IJITEB
Vol. 5
No. 1
2023

International Journal of Information Technology and Business

http://ejournal.uksw.edu/ijiteb

Conditional GAN Approach for Colorize Image on Potato Leaf

Radius Tanone [a,*]

[a] Department of Information Management, Chaoyang University of Technology, Taichung, Taiwan

Abstract: In taking pictures using a camera, for example, it often results in poor image quality. The resulting image quality can be in the form of black and white images. This can interfere with further processing of the image. Specifically, this can happen in the agricultural sector, such as the results of a bad potato leaf image (black and white). In fact, the images of leaves from agricultural land will be used for image processing which can help with the planting process, for example. This is of course very disturbing, so a way is needed to translate a gray scale image into an image that has a color resembling the shape of an actual leaf. In fact, the development of Deep Learning with various models has grown rapidly, one of which is the Conditional GAN which can process image to image translation. Seeing the purpose of the Conditional GAN, this paper implements the Pix2Pix model based on the Conditional GAN which aims to translate black and white images into images that have color. This experiment produces images with colors that match the original image with good quality. So that with the results of this experiment, it is hoped that problems with taking images that are sometimes not good enough can be solved by image translation using deep learning.

Keywords: Pix2pix, conditional GAN, image to image translation, colorize image, potato leaf

1. Introduction

Image processing is a branch of computer science that is expanding. Various methods are still being developed to achieve the best results when processing an image. Many researchers have studied concepts such as image fusion [1], image dehaze [2], [3], and so on. All of this is done to address the issues that arise in an image. In other hand, other areas of computer science, such as Deep Learning (DL) [4], have become a better solution for image processing.

Artificial Intelligence (AI) [5] is a research area in the field of computers that is currently advancing. Part of AI is Machine Learning (ML) [6] and DL [7], [8]. Particularly for DL, many methods have been developed to solve the image problem and one of them is Generative Adversarial Networks (GAN) [9]. The Generative Adversarial Network (GAN) is an artificial neural network architecture that aims to create completely new data from nothing. In general, image data is the primary target of GANs. In summary, the GAN network is trained to generate a new image based on a set of images seen previously during the training process.

In addition, research on this model is still intensifying. Image translation is one of the components of GAN. Image-to-image translation is the process of converting an image from one domain to another with the aim of determining the mapping between an input image and an output image. Traditionally, this task was solved using a training set of matching image pairs. Isola and others [10] attempt to create a new image using image-to-image translation using the ground truth image as a basis. Image-to-image translation is a technique used in GANs to transform an image from one domain to another, to learn the mapping between an input image and an output image. Traditionally, this task has been

* Corresponding authors
e-mail addresses: s11014903@gm.cyut.edu.tw

accomplished by using a training set of aligned image pairs. However, Isola et al. attempted to generate a new image based on the ground truth image, using image-to-image translation. The Pix2Pix GAN [11] is an approach that enables image-to-image translation. Pix2Pix GAN is a Conditional GAN implementation where image generation is conditional on a given image.

A degree of problem implementation utilizing this approach is necessary, considering the image processing trends covered above and DL ideas like Conditional GAN. In greater detail, the agricultural industry can benefit from this methodology. The agriculture industry presents a lot of challenges to farmers. Agriculture plants' leaves are afflicted with one of these diseases. Naturally, a number of computer science issues, like computer vision [12][13] and others, can be resolved to address this issue. The detail section still frequently experiences an image issue, though.

Leaf images are frequently taken for analysis on agricultural land by farmers or authorities. Furthermore, the intended analysis is similar to using deep learning to solve classification problems, object detection, and so on. However, image quality is an important consideration in this case. Often, the image captured does not match the actual one. For example, if the image is dark, the color of the leaves will differ from the original. Alternatively, an image captured with a camera that produces a grayscale image. This makes further image processing on agricultural land difficult.

Based on this context, a research question arises: how to process images that are not clear based on previously owned ground truth images? The goal is to create a new image that is similar to the original. To answer this question, the goal of this study is to conduct an experiment involving image translation on potato leaves. The goal is to create a new image that resembles the original. The results of this experiment are expected to be used for other purposes, such as image processing with deep learning.

More specifically, this paper will use Pix2Pix as part of image to image translation to generate potato leaves. In addition, a literature review from previous research and the concepts used to generate images will be explained in the following chapter. Furthermore, the following chapter discusses the method used and displays the results and discussion

2. Methods

2.1. Research Workflow

In conducting this research, a research workflow was created so that research goals could be achieved. In more detail, the stages can be seen in Fig. 1.

