

SURVEY

A Survey on Intelligent Gesture Recognition Techniques

JUAN JESUS OJEDA-CASTELO¹, MARIA DE LAS MERCEDES CAPOBIANCO-URIARTE²,
JOSE ANTONIO PIEDRA-FERNANDEZ¹, AND ROSA AYALA¹

¹Department of Informatics, University of Almeria, 04120 Almeria, Spain

²Department of Economics and Business, University of Almeria, 04120 Almeria, Spain

Corresponding author: Juan Jesus Ojeda-Castelo (juanje.ojeda@ual.es)

This work has been funded by the EU ERDF and the Andalusian Government (Spain) under the project I+D UrbanITA (ref. P20_00809).

ABSTRACT Gesture recognition is an ideal means of interaction because it allows users not to have to make contact with any surface, which is a safe and hygienic means, especially in the pandemic situation that is occurring worldwide. However, gesture recognition is not a new discipline and it has been researched for many years but this type of interaction has not succeeded in replacing the keyboard and mouse. It is very useful to know about the advances that are being made with artificial intelligence in gesture recognition to be able to perform a more robust and reliable gesture recognition with a low response time. As it is, deep learning is being integrated into various areas to increase improvement in performance and one such area is artificial intelligence. In this way, there is the possibility that in the future the recognition of gestures will be a viable option as a means of daily interaction for the user and the main objective of this paper is to contribute to that process. For this reason, this study has analyzed 571 papers related to gesture recognition and artificial intelligence. This analysis has extracted relevant information related to scientific production, such as the most productive authors and journals or the most pertinent articles on the subject. Furthermore, we have developed our own model, which shows the relationship between the types of gesture recognition and the artificial intelligence techniques that have been applied for this task.

INDEX TERMS Artificial intelligence, deep learning, gesture recognition, machine learning.

I. INTRODUCTION

Gesture recognition is a highly focused area of research where the user can interact with computerized systems, such as controlling a computer [1], a home automation [2] or an automobile [3]. This field can be classified depending on the characteristics or objectives that imply such recognition. Depending on the part of the body that is being recognized, you can differentiate between hand [4], body [5] and face [6]. However, if the recognition involves a movement then it will be dynamic whereas if it is a pose the gesture will be static [7]. Finally, depending on the method used to recognize the gesture, this recognition can be classified into vision-based [8] or sensor-based [9].

The associate editor coordinating the review of this manuscript and approving it for publication was Shovan Barma¹.

In recent years, a multitude of devices have been created for gesture recognition such as Microsoft Kinect [10], Leap Motion [11], Intel Real Sense [12], Myo [12], Camboard Pico [13] and so on. For instance, Tolentino *et al.* [14] developed a system to detect emergency situations through gesture recognition. Microsoft Kinect was integrated in order to obtain the location of the different joints in the user's body and mathematical information such as the distance or the angle between the distinct joints. Thus, it is possible to identify 14 gestures that will be the input of the decision tree in order to make a decision whether the user is in one of the emergency situation categories: medical, life-threatening condition, or disaster. In [15] hand gestures are recognized (instead of body gestures in the previous work) to identify American sign language. In this approach, the combination of Leap Motion and Hidden Markov Models (HMM) in a Virtual

Reality environment is applied to classify 24 gestures. The Leap Motion device is necessary to extract the 3D data from 11 joints that this sensor is able to recognize and the aim of the HMM is to classify the gestures by means of the feature vector sequence that contains the distance and angle between the different joints.

However, gesture recognition is not a strange or even novel technique for users due to the day-to-day use of mobiles. Mobile phones are used to check mail, surf the internet, interact with social networks, take photos and even play games. These mobile devices have become a fundamental part of users' lives because it is capable of executing all the tasks that can be carried out with a computer but with very small dimensions so they can be with the users 24 hours a day regardless of their locations or circumstances. The most relevant thing is that these tools are handled with gesture recognition. From tapping the device to opening an application to dragging their finger to reduce the size of an image or change the screen, all these movements are gestures that people make daily several times a day. Furthermore, these devices are not only used by adults, but the immersion of these tools in the daily routine has also enabled children to use them at an early age and to perform the gestures that involve the use of these devices in a natural and intuitive way [16].

Therefore, the use of gestures to replace traditional forms of interaction is not a surreal idea but has not been implemented well except in mobile devices mainly due to the fact that the precision and response time are not optimal to make the user feel confident and comfortable with this type of interaction. Hence, fields of research such as Artificial Intelligence (AI) have been used to improve these parameters and make gesture recognition the predominant means of interaction.

AI consists of the creation of systems or machines that have the objective of performing tasks in the same way that a human would. Its fields known as Machine Learning (ML) and Deep Learning (DL) are applied to numerous scientific fields such as Medicine [17], Natural Language Processing [18], Information Retrieval [19], Cybersecurity [20], Computer Vision [21], Internet-of-Things (IoT) [22], and so forth. Moreover, they have been applied to gesture recognition to achieve better results since the necessary performance for its general use has not been achieved through robust models that have been created in vision-based gesture recognition. In [23], a new Convolutional Neural Network (CNN) is designed to recognize 3D hand gestures through the hand joints' location without using the depth distance. The novelty of this neural network lies in the parallel processing and the use of residual connections for each signal. This approach leads to 91.28% accuracy in the dataset named DHG which contains a collection of dynamic gestures. *deepGesture* [24] is a deep learning framework that applies convolutional and gate recurrent unit (GRU), feeding the recurrent neural network (RNN) with data from a gyroscope and an accelerometer sensor. This methodology consists of four parts: the input

layer that obtains the data from the sensors, the convolutional layers which extract the features, the GRU layers that process the sequential information, and the fully connected layer which calculates the result. The reason for using this method is because it is faster and only needs a small amount of data compared to other methods such as long-short term memory (LSTM). In [25], a 3D ResNet has been created to extract spatiotemporal features from the dataset to combine it with a memory module in order to make a one-shot learning gesture recognition. Furthermore, another contribution of this work is the creation of a dataset with 3045 videos about hand gestures since it is sometimes difficult to find a dataset that fits the needs of the problem under study. In [26], a model to recognize hand gestures with forearm electromyography (EMG) is described. The EMG signals are produced by the device Myo armband as input for the system and the algorithms k-nearest neighbor and dynamic time warping (DTW) is involved in the classification process. The goal of the k-nearest neighbor algorithm is to estimate the conditional probabilities in the feature matrix while the DTW is in charge of calculating the distance function with the Manhattan distance. The particularity of this work resides in the fact that the authors affirm that this model is capable of learning any hand gesture through training.

The fact that this form of interaction is more robust and reliable will be a necessity in the times that pass since now cities are becoming Smart Cities, which opens a world of possibilities and facilities to the user with the immense volume of data to handle. Among these possibilities is that users can interact with the sensors on the same street, in a smart building [27] or in a smart classroom [28]. In fact, Ma et al. [29] have developed an algorithm to recognize the gestures of police officers when they control traffic using a spatiotemporal convolution neural network (ST-CNN). This proposal has been created in a virtual geographic interactive environment that could later be applied to a real environment in a smart city.

Finally, the motivation for this study is preceded by the fact that we wanted to develop a method based on natural interaction in order to control a computer system, since it would be a more natural and faster means of interaction for the user, with the additional advantage of which could also be used by people with special needs [30]. Among all the natural interaction modes that exist (gesture recognition, speech recognition, brain-computer interface, and so on), preference was shown for gesture recognition because it is intuitive, it does not interrupt the user's activity [31] and it does not require any additional devices since the command is performed only through the movement of a part of the body [32]. Moreover, the participants of this study [33] systematically preferred natural interaction input over mouse input and performed better regarding some parameters such as speed and accuracy. It is known that AI techniques got satisfactory results in the area of gesture recognition [34]. For this reason, there was a need to conduct a study on AI techniques applied to this type of interaction.

The contributions of the present study can be summarized as the following:

- The analysis of relevant papers from 2000 to 2020 to extract the most significant indicators.
- Numerous data related to scientific production such as the number of authors, journals, and institutions, among others, have been analyzed.
- The elaboration of our own model that expresses the relationship between ML and DL techniques and types of gesture recognition.
- The authors are not aware of the existence of a paper with the same characteristics or that analyzes the same parameters that address this issue.

The remainder of the paper is organized as follows: Section II describes the fundamental concepts of gesture recognition and AI and the most recent works related to these topics. Section III presents the methodology of Gesture Recognition. Section IV shows the results obtained through carrying out an evaluation and analyzing these results. Section V summarizes the conclusions and discusses future work.

II. BACKGROUND

In this section, the main topics of this study will be described, which are gesture recognition and AI.

A. GESTURE RECOGNITION

Gestures are a necessary means of communication between people and are defined in the category of non-verbal communication. This fact has led to this type of communication being transferred to interaction with computers, also because it is a more natural and faster form of communication. However, gestures can be performed with different parts of the body and these are mainly categorized into facial gestures, hand gestures, and body gestures.

1) FACIAL EXPRESSIONS

Facial expressions are characterized because they are capable of expressing a person's emotions during the communication process. The facial expression recognition process consists of three stages: preprocessing, feature extraction, and classification [35].

- **Preprocessing:** In this process, different transformations are carried out on the frames to improve the quality of the image. The most used techniques are cropping and scaling on the image of the face, normalization and histogram equalization to modify the illumination and the Viola-Jones algorithm is used to determine the size and location within the image.
- **Feature extraction:** The objective of this method is to identify the most relevant characteristics/descriptors that allow defining a face. Among the techniques to be highlighted are:
 - **Supervised Descent Method** which extracts the main positions of the face and estimates the distance between the different elements of it.

- **Line Edge Map** It is used to extract useful information from the image edge representation feature and is widely applied in face recognition because it detects lines as features from a face edge map.
- **Principal Component Analysis** it is applied to extract fundamentally global and low dimensional features.
- **Classification:** The most demanded classification techniques for facial expressions, as can be seen later in this work, are Support Vector Machine and Convolutional Neural Networks.

The first step is the identification of the central focus of this study (NUI - Natural User Interaction), in order to show the information relating to the scientific output (content analysis) and the analysis of the co-occurrence of keywords in this field of research (content analysis). The second step is the selection of the database. Bearing in mind that the results of the analysis may vary dep

In this study, facial expressions were detected in 64% of the analyzed papers. 89% of the papers that recognize facial expressions apply DL techniques, of which CNN accounts for 85.1%, LSTM 4.1%, RNN 3.4% and other techniques 7.4%. However, 42% of the papers use ML techniques (the percentages of DL and ML techniques are greater than 100% because in some investigations one type of technique was not used exclusively and it is applicable to the rest of the recognition types), of which SVM accounts for 49.6%, K-Means and K-Nearest Neighbors 14.4%, Artificial Neural Networks 11.2% and other techniques 24.8%.

2) HAND GESTURES

The hands are the part of the body that is most frequently related to the recognition of gestures since we use them, for example, to greet another person. In the recognition of hand gestures, there are two general perspectives: recognition based on vision and recognition based on sensors [36].

- **Recognition based on vision:** This type of recognition is the most used because it is the least invasive for the user since it is based exclusively on the use of one or more cameras. In terms of vision, basically, two types of recognition are usually distinguished: 3D model-based and appearance-based. 3D model-based approaches work with depth in order to obtain the 3D position of different points of the hand and thus track those points to detect gestures in relation to the relative position of other points that are taken as reference. On the other hand, appearance-based approaches make use of various properties of the hand to recognize gestures. These properties can be the color of the hand to detect movement through markers or some geometric characteristics such as perimeter, convexity, or elongation.
- **Recognition based on sensors:** This version makes use of sensors that are usually integrated into a glove and detect the position of the hand at all times. These sensors can be of different types: mechanical, magnetic, and

ultrasonic, among others. The most used technique to recognize hand gestures based on sensors is HMM.

In this study, it has been determined that 18% of the analyzed papers are related to hand gesture recognition. 95% of the papers that recognize facial expressions apply DL techniques, of which CNN accounts for 75%, RNN is 12.5%, LSTM is 7.5% and other techniques are 5%. However, 44% of the papers use ML techniques, of which SVM corresponds to 45.9%, PCA has 13.5% and Artificial Neural Networks has 10.8% as well as Decision Trees, being others techniques 29.8%.

3) BODY GESTURES

Facial gestures and hand gestures are usually the most used, but there is another type of gesture that should also be highlighted: body gestures. Body gestures refer to movements made with large muscle groups and are called gross motor skills. Unlike hand gestures, which are more precise and short-range movements, body gestures tend to be more indeterminate and long-distance gestures.

In the process of recognizing these gestures, it is important to take into account the type of model chosen to make the abstraction of the body. In general, two types of models are used: part-based models and kinematic models [37].

Body pose recognition involves three steps: background subtraction, body pose detection, and tracking. Body pose detection and body pose tracking are the most interesting methods because of the topic of this study. In the body pose detection, the posture of the human being is analyzed and the orientation of the different parts of the body is identified. Currently, ML and DL techniques are mainly applied in this task, SVMs are used to classify body positions from features extracted from a [38] dataset, CNNs are based on 2D features to detect the body [39] while the LSTM are included especially when you want to perform detection on videos since this type of network takes into account the time variable [40]. On the other hand, in body pose tracking, a series of parameters such as the position or the shape is estimated in each frame. Some techniques such as Expectation Maximization and Optical flow [41] are used to detect the human body in a frame, from the pixels located in different parts of the body or its contour. Also, for body pose tracking there are tools like MediaPipe [42] that make this process easier to implement.

In this study, it has been determined that 13% of the papers analyzed are related to body gesture recognition, while the remaining 5% correspond to other gestures. Regarding body gestures, 93% of the investigations that recognize this type of gesture use DL techniques while 37% use ML techniques. Among the DL techniques is CNN with 35.1%, RNN with 22.8%, and LSTM with 15.7%, with other techniques with 26.4%. Regarding ML techniques, there is SVM with 26.1%, K-Means and K-Nearest Neighbors at 21.7%, Artificial Neural Networks at 21.7%, and PCA 13%, where the remaining 17.5% belongs to other techniques.

4) APPLICATIONS

The main applications associated with gesture recognition are shown in Figure 1.

