Deteriorated Characters Restoration for Early Japanese Books Using Enhanced CycleGAN

Hayata Kaneko [1], Ryuto Ishibashi [1] and Lin Meng [2],*

1 Graduate School of Science and Engineering, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu 525-8577, Shiga, Japan; ri0099xx@ed.ritsumei.ac.jp (H.K.); ri0097fx@ed.ritsumei.ac.jp (R.I.)
2 College of Science and Engineering, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu 525-8577, Shiga, Japan
* Correspondence: menglin@fc.ritsumei.ac.jp

Abstract: Early Japanese books, classical humanities resources in Japan, have great historical and cultural value. However, Kuzushi-ji, the old character in early Japanese books, is scratched, faded ink, and lost due to weathering and deterioration over the years. The restoration of deteriorated early Japanese books has tremendous significance in cultural revitalization. In this paper, we introduce augmented identity loss and propose enhanced CycleGAN for deteriorated character restoration, which combines domain discriminators and augmented identity loss. This enhanced CycleGAN makes it possible to restore multiple levels of deterioration in the early Japanese books. It obtains the high readability of the actual deteriorated characters, which is proved by higher structural similarity(SSIM) and accuracy of deep learning models than standard CycleGAN and traditional image processing. In particular, SSIM increases by 8.72%, and the accuracy of ResNet50 for damaged characters improves by 1.1% compared with the competitive CycleGAN. Moreover, we realize the automatic restoration of pages of early Japanese books written about 300 years ago.

Keywords: deteriorated characters restoration; early Japanese books; enhanced CycleGAN; cultural heritage presevertion

1. Introduction

Early Japanese books are a cultural heritage, have stored the wisdom of our ancestors, and contain much information about Japanese politics, economy, culture, etc. These early Japanese books are described by Kuzushi-ji, a kind of old Japanese character style, and are also one of the essential factors to symbolize early Japanese books. However, Kuzushi-ji is not used in the present day, causing only few experts of classical Japanese can read Kuzushi-ji and understand the contents of the books. To re-organize and preserve this cultural heritage, researchers have digitalized the early Japanese books and applied the combining Kuzushi-ji and Optical Character Recognition(OCR) to recognize the Kuzushi-ji [1–3]. Humanities research intuition such as the Center for Open Data(CODH) [4] and Art Research Center(ARC) of Ritsumeikan [5] digitize the early Japanese books and re-organize them for the database to prevent degradation and prompt combine computer science with humanities. Currently, lots of researchers apply deep learning and machine learning methods for cultural heritage protection, organization, etc. Literature organization is one of the hot topics in this area, such as the re-organization of OBI [6,7], Kuzushi-ji [1,2,8], and Rubbing [9]. In detail, Yue et al. aim to achieve a good accuracy recognition for Oracle Bone Inscription, which is ancient characters described on the tortoises' shells and animals' bones. Zhang et al. combine simple deep learning models and lexical analysis to recognize rubbing characters described on the bones from 3000 to 100 years ago. Lyu et al. use deep learning and image processing method for detecting and recognizing Kuzushi-ji. It is also the target literature of this paper.

However, the early Japanese books have been damaged, and some Kuzushi-ji is scratched and faded due to deterioration over the years. The damage increases the difficulty of Kuzushi-ji recognition and prevents early Japanese book re-organization accuracy. In the worst case, it causes the loss of some cultural heritage. Hence, the damaged Kuzushi-ji restoration becomes an emergency research topic for preserving cultural heritage.

Nowadays, we can realize high-quality and objective restoration thanks to developing generative models using deep learning. Generative Adversarial Nets(GANs) [10] are generative models that balance sampling speed and quality well. Image-to-Image translation based on GANs is thereby effective, and it has been studied in many different cases. GAN-based ancient characters restoration is focused on supervised image-to-image translation for inpainting damaged characters [11,12]. Su et al. inpainted masked characters in the book of Qing dynasty and Yi, handwritten ancient Chinese characters, with unmasked images and applied them to the practical ancient text. Wenjun et al. inpainted large-area damaged characters in the ancient Yongle Encyclopedia with ground truth and examples of them. To our knowledge, this is the first time the practical restoration of Kuzushi-ji and the early Japanese books using deep learning. When applying supervised image-to-image translation for Kuzushi-ji, the various types of damaged and distorted letter styles make it hard to prepare train data. Moreover, most damaged characters are weak damages, which makes GANs training difficult.

The problems are shown in Figure 1a, a piece of Early Japanese Books, which has several levels of deterioration such as faded ink, global, partial, and scratches, are expanded listed in Figure 1c.
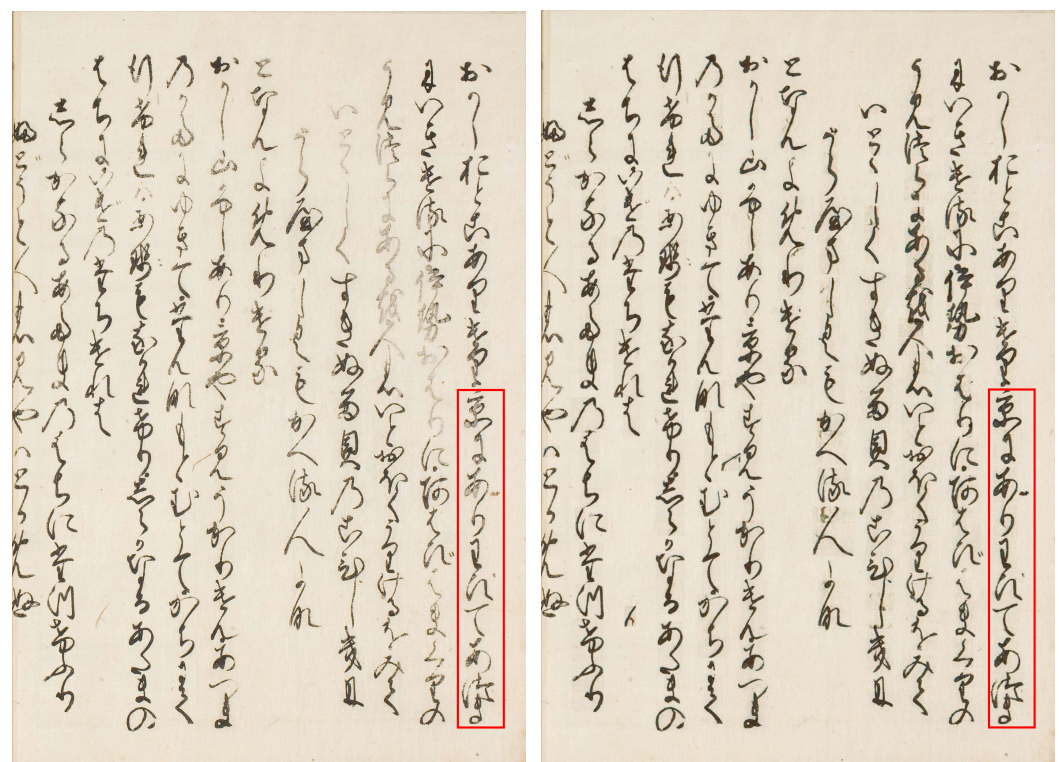
(**a**) A deteriorated page   (**b**) Restoration goals

