Understanding popularity of vines

1, 2
King's College London, UK
blah, bleh}@kcl.ac.uk

1. ABSTRACT

In this article we examine the relevence of several aesthetic, sentimental and social features with popularity of microvideos. We look at influence of social and aesthetic features on how a vine video performs in the social world. We also look at the affective component of the videos to see if there are any peculiar sentiments related to the success of a vine video

2. INTRODUCTION

The Art of story telling could be attributed to be one of the most ancient arts. It could be in the form of neolithic paintings to the egyptian hieroglyphs or among the elaborate epics of Illiad and oddyssey to the elaborate power plays of the Game of thrones, humans have always strived to record or create elaborate plots and stories. The human need of transferring experiences to others in different forms of creative arts, has ever so created the world as interesting as we see it.

The art of story telling has spawned and transformed several industries, including the very important entertainment industry. With progress of technology, entertainment industry has gone through several rejuvination cycles. Starting with plays to television, each technological advancement has created a new form of stories to be presented. The latest of these cycles was powered by the internet with the help of services like Youtube, Netflix, Hulu and Amazon.

Over the past few years, the idea of microvideos has taken over the social world. It started with Vine, a company that was found in 2012, which allowed users to upload micro videos, not larger than 6 seconds. These videos are then consumed and shared, and are ranked based on how many times they are looped over. The videos have a spectrum of generes, starting from a quick home video about domestic cats and dogs, to elaborate short skits done by now acclaimed vine stars [1] [2]

[1] www.vanityfair.com/hollywood/2016/02/king-bach-rocketjump-youtube-vine-stars

[2] http://newmediarockstars.com/2015/04/youtubers-viners-attend-the-white-house-correspondents-dinner-gallery/

WWW 2015, May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3469-3/15/05.
http://dx.doi.org/10.1145/2736277.2741671.

3. RELATED WORK

The work done in micro video analysis has been limited. Work by Miriam et.al [**?**] try to quantify and build on the notion of creativity. Work by [**?**] use textual sentiments to bring thousands of fiction novels to sentiment space and show that most novels follow 7 salient categories of stories. A paper by Nguyen et.al [**?**] collected more than 200 thousand micro videos from vine. A work done by Fontanini et.al [**?**] explore relevence of perceptual sentiments to popularity of a video, but the work done was on youtube viral videos, which have a much richer composition and structure. The problem of understanding what makes a visual media stick, has been a difficult one to solve. There are a few approaches to understand the aesthic and memorability aspects of an image [**?**] [**?**] [**?**]

4. DATASET

We crawled vine for over a month and did a snowball sampling to collect over 12000 videos which were ranked to be popular by the vine service over 2 weeks. The ranking was solely based on the number of loops the videos have gained over the period since creation of the video. We collected an additional 70,000 videos which were at a very early stage of their lifetime. These videos were not classified to be popular becasue of their nascent nature. We tracked these videos over 4 weeks for their popularity metrics and user metrics. Out of these 70 thousand videos about 3 thousand made it to the popular category just by using the metric of vine loops. This made the dataset to be an exhaustive list of 82 thousad videos with 15 thousand hitting the popular list. We also collected the related metadata about the videos themselves, and the metadata about the user profile which posted the videos.

The popularity distribution of the whole dataset follows as expected a zipf distribution. The Fig. 1 shows the distribution of likes and repost counts of the collected videos on al og scale. Videos with 0 likes or reposts were given a marginal 1 like to avoid undefined logarithms.

Over the course of this study, we also used several third party datasets, to corroborate different properties of vine videos, with datasets previously studied, ranked and evaluated by other research groups. For comparing the aesthetic qualitiy of our videos, we used the dataset provided by Datta et.al [**?**] which provided rated images , rated by a crowdsourced activity, on the scale of 0 to 7. Each image has more than 15 ratings, and the median ratings of these images were considered for our work. For our work with frame sentiments, we use the dataset provided by the work of sentibank [**?**] [**?**] for baselining the performance of our detector for detecting frame sentiments in vine. Finally for doing the exploratory work with facial expressions, we used the dataset opened up to the public

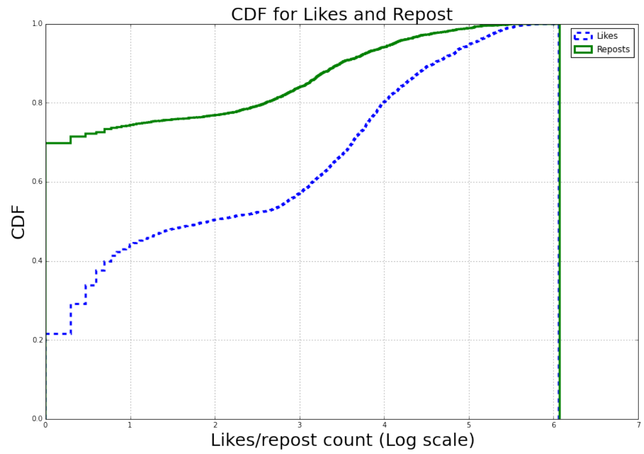

Figure 1: *CDF of Like count and Repost count.* **The values are normalized and on a Logarithmic scale. As expected from a long tail distribution of metrics like popularity, the dataset has a lot of videos with zero likes and reposts. Such videos are synthetically given one like and repost to avoid undefined values**

for the facial expression recognition competition 2013 [**?**] to train and test our convolutional network

5. METHODS

Over all the challege was to understand what makes a vine popular. And for that the there was a need to explore correlations of all the possible abstract midlevel features made available to us because of the rapid development in the fields of Deep machine learning. The main important contribution of machine learning is the ability of computationally extracting abstract higher level representations, which vaguely represent human perception.

