universidade de aveiro

Mass Surveillance in China - a data overhead analysis

Author: Tomás Costa

Date: 06/06/2020

Index

1 Introduction

This report was inspired in the documentary [1] by VICE on serves to study the massive impact of camera surveillance in China, and it addresses three major situations:

- The social/ethical impacts on the daily lives of Chinese citizens.
- Behind the Scenes on China's person recognition system
- The complexity of data flow over a video network for real time person recognition

This situation is important to study, as this mass-surveillance state in China is affecting other countries. As I will explain below, China's digital espionage methods are means of social control, since the new "scoring system" [2], allows you to categorize citizens based on their behaviour.

Figure 1: Person Recognition in a Crowd

2 China's reality

As stated above, China's reality is of a western civilization thinks of a dystopian future. As of 2020, the social credit system in China is completely implemented. This credit system places Chinese citizens in different levels according to how many points they have, some of its lower levels, withdraw your rights of getting public transportation tickets, denial of social services, restricted access to public services and public shaming, by exposure on TV screens in public spaces.

On paper, this could be a good idea, as it would promote good behaviour and reward citizens, but, this system only allows for a complete control of Chinese government, since some of the activities that can lead you to losing points are:

- Spreading rumors on the internet
- "Insincere" apologies for crimes committed
- Posting anti-government messages on social media
- "Illegally" protesting against authorities

As we see listed above, some of these "negative actions", are basic human rights and therefore, their standpoint is highly doubtful. On the other hand, there's some benefits that originate from a high score in the system, as:

- Priority for school admissions and employment
- Cheaper public transport
- Shorter wait times in hospitals
- Tax breaks
- Fast track promotion at work

But even then, there's a possibility to overpass the system and manipulate it, which is highly probable, according to this article [3]

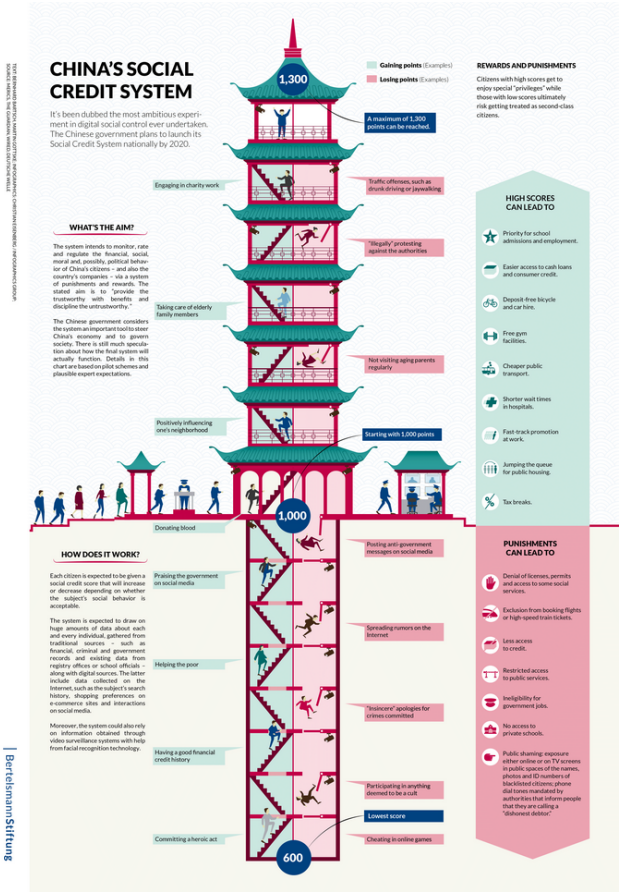

Figure 2: China's Social Credit System

And as quoted by Zhuang Daohe, a Chinese lawyer on VICE's video[1]:

> The Chinese government has very varied and complex methods of obtaining personal information, and these traps that aim to gather personal information dont have enough regulation and laws on them, so they should be used fairly. But, if there are flaws in the Chinese morality and its legal system, then the credit system will fall apart.

Below, I will detail more in depth some of the algorithms that make it work and some of its impacts, like racial profiling and exclusion.

