Developing of an Integrated Framework Leveraging Statistical Analysis and Deep Learning Techniquesto Optimize Transportation Servicesfor Arbaeen Visitors

Prof.Samira Naji Kadhim
University of Baghdad - College of Science for Women - Department of Mathematics
Samirank_math@csw.uobaghdad.edu.iq

Asst.Prof.Asraa Muayed Abdalah
University of Baghdad - College of Science for Women - Department of Computer Science
Asraa.m@csw.uobaghdad.edu.iq

Abstract

The Arbaeen pilgrimage continuous to show increase in the size of the crowed, emphasizing the essential demand for precise analytical system to optimize crowed organization, service managing, and decision support. The present work develops two frameworks that employ predictive statistical modeling combined with contemporary artificial intelligence based solutions for managing this extensive religious occasion. In this regrad, the first framework involves a thorough analysis on established and other types of forecasting methods such as Box–Jenkins time series models, nonparametric regressions (Nadaraya–Watson and Epanechnikov kernels), and so on. One-way ANOVA—combined with multilayer perceptron models applied to forecast health indicators such as pilgrim birth rates. Inspired by best practices in the classification of Arbaeen films; the second framework proposes Comparison Replicating best practices from the Arbaeen film scene classification, the second strand, a video scene recognition method is proposed based on transfer learning from Inception V3 and VGG-16. These neural networks trained to recognize and distinguish between images from the Arbaeen procession and other mass gatherings, this system improves deeper machine-based understanding of context and support the implementation of AI-driven surveillance systems.Using transfer learning results in ingesting and interpreting largescale pilgrimage data. Engineered to suit Arbaeen's distinct characteristics,this plan also provides a flexible foundation for managing similar large-scale events such as other eligious or public gatherings. To support real time monitoring we explore visual scene recognition after presenting forecasting.

Keywords: Arbaeen Pilgrimage, Inception V3, Nonparametric Regression, Pilgrim Services, Scene Classification, Time Series Analysis, Transfer Learning, VGG-16.

Introduction

One of the largest religiuos gatherings globally is the annual Arbaeen pilgrimage in Karbala, Iraq, where millions of Shia Muslims are drawing each year. With the growing number of attendees at this significant religious event, the relevant authorities are facing numerous and escalating challenges in crowd control, visitor transportation planning, and emergency response.

This study proposes a two-layered framework model that integrates two distinct models. The first model employs statistical methods to forecast the number of future visitors. The second model leverages transfer learning through neural Inception V3 and VGG-16 networks for recognizing and classifying images. The purpose of this model is developed to classify images of the Arbaeen pilgrimage and separating them from other types of crowed gatherings, for example the ones taking place during other religious events or large athletic competitions. This study facilitates both strategic foresight coupled with real time visual observations during the event.

Related Work

Multiple earlier studies have employed statistical approaches to predict visitor counts of the Arbaeen pilgrimage. Al-Razzaq, Mohammed, and Rasheed (2024), they employed geometric and exponential models on 2017-2022 data, predictiong over 35 million visitors by 2030 via the geometric model. Nevertheless, such approaches do not match the precision and flexibility of advanced AI-based forecasting techniques.

Khayyat and Mustafa (2025) in their statistical investigation analyzed the tourists nationalities, showing that Karbala during the Arbaeen attracted 53.3% of all tourist visits in Iraq during the event.

In their study (Mostafa et al., 2024) investigates the use of solar energy to power service points along the pilgrimage route from Baghdad to Karbala. By applying remote sensing techniques and GIS, the researchers identified optimal locations for photovoltaic installations using satellite imagery and multi-criteria analysis. The simulation of energy production was performed using MATLAB. The study highlights the feasibility of sustainable energy for Arbaeen processions and provides a framework for future renewable energy planning in similar contexts.

Recent developments in deep learning researches have demonstrated encouraging results in classifying visual contents related to high density events. Convolutional Neural Networks (CNNs), especially those applying pre-trained models via transfer learning, achieved notable accuracy levels in recognizing and classifying between different types of mass events based on visual cues.

