Assorted Techniques for Defining Image Descriptors to Augment Content Based Classification Accuracy

Rik Das, Mohammad Arshad, Pankaj Kumar Manjhi

*Abstract***:** *Image data has turned out to be a significant means of expression with the advancements of digital image processing technologies. Image capturing devices has now transformed to commodities due to smart integration with cell phones and other useful devices. Huge amount of images are getting accumulated daily in gigantic databases which requires categorization for prompt retrieval in real time. Content based image classification (CBIC) thus gained it's popularity in classifying images to their corresponding categories. Feature extraction techniques are the foundation of CBIC to represent the image data in the form of feature vectors. This work has implemented three different feature extraction techniques from spatial domain, transform domain and deep learning domain. The three different feature vectors feature vector are contrasted to investigate the robustness of descriptor definition for content based image classification*

*Keywords***:** *binarization, image transform, pretrained CNN, feature extraction, image classification*

I. INTRODUCTION

Classification of images based on content of the image data has revealed remarkable growth in recent times due to improved techniques of feature extraction and robust machine learning algorithms. Availability of open source architecture of neural networks for deep learning has further leveraged the research in the aforementioned domain [1]. Prior to automated feature extraction using deep neural networks, traditional techniques of handcrafted feature extraction has contributed immensely towards progressive development of content based image classification [2]. However, feature engineering has a major role in formulating robust content based features using hand crafted techniques [3]. In contrast, feature extraction using deep learning techniques hardly require any manual intervention for revealing credible feature patterns from content based image data. This work has carried out feature extraction using three different techniques. Two of the techniques are handcrafted ones and the last one is based on pretrained convolutional neural network [4]. Feature extraction using handcrafted techniques embrace two popular methods, namely,

* Correspondence Author

Rik Das*, PGPM – Information Technology, Xavier Institute of Social Service, Ranchi, India. Email: rikdas78@gmail.com

Mohammad Arshad, MCA Department, Vinoba Bhave University, Hazaribag, India. Email: arshadnel@gmail.com

Pankaj Kumar Manjhi, University Department of Mathematics, Vinoba Bhave University, Hazaribag, India. Email: 19pankaj81@gmail.com

binarization technique and image transform techniques. Both the techniques are highly effective to extract meaningful descriptors from image content. Bernsen's technique of local threshold selection for image binarization is one of the familiar techniques for extraction of features by separating the foreground of the image from it's background by selection of local threshold within a given window . Another useful technique for image descriptor definition is to use fractions of transform coefficients as feature vectors for classification purpose. In this work, slant transform is used for the purpose of feature vector generation using transform technique [5]. The process of selection of effective fractional coefficient is instrumental in reducing the dimension of image features extracted using transform technique. Finally, a pretrained convolutional neural network (*VGG 16*) is used for extraction of content based features from image data [6].

A well known open access dataset is used as test bed for the experimentation purpose. Extracted features are evaluated for classification accuracy by means of two diverse classifiers, specifically, support vector machine (SVM) and random forest. The results have revealed superior categorization of images with descriptors extracted using VGG 16 in contrast to traditional handcrafted method.

II. LITERATURE REVIEW

Mainstay of CBIC is principally reliant on significance of descriptor representation. Several techniques are designed to extract features manually in which feature engineering has a significant stake. Binarized statistical features are extracted using three different filters from iris images in [7]. Magnetic resonance imaging binarization is carried out in [8] that have computed the threshold based on mean, variance, standard deviation and entropy. Extremely randomized trees are used for binarization of degraded document images in [9]. SIFT features are transformed to binary representations in [10] for fast image retrieval.

Transform techniques are crucial for energy compaction in images. This attribute of the transform techniques can be well utilized to extract meaningful content based image features. Slant transform is used for feature extraction for iris image recognition [11]. A novel algorithm for feature extraction is designed in [12] using slant transform. Extended classification results are achieved using transform techniques for feature extraction in [13]. Fusion is carried out at score level and feature level using fractional coefficients of image transforms [14]. Local feature extraction of video data using

pretrained CNN is carried out in [15]. Effect of feature representation for selection of dictionary for video summarization is studied and carried out by means of deep feature extraction [16]. Learning of semantic feature map is also performed using automated feature extraction [17].

This work has carried out feature extraction using all the three techniques, namely, binarization, image transformation and pretrained CNN based feature extraction. The classification results with extracted features have shown improved performances compared to existing literature.

III. OUR TECHNIQUES

Three different techniques of feature extraction are implemented in this work. The following section has described each of the techniques in brief initiating with the binarization technique followed by image transform technique and automated feature extraction using pretrained CNN technique.

