Online-Journals.org

iJIM International Journal of Interactive Mobile Technologies

iJIM | eISSN: 1865-7923 | Vol. 19 No. 3 (2025) | OPEN ACCESS

https://doi.org/10.3991/ijim.v19i03.50935

PAPER

Deep Learning for Cultural Heritage: A Mobile App for Monument Recognition Using Convolutional Neural Networks

Md. Samaun Hasan[1–4](✉), Md. Salah Uddin[1], Kazi Jahid Hasan[1], Mizanur Rahman[1], Mohammad Ali[2]

[1]Daffodil International University, Dhaka, Bangladesh

[2]Rajshahi University, Rajshahi, Bangladesh

[3]Jahangirnagar University, Savar, Bangladesh

[4]Visva-Bharati University, Santiniketan, West Bengal, India

hasan.mct@diu.edu.bd

ABSTRACT

Convolutional neural networks (CNN) has multi-dimensional features that are inextricably interrelated to the extraordinary identification of visionary objects. Humans have a thirst of eagerness to know about ancient monuments. CNN can be a suitable tool for this purpose, in addition to some image processing approaches. This study has developed a mobile application for monument identification in two classes. The prominent contributions have been performed in the data preprocessing and feature extraction requirement using the Canny approach. In addition, CNN has been applied for the utmost level of recognition of the inputted visionary image to the mobile application using CNN. The proposed mechanism has obtained 90.68% accuracy in the testing period as its utmost accuracy result. The coagulation of CNN and other cutting-edge technology has enhanced the performance of the developed application. Moreover, these features have introduced this implementation as a comparatively desired and required app in daily utilization.

KEYWORDS

architectural heritage, convolutional neural network (CNN), mobile apps, deep learning

1 INTRODUCTION

South Asia's Indian subcontinent is a distinctive region. This continent contains many historic tourist attractions, including the Somapura Mahavihara, the Sona Mosque, the Kusumba Mosque, the Sixty Dome Mosque, the Kantajir Temple in Bangladesh, the Taj Mahal, the Qutub Minar in India, the Tomb of Jahangir in Pakistan, and others [1]. Numerous mysterious and unknown ancient structures are gradually coming to people's attention. Explorers are excited to observe the ongoing ancient building uncovering because of that [2]. Remote travelers occasionally struggle to find an informed visit manager [3]. In this situation, it is challenging for tourists to find out the exact date that mosques, shrines, and other old buildings were built. In this light, we have looked

Hasan, M.S., Uddin, M.S., Hasan, K.J., Rahman, M., Ali, M. (2025). Deep Learning for Cultural Heritage: A Mobile App for Monument Recognition Using Convolutional Neural Networks. *International Journal of Interactive Mobile Technologies (iJIM)*, 19(3), pp. 22–40. https://doi.org/10.3991/ijim.v19i03.50935

Article submitted 2024-07-05. Revision uploaded 2024-10-24. Final acceptance 2024-11-23.

into an investigation in which mobile application app users understood the development time frame or period by spotting the stunningly antiquated style [4]. Recently, certain research papers that use artificial intelligence (AI) and PC vision techniques in the section on outdated engineering and archaic exploration have been published [5].

1.1 Overview

In terms of quality and type, the city is currently creating fresh information about itself in the form of data that can be saved, categorized, and evaluated. Standardization of format, information accumulation, administration, and selection, data processing, and data visualization are all current technology needs [11]. This is driving the tools' development to unprecedented levels.

It is possible to quickly define new networks of linkages between documents following changing requirements and contents. The physical form of architecture continues to play a significant part in this complex information topology [23], which has been expanded by digital archives and can serve as the foundation for even the most modern elaborations [7]. A particular method of automated architecture recognition may alter how information about the urban environment is gathered, processed, and analyzed. It may also open up new avenues for access to a spatialized information network [8].

This study represents the era identification of historical buildings, as ancient buildings seek to strengthen the relationship between the real and digital worlds [9]. The development of an educational tool based on object identification would be well suited for the heritage sector given the number of projects on the digitalization of cultural assets and associated archives that have already been initiated. We are researching the AI to address these problems [10]. Convolutional neural network (CNN) approaches are specifically discussed [6]. These digital tools can progressively interact with urban circumstances because they are widely used, especially on mobile devices.

The study seeks to broaden its inquiries from this initial trial to broader AI-related issues. Can we create new types of maps using this technology's ability to sift and organize massive volumes of data? What are the drawbacks and dangers of automatic architectural recognition systems [11]?

We used a CNN-based program in a mobile application for the Archaeological Central Area in South Asia's Indian subcontinent, which is a distinctive region. that includes the Somapura Mahavihara, the Sona Mosque, the Kusumba Mosque, the Sixty Dome Mosque, the Kantajir Temple in Bangladesh, the Taj Mahal, the Qutub Minar in India, and the Tomb of Jahangir in Pakistan, as our first case study [8].

1.2 Related research

The noteworthy implementations are in the following:

- Travelers are highly eager about seeing the recently unveiled ancient architecture, and they want to know details about that [7], [12], [13].
- Sometimes, tourists cannot find an informative tour guide, and it has become difficult to acquire knowledge of the architecture [8], [9], [10].
- At present, if archaeological materials have been found network-based, it has become difficult to detect their period [8], [11], [12], [13].

1.3 Research scope

The existing literature did not introduce any well-developed approach that can be used as an identifying tool for understanding the status of the ancient architectures. This study has provided an effort to mitigate this study gap by developing a mobile-oriented application for the tourists in order to get the exact knowledge about their desires.

1.4 Research applicability

- Archaeological materials era detection [13], [14], [15].
- Mobile or web-based applications to check the era by capturing an image [15], [16], [17], [18].
- Architectural preservation and historical documentation [16], [17], [18], [19].
- Machine learning and image processing techniques [18], [19], [20], [21], [24].
- Comparative architectural studies [19], [20], [21], [22], [25], [26].

1.5 Research questions and answers

Research question 1: How effective is CNN in identifying and categorizing ancient monuments through mobile applications?

Answer: CNN proves to be highly effective in identifying and categorizing ancient monuments, achieving a 90.68% accuracy rate in testing. This effectiveness is largely attributed to the robust feature extraction and data preprocessing capabilities of CNN, particularly when combined with image processing techniques such as the Canny approach [6].

Research question 2: What challenges do travelers face in obtaining accurate information about ancient structures, and how does the mobile application address these challenges?

Answer: Travelers often struggle with the lack of informed tour guides and difficulty in accessing accurate historical information. The mobile application addresses these challenges by providing an easily accessible, accurate, and interactive platform for monument recognition and information retrieval, thus enhancing the travel experience [3], [4].

