

A Novel Cultural Crowd Model Toward Cognitive Artificial Intelligence

Fatmah Abdulrahman Baothman*, Osama Ahmed Abulnaja and Fatima Jafar Muhdher

Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, 21431, Kingdom of Saudi Arabia

*Corresponding Author: Fatmah Abdulrahman Baothman. Email: fbaothman@kau.edu.sa

Received: 05 February 2021; Accepted: 27 March 2021

Abstract: Existing literature shows cultural crowd management has unforeseen issues due to four dynamic elements; time, capacity, speed, and culture. Cross-cultural variations are increasing the complexity level because each mass and event have different characteristics and challenges. However, no prior study has employed the six Hofstede Cultural Dimensions (HCD) for predicting crowd behaviors. This study aims to develop the Cultural Crowd-Artificial Neural Network (CC-ANN) learning model that considers crowd's HCD to predict their physical (distance and speed) and social (collectivity and cohesion) characteristics. The model was developed towards a cognitive intelligent decision support tool where the predicted characteristics affect the estimated regulation plan's time and capacity. We designed the experiments as four groups to analyze the proposed model's outcomes and extract the interrelations between the HCD of crowd's grouped individuals and their physical and social characteristics. Furthermore, the extracted interrelations were verified with the dataset's statistical correlation analysis with a P-value < 0.05 . Results demonstrate that the predicted crowd's characteristics were positively/negatively affected by their considered cultural features. Similarly, analyzing outcomes identified the most influential HCD for predicting crowd behavior. The results also show that the CC-ANN model improves the prediction and learning performance for the crowd behavior because the achieved accepted level of accuracy does not exceed 10 to 20 epochs in most cases. Moreover, the performance improved by 90%, 93% respectively in some cases. Finally, all prediction best cases were related to one or more cultural features with a low error of 0.048, 0.117, 0.010, and 0.014 mean squared error, indicating a novel cultural learning model.

Keywords: Cultural crowds; learning model; artificial neural network; hHofstede cultural dimensions; predicting group behaviors; crowd management

1 Introduction

The escalated level of urbanization and expanding population has made crowded gatherings substantially prominent [1]. This phenomenon is evident from integrating crowd analytics algorithms across various domains [2–4]. Wijermans defined a crowd as “A group of individuals at the same physical location at the same time” [5], while crowd management (CM) refers to the



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

preparations and planning of an event involving forecasting and scheduling techniques [6]. The resulting event plan includes using operational measures and technical infrastructure to ensure crowd satisfaction and safety. Pre-planning for an event usually consumes approximately 90% of the entire CM effort of an event. Meanwhile, operational measures, including the potential emergencies that support the resulting plan's implementation, consume the remaining 10% [6]. The significance of event planning is observed from the incidents of Shanghai's Bund event on New Year's Eve 2015 with 300,000 crowded individuals around Chen Yi Square. A total of 36 people lost their lives due to inadequate management and control, while 47 people were severely injured [7].

Understanding and classifying the crowds using human factors (religious, social, and cultural) and employing these factors within the CM system could improve the accuracy and efficiency of the CM process. The values occur in improved decision-making, professional planning, and efficient strategies regarding crowd flow rate, density, and satisfaction level. Thus, each crowd and event have its characteristics and challenges, and *one solution cannot fit all*. Therefore, an optimal level of satisfaction with the devised CM plans must be ensured to manage and regulate people in open areas. Planning for an event usually begins well before the event, especially for events with a high uncertainty level. Similarly, hospitals may face overcrowded management issues during these events, particularly in a pandemic, such as in COVID-19. Models of crowd classification can increase doctors' productivity and hospital management systems' effectiveness, allowing appropriate utilization of their time. Time and capacity are the commonly reported issues of CM and regulation. Favaretto, Musse, and Costa [8] showed that individuals' culture affects their time to accomplish a specific task. Similarly, they reported that some nations are faster than others in completing an assigned job. Thus, time scheduling and estimated capacity are influenced by regulatory plans, which may be crucial in controlling sick or special needs individuals. The time element is essential in artificial intelligence design principles; the three-time scales are followed to explain behaviors affecting learning, development, and evolution [9].

The study aims to examine the interrelations between the three dimensions of the proposed cultural crowds (CC) learning model, which is inspired by the Big-Four Geometrical Dimensions (Big4GD) model [8]. An artificial neural network (ANN) is employed to apply the CC learning model. The three dimensions of ANN cover physical, social, and cultural characteristics using the proposed Cultural Crowd-Artificial Neural Network (CC-ANN) learning model to extract the knowledge of dimensions and characteristics. Similarly, the cultural context effect of a group's crowd, represented as Hofstede Cultural Dimensions (HCD), is determined using the proposed model. The results are verified using statistical significance to test the proposed model. The model is designed to grow into a cognitive intelligent decision support tool, which assists in crowd behavior prediction based on the "state-oriented" perspective. Human-agent behavior in a specific scenario is linked to the actual mechanism for CM systems. The objective is to improve CM efficiency, w.r.t crowd flowrate, density, collectivity, and satisfaction level. Such a learning model is limited to event locations and limited to event locations and includes facilities and environments, such as airports, hospitals, and hotels. The following subsections present some essential concepts related to the current study.

The following three points are summarizing the main contributions of this work as follows:

- The present research provides a CC learning model based on ANN toward cognitivism covering three cognitive processes, perception, learning, and knowledge acquisition. The perception nodes of ANN receive inputs of physical, social, and cultural characteristics.

The learning phase was achieved by designing four different groups of experiments. The knowledge acquisition phase is achieved using a comparative analytical technique of measuring the CC-ANN's model results, extracting interrelations between the crowd's HCD, and their behavior in terms of distance, speed, etc. collectivity, and cohesion. Accordingly, our understanding of crowds' micro-level has been enhanced with increasing individuals' satisfaction levels.

- This study designed and provides an exceptional CC-ANN learning model as the first crowd behavior learning model employed the six HCD identified by “power distance, individualism, masculinity, uncertainty avoidance, long-term orientation, and indulgence” to predict crowd's physical (distance and speed) and social (collectivity and cohesion) characteristics. Moreover, this work is pioneering in studying crowd behavior considering individuals' cultural background because it implemented the mentioned comparative analytical technique to measure the CC-ANN model's outcomes and extracted the most influential HCD (positively or negatively) in predicting crowd behavior, which leads to rich crowd management's state-of-the-art future works.
- The developed tool for crowd behavior successfully predicts individuals' behavior within a group using four characteristics considering individuals' six HCD. The CC-ANN model shows a low error of 0.048, 0.117, 0.010, and 0.014 Mean Squared Error in all prediction best cases. The Learning model uses a combination of hyperparameters; the developed CC-ANN learning model as a part of a cognitive intelligent decision support tool improves Crowd Management systems' performance within cross-cultural events. In most cases, the achieved accepted level of accuracy does not exceed 10 to 20 epochs. Moreover, the CC-ANN learning model improves the performance by 90% and 93%, respectively, in some cases.

The rest of the paper is structured as follows. Section 2 presented some essential concepts related to our study; HCD model, Big4D model, Cognitive Intelligent Decision Support Systems, and ANN. Section 3 surveys associated research with CM and the CC-based work. Section 4 explains the CC learning model. Section 5 discusses the proposed approach considering the dataset, experiments, and design and implementation of the CC-ANN learning model. In addition to the significant statistical investigation results, Section 6 presents the outcomes of the CC-ANN learning model, the limitations of the present study, and lists potential future works that could improve the proposed learning model. Finally, Section 7 summarizes the study and provides numerous recommendations to develop and enhance the current research work.

2 Background

2.1 Hofstede Cultural Dimensions (HCD) Model

Hofstede defined culture as “*the collective programming of the mind that distinguishes individuals of one group of people from another [10].*” The HCD model shows six cultural dimensions scored as integer values in the range of [0, 100], where the cultural differences between countries are apparent in the individual behavior when scoring with a difference of at least 10. [Tab. 1](#) presents the six dimensions according to the Hofstede Insights website [[11](#)].

The HCD dataset [[11](#),[12](#)] contained the total number of the cultural dimension scores for 111 countries and was included in the Hofstede study [[10](#)]. The dataset was derived from the Greet Hofstede website [[12](#)] and the Hofstede insights website (Country Comparison Tool) [[11](#)]. [Tab. 2](#) shows the scores of the six HCDs for the countries included in the current study.

Table 1: Definitions of HCD

HCD	Definition
Power distance (PDI)	Indicates power distribution of the society, where the high PDI scores reflect the unequal and hierarchical power distribution of privileges. Low PDI scores indicate that societies intend to have equal rights rather than privileges.
Individualism (IDV)	Indicates the identity of the individuals. High IDV scores indicate the individualist societies with a strong sense of “I,” where an individual’s identity is independent of others. Low IDV scores reflect the collectivist societies with a strong sense of “we,” where individuals are psychologically and practically in-person and in-group dependent.
Masculinity (MAS)	Indicates social concerns. High MAS scores indicate masculine societies, where individual interests lie in social status and personal achievement apart from material success. In addition to finding consensus and ensuring leisure time, low MAs scores reflect feminine societies concerning life quality and caring of less fortunate persons.
Uncertainty avoidance (UAI)	Indicates the acceptance of uncertainty, where high UAI scores reflect the tendencies to the written and unwritten rules of structure and predictability. Low UAI scores emphasize that uncertainty is normal, where each day is considered.
Long-Term orientation (LTO)	Indicates the orientation of individual life. High LTO scores indicate LTO societies, where individuals focus on thrift and perseverance because the world will never stop changing. Low LTO scores reflect short-term orientation, where the concerns are on the fulfillment and traditions of respecting social obligations.
Indulgence (IVR)	Indicates the attitude of the individual. High IVR scores reflect indulgent societies, where the individual can act happy. Low IVR scores indicate restraint societies, where social norms restrict the gratification of needs and reduce the importance of leisure.

Table 2: Scores of the six HCDs for the countries included in the current study

Country	PDI	IDV	MAS	UAI	LTO	IVR
Austria	11	55	79	70	60	63
China	80	20	66	30	87	24
France	68	71	43	86	63	48
Brazil	69	38	49	76	44	59
Japan	54	46	95	92	88	42
Germany	35	67	66	65	83	40
Spain	57	51	42	86	48	44
Turkey	66	37	45	85	46	49

2.2 Big-Four Geometrical Dimensions (Big4GD) Model

The Big4GD Model [8] characterizes four crowd’s related dimensions of features to describe pedestrians structured in video sequences as groups/crowds. The current study covers the following three dimensions.

