Old Photos Restoration by Using VAE

Pingyi Ma[1,*], Haozhe Kuang[2,a]

[1]Krieger School of Arts & Science, Johns Hopkins University, Baltimore, 21218, U.S.A
[2]Shanghai Jianping High School, Shanghai, 200135, China

Abstract. VAE is a generative model that "provides a probabilistic description of observations in potential Spaces". Put simply, this means that VAE stores potential attributes as probability distributions. The idea of variational auto-encoders or VAE is deeply rooted in the methods of variational DB Bayesian and graphical models. This piece of work will discuss VAE Structure, VAE Loss Function, VAE Translation, and our final effects.

1 Introduction.

Photos are important to people as they are taken to store memories and important moments. They are deemed as flashback points to people. As technology proceeds, contemporary people tend to take digital photos and store them online, which preserves the quality of photos and makes them easily accessible. However, many old photos taken from the last century or even earlier were taken by cameras and printed on paper. Although they were preserved, many of them are now degraded through time. Though some light degradation can be removed by specialists, photos experiencing severe scratches, loss of color, and holes are not likely to be restored. Hence, it requires the usage of the algorithm to restore old photos.

Currently, there are multiple algorithms for old photos in painting, among which the method [1], proposed by the Microsoft team, involves variational auto-encoders that render decent effects that can bring old photos back to high resolution and remove most scratches.

When This work thinks of machine learning, the first thing that most likely comes to mind is various algorithms. Discriminant models, which predict labels or categories of input data based on their features, are at the heart of all classification and prediction solutions. In contrast to these models, generative algorithms help us tell a story about the data and provide possible explanations for how the data was generated. Unlike various algorithms that map features to labels, generative models attempt to predict features given labels [2].

A standard auto-encoder consists of two similar networks, an encoder and a decoder. The encoder takes the input and transforms it into a smaller representation that the decoder can use to transform it back to the original input. The latent space into which they transform the input and the space in which their encoding vectors reside may not be continuous. This is a problem for generative models, since we all want to variations in the input image to be generated randomly from the latent spaces, or from consecutive latent spaces.

This work will share the progress about how two autoencoders work and the essence for the discriminant models.

2 Method for Old Photo Restoration

2.1 VAE Structure

The key factor in the method is that it stores synthesized old photos with degradation in one space X, real photos in one space R, and old photos without degradation in space Y. Figure 1 shows the whole procedure of VAE structure. In figure 1, it puts real photos and synthesized old photos into their own latent spaces that share the same domain because these types of photos are both corrupted, sharing some constraints. Then, real photos are put into one latent space. Then, it trains two VAEs. with VAEs, images are transformed to compact latent space [3].

[*]Pma14@jh.edu
[a]2814559307@qq.com

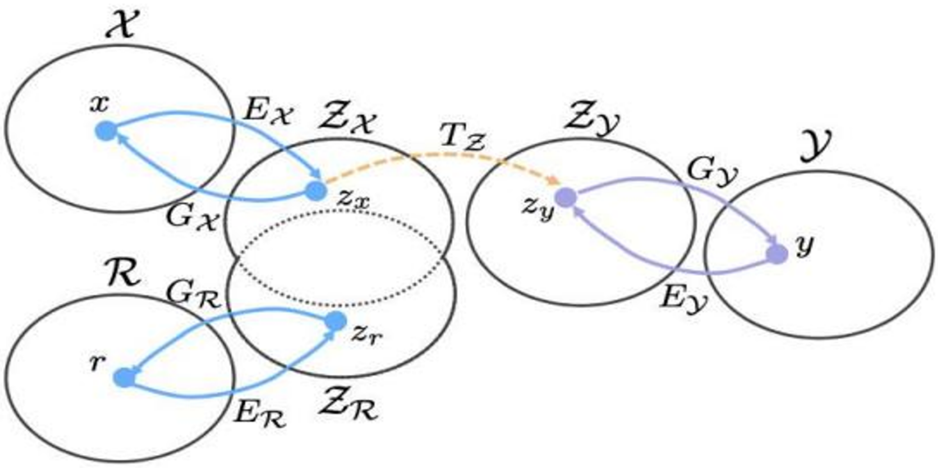

Fig. 1. The progress that translation methods using three domains from paper [3]

