Generation of a short narrative caption for an image using the suggested hashtag

Shivam Gaur
Master of Computer Science (Management)
The University of Queensland
Brisbane, Australia
s.gaur@uqconnect.edu.au

***Abstract*—Existing methods for image captioning aim to generate captions in natural language by using the image attributes. Though these captions are enough to explain what the image is about in most cases, yet sometimes more than a single sentence is desired to describe the context of an image especially when a good caption for the image draws more public attention or 'likes' on a social media post. Though some work has been done to develop models that can generate hashtags for images, but no research work exists that can use those hashtags to create story-like captions. A hashtag can be defined as a word preceded by the symbol '#' and is used to identify an image on social media sites like Instagram. The goal of this application paper is to explore the possibility of generating hashtags for an input image and leverage it to generate meaningful anecdotes connecting to the essence of the image. The experiment conducted, uses an attention-based encoder-decoder framework to produce hashtags for the raw image while a character-level language model, which is trained using a multi-layer RNN, is used to generate stories using one of the suggested hashtags. The model was then tested on HARRISON dataset of Instagram images and the results were qualitatively analyzed through a user study. After analyzing the outcomes of the experiment, it was concluded that this area of research has huge prospects.**

***Index Terms*—Image caption, Hashtag, Story, Attention model, Language model**

I. Introduction

The process of comprehending the relationship among the various components of an image and generating a sentence or a set of keywords in a natural language to describe those interdependencies is formally called as 'Image Captioning'. The various components that make up an image can be categorized as tangible objects such as people, animals, food items, etc. or intangible entities or concepts like a place, mood of the image, scene in the background, etc. For human beings, the ability to perceive the message conveyed by an image is innate and the skill to put the message into words is learnt with age without a lot of effort. However, the said tasks are not easy for a machine that only speaks the language of zeros and ones [1]. Image Captioning is one of the challenging problems of Artificial Intelligence and AI researchers have been spending a lot of time to develop models with improved efficiency day and night. The fact that this field of AI deals with the concepts of Computer Vision and Natural Language Processing (NLP), makes Image Captioning a difficult yet interesting task. The knowledge of Computer Vision is necessary to identify and segregate various attributes of an image while the NLP algorithms are required to find and represent the relationship between the attributes in a language understood by human beings. This process of automatic generation of captions for an image using AI finds various applications like automatic generation of annotations for images, hashtags and captions for social media posts, description for a scene in an image, etc.

Recently, a lot of focus has been given to generate captions for images posted by people on social media sites like Twitter, Facebook, Instagram, etc. Various models have been developed that can generate effective and meaningful sentences with high accuracy. Though these sentences are enough to tell the user what the image is about in most cases, yet sometimes more than just a sentence is required to describe the context of image especially when a good description of the image draws more public attention or 'likes' on social media post. In addition to this, various attempts have also been made to generate relevant hashtags as captions instead of a sentence in natural language as it is a widely known fact that good 'Hashtags' are responsible to make a post to reach to out to millions of people in the virtual world. A hashtag is any word preceded by a '#' and is used to identify images on social media sites like Instagram [2]. Hashtags are increasingly becoming popular on various social media platforms like Facebook, Twitter and Instagram to summarize user posts and attract engagement [2]. They can be used to indicate location, object or an organization, express emotions, categorize topics, infer context or situation [2]. Many businesses use hashtags for marketing purposes on Instagram. Hence, it will not be an exaggeration to say that hashtags have brought a revolution in today's virtual world. This work aims to explore the possibility of using hashtags to produce story-like captions for an image on social media which could add more meaning to the user post, and hence attract more followers. The objective of this paper is to demonstrate the outcomes of an experiment on image captioning wherein the goal is to generate relevant hashtags for an image and leverage those hashtags to generate succinct story-like caption for the image. This is achieved by using an attention-based [1] framework which will attend to different

2473-3490/19/$31.00 ©2019 IEEE
DOI 10.1109/ICDEW.2019.00060

attributes of an image and generate meaningful hashtags based on the set of feature vectors obtained and then use a character-level RNN language model to generate a short-story using one of the suggested hashtags as the seed text [5].

II. Related work

This section gives an overview of work done related to image captioning using attention mechanism. It then moves on to discuss some of the notable research done in generating hashtags as captions and finally throws light on the prospects of generating story-like captions for an image.

