"REVOLUTIONIZING HUMAN-COMPUTER INTERACTION: USING THE POWER OF MACHINE LEARNING AND DEEP LEARNING"

Research Paper

Amir Arshia Janighorban, Mohajer Technical University of Isfahan, Isfahan, Iran, amirarshiajanighorban@gmail.com

"Abstract"

The multidisciplinary discipline of human-computer interaction (HCI) is concerned with developing, assessing, and enhancing human-computer interaction. It includes investigating user interface design to develop user-friendly and functional technology. To create user-friendly interfaces and improve the user experience, it is necessary to understand human capabilities, limitations, and preferences. HCI takes into account things like usability, accessibility, ergonomics, and cognitive processes. Human-Computer Interaction (HCI) aims to facilitate smooth communication and cooperation between humans and computers, enabling users to carry out tasks successfully and efficiently. It entails creating user-friendly, responsive, and intuitive interface designs. Input and output devices, graphical user interfaces, voice and gesture recognition, virtual and augmented reality, mobile and wearable technologies, and more are all in the broad human-computer interaction (HCI) category.
Keywords: ML, DL, HCI.

1 Introduction

With the advent of machine learning (ML) and deep learning (DL), a new era of human-computer interaction (HCI) has emerged, revolutionizing the way humans interact with computers. ML and DL have made it possible for computers to interpret and respond to human input more intuitively and intelligently, making technology an extension of ourselves. In this study, we explore the transformative impact of ML and DL on HCI and delve into some remarkable applications that have changed the digital landscape. Human-computer interaction has come a long way since the days of punch cards and command-line interfaces. The overview of Human-Computer Interaction (HCI) history offers a vital context for comprehending the revolutionary influence of Machine Learning (ML) on this area. HCI originated in the early days of computers when most interactions were command-driven and text-based. These early interfaces had accessibility and user-friendliness issues. With the development of ML and AI, HCI has seen a notable progression toward more intuitive and user-centric design. ML algorithms have substantially impacted HCI as they have evolved from command-line interfaces to touchscreens, voice-activated systems, graphical user interfaces (GUIs), and more. Recognition is primarily reliant on tracking and detection accuracy. Not to be overlooked is the difficulty of making the entire system user-friendly. Resolving the other issues described above will largely determine how to handle this. With the help of HCI data, we can design more intelligent systems and learn more efficiently. One significant area of artificial intelligence is machine learning. It has shown considerable promise for research and development (R&D) and has advanced greatly in several disciplines. Machine intelligence is increasing thanks to machine learning technology in HCI. This Special Issue aims to include new research and review articles that address the most recent advancements in machine learning-based human-machine interaction.

2 Natural Language Processing (NLP) and Voice Recognition

Like many other AI issues, automatic voice recognition can be achieved by compiling a sizable pool of labeled data, training a model, and then using the trained model to classify fresh data reliably. The twist is that speech has many variations and is time-structured.(Bae and Song, 2017) We'll discuss our difficulties when translating spoken words and sentences into written language. We'll go deeper into the sound signal and different speech models to see how these difficulties might be overcome. Our information is the sound signal. We will discuss phonetics, signal analysis, and feature extraction for voice data representation. Voice recognition and natural language processing (NLP) aim to enable computers to understand and process spoken and written human language. As a result of their ability to facilitate enhanced understanding and natural interaction between humans and machines, they are essential in the evolution of human-computer interaction. Let's examine each of these areas in more detail. NLP, as a subset of AI, studies how computers process language. It entails creating models and algorithms that let computers comprehend, interpret, and produce natural language. NLP includes various tasks, such as:

2.1 Text understanding

NLP algorithms examine textual material and derive meaning from it. This includes parsing, named entity recognition, part-of-speech tagging, and semantic analysis, among other activities. It makes it possible for computers to understand documents' semantics and sentence structure.

2.2 Sentiment analysis

Thanks to NLP approaches, computers can recognize and comprehend any sentiment represented in text, whether neutral, positive, or negative. Applications for sentiment analysis include brand reputation management, customer feedback analysis, and social media monitoring.

2.3 Language translation

Machine translation programs, such as Google Translate, which translate text mechanically across languages, are powered by natural language processing (NLP). These systems interpret texts in one language and provide equivalent translations using statistical models or neural networks.

