



INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

VISUAL CONTENT LIKE IMAGE SENTIMENT ANALYSIS IN SOCIAL MEDIA: REVIEW

Shaikh Shafi Shadulla *, Tayade Pooja Mahendra

* M.Tech, Department of Computer Science and Information Technology, Dr. BAMU, Aurangabad

M.Tech, Department of Computer Science and Information Technology, Dr. BAMU, Aurangabad

DOI: 10.5281/zenodo.573539

ABSTRACT

In today's social media everything is online. People express themselves very openly and convey their message, opinion, emotion, sentiments, and attitudes towards entities such as products, services, organizations, individuals, issues, events, topics, and their attributes. To express on social media there is not only text message available but also you now can use images, emoticons, videos, likes and dislikes, graphics, stickers etc. As it becoming easier to click or capture and handle photo, image with handheld devices like camera, smartphones, phablets, tablets etc., Social media users are intending to say through visual and textual content rather than only text. Use of images, videos, with combination of text becoming so popular. Visual content speaks on your behalf very strongly. A single image replaces thousands of words. Visual content is combination on images videos, color, tone, texture, lines, shapes and many more. Human brain can sense the sentiment as soon as he visuals content. So the question arises here is that, can a machine do that? Is there any way to say what emotion, sentiment feeling user trying to say through images. AS machine learning, Artificial Intelligence is coming to play its part. Many researchers have tried and many are trying to explore visual content sentiment analysis. It's very challenging task ahead of researchers. Significant work has carried out in this area although it is in its basic form. So in this paper we are trying to explore different researchers work.

KEYWORDS: Sentiment, Images, Visual Content.

INTRODUCTION

Sentiment analysis is the field of study analyzing people's opinion, attitudes, emotions, regarding personal or individual, community, or group of people, activities, product, business, survey anything that includes people's interest directly or indirectly. People are expressing themselves by communicating each other personally, writing letters etc. But ever since internet and connecting devices are introduced people get connected through social media on various platform. People in large numbers like to express them about what they think? How they think? What are their perspectives? With rapid development in social media like Facebook, Twitter, Reddit, Instagram, Sina Webo, people interesting to express opinions in these applications. However in micro blogging site at start people were using only text messages to express themselves. But nowadays users are using combination of texts and images, videos, emoticons, and different symbols to share their opinion and lives.

Let's take an example shown in below figure 1.1. As figure is showing people tend to record their every moment in camera in explore it on social media. As well people use some pictures or images to show what they want to say. Images speak or convey very affectively. As in figure 1.1 both the images are suggesting sunset is there. And by looking at captions or textual comments we come to the conclusion that sunset is happening and nothing more. Bothe sentences do not clarify about user sentiment or user status in that particular text because it does not contain any type of emotion suggesting word like joyful, terror, happy, disgust, sad etc. So this short text based sentiment analysis system does not work here and limits its performance to convey any sentiment. Now just take a look at images or pictures that are combined with these short text comments. By considering first image we can say that user is in happy emotion as soon as we see that two young energetic persons actively jump into the air with full stretches and cheer. now the sentence will make sense that this sunset is special for these two youngsters as they are very 'happy' and commenting this sunset as special. And they are 'enjoying' this sunset at its very best.



(a)

"It's Another Sunset Time...Special..."



(b)

"It's Another Sunset!!!!"

Figure 1.1 Example Shows How Images Are Conveying Sentiments

Now take look at image in figure 1.1(b), by reading only text like it's another sunset text is somewhat similar with previous image caption. So is it conveying similar message to users about emotion. Might be right as well seeing text only, but image clarifies all as image is seen to be a guy with head down and sitting on bench looking so 'sad' or 'not so happy'. It's make a sense here right. So all the comments and images with them now clearly clarifies what user wants to say.

So visual content is having their own strength to directly perceive emotions with very strong impact. This user generated visual content comes to play key role in emerging applications such as education, entertainment, social media, advertisement, media analysis, etc. Considering the limitation of text based public sentiment analysis system, a new cross media (i.e. text +image) public sentiment analysis system should effectively fuse the results of natural language sentiment analysis and visual sentiment analysis. To uncover the secret of visual sentiment analysis some researcher tried to model the human emotions and sentiment through machine learning methods.

RELATED WORK

1. Cross-Modality Sentiment Analysis for Social Multimedia by Rongrong Ji, Donglin Cao, Dazhen Lin 2015 IEEE International Conference on Multimedia Big Data. In this work they implemented two multimodal-based visual sentiment analysis models. Both model works on multimodal content that is images from correlation and Hypergraph view. Experimental dataset they built a test dataset from the micro blogging site Sina Microblog. Collected 5000 posts containing texts, image and emotion. First created Multimodal Correlation Model (MCM). Features from each from each modality is extracted as from text feature is word; from emoticon feature is symbol; from image ANP (i.e. Adjective Noun Pair) is extracted as feature. These modalities are modeled into Markov Random Field to find correlations. Then Multimodal Hypergraph Learning Model is implemented using bag of words model for each modality. Each post is labeled with either positive, neutral or negative. In result this model outperforms the text based sentiment analysis.

