



OPEN

Biometric recognition through gait analysis

Claudia Álvarez-Aparicio[✉], Ángel Manuel Guerrero-Higuera,
Miguel Ángel González-Santamarta, Adrián Campazas-Vega, Vicente Matellán &
Camino Fernández-Llamas

The use of people recognition techniques has become critical in some areas. For instance, social or assistive robots carry out collaborative tasks in the robotics field. A robot must know who to work with to deal with such tasks. Using biometric patterns may replace identification cards or codes on access control to critical infrastructures. The usage of Red Green Blue Depth (RGBD) cameras is ubiquitous to solve people recognition. However, this sensor has some constraints, such as they demand high computational capabilities, require the users to face the sensor, or do not regard users' privacy. Furthermore, in the COVID-19 pandemic, masks hide a significant portion of the face. In this work, we present BRITTANY, a biometric recognition tool through gait analysis using Laser Imaging Detection and Ranging (LIDAR) data and a Convolutional Neural Network (CNN). A Proof of Concept (PoC) has been carried out in an indoor environment with five users to evaluate BRITTANY. A new CNN architecture is presented, allowing the classification of aggregated occupancy maps that represent the people's gait. This new architecture has been compared with LeNet-5 and AlexNet through the same datasets. The final system reports an accuracy of 88%.

User identification has become increasingly important in different research areas. For example, in the field of cybersecurity, to prevent access to critical facilities or in social and assistive robotics to improve Human-Robot Interaction (HRI).

In the cybersecurity field, it is necessary to take into account that the number of cyber-attacks in different environments has increased exponentially in recent years^{1,2}. On access control to critical infrastructures, cyber-criminals have many techniques to gain access to facilities either remotely, by obtaining credentials, exploiting vulnerabilities in some of the systems, or physically, by breaching access to the facilities^{3,4}. If the access is done remotely, they can be detected in the system when they start to perform malicious actions such as privilege escalation or lateral movements. These techniques are detected by analysing network traffic, system logs, or the system's overall behaviour. Nevertheless, if the access is physical, an attack named tailgating⁵, a cybercriminal only could be detected by building employees or security personnel if it exists. Thus, more and more companies nowadays implement biometric systems in their infrastructure. The standard Radio Frequency Identification (RFID) cards or passcodes have become outdated.

In robotics, the tasks that a mobile robot has to accomplish are becoming more complex, specifically if we focus on social or assistive robotics. These complex tasks are usually divided into little skills that the robot can solve. First, a robot needs to know its location in the environment where it operates⁶—this is known as localization. Then, it needs to calculate the best path avoiding obstacles or damaging people or objects in their trajectory⁷—aka navigation. Finally, the robot interacts with people and sometimes works with them on specific tasks⁸—aka Human-Robot Interaction (HRI). There are robust solutions for the first two skills because these areas have been extensively studied in the literature. The third one is probably the most complex skill, so many researchers are currently working on it. Human-Robot Interaction (HRI) has to be as similar as possible to human-human interaction⁹. Interaction refers to collision avoidance, but it also involves approaching skills or communication. Such elements are associated with two essential skills: people tracking and recognition. It is necessary to know where the people are every time, but it is also crucial to know who they are.

Tracking people helps to improve navigation skills in mobile robots and promote socially acceptable robots¹⁰. Many solutions in the literature attempt to solve this problem with Red Green Blue Depth (RGBD) cameras to detect people in the environments as shown in¹¹. Other researchers combine data from both Laser Imaging Detection and Ranging (LIDAR) and Red Green Blue Depth (RGBD) cameras^{12–14}. However, these approaches have a high computing demand, and it may be a drawback if the tracking runs onboard a robot. Therefore, some solutions have been proposed in the literature to solve the tracking problem using 2D Laser Imaging Detection

Department of Mechanical, Computer Science and Aerospace Engineering, University of León, 24071 León, Spain.
✉email: calvaa@unileon.es

and Ranging (LIDAR) sensors. For example, the work¹⁵ reviewed different methods for robot navigation in crowded indoor environments. Furthermore, the clustering and centre point estimation combined with the walking centre line estimation is used to detect people with or without a walker¹⁶.

The above approaches are not robust enough when dealing with occlusions or changes in gait speed. To address such issues, the authors proposed a Convolutional Neural Network (CNN)-based tool that allows for locating people within the robot surroundings using the data provided by a single Laser Imaging Detection and Ranging (LIDAR) sensor¹⁷. This tool, called People Tracking (PeTra), creates an occupancy map from the Laser Imaging Detection and Ranging (LIDAR) sensor's readings. Maps are processed by a Convolutional Neural Network (CNN) which returns segmented data belonging to the people in the robot's surroundings. A centre-of-mass calculation provides the people's location estimates from the segmented data. Several versions of the People Tracking (PeTra) have been released. Then, it was included a correlation method of location estimates using Kalman filters, as well as an optimization for the Convolutional Neural Network (CNN)¹⁸. Finally, a bootstrapping method was proposed to improve the accuracy of the tool in specific locations¹⁹. People Tracking (PeTra) is the tool selected to detect people in this work.

People recognition is a hot topic. It is not only required in mobile robotics, but also prevalent in our daily life. The use of biometry technologies is very extended. Biometry uses information about a specific part of the human body or behaviour, allowing us to distinguish people by analyzing such data²⁰. Biometric technology splits into two main groups. The physiological biometric technologies analyze a specific feature of the body²¹. It is a well-known method since a significant part of society uses it daily. Within this group of techniques, we can find fingerprint²², facial²³, or iris²⁴ identification. On the other hand, behavioural biometric technologies analyze actions carried out by people²⁵. In this second group, a time component is required since any action has a beginning, developing, and ending²⁶. Within this group, we can find voice²⁷, hand-writing²⁸, or gait identification²⁹. The last mentioned work presents a review of the methods used to capture gait information based on vision, sound, pressure, and accelerometry. Gait features can be extracted from a sequence of visual images or video, an underfoot pressure image sequence obtained using a pressure mat sensor, an acceleration trace recorded by an accelerometer in a wearable device, or an audio recording.

As mentioned above, a robot must recognize the person interacting with to promote socially acceptable robots. In robotics, most studies use physiological biometric technologies—specifically face recognition using a camera³⁰. However, this sensor has a critical drawback—data processing has a higher computational cost than other sensors. Furthermore, on face recognition, people should face the camera constantly. In addition, the worldwide COVID-19 situation and the consequent masks make this method even more complex. To meet such issues, some authors propose a multi-modal biometric identification combining face and voice identification^{31,32}. Finally, it is essential to point out that the use of cameras bounds the users' privacy³³. Regarding Laser Imaging Detection and Ranging (LIDAR) sensors, several proposals use these sensors to perform human detection and tracking. This sensor solves the user privacy problem as well as the usability of the system, since the user does not have to constantly look at the robot. Recent studies analyze people's gait using 2D and 3D Laser Imaging Detection and Ranging (LIDAR) sensors to improve the tracking of individuals.

The work³⁴ proposes an Long Short-Term Memory (LSTM)-based method for gait recognition using a multi-line Laser Imaging Detection and Ranging (LIDAR) sensor. The study proposes to create a silhouette from the point cloud extracted from the multi-line LIDAR and process it through an LSTM-based CNN network. The experimental results revealed that the proposed approach achieved a 60% of performance classification for 30 people.