2.4 Question answering

Computers can understand inquiries and deliver pertinent replies thanks to NLP algorithms. This covers duties including retrieving information, summarizing documents, and comprehending questions. Search engines, chatbots, and virtual assistants all use question-answering systems. Since voice input is frequently transformed into text for additional processing using NLP techniques, voice recognition and NLP are closely related fields. Combined, they allow for more conversational and organic interactions between people and computers, increasing accessibility and user-friendliness of technology. Human-computer interaction has been revolutionized by NLP and voice recognition technologies, which allow computers to comprehend and process spoken and written human language.(Lomte *et al.*, 2012) Voice assistants, sentiment analysis tools, language translation systems, and other innovations have been made possible by these breakthroughs, which have also increased the usability and accessibility of technology for a larger group of people.

3 Data Visualization

The graphical depiction of data and information using visual components, including charts, graphs, maps, and diagrams, is known as data visualization. It entails turning unprocessed data into visually appealing representations that are simpler to comprehend, evaluate, and share. Presenting intricate data linkages, patterns, and trends so that users can easily and rapidly understand visually is the main objective of data visualization. The following are some essential elements and concepts of data visualization.

3.1 Data representation

The first step in data visualization is to choose and represent the pertinent data. This entails determining which measurements, dimensions, and variables require visualization. The type of data can be numerical, category, chronological, or spatial, and the features of the data will determine which visual representations are most appropriate.

3.2 Visual encoding

Mapping the data qualities to visual properties, including location, length, angle, color, form, and size, is known as visual encoding. For instance, a scatter plot employs the position of dots to show the relationship between two variables, but a bar chart uses the length of the bars to express quantities. Selecting the right visual encodings is essential to effectively displaying the information and communicating the desired meaning.

3.3 Visual perception

Data visualization uses the ability of human vision to make information more accessible to interpret. It uses pre-attentive characteristics that are easily and immediately noticed, like color, size, and position. Comprehending the fundamentals of visual perception enables designers to produce visualizations that highlight significant trends, enable comparisons, and support the interpretation of data.

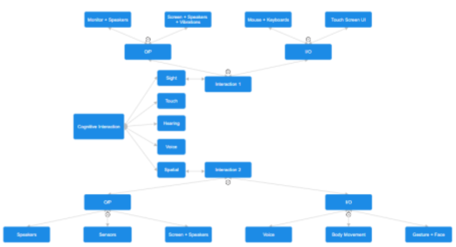

Figure 1. System analysis by HCI proposed method

3.4 Data abstraction

Large or complicated data sets are frequently summarized and aggregated for data visualization to provide a clear and succinct depiction. Averaging, grouping, and filtering are aggregation procedures that minimize complexity while maintaining important insights and patterns. When determining the level of abstraction, a balance between giving a thorough overview and preserving the essential details should be struck.

3.5 Interactive exploration

Interactive components improve data visualization by enabling users to explore and engage with visual representations. Users can examine specific data points, navigate between various visualizations, and delve into specifics using interactive capabilities like filtering, sorting, zooming, and connecting. Incorporating interactivity into data allows people to gain more insight and increases engagement.

4 Sliding Window Mechanism

The sliding window mechanism is used in computer science, data analysis, and signal processing to handle and examine a portion of a more extensive dataset. With this method, operations are carried out on only the part of the data inside the "window" at a time by moving a fixed-size "window" over the data. Its general operation and applications are as follows.

7	6	5	4	2	2	5	2
1	2	3	3	4	1	2	3
7	6	5	4	2	2	5	2
2	3	4	5	5	7	2	6
2	3	2	1	2	8	7	6
3	3	1	3	2	6	8	9
7	6	5	4	2	2	5	2

Figure 2. Slide from left to right

7	6	5	4	2	2	5	2
1	2	3	3	4	1	2	3
7	6	5	4	2	2	5	2
2	3	4	5	5	7	2	6
2	3	2	1	2	8	7	6
3	3	1	3	2	6	8	9
7	6	5	4	2	2	5	2

Figure 3. Slide from right to left

4.1 Left to right and right to left

The direction in which the window glides over the data set is referred to as "left to right" when discussing a sliding window technique. This method is frequently used in applications where data is handled sequentially from start to finish. This is how it usually works.