Research question 3: What role does the integration of cutting-edge technologies play in the performance and desirability of mobile applications for cultural heritage?

Answer: The integration of cutting-edge technologies such as CNN and advanced image processing significantly enhances the performance and desirability of mobile applications for cultural heritage. These technologies improve accuracy, user engagement, and the overall educational value of the application, making it a preferred tool for both casual users and scholars [10], [11].

2 METHODOLOGY

Using an application-based methodology, our research has shown how mobile application can determine the age of an old structure. An image was initially captured and submitted to the Canny Edge Detector. Following this methodology, two distinct dataset types were obtained—a training set and a test set. Then, to get the prediction outcome for earlier eras (the Sultanate period and the Mughal period), we developed a CNN model that depicts the proposed system's stages and design shown in Figure 7.

3 BUILDING STRUCTURE AND FEATURE IDENTIFICATION

The results of the suggested model construct of realizing the old era and a machine give likely data by learning the features of the ancient buildings. Theoretically, a programmable machine can learn the characteristics of various historical structures, including the dome/minaret/tower/jewel, door, pillar, arc, and front (see Figure 3). The Mughal and Sultanate architecture's dome, minaret, tower, jewel, pillar, arch, and front (see Figures 1 and 2) were obtained by utilizing the Canny Edge Detector method (see Figure 4) [7]. The structure, décor, and materials used during the Sultanate and Mughal periods varied significantly from one another.

The Mughal mosques are only ornamented with plaster in various styles, such as rectangle, arched panel, flat, four-centered, or multi-foiled arches. In contrast, the Sultanate Mosques in Bengal have a unique exposition of surface decoration, including stone-cut design, terracotta, colored marble, and pleasurable ornamentation due to the absence of readily available building materials, such as stone, within proximity.

4 COLLECTION OF DATA TRAINING

Based on a training procedure in which the AI learns from a set of manually labeled photographs, AI technology is capable of giving labels to input images. 17 monuments of varying sizes and conservation quality have been trained into the model. The Somapura Mahavihara, the Sona Mosque, the Kusumba Mosque, the 60 Dome Mosque, the Kantajir temple in Bangladesh, the Taj Mahal, the Qutub Minar in India, the Tomb of Jahangir in Pakistan, a poorly preserved temple of which not much can be seen except for the podium. Our dataset consists of an initial collection of approximately 150–200 images of each monument that were taken specifically for our purpose (see Figure 6). The stances were chosen to provide a thorough overview of the monument, and special consideration was paid to the visitor-friendly settings, taking both little details and expansive views into account.

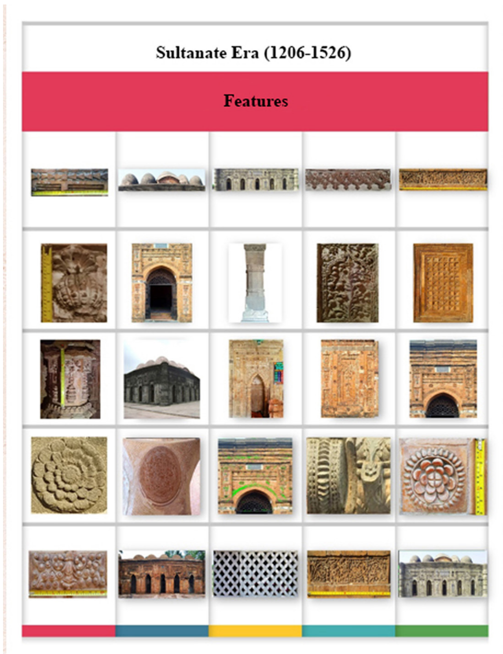

Fig. 1. Raw data of the sultanate feature

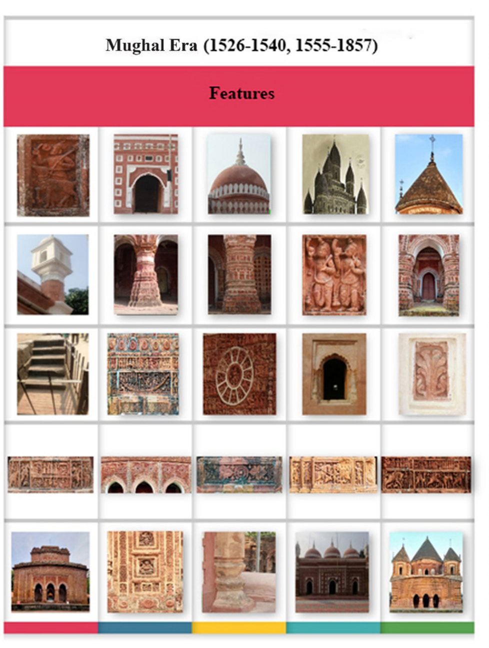

Fig. 2. Raw data of the Mughal feature

The first set of images was expanded, resulting in changed versions of the originals, to build a larger dataset. DL object recognition has found data augmentation strategies to be especially effective [8]. The operations carried out were transposition, rotation, scaling, noise injection, and color changes. Each landmark had 500 photos in our expanded dataset. The training method required fixed input dimensions; thus, all of the images were eventually cropped and down sampled to a size of 224 × 224 pixels. Images were arranged into named files, each one relating to a different monument, to be utilized in the training phases.

