An Efficienti Fabric Pattern Classification Using Transfer Learning Convolutional Network

Thomaskutty Reji, Jossy George, Athishay Kesan and Andrea Rodrigues

January 27, 2022

An Efficient Fabric Pattern Classification using Transfer Learning Convolutional Network

Thomaskutty Reji*, Jossy George, Athishay Kesan, Andrea Rodrigues
CHRIST (Deemed to be University), India
* thomaskutty.reji@science.christuniversity.in

Abstract. Since, fabric classification is a challenging task in the image classification domain, helping business people to quickly and efficiently categorize fabric patterns is very much needed. Many of the online fashion portals uses various machine learning or deep learning model for automated tagging or automated image classification of images in their websites. In recent years transfer learning techniques are very helpful for building task specific models from pretrained networks thus it has received a lot of attraction among researchers and data scientist. This paper proposes an efficient Transfer Learning Convolutional Network (TLCN) for detecting the fabric pattern classes. The proposed system modifies the pretrained network architecture in terms of pooling and convolution operations to make an optimized and a robust model. Four different networks were trained on a dataset which is scrapped from many websites. Experimental analysis showed that VGG16 model performed better than other transfer learning models.

Keywords: Convolutional Neural Networks, Transfer Learning, Fabric Classification

1. Introduction

Due to the increase in the number of images on various online fashion portals, the need for an effective and efficient image classification model is important. There are many use-cases where we can use data science solutions in the fashion industry. Some of these applications include apparel classification, fabric defects detection, pose estimation etc. The task of recognizing the image of clothing is helpful of online buyers who are interested in the keyword search for fashion fabric filters. In many online web portals, we can see such kind of applications, but sometimes we get irrelevant fabric designs or unmatching clothes with respect to our input keyword. And some websites are automated the task of image tagging. All of these use cases can be approached using data science.

Among various image classification use-cases in the fashion industry, fabric classification is one of the important topics. The fabric design plays an important role in sales. We have different kinds of fabric designs and among them floral, check, brocade, chinos, dotted, stripped and zigzag are the common one. There are various machine learning classification algorithms that can be used to categorize images which includes k nearest neighbors, support vector machines, random forest etc. Convolutional networks have been widely used in large-scale computer vision tasks. So, if we have a sufficient amount of training data CNN gives better results than using other machine learning algorithms. ImageNet competition has produced various high performing classification networks like VGG, ResNet, GoogleNet etc. Moreover, the revolution of transfer learning helps us even if we have less amount of training data. [1]

We have some challenges associated with the fabric and general fashion image classification problems for instance, different spatial positions and sizes of clothes can affect the image quality and prediction. It is difficult to get a good image dataset due to these reasons. Fabric styles can be easily manipulated by stretching and folding styles, and it can be ambiguous to identity the similar fabrics. There are very smaller number of works have done on fabric image classification using transfer learning. Mainly there are two ways through which we can make of transfer learning techniques. Some of the existing works modifies the original networks which is trained on ImageNet dataset. The next option is to make a hybrid model combining combing multiple techniques or algorithms. [2]

In this paper, both modified pre-trained networks and hybrid models are used to classify fabric designs, where the model receives an input image of fabric and the output will be the probabilities of being a distinct category in the training data. These probabilities can be further used to get the final prediction label. The existing networks are modified by adding multi-pooling technique and additional convolutional layers. The hybrid model used the pretrained network for feature learning and finally, classified the fabric using support vector machines. This work will try to find the best performing pre-trained network for the fabric classification use-cases. An entirely new fabric image dataset is scrapped from websites which includes seven categories' images, exhibit high variations.

A. Contribution: In this paper the Transfer learning CNN with additional convolutions and multiple pooling is introduced. The model is trained on a new fabric pattern dataset which consists of seven fabric designs. The model can be used to classify fabric images into one of the seven categories.
B. Organization: The introduction is given in section 1, the existing research papers are discussed in section 2, the proposed model is explained in section 3, the algorithm in section 4, the performance analysis is discussed in section 4 and finally the conclusion is given in section 5.