Figure 1. *Cont.*

(**c**) Expansion of a deteriorated page   (**d**) Expansion of restoration goals
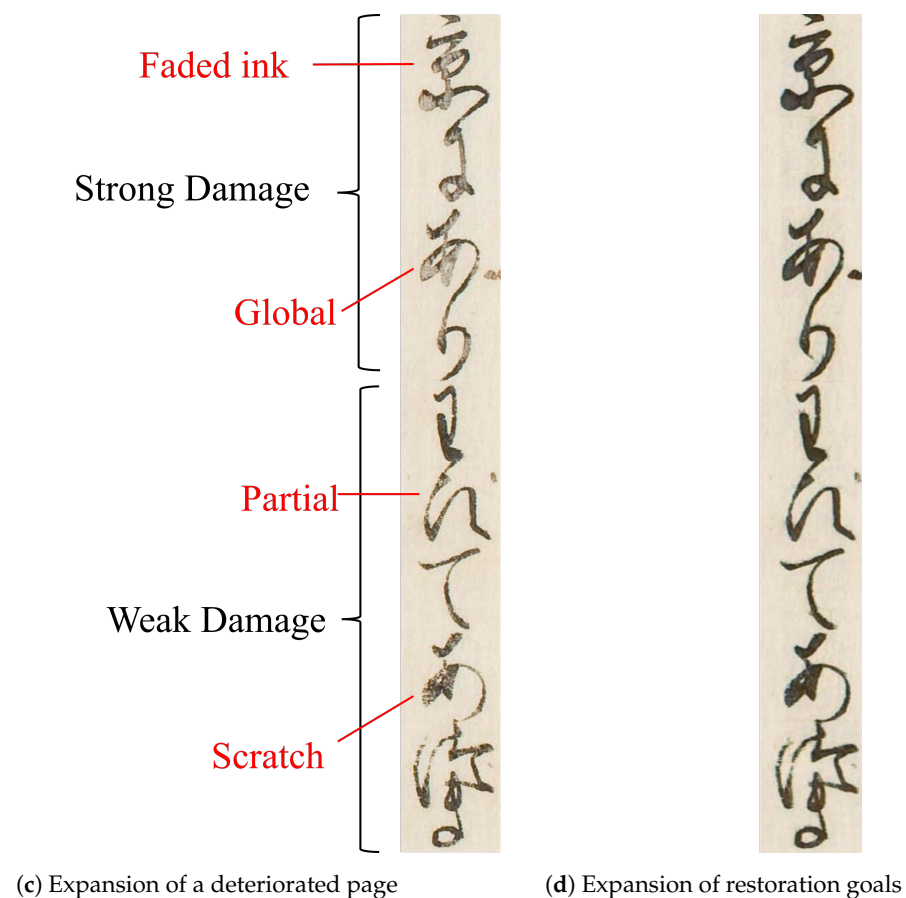
Figure 1. Restoration result of a page of the book by enhanced CycleGAN. The red square of (**a**) marks a part of deteriorated characters, and the red square of (**b**) marks restoration results.

To solve these problems, we propose an enhanced CycleGAN and aim to realize the restoration of deteriorated early Japanese books. An example of restoration through our proposal is shown in Figure 1b,d. In detail, we employ CycleGAN [13], unsupervised[1] image-to-image translation based on GANs, for the training with only authentic damaged images and undamaged images. Then, we propose a robust CycleGAN for weak damages, which combines a domain discriminator [14] and augmented identity loss. The augmented identity loss improves the identity mapping loss for CycleGAN. The enhanced CycleGAN provides truth restoration and high readability for early Japanese books. The contribution detail of this paper is:

- We employ a domain discriminator for the restoration of deteriorated characters. It leads to improving the degree of restoration.
- We propose the augmented identity loss, which lets the generator learn the identity mapping by a large number of images and generalized images. The enhanced CycleGAN, combining the domain discriminator and augmented identity loss, realizes high-quality damage restoration. It is proved by the quantitative results using PSNR and SSIM.
- We indicate that image enhancement using image processing and CycleGAN increases the accuracy of Kuzushi-ji recognition. The proposed method is more effective than traditional image processing for the Kuzushi-ji expressed as RGB.

2. Related Works

2.1. Image Processing for Ancient Characters

The automatic analysis of ancient documents is dramatically improving due to deep learning development. Nevertheless, image pre-processing is vital because ancient docu-

ments suffered from deterioration. Binarization is the general and effective pre-processing method for denoising, sharpening, and background separating [8,15,16]. Simultaneously, it provides readable documents by thresholding faded ink, thin, weak text, et. [17]. The Otsu method [6,18], an automatic threshold-determining algorithm, is frequently used because it handles efficiently for multiple ancient documents and characters. However, binarization causes the loss of character parts since it does thresholding. In this paper, we indicate that the enhanced CycleGAN has the advantage of more readable restoration and higher recognition accuracy for ancient characters than binarization.

2.2. Image-to-Image Translation Based on GANs

Generative Adversarial Nets(GANs) [10] consist of two networks, a generator, and a discriminator. The generator generates new images from random noise. The discriminator train to distinguish between real data and generated data. Conversely, the generator learns to have the discriminator in-distinguish between real data and generated images. As a result, the generator generates similar to real data. This development also contributes significantly to the evolution of image-to-image translation methods.

Isola et al. [19] propose the Pix2Pix, a supervised image-to-image translation based on conditional GAN, which enables high-quality image transformation. Pix2Pix uses paired images to train conditioning and its reconstruction. On the other hand, Zhu et al. [13] propose an unsupervised image-to-image translation with a cycle consistency loss, namely CycleGAN. The cycle consistency loss is a restriction that the original image via forward-backward translation returns to the original image. This is why CycleGAN allows capturing the input feature even though using unpaired images. CycleGAN is flexible and used in a variety of scenes such as super resolution [14], denoising [20], and repairing damaged images [21]. These CycleGAN-based applications for improving the quality of images contribute to accurate classification [22] or data augmentation [23].

2.3. Character Restoration Based on Image-to-Image Translation

Liu et al. [24] proposed characters inpainting using supervised image-to-image translation with feature extraction of a deep learning recognition model. Su et al. [11] restored the loss of ancient characters by two networks based on conditional GAN, the Shape restoration network and the Texture restoration network. Wenjun et al. [12] proposed GAN-based large damages inpainting for ancient characters by adding the attention module and the feature map of the example text to GANs. For early Japanese books, the actual deteriorated characters are faded ink and scratching, many are weak, global, or local as shown in Figure 1. To reproduce these paired images without omission is difficult and inefficient. Thus, we adopt CycleGAN to train with actual images, then propose a restoration model for the early Japanese books.

