INDISIGHT: A CNN Based Image Recognition Of Indian Monuments

Mrs Belji T [*1], Hemanth Mahesh [*2], K Preetham [*3], Kiran Chandrashekar Datawad [*4], J R Chandan [*5]

[*1]Assisstant Professor, [*2,3,4,5]Student, Department of CSE, KS School Of Engineering & Management, Bengaluru, India.

Abstract - In a world rich with cultural heritage and historical significance, preserving and celebrating the architectural wonders of India is of paramount importance. Our project, "IndiSight" introduces a groundbreaking solution for the seamless identification and appreciation of Indian monuments. This system utilizes convolutional neural networks to analyze images and provide detailed information about the monuments. With this system, we can explore India's rich cultural heritage and gain a deeper understanding of its history.

Index Terms - Artificial Intelligence, Machine Learning, Deep Learning, Image Classification, Architectural Heritage, CNN(Convolutional neural networks), Monument Recognition.

I.INTRODUCTION

In recent years, the proliferation of high-performance parallel computers equipped with fast GPUs has facilitated the simulation of complexly interconnected neural networks. However, these machine learning systems often appear as opaque black boxes accessible primarily to those with minimal theoretical knowledge. Despite numerous attempts, finding the optimal network topology for specific problems remains a challenge, resulting in a vast array of models forming a comparative evolutionary landscape.

From an epistemological perspective, the interdisciplinary nature of machine learning stands out prominently. It spans across diverse subjects like linguistics, theoretical physics, computer science, art, biology, and logic. This multidisciplinary essence aligns with its foundational goal of mimicking brain functions. Within this context, leveraging deep learning methods to generate queries from real-world data for database access becomes a notable cultural pursuit. The objective here is to utilize machine learning, particularly deep learning, to initiate queries based on real objects like architectural artifacts or monuments, aiming to extract comprehensive information pertaining to these entities.

II.MOTIVATION

Urban elements within a city form interconnected narratives, often concealing inaccessible information. Sites, whether ancient, modern, or archaeological, encapsulate a tapestry of historical events across time. This wealth of information constitutes a virtual network, interlinking various data points. Storing this information in a database demands queries that pivot on diverse parameters. For instance, encountering an ancient statue might spark interest in sculptural technique, historical context, archaeological significance, or the creator's biography. Simultaneously, one might seek information about analogous specimens across different geographical regions.

Traditional solutions like audio tours, informative panels, or QR codes fall short in addressing such diverse queries. Conversely, prevalent portable music recognition systems engage with environmental features to access data and services. Similarly, by connecting architectural objects, artworks, and spatial characteristics with digital content, mobile devices could serve as gateways enabling complex queries initiated by object recognition.

Cultural heritage management faces the challenge of resource optimization. Existing infrastructure in cultural sites demands tailored maintenance, posing challenges for both large sites unable to provide adequate services and smaller sites constrained by surveillance and maintenance costs. Mobile computing technologies present an avenue to surmount limitations, augmenting the visitor experience at cultural sites while enhancing accessibility to digital heritage collections

III. RELATED WORKS

Unprecedented computational resources and burgeoning datasets have propelled the evolution of deep learning (DL) techniques. These algorithms, a subset of machine learning (ML), empower machines to comprehend complex concepts by hierarchically interpreting input objects. Convolutional Neural Networks (CNNs), a prominent DL model for image processing, have relevance across multiple domains, such as facial authentication, handwriting recognition, and medical image analysis.

CNNs' adoption extends to diverse domains, including object detection in street-level imagery and architectural landmark classification. Entities like Google have released sizable datasets and challenges, underscoring the growing interest in landmark recognition. However, datasets specifically focused on architectural elements remain relatively limited.

IV.DATASET

The project focuses on recognizing Indian monuments, encompassing a diverse array of landmarks scattered across the country. Unlike typical object recognition datasets, capturing images of Indian monuments presents unique challenges. These monuments often remain fixed in position, resulting in visually similar images, making it arduous for models to discern them from their surroundings. Moreover, spatial and historical proximity among monuments contributes to recurrent occurrences of similar backgrounds and architectural features.

The dataset creation commenced from scratch, capturing images not only of renowned structures but also encompassing less prominent buildings. The dataset labels images of Indian monuments, providing a comprehensive overview of their architectural attributes. Rigorous selection of viewpoints ensured coverage from diverse angles, encompassing common visitor spots, intricate details, and panoramic views. The dataset comprises images taken from both mobile phones and professional cameras, encompassing varying lighting and exposure conditions to challenge the DL model

V. ARTIFICIAL NEURAL NETWORKS

Artificial Neural Networks (ANNs) serve as the backbone of machine learning, approximating general functions. Inspired by the biological nervous system, ANNs comprise nodes representing neurons interconnected by edges. The information transfer among neurons involves nonlinear operations, often encompassing activation functions like Rectified Linear Unit. The Universal Approximation Theorem asserts that a feedforward neural network has the capability to simulate or mimic any continuous function by minimizing the error between actual and predicted values using weights and biases.

Choice of neural network parameters like the number of units and activation functions typically relies on empirical experience, aiming to optimize network performance. Stochastic Gradient Descent (SGD) proves efficient for large datasets, updating network parameters by backpropagation from the output layer to minimize error.

VI. DEEP LEARNING AND CONVOLUTIONAL NEURAL NETWORKS

Designing the architecture of neural networks lacks a definitive blueprint, although increasing units often enhances the network's function approximation capabilities. Deep feedforward neural networks, comprising multiple hidden layers, form a class of deep neural networks. Deep learning emphasizes the optimization of tensor operations to bolster network performance.

Training CNNs involves similar processes as feedforward neural networks, leveraging backpropagation to update weights based on computed gradients. This encompasses inference, gradient computation, and weight updates using techniques like statistical gradient descent, coupled with optimization and regularization.

VII. METHODOLOGY

1. **DATA PREPARATION**: Fundamental to Machine Learning (ML) success is meticulous dataset preparation involving tasks such as data selection, augmentation, formatting, and cleaning. Manual verification of images, highlighted by feature extraction, facilitates effective learning.

2. **NETWORK TRAINING**: Convolutional networks, adept at classification, leverage convolutional layers to extract input image features, culminating in a fully connected network for classification. Training involves parameter tuning and optimization for the fully connected layer, ensuring compatibility with mobile devices.

3. **BUILDING THE MODEL**: Transitioning the inference model for use on websites or web apps involves converting the model using TensorFlow Lite, an open-source DL framework included in TensorFlow.

IX. CONCLUSIONS

Artificial Intelligence (AI) is transforming interactions with urban environments. Existing search engines predominantly rely on text-based or partially image/sound-based searches on unorganized datasets, creating limitations in accessing cultured or academic information related to archaeological or architectural objects.

The vast repository of Indian architectural and archaeological heritage merits accessible pathways, initiated by real objects, to access abundant knowledge, studies, and stories. By employing this method, we not only mend between virtual and physical environments but also serves as an educational tool, redefining cultural awareness and knowledge of ancient and modern architecture.

This fusion of AI technology and multidisciplinary fields lays the groundwork for a paradigm shift in societal norms, blurring boundaries between humanistic, scientific, and artistic disciplines.

X. REFERENCES

[1] https://www.tensorflow.org/tutorials.
[2] https://keras.io/
[3] https://pytorch.org/tutorials/
[4] https://developer.mozilla.org/en-US/
[5]https://www.coursera.org/specializations/deep-learning.
[6]https://www.coursera.org/learn/convolutional-neural-networks.