INTERNATIONAL JOURNAL OF COMPUTER ENGINEERING AND APPLICATIONS
VOLUME XVI, ISSUE IV, APRIL 22, WWW.IJCEA.COM ISSN 2321-3469

DEEP LEARNING APPROACH TO RESTORE OLD IMAGE BACK TO LIFE

Aishwarya Patil [1], Hritik Londhe [1], Shubham Kadave [1], Megha V. Gupta [2]

[1]Research Scholar, [2]Vice Principal, Department of Computer Engineering, New Horizon Institute of Technology and Management, University of Mumbai, Thane, India

ABSTRACT:

We propose using deep learning to recover historical old pictures that have suffered substantial deterioration. Unlike conventional traditional restoration tasks that can be handled using supervised learning, the deterioration in real world photos is complicated, and the domain gap between synthetic images and real-world old photos causes the network to fail to generalize. As a result, we present a unique novel triplet domain translation network by leveraging real photos along with massive synthetic image pairs. Specifically, we train two variational autoencoders (VAEs) to transform old photos and clean photos into two latent spaces, respectively. Synthetic paired data is used to learn the translation between these two latent spaces areas. Because the domain gap is closed in the compact latent space, this translation generalizes effectively to real photographs. Furthermore, to handle numerous degradations intermingled in a single old photo, we build a global branch with a partial nonlocal block targeting structured defects, like scratches and dust spots and a local branch targeting unstructured defects such as noises and blurriness. In the latent space, two branches are fused, resulting in increased potential to restore old photos that have various defects. In addition, we use another face refinement network to retrieve fine aspects of faces in the old photos, ultimately generating photos with improved perceptual quality. With comprehensive experiments, the proposed pipeline demonstrates superior performance over state-of-the-art methods as well as existing commercial tools in terms of visual quality for old photos restoration.

Keywords: Variational Autoencoder, Latent space, Structured Degradation, Unstructured Degradation, Non-local block, Residual Block

[1] INTRODUCTION

Photos are taken to freeze the wonderful moments that would otherwise be gone. Even though time passes, watching them might elicit recollections of the past. Nonetheless, vintage old photo prints deteriorate when stored in bad environmental conditions, resulting in irreversible destruction to the valuable photo information. Fortunately, as mobile cameras and scanners become more accessible and affordable, consumers may now digitalize the photos and send them to a competent professional for restoration. Manual retouching, on the other hand, is typically arduous and time-consuming, rendering stacks of old images impossible to restore. As a result, it is interesting to design automatic algorithms that can immediately repair old images for individuals who want to resurrect ancient photos to old life.

Prior to the advent of deep learning, various attempts were made to recover images by automatically recognizing the localized defects such as scratches and blemishes and filling up the damaged areas with inpainting techniques.[1] However, because none of these approaches can fix spatially homogeneous problems like as film grain, sepia effect, color fading, and so on, the photographs after restoration still seem antiquated when compared to current photographic images. With the emergence of deep learning, we will be addressing a variety of low-level image restoration problems by exploiting the powerful representation capability of convolutional neural networks, i.e., learning the mapping for a specific task from a large number of synthetic images. However, the same framework. does not apply to old photo restoration. As a result, the model trained from the synthetic data generalizes poorly to real-world images. Second, old photos are plagued with a compound of degradations and inherently requires different strategies for repair: unstructured defects that are spatially homogeneous such as film grain and color fading, should be restored by utilizing the pixels in the neighborhood, whereas the structured defects, such as scratches, dust spots, etc., should be repaired with a global image context.[2]

[1.1] DEGRADATION IN IMAGES:

In this project, we intend to design a old photo restoration model that can remove both structured defects such as scratches and dust spots, and also the unstructured defects, such as noises and blurriness. The following picture restoration approaches have been mentioned: In today's world of comfort and luxury, various high priced costly vehicles are available. Many of these vehicles have been launched with inbuilt security systems. However, even though a huge amount of capital is being invested in areas of vehicle security, the cases of vehicle theft is still rising. [1,2]

Single Degradation Image Restoration:

Image degradation can be roughly categorized into two groups: unstructured degradation such as noise, blurriness, color fading, and low resolution, and structured degradation such as holes, scratches, and spots. For the former unstructured ones, traditional works often impose different image priors, including non-local self-similarity, sparsity and local smoothness. Recently, a lot of deep learning-based methods have also been proposed for different image degradation, like image denoising, super-resolution and deblurring.[3] Compared to unstructured degradation, structured degradation is more challenging and often modeled as the "image painting" problem. Thanks to powerful semantic modeling ability, most existing best-performed inpainting methos are learning based.