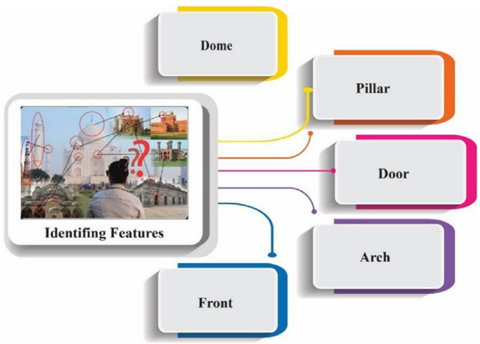

Fig. 3. Identifying five features

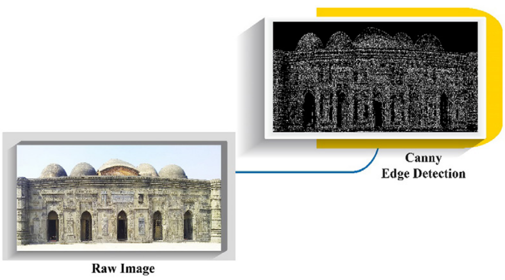

Fig. 4. The Canny Edge Detector method

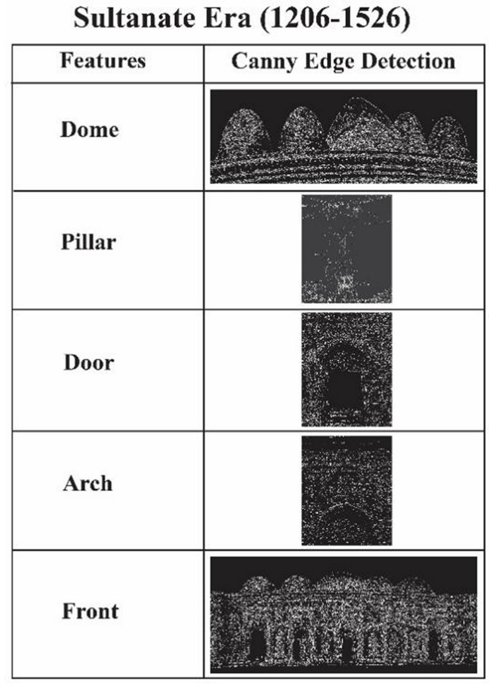

Fig. 5. Training dataset and classification of sultanate features

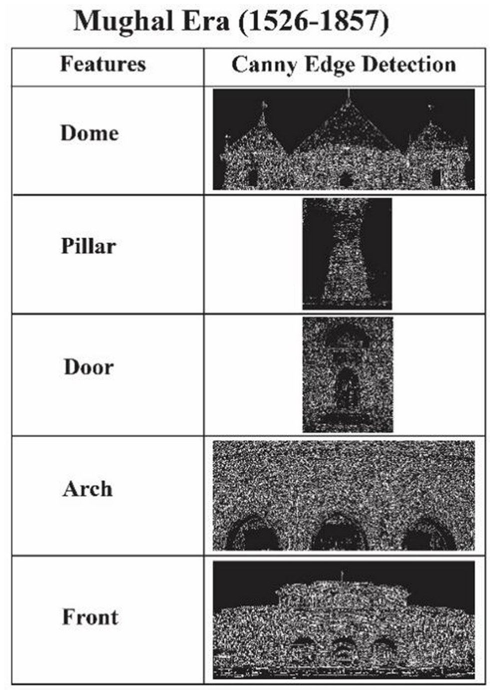

Fig. 6. Training dataset and classification of Mughal features

5 DESCRIPTION OF THE APP

Cultural sites and related archives that have been digitalized provide an organized and usable source of data that may be obtained using the adaptable features of DL object recognition algorithms.

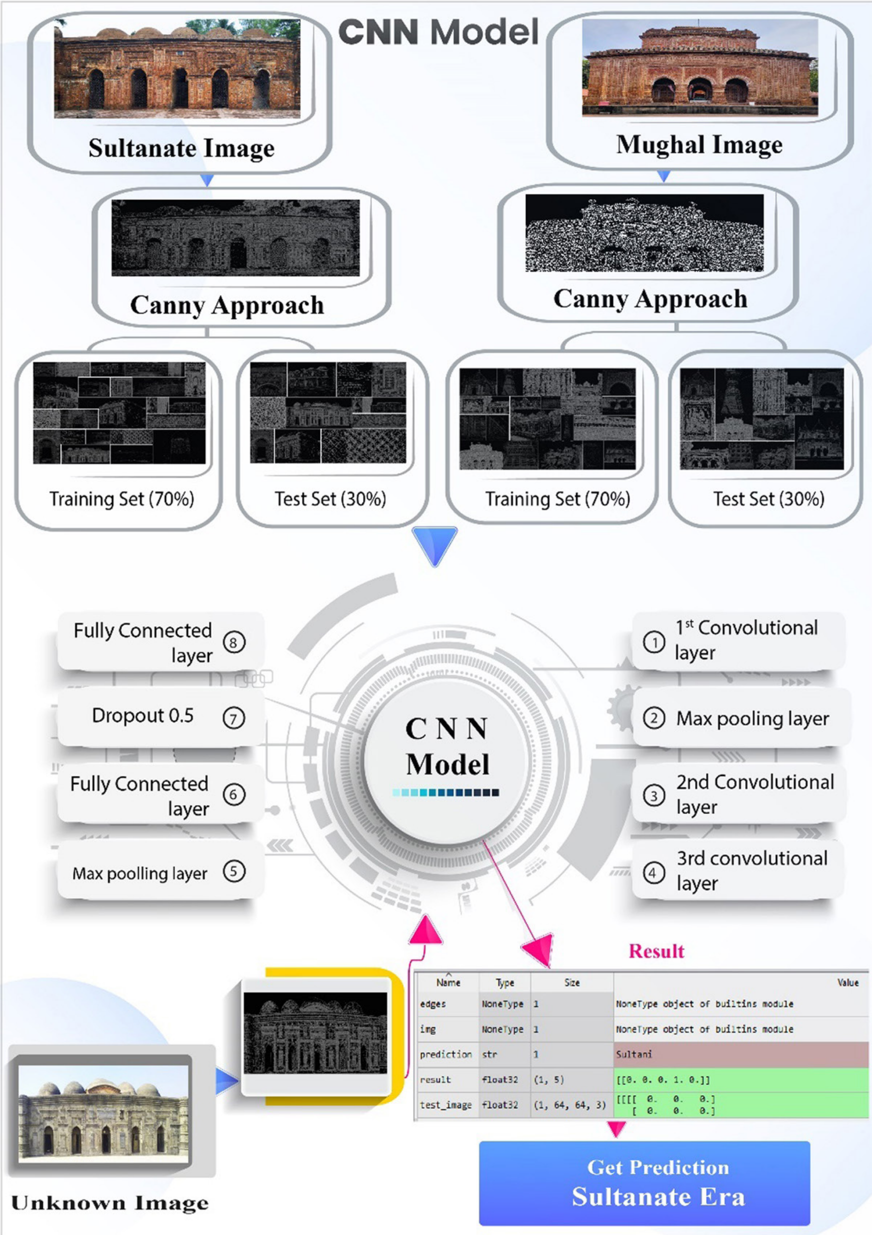

Fig. 7. CNN model for era identification of the old building

The current study created a mobile app that enables users to get information about a structure or piece of art by simply aiming the camera at the item. The end result can link a collection of papers, including images, texts, typography [13], 3D models, and other types of data, to the actual construction period. It expands on early advances of CNN methods for identifying architectural features and objects.

The app is built around two basic software building elements (see Figure 7). A database containing the following characteristics is the first one. It is accessible online, allows the upload of various document types together with any associated metadata, and is geographic-enabled, meaning that it can be examined and modified using GIS tools [12]. The ML component is a separate piece of software. It offers the following characteristics: It downloads monument data from an internet database, is kept on the device, uses a tiny amount of disk space, and localizes the user.

Figure 7 shows the layout of the app's two primary parts: the digital archive, which stores data and documents, and the AI engine, which communicates with the site via the app.