This paper aims to generate image captions using attention mechanism discussed in [1]. The attention mechanism is a powerful framework to produce captions as the model focuses on only the most important attributes of the image and discards all noise. The CNN-LSTM framework used in [1] forms the backbone of the experiment done in this work. Recently, researchers have started paying attention to use image captioning models for producing relevant hashtags for images due to ever increasing importance of hashtags in the social network platforms. Authors in [7], leverage hashtags from user posts and user's personal information such as age and gender to predict customized high-quality hashtags for images posted on Facebook. In [8], authors introduce a novel captioning model. They introduce, Context Sensitive Memory Network to store context information. They aim to generate personalized image captions and hashtags for user's post on Instagram by using the user's own vocabulary, writing styles and expressions [8]. In addition to this, another model was introduced in [9] where it makes use of image-user-hashtag triplet to predict hashtags specific to user's hashtag usage pattern. This is another example of personalization where relevant hashtags are produced that are suitable for a user [9]. Another example of generation of personalized image captioning is introduced in [10], where user tags are leveraged along with the attention on image attributes to produce good quality personalized image captions. By using dual attention mechanism, noisy tags are discarded while generating captions [10]. All the works, discussed above have tried to address the same issue, i.e., generation of personalized image captions and hashtags. Different from the rest, Ryan Kiros et al. built an artificial neural network based, 'neural-storyteller' [11], which makes use of skip-thought vectors to generate little stories as captions for images. The approach consists of keeping the 'thought' of caption but replacing the caption with the passage of a romance novel [11].Another work mentioned in [3], introduces a novel method of narrative story generation using the skeleton-based model [3]. The model looks for the key phrases in input captions [3]. These phrases are then used to generate a skeleton which then expands to generate fluent short stories using reinforcement learning method [3]. Another example of visual storytelling is introduced in [6], where narratives are generated from a sequence of image captions. The model takes into account, a series of images and associated captions and then leverage those to generate a new story-like caption [6].

This work is unique because it leverages the hashtags to generate stories for images as captions using character level language model introduced by Karpathy in [5].

III. Methodology

The objective of this work 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. The following sections discuss the two processes in brief.

A. Hashtag generation

To generate hashtags for the image, the attention-based framework introduced in [1] 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

$$\mid X \mid \leq 1$$

. X is encoded as a sequence of 1-of-K encoded words [1].

$$X = \{x_1, x_2, x_3, x_4.....x_N\}, x_i \epsilon R^K$$

where K is the size of the vocabulary and N is the number of hashtags generated for the image. The encoder extracts features from the image from lower layers of CNN [1] which helps the decoder in paying attention to different regions of the image by selecting different subset of the feature vectors [1]. The set of L feature vectors can be represented as:

$$A = \{a_1, a_2, a_3, a_4....a_L\}, a_i \epsilon R^D$$

and is the D-dimensional representation of a part of the image. The decoder uses a LSTM that is responsible for producing a hashtag (one word) at each time step t, which is conditioned on a context vector z_t, the previous hidden state h_{t-1} and the previously generated hashtag.
The LSTM framework used can be represented by the following equation [1].

$$\begin{pmatrix} \mathrm{i}_t \\ \mathrm{f}_t \\ \mathrm{o}_t \\ \mathrm{g}_t \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} T_{D+m+n,n} \begin{pmatrix} \mathrm{Ey}_{t-1} \\ \mathrm{h}_{t-1} \\ \mathrm{z}_t \end{pmatrix} \quad (1)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (2)$$

$$h_t = o_t \odot tanh(c_t) \quad (3)$$

In the above equations, i_t, f_t, o_t, h_t denote input, forget, output and hidden state of the LSTM, respectively [1].The context vector $z_t \epsilon R^D$ represents the visual information gathered from a particular image location at time t. E is the embedding dimension where $E \epsilon R^{mXK}$ [1]. σ is the logistic sigmoid activation and $\odot$ is the element-wise multiplication operator

[1]. In this paper, the soft attention mechanism is used where weighted image features are used as input to the the LSTM to account for attention. ϕ is a function to calculate context vector z_t from annotation vector a_i obtained from features extracted from the location i. Let α_i be a weight associated with each location i and intuitively it can be defined as the probability that the location i is the right place to pay attention for producing the next hahstag [1]. The weight α_i is calculated by the attention model f_{att} using a multilayer perceptron conditioned on the previous hidden state h_{t-1} [1].
In soft attention method, context vector $z_t = \phi(\{a_i\}, \{\alpha_i\})$ where $\phi(\{a_i\}, \{\alpha_i\}) = \sum_{i=1}^{L} \alpha_{t,i} a_i$.