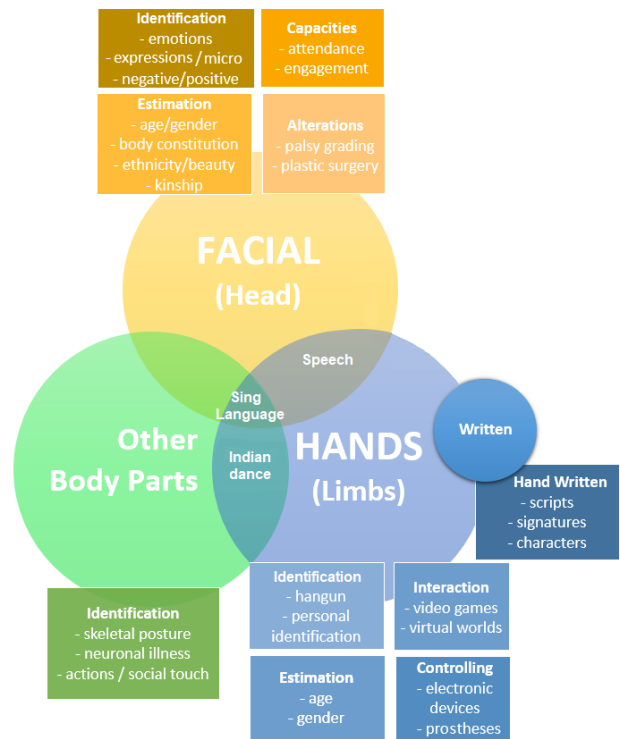


FIGURE 1. Main applications in gesture recognition.

The identified applications were classified as combined, those that are present in most of the main types of recognition (facial, hands, or other body parts), and specific, exclusive to each type of recognition. According to the analysis carried out on the universe of documents analyzed, facial recognition is the technique that presents the most applications together with hand recognition. Being the recognition of sign language the most complex application identified since it is a visual language that draws on all the main types of recognition through fingerspellings, facial and body gestures in different languages, such as Arabic [43], American [44], Chinese [45], Korean [46] and Indian [39]. Due to the difficulty of sign language recognition, device-based studies have been conducted to improve accuracy [47], [48], [49]. However, to achieve the speech recognition [50], [51] is needed for both language and speaker identification. Finally, recognition of Indian dance performance needs both body pose recognition and hand gesture recognition [52], [53].

Within the specific applications, the identification and estimation applications stand out, which are based on the three surveys studied but applied individually. The identify action uses facial recognition and is used to identify expressions [54], micro-expressions [55], emotions [56] and negative expression [57]. Hand recognition is used to identify weapons [58], [59] and people [60]. Finally, recognition

focused on the rest of body parts can be identified skeletal posture [61], neuronal illness or social actions [62], as social touch [63]. The estimation of age and/or gender can be done through facial recognition [64], [65] or through hand gestures [66], [67]. Apps estimating ethnicity [68], beauty [69] and body constitution [70] were also detected exclusively through facial recognition.

Finally, applications that exclusively use a single type of recognition were found. Facial recognition is used to identify capabilities such as attendance [71] or engagement [72] or to detect alterations such as palsy grading [73] or plastic surgery [74]. In the recognition of hand gestures, the applications of the skill of handwriting (scripts, characters, and signatures) stand out above the rest [75], [76], [77]. Man-machine interaction applications and their control use hand gesture recognition. For example, interactions for video games and virtual worlds [78] and control actions in the use of electronic devices [79] and specifically prosthetic [80].

B. ARTIFICIAL INTELLIGENCE

AI is defined as a discipline that seeks to imitate human behavior, such as being able to make decisions or differentiate between different objects. The most generic types of AI that are usually distinguished are weak AI and Artificial General Intelligence (AGI) [81]. The weak AI refers to the resolution of specific problems, such as the example of when a machine beat the chess expert Kasparov [82], but this algorithm is only dedicated to finding the solution of this task and could not solve others chores. However, the AGI would be more similar to how we act and this type of intelligence would be more adaptable and flexible, with which it could face situations independent of those for which it was created.

Some of the disciplines using AI are in education, where robots have been created to enhance student learning [83]; in agriculture in order to warn about adverse environmental situations that could ruin the harvest and improve agricultural yields [84] or in health with the use of virtual assistants that can make a diagnosis based on the symptoms that the patient has and thus serve a greater number of patients [85].

1) MACHINE LEARNING

ML is a branch of AI that allows machines to learn through statistical calculations. This discipline can be found today in everyday applications for users such as Netflix, Siri, or Gmail automated responses. The types of ML are supervised learning, unsupervised learning, and reinforcement learning.

- **Supervised learning:** in this type of learning, previous information is provided that helps in the predictions made by the system. An example would be if you want to classify animals, the input images of the algorithm will be labeled. In this way, the image can be associated in advance with the name of the animal.
- **Unsupervised learning:** in this case, you will not have prior information to help you associate the images and you will have to find patterns to help you organize them in some way. An example would be to provide the

algorithm with a series of flower images and have it group them into categories. The algorithm will identify the characteristics and will group in the same category the elements that have similar characteristics, this technique is known as clustering.

- **Reinforcement learning:** This model learns through trial and error, where correct actions will be reinforced to successfully complete the task. An example of this type of learning is autonomous car navigation systems.

2) DEEP LEARNING

Deep Learning has the objective of developing algorithms that are assimilated into the functioning of the brain. This discipline is so called because it uses deep neural networks that are made up of many layers and that have provided better results than other techniques, although it also consumes a lot of resources. These deep neural networks consist of a system of multiple layers each. These layers act as a filter, going from the most general elements to the most specific, and can iterate as many times as necessary on the data to analyze this data and draw conclusions. Therefore, these neural networks can identify patterns and classify different types of information.

The types of neural networks that are usually used in DL are:

- **Convolutional Neural Network (CNN):** These neural networks were designed for the processing of structured matrices such as images. This means that they can classify images based on patterns that appear in them, such as lines or circles. These neural networks use several convolutional layers that can be stacked on top of each other and thus recognize more complex shapes.
- **Recurrent Neural Networks (RNN):** This architecture uses sequential data or time series data. The difference between this neural network and the others is that it has an “artificial memory” generated by the fact that the neurons that make up its architecture receive the input from the previous layer which, together with its own output from the previous time instant, generates the output of this neuron. This operation allows you to make future predictions from historical data.
- **Generative Adversarial Network (GAN):** The philosophy of this technique is to use two artificial neural networks and oppose each other in order to create new content or synthetic data. One of the networks generates and the other works as a discriminator. The discriminating network has been trained to recognize real content and censors the other network’s content in order to generate content that looks real.

3) TRANSFER LEARNING

TL is a technique that has provided numerous advantages in general but also to gesture recognition. This statement is based on the advantages offered by this method, where the main ones are the fact that it is possible to perform a task with much less data than with the traditional procedure, it is faster in the execution of the training because it is not necessary to

train the complete model but only a part of it and its learning rate is usually higher because it has been previously trained in a similar task [86], [87], [88]. In this study, the occurrences of TL have been determined in the papers that comprise this work, where this technique has been used mostly in DL, especially in CNN, with a percentage of 89% of all papers that have applied TL and the RNN reaches 4%. Some of the most relevant studies in the field of TL applied to gesture recognition would be this framework [89] which is capable of identifying a series of gestures through a siamese architecture based on recurrent convolutional networks. In this architecture, the spatial characteristics are first learned from CNN and later the temporal characteristics are learned with a Recurrent neural network. The results of the experiments support the efficacy of this method, obtaining 89.5% precision. Another relevant work in this regard is WiADG [90] which is a gesture recognition system that works independently of any device since it uses the WiFi signal to perform this recognition. This system has the particularity that it can identify gestures in a dynamic environment with the use of a supervised adversarial domain adaptation, through the IoT devices found in it. These devices are the source of the channel state information that is necessary to analyze and detect the different gestures made by the user. The results show that this system achieves an accuracy of 98% in the gesture recognition process.

4) APPLICATIONS

In this section, the most recent studies on ML and DL will be described. In relation to ML, there is this study [91], in which a framework has been created that detects if a person has COVID-19 by answering a series of questions. In this way, hospitals are not so saturated and a larger number of the population can be treated. The ML part of this framework is formed by the gradient-boosting model, which has been trained with a total of 51,831 records of patients with and without COVID-19. This work [92] has the objective of predicting the concentration of polluted air in smart cities. In carrying out this task, the Multi-Layer Perceptron and Random Forest algorithms have been used, which have analyzed the air data from some smart cities. The results of the experiments have determined that Random Forest has obtained a higher precision than Multi-Layer Perceptron. A topic that is booming at the moment is Quantum Computing and this field has also been involved in AI with the name of Quantum Machine Learning, for example in [93] the authors propose a methodology to evaluate the advantages offered by AI. quantum computing in learning tasks. In this case, a set of experiments are carried out with classical models and with quantum models to determine in which situations Quantum Machine Learning can be an advantage over traditional models.

On the other hand, as for DL in [94] the goal is to do IoT image identification. To achieve this goal, a series of images are extracted from the IoT sensors to introduce them as input to the Principal Component Analysis (PCA) technique, which will be in charge of extracting features. These features will be trained on CNN in order to recognize images in the context

of IoT. In women, breast cancer is the second most common type of cancer and for this reason in this study [95] a methodology based on DL has been developed to improve the diagnosis of this type of cancer. This diagnosis is made through a CNN, which extracted characteristics from the input data of the multiparametric magnetic resonance imaging and then a support vector machine (SVM) has been trained with the characteristics extracted from the CNN to detect if the tumor is malignant or benign. In [96] a DL algorithm has been integrated to recognize human emotions through electroencephalogram (EEG) signals. The process has been divided into several stages, where the EEG signal has first been broken down into rhythms. Next, higher-order statistics have been used to analyze the signal in a higher dimensional space. However, this causes repeated information to be generated, which requires the use of a dimensionality reduction technique. Once the redundant information has been eliminated, the recurrent neural network called long-short term memory will be fed with the normalized attributes. This neural network will be in charge of discriminating the signals and thus classifying the emotions.

C. GESTURE RECOGNITION AND ARTIFICIAL INTELLIGENCE

In this section, unlike the previous ones, works that have combined gesture recognition with AI will be presented. In this case, the aspects that have been considered most outstanding will be described, which are the most significant approaches that have been created, the original and innovative architectures that have been created recently, and the cases that have stood out for their reliability.

1) RELEVANT APPROACHES

In terms of approaches, 24% corresponds to hybrid models, which are widely used because the objective is to make the most of the advantages of the methods involved and in this way achieve greater performance and accuracy, and they are also more flexible and robust than the non-hybrid approaches [97]. The most developed hybrid approach has been the combination of SVM and CNN, which accounts for 48% of the total of these hybrid models. An example of these methodologies is HandSense [98], which consists of a system that has the objective of recognizing dynamic gestures with dependent hands in various 3D CNNs that are dedicated to extracting spatio-temporal characteristics so that SVM can recognize the different gestures through characteristics. In [99] another work is described where dynamic gestures have been recognized by developing a 3D separable convolutional neural network so that this compact model can be used in augmented reality glasses. However, the separation made during the 3D convolution process has caused problems with the gradient that the authors have solved including the skip connection method and the layer-wise learning rate. This approach has proven to be more effective than other 3D CNN models with similar characteristics trained for the classification of dynamic hand gestures.

However, CNN is the most widely used popular technique, which has a percentage of 56% and has been applied mainly to recognize facial expressions, as in this study [100] where images are pre-processed before inserting them as input on CNN for learning and detecting facial expressions. Although in the process there was an issue that was the absence of enough data for the learning of CNN the authors solved it through data augmentation. RNN is another of the most used techniques, even though its percentage is lower than CNN, with 11% as in the following work [101] whose purpose is the recognition of certain actions by applying RNN. The recognition process is based on an RNN that receives a series of time-of-flight measurements from which it obtains a prediction by analyzing the complete temporal sequence. This system is capable of recognizing with high precision the following actions: walking forward, walking reverse, sitting down, standing up, and waving a hand.

2) NOVEL PROPOSALS

Despite these being the most popular approaches, we must also highlight new proposals such as pyramidal 3DCNN [102] or capsule network [103]. The pyramid 3DCNN architecture is composed of a pyramid input, pyramid fusion, multi-modality fusion, and two 3DCNNs. The pyramid input has the function of segmenting each of the videos it receives as input because they have different lengths and uses uniform sampling with temporal jitter to obtain the pyramid input. Next, this input feeds the pyramid fusion phase in order to join the features of the pyramid input, and finally, because the networks had trained the data with RGB and depth independently, the multi-modality fusion has to merge both modalities in order to enable the gesture recognition. The capsule network is made up of 5 convolutional layers, the primary capsule layers that are made up of 3872 capsules of 8 dimensions, and the layer named *GestureCaps* that contains 5 capsules of 16 dimensions. Each capsule is a set of neurons and has an activation probability and a matrix pose. The operation of this architecture is that the 5 convolutional layers are used to reduce the dimension and then the capsule layers have a squashing function to obtain the output of the capsule that is multidimensional. In this way, the output feeds the layer *GestureCaps* whose output will affect the decision of the output layer applied by softmax to obtain the corresponding label in the gesture classification.