- Physical dimension represents physical crowd features obtained by tracking pedestrians, such as speed, angular variations, and inter-pedestrian space.
- Social dimension represents crowd social features derived based on physical features and social interaction, such as the collectivity and cohesion of grouped individuals.
- Cultural dimension, represents the cultural features of the crowd based on the HCD model.
- The CC videos dataset is established through the Big4GD model, as illustrated in the following subsection.

The CC videos dataset, which investigates crowd behaviors' cultural aspects, is a video clips dataset from different crowded scenes of 11 different countries [8]. Moreover, this dataset contains 29 video clips varying from 100 to 900 frames per clip. The dataset contained some extracted features of walking grouped individuals classified by nationality according to their countries to investigate some aspects of their cultures. The CC dataset attributes were extracted from the videos and analyzed for the group individuals. The current study comprised the following physical and social features of the CC dataset.

- Physical features: Number of Frames, Number of people, Direction, Angular Variation, Area, Distance, Speed.
- Social features: Collectivity and Cohesion.
- Distance is related to the means of distances between the group members.
- Speed is related to the mean of group member speed.
- Collectivity indicates acting as a union degree of grouped individuals [13].
- Cohesion indicates the stability of the relationship within the group [14].

Regarding features data, the Number of Frames, the number of people features are integer values, while the other features are decimal values.

2.3 Cognitive Intelligent Decision Support Systems (IDSSs)

Crowd behavior models are recommended to develop decision support systems for CM from crowd behavioral models that support strategic decision making for crowd events planning, such as evaluating the consequences of the decisions [5]. Cognitive models comprise the following four processes: perception, learning, knowledge acquisition, and memory development. Similarly, in addition to knowledge models and methods based on decision analysis, IDSSs extensively use artificial intelligence techniques [15]. Furthermore, intellectual resources, databases, expert systems technologies, and model bases related to groups or individuals effectively support the decision-making process [15,16].

2.4 Artificial Neural Network (ANN)

ANN is useful for modeling cognitive psychology in its four processes: perception, learning, knowledge acquisition, and memory development. In addition to models' capability to classify unknown types, the high resilience to noisy information is one of these models' main advantages [17]. ANN models learn through the comparison of the classification process outcomes with the actual outputs. The calculated error improves with each backpropagation iteration, and the weights affect the hidden layer parameters for effective results. The ANN models allow adjustable hyperparameters during training to change the original design of the numbers and neurons of hidden layers, affecting the learning model's efficiency. Numerous methods and approaches are used to determine these hyperparameters, as covered by the study of Kanwar et al. [18]. Grid search is one of the approaches used to determine the best combination of ANN hyperparameters. The best combination was chosen per the least mean squared error (MSE) after testing each

combination, as illustrated in Section 4.3. Thus, the grid search could be used with the following parameters.

- Number of Epochs: It is related to iteration, where all the training examples pass one time forward and one time backward.
- Batch Size: It is related to the number of training examples in one epoch.
- Optimizer: It is related to the algorithm used to update every layer's weight after every iteration.
- Kernel Initialization: It is related to the method used to initial random weights of Keras layers.

ANN models could be implemented by using Python programming language with Keras Package [19]. Keras is a Python high-level ANN application programming interface for building fully functional deep learning models. TensorFlow is the most widely adopted usage of Keras as a low-level application programming interface backend.

3 Related Work

CM is a multidisciplinary problem that integrates different domains, such as psychology, sociology, theoretical physics, and computer sciences. Thus, this section contains a survey of existing works related to CM and CC in computer sciences. A review of recent literature on common CM practices shows that some researchers presented modeling of crowd psychology behavior [20,21], while others were linked to the behavior prediction based on physics-inspired models and behavior recognition using various sensors and analysis [22]. Some studies also highlight different approaches for studying crowd behavior through crowd behavior prediction models [23,24] using simulation for event planning. The prediction models usually relate to crowd dynamics and CM patterns, which are the “*coordinated movement of a large number of individuals to which is a semantically relevant meaning can be attributed, depending on the respective application [6],*” such as individual queues and groups at a specific location. One of the popular crowd behavior models is the social-force model. By contrast, the CROSS model represents crowd behavior through simulation, focusing on the general rather than specific crowd behaviors, such as fleeing or violence [5]. The CROSS model has multiple levels: 1) the group level where behavior patterns emerge, 2) the individual level where the behavior is generated, and 3) the cognitive level where the behavior is affected [5]. Furthermore, the study by Venuti et al. [25] and Bellomo et al. [26] provided an extensive survey on crowd simulation and modeling.

Computer vision enhances images and segmentation in many fields, such as diagnostic models in solving medical problems [27]. Yaseen et al. [28] and Zhan et al. [29] adapted computer-vision techniques to recognize crowd behaviors by characterizing and automatically detecting anomaly cases within a crowd. Atallah et al. [30] employed ubiquitously extensive pervasive technologies on social behavior monitoring, such as smartwatches and phones, by distributing sensing modalities, such as temperature, movement, and spatial proximity, over a wide range; for example, detecting crowd dynamics, such as social groups, pedestrian bottlenecks, and crowd flow, using smartphones [31]. Similarly, crowd textures, such as pedestrian clogging and lanes, are detected using proximity sensors because of the strong spatiotemporal nature [32]. Some studies also used microphones [31–33] to recognize places and locations and measure the crowd mood [34]. Finally, accelerometers are utilized to characterize individuals' queues and activities, such as walking and running [35]. Such approaches could be related to ambient intelligence, which is the “*electronic systems that are sensitive and responsive to the presence of people [6],*” to support and enhance crowd evacuation monitoring, including situation and social-aware computing devices and sensors.

Furthermore, some researches focused on managing emergencies and disasters [36]. Therefore, distinguishing between CM and crowd control is important. Crowd control involves adopting measures and strategies to manage crowds when getting out of control [6]. By contrast, CM is related to the correct and balanced allocation of resources concerning the preparations for avoiding crowd control, monitoring, predicting, and steering crowd behavior [6]. The CM strategy generally includes the following: 1) capacity planning in the short and long term, 2) understanding crowd behavior, 3) crowd control, and 4) stakeholder approach [37]. Moreover, the CM integrated approach is a continuous process of planning, organizing, coordinating, implementing, and monitoring. Some studies focused on specific cases, such as underground stations [38,39], and controlling air traffic [40,41], investigating how accidents could be avoided during events. These studies supported event planning and debriefing by adapting technologies to support decision-making. Moreover, a socio-technical system analyzed the crowd behaviors under normal situations and emergencies to guide and facilitate management and planning of crowded events [42]. The socio-technical systems are those “*designed and operated with a holistic approach for optimizing both technical and social factors* [6].” These systems are also considered in the organizational perspectives’ decision-making process through suitable compromises [42]. The decision based on classification tasks is complicated in cross-cultural events [43].

Crowd individuals and groups are a core element to CM [37]. These elements are strongly related to the event type, season, frequency, time, location, expected type, and the crowd’s anticipated number. The event type could be religious, educational, political, youth festival, or sport. The season of the event describes the frequency period during the year for each individual. The event time frames and locations include temporary/permanent, topography, remote/urban area, plain/hilly terrain, and close/open space. In addition to unwanted visitors as terrorists, thefts, and disruption, the crowd types and numbers include gender, age, local or visitors, and motivations, such as social, religious, political, entertainment, and economic [37]. The crowd classification tool uses a knowledge-based approach, ontology, and fuzzy rules to classify the crowd according to sociological theory [44]. Furthermore, crowd video analysis applications classify the crowds, in which descriptors define crowd properties fed through two-dimensional convolution neural networks for complex feature generation and classification [45]. Favaretto, Musse, and Costa [8] built the Big4GD Model to represent the following four dimensions: physical, social, cultural, personal, and emotional. The dimensions of this model describe grouped pedestrians in crowds’ video sequences. The established Big4GD model creates the CC project (Virtual Humans Simulation Laboratory), which provides an organized and classified CC dataset from crowded scenes, illustrating various aspects between people from different cultures [46]. The attributes of the CC dataset were extracted from the videos and analyzed for the individual groups. By contrast, the proposed model considers the HCD to predict the following four attributes; distance, speed, collectivity, and cohesion. Another work [47] added a virtual CC environment. Both studies [8,47] investigated the crowd’s culture according to their behavior and extracted the HCD crowds’ scores according to the crowd behavior under experiment. Alternatively, the current study considered HCD to predict physical and social crowd behaviors.

No available crowd behavior prediction tool considers the human frame of reference of crowd individuals (i.e., religious, social, and cultural) based on the literature review and the researchers’ best knowledge. The inclusion of human factors produces an efficient CM system, especially if integrated with an intelligent decision support tool. For example, this system can be used as a precaution phase in the cognitive CM systems to predict crowd movements and behaviors before arriving at the event location and ensure the precaution decision to avoid overcrowding

in hospitals. Moreover, this system could generate alternative plans as a backup in case of an accident or unforeseen circumstances, and the system could adapt according to the tool predictions. Information could be automatically collected by agents of cognitive platforms supported by sensors and the Internet of Things technologies. The current study targets high-density crowds, which usually contain grouped individuals of different nationalities and cultures. Thus, a cultural element is included, and the impact of the cultural frame of reference on crowd behavior prediction is tested. Significantly, the research question aims to discover whether using the crowd's cultural frame of reference would improve their physical and social behaviors' prediction efficiency. The HCD model [48] and the CC dataset [46] are respectively employed to represent the crowd's cultural frame and crowd characteristics. Thus, CC and HCD datasets are combined in this study, and training and testing experiments are designed.