$$\alpha_{t,i} = \frac{\exp(e_{ti})}{\sum_{k=1}^{L} \exp(e_{tk})} \tag{4}$$

$$e_{ti} = f_{att}(a_i, h_{t-1}) \tag{5}$$

B. Story generation

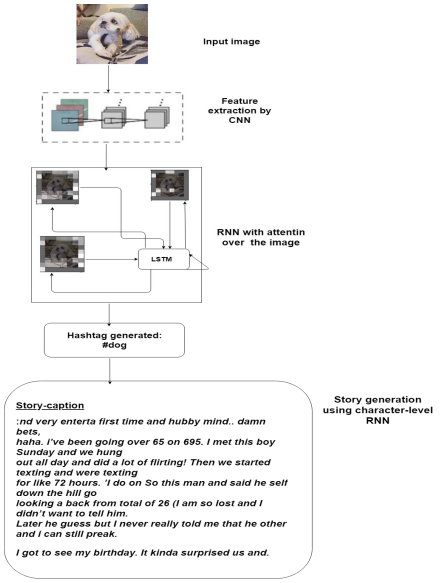
To generate story-like captions from the hashtags generated above, an character-level Recurrent (RNN) Neural Network is trained on a corpus of personal narratives. The RNN models the probability distribution of the characters in sequence given a sequence of previous characters. From the hashtags generated 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. The RNN is trained with mini-batch stochastic gradient and RMSProp optimizer. The RNN makes use of the recurrent network to remember the context before generating the next character in sequence. While generating the story, the character sequence in the seed hashtag helps decide the model the context and new text to be generated. The Fig 1. shows the work flow of story generation from an input image.

IV. Experiment

This section describes the experimental setting which can be divided into two phases. In the first phase, the objective is to generate relevant hashtags for the image whereas the second phase deals with the generation of story for the image from the hashtags obtained in the previous phase. Once the results are obtained, they are then discussed and analyzed qualitatively.

A. Phase 1: Hashtag generation

1) Dataset: The attention-based model is trained on HARRISON dataset. HARRISON is an acronym for ‘Hashtag Recommendation for Real-World Images in Social Networks [2]. The dataset consists of 57,383 images and around 260,000 hashtags [2]. Every image has minimum of 1 hashtag and maximum of 10 hashtags associated with it and the average number of hashtags per image is 4.5 [2]. The ground truth hashtags for each image are made up of 1000 most frequently occurring hashtags [2].


Input image
Feature extraction by CNN
RNN with attentin over the image
LSTM
Hashtag generated: #dog
Story generation using character-level RNN
Story-caption

:nd very enterta first time and hubby mind.. damn bets,
haha. i’ve been going over 65 on 695. I met this boy Sunday and we hung
out all day and did a lot of flirting! Then we started texting and were texting
for like 72 hours. ’I do on So this man and said he self down the hill go
looking a back from total of 26 (I am so lost and I didn’t want to tell him.
Later he guess but I never really told me that he other and i can still preak.

I got to see my birthday. It kinda surprised us and.

Fig. 1. Story generation using hashtag

2) Implementation details: The objective of this phase is to generate a caption in the form of hashtags for the a given image. This is achieved by using an attention-based encoder-decoder framework for image caption generation as mentioned in [1]. The reason for using attention model is that model focusses more on important attributes of the image while generating caption. The entire network is trained from end-to-end. InceptionV3 (pretrained on Imagenet) is used to classify images in the HARRISON dataset and features are extracted from the last convolutional layer. To do this, first image is resized to (299,299) and pixel range is set from -1 to 1. The feature vector of every image is obtained from the last convolutional layer whose output is of the shape: 8X8X2048. Before training the model, the hashtags associated with each image are tokenized to create a dictionary of unique tokens. The vocabulary size is limited to top 5000 hashtags and all other tokens are replaced by the word “UNK”. Next a word to index mapping is created and vice-versa. Next, the model is trained on 80 percent of data while the remaining is used for validation. The feature vector obtained from lower convolutional layer of InceptionV3 is converted into 64X2048 and then passed to the CNN encoder which is made up of single fully connected layer. The RNN then predicts next word for the caption using attention mechanism. Teacher forcing is used to determine the next input to the decoder. The table I summarizes the parameters of the model trained on a GEForce GTX Titan GPU to generate hashtag.

