Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 258 (2025) 3723–3730

Procedia Computer Science

www.elsevier.com/locate/procedia

International Conference on Machine Learning and Data Engineering

Explainable Artificial Intelligence for Computer Vision and Quantum Machine Learning

Asharul Islam Khan [a*], Ali Al Badi [b], Mohammed Alqahtani[c]

[a] *OCCI, HURC, Sultan Qaboos University, P.O.Box 33 AlKhodh, Muscat, P.C.123, Sultanate of Oman*
[b]*Dean, Gulf College, P.O.Box 885, Al Mabaila, Muscat, P.C.133, Sultanate of Oman*
[c]*Department of Computer Information Systems, Imam Abdulrahman Bin Faisal University, P.O. Box 1982, Dammam 31441, Saudi Arabia*

Abstract

Artificial Intelligence and Machine Learning (AI/ML) are the emulation of human intelligence by computer systems. The AI/ML models have made inroads into Computer vision and Quantum computing due to their tremendous ability for reasoning and problem-solving. Computer vision allows computers to interpret visual information (photos, videos) similar to humans while Quantum computers execute complex tasks more efficiently than regular computers. However, how to explain the reasons for specific outputs is an important challenge with AI/ML models when applied to Computer vision and Quantum computing. Without the explainability of the AI/ML, the reliability and validity of Computer vision and Quantum computing outputs, when AI/ML is applied to them, is of great concern for researchers. To solve the issue of trustworthiness and transparency researchers have come up with the idea of explainable AI/ML (XAI). Therefore, to understand the extent of research on XAI in the area of Computer vision and Quantum computing, this study used the existing literature from the Scopus database and statistically presented the results. The study has found that the research on XAI in Computer vision and Quantum computing has been growing particularly since 2019. For AI/ML in Computer vision, the maximum studies on the XAI are in the healthcare sector accounting for 57% followed by autonomous vehicles. As far as AI/ML in Quantum computing is concerned, the majority of XAI research is related to the healthcare sector accounting for 46% followed by cyber security (20%).

Keywords: Explainable Artificial Intelligence; explainable Machine learning; XAI; explainable healthcare; computer vision; Quantum computing

* Corresponding author.
E-mail address: a.khan2@squ.edu.om/ashar.367@gmail.com

1877-0509
10.1016/j.procs.2025.04.627

1. Introduction

AI/ML is a new industrial revolution. According to the Statista forecast by 2030, AI/ML will have a market volume of US$503.40bn while the market volume of Computer vision will be US$46.96bn by 2030 moving at an annual growth rate (CAGR 2024-2030) of 10.50%[†]. Computer vision uses AI/ML techniques and algorithms to detect and classify objects in multimedia files. Nowadays researchers are leveraging the benefits of Quantum computing to greatly accelerate classical machine learning tasks. Quantum machine learning is executed on quantum computers. It takes advantage of quantum computers' information processing efficiency [1]. Quantum computing uses counterintuitive quantum physics principles relying on quantum bits and qbits rather than 0 or 1 used in traditional computing.

Unfortunately, one of the challenges with AI/ML is the black-box nature of its algorithms. There is no explanation of what happens inside AI/ML classifiers like Recurrent Neural Network (RNN), Deep Neural Network (DNN), and the user community often suspects the outputs. As a result, many AI/ML-based outputs and judgments are difficult to understand by both researchers and common consumers. In safety-critical systems such as autonomous driving, the AI/ML model interpretability is important [2]. The AI/ML model for stock price prediction has limitations of generalization [3]. Even the Convolutional Neural Network (CNN), used in rescue robotics is limited by explainability [4]. Similarly, Computer vision in medical imaging has issues of explainability [5]. Likewise, despite Quantum Machine Learning's popularity, quantum neural networks (QNN) are highly counterintuitive and difficult to understand due to their architecture's unique quantum-specific layers (for example, data encoding and measurement). It prohibits QNN users and researchers from fully comprehending its inner workings and investigating the model's training status [6].