2. Related Works

Feng et al. [3] improved the performance of the famous Alexnet network by applying log ReLU activation and binary hash coding. They used the original architecture and optimized the entire network by applying modifications. Fabric categories of Image dataset is used to train the network. The accuracy was improved by around 5% than the original Alexnet. The work concluded that Log-Alexnet for classification of clothing fabric design is feasible and much efficient than the traditional state-of-the-art methods.

Performance of different CNN architectures were compared with respect to Fashion MNIST dataset by Saiharsha et al. [4]. Among all the architectures VGG like architecture which contains dense neural network with 16 or 19 layers, and 138 million parameters, performed well and got 90.77% test accuracy with normalization and various fine-tuning methods like dropouts, number of nodes in the hidden layers etc. ReLU activation was used for the five conv2D layers and for the last layer SoftMax activation was used. They applied dropout of 50% to avoid overfitting. But the network is computationally intensive and took high cost of training time. Keras library is used for implementing the proposed method

In transfer learning we need apparently less time to train the network and the resultant model is more generalizable. Li et al. [5] created Inception V3 model and VGG 19 transfer learning models, for fashion classification with an input image size of 224*224 and the training data consist of 444424 fashion images. The pre-trained models are accessed from the Keras library in Python. The only pre-processing includes the standardisation of pixels. Among the two models, VGG 19 achieved highest accuracy of 95%. For the VGG-19 network involves 3*3 stacked conv layers with pooling and stride of 2. Moreover, they added 5 extra layers to VGG. For the Inception they added 3 more layers to the network. The second part of the work consists of creating a visual search engine based on CNN autoencoder and ResNet autoencoder. For testing the visual search models, they use cosine similarity.

Boriya et al. [6] used image classification and image clustering for creating a visual search engine specifically for fashion images (ViSeR). ViSeR works in three phases, in the first step, image is transferred through deep neural network and predicting the top-n classes. In the second phase, features are extracted using neural networks and other computer vision techniques. In the last stage similar images were extracted based on the features obtained from the second phase and using other similarity metrics like cosine similarity or Manhattan distance similarity. In the first step they tried three neural networks namely, MobileNet, VGG16, VGG19 and in the second step they tried seven machine learning algorithms like, random forest, stochastic gradient descent, decision trees, k nearest neighbour, naïve bayes and extra tree models. They got the best accuracy of 84% when using a combination of MobileNet features with SGD classifier. They used Deep Fashion dataset for training the model and it has around 250k images with 18 categories.

Chen et al. [2] implemented clothing image classification using five different CNN architectures which includes conventional CNN, CNN with inception modules, Networks which contain both residual and inception blocks and two transfer learning methods. The conventional CNN composed of five convolutional blocks and three pooling layers. The second network which is CNN with Inception contains an inception module, which improves the classification accuracy moreover it computationally inexpensive and parameters than VGGNet and Alexnet. The inception module contains three branches where the first and second branch contains two convolutional layers and the third step is the max pooling operation. In the third network, they replaced convolutional blocks with residual block. For transfer learning methods they use Inception V3. They proved that transfer learning methods able to attain high accuracy of 92.02% and 89.25%. They used clothing image dataset which consists of 82631 samples.

Seo and Shin [7] used pretrained Google Net transfer learning model for classification of fine-grained fashion images. Pretrained models and transfer leaning techniques are heavily used if we have less amount of data and to shorten the training time. They have tuned the pretrained model on fine grained fashion dataset after removing the last fully-connected layer of GoogleLetNet, based on design attributes and achieved a final test accuracy of 62%. The dataset consists of just 1392images and the classification was on 24 categories. In the training part they implemented vanilla GoogleLetNet which is based fundamentally on Inception module and contains convolutional layers, average pooling layers, fully connected layers and SoftMax activations. Because of the pretraining and fine tuning they significantly reduced the training time.

Li et al. [8] presented two stream multi-task neural network for fashion image classification. They used DeepFashion dataset consists of 209222 images for training and 40,000 for the validation and 40,000 for testing purposes and used ResNet50 pretrained model for the classification. Their overall pipeline was consisting of landmark detection network, structure awareness and boundary awareness. For the landmark detection they deployed hourglass architecture which helps to extract robust structural representations. Boundary awareness detects the target edge and mark the line between selected landmarks and draw the target boundary to generate the attention map. In structure awareness feature maps are concatenated from the end of the landmark categorization network with the middle-level feature maps in fashion classification model. The suggested method outperformed for fashion landmark detection, category classification and attribute recognition.

