DESCULPT: Indian Temple Sculpture Iconography

Akshath Rao
Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Sindhu C
Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India, sindhucmaa@gmail.com

Aayush Mehta
Department of Computing Technologies, SRM Institute Science and Technology, Kattankulathur, India

Abstract

India is a land comprising of a lot a rich culture and heritage. There are a lot of rare and architectural sculptures which have been discovered and found in the recent years in India. Entity recognition of Indian sculptures can be viewed as one of the very challenging and difficult problems occurring in the very field of image recognition, classification as well as identification. In order to get this project started, it is constructed on a manually formed database con- sisting of pictures of Lord Ganesha and Lord Shiva Linga. In this work, Lord Ganesha and Lord Shiva Linga as the two entities which have to classified. Every image consists of a key characteristic which differentiates it from the other. Ori- entations, angles, size, colours of the images all play an important role in the processing of the model. The images have been taken from various and different angles, sizes, colours and orientations. The model will be trained on the image dataset consisting of the entities in Indian sculpture, in return signifying the rich heritage and culture. Deep learning plays a very important role in image classifi- cation and recognition. The effectiveness and accuracy of the suggested model can be evaluated using Support Vector Machine and Convolutional Neural Net- work respectively. Also many other important libraries such as NumPy, mat- plotlib , OpenCV have also been used in this process of image recognition and classification. The model which this project has helped trained has helped us reach an accuracy of 91% with the help of Deep Learning and Convolutional Neural Network along with the Sequential Model.

Keywords: *India, Deep Learning, CNN.*

1 Introduction

Times have changed drastically and so have people and the practices followed. Records containing details of all sorts of various monuments, buildings, sculptures, etc. would be kept in huge files and paperwork so one can access them later when needed. With the advent of the digital era we now dwell in, the process has become less arduous and therefore, less cumbersome. We can now exploit these technological advancements for our use. A key sense of perception - Vision, is pivotal for obtaining information sur- rounding us in any form. Digitally, these pictures are available in the form of images, videos, gifs, and many more which have been made easily attainable by the Internet. This has led to an enormous

amount of research projects concerning extracting vital insights from data pertaining to these visual information such as sentiment analysis, car model detection, human detection, and this research project – DeSculpt.

India is a country rich in culture and traditions. There are a myriad of temples(around 2 million), caves(2702), and monuments(3650) with several sculptures of deities, gods, past heroes, and beings engraved on them. Over the past years the findings of many ancient idols of gods have started to be found at regular intervals of findings. These numbers are growing day by day which makes the identification of such sculptures a very arduous task. Many archaeological models may seem similar but they are not re- ally so. Slight subtleties amongst these sculptures are nearly impossible to detect and recognize with the naked eyes. Every image consists of a key characteristic which dif- ferentiates it from the other. Orientation, angles, size, colour of the images all play an important role in the processing of the model. In order to make this task easy and effi- cient, many machine learning technologies have been created for helping archaeologists in detecting and classifying these sculptures. With the help of this paper, we elaborate upon an efficient recognition and detection machine learning model which would be able to do so. The primary aim is to classify the entity that is present in the provided or considered image of the sculpture. We intend to create a sculpture identification system that would be able to detect any sculpture fed to it by the user in an image format. This is the main motivation that drives our project forward. The paramount objective is to identify the entity in question with minimal precision error in the provided image of a sculpture. The most important goal is to accurately identify the object in question in the accompanying sculpture image. In light of our numerous discoveries, we have chosen to create a model based on the idea CNN(Convolutional Neural Network). Other meth- ods like ANN(Artificial Neural Network), SVM(Support Vector Machine) and SIFT(Scale Invariant Feature Transform) can also be used but CNN provides us with the most efficient and accurate working model. The sole purpose is to draw out the comparisons and then ultimately instituting an efficient model. Deep learning is going to be fundamental and essential in the design and hence CNN.

This work consists of a manually formed database consisting of numerous images of the Two Gods which are to be classified . These prime entities which are to be classified are – Lord Ganesha and Lord Shiva Linga. The image pool will contain shots from different angles and orientations leading to the highest possible precision output or re- sult. obtained here can be of a specific format like Jeg, Png or Jpeg for efficient function of the machine learning model. The acquired images in the database are later separated and divided into two parts namely each of training and testing purposes respectively. Here, a total of 10% of the total data to be categorized is used for testing purpose and the rest for the purpose of training respectively. The machine learning model is first loaded with the training data to get trained and then later we input the testing data in order to check if the model has been trained right and if it is working efficiently and producing the correct prediction.

