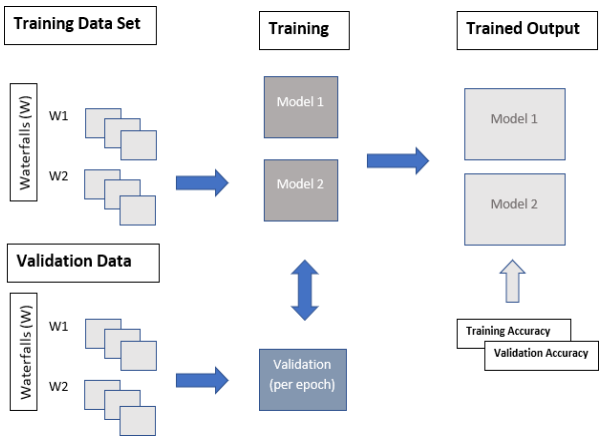

Fig. 1. Model implementation overview

III. Data Gathering

Data required for a research is unique from research to research. For this research the focus area is images of locations in Sri Lanka. Therefore, images of different tourist destinations are to be collected for the implementation of this research. (Fig.2)

Data are collected through different professional photographers of Sri Lanka, and some of the images were collected from the online sources, with their due permission, for the training and validation purposes.

Fig. 2. Sample image data

A. Constraints on data collection

There are many key destinations in the world for tourists. In Sri Lanka alone, there are hundreds of different places of attraction, importance and popularity**.** Therefore, as for the development of this research model, initially, Waterfall images are selected.

IV. Methodology

The main aim of this research is to find an effective solution for identifying untagged images in different social media and other online platforms. Expected outcome of this research methodology selection and implementation process would be to improve the tourism of the country by allowing people to identify unpopular places they discover in online platforms. One model would be to develop an accurate image recognition approach. Further the second model would be to recommend similar images for the related image. This second model would be the recommendation system for travellers. Fig.1 shows an overview of the research methodology and model training.

A. Model 1

Initial approach to this research would be to analyze the effectiveness of some existing applications in this regard, as a continuation of the findings in Literature review section. There are some platforms that do reverse image searches and other image recognitions [12] . But their accuracy in classification of images in complex data sets is questioned. Therefore, for this research image data set, a specific model with higher accuracy has to be trained.

Three different architectures are used for this research. Thereby, training and comparing the results of three models for the highest performing architecture for the model.

VGG 16 is one of the best Convoluted Neural Network (CNN) architectures for vision modeling developed recently [19]. VGG16 is named after its sixteen layers with weight. The highlightable difference between VGG16 from some of the popular CNN architectures would be absence of large number of hyper-parameters. Instead, its architecture is developed on convolutional layers of three by three filter with a stride one. Same padding is used every time. Further, a Maxpool layer of two by two filter with stride two is included. Then, throughout the whole architecture, this similar arrangement is repeated consistently. At the end of these layers, it includes a two Fully Connected Layers (FC). Then for the output of the architecture, a SoftMax is used. This is a very large network, which includes approximately 138 million parameters [19]. VGG19 consists of 19 deep layers and can be downloaded as a pre-trained model on some of the common images found on Image Net. Since the images used for this research are unique, this model can also be used for training and classifying new images. AlexNet is a much complex model to other CNN architectures, which has used multiple GPUs in developing the model. Other than that, there are some important features in AlexNet. This Architecture contains 8 layers, five layers are CNNs and three are Fully Connected (FC) layers. Some such unique features of AlexNet include ReLU Nonlinearity, multiple GPUs, Overlapping Pools etc. Major issue in AlexNet is Overfitting which are tried to be avoided by Data Augmentation and Dropout [17].

B. Model 2

The similar location identification model can be trained with a deep learning approach too. Here, through the researches reviewed, Siamese Net algorithms are used as high performing and accurate algorithms. Also, they are popular recently in facial recognition models as well [15]. Here, the output of the training process would be each image class being given a similarity score.

