

Person Identification by Evaluating Gait using 2D LiDAR and Deep Neural Network

Mahmudul Hasan (✉ hasan@hci.ics.saitama-u.ac.jp)

Saitama University <https://orcid.org/0000-0002-6274-4107>

Junichi Hanawa

Saitama University: Saitama Daigaku

Riku Goto

Saitama University: Saitama Daigaku

Hisato Fukuda

Gunma University: Gunma Daigaku

Yoshinori Kuno

Saitama University: Saitama Daigaku

Yoshinori Kobayashi

Saitama University: Saitama Daigaku

Research Article

Keywords: person identification, person recognition, gait estimation, 2D LiDAR, ResNet

Posted Date: March 22nd, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1425488/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Video-based recognition techniques are solemnly effective, and it comes to a new era of research nowadays. Yet again, it suffers some bottlenecks indeed. Situations, surroundings, and momentums may be disgraceful with all new inventions. So, to solve the drawbacks of technology is to imply a new technology on it. Biometric features are very authentic and high valued measures for human identifications. Most of the techniques are dependent on close contact with the subject. A gait is a pattern that performs by walking from the individual. Almost all studies of gait-based person identifications are performed by RGB or RGB-D cameras. Very few studies were done by using LiDAR data. Applying 2D LiDAR images for individual tracking and identification is superb when video surveillances fail to perform accurately due to environmental and imposed difficulties (i.e., disaster, rain, fog, smoke, snow, occlusion, cost, etc.). This research performed a comprehensive exhibition of 2D LiDAR data with a rigorous self-made dataset and customized residual neural network. We considered different experimental setups and found exciting precisions there. Our system is appropriate for recognizing a person based on his ankle level 2D LiDAR data.

1. Introduction

Person identification is a vast and ancient research field. Various techniques have been invented in this arena. Diversified biometric [1,2,27] characteristics enhanced the accuracy of this vital space. Camera-based applications eventually led to this research. Sophisticated innovations, modernized features, and computational capabilities make this exact day by day. But, privacy, natural calamities, etc. issues are a big concern for video-based processing. The emergence of new biometric features meticulously crafted the credibility of human recognition in different delicate applications. Analogously, the urgency of close contact with the devices down worth the performance of biometric identifications in some cases. Hence gait recognition is an apt alternate to person identification where subjects are not supposed to be nearer with devices. In all these innovations, video cameras [2] were used as key identifiers to demonstrate individuals' states and distinctiveness.

Some crucial circumstances and feasible benefits over traditional camera-based applications make gait more popular these days. Gait is a way of walking, and gait recognition refers to identifying a person based on their walking style. A very plausible gain 'remote access' over close camera contact is well accepted to all. Furthermore, without the cooperation of the subject, it can be recognized. Even if some biometric features (i.e., face, fingerprint, iris, etc.) are absent or cannot be identified, Gait plays a vital role in identifications. It is tough to impersonate the gait features, making it essential in crime analysis. However, video-based recognition suffers from low resolution capturing, natural disasters, computational complexities, etc. We propose a new modality with gait data analysis: ankle level 2D LiDAR-based person tracking and recognition. Recent enhancements on computational capabilities, a deep learning approach make this research more precise than ever.

We demonstrated a person tracking [5] system using 2D LiDAR in ankle level for gait analysis. Determining moving objects in front of the LiDAR sensor as a person was a challenging job. We fastidiously made the way of tracking by imposing density-based clustering over conventional approaches. This research started using multivariate density-based algorithms to fit the model best. A visualized tracking based on LiDAR data was undoubtedly challenging, and we prudently did the same. The development of density-based algorithms to augment the performance of the tracking was another challenge in this experiment. We proposed two new algorithms [6]: Enhanced Density Based Scan (EDBSCAN) and Enhanced Ordering Points To Identify the Clustering Structure (EOPTICS) to create the best cluster to determine an individuals' ankle positions and identify their way of walking. This approach enhanced the performance of our previous tracking system and showed a new person tracking system based on only 2D LiDAR data. This method also helped us find an alternative of surveillance cameras for which people were worried about their privacy and secrecy. Here, the necessity of tracking and confidentiality intent meets a unique solution. We found some influential features (i.e., age, height, etc.) can be measured by only ankle movement data, especially tracking data. We have broadened our study with person property estimation by a 2D LiDAR sensor [7]. Here we used a deep neural network for training and testing the model. We prepared a detailed dataset of the experiments with different ethnicities, gender, and height. A parametric formulation was performed to do the experiments, and results were distinctly identified. The outcomes of the trials were remarkable and trustworthy to compare with the actual ones.