2.2 VAE Loss Function

In the method, it utilizes two part of loss functions as matrix to identify the quality of the output. The first part is generative loss, which compares the model output with the model input. The second part is latent loss. This loss compares the latent vector with a zero main, unit variance Gaussian distribution. It penalizes the VAE if it starts to produce latent vectors that are not from the desired distribution [4].

2.3 VAE Translation

When each domain has photos, the process of VAE translation begins. Figure 2 shows the process of VAE Translation. In figure 2, VAE leverages data from three domains: real old photos, synthetic images, and restored photos. The translation is performed in latent space. The mapping between the two latent spaces is then learned with the synthetic image pairs, which restore the corrupted images to clean ones [5].

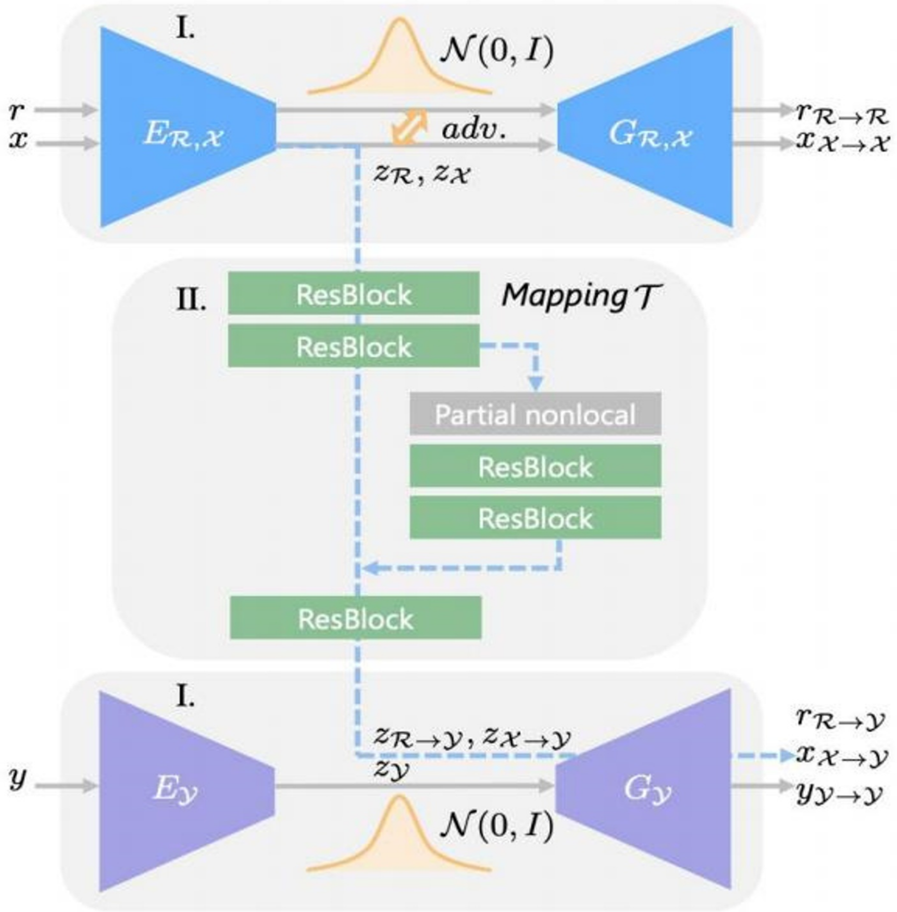

Fig. 2. A brief structure for old photo restoration

3 Outcome of VAE

The method synthesized old photos from the Pascal VOC data set and human-made degradation. After running, it outperforms most state-of-art algorithms. The detailed results are shown in Table 1. This work managed to reproduce their method by using their open-source code [6]. This work uses partial photos from the 2012 Pascal VOC data set and gathers our own old photos to train the model. The result of this work run is similar to the result provided by the team. However, the current deficit of the model is that it fails to restore photos with severe starches and discards some details of old photos. This work managed to change some parameters in the code, but no significant result was found [6].