3. Approach

We propose CycleGAN with a domain discriminator and augmented identity loss for early Japanese book restoration. Our method is a practical approach for applying existent data from early Japanese books. However, there is partial and weak deterioration rather than a highly noticeable deterioration. The standard CycleGAN has a limitation for weak damage restoration. To solve it, we adopt the domain discriminator to increase the degrees of transformation by penalizing non-transform clearly in Figure 2. Then, the augmented identity loss improves the preservation of character shapes because it is equal to the identity mapping loss with enough and sufficient data. This proposed loss function combined with the domain discriminator is our proposed enhanced CycleGAN for character restoration and its structure shown in Figure 3. The two additional functions make higher quality restoration.
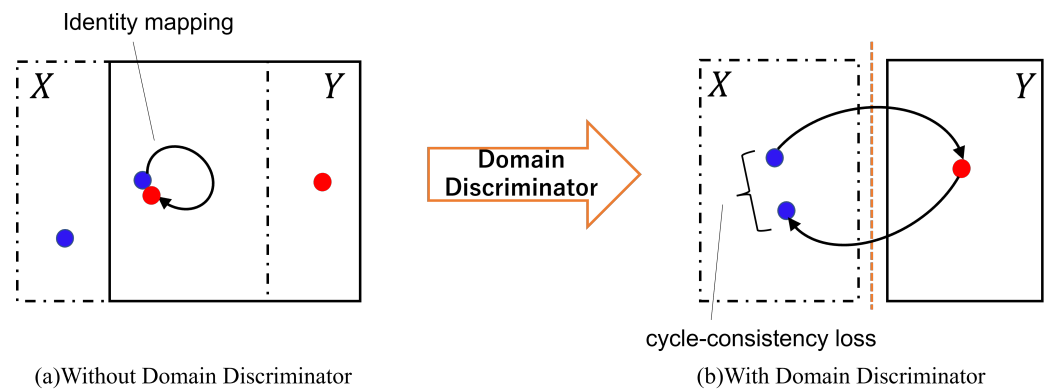
(a)Without Domain Discriminator (b)With Domain Discriminator

Figure 2. An overview of Domain Discriminator's work. When domain X is similar to domain Y, the discriminator cannot distinguish X and Y, so generators learn identity mapping (**a**) . The domain discriminator clearly classifies X and Y to help domain transformation (**b**). The blue bullets show images in domain X, and the red bullets show images in domain Y.
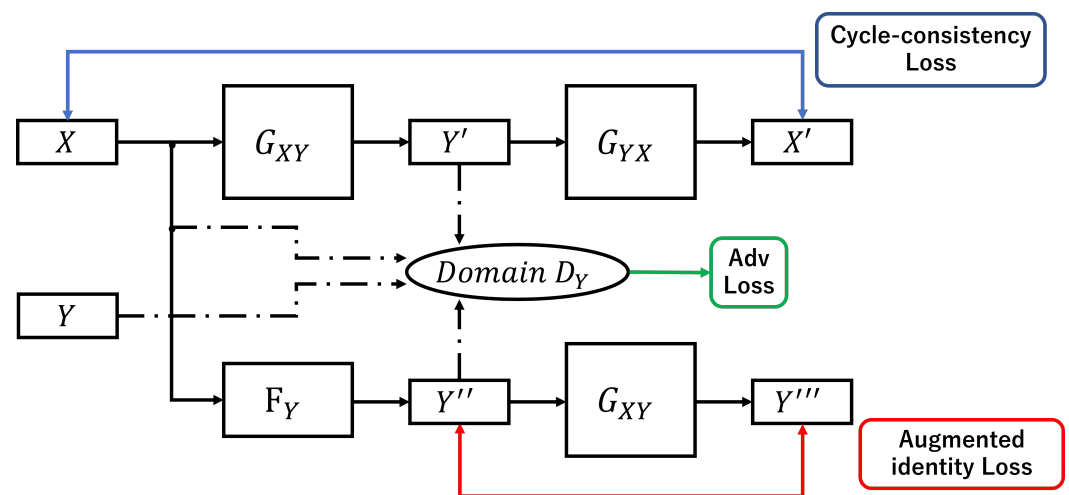

Figure 3. The overviews of training of $X \to Y$ by the proposed implementation. X is the inputs of D_Y in addition to the stadnadard inputs Y and Y' because D_Y is the domain-D. Moreover, Y'', the output of F_Y is also the input of D_Y to use aid loss. The output of F_Y is used for identity mapping loss of G_{XY} as the aid loss.

3.1. CycleGAN

We apply CycleGAN to restore damaged characters. CycleGAN is image transformation in two domains. In this paper, we describe domain X as the damaged domain and Y as the undamaged domain, and $X \to Y$ as the damage restoration. CycleGAN includes two discriminators D_X and D_Y and two generators $G_{XY} : X \to Y$ and $G_{YX} : Y \to X$. The adversarial loss with a least square error [13,25] between D_Y and G_{XY} is as follow:

$$\mathcal{L}_{adv}(G_{XY}, D_Y, X, Y) = \mathbb{E}_{y \sim p_{data}(y)} \left[(D_Y(y) - 1)^2 \right] + \mathbb{E}_{x \sim p_{data}(x)} \left[D_Y(G_{XY}(x))^2 \right] \quad (1)$$

$\mathbb{E}_{y \sim p_{data}(y)}[\cdot]$ denotes an expectation for data distribution $p_{data}(y)$. The object function of Equation (1) is:

$$min\mathcal{L}_D(D) = \mathbb{E}_{y \sim p_{data}(y)} \left[(D_Y(y) - 1)^2 \right] + \mathbb{E}_{x \sim p_{data}(x)} \left[D_Y(G_{XY}(x))^2 \right]$$
$$min\mathcal{L}_G(G) = \mathbb{E}_{x \sim p_{data}(x)} \left[D_Y(G_{XY}(x) - 1)^2 \right], \quad (2)$$

Cycle-consistency loss enforce $G_{YX}(G_{XY}(X)) \approx X$ and $G_{XY}(G_{YX}(Y)) \approx Y$ to correlate inputs and outputs. The cycle-consistency loss is defined as:

$$\mathcal{L}_{cyc}(G_{XY}, G_{YX}) = \mathbb{E}_{x \sim p_{data}(x)}[\|G_{YX}(G_{XY}(x)) - x\|_1] + \mathbb{E}_{y \sim p_{data}(y)}[\|G_{XY}(G_{YX}(y)) - y\|_1], \tag{3}$$

In addition, the identity mapping loss is employed in CycleGAN as a help objective function. Identity loss helps to preserve features between the input and output, such as the color composition [13,26]. The identity loss is:

$$\mathcal{L}_{id}(G_{XY}, G_{YX}) = \mathbb{E}_{y \sim p_{data}(y)}[\|G_{XY}(y) - y\|_1] + \mathbb{E}_{x \sim p_{data}(x)}[\|G_{YX}(x) - x\|_1] \tag{4}$$