- Initialization: Up to the specified window size, the window is initially positioned at the leftmost portion of the data set, encompassing the first segment of data points.(Welch, 2023)
- Sliding: After that, the window travels one step at a time to the right. Every time, a new data point from the right is added to the window, and the leftmost data point is removed.
- Continuous Operation: The data within the window is subjected to operations at every window location. These tasks could involve computations such as sums and averages or using particular algorithms for filtering or pattern recognition.

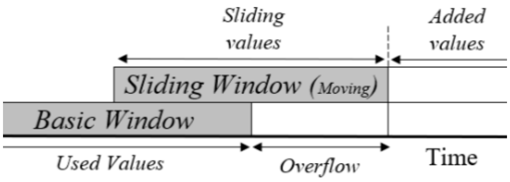

Figure 4. System analysis by HCI proposed method

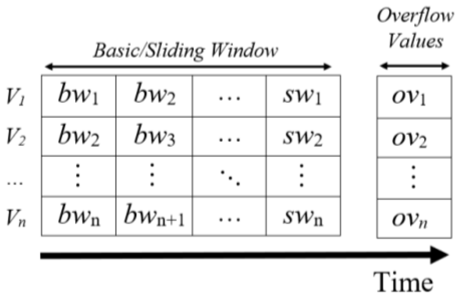

Figure 5. sliding window used in digital signal processing

- Text Analysis: A sliding window may be used in natural language processing to examine phrases or patterns in the context of surrounding words to move from left to right across a text.(Mistry, Maes and Chang, 2009)
  Time Series Analysis: To calculate moving averages or other statistical metrics to spot trends in financial or meteorological data, the window scrolls over time-stamped data items.
  Image processing: To apply detection algorithms in various sections of an image, a sliding window travels from left to right (and frequently top to bottom in a nested way) in tasks like object detection.

- Benefits of Sliding from Left to Right Natural Ordering: Processing data from left to right maintains the data's inherent chronological order for many forms, particularly time-series data. Real-time Processing: When new data is continuously entering from the right, this orientation facilitates real-time, or streaming, data processing. development Ease: Because it conforms to the standard architecture of data storage and retrieval systems intended to handle data sequentially, it frequently simplifies algorithm development.

- Things to Think About Edge Cases: Extra caution is required at the data's boundaries, especially at the beginning and end. The window might not initially fill unless padding or other handling is used. Backtracking: Strictly going from left to right may make it more difficult to include such input without additional methods if the analysis requires reviewing prior data points. (Mistry, no date)

In conclusion, repositioning the sliding window from left to right is a simple and efficient way to handle data in various fields. This method also fits well with the structure of most data handling systems and the natural flow of many different data types.

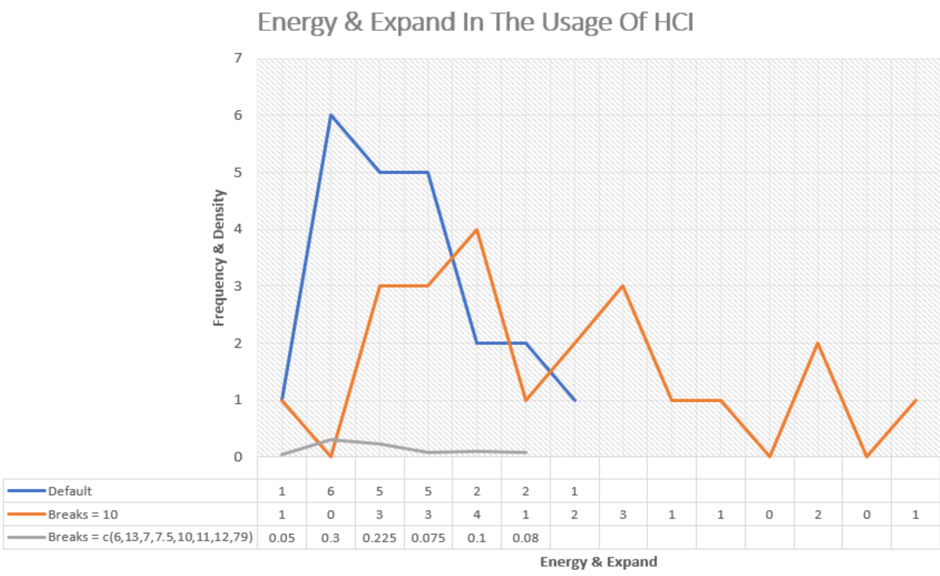

Figure 6. Narratives bar chart regressions

4.2 Storytelling and narratives

Storytelling and tales are frequently used in effective data visualization to make a point clearly and direct the user's understanding.(Wikipedia, no date) A well-designed visualization may lead the user through the data, highlighting significant trends and producing insightful conclusions using a logical flow and annotations.

