

A COMPREHENSIVE STUDY ON THE INTERDISCIPLINARY INTEGRATION OF COMPUTER SCIENCE: ADVANCEMENTS IN NATURAL SCIENCE EDUCATION, EYE-TRACKING TECHNOLOGIES, AND FEDERATED LEARNING FOR ANOMALY DETECTION

CRISTINA REYES¹ CLARISSE MENDOZA²

¹Department of Computer Science, University of the Cordilleras, Harrison Road, Baguio City, 2600, Benguet, Philippines.

²Department of Computer Science, Nueva Ecija Technological University, Burgos Avenue, Cabanatuan City, 3100, Nueva Ecija, Philippines.

Corresponding author: REYES N. H.P.

© REYES H.,P., Author. Licensed under CC BY-NC-SA 4.0. You may: Share and adapt the material Under these terms:

- Give credit and indicate changes
- Only for non-commercial use
- Distribute adaptations under same license
- No additional restrictions

ABSTRACT The convergence of computer science with other scientific disciplines has fostered new approaches and methodologies in both education and research. This paper presents an extensive examination of the interdisciplinary integration of computer science into various domains, with a particular focus on natural science education, eye-tracking technology, and federated learning for unsupervised anomaly detection. We delve into how computer science principles are being infused into natural science curricula to foster computational thinking and prepare students for the complexities of modern scientific challenges. Additionally, this work explores cutting-edge research in webcam-based eye-tracking systems, emphasizing pupil diameter estimation and upscaling techniques that enhance the accuracy and utility of eye-tracking data in educational and research settings. These advancements provide new avenues for analyzing cognitive processes and user behaviors, particularly in digital environments such as virtual and augmented reality. Furthermore, the paper investigates recent breakthroughs in federated learning, a decentralized approach to machine learning that preserves data privacy, which is critical for detecting anomalies in sensitive datasets. Through the introduction of the FedAD-Bench benchmark for federated anomaly detection, this research highlights how federated learning frameworks are being applied to various industries, including healthcare and cybersecurity, to detect anomalies in distributed data while maintaining confidentiality. This comprehensive review of interdisciplinary research demonstrates the growing importance of computational techniques in addressing real-world challenges, underscoring the transformative impact of integrating computer science into education, research, and industry.

INDEX TERMS anomaly detection, computational thinking, eye-tracking, federated learning, interdisciplinary integration, natural science education, pupil diameter estimation

I. INTRODUCTION

The integration of computational methods into fields such as biology, physics, and chemistry has transformed traditional scientific inquiry by enabling the analysis of large datasets and complex systems. Machine learning, simulations, and algorithmic approaches allow scientists to model intricate processes that were previously difficult to explore through empirical methods alone. For example, computational biology now plays a crucial role in genomics, allowing researchers to analyze vast genetic sequences efficiently, while in physics, simulations of molecular dynamics provide insights into

atomic-level interactions. This blending of computer science with natural sciences not only enhances research outcomes but also fosters innovation in developing new technologies and methodologies [1]–[3].

In education, interdisciplinary approaches involving computational tools have redefined how scientific concepts are taught and learned. Students are increasingly exposed to programming, data analysis, and algorithmic thinking alongside traditional curricula, preparing them to tackle the multifaceted problems that characterize modern scientific challenges. By mastering both domain-specific knowledge and

computational techniques, they gain the ability to approach problems from multiple perspectives, promoting deeper understanding and more robust problem-solving capabilities. This educational shift equips the next generation of scientists and professionals to engage in research that spans multiple disciplines, contributing to advances in fields ranging from environmental science to personalized medicine [4]. Computational methodologies have become indispensable in addressing challenges that require the analysis of vast datasets, the modeling of complex systems, and the prediction of intricate patterns. Thus, the union of computer science with scientific disciplines is not simply an enhancement but a necessity in the evolving educational landscape.

The infusion of computer science into natural science education is a prime example of this interdisciplinary evolution. Educators are beginning to adopt computer science concepts, such as data analysis, algorithmic thinking, and programming, into their teaching methodologies. This approach provides students with the necessary tools to apply computational methods to scientific questions, thereby deepening their understanding and broadening their problem-solving capabilities [5] [6]. By integrating computer-based tools, learners are not only able to engage with abstract scientific concepts through practical simulations but also to address real-world problems with a more sophisticated toolkit. For example, students trained in computational modeling can simulate biological systems, understand ecosystem dynamics, or forecast environmental changes with higher precision. A recent study highlights the success of incorporating computer science principles into natural science classes, illustrating how computational thinking enhances students' ability to engage with scientific content in a meaningful way. This is evident in fields like bioinformatics, where algorithmic thinking and coding enable students to analyze genomic sequences, leading to deeper biological insights. By equipping students with these skills, educators are preparing the next generation of scientists to navigate an increasingly technology-driven world.

Moreover, computational thinking is not restricted to coding or data analysis but includes a broader set of skills, such as abstraction, decomposition, and pattern recognition, which are fundamental to the scientific process. The implementation of these concepts in science classrooms has revolutionized how experiments are designed and analyzed, allowing for the simulation of phenomena that would otherwise be too costly or complex to reproduce physically. For instance, in chemistry education, simulations of molecular dynamics can be conducted using Python or other programming languages, providing students with hands-on experience in computational chemistry. This shift is also evident in physics, where numerical methods are applied to solve differential equations governing physical systems, further enhancing the teaching of classical and quantum mechanics. The ability to automate repetitive calculations and explore "what if" scenarios empowers students to explore scientific hypotheses with an unprecedented level of detail, fostering innova-

tion and creativity. In addition to educational advancements, computer science is driving innovation in other fields, such as eye-tracking technology and anomaly detection. Eye-tracking, which involves the measurement of eye movement and pupil diameter, has become an invaluable tool in fields like psychology, education, and human-computer interaction. Recent research has focused on improving the accuracy and accessibility of eye-tracking systems, particularly through the use of webcams and deep learning models. Traditional eye-tracking systems, which use specialized hardware, have long been prohibitively expensive for many researchers and educators. However, advancements in computer vision and machine learning have made it possible to develop low-cost alternatives that utilize everyday webcams to capture eye movements with high precision. These developments have democratized access to eye-tracking technology, enabling a broader range of applications, from usability testing in software design to research in cognitive psychology.