Benedek et al.³⁵ presents two approaches, one for gait analysis and the other one for activity analysis, both based on data streams of a Rotating Multi-Beam (RMB) Laser Imaging Detection and Ranging (LIDAR) sensor. The gait analysis is used to do the person re-identification during tracking and recognition of specific activity patterns. They use a silhouette-based approach to projecting the 3D point cloud of a person obtained through the Rotating Multi-Beam (RMB) Laser Imaging Detection and Ranging (LIDAR) sensor to an image 2D plane. This image is preprocessed and evaluated through a Convolutional Neural Network (CNN). The experimental results revealed that the proposed approach achieved an average of 87% of performance classification for 28 people. Finally, a proposal to preserve the user's privacy, the work³⁶, uses a 2D Laser Imaging Detection and Ranging (LIDAR) sensor located at ankle level for people tracking. The study also performed a gait analysis to get person height estimation using Laser Imaging Detection and Ranging (LIDAR) data.

In this work, we present Biometric RecognITion Through gAit aNalySis (BRITTANY), a tool that allows for identifying people through gait analysis using a 2D Laser Imaging Detection and Ranging (LIDAR) sensor and a Convolutional Neural Network (CNN) to process its data. A 2D Laser Imaging Detection and Ranging (LIDAR) sensor was selected because of its low computational requirements and its benefits for users' privacy. The gait analysis has been chosen because it is a behavioural biometry method. Physiological biometry-based systems only process data gathered in a specific time instant. On the other hand, behavioural biometry-based strategies collect input data during a time interval. It allows for analysing more than one sample of user data, so they get more robust evidence without using any other external identification source.

The remainder of the paper organises as follows: “Materials and methods” section describes the materials and evaluation methods used to carry out our research; the results are presented and discussed in “Results and discussion” section; finally, conclusions and future works are proposed in “Conclusion” section.

Materials and methods

A set of experiments were carried out to evaluate Biometric RecognITion Through gAit aNalySis (BRITTANY). In this section, the main elements of the research are in-depth depicted. Besides, we describe the methodology used to assess the accuracy of the proposed system.

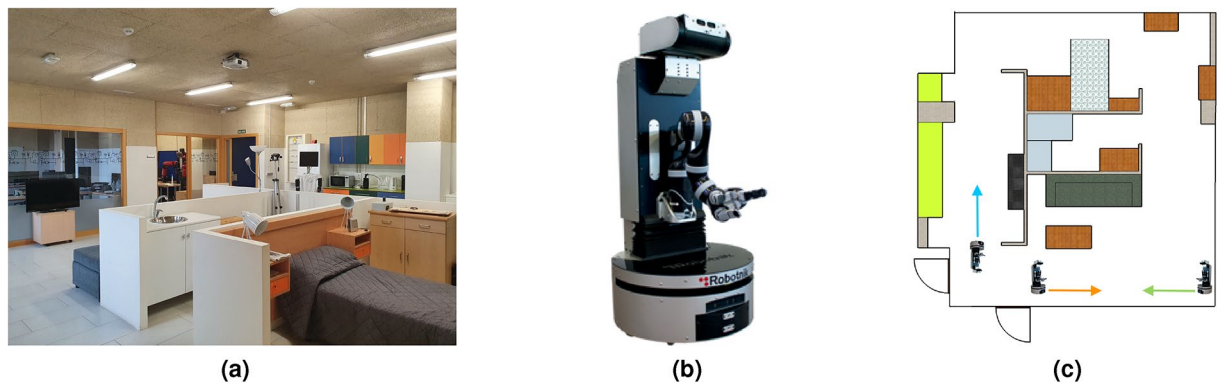


Figure 1. (a) Leon@Home Testbed. (b) Orbi-One Robot. (c) Leon@Home Testbed schema. The robot icons on the map point out the robot's location and orientation during the experiments; the arrows show the trajectory of people.

Leon@Testbed. The experiments have been done in the mock-up apartment Leon@Home Testbed³⁷, a certified testbed³⁸ of the European Robotics League (ERL) located in the Robotics Group's laboratory at the University of Leon—see Fig. 1a. It is used to test mobile service robots in a realistic environment. The apartment simulates a single-room home built in an 8 m × 7 m space. 60 cm-high walls—for allowing seeing—split the space into a living room, kitchen, bedroom, and bathroom.

Orbi-One robot. Orbi-One, the mobile service robot shown in Fig. 1b has been used to gather data. It is manufactured by Robotnik³⁹. It accommodates several sensors, such as a Red Green Blue Depth (RGBD) camera, a Laser Imaging Detection and Ranging (LIDAR) sensor, and an inertial unit. It also operates a six-degrees-of-freedom manipulator arm attached to its torso and a wheeled base for moving around the floor. Inside, an Intel Core i7 CPU with 8 GB of RAM allows it to run the software to control the robot hardware. The software runs the Robot Operating System (ROS) framework⁴⁰. Specifically, we used its onboard Laser Imaging Detection and Ranging (LIDAR) sensor to collect data for our research.

PeTra. People Tracking (PeTra)¹⁷ is developed by the Robotics Group at the University of León⁴¹ in recent years. People Tracking (PeTra) allows for locating people in the environment. It uses data provided by a Laser Imaging Detection and Ranging (LIDAR) sensor accommodated 20cm above the floor. It was evaluated by using an open-access dataset⁴². Starting from Laser Imaging Detection and Ranging (LIDAR) data People Tracking (PeTra) builds a 2D occupancy map that depicts the main features of the robot environment. Such occupancy map is then processed by a Convolutional Neural Network (CNN) that returns a new occupancy map containing only the points that belong to people. People Tracking (PeTra)'s Convolutional Neural Network (CNN) is based on the U-net architecture⁴³, commonly used to perform biomedical image segmentation⁴⁴. A centre-mass algorithm is computed over the new occupancy map to estimate the people's location. New versions of the tool were released in later years. People Tracking (PeTra) can correlate location estimates by using Euclidean distances⁴⁵ or using Kalman filters⁴⁶—a more robust correlation method. An optimized design for the Convolutional Neural Network (CNN) that allows People Tracking (PeTra) for working in real-time¹⁸. Finally, a bootstrapping-based method that improves the accuracy at specific locations, such as empty rooms or corridors¹⁹.

BRITTANY. Biometric RecognITion Through gAit aNaLYsis (BRITTANY) allows for recognizing people by their gait. The system is based on a Convolutional Neural Network (CNN) which uses an aggregation of occupancy maps provided by People Tracking (PeTra) as input. We pose that such aggregated occupancy maps are unique for each person and may be used to identify them.

Aggregated occupancy maps are processed to get probability values for each legitimate user. For instance, for five people, Biometric RecognITion Through gAit aNaLYsis (BRITTANY) might get [0.01, 0.96, 0.2, 0.24, 0.09] probability values, meaning that input data belongs to the first person with a 0.01 probability, to the second one with a 0.96 probability, to the third one with a 0.2 probability, to the fourth one with a 0.24 probability, and the fifth one with a 0.09 probability. Thus, we might assert that the input data belongs to the second person.

We consider several predictions during a time interval to prevent punctual errors. Thus, we evaluate the final estimation through a set of predictions by applying a most-voted item strategy.

Data gathering. Two datasets, available online⁴⁷, have been compiled. First, we gathered data in the mock-up apartment at Leon@Home Testbed described in “Leon@Testbed” section. Data was collected from the Laser Imaging Detection and Ranging (LIDAR) sensor onboard the Orbi-One robot mentioned in “Orbi-One robot” section. Both datasets consist of Rosbag files, a Robot Operating System (ROS) feature that allows for recording data during a time interval and playing them later.

