2022 IEEE World Conference on Applied Intelligence and Computing (AIC)

17 - 19 June 2022

Internet of Things, Information Security, Embedded Systems, Real-time Systems, Cloud Computing, Big Data Analysis, Quantum Computing, Automation Systems, Intelligent IoT eHealth, Bio-Inspired Intelligence, Brain Modeling and Simulation, Cognitive Systems, Cyber-Physical Systems, Data Analytics, Data/Web Mining, Data Science, Hybrid Systems, Intelligence for Security, Intelligent Decision Making Systems, Intelligent Information Processing, Intelligent Transportation Machine Vision Artificial Intelligence for Machine Vision, Imaging Sensors Technology, Features Extraction, Color and Texture Analysis, Image Segmentation, Convolutional Neural Network, Biometrics Recognition, Biomedical Imaging, Image/Video Classification, Image Restoration, Soft Computing for Machine Vision, Pattern Recognition, Gesture Recognition, Action Recognition, Intelligent Transport Systems, Surveillance, Human-Computer Interaction

Event Format: Virtual

Sponsors: Institute of Electronics & Telecommunication Engineers (India); Madhya Pradesh Section; Soft Computing Research Society

Call for Papers Deadline: 15 April 2022

A Study on Human-Computer Interaction based on Surveillance Tasks

Ruchi Jayaswal
Department of CSE
MITS
Gwalior, M.P., India
ruchi.jayaswal23@gmail.com

Manish Dixit
Department of CSE
MITS
Gwalior, M.P., India
dixitmits@mitsgwalior.in

2022 IEEE World Conference on Applied Intelligence and Computing (AIC) | 978-1-6654-7988-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/AIC55036.2022.9848839

Abstract— **The article presents a brief overview of the field of Human Computer Interaction (HCI). It discusses the multidisciplinary field of HCI and attempts to explain that the computer vision applications such as HCI, virtual reality, security, video surveillance, people monitoring is highly correlated to intelligent human video surveillance tasks. The roots and origin of HCI will be explained and how it evolved with the generations. Subsequently, the fidelity prototyping of HCI will be discussed, then later examines the human computer interaction related with intelligent video surveillance system through vision and movement mode generates the output according to the application requirement. The study aims to present an analysis of the surveillance application framework which is solved by the amalgamation of computer vision and HCI**.

Keywords— Human-Computer Interaction, Surveillance system, Prototyping, Computer vision, Cognitive Science

I. INTRODUCTION

Human-Computer Interaction (HCI) is a branch of computer science that studies the interaction of people (users) and computers, as well as the design, assessment, and development of user interfaces for computer systems that are responsive to the user's requirements and interests [1,2]. It is an interdisciplinary field including computer science, behavioural sciences, and design, as shown in figure 1. One of the primary goals of HCI is to make computer systems more user-friendly and useful. HCI is the study of the interaction between human (users) and computers. The interaction between users and computers is achieved via an interface-user interface. Users interact with computer systems via a user interface, which is made up of hard- and software components that offer input, enabling users to operate the system, and output, allowing the system to convey information to the user. HCI puts a focus on the design, implementation, and assessment of interfaces.

HCI may be applied in any discipline where computer installation is possible. Some of the areas where HCI may be applied with particular significance are listed below.

- Computer science technology is used for application design and engineering.
- Psychology is used for theory application and analysis.
- Sociology is concerned with the relationship between technology and organizations.
- Industrial design is taken for interactive items such as mobiles, microwave ovens and so on.

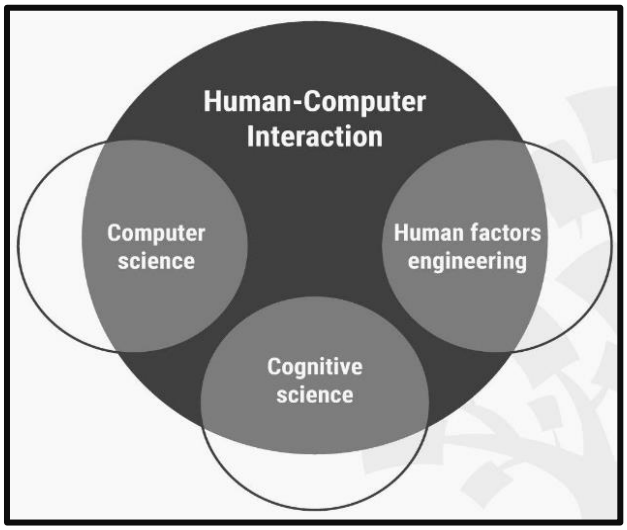

Figure 1. The Multidisciplinary field of HCI[1]