Instead of using a catalog or a predetermined path on a map, the app becomes the access point to an underutilized network of digital information simply by glancing at the urban setting. Questions about how to make the newly available information conveniently and sustainably accessible can be answered by linking a great amount of multimedia information to the components of a city that is part of our sub-content.

The lightweight "brain" takes advantage of a distinguishing feature of CNN models: the minimal amount of disk storage required [4]. In reality, DL models learn a way to recognize objects rather than obtaining their "knowledge" from hard-coded information (i.e., no pictures are retained). This method is especially suited to travelers in the suggested application since they do not have access to the Internet while traveling. The use of the part of the Indian subcontinent Archaeological Central Area as a test area presents an intriguing problem. On-site informational service provision is a challenge that is addressed. Informational panels, audio tours, or QR codes cannot easily complete this role since the required infrastructure may not be compatible with the priceless site and the volume of visitors.

The proposed app can make a lot of data available with just a smartphone. This material must therefore be accurate and comprehensive as well as suitable for a variety of consumers. In contrast to numerous other approaches that may be found in the literature, using CNNs on monuments inside a constrained boundary presents another issue. It might be challenging to tell one element from another when they are all present in the same cultural place or appear together in the same view. These circumstances, however, emphasize the practical utility of AI more so than the ability to recognize locations that are extremely distant from one another. Therefore, our research implies that it may be possible to use the most recent computer vision algorithms to obtain more detailed information.

6 DEVELOPMENT OF THE APP

Based on a training procedure in which the AI learns from a set of manually-labeled photographs, AI technology is capable of giving labels to input images. 10 monuments of varying sizes and conservation quality have been trained into the model. The model could potentially analyze input images of these monuments and provide labels such as Somapura Mahavihara, Sona Mosque, Kusumba Mosque, Bagha Mosque, Atiya Mosque, Sixty Dome Mosque, Khania Dighi Mosque, Bagerhat Kantajew Temple, Taj Mahal, Qutub Minar, or "Tomb of Jahangir," based on what it has learned during its training.

Preserving and understanding the historical significance of heritage buildings is of paramount importance. This app aims to provide users with a tool to easily identify the era of such buildings using their smartphone cameras. The utilization of digital image processing and machine learning enhances the accuracy and accessibility of this heritage identification process.

6.1 App design

The app is designed with a user-friendly interface that allows users to capture images, process them, and receive era-related information. The app's main components include the camera interface, image processing backend, and result display.

6.2 App development

Step 1: Capture image: Utilize Android's camera API to provide a seamless image capture experience. Implement a capture button within the app's user interface for users to capture images of heritage buildings. After successfully capturing the input images, all the images have been manually resized to 64×64 general basic input image sizes. The operation is especially for minimizing computational expenses and enhancing efficiency in high-resolution image processing with proper pixel understanding. This transformation is also for effectively implementing the obstacles in interference and fast training strategy with the applied convolutional neural model. One of the most important matters behind this processing is to avoid overfitting problems along with under-fitting problems in the collected dataset. Even this transformation provides aid to handle feature extraction along with the feature transformation mechanism indeed.

Step 2: Image processing with CNN mode: Pre-train a CNN model on a dataset of heritage building images from various eras. Integrate the pre-trained CNN model into the app's backend using a machine learning framework such as TensorFlow or PyTorch. Implement image preprocessing to resize and normalize captured images before feeding them into the CNN model. This effective method removes the requirement for manually designing features by enabling autonomous feature extraction from photos. For applications such as categorization, object identification, and segmentation, CNNs employ layers of CNNs to identify patterns such as edges, textures, and forms. They are very good at visual recognition because of their hierarchical nature, which enables the neural network to learn complicated representations. Because CNNs can interpret images quickly and reliably, they are frequently utilized in applications including recognizing images.

Step 3: Display era information: Design a result display screen to show the detected era of the heritage building. Integrate the CNN model's output with the display, presenting the era information in a user-friendly format. Early Islamic architecture, characterized by strong, fortress-such as constructions such as the Qutub Minar as well as the Tughlaqabad Fort, was introduced during the Sultanate century (13th–16th century). Grand, elaborate structures with elaborate marble inlays of diamonds, domes, and vast gardens, such as the Taj Mahal, along with the Mughal era (16th–18th century). Both eras had a profound impact on the architectural heritage of the area, influencing subsequent advancements in both art and architecture. These monuments continue to be recognizable representations of historical and cultural importance.

6.3 CNN model training

Collect a diverse dataset of heritage building images from different eras of the Indian subcontinent. Preprocess the dataset, including resizing images to a consistent resolution and applying data augmentation techniques. Design and train a CNN model using a suitable architecture (e.g., VGG16, ResNet) and transfer learning from pre-trained models if needed. Evaluate the model's performance using appropriate metrics and fine-tune hyperparameters as necessary.

6.4 App testing and optimization

Perform rigorous testing of the app's functionalities, covering various scenarios and image qualities. Optimize the app's performance by minimizing latency during image processing and result display. Ensure the app's compatibility with a wide range of Android devices and screen sizes represented in Figure 8.

Fig. 8. Application and user interaction outcome on sultanate era monuments

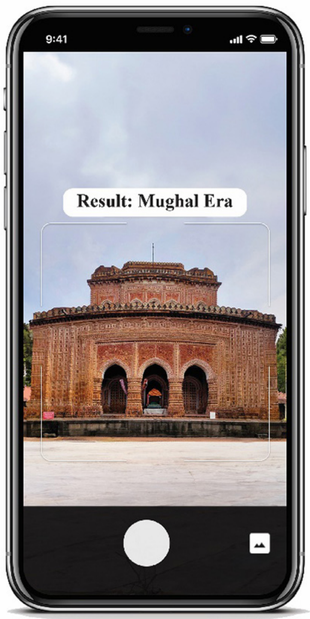

Fig. 9. Application and user interaction outcome on Mughal era monuments

6.5 User experience and interface design

Design an intuitive and visually appealing user interface to enhance user experience. Implement responsive design principles to ensure the app's usability on different devices represented in Figure 9. This study has arranged a survey activity against the utilization-based experience from the users, and the feedback against interface design has also been evaluated through this survey investigation. This investigation has also demonstrated the features that serve the needs of tourists and other users, along with that, navigate the features through experiences.

6.6 Deployment

Package the app for distribution on the Google Play Store. Provide clear documentation and user instructions for seamless onboarding. To differentiate between sculptures from each of these historical eras, the program examines architectural elements such as domes, arches, and elaborate patterns. The app allows users to upload pictures of monuments and offers precise details on the building's history and place of origin. In addition to being a useful tool for history buffs and scholars, this tool raises cultural awareness.