The full objective included the adversarial loss, the cycle-consistency loss, and the identity loss is :

$$\begin{aligned}\mathcal{L}(G_{XY}, G_{YX}, D_X, D_Y) &= \mathcal{L}_{adv}(G_{XY}, D_Y) + \mathcal{L}_{adv}(G_{YX}, D_X) \\ &+ \mathcal{L}_{cyc}(G_{XY}, G_{YX}) + \mathcal{L}_{id}(G_{XY}, G_{YX})\end{aligned} \tag{5}$$

3.2. Domain Discriminator

Kim et al. [14] introduce a domain discriminator (domain-d) to guide generated images following target domain distribution rather than the source domain. The domain-d learns the source domain as generated images. The reason why the standard CycleGAN occurs identity transformation is the images included with X and Y are satisfied with the loss functions of that CycleGAN as shown in the left of Figure 2. The high similarity between weak damaged and undamaged characters applies to them. The identity transformation of GXY and GYX is described as $G_{XY}(x) = x$ or $G_{YX}(y) = y$ mathematically. When $D_Y(x) = 1$, $D_Y(G_{XY}(x) = x)$ is not penalized by the adversarial loss, and $G_{XY}(x)$ does not transform x. Therefore, we employ domain discrimination which distinguishes $DY(x) = 0$ for the penalty for $D_Y(G_{XY}(x) = x)$.

The adversarial loss attached the domain-d with a least square error using D_Y and $G_{XY} : X \rightarrow Y$ is:

$$\begin{aligned}\mathcal{L}_{adv}(G_{XY}, D_Y, X, Y) &= \mathbb{E}_{y \sim p_{data}(y)}\left[(D_Y(y) - 1)^2\right] + \mathbb{E}_{x \sim p_{data}(x)}\left[D_Y(G_{XY}(x))^2\right] \\ &+ \mathbb{E}_{x \sim p_{data}(x)}\left[D_Y(x)^2\right],\end{aligned} \tag{6}$$

In this study, we employ the domain-d for both D_X and D_Y because this implementation improves the quality of generated images experimentally.

3.3. Augmented Identity Loss

In this section, we pay attention to the identity mapping loss(id loss) in CycleGAN. The id loss helps to keep the input's color composition and shapes [13,26]. For this reason, it is suitable to preserve an author's font styles and old book texture in the early Japanese books. According to Xu et al. [27], however, the id loss can constrain the degree of transformation. In addition, decreasing of id loss is less than other loss functions of CycleGAN, such as cycle-consistency loss and adversarial loss, when increasing the train data. We present the problem of id loss as follows:

- A larger amount of train data is necessary to decrease the id-loss sufficiently.
- When an image is appropriate, the identity mapping loss enforces the generator to learn incorrect identity transformation. In particular, the identity transformation learning of overlapping-domain images conflicts with image transformation to the other domain.

Therefore, We propose augmented identity loss(aid loss) as:

$$
\begin{aligned}
\mathcal{L}_{aid}(G_{XY}, G_{YX}) = \; & \mathbb{E}_{y \sim p_{data}(y)}[\gamma \|G_{XY}(F_Y(x)) - F_Y(x)\|_1 \\
& + (1 - \gamma)\|G_{XY}(y) - y\|_1] \\
& + \mathbb{E}_{x \sim p_{data}(x)}[\gamma \|G_{YX}(F_X(y)) - F_X(y)\|_1 \\
& + (1 - \gamma)\|G_{YX}(x) - x\|_1]
\end{aligned}
\tag{7}
$$

where $\gamma = epoch/number\,of\,epochs$. For the dynamic training of F_X and F_Y, generated images for the first epochs are not high-quality. For this reason, we decide γ increases linearly with epoch following with [28]. F_X and F_Y are new generators of GANs. F_X close input to the distribution of domain-X, and F_Y close input to the distribution of domain-Y. The generated images are abundant and gradually change every epoch because F_X and F_Y are trained simultaneously as the CycleGAN training.

Then, we use the two domain discriminators, domain-D_X and domain-D_Y, to learn the adversarial loss of F_X and F_Y. The domain discriminators guarantee that outputs of F_X and F_Y are independent domain images.

Thereby, the object function of domain-d combined with aid loss is as below:

$$
\begin{aligned}
\mathcal{L}_{adv}(G_{XY}, D_Y, F_Y, X, Y) = \; & \mathbb{E}_{y \sim p_{data}(y)}\left[(D_Y(y) - 1)^2\right] + \mathbb{E}_{x \sim p_{data}(x)}\left[D_Y(G_{XY}(x))^2\right] \\
& + \mathbb{E}_{x \sim p_{data}(x)}\left[D_Y(x)^2\right] + \mathbb{E}_{x \sim p_{data}(x)}\left[D_Y(F_Y(x))^2\right],
\end{aligned}
\tag{8}
$$

3.4. Full Objective

Based on our changes, the full objective of enhanced CycleGAN is:

$$
\begin{aligned}
\mathcal{L}(G_{XY}, G_{YX}, D_X, D_Y) = \; & \mathcal{L}_{adv}(G_{XY}, D_Y, F_Y, X, Y) + \mathcal{L}_{adv}(G_{YX}, D_X, F_X, X, Y) \\
& + \lambda_{cyc}\mathcal{L}_{cyc}(G_{XY}, G_{YX}) + \lambda_{id}\mathcal{L}_{aid}(G_{XY}, G_{YX})
\end{aligned}
\tag{9}
$$

The overall structure of enhanced CycleGAN is shown in Figure 3.

4. Experimentation

4.1. Datasets

In this paper, we use two Kuzushi-ji datasets to measure the performance of the proposal. One is KMNIST [29] which contains 28×28 image size and gray-scale character of 10 classes. KMNIST has 60,000 train data and 10,000 test data. The other is the "Nisemonogatari" dataset (NISE) [30] which is the Kuzushi-ji dataset cut from actual early Japanese books. The NISE contains 21,701 characters of images resized into 64×64 RGB images with 468 classes. We divide them into damaged characters and undamaged characters for training. In the case of KMNIST, we visually choose 700 damaged characters from the 60,000 train data. We also choose 203 damaged characters from the 10,000 test data. Then, we define the remaining characters as undamaged train and test characters. In regard to NISE, we select 1518 damaged characters, then define 787 train characters (about 70%) and 381 test data(about 30%). The remaining data is undamaged train and test data. Table 1 reports the full number of damaged and undamaged characters. Considering the unbalance between damaged and undamaged train data, we pick constant 2000 characters from undamaged training data for the CycleGAN-based method.

Table 1. Details the number of datasets. Only when CycleGAN-based methods training. The undamaged train data is fixed at 2000 while training CycleGAN.

	Damage		Undamage	
	Train	**Test**	**Train**	**Test**
KMNIST	700	203	19,433 (2000)	9797
NISE	787	381	15,066 (2000)	9902

Finally, we apply our proposal to the whole pages of "Nisemonogatari" [30] for observing the revival of an early Japanese book. This book is a parody book published about 300 years ago and written about Japanese society at that time. The acquisition conditions of the images are very important for this research. Since the conditions are changed depending on the literature's size, shape, etc. Therefore, a digitization manual is made for literature and published in Japan, which is used for image acquisition [31]. The images of literature are taken based on this manual.