To solve the issue of expansibility in Computer vision and Quantum Machine Learning, the researchers have come up with the idea of XAI. The XAI aims to improve trustworthiness, transparency, and confidence in the AI/ML outputs. XAI can provide reliable answers to questions regarding how data is evaluated and what criteria are considered during the decision-making process. Hence, this study aims to investigate the existing research and contributions made on the subject of XAI in the area of Computer vision and Machine Learning in Quantum computing besides suggesting relevant research directions. This study will encourage scholars and practitioners to devote financial and non-financial resources to the nascent field of XAI in the area of Computer vision and Quantum computing. There are five sections in this paper. Section 2 describes the basics of Computer vision and Quantum computing, and XAI. Section 3 presents the methods and findings. Section 4 presents the discussion. The last section summarizes the study.

2. Background

This section describes the basics of AI/ML, Computer vision, and Quantum computing. In today's age tons of multimedia information are available. The indexing and information extraction from such huge data sources requires AI/ML-based computing interventions. With the help of Computer vision, one can detect, identify, and extract relevant information from multimedia files. Computer vision uses AI/ML techniques either supervised, unsupervised, or semi-supervised to render the information from the multimedia files. The biggest advantages of such algorithms and their models are in classification mainly in healthcare such as if a disease is benign or malignant. Medical imaging is an emerging area with a number of applications in scientific and technological fields, including health sciences and engineering. The unsupervised algorithm looks for patterns in the input data rather than depending on the input labels. For instance, Generative Adversarial Networks (GAN) and Auto Encoders. The reinforcement algorithms depend on a reward function where neural network weight is optimized for the sake of maximizing the reward function. Computer vision is useful for object identification, classification, and extracting meaningful information from photos, graphic documents, and videos. For instance, figure 1 shows the Computer vision implementation in image processing.

[†] https://www.statista.com/outlook/tmo/artificial-intelligence/computer-vision/worldwide

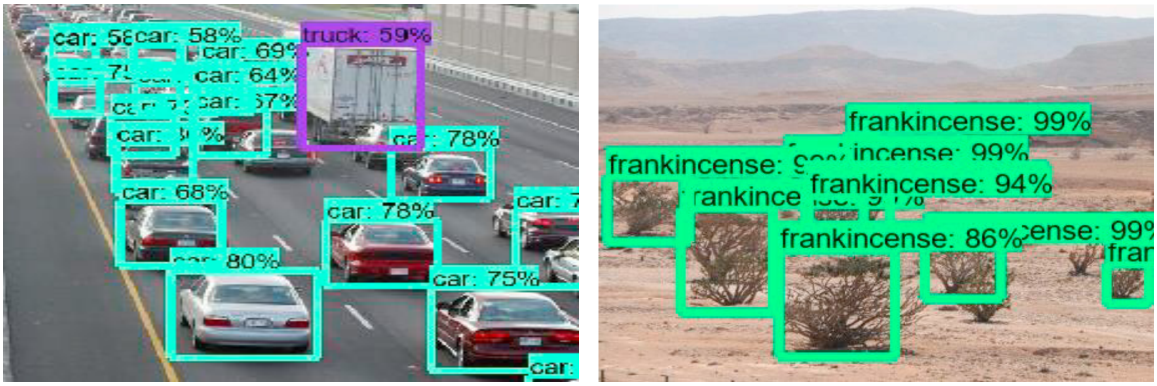

Figure 1. Computer vision implementation in image processing (Deep learning and Faster-RCNN-Inception-V2 model)

Quantum computers' huge computational capacity and speed have the potential to create a quantum leap in various fields. Quantum machine learning is now gaining popularity in a variety of fields because of its better performance and capabilities. However, Quantum machine learning requires comprehension by users which is a complex and challenging task [7] due to the black-box nature of AI/ML-based techniques. The AI/ML models created with AI/ML algorithms such as Support Vector Machines (SVM), Random Forests, and K-Nearest Neighbors (KNN), as well as Deep learning algorithms such as Artificial Neural Networks, CNN, and RNN, are difficult to interpret, even for an expert in AI/ML field. Therefore, the question arises to what extent the research in the XAI field has progressed and what segments of the Computer vision and Quantum computing sector it has impacted.

