Resume Paper

[redacted]

Research Area : Computer vision

Paper Title,Writer, Name of Journal/Conference and Publised Year :

Paper title : Making a Batik Dataset for Text to Image Synthesis Using Generative Adversarial Networks

Writer : Aifa Nur Amalia, Arief Fatchul Huda, Diena Rauda Ramdania, Mohamad Irfan

Name of journal /conference : 2019 IEEE 5th International Conference on Wireless and Telematics (ICWT)

Published year : 2019

Research Problem :

1. How to produce the **synthetic batik pattern** that is similar to the original without removing the **characteristics possessed** by each batik pattern using text.

Proposed Method :

Some steps that proposed by the research problems are follows :

- **Text To Image Synthesis**
- **Generative Adversarial Networks**
- **Data Collection**
- **Architecture System**

A. Text To Image Synthesis

Image synthesis basically an attempt to produce **image** of **objects**. But in this research rather than **image as the input**, **image + text were used** to be the inputs to the architecture system. In my perspective, this approach were used to make the **batik pattern** just by using the text in later application. **Before synthesize images** of an objects, the model need to understand the **pattern** and the **caption** therefore in this study **GAN was** used and some modification of the model were also applied.

B. Generative Adversarial Network

Gan consist of 2 models **Generator** and **Discriminator** in which both have a different purpose, one of the applications is **synthesize** image.

C. Dataset Collection

There are 4 steps involved in here :

- **Collect initial data**
- **Image Cropping and Image type preprocessing**
- **Adding Variations**
- **Giving Captions**

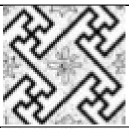
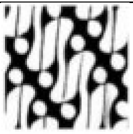
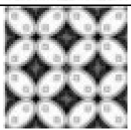
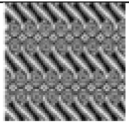
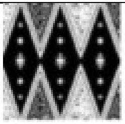
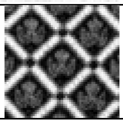
Dataset was collected from the **google Image,** and the image processing was done. The image information is shown by the table I :

TABLE I Image information

Image parameters	Value
Size	(500 x 500)
Image format	Grayscale

In my perspective, because **batik** patterns are important to generate and also to preserve it, therefore by using high resolution the **details of image were preseved. And** because the computation of **RGB** image was high therefore by just put the **the image to grayscale** will reduce the computation needed by **GAN** to generate the pattern later. Table II shows the batik patterns were used in the research

TABLE II Batik patterns

Batik pattern name	Image
Swastika/Banji	
Lereng	
Kawung	
Pilin	
Tumpal	
Ceplokan	

For each patterns, variations were applied such as, but the **motive characteristics were hold** :

- **Colors**
- **Shapes**

- **Patterns**

In my perspective, this were done because of **the color, and small variations were applied** by the batik creator to make the result more **enganging** and **beauty. Therefore,** by adding these variations, the generated **GAN** model will learn **additional** uncertainty, and that was a good approach. And last step was giving a caption. This was done because of the **research problem that was proposed by the researchers.** There are 15 variations of **captions**, again since the uncertainty were applied in real world, therefore by doing so the model will able to **generalize well** later.

D. Architecture System

There are stages that were used by the researchers :

- **Data loading**
- **Training Inference**
- **Training**

Data loading stage was to load and prepare the data needed to train the GAN model. Here are the data information used by the researchers

- **Vocab.pickle**
- **Image_train.pickle**
- **Image_test.pickle**
- **N.pickle**
- **Caption.pickle**

In my perspective, pickle was used because it is one of object that can stored information. **And also** it is making sure the **seed** / **distribution** stays the **same** for each testing in the future. **Training inference stage** purpose was to make a model learn to recognize the **objects in image and also the word recognition by using CNN for image and RNN for the word recognition.** This making sure that layers that would be used by the **generator** has a starting point . **Training Stage** was to do the actual job, by creating the generator and discriminator for creating the batik generated dataset system.

From the steps involved in the research, a lot of constraint already been taking account and the researchers seems understand the field of study, from here I can take a new insights to make GAN much better and faster result !

Experiment Result :

The System experiments were carried, and table III shows the **testing scenarios**

Table III Testing Scenarios

Scenario	Epochs	Repetition	Learning Rate
I	1000	1	0.0002
II	600	2	0.0002

No justification provided by the researchers to explain where the number comes from, and why using 2 scenarios. But in my perspective, the researcher just want to show when the model was too complex result for generating the **batik pattern.**

Table IV shows the summary of analysis result proposed by the researchers and figure 1 & 2 shows the generated images

TABLE IV Summary Scenarios

Scenario I	**Scenario II**
Batik patterns **look clear and contrast**	Batik patterns **look clearer and more contrast (Much** Better)
The appearance of baik patterns **is less uniform** dan hasil deskripsi nya kurang sesuai dengan batik pattern yang dihasilkan	The appearance of batik patterns is **more uniform** and hasil deskripsi nya sesuai dengan batik pattern yang dihasilkan
The generated batik pattern does not look significant	**The generated** batik pattern does **looks significant**