HCI quickly became the focus of significant academic research [3]. Those who studied and worked in HCI considered it as a critical tool for popularizing the notion that a computer's connection with a user should resemble a human-to-human, open-ended discussion. Initially, HCI researchers worked on improving the usability of desktop computers; that is, learners, concentrated on the usability of computers). Nevertheless, as Internet and smartphone technologies become more common, computer use will transition away from the pc and toward the mobile context. In addition, HCI has grown to cover new domains such as human action recognition systems, gesture observing systems, smart vision system, and many more. Various problems can be solved by amalgamating these three technologies, i.e., HCI, Computer Vision, and Deep learning. Some applications of HCI using computer vision include surveillance, face recognition, handwritten recognition, speech recognition, health monitoring, smart city etc.

With the fast technical development of computer devices and their increasing prevalence within telecommunications in recent years, the area has extended to include social and organisational academics who investigate

978-1-6654-7988-2/22/$31.00 ©2022 IEEE
DOI: 10.1109/aic.2022.64

the influence of the technologies on human-human interaction. This, in turn, leads to the discovery of new applications and the creation of new devices. Today HCI is a proactive, multidimensional, and methodologically diversified component of the subject under study, following and guiding technological advancement. To design an HCI system, three important 'USE' words can be used to make an interactive system.

1. Useful: It should be functional and does things
2. Usable: Make more friendly, easy to do things, does the right things, user experience
3. Used: It should be attractive, available and acceptable to organization.

There are several branches of HCI that are particularly important to it. The sub-areas of HCI are still in under research progress. Many applications of HCI are needed to work on them using computer vision technologies, cognitive methods and others. Figure 2 shows some areas of human computer interaction that make the computer more understandable for users.

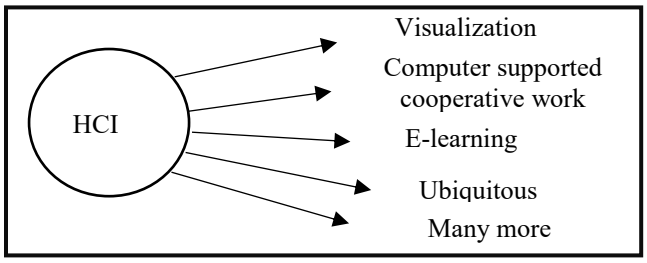

Figure 2. *Branches* of HCI[2]

The paper's main contribution is towards the overview of HCI and the usage of HCI in computer vision-based surveillance tasks. Moreover, it reflects mainly on the evolution of HCI, the process of HCI and how HCI is applicable on video surveillance applications.

The organization of the paper is described in four sections. Section I presents the introduction part about the human-computer interaction. Section II describes the history of the HCI. The methodology and process of HCI design are described in section III. The human-computer interaction in surveillance tasks is demonstrated in section IV. At last, Section V discusses the conclusion part.

II. HISTORY OF HCI

2.1 The Evolution of HCI

Only information technology specialists and enthusiastic fans used computers in the early 1970s. This changed dramatically with the introduction of the personal computer in the late 1970s. Personal computers, which contains specific software (like word processors and spreadsheets) as well as personal computer platforms (like OS, programming languages, and equipment), decided to make everyone in the world a potentially affected pc user. On the other hand, highlights the limitations of computers in terms of convenience for those who seek to be using desktops as equipment.

Personal computing became a concern at an inopportune moment. By the end of the 1970s, cognitive science had progressed to include cognitive psychology, AI, linguistics, cognitive ethnography, and philosophies of cognition. The creation of methodical and statistically based applications known as "cognitive engineering" was an objective of cognitive science. Cognitive science provided persons, notions, abilities, and a perspective for meeting those needs in the same manner that personal computing exposed the practical need for HCI through a daring synthesis of science and engineering. HCI was among the first uses of cognitive engineering. Parallel improvements aided this in engineering and design disciplines adjacent to and frequently overlapping with HCI, most significantly user factors design and documentation development. Human factors had established empirical and task-analytic methodologies for assessing human-system interactions in areas like aircraft and manufacturing and were now concentrating on interactive systems where human operators had greater problem-solving discretion. Documentation development evolved from its traditional function of creating systematic technical descriptions to a cognitive approach that combined composition, readability, and media theories with user testing. Documents and other items that can be used are also necessary.