The first dataset (\mathcal{D}_1) is composed of 90 five-second Rosbag files. Data recorded correspond to a person walking straight in front of the robot. Each Rosbag file contains the data gathered by the Laser Imaging Detection and

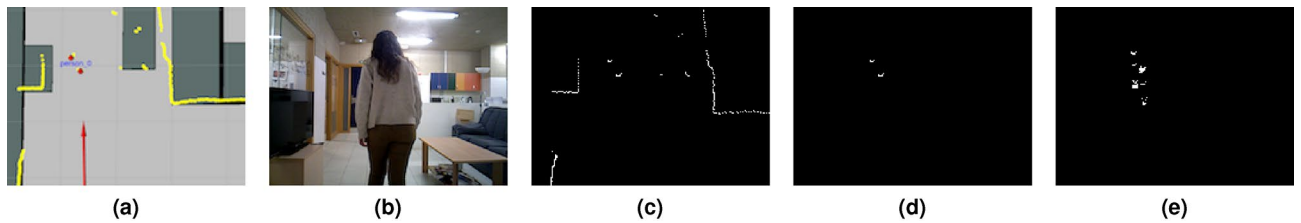


Figure 2. Data curation process: (a) Laser Imaging Detection and Ranging (LIDAR) and People Tracking (PeTra) data visualized on Rviz—The red arrow shows the robot's location and orientation, the yellow points show Laser Imaging Detection and Ranging (LIDAR) readings; (b) snapshot from Orbi-One robot camera; (c) occupancy map computed from Laser Imaging Detection and Ranging (LIDAR) data; (d) occupancy map computed by People Tracking (PeTra); and (e) occupancy map aggregation.

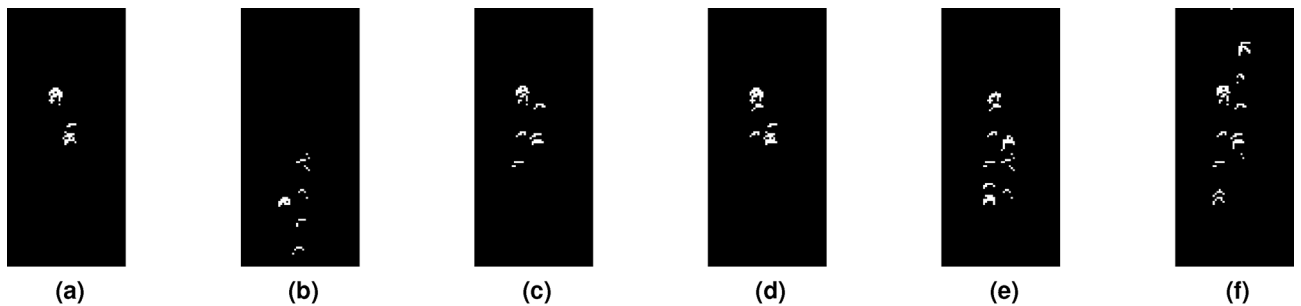


Figure 3. Aggregated occupancy maps for (a) $\mathcal{C}_{5 \times 0}$, (b) $\mathcal{C}_{5 \times 1}$, (c) $\mathcal{C}_{5 \times 2}$, (d) $\mathcal{C}_{10 \times 0}$, (e) $\mathcal{C}_{10 \times 1}$, and (f) $\mathcal{C}_{10 \times 2}$.

Ranging (LIDAR) sensor onboard the robot. We recorded data from five different people at three locations—shown in Fig. 1c on the testbed schema. We recorded six Rosbag files for each location and person.

The second dataset (\mathcal{D}_2) is composed of 108 Rosbag files. We recorded data from six people walking straight in front of the robot at three locations—see Fig. 1c. We recorded six Rosbag files for each location and person. Five out of those six people are the same in both datasets. The last one is not registered in the system. We use its data to evaluate the false-positive cases in the system.

Data curation. Once we have gait data recorded as Rosbag files, it is necessary to curate and tag them to use them for fitting the Convolutional Neural Network (CNN) in charge of people identification. \mathcal{D}_1 , depicted in “Data gathering” section, have been used to train the Convolutional Neural Network (CNN).

Since Biometric RecognITion Through gait AnaLYsis (BRITTANY) uses a Convolutional Neural Network (CNN) to identify people, we need to convert data from Rosbag files into images \mathcal{I} that describe people's gait. Image generation process is shown in Fig. 2. The Fig. 2a shows the Laser Imaging Detection and Ranging (LIDAR) data as yellow points for the real scene shown in Fig. 2b. The red arrow shows the location and orientation of the robot. First, we play each Rosbag file obtaining occupancy maps for each Laser Imaging Detection and Ranging (LIDAR) reading—see Fig. 2c. From the above occupancy maps, People Tracking (PeTra) provides a second occupancy map segmenting the points belonging to people—see Fig. 2d. Finally, People Tracking (PeTra)'s occupancy maps are aggregated. Such aggregation allows for depicting people's gait—see Fig. 2e.

An individual occupancy map is represented as an image (I), see Fig. 2d. In these images, white pixels represent a Laser Imaging Detection and Ranging (LIDAR) point belonging to a person, and black pixels represent 1) Laser Imaging Detection and Ranging (LIDAR) points belonging to objects, or 2) points where the Laser Imaging Detection and Ranging (LIDAR) sensor has not detected collisions. The final image (\mathcal{I}) is created by concatenating individual occupancy maps, see Fig. 2e. These occupancy maps are concatenated using the logical OR (\parallel) operation. Equation (1) represents the operation performed to create the final aggregate occupancy maps. In the Equation, \mathcal{I} represents the final image obtained from the concatenation of individual images (I), n is the number of I used in the aggregation, and s is the number of steps between successive images (I).

$$\mathcal{I} = \parallel_{i=0}^n I_{i(s+1)} \quad (1)$$

Different aggregation settings have been evaluated to select the best one. They are named as $\mathcal{C}_{n \times s}$, where $n \in \{5, 10\}$ is the number of Laser Imaging Detection and Ranging (LIDAR) readings used in the aggregation, and $s \in \{0, 1, 2\}$ is the number of steps between successive readings. For instance, in $\mathcal{C}_{5 \times 2}$, five Laser Imaging Detection and Ranging (LIDAR) readings are used by considering one out of three successive Laser Imaging Detection and Ranging (LIDAR) readings. Figure 3 show a sample of the resulting aggregated occupancy map for each setting schema.

