This work was supported in part by the U.S. Department of Commerce under Grant BS123456. *(Corresponding author: Bin. He).* Here you may also indicate if authors contributed equally or if there are co-first authors.

The next few paragraphs should contain the authors' current affiliations, including current address and e-mail. For example, First A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail: author@ boulder.nist.gov).

Second B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

Third C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba 305-0047, Japan (e-mail: author@nrim.go.jp).

Mentions of supplemental materials and animal/human rights statements can be included here.

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org

Face Privacy Protection and Self-decryption Method Based on Humanoid Association Mechanism

Zhongpan Zhu, Qiwei Du, Kaijing Ma, Bin He*, *Member, IEEE*, Zhipeng Wang, Gang Li ,Junze Zhu

***Abstract*— With the large-scale industry application of artificial intelligence and video surveillance, massive video data storage and personal privacy issues are highlighted, which restrict the application expansion. From the perspective of humanoid memory mechanism, we propose a video abstraction encryption and decryption algorithm based on high and low dimensional information association cognitive mechanism, which uses face recognition algorithm to locate human faces and encrypt the original video with mosaic encryption method, and then perform spatio-temporal index encoding, and further use abstract face feature memory to match and decode with the same identity person to construct a de-mosaic decryption key function. The core innovation is the humanoid memory mechanism for parsing and modelling, and combined with specific AI techniques such as YOLO and GAN for initial experimental validation in face encryption and decryption. The algorithm research will be important in the research of video information compression and storage, person re-identification and personal privacy protection.**

***Index Terms*—Enter keywords or phrases in alphabetical order, separated by commas. For a list of suggested keywords, send a blank e-mail to keywords@ieee.org or visit http://www.ieee.org/organizations/pubs/ani_prod/keywrd98.txt**

I. Introduction

The surveillance cameras distributed in all corners of the city play an indispensable role in the city security management. Surveillance video/image data collection and analysis based on AI and IOT technologies has been an important technical grip for different scenes in smart city development.

The extraction of trust information from surveillance data has attracted the interest of many researchers and has led to the analysis of images from numerous IoT vision sensors [1-3]. However, the large-scale deployment of vision sensors leads to a number of challenges: 1) First, the huge number of camera video images leads to a data disaster. At 30 frames per second and 5MB per image, a single camera generates a data storage requirement of 12,656.25Gb a day, while IHS research indicates that there will be over one billion surveillance cameras worldwide in future. These video stores take up a large amount of hardware resources, and no data center can withstand the daily growth of video data, which must be overwritten on a regular basis [2]. Secondly, information redundancy in massive camera video data leads to key information being overwritten and video-based information retrieval being difficult [3]. In addition, massive video transmission takes up a large amount of communication bandwidth, and communication costs are high, making it difficult to achieve widespread cameras for collaborative use to achieve mega-city governance [4]. Meanwhile surveillance cameras have led to the leakage of residents' biometric privacy, raising ethical and regulatory concerns. How to safeguard the functionality of surveillance cameras while improving the above challenges has become a research direction for a wide range of scholars.

In this paper, we take a humanoid cognitive perspective to carry out theoretical research for exploring new models of large-scale camera urban applications. We humans, from infants to the elderly, perceive a large amount of picture information with both eyes over decades and can have long-term clear memories of the people and things we experienced. However, we are often unable to reproduce all of the image information that occurred, but rather combine it with high-dimensional semantic abstraction to achieve coarse-grained picture recall. We also tend to remember familiar faces not through detailed facial features such as single or double eyelids, but into general impressions of higher-dimensional semantic information. In addition, the high-dimensional abstract semantics in our human brain memory plays an important role in blurring human decryption recognition. Humans can recognize acquaintances through blurred or partially blurred facial images, but not strangers. The process of humanoid perceptual memory mechanism to handle the massive amounts of video data is difficult to have a theoretical explanation. But the association between low-dimensional fine-grained information and higher-dimensional coarse-grained information for humanoid perceptual data compression and decryption has theoretical significance and practical value, which are worth using for processing massive surveillance vedio data. In this paper, we try to propose an autonomous face degradation encryption and decryption algorithm based on the above humanoid association memory mechanism.