Numerous significant historical events also contributed to HCI creation [3]. When software development became engrossed in uncontrollable software intricacy (the "software crisis") in the 1970s, It began to place a greater emphasis on non-functional needs such as ease of use and maintenance and experimental software design processes based on iterative prototypes and experimental verification. Computer animation and information retrieval first appeared in the 1970s, and it was evident that communication systems were crucial to building on early accomplishments.

HCI's primary technological focus has always been and will continue to be usability. The phrase "easy to learn, easy to use" was the first to represent this idea in a straightforward fashion. The raw ease of this notion gives HCI an edgy and forceful character in computing. It united the discipline and allowed it to have a more significant and more efficient effect on computing scientific and technological development.

There were various breakthroughs along the way, from the first batch processing machines to the user-centric design listed here[1,2].

Table 1. The Historical Evolution of HCI

Year	Name	Description
1946	Early computer (ENIAC)	Improvement in computing power by expanding the hardware technology
1950	Visual display unit	In the earlier version of VDU, a SAGE and air defence system of the USA are used.
1962	Development of the sketchpad	Sketchpad was created by Ivan Sutherland, that showed computers can be used for more than data processing.
1963	Programming toolkits	It is developed from smaller to larger systems and components.
1968	Word processor, mouse	Developed for online system.
1970	Personal computer Dynabook	Designed Smalltalk at Xerox PARC
1973	Windows and WIMP interfaces	Concurrent jobs on one workstation, moving between work and displays, and sequential interaction are all aspects of consecutive interaction.
1980	Metaphor	The first systems to incorporate metaphors were the Xerox Star and Alto, which resulted in interface spontaneity.
1982	Direct Manipulation	It was first employed in Apple's Macintosh computer in 1984, and it lowered the risks of syntactic mistakes.
1945 and 1980	Hypertext	To indicate a text's non-linear structure then multimodality introduced.
1988	Ubiquitous Computing	The most active research field in HCI at the moment. Pervasive computing is a type of sensor-based/context-aware computing.
1989	WWW	First graphical Mosaic browser is introduced
1990	Computer Supported Cooperative Work	Mediated communication

III. PROCESS OF H.C.I

A) The methodology of Human Computer Interaction design

The interaction between a person and a computer to perform various tasks and deliver updates to the user depends on the nature of information fed into the system by the human is critical to the development of human-computer interaction[3][4]. The interaction between machine and user is depicted in Figure 3.

The adoption of four techniques recommended by Ebert in building user-friendly and easy-to-use computer interfaces is part of the work of designing the HCI. The user interfaces should be logical and simple to use to get the intended results. The following are the methodologies used in creating computer user interfaces.

1. Anthropomorphic Technique

When the HCI is designed with the utmost care, it can deliver output comparable to that of a human.

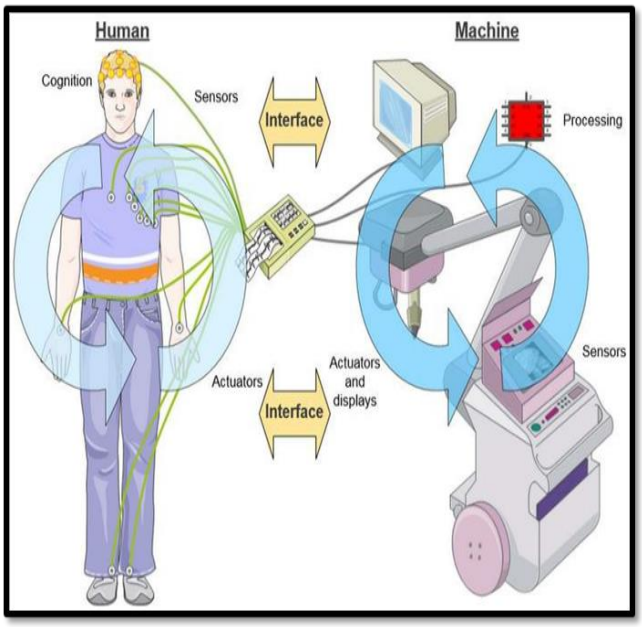

Figure 3. The interaction between computer and human[10]

2. Predictive Modeling

This assesses and analyses the user's experiences in connection to the time it will take them to create and finish a job on the computer successfully and efficiently. The usage of Goals, Operators, and Methods, as well as the section rules technique, allows the user to assess the system interface's suitability by analyzing the human's performance in determining the amount of time they may take to complete various activities on the computer interface.