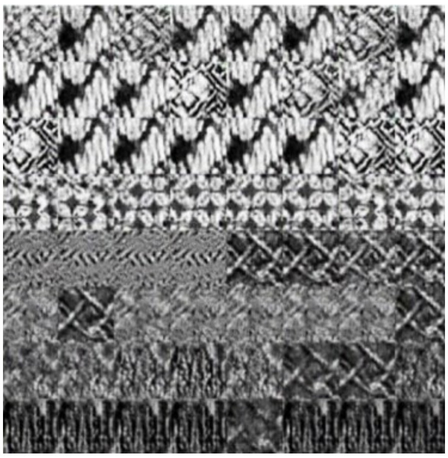


Figure 1 Scenario I Output

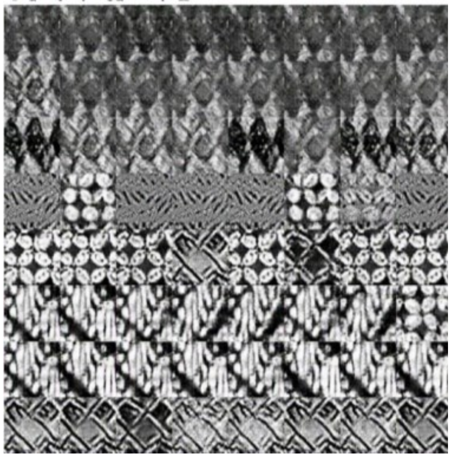


Figure 2 Scenario II Output

We can see, scenario II does have a better **result** compare with scenario I. My assumption is that because of the model too complex in scenario 1. It can be shown by the number of epochs, **1000 epochs** compared with the **scenario II.**

Also the researcher managed to do human level perfomance testing by 35 respondents. Table V shows the result of testing. 1 : similar; 2: less similar; 3:not similar

TABLE V Human Level Perfomance testing

Batik Pattern	Scenario I (average)	Scenario II (Average)
Swastika	1.63	1.86
Lereng	2.23	1.71
Kawung	1.71	1.57

Pilin	2.31	1.60
Tumpal	2.26	1.49
Ceplokan	2.43	1.43
Average	2.10	1.61

On average the result much better with scenario II, this shows that scenario II can be accepted to generate the batik pattern. But there are 3 patterns in which having > 1.5. In my opinion it is because the pattern was to complex. But it is already a great improvement to generate batik pattern

From the whole experiments, we can say that the proposed methods do a great job in making synthesize dataset, especially batik pattern in which the model too was able to capture the local pattern well. And the testing was also done by involving human into account. The result from Human Level Perfomance was promising !

Paper Contribution :

- **The contribution of this paper are as follows :**
    - It provides a better Generative solution for batik pattern compared with the previous paper []. It also shows that the computation is greatly reduced by using the grayscale method!
    - The learning rate adjusting is affecting the result of model. This is stated by researcher. Therefore, adjusting it at some points is greatly made the model learning much better.
    - Saving the data to be used in the future is greatly a good approach to **reproduce** the current result. Therefore this made the result for scenario I and II persistent.
    - The result was compared with the **Human level perfomance** in which make the result of GAN more reliable. And also the result average was **1.61** In which it was less similar with the actual pattern!

Paper Drawback/Limitation :

- **There is no Error analysis involved, only comparasion. Error analysis is important to understand the output of the model.**
- **There is no SME of batik involved in the HLP calculation.** This is important since they can look details on creating a beauty batik pattern !

Paper Title,Writer, Name of Journal/Conference and Publised Year :

Paper title : Robust Batik Image Classification using Multi Texton Co-Occurrence Descriptor and Support Vector Machine

Writer : Agus Eko Minarno, Yufis Azhar, Fauzi Dwi Setiawan Sumadi, Yuda Munarko

Name of journal /conference : 2020 3rd International Conference on Intelligent Autonomous Systems (ICoIAS)

Published year : 2020

Research Problem :

1. **How to create a realiable yet simple model for batik patterns classification ?**

Proposed Method :

The proposed method was **Multi texton Co-Occurrence descriptor**. In my opinion this was used because of batik relied on **local patterns or features that occur frequently** in images**,** just what the paper [] told. It means this was a good approach to make a reliable classification system, especially batik pattern. This method basically convert the actual input image to **textons representation. The algorithm** make sure the **local pattern** taking into account. The step involved in this algorithm are :

- **Texton extraction :** Basically **set of textons** is extracted **by using filter bank/ filters**. Each **filter** respond to specific **texture pattern (Edge for example)**
- **Texton assignment : Each pixel in the input image** is **assigned** a **texton label** based on the result of **filter** bank. **This labels** basically represent the **local texture pattern**
- **Co-occurrence Matrix computation : Basically calculate the probability of pairs of texton labels occuring together in image**. By doing so, spatial relationship between different textons can be captured. This is used to characterize the **texture pattern in the image**.
- **Descriptor computation : Statistical measures** (Entropy, contrast, and correlation) are computed **from co-occurrence matrix** to generate **a feature vector** that represents the **characteristics of the input image**.
- **Classification / Segmentation : The input were fed into the SVM/KNN in this case.**

There are 2 dataset were used in this research, table VI shows the information of each dataset

TABLE VI Dataset information

Dataset	Total images	Total classes
Batik300	**300**	50
Batik41K	**41,621**	355

The details of each class is not explained by the researchers, therefore I do not know what are 355 classes or 50 classes names are. This should be pointed by the researcher, by doing so, another researcher know what are this author done in the past, and could reproduce the research later !