Wang et al. (2019) proposed a model in which Inception V3 with transfer learning is used for pulmonary image classification. Features are automatically extracted by the model from chest X-ray images and applied various classifiers including Softmax, Logistic Regression, and SVM. Their results showed improved diagnostic accuracy and sensitivity compared to traditional DCNN models. To enhance overall system performance; data augmentation was used. The approach demonstrated the advantages of using transfer learning in medical image analysis.

Tammina, S. (2019) showed that transfer learning can be used for clustering, classification and regression. He used transfer learning based on VGG-16 to enhance the classification of binary image, and showed learning process in a comparative diagram between CNN and transfer learning using VGG-16 as shown in figure 1. The computational cost was reduced by removing VGG16's fully connected layers and customizing the network while enhancing classification performance. Table 1 shows the results obtained by both networks.

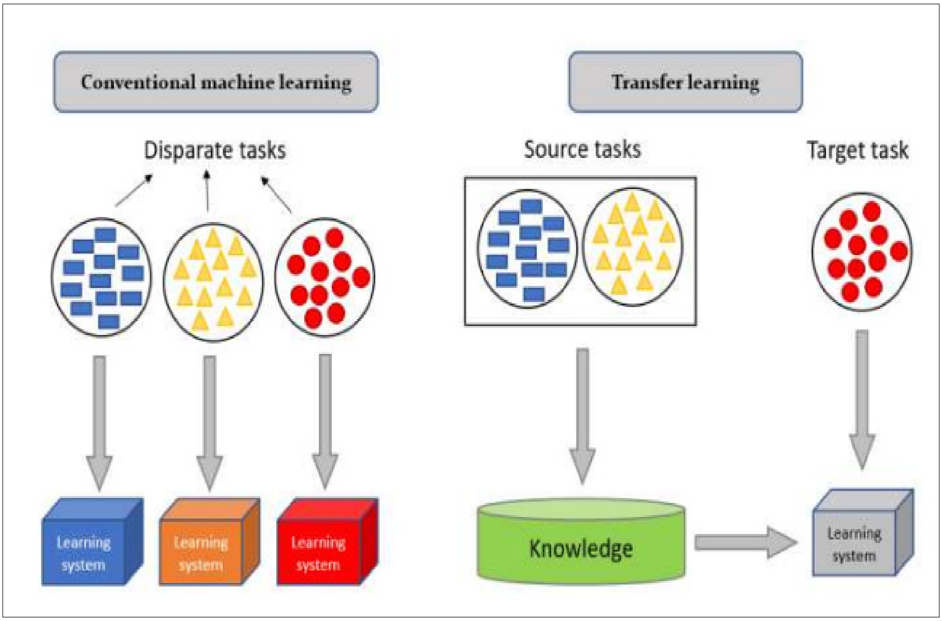

Figure 1. A diagram shows differences of Learning Processes between Conventional Machine Learning and Transfer Learning

Table 1: Accuracy for Different Neural Network Models.

	Training accuracy	Validation accuracy
Basic Convolutional neural network	98.20%	72.40%
Fine tuning CNN with image augmentation	81.30%	79.20%
Fine tuning CNN with pretrained VGG-16 model and image augmentation	86.50%	95.40%

According to our knowledge, this is the first work to combine SARIMA forecasting and CNN-based visual classification tailored to Arbaeen's logistical needs. This paper analyze different statistical methods by comparing them also it compares between classification accuracy obtained using transfer learning. The results from this study can be used to provide strategic planning and improves overall safety.