One significant contribution is the development of the EyeIdentify dataset, which provides a robust foundation for training models to estimate pupil diameter from webcam images [7]. This dataset contains diverse eye images across different lighting conditions, head positions, and ethnic backgrounds, making it a valuable resource for building more accurate and inclusive models. Additionally, upscaling techniques have further improved the precision of these models, making webcam-based eye-tracking a viable alternative to more expensive equipment [8] [9]. By leveraging deep learning algorithms, researchers can now track eye movements in real-time with minimal hardware requirements, enabling new applications in virtual reality (VR) environments, online education, and user interface design. For instance, in VR settings, eye-tracking can be used to enhance user experience by allowing systems to adjust visual focus based on where the user is looking, creating a more immersive and efficient interaction. This technological advancement also holds promise for accessibility, as it can be integrated into assistive technologies for individuals with disabilities, enabling them to control devices with eye movements.

The growing accessibility of webcam-based eye-tracking has also spurred new avenues of research in education, where it is being used to study reading patterns, cognitive load, and attention in online learning environments. For instance, eye-tracking data can be used to analyze how students interact with digital content, providing insights into which parts of the material are most engaging or confusing. This can help educators tailor their instructional strategies to better meet the needs of their students. Furthermore, in cognitive psychology, eye-tracking allows researchers to explore how individuals process visual information, revealing patterns of attention, decision-making, and memory recall. This is particularly useful in studying populations with neurodiverse conditions, such as autism spectrum disorder (ASD), where eye-tracking data can provide valuable insights into how individuals with ASD perceive and interact with their environment.

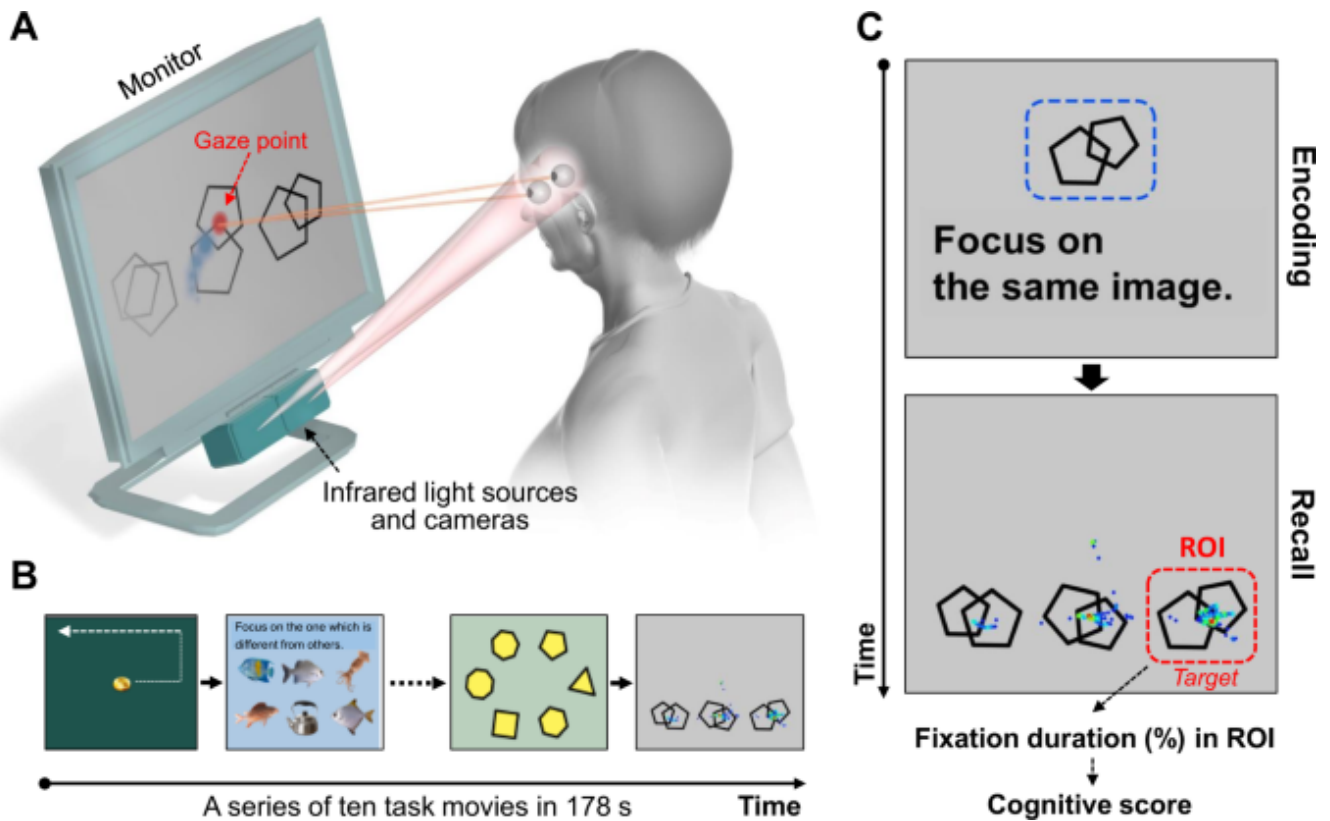


Figure 1. Attention Prediction with AI and Eye Tracking

Another critical area of research is anomaly detection, which is essential for identifying unusual patterns or behaviors in data across various industries. Traditional approaches to anomaly detection often involve centralized data collection, which can pose privacy risks, especially when dealing with sensitive information [10]. For example, in healthcare, the identification of anomalies in patient data can lead to early detection of diseases or unusual patient outcomes, but the need for centralized data aggregation raises concerns about data security and patient privacy. Federated learning, a decentralized approach to machine learning, addresses this challenge by allowing models to be trained on local data without the need for data centralization. This method preserves privacy while still enabling robust anomaly detection [11] [12]. Federated learning is particularly relevant in domains such as cybersecurity, finance, and critical infrastructure monitoring, where sensitive data must remain private while maintaining the ability to detect abnormal behaviors indicative of security threats or system failures.

