AN ATTENTION HIERARCHICAL MODEL FOR SOCIAL CONTEXTUAL IMAGE RECOMMENDATION IN CLOUD

K. Venkateswarlu

Assistant Professor, Dept. of Master of Computer Applications, Narayana Engineering College, Gudur, Andrapradesh

Ch. Gurunadh

PG Scholar, Dept. of Master of Computer Applications, Narayana Engineering College, Gudur. Andrapradesh

Abstract—Image based social networks are among the most popular social networking services in recent years. With tremendous images uploaded every day, understanding users' preferences on user-generated images and making recommendations have become an urgent need. In fact, many hybrid models have been proposed to fuse various kinds of side information (e.g., image visual representation, social network) and user-item historical behaviour for enhancing recommendation performance. However, due to the unique characteristics of the user generated images in social image platforms, the previous studies failed to capture the complex aspects that influence users' preferences in a unified framework. Moreover, most of these hybrid models relied on predefined weights in combining different kinds of information, which usually resulted in sub-optimal recommendation performance.We develop a hierarchical attention model for social contextual image recommendation.

Keyword: Hierarchical Attention, Social Contextual, Image Recommendation,Upload History, Social Influence.

I. Introduction

There is an word saying "a picture is worth a thousand words". When it comes to social media, it turns out that visual images are growing much more popularity to attract users. Especially with the increasing adoption of smartphones, users could easily take qualified images and upload them to various social image platforms to share these visually appealing pictures with others. Many image-based social sharing services have emerged, such as Instagram1, Pinterest2, and Flickr3. With hundreds of millions of images uploaded every day, image recommendation has become an urgent need to deal with the image overload problem. By providing personalized image suggestions to each active user in image recommender system, users gain more satisfaction for platform prosperity. E.g., as reported by Pinterest, image recommendation powers over 40% of user engagement of this social platform [30]. Naturally, the standard recommendation algorithm provides a direct solution for the image recommendation task.

In this paper, we study the problem of understanding users' preferences for images and recommending images in social image-based platforms. Fig. 1 shows an example of a typical social image application. Each image is associated with visual information. Besides showing likeness to images, users are also creators of these images with the upload behaviour. In addition, users connect with others to form a social network to share their image preferences.The rich heterogeneous contextual data provides valuable clues to infer users' preferences to images. Given rich heterogeneous contextual data, the problem of how to summarize the heterogeneous social contextual aspects that influence users' preferences to these highly subjective contents is still unclear. What's more, in the preference decision process, different users care about different social contextual aspects for their personalized image preference. E.g. Lily likes images that are similar to her uploaded images, while Bob is easily swayed by social neighbours to present similar preference as her social friends. In other words, the unique user preference for balancing these complex social contextual aspect makes the recommendation problem more challenging.

To address the challenges mentioned above, in this paper, we design a hierarchical attention model for social image recommendation. The proposed model is built on the popular latent factor-based models, which assumes users and items could be projected in a low latent space. In our proposed model, for each user, in addition to basic latent user interest vector, we identify three key aspects (i.e., upload history, social influence and owner admiration) that affect each user's preference, where each aspect summarizes a contextual factor from the complex relationships between users and images. Specifically, the upload history aspect summarizes each user's uploaded images to characterize her interest. The social influence aspect characterizes the influence from the social network structure, and the owner admiration aspect depicts the influence from the uploader of the recommended image. The three key aspects are combined to form the auxiliary user latent embedding. Furthermore, since not all aspects are equally important for personalized image recommendation, we design a hierarchical attention structure that attentively weight different aspects for each user's auxiliary embedding.

II. Related Work

General Recommendation. Recommender systems could be classified into three categories: content-based methods, Collaborative Filtering (CF) and the hybrid models. Among all models for building recommender systems, latent factor-based models from the CF category are among the most popular techniques due to their relatively high performance in practice. These latent factor-based models decomposed both users and items in a low latent space, and the preference of a user to an item could be approximated as the inner product between the corresponding user and item latent vectors.