For real old images, since they are often seriously degraded by a mixture of unknown degradation, the underlying degradation process is much more difficult to be accurately characterized. In other words, the network trained on synthetic data only, will suffer from the domain gap problem and perform badly on real old photos.[4] In our project, we will model real old photo restoration as a new triplet domain translation problem and some new techniques are adopted to minimize the domain gap.

Mixed Degradation Image Restoration:

In the real world, a corrupted image may suffer from complicated defects mixed with scratches, loss of resolution, color fading, and film noises.[5] However, research solving mixed degradation is much less explored. Existing methods still rely on supervised learning from synthetic data and hence cannot generalize to real photos. Besides, they only focus on unstructured defects and do not support structured defects like image inpainting.

Old Photo Restoration:

Old photo restoration is a classical mixed degradation problem, but most existing methods focus on inpainting only. They follow a similar paradigm i.e., defects like scratches and blotches are first identified according to low level features and then in painted by borrowing the textures from the vicinity. However, the hand-crafted models and low-level features used are difficult to detect and fix such defects well. Moreover, none of the previous models consider restoring some unstructured defects such as color fading or low resolution together with inpainting.[3] Thus, photos still appear old fashioned after restoration. In our work, we reinvestigate this problem by virtue of a data driven approach, which can restore images from multiple defects simultaneously and turn heavily-damaged old photos to modern style.

[2] LITERATURE REVIEW

Following are the various insights gathered from different sources which have proved helpful in our literature survey. Hence, the purpose of this review is to gain an understanding of the various approaches for old photo restoration, as well as their flaws that may be addressed.

[I] Image Restoration for Halftone Pattern Printed Pictures in Old Books

They proposed an image restoration technique for halftone pattern printed pictures. This approach recovered the original analogue halftoning of the old pictures into a digital clean halftoning.[1] the scanned image was transformed in its black and white version by applying the optimal threshold given in MATLAB by the gray thresh function. This technique successfully processes black and white spots with less than 20 pixels in composition, replacing them with digital regular shapes, which makes the image clearer and cleaner.

[II] Noise2Noise: Learning Image Restoration without Clean Data

They proposed a model that learns photographic noise removal, denoising synthetic Monte Carlo images, and reconstruction of under sampled MRI scans - all corrupted by different processes - based on noisy data only.[2] For denoising the images, they have made use of BM3D denoiser. They have shown that clean data is not necessary for denoising, for instance, the classic BM3D algorithm draws on self-similar patches within a single noisy image.

[III] Towards the Automation of Deep Image Prior

They proposed a stopping method, Orthogonal Stopping Criterion (OSC) that adds a pseudo noise to the corrupted image and measures the pseudo-noise component in the recovered image by the orthogonality between signal and noise. Their work is about performing useful restoration tasks with a Convolutional Neural Network. Their projects focus on restoring the input image with many blocky artifacts that are materialized during JPEG compression, image inpainting where some regions of input image are missing, super resolution, image denoising. [3]

[IV] CycleISP: Real Image Restoration via Improved Data Synthesis

They have proposed a data synthesis approach for denoising. Here to obtain noisy version of clean images they have added white Gaussian noise.[4] Their proposed CycleISP framework models camera imaging pipeline in forward and reverse directions. This framework allows them to generate realistic clean/noisy pairs in sRGB and RAW spaces.

To circumvent the shortcomings of existing systems, we will formulate the old photo restoration as a triplet domain translation problem.

[3] PROBLEM DEFINITION

After a certain period, old photo prints deteriorate when stored in poor environmental conditions. The goal of this project is to develop a system that can bring damaged photos back to life by removing structured degradation such as scratches and dust spots and unstructured degradation such as noises and blurriness.

[4] PROPOSED SYSTEM OVERVIEW

In contrast to conventional image restoration tasks, old photo restoration is more challenging. Firstly, old images have significantly more intricate deterioration that is difficult to be modeled realistically and there always exists a domain gap between synthetic and real photos. As a result, the network usually cannot generalize well to real photos by purely learning from synthetic data. [3] Second, the defects in old images are a compound of multiple degradations, thus essentially requiring different strategies for restoration. Unstructured defects such as film noise, blurriness and color fading, etc. can be restored using spatially homogeneous filters by utilizing surrounding pixels within the local patch; structured defects such as scratches and blotches, on the other hand, should be inpainted by considering the global context to ensure the structural consistency.