3. Method and Findings

This study has used a review of the existing literature from Scopus-indexed literature focused on XAI in Computer vision and Quantum computing to investigate the extent of the research in the field. Scopus is one of the most popular and largest databases. As a result, the research team conducted a study using the Scopus database. A number of keywords were used. They are "*explainable*", "*Machine learning*", "*XAI*", "*interpretable*", "quantum", and "*commuter vision*", and "*Artificial Intelligence.*" The keyword search yielded a total of 35 relevant publications for analysis. Moreover, statistical analysis was conducted on the collected information to make it presentable form.

The study has found that XAI in Computer vision and Quantum computing has been growing since 2019. The majority of the publications came in 2022, approximately 90%. Figure 2 shows the explainable AI/ML studies in Computer vision and Quantum computing.

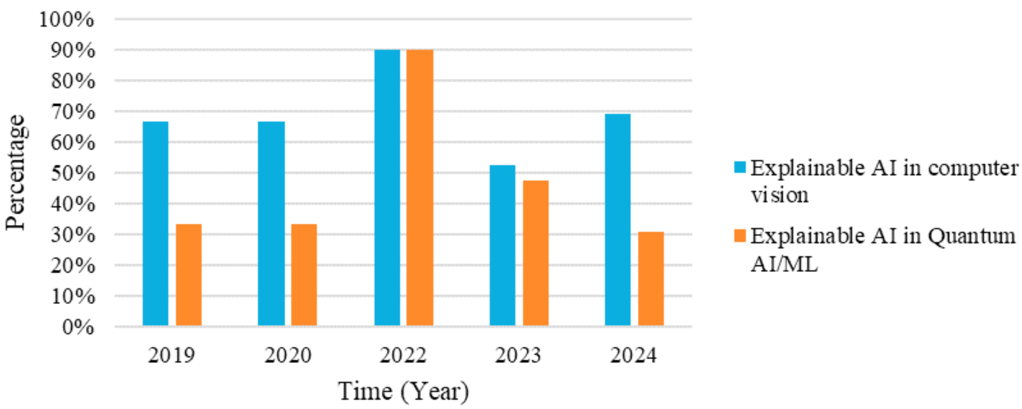

Figure 2. Explainable AI/ML studies in the area of Computer vision and Quantum computing

3.1. Explainable AI/ML in Computer vision and Quantum computing

The healthcare sector (57%) has received the maximum attention from researchers for explainable AI/ML in Computer vision followed by autonomous vehicles (22%). The application of AI/ML techniques to medical image processing has transformed the field of healthcare. However, classical AI models' lack of interpretability and transparency has hampered their broad use in clinical practice. XAI addresses this issue by offering interpretable explanations for AI/ML-generated conclusions. In the self-driving vehicles, the Computer vision plays critical role helping in the object detection and recognition on the streets. However, the AI/ML algorithms should be explainable in terms of their actions [8]. Figure 3 shows the explainable AI/ML in Computer vision.

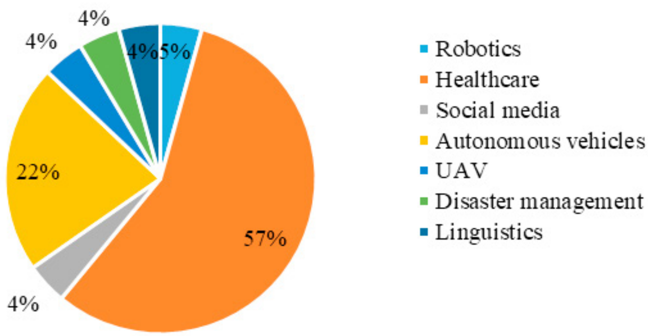

Figure 3. Explainable AI/ML in Computer vision

The healthcare sector (46%) has received the maximum attention of researchers for explainable AI/ML in Quantum computing followed by cyber security (20%). The outcomes of XAI models in medical image processing allow healthcare workers to comprehend reasons for AI/ML model outputs, resulting in increased diagnostic accuracy, and trust in AI/ML systems. Figure 4 shows the explainable AI/ML in Quantum computing.

