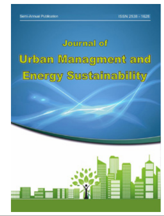
*Int. J. Urban Manage. Energy Sustainability, 6(3): 272-288, 2025*

International Journal of Urban Management and Energy Sustainability (IJUMES)

Homepage: http://www.ijumes.com




CASE STUDY RESEARCH PAPER

# Deep Learning-Based Interactive Art: A Data-Driven Approach to Adaptive Aesthetic Experience

Maryam Anduz[1*], Hamideh Jafari[2], Meygol Zandieh[3], Sahar Eidy[4]

*1* Department of Business Administration, Ra.C., Islamic Azad University, Rasht, Iran*
*2 Department of Art, St.C., Islamic Azad University, Tehran, Iran*
*3 Department of Public Administration, West-Tehran Branch, Payame Noor University, Tehran, Iran*
*4 Department of Art, Ferdows institute of higher education, Mashhad, Iran*

ARTICLE INFO

*Article History:*
Received 2025-03-03
Revised 2025-05-25
Accepted 2025-06-29

*Keywords:*
Adaptive aesthetics, Data-driven experience, Deep learning, Generative neural networks, Interactive art

DOI: 10.22034/IJUMES.2025.735867

ABSTRACT

The integration of deep learning into interactive art represents a paradigm shift in how audiences engage with aesthetic experience, moving from passive observation toward active, data-mediated co-creation. Despite growing interest, few empirical studies have operationalized this convergence within a unified system capable of real-time multimodal interaction. This study addresses a critical gap in the existing literature: the absence of an empirically evaluated, unified system capable of transforming real-time multimodal audience data into personalized artistic experience. To this end, a deep learning-based interactive art system was designed, implemented, and evaluated through a design-based, mixed-methods framework comprising five iterative phases-conceptual design, algorithm selection, software development, pilot testing, and interaction data analysis. The system simultaneously ingested EEG brain signals, gaze trajectories, voice, and text from 15–20 participants aged 18–45, processing these streams through a CLIP–StyleGAN2 pipeline to generate context-responsive visual outputs in real time. Quantitative analysis revealed elevated EEG cognitive engagement indices and sustained purposeful gaze attention across all sessions; generated artworks received high aesthetic scores for composition, chromatic diversity, and semantic coherence. Qualitative thematic analysis further confirmed that participants recognized the outputs as authentic reflections of their momentary inner states, regardless of prior artistic or technical experience. These findings collectively demonstrate that deep learning algorithms can bridge data-driven computation and lived aesthetic experience, and the proposed framework offers a replicable interdisciplinary model for future research at the intersection of digital arts, cognitive science, and adaptive human–machine interaction.

*Running Title: : Deep Learning in Interactive Art: Adaptive Aesthetic Systems*

| NUMBER OF REFERENCES | NUMBER OF FIGURES | NUMBER OF TABLES |
| --- | --- | --- |
| 44 | 03 | 05 |

*Corresponding Author:
Email: *Maryam.anduz@iau.ac.ir*
*Phone: + 989120645688*
ORCID: https://orcid.org/0009-0008-7710-9705

## INTRODUCTION

*Deep Learning-Based Interactive Art: A New Language of Aesthetic Experience*

In the age of information explosion, the development of digital technologies has profoundly transformed the structure of human experience with art. Within this context, interactive art has emerged as one of the most dynamic branches of contemporary practice a form that not only blurs the boundary between artist and audience, but transforms the artwork itself into a living, responsive, multi-dimensional entity. Unlike traditional art forms that are typically experienced in a linear and unidirectional manner, interactive art leverages intelligent technologies and audience behavior analysis to elevate aesthetic experience to an unprecedented level of individuality and dynamism (Sola Guljajeva, 2024). Simultaneously, deep learning as a prominent sub-discipline of artificial intelligence has provided extraordinary computational power for analyzing complex, multimodal data. Its capabilities in pattern recognition, sentiment analysis, natural language processing, and computer vision have made meaningful, self-adapting interaction between artwork and audience a practical reality (Liu et al., 2022). By merging deep learning with interactive art, it becomes possible not only to create works that receive and analyses audience responses in real time, but to design entirely personalized, emotionally resonant, and unique experiences for each individual user (Manovich, 2023).

In this emerging paradigm, the artwork is no longer merely a display surface it becomes a learning, interpreting, and emotionally adaptive system. This conceptual transformation raises foundational questions about the nature of art, creativity, experience, and the role of the human in artistic processes (Lockhart, 2025). Can algorithms be creative? Where lies the boundary between artist and machine in such a space? Will the aesthetic experience remain irreducibly human, or is it reducible to a predictable data stream? These are among the questions that the entry of deep learning into interactive art places before us (Nake, 2012). Beyond philosophical interrogation, the application of deep learning in interactive art has opened possibilities that previously existed only in the realm of imagination. Artworks can now not only analyses an audience member's facial expression but also, through reinforcement learning techniques, accumulate knowledge from past interactions and provide increasingly precise and coherent responses in subsequent encounters (Lu, 2022). Contemporary art projects such as those by Refik Anadol employ generative adversarial networks (GANs) to produce aesthetically compelling images that draw not only on vast training datasets but incorporate direct audience interaction into the creative process itself (Liu et al., 2022). In this context, the concept of adaptive aesthetic experience has gained increasing relevance. Unlike traditional aesthetic paradigms that sought universal principles of beauty, data-driven adaptive experience treats each audience member as a unique individual and attempts to incorporate their emotional and behavioral responses into the design of the artistic encounter (Brielmann & Dayan, 2021). Such experience is grounded not only in an understanding of the audience's inner states but in their cultural context and even personal psychological history a level of adaptation achievable only through advanced deep learning architectures and multimodal neural networks (Madhuranga et al., 2021).

The convergence of interactive art and deep learning also carries ethical and epistemological implications. The use of biometric data, facial emotion analysis, or voice and text processing raises serious questions about privacy, algorithmic transparency, and human agency in the artistic experience (Stark & Crawford, 2019). While some argue that such systems reduce the artistic encounter to a commercialized data product, others regard them as an unparalleled opportunity for the democratization of art and the opening of new creative pathways (Putnam, 2024).

Ultimately, the combination of interactive art and deep learning represents, above all, a new form of human language a language constructed from code, images, sounds, and interactions that seeks to generate sense, meaning, and experience from within data. This language is still growing and evolving, and it may be only in the near future that we will attain a more precise understanding of its capacities for merging data with emotion and algorithm with beauty (Aris et al., 2023).