5 Feature Learning

In HCI (Human-Computer Interaction), feature learning is the process of automatically recognizing and deriving pertinent and meaningful features from data gathered during user interactions with technology. Textual data, sensor data, behavioral data, and other input types might all fall under this category.(Baldauf and Frihlich, no date)

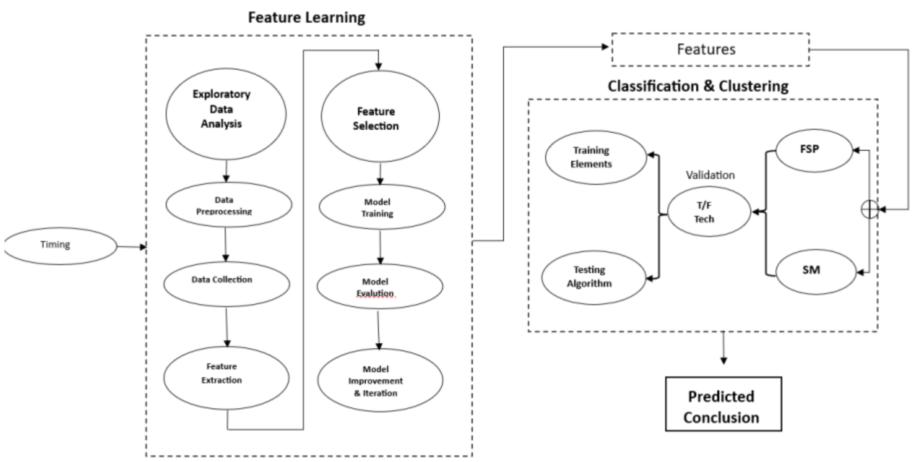

Figure 7. Classification of feature learning

Feature learning approaches in HCI can improve user experience by providing more context-aware and tailored interactions. Machine learning algorithms may adjust to user preferences, behavior, and

demands by identifying patterns in user data, improving interaction efficiency and effectiveness.
Feature learning techniques that are frequently used in HCI include:

- Algorithms for clustering and classifying data: These strategies combine comparable user actions or input patterns to help the system pick up on and adjust to user preferences.
- Dimensionality reduction: Techniques for decreasing data complexity can help identify the key characteristics that impact users' engagement with technology.
- Neural networks: The system can produce more precise predictions and suggestions thanks to deep learning neural networks' ability to extract complicated properties from data automatically.(Betke, Gips and Fleming, 2022)

All things considered, feature learning in HCI is essential to developing intelligent and flexible user interfaces that can predict user demands and offer tailored experiences.
Classification is the process of classifying objects or data into groups according to specific characteristics. Its primary objective is to arrange and streamline complicated data so that it is simpler to comprehend, evaluate, and use for decision-making.
Classification in machine learning is a form of supervised learning in which the algorithm predicts the class or category of new, unseen data by learning from labeled training data.(Hürst and Wezel, 2013)
Numerous categorization techniques exist, each with unique advantages and disadvantages, including decision trees, support vector machines, logistic regression, and neural networks.
Numerous industries use classification extensively, including business, marketing, finance, healthcare, and image identification. It is a vital tool for jobs like spam identification, sentiment analysis, customer segmentation, fraud detection, and medical diagnosis.
All things considered, categorization is an essential component of data organization and interpretation, making it useful for problem-solving and decision-making across a wide range of industries.