We summarize the related work in the following four categories.

General Recommendation: Recommender systems could be classified into three categories: content-based methods, Collaborative Filtering (CF) and the hybrid models.Specifically, BPR optimized a pairwise based ranking loss, such that the observed implicit feedbacks are preferred to rank higher than that of the unobserved ones. As users may simultaneously express their opinions with several kinds of feedbacks (e.g., click behaviour, consumption behaviour).SVD++ is proposed to incorporate users' different feedbacks by extending the classical latent factor based models, assuming each user's latent factor is composed of a base latent factor, and an auxiliary latent factor that can be derived from other kinds of feedbacks. Due to the performance improvement and extensibility of SVD++, it is widely studied to incorporate different kinds of information, e.g., item text, multi-class preference of users.

Image Recommendation: In many images based social networks, images are associated with rich context information, e.g., the text in the image, the hashtags. Researchers proposed to apply factorization machines for image recommendation by considering the rich context information. Recently, deep Convolutional Neural Networks (CNNs) have been successfully applied to analysing visual imagery by automatic image representation in the modelling process.

Social Contextual Recommendation: Social scientists have long converged that a user's preference is similar to or influenced by her social connections, with the social theories of homophily and social influence. With the prevalence of social networks, a popular research direction is to leverage the social data to improve recommendation performance. E.g., Ma et al. proposed a latent factor-based model with social regularization terms for recommendation. Since most of these social recommendation tasks are formulated as non-convex optimizing problems, researchers have designed an unsupervised deep learning model to initialize model parameters for better performance. Besides, Context MF is proposed to fuse the individual preference and interpersonal influence with auxiliary text content information from social networks.

Many network embedding models have been proposed. The network embedding could be used for the attention networks. We distinguish from these works as the focus of this paper is not to advance the sophisticated network embedding models. We put emphasis on how to enhance recommendation performance by leveraging various data embeddings.

Attention Mechanism: Neural science studies have shown that people focus on specific parts of the input rather than using all available information.Recently, the attention mechanism is also widely used for recommender systems. Given the classical collaborative filtering scenario with user-item interaction behaviour, NAIS extended the classical item-based recommendation models by distinguishing the importance of different historical items in a user profile. With users' temporal behaviour, the attention networks were proposed to learn which historical behaviour is more important for the user's current temporal decision. A lot of attention-based recommendation models have been developed to better exploit the auxiliary information to improve recommendation performance. E.g., ANSR is proposed with a social attention module to learn adaptive social influence strength for social recommendation. Given the review or the text of an item, attention networks were developed to learn informative sentences or words for recommendation.

The work that is most similar to ours is the Attentive Collaborative Filtering (ACF) for image and video recommendation. By assuming there exists item level and component level implicitness that underlines a user's preference, an attention-based recommendation model is proposed with the component level attention and the item level attention. Our work borrows the idea of applying attention mechanism for recommendation, and it differs from ACF and previous works from both the research perspective and the application point. From the technical perspective, we model the complex social contextual aspects of users' interests from heterogeneous data sources in a unified recommendation model. In contrast, ACF only leverages the image (video) content information. From the application view, our proposed model could benefit researchers and engineers in related areas when heterogeneous data are available.

III. Proposed Work

The prevalence of social networks, a popular research direction is to leverage the social data to improve recommendation performance.

- E.g., Ma et al. proposed a latent factor-based model with social regularization terms for recommendation.
- Since most of these social recommendation tasks are formulated as non-convex optimizing problems, researchers have designed an unsupervised deep learning model to initialize model parameters for better performance.