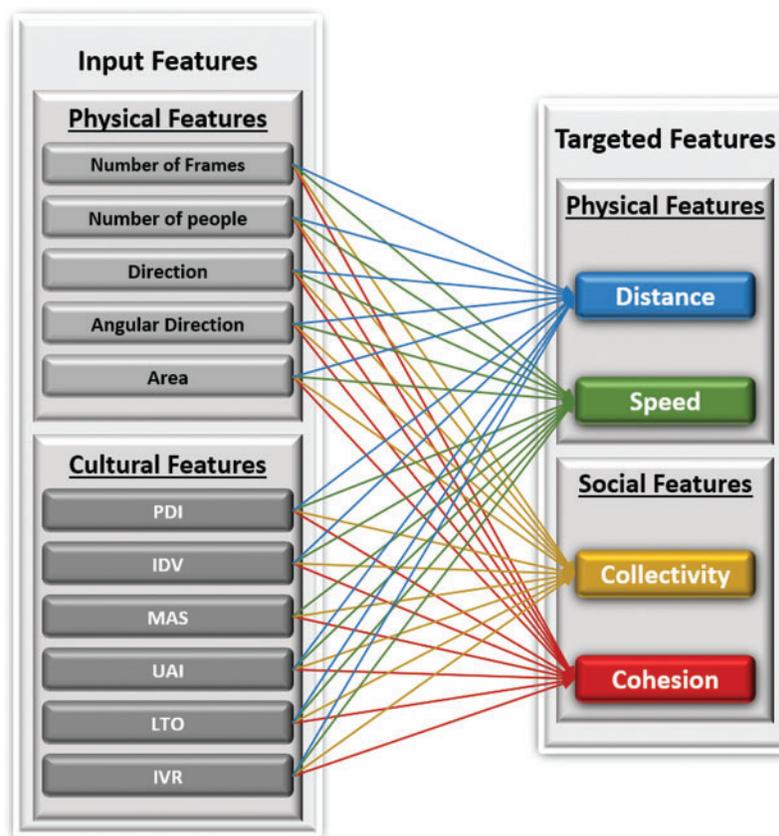


Figure 1: Proposed CC learning model

A model, which is represented as the HCD, is proposed to predict the physical (distance and speed) and social (collectivity and cohesion) characteristics of crowds considering their cultural frame of reference. Accordingly, the proposed CC learning model comprises physical, social, and cultural dimensions, as shown in Fig. 1. The concept of cultural dimensions is implemented based on the seminal work of Hofstede [48], as illustrated in Section 2.1. By contrast, the dimension of the three Big4GD models [8] is employed to represent crowd behavior characteristics, as illustrated in Section 2.2. Also, this study focused on grouped individuals of crowds. The cultural dimension

was successfully mapped to the characteristic parameters for the individuals and groups of crowds. Similarly, all input features were connected to all-out features to determine the importance of HCD in predicting the targeted attributes. The training and testing experiments were designed based on HCD feeding to the CC dataset to study the impact of predicting crowd physical and social targeted attributes, as explained in the results section.

The ANN is used as a CC learning model for an individual's emergent behaviors, group patterns from a dynamic system (called collective intelligence), and time changes for CM to create a cognitive decision support tool. ANN is applied in the CC learning model because it is useful for modeling cognitive psychology in perception, learning, knowledge acquisition, and memory development. As previously mentioned, the proposed learning model is developed as a part of IDSS. The work in this article is goal-oriented research. This study aims to discover the existing interrelations between the following three dimensions: 1) frame of reference of culture crowds (HCD), 2) physical (distance and speed), and 3) social (collectivity and cohesion) characteristics. The three cognitive processes, namely perception, learning, and knowledge acquisition, are covered in this study to achieve cognitivism in the model. The perception nodes receive inputs of physical, social, and cultural characteristics, as illustrated in the methodology section. The learning from the experiment process is explained in the results section. The knowledge acquisition phase is achieved using analysis and extraction processes. Such a method focuses on discovering the interrelations among the HCD of crowds, including distance, speed, collectivity, and cohesion, as illustrated in the results section.

4 Proposed Methodology

The methodology focuses on extracting the interrelations between the crowd's six HCDs and its four targeted physical (distance and speed) and social (collectivity and cohesion) attributes in predicting crowd behavior. The following sections describe the methodology considering the dataset, experiments, and design of the CC-ANN learning model and its implementation.

4.1 Dataset

Two datasets, namely the HCD and CC datasets, are integrated into the country to build the proposed model. The HCD dataset contains the scores (ranging from 0 to 100) of the following six HCDs: "power distance (PDI), individualism (IDV), masculinity (MAS), uncertainty avoidance (UAI), long-term orientation (LTO), and indulgence (IVR)." The CC dataset is suitable for the current study because it has the same interest and vision of a cross-cultural crowd investigation. Moreover, the CC dataset is suitable considering sample and population sizes with a 99% confidence level and a 0.5 confidence interval [49]. This study comprised an integrated dataset, which had the following input and targeted attributes.

- Input attributes: Number of frames, Number of people, Direction, Angular variation, Area, and HCD (PDI, IDV, MAS, UAI, LTO, and IVR).
- Targeted attributes: Distance, Speed, Collectivity, and Cohesion.

The gradual appending of the HCD is used to input features of the CC dataset to design the current research experiments (as discussed in the following section). The records related to Portugal and the United Kingdom were eliminated because they were less than three records and considered outliers. The records of the United Arab Emirates were also eliminated due to the unavailable scores of LTO and IVR features. Furthermore, the dataset was divided into two distinct subsets: 70:30 for the training and testing phases, respectively. Finally, a statistical correlation analysis of the dataset was performed to verify the model statistically. The analysis

aims to determine whether HCD has statistically significant relations with the targeted attributes through the designed experiments. If the relation has P-value < 0.05 , then the relation is generally statistically significant [50].

4.2 Experiments

All experiments were designed to predict group distance, speed, collectivity, and cohesion separately; thus, these experiments were conducted four times for each targeted attribute. Moreover, the 29 designed experiments were classified into four groups according to appending/removal of the six HCD to CC dataset features as inputs to the proposed CC-ANN learning model as follows.

- 1st group: “How the performance of ANN will be affected by adding all six HCDs.” Thus, two experiments were conducted: one experiment contained only crowd features from the CC dataset, and the other experiment comprised all six HCDs in addition to the CC dataset.
- 2nd group: “How the ANN’s performance will be affected by adding one of the HCDs.” This experiment indicates how each targeted attribute was affected by each HCD positively or negatively. Thus, six experiments, where each one appends only one of the HCD, were conducted.
- 3rd group: “How the ANN’s performance will be affected by removing one of the HCDs.” This experiment indicates that model inputs will contain CC dataset features and five HCDs and determine how each targeted attribute is affected by removing one of the HCDs positively or negatively. Thus, six experiments, where each contained all HCDs except one, were conducted.
- 4th group: “How the performance of ANN will be affected by adding two HCDs.” This experiment determines how each targeted attribute is affected positively or negatively by each couple (permutation) of HCDs. Thus, 15 experiments, where each appends one permutation couple, were conducted.

4.3 CC-ANN Learning Model Design

The present study describes the design of the CC-ANN learning model, as presented in Fig. 2. As mentioned before, the grid search technique is used to determine the best hyperparameter combination of ANN. Thus, the grid search is used with the following parameters.

- Number of Epochs: The available options at grid search were (50, 100, and 150).
- Batch Size: The available options at grid search were (5, 10, and 20).
- Optimizer: The available options at the grid search were (‘RMSProp’ and Adam).
- Kernel Initialization: The grid search options were (glorot_uniform, normal, and uniform).

The available options at the grid search were selected based on the trial and error approach during the current phase. The other hyperparameters, such as (weight and bias initialization, number of hiding layers and neurons, and activation functions), were determined using the trial and error technique method. All options and hyperparameters are optimized in the next phases of the current research to obtain effective results. The grid search was performed for each targeted attribute separately. Tab. 3 presents the best combinations of hyperparameters based on the least MSE, where the values of these combinations are used in all experiments related to their targeted attributes.

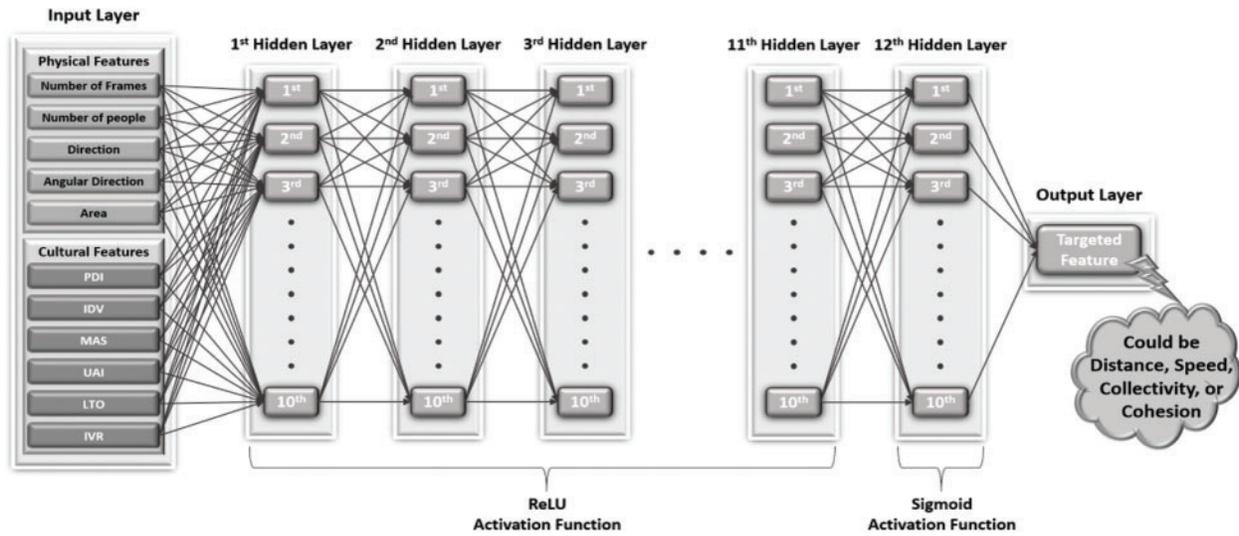


Figure 2: CC-ANN learning model design

Table 3: Best combinations of hyperparameters chosen by grid search

Attributes/parameter	Distance	Speed	Collectivity	Cohesion
Epochs	100	100	100	150
Batch	20	5	20	10
Optimizer	RMSProp	Adam	Adam	RMSProp
Kernel initialization	Normal	Normal	Normal	Uniform
Loss (MSE)	0.0487	0.1050	0.0101	0.0138

The neuron number was changed based on the number feature variation as per the appending of HCD to CC dataset features considering the number of neurons at the input layer. Furthermore, the output layer has only one neuron because the model has only one output, where the experiments were performed separately for each targeted attribute. The model also has 12 hidden layers, where each layer has 10 neurons. The ReLU activation function was applied at the first 11 hidden layers. By contrast, the Sigmoid activation function was applied at the last hidden layer. By contrast, no activation function was used at the output layer because the current study focuses on the direct prediction of a numerical value. The MSE regression metric was used to calculate the difference between the actual and predicted values for each instance [51] in evaluating the CC-ANN learning model. The following subsection describes the proposed model mathematically. Afterward, the describes the implementation of the proposed design in the next subsection.