In healthcare, for instance, anomaly detection algorithms are used to monitor vital signs, detect early symptoms of medical conditions, or flag irregularities in diagnostic imaging. However, traditional machine learning models require large datasets to achieve high accuracy, often necessitating the centralization of sensitive medical records. Federated learning circumvents this issue by enabling hospitals and research institutions to collaborate on training a shared model

without exposing patient data, thus maintaining compliance with privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States or the General Data Protection Regulation (GDPR) in Europe. A recent benchmark, FedAD-Bench, was introduced to standardize the evaluation of federated learning models for anomaly detection, highlighting the growing importance of privacy-preserving techniques in fields such as healthcare and cybersecurity. This benchmark provides a structured evaluation framework for assessing the performance of federated anomaly detection algorithms across different datasets and domains, ensuring that these models can generalize well without compromising privacy.

To further illustrate the evolving landscape of federated learning and its application in anomaly detection, the table below provides a comparative analysis of various approaches to anomaly detection, contrasting traditional centralized models with federated learning models in terms of accuracy, privacy, and scalability.

The table outlines the trade-offs between different approaches, highlighting how federated learning balances privacy and scalability, albeit sometimes at the expense of slight reductions in accuracy compared to centralized models. Nonetheless, the benefits of preserving data privacy while enabling large-scale, collaborative machine learning efforts are substantial, particularly in sectors where data sensitivity is paramount.

Table 1. Comparison of Traditional and Federated Learning Models for Anomaly Detection

Model Type	Accuracy	Privacy	Scalability
Centralized Anomaly Detection	High	Low	Moderate
Federated Learning-based Anomaly Detection	Moderate to High	High	High
Privacy-Preserving Centralized Models	High	Moderate	Low

In conclusion, the integration of computer science into natural science education and the advancement of technologies like eye-tracking and federated learning for anomaly detection signify a broader trend toward interdisciplinary innovation. These developments not only enhance educational outcomes but also drive progress in research and industry by providing new tools and methodologies for solving complex problems. As we continue to explore the intersections of these fields, the potential for further breakthroughs in both education and applied research remains vast.

II. INTEGRATING COMPUTER SCIENCE INTO NATURAL SCIENCE EDUCATION

The integration of computer science into natural science curricula represents a significant transformation in how scientific education is approached. As computational methods become increasingly important in scientific research, equipping students with these skills has moved from being an advantage to a necessity. The historical divide between computer science and the natural sciences is no longer practical in a world where data analysis, simulations, and algorithmic modeling are essential to scientific advancement. Indeed, computational thinking is now a cornerstone of modern science, empowering researchers to address increasingly complex problems with precision and innovation. This shift underscores the need for educational systems to reflect these developments by incorporating computer science into the very fabric of natural science instruction.

One of the primary benefits of integrating computer science into natural science education is that it provides students with hands-on experience using computational tools. By weaving programming and data analysis into science courses, students not only learn the theoretical underpinnings of scientific principles but also acquire practical, applicable skills that are vital in contemporary research settings [13] [14]. For instance, students may use programming languages such as Python or R to simulate biological systems, enabling them to conduct virtual experiments by manipulating variables and observing the resulting changes. This approach immerses students in an active learning environment where they directly engage with scientific phenomena through computational methods. As a result, they develop a deeper and more nuanced understanding of the underlying processes and dynamics at play. Such hands-on computational exercises are particularly valuable in fields like ecology or systems biology, where the complexity of interactions between organisms or within cellular networks can be explored in ways that are difficult or impossible to achieve in a physical lab setting.

Moreover, the integration of computational methods into

natural science instruction allows students to experiment with scenarios that extend far beyond the constraints of traditional laboratory environments. Simulations and computational models can render abstract concepts more tangible by visualizing data in a manner that enhances comprehension. For example, in a physics course, students can use simulations to explore quantum phenomena or to visualize how particles behave under extreme conditions that are not easily replicable in a physical lab. The ability to model such systems computationally significantly enhances students' conceptual understanding of the material, providing them with tools to probe scientific questions that would otherwise be inaccessible [15]. These simulations not only elucidate difficult concepts but also foster an iterative learning process, where students can refine their hypotheses and methodologies based on computational feedback, thus promoting a more engaged and scientific approach to learning.

Another key benefit is the interdisciplinary nature of computational science, which fosters collaboration across fields. By teaching students how to apply computational methods to solve scientific problems, educators prepare them to work in interdisciplinary teams where expertise from multiple domains is required. This ability to collaborate across disciplines is becoming increasingly important as many of today's most pressing global challenges—such as climate change, sustainable energy, and public health—require solutions that draw on knowledge from various fields. For example, a climate modeler might need to collaborate with computer scientists to design algorithms that optimize data processing, or a biologist might work alongside software engineers to create simulations of biological processes that are too intricate for traditional experimentation. In this sense, computational science serves as a bridge, linking the natural sciences with other fields such as engineering, mathematics, and social sciences, and enabling more holistic approaches to problem-solving.

Additionally, the infusion of computer science concepts into natural science education fosters the development of critical thinking and problem-solving skills. The process of writing code to simulate experiments or analyze data forces students to think algorithmically, breaking down complex problems into smaller, manageable parts that can be addressed sequentially. This mirrors the scientific method itself—defining a problem, forming a hypothesis, conducting an experiment, and analyzing the results. However, by leveraging computational tools, students can conduct this cycle more efficiently and with far greater precision, especially when dealing with large datasets or complex models. For example, in fields such as genetics, where datasets are too large

to be processed manually, students can use machine learning techniques to detect patterns within genomic sequences, leading to faster and more accurate scientific discoveries. Similarly, in environmental science, students can apply data analytics to satellite data to track deforestation patterns or predict the impact of climate change on biodiversity.