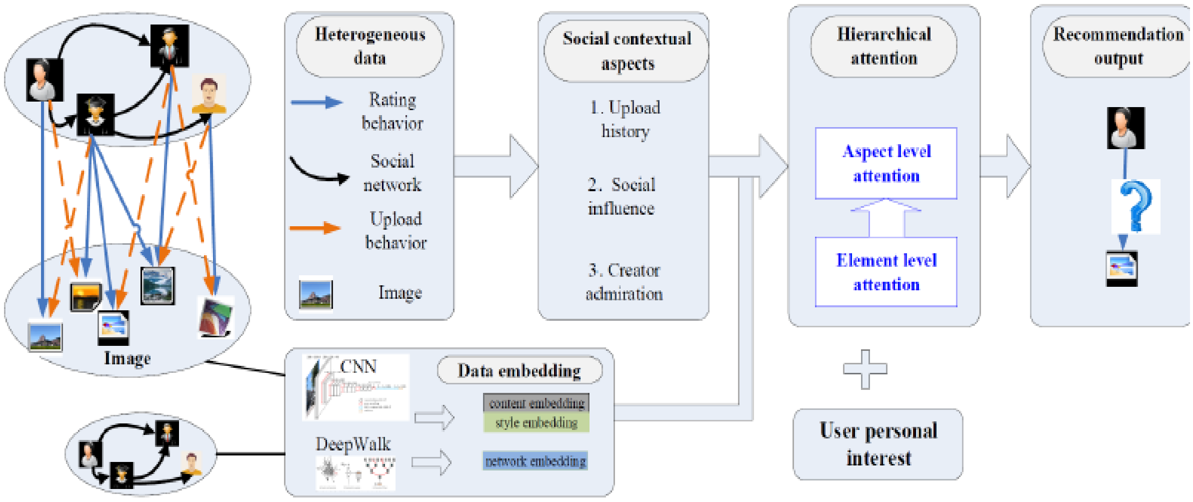

Figure.1: System overview

An overall framework of social contextual image recommendation, where the left part shows the data characteristics of the platform, and the right part shows our proposed model.

The system has following implementation modules:

Admin: In this module, the Admin has to login by using valid user name and password. After login successful he can perform some operations such as View All Users and Authorize, Add Category, Add Images, View All Heterogeneous Images with Rate, View Social Influence Attention Image, View Recommendations By Category, View All Reviewed Behavior Images, View All Searched History, View All Friend Req and Res,View Results.

Friend Request & Response: In this module, the admin can view all the friend requests and responses. Here all the requests and responses will be displayed with their tags such as Id, requested user photo, requested user name, user name request to, status and time & date. If the user accepts the request then the status will be changed to accept or else the status will remains as waiting.

Social Network Friends:In this module, the admin can see all the friends who are all belongs to the same site. The details such as, Request From, Requested user's site, Request to Name, Request To user's site.

All Recommended Images:In this module, the admin can see all the images which are shared among the friends in same and other network sites. The details such as post image, title, description, recommend by name and recommend to name.

Adding Images*:*In this module, the admin adds images details such as title, description and the image of the post. The post details such as title and description will be encrypted and stores into the database.

User*:*In this module, there are n numbers of users are present. User should register before performing any operations. Once user registers, their details will be stored to the database. After registration successful, he has to login by using authorized user name and password. Once Login is successful user can perform some operations like Search Friends, View Friend Requests, each Images and Recommend, View My Search History, View All Recommended Images, View Other User Recommended Images, View Top K Recommendation.

Searching Users*:*In this module, the user searches for users in Same Site and in Different Sites and sends friend requests to them. The user can search for users in other sites to make friends only if they have permission.

This proposed work follows the following algorithm

In this subsection, we would follow the bottom-up step to model the hierarchical attention networks in detail. Specifically, we would first introduce the two bottom layered attention networks: the upload history attention network and the social influence attention network, followed by the top layered aspect importance attention network that is based on the bottom layered attention networks.