Sarwo et al. [9] used simplified resnet50 backbone for logo detection and brand recognition. They showed that ResNet based backbone increased classification accuracy with less time. ResNet has got different variations such as resnet18, resnet50, resnet34, resnet50 etc and the main disadvantage of these models are they require more training time and the deep structure. So, they created a simplified version of ResNet with a smaller number of model hyper parameters in order to reduce the training and classification time but without sacrificing the accuracy. ROMYNY Logo 2016 dataset which consists of 20 categorizations were used for the training and implementation. The proposed network attained mean average precision of (0.408+-0.1050) with average training time of one hour and forty-one minutes which is less than the usual ResNet variants.

Multiple-Loss dual output CNN was implemented by Stephen et al. [10] in 2019. They created two convolutional blocks in the network where the first block is responsible for the feature extraction and determination of classes. It consists of 5-layer subblocks with dropout, pooling, batch normalization, regularization and padding and striding functions. Second block is for getting the colour features and it contains 4 smaller subblocks. Each block contains its own loss functions. In total the network model has 5,605,255 total parameters, and out of these parameters 5603,495 is trainable and 1760 are non-trainable. L2 Regularization was also applied in order to add extra penalty. 100 epochs were used with smaller batch size for improving the training time and computation. They have achieved 98% and 96 % for fashion and colour respectively. The training data set consists of 2167 fashion images and 6 categories (Deep Fashion dataset). So, the proposal showed that how classifications can be learned efficiently to improve the performance of the two tasks.

3. Proposed Model

The data set for the training is scrapped from websites using python beautiful soup and selenium libraries. It has a total of more than 4200 images, covering seven fabric classes include floral, check, brocade, chinos, dotted, stripped and zigzag. Each subdirectory of fabric classes contains more than 650 images. The entire images are splitted into train and testing folders. The images are scaled using Keras standardizer and generated 10 augmented samples. The final size of the processed image will be 256*256*3.

Figure 1 represents the entire model workflow, which includes the stages from data extraction to the final model predictions.