7 USER INTERFACE AND DATABASE CONNECTION

We worked on a user-friendly, quick-to-use app. The application has two primary displays or pages (see Figures 8 and 9).

The primary screen has a simple layout with backdrop images from the camera feed and a square drawn to enclose the target monument. The app only shows results when the computed prediction accuracy exceeds a set threshold value greater than 80 percent which is considered reliable. ML recognition runs continuously. Touching the squared frame when a monument name is displayed takes you to the monument's details.

The information set previously downloaded from the web database is laid forth on the monument details page.

The application communicates with the Cult database, where information about the monuments covered by the ML training dataset was added (see Figures 5 and 6). Web-based services are hosted by Cult and allow users to post texts, photos, and models as well as attach metadata to each file. It also enables content exploration via a website or through the use of GIS software. The service is powered by free and open-source software, including the Django web development framework, PostgreSQL for database management, and PostGIS for geographic expansion [9].

The Dublin Core Metadata Initiative provides the foundation for the organization of all metadata. Title, creator, subject, description, publisher, contributors, date, type, format, identifier, source, language, relation, coverage, and rights are just a few of the 15 entries of the DCMI Element set that can be inserted. Additional terms include the contributors' roles, spatial coverage (using WKT format geometries or GIS), and the type of relationship between one document and another. The software can recognize objects without a network connection. Only when opening for the first time or updating local data, as well as downloading remote data from the online database, is an internet connection necessary. The app makes a web service request to access data saved on Cult.

8 RESULT AND ANALYSIS

We used the information in the matrix to assess how well these systems performed. With our adjusted dataset as training data, we determined the accuracy of our CNN model (see Figure 7). Sample Dataset 1: First, we used a single image to make a single prediction in our CNN model, correctly predicting the Sultanate Table 1 and Mughal Table 2 period buildings. Finally, we achieved a single image accuracy rate of 95% [10].

The model's outputs include detecting the vintage of construction, for which a computer algorithm generates a likely result by studying the characteristics of historic structures. This work has demonstrated how computer software picks up on the various characteristics of antique structures, such as the front, dome, and minaret. Confusion matrices contain information about the expected classes that are established by a classification process and are mostly used to evaluate the performance of a classifier model. The information in the matrix has been used to assess the effectiveness of such systems. With the adjusted dataset, the CNN model was trained, and accuracy was determined. Figure 7 depicts the CNN model's makeup and illustrates how the method effectively identified the era from an image of an ancient or historic building.

Table 1. Example table for Sultanate Era

Name	Type and Size	Value
edges	uint8 (800,1200),	array ([[0,0, 0, ...0,0,0],
		[0, 0, 0,..., 0, 0, 0],
i	int (1)	1
img	uint8 800,1200,3),	Array ([[[210, 186, 144, ...
		213, 189, 147],
prediction	str (1)	Sultanate
\text{result}	float32 (1,2, 3),	Array ([[0.,1]],
		dtype=float32)
test-image	float32 (1,64,64,3),	Array ([[0.,0.,0.,],
		[0., 0., 0.],
edges	uint8 (800,1200),	Array ([[0,0,0, ..., 0,0,0],

Table 2. Example table for Mughal Era

Name	Type and Size	Value
edges	uint8 (800,1200),	array ([[0,0,0,..0,0,0],
		[0, 0, 0, ..., 0, 0, 0],
i	int (1)	1
img	uint8 (800,1200,3),	array ([[[210,186,144, ...
		213, 189, 147],
prediction	str (1)	Mughal
result	float32 (1, 2),	array ([[1.,0]],
		dtype=float32)
test-image	float32 (1,64,64,3),	Array ([[[0.,0.,0.,],
		[0., 0., 0.],

The proposed model has been applied to 409 images categorized into two classes, such as Mughal and Sultanate, during its training phase. After that, the model has been tested on 79 images, also following the same classification. During his training phase, it received the accuracy from 66% to 86%, depicted in Figure 10. The yellow line indicates the validation accuracy on the provided images of the proposed model. Similarly, the blue line indicates the training accuracy.

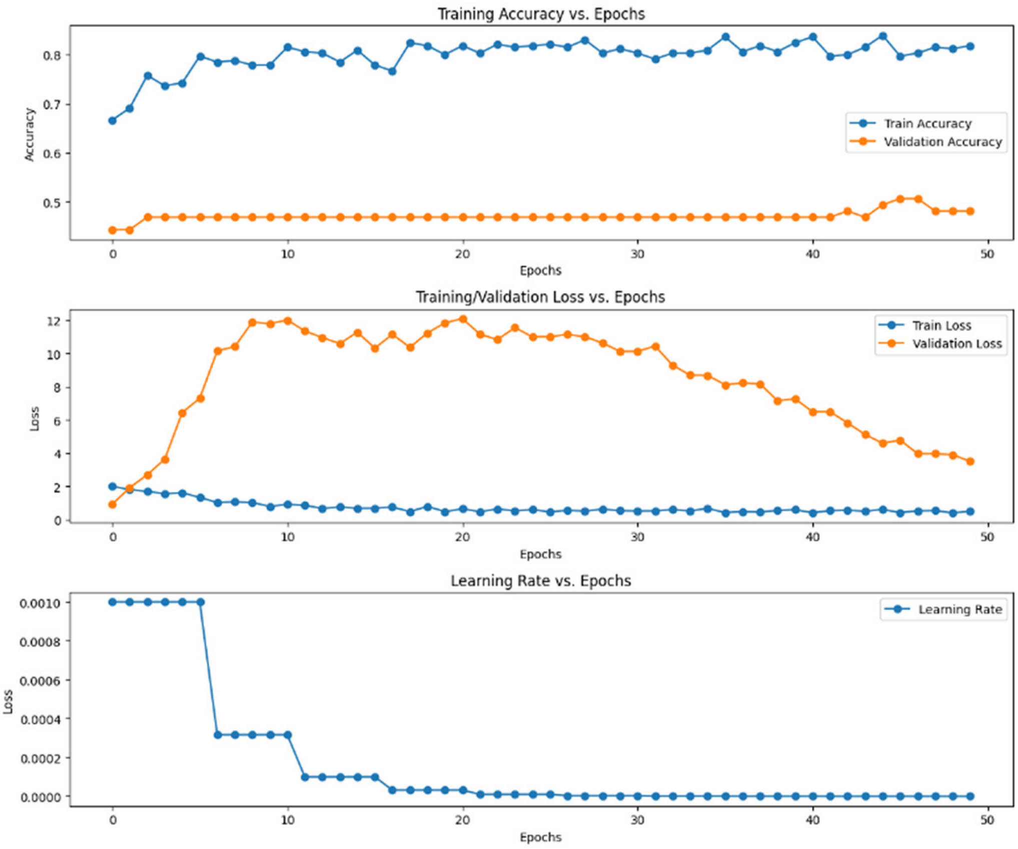

Fig. 10. The results on train accuracy vs. validation accuracy

Eventually, the model has been tested on the 79 images categorized into two classifications and received 90.68% accuracy. The model has provided an effort to build an application that can identify the archaeological architectures provided by the users as a tourist. Generally, this is the start of developing an application, it will be severely optimized and user-friendly day by day. Initially, this study has faced some of overfitting and under-fitting problems during building the mechanism for the application.