3. Empirical Method

An empirical technique is used to examine the design's efficiency by assessing and evaluating the interface's outputs to compare the designs' usability.

B) Fidelity Prototyping

It is described as the accuracy with which a product's output has been replicated and the extent to which it has been simplified. This concept is used in developing HCI designs to determine how good the interfaces are in reproducing the exact result as entered in a simple manner that the user can understand with little or no explanation or experience. There are two types of fidelity models in the world of human-computer interaction: low fidelity prototypes and high fidelity prototyping[3].

1. Low-Fidelity Prototyping

Prototyping is the most straightforward method of transforming product and concept designs into real and testable ideas, allowing humans to gather and evaluate data from the initial phases of the computer interface interaction. Users are encouraged to interact with interfaces when the tasks' analysis and formulation are simplified, reducing the time they may spend completing a job.

2. High-Fidelity Prototyping

More information on the product's functionality and the interface is included in prototyping, which is close to the final version. It is used to identify the primary concerns that may obstruct people's capacity to interface with computers and accomplish multiple roles efficiently due to the design.

IV. Human Computer Interaction in Image/Video Surveillance Tasks

The use of Human-Computer interface system for visual surveillance tasks[5-8] with the help of computer vision technology. Identifying and monitoring people or items in a scene and interpreting human postures for communication with amenities, devices, and systems in the digital house are among the goals for HCI. To develop a comprehensive picture of their surroundings and understand the parts of the scene, the system must discourse several perception tasks at many levels, such as representation or characterization, segmentation, analysis, and movement tracking.

The vision module must also be integrated into a worldwide system that works in a challenging situation by obtaining visuals from image collection equipment at video regularity and delivering findings to relatively high configuration systems. Later, these systems make real-time decisions and must fulfill various criteria, such as time limitations, high availability, resilience, high speed of processing, and reconfigurability.

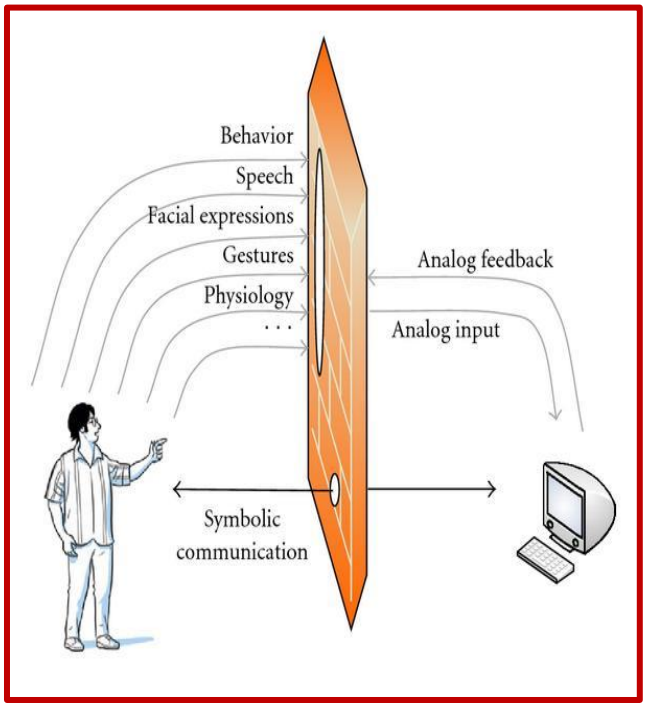

Figure 4. The communication way between machines and human in surveillance tasks[4].

Figure 4 presents the way of communication in between humans and machines. It stated that human gives input through their vision, sound, gesture, movement, behavior, expressions, psychology and more to the machine. Further, machines process the input as an analog signal and produce an output either in visual or audio-visual form.

Visual tracking is a hot topic in computer vision and HCI research[9]. Modeling the environment, motion detection, categorization of moving objects, tracking and recognizing activities are common phases in developing a visual surveillance process in dynamic settings. Most of the work focuses on applications that track the movement of human or vehicles and may be used for various purposes. Such as regulating access to restricted areas, identifying persons, network monitoring, anomalous identification and managing alarms, and interactive monitoring utilising numerous cameras[12,15].

