Machine Learning for the Arts
UCSD FALL 2019
FINAL PROJECT

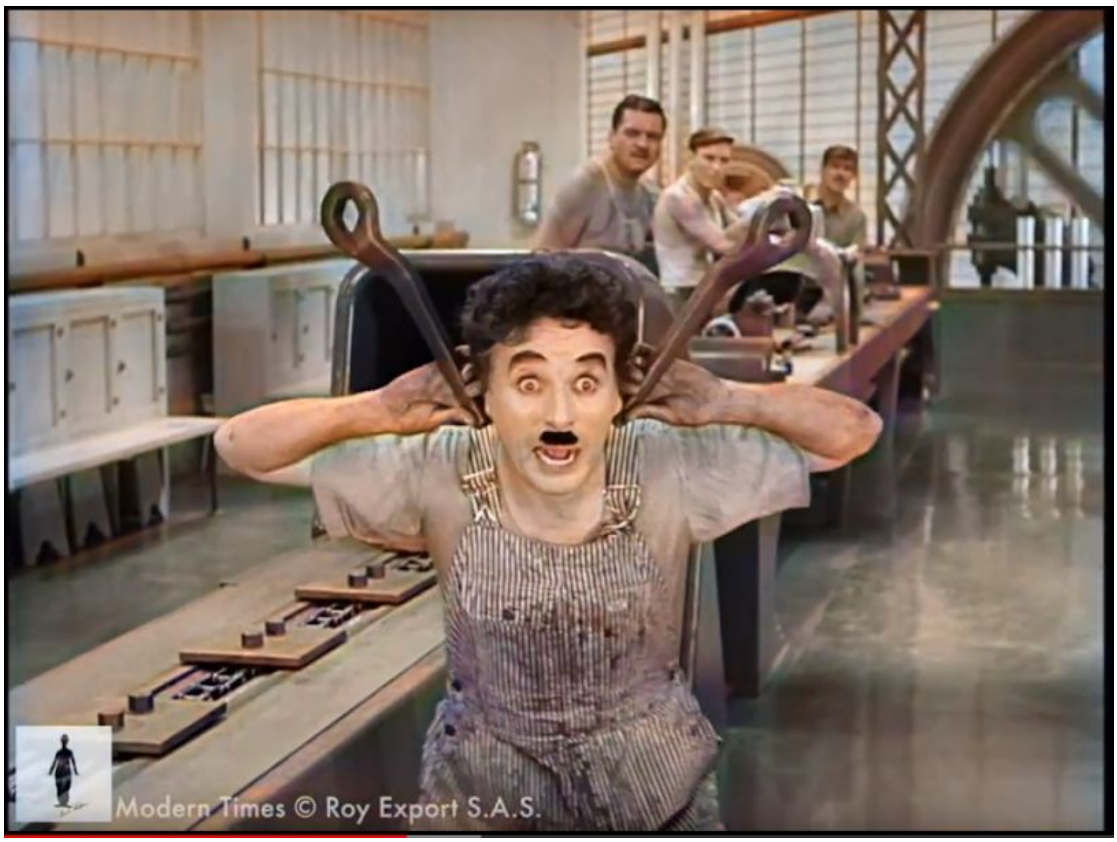
Chaplin Rejuvenated

Automatic Colorization and Tone Alteration of the Chaplin FIlms

Qiyuan Wu
Jiaye Wang
Yifan Hou


Press Esc to exit full screen


Modern Times © Roy Export S.A.S.


Modern Times © Roy Export S.A.S.

DESCRIPTION

For this project, we plan to convert a black and white classic clip from Modern Times by Charlie Chaplin into a colored video using the Deep Coloring neural network. We will also convert the background music and effects into one with modern tune using MIDI analysis. We will revisit the neural network as well as style transfer knowledge we learned during the latter part of the class. The creative goal is to rejuvenate the classic masterpieces by colorizing the clip intelligently and also innovating the background theme music to give it a more modern perception. This is also a very good way to recommend Chaplin's artwork to the audience who are not interested in watching black and white films, since they now have vivid colors and are even more attractive. We will present the work on Wednesday by going over the background of the movie, and then show the audience the original clip as well as the colorized clip. We will also walk through the procedures and rationale of the neural network model we have used.

Concept

The inspiration of re-newing Chaplin's masterpiece is that its essence is still so relatable and relevant even after 83 years! The mundane, repeated and redundant work in pipeline resembles exactly the routine of most office and service workers of the current societies. Like the preface of the movie says, this is a story of industry, of individual enterprise - humanity crusading in the pursuit of happiness. This universal and eternal theme is what makes this movie an everlasting classic. Modern Times' factory scene extremely relatable to me, because sometimes I have obsessive-compulsive disorder symptoms as well: sometimes I simply could not get over a thing just like the protagonist got stuck in the action of turning screws, which is why this scene struck me so deeply.