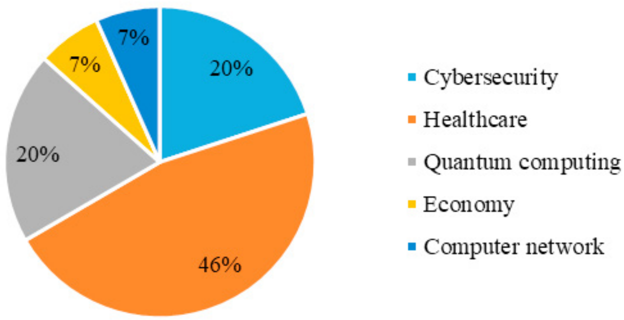

Figure 4. Explainable AI/ML in Quantum computing

There are various research themes related to explainable AI/ML in Computer vision and Quantum computing. For instance, robots for search and rescue, face expression classification, object detection, etc. However, the majority corresponds to Malware detection (15.4%) and image processing (38%). Table 1 shows XAI research themes in Computer vision and Quantum computing.

Table 1. XAI research themes in Computer vision and Quantum computing

XAI research in Quantum AI/ML with key themes	Percentage	XAI research in Computer vision with key themes	Percentage
Breast cancer images	7.7%	Robot for search and rescue	3.4 %
COVID-19 detection and classification	7.7%	Face expression classification	13.8 %
False information detection	7.7%	Image processing	34.5 %
Malware detection	15.4%	Object detection	37.9 %
MRI-Radiomic Quantum Neural Network	7.7%	Review	3.4 %
Predict the risk of heart disease	7.7%	Alzheimer's disease	3.4 %
Quantum clustering method	7.7%	Glaucoma diagnosis	3.4 %
Quantum neural networks	7.7%		
Signal modulation classification	7.7%		
Small quantum states	7.7%		
Stock price prediction	7.7%		
Wearable electronic optical data analysis	7.7%		

3.2. Explainable AI/ML approaches

There are different XAI approaches developed by researchers such as Intrinsic, Post-hoc, Local, Global, etc. [9, 10]. Posthoc looks into "what went wrong" in a black box environment [11]. Model agnostic explainable approaches are appropriate for clinical application [12]. The four important XAI approaches are explained below.

Local and Global Explanation: The local approaches explain a single data point or forecast while the Global method interprets either a number of data points or the complete model [12]. Global approaches seek to describe in detail the model's overall performance.

Model-specific and Model-agnostic: Model-specific methods function with only one data type or specific model types (e.g., SHAP) but Model-agnostic techniques work with any data type. The Model agnostic approach is based on the principle of looking for change in the output as a consequence of input changes [12].

Intrinsic and Post-hoc: Intrinsic has a structure that is simple enough and inherently interpretable. In such a case linear model or the splits learned with a Decision Tree can be used to deduce why a model makes the predictions it does. Post-hoc explanation strategies, on the other hand, provide explanations after the classification has been formed.

Attribution and Non-attribution: In order to determine which of the neural network's inputs are the most relevant in terms of the network's output, attribution methods are used. Here the important neural network characteristics are determined on the basis of class score when a feature is removed [12].

3.3. Explainable AI/ML frameworks

The working of XAI includes a number of elements. Figure 5 shows a simple illustration of explainable XAI working.