A) Input/Output channels for HCI surveillance systems

- Information is received and provided in a person's interaction with the outside world i.e., input and output.
- In a machine interaction, the user gets data from the computer and reacts by supplying input to the computer — the human's output becomes the machine's input and vice versa.

• The senses supply the bulk of intake in humans, whereas the motor function of the effectors provides the most of the output.

The five major sensory inputs to the system are vision, behavior, hearing, touch, taste, and smell. The first three are the most important in terms of HCI, though we discuss the role of human vision output, which takes as an input for a surveillance system[10].

1. Role of Vision input for surveillance system

Even though vision is the average person's primary source of information, it is a difficult task with a variety of perceptual limitations.

The human eye is a complex organ. Light is the starting point for seeing. The eye is a device that receives light and converts it to electrical energy. The picture of items in the environment is focused upside down on the back of eye as light is reflected off them. The eye's receptors convert it into electrical impulses, which are then sent to the brain.

A variety of critical components make up the eye.

- The front of the eye's cornea and lens concentrate light into a crisp image on the retina, which is located in the rear of eye.
- Rods and cones are two kinds of photoreceptors found in the retina, which are both light-sensitive.
- Rods are very delicate to bright and let humans grasp in dim light.
- The rods have also inhibited the cones, so they don't work.
- Cones are the second type of receptor in the eye. They can endure lighter since they are less sensitive to light than rods.
- Cones are divided into three categories, each of which is subtle to a particular wavelength of light. This enables color perception.
- The eye has around 6 million cones, most of which are focused in the fovea, a tiny region of the retina where pictures are fixed.

The data obtained by the visual apparatus must be filtered and transmitted to processing units that allow us to recognize coherent scenes, disambiguate relative distances, and distinguish colors.

2. Role of human movement in surveillance tasks

Likewise, human vision, the movement of humans, is also an important input for surveillance applications such as human behavior recognition. In the design of interactive systems, speed and precision of movement are essential factors, particularly in regard to time it takes to travel to a specific object on a screen. The human movement is recorded by a device and gives input to a designed system for processing the gesture. The target might be a button, a menu item, or an icon [11,13].

Similarly, other sensors also play a crucial role for different surveillance tasks. A superordinated control loop of the technical system can be interpreted as human supervisory control of complicated systems or processes [10]. The essential identification and classification functions of images/videos in the surveillance system are based on numerous sensors, gestures, vision, sound, movement and many more[16,17]. A number of other factors (such as the occlusion, illumination effect etc.) must be taken into account. The design of ergonomic human-computer interaction is a significant aspect in assisting operators in surveillance activities, as well as in time-critical scenarios and at a computer workstation with a high informational workload.

The basic block diagram of the surveillance system which classifies the features into suspicious and non-suspicious patterns as shown in figure 5. Several steps are followed to develop an intelligent surveillance system[15]. The system takes input from a human sense in the image or video format then, and then preprocessing techniques are applied to fine the dataset. Further, feature extraction techniques are able to pertain to necessary features and then classify all the features through classifiers based on the requirement of applications.

3. Application areas based on Image/video surveillance systems

Numerous applications are present that are based on Image/video surveillance systems solved by the HCI. The most prominent areas are human abnormal behavior analysis[20][23], face recognition system[21][24-25], face mask detection[18], social distance monitoring[19], fall detection[27], shoplifting system[26] and many more. The block diagram of such applications is same. These applications are oriented with human involvement and take the input to the system then perform all the steps to obtain the results.

In pandemic times[22], the role of HCI is vital to develop such a system for all human beings. These are highly effective systems or devices currently used in pandemics such as Face mask detection, Observing social distancing, Digital thermometers, X-Ray machines that classify the report as infected or not, etc.