Upload History Attention

The goal of the upload history attention is to select the images from each user a's upload history that are representative to a's preferences, and then aggregate this upload history contextual information to characterize each user. Given each image j that is uploaded by a, we model the upload history attentive score αaj as a three-layered attention neural network:

$$\alpha_{aj} = \mathbf{w}^1 \times \sigma(\mathbf{W}^1[\mathbf{p}_a, \mathbf{q}_a, \mathbf{x}_j, \mathbf{w}_j, \mathbf{e}_a, \\ \mathbf{W}^c\mathbf{f}_j^c, \mathbf{W}^s\mathbf{f}_j^s, \mathbf{W}^c\mathbf{f}_a^c, \mathbf{W}^s\mathbf{f}_a^s])$$

Social Influence Attention

The social influence attention module tries to select the influential social neighbours from each user a's social connections, and then summarizes these social neighbours' influences into a social contextual vector. If user a follows b, we use βab to denote the social influence strength of b to a. Then, the social attentive score βab could be calculated as:

$$\beta_{ab} = \mathbf{w}^2\sigma(\mathbf{W}^2[\mathbf{p}_a, \mathbf{p}_b, \mathbf{q}_a, \mathbf{q}_b, \mathbf{e}_a, \mathbf{e}_b, \mathbf{f}_a^c, \mathbf{f}_a^s]),$$

Aspect Importance Attention Network

The aspect importance attention network takes the contextual representation of each aspect from the bottom layered attention networks as input, and models the importance of each aspect in the user's decision process. Specifically, for each pair of user a and image i, we have two contextual representations from the bottom layer of HASC as: upload history contextual representation e xa, the social influence contextual representation e.q a, and the owner appreciation contextual representation qCi. Then, the aspect importance score γal (l=1, 2, 3) is modelled with an aspect importance attention network as:

$$\gamma_{al} = \mathbf{w}^3 \sigma(\mathbf{W}^3 \mathbf{a}_l),$$

IV. Experiment Results

In this section, we show the effectiveness of our proposed HASC model. Specifically, we would answer the following questions:

Q1: How does our proposed model perform compared to the baselines?

Q2: How does the model perform under different sparsity?

Q3: How does the proposed social contextual aspects and the hierarchical attention perform?

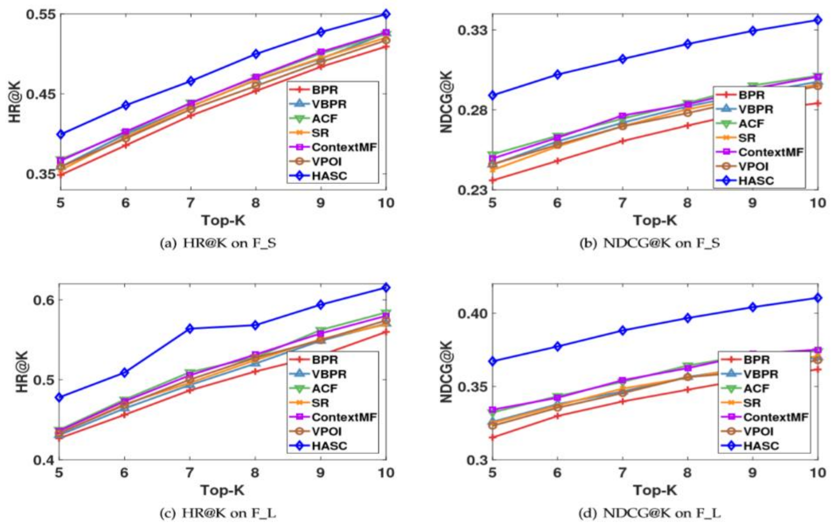

Fig. 4. Overall performance of different models on the two datasets

Overall Performance

Fig.4showstheoverallperformanceofallmodelsonHR@K and NDCG@K on the two datasets with varying sizes of K, where the top two subfigures depict the results on F S dataset and the bottom two subfigures depict the results on F L dataset. As shown in this figure, our proposed HASC model always performs the best. With the increase of the top-K list size, the performance of all models increases. The performance trend is consistent over different top-K values and different metrics. We find that considering either the social network or the visual image information could all deviate the data sparsity problem and improve recommendation performance.