Here, in the subsequent section of the research paper contains the structure of it con- sisting of Literature survey, Methodology, Experimentation and Results. The main hope is that this project would become vital in the education sector for analysis of var- ious Indian sculptures.

The hybrid of algorithms and various experiments conducted in future would be the key in providing essential data for further research in the field of object analysis and categorization in the domain of image processing. This research has the potential to extend its utility to numerous fields possessing

the classical entity identification issue. Medical, Astronomy and Military are just the tip of the iceberg that can make excellent use of this tool in solving real life problems.

Fig. 1. Sample images of Lord Ganesha

Fig.2. Sample images of Lord ShivaLinga

2 Literature Survey

Sculpture dectection is an idea not been given much attention yet. It is a noval one, where much work has Not been devoted to it. In this field, a lot of research has not yet been made. There have been many different approaches and strategies shown and implemented by various researchers and authors. The section below shows the given out findings of the key and important aspects each used by the authors in various Research papers.

A member from IACSIT worked on a sculpture detection model which he called the Thai Buddhist Sculpture Recognition System[1]. The main target of this system was the recognition of monuments relating to Buddhism in Thailand. This system consists of five parts viz. Image acquisition (via a camera), image pre-processing, Feature Ex- traction, Image recognition(applied using the Euclidean Distance technique), and dis- playing the output through a graphic user interface.

Another team worked on the similar concepts, i.e, Sculpture Recognition using a different algorithm[2]. They combined and compared various algorithms which gave out the same result to check the most efficient one.The SIFT algorithm was paired with K-Nearest Neighbours, Support Vector Machine, and Artificial Neural Network using four different approaches – Min, max, mean, and median key paddings. CNN wax`s used as well. They used Contrast Limited Adaptive Histogram Equalization to improve the efficiency of the overall system. The most efficient and accurate model they worked out was that of CNN with the incorporation of CLAHE. A total of around 2500 images (224*224 pixels) from 15 different entity classes(Indian deities) were used training and 20 for testing.

A group of six people from IIT Roorkee attempted to tackle a similar problem. They made a model that recognized monuments in India using Deep Convolutional Neural Networks[3]. Several images of Indian monuments from various angles were fed into the model. They acquired a very high amount of accuracy. Feature extraction was car- ried out using the Histogram of oriented gradients. Training of the data was done using Local binary patterns and GIST features.

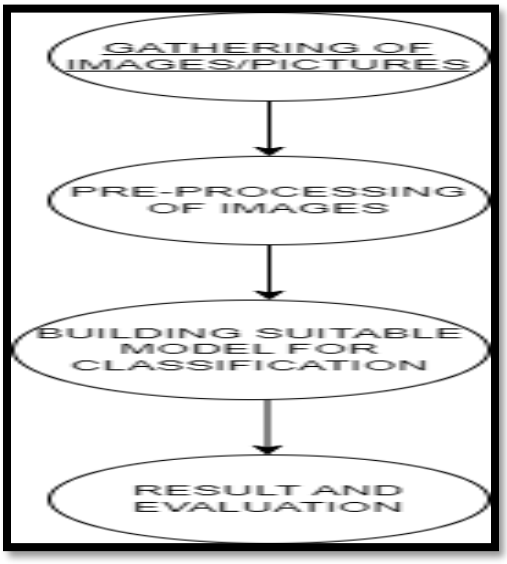
Table. 1. Literature Survey Table

No	Methodology/ Model used	Merits And Findings
[1]	• Acquisition: images -> train • Preprocessing: Cropping, Resizing, Grey-scale conver-sion • Feature Extraction: Edge, Eye,Nose detection, Bottom Neck,Top Hair • Recognition: 14 features using Euclidean Distance method	• High precision rate(90%)
[2]	• SVM, ANN, and CNN modelsusing 4 approaches.	• CNN + CLAHE was the most accurate one (27.66%)
[3]	• DCNN • HOG,GIST,LBP Used	• High accuracy(92.7%)
[4]	• OpenCV implementation of Viola-Jones Face Detector, • Cascade classifiers trained on images	• High accuracy and low false positive rate
[5]	• 3D data capturing and extrac-tion of global shape features. • Spherical Harmonic 3D shape descriptor	• Highly applicable for dam-aged Buddhist statue resto-ration and sculpture stylis-tic analysis