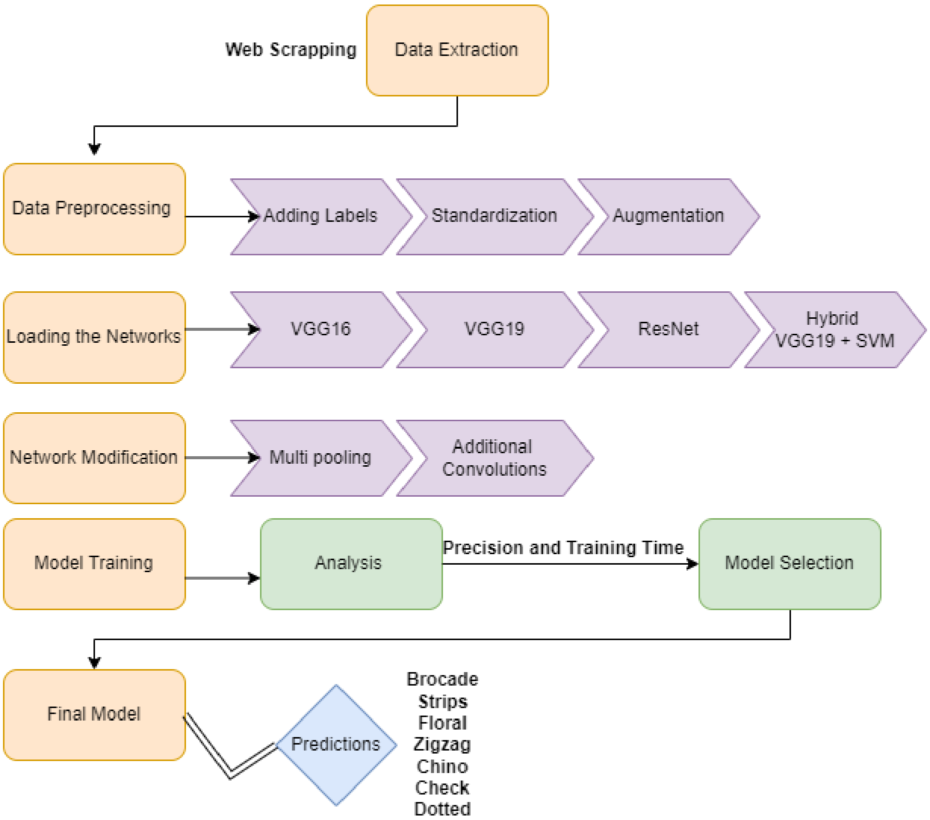

Figure 1: Model Workflow

Four different networks were trained based on the preprocessed training data. Three of them directly uses the pre-trained network with modifications in the convolutional layers and pooling. The fourth one is a hybrid model of VGG19 and Support Vector Machines. In this model we learn the feature representation using VGG19 and for the classification part we use Support vector machines. SVM takes the data points and separates them using a hyperplane. For reducing the complexity and limitations of the traditional computer vision methods, this paper proposes an efficient VGG16 framework with hidden layers for the fabric pattern classification. Batch Normalization and dropout methods are used to avoid overfitting. Moreover, two additional convolutional layers are added with multiple pooling. In multiple pooling we concatenate the minimum and maximum together. The additional convolutional layers use 32 filters of size of 3*3 with Relu activation. After the feature extraction part, the one single fully connected layer of 32 nodes is used followed by the flattening layer. The pre-trained networks' learned weights are freezed during the training time. For the additional convolutional layers, the dropout is set to 0.2. The networks with modifications will be trained in the first phase, then we see the difference in performance. For all the model Adam optimizer is used for during the model compilation. Four evaluation metrices were used for the final model selection namely accuracy, recall, precision and f1 score. Testing data is passed as the validation data in the training period.

For the loading the image and the data augmentation Keras image data generator is used. The pre-trained networks are available in Keras applications. Python beautiful soup and selenium are used for the image scrapping. For the data visualization matplotlib library is used. TensorFlow framework and python is used to implement the fabric pattern classification model. In the fabric images 80% of the data is used as the training set and the rest as the testing set. Though more than twenty epochs are fairly sufficient for learning the feature representations, we trained all the models for thirty epochs.

4. Experimental Results and Analysis

Four different accuracy metrices such as accuracy, precision, f1 score and recall were used for the model selection. All the networks are trained for 30 epochs with Adam as the optimizer. For the hybrid model we selected the hinge loss as the loss function and for other models' categorical cross entropy is used as the loss. For the final model selection six evaluation criteria were used namely loss rate, accuracy, recall, f1 score, precision and training time. Precision refers to the ratio of predicted positives to the actual positives. Since fabric pattern identification is a problem statement where we need high precision rate, we use precision as the first criteria in the model selection. The second criteria will be test loss. Table 1 shows the accuracy metrices of each model in the training phase.

Table 1. Model Train Accuracies

Model	Loss	accuracy	Recall	F1	Time	Precision
VGG16	0.31	0.90	0.88	0.90	43.3	0.92
VGG19	0.32	0.89	0.87	0.89	43	0.91
RESNET	1.13	0.60	0.45	0.56	44.3	0.77
VGG19+SVM	0.28	0.91	0.89	0.91	66	0.93
VGG16-C	0.12	095	0.95	0.95	50.73	0.96

In any Machine learning or deep learning the final model selection will be taken after analysing the test accuracy. Table 2 shows the various test evaluation metrices of five models. Test accuracies usually gives an estimate for the generalization error of any machine learning model.

Table 2. Model Test Accuracies

Model	Loss	accuracy	Recall	F1	Precision
VGG16	0.56	0.84	0.83	0.84	0.85
VGG19	0.73	0.81	0.79	0.81	0.84
RESNET	1.06	0.66	0.58	0.66	0.76
VGG19+SVM	0.71	0.84	0.84	0.85	0.84
VGG16-C	0.73	0.81	0.81	0.82	0.81

From the Table 1, VGG like architecture achieved above 90% training precision with optimizer as Adam and 30 number of epochs. A maximum test precision of 85 percent is achieved by using around 4000 images. The metrics which is showed in the above tables are correspond to the last epoch.

Figure 2 shows the performance difference between vgg16 original and the vgg16 modified with additional convolutions and the multiple pooling.