There are two Convoluted Neural Networks in this. These are two copies of the same network. They are sharing same parameters. The two input images are passed through the two ConvNets to generate feature vectors of fixed lengths for each image. When images are considered, if two images are similar in features, they must have same vector value in a properly trained model [20]. Therefore, in the next layer, a difference in the vector values is taken, absolute difference between two vectors when images are similar and when images are different must also be very different. Then through a Sigmoid layer, a similarity score is generated. It also depends on the similarity of the images [15]. Then the similarity score gained from this model would be mapped through a Vector space to a dimensional space(128 dimnesions). Through this all the different image classes are mapped with their Similarity Scores. Thereby, when an input image is gaining a similarity score, the images with features similar to that image can be identified with values that care close to it in the mapped space. Thereby, the final output would be three similar images to the image that is inputted to the system.

TABLE II. Results Evaluation Model 2

Model Arch.	Training Accuracy	Val. Accuracy
Siamese N. N.	0.8860	0.9126

TABLE III. RESULTS EVALUATION MODEL 1

Model Arch.	Training Accuracy	Val. Accuracy
VGG16	0.8566	0.8820
VGG19	0.7921	0.8431
AlexNet	0.8021	0.8456

V. DISCUSSION OF RESULTS

The results of training of the above two models can be interested through this section. First, the results of model 1 would be discussed. In the first model a comparative analysis of the three architectures were done, of which the results are so far, the training accuracy and the validation accuracy for image data sets of 40 images of 30 classes. 30 classes used for training are images of the 30 waterfalls in Sri Lanka.

Each model was trained with a training data set as well as some validation data, in order to receive the validation accuracy per epoch, after every training epoch. Number of epochs used for the training was 10 for each model. Each model was trained thrice with the same dataset for valid results of the final model. Obtained output of the training and validation accuracies for the three architectures are included in Table III. Through the above initial training results of the models, it is evident that VGG16 based architecture shows higher accuracy for the provided data set.

When we focus on the second model, it was implemented with just one architecture. A pre trained model was not used for this model. Therefore, the development and training were complicated and time consuming than the previous model. It was also trained for the same data set of 30 classes of waterfall images. It gave a final training accuracy and validation accuracy as shown in Table II. As it is evident, this model also provides a reliable accuracy.

VI. CONCLUSION

In conclusion, since a higher accuracy and performance is shown in the research done so far, the two models would be able to use in real time implementations, for the aim of promotion and improvement of tourism in Sri Lanka. Also, the models can be further improved in the future.

VII. FURTHER IMPROVEMENTS

There are further trainings to be done in order to show the accuracy of the model with many different types of Sri Lankan destinations. It is expected to train the models again with an image data set of higher number of images and changing the number of classes to be trained, to check the withholding of the conclusion arrived during the above results discussion. Thereafter, the model can be improved into a real time application, which can be integrated with existing tools or as a separate new application. (fig. 3)

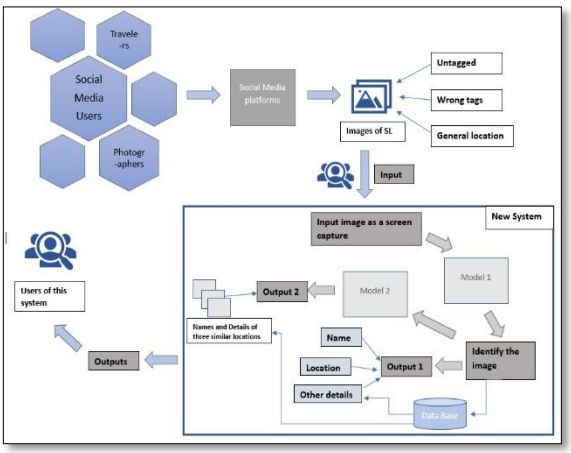

Fig. 3. Model implemetation

REFERENCES

[1] Sltda.gov.lk. 2019. SLTDA | Sri Lanka Tourism Development Authority. https://www.sltda.gov.lk/annual-statistical-report, accessed 12/07/2020