3) RELIABLE CASES

Albeit, it is significant to learn about the most novel approaches but they are not useful if they are not reliable. That is the reason why a few of the most dependable cases are going to be described. In this work, [104] facial expressions and body gestures are recognized using a model that combines CNN, long short-term memory, and principal component analysis (PCA) to extract spatio-temporal features. This proposal has reached an accuracy of 99.57% on a face and body dataset (FABO) [105]. In this study, [106] a hybrid model has been created which combines the features

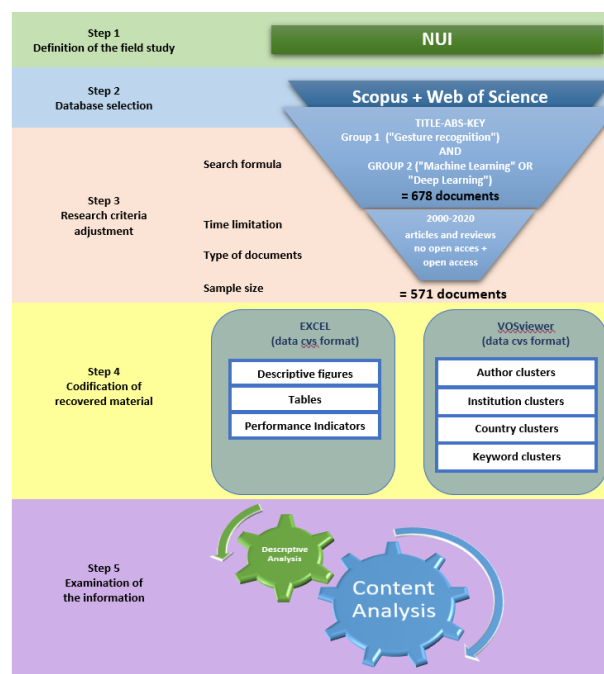


FIGURE 2. Bibliometric methodology flow chart.

extracted from the scale-invariant feature transform (SIFT) with the CNN features to recognize facial expressions. This combination has achieved a precision of 94.82% in the Cohn-Kanade (CK+) database [107]. In [108] a total of 27 gestures have been recorded with a linear optical sensor to detect hand gestures. Three types of gesture representation have been recorded: raw, simple features, and high-level features; which have been the input data of the RNN to classify hand gestures. Of these three types of representation, raw data has been the one that has obtained the highest hit rate and the average accuracy has been 96.89%. This job [109] is intended to make static gesture recognition applied to the sign language. An improved convolutional neural network (ICNN) has been developed and dropout and L2 regularization has been applied to avoid overfitting. The efficiency of this method for the recognition of sign language is reflected in the classification accuracy obtained of 99.96%.

III. MATERIALS AND METHODS

The identification of the key elements of the research field of gesture recognition and AI through machine learning and deep learning tools follows the bibliometric analysis technique that shows relevant information about authors, institutions, documents and keywords [110], [111]. This bibliometric study is a five-step process:

- 1) Definition of the research field.
- 2) Selection of the database.
- 3) Setting of research criteria.
- 4) Coding of the retrieved material.
- 5) Examination of the information.

In this way, the process becomes clearer and could be reproducible (see Figure 2).

The first step is the identification of the central focus of this study (NUI - Natural User Interaction), in order to show the information relating to the scientific output (continent analysis) and the analysis of the co-occurrence of keywords in this field of research (content analysis). The second step is the selection of the database. Bearing in mind that the results of the analysis may vary depending on the database selected, and in line with [112], this study uses the two most widely used bibliometric data sources, namely Web of Science (WoS, produced by Clarivate Analytics) and Scopus (created by Elsevier). Although the Google search engine could offer additional coverage to WoS and Scopus, it has certain problems. First, it lists a large number of non-academic sources, including grey literature that is not peer-reviewed [113]. Second, the search algorithm is not reproducible, as the results are displayed on the basis of previous searches and interactions [114]. Third, it is difficult to use for large-scale analysis [115]. Therefore, the limitations above have deterred us from including it in our analysis.

Once the databases have been selected, the next step is the adjustment of the research criteria. At this stage, the research criteria are set with Boolean operators to obtain an accurate search and to facilitate the capture of large database. The parameters used to retrieve the search were: TITLE-ABS-KEY (“gestur* recognition” OR “hand* recognition” OR “bod* recognition” OR “fac* recognition” OR “hand* gestur*” OR “bod* ges- tur*” OR “fac* gestur*” OR “gestur* interact*” OR “gestur* based interact*” OR “hand* interact” OR “bod* interact*” OR “fac* interact*” OR “gestur* detect*” OR “hand* de- tect*” OR “bod* detect*” OR “fac* detect*” OR “gestur* model*” OR “gestur* classif*”) AND (“machine learning” OR “support vector machine” OR “decision tree” OR “k- nearest neighbor” OR “naive bayes” OR “random forest” OR “hidden markov model” OR “dynamic time warping” OR “bayesian network” OR “k-means” OR “artificial neural network” OR “deep learning” OR “convolutional neural network” OR “recurrent neural network” OR “long short-term memory network” OR “deep belief Network”) from the title, abstract and keywords. The search was limited to the period 2000-2020. The first paper on this topic is the article by Ng C.W. and Ranganath S. entitled “Gesture recognition via pose classification”, published in 2000. This paper describes how a gesture recognition system can be trained by estimating hand poses [116]. This strategy, which includes hand poses in the training of the system, reduces training and recognition times by offering the possibility of applying the system in real time. However, these authors were not only the pioneers in combining both fields of research in a single article, but later in 2002, they published “Hand pose training in real time”. In the same year, they also published “Real-time gesture recognition system and application”, in which they used hand poses to create a vision-based system which allowed objects in an interface to be controlled by gestures with a 91.9% hit rate [117].

TABLE 1. Summary of data used.

Data	Gesture Recognition + Artificial Intelligence
Number of articles	571
Number of citations	9,366
Number of keywords	2,897
Number of journals	244
Number of authors	2,167
Number of institutions	721
Number of countries	66
Study time	2000 - 2020
Data sources	Scopus - Web of Science

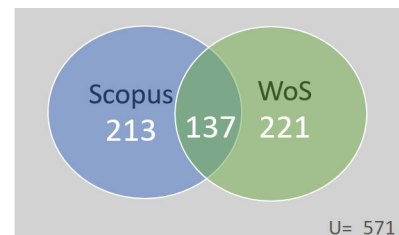


FIGURE 3. Distribution of the articles published in the databases used.

The search in both databases (Scopus + Web of science) was carried out at the end of September 2020. In terms of inclusion and exclusion criteria, only articles, books and book chapters were considered, including open access documents [118].

Table 1 summarises and quantifies in global form the content variables that will be analyzed in more depth later on. The number of documents found exclusively in WoS was 213 and 221 in Scopus, although there were 137 in common. Therefore, the final sample consisted of 571 documents (Figure 3).

The fourth step is the coding of the retrieved material, which was downloaded in csv format and coded using Excel (version 2013) and VOSviewer (version 1.6.9). The data were pre-processed for further analysis. At first, duplicate documents in both databases were removed. Second, the abstract and title of each document were checked to ensure that they met the search criteria. Third, documents with missing information were corrected.

Finally, the last step is the examination of the information. This phase is carried out using two bibliometric analysis techniques: performance analysis and scientific mapping [110]. Firstly, following previous studies [119], [120], the performance analysis is based on productivity, taking into account the number of publications as the main indicator. In addition, the number of citations and the h-index are used to enrich the performance analysis of authors, journals, institutions, and countries. Their main objective is to provide an up-to-date overview of the research field by identifying the works that constitute its intellectual base [121]. Secondly, scientific mapping aims to reveal the structure and dynamics of scientific fields [122]. It is a spatial representation of how disciplines, fields, authors, or works relate to each other [121]. This methodological approach suits the purposes of this study. Therefore, in order to examine various interesting

aspects of the research field, we conducted a scientific mapping based on co-authorship and co-writing analyses. On the one hand, co-authorship analysis allows us to identify the social network of a research field through the links between its most relevant authors and the subgroups that emerge from collaborations at the level of institutions and countries [123]. This technique captures stronger social links than other measures of relatedness, making it ideal for examining social networks [122]. On the other hand, the analysis of keywords in aggregate or in concurrence makes it possible to establish the intellectual structure of a scientific field by dividing the words into different groups [122]. Specifically, this method quantitatively relates and ranks the conceptual content of publications according to the content of publications based on the occurrence of similar word pairs [124]. In other words, keyword co-occurrence helps to identify a research domain through the specific connections made between its keywords [125], [126]. In fact, keyword co-occurrence analysis is particularly appropriate for this purpose compared to other bibliometric methods, such as co-citation analysis and bibliometric linkage, as the latter are not optimal for mapping research fronts due to, among other reasons, the fact that citations take time to accumulate and it is, therefore, more difficult to directly connect the most recent publications across clusters of knowledge bases [122]. In addition, the keywords of an article reflect its main content and the frequency of occurrence and co-occurrence represent the most important topics addressed by papers in a research area and how they link to each other [121]. The sum of all joint co-occurrences between keywords allows for the establishment of a network map of the co-occurrence matrix, which allows the recognition of several thematic groups or clusters. Finally, maps are produced for different time periods in order to map changes in the conceptual space of the field [127] and future avenues of research.

Despite the emergence of this field of research since the beginning of the current millennium, its significance in scientific production can be traced back to 2015, as it is shown in Figures 4 and 5. In this paper, more than 22 review articles in the scientific production are listed in Table 2. It should be noted that none of them conducted a bibliometric study.

IV. RESULTS AND DISCUSSION

This section describes the results obtained from the methodology carried out in the previous section, where different important aspects are analyzed such as the authors who have written the most publications and the institutions they belong to or which country has the highest scientific production in relation to the subject studied. This information is intended to represent the current state of AI techniques used in gesture recognition and to have an idea of how this research is going to continue.

A. DATA ANALYSIS

This analysis is going to start with the main productive indicators related to the documents published per year (see Table 3),

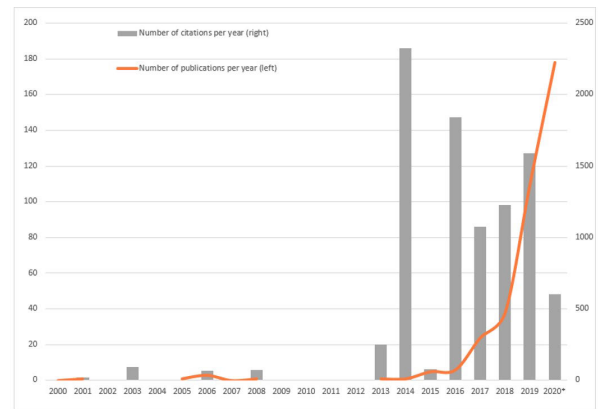


FIGURE 4. Evolution of published articles and citations from 2000 to 2020.

such as the number of papers published per year (A), number of citations per year (C), the average number of citations per paper (C/A), number of authors per year (AU), number of authors that published at least one paper in a specific year (AUA), number of journals that published at least one paper in a specific year (JA) and number of countries that published at least one paper in a specific year (COA). Nevertheless, the rest of the tables contain these parameters: Number of total papers (A), number of citations for all papers (C), average citation per paper (C/A), year of first published paper (1st A), year of last published paper (Last A), population (million inhabitants) (P) and number of papers per one million inhabitants (AP).

Regarding the number of articles, it can be seen in Figure 4 that since 2014 there has been an interest in this area of knowledge on the part of the scientific community, highlighting the massive increase in documents in the recent years wherein 2018 less than 50 publications were written and only 180 papers were published in 2020. Furthermore, the analysis of the number of citations revealed that 2014 is the year with the highest number of citations (180). In addition, Figure 5 shows that the number of authors who have published articles on the subject (AUA) has increased exponentially since 2017 with 200 authors involved in 2020, which shows the growing interest and an increasing number of collaborations between authors trying to fill the research gap in this field. Moreover, this broad field of research has been accompanied by steady growth in the number of journals and countries that publish relevant papers.

In the distribution of the documents included in this bibliometric analysis; out of 93.52% of papers, 3.85% are reviews and the rest is made up of papers published in special editions of presentations, books, and book chapters (Figure 6).

Regarding journals, Table 4 shows additional bibliometric indicators, such as citations, the average number of citations per article, the year of the first publication, the year of the last publication, and the h index. The most relevant journal in terms of number of articles is IEEE Access with

TABLE 2. Most significant reviews related to gesture recognition and AI.

Authors Year	Title	Contribution to the fields of	
		Gesture Recognition	Artificial Intelligence
Arac A, 2020	A review of recent approaches for emotion classification using electrocardiography and electrodermography signals	Facial expression recognition	Support Vector Machine; K-Nearest Neighbour; Random Forest
Bulagang AF et al, 2020	Facial Expression Recognition Using Computer Vision: A Systematic Review	Facial expression recognition	Convolutional Neural Network; Support Vector Machine; K-Nearest Neighbour; Naive Bayes; Hidden Markov Model; Decision Tree; Random Forest
Canedo D & Neves AJR, 2019	Continuous authentication using biometrics: An advanced review	Face recognition	Convolutional Neural Network
Dahia G et al, 2020	Convolutional neural network: a review of models, methodologies and applications to object detection	Human activity recognition	Convolutional Neural Network
Fang YT et al, 2020	Visual Object Recognition: Do We (Finally) Know More Now Than We Did?	Face recognition	Convolutional Neural Network
Hu TH et al, 2017	Memristor devices for neural networks	None	Spiking Neural Network; Artificial Neural Network
Jeong H & Shi LP, 2019	A review of image-based automatic facial landmark identification techniques	Facial recognition	Convolutional Neural Network
Johnston B & de Chazal P, 2018	A survey on security threats and defensive techniques of machine learning: A data driven view	None	Naive Bayes; Logistic Regression; Decision Tree; Support Vector Machine; Deep Neural Network
Liu Q et al, 2018	A Review on Automated Facial Nerve Function Assessment from Visual Face Capture	Facial expression recognition	Support Vector Machine; K-Nearest Neighbour; Artificial Neural Network
Lou J et al, 2020	Artificial Intelligence in Medical Practice: The Question to the Answer?	None	Convolutional Neural Network
Miller DD & Brown EW, 2018	Face Space Representations in Deep Convolutional Neural Networks	Face recognition	Deep Convolutional Neural Network
O'Toole AJ et al, 2018	Primer on machine learning: Utilization of large data set analyses to individualize pain management	Facial expression recognition	Deep Neural Network
Rashidi P et al, 2019	A multimodal approach for automatic detection of infant pain using facial expression and crying	Facial expression recognition	Convolutional Neural Network; Long Short Term Memory
Sandeep PVK & Suresh Kumar N, 2020	Artificial intelligence in radiation oncology: A specialty-wide disruptive transformation?	None	None
Thompson RF et al, 2018	Deep Learning for Computer Vision: A Brief Review	Face recognition; Activity recognition; Human pose estimation	Convolutional Neural Network; Deep Belief Network; Deep Boltzmann Machine; Stacked Denoising Autoencoders
Voulodimos A et al, 2018	Face Feature Extraction: A Complete Review	Face recognition	Convolutional Neural Network; Deep Belief Network
Wang HJ et al, 2018	Hand Gesture Recognition in Automotive Human-Machine Interaction Using Depth Cameras	Hand gesture recognition	Long Short-Term Memory; Convolutional Neural Network
Zhou FY et al, 2017	Deep Learning: The Good, the Bad, and the Ugly	Face recognition; Action recognition	Convolutional Neural Network; Generative Adversarial Network

62 articles (A), while if we refer to citations and citations by articles, the most prominent journal is IEEE Transactions on Pattern Analysis and Machine Intelligence with 4,440 citations (C) and 444.00 citations per article (C/A) during the period from 2013 to 2020, the year of the first and last published article. In second place is Neurocomputing with 45 articles, followed by Multimedia Tools and Applications with 23. However, taking into account h-index,

Multimedia Tools and Applications becomes the magazine with the greatest impact with an h-index of 43.