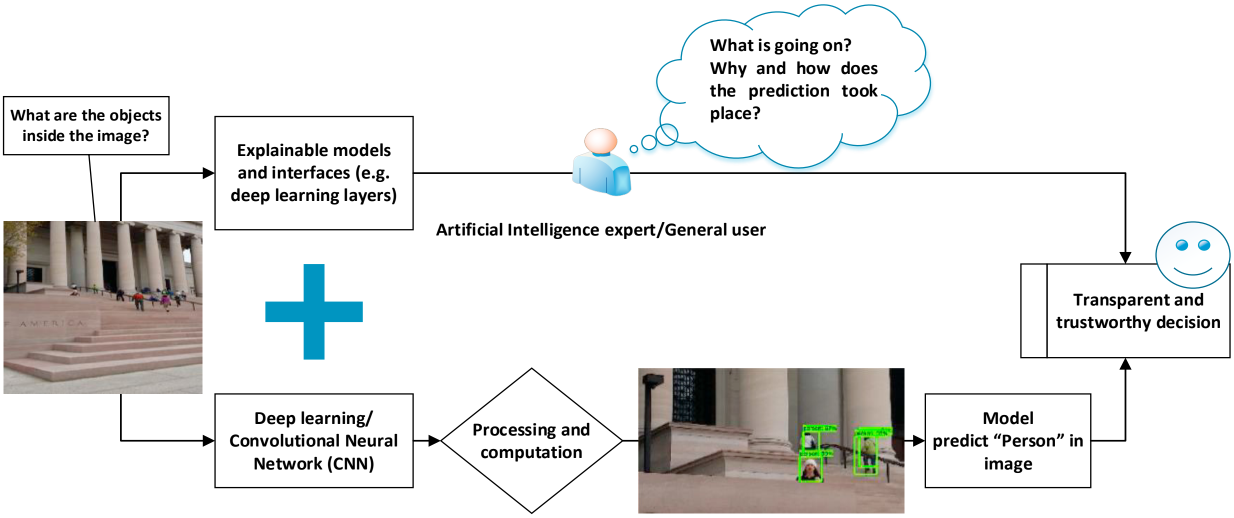

Figure 5. Simple illustration of XAI working (Authors own work)

There are a variety of frameworks and techniques used in the XAI. The Accumulated local effects (ALE), Contrastive Explanation Method (CEM), Partial Dependence Plot (PDP or PD plot), SHapley Additive exPlanations (SHAP), Local interpretable model-agnostic explanations (LIME), and Global Interpretation with Recursive Partitioning (GIRP) are the XAI frameworks. The prominent ones are listed in the table2.

Table 2. Frameworks and techniques used in the XAI

XAI frameworks	Description	Global/ Local
SHAP [13]	SHAP is based on Shapley values, which are commonly used for optimal credit allocation. It is a paradigm for explaining the outcome of any black-box model. However, it is more efficient on certain model types (such as tree ensembles)	Global and Local
LIME [14]	This framework generates or automates interpretable models for the AI/ML prediction models.	Local
PDP [15]	This framework measures the small changes in the outputs of AI/ML models as a consequence of one or two factors.	Global
ALE [16]	This technique works on classification and regression models for measuring the effects of features.	Global
CEM [17]	CEM interprets the classification model outputs	Local
GIRP [18]	GIRP works on the most important decision rules of the AI/ML model for interpretations	Global

4. Discussion

Gesture and facial identification, recognition of handwritten letters and digits, sophisticated driver assistance systems, and behavioral investigations are some of the applications of Computer vision [19]. In healthcare, it is used in medicine, surgery, and diagnostics. In the industry it is used to perform predictive maintenance, therefore helping the industries to replace parts and machines before breakdowns leading to efficiency and efficiency of producing units. Object identification models such as DNN used in UAVs or self-driving cars suffer from a lack of transparency [20]. This could lead to safety issues in many instances.

People find it challenging to understand how these algorithms get their findings. It is not only necessary to examine the explanation, but also to assess its quality and relevance to the given setting. Even professionals in the fields find it challenging to grasp the outcomes of AI/ML models since they are black boxes. It is critical to balance AI/ML accuracy and interpretability because the best-performing systems, such as Deep Learning, are typically the least transparent in explanation, as opposed to Decision Trees. Such a difficult scenario demands an extra layer of interpretation. XAI solves these problems by incorporating extra explanations into the AI/ML outputs [21].

Object detection can be used to avoid traffic collisions, recognize facial expressions, and identify emotions based on human postures. In applications like Face Verification, it is critical to ensure decision transparency, justice, and accountability. Therefore, more research on XAI is needed [22].