II. RELATED WORK.

A. face recognition of video surveillance

In the perspective of recent advances in the field of AI-driven face recognition of video surveillance, the human face object tracing for video surveillance has gained widespread adoption in urban security and community management. A lot of scholars are committed to the research of computer vision technique with promising accuracies and efficiencies for face recognition and object detection [5-7]. The face recognition methods 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, 14]. With the development of deep learning technology, the application boundary of face recognition will be gradually opened. 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. Suandi proposed fuzzy ARTMAP neural networks to solve the open-set single-sample face recognition problem and an automatic pose normalization technique without model fitting and human intervention, which greatly improves the performance of open-set single-sample face recognition methods in surveillance environments [9,11]. The "open-set" face recognition prone to increase the human privacy exposure degree in the ubiquitous city surveillance network.

The low resolution of urban monitoring picture and the difficulty of small face feature extraction are being changed. Even though the surveillance cameras are usually placed far away from the objects and the resolution of the captured face images is low due to distance, extensive research has been carried out for recognizing acceptable recognition features at low quality video frames. Zhao et al. took an end-to-end approach to match high-resolution (HR) images with low-resolution (LR) images in surveillance videos[8]. Singh et al. improved the number of descriptors in the image and mitigates the effects of noise based on super-resolution faces[12]. Dharrao et al. used the Viola-Jones algorithm to detect the face part in the video sequential frames and improved the quality of the face part by applying a super-resolution scheme based on bicubic interpolation[15]. In addition, the multi-resolution convolutional neural networks (MRCNN) and anti-aliasing techniques were adopted to solve the low-resolution problems[16].

The development trend of face recognition technologies shows that the challenge of citizen's face privacy feature under the ubiquitous cameras is more and more serious. How to explore a new paradigm for large-scale camera urban applications from the perspective of humanoid cognition by performing face reduction encryption on the recognized video images are meaningful.

B. Face encryption and decryption algorithm

The problem of privacy leakage has aroused widespread concern. Face recognition of video surveillance have become ubiquitous in daily lives, but it is difficult to balance between intelligent vision applications and personal privacy protection. In addition to improving relevant laws and regulations to regulate the acquisition, storage and use of videos, corresponding technical measures are needed to protect personal privacy. The cryptography-based face privacy protection scheme selectively encrypts the face region in the video that shows the identity and can be decrypted to recover the original video in case of future legitimate demand. How to integrate the autonomous face degradation encryption and decryption algorithm of humanoid association memory mechanism into AI face recognition algorithm is an urgent breakthrough direction.

Most of the existing face encryption schemes are homomorphic-based[18-26]. There are three different types of homomorphic encryption schemes: (1) partially homomorphic encryption, (2) somewhat homomorphic encryption and (3) fully homomorphic encryption (FHE).Tamiya et al. proposed a successful homomorphic encryption-based face template protection scheme by computing the squared Euclidean distance between facial features with a single homomorphic multiplication method[20]. Román et al. suggested using the Kyber and Saber public key encryption (PKE) algorithms along with homomorphic encryption (HE) in facial recognition systems to achieve smaller protected template and key sizes and faster execution times than other HE schemes that use lattices[21]. The use of fully homomorphic encryption algorithms provides a higher level of privacy authentication for the queried face. Huang offered a successful, privacy-preserving face verification method based on a corrupted circuit and fully homomorphic encryption[22]. Some researchers used CKKS fully homomorphic encryption to encrypt the normalised facial feature vector [18,23].

Due to the low computational efficiency of using homomorphic encryption, other studies tried to find lightweight algorithms to encrypt faces. Tan et al. proposed a novel approach to implement video-based ring-learning (ring-LWE) cryptography for face encryption and decryption on a graphics processing unit (GPU)[29].Duong-Ngoc et al. proposed a novel method to comprehensively protect facial images extracted from videos based on NewHope cryptography for post-quantum cryptosystems, greatly reducing the time for encryption and decryption [27]. Zhao et al. proposed and implemented a simple and efficient speckle-based optical cryptosystem to encrypt face images by seemingly random optical speckles at the speed of light, by training an cryptographic neural network to decrypt face images from random speckles [28]. A fast block scrambling method was used to scramble the detected faces [31,32]. In addition, an encryption technique using face biometrics to generate random phase masks [33]. A THM (Tent-Henon Map) chaotic encryption of faces was proposed in combined with the properties of tent chaos and Henon chaos[34]. Liu proposed a RGB image encryption algorithm based on DNA encoding and chaos map [35]. Wu proposed a Generative Adversarial Network (GAN)-based method to encrypt facial features using Wasserstein Generative Adversarial Network Encryption (WGAN-E) [36]. Ashiba used a graph theory-based graph first decomposition mask (GFH) coding algorithm[37]. There are still room for improvements in terms of computational communication efficiency and privacy-preserving effects. Active perception of key privacy features for target encryption based on humanoid cognitive mechanism provides a preliminary exploration in this direction.