[2] Tuazon, B., O'Key, S. and Simon, D., 2020. *Terror In Sri Lanka*, https://edition.cnn.com/interactive/2019/04/world/sri-lanka-attacks/

[3] Lasanda, K., 2020. *Sri Lanka'S Covid-19 Response And The Challenge Ahead*. http://www.dailymirror.lk/opinion/Sri-Lankas-Covid-19-response-and-the-challenge-ahead/172-189816

[4] StatCounter Global Stats, 2019. *Social Media Stats Sri Lanka*. https://gs.statcounter.com/social-media-stats/all/sri-lanka

[5] Liu, Bo & Yuan, Quan & Cong, Gao & Xu, Dong., 2014. Where your photo is taken: Geolocation Prediction for Social Images. Journal of the Association for Information Science and Technology. https://www.researchgate.net/publication/264716139_Where_your_photo_is_taken_Geolocation_Prediction_for_Social_Images

[6] Sltda.gov.lk. 2020. *SLTDA | Sri Lanka Tourism Development Authority*. https://www.sltda.gov.lk/projects

[7] Google Maps, 2020. https://www.google.com/maps

[8] Umashanthi Pavalanathan & Jacob Eisenstein, 2015. Confounds and Consequences in Geotagged Twitter Data. School of Interactive Computing, Georgia Institute of Technology.

[9] Jaeyoung Choi, Istemi Ekin Akkus, Serge Egelman, Gerald Friedland, Robin Sommer, Michael Carl Tschantz and Nicholas Weaver, 2018. Cybercasing 2.0: You Get What You Pay For. Cornell University.

[10] Claudia Hauff, 2013. A study on the accuracy of Flickr's geotag data. Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval.

[11] Chutel, P. and Sakhare, A., 2020. Evaluation Of Compact Composite Descriptor Based Reverse Image, https://www.semanticscholar.org/paper/Evaluation-of-compact-composite-descriptor-based-Chutel-Sakhare/b285fc8182bdec386ad8e12d85c442ee6ced51a7

[12] Julia D. Ransohoff, Shufeng Li, Kavita Y. Sarin, 2016. *Assessment of Accuracy of Patient-Initiated Differential Diagnosis Generation by Google Reverse Image Searching,* https://jamanetwork.com/journals/jamadermatology/article-abstract/2529572

[13] Sharma N., Jain V., Mishra A., 2018. An Analysis Of Convolutional Neural Networks For Image Classification. https://www.sciencedirect.com/science/article/pii/S1877050918309335

[14] Gregory Koch, 2016. Siamese Neural Networks for One-shot Image. https://www.cs.cmu.edu/~rsalakhu/papers/oneshot1.pdf.

[15] Abeysinghe, Chamath W, Anuradha P, Indika. (2019). Snake Image Classification using Siamese Networks.

[16] Shallu, Mehra R, 2018. Breast cancer histology images classification: Training from scratch or transfer learning? https://www.sciencedirect.com/science/article/pii/S2405959518304934

[17] Wei J., 2019. *AlexNet: The Architecture that Challenged CNNs. https://towardsdatascience.com/alexnet-the-architecture-that-challenged-cnns-e406d5297951*

[18] [Almryad, & Kutucu, 2020]Almryad, A & Kutucu, H. (2020). Automatic identification for field butterflies by convolutional neural networks.

[19] Thakur, R., 2019. Step By Step VGG16 Implementation In Keras For Beginners. [online] Medium.

[20] Lamba, H., 2019. *One Shot Learning with Siamese Networks using Keras*. https://towardsdatascience.com/one-shot-learning-with-siamese-networks-using-keras-17f34e75bb3d.

[21] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks

[22] Bhangu G, 2020. 11 *best reverse image search engines. https://otechworld.com/reverse-image-search-engine*