4.4 Mathematical Background and Description of the Model

This subsection describes the proposed CC-ANN learning model mathematically. Fig. 3 shows an abstracted design of the proposed model.

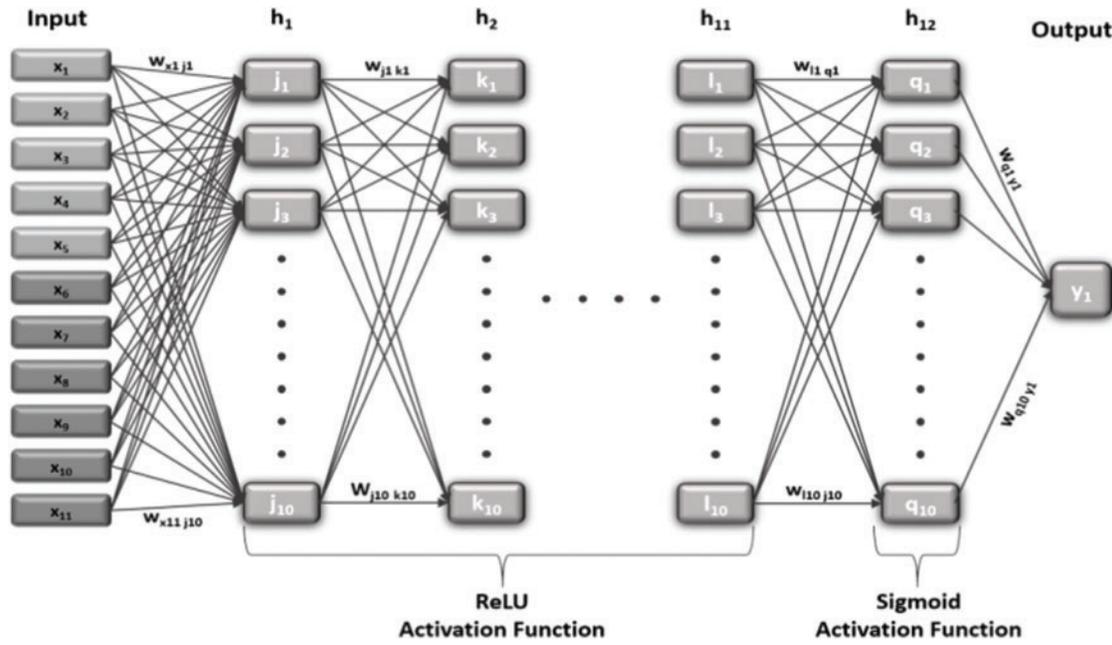


Figure 3: CC-ANN learning model abstracted design

First, the forward propagation of the proposed model will be explained. The input matrix includes n elements, while the maximum number of inputs is 11. The number of inputs is not fixed because it changes according to the experiment group, as previously mentioned in Section 4.2. The output matrix has only one element, as shown in the following equations:

$$\text{Input} = [i_1 \ i_2 \ \dots \ i_n], \quad (1)$$

$$\text{Output} = [y_1]. \quad (2)$$

As previously mentioned, the ReLU activation function is used at the first 11 hidden layers. The following equation describes the ReLU activation function [52]:

$$\text{ReLU} = \max(0, x). \quad (3)$$

Meanwhile, the following equations describe the matrix operation of the first hidden layer (h_1):

$$[h_{1_{in1}} \ h_{1_{in2}} \ \dots \ h_{1_{in10}}] = [i_1 \ i_2 \ \dots \ i_n] \times \begin{bmatrix} w_{i_1 j_1} & w_{i_1 j_2} & \dots & w_{i_1 j_{10}} \\ w_{i_2 j_1} & w_{i_2 j_2} & \dots & w_{i_2 j_{10}} \\ \vdots & \vdots & \vdots & \vdots \\ w_{i_n j_1} & w_{i_n j_2} & \dots & w_{i_n j_{10}} \end{bmatrix} + [b_{j_1} \ b_{j_2} \ \dots \ b_{j_{10}}]. \quad (4)$$

Thus, the ReLU operation at h_1 layer is described as follows:

$$[h_{1_{out1}} \ h_{1_{out2}} \ \dots \ h_{1_{out10}}] = [\max(0, h_{1_{in1}}) \ \max(0, h_{1_{in2}}) \ \dots \ \max(0, h_{1_{in10}})]. \quad (5)$$

Afterward, the following equation describes the matrix operation of the second hidden layer (h2):

$$[h2_{in1} \ h2_{in2} \ \dots \ h2_{in10}] = [h1_{out1} \ h1_{out2} \ \dots \ h1_{out10}] \times \begin{bmatrix} w_{j1k1} & w_{j1k2} & \dots & w_{j1k10} \\ w_{j2k1} & w_{j2k2} & \dots & w_{j2k10} \\ \vdots & \vdots & \vdots & \vdots \\ w_{j10k1} & w_{j10k2} & \dots & w_{j10k10} \end{bmatrix} + [b_{k1} \ b_{k2} \ \dots \ b_{k10}]. \tag{6}$$

Then, the ReLU operation at the h2 layer is described as follows:

$$[h2_{out1} \ h2_{out2} \ \dots \ h2_{out10}] = [\max(0, h2_{in1}) \ \max(0, h2_{in2}) \ \dots \ \max(0, h2_{in10})]. \tag{7}$$

The rest of the ReLU hidden layers (from h3 to h11) function according to the same concepts of the operation matrix and ReLU operation of the h2 layer. Conversely, the last hidden layer works with the Sigmoid activation function. The following equation describes the Sigmoid activation function [53]:

$$Sigmoid = \frac{1}{(1 + e^{-x})}. \tag{8}$$

Meanwhile, the matrix operation of the last hidden layer (h12) is:

$$[h12_{in1} \ h12_{in2} \ \dots \ h12_{in10}] = [h11_{out1} \ h11_{out2} \ \dots \ h11_{out10}] \times \begin{bmatrix} w_{l1q1} & w_{l1q2} & \dots & w_{l1q10} \\ w_{l2q1} & w_{l2q2} & \dots & w_{l2q10} \\ \vdots & \vdots & \vdots & \vdots \\ w_{l10q1} & w_{l10q2} & \dots & w_{l10k10} \end{bmatrix} + [b_{q1} \ b_{q2} \ \dots \ b_{q10}]. \tag{9}$$

Accordingly, the Sigmoid operation at the h12 layer is described as follows:

$$[h12_{out1} \ h12_{out2} \ \dots \ h12_{out10}] = \left[\frac{1}{(1 + e^{-h12_{in1}})} \ \frac{1}{(1 + e^{-h12_{in2}})} \ \dots \ \frac{1}{(1 + e^{-h12_{in10}})} \right]. \tag{10}$$

Second, the backpropagation of the proposed model will be explained. The following equation describes the change in error considering the weight:

$$\frac{\partial error}{\partial w_{ij}}. \tag{11}$$

Moreover, the following equation describes updating of the weights through the backpropagation:

$$w'_{ij} = w_{ij} - \alpha \frac{\partial error}{\partial w_{ij}}, \tag{12}$$

where α is the learning rate.

The derivative of the Sigmoid activation function is generally presented as follows [53]:

$$\partial \text{Sigmoid}(x) = \text{Sigmoid}(x) \times (1 - \text{Sigmoid}(x)). \quad (13)$$

Thus, the derivative of the Sigmoid layer (h12) is:

$$\frac{\partial h12_{out1}}{\partial h12_{in1}} = \frac{\partial \text{Sigmoid}(h12_{in1})}{\partial h12_{in1}}, \quad (14)$$

$$\frac{\partial h12_{out1}}{\partial h12_{in1}} = \text{Sigmoid}(h12_{in1}) \times (1 - \text{Sigmoid}(h12_{in1})), \quad (15)$$

$$\begin{bmatrix} \frac{\partial h12_{out1}}{\partial h12_{in1}} \\ \frac{\partial h12_{out2}}{\partial h12_{in2}} \\ \vdots \\ \frac{\partial h12_{out10}}{\partial h12_{in10}} \end{bmatrix} = \begin{bmatrix} \text{Sigmoid}(h12_{in1}) \times (1 - \text{Sigmoid}(h12_{in1})) \\ \text{Sigmoid}(h12_{in2}) \times (1 - \text{Sigmoid}(h12_{in2})) \\ \vdots \\ \text{Sigmoid}(h12_{in10}) \times (1 - \text{Sigmoid}(h12_{in10})) \end{bmatrix}. \quad (16)$$

By contrast, the derivative of the ReLU activation function is [52]:

$$\partial \text{ReLU}(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x > 0. \\ \text{undefined} & \text{if } x = 0 \end{cases} \quad (17)$$

Thus, the derivative of the ReLU layer (h11) is:

$$\frac{\partial h11_{out1}}{\partial h11_{in1}} = \frac{\partial \text{ReLU}(h11_{in1})}{\partial h11_{in1}}, \quad (18)$$

$$\frac{\partial h11_{out1}}{\partial h11_{in1}} = \begin{cases} 0 & \text{if } h11_{in1} < 0 \\ 1 & \text{if } h11_{in1} > 0, \\ \text{undefined} & \text{if } h11_{in1} = 0 \end{cases} \quad (19)$$

$$\begin{bmatrix} \frac{\partial h_{11_{out1}}}{\partial h_{11_{in1}}} \\ \frac{\partial h_{11_{out2}}}{\partial h_{11_{in2}}} \\ \vdots \\ \frac{\partial h_{11_{out10}}}{\partial h_{11_{in10}}} \end{bmatrix} = \begin{bmatrix} \begin{cases} 0 & \text{if } h_{11_{in1}} < 0 \\ 1 & \text{if } h_{11_{in1}} > 0 \\ \text{undefined} & \text{if } h_{11_{in1}} = 0 \end{cases} \\ \begin{cases} 0 & \text{if } h_{11_{in2}} < 0 \\ 1 & \text{if } h_{11_{in2}} > 0 \\ \text{undefined} & \text{if } h_{11_{in2}} = 0 \end{cases} \\ \vdots \\ \begin{cases} 0 & \text{if } h_{11_{in10}} < 0 \\ 1 & \text{if } h_{11_{in10}} > 0 \\ \text{undefined} & \text{if } h_{11_{in10}} = 0 \end{cases} \end{bmatrix} \cdot \quad (20)$$

The rest of the ReLU hidden layers (from h10 to h1) function according to the same derivative of the ReLU activation function of the h11 layer.