2. Learning to Predict Image Affects in Social Networks by Sen Wu, Zhihao Jia. In this work researcher tries to overcome the downside of image analysis that is similar images may cause totally different affects. So how to predict the affects for these seemingly ambiguous images.so to overcome this issue they introduced homophily phenomenon into predictor and refining prediction by constructing affect homophily among users in the same community. It works on principle that individuals tends to choose friends with similar friends with similar tests and performances. So social image network is constructed in Community-aware Factor Graph Model. It extracted two features from image one is color and another is social correlation containing uploading time and community ID. Dataset used from Flickr with 150,000 images from different groups and individuals. This model gives the accuracy near about 45% and precision near about 50%.

3. Large-scale Visual Sentiment Ontology and Detectors Using Adjective Noun Pairs by Damian Borth Rongrong Ji Tao Chen Thomas Breuel Shih-Fu Chang 2013 ACM. In this work researchers proposed novel

[Shaikh* *et al.*, 6(5): May, 2017]
ICTM Value: 3.00

approach based on understanding of visual concepts related to sentiment. Key points focused are: first, a method was built upon psychological theories and web mining to automatically construct a large-scale Visual Sentiment Ontology (VSO) consisting of more than 3,000 Adjective Noun Pairs (AN-P). Second, they proposed SentiBank, a novel visual concept detector library that can be used to detect the presence of 1,200 ANPs in an image.

4. Deep Multimodal Learning for Affective Analysis and Retrieval by Lei Pang, Shiai Zhu, and Chong-Wah Ngo in 2008 IEEE Transactions On Multimedia. In this work instead of directly combining different modalities works carried on highly nonlinear relationships existing among low level features across different modalities for emotion prediction. They used deep Bozamaan machiene (DBM) a joint density model over the space of multimodal inputs, including visual, auditory, and textual modalities, is developed. Dataset is constructed on E-Flicker and E-YouTube. DBM is aiming to learn features coupling emotion and semantic signals buried in multimodal signals and it is unsupervised learning in nature.

5. Cross-modality Consistent Regression for Joint Visual-Textual Sentiment Analysis of Social Multimedia by Quanzeng You and Jiebo Luo, Hailin Jin, Jianchao Yang in 2016 ACM. In this work researcher's addresses work in two major challenges firstly, proposing a novel multi-modality regression model, which can integrate different modality features and secondly, demonstrating the feasibility of using weakly labeled data and how to easily transfer models from one domain to another domain. They adopted Convolutional Neural Networks to visual sentiment analysis. After that they proposed a novel multi-modality regression model, CCR, which tries to impose consistent constraints across related but different modalities. Accuracy for Twitter dataset is near about 0.727.

6. Discovering Affective Regions In Deep Convolutional Neural Networks For Visual Sentiment Prediction by Ming Sun, Jufeng Yang, Kai Wang, Hui Shen in Multimedia and Expo (ICME), IEEE International Conference on 2016. In this work researchers are addressing the problem of automatically recognizing emotions in still images. This work is different in regarding while most of current work focuses on improving whole-image representations using CNNs, rather than that it focused that discovering affective regions and supplementing local features will boost the performance, which is inspired by the observation that both global distributions and salient objects carry massive sentiments. They proposed an algorithm to discover affective regions via deep frame-work, in which we use an off-the-shelf tool to generate N object proposals from a query image and rank these propos-ales with their abjectness scores. Then, each proposal's sentiment score is computed using a pre-trained and fine-tuned CNN model. We combine both scores and select top K regions from the N candidates. These K regions are regarded as the most affective ones of the input image.

7. Can We Understand van Gogh's Mood? Learning to Infer Affects from Images in Social Networks by Jia Jia, Sen Wu, Xiaohui Wang, Peiyun Hu#, Lianhong Cai, Jie Tang in 2012 ACM. This is somewhat basic covering paper research is carried on Van Gogh's artwork and sentimental influence is counted. This result is used find the answer of that what are the fundamental features that reflect the affects of authors in image. Experiment is carried on more than 20,000 images randomly downloaded from Flickr. It gives the precision of 49% and recall of 24%of inferring author' affects.so for that researchers used a Partially Labeled Factor Graph Model(PFG), All images uploaded by users can be formalized as variables and observation factor functions in a factor graph.

