Amazon Rekognition Should Diversity Data and Strengthen Management

Amazon Rekognition is a pre-trained computer vision tool which is used to extract data and information from videos or images, and the settings can be changed to fit particular customer needs. It is a powerful tool since it can be used in many different aspects, such as face compare and analysis, celebrity recognition, text recognition and so on[1]. However, serious ethical problems exist. This report will analyse three of these issues, and provide suitable recommendations to solve them.

Bias and Discrimination

Firstly, Amazon Rekognition has been criticized for higher face recognition error rates for certain groups of people, which clearly shows bias. According to the researches, Rekognition performs poorly when recognizing images of women[2] and people with darker skin[3]. Besides biometric characteristics, wearing a headgear is also proved to increase the probability of being misidentified[4]. When considering emotion detection, more failure occurs for older adults[5]. These show indisputably that Rekognition is unethical. Amazon Rekognition is not a fair tool that shares an average error rate for everyone, with or without headwear. Instead, its discrimination has caused public indignation.

The main cause for large difference in error rates is the bias of underlying training data. When collecting image or video data for Rekognition, Amazon mainly focuses on men with white skin. In contrast, the number of image records of women or people with dark skin is very limited. Since the algorithm learns based on the dataset it is provided, it is not surprising that the algorithm would perform badly on latter group. The headgear and age problems are the same, where Amazon lacks dataset of people wearing headgear, and the emotion detection is not targeted on the elders.

Privacy

In addition, Amazon Rekognition is accused of intruding human rights. It was used by governments and law enforcement agencies as a surveillance tool[6]. The body cameras worn by policemen could automatically identify and record the images of citizens into a database, even if they were just walking by. Without informing or getting consent from citizens, the police agencies were certainly abusing human rights. The impact is enormous, because it is estimated that in the USA, half of the population are recorded without notice. Moreover, the law enforcement agencies are using Rekognition against immigrants and people of color[7]. They can keep track of the immigrants' location and discern their actions after applying Rekognition to surveillance cameras.

It is true that the ones who destroy privacy are governments and law enforcement agencies, not directly Amazon Rekognition, but this does not mean Rekognition is not responsible. Knowing the potential ability to surveil people, Amazon does not limit user access effectively, especially for law enforcement agencies, which as a result, causes public panic.

Safety

The third ethical issue of Amazon Rekognition is the safety of its system. In GeekPwn, an international cyber security competition, a team led by an employee of Facebook breached Amazon Rekognition in only 20 minutes[8]. This caused the host of the competition to be identified as Schwarzenegger by Rekognition. This is really dangerous. If Rekognition can be broken in 20 minutes in the competition, it would be easy for hackers to invade the system as well. Given the massive use of Rekognition nowadays, hackers can masquerade as someone else. They can log in others' online banking accounts, and cause property loss. Or hackers can pretend to be celebrities or politicians, giving fake speeches and collapsing their reputation. Thus, the weak system with safety issue leads to collapse of trust.

On the other hand, specific biometric information extracted by Rekognition is stored into its database as binary numbers, series of 0s and 1s. Since the system of Amazon Rekognition is not robust enough, this database containing secret biometric data is also at risk of being stolen[9]. In this context, it is impossible to persuade the public to trust any authentication based on these data. The society would degenerate, and people would start to use more primitive attestation methods, which is not what we want to see.

Recommended Improvements

Many actions have to be taken to make Amazon Rekognition more ethical. For the bias issue, Amazon needs to diversity training data and provide regular test for potential bias. The performance of Rekognition on certain group of people is based on the amount of training data it gets. Therefore, to avoid discrimination, Amazon should get equal amount of training data for all gender, ethnic groups and age groups, both with and without headwear. To achieve this, Amazon could recruit equal numbers of volunteers from all groups, and use their images and videos as the dataset. Same number of random samples for every group is extracted as the test set, while others are used to train Rekognition.

During regular tests, if any bias is detected, for example, the error rate for a specific ethnic group is remarkably higher, Amazon should take action immediately by getting more data from people of that particular group. At the same time, more complex algorithms can be developed to achieve a higher overall accuracy.

Amazon had already tried to solve the privacy issue by stopping the use of Rekognition by police. At first, it was a one-year ban[10], while nearly a year later,

Amazon announced that the police would be blocked from Rekognition indefinitely[11]. While this can be useful to some degree, it does not solve the fundamental problem, since users other than law enforcement agencies may also misuse Rekognition and intrude privacy. Therefore, Amazon should strengthen their user access management as a whole. Before allowing any users to apply Rekognition, Amazon should provide them with detailed ethical guidelines, getting their consent, and ask for their purpose. The service is only available after approval. Meanwhile, Amazon should inspect users' actions regularly after informing them. Whenever potential malicious behaviors are discovered, stop user access at once. By doing this, Amazon can better prevent anyone from misusing Rekognition, not just the police or governments.

