**Structure Idea**

Abstract

1. Introduction
2. Related Works
3. Dataset
4. Proposed Method
5. Experiments
6. Conclusions
7. References

**Working Hypothesis**

Improving automated video editing by dividing each task to a distinct network, such as VSRN(improve resolution), VDN(removing noise), or VCCGAN(color grading), etc. and merging the results.

**Methodology**

1. Creating the dataset
   1. For the scene detection network, collect a large dataset of videos labeled with different scene types, such as indoor, outdoor, day, night, etc.
   2. For the object detection network, collect a large dataset of images labeled with different object types, such as people, cars, animals, etc.
   3. For the sound detection netowrk, collect a large dataset of audios labeled with different actions such as walking, opening a door, etc.
2. Feature Extraction
   1. Extracting the audio from the video to used it for the color grading.
   2. Diving the video in scenes.
3. Network Selection
   1. Video super-resolution (VSRN): If the input video is low resolution, use a VSRN to upscale the video to high resolution.
   2. Video denoising (VDN): If the input video is noisy, use a VDN to remove the noise.
   3. Video color correction (VCCGAN): If the input video has poor color quality, use a VCCGAN to correct the colors.
4. Enhancement – for each scene in the video
   1. Video super-resolution (VSRN): If the input video is low resolution, use a VSRN to upscale the video to high resolution.
   2. Video denoising (VDN): If the input video is noisy, use a VDN to remove the noise.
   3. Video color correction (VCCGAN): If the input video has poor color quality, use a VCCGAN to correct the colors.
5. Output video generation
   1. Combine the enhanced scenes to generate the output video.
6. Evaluation
   1. Evaluate the quality of the output video using a variety of metrics, such as perceptual similarity index (PSI), structural similarity index (SSIM), and video quality metric (VQM).

**Original Approach**

The use of GAN for the integration of audio in video editing based on each objects that appear in a scene.

**Experiments Description**

* Parallelization of the training for each network.
* Evaluate the performance of individual networks.
* Evaluate the performance of the proposed approach.
* Combining the enhanced scenes.

**Possible Original Contribution**

How to optimise real-time video editing? This method can be applied to a range of video applications, including surveillance, providing sharper images for security to monitor, and filmmaking, reducing the time spent on manual video editing.

**References**

[1] Dawit Mureja Argaw, Fabian Caba Heilbron, Joon-Young Lee, Markus Woodson and In So Kweon - The Anatomy of Video Editing: A Dataset and Benchmark Suite for AI-Assisted Video Editing*; Computer Vision – ECCV*, 2022.  
[2] Mark Weber, Jun Xie, Maxwell Collins, Yukun Zhu, Paul Voigtlaender, Hartwih Adam, Bradley Green, Andreas Geiger, Bastian Leibe, Daniel Cremers, Aljosa Osep, Laura Leal-Taixe, Liang-Chieh Chen - STEP: Segmenting and Tracking Every Pixel; *NeurIPS,* 2021.   
[3] Renjie Liao, Xin Taom Ruiyu Li, Ziyang Ma, Jiaya Jia - Video Super-Resolution via Deep Draft-Ensemble Learning;  *IEEE International Conference on Computer Vision,* 2015*.*   
[4] Jianqi Chen, Yilan Zhang, Zhengxia, Keyan Chen, Zhenwei Shi - Dense Pixel-to-Pixel Harmonization via Continuous Image Representation; *ArXiv*, 2023.

[5] Ning Xu, Brian Price, Scott Cohen, Jimei Yang, Thomas Huang - Deep Interactive Object Selection; *IEEE, 2016.*

[6] John L. Gibbs - Video Color Grading Via Deep Neural Networks; *International Journal on Computer Science and Information Systems*, 2018.

[7] Jiandan Zhong, Tao Lei, Guangle Yao - Robust Vehicle Detection in Aerial Images Based on Cascaded Convolutional Neural Networks; *Sensors,* 2017.

[8] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang - Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising; I*EEE Transactions on Image Processing, Volume 26, Issue 7*, 2016.

[9] Antonia Creswell, Tom White, Vincent Dumoulin, Kai Arulkumaran, Biswa Sengupta and Anil A Bharath - Generative Adversarial Networks: An Overview; *IEEE Signal Processing Magazine,* 2017.

[10] Jie Gui, Zhenan Sun, Yonggang Wen, Dacheng Tao, Jieping Ye - A Review on Generative Adversarial Networks: Algorithms, Theory, and Applications; *Journal Of Latex Class Files, Vol. 14, No. 8*, 2015

[11] Kieran G. Larkin – Structural Similarity Index SSIMplified: Is there really a simpler concept at the heart of image quality measurement?; 2015.

[12] Alice Lucas, Santiago Lopez-Tapia, Rafael Molina and Aggelos K. Katsaggelos - Generative Adversarial Networks and Perceptual

Losses for Video Super-Resolution; *IEEE International Conference on Image Processing,* 2018.

[13] Ming-Jun Chen, Alan C. Bovik - Fast structural similarity index algorithm; *Springer-Verlag,* 2010.

[14] Panagiotis Kouzouglidis, Giorgos Sfikas, Christophoros Nikou – Automatic Video Colorization Using 3D Conditional Generative Adversarial Networks; *Bebis, G., et al. Advances in Visual Computing. ISVC*, 2019.

[15] Hongying Liu, Zhubo Ruan, Peng Zhao, Chao Dong, Fanhua Shang, Yuanyuan Liu, Linlin Yang, Radu Timofte - Video Super-Resolution Based on Deep Learning: A Comprehensive Survey; *Artif Intell Rev 55, 5981–6035*, 2022.