B. Phase 2: Story Generation

1) Dataset: The character -level RNN model is trained on ‘PersonaBank’ corpus which is a collection of 108 personal narratives from various weblogs [4]. These stories cover a wide range of topics from romance and wildlife to travel and sports.

TABLE I
PARAMETERS OF MODEL FOR HASHTAG GENERATION

Batch Size	1
Buffer Size	1000
Embedding Dimension Size	256
GRU Units	512
Optimizer	Adam
Epochs	10
Time per epoch	2.5 hours approx.

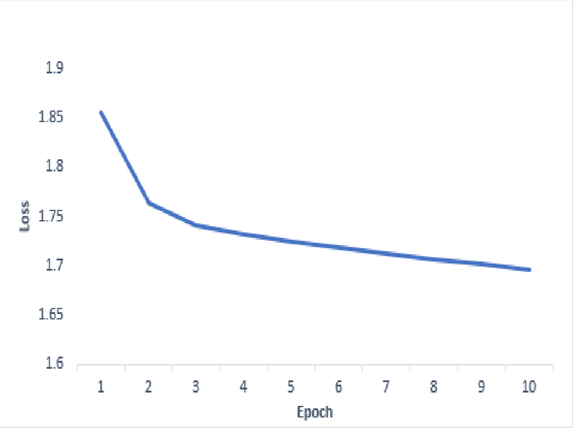


Fig. 2. Plot of Epochs vs. Loss for hashtags generation

Out of 108 stories, 55 are positive stories while the remaining are negative [4]. Average length of story in the corpus is 269 words [4]. The reason for choosing such a dataset is that the stories are real from real people and can connect well to the images on social networking sites which are real posts from real users, thereby conveying meaning to the caption generated from the seed hashtag.

2) Implementation details: A Recurrent Neural Network can be trained over a large corpus of text to build a probability distribution over the next character in the string of text, given the previous characters of the string [5]. The idea is to build a character-level language model by training the RNN on the ‘PersonaBank’ corpus so that it models the probability distribution over the next character in the caption (narrative from the corpus) given a sequence of previous characters in the form of hashtag. This is purely a classification task where the input is a series of one hot vector encoded characters, each fed into RNN one at a time and the output is a class in the form of one hot vector encoded target character in the vocabulary [5].

The language model is trained using a standard categorical cross-entropy loss. Let ‘t’ be the vector representing the one hot encoding of the output character and ‘v’ be the vector of probability distribution over ‘n’ characters in the ’PersonaBank’ vocabulary, then the loss for single character is equal to $-\sum_{i=1}^{n} t_i \log(v_i)$

The aim is to generate the model that can capture the writing style of the authors present in the ’PersonaBank’. Hence, the story like captions generated using this model will be very close to real life writing styles of real people. The table II summarizes the parameters of the language model trained on a GEForce GTX Titan GPU to generate stories with 500 characters in length.

TABLE II
PARAMETERS OF LANGUAGE MODEL FOR STORY GENERATION

Batch Size	5
Sequence Length	1
Embedding Dimension Size	1024
LSTM Layers	2
Optimizer	RMSProp
Activation Function	Softmax
Epochs	20
Steps per epoch	100
Learning Rate	0.01

V. PERFORMANCE ON HARRISON DATASET

A. Positive results

Following are some of the examples of good captions produced for images in validation dataset. These captions are either closely related or slightly related to the context of the image.

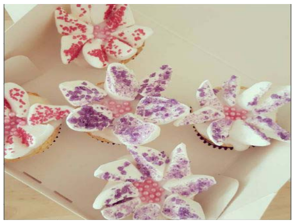


Fig. 3. Hashtag generated: #cake
Caption from hashtag: We sat the fact that I had not been lucky enough in the magic is on the horizon and continued talking with him, and he thought she asked me out, he is hours away as we spend of the night John and Anne# e# mame-as a love you with a violinist playing in the corner the hall. It was sit of it and set it up idea forth the snow and else thought of getting a small one, but my budget is rather tight and in the elevator at Gillman and shouted ’Oh, crumbles!!!’