4.2. Training Condition

In this experiment, we train the CycleGAN-based methods according to the original implementation of [19]. The learning rate of G and D is 0.0002, but the learning rate of F is 1×10^{-6} to avoid the mode collapse for aid loss. We describe that the aid loss is affected by the generated image quality of F in Section 4.5.1. These learning late of G and F decay linearly from 100 to 200 epoch. We employ two data augmentation methods: center cropping and horizontal flipping at 50%. The network architecture is similarly following [19] of the style transfer network from [32]. The network architecture of F is the same as G. We decide $\lambda_{cyc} = 10$ and $\lambda id = 5$.

4.3. Comparing Conditions

To our knowledge, there is no experiential standard for deterioration restoration for faded ink and scratches. Hence, we measure Peak Signal-to-Noise Ratio(PSNR) [33] and Structure Similarity(SSIM) [34] between artificial damaged and restored images. We also compare the recognition accuracy of deep learning on restored images. Our proposed method is compared with the standard CycleGAN and binarization by the Otsu method. Regarding NISE, we preserve the 3-channels background by multiplying the binarized images and the original images.

4.3.1. PSNR/SSIM

To evaluate quantitatively, we reproduce the deteriorated images from undamaged images of the KMNIST and NISE in Figure 4. The deteriorated images are duplicated by adding $\alpha \times R$ to the pixel under a specific value $\beta \times R$. Here, α is the damage level, and β is the range of pixels value adding damage. R denotes the data range of the pixel value. We experiment with $\alpha = 0.1, 0.2, 0.3, 0.4, 0.5$ and $\beta = 0.5$ for NISE, and $\alpha = 0.5, 0.6, 0.7, 0.8, 0.9$ and $\beta = 0.1$ for KMNIST. Because a character part of KMNIST is a high pixel value, different from NISE, α is added after inverting KMNIST pixels, the manually reproduced deteriorated image is inverted again. These artificially damaged images are made from undamaged test data.
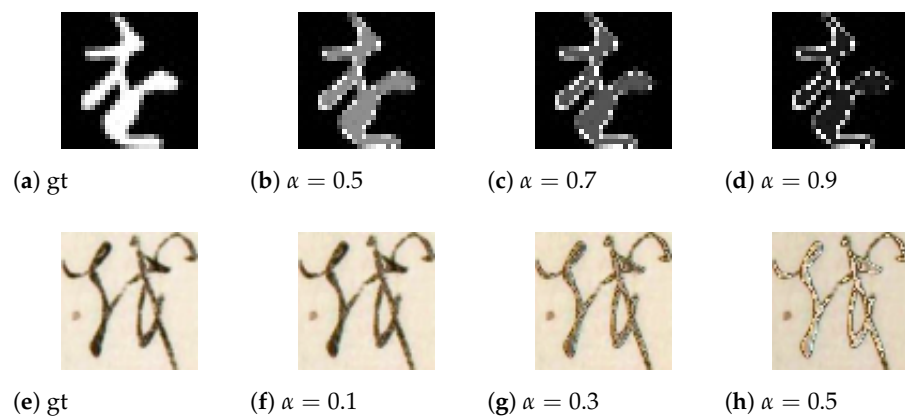
(**a**) gt　　　　　　(**b**) $\alpha = 0.5$　　　　　　(**c**) $\alpha = 0.7$　　　　　　(**d**) $\alpha = 0.9$

(**e**) gt　　　　　　(**f**) $\alpha = 0.1$　　　　　　(**g**) $\alpha = 0.3$　　　　　　(**h**) $\alpha = 0.5$

Figure 4. Manual deteriorated images by image processing.

4.3.2. Accuracy Measurement of Deep Learning Model

Moreover, we metric the recognition accuracy of the restored images by pre-trained deep learning model for each dataset. We employ four models consisting of multiple layers such as LeNet [35], ResNet18, ResNet34, and ResNet50 [36]. We train these models using all of the damaged and undamaged train data and compare the accuracy.

Table 2 is the accuracy of damaged and undamaged test data. It proves accurate recognition is hard due to the damages, which is the direct cause of the decreasing accuracy of the early Japanese books.

Table 2. Acc by the 4 types of classifiers on the test data of KMNIST and NISE.

	KMNIST		NISE	
	Damaged	**Undamaged**	**Damaged**	**Undamaged**
LeNet	0.788	0.958	0.622	0.797
ResNet18	0.906	0.989	0.937	0.978
ResNet34	0.887	0.984	0.942	0.978
ResNet50	0.901	0.986	0.934	0.973

4.4. Experimental Results

4.4.1. KMNIST

Figure 5 shows restoration results. The completely lost part is not restored by the binarization, but it succeeded in sharpening and clarification as shown in Figure 5b. Whereas, CycleGAN-based methods can connect the lost part (Figure 5c–e). Domain-d (Figure 5d) increases the degree of transformation more than the standard CycleGAN; however, it misses the characteristics. Our proposal, combining the domain-D and the augmented identity loss (Figure 5e), can restore satisfied both the degree of transformation and preserve the input's features. Other restoration results of KMNIST are shown in Figure 6.

Table 3 reports the experimental result by PSNR and SSIM. Our method is the best result for both PSNR and SSIM at $\alpha = 0.7, 0.8, 0.9$. When it comes to the result of $\alpha = 0.7$, SSIM increases by 10.93% compared with the results of inputs, and it is 3.13% higher than the standard CycleGAN. PSNR/SSIM of binarization is lower than CycleGAN-based methods with every α. Table 4 reports the accuracy of restored characters. Binarization is the highest accuracy on all the classifiers even though the PSNR/SSIM is the lowest. On the other hand, the accuracy of binarized undamaged is lower than CycleGAN-based methods. Our proposed method is higher for damaged characters except for ResNet50 than the standard CycleGAN.
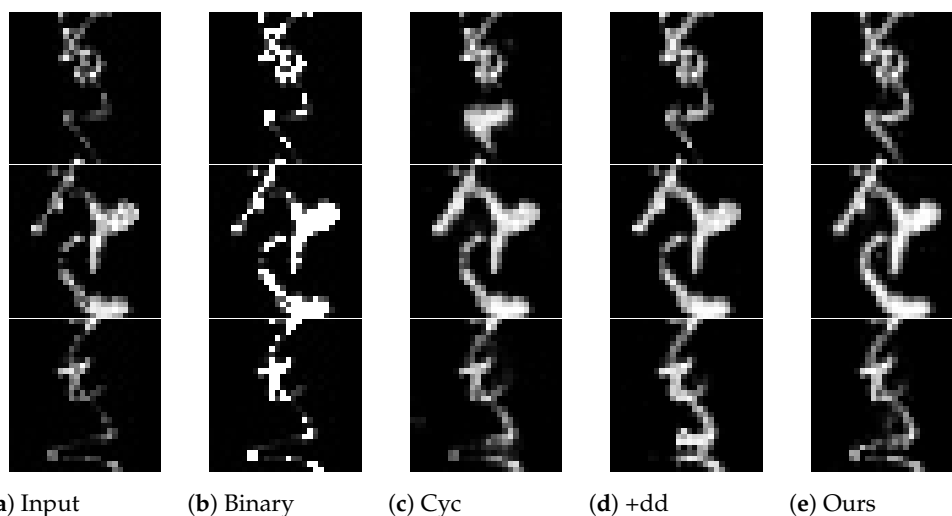
| (**a**) Input | (**b**) Binary | (**c**) Cyc | (**d**) +dd | (**e**) Ours |

Figure 5. Restoration result of the binarization and CycleGAN-based method for KMNIST. (**a**) The input data. (**b**) Binarization by Otsu method. (**c**) The standard CycleGAN. (**d**) The CycleGAN with domain discriminators. (**e**) The CycleGAN with combining domain discriminators and augmented identity loss.