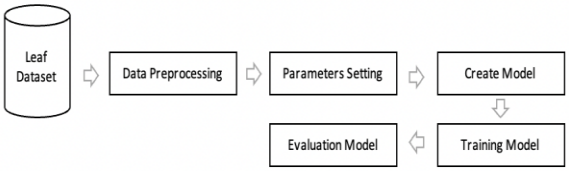


Fig. 1. Research Workflow

According to Fig. 1, the research stages begin with the preparation of datasets in the form of leaf datasets. The leaf dataset used in this experiment is potato leaf. In addition, the data is preprocessed so that it can be processed later. In preprocessing, image size is set to 256x256. Next, color images and grayscale images are the two types of leaf image data that prepared in this stage. The goal of preparing a gray scale image is to use it as an input image later on when using the Pix2Pix model. Next, create the Pix2Pix model from scratch. Following that, a training model will be run, and finally, a model evaluation will be run to see the results of the training model.

2.2. Pix2Pix Architecture

In using the Pix2Pix model for image translation, there is an architecture that must be understood in carrying out this experiment. The architecture of the Pix2Pix used can be seen in Fig. 2.

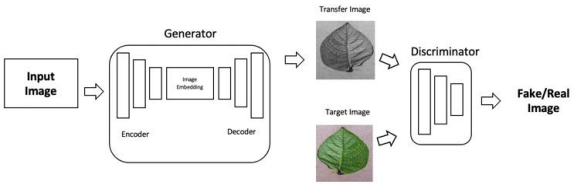


Fig. 2. Pix2Pix Architecture on Potato Leaf

According to Fig. 2, the Pix2Pix architecture for image translation begins with an input image in the shape of a potato leaf. The prepared dataset, in accordance with the preprocessing, is a gray scale image as the input image. Furthermore, the generator will use an encoder and decoder to process the input image. The image transfer and target image will then be processed before being further processed in the discriminator. The discriminator then generates an output image that resembles the input image in the form of a potato leaf.

Regarding Pix2Pix, the process can be described mathematically by first understanding the basics of GAN. The generator (G) essentially acquires knowledge of the relationship between the actual data and the random variations.

$$G: \{x, y\} -> \ y \tag{1}$$

Likewise, the discriminator (D) learns representations from labels and actual data.

$$D(x, y) \tag{2}$$

CGAN can be effectively used for image-to-image translation tasks by enabling a setting that conditions the generator on an input image. This means that the generator uses a condition distribution, such as a blueprint or guidance, to generate a target image that corresponds to the input image. This approach allows for more precise and accurate image generation in various applications.

$$L_D^{CGAN} = E[\log(D(x,c))] + E[\log(1 - D(G(z),c))] \quad (3)$$

$$L_D^{CGAN} = E\left[\log\left(D(G(z),c)\right)\right] \quad (4)$$

Furthermore, mathematically, the Pix2Pix loss function consists of two components (one for the discriminator and the other for the generator) which can be defined as follows.

$$L_{cGAN}(G,D) = E_{x,y}[logD(x,y)] + E_{x,z}[\log(1 - D(x,G(x,z)), \quad (5)$$

with the loss function for generator defined as follows.

$$L_{L1}(G) = E_{x,y,z}[||y - G(x,z)||_1] \quad (6)$$

Therefore, the final loss defined as follows.

$$G^* = \arg min_G max_D L_{cGAN}(G,D) + \lambda L_{l1}(G) \quad (7)$$

2.3. Potato Leaf Dataset

The description of the dataset we use is from Kaggle [14] and corresponds to the problem in this research, which is the identification of the type of disease in potato leaves. In total, there are 8286 images with 256x256 original sizes in the potato leaf datasets. This total is divided into three classes: Early Blight, Healthy, and Late Blight. Table I contains a detailed description of the dataset as well as the number of classes.

TABLE I. POTATO LEAF DATASET

Label of Class	Class	Example of Image	Total per class
"0"	Potato Leaf Early Blight		3149
"1"	Potato eaf Healthy		2006
"2"	Potato Leaf Late Blight		3131

After the class is divided into three sections, as shown in Table 1, the dataset must be preprocessed so that it can be used in the selected deep learning model. Furthermore, we generate a gray scale image based on a color image from the dataset so that it can be used as an input image later when applying the Pix2Pix model.

3. Result and discusion

3.1 Results

This sub-chapter contains the results of training with the Pix2Pix model. The first step is to prepare the color image and gray scale image color, as shown in Fig. 3.

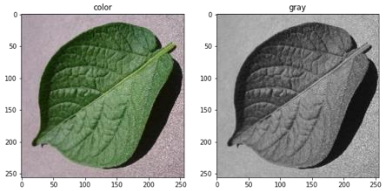


Fig. 3. Color Image and Gray Scale Image.