C.Humanoid memory cognition

Human brain is a typical encryption and decryption processing device with low energy consumption and high efficiency. The brain can store learned concepts in memory and recall them when it sees partial or broken patterns. Franklin et al. proposed a structured event memory model (SEM) of event cognition, illustrating human abilities in event segmentation, memory and generalization. SEM can be extended to a high-dimensional input space to produce humanoid event segmentation for natural video data, and illustrates a wide range of memory phenomena [38]. Sun et al. proposed a new model humanoid visual cognitive and language-memory network for visual dialog (HVLM) to simulate global and local dual-view cognition in the human visual system to comprehensively understand images [39]. Inspired by humanoid perception and memory we explored a new model of face privacy protection for urban large-scale camera monitoring with . The research of this algorithm is of great significance to the research of video information compression and storage, character recognition and personal privacy protection.

III. PROPOSED APPROACH

A. Problem description.

The process of human face perception and identity recognition based on fuzzy impression memory association is highly complex. Each of us sees many faces in daily life scenarios, however, not all the information about faces are remembered. As shown in figure 1 for example, when some people meet with each other unintentionally, their mind will unconsciously recall that they have seen such a face at a certain time, place and event. Moreover, they can recall the memory of more detailed scene and clearer features. The process can actually be simplified as the human brain perceives the concrete face image information seen by the eyes to extract high-dimensional abstract semantic features. The high-dimensional abstract semantic features are retrieved and matched with the high-dimensional semantic information indexed in memory combining person, event, time and place, and the past feature-blurred memory scene is clearly reproduced in combination with the current perceived face image.

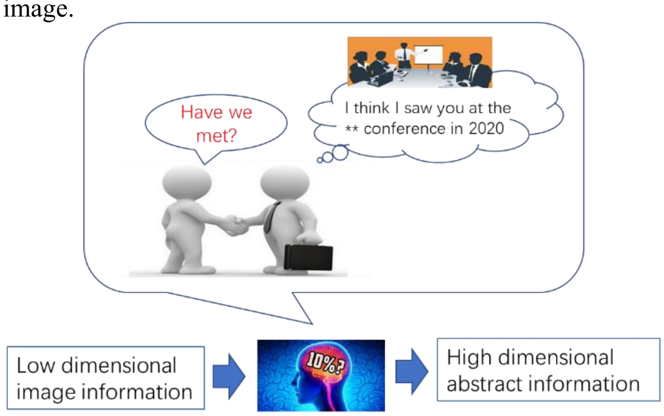

Fig. 1. Case: Humanoid abstract associations triggered by perceptual features.

The general expression of the Humanoid Association is as follows.

$$I(A_i, A_{ip}, B_i) \rightarrow I(A'_i, A'_{ip}, B'_i) \tag{1}$$

where *I is the image*, A_i is the set of low-dimensional full-dimensional information about the i_{th} people's face perceived by the brain in the first stage, B_i is the set of high-dimensional abstract semantic features of i_{th} people's face formed by the brain in the mind based on A_i, $A_{ip}$is the encrypted data set of A_i, A'_{ip} is the decryption set partially from A_{ip} and B_i, A'_i is the set of low-dimensional full-dimensional information about the face perceived by the brain in the second stage, and B'_i is the set of high-dimensional abstract semantic features formed by the brain in the mind based on A'_i . The algorithm for solving the above expression is as follows.

Algorithm: encryption and decryption

Input: A_i, B_i, A'_i,
Output: A'_{ip}
For A_i in Brain Do encryption Key matching
Using cerebral neural network for high dimensional semantic abstraction
$f_1(A_i) \rightarrow A_{iP}$, where $A_{iP} \subsetneq A_i$, $i=1, 2, \ldots\ldots$
For B_i in Brain Do decryption
$f_1(A'_i) \rightarrow B'_i$
Find matching key B'_i to B_i, where is the maximum face similarity
$P(A'_{ip} \mid A_{ip}) = f_2(B'_i \cap B_i) \rightarrow 1$
For A_{iP} in Brain Do encryption
$f_3(A_{iP}, B_i) \rightarrow A'_{ip}$
return A'_{ip}

This paper combines the above humanoid perceptual associative memory algorithm with the face encryption and decryption requirements of surveillance video to solve the following problems.