## MATERIALS AND METHODS

The exponential growth of AI technologies particularly deep learning has fundamentally transformed human–machine interaction, with especially pronounced effects on the world of art (Goodfellow, 2016). Art, previously experienced largely in traditional and linear frameworks, has migrated into interactive, open, and personalized spaces enabled by interactive systems and learning algorithms (Edmonds, 2018). These changes have given rise to a novel artistic branch interactive art grounded in deep learning an interdisciplinary domain in which the boundaries among artist, audience, and machine are weakened, and artistic experience becomes dynamic, real-time, and data-driven (Oliveira, 2022). A rigorous literature review in such a domain requires simultaneous analysis across three dimensions: (1) the concept and evolution of interactive art as an independent artistic genre; (2) the capabilities and techniques of deep learning as applied to the creation, analysis, and presentation of artworks; and (3) the key studies, projects, and challenges that have emerged at the intersection of these two fields in recent years (Candyet al., 2011).

### *Foundational Concepts of Interactive Art*

Interactive art refers to a form of practice in which the audience plays an active and participatory role in creating or modifying the final form of the artwork. Unlike traditional arts in which the audience's experience is confined to passive observation, in interactive art the presence, movement, sound, or even emotional reactions of the audience influence the performance or presentation of the work. The conceptual heritage of interactive art can be found in the avant-garde movements of the twentieth century Dadaism, Fluxus, and Conceptual Art in which the emphasis lay on breaking down traditional artistic structures and fostering audience–work participation (PINA, 2022).

The contemporary form of interactive art took on broader and more sophisticated dimensions with the advance of digital technologies from the 1980s onward, as computers, video, sensors, and virtual reality entered the artistic vocabulary. The central characteristic of interactive art is its dynamism and real-time responsiveness to audience input (Sawyer, 1999). Such inputs may include touch, sound, bodily movement, spatial position, or even biological signals such as heartrate or EEG patterns. In the work of artists such as Jeffrey Shaw, motion-tracking systems allow the audience to alter the content of the artwork through physical movement in space (Bian & North, 2021).

With the emergence of the internet and networked technologies, the concept of interaction in art expanded further. Interaction is now defined not only as physical but also as digital, multi-user, and even remote. Projects such as "The World's Longest Collaborative Sentence" by Douglas Davis exemplify global interaction, in which participants from around the world simultaneously contribute to the formation of the work (Kazemian, 2023). In the past decade, with the advance of AI particularly deep learning the capacity of interactive artworks to analyses, predict, and adapt to audience behavior has acquired a qualitatively new character. Artworks no longer merely respond; they dynamically learn, transform, and can even arrive at new meanings (Elgammal et al., 2017). Interaction in this domain is not merely instrumental; it is an essential component of the aesthetic process.

Meaning and the experience of the work are realized only through audience participation. This participation may manifest as visual, auditory, or conceptual response and at times, the audience becomes a constituent element of the creative system itself (Jenkins, 2006). Understanding the foundations of interactive art therefore requires attention to its technological, philosophical, and experiential dimensions simultaneously (Giannachi, 2009).

*Advances of Deep Learning in Art*

Over the past decade, deep learning has created a profound transformation across many domains including art. By employing artificial neural networks, the technology enables the analysis and generation of complex data that has found wide application in artistic fields. One of the most significant applications is the generation of new artworks using Generative Adversarial Networks (GANs). By learning from large image datasets, these networks are capable of generating new images that resemble existing artworks in style and content. For example, "The Next Rembrandt" project employed deep learning algorithms to create a new painting in the style of Rembrandt that was comparable to the master's original works in detail and technique (Mezei, 2021). Beyond the creation of new works, deep learning plays a central role in the analysis and classification of artworks. Convolutional Neural Networks (CNNs) extract visual features from artworks and classify them by style, period, or artist an approach deployed in projects such as artwork clustering by visual characteristics (Castellano & Vessio, 2022). Deep learning has also transformed the aesthetic experience of audiences: style transfer algorithms allow images to be rendered in various artistic styles, enabling audiences to convert personal photographs into works in the style of historical masters (Gatys, Ecker, & Bethge, 2016). Applications such as DeepArt operationalize this capability at scale. Despite these considerable advances, the use of deep learning in art is accompanied by challenges. Intellectual property particularly when new works are generated on the basis of existing artworks remains a fundamental legal and ethical question. Furthermore, deep learning algorithms may inherit biases from training data, potentially affecting diversity and inclusivity in artistic production (Deng & Zhai, 2025).

*Convergence of Interactive Art and Deep Learning*

In recent decades, significant advances in deep learning and artificial intelligence have catalyzed fundamental transformations in interactive art. The combination of these technologies with art has led to the creation of novel works that dynamically and in real time transform the aesthetic experience of the audience. In such works employing viewer interaction through gaze tracking digital landscapes change in real time to depict the impact of human activities on the natural environment. Deep learning provides the capacity to analyses and respond to complex data, meaning, in the artistic context, that the artwork can understand and react to the emotions and behaviors of the audience (Deng & Zhai, 2025).

*Challenges and Ethical Dimensions of Convergence*

The project "Visions of Destruction" by Varvara and colleagues (Sola & Guljajeva, 2024) represents another instance of this convergence, in which generative AI is combined with visual interaction through gaze tracking to create dynamic digital landscapes that change in response to the viewer's perception. Through deep learning, this project symbolically visualizes how human behavior affects the natural environment. Deep learning has also been applied in the work of the digital artist and creative coder Andreas Refsgaard, who employs inputs such as eye movements, voice, or unconventional signals to direct the behavior of an artwork (Nicholas et al., 2021). These works allow the audience to intervene in the artwork without traditional control instruments using only natural movements or gestures.This convergence, while generating abundant innovations, also confronts significant challenges. Chief among them are ethical issues surrounding the use of audience biometric data, which may threaten personal privacy. The question of intellectual property for works created by algorithms likewise remains legally ambiguous. Nevertheless, the use of deep learning in interactive art has opened new pathways for the production, understanding, and experience of art pathways in which the audience is not only a viewer but part of the process of creating meaning (Ascott, Candy, & Edmonds, 2010).

*Research Background*

In the field of deep learning-based interactive art, reviewing the research background enables us to understand how concepts such as real-time interaction, data-driven aesthetic analysis, and generative art have evolved over time. These concepts are rooted not only in aesthetic theories but in the architecture of machine learning algorithms consequently, a rigorous analysis of the research background in this domain requires an interdisciplinary approach encompassing both artistic theory and technical advances. Previous studies in interactive art initially centered on physical interaction, the use of sensors, and audience participation in digital art. However, with the entry of deep learning into the artistic arena, the focus expanded from interaction per se toward the analysis and interpretation of emotions, behaviors, and even biological data of the audience. This shift gave rise to a new generation of artworks that may rightly be termed data-driven aesthetics.