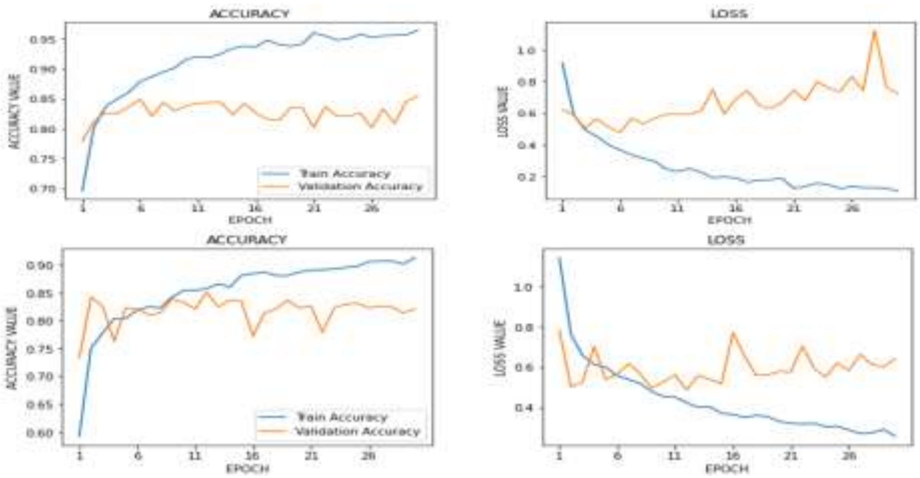

Figure 2: VGG16 Original (top) and VGG16 Modified with Convolutions and Pooling(down)

From the Figure 2 of vgg16 original and the modified it can be observed a significant decrease in the validation accuracy without sacrificing the validation accuracy. From the experimental analysis it could see that the precision value and other measures are almost similar for the hybrid model and the VGG16. But considering the test loss VGG16 is much better. From Table 2, the classification score obtained by ResNet is lower compared to the other models. The test loss is also much higher for the ResNet model. We could see that the method we proposed has sufficiently improved the classification precision. The following Figure 3 shows the comparative analysis of different networks in terms of evaluation criterion.

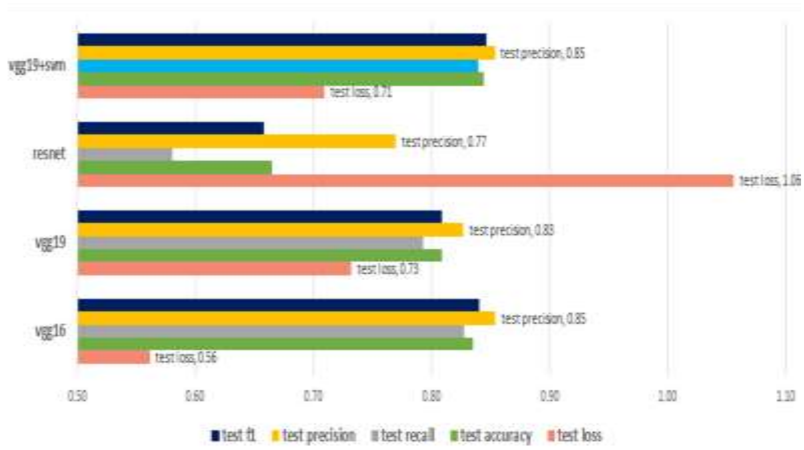

Figure 3. Model Comparative Analysis.

It is clearly evident that the loss of ResNet architecture is the worst among all other models. We can also see that the hybrid and the VGG16 attained almost same level of accuracies. Even though we trained the network for 30 epochs the highest precision obtained is 86%. This is mainly due to the less training data. So, if we improved the training dataset, we may get better result than this. Compared to classical VGG16, the modified network performed better. Transfer learning process is less time consuming. It takes around one hour even for the high depth networks. From the results of VGG 16 network without additional layers and multiple pooling we can see that accuracy and precision is reduced by around 4%. This significant difference shows that the proposed network with additional modifications performed better and based on various evaluation metrices we can see that VGG 16 is the best model for the use case of fabric pattern classification.