6 The Landscape

The interdisciplinary topic of Human-Computer Interaction (HCI) in Machine Learning (ML) and Deep Learning (DL) is known as the "brain-shape landscape." It centers on how people interact with systems driven by these algorithms. This entails researching how these systems are made to fit humans' cognitive capacities better and comprehending how people behave when engaging with them.(Wang, 2008)
In this context, researchers aim to develop user-friendly, intuitive interfaces that are sensitive to people's needs and preferences. This entails building interfaces that can adjust to users' different skill levels and cognitive capacities and including functionalities that facilitate human decision-making and problem-solving processes.
Furthermore, the brain-shape landscape of HCI in ML and DL examines the ethical ramifications of utilizing these technologies in human-computer interactions.(Chawalitsittikul and Suvonvorn, 2020) This covers factors including algorithmic bias, data privacy, and the effects of automation on human work and decision-making.
Finally, the brain-shaped landscape of HCI in ML and DL aims to optimize the synergy between human cognition and machine intelligence. This will lead to the creation of more effective, practical, and morally good systems for human-computer interaction.

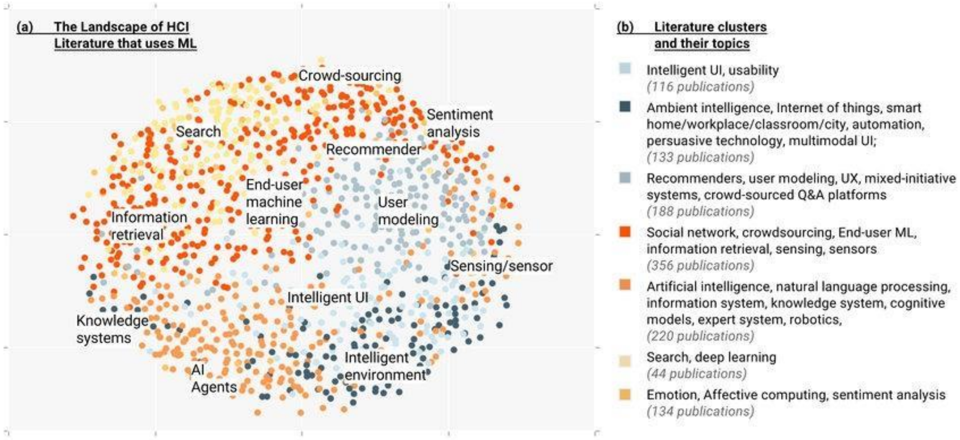

Figure 8. Landscape of HCI in ML and DL

7 Custom Fully Tested Hand Gestures

In human-computer interaction (HCI), "custom hand gestures" refers to unique hand gestures or movements a user makes to communicate with a computer or other device. More naturally and intuitively, digital material or interfaces can be controlled, navigated, or manipulated with these motions.(Wang and Popović, 2019) Custom hand motions used in HCI include the following examples:

- Pinching: To zoom in or out on a map or image, press and hold down two fingers on a touchscreen.
- Swiping: Use your finger to move across a touchscreen to navigate between pages or scroll up and down on a webpage.
- Rotating: To rotate an image or object on a touchscreen, twist two fingers.
- Tapping: Selecting or activating an item on a screen by quickly touching it with the finger.
- Three-finger swipe: Swipe three fingers over a trackpad to move between open programs or desktop areas.(Sekora and Riege, 2008)
- Peace sign: Placing two fingers together to initiate a certain action, such as launching a new tab or snapping a screenshot.
- Fist bump: When you make a payment or confirm a selection, you make a fist and tap the surface.

These personalized hand gestures can change depending on the device or application being used. Users can also program or alter them to fit their own needs and preferences. By including these gestures in interface and interaction designs, designers may improve user experience and make interactions more effective and interesting.(Yu and Leeser, 2021)

8 Gesture Recognition (Software & Hardware)

With gesture recognition technology, computers can now recognize and process human movements and gestures as input or orders. They record and analyze hand gestures, body postures, facial expressions, and other actions in real time using sensors, cameras, or other devices.(Manresa *et al.*, 2005)
Vision-based and sensor-based gesture recognition are the two primary categories. Whereas sensor-based gesture recognition employs accelerometers, motion sensors, or gloves to detect motions, vision-based gesture recognition uses cameras to record gestures. Interactive systems like video games, virtual reality apps, smart TVs, and smartphones frequently incorporate gesture recognition.(Plataniotis and Venetsanopoulos, 2015) It can also be applied to robots, security systems, and healthcare. Gesture recognition is frequently used to control devices without requiring physical contact, interact with virtual objects in three dimensions, and create more intuitive and natural user interfaces. Nevertheless, there

are certain drawbacks to gesture recognition as well, such as making sure it is reliable and accurate in a variety of settings. It can identify a large variety of movements.