Data Description for Statistical Methods

Table 2 used by Al-Razzaq, Mohammed, and Rasheed (2024) which estimates visitor numbers from 2017 to 2022 then show prediction of number of visitors in 2030. Predictive Methods (Geometric Growth and Exponential Growth) for Estimating Visitor Numbers. Geometric Growth method predicted that the number of visitors will be 35,403 million; while the Exponential Growth predicted the number of visitors to be 25,399 million in 2030. These can provide baseline data for evaluating more advanced forecasting models such as SARIMA

Table 2: Estimated Visitor Numbers 2017–2030)).

year	Actual / Estimated Visitor
2017	9900
2018	10714
2019	10200
2020	10367
2021	11426
2022	12452
2030	35403

The used data comprises detailed table on tourist attractions in Karbala, visitor demographics by nationality as in Table 3 and continent and hotel occupancy rates during the 2021 Arbaeen pilgrimage.

Table :3 Distribution of Visitors by Nationality.

Rank	Nationality	Number of Visitors	Percentage (%)
1	Afghanistan	852	0.51
2	UAE	314	0.19
3	Armenia	3	0.0018
4	Bahrain	1,651	0.99
5	Hong Kong	1	0.0006
6	Indonesia	27	0.02
7	India	361	0.22
8	Iran	136,225	82.33
9	Jordan	226	0.14
10	Kuwait	2,868	1.73
11	Lebanon	11,138	6.73
12	Palestine	22	0.01
13	Oman	444	0.27
14	Pakistan	9,406	5.68

15	Qatar	93	0.06
16	Saudi Arabia	1,335	0.81
17	Syria	361	0.22
18	Yemen	46	0.03
19	Cambodia	3	0.0018
20	Kyrgyzstan	7	0.0042
21	Myanmar	3	0.0018
22	Malaysia	32	0.02
23	Uzbekistan	44	0.03
Total		165,462	%100

The data shows that 93.64% of visitors originate from Asia, 82% of visitors came from Iran. 3.44% of visitors came from Europe while Africa visitors constitute 0.78%. Other continents show varied but smaller contributions, these data serve as the foundation for statistical modeling and predictive analytic. Sophisticated analytical models were used to perform comparative statistical analysis to obtain meaningful insights from visitor data, we employed a suite of advanced statistical modeling techniques—Box–Jenkins time series modeling, nonparametric kernel regression, and one-way ANOVA. Time-series as well as cross-sectional data are used to analyze behavioral patterns of visitors and regional focus. These methods were applied to two types of data:

- Time-series which estimates the total annual visitors from 2017 to 2030 shown by Table 2.
- Cross-sectional distribution of visitors by nationality in a given year shown by Table 3.

Each method provides a unique perspective on the data, enabling a multi-dimensional understanding of visitor behavior, trends, and group-level differences.

Semi-Annual Scientific Journal
www.arbnj.com
Ramadan 1447 A.H, March 2026 A.D

Methodological Overview and Comparative Insights

1. **One-Way ANOVA (Bittner, 2022):** Regional Disparities in Visitor Distribution Using Table ,3 nationalities were grouped into geographic regions:

- Iran, Lebanon, Saudi Arabia, Jordan and Kuwait considered to be Middle East countries.
- Pakistan and India are considered to be South Asia countries.
- Afghanistan and Uzbekistan considered to be Central Asia countries.
- Others :Remaining nationalities

To investigate if there are statistically significant differences in the average number of visitors per nationality across these regions an ANOVA test was used.

Hypotheses:

- H0: All regional means are equal.
- H1: At least one regional mean differs significantly.

The null hypothesis was rejected because of the large F-statistic, and the smaller than 0.01 p-value. A post-hoc Tukey's HSD test showed that the number of Iranian visitors (from Middle Eastern countries) is greater than the number of visitors from other regions. This finding underscores the importance of tailoring infrastructure and services to meet the requirements of these prominent visitor groups. Especially during crowded times, also it is very important to analyze regional differences in attendance to allocate resources effectively and ensure a smooth, well-supported experience for everyone.

