

Study of Human affects in Vine videos

sagar , nishanth
King's College London, UK
sagar.joglekar, nishanth.sastry}@kcl.ac.uk

1. ABSTRACT

In the past few years we witnessed the rise of the *Selfie* phenomena. The world of online social networks was taken by a storm. There were several social, psychological and social computing studies trying to understand this phenomena. In late 2012, a company was found on the basis of adding a new dimension to the concept of social media called Vine. Vine was solely based on the premise of sharing a short high impact video that delivers a message. Soon the service became popular and the bigger and more popular services like Instagram and Twitter started following their footsteps. They have now enabled sharing of limited duration video clips. Our paper tries to measure this phenomenon and makes certain studies about how a particular vine-like video gains popularity. We further look at this through the lens of affective computing and machine learning, and propose a new framework to understand human affects in the budding research field of social *Artificial Intelligence*. We do all this by collecting a 1 month long dataset of all the popular videos overall and amongst popular channels.

2. INTRODUCTION

Online social networks (OSNs) have seen a massive surge in usage over the past decade. The surge is going hand in hand with the explosion of smart phone industry. More and more social interactions are now driven by media contents like selfies, group selfies and videos 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 [5]. With the rise of social media networks like Vine and Instagram, human to human non-verbal interactions have another dimension to manifest. One of the predominant modality of self expression arose from this boom in social media, and that was the Selfie. The [5] paper explores several of the properties of selfie amongst Instagram users, where they explore correlation of facial orientation, poses and smiles with parameters like country of origin of the selfie user, post frequency, likes received, number of faces in the pictures, gender and smile scores. Such studies give us interesting insights about the sharp rising OSN phenomena of selfies. The study also states that more than

50 percent of photos shared on Instagram, fall under the category of selfies.

A major change in these behaviours was seen when the social network called Vine was launched in 2012. Vine adds another dimension to the act of self expression, where the users can record a 7 second long video and post it online and get engagement from peers. This service got so popular that the larger services like Instagram and Twitter also added the feature of short videos to their services. These developments just indicate that the fact that videos add more dimensions to the act of self expressions, they are being used heavily by the millennials and the young adult generation. Hence our paper asks the following questions regarding this medium of expressions

1. Video Selfie: Is there something analogous to a still selfie in this new medium? How can we define it?
2. The Influence: Do followers react differently to video selfies compared to non-selfies?
3. Does the face says it all: Do human affects influence the popularity of a video selfie?
4. What do the vines say: What are the most common ontological descriptions that describe vines?

To our best of knowledge, this would be a first of its kind investigation into this type of medium of expression. More over our paper tries to employ techniques like adjective noun pairs [1] and convolutional neural networks [2] to understand human affects in video. We work on real data collected from a popular social media service called Vine over one month. In the following sections we will try to describe our approach

3. RELATED WORK

There have been several attempts to understand the phenomena of self expression. With the rise of selfie, the expression exposed several facets of human nature. [5] looked at the phenomenon of selfie as a whole. There the authors explain social structures, temporal dynamics, demographics and memes using Instagram datasets of selfies. There were other works in this area [3] which explore the content itself. They talk about what kind of content is posted by specific categories of users. Also they try to understand how different types of content relate to the number and types of followers an account gathers. Another interesting work [4] goes in a different direction and tries to understand whys and whats of a perfect selfie. They try to analyse adjective noun pairs using Sentibank [1] and discuss the salient characteristics of popular selfies. These works ask a bigger question in the realm of social media analysis. What aspects of media appeal to humans. What makes a media more or less consumable. We look at this problem from an angle

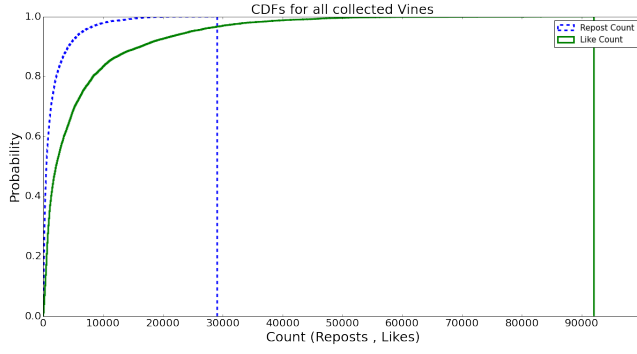


Figure 1: CDF of likes and repost count across collected Vines.

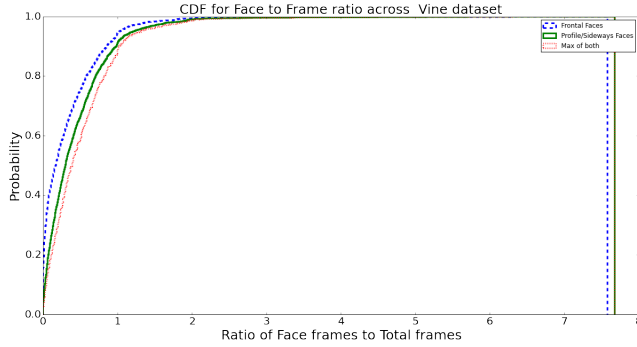


Figure 2: CDF of Face frames to total frames calculated for Frontal and Profile faces

of machine intelligence question. Can you learn the higher level representation that makes a video or an image disseminate faster in the network. To begin our work, we try analysing Vine social network and the content uploaded over it. The following section would delve deeper into our work.