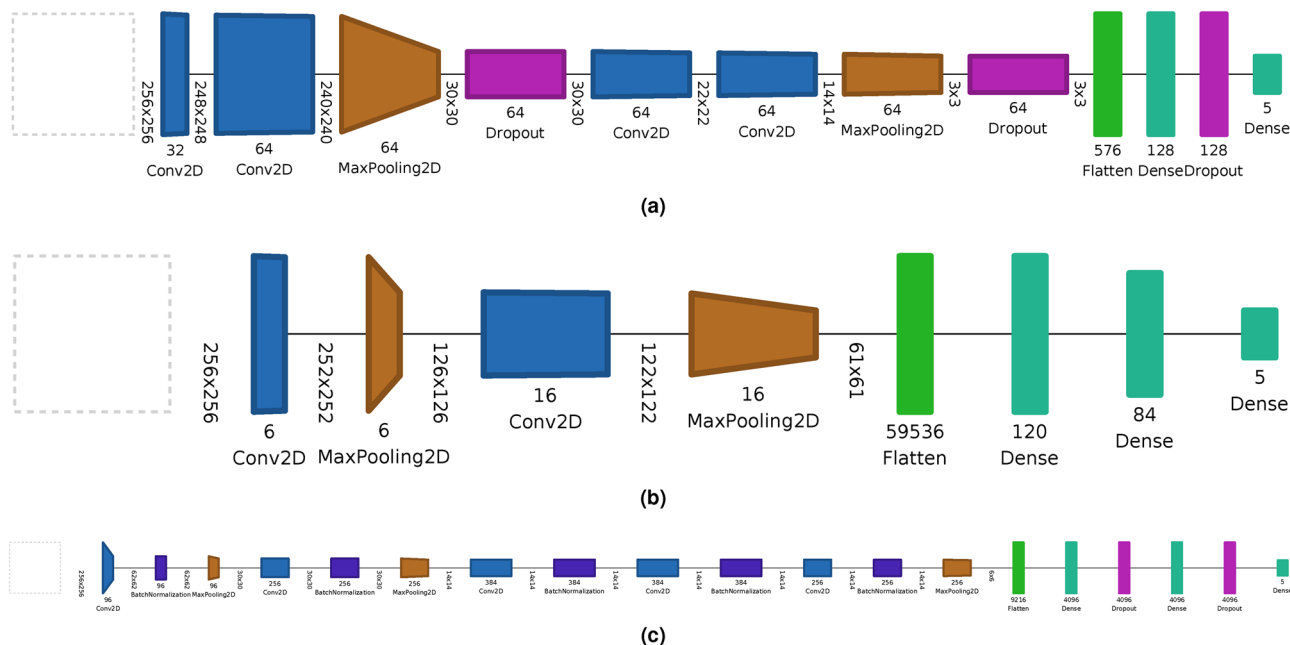


Figure 4. CNN architectures of (a) Custom, (b) LeNet, (c) AlexNet.

Architecture	Trainable parameters	Non-trainable parameters	Total parameters
Custom	906,757	0	906,757
LeNet	7,157,601	0	7,157,601
AlexNet	58,281,349	2752	58,284,101

Table 1. Parameters for each architecture, first column represents the architecture, the rest represents the trainable, non-trainable and total parameters of each architecture.

Convolutional neural network design. Convolutional Neural Network (CNN) usage has increased in recent years since a large number of systems integrate them⁴⁸. Convolutional Neural Network (CNN) identify features from tagged datasets. Such features are often imperceptible to humans. Biometric Recognition Through gAit aNalysis (BRITTANY) uses a classification Convolutional Neural Network (CNN) to identify people. It receives an occupancy map (\mathcal{I}) as input and returns a success rate for each person it was trained. To define the Convolutional Neural Network (CNN), Keras API⁴⁹ have been used using TensorFlow⁵⁰ as back-end.

To carry out the image classification to be performed by BRITTANY, we propose a new neural network architecture, hereafter "custom". To evaluate its performance, we have selected two other well-known architectures to perform image classification LeNet-5⁵¹ and AlexNet⁵². These two architectures have been selected because they are the first well-known architectures that solved image classification problems. Both architectures have been modified to process as input a 256x256 image and as output a Dense layer of 5. Other architectures such as VGG16⁵³ have been tested, but due to the complexity in the deep layers, the model created did not generalize correctly to perform the image classification used by Biometric Recognition Through gAit aNalysis (BRITTANY).

Figure 4, generated with Net2vis tool⁵⁴, illustrates the three architectures, Fig. 4a represents the custom architecture proposed in this work, and Fig. 4b and c represents the architecture of LeNet and Alexnet respectively. In addition, Table 1 shows the number of trainable and non-trainable parameters for each proposed architectures. As can be seen, the custom architecture has a much lower number of trainable parameters than the other two well-known models.

The Convolutional Neural Networks (CNNs) models were trained on Caléndula the parallel computing cluster of Supercomputación Castilla y León (SCAYLE)⁵⁵. Supercomputación Castilla y León (SCAYLE) is a public research centre dependent on the Community of Castilla y León (Spain) whose main activity is to support the improvement of R & D & I tasks. Six Convolutional Neural Network (CNN) models have been created for each architecture, one for each out of 6 aggregation settings ($\mathcal{C}_{5 \times 0}$, $\mathcal{C}_{5 \times 1}$, $\mathcal{C}_{5 \times 2}$, $\mathcal{C}_{10 \times 0}$, $\mathcal{C}_{10 \times 1}$, and $\mathcal{C}_{10 \times 2}$) depicted in "Data curation" section. In this way, a total of 18 Convolutional Neural Network (CNN) models have been trained.

Evaluation. The evaluation was carried out using the \mathcal{D}_2 —see "Data gathering" section. We played each Rosbag file using Biometric Recognition Through gAit aNalysis (BRITTANY) with different setting schemas to evaluate the accuracy. We need to know whether or not the user has been recognized properly for each run. Such data allow us to obtain the confusion matrix that allows visualization of the performance of our tool with the six

setting schemas defined and the three Convolutional Neural Network (CNN) architectures used, described in “Convolutional neural network design” section.

Moreover, to evaluate the overall Biometric Recognition Through gait analysis (BRITTANY)’s performance the following Key Performance Indicators (KPI)’s obtained through the confusion matrix are considered: Accuracy (\mathcal{A}), Precision (\mathcal{P}), Recall (\mathcal{R}), and F₁-score (\mathcal{F}). As the proposed method is a multi-class classification, it is necessary to calculate Key Performance Indicators (KPI)’s for each class. Then an arithmetic average is calculated to obtain the Key Performance Indicators (KPI)’s of the global system⁵⁶. \mathcal{A} —see Eq. (6)—measures the proportion of correct predictions, both positive and negative cases, among the total number of cases examined. \mathcal{A}_k is calculated as shown Eq. (2). The \mathcal{P} score—see Eq. (7)—shows the fraction of relevant instances among the retrieved instances. \mathcal{P}_k is calculated as shown Eq. (3). The \mathcal{R} score—see Eq. (8)—shows the rate of positive cases that were correctly identified by the algorithm. \mathcal{R}_k is calculated as shown Eq. (4). Finally, the \mathcal{F} score—see Eq. (9) is the harmonic mean of precision and recall. \mathcal{F}_k is calculated as shown Eq. (5). In equations, TP represents the true-positive rate, TN is the true-negative rate, FP is the false-positive rate and FN the number of classes into which the model classifies.

$$\mathcal{A}_k = \frac{TP_k + TN_k}{TP_k + FP_k + TN_k + FN_k} \quad (2)$$

$$\mathcal{P}_k = \frac{TP_k}{TP_k + FP_k} \quad (3)$$

$$\mathcal{R}_k = \frac{TP_k}{TP_k + FN_k} \quad (4)$$

$$\mathcal{F}_k = 2 \frac{\mathcal{P}_k \times \mathcal{R}_k}{\mathcal{P}_k + \mathcal{R}_k} \quad (5)$$

$$\mathcal{A} = \frac{\sum_{k=0}^K \mathcal{A}_k}{K} \quad (6)$$

$$\mathcal{P} = \frac{\sum_{k=0}^K \mathcal{P}_k}{K} \quad (7)$$

$$\mathcal{R} = \frac{\sum_{k=0}^K \mathcal{R}_k}{K} \quad (8)$$

$$\mathcal{F} = \frac{\sum_{k=0}^K \mathcal{F}_k}{K} \quad (9)$$

To evaluate the tradeoff between the TP and FP rates of each class, we have computed the Receiver Operating Characteristic (ROC) curve⁵⁷ for each one of the six setting schemas defined and the three proposed architectures. Besides, the Area Under the Curve (AUC) has been calculated to depict how much the models can distinguish between classes.