2. Nonparametric Regression (Salibian-Barrera, 2023): Capturing Trends Without Parametric Assumptions

To smooth the number of visitors and extract data and to estimate underlying trends; a Nadaraya–Watson estimator with an Epanechnikov kernel smoothing was using Table 2.

$$\hat{y}_t = \frac{\sum_{i=1}^{n} K_h(t-i) y_i}{\sum_{i=1}^{n} K_h(t-i)} \quad \ldots\ldots\ldots (1)$$

Where:

- Kh(.): Epanechnikov smoothing kernel function,
- h: used to control smoothness. It is the bandwidth parameter,

From this study it is found that growth in visitor numbers is not linear nor exponential, but rather follows a more complex trajectory, because it is affected by external events such as geopolitical trends or policy shifts. This model is useful when parametric assumptions fail or incorrect, providing a flexible replacement to classical curve fitting. Th adaptability of this method helps facilitate the recognition minor changes, providing more responsive forecasting methods.

3. Box–Jenkins Time Series Modeling (SARIMA) (Gorgess & Ibrahim, 2013; Akpanta et al., 2015): Forecasting Visitor Counts

A SARIMA approach was applied to the historical data in Table 3, capturing both trend and seasonality (if available):

$$\mathrm{SARIMA}(p,d,q)(P,D,Q)_m \quad \ldots\ldots\ldots\ldots (2)$$

Where:

- p,d,q: Non-seasonal autoregressive, differencing, and moving average components,
- Q,D,P: Seasonal counterparts,
- m: Season-specific timeframe (e.g., annual seasonality).

Although the limitation of historical data (only six years), the model effectively detected the increasing trend and the generated outcomes for the next decade (2023–2030). These forecasts match well with the values in Table 3, supporting the model's predictive performance. Strategic planning is enabled by using time series modeling, allowing authorities can predict increasing number of visitors and can allocate resources accordingly. Table 4 shows methodological overview and comparative insights. While Table 5 shows summary of analytical approaches for regional visitor data.

Table 4: Methodological Overview and Comparative Insights

Method	Objective	Data Used	Key Findings
One-Way ANOVA	Compare average visitor counts across nationalities grouped into regions	Table 3	Significant regional disparities found; Middle Eastern countries dominate
Nonparametric Regression(Nadaraya–Watson, Epanechnikov Kernel)	Smooth noisy data without assuming functional form	Table 2	Revealed nonlinear growth trend in visitor numbers over time
Box–Jenkins (SARIMA)	Forecast visitor numbers using historical patterns	Table 2	Accurately modeled increasing trend with seasonal components

Table 5: Evaluation √ X Analytical Methods for Tourism Data Analysis.

Method	Type	Temporal?	Handles Complexity?	Provides Forecast?	Regional Insights?	Best Use Case
SARIMA	Predictive	√	√	√	X	Forecasting future visitor numbers
Nonparametric Reg.	Exploratory	√	√	X	X	Smoothing noisy data, detecting hidden trends
One-Way ANOVA	Inferential	X	X	X	√	Comparing regional visitation levels

This comparative study reveals how the integration of different statistical techniques can offer a broad perspective of visitor dynamics; where One-way ANOVA detects regional differences and facilitates precise interventions. Nonparametric regression show unseen structures in noisy data, especially when relationships are unclear. SARIMA modeling enables accurate forecasting, supporting long-term planning and resource allocation. Collectively, these strategies build solid analytical structure framework leveraging mathematics and Artificial Intelligence, applicable to a diverse spectrum of demographic, economic, and social forecasting tasks.

Arbaeen pilgrimage Dataset for Scene Classification

A custom dataset of 400 annotated images was created from various films and documentaries:

- Arbaeen Scenes (200 images): Processions, tents, flags, rituals.
- Non-Arbaeen Mass Gatherings (200 images): Concerts, Hajj, protests, sports events.

Each image was labeled binary (1 for Arbaeen, 0 otherwise) and splitted into Training: 80%, Validation: 10% and Testing: 10%.