Regarding computational complexity, there is an activation function $g(x)$ for each layer and a matrix multiplication calculated as mentioned above. The asymptotic runtime of the naive matrix multiplication is $O(n^3)$ [54]. The run time of $g(x)$ function is $O(n)$ because it is an elementwise function. In general, analyzing feedforward dimensions as a recurrent fully connected network could be calculated as the following equations [55]:

$$n_{mul} = \sum_{k=2}^{n_{layers}} (n_k^3) + (n) \quad (21)$$

$$n_g = \sum_{k=1}^{n_{layers}} (n_k) \quad (22)$$

where n_k is the neuron number in layer k that including the bias unit, n_{mul} is the number of the performed multiplications, and n_g is related to how many the activation function was applied. Thus, the time complexity is:

$$time = n_{mul} + n_g \quad (23)$$

$$time = \sum_{k=2}^{n_{layers}} (n_k^3) + (n) + \sum_{k=1}^{n_{layers}} (n_k) \quad (24)$$

Similarly, the total runtime with L layers in backpropagation with the delta error is [55]:

$$O(time_{error}) = n^4 \quad (25)$$

Therefore, the time to calculate all weights between n layers is [55]:

$$O(time_{weight}) = n^4 \quad (26)$$

Concerning the computational complexity of the CC-ANN learning model, due to the variation of the number of input features in each experimental group, the whole work's computational complexity is not unified. Therefore, we applied Eqs. (24)–(26) to calculate the layers' computational complexity because it is the most complicated and fixed part of the proposed design between all experiments. Relating to forward propagation, the time complexity of the 12 hidden layers with 10 neurons per each was calculated as the following:

$$time = \sum_{k=2}^{12} (10^3) + (10) + \sum_{k=1}^{12} (10) = 12,240 \quad (27)$$

Regarding backpropagation, the total runtime is:

$$O(time_{error}) = n^4 = 12^{10} = 6.19174e10 \quad (28)$$

Therefore, the time to calculate all weights is:

$$O(time_{weight}) = n^4 = 12^{10} = 6.19174e10 \quad (29)$$

On the other hand, Tab. 3 shows the best number of epochs extracted by grid search as 100 and 150 epochs. In contrast, the proposed CC-ANN learning model achieved an acceptable level of accuracy within 10 to 20 epochs in most cases. Accordingly, the CC-ANN learning model improves the performance in percentage, as shown in the following Tab. 4.

Table 4: CC-ANN performance improvement (in percentage)

	100 Epochs (selected by grid search) (%)	150 Epochs (selected by grid search) (%)
10 Epochs (best case)	90	93.33
20 Epochs (worst case)	80	86.67

4.5 Model Implementation

The ANN is implemented using the Python programming language with Keras Package [19]. The establishment and evaluation of the model went through the following steps.

- The constructed model is sequential, where layers are linearly stacked, thereby adding more than one layer to build the model. The dense layer is used to have fully-connected layers [56].
- The essential computation included the activation functions (ReLU) and (Sigmoid) plus Kernel Initializer (normal) and (uniform) initializers.
- The model went through compiling with the loss function and regression metric (MSE) after its establishment. Moreover, optimizers (Adam) and (RMSProp) are included.
- Finally, the model is fitted with some data for its evaluation by providing the suitable training/testing split [70:30] splitting percentage. The 10-fold cross-validation is used in such a phase.

5 Results and Discussion

The most significant results of all the executed experiments were based on appending one or more HCDs to the CC dataset and the dataset’s statistical correlation analysis. The observations covered the effect of appending and removing HCD on predicting targeted attributes (distance, speed, collectivity, and cohesion). The best and worst cases for all experiments according to their MSE are presented in [Tab. 5](#) and [Fig. 4](#). The analysis of results determined that the focus is not on value addition to all HCDs at once. This study focused on “how targeted attributes are affected positively/negatively by the HCD,” that is, defining the interrelations between targeted attributes and each HCD feature.

Table 5: Predicting best and worst cases of targeted attributes

Target feature	Best case		Worst case	
	Experiment	MSE	Experiment	MSE
Distance	CC + (HCD-MAS)	0.048	CC + (MAS&IVR)	0.054
Speed	CC + (HCD-IVR)	0.117	CC + (IDV)	0.250
Collectivity	CC + (MAS<O)	0.010	CC + (HCD-LTO)	0.011
Cohesion	CC + (PDI&IVR)	0.014	Only CC	0.015

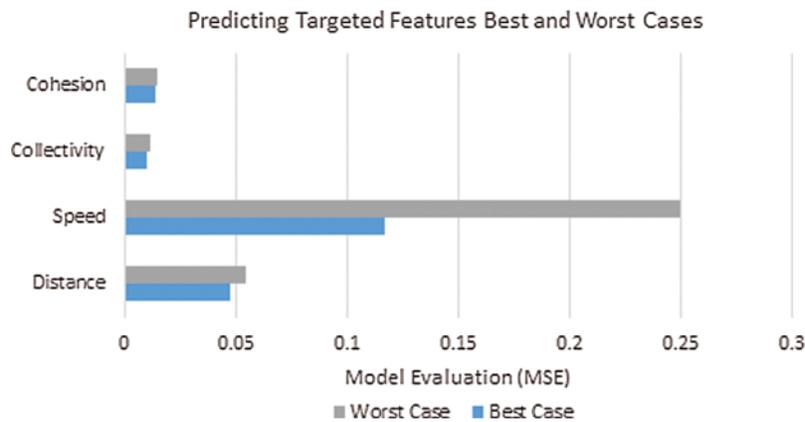


Figure 4: Predicting best and worst cases of targeted attributes

By contrast, the difference between MSE values in the best and worst cases is small, with a minimal range value between the maximum and minimum targeted attributes, as shown in [Tab. 6](#). CM plans would be influenced by a slight difference regarding large and high-density crowds. Thus, if the proposed model is applied for a crowd of three million individuals, then the small difference will significantly impact its regulation plan, as shown in [Tab. 7](#). The conducted statistical correlation analysis provided that all the designed experiments’ input features were significant because their P-value < 0.05. Otherwise, not all HCDs were significant in all experiments, while the significant HCD verified the CC-ANN outcomes, as explained in the section’s remainder.

Table 6: Ranges of actual values of targeted attributes

Distance range	Speed range	Collectivity range	Cohesion range
1.02	1.87	0.96	0.90

Table 7: Impact of the small difference between best and worst MSEs on a crowd of three million individuals

Targeted attribute	Best MSE (for 3,000,000 crowd)	Worst MSE (for 3,000,000 crowd)	MSE difference (for 3,000,000 crowd)
Distance	143,130	162,396	19,266
Speed	351,564	748,815	397,251
Collectivity	29,532	33,963	4431
Cohesion	41,532	44,604	3072

Notably, predicting the best case for distance was related to the appending of all HCDs, except the MAS dimension, as shown in Fig. 5. Thus, the distance between the individuals of a group has a close relationship with the appended HCD. By contrast, the absence of the MAS cultural dimension afforded the model an improved capability to predict distance value. The MAS dimension shows a poor correlation between the individuals from masculine/feminine societies and their distance, as mentioned in Tab. 1. Moreover, the results showed that predicting worst-case distance was related to appending MAS and IVR dimensions. Thus, the MAS dimension's existence negatively affects the model performance because its absence through all HCDs improved the model's capability to predict distance values (predicting best case distance). Moreover, the combination of the MAS dimension with IVR confused the model in predicting distance value. Concerning MAS and IVR dimensions, a poor correlation is observed between the individuals from masculine/feminine and indulgence/restraint societies, as mentioned in Tab. 1. Thus, a poor correlation existed even between such a combination and the distance of individuals. The statistical correlation analysis of the dataset regarding distance indicates that only four experiments contained HCDs as a statistically significant predictor, as shown in Tab. 8. Three experiments included the MAS dimension as the only culturally significant predictor. Accordingly, the previously mentioned findings for best and worst predicting distance cases with the proposed CC-ANN learning model are verified.

Predicting the best case of speed is related to appending all HCDs, except IVR, as shown in Fig. 6. Consequently, the walking speed of grouped individuals is different according to their culture. Conversely, the absence of the IVR dimension improved the capability of the model to predict speed value. The IVR dimension may be related to poor correlation between the individuals from Indulgence/Restraint societies and their speed. Notably, in addition to the CC dataset, predicting the worst case of speed is associated with the IDV dimension, which means that the existence of such a dimension confuses the model in speed prediction. The IDV dimension may be due to the poor correlation between the individuals from individualist/collectivist societies and their speed. In addition, speed is the only targeted attribute with one and more HCDs as statistically significant predictors in most of the experiments, as shown in Tab. 9. This condition reflects the correlation of speed prediction to HCD. By contrast, predicting the best cases of speed did not have any culturally significant predictors.