5. Conclusion

A fashion fabric pattern dataset is made with around 4200 images. Based on this dataset, a new fabric pattern classification model using transfer learning technique is created by improving by adding new convolutional layers and multiple pooling. The four different models were trained and the results showed that the VGG16 model with the multiple pooling and additional two convolutional layers outperformed the other networks and received the high precision of 85% in the testing phase. The selected model exhibits lowest training time among others. The proposed network shows high performance if we have more amount of training data. Future directions of this work include creating various fabric patterns using generative adversarial networks and building a system for automatically detecting the fabric pattern designs in social medias thereby we can know the current trends and customer likes and dislikes.

References

[1] Y. Xu, W. Liu and F. K. Kelly, "Compressed Domain Image Classification Using a Dynamic-Rate Neural Network," *IEEE Access,* vol. 8, pp. 217711-217722, 2020.

[2] L. Chen, R. Han, S. Xing and S. Ru, "Research on Clothing Image Classification by Convolutional Neural Networks," *Proceedings - 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2018,* vol. 2019, pp. 1-5, 2019.

[3] M. Feng, D. Chen and X. Wang, "Classification of Fabric Patterns Image Based on Improved Log-AlexNet," *Proceedings - 2018 11th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics, CISP-BMEI 2018,* no. 61572381, pp. 1-6, 2019.

[4] S. B, A. L. A, D. B, K. R and G. M, "Evaluating performance of deep learning architectures for image classification," *Proceedings of the 5th International Conference on Communication and Electronics Systems, ICCES 2020,* no. Icces, pp. 917-922, 2020.

[5] F. Li, S. Kant, S. Araki, S. Bangera and S. S. Shukla, "Neural Networks for Fashion Image Classification and Visual Search," *SSRN Electronic Journal,* 2020.

[6] A. Boriya, S. S. Malla, R. Manjunath, V. Velicheti and M. Eirinaki, "ViSeR: A visual search engine for e-retail," *Proceedings - 2019 1st International Conference on Transdisciplinary AI, TransAI 2019,* pp. 76-83, 2019.

[7] Y. Seo and K.-s. Shin, "Image classification of fine-grained fashion image based on style using pre-trained convolutional neural network," *2018 IEEE 3rd International Conference on Big Data Analysis, ICBDA 2018,* pp. 387-390, 2018.

[8] P. Li, Y. Li, X. Jiang and X. Zhen, "Two-Stream Multi-Task Network for Fashion Recognition," *Proceedings - International Conference on Image Processing, ICIP,* Vols. 2019-Septe, pp. 3038-3042, 2019.

[9] S. S, Y. Heryadi, E. Abdulrachman and W. Budiharto, "Logo detection and brand recognition with one-stage logo detection framework and simplified resnet50 backbone," *2019 International Congress on Applied Information Technology, AIT 2019,* no. 1, 2019.

[10] O. Stepen, U. J. Maduh, S. Ibrokhimov, K. L. Hui, A. A. Al-Absi and M. Sain, "A Multiple-Loss Dual-Output Convolutional Neural Network for Fashion Class Classification," *International Conference on Advanced Communication Technology, ICACT,* Vols. 2019-Febru, pp. 408-412, 2019.

[11] A. J. Medeiros, L. Stearns, C. Chen and E. J. Froehlich, "Recognizing clothing colors and visual T extures using a finger-mounted camera: An initial investigation," *ASSETS 2017 - Proceedings of the 19th International ACM SIGACCESS Conference on Computers and Accessibility,* pp. 393-394, 2017.

[12] J. Xiang, T. Dong, R. Pan and W. Gao, "Clothing Attribute Recognition Based on RCNN Framework Using L-Softmax Loss," *IEEE Access,* vol. 8, pp. 48299-48313, 2020.

[13] S. C. Hidayati, K. L. Hua, Y. Tsao, H. H. Shuai, J. Liu and W. H. Cheng, "Garment Detectives: Discovering Clothes and Its Genre in Consumer Photos," *Proceedings - 2nd International Conference on Multimedia Information Processing and Retrieval, MIPR 2019,* pp. 471-474, 2019.

[14] Y. Zhang, P. Zhang, C. Yuan and Z. Wang, "Texture and shape biased two-stream networks for clothing classification and attribute recognition," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 13535-13544, 2020.