1) To modeling the humanoid cognitive mechanism, the high-dimensional abstract memory and compressed perception process f_1 function need to be solved. and propose an artificial intelligence algorithm for solving A_{ip} and B_i to identify and locate faces in videos, extract high-dimensional semantic features while encrypting video faces with reduced resolution.

2) Drawing on humanoid associative memory mechanism, the algorithm models the memory storage of high-dimensional semantic features and associative matching f2, and proposes a recall-triggered matching index mechanism to achieve associative memory matching based on B1 and B2.

3) Drawing on the humanoid perception-triggered recall mechanism, the associative recall of high-dimensional semantic features A1 and B2 low-resolution video is modelled to solve f3 for indexing location as well as high-resolution decryption.

B. AI Methodologies

Inspired by humanoid perception, compressed memory, and associative recall, we propose an algorithmic framework that can be used to encrypt/decrypt surveillance video faces as shown in the figure below.

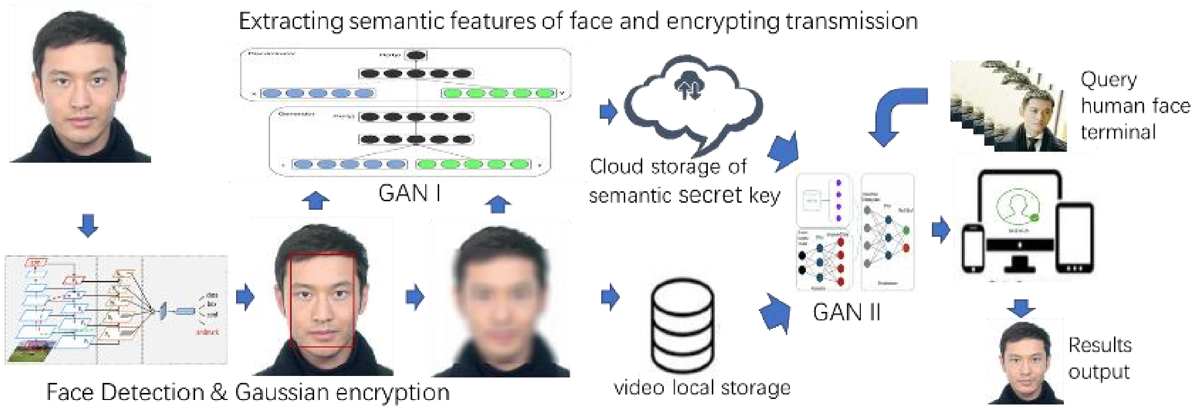

Fig. 2. Framework of personal self-decryption.

1) Encryption method

For the video frame input V, the YOLO5-face deep learning model ϕ is used to achieve the recognition and localisation of faces by the surveillance cameras at the edge end, into obtaining $A_1 = \phi(V)$, YOLO5-face is chosen because the model targets the face recognition segmentation needs, adds landmark branches in YOLOV5, and improves the accuracy of face detection and localisation by regressing the wing loss function through five facial key points.

$$loss(s) = loss_O + \lambda_L \cdot loss_L$$

$$wing(x) = \begin{cases} w \cdot ln(1+|x|/e), & \text{if } x < w \\ |x| - C, & \text{otherwise} \end{cases}$$

After completing face target detection, the face in the recognition frame is subjected to Gaussian blurring, i.e. a Gaussian convolution budget is applied to the face image with the probability density distribution function shown below.

$$\mathrm{f(x,y)} = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{d_x^2+d_y^2}{2\sigma^2}}$$

2) Key storage and matching

Unlike traditional video surveillance systems, this method no longer stores the original video, but chooses to locally store the encrypted video, while uploading the high-dimensional abstract semantics to the cloud for subsequent processing and analysis. For example, for the face retrieval service of post-surveillance, as the local storage of encrypted faces loses a large amount of face feature information, the video cannot be retrieved for review, but needs to be indexed for the high-dimensional semantic B for query service, which is similar to the human perceptual memory. In order to make the AI system capable of humanoid high-dimensional abstract computation, we establish an adversarial learning network GAN cognitive model for subsequent analysis of YOLO5-face localised faces, extract high-dimensional abstract semantic features B, and fuse video frame time series, edge camera's own latitude and longitude and pixel coordinates as the sign bit encoding of high-dimensional abstract semantic features into the cloud database, high-dimensional abstract The mapping between the high-dimensional abstract semantic feature B and the encrypted face image frame can be associated with the above marker bit encoding to facilitate the subsequent video decryption work.