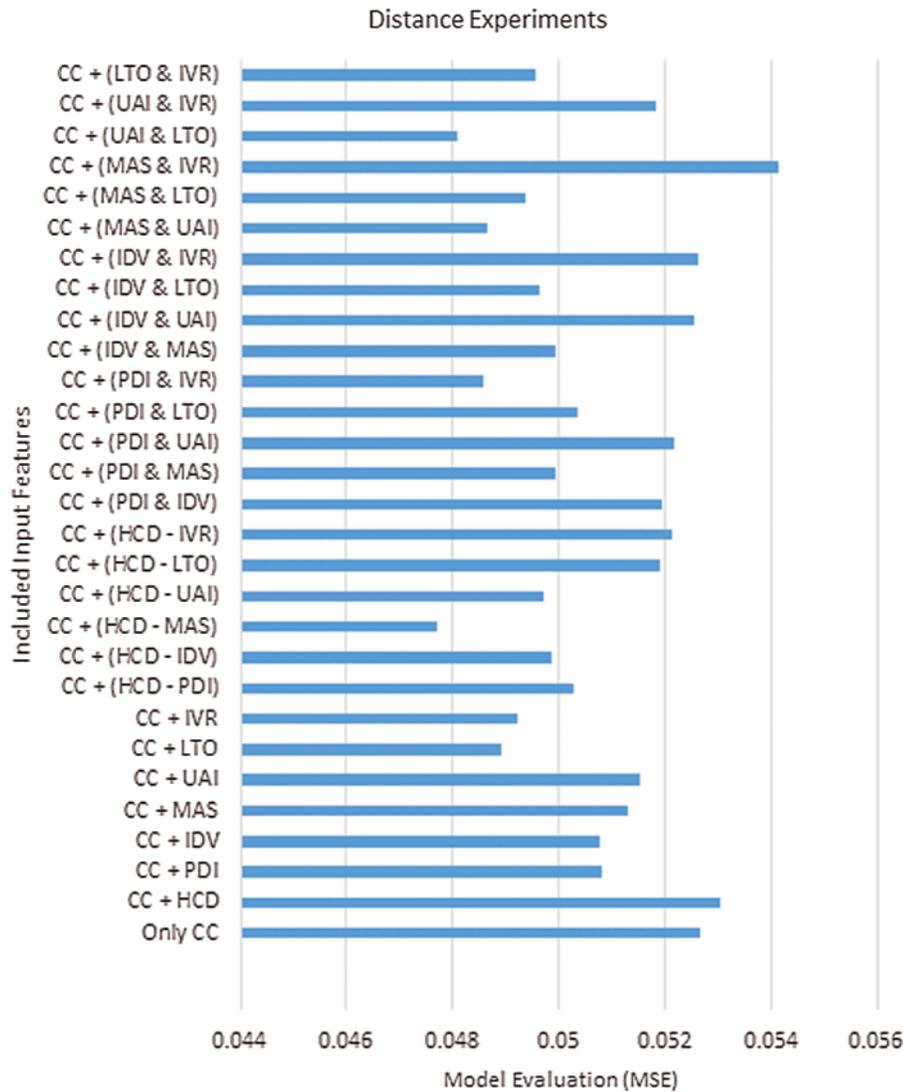


Figure 5: Model evaluation (MSE) of predicting distance experiments

Table 8: Statistically significant relations of HCD with targeted attributes of distance

HCD predictors	HCD significant	HCD predictors	HCD significant
HCD-IDV	MAS	MAS & LTO	MAS
HCD-MAS	IDV	MAS & IVR	MAS

Notably, predicting collectivity best case is related to the appending of MAS and LTO dimensions, as observed in Fig. 7. Therefore, the combination provides the model an improved capability to predict the collectivity of grouped individuals. The definitions of MAS and LTO dimensions indicate their possible association with a strong correlation between the individuals from masculine/feminine and long-/short-term oriented societies, as mentioned in Tab. 1. Moreover, a

strong correlation is found between the combinations and the individuals and their collectivity. By contrast, the prediction of the worst case of collectivity is related to appending all HCDs, except the LTO dimension. Thus, the absence of LTO negatively affects the model performance because its combination with the MAS dimension gave the best result in predicting collectivity. The LTO dimension indicates a strong correlation between the individuals from Long/Short-term oriented societies and their collectivity.

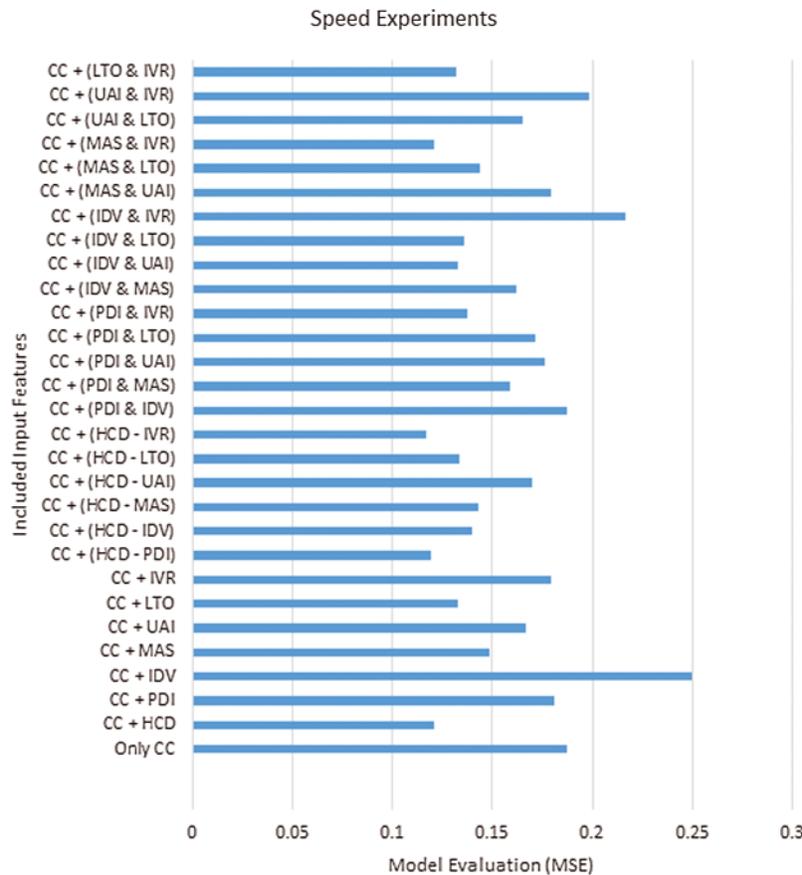


Figure 6: Model evaluation (MSE) of predicting speed experiments

Table 9: Statistically significant relations of HCD with targeted attributes of speed

HCD predictors	HCD significant	HCD predictors	HCD significant	HCD predictors	HCD significant
IDV	IDV	HCD-UAI	PDI, IDV, LTO, IVR	IDV&UAI	UAI
MAS	MAS	HCD-LTO	PDI, MAS, UAI, IVR	IDV<O	IDV, LTO
UAI	UAI	PDI&IDV	PDI, IDV	IDV&IVR	IVR
LTO	LTO	PDI&MAS	PDI, MAS	MAS&UAI	MAS, UAI
IVR	IVR	PDI&UAI	UAI	MAS<O	LTO
HCD-PDI	IDV, UAI, LTO, IVR	PDI<O	LTO	MAS&IVR	MAS, IVR
HCD-IDV	PDI, UAI, IVR	PDI&IVR	IVR	UAI<O	UAI, LTO
HCD-MAS	LTO, IVR	IDV&MAS	IDV, MAS	LTO&IVR	LTO

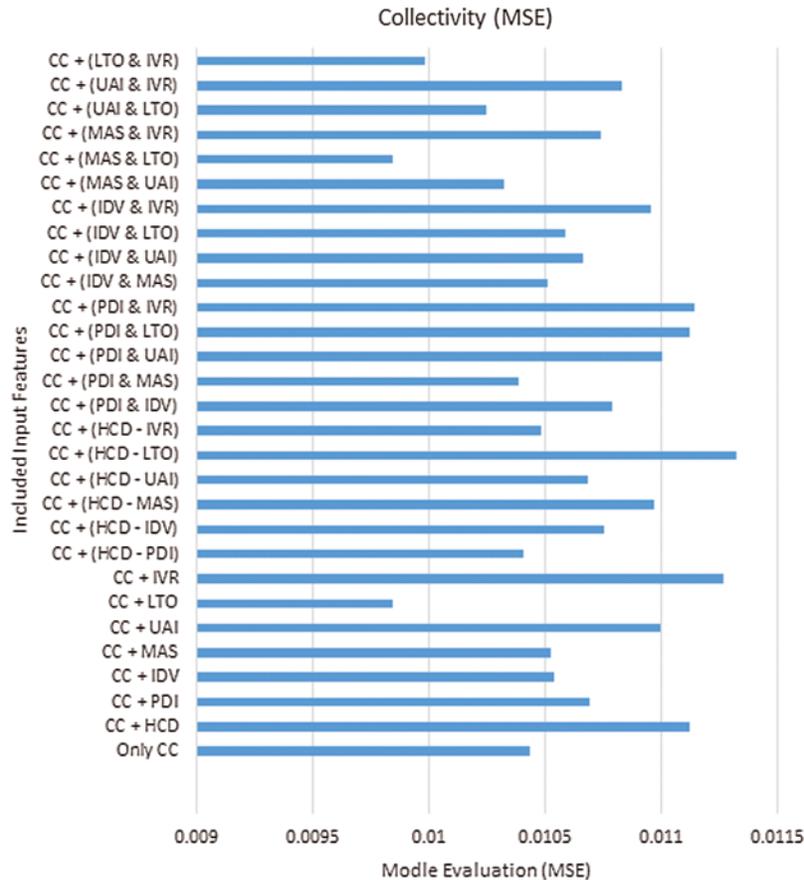


Figure 7: Model evaluation (MSE) of predicting collectivity experiments

The absence of such dimension confused the model in predicting collectivity values despite the existence of other HCDs. On the contrary, collectivity is the only targeted attribute with no HCD as a statistically significant predictor in any experiment. This finding shows a strong contribution to the proposed model in extracting some interrelations that were not discovered statistically.

Fig. 8 shows that predicting the best case of cohesion is related to the appending of PDI and IVR dimensions. Therefore, the combination provides the model with improved capability to predict cohesion value. The PDI and IVR dimensions show that this condition may be related to a strong correlation between the individuals who came from low/high power distance and indulgence/restraint societies. Moreover, a robust correlation is present between such combination and the cohesion of individuals. Furthermore, predicting the worst case of cohesion was related to the absence of all HCDs. This condition indicates a strong relationship between individual cohesion of crowd and their culture, wherein the lack of culture dimensions confuses the model in predicting cohesion values. Similarly, the statistical correlation analysis of the dataset regarding cohesion indicates that only six experiments contained HCD as statistically significant predictors, as shown in Tab. 10. The six experiments included only one HCD significant predictor (UAI, LTO, and IVR). Accordingly, verifying our findings, as mentioned before, predicts cohesion best case with our CC-ANN learning model.