The rise of next-generation AI/ML has recently brought responsible, ethical, and trustworthy decision-making to the forefront as one of society's most urgent concerns [23]. The laws and regulations in many countries require an explanation of the black box algorithms of AI, besides the ethical requirements. For example, regulations such as Canada's Personal Information Protection and Electronic Documents Act (PIPEDA), the USA's Health Insurance Portability and Accountability Act (HIPAA), and the European Union's General Data Protection Regulation (GDPR) have been established to address privacy concerns [24]. Additionally, the Defense Advanced Research Projects Agency (DARPA) has shown interest in explainable AI (XAI) and has developed the DARPA-XAI program [24].

In the end, the XAI paradigm emphasizes justice and accountability. XAI is not only beneficial and critical in the healthcare environment, but it also opens a wide range of possibilities for AI/ML solutions overall. As a result, the opacity of AI/ML, which has been widely criticized, can be decreased, and critical confidence can be developed. This is exactly what will boost future customer acceptance over time.

5. Conclusion

The AI/ML has already emerged as one of the century's most crucial technologies. However, questions regarding its safe and dependable use are increasing. For example, self-driving automobiles that use Computer vision have long been a subject of contention in advanced countries such as the United States. The same might be said for whether Computer vision can or should help with, or even make, healthcare decisions. Despite significant breakthroughs, the lack of transparency and interpretation abilities in AI/ML-based systems has remained a major barrier to adoption. This has sparked a new discussion about the value of AI/ML in the lifesaving and vital decision-making systems mainly utilizing technologies such as Computer vision and Quantum computing. This study found that the XAI in Computer vision and Quantum computing is an emerging area of research. The interpretation and explanations of the outputs of AI/ML in Computer vision and Quantum computing are moving forward with the development of new processes, models, and methodologies. Much of the research in the XAI area is in the healthcare sector and cybersecurity. The XAI is essential in healthcare due to its credibility, explanation, application, dependability, equality, openness, interaction, and greater sense of privacy. However, other areas also need attention including disaster management and robotics. One of the major issues is that XAI techniques are designed for general usage without considering the context of the difficulties. The findings can serve as a foundation for future research developments and encourage experts and professionals from various fields to embrace the benefits of explainable AI (XAI) within their industries.

References

[1] Schuld, Maria, and Petruccione, Francesco. (2021)" *Machine learning with quantum computers."* Springer. **676** (2021)

[2] Gierse, Daniel, Neubürger, Felix, and Kopinski, Thomas. (2023). A Novel Architecture for Robust Explainable AI Approaches in Critical Object Detection Scenarios Based on Bayesian Neural Networks. In *World Conference on Explainable Artificial Intelligence*. Springer 126-147

[3] Gandhudi, M., P.J.A, A., Fiore, U., and G.R, G. (2024) "Explainable hybrid quantum neural networks for analyzing the influence of tweets on stock price prediction." *Computers and Electrical Engineering*. **118** (2024): 1-9

[4] Hönig, Peter, and Wöber, Wilfried. (2023). Explainable Object Detection in the Field of Search and Rescue Robotics. In *International Conference on Robotics in Alpe-Adria Danube Region*. Springer. 2023: 37-44

[5] Sohail, A., Fahmy, M. A., and Khan, U. A. (2023) "XAI hybrid multi-staged algorithm for routine & quantum boosted oncological medical imaging." *Computational Particle Mechanics*. **10** (2023): 209-219

[6] Ruan, S., Liang, Z., Guan, Q., Griffin, P., Wen, X., Lin, Y., and Wang, Y. (2024) "VIOLET: Visual Analytics for Explainable Quantum Neural Networks." *IEEE Transactions on Visualization and Computer Graphics.* **30**(2024): 2862-2874

[7] Frohnert, F., and van Nieuwenburg, E. (2024) "Explainable representation learning of small quantum states." *Machine Learning: Science and Technology*. ***5*** (2024)

[8] Nowak, T., Nowicki, M. R., Cwian, K., and Skrzypczynski, P. (2019). How to improve object detection in a driver assistance system applying explainable deep learning. In *IEEE Intelligent Vehicles Symposium, Proceedings*. **2019** (6): 226-231