Furthermore, 100 epochs were used in training with the Pix2Pix model that we created from scratch. The model then generates output in the form of a generated image, as shown in Fig. 4.

(a)

(b)

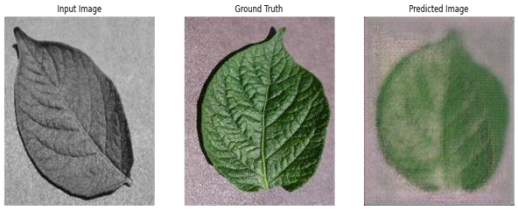
Fig. 4(a) and (b). Generated Image by using Pix2Pix Model

Fig 4. (a and b) show the Pix2Pix model results. As can be seen from this output, Pix2Pix generates an image that is identical to the original image. The gray scale image used as input can be generated correctly by the

model to look like the original image (color image). This is a perfect way for image translation implementation. The gray image is converted into an image color that is identical to the original image. Image translation can be used to solve problems such as poor image quality, a lack of datasets, and other issues.

3.2 Discussion

In this sub-chapter, we will go over important aspects of implementing the image to image translation model with Pix2Pix. The transfer image and target image must have the same shape characteristics, according to the architecture. It can be seen from the training results that if the transfer image and the target image are significantly different, the discriminator results will follow the target image rather than the input image. If this occurs, the image translation will not properly process the input image. An example of this explanation can be seen in Fig. 5.



(a)

(b)

Fig. 5(a) and (b). Different Transfer Image and Target Image

4. Conclusion

From the experiments we did using Pix2Pix for image translation on potato leaves, it can be concluded that the GAN-based model can work well. The output in the form of a generated image based on target image can be produced properly by the Pix2Pix model. From our observations, the key to generating a good image is in the transfer image and the target image which must be similar in shape so that the output is good.

Another finding from this experiment is that good parameter settings aid the generate process in producing an output image. The results of this exercise can undoubtedly provide solutions to problems such as improving image quality when translating images using DL models such as GAN. This experiment can be used in the future to duplicate datasets when processing problems in computer vision require a large number of datasets. Pix2Pix image translation may be one of the best solutions.

References

[1] C.-H. Hsieh and Q. Zhao, "Image Enhancement and Image Hiding Based on Linear Image Fusion," in *Image Fusion*, O. Ukimura, Ed. Rijeka: IntechOpen, 2011.

[2] W. Wang and X. Yuan, "Recent advances in image dehazing," *IEEE/CAA J. Autom. Sin.*, vol. 4, no. 3, pp. 410–436, 2017, doi: 10.1109/JAS.2017.7510532.

[3] C.-H. Hsieh, "Dehazed Image Enhancement by a Gamma Correction with Global Limits," in *2019 IEEE 10th International Conference on Awareness Science and Technology (iCAST)*, 2019, pp. 1–4, doi: 10.1109/ICAwST.2019.8923309.

[4] "What is Deep Learning? | IBM." https://www.ibm.com/cloud/learn/deep-learning (accessed Apr. 21, 2022).

[5] "What is Artificial Intelligence (AI)? | IBM." https://www.ibm.com/cloud/learn/what-is-artificial-intelligence (accessed Apr. 21, 2022).

[6] "What is Machine Learning? | IBM." https://www.ibm.com/cloud/learn/machine-learning (accessed Apr. 21, 2022).

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015, doi: 10.1038/nature14539.

[8] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278 – 2323, 1998, doi: 10.1109/5.726791.

[9] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," *Sci. Robot.*, vol. 3, no. January, pp. 2672–2680, Jun. 2014, doi: 10.48550/arxiv.1406.2661.

[10] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, and B. A. Research, "Image-to-Image Translation with Conditional Adversarial Networks," Accessed: Dec. 27, 2022. [Online]. Available: https://github.com/phillipi/pix2pix.

[11] "pix2pix: Image-to-image translation with a conditional GAN | TensorFlow Core." https://www.tensorflow.org/tutorials/generative/pix2pix (accessed Dec. 27, 2022).

[12] L. Deng and D. Yu, "Deep Learning Methods and Applications ," 2014, doi: 10.1561/2000000039.

[13] M. Tariqul Islam and A. N. Tusher, "Automatic Detection of Grape, Potato and Strawberry Leaf Diseases Using CNN and Image Processing," *Lect. Notes Networks Syst.*, vol. 238, pp. 213–224, 2022, doi: 10.1007/978-981-16-2641-8_20/TABLES/3.

[14] "Bangladeshi Crops Disease Dataset | Kaggle." https://www.kaggle.com/datasets/nafishamoin/bangladeshi-crops-disease-dataset (accessed Mar. 29, 2022).