Table 3. PSNR /SSIM of restored images on damage adding test data of KMNIST. The input is artificially processed damaged characters. The best results are shown in boldface.

	$\alpha = 0.5$	$\alpha = 0.6$	$\alpha = 0.7$	$\alpha = 0.8$	$\alpha = 0.9$
Input	16.36/0.6871	14.78/0.6509	13.49/0.6148	12.28/0.5764	11.35/0.5350
Binary	16.73/0.6698	16.24/0.6482	14.35/0.6504	11.23/0.5692	10.28/0.5297
Cyc	**19.15/0.8701**	17.15/**0.7514**	15.27/0.6928	13.68/0.6318	12.12/0.5550
Ours	18.44/0.6814	**17.88**/0.7032	**16.86/0.7241**	**15.32/0.6948**	**13.46/0.6375**

Table 4. The accuracy of restoration on damaged and undamaged KMNIST. The best results are shown in boldface.

	Damage				UnDamage			
	LeNet	**Res18**	**Res34**	**Res50**	**LeNet**	**Res18**	**Res34**	**Res50**
input	0.788	0.906	0.887	0.901	**0.958**	**0.989**	**0.984**	**0.986**
Binary	**0.847**	**0.926**	**0.926**	**0.921**	0.943	0.985	0.979	0.985
Cyc	0.827	0.897	0.887	0.892	0.957	0.987	0.982	0.985
+dd	0.837	0.882	0.892	0.877	0.955	0.986	0.981	0.984
Ours	**0.847**	0.901	0.901	0.891	0.954	0.986	0.981	0.984

(**a**) Input

(**b**) Binary

(**c**) Cyc

(**d**) Ours

Figure 6. Restoration results of KMNIST.

4.4.2. Nisemonogatari

Figure 7 displays the restoration results for NISE. Binarization is hard to fix the lost part of the data as same as KMNIST(Figure 7b), and even CycleGAN cannot fix dilute and partially lost parts(Figure 7c). The domain discriminator(Figure 7d) obtains the same results as KMNIST which has huge changes and feature missing. ur proposal achieves superior restoration(Figure 7e). Other restoration results of NISE are shown in Figure 8.

Table 5 reports the results of PSNR/SSIM on the NISE. Our method obtains the best results on $\alpha = 0.2, 0.3, 0.4$. It is higher than the standard CycleGAN except for $\alpha = 0.1$, moreover, the differences from the standard CycleGAN are just 0.55 for PSNR and 0.0019% for SSIM. In Table 6, the accuracy of binarization is lower than CycleGAN-based methods and inputs. Our method is higher accuracy for damaged characters than traditional methods and only one exceeds the inputs among the three classifiers.
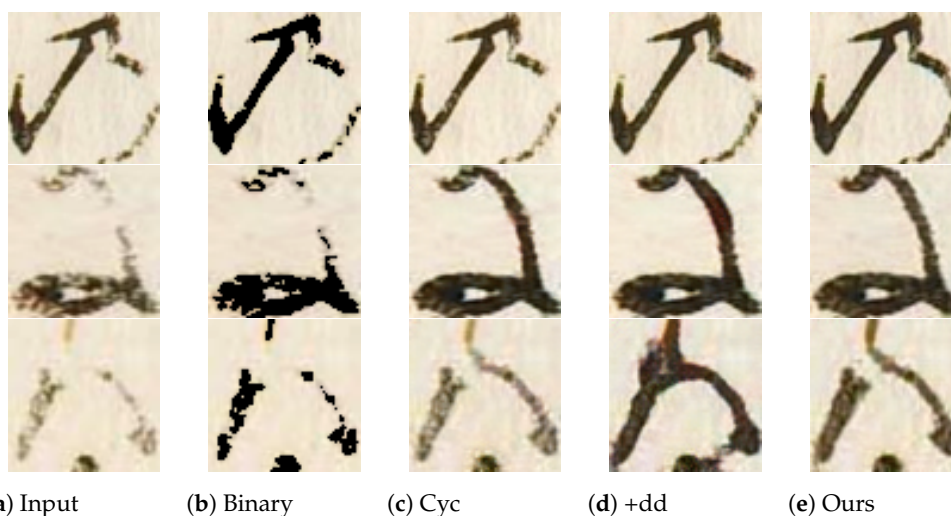
| (**a**) Input | (**b**) Binary | (**c**) Cyc | (**d**) +dd | (**e**) Ours |

Figure 7. Restoration result of the binarization and CycleGAN-based method for NISE. (**a**) The input data. (**b**) Binarization by Otsu method. (**c**) The standard CycleGAN. (**d**) The CycleGAN with domain discriminators. (**e**) The CycleGAN with combining domain discriminators and augmented identity loss.

Table 5. PSNR/SSIM of restored images on damage adding test data of NISE. The input is artificially processed damaged characters. The best results are shown in boldface.

	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.5$
Input	27.06/0.8378	21.04/0.6769	17.52/0.5917	15.02/0.5615	13.08/**0.5387**
Binary	14.69/0.5585	17.08/0.7382	13.54/0.4859	13.29/0.4776	12.00/0.4887
Cyc	**31.55/0.9195**	25.69/0.8435	20.57/0.6834	18.03/0.5651	15.92/0.4842
Ours	31.00/0.9176	**25.96/0.8704**	**21.83/0.7706**	**18.95/0.6413**	**16.43**/0.5309

Table 6. The accuracy of restoration on damaged and undamaged NISE. The best results are shown in boldface.