5.1 Low level Aesthetic Features

: There are some well known computationally evaluatable aesthetic features like Rule of thirds, Sharp Pixel proportion, Contrast, Simplicity etc [**?**]. The parameters basically compute perceptual features of an image based on well know heurestic rules set by photographers. For baseline and comparison, we compare the features with images taken from the dataset from photo.net [**?**]. We only choose images with median ratings of 6 or above on aesthetic scale of 0 to 7. Because of the very nature of Vine videos, it was possible to sample 6 images , one for each second of the video, to get a good approximation of the aesthetic quality of the whole video. So in our processing pipeline, we sample one image per second from the 6 second long vine, and then take a median score of the aesthetic parameter across the sampled image. This score is assigned to the whole video to signify the value of that particular aesthetic parameter.

5.2 Presence of Faces

: One important aspect of micro videos is the presence of user as an actor in the video. When you look at viral vine videos, most videos seem to have a lead actor performing a skit. The hypothesis here was that vine has become a social media network, where actors have gained prominence and become a reason for popularity. So we did a small experiment, where we sample one image every second from all the videos collected. Then we calculate the percentage of frames across each video which contained at least one face in it.

Table 1: List of Aesthetic parameters computed for highly rated aesthetic images, Popular videos and unpopular videos. Most parameters have no bias towards either popular or unpopular videos

Parameter	Aesthetic Images		Popular Vines		Unpopular Vines	
_	Mean	Median	Mean	Median	Mean	Median
Color Contrast	51.05	30.22	29.88	16.43	20.23	8.83
Intensity Balance	0.11	0.08	0.16	0.13	0.17	0.14
Luo Simplicity	0.009	0.005	0.013	0.012	0.015	0.014
Sharp pixel proportion	0.103	0.098	0.090	0.085	0.089	0.081
Image Saturation	0.943	0.974	0.672	0.678	0.615	0.646
Avg. Brightness	0.148	0.141	0.137	0.130	0.139	0.124
Rule of Thirds	0.879	0.899	0.883	0.883	0.878	0.882
ROI Proportion	0.316	0.089	0.175	0.112	0.165	0.110

We use the well tested Viola Jones dectector for frontal and profile face detection [**?**].

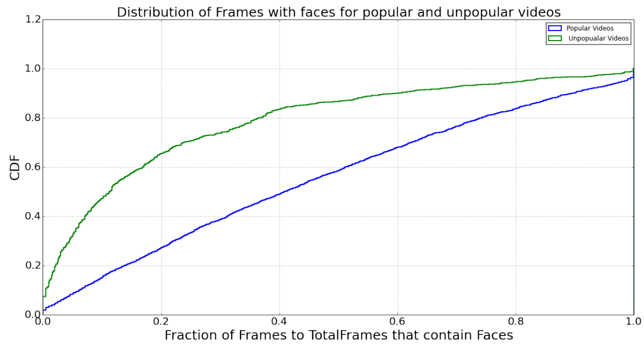

Figure 2: *CDF for popular and unpopular videos. The CDF signifies the cumulative distribution of percentages of face containing frames in a vine video. The observation here is popular videos tend to have higher percentages than unpopular videos*

6. REFERENCES

[1] Borth, D., Ji, R., Chen, T., Breuel, T., and Chang, S.-F. Large-scale visual sentiment ontology and detectors using adjective noun pairs. In *Proceedings of the 21st ACM International Conference on Multimedia* (New York, NY, USA, 2013), MM '13, ACM, pp. 223–232.

[2] Datta, R., Li, J., and Wang, J. Z. Algorithmic inferencing of aesthetics and emotion in natural images: An exposition. In *2008 15th IEEE International Conference on Image Processing* (2008), IEEE, pp. 105–108.

[3] Fontanini, G., Bertini, M., and Del Bimbo, A. Web video popularity prediction using sentiment and content visual features. In *Proceedings of the 2016 ACM on International Conference on Multimedia Retrieval* (2016), ACM, pp. 289–292.

[4] Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., Cukierski, W., Tang, Y., Thaler, D., Lee, D.-H., et al. Challenges in representation learning: A report on three machine learning contests. In *International Conference on Neural Information Processing* (2013), Springer, pp. 117–124.

[5] Isola, P., Xiao, J., Torralba, A., and Oliva, A. What makes an image memorable? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2011), pp. 145–152.

[6] JOU, B., CHEN, T., PAPPAS, N., REDI, M., TOPKARA, M., AND CHANG, S.-F. Visual affect around the world: A large-scale multilingual visual sentiment ontology. In *Proceedings of the 23rd ACM international conference on Multimedia* (2015), ACM, pp. 159–168.

[7] KALAYEH, M. M., SEIFU, M., LALANNE, W., AND SHAH, M. How to take a good selfie? In *Proceedings of the 23rd ACM International Conference on Multimedia* (New York, NY, USA, 2015), MM '15, ACM, pp. 923–926.

[8] NGUYEN, P. X., ROGEZ, G., FOWLKES, C., AND RAMAMNAN, D. The open world of micro-videos. *arXiv preprint arXiv:1603.09439* (2016).

[9] REAGAN, A. J., MITCHELL, L., KILEY, D., DANFORTH, C. M., AND DODDS, P. S. The emotional arcs of stories are dominated by six basic shapes. *arXiv preprint arXiv:1606.07772* (2016).

[10] REDI, M., O'HARE, N., SCHIFANELLA, R., TREVISIOL, M., AND JAIMES, A. 6 seconds of sound and vision: Creativity in micro-videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2014), pp. 4272–4279.

[11] YEH, C.-H., HO, Y.-C., BARSKY, B. A., AND OUHYOUNG, M. Personalized photograph ranking and selection system. In *Proceedings of the 18th ACM international conference on Multimedia* (2010), ACM, pp. 211–220.