The calculation to obtain the high-dimensional abstract semantic B is as follows. Through the previous original video A_1 and encrypted video A_{1p} form a paired data set, A_{1p} coupled with random noise э as the input to the generator for training and learning, while A_1 is used as the discriminator input for judgement, making $A_{1p} + э \rightarrow A_1$. In the training process, we embed a subspace model with orthogonal bases in each generative network layer used to obtain the hierarchical semantics of the training model, which in turn uses the abstract semantic feature B_1 as the key for decoding A_{1p} and stores it in the cloud-based key repository. The solving process of $f_1(A_i)$ is thus completed.

Since the high-dimensional abstract semantic features learned by GAN networks often do not have interpretability, in order to study how to match face retrieval by abstract semantics B_1 and B_2 also need to conduct in-depth research on abstract semantic B. Currently, in the field of face recognition, the technology of matching features of faces through deep neural networks to determine the identity of faces is more mature, while the abstract semantics B_1 and B_2 as identity keys to determine the identity of faces becomes more challenging. The other part is classified as impressionistic abstract semantics, which is like the general impression of a human face. identity, rather than simply by matching specific features. It is difficult not to give the formula f_2 for solving the feature calculation for B_1 and B_2 in the context of a specific example, so f_2 will be described specifically in the later experimental chapters.

3) Decryption process

The previous section implements the original video encryption and the process of key storage and query matching. This section discusses how to output the the decrypted video A_1 by the encrypted video A_{1p}, high-dimensional abstract semantics B_1 and the face A_2 in the second video that matches the identity of a person in the first video.

This process is equivalent to associative memory, where we can associate images from the past with the current image, and the blurred features of a face can often be made clear again.

In this regard, we build a open-set face re-identification and cGAN-based decryption model. Firstly, through the method described in the previous section, the high-dimensional abstract semantics B_2 is extracted from A_2. Then, the similarity is calculated between B_2 and all the high-dimensional abstract semantics B_{1i} in the key pool corresponding to the encrypted video A_{1p}, and the B_{1i} with the highest similarity is taken. If the similarity is lower than a certain threshold, the face is judged to be strange and further decryption is rejected; if the similarity exceeds a certain threshold, the high-dimensional abstract semantics B_{1i} is added to the generator with A_{1p} as input as a constraint, and the decrypted video A_1 is output.

【此处不知道是否有矛盾，是否还需要训练，直接通过B_1与A_{1P}求得A_1】【感觉不需要再次训练，因为之前已经训练好了生成器，由于这次新出现的人脸之前也出现过，其高维语义特征应该已经被生成器“消化”过，所以只要根据相似度找出该人脸之前存储的特征密钥，喂入生成器直接生成即可】有道理，这里表述改一下不是进一步训练，而是调用提取特征。

IV. EXPERIMENTS AND RESULTS

A. Dataset

We acquired face images based on temporal head pose changes of experimental subjects of different genders and ages, and obtained a dataset with Gaussian encryption paired with the original images using a method based on YOLO5-face with Gaussian encryption. 1000 images were acquired for each person, for a total of 20 people with a total of 20,000 images, in order to exclude other background features in the images from the subsequent GAN network high-dimensional semantic abstraction In order to exclude other background features in the images from the subsequent GAN network high-dimensional semantic abstraction, all faces were collected in the same background for the face collection process. The figure below shows a portion of the extracted paired dataset. This data encryption process also validates the feasibility of the video encryption method.

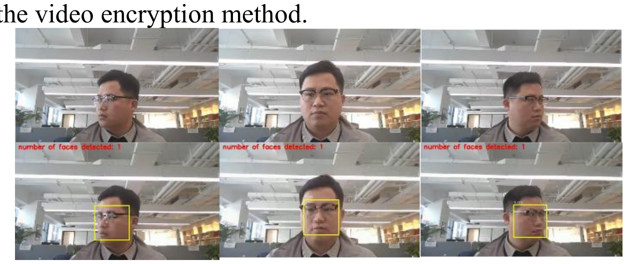

Fig. 3. Encryption method test

B. GANI Training and Encryption Process

We refer to the PULSE model to improve the GAN network and complete the training on the above dataset. Based on NVIDIA's StyleGAN algorithm, the PULSE model uses an unsupervised image super-resolution method to transform low-resolution images into high-quality, high-resolution images that can reproduce image detail features such as skin tone, eyes, lips, etc. However, the generated high-resolution face images do not resemble the real looks of the photo subjects. To this end, this paper carries out semi-supervised learning by embedding a subspace model with orthogonal bases in each generative network layer to obtain the hierarchical semantic B_1 of the training model, and its network architecture is shown in the following figure.