Service	Markets Served	Services Provided
Tableau Software	Tableau e Colleges and Universities e Manufacturing Analytics e Financial Services and Insurance	e Student performance management e Enrollment movement e Donor Relations and Federal Funding e Improve process efficiency, centralize production
Detica	e Insurance and healthcare solutions e Banking market	e Credit scoring e Fraud solutions
Fair Issac Corporation	FICO® Score e Banking and financial	e Credit scoring e Fraud solutions
ViPS	e Insurance and healthcare solutions	e Sarah Lawrence College
KXEN	InfiniteInsight™ e Customer lifecycle	e Help companies to optimize every step in customer acquisition, cross-sell, up-sell, retention and next best
IBM	IBM Business Analytics e Fraud solutions e Healthcare market	e Delivers information that decision-makers trust to improve business performance e Provides clear, immediate and actionable insights into
SAS Institute Inc.	Visual Analytics e Fraud solutions e Insurance & healthcare solutions e property & casualty	e Interactively visualize, explore and communicate data discoveries e Transform structured and unstructured data into fact- based insights for better decisions
Microsoft	Visual Analytics	e Sinclair Community College
TIBCO	College Scheduler	e University of Arizona
SAP AG	e Customer lifecycle e Campaigns and promotions optimization e Sales and purchasing planning	e Customer acquisition, cross-sell, up-sell, retention and next best activity e More targeted and effective promotions and campaigns

Table 1. Prostheses Hand Gesture Recognition Using Surface Electromyography and Deep Learning

9 Linear Analytics

Analyzing data using a linear approach by looking at the relationships between various variables is known as linear analytics. To put it another way, the goal is to understand the relationship between changes in one variable and changes in another, assuming a linear relationship between the two.(Gonzale and Woods, 2022) Techniques like linear regression, a statistical approach to describe the connection between a dependent variable and one or more independent variables, can be utilized.

- In many disciplines, including business, finance, economics, and social sciences, linear analytics is frequently used to forecast trends, find patterns in data, and make predictions. Researchers and analysts can obtain critical insights into the underlying structure of the data and make defensible conclusions by examining the linear correlations between variables.(Suandi *et al.*, 2010)
- Linear analytics' main advantages include its simplicity and interpretability. It offers an easy-to-understand method for analyzing data and informing stakeholders of the findings. Furthermore, linear analytics can find outliers and abnormalities in the data, which can improve the analysis's accuracy and dependability.
- Overall, linear analytics is a potent instrument for deciphering the correlations between variables and drawing insightful conclusions from data, making it a crucial component of many professionals' and businesses' analytical toolboxes.(Derpanis, no date)

Possible Errors in Front of Velocity And The Detection of Them Using Analytics

frame no	**experimental X**	**experimental Y**	**Real X**	**Real Y**	**Error(pixel)**	**σ of error**
1	476	13	469	46	33.73	
2	470	25	462	65	40.79	
3	261	416	231	436	36.05	8.27
4	224	451	194	449	30.06	
5	295	310	320	350	32.02	

Table 2. Error rate between experimental detection and actual position

frame no	**without boundary box(s)**	**with boundary box(s)**
1	0.124	0.083
2	0.097	0.096
3	01	0.11
4	0.095	0.08
5	0.117	0.082
6	0.089	0.085
7	0.1	0.087
8	0.127	0.081
9	0.08	0.09
10	0.1	0.086
Avg	0.1029	0.0884

Table 3. Time required to detect after losing

10 User-Friendliness

Five users were invited to utilize our system and carry out the cursor movement, left and right clicks, zoom in and out, forward and backward, and zoom in and out to assess user-friendliness. Every user was instructed to carry out each command thirty times. Table II displays the overall accuracy in the first ten attempts, 59.38%. The average correctness of each command, which appears at the bottom of the table, is another feature of Table II. Additionally, we can see that for users 1, 2, 3, 4, and 5, the average

accuracy for all commands is 64.2%, 61.4%, 55.7%, 62.8%, and 52.8%, respectively. Tables I and II also show comparables.
The data presented in Table II pertains to the subsequent 10 attempts and the subsequent 10 attempts, respectively. Tables II, I, and II show that the average performance of each command increases as the user tries the same command more than once; after 30 attempts, the average accuracy is 79.4% overall. The system will perform better as the user gets more used to it.(Nadir, 2019)