9 EVALUATION

Based on a number of important criteria, the software for identifying Mughal imperial Sultanate periodic architecture is capable of being contrasted with current methods. The software uses AI-based image recognition to streamline this process, providing faster and more reliable results than traditional approaches that rely on expert manual identification, which may be laborious and prone to human mistake. This program is more accurate in differentiating among Mughal as well as Sultanate elements because it is specifically designed for South Asian construction, compared to universal historical recognition programs.

It also makes use of deep learning models, which are constantly enhanced by fresh data, a major benefit over rule-based or static databases. The software makes monument identification easier for scientists and the general public by facilitating simple utilization via mobile devices, whereas existing methods may necessitate extensive background information or in-person trips to archives. It is also better than conventional manual approaches due to its scalability and capacity to process enormous amounts of image information. This study has arranged a survey activity against the utilization-based experience from the users, and the feedback against interface design has also been evaluated through this survey investigation. This investigation has also demonstrated the features that serve the needs of tourists and other users (see Figure 11), along with that, navigate the features through experiences (see Figure 12).

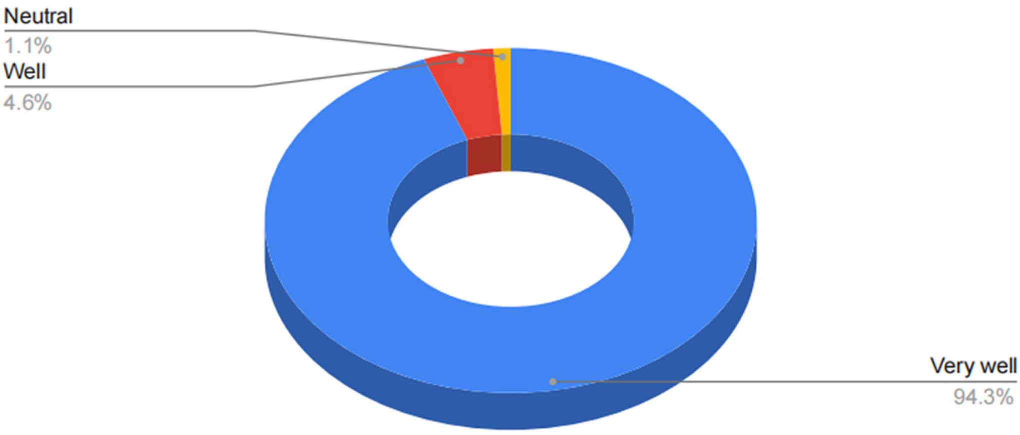

Fig. 11. Demonstration of features that serve the needs

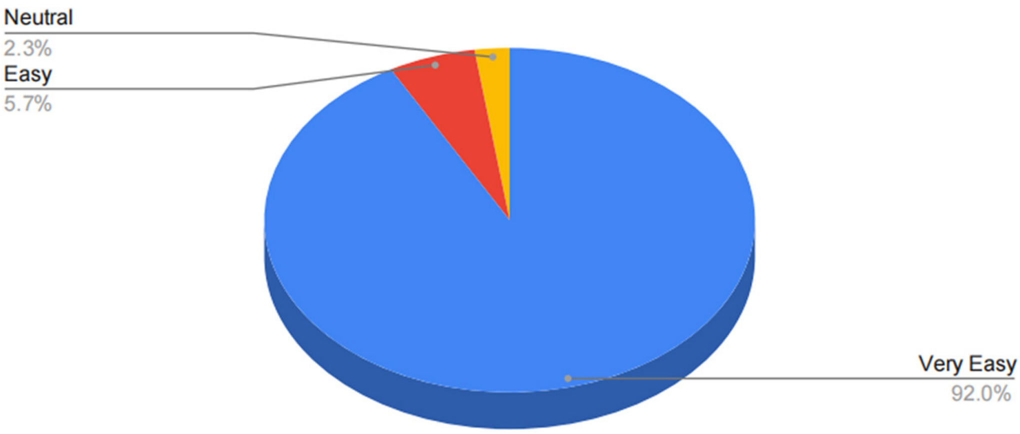

Fig. 12. Demonstration of features that navigate experiences

By employing convolutional structures that apply tiny filters over the image and utilizing fewer parameters than fully linked networks, CNNs lower the computing costs associated with the processing of images. These layering' parameter sharing reduces the need for duplicate calculations by enabling the exact same filter to identify features in several areas of the image. In order to reduce the image's spatial dimensions and, thus, its computing cost, CNNs also use layers to combine to reduce the sample size of the image. Furthermore, CNNs are more effective for handling vast amounts of picture data because of their intrinsic sparse connections, which simplify processes. With fewer resources, this structure guarantees quicker learning and deduction. The precision as well as the efficiency of the memorial detecting software can be greatly increased by using transformer-based models or combination methods that combine CNNs. transformers can use self-attention methods to capture global context and long-range dependencies, while CNNs are best at extracting local information from images. By combining the two, the application may be able to improve classification performance by utilizing the Transformers' comprehension of relationships throughout the image and CNNs' spatial feature extraction. Additionally, by allowing the system to identify monuments under various illumination or partially obscured circumstances, hybrid models can improve generalization and increase the robustness and dependability of identification. This combination may help the software become more flexible and scalable in real-world settings.

The improved system might employ a multi-class classification strategy, making use of hybrid models or deeper CNNs to more precisely differentiate between different architectural styles. To enable the program to detect minute architectural differences, this would entail training on a varied dataset that represents many historical eras and geographical locations. To increase classification accuracy, the system might also include information, including information about the monument's geography and composition. The software would also be a potent teaching tool, encouraging the preservation of cultural assets by providing deeper insights. The software would be positioned as a complete system for historic preservation studies with this expansion.

Customers would be able to experience historical landmarks in a more dynamic and captivating manner if the app included augmented reality (AR) technologies. Users could examine monuments with more context, putting the past to life, by overlaying current photographs with comprehensive historical details, architectural points of interest, and restoration images. This function might lead users on chronological tours, highlighting important design features or historical shifts right on the screens of their devices. Vacationers, learners, and history enthusiasts would find the app more enticing due to its immersive AR experience, which would promote deeper investigation and study. AR may also encourage a better understanding of history and culture, which would support efforts to save monuments.