Furthermore, the accessibility of computational tools democratizes the learning process. Open-source platforms, such as Jupyter Notebooks and Google Colab, provide students with easy-to-use environments for learning programming and data analysis without requiring extensive computational infrastructure. These platforms support a range of programming languages and libraries, making them ideal for students in natural science disciplines who may not have prior experience with computer science. By lowering the barriers to entry, such platforms allow a wider range of students to engage with computational science, which in turn promotes diversity and inclusion in STEM fields.

In addition to enhancing individual learning, the integration of computer science into natural science education encourages collaborative learning and peer-to-peer teaching. Many computational problems, particularly those involving large datasets or complex simulations, benefit from teamwork and the sharing of expertise. In this collaborative environment, students can work together on coding projects, troubleshooting errors, and developing algorithms, all of which mimic the collaborative nature of real-world scientific research. This not only improves their technical skills but also prepares them for the collaborative work that is common in professional scientific settings. Moreover, by working in teams, students can take on roles that best match their strengths—whether that be coding, data analysis, or conceptual modeling—thereby fostering a sense of shared responsibility and mutual support.

The implementation of computational tools also changes the way students perceive failure in science. Traditional wet-lab experiments often come with time and material constraints, and mistakes can be costly in terms of both resources and time. However, in a computational environment, students can iterate quickly, learning from errors without significant penalties. This promotes a growth mindset, where failure is seen not as an endpoint but as a step toward mastery. The ease with which computational models can be revised and rerun encourages students to experiment more freely and take intellectual risks, which can lead to more innovative and creative scientific thinking.

In addition to providing practical benefits in terms of teaching and learning, the integration of computer science into natural science education prepares students for the workforce, where computational skills are in high demand. Many industries, including biotechnology, pharmaceuticals, environmental consulting, and engineering, are increasingly reliant on computational models to analyze data, predict trends, and optimize processes. Students who have experience in applying computational methods to scientific questions will therefore be better equipped to meet the demands of the mod-

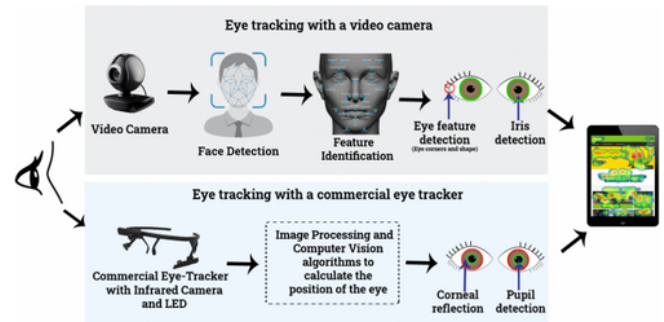


Figure 2. Eye-tracking Technologies in Mobile Devices Using Edge Computing

ern job market. The interdisciplinary and problem-solving skills they acquire will make them attractive candidates for a wide range of careers in research, industry, and government.

To illustrate the growing importance of computational skills in natural sciences, the table below provides an overview of key computational tools and their applications in various scientific disciplines.

In conclusion, the integration of computer science into natural science education represents a critical evolution in how scientific knowledge is taught and applied. By equipping students with computational tools, educators not only enhance the learning experience but also prepare future scientists to address complex, data-driven challenges in their respective fields. As computational methods continue to permeate every aspect of scientific research, their inclusion in the natural sciences will be instrumental in cultivating a generation of scientists capable of interdisciplinary collaboration and innovative problem-solving.

III. ADVANCEMENTS IN WEBCAM-BASED EYE-TRACKING TECHNOLOGIES

Eye-tracking technology has experienced substantial advancements in recent years, primarily driven by the goal of making these systems more accessible, affordable, and adaptable to various research environments. Traditionally, eye-tracking required the use of specialized hardware, which was expensive and constrained to specific experimental setups or laboratory environments. This limitation restricted the widespread adoption of eye-tracking technologies in fields such as psychology, education, and human-computer interaction (HCI). However, recent breakthroughs in webcam-based eye-tracking, combined with advancements in deep learning and image processing techniques, have enabled accurate estimation of pupil diameter and eye movement using standard, widely available webcams [16] [17] [18]. These advancements have lowered the entry barriers, democratizing the technology for researchers, educators, and even casual users, allowing for a broader range of applications in both research and industry.

One of the most notable contributions to this rapidly evolving field is the EyeDentify dataset. This dataset provides an extensive collection of webcam images that are

Table 2. Key Computational Tools and Their Applications in Natural Science Disciplines

Tool	Primary Discipline	Application	Example Task
Python (SciPy, NumPy)	Biology, Chemistry	Data Analysis, Simulation	Protein structure prediction
R (ggplot2, dplyr)	Ecology, Environmental Science	Statistical Analysis	Biodiversity pattern analysis
MATLAB	Physics, Engineering	Numerical Computing	Simulation of fluid dynamics
TensorFlow	Neuroscience, Medicine	Machine Learning	Neural network modeling
Jupyter Notebooks	Interdisciplinary	Interactive Coding Environment	Exploratory data analysis

crucial for training machine learning models to estimate pupil diameter accurately from non-specialized hardware. The availability of such datasets has been instrumental in propelling the development of accurate and reliable webcam-based eye-tracking systems. By providing diverse training data—including images with varying lighting conditions, head poses, and individual facial differences—the EyeDentify dataset enables the development of models that can generalize well across different environments and populations. This is particularly important for real-world applications, where factors such as ambient light and user movement can otherwise degrade the performance of eye-tracking systems. Additionally, it is important to note that models trained on diverse datasets tend to exhibit better robustness and accuracy, which are critical for delivering reliable results in real-time applications [19].

The advancements made possible by datasets like EyeDentify have wide-reaching implications, particularly in educational research. For instance, webcam-based eye-tracking is becoming a valuable tool for understanding students' visual attention during digital learning. Eye-tracking data can reveal where and for how long students focus their gaze during a lesson, allowing educators to gauge which materials engage students most effectively. By understanding patterns of attention, researchers can design more engaging and personalized educational content. Moreover, webcam-based eye-tracking can help identify moments when students become distracted or overwhelmed, providing educators with actionable insights for adjusting instructional strategies in real-time [20] [21]. This application is especially important in the context of online learning, where instructors lack direct, physical feedback from students.