4. DATASET

For our work we collected posts from Vine social network, about top 100 popular posts over all and also over specific channel categories. We collected the metadata every 6 hours to make sure there is minimal overlap in rankings and we continued this exercise for over a month. Finally we filtered all the uniques posts out and collected the actual vine mp4 files. In total we have 16504 unique vine clips collected over a month that ranked in the top 100 posts across vine. We also store the individual post metadata and the profile of the post creator. Some statistics for our dataset are as follows

Parameter	mean Value	Median Value
Reposts	1558	552
Likes	5754	2193
Loops	205504	76895

From the statistics of reactions to vines, it seems loops or the number of times a video is replayed is not necessarily a good measure to quantify popularity. However the likes and repost count could act as a good descriptor for a Vine's popularity.

Figure 4 shows that the behaviour of reposts could be used as a viable metric for user's interaction with a vine video and dissemination of the video in the network.

5. VIDEO SELFIE

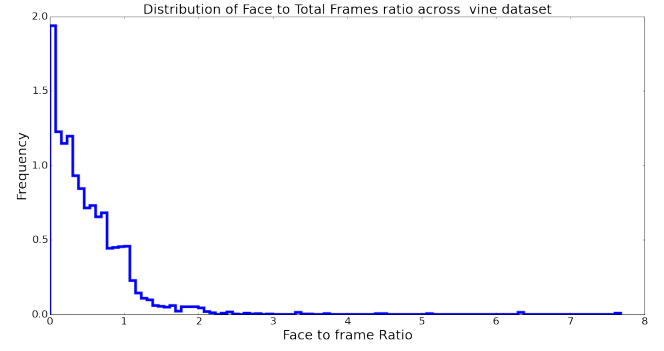


Figure 3: Distribution of face frames compared to total frames for all the Vine videos in the dataset

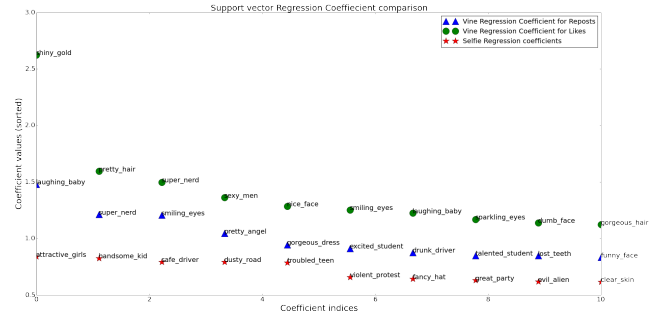


Figure 4: Top 10 Support Vector regression coefficients for Instagram selfies, Vine Likes and Vine reposts

The first question we ask about this dataset, to understand the contribution of human affects was, is there any such thing as a video selfie. To understand this better, we processed each video using the well known framework for face detection by Viola and Jones [6] to detect profile and frontal faces in a video. We collected these measurements across all the collected vine videos and found that more than 50 percent of popular vines which ranked in top 100 that day, had more than 50 percent of the face to frame ratio. Moreover orientation of faces seemed to be mixed between frontal and side ways. The mean ration of face frames to total frames was 54.74 percent, across the collected samples. This says a lot about the nature of this social media. There seems to be a preference among Vine consumers towards videos with obvious face protagonist over others.

6. PREDICTION OF LIKES AND REPOSTS IN VIDEO SELFIES

7. 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] CAMPOS, V., SALVADOR, A., GIRO-I NIETO, X., AND JOU, B. Diving deep into sentiment: Understanding fine-tuned cnns for visual sentiment prediction. In *Proceedings of the 1st International Workshop on Affect & #38; Sentiment in*

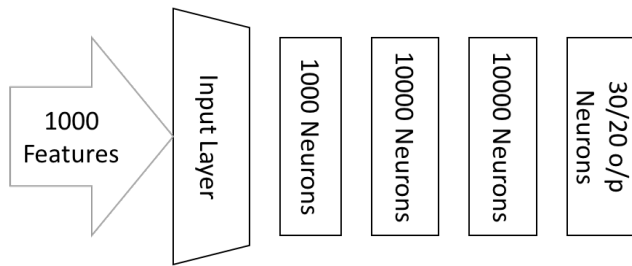


Figure 5: **Neural network architecture for vine repost and like prediction**

Multimedia (New York, NY, USA, 2015), ASM '15, ACM, pp. 57–62.

- [3] HU, Y., MANIKONDA, L., AND KAMBHAMPATI, S. What we instagram: A first analysis of instagram photo content and user types. In *ICWSM* (2014), AAAI.
- [4] 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.
- [5] 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.
- [6] VIOLA, P., AND JONES, M. J. Robust real-time face detection. *Int. J. Comput. Vision* 57, 2 (May 2004), 137–154.