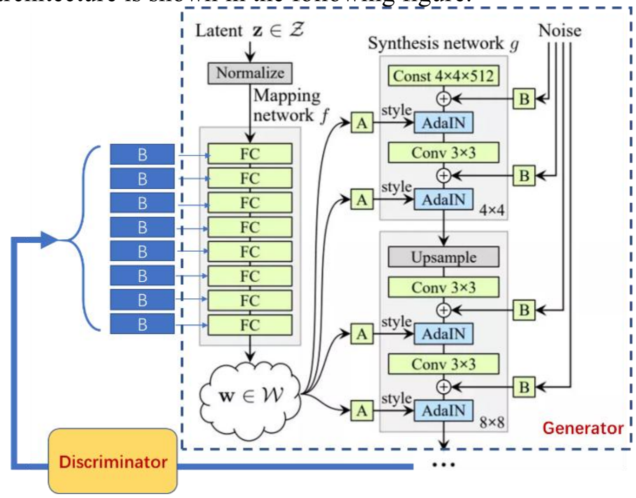

Fig. 4.此图是网上的，需要修改

We first extract the original face P_1 from the training set, form P_2 after mosaic encryption of the original face, and feed P_1 and P_2 into GAN for training after stitching them together. At the same time, P_1 is convolved several times to extract multi-dimensional face features from low-dimensional to high-dimensional, and this information is integrated and encrypted to form a key, which is bound to the identity ID of the processed face and added to the face key pool.

C. GANI Decryption Process

When decrypting, a newly captured face is input. Firstly, determine whether it has appeared in the dataset based on the similarity of face features. If it has not appeared before, the decryption is rejected. If it has appeared, its identity ID is confirmed and the previously stored feature key is used to guide GAN to decrypt the face with mosaic with the specified id and output the decrypted face.

D. Results Discussion

The encryption and decryption effects are shown below. We can see that GAN realistically restores the encrypted face guided by the previously stored face feature key.

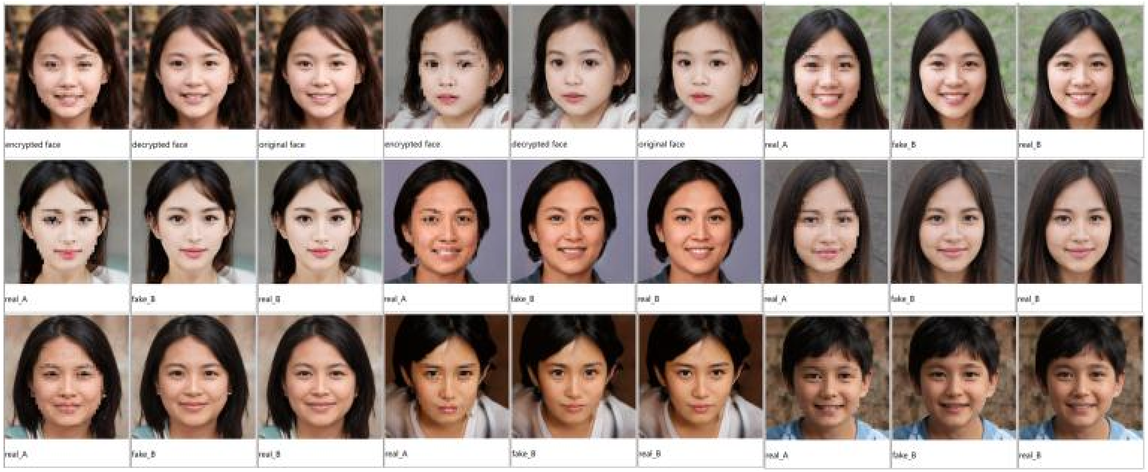

Fig. 5.