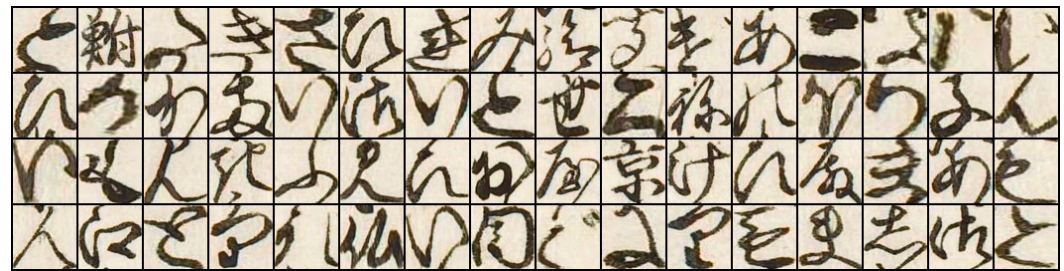|	Damage				UnDamage			
	LeNet	**Res18**	**Res34**	**Res50**	**LeNet**	**Res18**	**Res34**	**Res50**
Input	0.622	**0.937**	**0.942**	0.934	0.797	0.978	**0.978**	**0.973**
Binary	0.619	0.871	0.848	0.882	0.797	0.978	**0.978**	**0.973**
Cyc	0.646	0.929	0.932	0.929	0.791	0.978	0.977	0.972
+dd	0.646	0.919	0.921	0.924	**0.799**	0.978	0.976	0.972
Ours	**0.651**	**0.937**	0.929	**0.945**	0.793	0.978	0.976	**0.973**

(**a**) Input

(**b**) Binary

(**c**) Cyc

(**d**) Ours

Figure 8. Restoration results of NISE.

4.5. Full Page Restoration of the Early Japanese Book

Characters are extracted from the pages and divided into proper image sizes, then input to the Generator. Figure 9 shows the restoration results for the full pages (the Left is the damaged images, and the right is the restoration results). We observe a vivid ink color and the reconnection of scratched and lost parts. Through the restoration of the weak damages, whole characters are clarified on the page. Table 7 reports the processing times of full-page restoration by binarization and enhanced CycleGAN. For the CycleGAN-based method, we measure the times on the CPU and the GPU, which AMD Ryzen 9 3590X 16-Core Processor CPU and Nvidia GeForce GTX 1080ti GPU. Deep learning methods need longer processing time than traditional image processing. However, the restoration is automatic and far faster than human work. Furthermore, proposal achieves a better improvement than image processing method.

Table 7. The computation time of full-page restoration. Page 1 is showed in Figure 1 and page 2 and page3 are showed in Figure 9. The items are the number of extracted characters from each page.

	Page 1 (184 Items)	Page 2 (130 Items)	Page 3 (205 Items)
Binary	0.3653	0.3377	0.4175
Ours(CPU)	3.7580	2.8321	4.0769
Ours(GPU)	0.7612	0.6183	0.8538

Figure 9. Full pages restoration of Nisemonogatari by enhanced CycleGAN.

4.5.1. Effect of the Augmented Identity Loss

The aid loss is affected by generated image quality. Figure 10 shows output images by FY without domain-d and in different learning rates (lr). Without the domain-d, F_Y should generate undamaged characters, but we observe those characters contain faded ink and weak scratches. As shown in Table 8, the outputs of F without domain-d are not suitable for aid loss. Figure 10b shows that the generator at $lr = 5 \times 10^{-6}$ occurs the mode collapse, which outputs identical images. It decreases the accuracy of aid loss at $lr = 5 \times 10^{-6}$ compared with $lr = 1 \times 10^{-5}$. These results indicate that aid loss works as data augmentation and data generalization for identity mapping loss.
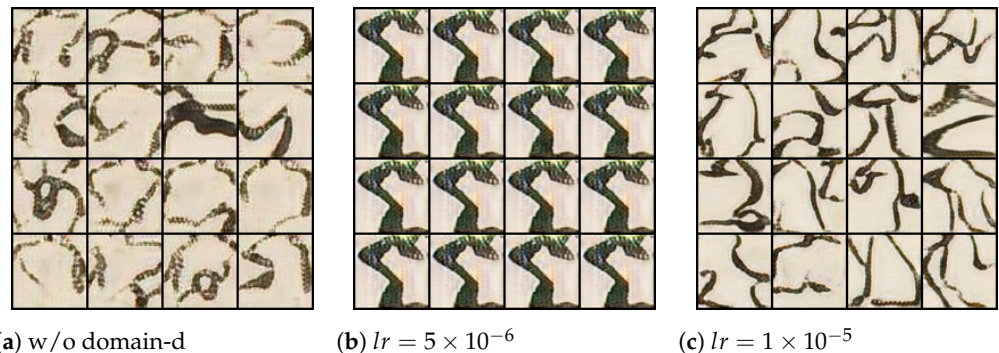

(**a**) w/o domain-d (**b**) $lr = 5 \times 10^{-6}$ (**c**) $lr = 1 \times 10^{-5}$

Figure 10. Outputs by F_Y in the cases of (**a**) Without domain-d (w/o d-d) and $lr = 5 \times 10^{-6}$. (**b**) $lr = 5 \times 10^{-6}$. (**c**) $lr = 1 \times 10^{-5}$ and with domain-d.

Table 8. The detailed result by aid loss. The implementation Without domain-D(w/o dd) is not employed domain-D, but aid loss is used. W/o aid is only employed domain-d. The best results are shown in boldface.

	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$	$\alpha = 0.4$	$\alpha = 0.5$
w/o dd	21.60/0.7281	19.87/0.6722	18.07/0.6019	15.83/0.4964	13.23/0.3474
w/o aid	31.56/0.9098	24.84/0.8045	19.72/0.6574	16.55/0.5426	12.83/0.3904
$lr = 5 \times 10^{-6}$	28.21/0.8715	24.84/0.8143	21.04/0.7096	18.29/0.6303	**16.45/0.5354**
$lr = 1 \times 10^{-5}$	**31.00/0.9176**	**25.96/0.8704**	**21.83/0.7706**	**18.95/0.6413**	16.43/0.5309

5. Conclusions

We proposed deteriorated characters restoration using enhanced CycleGAN with the domain discriminator and the augmented identity loss. This enhanced CycleGAN enables high-quality restoration compared with the standard CycleGAN and binarization in the actual early Japanese book. Furthermore, it provides the revitalization of ancient documents and high readability. Our proposal encourages more accurate recognition of deep learning models in the actual book. In other words, high-quality restoration makes them readable for us as well as artificial intelligence. In this paper, we achieve slight-to-moderate damage restoration on a specific early Japanese book. However, our method has a limitation for complete damage because the damaged character is few, and it is based on CycleGAN. Furthermore, the early Japanese books have multiple types of deterioration and color configuration, and paper texture. In terms of future work, we aim to realize a high degree of freedom restoration for ancient documents.

Author Contributions: Methodology, H.K.; Software, R.I.; Project administration, L.M. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

[1]   Unsupervised means paired teaching images is not necessarily.

