Comparative study of Emotion recognition in social media

sagar , nishanth

King's College London, UK

sagar.joglekar, nishanth.sastry}@kcl.ac.uk

1. ABSTRACT

Sentiment analysis in Online social networks (OSNs) is a very active field of research. The prime research motive of sentiment analysis in OSNs has been analysing textual data shared over popular networks like Twitter and Facebook. But with the explosive growth of smartphone industry, inclusion of other media factors are on the rise. Mediums like videos, photos and audio convey much more information about the context of a social interaction than plain text. This has posed an interesting prospect for computer science i.e. inclusion of human affects conveyed through these mediums to allow another dimension for social interactions. These mediums however pose a higher complexity in problem space. Our paper tries to explore this space by including neural network based approaches and also shows that using heterogeneous network designs together can give us a much higher precision in understanding the sentimental context of a media. In this paper we look at commercially available techniques, and compare results with custom designed neural networks. Finally we explore possibility of combining different approaches and benchmark them against some popular datasets in the wild

2. INTRODUCTION

Online social networks (OSNs) have seen a massive surge in usage over the past decade. The surge is also going hand in hand with the explosion of smart phone industry. More and more social interactions are now driven by media content like selfies and group selfies because of the ubiquitous nature of cameras. A sharp change in cultural aspects of online social interactions are evident and have also been studied in detail in papers like [6].

3. A SOCIAL APPROACH TO SENTIMENTS:

Sentiments are fundamental part of our day to day social interactions. A face to face social interaction is generally augmented with facial expression, body language and linguistic sentiment to convey the exact meta information. These properties are very human in nature and are mimicked in the social networks as well. Studies like [3] have explored the world of linguistic sentiment in social networks, by comparing several popular textual sentiment analysis

methods used for analysing tweets. Our paper tries to explore a similar exercise for perceptual sentiment in social media.

When it comes to perceptual sentiments, there are two broad categories that could be explored. The first category looks at the perceptual sentiment evoked by a social media content. The second category talks about the actual latent perceptual sentiment that comes with the context of the content itself. We will discuss about the research problems about both these categories.

3.1 Evoked perceptual sentiment

Several works have done in depth studies using methods like crowdsourcing to understand the different shades of a particular evoked emotion. Works like UrbanGems [1] and StreetScore [5] use crowdsourcing methods to understand degrees of human sentiment evoked because of pictures of real urban neighbourhoods. Sentiments like the feeling of safety and aesthetics are especially hard to quantify and crowdsourcing helps the authors to do some interesting modelling. On the other hand there are papers like [2] by L. Jeni et.al. describe utility of actual facial expression detection for understanding content consumer reaction. Such approaches help us understand the very effect of a particular content on the consumer.

3.2 Latent perceptual sentiment

This approach is what this paper stresses on. By latent perception, we mean the hidden parameters, which are part of the very content. Social networks like reddit have specific sub-reddits that work on appealing to these types media sentiments that evoke emotions like empathy, disgust, contempt and love . One such popular sub-reddit is labelled R/aww which contains images and GIFs that showcase cute animals and animal behaviours. Another one called R/cringe appeals to the sentiments of awkwardness and discomfort by exhibiting videos and Gifs about people in awkward situations. These specific social channels are popular because the content shared over these channels have a certain type of latent sentimental response, which the consumers of these channels resonate with.

Our paper focuses on this part of the story, and tries to survey and benchmark certain state of the art methodologies out there. We also propose certain hybrid approaches, which show that we can attain much better performance if a heuristic approach to combine certain methods is taken.

4. SENTIMENT ANALYSIS METHODS

To the best of our knowledge we have evaluated certain popular approaches in solving the problem of extracting latent sentiment in a media content. The sentiment analysis methods broadly fall into two bins. One is the Content based Image retrieval (CBIR) [4] set

of approaches, which actually analyse the image structure and contents to extract features and inferences about the image. The second bin is emotional semantic image retrieval (ESIR) [7] which aim at trying to extract the semantic gist of a particular image. Human brain is great at extracting such semantics. For example it is very natural for a person to describe a particular image as "picturesque" or "scenic" or to describe someone's clothing as "tacky" , "classy" or "elegant". These semantic classes, no matter how subjective, are also sufficiently descriptive for another human being to process. Sentiments also

5. REFERENCES

[1] ADAM BARWELL, DANIELE QUERCIA, J. C. http://www.cam.ac.uk/research/news/how-to crowdsource-your-happy-space, 2012.

[2] JENI, L. A., LŐRINCZ, A., NAGY, T., PALOTAI, Z., SEBŐK, J., SZABÓ, Z., AND TAKÁCS, D. 3d shape estimation in video sequences provides high precision evaluation of facial expressions. *Image and Vision Computing 30*, 10 (2012), 785–795.

[3] JOO, J., LI, W., STEEN, F. F., AND ZHU, S. C. Visual persuasion: Inferring communicative intents of images. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2014), pp. 216–223.

[4] LIU, Y., ZHANG, D., LU, G., AND MA, W.-Y. A survey of content-based image retrieval with high-level semantics. *Pattern Recognition 40*, 1 (2007), 262 – 282.

[5] NAIK, N., PHILIPOOM, J., RASKAR, R., AND HIDALGO, C. Streetscore – predicting the perceived safety of one million streetscapes. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on* (June 2014), pp. 793–799.

[6] SOUZA, F., DE LAS CASAS, D., FLORES, V., YOUN, S., CHA, M., QUERCIA, D., AND ALMEIDA, V. Dawn of the selfie era: The whos, wheres, and hows of selfies on Instagram. In *Proceedings of the 2015 ACM on Conference on Online Social Networks - COSN '15* (2015), pp. 221–231.

[7] WANG, W., AND HE, Q. A survey on emotional semantic image retrieval. In *Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on* (Oct 2008), pp. 117–120.