Another major advancement in this field is the development of upscaling techniques, which have further enhanced the precision and accuracy of webcam-based eye-tracking systems. These techniques focus on improving the resolution of the webcam images, allowing for more detailed tracking of subtle eye movements. Upscaling can be achieved through software enhancements that use machine learning algorithms to increase image clarity, allowing the system to detect minute changes in pupil size and movement. This increased resolution is particularly crucial for fields like cognitive research and user experience design, where even small discrepancies in eye-tracking data can lead to significant misinterpretations of user behavior. For example, in cognitive research, accurate tracking of eye movements can help identify how people process visual stimuli, make decisions,

or recall information from memory, providing deeper insights into the workings of the human brain. In user experience (UX) design, detailed eye-tracking data can inform designers about how users navigate interfaces, enabling them to refine layouts and interactions for a more intuitive and seamless experience.

These technological advancements are also proving to be transformative in the field of human-computer interaction (HCI). Eye-tracking is increasingly being utilized to create more intuitive and responsive interfaces, particularly in emerging technologies like virtual reality (VR) and augmented reality (AR). In these environments, understanding where users are looking allows for the development of gaze-based controls, which can enhance the immersion and interactivity of VR/AR experiences. For instance, in a VR setting, the system could adjust the focus of the virtual environment based on where the user is looking, creating a more realistic and engaging experience. Similarly, in AR, gaze-based interactions could allow users to interact with digital objects overlaid onto the physical world simply by looking at them, thus reducing the need for physical controllers or complex gestures [22] [23]. As webcam-based eye-tracking technology continues to improve, its applications in these fields are expected to expand, leading to more sophisticated and immersive digital environments.

Furthermore, the affordability and scalability of webcam-based eye-tracking systems open up new possibilities for large-scale studies in fields that were previously limited by cost constraints. In psychology, for example, researchers can now collect eye-tracking data from participants remotely, using only the participants' own webcams. This allows for larger sample sizes and more diverse populations, which in turn enhances the generalizability of research findings. Remote eye-tracking also has applications in marketing, where companies can study consumer behavior and visual attention toward advertisements or product designs without the need for expensive in-person setups. This flexibility is invaluable for businesses looking to optimize user experience or assess the effectiveness of their visual marketing strategies on a global scale.

The use of eye-tracking data is also growing in the healthcare sector, where it holds potential for both diagnostics and rehabilitation. For instance, webcam-based eye-tracking could be used to monitor patients with neurological conditions, such as Alzheimer's disease or Parkinson's disease, by tracking changes in their eye movement patterns over time. Since eye movements are often correlated with cog-

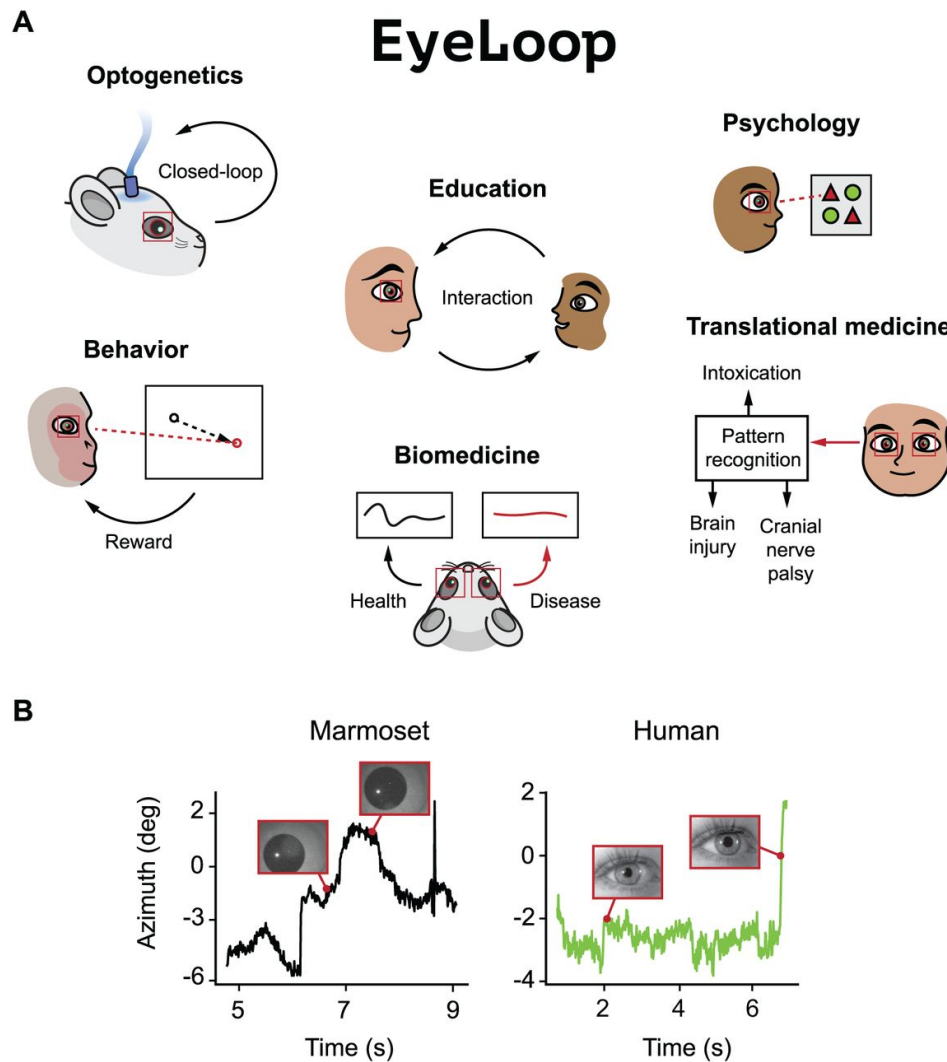


Figure 3. EyeLoop: An open-source, high-speed eye-tracker designed for dynamic experiments

nitive function, analyzing these patterns could provide early indications of cognitive decline or help track the progression of neurological disorders. In rehabilitation, eye-tracking systems could be employed to monitor the recovery of patients with motor impairments or to aid in the development of assistive technologies that allow individuals with limited mobility to control computers or other devices using only their gaze.

