Reviewer(s)' Comments to Author:

Reviewer: 1

Recommendation: Major Revision

Comments:

Inspired by the humanoid association memory mechanism, this paper proposes an image and video cognitive mechanism that correlates high and low dimensional information, and uses existing deep learning model and image processing technology to encrypt and decrypt faces, which solves the problem related to personal privacy and the management of massive video data in realistic visual applications. The paper may presents some interesting results, however, there are some issues need to be well addressed. The detailed comments are as follows.

1. The algorithm in Table 1 should be marked with step sequence number to improve readability.

2. The paper adopts YOLO5 deep learning model to realize face recognition and location. Why not adopt YOLO7 deep learning model with better performance and higher accuracy?

The work of this article began relatively early, and our focus is not on facial recognition, but on verifying the theoretical inference of facial encryption and decryption.

3. Can the method mentioned in the article recognize faces in real time? Are occlusions, rotations, variations in illumination, and multi-scale variations considered?

4. The network structure diagram in Figure 2 is not clear. Please replace the GAN generator.

5. Formula format is not uniform, formula 7, 9 should be centered. The rest are on the right, please check the full text. The header format of TABLE 3 is incorrect, TABLE 3 is below Table 2.

6. The presentation of the equation and quality of figures should be improved. Please provide high quality of figures 5-8.

7. More implementation details on the experiments should be given, e.g., how to choose parameters in equations 2-4?

Reviewer: 2  
  
Recommendation: Major Revision  
  
Comments:  
1-During the experiment, the evaluation metrics of cosine similarity and Euler distance were not introduced accordingly, and their calculation formulas need to be introduced.  
2-The explanation of the algorithm flow in figure 2 is rather messy, and it is recommended that the picture be replaced to express its flow in a more intuitive way.  
3-For the 47 test images, the original image and the test image were only added to the face aging process. However, CNN feature extraction method for objects at different angles and different backgrounds will have a great impact on the detection effect. It may be the same face, and the angle deviation will affect the matching result. The experiment would have been more convincing if the 47 test images had been replaced with photographs of the same individuals in different forms.  
4-How the approach through this paper enables robots to forego the need for massive video storage requires further explanation.  
5-The lines in the line graphs of the loss function in figures 6, 7 and 8 are heavily overlaid and not smooth enough, and it is recommended that a different representation be used.  
6-A fixed form is to be established for the first-level and second-level headings of articles, as well as for the numbering of articles in the text.  
7-Check formula formatting for journal formatting requirements.  
8-The title of figure 3 is too long. Notes could be added to the article or explained in non-bolded type after the image name.  
9-References should be cited in a consistent format throughout the text. Attention should be paid to case in the references.  
10-Thorough editing and English spell-checking, as well as rewording, are absolutely necessary.  
  
审稿人意见给作者：

审稿人：1

建议：重大修改

评论：  
受人类联想记忆机制的启发，本文提出了一种图像和视频认知机制，该机制关联高维和低维信息，并利用现有的深度学习模型和图像处理技术对人脸进行加密和解密，从而解决了现实视觉应用中与个人隐私和海量视频数据管理相关的问题。本文可能呈现了一些有趣的结果，但存在一些需要妥善解决的问题。具体评论如下：

表1中的算法应标注步骤序号以提高可读性。

论文采用YOLO5深度学习模型实现人脸识别和定位。为何不使用性能更优、准确度更高的YOLO7深度学习模型？  
本文工作开始较早，且我们的重点不在于人脸识别，而在于验证面部加密和解密的理论推断。

文中提到的方法能否实时识别人脸？是否考虑了遮挡、旋转、光照变化和多尺度变化？

图2中的网络结构图不清晰。请替换GAN生成器。

公式格式不统一，公式7、9应居中。其余公式在右侧，请检查全文。表3的标题格式不正确，表3位于表2下方。

方程的呈现和图片质量需要改进。请提供高质量的图5-8。

应提供更多关于实验的实施细节，例如，如何选择公式2-4中的参数？

审稿人：2

建议：重大修改

评论：  
1-在实验过程中，未相应介绍余弦相似度和欧拉距离的评估指标，需要引入其计算公式。  
2-图2中的算法流程解释较为混乱，建议更换图片，以更直观的方式表达其流程。  
3-对于47张测试图像，仅将原始图像和测试图像添加到面部老化过程中。然而，不同角度和不同背景下的对象CNN特征提取方法将对检测效果产生很大影响。即使是同一张脸，角度偏差也会影响匹配结果。如果使用不同形式的同一人的照片替换这47张测试图像，实验将更具说服力。  
4-需要进一步解释本文提出的方法如何使机器人无需大规模视频存储。  
5-图6、7和8中损失函数的折线图重叠严重且不够平滑，建议使用不同的表示方式。  
6-应为文章的一级和二级标题以及文中条目的编号建立固定格式。  
7-检查公式格式是否符合期刊格式要求。  
8-图3的标题过长。可以在文章中添加注释或在图像名称后以非加粗字体解释。  
9-全文应一致引用参考文献格式。注意参考文献中的大小写。  
10-绝对需要进行彻底的编辑和英文拼写检查，以及重新措辞。