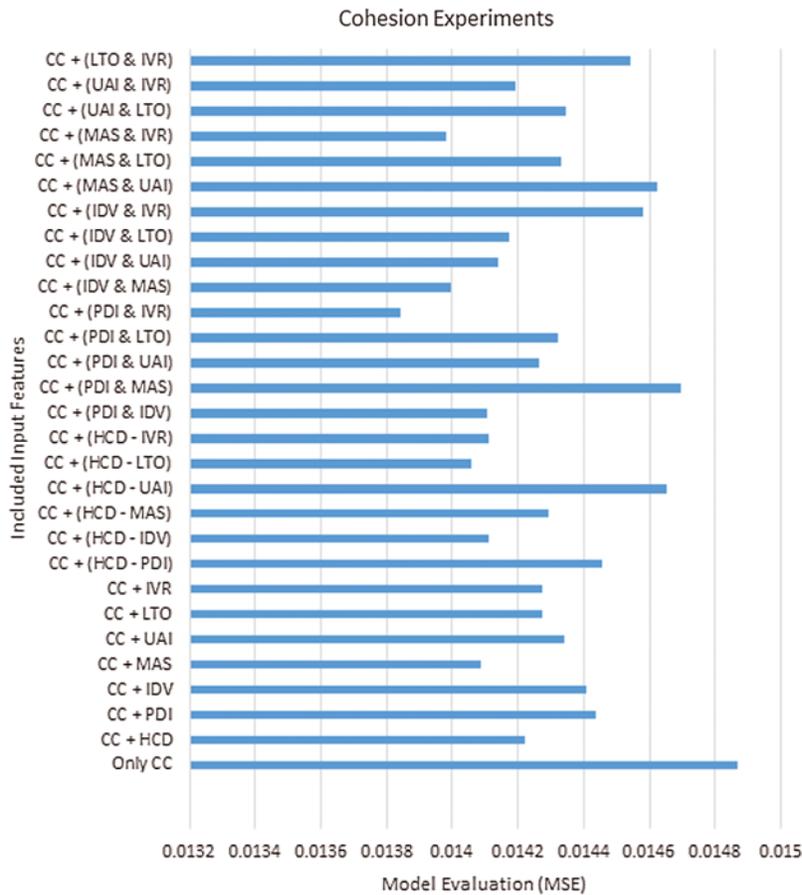


Figure 8: Model evaluation (MSE) of predicting cohesion experiments

Table 10: HCD statistically significant relations with cohesion targeted attribute

HCD predictors	HCD significant	HCD predictors	HCD significant
HCD-UAI	UAI	PDI & UAI	UAI
LTO	LTO	PDI & LTO	LTO
IVR	IVR	PDI & IVR	IVR

As discussed above, the results ensure that crowd characteristics are positively/negatively affected by crowds’ individual cultures in various ways and different cultural dimensions. Moreover, in addition to other crowd characteristics, all results in best-case predictions were related to the existence of one or more HCDs. Adding countries and scenarios to the dataset may provide an effective vision for defining those relations. By contrast, the CC-ANN learning model extracted some interrelations that were not discovered statistically, such as the interrelations related to collectivity’s targeted attribute. Furthermore, the flow rate and capacity calculations, space, and speed assigned to each crowd individual are different during AI-based CM systems’ regulatory plans. In addition to other crowd characteristics, such values depend on their culture.

However, our study had several limitations. The major limitation as per literature is the lack of CC similar and related works and resources. Thus, we suggest future works to provide more CC resources such as virtual environments, simulations, and feature extraction tools. Nevertheless, they are beyond the scope of this study. Moreover, this work could not provide a memory development cognitive process due to limited hardware capability and capacity. Consequently, we suggest adding a memory development process and adapting the Internet of Things technologies as future works to improve cognitive modeling. Finally, due to lack of time, the study could not optimize the other CC-ANN hyperparameters and could not provide more ANN designs and types such as Convolutional Neural Network (CNN) and deep learning. Accordingly, as future works, we suggest developing more ANN types and designs besides applying optimization of the other CC-ANN hyperparameters, which will reflect the proposed learning model's better performance.

These results warrant further investigation with experts from other disciplines, such as psychologists and sociologists, to conduct investigations on related crowd characteristics and psychological factors to have a clear vision of the mentioned results and an accurate definition of the discovered relations. Furthermore, further studies with additional experiments are required to analyze the effect of increasing/decreasing HCD scores in predicting targeted attributes and investigating the CC-ANN learning model with different CC types. Thus, additional cultural factors and countries must be studied along with various crowd events, types, densities, and scenarios to extend the research for big data. Finally, future investigations might be possible to examine the reflection of these results on real crowds with real-time monitoring.

6 Conclusion

The complexity level of crowds increases within cross-cultural events because each mass and event have different characteristics and challenges in different environments. Predicting crowd behavior by considering the cultural frame of reference of individuals is a challenging and unexplored problem. Therefore, the present study developed the CC-ANN learning model that employed the six HCD to predict physical (distance and speed) and social (collectivity and cohesion) crowd characteristics. Up to our knowledge, this work is a pioneering study of crowd behavior considering individuals' cultural backgrounds for the first time. Moreover, the proposed learning model developed based on ANN toward cognitivism covering three cognitive processes, perception, learning, and knowledge acquisition. Furthermore, the comparative analytical technique applied to the CC-ANN's model outcomes to extract interrelations between the crowd's HCD and physical and social characteristics. Additionally, to extract the most influential HCD (positively or negatively) in predicting crowd behavior. Consequently, our understanding of crowds' micro-level enhanced with increasing individuals' satisfaction level. The proposed CC-ANN learning model successfully predicts individuals' behavior within a group using four characteristics considering individuals' six HCD. It shows a low error of 0.048, 0.117, 0.010, and 0.014 Mean Squared Error in all prediction best cases. Moreover, it improves Crowd Management systems' performance within cross-cultural events. In most cases, the achieved accepted level of accuracy does not exceed 10 to 20 epochs. Furthermore, the CC-ANN learning model improves the performance by 90% and 93%, respectively, in some cases.

The major limitation as per literature is the lack of CC similar and related works and resources. Additionally, due to limited hardware capability and capacity, we could not provide a memory development cognitive process we could not provide a memory development cognitive process due to limited hardware capability and capacity. Moreover, the present work could not

optimize the other CC-ANN hyperparameters and could not provide more ANN designs and types due to lack of time. Further studies need to add a memory development process and adapting the Internet of Things, the IoT technologies to improve cognitive modeling. As well, we suggest applying optimization of the other CC-ANN hyperparameters in addition to developing more ANN designs and types. Psychologists and sociologists will be included in future work to obtain a comprehensive insight into the result and relations. Such a suggestion would require additional experiments to analyze the effect of increasing/decreasing HCD scores in predicting targeted attributes and different CC types. Moreover, big data and additional cultural factors and countries may improve the model using various crowd events, types, densities, and scenarios.

Acknowledgement: The authors acknowledge the Deanship of Scientific research (DSR) at King Abdulaziz University, Jeddah, for their technical and financial supports.

Funding Statement: This project is funded by the Deanship of Scientific research (DSR), King Abdulaziz University, Jeddah, under Grant No. (DF-593-165-1441). Therefore, the authors gratefully acknowledge the technical and financial support of the DSR.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] V. A. Sindagi, R. Yasarla and V. M. Patel, "Pushing the frontiers of unconstrained crowd counting: New dataset and benchmark method," in *Proc. of the IEEE Conf. on Computer Vision*, Seoul, Korea (South), pp. 1221–1231, 2019.
- [2] T. Li, H. Chang, M. Wang, B. Ni, R. Hong *et al.*, "Crowded scene analysis: A survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 3, pp. 367–386, 2014.
- [3] W. Li, V. Mahadevan and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 1, pp. 18–32, 2013.
- [4] M. Marsden, K. McGuinness, S. Little, C. E. Keogh and N. E. Connor, "People, penguins and petri dishes: Adapting object counting models to new visual domains and object types without forgetting," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 8070–8079, 2018.
- [5] M. Mohammed, S. Gunasekaran, S. Mostafa, A. Mustafa and M. Ghani, "Implementing an agent-based multi-natural language anti-spam model," in *Int. Symp. on Agent, Multi-Agent Systems and Robotics*, Putrajaya, Malaysia, IEEE, pp. 1–5, 2018.
- [6] S. Mostafa, A. Mustapha, A. Hazeem, S. Khaleefah and M. Mohammed, "An agent-based inference engine for efficient and reliable automated car failure diagnosis assistance," *IEEE Access*, vol. 6, pp. 8322–8331, 2018.
- [7] M. Mohammed, M. Ghani, S. Mostafa and D. Ibrahim, "Using scatter search algorithm in implementing examination timetabling problem," *Journal of Engineering and Applied Sciences*, vol. 12, pp. 4792–4800, 2017.
- [8] S. Mostafa, M. Ahmad, A. Mustapha and M. Mohammed, "Formulating layered adjustable autonomy for unmanned aerial vehicles," *International Journal of Intelligent Computing and Cybernetics*, vol. 10, no. 4, pp. 430–450, 2017.
- [9] S. Mostafa, S. Gunasekaran, A. Mustapha, M. Mohammed and W. Abdulllah, "Modelling an adjustable autonomous multi-agent internet of things system for elderly smart home," in *Int. Conf. on Applied Human Factors and Ergonomics*, Washington, USA, Springer, pp. 301–311, 2019.
- [10] G. Hofstede, "Dimensionalizing cultures: The Hofstede model in context," *Online Readings in Psychology and Culture*, vol. 2, no. 1, pp. 2307–0919, 2011.

