In contrast to the traditional method of using the human eye to examine the surveillance record to detect and gather reliable information, face recognition for video surveillance aims to accurately identify the presence of a target person through surveillance cameras placed throughout the city, greatly freeing up labor and improving efficiency.Face recognition is already relatively efficient and has gained widespread adoption. However, problems such as low resolution, high noise, angle, pose, lighting and occlusion still bring challenges for face recognition in video surveillance.

Among the many problems, recognising faces from low-quality video frames is one of the challenges prevalent in video surveillance systems. On the one hand, in surveillance environments where cameras are usually placed far away from the subject, the resolution of the captured face images is low due to distance, lighting and shooting angle; on the other hand, in the age of dense urbanization, cameras take pictures of scenes with dense crowds in the background, which severely distorts the faces that are photographed; all of which leads to poor accuracy of most face recognition algorithms under video surveillance applications. Numerous studies have made an effort to solve this issue by creating algorithms that can deliver acceptable recognition performance at low quality video frames. [8] takes an end-to-end approach to match high-resolution (HR) images with low-resolution (LR) images in surveillance videos. [12, 15] use super-resolution methods to improve images. [12] improves the number of descriptors in the image and mitigates the effects of noise based on super-resolution faces. In [15] the video frames are provided sequentially to the Viola-Jones algorithm to detect the face part and the quality of the face part is improved by applying a super-resolution scheme based on bicubic interpolation. In addition, [16] used multi-resolution convolutional neural networks (MRCNN) and anti-aliasing techniques to solve the low-resolution problem.

The majority of face recognition in video surveillance today is "closed-set," which only recognizes the identity of previously registered objects. However, "open-set" has gained popularity as a result of the differences between the source and target domains, which make it less effective when transferring face recognition systems from controlled environments to uncontrolled scenes. [9] proposed the use of fuzzy ARTMAP neural networks to solve the open-set single-sample face recognition problem in real-world video surveillance scenes. Focusing on solving the pose change problem, [11] proposed an automatic pose normalisation technique without model fitting and human intervention, which greatly improves the performance of open-set single-sample face recognition methods in surveillance environments.

Approaches to face recognition mainly include 1) traditional methods ,which rely on hand-crafted feature extraction techniques and a pre-trained classifier along with fusion, and 2) deep learning methods, which automatically learn features and classifiers together utilizing enormous quantities of data[10,13]. Some recent studies have shown that patch-attention is more powerful than convolution in recognition models, and [14] designed a patch-attention generative adversarial network (PA-GAN) model to aggregate some robust features on behalf of a set of raw surveillance frames, which not only improves recognition accuracy, but also reduces the computational cost of face matching.

With the development of artificial intelligence technology, face recognition technology combined with surveillance camera technology has become an important tool for obtaining effective information and improving the efficiency of analysis in mega-city governance. While being applied on a large scale, the problem of privacy leakage has aroused widespread concern. This paper attempts to explore a new paradigm for large-scale camera urban applications from the perspective of human-like cognition by performing face reduction encryption on the recognised video images.