V. Conclusion

In this paper, we propose a video abstraction encryption and decryption algorithm based on high and low dimensional information association cognitive mechanism for massive video data disaster and personal privacy problems. When storing the video, the YOLO-based face recognition and face encryption algorithm is used to encrypt the video, while the high and low dimensional semantic information of the face is extracted to form a feature key, and index association is established with the face in the video and stored in the key pool corresponding to the video. When decrypting, the features extracted from the specified new faces are used to search in the key pool, and the matched faces are reduced to clear faces using the GAN model. We initially validate the feasibility of this video encryption and decryption algorithm on a self-built dataset. Our research has important implications in terms of how to strike a balance between privacy protection and machine vision research. In the future, this algorithm is promising to play an important role in privacy protection and big data storage, face re-identification, and other fields.

Appendix

Appendixes, if needed, appear before the acknowledgment.

Acknowledgment

The preferred here.

IEEE Guidelines and Policies

A full.

References

[1] Z. Gao, C. Xu, H. Zhang, S. Li and V. H. C. de Albuquerque, "Trustful Internet of Surveillance Things Based on Deeply Represented Visual Co-Saliency Detection," IEEE Internet of Things Journal, vol. 7, no. 5, pp. 4092-4100, May 2020.

[2] Ș. Kolozali et al., "Observing the Pulse of a City: A Smart City Framework for Real-Time Discovery, Federation, and Aggregation of Data Streams," IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2651-2668, April 2019.

[3] O. Styles, T. Guha and V. Sanchez, "Multi-Camera Trajectory Forecasting with Trajectory Tensors," IEEE Transactions on Pattern Analysis and Machine Intelligence. 2021, 44(11): 8482-8491.

[4] C. W. Chen, "Internet of Video Things: Next-Generation IoT With Visual Sensors," IEEE Internet of Things Journal, vol. 7, no. 8, pp. 6676-6685, Aug. 2020.

[5] Zou, Zhengxia, et al. "Object detection in 20 years: A survey." arXiv preprint arXiv:1905.05055 (2019).

[6] 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.

[7] 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.

[8] 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.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] S. Guo, T. Xiang, and X. J. S. P. Li, "Towards efficient privacy-preserving face recognition in the cloud," vol. 164, pp. 320-328, 2019.

[18] Y. Yang, Q. Zhang, W. Gao, C. Fan, Q. Shu, and H. J. W. P. C. Yun, "Design on Face Recognition System with Privacy Preservation Based on Homomorphic Encryption," vol. 123, no. 4, pp. 3737-3754, 2022.

[19] T. Yang, Y. Zhang, J. Sun, and X. J. N. P. L. Wang, "Privacy enhanced cloud-based facial recognition," vol. 54, no. 4, pp. 2717-2725, 2022.