- [11] G. Hofstede and M. H. Bond, "Hofstede's culture dimensions: An independent validation using rokeach's value survey," *Journal of Cross-Cultural Psychology*, vol. 15, no. 4, pp. 417–433, 1984.
- [12] R. Baskerville, "Hofstede never studied culture," *Accounting, Organizations and Society*, vol. 28, no. 1, pp. 1–14, 2003.
- [13] B. Zhou, X. Tang and X. Wang, "Measuring crowd collectiveness," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Portland, USA, pp. 3049–3056, 2013.
- [14] K. Abdulkareem, M. Mohammed, S. Gunasekaran, M. Al-Mhiqani, A. Mutlag *et al.*, "A review of fog computing and machine learning: Concepts, applications, challenges, and open issues," *IEEE Access*, vol. 7, pp. 153123–153140, 2019.
- [15] S. Albahli, A. Alsaqabi, F. Aldhubayi, H. T. Rauf, M. Arif *et al.*, "Predicting the type of crime: Intelligence gathering and crime analysis," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 2317–2341, 2021.
- [16] S. Wan and T. C. Lei, "A knowledge-based decision support system to analyze the debris-flow problems at Chen-yu-lan River," *Knowledge-Based Systems*, vol. 22, no. 8, pp. 580–588, 2009.
- [17] M. A. Mohammed, B. Al-Khateeb, A. N. Rashid, D. A. Ibrahim, M. K. Abd Ghani *et al.*, "Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images," *Computers & Electrical Engineering*, vol. 70, pp. 871–882, 2018.
- [18] N. Kanwar, A. K. Goswami and S. P. Mishra, "Design issues in artificial neural network (ANN)," in *4th Int. Conf. on Internet of Things: Smart Innovation and Usages*, Ghaziabad, India, IEEE, pp. 1–4, 2019.
- [19] N. Ketkar, "Introduction to Keras," in *Deep Learning with Python*, Berkeley, CA, United States: Apress, Springer, pp. 97–111, 2017.
- [20] H. Gayathri, P. M. Aparna and A. Verma, "A review of studies on understanding crowd dynamics in the context of crowd safety in mass religious gatherings," *International Journal of Disaster Risk Reduction*, vol. 25, no. 1, pp. 82–91, 2017.
- [21] Y. Chai, C. Miao, B. Sun, Y. Zheng and Q. Li, "Crowd science and engineering: Concept and research framework," *International Journal of Crowd Science*, vol. 1, no. 1, pp. 2–8, 2017.
- [22] S. Mohammadi, H. K. Galoogahi, A. Perina and V. Murino, "Physics-inspired models for detecting abnormal behaviors in crowded scenes," in *Group and Crowd Behavior for Computer Vision*, Cambridge, MA, United States: Academic Press, Elsevier, pp. 253–272, 2017.
- [23] S. Al Bosta, "Crowd management based on scientific research to prevent crowd panic and disasters," *Pedestrian and Evacuation Dynamics*, pp. 741–746, 2011. https://link.springer.com/chapter/10.1007/978-1-4419-9725-8_66.
- [24] A. A. Siddiqui and S. M. V. Gwynne, "Employing pedestrian observations in engineering analysis," *Safety Science*, vol. 50, no. 3, pp. 478–493, 2012.
- [25] F. Venuti, L. Bruno and N. Bellomo, "Crowd dynamics on a moving platform: Mathematical modelling and application to lively footbridges," *Mathematical and Computer Modelling*, vol. 45, no. 3–4, pp. 252–269, 2007.
- [26] N. Bellomo and C. Dogbe, "On the modeling of traffic and crowds: A survey of models, speculations, and perspectives," *SIAM Review*, vol. 53, no. 3, pp. 409–463, 2011.
- [27] M. A. Mohammed, K. H. Abdulkareem, B. Garcia-Zapirain, S. A. Mostafa, M. S. Maashi *et al.*, "A comprehensive investigation of machine learning feature extraction and classification methods for automated diagnosis of covid-19 based on x-ray images," *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3289–3310, 2021.
- [28] S. Yaseen, A. Al-Habaibeh, D. Su and F. Otham, "Real-time crowd density mapping using a novel sensory fusion model of infrared and visual systems," *Safety Science*, vol. 57, no. 6, pp. 313–325, 2013.
- [29] B. Zhan, D. N. Monekosso, P. Remagnino, S. A. Velastin and L. Q. Xu, "Crowd analysis: A survey," *Machine Vision and Applications*, vol. 19, no. 5–6, pp. 345–357, 2008.
- [30] L. Atallah and Z. G., "The use of pervasive sensing for behavior profiling: A survey," *Pervasive and Mobile Computing*, vol. 5, no. 5, pp. 447–464, 2009.

- [31] M. Wirz, E. Mitleton-Kelly, T. Franke, V. Camilleri, M. Montebello *et al.*, “Using mobile technology and a participatory sensing approach for crowd monitoring and management during large-scale mass gatherings,” in *Co-Evolution of Intelligent Socio-Technical Systems*, Berlin, Heidelberg, Germany: Springer, pp. 61–77, 2013.
- [32] C. Martella, M. Steen, A. Halteren, C. Conrado and J. Li, “Crowd textures as proximity graphs,” *IEEE Communications Magazine*, vol. 52, no. 1, pp. 114–121, 2014.
- [33] M. Wirz, T. Franke, D. Roggen, E. Mitleton-Kelly, P. Lukowicz *et al.*, “Inferring crowd conditions from pedestrians’ location traces for real-time crowd monitoring during city-scale mass gatherings,” in *21st Int. Workshop on Enabling Technologies: Infrastructure for Collaborative Enterprises*, Toulouse, France, IEEE, pp. 367–372, 2012.
- [34] Cinimod studio, “Peru National Football Stadium on Cinimodstudio,” 2011. [Online]. Available: <http://cinimodstudio.com/project/peru-national-football-stadium/>.
- [35] F. T. Illiyas, S. K. Mani, A. P. Pradeepkumar and K. Mohan, “Human stampedes during religious festivals: A comparative review of mass gathering emergencies in India,” *International Journal of Disaster Risk Reduction*, vol. 5, pp. 10–18, 2013.
- [36] T. X. Bui and S. R. Sankaran, “Design considerations for a virtual information center for humanitarian assistance/disaster relief using workflow modeling,” *Decision Support Systems*, vol. 31, no. 2, pp. 165–179, 2001.
- [37] A. Saleem, “Managing crowd at events and venues of mass. gathering: A guide for state government, local authorities, administrators and organizers,” *National Disaster Management Authority (Ndma), Government of India*, 2014. http://southwestgarohills.gov.in/pdf/ndmg_crowd.pdf.
- [38] L. Suchman, “Centers of coordination: A case and some themes,” *Discourse Tools and Reasoning*, vol. 160, pp. 41–62, 1997.
- [39] P. Luff and C. Heath, “The collaborative production of computer commands in command and control,” *International Journal of Human-Computer Studies*, vol. 52, no. 4, pp. 669–699, 2000.
- [40] R. Bentley, J. A. Hughes, D. Randall, T. Rodden, P. Sawyer *et al.*, “Ethnographically-informed systems design for air traffic control,” in *Proc. of the ACM Conf. on Computer-Supported Cooperative Work*, Toronto Ontario, Canada, pp. 123–129, 1992.
- [41] W. E. Mackay, A. L. Fayard, L. Frobert and L. Médini, “Reinventing the familiar: Exploring an augmented reality design space for air traffic control,” in *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems*, Los Angeles California, USA, pp. 558–565, 1998.
- [42] M. C. Davis, R. Challenger, D. N. Jayewardene and C. W. Clegg, “Advancing socio-technical systems thinking: A call for bravery,” *Applied Ergonomics*, vol. 45, no. 2, pp. 171–180, 2014.
- [43] M. Mohammed, K. Abdulkareem, A. Al-Waisy, S. Mostafa, S. Al-Fahdawi *et al.*, “Benchmarking methodology for selection of optimal covid-19 diagnostic model based on entropy and topsis methods,” *IEEE Access*, vol. 8, pp. 99115–99131, 2020.
- [44] S. Bandini, L. Manenti, S. Manzoni and F. Sartori, “A knowledge-based approach to crowd classification,” *Pedestrian and Evacuation Dynamics*, pp. 515–525, 2011. https://link.springer.com/chapter/10.1007%2F978-1-4419-9725-8_46.
- [45] A. Burney and T. Q. Syed, “Crowd video classification using convolutional neural networks,” in *Int. Conf. on Frontiers of Information Technology*, Islamabad, Pakistan, IEEE, pp. 247–251, 2016.
- [46] R. M. Favaretto, L. Dihl, R. Barreto and S. R. Musse, “Using group behaviors to detect hofstede cultural dimensions,” in *IEEE Int. Conf. on Image Processing*, Phoenix, AZ, USA, pp. 2936–2940, 2016.
- [47] D. Lala, S. Thovuttikul and T. Nishida, “Towards a virtual environment for capturing behavior in cultural crowds,” in *Sixth Int. Conf. on Digital Information Management*, Melbourne, Australia, IEEE, pp. 310–315, 2011.
- [48] G. Hofstede, “Culture’s consequences: Comparing values, behaviors, institutions and organizations across nations,” *Sage Publications*, vol. 48, no. 1, pp. 127–131, 2003.
- [49] A. Hazra, “Using the confidence interval confidently,” *Journal of Thoracic Disease*, vol. 9, no. 10, pp. 4125–4129, 2017.

- [50] I. J. Hussein, M. A. Burhanuddin, M. A. Mohammed, M. Elhoseny, B. Garcia-Zapirain *et al.*, “Fully automatic segmentation of gynaecological abnormality using a new viola-jones model,” *Computers, Materials & Continua*, vol. 66, no. 3, pp. 3161–3182, 2021.
- [51] M. Mohammed, K. Abdulkareem, S. Mostafa, M. Ghani, M. Maashi *et al.*, “Voice pathology detection and classification using convolutional neural network model,” *Applied Sciences*, vol. 10, no. 11, pp. 3723, 2020.
- [52] A. A. Mutlag, M. k. Ghani, M. A. Mohammed, M. Maashi S. A. Mostafa *et al.*, “MAFC: Multi-agent fog computing model for healthcare critical tasks management,” *Sensors*, vol. 20, no. 7, pp. 1853, 2020.
- [53] M. K. Ghani, M. A. Mohamed, S. A. Mostafa, A. Mustapha, H. Aman *et al.*, “The design of flexible telemedicine framework for healthcare big data,” *International Journal of Engineering & Technology*, vol. 7, no. 3. 20, pp. 461–468, 2018.
- [54] K. H. Abdulkareem, M. A. Mohammed, A. Salim, M. Arif, O. Geman *et al.*, “Realizing an effective covid-19 diagnosis system based on machine learning and iot in smart hospital environment,” *IEEE Internet of Things Journal*, pp. 1, 2021.
- [55] J. Schmidhuber, “A fixed size storage $O(n^3)$ time complexity learning algorithm for fully recurrent continually running networks,” *Neural Computation*, vol. 4, no. 2, pp. 243–248, 1992.
- [56] M. Awan, M. Rahim, N. Salim, M. Mohammed, B. Garcia-Zapirain *et al.*, “Efficient detection of knee anterior cruciate ligament from magnetic resonance imaging using deep learning approach,” *Diagnostics*, vol. 11, no. 1, pp. 105, 2021.