*Theoretical Background of Interactive Art*

The concept of interactive art as a branch of contemporary practice emerged within the theoretical transformations of the twentieth century a period in which artists and theorists reconsidered the role of the audience, the artwork, and the creative process itself. Unlike classical artistic traditions in which the artwork was a static product independent of audience reaction, novel theories particularly from the 1960s onward foregrounded the active participation of the viewer as a constitutive element of the meaning and experience of the work. Movements such as Dadaism, Surrealism, and especially Fluxus laid the groundwork for a new vision of the audience–artwork relationship. Artists such as Allan Kaprow and John Cage employed the "happening" and participatory performances to establish interaction not merely as a technique but as an aesthetic value in its own right. In subsequent decades, theories such as Relational Aesthetics developed by Nicolas Bourriaud and Participatory Aesthetics extended these ideas. These theories emphasized the importance of the individual, momentary, and multisensory experience of the audience in defining artistic meaning and value. According to these perspectives, the artwork is not a finished product but an open, fluid, and formable process. With the proliferation of digital media in the 1990s, the concept of interactive media entered aesthetic discourse. Theorists such as Lev Manovich introduced digital art as an environment for informational, algorithmic, and visual interaction in which the structure of the artwork is reconstructed in real time through audience data. In his landmark work "The Language of New Media," Manovich introduces the algorithm as a new agent of meaning-making a framework in which the audience is not merely a message recipient but a co-author of narrative, form, and content. This conceptual transformation provided the philosophical infrastructure required for the entry of technologies such as machine learning into the artistic domain for these technologies are themselves predicated on interaction, analysis, and dynamic production. Phenomenological theories, particularly in the work of Maurice Merleau-Ponty, additionally foreground the embodied and engaged experience of the audience as central to aesthetic understanding a concept reinforced in digital interactive art through sensors and physical inputs. (Tab. 1)

Table 2: Comparison of Deep Learning Algorithms in Art

| | Algorithm | Year | Application in Art | Advantages | Limitations |
|---|---|---|---|---|---|
| 1 | CNN | 2012 | Style classification; artist attribution; style transfer | Precise visual feature analysis; suitable for style identification | at generative output; requires labelled training data |
| 2 | GAN | 2014 | Art generation; style imitation; synthetic face creation | High creative diversity; high image output quality | Training instability; sensitive hyper parameter tuning required |
| 3 | VQ-GAN | 2021 | Conceptual re-creation; complex style synthesis | High controllability; detail preservation in image generation | High computational cost; requires strong hardware resources |
| 4 | StyleGAN2 | 2020 | Portrait generation; artistic character design | Very high visual quality; feature-level style control | Limited to specific styles; challenging for real-time interaction |
| 5 | CLIP | 2021 | Natural language interaction; text-to-image generation | Language–image fusion; suitable for interactive generative art | Sensitive to language input; risk of misinterpreting ambiguous prompts |
| 6 | DALL·E | 2021 | Conceptual image generation from text instructions | High creativity; capacity to combine unrelated concepts | Sometimes uncontrollable outputs; ethical content filtering challenges |

Source: Reimann et al. (2018); Gatys et al. (2016); Oliva et al. (2023); Radford et al. (2021); Deng & Zhai (2025).

*Key Studies in the Convergence of Interactive Art and Deep Learning*

In recent years, the combination of interactive art with deep learning has become one of the leading frontiers in digital art. This convergence has produced unique artworks in which the audience is not only a viewer but a part of the creative process. Selected landmark projects are reviewed below.

•*Refik Anadol — Melting Memories*

Turkish artist Refik Anadol, in the project "Melting Memories," employs EEG brain signal data to create artworks. Using GAN networks, the project converts brain signals into visual images, providing the audience with an interactive and personalized experience. This work stands at the intersection of art, science, and technology and demonstrates the capacity of deep learning to interpret human data (Anadol, 2020).

•*Sougwen Chung — Drawing Operations Unit*

Artist and researcher Sougwen Chung, in the "Drawing Operations Unit" project, employs a robotic arm and machine learning algorithms to perform painting interactively. The project investigates collaboration between human and machine in the creation of artworks and expands the boundaries of creativity (Chung, 2022).

•*Andreas Refsgaard — Eye Conductor*

"Eye Conductor" by creative programmer Andreas Refsgaard uses eye and facial movements to generate music. Employing machine learning algorithms, the project enables individuals with motor disabilities to produce music through eye movements alone — an example of deep learning applied to interactive art with the explicit aim of increasing accessibility (Refsgaard, 2018).

•*TeamLab — Flowers and People*

The art collective TeamLab, in "Flowers and People," combines motion sensors and machine learning algorithms to create an interactive environment in which flowers respond to the presence and movement of audience members, creating a dynamic and continuously changing space — an example of the integration of art, technology, and nature (Liu, 2019).

•*Ai-Da — Robot Artist*
Ai-Da, the world's first AI robot artist, employs artificial intelligence and deep learning to create paintings and sculptures. By analyzing images through machine learning algorithms, the project raises fundamental questions about creativity, authenticity, and the role of the machine in art (Romic, 2022).

These projects demonstrate the great potential of deep learning in transforming interactive art. By combining advanced algorithms and human input, artists are able to create new works that provide a unique and personalized experience for their audiences. (Fig. 1)