Table 1. The outcome of the implementation of VAE compared with other methods

Method	Top 1	Top 2	Top 3	Top 4	Top 5
DIP [**43**]	2.75	6.99	12.92	32,63	69.70
CycleGAN [**44**]	3.39	8.26	15.68	24.79	52.12
Sequential [**56, 57**]	3.60	20.97	51.48	83.47	93.64
Attention [**42**]	11.22	28.18	56.99	75.85	89.19
Pix2Pix [**55**]	14.19	54.24	72.25	86.86	96.61
Ours	**64.83**	**81.35**	**90.68**	**96.40**	**98.72**

4 Conclusion

In recent years, the craze of artificial intelligence has swept the world, and many excellent and iconic achievements have been achieved. However, the development of AI has to rely on huge databases. In AI computing, the quality of the database sometimes outweighs the importance of the algorithm itself. At present, there is a big problem in the database: how to update the samples so that the computer can recognize the old samples and the new samples at the same time, which has become a challenging and significant research topic and task.

In order to solve the above problems, people start to zero-sample learning. The so-called zero sample learning is to identify the categories that do not exist in the training set. For example, in order to identify cats, dogs, and pigs, it is necessary to provide a large number of pictures of cats, dogs, and pigs for model training, and then given a new picture, it is possible to determine which category of cat, dog, or pig belongs to [7]. However, for the categories of cattle and tigers that did not appear in the previous training pictures, this model could not identify the cattle and tigers. In the past decade or so, people have begun to work on zero-sample learning, including the basic definition of the subject, the evaluation of algorithm performance, algorithm improvement and innovation. However, most of the previous zero-sample learning algorithms only test the new categories, and there is no exact evaluation of the old categories. That is, when testing, the old categories of cat, dog and pig are not tested, but the new categories of cow and tiger are tested. This does not fit the real-life scenario of both old and new categories being tested. Therefore, people further put forward the more realistic concept of generalized zero-sample learning, that is, the old category of assessment was added to the test, and the cat, dog, pig, cow and tiger were put together for the test.

The main method of zero-sample learning is the connection between visual features and semantic features. Generalized zero-sample learning also forms three methods to solve this problem on the basis of previous research work. The first is the mapping from visual features to semantic features. In this method, visual features are generally extracted through simple features, then mapped to semantic features through the full connection layer, and finally, the target recognition is realized. However, the mapping of information from high latitude to low latitude inevitably leads to the loss of information. In order to alleviate this problem, a second method is proposed, that is, the mapping of semantic features to visual features. The method mainly maps semantic features to visual features through data generators. The third method is the cross-mapping of two characteristic variables. The typical method is to map the two variables to a hidden space. In the same hidden space, the connection between the two characteristics is realized [7].