[20] H. Tamiya, T. Isshiki, K. Mori, S. Obana, and T. Ohki, "Improved Post-quantum-secure Face Template Protection System Based on Packed Homomorphic Encryption," in *2021 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2021, pp. 1-5: IEEE.
[21] R. Román, R. Arjona, P. López-González, and I. Baturone, "A Quantum-Resistant Face Template Protection Scheme using Kyber and Saber Public Key Encryption Algorithms," in *2022 International Conference of the Biometrics Special Interest Group (BIOSIG)*, 2022, pp. 1-5: IEEE.
[22] H. Huang, L. J. J. o. I. S. Wang, and Applications, "Efficient privacy-preserving face verification scheme," vol. 63, p. 103055, 2021.
[23] L. Jiasen, W. X. An, C. Bowei, T. Zheng, Z. J. I. J. o. M. C. Kaiyang, and M. Communications, "Outsourced Secure Face Recognition Based on CKKS Homomorphic Encryption in Cloud Computing," vol. 12, no. 3, pp. 27-43, 2021.
[24] D. Sun *et al.*, "Face Security Authentication System Based on Deep Learning and Homomorphic Encryption," vol. 2022, 2022.
[25] P. Drozdowski, N. Buchmann, C. Rathgeb, M. Margraf, and C. Busch, "On the application of homomorphic encryption to face identification," in *2019 international conference of the biometrics special interest group (biosig)*, 2019, pp. 1-5: IEEE.
[26] Q. Wang, L. Gao, H. Wang, and X. J. I. A. Wei, "Face detection for privacy protected images," vol. 7, pp. 3918-3927, 2019.
[27] P. Duong-Ngoc, T. N. Tan, and H. J. I. A. Lee, "Efficient NewHope cryptography based facial security system on a GPU," vol. 8, pp. 108158-108168, 2020.
[28] Q. Zhao et al., "Speckle-based optical cryptosystem and its application for human face recognition via deep learning," 2022.
[29] T. N. Tan, Y. Hyun, J. Kim, D. Choi, and H. Lee, "Ring-LWE based face encryption and decryption system on a GPU," in 2019 International SoC Design Conference (ISOCC), 2019, pp. 15-16: IEEE.
[30] X. Wang, N. Guan, and P. J. O. Liu, "A selective image encryption algorithm based on a chaotic model using modular sine arithmetic," vol. 258, p. 168955, 2022.
[31] K. Nakai, M. Kuribayashi, and N. Funabiki, "A Study of Privacy Protection of Photos Taken by a Wide-angle Surveillance Camera," in 2021 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC), 2021, pp. 1865-1871: IEEE.
[32] K. M. Hosny, M. A. Zaki, H. M. Hamza, M. M. Fouda, and N. A. J. I. A. Lashin, "Privacy Protection in Surveillance Videos Using Block Scrambling-Based Encryption and DCNN-Based Face Detection," vol. 10, pp. 106750-106769, 2022.
[33] Y. Shen, C. Tang, M. Xu, Z. J. O. Lei, and L. Technology, "Optical selective encryption based on the FRFCM algorithm and face biometric for the medical image," vol. 138, p. 106911, 2021.
[34] Z. Liu, J. Li, and J. J. M. B. E. Liu, "Encrypted face recognition algorithm based on Ridgelet-DCT transform and THM chaos," vol. 19, pp. 1373-1387, 2022.
[35] Liu, Y. , J. Tang , and T. Xie . "Cryptanalyzing a RGB image encryption algorithm based on DNA encoding and chaos map." Optics & Laser Technology 60(2014):111-115.
[36] C. Wu, B. Ju, Y. Wu, N. N. Xiong, and S. J. E. Zhang, "WGAN-E: A generative adversarial networks for facial feature security," vol. 9, no. 3, p. 486, 2020.
[37] H. J. M. T. Ashiba and Applications, "Presented cancelable face recognition system using graph theory," pp. 1-22, 2022.
[38] N. Franklin, K. A. Norman, C. Ranganath, J. M. Zacks, and S. J. Gershman, "Structured event memory: a neuro-symbolic model of event cognition," 2019.
[39] K. Sun, C. Guo, H. Zhang, and Y. Li, "HVLM: Exploring Humanoid Visual Cognition and Language-Memory Network for Visual Dialog," *Information Processing & Management,* vol. 59, no. 5, p. 103008, 2022/09/01/ 2022.

First A. Author (Fellow, IEEE)
.

Second B. Author, photograph and biography not available at the time of publication.

Third C. Author, Jr. (Member, IEEE), photograph and biography not available at the time of publication.

TABLE I
MATCHING FROM HIGH DIMENSIONS.

LATENT CODE LAYERS USED	Accuracy	
	Cosine similarity	Euclidean distance
1-1	0.2553	0.2766
1-2	0.4681	0.4255
1-3	0.7447	0.7447
1-4	0.7660	0.7660
1-5	0.7872	0.7660
1-6	0.7872	0.8085
1-7	0.8298	0.8085
1-8	0.8298	0.8298
1-9	0.8085	0.7660
1-10	0.9362	0.8936
1-11	0.8298	0.8298
1-12	0.9362	0.8936
1-13	0.8723	0.8723
1-14	0.8936	0.8723
1-15	0.8723	0.8723
1-16	0.9149	0.9149
1-17	0.8511	0.8511
1-18	0.9149	0.9149

TABLE II
MATCHING FROM LOW DIMENSIONS.

LATENT CODE LAYERS USED	Accuracy	
	Cosine similarity	Euclidean distance
18-18	0.1915	0.1915
18-17	0.1702	0.1702
18-16	0.2766	0.2766
18-15	0.3830	0.3830
18-14	0.6383	0.6383
18-13	0.7234	0.7021
18-12	0.7021	0.6809
18-11	0.8298	0.8085
18-10	0.8298	0.8298
18-9	0.8936	0.8936
18-8	0.7872	0.7872
18-7	0.8298	0.8298
18-6	0.8085	0.8085
18-5	0.7660	0.7660
18-4	0.8298	0.7660
18-3	0.8936	0.8723
18-2	0.8511	0.8298
18-1	0.9149	0.9149