Results and discussion

Biometric Recognition Through gait analysis (BRITTANY)’s evaluation was done as described in “Evaluation” section. We obtained a confusion matrix for each model (custom, LeNet and AlexNet) and each setting schema ($\mathcal{C}_{5 \times 0}$, $\mathcal{C}_{5 \times 1}$, $\mathcal{C}_{5 \times 2}$, $\mathcal{C}_{10 \times 0}$, $\mathcal{C}_{10 \times 1}$, and $\mathcal{C}_{10 \times 2}$). Figure 6, shows the confusion matrices for each model and setting schema. In the first row are the confusion matrices, in blue, for the custom models, in brown (second row) the confusion matrices for LeNet and in the third row, in green, the confusion matrices for AlexNet. From left ($\mathcal{C}_{5 \times 0}$) to right ($\mathcal{C}_{10 \times 2}$) are the different configurations used. Every confusion matrix consists of rows and columns representing user identifiers. The matrices check situations where BRITTANY provided a correct or wrong outcome using each model respectively. A perfect system would have all the values on the main diagonal. According to the values shown in Fig. 6, the Key Performance Indicators (KPI)’s, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) have been calculated for each of the architectures and setting schemas. These Key Performance Indicators (KPI)’s are presented in Fig. 5 that show the Accuracy (\mathcal{A})—Fig. 5a, Precision (\mathcal{P})—Fig. 5b, Recall (\mathcal{R})—Fig. 5c, and F₁-score (\mathcal{F})—Fig. 5d, for each model proposed and each setting schema. Moreover in Fig. 7 are presented the Receiver Operating Characteristic (ROC) curves for each model and setting schema. In the first row Receiver Operating Characteristic (ROC) curves for the custom models, in the second row Receiver Operating Characteristic (ROC) curves for LeNet and in the third row Receiver Operating Characteristic (ROC) curves for AlexNet. From left ($\mathcal{C}_{5 \times 0}$) to right ($\mathcal{C}_{10 \times 2}$) are the different configurations used. The Receiver Operating Characteristic (ROC) curves have been created for each user in the system (U0 - U4) and the not registered user (!U), in this way, it is also calculated the Area Under the Curve (AUC) for each user, model and setting schema.

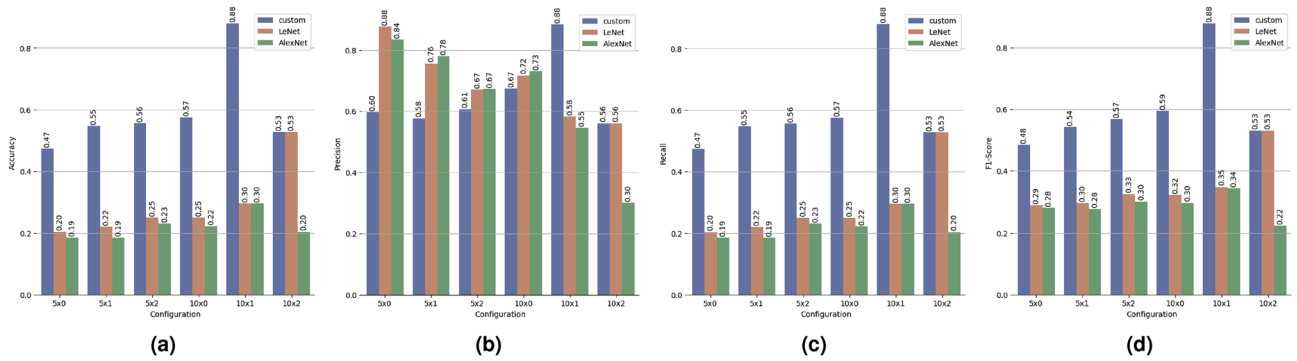


Figure 5. Key Performance Indicators (KPI): Accuracy, Precision, Recall and F1-score, for each of the trained models and each configuration proposed ($\mathcal{C}_{5 \times 0}$, $\mathcal{C}_{5 \times 1}$, $\mathcal{C}_{5 \times 2}$, $\mathcal{C}_{10 \times 0}$, $\mathcal{C}_{10 \times 1}$, and $\mathcal{C}_{10 \times 2}$).

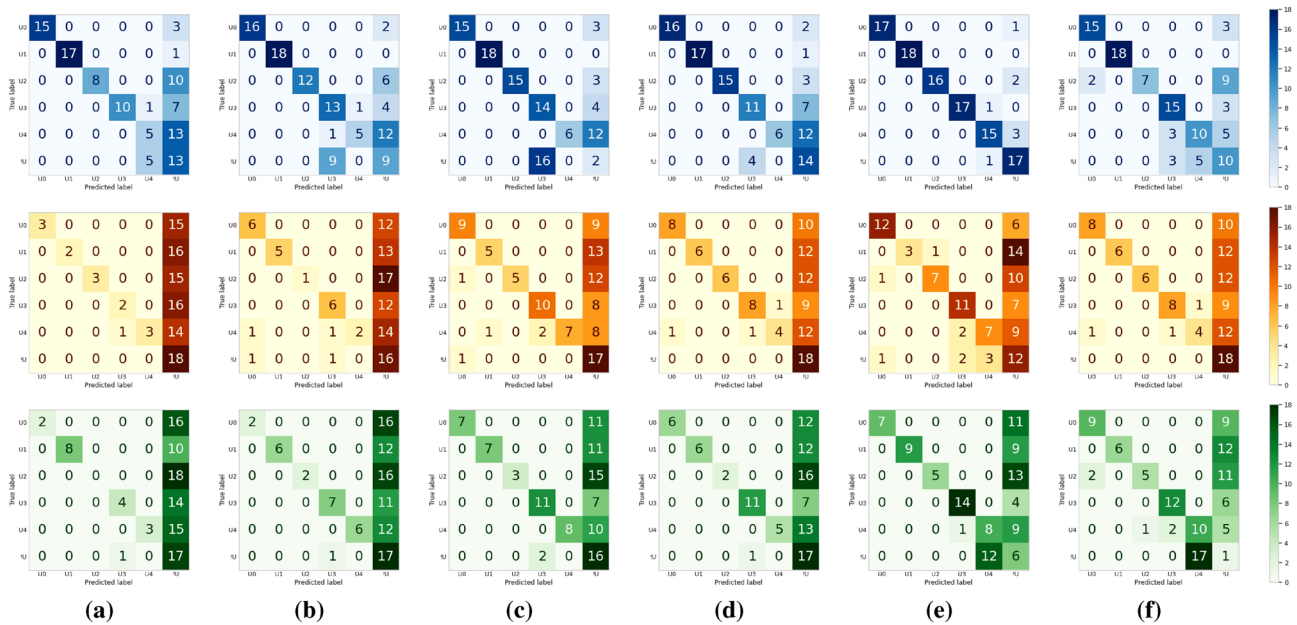


Figure 6. Confusion matrices for each architecture: Custom in blue (first row), LeNet in brown (second row) and AlexNet in green (third row). From left to right the different configurations used: (a) $\mathcal{C}_{5 \times 0}$, (b) $\mathcal{C}_{5 \times 1}$, (c) $\mathcal{C}_{5 \times 2}$, (d) $\mathcal{C}_{10 \times 0}$, (e) $\mathcal{C}_{10 \times 1}$, and (f) $\mathcal{C}_{10 \times 2}$.