PROPOSED METHODOLOGY

So after going through all these methods for evaluating visual sentiment analysis one thing is for sure that existing research efforts on affective computing are mostly dedicated to single media, either text caption or visual content.it is easily a liner relationship among modalities. So for better and accurate prediction of visual content it is required to explore the learning of highly non-linear relationship that exist among low-level features across different modalities for emotion prediction. Another aspect is that while current work is focusing on improving whole-image representations it is noted that instead of whole image discovering affective regions and just focusing on supplementing local features will boost performance.

So the proposed methodology is tend to work on combining the different modalities into one group an focusing in object in an image rather than whole image.

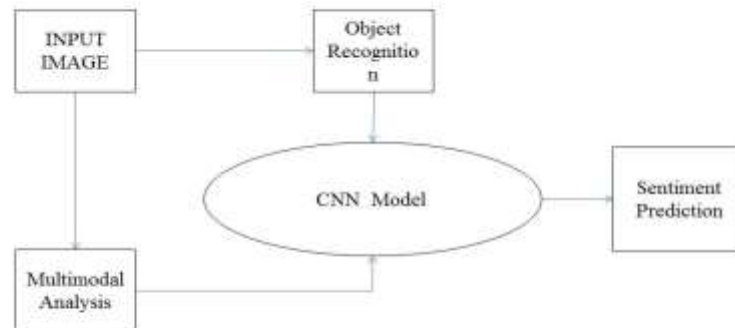


Figure 3.1 Working Block Diagram

CONCLUSION

Sentiment analysis based on visual content and textual content multimodality content is an interesting and challenging task. In this work, we leverage recently developed deep learning models to extract both textual and visual features to analyze the sentiment expressed in microblog.

The classification of emotions that are addressed by images depends on the opinion of each person individually, but we propose and investigate an idea how to compute image sentiment scores using external resources. Most methods or algorithms keep focus on either text or image. It's very daunting task to combine both modalities and draw any sentiment. For human it's very easy to see at an image any sense its color, graphics, lines, curves and lets decide what that particular image wants to say. But for computer or any computing machine it's not an easy task to force it to think or analyze like human intelligence.

We are trying to address the problem of automatically recognizing emotions in images. Considering many classification CNN works very fine with image as it tries to self-learn. Object recognition is also another area where image sentiment detection is coming to play a vital role.

ACKNOWLEDGEMENTS

We are glad to express our sentiments of gratitude to all who rendered their valuable guidance to us. We would like to express our appreciation and thankful to Prof. R. R. Deshmukh, Head of Department, Department of Computer Science and Information Technology, Dr. BAMU, Aurangabad. We are also thankful to our guide Prof. S. N. Deshmukh, Department of Computer Science and Information Technology, Dr. BAMU, Aurangabad. We also thank the anonymous reviewers for their comments.

REFERENCES

- [1] Donglin Cao, Rongrong Ji, Dazhen Lin, and Shaozi Li, "A Cross-Media Public Sentiment Analysis for Social Media", Lin 2015 IEEE International Conference on Multimedia Big Data.
- [2] Sen Wu and Zhihao Jia, "Learning to Predict Image Affects in Social Networks", In cs229.stanford.edu 2013.
- [3] Damian Borth , Rongrong Ji , Tao Chen , Thomas Breuel and Shih-Fu Chang "Large-scale Visual Sentiment Ontology and Detectors Using Adjective Noun Pairs",in MM '13 Proceedings of the 21st ACM international conference on Multimedia, October 2013.
- [4] Lei Pang, Shiai Zhu, and Chong-Wah Ngo," Deep Multimodal Learning for Affective Analysis and Retrieval", in IEEE Transactions On Multimedia, VOL. 17, NO. 11, NOVEMBER 2015.
- [5] Quanzeng You and Jiebo Luo, Hailin Jin, Jianchao Yang, "Cross-modality Consistent Regression for Joint Visual-Textual Sentiment Analysis of Social Multimedia", in 2016 ACM.



-
- [6] Ming Sun, Jufeng Yang, Kai Wang and Hui Shen, "Discovering Affective Regions In Deep Convolutional Neural Networks For Visual Sentiment Prediction" in Multimedia and Expo (ICME), IEEE International Conference on 2016.
- [7] Jia Jia, Sen Wu, Xiaohui Wang, Peiyun Hu, Lianhong Cai, and Jie Tang, "Can We Understand van Gogh's Mood? Learning to Infer Affects from Images in Social Networks", In MM '12 Proceedings of the 20th ACM international conference on Multimedia , November 2012

CITE AN ARTICLE

Shadulla, S. S., & Tayade, P. M. (2017). VISUAL CONTENT LIKE IMAGE SENTIMENT ANALYSIS IN SOCIAL MEDIA: REVIEW .INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY, 6(5), 364-368. doi:10.5281/zenodo.573539