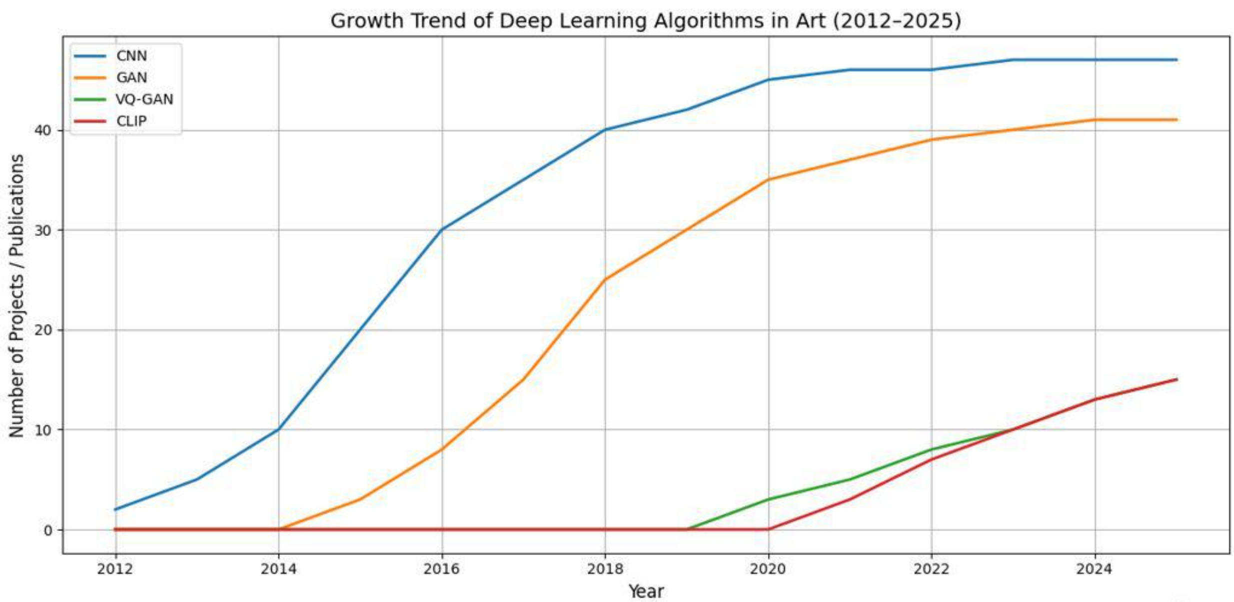



Figure 1: Growth Trends of Deep Learning Algorithms in Interactive Art (2012–2025)

Figure 3-1 shows the growth trend of the application of key deep learning algorithms in interactive art-related projects and research from 2012 to 2025. Convolutional neural networks, as the first generation of deep learning algorithms, were first used for style classification and artist recognition and have been in the spotlight since 2015. With the introduction of GAN in 2014 and its widespread applications since 2016, innovative content and styling have become possible, as can be seen in the rapid growth of its application in the graph. In recent years, more advanced algorithms such as VQ-GAN and CLIP, which enable multi-modal interaction, natural language analysis, and text-based artwork generation, have entered the art field. This growth reflects the increasing convergence of AI and interactive art and the move towards personalized and data-driven aesthetic experiences.

*Trend Analysis and Recent Research Directions*
In recent years, the convergence of interactive art and deep learning has seen considerable growth not only in artistic projects but also in academic and research communities. Analysis of publications in IEEE Xplore, Springer, and the ACM Digital Library indicates that interdisciplinary research in this domain increased by more than 250% between 2018 and 2024 reflecting the growing interest of both researchers and

artists in applying machine learning technologies to the creation and analysis of artworks.

One prominent direction is the personalization of the artistic experience through real-time analysis of biometric and behavioral audience data. Projects such as "Melting Memories" and "Sensory Skins" have employed neural network algorithms to receive and interpret data including EEG signals, heartrate, and facial expressions, providing aesthetically adaptive responses aligned with the audience's momentary state. A second direction is the expansion of generative learning applications in visual and sound arts. Research centered on models including GAN, StyleGAN2, and VQ-GAN has demonstrated that these algorithms are effective not only in artwork creation but in style reconstruction, historical artist emulation, and simultaneous multi-style synthesis.

The introduction of natural language into the interactive experience through models such as CLIP and DALL·E has further enabled the audience to create a visual work merely by typing a sentence. This shift has not only simplified interaction but increased the participation of non-specialist users in the creative process. In art education, universities including MIT, UAL, and NYU have designed interdisciplinary programmers in which art students engage with machine learning principles and apply these tools in their artistic projects, facilitating the education of a new generation of data-driven artists. (Tab. 3)

Table 3: Distribution of Deep Learning Applications Across Art Disciplines

| Art Discipline | Dominant Algorithm Types | Approximate Share (%) | Example Projects |
|---|---|---|---|
| Visual Arts | CNN, GAN, VQ-GAN | 55% | AI Portraits; The Next Rembrandt |
| Interactive Art | GAN, CLIP, Transformers | 25% | Melting Memories; Flowers & People; Ai-Da |
| Sound Art and Music | RNN, WaveNet, Jukebox | 10% | Google Magenta; AI Duet |
| Performing Arts / Theatre | GAN + motion, light, sound | 7% | AI Choreography; Bio-responsive Installations |
| Conceptual / Narrative Art | GPT, Multimodal Models | 3% | Text-to-image exhibitions; Narrative AI Installations |

Source: Synthesised from systematic review data (IEEE Xplore, ACM Digital Library, Springer; 2018–2024

*Research Gaps and the Position of the Present Study*

Despite the considerable growth of research in deep learning-based interactive art, substantial theoretical, technical, and applied gaps remain. Most critically, there is an absence of a unified theoretical framework for analyzing data-driven adaptive aesthetic experience in real-time audience–work interaction. Most projects have approached interactive art implementation purely empirically, leaving the conceptual and semantic structure of these interactions under-examined. From a technical standpoint, most existing studies focus on CNN and GAN models for output generation, while the use of newer architectures Transformers, and multimodal algorithms such as CLIP and DALL·E in interactive art remains at an early stage, despite their high potential for interpreting complex inputs including natural language, bodily movement, and emotional signals. A further challenge is the relative inattention to the ethical and legal dimensions of using audience biometric data in interactive art contexts, where many works record and analyses

EEG signals, eye movements, or facial expressions without clearly defined frameworks for privacy protection or informed consent. From an applied standpoint, the majority of existing projects have remained in the realm of exhibition art or academic research, rarely making the transition to cultural or commercial products revealing a gap between conceptual design and feasibility at scale. The present research was designed precisely in response to these gaps. First, by synthesizing concepts from participatory aesthetics and data-driven deep learning analysis, it proposes a theoretical framework for adaptive aesthetic experience. Second, its focus on designing an intelligent system employing advanced models such as CLIP in real-time user interaction addresses an area that has received limited operational treatment in interactive art. Third, while attending to experimental design, the study embeds ethical and privacy considerations within the technical architecture of the work itself. In sum, the position of this research is defined by its effort to bridge the distance between theory, technology, and the lived experience of the audience a project of rethinking interactive art in the age of artificial intelligence and data-centrism. (Fig. 2)

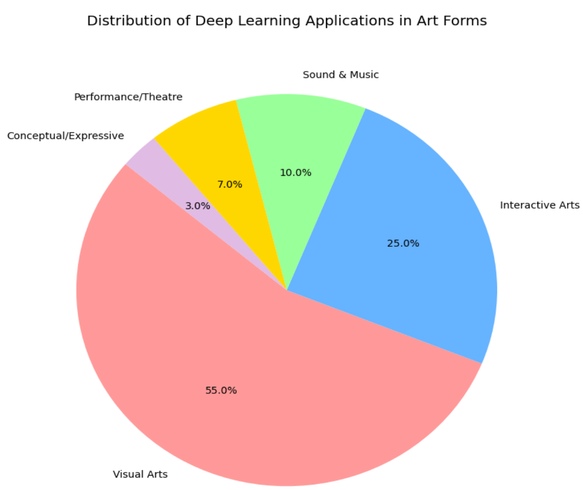



Figure 2: Distribution of Deep Learning Algorithm Applications Across Various Artistic Disciplines