3 DeSCULPT Methodology

In the methodology section, the work address the method used in our process. It consists of Flow Chart and a block diagram explaining the process. The procedure is broken down into different block diagrams for reference. The approach that is expected to be used for image processing has been presented by this work. The section described be- low contains the supposed method of image processing. These are namely the: DeS- culpt Flow diagram as shown in fig.3 and The DeSculpt Architectural block diagram as represented in fig.4.

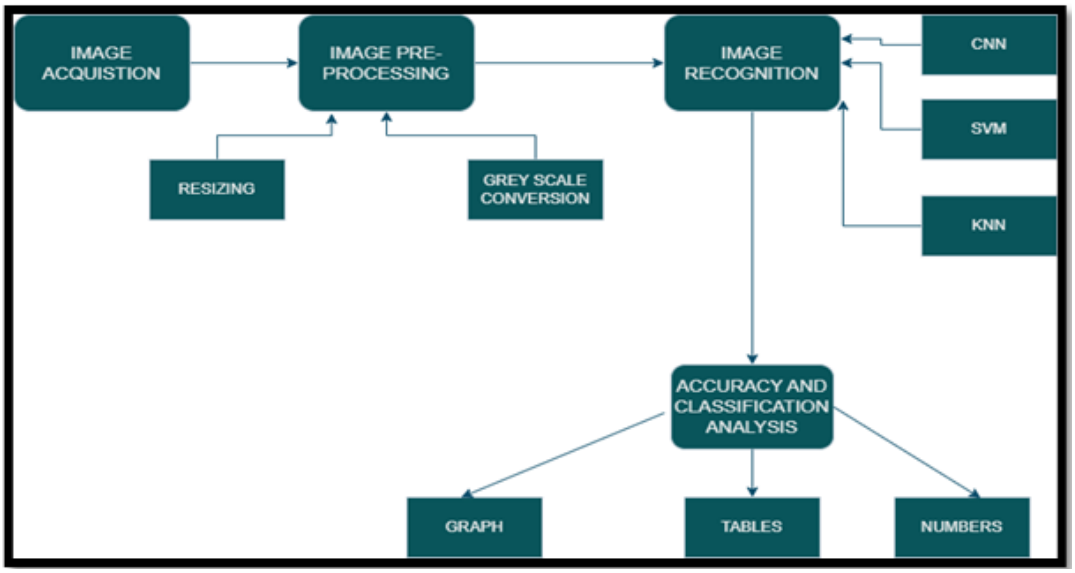
Fig. 3. DeSculpt Flow diagram



The diagram in fig.3 primarily outlines the general flow of steps involved in any basic image processing. A total of 4 steps have been primarily indicated. It contains: Gathering of images, Pre-processing of those images, Model

for classification and recognition, Results and evaluation for representation of statistics related to the project.

Fig. 4. DeSculpt Architectural Block Diagram



The block diagram in fig.4 is the advanced and detailed representation of the major steps involved in the process. It consists of 4 main components namely: Image Acqui- sition, Image Pre-Processing, Image Recognition and Accuracy and Classification Analysis.

3.1 Image Acquisition

Here the task start off with the image acquisition process which involves the collection or acquiring of the required images to form the database. “Image acquisition”, entails the gathering or purchasing of the necessary images to create the database. The database images consists of the two gods respectively. Plenty of images of these gods have been collected respectively. Later, the pictures have been divided into two parts for Training and Testing respectively. This is manually formed database from where the pictures have been primarily taken from google images.

Nearly 10% of the acquired images have been used for testing purposes while the remaining have been used in the process of training the model.

3.2 Image Pre-Processing

Next on to the image pre-processing stage, mainly dealing with two main functions namely – resizing the images and conversion of the images into the required grey-scale format in order to apply the suitable machine learning algorithm and model. To get started with this procedure, start off by first resizing all of the images in the database such that they have a resolution of 100 x 100 pixels on each side. This brings the pho- tographs up to a consistent standard, which allows the process of conversion to proceed in a cohesive fashion. This also standardizes the pictures so that the conversion process can take place in an unified manner.