As for authors, Table 5 displays the most productive authors in the area. These authors come mainly from South Korea and the institution called Dongguk University. It is worth mentioning that the affiliation indicated in Table 5 belongs to the one indicated at the time of publication of the last document. All the authors in this table have published the

TABLE 3. Main characteristics of the data used.

Year	A	C	C/A	AU	AUA	IA	JA	COA
2000	1	20	20.00	2	2	1	1	1
2001	0	-	-	-	-	-	-	-
2002	1	96	96.00	4	4	1	1	1
2003	0	-	-	-	-	-	-	-
2004	1	3	3.00	6	6	1	1	1
2005	3	65	21.67	12	12	3	3	3
2006	0	-	-	-	-	-	-	-
2007	1	73	73.00	5	5	1	1	1
2008	0	-	-	-	-	-	-	-
2009	1	5	5.00	8	8	1	1	1
2010	0	-	-	-	-	-	-	-
2011	0	-	-	-	-	-	-	-
2012	1	253	253.00	4	4	1	1	1
2013	1	2,326	2,326.00	8	8	8	2	4
2014	5	80	16.00	14	14	6	4	5
2015	6	1,844	307.33	17	17	13	6	2
2016	24	1,077	44.88	104	103	40	19	15
2017	38	1,226	32.26	151	149	60	28	14
2018	110	1,588	14.44	439	434	159	65	28
2019	178	603	3.39	706	672	289	115	48
2020*	198	107	0.54	794	727	319	95	47
2021**	1	0	0.00	2	2	1	1	1

*until July inclusive

**early publication

TABLE 4. Ten most productive journals.

Rank	Journal	A	C	C/A	1st A	Last A	h-Index
1	IEEE Access	62	147	2.37	2018	2020	6
2	NEUROCOMPUTING	45	819	18.20	2014	2020	14
3	MULTIMEDIA TOOLS AND APPLICATIONS	23	43	1.87	2017	2020	43
4	SENSORS	18	545	30.28	2016	2020	6
5	IEEE Transactions on image processing	13	250	19.23	2017	2020	5
6	COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE	10	230	201.60	2016	2019	4
7	IEEE Transactions on Pattern Analysis and Machine Intelligence	10	4,440	444.0	2013	2020	9
8	IEEE Transactions on multimedia	9	76	8.44	2016	2020	4
9	APPLIED SCIENCES-BASEL	8	12	1.33	2019	2019	2
10	IEEE Transactions on Information Forensics and Security	8	34	4.25	2014	2020	2

TABLE 5. Most productive authors.

Rank	Authors	A	C	C/A	1st A	Last A	h-Index	Country	Affiliation
1	Zhan, Shu	4	63	15.75	2016	2020	2	Peoples R China	Hefei Univ Technol
2	Park, Kang Ryoung	4	39	9.75	2017	2019	3	South Korea	Dongguk Univ
3	Pham, Tuyen Danh	4	39	9.75	2017	2019	3	South Korea	Dongguk Univ
4	Nguyen, Dat Tien	4	39	9.75	2017	2019	3	South Korea	Dongguk Univ

same number of articles with a total of 4 articles, with Zhan, Shu being the most cited author and the one with the highest number of citations per article between 2016 and 2020. He was affiliated with the Hefei University of Technology.

Figure 7 shows the international networks of academic institutions, which have in common more than 5 scientific studies among researchers, with two groups identified. The main group (red) is the most productive. It includes three universities of Chinese origin, highlighting the most productive academic institution (Table 6) Chinese Academy of Science, along with Xidian University and the University of Chinese Academy of Science. The other group includes Tsinghua

University (Peoples R China), Shenzhen University (Peoples R China), Carnegie Mellon University (USA), and Nanyang Technological University (Singapore).

The analysis of the institutions is shown in Table 6, which shows the ten most productive institutions in this area analyzed from 2000 to 2020. These institutions are all located in China, a fact that is not strange due to the fact that most of the authors who publish in this field as well as the most relevant scientific production come from China. According to the number of articles, the most productive is the Chinese Academy of Science, which has had 20 articles since 2016. The second most productive institution is the University of

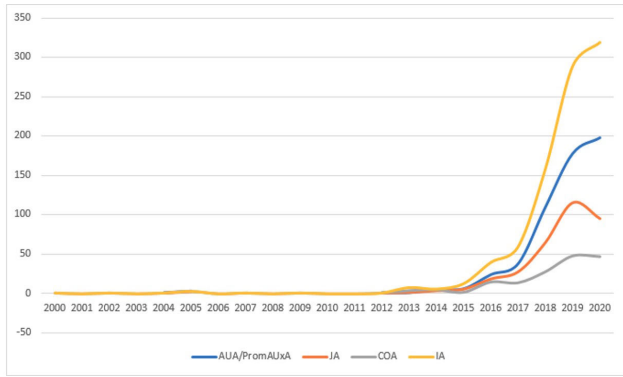


FIGURE 5. Evolution of the number of authors, journals, countries and institutions.

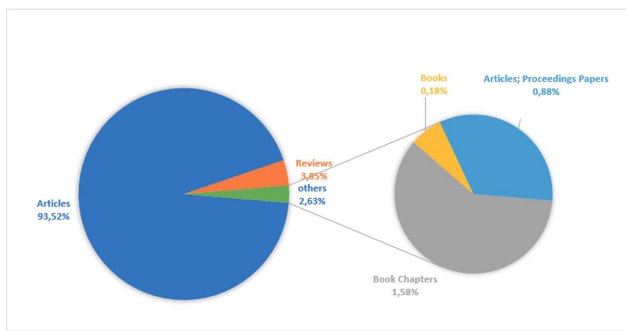


FIGURE 6. Document types included.

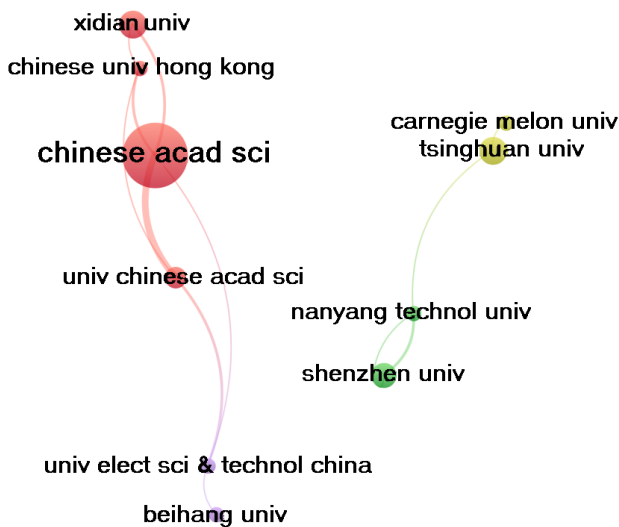


FIGURE 7. Network of the co-authorship-based cooperation between institutions.

Chinese Academy of Science with 10 articles since 2017. The third academic institution is Tsinghua University which also has 10 articles. According to the number of citations per article (C/A), the first place is occupied by the University of Chinese Academy of Science with a ratio of 24.80 citations per article. In this regard, Tsinghua University ranks second

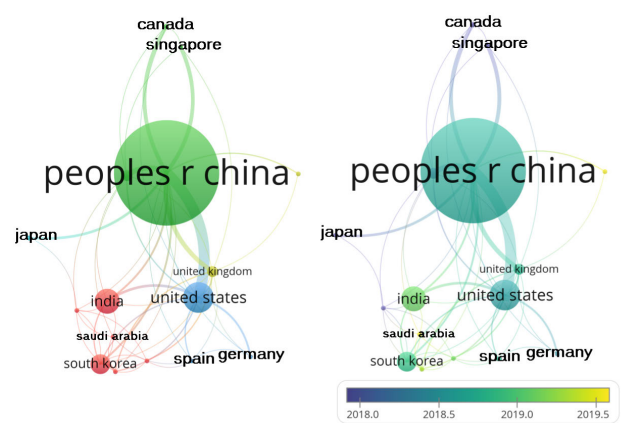


FIGURE 8. Network of the co-authorship-based between countries.

with 19.60 citations per article. From the point of view of the h index, Tsinghua University and Nanjing University of Science & Technology share a position since they have a value of 6 for this indicator. As can be deduced from the data shown in the table, the institution that has been a pioneer in this field of research has been Sun Yat Sen University.

Table 7 presents the countries that have published a greater number of articles on gesture recognition using techniques from the field of AI, which includes information related to citations, articles, the index h, and the year of the first and last publication. It should be noted that an article may represent more than one country, as the countries are established by the affiliated institutions of the researchers involved. The most influential countries in terms of the number of articles are China and the United States with 252 and 74 documents respectively, followed by India (56) and South Korea (47), as can be seen in Figure 8. In addition, taking into account the number of articles per million inhabitants (PA), Australia ranks first with 1.04. Considering the number of citations (C), the documents from China (6,325) stand out, although the highest relation of citations per article is obtained by the United States, with 45.12 citations per document. Considering the h-index, China leads the ranking with an index of 26.

Table 8 shows the ten most relevant papers associated with the field of research analyzed in this work. These papers have been ranked according to the number of citations and the year of publication to avoid the deficiency that may be caused by the exclusive fact of considering the citations of a paper. The most relevant data from each of these papers will be presented below:

- 1) In [128], an architecture based on 3D convolutional neural networks for human activity recognition is presented. This architecture extracts the spatial and temporal characteristics of the videos it analyzes and also has the particularity that it generates several information channels from different videos to finally combine all that information and obtain the features.
- 2) The novelty of this work [129] is to train a deep network with a spatial pyramid pooling layer, which shows good

TABLE 6. Ten most productive institutions.

Rank	Institution	Country	A	C	C/A	1st A	Last A	h-Index
1	Chinese Acad Sci	Peoples R China	20	320	16.00	2016	2020	4
2	Univ Chinese Acad Sci	Peoples R China	10	248	24.80	2017	2020	3
3	Tsinghua Univ	Peoples R China	10	196	19.60	2015	2020	6
4	Shenzhen Univ	Peoples R China	10	72	7.20	2018	2020	4
5	Nanjing Univ Sci & Technol	Peoples R China	9	97	10.78	2017	2020	6
6	Zhejiang Univ	Peoples R China	9	71	7.89	2018	2020	3
7	Xidian Univ	Peoples R China	9	21	2.33	2018	2020	3
8	Sun Yat Sen Univ	Peoples R China	8	26	3.25	2014	2020	3
9	South China Univ Technol	Peoples R China	7	131	18.71	2015	2020	4
10	Tianjin Univ	Peoples R China	7	32	4.57	2018	2020	4

TABLE 7. Ten most productive countries.

Rank	Country	A	P	AP	C	C/A	1st A	Last A	h-Index
1	Peoples R China	252	1,420,062,022	0.18	6325	25.10	2013	2020	26
2	United States	74	329,093,110	0.22	3339	45.12	2013	2020	17
3	India	56	1,368,737,513	0.04	129	2.30	2016	2021	7
4	South Korea	47	51,339,238	0.92	189	4.02	2016	2020	7
5	Australia	26	25,088,636	1.04	398	15.31	2005	2020	8
6	United Kingdom	26	66,959,016	0.39	709	27.27	2013	2020	7
7	Germany	12	82,438,639	0.15	457	38.08	2016	2020	6
8	Canada	12	37,279,811	0.32	374	31.17	2012	2020	5
9	Taiwan	12	23,758,247	0.51	56	4.67	2005	2020	3
10	Pakistan	12	204,596,442	0.06	34	2.83	2018	2020	3

results in detection and classification tasks, as well as a greater speed. This approach is faster than R- CNN and makes it suitable for real-world applications.

- 3) In [130], the authors have developed a Deep Learning-based framework called DeepConvLSTM that combines convolutional layers and recurring LSTMs, in order to recognize human activity from data extracted by wearable sensors. Among the main characteristics of this union are that it allows us to extract the characteristics automatically and to model the temporal dependencies of the activation of their neurons.
- 4) In [131], a hybrid model formed by the convolutional neural network (CNN) and Support Vector Machine (SVM) classifiers for handwritten digit recognition is shown. This model has been created by replacing the last layer of the CNN with an SVM classifier in order to recognize unknown patterns. This architecture has presented promising results mainly due to 3 reasons: the automatic extraction of salient features, this model has the advantages of CNN and SVM techniques, and the reduction of complexity in the decision process compared to other models.
- 5) In [132], three new architectures for CNN of sEMG-based gesture classification are presented. Furthermore, it describes a novel Transfer Learning (TL) paradigm that improves the performance of these convolutional networks. This scheme is based on the learning of two CNNs at the same time. The network

that the authors refer to as the “source” shares information with the second network through element-wise summation with the advantage that reduces the issues of connecting both networks. Moreover, it implements a method that fosters residual learning, therefore the second network only has to learn a limited number of weights. Two datasets have been tested where it has been shown that this TL design has improved in each of the proposed CNN, where one of the networks has achieved an accuracy of 98.31%.

- 6) In [100], a facial expression recognition system has been created using traditional Deep Learning techniques such as convolutional neural networks with the aim of making it a simple method. Despite its simplicity, this methodology has managed to obtain an accuracy rate of 96.76%.
- 7) In [133], a review focused on Deep Learning is presented since this area of knowledge has been in high demand due to the promising results that it has had in tasks such as object recognition and face recognition. This study shows the structure and characteristics of the components of the Convolutional Neural Networks as well as the various applications in which it has been used. Finally, some experiments have been carried out where Convolutional Neural Networks that had different structures have been evaluated and the advantages and disadvantages of these Deep Learning techniques have been described.