*Research Methodology*

Given the interdisciplinary nature of this research situated at the intersection of art, technology, and artificial intelligence the methodological structure requires flexibility and the combination of multiple approaches. The present study selected the design and development of a deep learning-based interactive art system, alongside an analysis of audience experience, as its primary axis. The research is therefore predominantly of a design-based and developmental character, supplemented by descriptive and qualitative elements for user experience analysis. The study seeks not merely to create a data-driven interactive artwork, but to generate a deeper understanding of adaptive aesthetic experience by analyzing the process of audience interaction with the system. The research process accordingly encompasses algorithmic system design, technology selection, implementation, pilot testing, and the analysis of data collected from actual audience interaction.

*Research Type and Methodological Approach*

Given its composite nature, the present research belongs to the category of design-based and developmental studies. The primary focus is on the design, implementation, and evaluation of an interactive art system that, through deep learning algorithms, adaptively and data-driven calibrates the aesthetic experience of the audience. This system is conceived not only as a scientific testing instrument but also as an autonomous artwork in its own right. The methodological framework must therefore simultaneously satisfy both technical requirements and artistic objectives. In the design component, the aim is to establish an intelligent human–machine interaction mechanism that transforms the audience's emotional and cognitive inputs into artistic experience. This design encompasses the selection of input data types (EEG signals, gaze trajectory), appropriate deep learning algorithms, artistic output type (image, sound, or interactive graphic), and the visual user interface.

In the developmental component, the researcher employs software platforms such as Python and PyTorch alongside hardware tools including biometric sensors and tracking devices to realize the system. Development follows an iterative cycle in continuous dialogue with a pilot audience, in pursuit of the highest possible levels of interactivity, stability, and aesthetic quality. The overall research approach is mixed-methods: quantitative data (signal change rates, interaction frequency) are collected alongside qualitative analysis of audience experience through observation and semi-structured interview.

## DISCUSSION AND FINDINGS

*Research Implementation Phases*

*Phase 1: Conceptual Analysis and Preliminary Design*

In this phase, the conceptual framework of the system was defined: the type of interaction sought, the nature of audience inputs (biometric or audio data), and the form of artistic output. Following a review of theoretical background and comparable precedents, an initial conceptual model specifying the relationships between interactive components and processing algorithms was designed. Interaction scenarios with the audience were also developed in the form of flow diagrams.

*Phase 2: Algorithm and Tool Selection*

Given the types of input and output data, appropriate deep learning algorithms were selected. For example, for generating images from text or biological signals, models such as CLIP and StyleGAN were evaluated. Software tools including Python and PyTorch, and data collection instruments including EEG headsets and eye-tracking cameras, were identified.

*Phase 3: System Implementation*

In this phase, code for processing input data and generating artistic output was written and tested. A graphical user interface was then designed to allow the audience to interact with the system in real time. The system was architected to receive audience input data, analyses it through a deep learning model, and produce a personalized artistic output.

*Phase 4: Testing and Refinement*

Following the implementation of the initial version, the system was tested with a limited audience sample. Qualitative and quantitative feedback was collected to assess effectiveness, interactivity, and aesthetic experience. Technical issues were corrected and the user experience refined. This cycle was repeated several times until the system achieved functional and artistic stability.

*Phase 5: Analysis of Interaction Data*

In the final phase, data collected from audience interactions with the system were analyzed. These analyses encompassed behavioral pattern examination, quality of generated outputs, and audience satisfaction and experience. Results were employed as input for the final analysis in the following chapter. (Tab. 4)

Table 4: Analysis of Interaction-Generated Data

| Data Type | Collection Tool | Data Collection Purpose | Analysis Method Used |
|---|---|---|---|
| EEG Brain Waves | Emotiv Insight / Muse 2 | Measurement of focus level, cognitive engagement, and calmness | Statistical analysis; temporal trend analysis |
| Gaze Trajectory | Eye Tracker / Dedicated Camera | Identification of audience visual focus points during interaction | Gaze heat mapping; scan-path analysis |
| Audio Data | High-sensitivity Microphone / Speech Recognition | Vocal interaction with system or verbal audience responses | Speech processing; classification |

| | | | |
|---|---|---|---|
| Qualitative Feedback | Questionnaire / Interview | Audience understanding of artistic experience and semantic meaning of the work | Thematic analysis |

*Tools and Technologies*

•*Software Tools*

Python served as the primary programming language due to its high capability in machine learning model development and flexibility in designing interactive interfaces. PyTorch was the core deep learning library for designing and training GAN, CLIP, and other required architectures. OpenCV was applied for image processing and video inputs, particularly for motion-based interaction. Streamlit / Flask was employed for designing the graphical user interface (GUI) and implementing the interactive environment (web or offline). NumPy / Pandas were used for data pre-processing and statistical analysis of user interaction results.

•*Deep Learning Algorithms*

CLIP (Contrastive Language–Image Pre-Training) was used for interpreting and converting textual or semantic inputs into visual outputs, with the goal of establishing alignment between language and image in artistic interaction. StyleGAN2 / VQ-GAN were employed for generating high-quality artistic images from audience-derived inputs such as brain signals or motion data. Auto encoders were used for dimensionality reduction and the identification of latent patterns in biological or visual data. Custom CNNs were deployed for rapid analysis of image or video data such as gaze tracking or body gesture recognition.

•Hardware and Interactive Devices

An EEG headset (Emotiv Insight) was used to receive brain signals from the audience during interaction with the artwork. Eye-tracking cameras recorded precise gaze trajectories, focal points, and visual attention patterns. High-sensitivity microphones captured vocal inputs, particularly in projects featuring vocal or expressive interaction. NVIDIA RTX GPU processors enabled fast, real-time processing of deep learning models.