In addition to this, shift it from colour to grayscale so that it can more easily be applied to the algorithms, and then, if necessary, also crop the picture or modify its orientation in order to get it into the appropriate format. The

RGB or GRB colour space is transformed into grayscale, depending on the kind of greyscale conversion being per- formed.

3.3 Image Classification

In the recognition phase , we discuss about the various models which can be used to classify the images into its respective class. Models such as CNN(Convolutional Neural Network),SVM(Support Vector Machine) can be used. Deep Learning with sequential model or CNN with SVM both be used for the building , recognition and testing pur- poses. During the recognition phase, the work talks about the several modelsthat may be used to the process of sorting photos into their appropriate categories. There are a variety of models that may be used, including CNN, SVM and KNN (k-Nearest Neigh- bors). CNN with SVM or Deep Learning with Sequential Models (DLSM) may both be utilized for the construction, recognition, and testing processes. There are just two types of representations that need to be categorized and identified in this context, and those are of Lord Ganesha and the Lord Shiva Linga. Because of this, the project is able to employ any of the two methods as they are both linear and binary and can accurately categorize the information. In the form of integers such as 0 and 1, an array may be used to discern whether the picture depicts Lord Ganesha or the Lord Shiva Linga

3.4 Classification Analysis

Moving onto the accuracy and classification analysis phase, we determine the accuracy of prediction of the model, calculate the loss or error. Loss or error refers to the infor- mation that is lost during the processing of the images in the prior stages. We also use matplotlib as one the of libraries, which would help in plotting a graph representing the accuracy and loss of the training data vs the testing data. Libraries like NumPy are taken into consideration since they are particularly effective in terms of the speed at which the processed input is returned in the form of arrays. In this phase, we also show and analyze the outcome of the numbers obtained during the testing phase and represent it in the form of graphs, statistical numbers and output pictures. Throughout this phase, we also represent and analyze the results of the testing phase. The end result would be primarily in the form of numbers from 0 and 1 based on its classification labelling and would also include the percentage indicating the level of accuracy determined by the used model. Below is shown a brief overview of the results obtained in fig.5 and fig.6.

Fig. 5. Accuracy and Loss Values

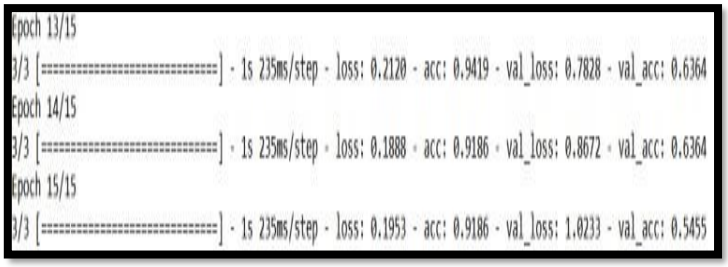
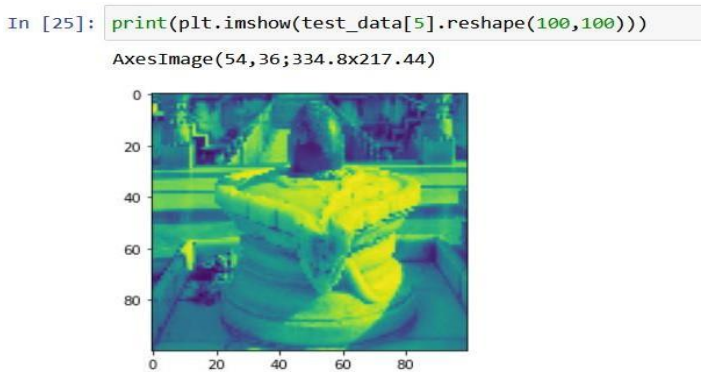


Fig. 6. Predicted Output



4 Implementation

In phase of the research paper primarily speaks about the implementation of the model. The work has been implemented with the current image processing and classification model using Deep Learning. CNN is used along with a Sequential Model in the classi- fication process.

4.1 Libraries Used

Multiple machine learning libraries have been used in order to train the model for identification and classification of the images. Tensorflow along with keras is widely used to implement the models and the required libraries for training and testing purposes. Open CV is also used here.