TABLE 8. Ten most cited articles published in the last decade taking into consideration the year of publication.

Rank	Title	Author/s	Journal	C	Year	C/A
1	3D Convolutional neural networks for human action recognition	Ji S, Xu W, Yang M, Yu K	IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE	2,317	2013	331.0
2	Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition	He K, Zhang X, Ren S, Sun J	IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE	1,739	2015	347.8
3	Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition	Ordonez FJ, Roggen D	SENSORS	461	2016	115.3
4	A novel hybrid CNN-SVM classifier for recognizing handwritten digits	Niu XX, Suen CY	PATTERN RECOGNITION	253	2012	31.6
5	Deep learning for electromyographic hand gesture signal classification using transfer learning	Cote-Allard U, Fall CL, Drouin A, Campeau-Lecours A, Gosselin C, Glette K, Laviolette F, Gosselin B	IEEE Transactions on Neural Systems and Rehabilitation Engineering	241	2019	30.1
6	Facial expression recognition with Convolutional Neural Networks: Coping with few data and the training sample order	Lopes AT, De Aguiar E, De Souza AF, Oliveira-Santos T	PATTERN RECOGNITION	207	2017	69.0
7	Review of Convolutional Neural Network	Zhou FY, Jin LP, Dong J	CHINESE JOURNAL OF COMPUTERS	174	2017	87.0
8	Deep learning and process understanding for data-driven Earth system science	Reichstein M, Camps-Valls G, Stevens B, Jung M, Denzler J, Carvalhais N, Prabhat	NATURE	159	2019	159.0
9	Deep Learning for Computer Vision: A Brief Review	Voulodimos A, Doulamis N, Doulamis A, Protopapadakis E	COMPUTATIONAL INTELLIGENCE AND NEUROSCIENCE	149	2018	74.5
10	Deep Dynamic Neural Networks for Multimodal Gesture Segmentation and Recognition	Wu D, Pigou L, Kindermans PJ, Le NDH, Shao L, Dambre J, Odohez JM	IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE	142	2016	35.5

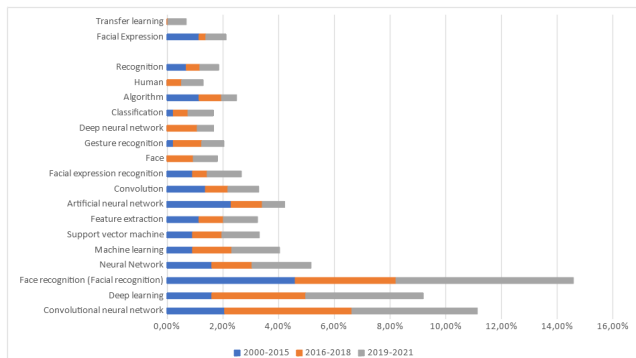


FIGURE 9. Most cited keywords from 2000 to 2020.

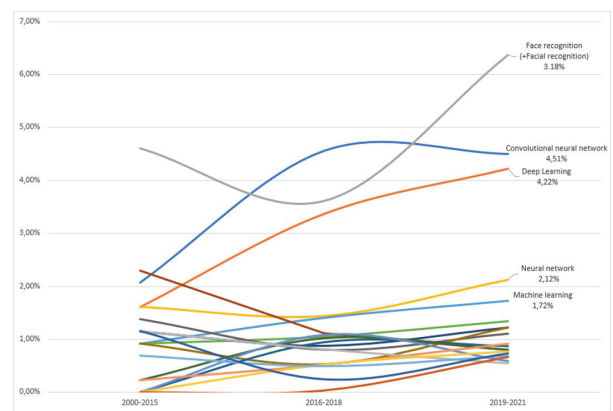


FIGURE 10. Normalized frequency of occurrence for each keyword among papers published in three time period considered.

- 8) In [134], a description of Machine Learning models as replacement of the traditional physical models in Earth sciences is given since more and more data is being extracted from this field of knowledge and data science is becoming necessary to deal with and analyze this huge amount of data volume.
- 9) In [21], the authors have reviewed the main developments in Deep Learning architectures and algorithms for Computer Vision. This work has been based mainly on three types of Deep Learning models: Convolutional Neural Networks (CNN), Boltzmann-type models, and Stacked Autoencoders. These techniques have been

- 10) In [135], a framework has been developed that combines deep neural networks with Hidden Markov Models (HMM) for gesture recognition. The deep neural networks used are a Gaussian-Bernoulli Deep Belief Network to extract skeletal joint features and a 3D Convolutional Neural Network to extract features such

TABLE 9. Top twenty of keywords used (Part I).

2000-2015		2016-2018		2019-2020	
Keywords	A %	Keywords	A %	Keywords	A %
Face recognition (+ Facial recognition)	20 4,61%	Face recognition (+ Facial recognition)	103 3,62%	Face recognition (+ Facial recognition)	270 6,37%
Artificial neural network	10 2,30%	Artificial neural network	32 1,12%	Artificial neural network	34 0,80%
Convolutional neural network	9 2,07%	Convolutional neural network	130 4,56%	Convolutional neural network	191 4,51%
Deep learning	7 1,61%	Deep learning	96 3,37%	Deep learning	179 4,22%
Neural network	7 1,61%	Neural network	41 1,44%	Neural network	90 2,12%
Convolution	6 1,38%	Convolution	23 0,81%	Convolution	47 1,11%
Feature extraction	5 1,15%	Feature extraction	25 0,88%	Feature extraction	52 1,23%
Algorithm	5 1,15%	Algorithm	23 0,81%	Algorithm	23 0,54%
Facial expression	5 1,15%	Facial expression	7 0,25%	Facial expression	31 0,73%
Machine learning	4 0,92%	Machine learning	40 1,40%	Machine learning	73 1,72%

as depth and RGB images. The role of the HMM is to take into account time dependencies in the learning process.

In Tables 9 and 10 you can see the twenty most representative keywords in different time intervals. The common factor in these different periods is characterized by face recognition, which is among the highest positions in each of the defined intervals. This trend is due to the fact that face recognition is a common problem in Computer Vision that continues to

TABLE 10. Top twenty of keywords used (Part II).

2000-2015		2016-2018		2019-2020	
Keywords	A %	Keywords	A %	Keywords	A %
Support machine	4 0,92%	Support machine	30 1,05%	Support machine	57 1,34%
Facial expression recognition	4 0,92%	Facial expression recognition	15 0,53%	Facial expression recognition	52 1,23%
Recognition	3 0,69%	Recognition	14 0,49%	Recognition	29 0,68%
Recurrent neural network	2 0,46%	Recurrent neural network	13 0,46%	Recurrent neural network	20 0,47%
Gesture recognition	1 0,23%	Gesture recognition	29 1,02%	Gesture recognition	34 0,80%
Classification	1 0,23%	Classification	15 0,53%	Classification	39 0,92%
Transfer learning	0 0,00%	Transfer learning	1 0,04%	Transfer learning	28 0,66%
Face	0 0,00%	Face	27 0,95%	Face	37 0,87%
Deep neural net-work	0 0,00%	Deep neural net-work	31 1,09%	Deep neural net-work	25 0,59%
Human	0 0,00%	Human	15 0,53%	Human	20 0,47%

persist and also the identification of the face and its respective verification have become very popular for access control and security issues, as well as aspects of biometrics [136]. However, it is necessary to make a special mention of the Convolutional Neural Network and Deep Learning keywords that are also in the first positions along with Face recognition since Deep Learning has made important contributions to Computer Vision in recent years. Finally, in relation to gesture recognition, which is one of the most representative keywords for this study, it should be noted that it has begun to arouse

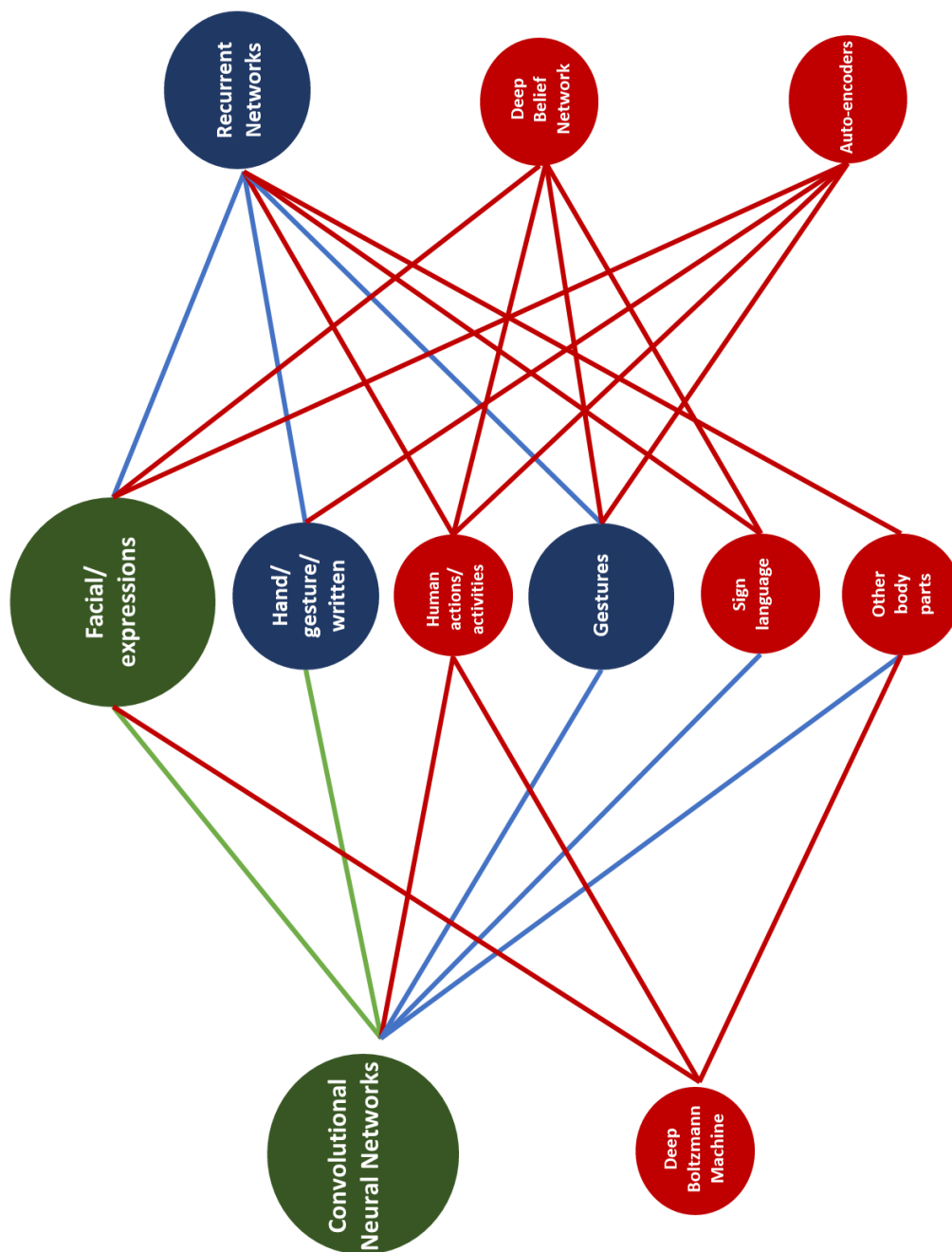


FIGURE 11. Relationship between the DL techniques and the gesture recognition.

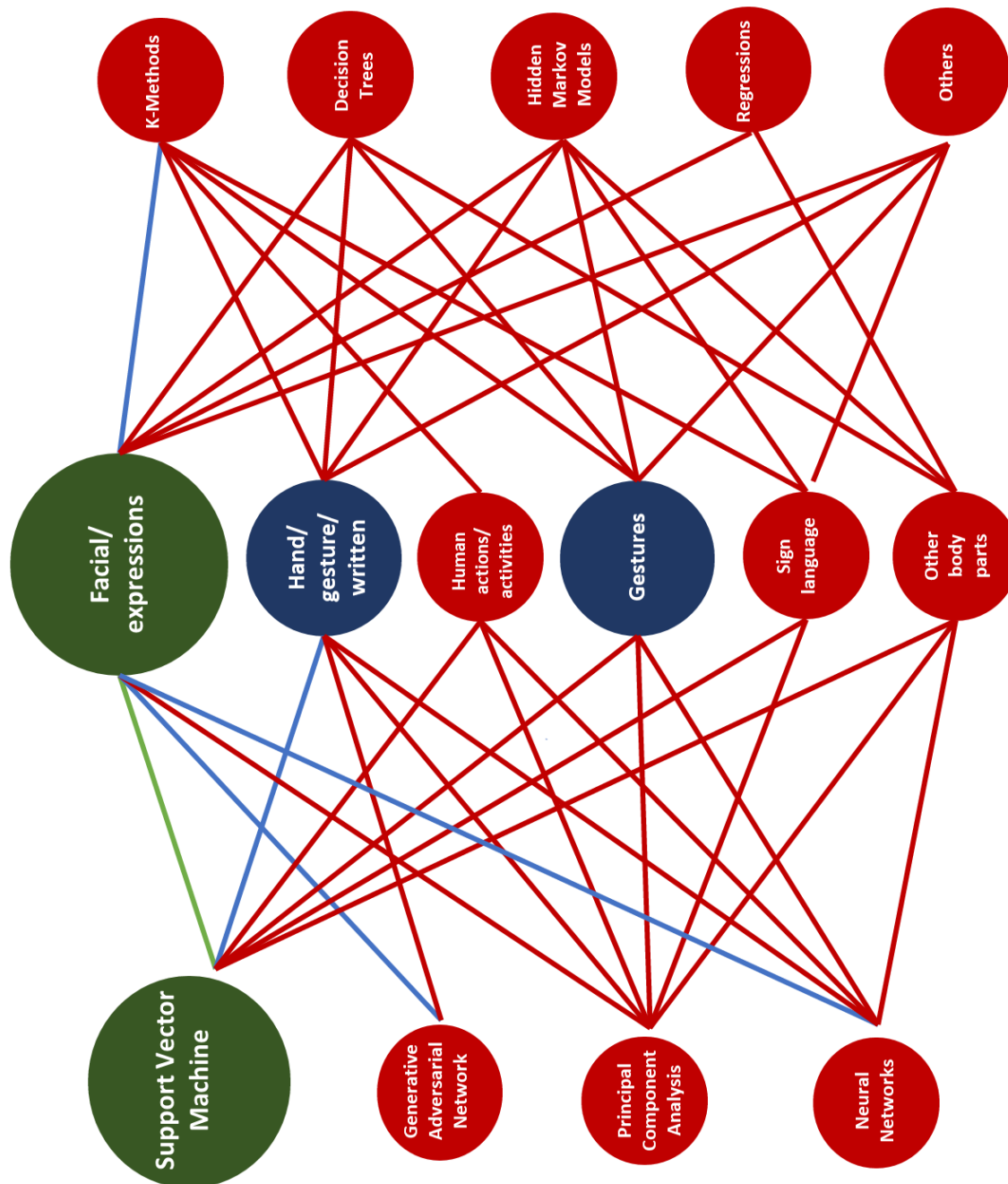


FIGURE 12. Relationship between the ML techniques and the gesture recognition.

more interest from 2016 when it is among the 10 most cited keywords in the papers of that time. This statement is confirmed in Figure 9 where it can be seen that the percentage of frequency of said keyword is higher between the period 2016-2020.