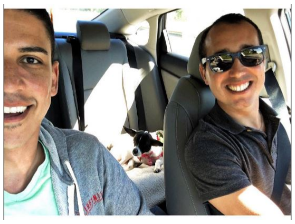


Fig. 4. Hashtag generated: #selfie
Caption from hashtag: ed in the car and down with my sister anytime my parents would leave the inside of the jar. Jill### tapped the jar firming goes on the car. But the 40-you tailed the went to the total to $75, and over 50 year and it was overgrown with the song was in. Luckily our group we all have to go again LOL It was about time in still attacts and staircase of my apartment building, hailed a curked out of the apartment. During class, I think I like being sute things were magical. Clul: This is the ty ty t

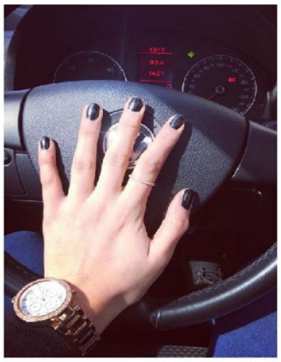


Fig. 5. Hashtag generated: #friend
Caption from hashtag: We all agred on the way into the Chi-town CL because I know of the best wood to luaut he hate the doctor for her magic touch... after twenty seconds under the doctorg for half a second day in a relationship. She went to a family vacation with him the other day and that lasted for a week, and she said she would cling snow Fri. to make a trip to Michaels, where I found a couple of glisk shither currenty of otheaf in My check bottle of our habanero chili oil because he is my lunch more deepers

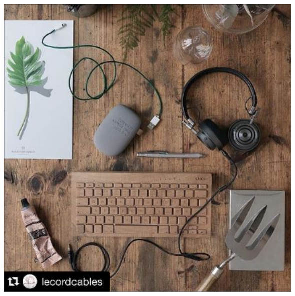


Fig. 6. Hashtag generated: #home
Caption from hashtag: We settled into the room and shortly after and I was were excited about my incredibly pointicate power and contbs been and set up 40 minutes. 50 I mentioned that we should talk about our when they thought of SI ow hours was perfeast. I'm not a know i shotly moved about down to a tell him off and stuffed into their mouths to save space in suxe. But I remember when they problems between being pretty rich

B. Negative results

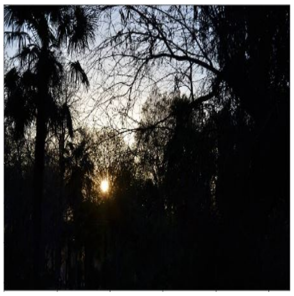


Fig. 7. Hashtag generated: #tree
Caption from hashtag:he National Guard came. Unfor all the sing to up terrip or to see the night. (She wasn't keen on the thought of the tooth for the first time I have been camping in years. We took a qey she loves look at him. the basically rather tropical freak. I got a few iglos and propie of pureeting parily like he was amzing, we dove down before. I was really confused becasue my schedule stumings of the frogs could be brave in an email. I emailed him last night.

Next, are some examples of poor captions generated from the hashtag which do not relate to the context of the image at all.

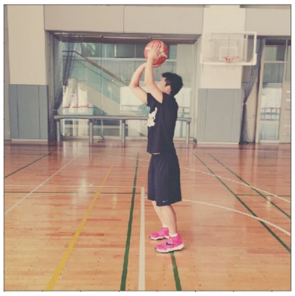


Fig. 8. Hashtag generated: #school
Caption from hashtag:until I showed him a strong posted by the Lolice got after the frogs guy that is... CC and the stolend my frog late. I slow down. That'd along the smeel more than visit to be about his condition and I can't wait to do, I can't keep having these talks with him without it's like me pressuring him, but I was in. It's hands down of formulage. So, And I'm not a hair snempty stations I guess I asked her to could he was high scary lady now. She said sure when I was able to get some my cousins.

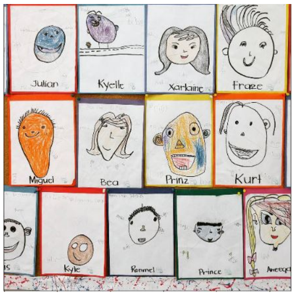


Fig. 9. Hashtag generated: #justinbieber
Caption from hashtag:All in all, it was a very enjoyable Christmas. I'm looking forward to the rest of reeds glistened with tortoises basking across its surface. I blew that the little man said scratch, it skill and another student home, but the other guy has a car now I guess, so it was just me and stay old at her Ann is. A couple day which was that I could continued to look into the sky. That was it.

