

# A STUDY OF DEEP LEARNING TECHNIQUES FOR CULTURAL EVENTS RECOGNITION

Aman Swaraj<sup>1</sup>, Harshita Sahni<sup>2</sup>, Archit Agarwal<sup>3</sup>, Neeraj Kumar Pandey<sup>4</sup>,  
Supriya Shukla<sup>5</sup>

<sup>1,2,4,5</sup>Department of Computer Science and Engineering,  
College of Engineering Roorkee, Roorkee, Uttarakhand  
<sup>3</sup>Department of Information Technology,  
College of Engineering Roorkee, Roorkee, Uttarakhand

## **ABSTRACT**

*Indian Culture exceeds beyond the mere definition of 'simply how people live' as it scientifically operates according to specific, detailed knowledge of the eventual aim of life and the means to attain it. Over the years, a lot of work has been done on object recognition and scene recognition but Event Recognition is still one of the fields wherein lies a huge potential for lots of research work and so with this paper, we put forward our best step to preserve the culture of India. More than 150 images of near about 20 cultural events are collected. Results are derived from support vector machine classifier using features extracted by a pre-trained convolutional neural network- Alex Net. In most visual recognition tasks, it strongly suggested that features obtained from deep learning with convolutional nets should be the chief candidate. Our proposed framework has classified images with a comparable accuracy of 77.72 %.*

## **KEYWORDS**

*Indian cultural events recognition, Convolutional neural network, Local Binary Patterns, Support Vector Machine.*

## **1. INTRODUCTION**

Indian Culture has always been an identity mark for India. Cultural heritage is undoubtedly India's Golden Feather. The Indian culture includes various forms of traditional expressions such as literature, dance, music, rituals etc. Cultural also connects with agriculture, fisheries, forestry etc.

Culture enclose our glorious history that outlines our past, molds who we are today and who we are likely to become. There are many international organizations that aim at protecting cultural heritage, some of them are UNESCO, British Council, and many others. Even people from small towns promote their culture in an antiquated way [1]. Indian heritage foundation in India has also shouldered important projects for promoting the Indian culture. Today, in the contemporary society many festivities are not being celebrated the way they should be, the actual meaning behind the events are getting lost, their purpose getting defeated. Therefore, it needs to be protected from getting into the dormant state.

Deep learning is a division of machine learning which is stimulated by the functioning of the brain. It is considered to be most powerful and proficient model that performs outrageous even on laborious pattern recognition in vision and speech [2] [3]. Indian Government has tried to deal with this issue on different levels but they have failed to identify and preserve India's intangible cultural heritage. And so the need of the hour is to promote in such a way that every group, individual and institutions should get involved in preserving India's Cultural Heritage.

And realizing that, the Ministry of Culture has formed a Coordination Committee on the Living and Diverse Traditions of India to search out new ways to preserve India's insurmountable cultural heritage. The Committee is dealing with the issue nicely by constituting a subgroup whose purpose is to maintain a proper database, keeping records of various traditions of India. And to make sure that the information reaches a larger audience, it is recommending a digital presence of the same.

With this motivation, we hope to fulfill the gaps where the government is lagging on the 'Digital Promotion part' and contribute our bit to the nation so that it rises from the ashes like a phoenix and emerges as a 'GOLDEN BIRD' once again.

## 2. RELATED WORK

Previous work on cultural event recognition has been done by exploiting temporal model. Cultural events were detected and classified using visual features extracted from neural network with temporal statistics through a hierarchical classifier scheme in the Chalearn Challenge 2015. [1]

## 3. DATASET COLLECTION

Dataset collection is one of the most important phases. When deep learning approach is employed a strong and larger database is required for high performance. Dataset mostly compromises of images collected from Google. Dataset consists of images from 20 different festivals each consisting of more than 150 images.



Figure 1. Images of different cultural events

## **4. METHODOLOGY**

### **4.1. Support Vector Machines**

SVM is a supervised learning model, which has recently come into concern in the field of machine learning and solving pattern recognition problem. Their basis is the 'margin maximization principle'. They perform structural risk minimization, which improves the complexity of the classifier with the aim of achieving excellent generalization performance. In a higher dimensional space, SVM classifies by constructing a hyperplane which optimally segregates the data into two categories. Standard numeric techniques for QP has become impractical for very large datasets. SVM works well in high-dimensional spaces and gives excellent results in texture classification [4].

### **4.2. K-Nearest Neighbours**

In case of lack of knowledge about the distribution of the data, this classification technique works very smoothly. For the classification, the Euclidean distance is calculated between the new instance and the stated training samples [5]. KNN editing has the endowment of removing noisy instances from the training set [6]. Predictions in KNN are directly made using the training data set only.

### **4.3. Random Forrest**

Random forest is a classifier method used for classification which is fast to implement tree-structured predictors are the base constituent of the ensemble and is constructed using interpolation of randomness, therefore, termed as random forest[7]. Random Forest classifier creates many classification trees.

Many technical reports, Bremen (1996, 2000, 2001, 2004), demonstrated the steady gain in the classification and regression accuracy after using this classifier. It is basically used in unsupervised machine learning, used to describe unlabeled data.

### **4.4. Histogram of Oriented Gradients**

HOG is an image descriptor which is used for object detection. The fundamental concept behind HOG is to describe the properties of the image like appearance, shape, color, etc. by distributing the intensity gradients properly. It divides the image into a small connected region called cells and a histogram of gradient direction is compiled in the pixel in each cell.