B. GESTURE RECOGNITION MODEL

The models in Figures 11 and 12 represent the relationship between AI techniques and the various types of gesture recognition. This model has been created by reviewing each paper and extracting the type of gesture recognition and AI techniques that were used to identify the gestures. As a result, the graph has been obtained, where it can be observed how the nodes in the center represent the types of gestures while the peripheral nodes display the different AI techniques that have been identified. Figure 11 shows Deep Learning techniques while Figure 12 displays Machine Learning techniques. This relationship between these two concepts has been determined by the edges that form the connection between the central and peripheral nodes, as well as the size of the nodes and the color of these edges and nodes. On the one hand, it is necessary to comment that there are three different sizes: small, medium and large. Large-sized nodes are those with a number of occurrences greater than 80, medium-sized ones are in the (30-80] range, and small ones are in the (0-30] range. On the other hand, there are 3 different colors: green, blue and red. Regarding the nodes, these were painted according to size, the nodes with large size were painted green, with medium size blue and the small size nodes red. Regarding the edges, the thickness of the edges is the same for all and the occurrences have been reflected by their color. The green edges mean that they have a number of occurrences greater than 30, the blue edges are in the range of (10-30] and the red edges are in the range (0-10]. However, these edges not only inform about the number of occurrences but also show which AI techniques have been used for each type of recognition since if there are no occurrences then the edge is not painted in the figure.

From the point of view of success cases, this model verifies the techniques that have been most successful since there have been more published papers that have used this AI technique in a type of recognition. Some of the techniques have been included in studies of each of the types of gesture recognition such as Convolutional Neural Networks, Recurrent Neural Networks and Support Vector Machine, while others have only been used in a couple of them. Furthermore, Convolutional Neural Networks, Recurrent Neural Networks, and Support Vector Machine have been featured the most in scientific publications. This fact could be interpreted as that these methods applied to gesture recognition have obtained good results and researchers have chosen to use them to a greater extent. On the other hand, the techniques that have not been so successful have been: Hidden Markov Models, Deep Boltzmann Machine and Regressions.

V. CONCLUSION

Gesture recognition techniques have aroused the interest of the scientific community in recent years, and have been used in various fields such as education, vehicles, virtual reality, sign language translation and so forth. However, the application of AI techniques to this field of research has been a determining factor in its progress. This fact has been the origin to conduct this study where the most relevant indicators regarding research on the subject of AI techniques applied to gesture recognition that span the last twenty years, from 2000 to 2020 have been collected.

From this study, it has been found that there has been a substantial increase in the number of articles produced and citations referring to the subject of this bibliometric analysis as of 2016, while previously it went unnoticed. A noteworthy fact is that the academic institutions that have shown a greater interest in the subject and have contributed the most to research on gesture recognition and AI come mainly from China. This fact is probably prompted because Deep Learning methods are being applied to numerous areas of knowledge including Computer Vision. This combination is producing great advances in the field of Computer Vision, which is closely related to gesture recognition because in the classification of this type of recognition there is vision-based gesture recognition. Therefore, it is not a surprise that there has been a recent interest in this subject in recent years to obtain promising results in this field with the use of AI techniques.

In this bibliometric analysis, we have worked with approximately 571 articles extracted from the WoS and Scopus scientific databases. The research criteria are the combination of different types of gesture recognition and ML and DL techniques, in order to select the papers that have specifically used an AI technique to include gesture recognition in their development. Based on this information, the most relevant aspects related to scientific production on the subject of this study have been described, where in the first place the process that has been followed to prepare this bibliometric study has been explained in section III and the results obtained from the said process have been shared below in the IV section. In section IV, the indicators that are relevant for this study have been shown: the evolution of the articles published during the years that this study comprises, the journals with the greatest number of articles, the authors that have been cited the most, the countries where they have produced the largest number of articles and with the highest h-index. Regarding the indicators indicated previously, the results show that the journals with the highest number of published papers are IEEE Access, Neurocomputing and Multimedia Tools and Applications, although the journal with the most citations is IEEE Transactions on Pattern Analysis and Machine Intelligence. The countries that have been the most productive are China and United States, with the institutions from China being the ones that have produced a greater number of articles on the topic of interest. Nevertheless, the majority of authors

who have published a greater number of articles are from South Korea. Furthermore, a brief description of each of the ten articles with the most citations, as well as the keywords that are most meaningful when searching on the terms concerning this work have been described.

From our models, it can be concluded that facial gestures have been the most demanded in the researches, followed by hand gestures. Regarding DL, CNN and Recurrent networks are the most used, CNN being the most preferred technique to use both in facial expressions and in hand gestures and it should be noted that this method is one of the few that has been applied in all types of gestures presented. In the model referring to ML, the SVM technique has stood out, where it has obtained the highest number of occurrences in the works that involve facial expressions. Moreover, the combination of CNN and SVM has been highly acclaimed by researchers.

Despite the research that is being carried out, the interest in including AI algorithms in gesture recognition is recent and there are still quite a few limitations to overcome. However, the advances are promising, as can be seen in the description of the ten most cited articles that have been included in this analysis.

REFERENCES

- [1] P. Xu, "A real-time hand gesture recognition and human-computer interaction system," 2017, *arXiv:1704.07296*.
- [2] H. Zou, Y. Zhou, J. Yang, H. Jiang, L. Xie, and C. J. Spanos, "Wi-Fi-enabled device-free gesture recognition for smart home automation," in *Proc. IEEE 14th Int. Conf. Control Autom. (ICCA)*, Jun. 2018, pp. 476–481.
- [3] S. Tateno, Y. Zhu, and F. Meng, "Hand gesture recognition system for in-car device control based on infrared array sensor," in *Proc. 58th Annu. Conf. Soc. Instrum. Control Eng. Jpn. (SICE)*, Sep. 2019, pp. 701–706.
- [4] G. Li, H. Tang, Y. Sun, J. Kong, G. Jiang, D. Jiang, B. Tao, S. Xu, and H. Liu, "Hand gesture recognition based on convolution neural network," *Cluster Comput.*, vol. 22, no. S2, pp. 2719–2729, Mar. 2019.
- [5] R. Li, K. Zou, and W. Wang, "Application of human body gesture recognition algorithm based on deep learning in non-contact human body measurement," *J. Ambient Intell. Humanized Comput.*, vol. 11, pp. 1–11, Apr. 2020.
- [6] J. Li and D. Zhang, "Face gesture recognition based on clustering algorithm," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Jun. 2019, pp. 2008–2012.
- [7] Y. Zhang, Y. Huang, X. Sun, Y. Zhao, X. Guo, P. Liu, C. Liu, and Y. Zhang, "Static and dynamic human arm/hand gesture capturing and recognition via multiinformation fusion of flexible strain sensors," *IEEE Sensors J.*, vol. 20, no. 12, pp. 6450–6459, Jun. 2020.
- [8] K. Sinha, R. Kumari, A. Priya, and P. Paul, "A computer vision-based gesture recognition using hidden Markov model," in *Innovations in Soft Computing and Information Technology*. Cham, Switzerland: Springer, 2019, pp. 55–67.
- [9] M. Kim, J. Cho, S. Lee, and Y. Jung, "IMU sensor-based hand gesture recognition for human-machine interfaces," *Sensors*, vol. 19, no. 18, p. 3827, Sep. 2019.
- [10] Y. Li, "Hand gesture recognition using Kinect," in *Proc. IEEE Int. Conf. Comput. Sci. Autom. Eng.*, Jun. 2012, pp. 196–199.
- [11] W. Lu, Z. Tong, and J. Chu, "Dynamic hand gesture recognition with leap motion controller," *IEEE Signal Process. Lett.*, vol. 23, no. 9, pp. 1188–1192, Sep. 2016.
- [12] S. He, C. Yang, M. Wang, L. Cheng, and Z. Hu, "Hand gesture recognition using MYO armband," in *Proc. Chin. Autom. Congr. (CAC)*, Oct. 2017, pp. 4850–4855.
- [13] F. Zoghalmi, H. Heinrich, G. Schneider, and M. A. Hamdi, "Tracking body motions in order to guide a robot using the time of flight technology," in *Proc. IPIN (Short Papers, Work Prog. Papers)*, 2019, pp. 48–55.
- [14] R. E. Tolentino, P. M. F. Guinto, and D. Y. B. Maypa, "Recognition of different emergency situation through body gesture using microsoft Kinect sensor," in *Proc. IEEE 11th Int. Conf. Humanoid, Nanotechnol., Inf. Technol., Commun. Control, Environ., Manage. (HNICEM)*, Nov. 2019, pp. 1–5.
- [15] A. Vaitkevičius, M. Taroza, T. Blažauskas, R. Damaševičius, R. Maskeliūnas, and M. Woźniak, "Recognition of American sign language gestures in a virtual reality using leap motion," *Appl. Sci.*, vol. 9, no. 3, p. 445, Jan. 2019.
- [16] A. O. Kılıç, E. Sari, H. Yucel, M. M. Oğuz, E. Polat, E. A. Acoglu, and S. Senel, "Exposure to and use of mobile devices in children aged 1–60 months," *Eur. J. Pediatrics*, vol. 178, no. 2, pp. 221–227, Feb. 2019.
- [17] J. Ker, L. Wang, J. Rao, and T. Lim, "Deep learning applications in medical image analysis," *IEEE Access*, vol. 6, pp. 9375–9389, 2018.
- [18] H. Li, "Deep learning for natural language processing: Advantages and challenges," *Nat. Sci. Rev.*, vol. 5, no. 1, pp. 24–26, Jan. 2018.
- [19] W. Zhang, X. Zhao, L. Zhao, D. Yin, G. H. Yang, and A. Beutel, "Deep reinforcement learning for information retrieval: Fundamentals and advances," in *Proc. 43rd Int. ACM SIGIR Conf. Res. Develop. Inf. Retr.*, Jul. 2020, pp. 2468–2471.
- [20] G. Apruzzese, M. Colajanni, L. Ferretti, A. Guido, and M. Marchetti, "On the effectiveness of machine and deep learning for cyber security," in *Proc. 10th Int. Conf. Cyber Conflict (CyCon)*, May 2018, pp. 371–390.
- [21] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–13, Feb. 2018.
- [22] F. Zantalis, G. Koulouras, S. Karabetos, and D. Kandris, "A review of machine learning and IoT in smart transportation," *Future Internet*, vol. 11, no. 4, p. 94, Apr. 2019.
- [23] G. Devineau, F. Moutarde, W. Xi, and J. Yang, "Deep learning for hand gesture recognition on skeletal data," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit. (FG)*, May 2018, pp. 106–113.
- [24] J.-H. Kim, G.-S. Hong, B.-G. Kim, and D. P. Dogra, "DeepGesture: Deep learning-based gesture recognition scheme using motion sensors," *Displays*, vol. 55, pp. 38–45, Dec. 2018.
- [25] L. Li, S. Qin, Z. Lu, K. Xu, and Z. Hu, "One-shot learning gesture recognition based on joint training of 3D ResNet and memory module," *Multimedia Tools Appl.*, vol. 79, nos. 9–10, pp. 6727–6757, Mar. 2020.
- [26] M. E. Benalcázar, A. G. Jaramillo, A. Zea, A. Páez, and V. H. Andaluz, "Hand gesture recognition using machine learning and the Myo armband," in *Proc. 25th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2017, pp. 1040–1044.
- [27] J. A. Dakheel, C. D. Pero, N. Aste, and F. Leonforte, "Smart buildings features and key performance indicators: A review," *Sustain. Cities Soc.*, vol. 61, Oct. 2020, Art. no. 102328.
- [28] M. K. Saini and N. Goel, "How smart are smart classrooms? A review of smart classroom technologies," *ACM Comput. Surveys*, vol. 52, no. 6, pp. 1–28, Nov. 2020.
- [29] C. Ma, Y. Zhang, A. Wang, Y. Wang, and G. Chen, "Traffic command gesture recognition for virtual urban scenes based on a spatiotemporal convolution neural network," *ISPRS Int. J. Geo-Inf.*, vol. 7, no. 1, p. 37, Jan. 2018.
- [30] J. J. Ojeda-Castelo, J. A. Piedra-Fernandez, and L. Iribarne, "A device-interaction model for users with special needs," *Multimedia Tools Appl.*, vol. 80, no. 5, pp. 6675–6710, Feb. 2021.
- [31] O. Bau and W. E. Mackay, "OctoPocus: A dynamic guide for learning gesture-based command sets," in *Proc. 21st Annu. ACM Symp. User Interface Softw. Technol.*, 2008, pp. 37–46.
- [32] T. Baudel and M. Beaudouin-Lafon, "CHARADE: Remote control of objects using free-hand gestures," *Commun. ACM*, vol. 36, no. 7, pp. 28–35, Jul. 1993.
- [33] D. Watson, M. Hancock, R. L. Mandryk, and M. Birk, "Deconstructing the touch experience," in *Proc. ACM Int. Conf. Interact. Tabletops Surf.*, Oct. 2013, pp. 199–208.
- [34] T. Mo and P. Sun, "Research on key issues of gesture recognition for artificial intelligence," *Soft Comput.*, vol. 24, no. 8, pp. 5795–5803, Apr. 2020.
- [35] I. M. Revina and W. R. S. Emmanuel, "A survey on human face expression recognition techniques," *J. King Saud Univ.-Comput. Inf. Sci.*, vol. 33, no. 6, pp. 619–628, Sep. 2018.
- [36] A. S. Al-Shamayleh, R. Ahmad, M. A. M. Abushariah, K. A. Alam, and N. Jomhari, "A systematic literature review on vision based gesture recognition techniques," *Multimedia Tools Appl.*, vol. 77, no. 21, pp. 28121–28184, Nov. 2018.