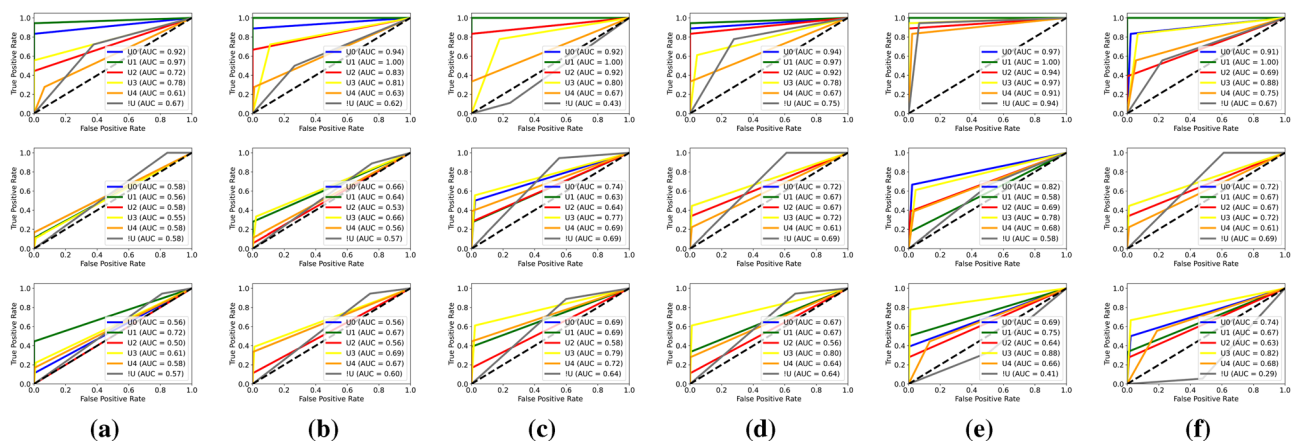


Figure 7. Receiver Operating Characteristic (ROC) curves and Area Under the Curve (AUC) for custom (first row), LeNet (second row) and AlexNet (third row). From left to right the different configurations used: (a) $\mathcal{C}_{5 \times 0}$, (b) $\mathcal{C}_{5 \times 1}$, (c) $\mathcal{C}_{5 \times 2}$, (d) $\mathcal{C}_{10 \times 0}$, (e) $\mathcal{C}_{10 \times 1}$, and (f) $\mathcal{C}_{10 \times 2}$.

Focusing on Fig. 5a, we see that accuracy scores that measure the proportion of correct (both positive and negative) predictions among the total number of cases examined. The accuracy score is only higher than 0.8, specifically 88%, with the $\mathcal{C}_{10 \times 1}$ schema using the custom model. The accuracy scores in the schemas for the custom model are between 47% and 57%. The accuracy of LeNet and AlexNet for each schema is much lower than that obtained with the custom model, obtaining values between 19% and 30%. Except for the $\mathcal{C}_{10 \times 2}$ schema where LeNet obtains the same value as custom (53%).

The precision (\mathcal{P}), see Fig. 5b, represents the fraction of relevant instances among the retrieved samples. In this case, most of the values are higher than 50% for all the models; only the model AlexNet with schema $\mathcal{C}_{10 \times 2}$ reports a precision of 30%. The two models with a better precision are LeNet with the schema $\mathcal{C}_{5 \times 0}$ and custom with the schema $\mathcal{C}_{10 \times 1}$, for both models, the precision score is 88%.

The recall score (\mathcal{R})—also called sensitivity or true positive rate—, see Fig. 5c is the ratio of positive instances correctly detected by the algorithm. In this score, the maximum value, 88%, is again for the custom model and $\mathcal{C}_{10 \times 1}$ schema. The recall scores in the schemas for the custom model are between 47% and 57%. The recall of LeNet and AlexNet for each schema report values between 19% and 53%.

It is often convenient to combine precision and recall into a single metric, see Fig. 5d. The F_1 score (\mathcal{F}) is the harmonic mean of \mathcal{P} and \mathcal{R} . Whereas the regular mean treats all values equally, the harmonic mean gives much more weight to low values. As a result, the classifier will only get a high F_1 score if both recall and precision are high. The model with the best F_1 score is the custom using the schema $\mathcal{C}_{10 \times 1}$, specifically 88%. The remaining schemas for the custom model have a F_1 score between 48% and 59%. The F_1 score of LeNet and AlexNet for each schema is much lower than that obtained with the custom model, obtaining values between 22% and 35%. Except for the $\mathcal{C}_{10 \times 2}$ schema where LeNet obtains the same value as custom (53%).

The model with the best Key Performance Indicators (KPI)'s is custom using the schema $\mathcal{C}_{10 \times 1}$, all the Key Performance Indicators (KPI)'s reports a score of 88%. Then, focusing on the confusion matrix of the custom model and the best schema, $\mathcal{C}_{10 \times 1}$, shown in Fig. 6 (first row, fifth column), we see that most of the values are in the main diagonal. However, there are some errors. U1 was the only correctly identified user in all cases—we have 18 cases, six possible users at three different locations. U0 user was recognised as U0 in 17 cases. However, there is one case where U0 was not recognised as a registered user (!U). U2 was identified correctly in 16 cases, and he was not recognised (!U) in 2. U3 was correctly identified in all cases but one, where he was recognised as U4. It is important to point out that this is the only case in the evaluation of the $\mathcal{C}_{10 \times 1}$ schema and the custom model where two registered users were mistaken. The results for U4 are the worst in the evaluation. In 15 cases, the user was correctly identified, but in 3 cases, he was not recognised (!U). Finally, a non-registered user was used to test the Biometric Recognition Through gait analysis (BRITTANY)'s robustness in front of unknown users. Such a user was wrongly recognised as U3 in 1 case. In the remaining 17 cases, he was correctly identified as non-registered.

Focusing on the Receiver Operating Characteristic (ROC) curves for each of the models obtained, see Fig. 7. It can be seen that the best representation of the Receiver Operating Characteristic (ROC) curves corresponds to the custom model and the $\mathcal{C}_{10 \times 1}$ schema (first row, fifth column). The quality of the model increases when the curve moves towards the upper left corner of the graph. This is because it improves its TP rate, also minimising the FP rate. Moreover, the Area Under the Curve (AUC) values are used as a summary of the model's performance. The more curve moves towards the upper left corner of the graph, the more area is contained under it and therefore, the classifier is better. A perfect classifier has an Area Under the Curve (AUC) of 1. The Area Under the Curve (AUC)'s for each user obtained from BRITTANY using the custom model and the $\mathcal{C}_{10 \times 1}$ schema are (0.97, 1.00, 0.94, 0.97, 0.91 and 0.94) respectively, all of them higher than 90%.