Performance under Different Data Sparsity

A key characteristic of our proposed model is that it alleviates the data sparsity issue with various social contextual aspects modelling. In this subsection, we investigate the performance of various models under different data sparsity. We mainly focus on the F L dataset as it is more challenging with sparser user rating records compared to the denser F S dataset. Specifically, we bin users into different groups based on the number of the observed feedbacks in the training data, and then show the performance under different groups.

Fig. 5 shows the results, where the left part summarizes the user group distribution of the training data and the right part depicts the performance with different data sparsity.

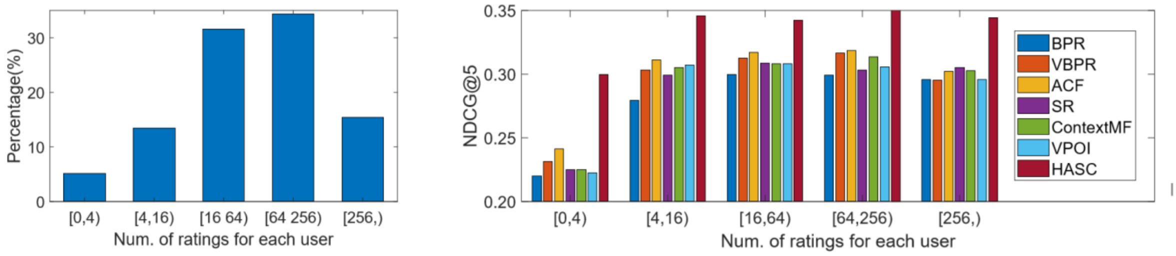

Fig. 5. Performance under different sparsity

As shown in the left part, more than 5% users have less than 4 ratings, and 20% users have less than 16 ratings with more than 100 thousand images on the F L dataset. When the rating scale is very sparse, theBPR baseline cannot work well under this situation as it only modelled the sparse user-image implicit feedbacks. Under this situation, the improvement is significant for all modelsoverBPRasthesemodelsutilizeddifferentauxiliary data for recommendation. E.g., when users have less than 4 ratings, our proposed HASC model improves over BPR by more than 35%. As user rating scale increases, the performance of all models increase quickly with more training rating records, and HASC still consistently outperforms the baselines.

Attention Analysis

In this part, we conduct experiments to give more detailed analysis of the proposed attention network. We would evaluate the soundness of the designed attention structure and the superiority of combining the various data embeddings for attention modelling.

In the experiments, we use the LeaklyReLU as the activation function σ(x) for attention modelling, and then attentively combine the elements of each set with a soft attention. Alternately, instead of attentively combining all the elements, a direct solution is to use the hard attention with MAX operation that selects the element with the largest attentive score at each layer of the hierarchical attention network.

V. Case Study

In order to better understand the proposed model, we visualize several typical users and the experimental results of different recommendation models in Fig.7. In this figure, each row represents a user. The first column shows the images liked by the user in the training data, and the second column shows the test image of each user in the test data. Please note that, due to page limit, we only show six typical training images of each user if she has rated more than 6 images in the training data. The third column shows the NDCG@5 results of different models. Specifically, to validate the effectiveness of different aspects in the modelling process, we use U, S, and C to denote the three simplified versions of our proposed HASC model that only consider the upload history aspect (i.e., $\gamma a2 = \gamma a3 = 0$), the social influence aspect(i.e., $\gamma a1 = \gamma a3 = 0$), and the owner admiration aspect(i.e., $\gamma a1 = \gamma a2 = 0$). We present the learned attention weights of different aspects of our proposed HASC model in the fourth column. The last column gives some intuitive explanations of the experimental results. As shown in this figure, by learning he importance of different aspects with attentive modelling, HASC could better learn each user’s preference from various social contextual aspects. Thus, it shows the best performance for the users in the first three rows. In the fourth row, we present a case that all the models do not perform well expect than the simplified C model from HASC that leverages the single creator admiration aspect into consideration. We carefully

ISSN: 2278-4632
Vol-10 Issue-5 No. 1 May 2020

analyse this user's records and guess a possible reason is that: the style and the content of the test image has rarely appeared in the user's training data. As this test image differs from the distribution of the training images of this user, most models could not perform well. However, the C model that leverages the owner admiration shows better results than the remaining models, as this user has liked several images uploaded by the owner. This example gives us an intuitive explanation that shows when our proposed model may not perform very well. Nevertheless, we must notice that this case is caused by the situation that the test pattern is not consistent with the patterns in the training data, which is uncommon. Therefore, we could empirically conclude that our proposed model shows the best results for most cases.