- [37] F. Noroozi, C. A. Corneanu, D. Kamińska, T. Sapiński, S. Escalera, and G. Anbarjafari, "Survey on emotional body gesture recognition," *IEEE Trans. Affect. Comput.*, vol. 12, no. 2, pp. 505–523, Apr. 2021.
- [38] S. Akbarian, G. Delfi, K. Zhu, A. Yadollahi, and B. Taati, "Automated non-contact detection of head and body positions during sleep," *IEEE Access*, vol. 7, pp. 72826–72834, 2019.
- [39] S. Sharma, P. T. Varigonda, P. Bindal, A. Sharma, and A. Jain, "Monocular 3D human pose estimation by generation and ordinal ranking," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 2325–2334.
- [40] I. Katircioglu, B. Tekin, M. Salzmann, V. Lepetit, and P. Fua, "Learning latent representations of 3D human pose with deep neural networks," *Int. J. Comput. Vis.*, vol. 126, no. 12, pp. 1326–1341, Dec. 2018.
- [41] J. Romero, M. Loper, and M. J. Black, "FlowCap: 2D human pose from optical flow," in *Proc. German Conf. Pattern Recognit.* Cham, Switzerland: Springer, 2015, pp. 412–423.
- [42] C. Lugaresi, J. Tang, H. Nash, C. McClanahan, E. Uboweja, M. Hays, F. Zhang, C. L. Chang, M. Yong, J. Lee, and W. T. Chang, "MediaPipe: A framework for perceiving and processing reality," in *Proc. 3rd Workshop Comput. Vis. AR/VR IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 1–4.
- [43] Y. Saleh and G. F. Issa, "Arabic sign language recognition through deep neural networks fine-tuning," *Int. J. Online Biomed. Eng.*, vol. 16, no. 5, pp. 71–83, 2020, doi: 10.3991/ijoe.v16i05.13087.
- [44] B. Xie, X. He, and Y. Li, "RGB-D static gesture recognition based on convolutional neural network," *J. Eng.*, vol. 2018, no. 16, pp. 1515–1520, Nov. 2018.
- [45] Y. Yu, X. Chen, S. Cao, X. Zhang, and X. Chen, "Exploration of Chinese sign language recognition using wearable sensors based on deep belief net," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 5, pp. 1310–1320, May 2020.
- [46] S.-K. Ko, C. J. Kim, H. Jung, and C. Cho, "Neural sign language translation based on human keypoint estimation," *Appl. Sci.*, vol. 9, no. 13, p. 2683, Jul. 2019.
- [47] M. A. Ahmed, B. B. Zaidan, A. A. Zaidan, M. M. Salih, and M. M. B. Lakulu, "A review on systems-based sensory gloves for sign language recognition state of the art between 2007 and 2017," *Sensors*, vol. 18, no. 7, p. 2208, 2018.
- [48] M. A. Ahmed, B. B. Zaidan, A. A. Zaidan, M. M. Salih, Z. T. Al-Qaysi, and A. H. Alamoodi, "Based on wearable sensory device in 3D-printed humanoid: A new real-time sign language recognition system," *Measurement*, vol. 168, Jan. 2021, Art. no. 108431.
- [49] M. A. Ahmed, B. B. Zaidan, A. A. Zaidan, A. H. Alamoodi, O. S. Albahri, Z. T. Al-Qaysi, A. S. Albahri, and M. M. Salih, "Real-time sign language framework based on wearable device: Analysis of MSL, DataGlove, and gesture recognition," *Soft Comput.*, vol. 25, no. 16, pp. 11101–11122, Aug. 2021.
- [50] M. R. Alam, M. Bennamoun, R. Togneri, and F. Soheli, "A joint deep Boltzmann machine (jDBM) model for person identification using mobile phone data," *IEEE Trans. Multimedia*, vol. 19, no. 2, pp. 317–326, Feb. 2017.
- [51] A. Lozano-Diez, J. Gonzalez-Dominguez, R. Zazo, D. Ramos, and J. Gonzalez-Rodriguez, "On the use of convolutional neural networks in pairwise language recognition," in *Proc. Adv. Speech Lang. Technol. Iberian Lang.* Cham, Switzerland: Springer, 2014, pp. 79–88.
- [52] A. Mohanty, P. Vaishnavi, P. Jana, A. Majumdar, A. Ahmed, T. Goswami, and R. R. Sahay, "Nrityabodha: Towards understanding Indian classical dance using a deep learning approach," *Signal Process., Image Commun.*, vol. 47, pp. 529–548, Sep. 2016.
- [53] B. S. Anami and V. A. Bhandage, "A comparative study of suitability of certain features in classification of bharatanatyam mudra images using artificial neural network," *Neural Process. Lett.*, vol. 50, no. 1, pp. 741–769, Aug. 2019.
- [54] X. Zhao, X. Shi, and S. Zhang, "Facial expression recognition via deep learning," *IETE Tech. Rev.*, vol. 32, no. 5, pp. 347–355, 2015.
- [55] Q. Li, J. Yu, T. Kurihara, H. Zhang, and S. Zhan, "Deep convolutional neural network with optical flow for facial micro-expression recognition," *J. Circuits, Syst. Comput.*, vol. 29, no. 1, Jan. 2020, Art. no. 2050006.
- [56] A. Ruiz-Garcia, M. Elshaw, A. Altahhan, and V. Palade, "A hybrid deep learning neural approach for emotion recognition from facial expressions for socially assistive robots," *Neural Comput. Appl.*, vol. 29, no. 7, pp. 359–373, 2018.
- [57] Z. Yue, F. Yanyan, Z. Shangyou, and P. Bing, "Facial expression recognition based on convolutional neural network," in *Proc. IEEE 10th Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Oct. 2019, Art. no. 012100.
- [58] F. Pérez-Hernández, S. Tabik, A. Lamas, R. Olmos, H. Fujita, and F. Herrera, "Object detection binary classifiers methodology based on deep learning to identify small objects handled similarly: Application in video surveillance," *Knowl.-Based Syst.*, vol. 194, Apr. 2020, Art. no. 105590.
- [59] H. Wu, Z.-W. Chen, G.-H. Tian, Q. Ma, and M.-L. Jiao, "Item ownership relationship semantic learning strategy for personalized service robot," *Int. J. Autom. Comput.*, vol. 17, no. 3, pp. 390–402, Jun. 2020.
- [60] T. Huynh-The, C.-H. Hua, N. A. Tu, and D.-S. Kim, "Learning 3D spatiotemporal gait feature by convolutional network for person identification," *Neurocomputing*, vol. 397, pp. 192–202, Jul. 2020.
- [61] X. Jiang, K. Xu, and T. Sun, "Action recognition scheme based on skeleton representation with DS-LSTM network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 30, no. 7, pp. 2129–2140, Jul. 2019.
- [62] A. Pampouchidou, M. Padiaditis, E. Kazantzaki, S. Sfakianakis, I. A. Apostolaki, K. Argyraki, D. Manousos, F. Meriaudeau, K. Marias, F. Yang, M. Tsiknakis, M. Basta, A. N. Vgontzas, and P. Simos, "Automated facial video-based recognition of depression and anxiety symptom severity: Cross-corpus validation," *Mach. Vis. Appl.*, vol. 31, no. 4, pp. 1–19, May 2020.
- [63] S. Albawi, O. Bayat, S. Al-Azawi, and O. N. Ucan, "Social touch gesture recognition using convolutional neural network," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–10, Oct. 2018.
- [64] M. Nimbarte and K. K. Bhojar, "Biased face patching approach for age invariant face recognition using convolutional neural network," *Int. J. Intell. Syst. Technol. Appl.*, vol. 19, no. 2, pp. 103–124, 2020.
- [65] S. L. Phung and A. Bouzerdoum, "A pyramidal neural network for visual pattern recognition," *IEEE Trans. Neural Netw.*, vol. 18, no. 2, pp. 329–343, Mar. 2007.
- [66] Y. Fang, Q. Lan, T. Xie, Y. Liu, S. Mei, and B. Zhu, "New opportunities and challenges for forensic medicine in the era of artificial intelligence technology," *Fa yi xue za zhi*, vol. 36, no. 1, pp. 77–85, 2020.
- [67] M. Afifi, "11K Hands: Gender recognition and biometric identification using a large dataset of hand images," *Multimedia Tools Appl.*, vol. 78, pp. 20835–20854, Aug. 2019.
- [68] I. Anwar and N. U. Islam, "Learned features are better for ethnicity classification," 2017, *arXiv:1709.07429*.
- [69] Y. Zhai, Y. Huang, Y. Xu, J. Gan, H. Cao, W. Deng, R. D. Labati, V. Piuri, and F. Scotti, "Asian female facial beauty prediction using deep neural networks via transfer learning and multi-channel feature fusion," *IEEE Access*, vol. 8, pp. 56892–56907, 2020.
- [70] E.-Y. Huan, G.-H. Wen, S.-J. Zhang, D.-Y. Li, Y. Hu, T.-Y. Chang, Q. Wang, and B.-L. Huang, "Deep convolutional neural networks for classifying body constitution based on face image," *Comput. Math. Methods Med.*, vol. 2017, pp. 1–9, Oct. 2017.
- [71] T. Sutabri, P. Pamungkur, A. Kurniawan, and R. E. Saragih, "Automatic attendance system for university student using face recognition based on deep learning," *Int. J. Mach. Learn. Comput.*, vol. 9, no. 5, pp. 668–674, Oct. 2019.
- [72] W.-H. Yun, D. Lee, C. Park, J. Kim, and J. Kim, "Automatic recognition of children engagement from facial video using convolutional neural networks," *IEEE Trans. Affect. Comput.*, vol. 11, no. 4, pp. 696–707, Oct. 2020.
- [73] G. Storey, R. Jiang, S. Keogh, A. Bouridane, and C.-T. Li, "3DPalsyNet: A facial palsy grading and motion recognition framework using fully 3D convolutional neural networks," *IEEE Access*, vol. 7, pp. 121655–121664, 2019.
- [74] A. H. Sable and S. N. Talbar, "An adaptive entropy based scale invariant face recognition face altered by plastic surgery," *Pattern Recognit. Image Anal.*, vol. 28, no. 4, pp. 813–829, Oct. 2018.
- [75] E. F. B. Tasdemir and B. Yanikoglu, "A comparative study of delayed stroke handling approaches in online handwriting," *Int. J. Document Anal. Recognit.*, vol. 22, no. 1, pp. 15–28, Mar. 2019.
- [76] N. Bi, J. Chen, and J. Tan, "The handwritten Chinese character recognition uses convolutional neural networks with the GoogLeNet," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 33, no. 11, Oct. 2019, Art. no. 1940016.
- [77] S. Lai and L. Jin, "Recurrent adaptation networks for online signature verification," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 6, pp. 1624–1637, Jun. 2019.