Methodology for Deep Learning for Image Scene Classification

As illustrated in the figure ,2 the methodology consists of several stages, beginning with image preprocessing, which includes frame extraction (2-second intervals), resizing(resize images to 299×299 for Inception V3 and 224x224 for VGG-16), normalization([[0,1scaling), augmentation(flip, rotation, zoom, brightness) and Label encoding (Binary labels assigned). Following this step, the images are divided into three sets: 80% for training, 10% for testing, and 10% for validation.

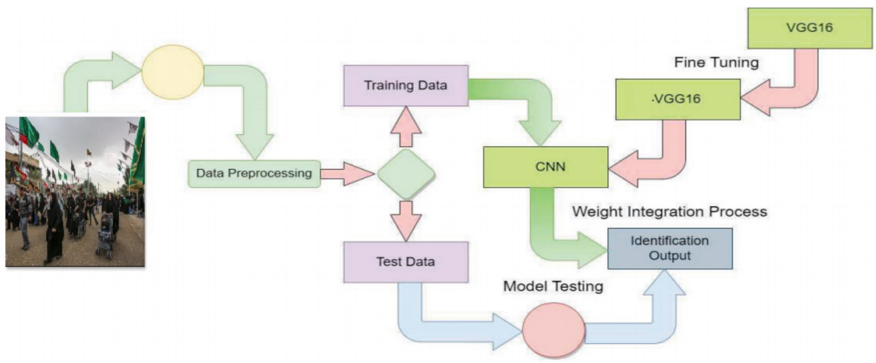

Figure 2: Methodology of proposed system

The training efficiency and accuracy is improved by using transfer learning technique, especially when working with small datasets. ImageNet is a large dataset used to train models so they can learn useful features; and then these models can be used by other networks instead of training from scratch, which can be time-consuming and computationally expensive (Ali & Abdulazeez, 2024). This pre-trained model is then adapted to our specific task in training and testing stage. Typically, we keep the early layers (which detects general features such as textures and edges) and either fine-tune the later layers or replace them with new layers tailored to our classification problem. This process significantly reduces training time and helps achieve higher performance, even with smaller datasets. In essence, transfer learning allows the usage of prior knowledge to accelerate learning and improve model accuracy in new but related tasks. The final stage is the recognition (classification) stage in which the system is required to recognize or classify images into Arbaeen Scenes and non Arbaeen scenes.

Model Architectures

The Inception V3 architecture overview is shown in Table 6 below. While Table 7 gives architecture overview of the VGG-16 network. Table 8 shows a comparison summarizing the key features of Inception V3 and VGG-16 networks.

Table 6: Inception V3 Architecture Overview

Stage	Description	Input Dimensions	Output Dimensions
Stem	Initial Conv layers + MaxPooling	299 × 299 × 3	35 × 35 × 192
Inception Module A ×3	Three Inception-A blocks (multi-branch convolution paths)	35 × 35 × 192	35 × 35 × 288
Reduction A	Downsampling via convolutions and pooling	35 × 35 × 288	17 × 17 × 768
Inception Module B ×5	Five Inception-B blocks (deeper feature extraction)	17 × 17 × 768	17 × 17 × 768
Reduction B	Further downsampling with Conv + Pooling	17 × 17 × 768	8 × 8 × 1280
Inception Module C ×2	Two Inception-C blocks (wide convolutional filters)	8 × 8 × 1280	8 × 8 × 2048
Global Average Pooling	Averages feature maps across spatial dimensions	8 × 8 × 2048	1 × 1 × 2048
Dense Output Layer	Fully connected layer with softmax activation	2048	2 (binary classification)