A. Defining Descriptor using Image Binarization

A local threshold based technique named Bernse's local threshold selection is used for adaptive binarization of images and separating the area of interest for feature extraction from it's background [18] as in Fig. 1. Technique of threshold selection separates the image to foreground and background sections by comparing each and every pixel with a computed value as in equation 1.

$$I_{pixel\_value} = \begin{cases} 1, & pixel_{value} \leq Thresh \\ 0, & otherwise \end{cases} \qquad (1)$$

I_{pixel_value} is the binarized image and $Thresh$ is the value of the computed threshold for binarization. Bernsen's threshold selection uses equation 1 locally within a predefined window of dimension $s \times s$ where the value of s for satisfactory output is observed to be 31. This local threshold selection takes care of illumination effect and other adversities for binarization and supports in extraction of significant features. The pixel values less than or equal to one are designated as higher intensity group and the rest are designated as lower intensity group. Mean value for both the groups are calculated for individual color space Red (R), Green (G) and Blue (B). Thus we have two mean values for each color space which sums up to a feature vector of dimension 6.
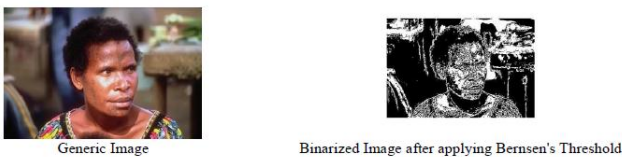

Fig. 1 Adaptive Thresholding for binarization

B. Defining Descriptor using Image Transforms

Slant image transform is applied on the image data to calculate the transform coefficient of the image. The image is resized to a dimension of 256 x 256 before applying the transform technique. Transform coefficient of same dimension of the image is received after applying slant transform. Henceforth, the higher energy coefficients are extracted as feature vectors in stages as illustrated in Fig. 2.
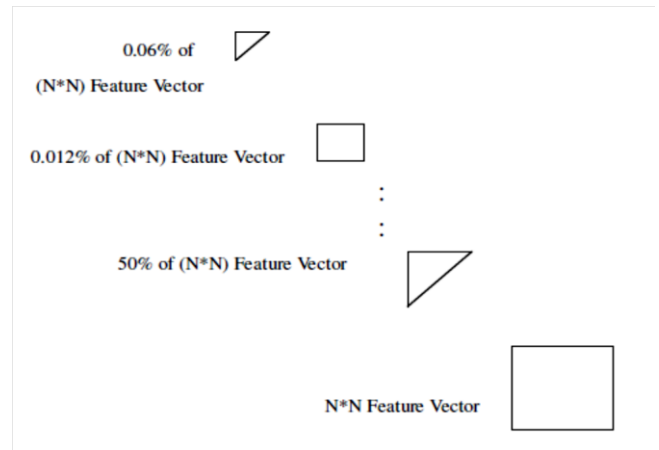

Fig. 2 Feature Extraction using Slant Transform

C. Defining Descriptor using Pre-trained CNN

A pre-trained CNN named *VGG 16* is used for feature extraction from the image data as in Fig.3. The input size of images for VGG 16 is *224 x 224*. All the images are resized to this dimension and is provided as an input. The features are extracted from an intermediate layer which has provided a significant descriptor definition to the classifiers to evaluate classification accuracy.
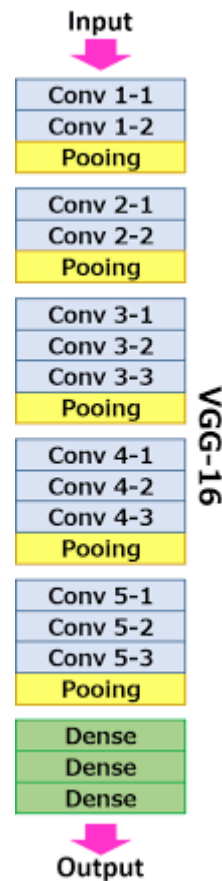

Fig. 3 *VGG 16* architecture

IV. DATASET DESCRIPTION

A widely used image dataset named Wang dataset is used for the experimentation purpose [18]. The dataset spans over ten varied classes of images spread across one thousand images. Individual category has equal number of images which has made it a balanced dataset. An illustration image is provided in Fig. 4.

.

Fig. 4 Sample Image Dataset

V. RESULTS AND DISCUSSIONS

Setup for the experiment involves an *Intel core i5* system having *4 GB* Ram using Python 3.6 programming language. Table 1 has given the results for classification with two separate classifiers using three different descriptors. The graphical comparison is shown in Fig. 5.