- [78] K. Nagano, H. Luo, Z. Wang, J. Seo, J. Xing, L. Hu, L. Wei, and H. Li, "Deep face normalization," *ACM Trans. Graph.*, vol. 38, no. 6, pp. 1–16, Dec. 2019.
- [79] L. Lu, J. Yu, Y. Chen, H. Liu, Y. Zhu, L. Kong, and M. Li, "Lip reading-based user authentication through acoustic sensing on smartphones," *IEEE/ACM Trans. Netw.*, vol. 27, no. 1, pp. 447–460, Feb. 2019.
- [80] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers Neurosci.*, vol. 11, p. 379, Jul. 2017.
- [81] G. Velarde, *Artificial Era*. Germany: PRICA, 2021, p. 273.
- [82] F.-H. Hsu, "IBM's deep blue chess grandmaster chips," *IEEE Micro*, vol. 19, no. 2, pp. 70–81, Mar. 1999.
- [83] L. Chen, P. Chen, and Z. Lin, "Artificial intelligence in education: A review," *IEEE Access*, vol. 8, pp. 75264–75278, 2020.
- [84] N. N. Misra, Y. Dixit, A. Al-Mallahi, M. S. Bhullar, R. Upadhyay, and A. Martynenko, "IoT, big data, and artificial intelligence in agriculture and food industry," *IEEE Internet Things J.*, vol. 9, no. 9, pp. 6305–6324, May 2022.
- [85] L. Athota, V. K. Shukla, N. Pandey, and A. Rana, "Chatbot for healthcare system using artificial intelligence," in *Proc. 8th Int. Conf. Rel., Infocom Technol. Optim. (Trends Future Directions) (ICRITO)*, Jun. 2020, pp. 619–622.
- [86] E. S. Olivias, J. D. M. Guerrero, M. Martinez-Sober, J. R. Magdalena-Benedito, and L. Serrano, *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques: Algorithms, Methods, and Techniques*. Hershey, PA, USA: IGI Global, 2009.
- [87] T. Kocmi, "Exploring benefits of transfer learning in neural machine translation," 2020, *arXiv:2001.01622*.
- [88] D. Sarkar, R. Bali, and T. Ghosh, *Hands-on Transfer Learning With Python: Implement Advanced Deep Learning and Neural Network Models Using TensorFlow and Keras*. Birmingham, U.K.: Packt, 2018.
- [89] J. Yang, H. Zou, Y. Zhou, and L. Xie, "Learning gestures from WiFi: A Siamese recurrent convolutional architecture," *IEEE Internet Things J.*, vol. 6, no. 6, pp. 10763–10772, Dec. 2019.
- [90] H. Zou, J. Yang, Y. Zhou, L. Xie, and C. J. Spanos, "Robust WiFi-enabled device-free gesture recognition via unsupervised adversarial domain adaptation," in *Proc. 27th Int. Conf. Comput. Commun. Netw. (ICCCN)*, Jul. 2018, pp. 1–8.
- [91] Y. Zoabi, S. Deri-Rozov, and N. Shomron, "Machine learning-based prediction of COVID-19 diagnosis based on symptoms," *NPJ Digit. Med.*, vol. 4, no. 1, pp. 1–5, Dec. 2021.
- [92] R. Murugan and N. Palanichamy, "Smart city air quality prediction using machine learning," in *Proc. 5th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2021, pp. 1048–1054.
- [93] H.-Y. Huang, M. Broughton, M. Mohseni, R. Babbush, and S. Boixo, "Power of data in quantum machine learning," *Nature Commun.*, vol. 12, no. 1, pp. 1–9, May 2021.
- [94] I. J. Jacob and P. E. Darney, "Design of deep learning algorithm for IoT application by image based recognition," *J. ISMAC*, vol. 3, no. 3, pp. 276–290, Aug. 2021.
- [95] Q. Hu, H. M. Whitney, and M. L. Giger, "A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI," *Sci. Rep.*, vol. 10, no. 1, pp. 1–11, Dec. 2020.
- [96] R. Sharma, R. B. Pachori, and P. Sircar, "Automated emotion recognition based on higher order statistics and deep learning algorithm," *Biomed. Signal Process. Control*, vol. 58, Apr. 2020, Art. no. 101867.
- [97] R. Aziz, C. K. Verma, and N. Srivastava, "Dimension reduction methods for microarray data: A review," *AIMS Bioeng.*, vol. 4, no. 2, pp. 179–197, 2017.
- [98] Z. Zhang, Z. Tian, and M. Zhou, "HandSense: Smart multimodal hand gesture recognition based on deep neural networks," *J. Ambient Intell. Humanized Comput.*, vol. 9, pp. 1–16, Aug. 2018.
- [99] Z. Hu, Y. Hu, J. Liu, B. Wu, D. Han, and T. Kurfess, "3D separable convolutional neural network for dynamic hand gesture recognition," *Neurocomputing*, vol. 318, pp. 151–161, Nov. 2018.
- [100] A. T. Lopes, E. de Aguiar, A. F. De Souza, and T. Oliveira-Santos, "Facial expression recognition with convolutional neural networks: Coping with few data and the training sample order," *Pattern Recognit.*, vol. 61, pp. 610–628, Jan. 2017.
- [101] I. Ofodile, A. Helmi, A. Clapés, E. Avots, K. M. Peensoo, S.-M. Valdma, A. Valdmann, H. Valtma-Lukner, S. Omelkov, S. Escalera, C. Ozcinar, and G. Anbarjafari, "Action recognition using single-pixel time-of-flight detection," *Entropy*, vol. 21, no. 4, p. 414, Apr. 2019.
- [102] G. Zhu, L. Zhang, L. Mei, J. Shao, J. Song, and P. Shen, "Large-scale isolated gesture recognition using pyramidal 3D convolutional networks," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 19–24.
- [103] A.-R. Lee, Y. Cho, S. Jin, and N. Kim, "Enhancement of surgical hand gesture recognition using a capsule network for a contactless interface in the operating room," *Comput. Methods Programs Biomed.*, vol. 190, Jul. 2020, Art. no. 105385.
- [104] B. Sun, S. Cao, J. He, and L. Yu, "Affect recognition from facial movements and body gestures by hierarchical deep spatio-temporal features and fusion strategy," *Neural Netw.*, vol. 105, pp. 36–51, Sep. 2018.
- [105] H. Gunes and M. Piccardi, "A bimodal face and body gesture database for automatic analysis of human nonverbal affective behavior," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 1, 2006, pp. 1148–1153.
- [106] X. Sun and M. Lv, "Facial expression recognition based on a hybrid model combining deep and shallow features," *Cognit. Comput.*, vol. 11, no. 4, pp. 587–597, Aug. 2019.
- [107] P. Lucey, J. F. Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, "The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 94–101.
- [108] K. Czuszyński, J. Rumiński, and A. Kwasniewska, "Gesture recognition with the linear optical sensor and recurrent neural networks," *IEEE Sensors J.*, vol. 18, no. 13, pp. 5429–5438, May 2018.
- [109] S. K. Swarnkar and A. Ambhaikar, "Improved convolutional neural network based sign language recognition," *Int. J. Adv. Sci. Technol.*, vol. 27, p. 302–317, Aug. 2019.
- [110] M. J. Cobo, A. G. López-Herrera, E. Herrera-Viedma, and F. Herrera, "Science mapping software tools: Review, analysis, and cooperative study among tools," *J. Amer. Soc. Inf. Sci. Technol.*, vol. 62, no. 7, pp. 1382–1402, Jul. 2011.
- [111] S. A. Morris and B. Van der Veer Martens, "Mapping research specialties," *Annu. Rev. Inf. Sci. Technol.*, vol. 42, no. 1, pp. 213–295, Nov. 2009.
- [112] L. F. Agramunt, J. M. Berbel-Pineda, M. M. Capobianco-Uriarte, and M. P. Casado-Belmonte, "Review on the relationship of absorptive capacity with interorganizational networks and the internationalization process," *Complexity*, vol. 2020, pp. 1–20, Mar. 2020.
- [113] S. Kraus, H. Li, Q. Kang, P. Westhead, and V. Tiberius, "The sharing economy: A bibliometric analysis of the state-of-the-art," *Int. J. Entrepreneurial Behav. Res.*, vol. 26, no. 8, pp. 1769–1786, Aug. 2020.
- [114] M. Gusenbauer and N. R. Haddaway, "Which academic search systems are suitable for systematic reviews or meta-analyses? Evaluating retrieval qualities of Google scholar, PubMed, and 26 other resources," *Res. Synth. Methods*, vol. 11, no. 2, pp. 181–217, Mar. 2020.
- [115] L. Waltman and E. Noyons, "Bibliometrics for research management and research evaluation," Leiden Univ., Leiden, The Netherlands, Tech. Rep., 2018.
- [116] C. Wah Ng and S. Ranganath, "Gesture recognition via pose classification," in *Proc. 15th Int. Conf. Pattern Recognit. (ICPR)*, vol. 3, 2000, pp. 699–704.
- [117] C. W. Ng and S. Ranganath, "Real-time gesture recognition system and application," *Image Vis. Comput.*, vol. 20, nos. 13–14, pp. 993–1007, Dec. 2002.
- [118] M. D. L. M. Capobianco-Uriarte, M. D. P. Casado-Belmonte, G. M. Marín-Carrillo, and E. Terán-Yépez, "A bibliometric analysis of international competitiveness (1983–2017)," *Sustainability*, vol. 11, no. 7, p. 1877, Mar. 2019.
- [119] H. Baier-Fuentes, M. H. González-Serrano, M. A.-D. Santos, W. Inzunza-Mendoza, and V. Pozo-Estrada, "Emotions and sport management: A bibliometric overview," *Frontiers Psychol.*, vol. 11, p. 1512, Jul. 2020.
- [120] E. Terán-Yépez, G. M. Marín-Carrillo, M. D. P. Casado-Belmonte, and M. D. L. M. Capobianco-Uriarte, "Sustainable entrepreneurship: Review of its evolution and new trends," *J. Cleaner Prod.*, vol. 252, Apr. 2020, Art. no. 119742.
- [121] M. Alayo, T. Iturralde, A. Maseda, and G. Aparicio, "Mapping family firm internationalization research: Bibliometric and literature review," *Rev. Managerial Sci.*, vol. 15, pp. 1–44, Aug. 2020.
- [122] I. Zupic and T. Čater, "Bibliometric methods in management and organization," *Organizational Res. Methods*, vol. 18, no. 3, pp. 429–472, Jul. 2015.
- [123] F. J. Acedo, C. Barroso, C. Casanueva, and J. L. Galan, "Co-authorship in management and organizational studies: An empirical and network analysis," *J. Manage. Stud.*, vol. 43, no. 5, pp. 957–983, Jul. 2006.

- [124] S. Bhattacharya and P. K. Basu, "Mapping a research area at the micro level using co-word analysis," *Scientometrics*, vol. 43, no. 3, pp. 359–372, Nov. 1998.
- [125] M. Callon, J.-P. Courtial, W. A. Turner, and S. Bauin, "From translations to problematic networks: An introduction to co-word analysis," *Social Sci. Inf.*, vol. 22, no. 2, pp. 191–235, 1983.
- [126] M. C. López-Fernández, A. M. Serrano-Bedia, and M. Pérez-Pérez, "Entrepreneurship and family firm research: A bibliometric analysis of an emerging field," *J. Small Bus. Manage.*, vol. 54, no. 2, pp. 622–639, Apr. 2016.
- [127] N. Coulter, I. Monarch, and S. Konda, "Software engineering as seen through its research literature: A study in co-word analysis," *J. Amer. Soc. Inf. Sci.*, vol. 49, no. 13, pp. 1206–1223, 1998.
- [128] S. Ji, W. Xu, M. Yang, and K. Yu, "3D convolutional neural networks for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 1, pp. 221–231, Jan. 2012.
- [129] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1904–1916, Sep. 2015.
- [130] F. J. Ordóñez and D. Roggen, "Deep convolutional and LSTM recurrent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, p. 115, Jan. 2016.
- [131] X.-X. Niu and C. Y. Suen, "A novel hybrid CNN-SVM classifier for recognizing handwritten digits," *Pattern Recognit.*, vol. 45, no. 4, pp. 1318–1325, Apr. 2012.
- [132] U. Côté-Allard, C. L. Fall, A. Drouin, A. Campeau-Lecours, C. Gosselin, K. Glette, F. Lavolette, and B. Gosselin, "Deep learning for electromyographic hand gesture signal classification using transfer learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Oct. 2019.
- [133] F. Y. Zhou, L. P. Jin, and J. Dong, "Review of convolutional neural network," *Chin. J. Comput.*, vol. 40, no. 6, pp. 1229–1251, Jun. 2017.
- [134] M. Reichstein, G. Camps-Valls, B. Stevens, M. Jung, J. Denzler, N. Carvalhais, and M. Prabhat, "Deep learning and process understanding for data-driven earth system science," *Nature*, vol. 566, pp. 195–204, Feb. 2019.
- [135] D. Wu, L. Pigou, P.-J. Kindermans, N. D.-H. Le, L. Shao, J. Dambre, and J.-M. Odobez, "Deep dynamic neural networks for multimodal gesture segmentation and recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 38, no. 8, pp. 1583–1597, Aug. 2016.
- [136] G. Guo and N. Zhang, "A survey on deep learning based face recognition," *Comput. Vis. Image Understand.*, vol. 189, Dec. 2019, Art. no. 102805.



JUAN JESUS OJEDA-CASTELO received the degree in computer science engineering from the University of Almería, the master's degree in systems and languages programming about computer science from UNED, and the Ph.D. degree from the University of Almería. He is currently a Substitute Interim Professor with the University of Almería, where he has been a Researcher with the Department of Mathematics for 2 years and a Collaborator of the research project titled "Researching the educational use of the Kinect at C.E.E.E. Princesa Sofia Almería"

which was supported by the Government of Andalusia. His research interests include human-computer interaction particularly natural interaction, natural user interface, computer vision, artificial intelligence, and deep learning. The devices that he usually attempts to include in his personal projects are Microsoft Kinect, Leap Motion, Intel RealSense, and Oculus Rift.



MARIA DE LAS MERCEDES CAPOBIANCO-URIARTE received the degree in biomedical engineering from the National University of Entre Ríos, the degree in business administration from the University of Miguel Hernández, and the Ph.D. degree in the area of sector economics, agriculture and development from the University of Almería. She is currently an Interim Substitute Professor with the University of Almería in world economics, macroeconomics, and Spanish economics, with publications on international trade and internationalization of companies oriented to the competitiveness and sustainability of economies. Her research interest includes the application of artificial intelligence tools in the analysis of economic indicators.



JOSE ANTONIO PIEDRA-FERNANDEZ received the Ph.D. degree in computer science from the University of Almería, in 2005. He is currently an Assistant Professor with the Department of Informatics, University of Almería. He is a member of the Research Applied Computing Group (TIC-211). He has been the Coordinator of the Master in Computer Engineering, since 2014, and the Director of the Quality Secretariat, since 2017. He works closely with the James Wang's Research

Group, The Pennsylvania State University. He received the Luis Labrandero Prize by the Spanish Association of Remote Sensing, in 2007. He got a patent in the field of recognition of cancer cells using a fuzzy robot vision system. He participates in various national, international, and regional projects. He is the author and coauthor in several scientific publications, among journal articles, national and international book chapters, and publications in national and international congress proceedings. He is a Reviewer of scientific articles in several international JCR impact journals, such as IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING or *International Journal of Remote Sensing*. His research interests include computer vision, artificial intelligence, natural interaction systems, and serious games applied to the field of education and health.



ROSA AYALA received the B.S. and M.S. degrees in computer science from the University of Granada, and the Ph.D. degree in computer science from the University of Almería, Spain, and conducted from the University of TU Delft (The Netherlands). From 1993 to 1999, she worked as a Research Fellow with the FIAPA Research Center, where she worked in several national and international research projects on geographic information systems (GIS) and satellite images.

From 1993 to 2003, she worked as a Lecturer at the University of Almería. She has participated in several projects funded by the Spanish Ministry of Science and Technology, and the Andalusian Ministry ST. From 2007 to 2015, she was the Deputy Director of the Higher Engineering School, University of Almería. Since 2015, she has been the Chief in Charge of the Campus and Infrastructure Secretariat in the General Direction of Campus, Infrastructure and Sustainability of the University of Almería, where she is currently an Associate Professor with the Department of Informatics. She is the author and coauthor of several scientific publications, among journal articles, national and international book chapters, and publications in national and international congress proceedings. She belongs to the Applied Computing Group (TIC-211), University of Almería. Her research interests include satellite classification, GIS, cellular automata, and smart cities.

• • •