The app's study base can be greatly enhanced through the incorporation of transdisciplinary and application-oriented sources of information, guaranteeing a more comprehensive comprehension of the types of buildings under analysis. Deeper understanding of the monuments' historical context can be gained using the app by combining information from disciplines including past times, archaeological investigation, and cultural anthropology. Additionally, this would increase the precision of architectural assessments, enabling people to comprehend the subtle differences between styles such as Sultanate, Mughal, and others. Improving participation requires concentrating on the user interface; readability could be improved by features such as tailored suggestions, engaging lessons, and easy browsing. Furthermore, using contemporary construction-related apps such as 3D modeling and AR would improve the app's usefulness. In the end, this strategy would guarantee that the app stimulates individuals in important manners using transdisciplinary understandings in addition to being a potent instructional tool.

10 CONCLUSION

This study introduces a pioneering approach, utilizing CNNs to detect architectural heritage eras in the Indian subcontinent. The mobile app developed by this study showcases the practical application of deep learning, providing a valuable tool for travelers, scholars, and enthusiasts to accurately identify the construction periods of historic buildings. This study bridges the gap between physical and digital realms, enhancing our understanding of architectural history and enabling innovative applications in heritage preservation.

The app's usefulness would be significantly increased if it could identify a greater variety of historical eras and designs in architecture in addition to binary classification. Users could learn more about a wider range of historical sites by adding more varied data and teaching the model to categorize other architectural styles, such as Gupta, Rajput, or British colonialism. This would create a more complete repository for artistic and chronological investigation, making the app useful not only for researchers who are interested in particular eras but also for visitors, historians, and architecture fans. This adaptability would also draw in more users, increasing the app's effect and reach.

This study is an effort to develop an application that can be a realm of information for tourist users. This study has a future plan of enhancing the area of identification for the archaeological architecture broadly. In the concern of technology, this study will integrate AI based techniques into the application for making the apps higher level, user-friendly, and informative. More specifically, sentiment analysis and recommendation systems will be integrated into the proposed application.

11 REFERENCES

[1] S. Su and T. Nawata, "Demolished building detection from aerial imagery using deep learning," in *Proceedings Int. Cartogr. Assoc.*, vol. 2, 2019, p. 112. https://doi.org/10.5194/ica-proc-2-122-2019

[2] A. J. Ghandour and A. A. Jezzini, "Autonomous building detection using edge properties and image color invariants," *Buildings*, vol. 8, no. 5, p. 65, 2018. https://doi.org/10.3390/buildings8050065

[3] B. Wen *et al.*, "Monitoring the green evolution of vernacular buildings based on deep learning and multi-temporal remote sensing images," *Building Simulation*, vol. 16, pp. 151–168, 2023. https://doi.org/10.1007/s12273-022-0927-7

[4] V. Palma, "Towards deep learning for architecture: A monument recognition mobile app," in *the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XLII, pp. 551–556, 2019. https://doi.org/10.5194/isprs-archives-XLII-2-W9-551-2019

[5] M. S. Hasan *et al.*, "Identification of construction era for Indian subcontinent ancient and heritage buildings by using deep learning," in *Proceedings of Fifth International Congress on Information and Communication Technology, ICICT 2020,* Advances in Intelligent Systems and Computing, X. S. Yang, R. S. Sherratt, N. Dey, and A. Joshi, Eds., Singapore: Springer, vol. 1183, 2020, pp. 631–640. https://doi.org/10.1007/978-981-15-5856-6_64

[6] S. Katsigiannis, S. Seyedzadeh, A. Agapiou, and N. Ramzan, "Deep learning for crack detection on masonry façades using limited data and transfer learning," *Journal of Building Engineering*, vol. 76, no. 1, p. 107105, 2023. https://doi.org/10.1016/j.jobe.2023.107105

[7] H. C. Nguyen, T. H. Nguyen, R. Scherer, and V. H. Le, "Deep learning for human activity recognition on 3D human skeleton: Survey and comparative study," *Sensors*, vol. 23, no. 11, p. 5121, 2023. https://doi.org/10.3390/s23115121

[8] N. I. Nife and M. Chtourou, "Improved detection and tracking of objects based on a modified deep learning model (YOLOv5)," *International Journal of Interactive Mobile Technologies (iJIM)*, vol. 17, no. 21, pp. 145–160, 2023. https://doi.org/10.3991/ijim.v17i21.45201

[9] M. S. Hasan *et al.*,"Heritage building era detection using CNN," in *IOP Conference Series: Materials Science and Engineering*, vol. 617, 2019, pp. 1–6. https://doi.org/10.1088/1757-899X/617/1/012016

[10] K. I. Jabbarova, "Application of WASPAS method to data platform selection under Z-valued information," in *15th International Conference on Applications of Fuzzy Systems, Soft Computing and Artificial Intelligence Tools – ICAFS-2022*, ICAFS 2022, Lecture Notes in Networks and Systems, R. A. Aliev, J. Kacprzyk, W. Pedrycz, M. Jamshidi, M. B. Babanli, and F. Sadikoglu, Eds., Cham: Springer, vol. 610, 2023, pp. 557–563. https://doi.org/10.1007/978-3-031-25252-5_73

[11] S. M. Allayear, M. F. A. Bhuiyan, M. M. Alam, S. R. Kabir, M. T. A. Munna, and M. S. Hasan, "Human face detection in excessive dark image by using contrast stretching, histogram equalization and adaptive equalization," *Int. J. Eng. Technol.,* vol. 7, pp. 3984–3989, 2018. https://www.sciencepubco.com/index.php/ijet/article/view/13713

[12] M. Rahman *et al.*, "Anatomical analysis between two languages alphabets: Visually typographic test transformation in morphological approaches," in *Advances in Computing and Data Sciences: 4th International Conference*, ICACDS 2020, Valletta, Malta, Revised Selected Papers 4., pp. 330–339, 2020. https://doi.org/10.1007/978-981-15-6634-9

[13] W. Wang *et al.*, "Traditional village building extraction based on improved mask R-CNN: A case study of Beijing, China," *Remote Sensing*, vol. 15, no. 10, p. 2616, 2023. https://doi.org/10.3390/rs15102616

[14] F. Zhang, N. Saeed, and P. Sadeghian, "Deep learning in fault detection and diagnosis of building HVAC systems: A systematic review with meta-analysis," *Energy and AI*, vol. 12, p. 100235, 2023. https://doi.org/10.1016/j.egyai.2023.100235