It is evident that feature extraction using pre-trained CNN has revealed the highest accuracy for classification contrasting rest of the handcrafted descriptors. Therefore, automated technique of feature extraction using pre-trained CNN (*VGG 16*) has outclassed the traditional handcrafted techniques in terms of classification accuracy and has resulted in robust descriptor definition compared to conventional means of feature engineering based approach.

Table- I Comparison of results for classification with two separate classifiers using three different descriptors

Techniques of Feature Extraction	Accuracy (%)	
	SVM	*Random Forest*
Defining Descriptors with Binarization using Bernsen's Local Threshold Selection	86.38	88.38
Defining Descriptors with Image Transforms using Slant Transform	89.29	91.26
Defining Descriptors Using Pre-trained CNN (VGG_16)	97.8	98.4

Descriptors defined using pre-trained CNN are able to create detailed and significant representation of the content in image data which is otherwise weaker in hand-crafted techniques. The domain specific nature of defining descriptors with binarization and slant transform has ignored necessary details that are being considered during definition of descriptors using pre-trained CNN (*VGG_16*). Hence the automated technique of defining content based image descriptors using *VGG_16* has reported higher classification accuracy in different classifier environments.
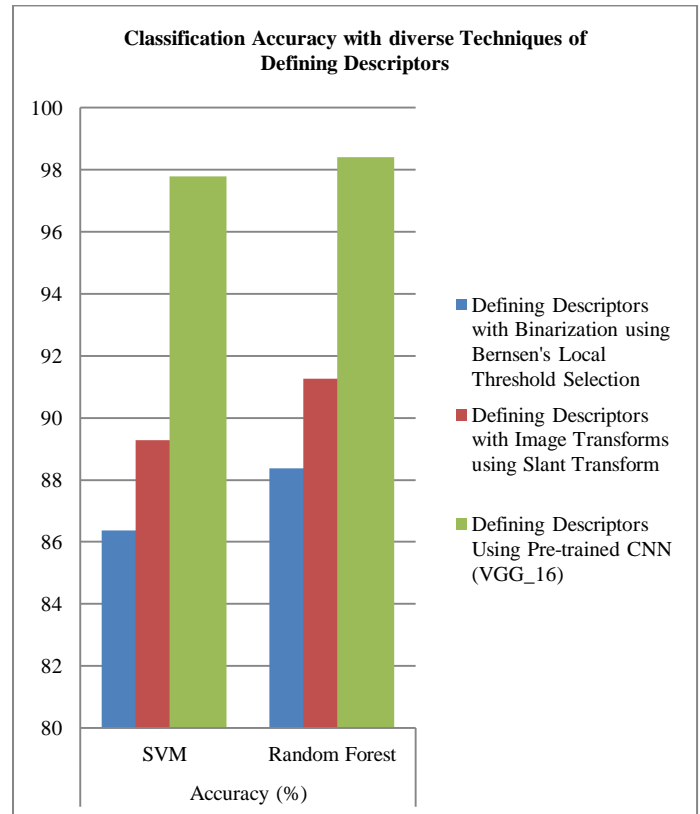

Fig. 5. Comparison of Classification Accuracies

VI. CONCLUSION

The research work has implemented three different methods for defining image descriptors based on content and has investigated the robustness of contribution of each of descriptors for boosting accuracy of classification. Results of the experiments have shown the supremacy of the automated feature extraction technique using pre-trained CNN over the conventional feature engineering based descriptor definition. This has established the fact that feature extraction with deep neural networks results in revelation of unknown patterns which is otherwise not perceived by traditional handcrafted methods. The research can be extended towards amalgamation of neural net layers for feature extraction in hybrid space. The extracted features can then be further contrasted to state-of-the-art to evaluate the classification efficacy in real time environment.

REFERENCES

1. Srdjan Sladojevic, Marko Arsenovic, Andras Anderla, Dubravko Culibrk, and Darko Stefanovic. "Deep neural networks based recognition of plant diseases by leaf image classification." *Computational intelligence and neuroscience* 2016 (2016).
2. Loris Nanni, Stefano Ghidoni, and Sheryl Brahnam. "Handcrafted vs. non-handcrafted features for computer vision classification." *Pattern Recognition* 71 (2017): 158-172.
3. Jianpeng Zhang, Yong Xia, Yutong Xie, Michael Fulham, and David Dagan Feng. "Classification of medical images in the biomedical literature by jointly using deep and handcrafted visual features." IEEE journal of biomedical and health informatics 22, no. 5 (2017): 1521-1530.

4. Le Hou, Dimitris Samaras, Tahsin M. Kurc, Yi Gao, James E. Davis, and Joel H. Saltz. "Patch-based convolutional neural network for whole slide tissue image classification." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2424-2433. 2016.