Finally, the accuracy obtained from Biometric Recognition Through gait analysis (BRITTANY) using the custom model and the $\mathcal{C}_{10 \times 1}$ schema is 88%. Focussing on works presented in the Introduction that applied biometry technologies, the work³⁴ creates a silhouette from the point cloud extracted from the multi-line LIDAR and processes it through an LSTM-based CNN network obtaining an accuracy of 60%. The work³⁵ uses gait analysis to do person re-identification, using the silhouette obtained from the projection of the 3D point cloud of a person to a 2D image. This method obtains an accuracy of 87%. It should be noted that those works use a 3D LIDAR instead of a 2D LIDAR. The number of values provided by a 2D LIDAR sensor is much smaller than the ones provided by a 3D LIDAR, so it is easier to process data. This fact facilitates the use of the final system in real-time. Moreover, the models used in those works have not been evaluated against users outside the system. Therefore, people who are not registered in the system could be classified as legitimate users of it.

Conclusion

This paper presents Biometric Recognition Through gait analysis (BRITTANY), a system that identifies people by analyzing their gait. Thus, the system is based on behavioural biometric technologies characterized for analyzing the features of a specific action performed by a person. This system processes sensor data obtained in real-time. A Laser Imaging Detection and Ranging (LIDAR) sensor was chosen because of its low computational demand as well as the privacy it provides. The Laser Imaging Detection and Ranging (LIDAR) readings are processed by People Tracking (PeTra) to create an occupational map segmenting the points that belong to people. Then, such segmented occupational maps are aggregated to build an image of people's gait. Aggregated occupancy maps are processed by a Convolutional Neural Network (CNN) model that outcomes the user identifier that corresponds to the person in front of the robot. The final prediction depends on several estimations by applying a voting strategy to prevent errors. Biometric Recognition Through gait analysis (BRITTANY) can be used in several applications on cooperative robotics and raises Human-Robot Interaction (HRI) since it allows the robot to “know” the people around him. Moreover, at indoor environments, the robot could detect foreigners

and alert their presence. Besides, it also can be used to control access to critical infrastructures, making it more difficult for cybercriminals to carry out tailgating attacks.

The evaluation was done by analyzing several setting schemas named $\mathcal{C}_{n \times s}$ with three Convolutional Neural Network (CNN) architectures, one of them is proposed in this work, called “custom” and the other two are well-known architectures (LeNet and AlexNet). The dataset \mathcal{D}_2 was used to measure the performance. It compiles 108 Rosbag files—Six recordings for six users at three different locations. Five users (U0–4) are well-known people. The last one (!U) is unknown to the system and was used to evaluate the system’s performance in front of strange people. The $\mathcal{C}_{10 \times 1}$ schema using the custom model gets the best results in all the Key Performance Indicators (KPI)’s computed. Such schema builds images by aggregating ten segmented occupancy maps taking one out of two. This schema provides an accuracy score of 88%.

As previously mentioned, this work is a Proof of Concept (PoC) for indoor environments with few people. We aim to determine if it is possible to identify people by their gait using a 2D LIDAR sensor to maintain the user’s privacy and reduce computational load. In future work, we propose to collect data from more users to evaluate Biometric Recognition Through gait analysis (BRITTANY) as an authentication tool.

We want to point out that all the datasets generated are available online⁴⁷. Besides, The source code developed during the current study is available under an open-source license in the GitHub repository⁵⁸. Finally, a docker image with all required software to double-check the evaluation posed in this paper is also available online⁵⁹.

Data availability

The datasets generated and/or analysed during the current study are available in the “Dataset for train and test BRITTANY (Biometric Recognition Through gait analysis)” repository, 10.5281/zenodo.5825885.

Received: 3 May 2022; Accepted: 19 August 2022

Published online: 25 August 2022

References

- Ulsch, M. *Cyber Threat!: How to Manage the Growing Risk of Cyber Attacks* (Wiley Online Library, 2014).
- Raiyn, J. et al. A survey of cyber attack detection strategies. *Int. J. Secur. Appl.* **8**, 247–256 (2014).
- Ogie, R. I. Cyber security incidents on critical infrastructure and industrial networks, in *Proceedings of the 9th International Conference on Computer and Automation Engineering*, 254–258 (2017).
- Salahdine, F. & Kaabouch, N. Social engineering attacks: A survey. *Future Internet* **11**, 89 (2019).
- Maan, P. & Sharma, M. Social engineering: A partial technical attack. *Int. J. Comput. Sci. Issues* **9**, 557–559 (2012).
- Panchpor, A. A., Shue, S. & Conrad, J. M. A survey of methods for mobile robot localization and mapping in dynamic indoor environments, in *2018 Conference on Signal Processing and Communication Engineering Systems (SPACES)*, 138–144 (IEEE, 2018).
- Cheng, J., Cheng, H., Meng, M. Q.-H. & Zhang, H. Autonomous navigation by mobile robots in human environments: A survey, in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 1981–1986 (IEEE, 2018).
- Richert, A., Müller, S., Schröder, S. & Jeschke, S. Anthropomorphism in social robotics: Empirical results on human–robot interaction in hybrid production workplaces. *AI Soc.* **33**, 413–424 (2018).
- Fox, J. & Gambino, A. Relationship development with humanoid social robots: Applying interpersonal theories to human–robot interaction. *Cyberpsychol. Behav. Soc. Netw.* **24**, 294–299 (2021).
- Lobato, C., Vega-Magro, A., Núñez, P. & Manso, L. J. Human-robot dialogue and collaboration for social navigation in crowded environments, in *2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, 1–6 (IEEE, 2019).
- Wang, Y., Wei, X., Shen, H., Ding, L. & Wan, J. Robust fusion for RGB-D tracking using CNN features. *Appl. Soft Comput.* **92**, 106302 (2020).
- Matti, D., Ekenel, H. K. & Thiran, J.-P. Combining lidar space clustering and convolutional neural networks for pedestrian detection, in *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 1–6 (IEEE, 2017).
- Bu, F., Le, T., Du, X., Vasudevan, R. & Johnson-Roberson, M. Pedestrian planar lidar pose (PPLP) network for oriented pedestrian detection based on planar lidar and monocular images. *IEEE Robot. Autom. Lett.* **5**, 1626–1633 (2019).
- Linder, T. & Arras, K. O. People detection, tracking and visualization using ROS on a mobile service robot. In *Robot Operating System (ROS)* (ed. Koubaa, A.) 187–213 (Springer, 2016).
- Rios-Martinez, J., Spalanzani, A. & Laugier, C. From proxemics theory to socially-aware navigation: A survey. *Int. J. Soc. Robot.* **7**, 137–153. <https://doi.org/10.1007/s12369-014-0251-1> (2015).
- Duong, H. T. & Suh, Y. S. Human gait tracking for normal people and walker users using a 2D lidar. *IEEE Sens. J.* **20**, 6191–6199 (2020).
- Guerrero-Higuera, Á. M. et al. Tracking people in a mobile robot from 2D lidar scans using full convolutional neural networks for security in cluttered environments. *Front. Neurobot.* **12**, 85 (2018).
- Álvarez-Aparicio, C. et al. People detection and tracking using LIDAR sensors. *Robotics* **8**, 75 (2019).
- Álvarez-Aparicio, C. et al. Convolutional neural networks refitting by bootstrapping for tracking people in a mobile robot. *Appl. Sci.* **11**, 10043 (2021).
- Jain, A., Hong, L. & Pankanti, S. Biometric identification. *Commun. ACM* **43**, 90–98 (2000).
- Adeoye, O. S. A survey of emerging biometric technologies. *Int. J. Comput. Appl.* **9**, 1–5 (2010).
- Valdes-Ramirez, D. et al. A review of fingerprint feature representations and their applications for latent fingerprint identification: Trends and evaluation. *IEEE Access* **7**, 48484–48499 (2019).
- Mane, S. & Shah, G. Facial recognition, expression recognition, and gender identification. In *Data Management, Analytics and Innovation* (eds Balas, V. et al.) 275–290 (Springer, 2019).
- Patil, C. M. & Gowda, S. An approach for secure identification and authentication for biometrics using iris, in *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*, 421–424 (IEEE, 2017).
- Obaidat, M. S., Traore, I. & Woungang, I. *Biometric-Based Physical and Cybersecurity Systems* Vol. 368 (Springer, 2019).
- Xiao, Q. Security issues in biometric authentication, in *Proceedings from the Sixth Annual IEEE SMC Information Assurance Workshop*, 8–13 (IEEE, 2005).
- Menshakou, P. & Murashko, I. Voice user identification in access control systems, in *Proceedings of International Conference on Open Semantic Technologies for Intelligent Systems (OSTIS-2017)*. Minsk: BSUIR, 175–178 (2017).
- Mohammed, A. A., Abdul-Hassan, A. K. & Mahdi, B. S. Authentication system based on hand writing recognition, in *2019 2nd Scientific Conference of Computer Sciences (SCCS)*, 138–142 (IEEE, 2019).
- Connor, P. & Ross, A. Biometric recognition by gait: A survey of modalities and features. *Comput. Vis. Image Underst.* **167**, 1–27 (2018).