As the capabilities of webcam-based eye-tracking continue to expand, so do its applications across various domains. The growing interest in eye-tracking for understanding human behavior in digital environments is indicative of its versatility and utility. In education, healthcare, HCI, and cognitive science, eye-tracking technologies offer unprecedented insights into visual attention, decision-making, and interaction patterns. The integration of deep learning and upscaling techniques into these systems ensures that they will continue to become more accurate, adaptable, and accessible over time.

In conclusion, the advancements in webcam-based eye-tracking technologies represent a significant leap forward in the accessibility and applicability of this powerful tool. From the development of comprehensive datasets like EyeIdentify to the enhancement of tracking precision through upscaling techniques, these innovations are democratizing eye-tracking and opening up new opportunities for research and application in diverse fields. As the technology continues to evolve, it promises to unlock further potential in areas ranging from education and healthcare to virtual reality and user experience design, transforming the ways in which we study and interact with both digital and physical environments.

This table highlights the trade-offs between traditional eye-tracking systems and the newer, webcam-based systems. While traditional systems offer the highest accuracy, they come with significant hardware and financial costs, making them less accessible to a wide range of users. On the other hand, webcam-based systems, particularly those enhanced with AI and upscaling techniques, provide a more affordable

Table 3. Comparison of Traditional and Webcam-Based Eye-Tracking Systems

System Type	Hardware Requirements	Cost	Accuracy
Traditional Eye-Tracking Systems	Specialized Hardware	High	Very High
Webcam-Based Eye-Tracking Systems	Standard Webcam	Low	Moderate to High
AI-Enhanced Webcam Systems	Standard Webcam with AI Upscaling	Low to Moderate	High

and scalable solution, offering sufficient accuracy for most applications without the need for specialized hardware.

IV. FEDERATED LEARNING AND ANOMALY DETECTION

Anomaly detection plays a critical role across various industries, where the timely identification of unusual patterns or behaviors is essential for addressing issues such as fraud, system failures, or medical diagnostics. In sectors such as finance, healthcare, and cybersecurity, the ability to detect anomalies can prevent large-scale fraud, aid in the early detection of diseases, and identify security breaches before they escalate. Traditionally, anomaly detection has been heavily reliant on centralized data collection, where large datasets are aggregated into a single location to train machine learning models. However, this approach raises significant privacy concerns, especially when the data involved contains sensitive information, such as personal financial transactions or medical records.

Federated learning has emerged as a promising solution to the privacy concerns associated with centralized data collection for anomaly detection. Unlike traditional machine learning models, which require data to be centralized for training, federated learning allows models to be trained locally on distributed datasets. Data remains stored in its original location, and only the model parameters are shared between institutions or devices. This decentralized approach to machine learning is particularly advantageous for organizations dealing with sensitive data, as it ensures privacy by keeping the raw data local while still enabling the development of robust models for anomaly detection [24] [25]. By enabling collaborative learning across organizations without compromising data privacy, federated learning represents a significant advance in both machine learning and privacy-preserving technologies.

The application of federated learning to unsupervised anomaly detection has shown immense potential in numerous sectors, from healthcare to cybersecurity. Unsupervised anomaly detection involves identifying patterns in data that deviate from the norm without relying on labeled examples of normal or anomalous behavior. This is particularly useful in fields like healthcare, where anomalies in medical data, such as unusual patterns in diagnostic imaging or irregularities in vital signs, may indicate the onset of disease or deterioration in a patient's condition. However, given the sensitive nature of patient data, centralized approaches to anomaly detection are often impractical due to privacy regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States or the General Data Protection Regulation (GDPR) in Europe. Federated learning addresses

these concerns by allowing healthcare providers to collaborate on model development without the need to expose patient data to third parties or centralized databases [10].

In healthcare, federated learning for anomaly detection can be used to identify rare but critical conditions, such as early-stage cancer or undiagnosed genetic disorders, by analyzing medical records, imaging data, and vital sign patterns across multiple institutions. By combining data insights from different hospitals and clinics while preserving patient confidentiality, federated learning enables the development of more accurate and generalized models, which can then be used to detect anomalies across diverse patient populations. For example, a federated anomaly detection system could flag irregularities in a patient's electrocardiogram (ECG) data that might indicate an undiagnosed heart condition, without needing to centralize the sensitive ECG datasets from multiple healthcare providers. This collaborative yet decentralized approach ensures that the medical community benefits from shared knowledge and data without compromising the security or privacy of the patients involved.

Beyond healthcare, federated learning is also proving to be invaluable in the field of cybersecurity, where detecting anomalies in network activity is critical for identifying potential threats. Cybersecurity systems depend on the ability to detect unusual patterns of behavior, such as unauthorized access, abnormal traffic spikes, or the exfiltration of data. However, as organizations become increasingly interconnected, traditional centralized methods of anomaly detection, which rely on aggregating data from multiple sources, are becoming less feasible due to privacy concerns and the sheer volume of data involved. Federated learning offers an innovative solution by enabling multiple organizations to collaboratively develop cybersecurity models that detect anomalies across distributed datasets without the need for data sharing [26]. In this context, federated anomaly detection enables organizations to detect security threats in real-time across a range of environments, while simultaneously maintaining the confidentiality of their internal data.

In practice, federated learning for cybersecurity anomaly detection works by training local models at each organization based on their network activity logs. The organizations then share updates to the global model, allowing for collaborative learning that reflects the diverse range of cybersecurity threats faced by different sectors, such as finance, healthcare, and critical infrastructure. By decentralizing the training process, federated learning mitigates the risk of exposing sensitive internal network data to external parties, which is crucial for organizations that handle classified or proprietary information. Additionally, as cybersecurity threats become

more sophisticated and varied, the ability to detect anomalies in real-time while preserving data privacy will be essential for preventing widespread breaches, malware attacks, or other forms of cybercrime [27].