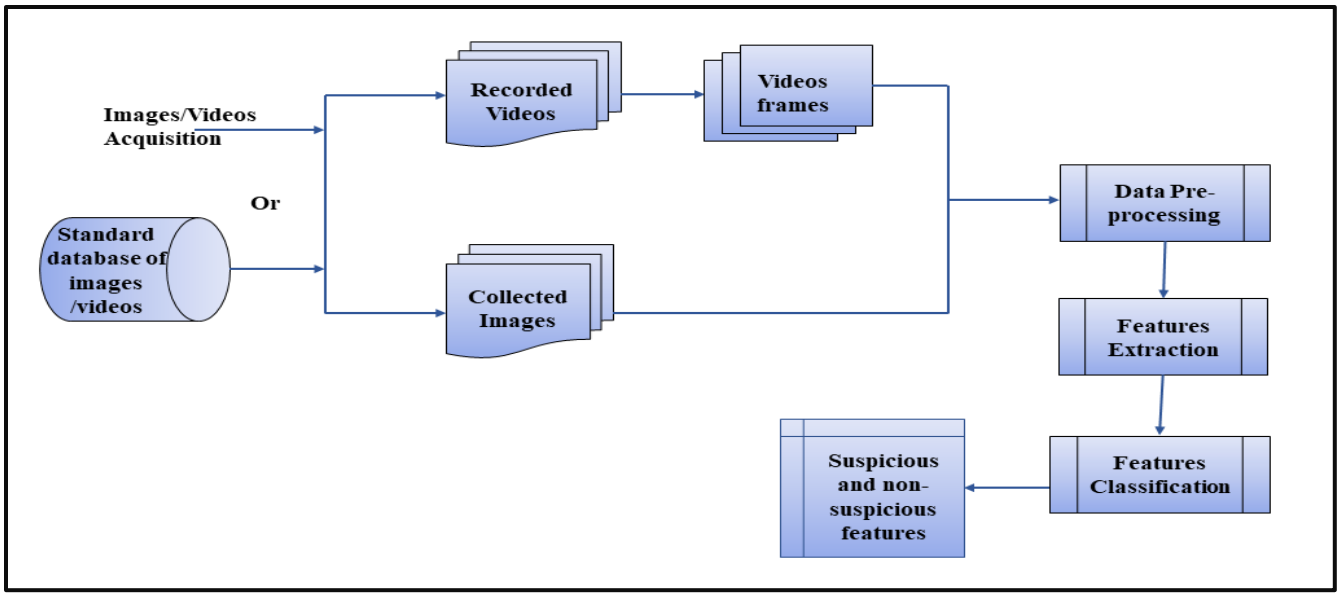

Figure 5. Framework of Human Computer interaction for surveillance tasks.

V. CONCLUSION

The article briefly overviews human-computer interaction in video/image surveillance tasks. The historical evolution of HCI is discussed in this study and covers the process or methodology of the HCI system. This work's main contribution is to discuss the surveillance tasks using human-computer interaction systems. It stated that the way of interaction between a machine and a human. There are many ways through which machines can take input from humans, process the model, and then later generate the output. In this work, vision and movement sensors are elaborated. At last, the surveillance applications are developed using HCI discussed. This paper described a model-based human-computer interaction design for surveillance tasks at public places such as organizations, schools, colleges, banks, public roads, stations, and more.

REFERENCES

[1] J. May, "Human–Computer Interaction," *International Encyclopedia of the Social & Behavioral Sciences*, pp. 7031–7035, Jan. 2001, doi: 10.1016/B0-08-043076-7/01422-4.

[2] Human Computer Interface - Quick Guide (tutorialspoint.com)

[3] Brey, Philip, and Johnny Hartz Søraker. "Philosophy of computing and information technology." In *Philosophy of technology and engineering sciences*, pp. 1341-1407. North-Holland, 2009.

[4] Schmidt, Ludger, and Daniel Ley. "Human-Computer Interaction in Aerial Surveillance Tasks." In *Industrial Engineering and Ergonomics*, pp. 511-521. Springer, Berlin, Heidelberg, 2009.

[5] Ibrahim, Sutrisno Warsono. "A comprehensive review on intelligent surveillance systems." *Communications in science and technology* 1, no. 1 (2016).

[6] X. Wang, Intelligent multi-camera video surveillance: A review, Pattern Recognition Letters, 34 (2013) 3-19

[7] Q. Cai and J.K. Aggarwal, Tracking human motion in structured environments using a distributed-camera system, IEEE Transactions on Pattern Analysis and MachineIntelligence21 (1999) 1241-1247

[8] R.T. Collins, A.J. Lipton, H. Fujiyoshi and T.Kanade, Algorithms for cooperative multisensory surveillance, Proceedings of the IEEE89 (2001) 1456-1477

[9] Dash, Samarendra Chandan Bindu, Soumya Ranjan Mishra, K. Srujan Raju, and L. V. Narasimha Prasad. "Human action recognition using a hybrid deep learning heuristic." *Soft Computing* 25, no. 20 (2021): 13079-13092.