3 Behind the Scenes on China's recognition system

There's several ways to track down citizens personal information, and all this is achieved with the camera setups that the major chinese cities are harboring. According to Comparitech's study, China takes 8 of the 10 most surveilled cities. And at the top of the list is Chongqing, with roughly 170 per 1000 citizens.

Figure 3: 10 cities with most surveillance cameras per person

There's several companies in the race for improving the quality of person and face recognition algorithms, for real time citizen tracking. According to the NYTimes [4], most of these companies are receiving big investments for their facial recognition software, which only speeds up the technology to a point of total efficient and fast algorithm. Below is some of the technology to achieve all this.

3.1 Person Recognition

As for person recognition, the most recent work from CloudWalk [5] (the company behind most of the work produced in facial and body recognition), is DenseBody[6], which we will further explain with a code example, but can be summarized as:

> Recovering 3D human body shape and pose from 2D images by directly regressing the 3D human mesh from a single color image using a Convolutional Neural Network(CNN). The proposed method achieves state-of-the-art performance on several 3D human body datasets including Human3.6M, SURREAL and UP-3D with even faster running speed

This person recognition software is present in most cameras, and it allows, for example, police units to find about any person in a little over 5 minutes as was once the case with a Chinese Student that was targeted for writing against SkyNet's [7] vigillance system.

SkyNet is China's surveillance system, with a little over 20 million cameras across China, and plans to add hundreds more, according to state media [8]. It is a bit ironic that the system is named after a fictional artificial

Figure 4: DenseBody's 3D object modeling

neural network-based conscious group mind and artificial general intelligence system that features centrally in the Terminator franchise, but in some way, it represents one of the paths China could head to, if these surveillance protocols don't start getting regulated.

3.2 Hotspot Recognition

Besides people and facial recognition, China is developing algorithms to recognize hotspots of people, as shown in the image below:

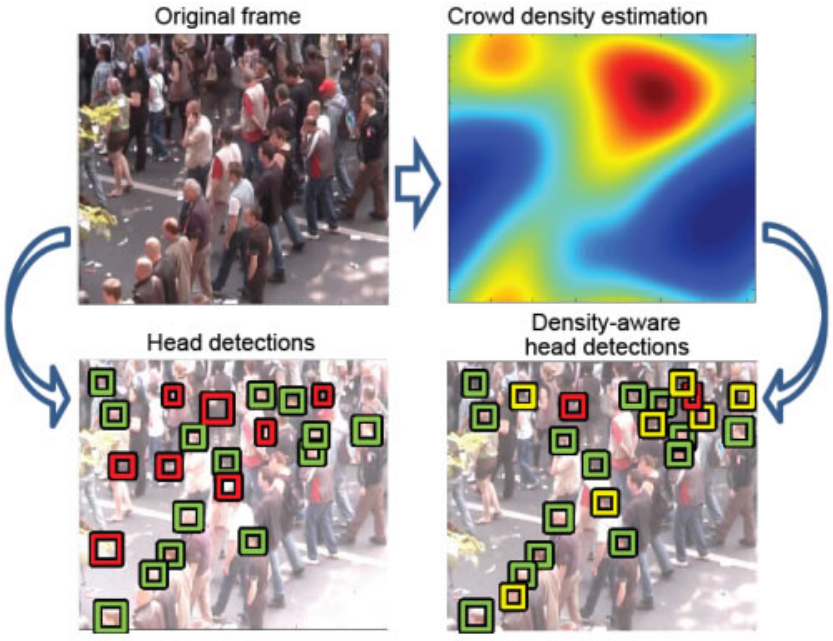

Figure 5: DenseBody's 3D object modeling

This crowd recognition technology shows areas where people have stayed for a long time and can catch suspicious activity within a large crowd, so that "action can be taken right away, with minimal damage", as further explained in VICE's video [1]

4 Computational Complexity analysis

Because video surveillance can consume a huge amount of bandwidth and because the load of bandwith varies so much, assuring good bandwith is fundamental for video surveillance systems. To analyze this complexity I will focus on these topics, as cited here [9]:

- Compression and Bandwith
- Bandwith per Camera
- Bandwith variance over time
- Sizing networks for video surveillance

4.1 Compression and Bandwith

All video surveillance, in essence, is sent to an IP network compressed, and even though the cameras can produce uncompressed video, they are almost always compressed before being sent to a network. And even thought it is possible to send it uncompressed the enourmous bitrate makes it unjustifiable and therefore impractical for almost all applications.