[15] A. Chopra, A. Sinha, H. Gupta, M. Sarkar, K. Ayush and B. K, "Powering robust fashion retrieval with information rich feature embeddings," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops,* Vols. 2019-June, pp. 326-334, 2019.

[16] T. H. Tran, H. X. Tran, T. V. Nguyen and K. N. An, "Building an Automatic Image Tagger with DenseNet and Transfer Learning," *Proceedings - 2019 International Conference on Advanced Computing and Applications, ACOMP 2019,* pp. 34-41, 2019.

[17] S. P. Kannojia and G. Jaiswal, "Ensemble of Hybrid CNN-ELM Model for Image Classification," *2018 5th International Conference on Signal Processing and Integrated Networks, SPIN 2018,* pp. 538-541, 2018.

[18] M. Ebrahim, A. Ayyoub and M. A. Alsmirat, "Will Transfer Learning Enhance ImageNet Classification Accuracy Using ImageNet-Pretrained Models?," *2019 10th International Conference on Information and Communication Systems, ICICS 2019,* pp. 211-216, 2019.

[19] C. Lu, D. Li and D. Zeng, "Logo Detection Based on Convolutional Neural Networks," *ICSIDP 2019 - IEEE International Conference on Signal, Information and Data Processing 2019,* 2019.

[20] A. Mahajan and S. Chaudhary, "Categorical Image Classification Based on Representational Deep Network (RESNET)," *Proceedings of the 3rd International Conference on Electronics and Communication and Aerospace Technology, ICECA 2019,* pp. 327-330, 2019.

[21] S. Shubathra, P. Kalaivaani and S. Santhoshkumar, "Clothing Image Recognition Based on Multiple Features Using Deep Neural Networks," *Proceedings of the International Conference on Electronics and Sustainable Communication Systems, ICESC 2020,* no. Icesc, pp. 166-172, 2020.

[22] J. Zhang and Z. Feng, "Inception densenet with hybrid activations for image classification," *2019 6th International Conference on Systems and Informatics, ICSAI 2019,* no. Icsai, pp. 1295-1301, 2019.

[23] J. Martinsson and O. Mogren, "Semantic segmentation of fashion images using feature pyramid networks," *Proceedings - 2019 International Conference on Computer Vision Workshop, ICCVW 2019,* pp. 3133-3136, 2019.

[24] T. Yamamoto and A. Nakazawa, "FASHION STYLE RECOGNITION USING COMPONENT-DEPENDENT CONVOLUTIONAL Artificial Intelligence Laboratory , Fujitsu Laboratories Ltd ., Kawasaki , Japan Graduate School of Informatics , Kyoto University , Kyoto , Japan," *2019 IEEE International Conference on Image Processing (ICIP),* pp. 3397-3401, 2019.

[25] M. Inoue, C. H. Forster and d. A. C. Santos, "Semantic Hierarchy-based Convolutional Neural Networks for Image Classification," *Proceedings of the International Joint Conference on Neural Networks,* 2020.

[26] F. H. Huang, H. M. Lu and Y. W. Hsu, "From Street Photos to Fashion Trends: Leveraging User-Provided Noisy Labels for Fashion Understanding," *IEEE Access,* vol. 9, no. 49189-49205, 2021.

[27] S. Das, A. Wahi, S. S, T. N and K. S, "Classification of knitted fabric defect detection using Artificial Neural Networks," *Proceedings of the 2019 International Conference on Advances in Computing and Communication Engineering, ICACCE 2019,* 2019.

[28] C. Y. Hsu, C. C. Tseng, S. L. Lee and B. Y. Xiao, "Image Classification Using Convolutional Neural Networks with Different Convolution Operations," *2020 IEEE International Conference on Consumer Electronics - Taiwan, ICCE-Taiwan 2020,* vol. 2, pp. 20-21, 2020.

[29] R. Aoki, T. Nakajima, T. Oki and R. Miyamoto, "Accuracy improvement of fashion style estimation with attention control of a classifier," *IEEE International Conference on Consumer Electronics - Berlin, ICCE-Berlin,* Vols. 2019-Septe, pp. 289-294, 2019.

[30] Y. Dafu, "Classification of Fabric Defects Based on Deep Adaptive Transfer Learning," *Proceedings - 2019 Chinese Automation Congress, CAC 2019,* pp. 5730-5733, 2019.