[10] Srujan Kumar Palluri, Human Computer Interaction, Research Methods and Professional Issues, Bournemouth University.

[11] Chowdhury, Mozammel, Junbin Gao, and Rafiqul Islam. "Human surveillance system for security application." In *International Conference on Security and Privacy in Communication Systems*, pp. 711-724. Springer, Cham, 2015.

[12] A. H. Ahmad *et al.*, "Real time face recognition of video surveillance system using haar cascade classifier," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 21, no. 3, 2021, doi: 10.11591/ijeecs.v21.i3.pp1389-1399.

[13] G. Sreenu and M. A. Saleem Durai, "Intelligent video surveillance: a review through deep learning techniques for crowd analysis," *Journal of Big Data*, vol. 6, no. 1, pp. 1–27, 2019, doi: 10.1186/s40537-019-0212-5.

[14] Jayaswal, Ruchi, and Manish Dixit. "A Framework for Anomaly Classification Using Deep Transfer Learning Approach." *Revue d'Intelligence Artificielle* 35, no. 3 (2021): 255-263.

[15] Jayaswal, Ruchi, and Manish Dixit. "Detection of Hidden Facial Surface Masking in Stored and Real Time Captured Images: A Deep Learning Perspective in Covid Time." *Traitement du Signal* 38, no. 6 (2021).

[16] Ansari, Mohd Aquib, and Dushyant Kumar Singh. "An approach for human machine interaction using dynamic hand gesture recognition." In *2019 IEEE Conference on Information and Communication Technology*, pp. 1-6. IEEE, 2019.

[17] Ansari, Mohd, and Dushyant Kumar Singh. "Monitoring social distancing through human detection for preventing/reducing COVID spread." *International Journal of Information Technology* 13, no. 3 (2021): 1255-1264.

[18] Bhuiyan, Md Rafiuzzaman, Sharun Akter Khushbu, and Md Sanzidul Islam. "A deep learning based assistive system to classify COVID-19 face mask for human safety with YOLOv3." In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1-5. IEEE, 2020.

[19] R. Jayaswal and M. Dixit, "Monitoring Social Distancing based on Regression Object detector for reducing Covid-19," 2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT), 2022, pp. 635-640, doi: 10.1109/CSNT54456.2022.9787660.

[20] ZhanLi, Li, and Ye JiaWei. "Abnormal Behavior Recognition Based on Transfer Learning." In Journal of Physics: Conference Series, vol. 1213, no. 2, p. 022007. IOP Publishing, 2019.

[21] Jayaswal, R., & Dixit, M. 2020. Comparative Analysis of Human Face Recognition by Traditional Methods and Deep Learning in Real-Time Environment. In 2020 IEEE 9th International Conference on Communication Systems and Network Technologies. 66-71. April 20.

[22] Saher, R., & Anjum, M. (2021). Role of technology in COVID-19 pandemic. *Researches and Applications of Artificial Intelligence to Mitigate Pandemics*, 109–138. https://doi.org/10.1016/B978-0-323-90959-4.00005-5.

[23] Khan NS, Ghani MS (2021) A Survey of deep learning based models for human activity recognition. Wireless Pers Commun. https://doi.org/10.1007/s11277-021-08525-w

[24] Zheng, Jingxiao, Rajeev Ranjan, Ching-Hui Chen, Jun-Cheng Chen, Carlos D. Castillo, and Rama Chellappa. "An automatic system for unconstrained video-based face recognition." *IEEE Transactions on Biometrics, Behavior, and Identity Science* 2, no. 3 (2020): 194-209.

[25] R. Ranjan *et al.*, "A fast and accurate system for face detection, identification, and verification," *IEEE Trans. Biometrics Behavior Identity Sci.*, vol. 1, no. 2, pp. 82–96, Apr. 2019

[26] Ansari, Mohd Aquib, and Dushyant Kumar Singh. "ESAR, An Expert Shoplifting Activity Recognition System." *Cybernetics and Information Technologies* 22, no. 1 (2022): 190-200.

[27] Ren, Lingmei, and Yanjun Peng. "Research of fall detection and fall prevention technologies: A systematic review." *IEEE Access* 7 (2019): 77702-77722