### **4.5. Local Binary Patterns**

LBP is a type of image descriptor which is used for classifying images in computer vision. Though simple yet is very effective texture operator. It classifies the image by thresholding each pixel's neighborhood and a result is a binary number. In various applications, this texture operator has become popular because of its computational simplicity and high discriminating power, which makes it possible to analyze images in challenging real-time settings.

### **4.6. Gist**

In recent times, the GIST descriptor received a lot of demand when it comes to scene recognition. It is a popular approach for web-scale image indexing [8]. It is based on Spatial Envelope, which

does not require any type of segmentation. And it has also produced significant results in case of image search. GIST is primarily used to retrieve set of images of similar landmarks, for example, Taj Mahal. And the image point based matching technique refines the result and builds a 3D model of the same. They define the features that separate a scene from the rest.

#### 4.7. Convolutional Neural Network (CNN)

CNN is a class of machine learning that is composed of multiple processing layers .it uses feed-forward artificial neural network or backpropagation algorithm. The first convolutional network was presented in 1989 [9] [10].

The primary eminence of CNN over other classifiers is that it requires little pre-processing as compared to others which have brought further enrichment in processing images, video, speech, and audio.

### 5. RESULT AND DISCUSSION

#### 5.1. Experimental Scenario 1

In this scenario, we calculate the basic features by using HOG, LBP and GIST feature separately with SVM, KNN, and Random Forest classifier. The Purpose of these computations is to find out the best distinguishing feature that would later add up to our proposed methodology CNN. (Table 1)

#### 5.2. Experimental Scenario 2

We proceed towards CNN, where we first calculate the accuracy of CNN alone with the SVM, KNN, Random Forest classifiers separately. The Purpose is to identify the best classifier to be used for the cultural events along with CNN. (Table 2)

#### 5.3. Our Proposed Approach

Our proposed approach uses the pre-trained Alex net architecture. As per the need of CNN, pre-processing of images is done afterwards. For the feature extraction, Alex net is used In order to make computations less expensive, we have used two layers (FC6 AND FC7) of the Alex net architecture. The mentioned two layers are fully connected layers out of the total of 8layers. 5 layers are convolutional layers. (Table 3)

Table 1. Experimental Scenario 1

Features/Models	SVM	KNN	Random Forrest
HOG	22.5476	16.4598	14.5698
LBP	25.3658	17.5486	18.3569
GIST	22.5789	20.6598	18.2578

Table 2. Experimental Scenario 2

Features/Models	SVM	KNN	Random Forrest
CNN	60.5486	55.6587	52.3698

Table 3. Experimental Scenario 3

Features/Models	HOG	LBP	GIST
CNN+SVM	65.5679	77.7200	62.3569

Out of all experimental scenarios, CNN+SVM+LBP gave the best result with an accuracy of 77.72%.

## 6. CONCLUSION

In this paper, we employed a deep learning technique for the classification of different cultural events in their respective domains with a comparable accuracy of 77.72%. Indeed, this paper is a contribution to the project lead by Indian government “parampara” for the digital promotion of Indian cultural events.

## ACKNOWLEDGMENTS

We would like to express our sincere gratitude to our supervisors Dr. Himanshu Chauhan (HOD CS Department), Ms. Supriya Shukla (Assistant Professor CSE Department), and Mr. Neeraj Kr Pandey (Assistant Professor CSE Department) for providing their invaluable guidance, comments, and suggestions throughout the course of the project. We would specially thank Mrs. Supriya Shukla for constantly motivating us to work harder.

## REFERENCES

- [1] Salvador, Amaia, et al. "Cultural event recognition with visual convnets and temporal models." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. 2015.
- [2] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
- [3] Hinton, Geoffrey, et al. "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups." IEEE Signal Processing Magazine 29.6 (2012): 82-97.
- [4] Kim, Kwang In, et al. "Support vector machines for texture classification." IEEE transactions on pattern analysis and machine intelligence 24.11 (2002): 1542-1550.
- [5] Peterson, Leif E. "K-nearest neighbor." Scholarpedia 4.2 (2009): 1883.
- [6] Wilson, D. L. Asymptotic Properties of Nearest Neighbor Rules Using Edited Data. IEEE Transactions on Systems, Man, and Communications 2, 3 (1972), 408–421.
- [7] G´erardBiau, “Analysis of a Random Forests Model “, Journal of Machine Learning Research 13 (2012)
- [8] Douze, Matthijs, et al. "Evaluation of gist descriptors for web-scale image search." Proceedings of the ACM International Conference on Image and Video Retrieval. ACM, 2009.
- [9] Guan, Donghai, et al. "Nearest neighbor editing aided by unlabeled data." Information Sciences 179.13 (2009): 2273-2282.
- [10] [https://en.wikipedia.org/wiki/Convolutional\\_neural\\_network](https://en.wikipedia.org/wiki/Convolutional_neural_network)

**AUTHORS****Aman Swaraj**

B.TECH -IV, Department of Computer Science and Engineering, College of Engineering Roorkee, amanswaraj007@gmail.com.

**Harshita Sahni**

B.TECH -IV, Department of Computer Science and Engineering, College of Engineering Roorkee, harshitasahni777@gmail.com.

**Archit Agarwal**

B.TECH -IV, Department of Information Technology, College of Engineering Roorkee, aavicky995@gmail.com.

**Supriya Shukla**

Assistant Professor, Department of Computer Science and Engineering, College of Engineering, Roorkee, supriyaluthra@yahoo.co.in