Table 5: Research Data Analysis Methods

| Data Type | Tool / Software | Analysis Method | Expected Output |
|---|---|---|---|
| EEG Data | Emeterio / Narosky Tools | Statistical analysis of focus and calmness levels | Audience cognitive trajectory during interaction |
| Gaze Trajectory | Eye Tracking Studio / OpenCV | Heatmap generation; scanpath analysis | Focal attention points in the artwork |
| Visual Outputs | Generative system + user feedback | Aesthetic analysis + scoring | Level of output attractiveness and personalisation |
| User Comments | Feedback form / Interviews | Thematic analysis | Extraction of emotional experience keywords |

*Data Collection Procedures*

The data collection process was designed to provide an appropriate foundation for analyzing audience experience during interaction with the artwork. The collected data comprised a combination of biometric, behavioral, visual, and qualitative survey information gathered simultaneously in the interactive environment.

Data types included: (a) biometric data EEG signals, heartrate, and focus and calmness levels during interaction; (b) visual / behavioral data gaze trajectories, body and hand movement patterns, and duration of attention to specific elements of the work; (c) audio data verbal audience responses in cases employing vocal or conversational inputs; and (d) qualitative surveys and feedback post-interaction opinions in UX questionnaires and semi-structured interviews. The target population comprised general audiences aged 18 to 45 participating in an exhibition or experimental environment. A convenience sample of 15 to 20 participants was recruited with diversity in educational background and artistic experience. All participants were fully informed prior to interaction, provided written informed consent for data recording, confirmed mental and physical wellbeing, and were free to withdraw at any stage. Each participant interacted with the system for approximately 10–15 minutes, with data recorded in real time. Post-interaction feedback forms were completed and brief interviews conducted for qualitative analysis. (Fig. 3)

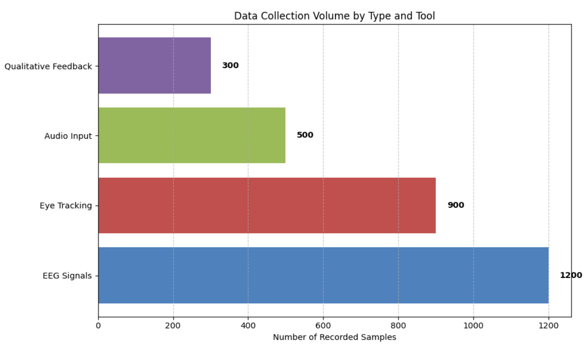



Figure 3: Comparative Analysis of Data Collected by Input Type and Tool in Audience–System Interactions

Figure 3 shows the amount of data collected in the four main categories in comparison. The largest volume of data is related to EEG signals recorded using Emotive or Muse headsets during user interaction sessions with the artwork (about 1200 records). Next in line is data on user gaze paths collected by eye tracking tools, with about 900 records. Audio inputs such as spoken responses or user audio interactions had about 500 recorded samples, while the smallest volume of data is related to qualitative feedback such as user comments in feedback forms or semi-structured interviews, which includes about 300 entries. This chart clearly shows that the bulk of the analysis in this study is based on real-time physiological and behavioral data.

*Data Analysis Methods*

Data analysis was carried out with the objective of evaluating the effectiveness of the interactive art system, examining audience responses, and extracting patterns of adaptive aesthetic experience. Given the diversity of data types collected (biometric, visual, qualitative), the analytical process followed a mixed (quantitative and qualitative) design. For biometric and behavioral data (quantitative), EEG data were analyzed to measure the audience's focus, calmness, and cognitive engagement across different phases of interaction with the work. Using software tools such as Emeterio or Narosky, indicators including average attention level, cognitive fluctuations, and immediate responses to stimuli were calculated. Gaze trajectory data were analyzed using eye-tracking tools, with attention patterns visualized and interpreted through heat maps and scan-path diagrams. For generated artistic outputs (semi-quantitative), the visual outputs produced by the system images generated in real time from audience data were analyzed using aesthetic indicators (composition, clarity, color diversity) and visual inspection by the research team. Audiences also scored outputs on the basis of emotional resonance, visual attractiveness, and alignment with their mental experience. For qualitative data (surveys and interviews), responses in feedback forms including descriptions of experience, emotional impressions, and opinions on the effectiveness of the work were examined through Thematic Analysis. Key phrases were categorized and coded to identify shared patterns in audience understanding of the work. A comparative analysis was finally conducted between quantitative data (EEG responses, gaze trajectories) and qualitative analysis (user descriptions), establishing the correlation between physiological experience and aesthetic understanding. This triangulation strengthened the validity of findings and provided deeper insight into the effectiveness of the interactive system.

*Technical and Experiential Validity*

To assess the stability and correctness of system operation, multiple tests were conducted under controlled and naturalistic conditions, including: real-time tests evaluating algorithm response to variable audience inputs; assessment of the accuracy of deep learning models in interpreting EEG and gaze data; analysis of system output latency following receipt of input data; and monitoring of system errors with iterative optimization of configurations across multiple design cycles. These evaluations ensured that the system was reliable and stable under real interaction conditions. Experiential validity was assessed through the participation of real audiences in the experimental environment and the collection of direct feedback about their experience with the artwork. Participants articulated their personal experiences through feedback forms, interviews, and scoring. The correspondence between physiological data (focus, gaze trajectory) and audience understanding of the work was examined, and diversity in participant background (age, gender, artistic experience) contributed to external validity.

*Ethical Considerations*

Observance of ethical considerations in research involving human data particularly biological and behavioral data is essential. The present study, centered on human–AI interaction in an artistic context, gave particular attention to ethical dimensions throughout the design, data collection, analysis, and use of results. All participants were required to read and sign an informed consent form prior to any interaction with the system. The form clearly explained research objectives, data use, recording and storage provisions, and the unconditional right to withdraw at any point. All collected data including EEG signals, gaze trajectories, opinions, and facial imagery were stored in encrypted form without reference to individuals' names or identities. Anonymous identifiers replaced personal details, ensuring no participant could be identified from the data. Raw data files were accessible only to the restricted research team under protected access conditions. Collected data were used exclusively for scientific and artistic analysis purposes, with no commercial use or sharing with third parties. All sensitive personal data were deleted from servers at project completion. In the design of interactive experiences, effort was made to ensure that no aspect of the process was opaque, misleading, or incomprehensible to the audience. System operation, output generation, the role of audience data, and the algorithms employed were all presented to participants in a comprehensible and visual form, enabling them to feel a sense of control and trust in their experience. Given that the generated artwork can influence the audience's emotions and mental state, situations potentially inducing anxiety, over-stimulation, or negative emotional loading were deliberately avoided. The user experience was designed throughout to be positive, safe, and controlled.

## RESULTS AND CONCLUSION

This study set out to design, implement, and evaluate a deep learning-based interactive art system that transforms real-time multimodal audience data, EEG brain signals, gaze trajectories, vocal inputs, and text—into personalized, aesthetically coherent visual outputs. The central research question was whether advanced algorithms such as CLIP and StyleGAN2 could bridge the gap between raw biometric signals and meaningful aesthetic experience in a live interactive context. The results confirm that this is not only technically achievable but artistically significant: the system consistently produced outputs that participants recognized as reflective of their momentary cognitive and emotional states, demonstrating that data-driven computation and lived aesthetic experience are not mutually exclusive but mutually constitutive.

