Big Data Application (BD73548) - Final Project

MLBatikNet: Deep Learning Based Multi-label Batik Impression Classification

1 INTRODUCTION

Batik has been recognized as an element of global cultural heritage produced by Indonesians. The wide diversity of patterns reflects a variety of influences, one example such as human characters or so called impressions. Ordinary people even most common Indonesians have the affliction to understand what kind of characters that are represented in a pattern of Batik. Utilizing the deep learning methods for helping people to understand the pattern impression of Batik is challenging computer vision task. Since each pattern is not only belong to one impression, but can reflect two or more impressions.

Multi-label classification tasks which are a well-known challenging issue due to the complexity of data are sometimes also intricate to define information concerning the classes that are not mutually exclusive. In medical field, Maxwell et al. exercises Deep Neural Networks (DNN) to identify patterns in patient data for intelligent health risk prediction Maxwell et al. (2017). Wehrmann & Barros identify the movie genre using Deep Convolutional Neural Networks (ConvNets) Wehrmann & Barros (2017). One interesting previous work to identify the pattern of Batik has been developed by exercising ensemble of multi-label classifier Ramadhan et al. (2014). However, this work does not implement deep learning techniques to address the issue.

In this work, the Convolutional Neural Networks (CNN) as one prominent deep learning technique are aimed to be exercised to identify the pattern of Batik that reflects impressions. This is distinguishable issue that very limited few works that has tried to address this issue especially for multi-label classification. In order to address the above-mentioned issue, this work proposes a CNN architecture, benefiting from transfer learning of pretrained models, that is particularly devised to identify multi-label classification of Batik called MLBatikNet.

2 RELATED WORKS

Our proposed work exploits several studies in the past. The base of our neural network architecture make use of Residual Network (ResNet) He et al. (2015); Veit et al. (2016). To minimize the computing effort, we also incorporate transfer learning Goodfellow et al. (2016). The two approaches then will be modified to be adapted our multilabel batik dataset and compared to previous works on multi-label classification Ramadhan et al. (2014).

2.1 RESIDUAL NETWORK

Residual Network (ResNet) He et al. (2015) is an architecture of deep learning that employs residual module and skip connection. In some manners, it behaves like ensembles of relatively shallow networks Veit et al. (2016). The principle of the residual network is to branch the paths of the neural network by skipping a single or more convolutional layers to the intermediate output of next convolutional layer.

The performance of (ResNet) had achieved the best score in ImageNet object detection and classification problem in 2015. However, it is not been tested on smaller dataset and different task. In this study, the performance will be discussed according to multi-label batik classification task.

2.2 TRANSFER LEARNING

Simply said, transfer learning is one method to benefit pretrained model to perform different learning task (Goodfellow et al., 2016). For example, a model already trained to classify the detection of the cat in some image dataset. This model can be used to detect dogs in other image dataset with several parameter adjustments and relearning. However, the cost of the adjustments and relearning is not much as training model from zero.

The interesting part of this study is that there is currently minimum number of study of batik classification using deep learning. Thus, there is no accessible model to be used as pretrained model that corresponds to batik image features. In this study, we will discuss the performance of object detection pretrained model such as ImageNet to learn scarce multi-label batik dataset compared to training using multi-label batik dataset from zero.

3 METHODS

In MLBatikNet, we use pretrained models to overcome the work to build the model from the beginning. This work is limited to usage of transfer learning to perform the classification. The proposed pretrained model is the family of ResNet for image classification He et al. (2015) and taken from official TensorFlow site ten (2018).

To adapt our dataset, the parameter of the last layer (before the output layer) of the pretrained ResNet will be reset and the output layer will be changed to six output layers instead of one. For the activation function, instead of softmax or ReLU, which used for multi-class classification

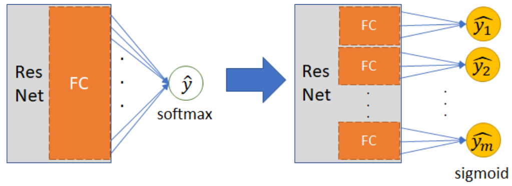

Figure 1: Our proposed change from existing ResNet architecture

Table 1: Distribution of Multilabel Batik Cloth Image Labelset

Impression labelset	Total
feminine	51
dynamic, masculine	23
mature, calm	13
mature, calm, warm	7
mature, calm, feminine	4
mature, calm, warm, feminine	2
mature, calm, warm, masculine	1
Total	102

problem, we will use sigmoid with cross entropy Maxwell et al. (2017). Our proposed change is depicted in Figure 1

4 Experimental Setup

4.1 Dataset

The provided dataset in this project are batik cloth images which consists of 102 labeled multilabel images and 11 unlabeled image. The dimension of single image is 500 x 500 pixel. The labels of the images are defined by set of six distinct impressions: mature, calm, warm, dynamic, masculine, feminine. The distribution of labelset, which is also imbalanced, can be seen in Table 1. Some of the example images can be seen in Figure 2

Regarding the limited and imbalanced dataset, which is not in deep learning's favor, several techniques of data augmentation will be configured. The data augmentation will lead to slightly smaller images but still capturing the prominent patterns of batik images. The data augmentation process is lead by random crops on the original image, several rotated images, and horizontal-vertical flipped images.

(a) (b) (c)

Figure 2: The sample batik images and their labelset: (a) dynamic, masculine, (b) mature, calm, (c) feminine

4.2 Evaluation Metric

Multilabel classification performance cannot be measured using accuracy like single label classification, thus in this project the performance is measured by two different metrics: Hamming Loss and Average Precision Tsoumakas et al. (2009); Zhang & Zhou (2007).

4.3 System Setup

This work is conducted in the system with Intel Xeon processor E5-2600 v4 @ 1.70 GHz x 16, 64 GB RAM, GeForce GT730 1 GB and OS Ubuntu 16.04.4 LTS 64 bit.

References

Models and examples built with TensorFlow, May 2018. URL `https://github.com/tensorflow/models`. original-date: 2016-02-05T01:15:20Z.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. `http://www.deeplearningbook.org`.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. URL `http://arxiv.org/abs/1512.03385`.

Andrew Maxwell, Runzhi Li, Bei Yang, Heng Weng, Aihua Ou, Huixiao Hong, Zhaoxian Zhou, Ping Gong, and Chaoyang Zhang. Deep learning architectures for multi-label classification of intelligent health risk prediction. *BMC bioinformatics*, 18(14):523, 2017.

Hani Ramadhan, Isye Arieshanti, Anny Yuniarti, and Nanik Suciati. Impression determination of batik image cloth by multilabel ensemble classification using color difference histogram feature extraction. *Kursor*, 7(4), 2014.

Grigorios Tsoumakas, Ioannis Katakis, and Ioannis Vlahavas. Mining multi-label data. In *Data mining and knowledge discovery handbook*, pp. 667–685. Springer, 2009.

Andreas Veit, Michael J. Wilber, and Serge J. Belongie. Residual networks are exponential ensembles of relatively shallow networks. *CoRR*, abs/1605.06431, 2016. URL `http://arxiv.org/abs/1605.06431`.

Jônatas Wehrmann and Rodrigo C Barros. Movie genre classification: A multi-label approach based on convolutions through time. *Applied Soft Computing*, 61:973–982, 2017.

Min-Ling Zhang and Zhi-Hua Zhou. Ml-knn: A lazy learning approach to multi-label learning. *Pattern recognition*, 40(7):2038–2048, 2007.