[15] A. Copiaco *et al.*, "An innovative deep anomaly detection of building energy consumption using energy time-series images," *Engineering Applications of Artificial Intelligence*, vol. 119, p. 105775, 2023. https://doi.org/10.1016/j.engappai.2022.105775

[16] Z. Xu, F. Zhang, Y. Wu, Y. Yang, and Y. Wu, "Building height calculation for an urban area based on street view images and deep learning," *Computer-Aided Civil and Infrastructure Engineering*, vol. 38, no. 7, pp. 892–906, 2022. https://doi.org/10.1111/mice.12930

[17] X. Deng, Y. Liang, X. Li, and W. Xu, "Recognition and spatial distribution of rural buildings in Vietnam," *Land*, vol. 12, no. 12, p. 2142, 2023. https://doi.org/10.3390/land12122142

[18] L. Zhuang *et al.*, "RSI-YOLO: Object detection method for remote sensing images based on improved YOLO," *Sensors*, vol. 23, no. 14, p. 6414, 2023. https://doi.org/10.3390/s23146414

[19] Y. Ogawa, C. Zhao, T. Oki, S. Chen, and Y. Sekimoto, "Deep learning approach for classifying the built year and structure of individual buildings by automatically linking street view images and GIS building data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 1740–1755, 2023. https://doi.org/10.1109/JSTARS.2023.3237509

[20] B. A. Alabsi, M. Anbar, and S. D. A. Rihan, "CNN-CNN: Dual convolutional neural network approach for feature selection and attack detection on Internet of Things networks," *Sensors*, vol. 23, no. 14, p. 6507, 2023. https://doi.org/10.3390/s23146507

[21] L. Xia, S. Mi, J. Zhang, J. Luo, Z. Shen, and Y. Cheng, "Dual-stream feature extraction network based on CNN and transformer for building extraction," *Remote Sensing*, vol. 15, no. 10, p. 2689, 2023. https://doi.org/10.3390/rs15102689

[22] M. Wahi *et al.*, "A deep learning classification approach using high spatial satellite images for detection of built-up areas in rural zones: Case study of Souss-Massa region-Morocco," *Remote Sensing Applications: Society and Environment,* vol. 29, p. 100898, 2023. https://doi.org/10.1016/j.rsase.2022.100898

[23] S. Rawas and A. D. Samala, "Revolutionizing brain tumor analysis: A fusion of ChatGPT and multi-modal CNN for unprecedented precision," *Int. J. Onl. Eng. (iJOE),* vol. 20, no. 8, pp. 37–48, 2024. https://doi.org/10.3991/ijoe.v20i08.47347

[24] K. Parmar and I. K. Varma, "A comprehensive approach to enhancing doctor-patient interaction: Bridging the gap for better healthcare," *International Journal of Online and Biomedical Engineering (iJOE)*, vol. 20, no. 13, pp. 4–23, 2024. https://doi.org/10.3991/ijoe.v20i13.50345

[25] B. Fernandez and M. Javier, "Michael Hansmeyer's algorithmic architecture: The transformative impact of 3D printing on architectural design and education," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 19, no. 7, pp. 44–58, 2024. https://doi.org/10.3991/ijet.v19i07.50845

[26] E. Saliha and Ö. Y. Arzu, "New building design in historical environment: Production process with optimization-based algorithm," *International Journal of Architectural Computing*, 2024. https://doi.org/10.1177/14780771241299630

12 AUTHORS

Dr. Md. Samaun Hasan, an Assistant Professor in the Department of Multimedia and Creative Technology, Daffodil International University, brings a diverse background as an artist, art critic, connoisseur, historian, and faculty member with a passion for fostering creativity and critical thinking, Dr. Hasan is committed to enhancing student performance through innovative teaching methods. Eager to contribute to the academic sphere, he seeks challenging opportunities to showcase his expertise and dedication. Driven by a desire to cultivate a dynamic learning environment, he aims to inspire students to explore and push the boundaries of their creative potential. With a focus on interdisciplinary research, Dr. Hasan aims to bridge the gap between art and technology such as machine learning, computer vision and deep learning, contributing valuable insights to the field (E-mail: hasan.mct@diu.edu.bd).

Md. Salah Uddin serves as an Assistant Professor within the Multimedia and Creative Technology Department at Daffodil International University in Dhaka, Bangladesh. His academic credentials include a Master's degree in System and Software Engineering obtained from the National Research University – Higher School of Economics (NRU HSE) in Moscow, Russia. Currently, he is engaged as a PhD Research Fellow at Graz University of Technology in Austria. Mr. Uddin boasts a commendable track record of publications across various domains, including High-Performance Computing, the iSCSI Protocol, Big Data, Cloud Computing, Network Storage Solutions (including NAS, SAN, iSCSI), Android Games, and Musical Applications. Additionally, he has contributed significantly as a Google Mentor.

Kazi Jahid Hasan holds the position of an Assistant Professor within the Multimedia and Creative Technology Department at Daffodil International University in Dhaka, Bangladesh. With a Master's degree in Computer Science and Engineering from the same institution, he brings a wealth of expertise to his role. Mr. Hasan's diverse background includes proficiency as a 3D artist, Animator, and CG Artist, alongside his responsibilities as a faculty member. He has established a commendable track record of publications spanning various domains such as High-Performance Computing, Animation, Android Games, AR, VR UI and UX Design, and Musical Applications. Additionally, his contributions as a Google Mentor underscore his dedication to sharing knowledge and fostering talent in the field.

Mizanur Rahman is an Assistant Professor at Daffodil International University and a researcher pursuing his PhD at Jagannath University in Dhaka with a wide-ranging research focus encompassing image processing, typography, art, education, and animation, Mizanur demonstrates a passion for exploring diverse aspects of visual communication. Prior to his academic pursuits, he honed his skills in the fast-paced environment of an advertising agency, gaining invaluable industry experience. Mizanur Rahman's remarkable contributions to graphic design have earned him numerous awards and recognition, underscoring his commitment to excellence in the field.

Dr. Mohammad Ali, a distinguished Professor in the Department of Graphic Design, Crafts, and History of Art at Rajshahi University, Bangladesh, boasts a multifaceted background as an artist, art critic, connoisseur, historian, and esteemed faculty member. Committed to nurturing creativity and critical thinking, Dr. Ali is dedicated to elevating student performance by employing innovative teaching methodologies. He actively seeks opportunities to contribute to the academic community, driven by a passion for fostering a vibrant learning environment. With a keen focus on interdisciplinary research, Dr. Ali endeavors to bridge the gap between art and technology, offering valuable insights to the field. His ultimate goal is to inspire students to delve into their creative potential and explore new horizons. Eager to push the boundaries of knowledge, Dr. Ali remains steadfast in his pursuit of challenging opportunities to showcase his expertise and unwavering dedication.