1. S. Bashbaghi, E. Granger, R. Sabourin, and M. Parchami, "Deep Learning Architectures for Face Recognition in Video Surveillance," in *Deep Learning in Object Detection and Recognition*, X. Jiang, A. Hadid, Y. Pang, E. Granger, and X. Feng, Eds. Singapore: Springer Singapore, 2019, pp. 133-154.
2. W. N. I. Al-Obaydy and S. A. Suandi, "Open-set face recognition in video surveillance: a survey," in *InECCE2019*: Springer, 2020, pp. 425-436.
3. A. H. Ahmad et al., "Real time face recognition of video surveillance system using haar cascade classifier," vol. 21, no. 3, pp. 1389-1399, 2021.
4. X. Zhao, Y. Chen, E. Blasch, L. Zhang, and G. Chen, "Face recognition in low-resolution surveillance video streams," in *Sensors and Systems for Space Applications XII*, 2019, vol. 11017, pp. 147-159: SPIE.
5. W. N. I. Al-Obaydy, S. A. J. N. C. Suandi, and Applications, "Open-set single-sample face recognition in video surveillance using fuzzy ARTMAP," vol. 32, no. 5, pp. 1405-1412, 2020.
6. C. Shirley, N. Ram Mohan, B. J. M. S. Chitra, and S. Processing, "Gravitational search-based optimal deep neural network for occluded face recognition system in videos," vol. 32, no. 1, pp. 189-215, 2021.
7. W. N. I. Al-Obaydy, S. A. J. M. T. Suandi, and Applications, "Automatic pose normalization for open-set single-sample face recognition in video surveillance," vol. 79, no. 3, pp. 2897-2915, 2020.
8. N. Singh, S. S. Rathore, S. J. M. T. Kumar, and Applications, "Towards a super-resolution based approach for improved face recognition in low resolution environment," pp. 1-33, 2022.
9. Z. Lei, X. Zhang, S. Yang, Z. Ren, and O. F. J. E. I. S. Akindipe, "RFR-DLVT: a hybrid method for real-time face recognition using deep learning and visual tracking," vol. 14, no. 9-10, pp. 1379-1393, 2020.
10. M. Liu, J. Liu, P. Zhang, and Q. J. I. A. Li, "PA-GAN: A patch-attention based aggregation network for face recognition in surveillance," vol. 8, pp. 152780-152789, 2020.
11. D. S. Dharrao, N. J. J. I. J. o. C. I. Uke, and Applications, "Fractional Krill–Lion algorithm based actor critic neural network for face recognition in real time surveillance videos," vol. 18, no. 02, p. 1950011, 2019.
12. M. J. I. J. o. A. C. S. Imandito and Applications, "Face Recognition on Low-Resolution Image using Multi Resolution Convolution Neural Network and Antialiasing Method," vol. 10, no. 12, 2019.

视频监控的人脸识别通过分布在城市各个角落的监控摄像头试图精确识别目标人物的存在，与传统人眼观看监控纪录来识别和捕获有效信息的方法相比，解放了人力，提高了分析效率。目前人脸识别的识别效率已经较高，且已经获得了广泛的应用。但是低分辨率、高噪音、角度、姿势、光照、遮挡等问题仍为视频监控的人脸识别带来挑战。

在众多问题中，从低质量的视频帧中识别人脸是视频监控系统普遍存在的挑战之一。一方面，在监控环境中，摄像头通常放置在离被拍摄者很远的地方，由于距离、光线和拍摄角度的原因，拍摄到的人脸图像分辨率低；另一方面，在密集的城市化时代，摄像头拍摄画面内人群密集；导致捕获到的人脸图像严重失真；这都导致大多数人脸识别算法在视频监控应用下的准确性较差。目前已有一些研究尝试解决该问题，开发能够在低分辨率视频帧下获得满意识别性能的算法。[8]采取一种端到端的方法来匹配监控视频中的高分辨率（HR）图像和低分辨率（LR）图像。[12，15]使用超分辨率方法来改进图像。[12]在超分辨率人脸的基础上提高了图像的描述符数量并减轻了噪声的影响。[15]中将视频帧依次提供给Viola-Jones算法来检测人脸部分，并通过应用基于双立方插值的超分辨率方案来提高人脸部分的质量。除此之外，[16]采用了多分辨率卷积神经网络（MRCNN）和抗锯齿技术来解决低分辨率问题。

目前视频监控中的大多数人脸识别都使用封闭集格式，只识别先前注册的对象的身份，但是由于源域和目标域之间的差异，在受控环境中表现良好的人脸识别系统在迁移至非受控场景中时表现并不令人满意，开放集格式已经变得更加流行。 [9]提出使用模糊ARTMAP神经网络来解决现实世界视频监控场景中的开放集单样本人脸识别问题。[11]关注解决姿态变化问题，提出一种无需模型拟合和人工干预的自动姿势归一化技术，极大地提高了监控环境中开集式单样本人脸识别方法的性能。

人脸识别的方法主要传统方法。。。。，和深度学习方法。。。[13][10]。最近的一些研究表明，patch-attention在识别模型中比卷积更加强大，[14]设计了一个补丁关注生成对抗网络（PA-GAN）模型来代表一组原始监控帧聚集一些健壮的特征，这不仅提高了识别精度，也降低了人脸匹配的计算成本。

随着人工智能技术的发展，人脸识别技术与监控摄像技术相结合，已成为特大型城市治理中获取有效信息、提高分析效率的重要手段。在大规模应用的同时，隐私泄露的问题引起了人们的广泛关注，本文试图从类人认知的角度探索大规模摄像头城市应用的新模式，对识别到的视频图像进行人脸降维加密。