Experiment Result :

From the experiment result, the researchers compared each methods for each dataset. Table VII shows the result for KNN method applied to **Batik300** dataset and table VIII Applied to **Batik41K**. While table IX and table X Shows the result for SVM method applied to **Batik 300** and **Batik41K** respectively.

TABLE VII KNN BATIK300

K	Method	Accuracy	F1	Training Time(s)	Testing Time(s)
1	GLCM	0.80	0.73	0.00	0.00
	MTH	0.89	0.89	0.00	0.01
	MTCD	**0.96**	**0.95**	**0.00**	**0.01**
2	GLCM	0.72	0.65	0.00	0.00
	MTH	0.83	0.82	0.00	0.01
	MTCD	0.92	0.89	0.00	0.01
3	GLCM	0.62	0.55	0.00	0.00
	MTH	0.79	0.78	0.00	0.01
	MTCD	0.90	0.87	0.00	0.01

TABLE VIII KNN BATIK41K

K	Method	Accuracy	F1	Training Time(s)	Testing Time(s)
1	GLCM	0.73	0.72	0.11	0.19
	MTH	0.94	0.94	0.41	0.43
	MTCD	**0.99**	**0.99**	**0.49**	**0.46**
2	GLCM	0.70	0.69	0.11	0.20
	MTH	0.94	0.94	0.42	0.50
	MTCD	0.99	0.99	0.48	0.55
3	GLCM	0.70	0.67	0.11	0.24
	MTH	0.94	0.94	0.40	0.57
	MTCD	0.99	0.99	0.49	0.64

From 2 tables above we can see that for K 1 with method **MTCD** the result was better. In my opion because of K=1 equals its own **texton representation**. It means that the model only able to get **THE closest one that having nearly same texton representation**. Although this statement does need the evidience about the **Knn outputs** but this is can be shown from the **f1 score for K > 1**. **The** model perfomance is reduced. **This is does not mean** the model **generalize the data well** in the future, since **Knn using distance**, therefore the **algorithm really affected by the outliers !**

TABLE IX SVM BATIK300

Kernel	Method	Accuracy	F1	Training Time(s)	Testing Time(s)
Linear	GLCM	0.80	0.79	0.00	0.00
	MTH	0.89	0.88	0.00	0.00
	MTCD	**0.96**	**0.95**	**0.00**	**0.01**
RBF	GLCM	0.50	0.44	0.00	0.01
	MTH	0.53	0.52	0.00	0.01
	MTCD	0.58	0.56	0.00	0.01

TABLE X SVM BATIK41K

Kernel	Method	Accuracy	F1	Training Time(s)	Testing Time(s)
Linear	GLCM	0.92	0.92	33.87	11.43
	MTH	0.97	0.97	36.29	13.65
	MTCD	**1.00**	**1.00**	**36.50**	**14.32**
RBF	GLCM	0.25	0.29	116.07	65.34
	MTH	0.48	0.50	253.57	81.68
	MTCD	0.58	0.60	256.81	86.71

From the 2 tables above IX and X we can see that the SVM with linear kernel does a great job, its mean the **representation of the feature was linear**, but without proper visualization of the actutal **textons representation and their output probability** I am having trouble to understand the outputs. Because as we know the model could be overfit !

From the whole experiments we can say that the MTCD method does a great work to classify the batik patterns dataset. But the explanations both from the proposed method to result are lacks. Therfore the reader should understand every possibilities. There are no error analysis, there are no probability distribution for the predicted class, and there are no details of the class used. From what I get, I could not interpret whether the method is actually get the essential part of **local pattern** from batik pattern without looking at the actual outputs evidiences (Image from the texton representation).

Paper Contribution :

- **It gaves a new method to preserve the** local feature from input image, this is useful especially for batik case ! Since batik mostly relies on the local pattern.
- **The method of MTCD** hopefully can be applied on the **Deep generator case !** To generate a reliable yet detail batik pattern. **Because from the evidence of this paper** MTCD does do the work of **taking a local pattern into account**.

Paper Drawback/Limitation :

- **No error analysis, therefore** we don't know which are the missclassified image from the **KNN** and **SVM** methods. If we doing so, we can make the result and more interpretable and understandable in the future when doing the reproducible research.
- **Too good to be true, with Accuracy 100% and F1-Score 100%**. There are no evidience from the researchers. They are not showing the testing dataset result in terms of image or probability distributions. Therefore, I am not sure if the result were actuallly 1.