Table 7: VGG-16 Architecture Overview

Stage	Description	Input Dimensions	Output Dimensions
Block 1	Two Conv2D layers with 64 filters + MaxPooling	224 × 224 × 3	112 × 112 × 64
Block 2	Two Conv2D layers with 128 filters + MaxPooling	112 × 112 × 64	56 × 56 × 128
Block 3	Three Conv2D layers with 256 filters + MaxPooling	56 × 56 × 128	28 × 28 × 256
Block 4	Three Conv2D layers with 512 filters + MaxPooling	28 × 28 × 256	14 × 14 × 512
Block 5	Three Conv2D layers with 512 filters + MaxPooling	14 × 14 × 512	7 × 7 × 512
Flatten	Flattening of feature maps	7 × 7 × 512 = 25088	25088
FC-1	Fully connected layer with 4096 units (Dense)	25088	4096
FC-2	Fully connected layer with 4096 units (Dense)	4096	4096
Output	Fully connected layer + Soft-max for binary output	4096	2 (binary classifi-cation)

Table 8: A comparison summarizing the key features of Inception V3 and VGG-16:

Feature	Inception V3	VGG-16
Model Depth[11]	~48 layers (including auxiliary classifiers and inception modules)	16 layers (13 Conv + 3 Fully Connected)
Design Philosophy[11]	Inception modules use multi-scale filters (1×1, 3×3, 5×5) in parallel	Uniform 3×3 filters across all convolutional layers
Feature Extraction[11]	Captures features at multiple scales simultaneously	Focuses on fine-grained spatial details using small kernels
Input Size[11, 12]	299 × 299 × 3	224 × 224 × 3
Parameter Count[12]	~23 million	~138 million
Computational Cost[13]	More efficient, optimized with dimensionality reduction	High memory usage, especially in fully connected layers
Transfer Learning Support[12, 14]	Pre-trained on ImageNet; strong generalization with fine-tuning	Pre-trained on ImageNet; widely used for transfer learning
Accuracy (ImageNet Top-1)[14]	~78.8%	~71.5%
Strengths[12]	Good trade-off between accuracy and efficiency; scalable architecture	Simplicity, strong baseline performance
Weaknesses[12]	More complex structure; harder to customize	Very high number of parameters; slower to train on large datasets

Both models were fine-tuned with Adam optimizer, binary cross-entropy loss, number of epochs = 20, batch size = 32, and learning rate =0.0001 .

Results

The performance of the system using the two models are shown in Table 9. Inception V3 shows slightly higher performance in all phases, indicating better generalization and learning capacity. VGG-16, despite its older architecture and higher parameter count, still achieves high accuracy, proving its robustness for binary classification tasks.

Table 9: Model performance comparison presented with labeled metrics.

Model	Training Accuracy (%)	Validation Accuracy (%)	Test Accuracy (%)
VGG-16	96.3	93.5	92.0
Inception V3	97.7	95.0	94.1

Inception V3 achieved higher accuracy due to its multi-scale feature extraction modules; while VGG-16 showed faster convergence and lower computational cost, making it suitable for lightweight deployment.

Conclusion

The Arbaeen pilgrimage is considered one of the largest mass gatherings in the world. To ensure its successful management, it is essential to draw on scientific studies grounded in statistics and artificial intelligence. Statistical models provide understanding of relationships and a conceptual framework for predictive analysis and regional comparison, whereas AI-based models support high-level deep learning mechanisms that support video tracking and scene recognition throughout the Arbaeen pilgrimage—a crucial part of real-time observation system and real time decision-making process. The applied SARIMA approach ensures precise long-term forecasts, while transfer learning using Inception V3 achieves over 94% test accuracy in scene classification. Findings show that the integration of statistical and intelligent techniques optimizes visitor flow control and management, enhances strategic planning, and guarantees more effective management of resources.

The SARIMA model used for predicting visitors number. While the ANOVA model used for analyzing different regions. In addition, learning using pre-trained convolutional neural networks (CNNs) to classify images and scenes captured from security camera footage and YouTube videos. The obtained results can help authorities in allocating infrastructure and personnel more efficiently during crowded times. The outcomes of this study can be leveraged to support crowd management, allocate necessary resources, and ensure proper care and services for Arbaeen pilgrims.