Technique

We used NoGAN to automatically color the videos. This is a new type of GAN training that has been developed to solve some key issues in the previous DeOldify model. It provides the benefits of GAN training while spending minimal time processing direct GAN training. Instead, most of the training time is spent pretraining the generator and critic separately with more straight-forward, fast and reliable conventional methods. A key inspiration here is that those more "conventional" methods generally get us most of the results we need, and that GANs can be used to close the gap on reality. During the very short amount of actual GAN training the generator not only gets the full realistic colorization capabilities that used to take days of progressively resized GAN training, but it also doesn't accrue nearly as much of the artifacts and other ugly baggage of GANs. In fact, we can pretty much eliminate glitches and artifacts almost entirely depending on our approach. This is a new and incredibly effective technique.

The video is supposed to be stable because NoGAN training is crucial to getting the kind of stable and colorful images seen in this iteration of DeOldify. NoGAN training combines the benefits of GAN training (wonderful colorization) while eliminating the nasty side effects (like flickering objects in video). In fact, videos are rendered using isolated image generation without

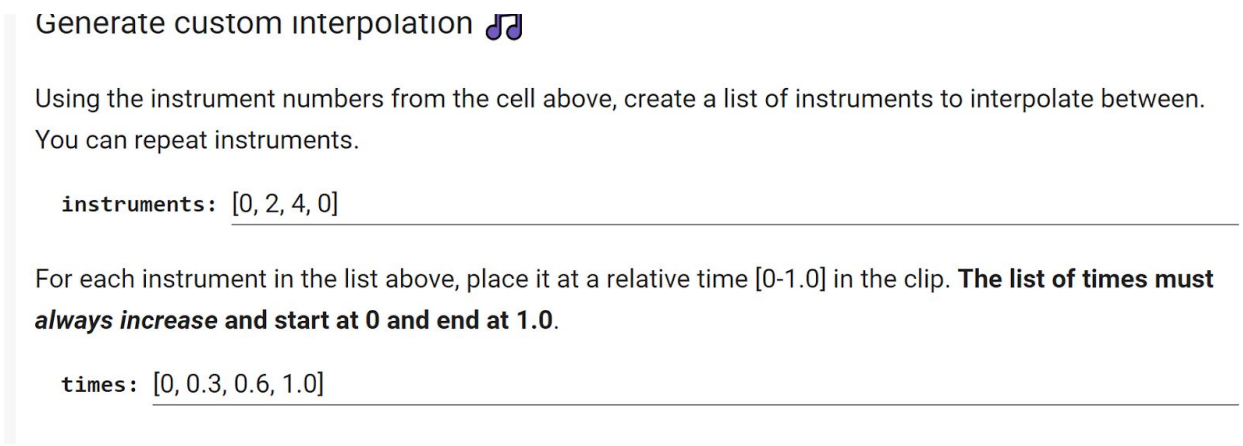
any sort of temporal modeling tacked on. The process performs 30-60 minutes of the GAN portion of "NoGAN" training, using 1% to 3% of imagenet data once. Then, as with still image colorization, we "DeOldify" individual frames before rebuilding the video.

Other than improved video stability, there is an intriguing matter going on here worth mentioning. It turns out the models we run, even different ones and with different training parameters and structures, keep arriving at more or less the same solution. That's even the case for the colorization of things you may think would be arbitrary and unknowable due to the lack of information, like the color of clothing, cars, and even special effects (as seen in "Metropolis").

For the audio part, we use online mp3 to midi transfer software to extract the midi file we want and pass the new midi file in to Gansynth. Then combine the new audio file with the colored video.

The midi file will record the music tone and duration as well as the strength, loudness of the music. By using different software, we can put those notes into different timbre. However, for most software, they can only use a certain timbre at one time.

In this project, I choose Gansynth to play the midi file with combination of different music instrument timbre. In Gansynth's colab section we can play with different parameters to manipulate the audio output.



Process

For the colorized part, we try different videos to compare the results. We find that when the video contains more objects and many of them are overlapped, the model will give these objects almost the same colors. But for a scene only contains limited objects and they are easy for the model to detect, the result is pretty good. There is another important parameter called render factor influences the video vivid. For getting a vivid colorized video, we try to set render-factor from 10 to 44. After comparing the results using different render factor, we find 21 is the best render factor.