It plays a very important role. It can be considered as one of the best image processing libraries used for classification and identification purposes. Numpy is also included. We use numpy as it provides us with a very fast and efficient speed during the model training and testing phase as well as during the processing of the images. Mat- plotlib is used in order to plot the results obtained. This is generally represented in the form of graphs with colours indicating the trends of the various components like accu- racy, loss, variable accuracy and variable loss.

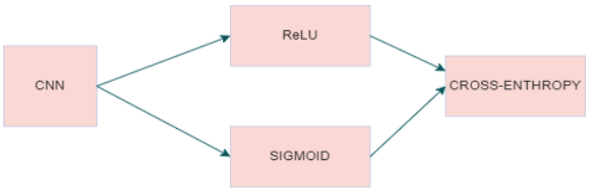
4.2 CNN-Sequential Model

CNN plays a very important role in the process of image classification. Here, in this model, we primarily deal with only two classes namely Lord Ganesh and Lord Shiva Linga. Hence to classify it, we can make use of a binary and linear model called a Sequential model consisting of some activation functions in order to determine the out- put. The CNN architecture consists of hidden layers which has been indicated in the project by the keyword 'layer'. This work has taken into consideration four layers of Convolution2D and Maxpooling2D followed by Flatten and Dense Layer supported by Activation function ReLU(Rectified Linear Unit) and Sigmoid Function which is used to fire up the neuron in the hidden layers in order to reach the output. Convolution2D layer is primarily used to produce a wide variety of tensor of outputs by creating a convolutional kernel with the input layer. Maxpooling2D is also used. We use this to extract and select the maximum number produced by the input layer. Dense layer is used in the implementation part. The primarily use of the dense layer is that it has the ability to receive the neuron which get fired from the previous layers. Dense Layer is used to classify image based on output from convolutional layers. Flatten layer is used our model in order to convert all the 2-D array obtained from the pooled maps into a single vector. Both dense layer and flatten layer play a significant role in determining the output of the model.

4.3 Activation Function

The CNN model has into four layers of Convolution2D and Maxpooling2D followed by Flatten and Dense Layer supported by Activation function ReLU and Sigmoid is used to fire up the neuron in the hidden layers in order to reach the output. ReLU is the activation function which has been used in this model for supporting and assisting each layer of the CNN. It is a linear function which will output the input directly of it is positive or else it will return zero. Sigmoid activation function is also used here. It is also used as it deals in identifying and classifying only two entities and the output re- quired would be in the form of 0 and 1. Binary cross -entropy function is also used here. This function was used as a loss function primarily active when in a binary classification task. These return a value of 0 or 1 respectively. The sigmoid function, which is a variant of the logistic function, is typically indicated by sig(x). The typical relationship between the various functions used in CNN is represented in fig.7.

Fig. 7. Activation Function Relationship.



5 Results and Discussion

After implementing this model, accuracy of the model was determined. Accuracy can be here defined as the ability of the model to correctly predict if the output belong either of the two entity class of Lord Ganesha and Lord Shiva Linga. The CNN – SEQUENTIAL model which was used was able to rightly and efficiently predict the output when an input as per the user was given

Table. 2. Input and Output Table

INPUT	OUTPUT
0	LORD GANESHA
1	LORD SHIVA LINGA

The images were separated and then classified into training and testing images. 10% of the database was used of testing while the remaining was used for training the model. Checkpoints were created as the ModelCheckpoint function which was used to determine the functioning of the model at the given reference points. It was also used to monitor the loss or error taking place during the functioning. The Epoch value was set to 15. Excessive epoch values was not set in order to prevent inaccurate and overfitting o0f the data. The validation split was set to 20% or 0.2 and later matplotlib was used to plot the data against the respective components of loss and accuracy.

5.1 Epoch Value And Graph

The epoch value was set to 15 . When the epoch cycle ended, we found the following parameters and its values : loss -0.19 , acc-0.91, val loss -1.02 and val acc-5.44.

Table. 3. Epoch table evaluation where cycle =15

EPOCH CYCLE NO	LOSS	ACCURACY
1/15	0.23	0.90
3/15	0.21	0.91
6/15	0.26	0.89
9/15	0.19	0.90
12/15	0.21	0.90
15/15	0.19	0.91

Matplotlib was also used to plot the graph . We obtain a graph showing the trend of val , acc, val loss , val acc respectively.