[9] Arrieta, Alejandro Barredo, Díaz-Rodríguez, Natalia, Del Ser, Javier, Bennetot, Adrien, Tabik, Siham, Barbado, Alberto, García, Salvador, Gil-López, Sergio, Molina, Daniel, and Benjamins, Richard. (2020) "Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI." *Information Fusion* ***58***: 82-115,

[10] Adadi, Amina, and Berrada, Mohammed. (2018) "Peeking inside the black-box: a survey on explainable artificial intelligence (XAI)." *IEEE Access.* **6**(2018): 52138-52160

[11] Kumar, K. P., Thiruthuvanathan, M. M., K.K, S., and Chandra, D. R. (2023)" Human AI: Explainable and responsible models in computer vision" In *Emotional AI and Human-AI Interactions in Social Networking*. **2023:** 237-254

[12] Gulum, Mehmet A., Trombley, Christopher M., and Kantardzic, Mehmed. (2021) "A Review of Explainable Deep Learning Cancer Detection Models in Medical Imaging." *Applied Sciences.* **11**(2021): 1-21.

[13] Lundberg, S. (2017).”A unified approach to interpreting model predictions.” 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA: **2017:** 1-10

[14] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). " Why should i trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining. **2016:** 1135-1144

[15] Friedman, J. H. (2001). “Greedy function approximation: a gradient boosting machine.” *Annals of statistics*. **29**: 1189-1232

[16] Apley, D. W., & Zhu, J. (2020). “Visualizing the effects of predictor variables in black box supervised learning models.” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, **82**(4), 1059-1086

[17] Dhurandhar, A., Chen, P. Y., Luss, R., Tu, C. C., Ting, P., Shanmugam, K., & Das, P. (2018). “Explanations based on the missing: Towards contrastive explanations with pertinent negatives.” *Advances in neural information processing systems*, **31:** 1-12

[18] Yang, C., Rangarajan, A., & Ranka, S. (2018). “Global model interpretation via recursive partitioning.” In 20th International Conference on High Performance Computing and Communications, IEEE. **20:** 1563-1570

[19] Al-Badi, A., Sharma, S. K., Jain, V., & Khan, A. I. (2020). “Investigating emerging technologies role in smart cities’ solutions.” In *Re-imagining Diffusion and Adoption of Information Technology and Systems: A Continuing Conversation: IFIP WG 8.6 International Conference on Transfer and Diffusion of IT, TDIT 2020, Tiruchirappalli, India, December 18–19,* Springer. **2020:** 230-241

[20] Hogan, M., Aouf, N., Spencer, P., and Almond, J. (2022). Explainable Object Detection for Uncrewed Aerial Vehicles using KernelSHAP. In *2022 IEEE International Conference on Autonomous Robot Systems and Competitions, ICARSC 2022.* **2022**: 136-141

[21] Borriero, Alessio, Milazzo, Martina, Diano, Matteo, Orsenigo, Davide, Villa, Maria Chiara, DiFazio, Chiara, Tamietto, Marco, and Perotti, Alan. (2024). Explainable Emotion Decoding for Human and Computer Vision. In *World Conference on Explainable Artificial Intelligence.* Springer. **2024:** 178-201

[22] Liu, G., Zhang, J., Chan, A. B., and Hsiao, J. H. (2024) "Human attention guided explainable artificial intelligence for computer vision models." *Neural Networks*. **177**(2024): 1-16

[23] Anzum, F., Asha, A. Z., Dey, L., Gavrilov, A., Iffath, F., Ohi, A. Q., Pond, L., Shopon, M., and Gavrilova, M. L. (2024)" A comprehensive review of trustworthy, ethical, and explainable computer vision advancements in online social media" In *Global Perspectives on the Applications of Computer Vision in Cybersecurity*. **2024**: 1-46

[24] Morphy, Erika. (2018). “What Is Explainable AI (XAI).” *CMS Wire*. Avaibale at: https://www.cmswire.com/digital-experience/what-is-explainable-ai-xai/. Acceessed on: 30 Sept 2024