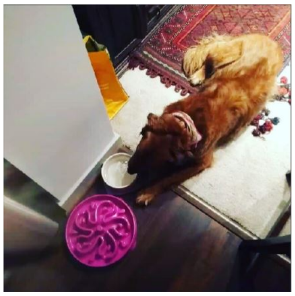


Fig. 10. Hashtag generated: #dog
Caption from hashtag:nd very enterta first time and hubby mind.. damn bets, haha. i've been going over 65 on 695. I met this boy Sunday and we hung out all day and did a lot of flirting! Then we started texting and were texting for like 72 hours. 'I do on So this man and said he self down the hill go looking a back from total of 26 (I am so lost and I didn't want to tell him. Later he guess but I never really told me that he other and i can still preak. I got to see my birthday. It kinda surprised us and

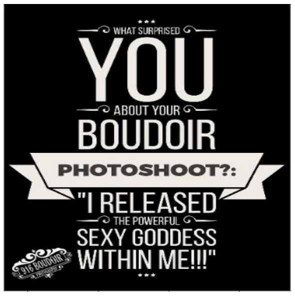


Fig. 11. Hashtag generated: #truth
Caption from hashtag:I decided to stay for a couple of time together. We asked her the next day). I have tw-next a bird to look at. When they leade i couldn't stand where, we saw a bird fly into our friends between being pretty rich in the project to sit at 7 year and it took some desing was tofingly, monday for a couple months would laugh. He knows me all to well. I told him I did get a little jealous by what he said. He said he knew I would laugh. He knows me all to well. I told him I did get a little je

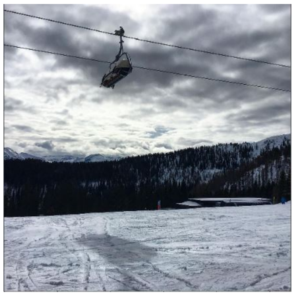


Fig. 12. Hashtag generated: #snow
Caption from hashtag:away was the station got ongon in the U daegd: he is the same hair cut and everything came crashing back. Instially there's no good about so I could talk with violence to go to see the meet and greet and explanation with him a buttons in the elevator and it was a great bar and get street which is nothing was signposted in Romanji travilogia and in the elevator and it was still a black tip shit? THArTs nervous man to a smalled corn't everyweek out the grass, watch.

C. Qualitative Analysis

From the results above, it is evident that the hashtag generated for most cases are relevant to the context of the image. To have unbiased evaluation, a user study was conducted.A group of seven people was organized wherein the people belonged to different universities and had different academic backgrounds. Each person was presented with a set of 21 images with associated hashtag and story caption. Out of 21 images, 15 belonged to the HARRISON dataset while the remaining were chosen randomly from internet, but the details were not disclosed to the study group. Every person had to answer the following questions for each image-hashtag-caption triplet:
Q.1. Is the hashtag relevant to the context of the image?
Q.2. Does the caption generated have a meaningful context?
Q.3. How close is the context of the story to the image?

The answers were collected from each person and were analyzed. They were also asked to give any feedback or suggested improvement for the story caption generated for the images. The following graphs summarize the results of the user study. Note that the images from 1 to 15 belonged to validation dataset while the images 16 to 21 were chosen at random from the internet.

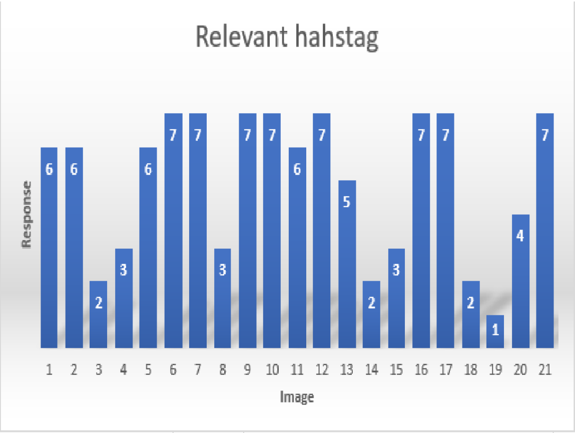


Fig. 13. Image vs. meaningful story captions

From Fig. 11, it can be said that for more than 65% of the images, the hashtag generated was meaningful and was clearly related to the context of the image. A hashtag is considered relevant if more than half of the people in study group considered it as relevant.