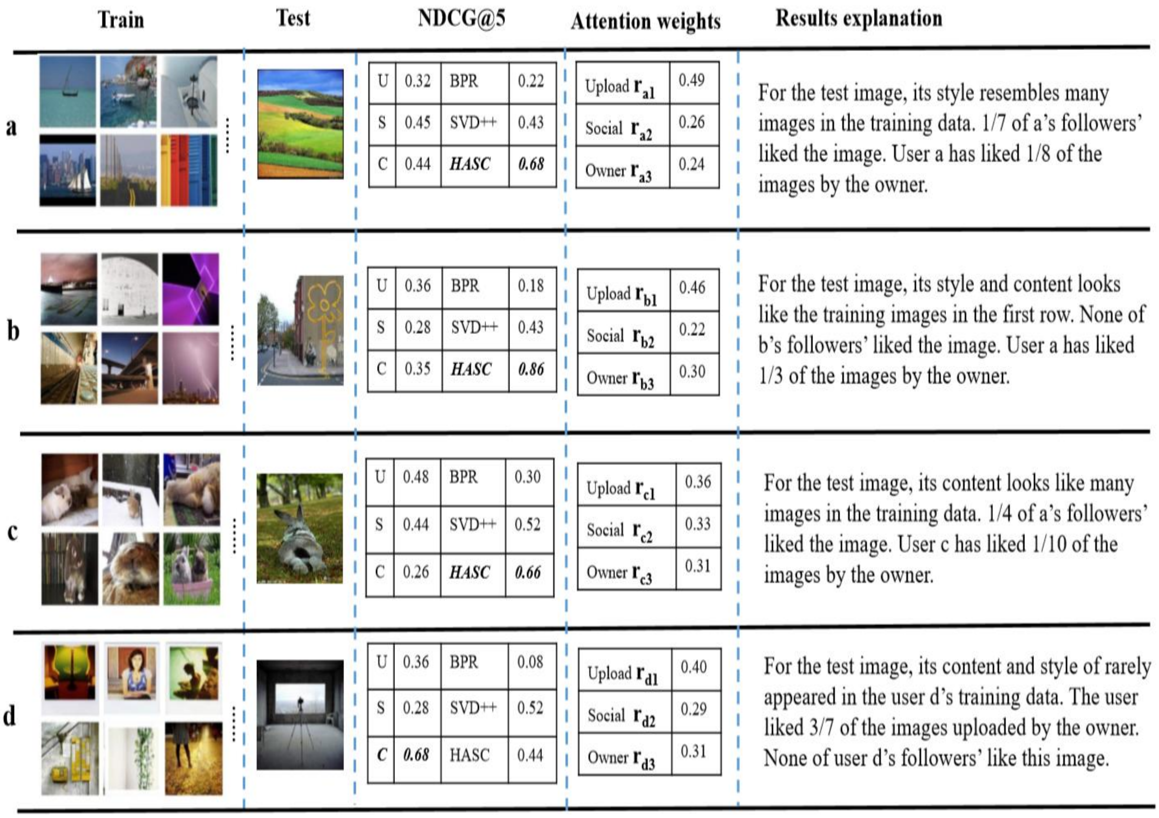

Fig. 7. The case study of several typical users.

In this figure, each row represents a user. The first and the second column are the training and test images of the user. The Top-5 recommendation results of NDCG@5 are shown in the third column. In the third column, the left three models are simplified versions of our proposed HASC model that only leverage one aspect, and the model with best performance is shown with bold italic letters.