Fig. 8. Graph indicating trends

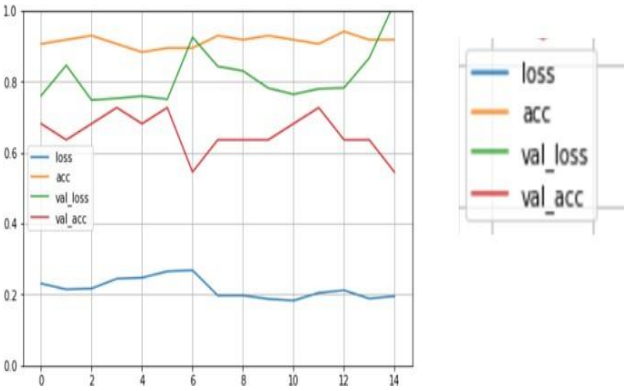


Fig. 9. Resulting output-Shiva Lingam

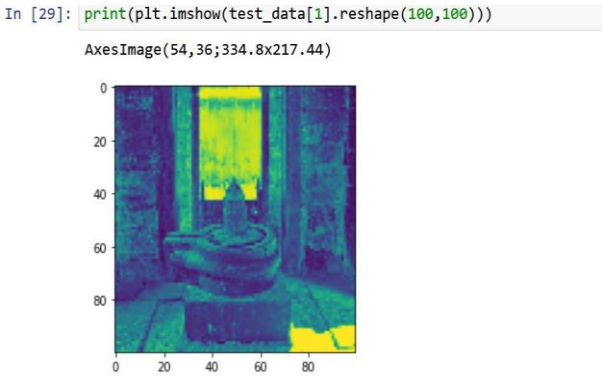
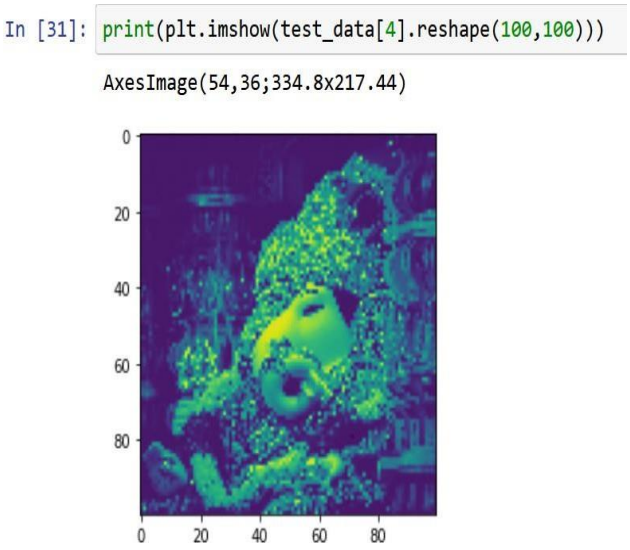


Fig. 10. Resulting output-Lord Ganesha



From the above, it can be noted that when the input of 1 was given, the model successfully returned the image of lord Shiva Linga and when the input of 0 , we successfully obtained lord Ganesha image. Thus, this also proves that the accuracy obtained of 91% is verified and valid.

6 Conclusion

The model uses a CNN model to actively, effectively, and accurately categorize if the provided image belongs to either of the two entity classes, which are those of Lord Shiva Linga or Lord Ganesha. The CNN model was utilized with sequential modelling to achieve an output of 0 and 1 accordingly to get the correct level of classification. ReLU, Sigmoid, and Binary Crossentropy are important activation functions supporting the model .

After effectively building the CNN model, epoch cycle of 15 was able to set and run successfully. After the epoch cycle ended , this model was able to obtain an accuracy of 91%. This implies that any image belonging to Lord Shiva Linga is not misclassified the other, thus leading to a good classification and identifying model.

In the future, similar model can used for multi-class or multi-labelling concept in the project and can be used in order to identify and classy more number of entities (Gods). With more enhancement to the CNN model and advancements to the functions, we can build a model having the ability to efficiently and correctly identify more than two entities with a good level of accuracy. To train and use high-quality diagnostic models, it is also anticipated that neural network models can be combined with ensemble, trans- fer learning, and other ground-breaking ideas.

References

Thai Buddhist Sculpture Recognition System (TBuSRS) Chomtip Porn- panomchai, Member, IACSIT, Vachiravit Arpapong, Pornpetch Iamvi- setchai, and Nattida Pramanus IACSIT International Journal of Engineering and Technology, Vol.3, No.4, August 2011.