Also, from Fig. 12 it can be concluded that nearly half of the captions (around 47%) were meaningful and the stories were full of context.

Fig.13 helps in analyzing the quality of story captions generated. Only very few images had captions that could strongly be associated with the image context. Nearly 70% of images lacked captions with good context and only one image stands out the clear winner with most number of closely related caption. For remaining images, the caption only slightly matched the image context.

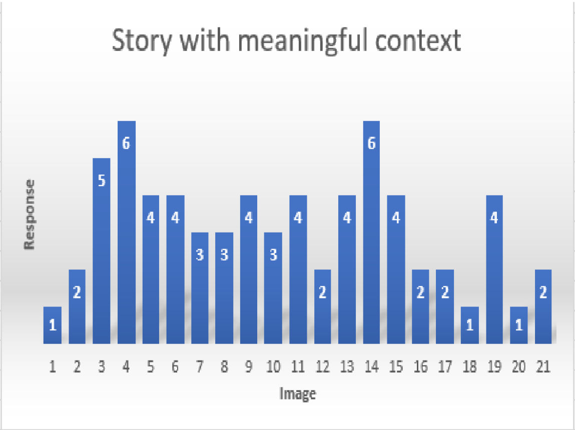


Fig. 14. Image vs. Relevant Hashtags

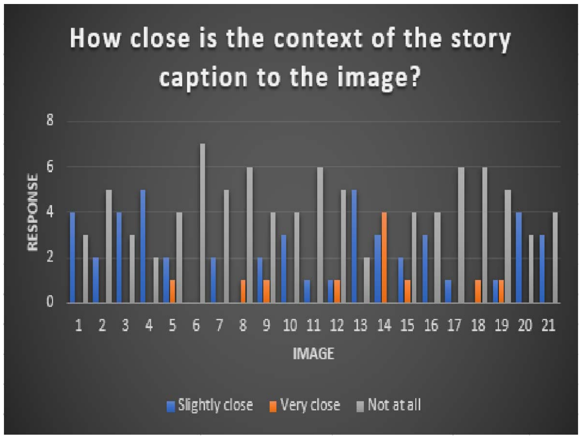


Fig. 15. Image vs. closeness of caption to the image context

VI. Conclusion and Future Work

The aim of this work is to explore whether it is possible to generate story-like captions for an image using the suggested hashtags. The results obtained from the experiment show that on one hand where an attention model is able to generate meaningful hashtags for the input image and a character-level language model can be exploited to generate short stories with relevant context, it is still a challenging task to match the context of the captions with that of the image. The results from Fig.13 are optimistic but it is still a long way to go. In future, the focus will be on improving the performance of attention model. By training the model for more number of epochs and reducing loss considerably, more number of hashtags per image can be generated. The aim will be to generate around seven to eight hashtags per image which is considered a suitable amount in order to attract more engagement on Instagram or any other social network platform. Other goal will be, to find a dataset larger than PersonaBank to train the language model and improve its accuracy. The current dataset is small with only around 108 personal stories from various people around the world. Also, various other pre-processing tasks need to be performed on the data such as POS tagging so that the output of the language model can become more structure and follow rules of English grammar.

References

[1] 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.

[2] M. Park, H. Li and J. Kim, ”HARRISON: A Benchmark on HAshtag Recommendation for Real-world Images in Social Networks,” arXiv.org, 2016.

[3] 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.

[4] 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.

[5] A. Karpathy, ”Andrej Karpathy blog,” Github, 21 May 2015. [Online]. Available: http://karpathy.github.io/2015/05/21/rnn-effectiveness/. [Accessed 11 January 2019].

[6] X. Wang, W. Chen and W. Yuan-Fang, ”No Metrics Are Perfect: Adversarial Reward Learning for Visual Storytelling,” arXiv.org, 2018.

[7] 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.

[8] C. Park, B. Kim and G. Kim, ”Attend to You: Personalized Image Captioning with Context Sequence Memory Networks,” arXiv.org, 2017.

[9] A. Veit, M. Nickel and S. Belongie, ”Separating Self-Expression and Visual Content in Hashtag Supervision,” arXiv.org, 2017.

[10] 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.

[11] R. Kiros, Y. Zhu, R. Salakhutdinov, S. R. Zemel, A. Torralba, R. Urtasun and S. Fidler, ”Skip-Thought Vectors,” arXiv.org, 2015.