4.2 Bandwith per Camera

The amount each camera consumes is varied, and for a network total bandwith consumption, you should sum up all the cameras bandwith consumption, as demonstrated in the formula (where X equals the total bandwith consumption and m equals the total amount of cameras) and image below:

$$\sum_{n=1}^{m} n = x$$

CAMERA	BANDWITH CONSUMPTION
Camera 1	4 Mb/s
Camera 2	4 Mb/s
Camera 3	4 Mb/s
Camera 4	2 Mb/s
Camera 5	2 Mb/s
Camera 6	2 Mb/s
Camera 7	2 Mb/s
Camera 8	1 Mb/s
Camera 9	1 Mb/s
Camera 10	1 Mb/s
Total Network Load : 23 Mb/s	

Figure 6: Sum of bandwith consumption of a network of 10 cameras

4.3 Sizing networks for video surveillance

These are the usual values for camera bandwith:

- 720P 30FPS Intersection: 4 Mb/s
- 1080p 10FPS Conference Room: 0.5 Mb/s
- 1080P 10FPS Conference Room: 0.625 Mb/s
- 1080P 30FPS IR On Intersection: 5 Mb/s
- 5MP 15FPs Panoramic Office: 3.5 Mb/s
- 4K 30FPS Intersection: 7 Mb/s
- 4K 10 FPS Night Outdoors: 24 Mb/s

We usually refer to the average consumption of a 720P 30FPS camera, which consumes about 4Mb/s. So taking in the formula above, and the looking at a global network for China (accounting for 20 million camera, using the average of 4Mb/s), we get these numbers.

$$\sum_{n=1}^{2x10^7} n = 2000000100000000 Mb/s$$

So, taking in a more global perspective, per second, China needs to process about *two hundred trillion ten million Mb*. And since they plan to add in hundreds more cameras by 2020, these numbers will vastly increase.

5 Practical analysis on recognition

Besides taking a look at a more theoretical analysis, I took a dive at an implementation of DenseBody [6], which is implemented in this GitHub [10]. To summarize, this densebody implementation takes an image as an input and estimates an UV position map.

Figure 7: Left: Image input; Right: Created UV Map

They already offer a pre-trained model, you can just create your own UV maps and use the trained model. Unfortunately, I wasnt able to run the model, since it was very computationally demanding, as explained in the section above.

6 Impacts

6.1 Social Impacts

How they use ML to recognize every citizen in real-time. This is how you cite something in the references [**?**]. A footnote[1]

6.2 Ethical Impacts

Levels of freedom based on ethnicity, etc..

7 Conclusions

What we've come to conclude

8 References

[1] Ella Reeve, *How China Tracks Everyone, VICE on HBO:* `https://www.youtube.com/watch?v=CLo3e1Pak-Y`
30, jan. 2019.

[2] Antoine Boquen, `https://nhglobalpartners.com/chinas-social-credit-system-explained/`
May 19, 2020.

[3] Gulizar Haciyakupoglu, Wu Shang-Su, `https://www.lowyinstitute.org/the-interpreter/china-social-credit-system-black-market-and-inequalities`
March 18, 2019.

[4] `https://www.nytimes.com/2018/07/24/business/dealbook/china-facial-recognition.html`
July 24, 2018.

[5] `https://www.cloudwalk.cn/`

[6] `https://arxiv.org/abs/1903.10153v3`

[7] `https://www.bgp4.com/2019/05/10/what-is-chinas-skynet-yes-it-is-what-you-think-it-is/`

[8] `paper.people.com.cn/rmzk/html/2017-11/20/content_1825998.htm`

[9] `https://ipvm.com/reports/bandwidth-guide-for-video-networks`

[10] `https://github.com/Lotayou/densebody_pytorch`

9 Appendix

9.1 Important Notes

[1] The algorithm called *NameHere* has X modules, basing off: `Example Code usage`