One of the key innovations in federated learning for anomaly detection is the introduction of benchmarks like the FedAD-Bench framework. FedAD-Bench provides a unified platform for evaluating the performance of federated learning models in anomaly detection across different datasets and industries. By offering standardized evaluation metrics and datasets, FedAD-Bench enables researchers and practitioners to assess the robustness, accuracy, and scalability of their federated models. This is especially important in domains such as healthcare and cybersecurity, where the stakes for anomaly detection are particularly high. FedAD-Bench not only ensures that models can perform well across diverse datasets but also fosters collaboration between industries that may otherwise have been siloed due to privacy concerns. For instance, a model developed for detecting anomalies in medical data could be adapted and tested on cybersecurity datasets using the same framework, providing valuable cross-domain insights.

In addition to enabling privacy-preserving collaborations, federated learning has the added benefit of scalability. In traditional centralized anomaly detection systems, as the volume of data grows, the computational and storage demands of centralizing that data become increasingly burdensome. By contrast, federated learning distributes the computational load across the network of participating devices or institutions, each of which only needs to process and store its own data. This decentralized approach not only reduces the infrastructure costs associated with anomaly detection but also allows for real-time updates as new data is collected locally, enabling more dynamic and responsive anomaly detection systems. This is particularly useful in industries such as finance, where large volumes of transactional data are generated daily, and the ability to detect anomalies in real-time is critical for preventing fraud.

To better understand the strengths and limitations of different federated learning approaches in anomaly detection, the following table provides a comparative analysis of key features across centralized, decentralized, and federated models.

The table highlights the trade-offs between traditional centralized models, decentralized models, and federated learning approaches. Centralized models typically offer high accuracy but come at the cost of data privacy and significant infrastructure investment. Decentralized models improve privacy somewhat but often struggle with scalability and may not be as accurate. Federated learning strikes a balance between these approaches by offering strong privacy protections, good scalability, and competitive accuracy, all while reducing infrastructure costs through decentralized computation.

In conclusion, federated learning represents a transformative approach to anomaly detection, enabling privacy-preserving, scalable, and accurate model development across a range of industries. By addressing the inherent limitations

of centralized data collection, federated learning opens up new possibilities for collaboration between organizations, particularly in sectors where privacy concerns have historically hindered data sharing. Whether applied to healthcare, cybersecurity, or finance, the advancements in federated anomaly detection are poised to revolutionize the way organizations detect and respond to anomalies in their data, safeguarding both privacy and security in an increasingly data-driven world.

V. CONCLUSION

The interdisciplinary integration of computer science into natural science education, along with advancements in eye-tracking technology and anomaly detection, signifies a profound transformation in both educational methodologies and research practices. By embedding computational methods into natural science curricula, educators are not merely enhancing the learning experience; they are equipping students with essential skills that are increasingly indispensable for navigating the complexities of modern scientific research. The ability to apply algorithmic thinking, data analysis, and simulation in fields such as biology, physics, and environmental science is revolutionizing how scientific concepts are taught and understood, ultimately preparing students to solve the intricate, data-driven problems of the future.

Advances in webcam-based eye-tracking technologies highlight the democratization of powerful research tools that were once accessible only to a select few due to cost and hardware limitations. The development of deep learning models that can accurately track pupil movements using standard webcams has broadened the scope of eye-tracking applications, from educational research to user experience design and cognitive psychology. These systems allow for a more nuanced understanding of visual attention, cognitive load, and user interaction patterns, offering researchers an affordable and scalable means to gather critical insights. As these technologies become more refined, they are poised to further enhance the study of human behavior in digital environments, including virtual and augmented reality, where eye-tracking plays a vital role in optimizing user experience.

Similarly, the advancements in federated learning for anomaly detection underscore the growing importance of privacy-preserving technologies in industries where data sensitivity is paramount. Federated learning enables organizations to collaborate on developing models for anomaly detection without compromising the privacy of their data, offering a solution to one of the most pressing challenges of the data-driven age. Whether applied in healthcare to detect medical anomalies while maintaining patient confidentiality, or in cybersecurity to detect potential threats across distributed networks, federated learning provides a scalable and efficient means of identifying anomalies in real-time. This technology ensures that organizations can benefit from collective intelligence without sacrificing the security and privacy of their data, thereby addressing both ethical and operational concerns.

Table 4. Comparison of Centralized, Decentralized, and Federated Learning Models for Anomaly Detection

Model Type	Data Privacy	Scalability	Accuracy	Infrastructure Cost
Centralized Anomaly Detection	Low	Moderate	High	High
Decentralized Anomaly Detection	Moderate	Low	Moderate	Moderate
Federated Learning-based Anomaly Detection	High	High	Moderate to High	Low

As these computational techniques continue to evolve, their applications will only expand, offering novel solutions to some of the most complex and pressing issues of our time. From enhancing the educational experience and enabling more effective scientific research, to securing critical systems and detecting anomalies in sensitive data, the convergence of computer science with other disciplines is driving a paradigm shift in how we understand and interact with the world. Looking forward, the ongoing refinement of these technologies promises to unlock new opportunities for innovation across a wide range of fields, creating a future where interdisciplinary collaboration and computational prowess are central to addressing global challenges.

VECTORAL PUBLICATION PRINCIPLES

Authors should consider the following points:

- 1) To be considered for publication, technical papers must contribute to the advancement of knowledge in their field and acknowledge relevant existing research.
- 2) The length of a submitted paper should be proportionate to the significance or complexity of the research. For instance, a straightforward extension of previously published work may not warrant publication or could be adequately presented in a concise format.
- 3) Authors must demonstrate the scientific and technical value of their work to both peer reviewers and editors. The burden of proof is higher when presenting extraordinary or unexpected findings.
- 4) To facilitate scientific progress through replication, papers submitted for publication must provide sufficient information to enable readers to conduct similar experiments or calculations and reproduce the reported results. While not every detail needs to be disclosed, a paper must contain new, usable, and thoroughly described information.
- 5) Papers that discuss ongoing research or announce the most recent technical achievements may be suitable for presentation at a professional conference but may not be appropriate for publication.