The safety issue can be solved by consolidating firewall and adding encryption. These are ways to make the system more robust, so that it would be harder for hackers to invade the system and obtain biometric information. It is possible for Amazon to divide its employees into two groups. One group is responsible for building a more robust system, while the other tries hard to breach it. Whenever the hacker group succeeds, they can inform the developer group the weakness of the system, and the developer group can make improvements accordingly. After several turns of the adversarial practices, the system should be much better. In the meantime, Amazon should provide training on cyber security for all employees. Equipped with these knowledge, employees can work more efficiently, and produce better outcomes.

The solutions listed above may be difficult and costly. Getting enough volunteers from all groups of people is time-consuming, and if the number is not achieved after a long time, Amazon needs to add pay to attract participants. Retraining the system with new dataset takes huge efforts as well. Specific staff need to be allocated to implement regular tests for bias and malicious behaviors. Training workers costs money, and employees may fail to build a system which is robust enough. Despite these obstacles, Amazon still needs to take actions straight away. Otherwise, it will continue being criticized for abusing human rights and may even face legal disputes from people being discriminated, or the ones who lose money because of insecurity of Amazon system.

Decision-making process

The reason why these ethical issues arise is that Amazon lacks careful consideration before deploying Rekognition. The mainly used training data are males with white skin, the user management is poor, and the system is insecure. In the future, Amazon should brainstorm possible ethical issues and test for their presence before deploying applications. If Amazon did so in the past, the issues demonstrated above would not arise in the first place, and the criticism would not exist.

Conclusion

Although Amazon Rekognition is powerful, there are several serious ethical issues

related to it, including discriminating certain groups of people, intruding privacy, and facing safety crisis. These could be solved if Amazon takes more effort to diversity its training data, strengthen user management, and build more robust systems. By doing this, Amazon is able to get rid of the risk of legal disputes and abusing rights. If Amazon takes actions in time, Rekognition will be an accurate tool to make people's accounts safer, and their life more convenient.

Bibliography

[1] Amazon. (2023). What can Amazon Rekognition do?. [Online]. Available from: https://aws.amazon.com/rekognition/?nc1=h_ls [Accessed: 24 February 2023].

[2] Schwemmer, C., Knight, C., Bello-Pardo, E. D., Oklobdzija, S., Schoonvelde, M., & Lockhart, J. W. (2020). Diagnosing gender bias in image recognition systems. *Socius*, *6*, 2378023120967171.

[3] Wang, M., Deng, W., Hu, J., Tao, X., & Huang, Y. (2019). Racial faces in the wild: Reducing racial bias by information maximization adaptation network. In *Proceedings of the ieee/cvf international conference on computer vision* (pp. 692-702).

[4] Tol, J. (2019). *Ethical Implications of Face Recognition Tasks in Law Enforcement* (Doctoral dissertation, Informatics Institute).

[5] Kim, E., Bryant, D. A., Srikanth, D., & Howard, A. (2021, July). Age bias in emotion detection: An analysis of facial emotion recognition performance on young, middle-aged, and older adults. In *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 638-644).

[6] Godfrey, C. (2020). Legislating Big Tech: The Effects Amazon Rekognition Technology Has on Privacy Rights. *Intell. Prop. & Tech. LJ*, *25*, 163.

[7] Crist, R. (2019). Amazon's Rekognition software lets cops track faces: Here's what you need to know. [Online]. Available from: https://www.cnet.com/home/smart-home/what-is-amazon-rekognition-facial-recognition-software/ [Accessed: 24 February 2023].

[8] Duan, Q. (2018). The post-90s Chinese programmers break through Amazon face recognition, which can make Jiang Changjian be mistaken for Schwarzenegger. [Online]. Available from: https://www.yicai.com/news/100045603.html [Accessed: 24 February 2023].

[9] AiLab. (2017). Is there a vulnerability in "face swiping login"? The theft of biometrics is more troublesome. [Online]. Available from: http://ailab.cn/20170322 83500_1/ [Accessed: 24 February 2023].

[10] Levy, A. & Hirsch, L. (2020). Amazon bans police use of facial recognition technology for one year. [Online]. Available from: https://www.cnbc.com/2020/06/10/amazon-bans-police-use-of-facial-recognition-technology-for-one-year.html [Accessed: 24 February 2023].

[11] Metz, R. (2021). Amazon will block police indefinitely from using its facial-recognition software. [Online]. Available from: https://edition.cnn.com/2021/05/18/tech/amazon-police-facial-recognition-ban/index.html [Accessed: 24 February 2023].