We formulate the old photo restoration as a triplet domain translation problem. Different from previous image translation methods, we will leverage data from three domains (i.e., real old photos, synthetic images and the corresponding ground truth), and the translation is performed in latent space. We will be proposing a novel triplet domain translation network by leveraging real photos along with massive synthetic image pairs. Specifically, we train two variational autoencoders (VAEs) to respectively transform old photos and clean photos into two latent spaces. And the translation between these two latent spaces is learned with synthetic paired data. This translation generalizes well to real photos because the domain gap is closed in the compact latent space. Besides, to address multiple degradations mixed in one old photo, we design a global branch with a partial nonlocal block targeting to the structured defects, such as scratches and dust spots, and a local branch targeting to the unstructured defects, such as noises and blurriness. Two branches are fused in the latent space, leading to improved capability to restore old photos from multiple defects.

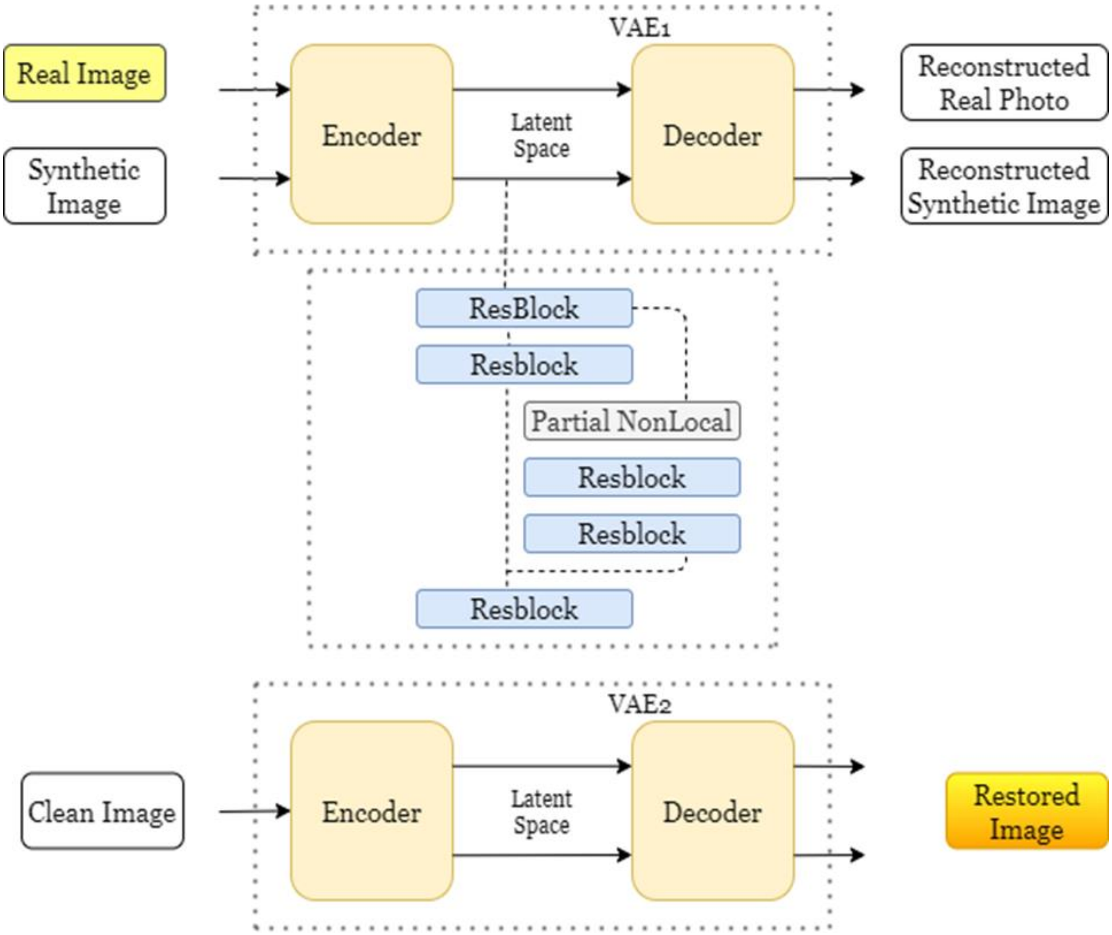

Figure: 1. Block Diagram of Proposed System.