11 Conclusion

The integration of machine learning (ML) and deep learning (DL) into human-computer interaction (HCI) has significantly transformed the way we interact with technology. These advancements have led to more intuitive and adaptive interfaces, improving the naturalness of communication through voice, gesture, and text recognition. Enhanced data visualization and feature learning techniques have made complex data more accessible and actionable, while custom gestures and gesture recognition technology have introduced more natural ways to interact with digital environments.

As HCI evolves, user-friendliness remains a key focus, with continuous improvements driven by user feedback and engagement. The convergence of ML, DL, and HCI promises to create more effective, user-centric interfaces, ensuring that technology remains accessible and aligned with human needs.

References

Bae, J.S. and Song, T.L. (2017) ‘Image tracking algorithm using template matching and psnf-m’, *International Journal of Control, Automation, and Systems*, 6(3), pp. 413–423.

Baldauf, M. and Frihlich, P. (no date) ‘Supporting hand gesture manipulation of projected content with mobile phones.’

Betke, M., Gips, J. and Fleming, P. (2022) ‘The camera mouse: Visual tracking of body’, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 10(1).

Chawalitsittikul, P. and Suvonvorn, N. (2020) ‘Real time hand marker tracking as a user input device’, in *Proceedings of ECTI Conference on Application Research and Development (ECTI-CARD)*. Chonburi, Thailand, pp. 427–431.

Derpanis, K.G. (no date) *Relationship between the sum of squared difference (ssd) and cross correlation for template matching*. Torronto.

Hürst, W. and Wezel, C. V. (2013) ‘Gesture-based interaction via finger tracking for mobile augmented reality’, *Multimedia Tools and Applications*, 62(1), pp. 233–258.

Lomte, V.M. *et al.* (2012) ‘Robust barcode recognition using template matching’, *International Journal of Engineering Sciences and Emerging Technologies*, 2(2), pp. 56–65.

Manresa, C. *et al.* (2005) ‘Real time hand tracking and gesture recognition for human-computer interaction’.

Mistry, P. (no date) *Sixthsense integrating information with the real world, http://www.pranavmistry.com/projects/sixthsense/.*

Mistry, P., Maes, P. and Chang, L. (2009) ‘Wuw - wear ur world: a wearable gestural interface’, in *Proceedings of the 27th international Conference Extended Abstracts on Human Factors in Computing Systems*. Boston, MA, USA.

Nadir, N.D. (2019) ‘Fast template matching method based optimized sum of absolute difference algorithm for face localization’, *International Journal of Computer Applications (0975 – 8887)* , 18(8).

Sekora, D. and Riege, K. (2008) ‘Real-time color ball tracking for augmented reality’, in *Proceedings of the 14th Eurographics Symposium on Virtual Environments*. Eindhoven, The Netherlands, pp. 9–16.

Suandi, S.A. *et al.* (2010) ‘ A study on similarity computations in template matching technique for identity verification’, *(IJCSE) International Journal on Computer Science and Engineering*, 2(8), pp. 2659–2665.

Wang, R.Y. (2008) ‘Real time hand tracking as a user input device’, in *Proceedings of the 21st annual ACM symposium on User interface software and technology*. Monterey, CA: ACM.

Wang, R.Y. and Popović, J. (2019) ‘Real-time hand-tracking with a color glove’, in H. H. (ed.) *ACM SIGGRAPH 2009 Papers*. New Orleans, Louisiana: ACM, pp. 1–8.

Welch, G. (2023) *An introduction to the kalman filter*. Available at: http://www.cs.unc.edu/~welch/kalman/ (Accessed: 1 August 2023).

Wikipedia (no date) *Kinnect*. Available at: http://en.wikipedia.org/wiki/Kinnect (Accessed: 2 August 2022).

Yu, H. and Leeser, M. (2021) ‘Optimizing data intensive window-based image processing on reconfigurable hardware boards’, in *IEEE Workshop on Signal Processing Systems Design and Implementation*.