References

1. Rasheed, H. A.-R., Mohammed, R. S., & Rasheed, A. A.-R. (2024). Estimating the number of arrivals during the Arbaeen pilgrimage to the holy city of Karbala. Al-Arba'een Journal, 2(2, Pt. 3), 132-148.
2. Khayyat, H. A. J. I., & Mustafa, Z. S. (2025). Analysis of the nationalities of tourists visiting the holy city of Karbala during the 2021 (1443 AH) Arbaeen pilgrimage: An inductive study [In Arabic]. Al-Arbaeen Journal, 2(2, Pt. 5, Sec. 2), 133-158.
3. Mostafa, S. S., Alatta, H. J. M., Khanjer, E. F., & Abul Al-Razak, B. (2024). Equip Ziyarte Al-Arba'een service points with clean energy through solar radiation using remote sensing techniques: A case study along the path from northern Baghdad to the holy Karbala. 7th International Scientific Conference of Ziyarat Al-Arba'een.
4. Wang, C., Lin, H., Liu, B. X., Zeng, C. Y., Chen, J., & Zhang, G. (2019). Pulmonary image classification based on Inception-v3 transfer learning model. IEEE Access, 7, 146533–146541. https://doi.org/10.1109/ACCESS.2019.2946000.
5. Tammina, S. (2019). Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images. International Journal of Scientific and Research Publications, 9(10),143-150.http://dx.doi.org/10.29322/IJSRP.9.10.2019.p9420.
6. Bittner, A. (2022). Analysis-of-variance (ANOVA) assumptions review: Normality, variance equality, and independence (pp. 28–33). https://doi.org/10.47461/isoes.2022_bittner.
7. Salibian-Barrera, M. (2023). Robust nonparametric regression: Review and practical considerations. Econometrics and Statistics. https://doi.org/10.1016/j.ecosta.2023.04.004.
8. Gorgess, H. M., & Ibrahim, R. (2013). Time series forecasting by using Box-Jenkins models. Ibn Al-Haitham Journal for Pure and Applied Science, 26(1), 337-345. University of Baghdad, College of Education for Pure Science (Ibn Al-Haitham) .

9. Akpanta A. C, I. E. Okorie1 and N. N. Okoye (2015). SARIMA Modelling of the Frequency of Monthly Rainfall in Umuahia, Abia State of Nigeria. American Journal of Mathematics and Statistics; 5(2): 82-88.

10. Ali, A. H., & Abdulazeez, A. M. (2024). Transfer learning in machine learning: A review of methods and applications. Indonesian Journal of Computer Science, 13(3).https://doi.org/10.33022/ijcs.v13i3.4068 .

11. Constantinescu, E. C., Udriștoiu, A.-L., Udriștoiu, Ș. C., Iacob, A. V, Gruionu, L. G., Gruionu, G., Săndulescu, L., & Săftoiu, A. (2021). Transfer learning with pre-trained deep convolutional neural networks for the automatic assessment of liver steatosis in ultrasound images. Medical Ultrasonography, 23(2), 135–139. https://doi.org/10.11152/mu-2746 .

12. Mattheuwsen, L., & Vergauwen, M. (2020). Manhole cover detection on rasterized mobile mapping point cloud data using transfer learned fully convolutional neural networks. Remote Sensing, 12(22), 3820. https://doi.org/10.3390/rs12223820.

13. Suganda Girsang, A. ., Dharma Saputra, A. ., & Yanrie, V. . (2023). Performance Comparison between VGG16 and Inception V3 for Organic Waste and Recyclable Waste Classification. International Journal of Intelligent Systems and Applications in Engineering, 11(2), 557–563. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2711.

14. Firmansyah, I., Rosnelly, R., & Wanayumini. (2023, February 25). Inception-V3 versus VGG-16: In rice classification using multilayer perceptron. In Proceedings of the 2nd International Conference on Information Science and Technology Innovation (ICoSTEC), Yogyakarta, Indonesia. https://prosiding-icostec.respati.ac.id/index.php/icostec/article/download/24/24.