30. Petrescu, R. V. Face recognition as a biometric application. *J. Mechatron. Robot.* **3**, 237–257 (2019).
31. Ouellet, S., Grondin, F., Leconte, F. & Michaud, F. Multimodal biometric identification system for mobile robots combining human metrology to face recognition and speaker identification, in *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 323–328 (2014). <https://doi.org/10.1109/ROMAN.2014.6926273>.
32. Al-Qaderi, M. K. & Rad, A. B. A multi-modal person recognition system for social robots. *Appl. Sci.* **8**, 387 (2018).
33. Karri, C., Cheikhrouhou, O., Harbaoui, A., Zaguia, A. & Hamam, H. Privacy preserving face recognition in cloud robotics: A comparative study. *Appl. Sci.* **11**, 6522 (2021).
34. Yamada, H., Ahn, J., Mozos, O. M., Iwashita, Y. & Kurazume, R. Gait-based person identification using 3D lidar and long short-term memory deep networks. *Adv. Robot.* **34**, 1201–1211 (2020).
35. Benedek, C., Gálai, B., Nagy, B. & Jankó, Z. Lidar-based gait analysis and activity recognition in a 4D surveillance system. *IEEE Trans. Circuits Syst. Video Technol.* **28**, 101–113 (2016).
36. Hasan, M. *et al.* Tracking people using ankle-level 2D lidar for gait analysis, in *International Conference on Applied Human Factors and Ergonomics*, 40–46 (Springer, 2020).
37. Robotics Group of Universidad de León. Leon@home Testbed (accessed 28 April 2022). <https://robotica.unileon.es/index.php?title=Testbed> (2022).
38. EU Robotics. ERL Certified Test Beds (accessed 28 April 2022). https://www.eu-robotics.net/robotics_league/erl-service/certified-test-beds/index.html (2022).
39. Robotnik. Robotnik homepage (accessed 28 April 2022). <https://robotnik.eu/es/> (2022).
40. Quigley, M. *et al.* ROS: An open-source robot operating system, in *ICRA Workshop on Open Source Software*, vol. 3, 5 (Kobe, Japan, 2009).
41. Robotics Group of Universidad de León. Robotics Group homepage (accessed 28 April 2022). <https://robotica.unileon.es/> (2022).
42. Álvarez-Aparicio, C. *et al.* Benchmark dataset for evaluation of range-based people tracker classifiers in mobile robots. *Front. Neurobot.* **11**, 72 (2018).
43. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, vol. 9351 of LNCS, 234–241 (Springer, 2015).
44. Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431–3440 (2015). <https://doi.org/10.1109/CVPR.2015.7298965>.
45. Álvarez-Aparicio, C. *et al.* Lidar-based people detection and tracking for home competitions, in *2019 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, 1–6 (IEEE, 2019).
46. Álvarez-Aparicio, C., Guerrero-Higueras, A. M., Rodr, F. J. & Mart, F. Detección y seguimiento de personas utilizando sensores lidar, in *Jornadas Nacionales de Robótica Spanish Robotics Conference*, 234–240 (2019).
47. Dataset 10.5281/zenodo.5825885 (accessed 12 January 2022). <https://doi.org/10.5281/zenodo.5825885> (2022).
48. Mane, D. & Kulkarni, U. V. A survey on supervised convolutional neural network and its major applications, in *Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications*, 1058–1071 (IGI Global, 2020).
49. Keras. Keras homepage (accessed 28 April 2022). <https://keras.io/> (2022).
50. TensorFlow. TensorFlow homepage (accessed 28 April 2022). <https://www.tensorflow.org/> (2022).
51. LeCun, Y. *et al.* Lenet-5, convolutional neural networks. <http://yann.lecun.com/exdb/lenet20>, 14 (2015).
52. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **25**, 1–9 (2012).
53. Simonyan, K. & Zisserman, A. *Very deep convolutional networks for large-scale image recognition*. arXiv preprint [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2014).
54. Bäuerle, A., van Onzenoort, C. & Ropinski, T. Net2vis—A visual grammar for automatically generating publication-tailored CNN architecture visualizations. *IEEE Trans. Visual Comput. Graph.* **27**, 2980–2991. <https://doi.org/10.1109/TVCG.2021.3057483> (2021).
55. SCAYLE. Scayle (accessed 28 April 2022) (2022). <https://www.scayle.es/>.
56. Grandini, M., Bagli, E. & Visani, G. Metrics for multi-class classification: An overview. arXiv preprint [arXiv:2008.05756](https://arxiv.org/abs/2008.05756) (2020).
57. Fawcett, T. An introduction to roc analysis. *Pattern Recogn. Lett.* **27**, 861–874 (2006).
58. C.Álvarez-Aparicio. BRITTANY (Biometric Recognition Through gAitaNaLYsis) (accessed 28 April 2022) (2022). <https://github.com/ClaudiaAlvarezAparicio/brittany>.
59. C.Álvarez-Aparicio. BRITTANY Docker image (accessed 28 April 2022) (2022). <https://hub.docker.com/r/claudiaalvarezaparcio/brittany>.

Acknowledgements

The research described in this article has been funded by the Instituto Nacional de Ciberseguridad de España (INCIBE), under the grant "ADENDA 4: Detección de nuevas amenazas y patrones desconocidos (Red Regional de Ciencia y Tecnología)", addendum to the framework agreement INCIBE-Universidad de León, 2019-2021. Miguel Ángel González-Santamarta would like to thank Universidad de León for its funding support for his doctoral studies.

Author contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by C.Á-A, Á.M.G-H, M.Á.G-S and A.C-V. The first draft of the manuscript was written by C.Á-A and Á.M.G-H and all authors commented on previous versions of the manuscript. The supervision was done by V.M. The project management and the funding acquisition were supplied by C.F-L. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.Á.-A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022