We could see how render factor influences the vedio

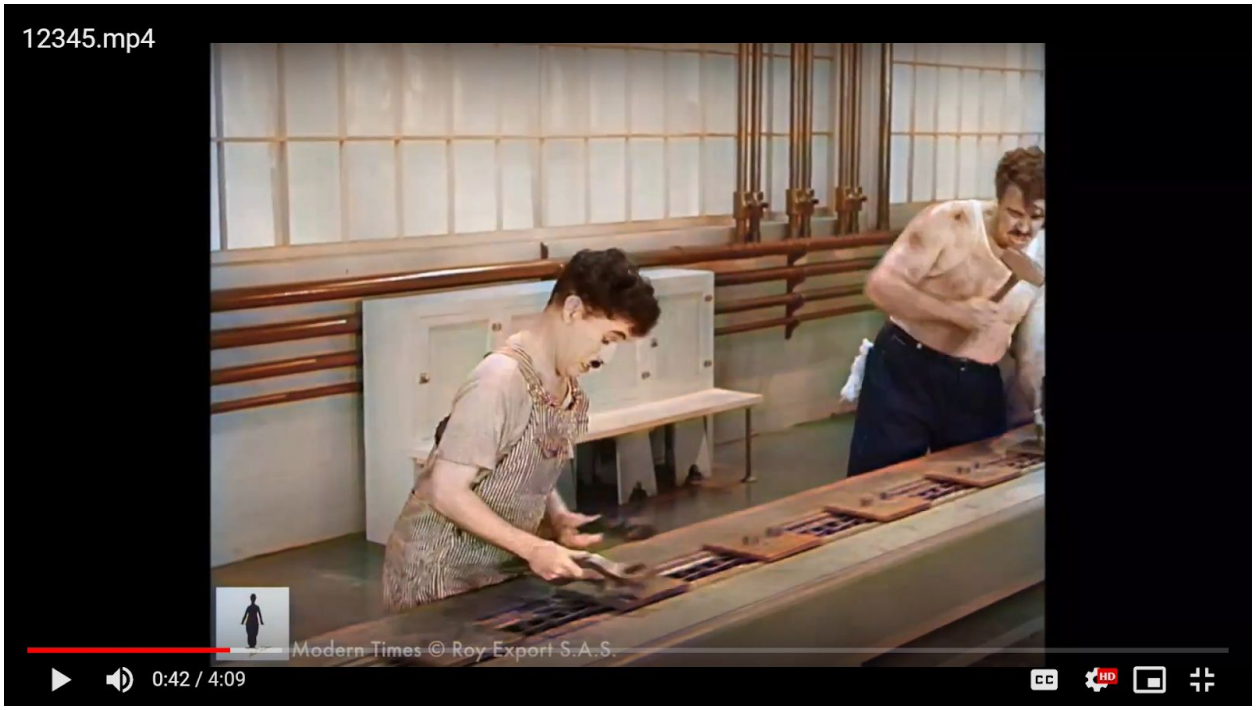
```
for i in range(10,45,2):
    colorizer.vis.plot_transformed_image('video/bwframes/video/00001.jpg', render_factor=i, display_render
_factor=True, figsize=(8,8))
```

To transfer the timbre of the audio from the video, we need to first extract the audio from the video. However, the extracted audio is in the form of mp3, which is a wave file. The Gansynth only transfer the midi audio file. In order to put the audio into gansynth, we need to first transfer the mp3 to midi file. But midi file only consist music instrument notes. In the movie, there are many lyrics included which cannot be transferred into midi file. And the result audio file will have different phase compared to the original audio file. We cannot solve this problem because of the content of midi file can only be music note. We cannot use music note to substitute human voice. Finally, instead of choosing movies with human audio, we choose the one that only contain music. After we transfer the mp3 file into midi file and output the audio file with the same phase with the original audio file. I put the new midi file into Gansynth. The Gansynth give me a newly formed audio file with different timbre inside it. Then We use Adobe Premiere pro to combine the colored video with the new audio file.

Result

We are very satisfied with the results of this project, especially the shaving clip, because it even colorizes the painting at the background vividly and convincingly, even the painting does not even have too many details. Also for this shaving clip, the colorization is quite stable, even when the characters are moving rapidly. For the Modern Times factory scene, we are satisfied that it preserves and reflects the dim and melancholy theme color of the movie, which gives the clip a repressive and stressful perception. Also the colorization of the dirty hands and clothes are quite satisfactory (as shown in the above picture). However, for this clip, the colorization is not very stable, so when the characters are moving, the colorization would simply disappear for a moment and then reappear. This instability is a side-effect of the GAN system, and the fact that we did not use tack on temporal modeling. Also for this factory scene, the colorization does not contain too many colors, and only mostly grey and brown. Although it might be the true colors of the factories, the colorization seems to be not extremely effective (not enough contrast in generated colors).

This is a good example of colorized movie clip.



Due to monotonous of industry scene, objects in colorized video have almost the same colors.

For the music part, we change the original music and give it a new tune. Since music is an important part of music, what we do could change the old movie style and renew them. Using some modern music style, audience will more likely to appreciate these old classic movies.

Reflection

As for the future work, we hope we could improve the stability when characters are moving. This is the hard part for this project. Since many colorized models turn grayscale video into colorful video, we hope we could find the difference between using black and white video and grayscale movie. In this project, we only change the tone of the music. If we have enough time, we want to change the tone of actor audios. Compared to changing music tones, changing audio might be more powerful to transfer movie style.

REFERENCE:

GANSynth: https://magenta.tensorflow.org/gansynth
Colab:
https://colab.research.google.com/notebooks/magenta/gansynth/gansynth_demo.ipynb

CODE: https://github.com/ucsd-ml-arts/ml-art-final-qiyuan_yifan_jiaye

RESULT:

All of our colorized videos merged with the new generated theme music are in this Google folder.
https://drive.google.com/drive/u/1/folders/0AM7acZ7fifDSUk9PVA