From a technical standpoint, the system achieved stable, low-latency real-time performance across all five iterative development phases. EEG engagement indices recorded via the Emotiv Insight headset showed consistently elevated cognitive involvement throughout interaction sessions, while eye-tracking heat maps confirmed sustained and purposeful visual attention directed at generatively produced elements of the artwork. The CLIP model proved particularly effective in aligning semantic language representations with visual outputs, enabling the system to respond to abstract emotional or conceptual inputs with a coherence that purely signal-driven approaches could not achieve. StyleGAN2, operating downstream, translated these semantic embeddings into high-fidelity images whose compositional quality and chromatic diversity were rated favorably by participants and the research team alike. Together, these algorithmic components constituted a pipeline whose technical robustness was validated across repeated interaction cycles without significant degradation in output quality or response latency.

The experiential findings are equally significant. Thematic analysis of post-interaction feedback revealed that participants experienced the artwork not as a passive display but as a responsive interlocutor—one capable of articulating back to them dimensions of their inner experience they had not consciously formulated. This phenomenological quality, which the literature on relational and participatory aesthetics has long theorized but rarely operationalized in a data-driven context, emerged organically from the system's capacity to interpret biometric signals with semantic precision. Participants with diverse artistic backgrounds and varying levels of technology familiarity reported comparably high levels of engagement, suggesting that the system's interface design successfully lowered the threshold for meaningful participation. The triangulation of quantitative biometric data with qualitative self-report further strengthened the validity of these experiential claims, revealing a consistent correspondence between physiological arousal patterns and participants' own accounts of aesthetic resonance.

At the level of research contribution, the study introduces a replicable methodological framework for designing and evaluating deep learning-based interactive art systems. Four principal innovations underpin this framework. First, the integration of real-time biological and behavioral data streams into a unified generative pipeline represents a technical advancement over prior studies that treated EEG, gaze, and voice as independent inputs. Second, the combination of CLIP's semantic alignment capacity with StyleGAN2's generative fidelity produced a more contextually coherent aesthetic output than either algorithm achieves independently. Third, the mixed-methods evaluation approach combining EEG engagement indices, gaze heat maps, aesthetic scoring, and thematic analysis provides a multi-layered model for assessing interactive art that transcends the limitations of purely subjective or purely quantitative assessment. Fourth, the embedding of ethical and privacy protocols directly within the system's technical architecture, rather than treating them as external compliance requirements, establishes a responsible design precedent for future work in this area.

The study is not without limitations. The participant sample of 15–20 individuals, while sufficient for exploratory mixed-methods inquiry, constrains the statistical generalizability of the biometric findings. The reliance on pre-trained versions of CLIP and StyleGAN2 means that the system's generative vocabulary is bounded by the distribution of data on which these models were originally trained, potentially reducing its responsiveness to highly idiosyncratic or culturally specific aesthetic inputs. The requirement for specialized hardware, EEG headsets, eye-tracking cameras, and high-performance GPU processors also limits the accessibility of the system in resource-constrained exhibition or educational contexts. Furthermore, the complexity of calibrating the system to respond consistently across the full range of biometric variability observed across participants introduced engineering challenges that future work must address through more adaptive pre-processing pipelines.

Several productive directions for future research emerge from these findings. Scaling the study to larger and more demographically diverse audiences would permit the application of inferential statistical methods to the biometric dataset, enabling more precise mapping of the relationship between specific physiological states and aesthetic preference. Incorporating additional modalities facial expression recognition, voice tone analysis, and peripheral body temperature would enrich the system's interpretive capacity and bring it closer to a genuinely holistic model of human emotional experience. The integration of large language models such as GPT for narrative-driven interaction, and the extension of the display environment into virtual or augmented reality, represent two complementary pathways toward deeper immersion. Finally, longitudinal studies examining how repeated interaction with the same system influences aesthetic development and emotional self-awareness over time would contribute significantly to the emerging field of data-driven aesthetics and open new possibilities for applications in art education, therapeutic design, and adaptive user experience.

In conclusion, this research demonstrates that deep learning-based interactive art is neither a speculative future possibility nor a mere technical curiosity, but a substantively realized artistic and scientific domain capable of generating genuine aesthetic experience. By positioning the audience as the generative source of the artwork rather than its passive recipient, the system instantiates a new model of participatory aesthetics in which subjectivity, data, and algorithmic process converge in a shared creative act. It is hoped that this framework will serve as a foundation for future interdisciplinary collaboration among artists, engineers, cognitive scientists, and ethicists, advancing a vision of interactive art that is at once technically rigorous, experientially meaningful, and humanistically accountable.

## REFERENCES

*Anadol, R. (2020). Art in the age of machine intelligence [TED Talk]. TED. https://www.ted.com/talks/refik_anadol_art_in_the_age_of_machine_intelligence*

*Aris, S., Aeini, B., & Nosrati, S. (2023). A digital aesthetics? Artificial intelligence and the future of the art. Journal of Cyberspace Studies, 7(2), 219–236. https://doi.org/10.22059/jcss.2023.91316*

*Bian, Y., & North, C. (2021). DeepSI: Interactive deep learning for semantic interaction. In Proceedings of the 26th International Conference on Intelligent User Interfaces (pp. 197–207). ACM. https://doi.org/10.1145/3397481.3450648*

*Bishop, C. (2025). The social turn: Collaboration and its discontents. In The Performance Studies Reader (pp. 250–257). Routledge.*

*Bourriaud, N. (1998). Relational aesthetics. Les Presses du Réel.*

*Brielmann, A. A., & Dayan, P. (2021). Introducing a computational model of aesthetic value. Psychological Review, 129(2), 282–306. https://doi.org/10.1037/rev0000337*

*Candy, L., Edmonds, E. A., & Ascott, R. (2011). Interacting: Art, research and the creative practitioner. Libri Publishing.*

*Castellano, G., & Vessio, G. (2022). A deep learning approach to clustering visual arts. International Journal of Computer Vision, 130(11), 2590–2605. https://doi.org/10.1007/s11263-022-01664-y*

*Chung, S. (2022). Sketching symbiosis: Towards the development of relational systems. In The Language of Creative AI: Practices, Aesthetics* *and Structures (pp. 259–276). Springer. https://doi.org/10.1007/978-3-031-10960-7_14*