5. N.F. Hassan, and Reem Majeed Ibrahim. "Classification of Gender Face Image Based on Slantlet Transform." Engineering and Technology Journal 34, no. 4 Part (B) Scientific (2016): 566-577.

6. Cesare Alippi, Simone Disabato, and Manuel Roveri. "Moving convolutional neural networks to embedded systems: the alexnet and VGG-16 case." In *Proceedings of the 17th ACM/IEEE International Conference on Information Processing in Sensor Networks*, pp. 212-223. IEEE Press, 2018.

7. R. Raghavendra and Christoph Busch. "Robust scheme for iris presentation attack detection using multiscale binarized statistical image features." *IEEE Transactions on Information Forensics and Security* 10, no. 4 (2015): 703-715.

8. Sudipta Roy, Debnath Bhattacharyya, Samir Kumar Bandyopadhyay, and Tai-Hoon Kim. "An improved brain MR image binarization method as a preprocessing for abnormality detection and features extraction." *Frontiers of Computer Science* 11, no. 4 (2017): 717-727.

9. Yue Wu, Premkumar Natarajan, Stephen Rawls, and Wael AbdAlmageed. "Learning document image binarization from data." In *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 3763-3767. IEEE, 2016.

10. Chun-Che Chen and Shang-Lin Hsieh. "Using binarization and hashing for efficient SIFT matching." *Journal of Visual Communication and Image Representation* 30 (2015): 86-93.

11. Tejas H Jadhav, and Jaya H. Dewan. "Iris Recognition using Self Mutated Hybrid Wavelet Transform using Cosine, Haar, Hartley and Slant Transforms with Partial Energies of Transformed Iris Images." *International Journal of Computer Applications* 139, no. 11 (2016).

12. Jinglong Zuo, Delong Cui, and Qirui Li. "A novel slant transform-based image feature extract algorithm." *International Journal of Embedded Systems* 9, no. 3 (2017): 202-210.

13. Sudeep D Thepade and Madhura M. Kalbhor. "Extended performance appraise of Bayes, Function, Lazy, Rule, Tree data mining classifier in novel transformed fractional content based image classification." In *2015 International Conference on Pervasive Computing (ICPC)*, pp. 1-6. IEEE, 2015.

14. Sudeep D Thepade, Rupali K. Bhondave, and Ashish Mishra. "Comparing Score Level and Feature Level Fusion in Multimodal Biometric Identification Using Iris and Palmprint Traits with Fractional Transformed Energy Content." In *2015 International Conference on Computational Intelligence and Communication Networks (CICN)*, pp. 306-311. IEEE, 2015.

15. Zhenzhong Lan, Yi Zhu, Alexander G. Hauptmann, and Shawn Newsam. "Deep local video feature for action recognition." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 1-7. 2017.

16. Mingyang Ma, Shaohui Mei, Jingyu Ji, Shuai Wan, Zhiyong Wang, and Dagan Feng. "Exploring the influence of feature representation for dictionary selection based video summarization." In 2017 IEEE International Conference on Image Processing (ICIP), pp. 2911-2915. IEEE, 2017.

17. Rui-Wei Zhao, Zuxuan Wu, Jianguo Li, and Yu-Gang Jiang. "Learning semantic feature map for visual content recognition." In Proceedings of the 25th ACM international conference on Multimedia, pp. 1291-1299. ACM, 2017.

18. Y. O. Jaafri, Mohammed Omari, and Rekia Dlim. "Evaluation of Bernsen enhancement method applied to Arabic manuscripts." In 2017 International Conference on Mathematics and Information Technology (ICMIT), pp. 1-7. IEEE, 2017.

19. Rik Das, and Ekta Walia. "Partition selection with sparse autoencoders for content based image classification." *Neural Computing and Applications* 31, no. 3 (2019): 675-690.

AUTHORS PROFILE

Dr. Rik Das is serving as an Assistant Professor in PGRM-IT at XISS, Ranchi. He is an experienced academician with over 15 years of experience in his domain. He has several International publications and is resource person for ICT initiatives. Rik is also instrumental in arranging Faculty Development Programs and International Collaborations for quality research work.

Mohammad Arshad is research scholar at Department of MCA, Vinoba Bhave University Hazaribag, Jharkhand. Prior to his current assignment, he has secured bachelor's degree in Mathematics and MCA. His current research interest covers machine learning techniques for image classification. Arshad is currently engaged with several working papers to carry out quality research work.

Dr. Pankaj Kumar Manjhi is an Associate Professor at University Department of Mathematics, Vinoba Bhave University, Hazaribag, Jharkhand. He has extensive research experience in Mathematics and Computer Science. He is also serving as Member of Editorial Board and as a Reviewer of a number of renowned journals.