VI.CONCLUSION

In this paper, we have proposed a hierarchical attentive social contextual model of HSC for social contextual image recommendation. Specifically, in addition to user interest modelling, we have identified three social contextual aspects that influence a user's preference to an image from heterogeneous data: the upload history aspect, the social influence aspect, and the owner admiration aspect. We designed a hierarchical attention network that naturally mirrored the hierarchical relationship of users' interest given the three identified aspects. In the meantime, by feeding the data embedding from rich heterogeneous data sources, the hierarchical attention networks could learn to attend differently to more or less important content. Extensive experiments on real-world datasets clearly demonstrated that our proposed HASC model consistently outperforms various state-of-the art baselines for image recommendation.

REFERENCES

1. Flickr Statistics. https://expandedramblings.com/index.php/ flickr-stats/, 2017. [Online; accessed 20-Jan-2018].
2. G. Adomavicius and A. Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. TKDE, 17(6):734–749, 2005.
3. A. Anagnostopoulos, R. Kumar, and M. Mahdian. Influence and correlation in social networks. In KDD, pages 7–15. ACM, 2008.
4. D. Bahdanau, K. Cho, and Y. Bengio. Neural machine translation by jointly learning to align and translate. In ICLR, 2015.
5. J. Chen, H. Zhang, X. He, L. Nie, W. Liu, and T.-S. Chua. Attentive collaborative filtering: Multimedia recommendation with itemand component-level attention. In SIGIR, pages 335–344. ACM, 2017.
6. T. Chen, X. He, and M.-Y. Kan. Context-aware image tweet modelling and recommendation. In MM, pages 1018–1027. ACM, 2016.
7. T.-S. Chua, J. Tang, R. Hong, H. Li, Z. Luo, and Y. Zheng. Nuswide: a real-world web image database from national university of singapore. In MM, page 48. ACM, 2009.
8. P. Cui, X. Wang, J. Pei, and W. Zhu. A survey on network embedding. TKDE, 2018.
9. S. Deng, L. Huang, G. Xu, X. Wu, and Z. Wu. On deep learning for trust-aware recommendations in social networks. TNNLSS, 28(5):1164–1177, 2017.
10. L. Gatys, A. S. Ecker, and M. Bethge. Texture synthesis using convolutional neural networks. In NIPS, pages 262–270, 2015.
11. L. A. Gatys, A. S. Ecker, and M. Bethge. Image style transfer using convolutional neural networks. In CVPR, pages 2414–2423, 2016.
12. L. A. Gatys, A. S. Ecker, M. Bethge, A. Hertzmann, and E. Shechtman. Controlling perceptual factors in neural style transfer. In CVPR, pages 3985–3993, 2017.
13. F. Gelli, X. He, T. Chen, and T.-S. Chua. How personality affects our likes: Towards a better understanding of actionable images. In MM, pages 1828–1837. ACM, 2017.
14. F. Gelli, T. Uricchio, X. He, A. Del Bimbo, and T.-S. Chua. Beyond the product: Discovering image posts for brands in social media. In MM. ACM, 2018.
15. Y. Gong and Q. Zhang. Hashtag recommendation using attentionbased convolutional neural network. In IJCAI, pages 2782–2788, 2016..

Author's Profile:

K. Venkateswarlu has received his MCA degree from Saraswathi Valu College of Engineering, Vellore affiliated to *Anna University in 2010* and M.Tech degree in *Computer Science* from PBR Vits, Kavali affiliated to *JNTU, Ananthapur* in 2014 respectively. He is dedicated to teaching field from the last 6years. He has guided 25 P.G students. At present he is working as an *AssistantProfessor* in Narayana Engineering College, *Gudur,* AndhraPradesh, India.

Ch. Gurunadh has Received hisB.Sc Degree in*Computer Science* from ESS Degree College,Venkatagiri affiliated to *VikramaSimhapuri University,Nellore* in 2017 and pursuing PG Degree in *Master of Computer Applications*(M.C.A)from NarayanaEngineering College,Gudur affiliated to JNTU Anantapur, Andhra Pradesh, India.