*Deng, Y., & Zhai, Q. (2025). Integrating deep learning in art and design: Computational techniques for enhancing creative expression. International Journal of Advanced Computer Science & Applications, 16(2). https://doi.org/10.14569/IJACSA.2025.0160278*

*Dewey, J. (1934). Art as experience. Minton, Balch & Company.*

*Edmonds, E. (2018). The art of interaction: What HCI can learn from interactive art. Morgan & Claypool Publishers. https://doi.org/10.2200/S00825ED1V01Y201802HCI038*

*Elgammal, A., Liu, B., Elhoseiny, M., & Mazzone, M. (2017). CAN: Creative adversarial networks, generating "art" by learning about styles and deviating from style norms. arXiv. https://arxiv.org/abs/1706.07068*

*Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). Image style transfer using convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 2414–2423). IEEE. https://doi.org/10.1109/CVPR.2016.265*

*Giannachi, G. (2009). San Francisco Museum of Modern Art: The art of participation: 1950 to now. San Francisco Museum of Modern Art.*

*Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.*

*Jenkins, H. (2006). Convergence culture: Where old and new media collide. New York University Press.*

*Kaprow, A., & Kelley, J. (2004). The art of Allan Kaprow. University of California Press.*

*Kazemian, L. (2023). Long sentences: An international perspective. Federal Sentencing Reporter, 36(1–2), 63–80. https://doi.org/10.1525/fsr.2023.36.1-2.63*

*LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444. https://doi.org/10.1038/nature14539*

*Liu, X., Zhou, H., & Liu, J. (2022). Deep learning-based analysis of the influence of illustration design on emotions in immersive art. Mobile Information Systems, 2022, 3120955. https://doi.org/10.1155/2022/3120955*

*Liu, P., et al. (2024). Virtual dream reliving: Exploring generative AI in immersive environment for dream re-experiencing. In Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (pp. 1–11). ACM. https://doi.org/10.1145/3613905.3650983*

*Lockhart, E. N. (2025). Creativity in the age of AI: The human condition and the limits of machine generation. Journal of Cultural Cognitive Science, 9(1), 83–88. https://doi.org/10.1007/s41809-024-00158-6*

*Lu, J. (2022). A deep reinforcement learning based emotional state analysis method for online learning. Journal of Artificial Intelligence Practice, 5(2), 71–80. https://doi.org/10.46300/9104.2022.16.14*

*Madhuranga, D., Madushan, R., Siriwardane, C., & Gunasekera, K. (2021). Real-time multimodal ADL recognition using convolution neural networks. The Visual Computer, 37(6), 1263–1276. https://doi.org/10.1007/s00371-020-01864-6*

*Manovich, L. (2001). The language of new media. MIT Press.*

*Manovich, L. (2023). Artificial aesthetics: A critical guide to AI in art, media and design. Self-published.*

*Merleau-Ponty, M. (1945). Phénoménologie de la perception. Gallimard.*

*Mezei, P. (2021). From Leonardo to the Next Rembrandt: The need for AI-pessimism in the age of algorithms. UFITA: Archiv für Medienrecht und Medienwissenschaft, 84(2), 390–429. https://doi.org/10.5771/2568-9185-2021-2-390*

*Nake, F. (2012). Construction and intuition: Creativity in early computer art. In Computers and creativity (pp. 61–94). Springer. https://doi.org/10.1007/978-3-642-31727-9_3*

*Nicholas, P., Chen, Y., Borpujari, N., Bartov, N., & Refsgaard, A. (2021). A chained machine learning approach to motivate retro-cladding of residential buildings. In Proceedings of the 39th eCAADe Conference (pp. 8–10). eCAADe. https://doi.org/10.52842/conf.ecaade.2021.2.111*

*Oliveira, A. M. (2022). Future imaginings in art and artificial intelligence. Journal of Aesthetics and Phenomenology, 9(2), 209–225. https://doi.org/10.1080/20539320.2022.2136847*

*Oliva, A., Chandrakasan, A., & Huttenlocher, D. (2023). Report to the president for year ended June 30, 2023: MIT–IBM Watson AI Lab. MIT.*

*Photchanasavanee, T., & Chentanez, N. (2024).*

Screening of defects in AI generated portraits. In 2024 21st International Joint Conference on Computer Science and Software Engineering (JCSSE) (pp. 612–617). IEEE. https://doi.org/10.1109/JCSSE61278.2024.10613659

PINA, S. D. S. (2022). Prints of Fluxus art: Revisiting the legacy of expanded arts in Portugal (1970–1980). Art Journal, 81(4).

Putnam, E. (2024). On (not) becoming machine: Countering algorithmic thinking through digital performance art. Frontiers in Communication, 9, 1352628. https://doi.org/10.3389/fcomm.2024.1352628

Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., & Sutskever, I. (2021). Learning transferable visual models from natural language supervision. In Proceedings of the 38th International Conference on Machine Learning (pp. 8748–8763). PMLR. https://proceedings.mlr.press/v139/radford21a.html

Reimann, M., Klingbeil, M., Pasewaldt, S., Semmo, A., Trapp, M., & Döllner, J. (2018). MaeSTrO: A mobile app for style transfer orchestration using neural networks. In 2018 International Conference on Cyberworlds (CW) (pp. 9–16). IEEE. https://doi.org/10.1109/CW.2018.00012

Refsgaard, A. (2018). Eye conductor [Interactive installation]. Andreas Refsgaard Studio.

Romic, B. (2022). Negotiating anthropomorphism in the Ai-Da robot. International Journal of Social Robotics, 14(10), 2083–2093. https://doi.org/10.1007/s12369-022-00289-8

Sawyer, R. K. (1999). The emergence of creativity. Philosophical Psychology, 12(4), 447–469. https://doi.org/10.1080/095150899105656

Sola, M. C., & Guljajeva, V. (2024). Visions of destruction: Exploring a potential of generative AI in interactive art. arXiv. https://arxiv.org/abs/2408.14644

Stark, L., & Crawford, K. (2019). The work of art in the age of artificial intelligence: What artists can teach us about the ethics of data practice. Surveillance & Society, 17(3/4), 442–455. https://doi.org/10.24908/ss.v17i3/4.13338

Wilson, S. (2003). Information arts: Intersections of art, science, and technology. MIT Press.



HOW TO CITE THIS ARTICLE

Anduz,M. , Jafari,H. , Zandieh,M. and Eidy,S. (2025). Deep Learning-Based Interactive Art: A Data-Driven Approach to Adaptive Aesthetic Experience. International Journal of Urban Management and Energy Sustainability, 6(3), 256-274.

DOI:10.22034/IJUMES.2025.735867