A. Dalara, C. Sindhu and R. Vasanth, "Entity Recognition in Indian Sculp- ture using CLAHE and machine learning," 2022 First International Confer- ence on Electrical, Electronics, Information and Communication Technolo- gies (ICEEICT), 2022, pp. 1-12, doi: 10.1109/ICEEICT53079.2022.9768565.

A. Saini, T. Gupta, R. Kumar, A. K. Gupta, M. Panwar and A. Mittal, "Im- age based Indian monument recognition using convoluted neural networks," 2017 International Conference on Big Data, IoT and Data Science (BID), 2017, pp. 138-142, doi: 10.1109/BID.2017.8336587.

G. Shalunts, M. Cerman and D. Albertini, "Detection of sculpted faces on building

facades," 2017 Asia-Pacific Signal and Information Processing As- sociation Annual Summit and Conference (APSIPA ASC), 2017, pp. 677-685, doi: 10.1109/APSIPA.2017.8282119.

Desai, P., Pujari, J., Ayachit, N. H., & Prasad, V. K. (2013). "Classification of archaeological monuments for different art forms with an application to CBIR". 2013 International Conference on Advances in Computing, Commu- nications and Informatics (ICACCI). doi:10.1109/icacci.2013.6637332.

Kampel, M., Huber-Mörk, R., & Zaharieva, M. (2009). "Image-Based Re- trieval and Identification of Ancient Coins". IEEE Intelligent Systems, 24(2), 26–34. doi:10.1109/mis.2009.29.

Relja Arandjelović, Andrew Zisserman, "Name that sculpture", ICMR '12: Proceedings of the 2nd ACM International Conference on Multimedia Re- trieval June 2012 Article No.: 3 Pages 1–8 https://doi.org/10.1145/2324796.2324800.

B. Ghildiyal, A. Singh and H. S. Bhadauria, "Image-based monument classi- fication using bag-of-word architecture," 2017 3rd International Conference on Advances in Computing,Communication & Automation (ICACCA) (Fall), 2017, pp. 1-5, doi: 10.1109/ICACCAF.2017.8344728. 30.

Shalunts G., Haxhimusa Y., Sablatnig R. (2011) "Architectural Style Classi- fication of Building Facade Windows". In: Bebis G. et al. (eds) Advances in Visual Computing. ISVC 2011. Lecture Notes in Computer Science, vol 6939. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3- 642-24031-7_28.

P. Shukla, B. Rautela and A. Mittal, "A Computer Vision Framework for Automatic Description of Indian Monuments," 2017 13th International Con- ference on Signal-Image Technology & InternetBased Systems (SITIS), 2017, pp. 116-122, doi: 10.1109/SITIS.2017.29.

K. -Z. Liu, P. -J. Lee, G. -C. Xu and B. -H. Chang, "SIFT-Enhanced CNN Based Objects Recognition for Satellite Image," 2020 IEEE International Conference on Consumer Electronics - Taiwan (ICCETaiwan), 2020, pp. 1- 2, doi: 10.1109/ICCETaiwan49838.2020.9258037.

S. K. Raza, T. Rajab, S. A. Ahmed and M. Khurram, "Visual Landmarks Recognition of Urban Structures using Convolutional Neural Network," 2020 3rd International Conference on Computing, Mathematics and Engi- neering Technologies (iCoMET), 2020, pp. 1-9, doi: 10.1109/iCoMET48670.2020.9073795.

Fix, Evelyn, and J. L. Hodges. "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties." International Statistical Review / Revue Internationale de Statistique, vol. 57, no. 3, [Wiley, International Sta- tistical Institute (ISI)], 1989, pp. 238–47, https://doi.org/10.2307/1403797.

T. Cover and P. Hart, "Nearest neighbor pattern classification," in IEEE Transactions on Information Theory, vol. 13, no. 1, pp. 21-27, January 1967, doi: 10.1109/TIT.1967.1053964.

Sindhu C and G. Vadivu, "Fine grained sentiment polarity classification us- ing augmented knowledge sequence-attention mechanism", in the Journal of Microprocessors and Microsystems, Vol 81, 2021.

https://doi.org/10.1016/j.micpro.2020.103
365.