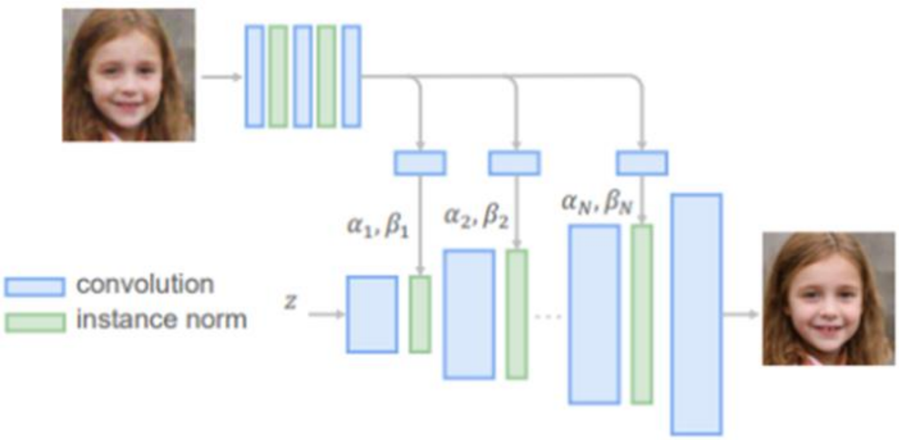

Figure: 2. The progressive generator network of face enhancement.

[I] Training Dataset:

Using images from the Pascal VOC dataset, we have synthesized old photos. We also gathered scratch and paper textures, which are then enhanced with elastic distortions, in order to render realistic defects. To blend the scratch textures over the real images from the dataset, we have utilized layer addition, lighten-only and screen modes with random level of opacity. To simulate large-area photo damage, we generated holes with feathering and random shape where the underneath paper texture is unveiled. Finally, to simulate the unstructured defects, film grain noises and blurring with random amount are included.

[II] Scratch Detection:

We train another network with Unet architecture to detect structured region for the partial nonlocal block. The detection network is first trained using only synthetic images. To correct the imbalance of positive and negative detections, we use the focal loss. To increase detection performance on real old photos, we annotate 783 gathered old photos with scratches, of which 400 are used to fine-tune the detection network.

[4.1] ALGORITHM

[I] Latent Translations with VAE's:

- Pix2Pix which learns the translation in image-level. The model is trained using only synthetic pairs.
- Two VAEs with an additional KL loss to penalize the latent space. The VAEs and latent mapping are all trained simultaneously.
- VAEs with two-stage training (VAEs-TS): the two VAEs are individually trained firs, then the latent mapping is learnt using the two fixed VAEs, ensuring that the translation is performed in two fixed domains.
- A Full model with latent adversarial loss is used.

[II] Partial Non-Local Block:

- A partial nonlocal block is proposed to make the triplet domain translation network support the restoration of structured degradation.
- We found that the partial nonlocal block could also ensure that the inpainting is only applied in the localized defect areas.

[III] Face Enhancement Network:

- The face enhancement network is jointly trained with the triplet domain translation network, i.e., input corrupted faces will first pass through this translation network, and then be reconstructed into a high-resolution version with the proposed enhancement network.
- We could observe that without joint training, unnatural redundant textures and artifacts are visible in the generated faces.
- One reason may be there existing some distribution bias between generated faces of the first stage and real degraded faces. By introducing the joint training, this kind of gap could be alleviated, leading to more pleasant and consistent results.

- To reconstruct a high-resolution face from real photos meanwhile maintaining underlying structure and style information, we propose to modulate the features of the coarse-to-fine generator in a hierarchical spatial condition manner.
- Hierarchical spatial injection achieves natural restoration results with the right structures and styles.
- By introducing the method of hierarchical injection, our enhancement network obtains the best scores on all four metrics.

[5] RESULTS AND DISCUSSION

To better illustrate the subjective quality, we conducted a study to compare our method with available services. We randomly selected some images having both structured and unstructured degradation and compared the restoration quality of other methods with ours. Some commercial software and applications, such as MintAI-Photo Enhancer and Remini Photo Enhancer, have begun to provide automated old photos restoration services. We also compare the restoration performance with them in Figure to highlight the efficacy of our approach. From their results, it was feasible to determine that their approaches overlook existing structural degradations and color fading. Our method, on the other hand, solves these issues and produces more visually appealing outcomes, such as the first and third rows in Figure. Meanwhile, because the proposed latent domain translation network has a smaller domain gap, it is better at dealing with real-world local defects such as noise. Finally, the enhanced human face has additional features due to a dedicated face enhancement network. In comparison to commercial equivalents, our approach could provide more clear, crisp, and vivid outcomes.