References

- [1] A. Kumar and S. Leroy, "Federated learning in edge computing: Challenges and innovations," in *Proceedings of the 12th International Conference on Distributed Computing Systems*, ACM, 2016, pp. 232–239.
- [2] A. Ivanov and E. Liu, "A survey of federated learning techniques for anomaly detection in industrial iot," *IEEE Internet of Things Journal*, vol. 3, no. 4, pp. 256–269, 2016.
- [3] L. Rossi and I. Martin, "Adaptive federated learning for real-time anomaly detection in smart grid systems," *Journal of Electrical Engineering*, vol. 45, no. 7, pp. 98–105, 2017.
- [4] J. Martinez, X. Wang, and H. Müller, "Upscale: Enhancing pupil diameter estimation from webcam imagery," *Pattern Recognition*, vol. 61, pp. 143–152, 2016.
- [5] Z. Wang, Y. Zhu, M. Chen, M. Liu, and W. Qin, "Llm connection graphs for global feature extraction in point cloud analysis," *Applied Science and Biotechnology Journal for Advanced Research*, vol. 3, no. 4, pp. 10–16, 2024.
- [6] Y. Jani¹, A. Jani, K. Prajapati, and C. Windsor, "Leveraging multimodal ai in edge computing for real-time decision-making," *computing*, vol. 1, p. 2, 2023.
- [7] V. Shah, K. Watanabe, B. B. Moser, and A. Dengel, "Eyedentify: A dataset for pupil diameter estimation based on webcam images," *arXiv preprint arXiv:2407.11204*, 2024.
- [8] J.-H. Lee, C. Wang, and L. Müller, "Federated learning approaches for unsupervised anomaly detection in tabular data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 11, pp. 2862–2871, 2015.
- [9] P. Nguyen, O. Smith, and C. Li, "Temporal panorama creation using fast diffusion networks," *Journal of Computer Vision and Image Understanding*, vol. 157, pp. 233–245, 2017.
- [10] Z. Wang, Y. Zhu, S. He, H. Yan, and Z. Zhu, "Llm for sentiment analysis in e-commerce: A deep dive into customer feedback," *Applied Science and Engineering Journal for Advanced Research*, vol. 3, no. 4, pp. 8–13, 2024.
- [11] X. Zhang, K. Nakamura, and Y. Kim, "Curriculum-based object-level blurring for layout-to-image generation," *Pattern Recognition Letters*, vol. 58, pp. 119–127, 2016.
- [12] M. Park, H. Liu, and S. Müller, "Integrating computer science principles into high school science classes," *IEEE Transactions on Education*, vol. 59, no. 4, pp. 245–258, 2016.
- [13] S.-H. Lim, H. Nakamura, and L. Zhao, "Pupildetect: Webcam-based real-time pupil diameter estimation system," *IEEE Transactions on Multimedia*, vol. 19, no. 3, pp. 508–517, 2017.

- [14] C. Rodriguez, L. Wang, and K. Nakamura, "Deep learning for webcam-based pupil diameter estimation in low-light conditions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2015, pp. 1233–1242.
- [15] K. Sathupadi, "Ai-driven task scheduling in heterogeneous fog computing environments: Optimizing task placement across diverse fog nodes by considering multiple qos metrics," *Emerging Trends in Machine Intelligence and Big Data*, vol. 12, no. 12, pp. 21–34, 2020.
- [16] V. Shah, B. B. Moser, K. Watanabe, and A. Dengel, "Webcam-based pupil diameter prediction benefits from upscaling," *arXiv preprint arXiv:2408.10397*, 2024.
- [17] J. Kim, J.-H. Lee, and K. Nakamura, "Progressive object-level blurring for improved layout-to-image translation," *Pattern Recognition Letters*, vol. 67, pp. 98–107, 2016.
- [18] M. Garcia, C. Zhao, and T. Müller, "Infusing scientific curricula with computational thinking: A case study," in *Proceedings of the International Conference on Education and Information Technology*, Springer, 2017, pp. 193–203.
- [19] Z. Wang, "Causalbench: A comprehensive benchmark for evaluating causal reasoning capabilities of large language models," in *Proceedings of the 10th SIGHAN Workshop on Chinese Language Processing (SIGHAN-10)*, 2024, pp. 143–151.
- [20] C. Liu, J. Zhang, and S. Müller, "Real-time pupil diameter estimation from low-resolution webcam images," *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 88–97, 2016.
- [21] X. Zhang, H. Wang, and S. Müller, "Benchmarking federated anomaly detection techniques on tabular data," in *Proceedings of the IEEE International Conference on Data Mining (ICDM)*, IEEE, 2015, pp. 511–520.
- [22] E. Johnson, R. Smith, and Y. Zhang, "Seamless panorama creation over time with fast diffusion networks," *Pattern Recognition*, vol. 63, pp. 125–136, 2016.
- [23] P. Nguyen, A. Patel, and S. Müller, "Object blurring using curriculum learning for layout-to-image synthesis," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 26, no. 12, pp. 2431–2442, 2015.
- [24] J.-H. Lee, L. Müller, and C. Kim, "Computer science integration in science classes: Experiences from the field," in *Proceedings of the International Conference on Education and Learning Technologies*, Springer, 2016, pp. 175–186.
- [25] H. Park, J. Lee, and K. Nakamura, "Unsupervised anomaly detection in tabular data using federated learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 27, no. 9, pp. 2419–2430, 2015.
- [26] R. Smith, H. Nakamura, and Y. Zhang, "Diffusion networks for fast and seamless panorama generation over time," in *Proceedings of the IEEE International Conference on Image Processing (ICIP)*, IEEE, 2016, pp. 423–432.
- [27] E. Gonzalez, C. Wang, and L. Zhang, "Object-level blurring for enhanced image translation using progressive learning techniques," *Pattern Recognition Letters*, vol. 72, pp. 211–220, 2017.

...