References

1.   Lyu, B.; Li, H.; Tanaka, A.T.; Meng, L. The Early Japanese Books Reorganization by Combining Image processing and Deep Learning. *CAAI Trans. Intell. Technol.* **2022**, *7*, 627–643.
2.   Lamb, A.; Clanuwat, T.; Kitamoto, A. KuroNet: Regularized residual U-Nets for end-to-end Kuzushiji character recognition. *SN Comput. Sci.* **2020**, *1*, 1–15.
3.   Chen, L.; Bing, L.; Hiroyuki, T.; Meng, L. A Method of Japanese Ancient Text Recognition by Deep Learning. *Procedia Comput. Sci.* **2020**, *174*, 276–279.
4.   Available online: http://codh.rois.ac.jp/ (accessed on 10 May 2023).
5.   Available online: https://www.arc.ritsumei.ac.jp/ (accessed on 10 May 2023).
6.   Meng, L. Two-Stage Recognition for Oracle Bone Inscriptions. In *Proceedings of the Image Analysis and Processing—ICIAP 2017*; Battiato, S., Gallo, G., Schettini, R., Stanco, F., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 672–682.
7.   Yue, X.; Li, H.; Fujikawa, Y.; Meng, L. Dynamic Dataset Augmentation for Deep Learning-based Oracle Bone Inscriptions Recognition. *J. Comput. Cult. Herit.* **2022**, *15*, 1–20.
8.   Lyu, B.; Tanaka, A.; Meng, L. Computer-assisted Ancient Documents Re-organization. *Procedia Comput. Sci.* **2022**, *202*, 295–300.
9.   Zhang, Z.; Wang, Z.; Tomiyama, H.; Meng, L. Deep Learning and Lexical Analysis Combined Rubbing Character Recognition. In Proceedings of the 2019 International Conference on Advanced Mechatronic Systems (ICAMechS), Kusatsu, Japan, 26–28 August 2019; pp. 57–62.
10.   Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144.
11.   Su, B.; Liu, X.; Gao, W.; Yang, Y.; Chen, S. A restoration method using dual generate adversarial networks for Chinese ancient characters. *Vis. Inform.* **2022**, *6*, 26–34.
12.   Wenjun, Z.; Benpeng, S.; Ruiqi, F.; Xihua, P.; Shanxiong, C. EA-GAN: Restoration of text in ancient Chinese books based on an example attention generative adversarial network. *Herit. Sci.* **2023**, *11*, 1–13.
13.   Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.
14.   Kim, G.; Park, J.; Lee, K.; Lee, J.; Min, J.; Lee, B.; Han, D.K.; Ko, H. Unsupervised real-world super resolution with cycle generative adversarial network and domain discriminator. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, New Orleans, LA, USA, 19–20 June 2020; pp. 456–457.
15.   Chamchong, R.; Fung, C.C.; Wong, K.W. Comparing binarisation techniques for the processing of ancient manuscripts. In Proceedings of the Cultural Computing: Second IFIP TC 14 Entertainment Computing Symposium, ECS 2010, Held as Part of WCC 2010, Brisbane, Australia, 20–23 September 2010; pp. 55–64.
16.   Aravinda, C.; Lin, M.; Masahiko, A.; Amar, P.G. A complete methodology for kuzushiji historical character recognition using multiple features approach and deep learning Model. *Int. J. Adv. Comput. Sci. Appl.* **2020**, *11*, pp. 694–700 .
17.   Sulaiman, A.; Omar, K.; Nasrudin, M.F. Degraded historical document binarization: A review on issues, challenges, techniques, and future directions. *J. Imaging* **2019**, *5*, 48.
18.   Bangare, S.L.; Dubal, A.; Bangare, P.S.; Patil, S. Reviewing Otsu's method for image thresholding. *Int. J. Appl. Eng. Res.* **2015**, *10*, 21777–21783.
19.   Isola, P.; Zhu, J.Y.; Zhou, T.; Efros, A.A. Image-to-image translation with conditional adversarial networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, USA, 18–20 June 2017; pp. 1125–1134.
20.   Gu, J.; Yang, T.S.; Ye, J.C.; Yang, D.H. CycleGAN denoising of extreme low-dose cardiac CT using wavelet-assisted noise disentanglement. *Med Image Anal.* **2021**, *74*, 102209.
21.   Luleci, F.; Catbas, F.N.; Avci, O. CycleGAN for Undamaged-to-Damaged Domain Translation for Structural Health Monitoring and Damage Detection. *arXiv* **2022**, arXiv:2202.07831.
22.   Guo, X.; Liu, X.; Królczyk, G.; Sulowicz, M.; Glowacz, A.; Gardoni, P.; Li, Z. Damage detection for conveyor belt surface based on conditional cycle generative adversarial network. *Sensors* **2022**, *22*, 3485.
23.   Sandfort, V.; Yan, K.; Pickhardt, P.J.; Summers, R.M. Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability in CT segmentation tasks. *Sci. Rep.* **2019**, *9*, 16884.
24.   Liu, R.; Wang, X.; Lu, H.; Wu, Z.; Fan, Q.; Li, S.; Jin, X. SCCGAN: Style and characters inpainting based on CGAN. *Mob. Networks Appl.* **2021**, *26*, 3–12.
25.   Mao, X.; Li, Q.; Xie, H.; Lau, R.Y.; Wang, Z.; Paul Smolley, S. Least squares generative adversarial networks. In Proceedings of the IEEE Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2794–2802.
26.   Taigman, Y.; Polyak, A.; Wolf, L. Unsupervised cross-domain image generation. *arXiv* **2016**, arXiv:1611.02200.

27. Xu, K.; Zhou, Z.; Wei, L. Effect analysis of image translation by controlling size of dataset and adjusting parameters for cyclegan. In Proceedings of the 2021 International Conference on Communications, Information System and Computer Engineering (CISCE), Beijing, China, 14–16 May 2021; pp. 458–461.
28. Wang, T.; Lin, Y. CycleGAN with Better Cycles. 2019. https://ssnl.github.io/better_cycles/report.pdf (accessed on 10 May 2023).
29. Clanuwat, T.; Bober-Irizar, M.; Kitamoto, A.; Lamb, A.; Yamamoto, K.; Ha, D. Deep learning for classical japanese literature. *arXiv* **2018**, arXiv:1812.01718.
30. Kuzushiji Dataset (Center for Open Data in the Humanities, National Institute of Japanese Literature). Available online: https://doi.org/10.20676/00000340 (accessed on 10 May 2023).
31. National Institute of Japanese Literature. Manual for the Digitization of Japanese Historical Literature. 2022 . Available online: https://www.nijl.ac.jp/pages/cijproject/images/digitization-manual/digitization-manual_NIJL-202205.pdf (accessed on 10 May 2023).
32. Johnson, J.; Alahi, A.; Fei-Fei, L. Perceptual losses for real-time style transfer and super-resolution. In Proceedings of the Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, 11–14 October 2016; pp. 694–711.
33. Huynh-Thu, Q.; Ghanbari, M. Scope of validity of PSNR in image/video quality assessment. *Electron. Lett.* **2008**, *44*, 800–801.
34. Wang, Z.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612.
35. LeCun, Y.; Haffner, P.; Bottou, L.; Bengio, Y. Object recognition with gradient-based learning. In *Shape, Contour and Grouping in Computer Vision*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 319–345.
36. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.