Figure: 4. Qualitative comparison with commercial software. It shows that our method can restore both unstructured and structured degradation and recovered results are significantly better than other methods.

[6] CONCLUSION

The existing systems restored photos by automatically detecting the localized defects such as scratches and blemishes, and filling in the damaged areas with inpainting techniques. Yet these methods focus on completing the missing content and none of them can repair the spatially-uniform defects such as film grain, sepia effect, color fading, etc., so the photos after restoration still appear outdated. Early works attempt to deblur faces by the guidance of an external reference, but an exemplar image with suitable texture for transfer is inconvenient to retrieve and the requirement of an external face database makes it cumbersome for practical usage.[5] The problem with previous conventional restoration techniques is that they are not able to generalize this is because they are using supervised learning which is a problem caused by domain gap between the old picture and the ones that are synthesized for training. The existing systems either remove the structured defects, such as scratches and dust spots, or else unstructured defects, such as noises and blurriness. No method is able to remove mixed degradation simultaneously. To restore the mixed degradation in old photos, we proposed a novel triplet domain translation network. The domain gap is reduced between old photos and synthetic images, and the translation to clean images is learnt in latent space. When compared to previous approaches, our method has less generalization issues. We proposed a partial nonlocal block that uses global context to recover the latent features, allowing scratches to be inpainted with better structural consistency. Our method has shown to be effective in restoring severely deteriorated old photos. Our approach, on the other hand, is incapable of handling sophisticated shading. This is due to the fact that our dataset contains few old photos with such defects. This constraint might be addressed using our approach by explicitly addressing shading effects during synthesis or by adding extra shades. This project can also be extended by offering the users an option whether they want the old photos to be restored as black and white image or a color image.

REFERENCES

[1] Adrian Ciobanu, Tudor Barbu, Mihaela Luca, "Image Restoration for Halftone Pattern Printed Pictures in Old Books" 2018, 10th International Conference of Electronics, Computers and Artificial Intelligence, Iasi, Romania, ECAI 2018, DOI:10.1109/ECAI.2018.8678981

[2] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, Timo Aila, "Noise2Noise: Learning Image Restoration without Clean Data" 2018, Proceedings of the 35th International Conference on Machine Learning, ICML 2018

[3] Qianwei Zhou, Chen Zhou, Haigen Hu, Yuhang Chen, Shengyong Chen, Xiaoxin Li "Towards the Automation of Deep Image Prior" 2019, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 9446–9454.

[4] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, Ling Shao, "CycleISP: Real Image Restoration via Improved Data Synthesis",2020, Proceedings of the IEEE Conference on Image and Video Processing, Computer Vision and Pattern Recognition.

[5] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-resolution image synthesis and semantic manipulation with conditional GANs" in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 8798–8807.

[6] "MintAI-PhotoEnhancer," https://apps.apple.com/app/id1573527767?ign-mpt=uo%3D4.

[7] "Remini photo enhancer," https://www.bigwinepot.com/indexen.html.

Author[s] brief Introduction

Aishwarya Patil

Aishwarya Patil is a Computer Engineering Research Scholar at New Horizon Institute of Technology and Management, Thane. Her research interests include Machine Learning, Deep Learning and Full Stack Development. She is the member of CSI.

Hritik Londhe

Hritik Londhe is a Computer Engineering Research Scholar at New Horizon Institute of Technology and Management, Thane. His research interests include Machine Learning and Deep Learning.

Shubham Kadave

Shubham Kadave is a Computer Engineering Research Scholar at New Horizon Institute of Technology and Management, Thane. His research interests include Deep Learning and Development of Gaming apps.

Megha V. Gupta

Mrs. Megha V Gupta is a Computer Engineering Research Scholar at the Research Centre, Datta Meghe College of Engineering in Airoli, Navi Mumbai. She is the Vice-Principal at the New Horizon Institute of Technology and Management, Thane, besides being an Assistant Professor at the Dept. of Computer Engineering. Her research interests include Artificial Intelligence and Machine Learning. She has also published a patent and over 15 published research papers in international journals/conferences. Mrs. Gupta pursued her engineering education at the University of Mumbai and RTM Nagpur University from where she has ME, Computer Engineering, and B.E Computer Technology degrees respectively. She is a member of IEEE and CSI.