

|              |                                                                                   |
|--------------|-----------------------------------------------------------------------------------|
| Title        | 敵対的生成ネットワークによる講義アーカイブシステムの板書鮮明化                                                   |
| Author(s)    | 崔, 博文                                                                             |
| Citation     |                                                                                   |
| Issue Date   | 2021-03                                                                           |
| Type         | Thesis or Dissertation                                                            |
| Text version | author                                                                            |
| URL          | <a href="http://hdl.handle.net/10119/17153">http://hdl.handle.net/10119/17153</a> |
| Rights       |                                                                                   |
| Description  | Supervisor: 長谷川 忍, 先端科学技術研究科, 修士 (情報科学)                                           |

# White Board Sharpening for Lecture Archive System by Generative Adversarial Networks

1810074 CUI BOWEN

In recent years, with the process of video recording, speeding up the Internet, and the capacity of data storage devices increasing, we have entered an era in which the captured video can be stored in the cloud and rebroadcast as well as live on the Internet. In 2017, the domestic Internet usage rate reached the 80% level at 80.9%, and by terminal, the usage rate on smartphones was 59.7%, which exceeded personal computers (52.5%). In addition, a network that can distribute and receive high-quality video content has been established, and an infrastructure that allows users to enjoy various video distribution services has been established. Furthermore, "hobbies / practical use / education" (30%) was ranked after "music" (45%) as a genre to be viewed on the video distribution service. From these facts, it can be said that video distribution using networks has become familiar to us.

Since the 1990s, e-learning, which means computer-based learning and education, has attracted attention. With the spread of the Internet, e-learning has become easier to realize. Also its educational value is drawing attention. Such e-learning system teaching materials are in a multimedia form that combines video, audio, board writing, and power-point. The users are supposed to be "learners" and "teachers". However, the traditional relationship between learners and teachers has changed. Learners can study in their free time and at any place. However, it is difficult to communicate in Two-way, such as ask questions. Teachers can work efficiently, but it is difficult to grasp the situation of learners. The lecture archive system is one of the e-learning systems. It is also an approach that records the video and audio of lectures held in the lecture room as digital data and used as the content of the asynchronous remote learning system.

In this research, we focused on the problem that the whiteboard board is difficult to read in the lecture archive system introduced in higher education institutions. The hardware solution proposed in the previous research takes time and cost to introduce new equipment to all rooms. It is not realistic because when delivering high-definition video such as 4K, the network bandwidth becomes very large. Also, the existed lecture video in the system cannot be clarified. JAIST has introduced a lecture archive system since 2006, and there are already many recorded videos, the sharpening method on software is more meaningful.

In addition, there are several super-resolution processing software methods that are applicable to lecture archive systems. In recent years, deep learning techniques such as SRCNN and SRGAN have become popular and have shown sufficient effects as a general image super-resolution processing method. However, the same effect cannot always be obtained, even in writing on a whiteboard. Furthermore, the characteristic of the lecture is that the changes in the writing on the board are not drastic. Compared to movies and animations, the written characters remain for a long time. Therefore, as the target of super-resolution processing, we targeted screenshots of the board writing area extracted from the videos. In this research, we expect when the SRGAN processing is applied with the expectation, a sharper image could be generated, and a high sharpening effect would be

obtained.

In this research, when applying SRGAN to the lecture archive system, some improvement measures were applied. As a model, the Subpixel layer, Convolution layer, ReLU layer were deleted for 2x upscaling factors. Sigmoid layer was deleted as the cause of the disappearance of the learning gradient to perform super-resolution processing. Furthermore, to improve the accuracy, a Skip connection was added to Discriminator. In performing general preprocessing, we also devised the method improvement which suitable for the lecture archive system. We realized in previous studies, the characters on the board were often binarized. By performing binarization processing, it is possible to increase the characteristics of the surplus on the board easily. However, in the lecture archive system, the character color of the board may be important. For example, by changing the character color, the part that the teacher wants to pay attention to and the part that the teacher does not want to pay attention to is often expressed. To retain such color information, we decided to clarify the color characteristics.

The contrast between the characters and the background is quite low in the hard-to-read board. Therefore, we thought about identifying the characteristics of the board writing and treating the board writing and the margins separately, but it is very difficult to separate them. Furthermore, the actual board writing environment also differs each time due to various factors such as the color of sunlight, the brightness of the classroom, and the color of letters. In other words, such a pretreatment method is by no means highly versatile, and if the shooting environment and the color of the pen used for writing on the board are not restricted. It is necessary to consider other pretreatment methods.

So, we used Contrast Limited Adaptive Histogram Equalization and found effective parameters to writing data. The noise was slightly higher with such adjustments than that of the SRGAN method, but the contrast and color depth could be significantly increased. In the evaluation experiment, the evaluation of readability by the subject exceeded the methods such as SRCNN and BICUBIC, and the best result showed the validity of the proposed improvement measure.

However, GAN-based super-resolution processing may generate new undesired features, although it can produce sharp images. This hinders the reading of small text details. Comparing the proposed method with SRCNN, the CNN-based super-resolution processing method has higher reproducibility, and the detailed features can be almost retained. However, from the detailed objective evaluation of this study, the correct answer rate in all the super-resolution methods is less than 40%. So, it is necessary to consider a better super-resolution processing method for reading characters with a small number of features.

At the end of this research, we have some findings about the future tasks. First, one of them is to improve the recall rate of high-definition board writing with SRGAN. In this proposed method, the SRGAN method process the writing on the board in an easy-to-read manner, but for details, we think that the high reproducibility point of SRCNN can be referred to. In the evaluation experiment, it was confirmed that the CLAHE pretreatment had a good effect. However, it will deteriorate the image quality, generate noise, and was not suitable for subjects who were concerned about noise. Therefore, the second direction is to consider a method that sets noise elimination as a set. Finally, VGG and MSE loss functions are used in SRGAN. The loss generated by comparing the features of the generated image and the original high-resolution image was affected by the margins of the whiteboard area, and it will cause a problem that the learning depth

cannot be deepened. It is considered necessary to set an innovation loss function that suitable for the writing on the whiteboard.

Keywords: lecture archive system, whiteboard, super-resolution, GAN, CLAHE.