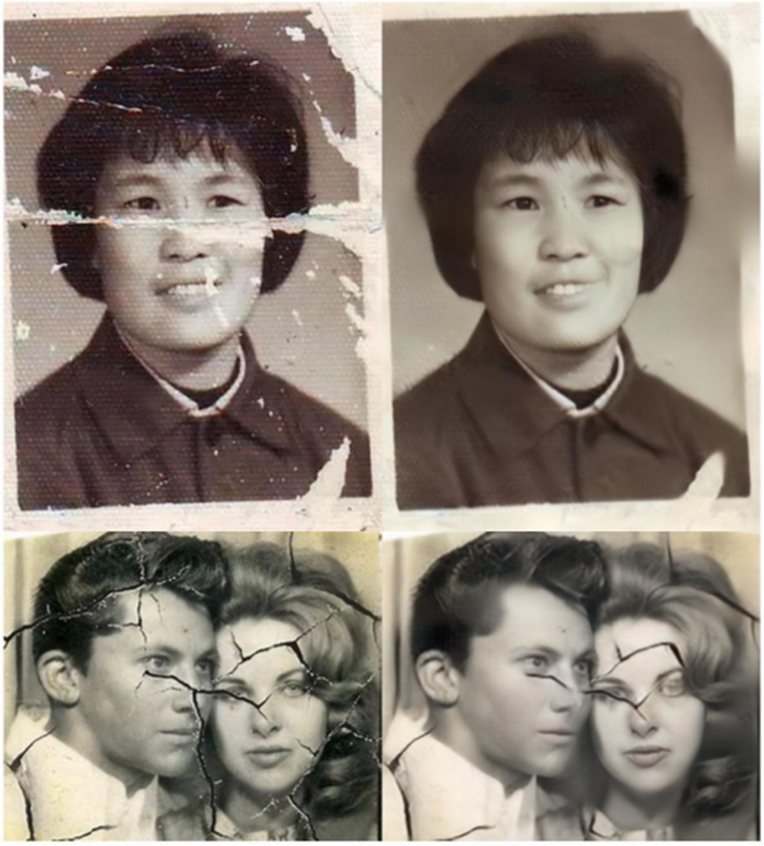

Fig. 3. shows two photos after restoration using our training models

Therefore, to sum up, there are few methods for generalized zero sample identification at present, and the accuracy and speed of the existing relevant identification methods are not good, and a large number of methods appear to have uneven accuracy and comprehensive performance evaluation results. Figure 3 shows partial outcomes after the implementation of VAE.

Old photos restoration using VAEs is a significant method that can successfully restore many old photos. It provides a valid result that shows it outperforms many state-of-art methods. Through reproduction, this work confirmed their result. However, the method fails to restore photos containing significant starches and discards some important details. Although this work tried to adjust some parameters in the code, no significant changes were made.

Reference

1. Wan, Z., Zhang, B., Chen, D., Zhang, P., Chen, D., Liao, J., & This work n, F. (2020). Bringing old photos back to life. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
2. Microsoft. (n.d.). Microsoft/bringing-old-photos-back-to-life: Bringing Old Photo Back to life (CVPR 2020 oral). GitHub. Retrieved December 3, 2022. medium.com [Internet Source]
3. Yeung Gingfung; Borowiec Damian; Yang Renyu; Friday Adrian; Harper Richard; Garraghan Peter (2022) Horus: Interference-Aware and Prediction-Based Scheduling in Deep Learning Systems. IEEE Transactions on Parallel and Distributed Systems, the first season, 53-59
4. Hasmat Malik; Smriti Srivastava (2021) Digital transformation through advances in artificial intelligence and machine learning, 103-108
5. Tao Bai; Pejman Tahmasebi (2022) Sequential Gaussian simulation for geosystems modeling: A machine learning approach, Geoscience frontiers, the first season, 5-7
6. N. Chandrasekaran; Radhakhrishna Somanah; Dhirajsing Rughoo; Raj Kumar Dreepaul; Tyagaraja S. Modelly Cunden; Mangeshkumar Demkah (2019) Digital Transformation from Leveraging Blockchain Technology, Artificial Intelligence, Machine Learning and Deep Learning, International Conference on Information Systems Design and Intelligence Applications, 12-13
7. Singh Astha; Prakash Shiv; Kumar Ankit; Kumar Divya (2022). A proficient approach for face detection and recognition using machine learning and high-performance computing, Concurrency and computation: practice and experience, the third season, 1-3