Deep Learning-Driven Feature Extraction for Classification of Indian Tribal Paintings

Nancy Hada[1], Kavita Vemuri[2]

[1,2])International Institute of Information Technology, Hyderabad, India

nancy.hada@research.iiit.ac.in, kvemuri@iiit.ac.in

Abstract

The proposed work endeavors to support the preservation of tribal arts in their original art form with their unique colors, shapes, and objects. Tribal arts of each region of India is a 'message' of their life, traditions, and history, requiring preservation in its pristine, unique form. For example, dotted patterns in Bhil, geometric figures in Warli, dashes in Gond, deities' expressions in Pichwai paintings, etc., are some notably diverse visual patterns that depict the profound stories of their wholesome culture. In this preliminary work, a dataset is collected by mining the websites, which are direct outlets of tribal arts, thereby ensuring the authenticity of the images included. The collected dataset includes 863 paintings across five different classes of tribal arts viz Warli, Bhil, Gond, Pichwai, and Kalighat, as described in Figure 1. Furthermore, image classification baselines on the collected dataset are developed by fine-tuning the pre-trained ResNet50, InceptionV3, and EfficientNetB0 models. The classification accuracy for the aforementioned baselines is reported in Table 1. Among the baselines, the best testing accuracy of 91.57 % is observed for EfficientNetB0; the confusion matrix for which is shown in Figure 2. The associated F1-scores for this baseline model across all five classes, namely Warli, Bhil, Gond, Pichwai, and Kalighat, were found to be 0.958, 0.882, 0.875, 0.909, and 0.958, respectively which suggests that the dataset is balanced in nature. Furthermore, class activation maps are generated for each image of the collected dataset across all baselines to ascertain the alignment of results generated with the ground truth, which offers a means to comprehend and compare the regions of focus in decision-making. Figure 3 provides an example of a class activation map revealing the salient focus areas, such as human figures represented by triangles and circles in a Warli painting. This analysis can also be used by artists for designing the artwork. The extension of the classification work is to detect counterfeit tribal paintings and quantify the extent of deviation from the original art form, ultimately contributing towards tribal arts preservation and safeguarding their invaluable cultural heritage.

Keywords— Tribal Art, Neural Network, Feature Extraction

Table 1: Accuracies for fine-tuned baseline models for Training, Validation, and Testing split on the collected dataset.

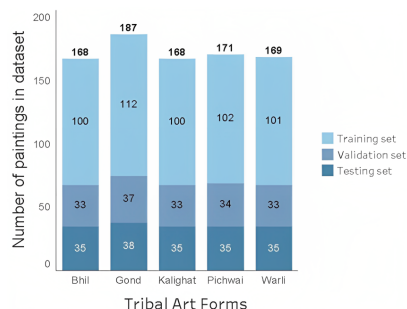| Models	Accuracy Percentage		
	Training	Validation	Testing
EfficientNetB0	98.25	93.75	91.57
ResNet50	98.07	89.39	85.91
InceptionV3	96.89	68.75	84.27

Figure 1: Data distribution with train, test, and validation splits
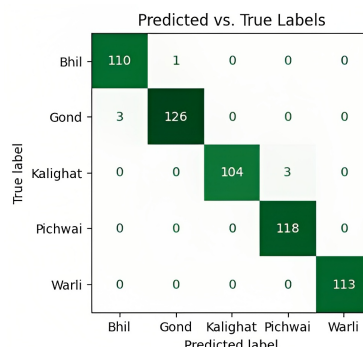

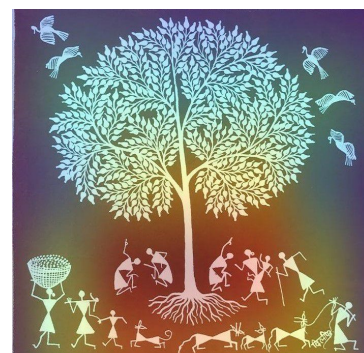Figure 2: Classification Co-variance matrix for EfficientNetB0

Figure 3: Class Activation Map for a sample Warli painting

References

[1] https://tribalartindia.com

[2] https://www.memeraki.com

[3] He, K., Zhang, X., Ren, S., Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[4] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z., 2016. Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

[5] Tan, M., Le, Q. (2019, May). Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning (pp. 6105-6114). PMLR.