Now we go up with some practical instincts that suffer from RGB/RGB-D cameras for illusion, illumination, natural calamities, and even real-time computational inaccuracies. LiDAR-based person identification is our new research goal that comprehensively deals with all inadequacies of visual imaging. We placed multiple 2d LiDAR sensors at ankle level to acquire the data. LiDARs emanate pulsed light on the surrounding objects and compute the distance it traveled to receive the sensor again. We set an experimental model where different LiDAR sensors were placed at ankle level, and persons were moving in front of the sensors. A Robot Operating System (ROS) was used for capturing the time series data in a bag file. We plotted these distance data in images with a specific rate to create motion history images (MHI). With these MHI, we made our datasets to perform the experiments. Continuous ankle movements on a surface develop a path of walking, which determines the tracking system efficiently. As all persons have distinct gestures and ways of walking, especially activities of ankles are significantly unique, diverse us to develop a gait-based person recognition system.

In figure 1, we have shown a block diagram that clarifies the system overview in brief. A sensing module collects data from the LiDAR sensor and passes these distance data to create Motion history Images. These are our input datasets. These Image datasets were used to estimate the

person property [7] through a residual neural network. We previously developed a tracking system [5] based on our modified density-based clustering approaches EDBSCAN and EOPTICS [6]. Now we developed our recognition system to identify a person through gait analysis.

2. Related Works

A thorough overview of recent innovations in person identification will be discussed in this section. We studied many recent articles on individual recognition. Most of the research was performed with video-based analysis. We tried to figure out the techniques from there. A ubiquitous approach is face recognition. Many research used this biometric feature over the years for person identification. Another biometric invention made this accurate study day by day. We have gone through gait-based identification also. As not many efforts have been put into LiDAR-based identification, we found very few on it.

2.1: Person recognition by face

Person identification research started significantly earlier. Some misleading information deteriorated its performance [8]. Sometimes prior knowledge misleads the recognition. The research continues. Many new approaches applied to it. A multimodal audio-video-based approach [9] was also used for person recognition. Individuals' face recognition and speaker identification technique were merged to improve the accuracy of the recognition. Facial video information was also applied to recognize a person [10]. Temporal information (i.e., eigenface, fisherface, elastic graph matching, etc.) was used for the purpose. The Grassmann manifold technique is a recent approach to recognize a person from a video [11]. Here a geometry-aware dimension reduction was achieved from the original Grassmann manifold that enabled propitious classifications. A Neural Person Search Machine [12] was invented for the person searching using recursive localization. A better optimization was possible for this local searching on the image. Still, ideal normalizations are needed in face recognition research. A dataset independent and automated standardization was desirable that eventually became less information lossy. Xiangyu Zhu et al. [13] depicted a method (HPEN: High-Fidelity Pose and Expression Normalization) to create a front view typical facial image. In time, this approach improved the performance of face recognition to identify a person. With the emergence of deep learning [26], face recognition research is now more advanced and accurate. The major invention was DeepFace [14], where a nine-layer deep neural network was used for piecewise transformation to represent a face. A specific dataset achieved about 97% accuracy, which was a great initiation of face recognition. Later, FaceNet [15] presented a concept of automatic learning, a mapping method from face images to measure face similarity. A deep convolutional network was used for training the model. A new record accuracy of 99% was achieved with the same dataset used previously. MagFace [16] is another global representation to recognize a face. It categorized the losses which learn general feature embedding of an input image even though the leading benchmark was not achieved.

2.2: Person recognition by Gait

Every human being has some different characteristics. It is this distinctive feature that sets people apart from others. Although people usually don't see faces from a distance, they can recognize people based on their walking. This method is commonly known as gait. Gait is an active method that distinguishes each animal according to its body movements and walking patterns. Many scientists have tried to justify this discovery. We will try to describe something like this here.

Gait research is not a new initiation in today's research. It is approximately more than 400 years prior idea. Gait itself can be categorized into different types. Human gait representation is usually categorized into model-based and the id model-free. It also can be different based on human size, age, clothes, religion, shoes, ethnicity, accident, etc. With the development of information technology, especially computer vision, automated gait recognition features were developed, increasing its performance rapidly. Silhouette analysis [17] was a tricky but straightforward way of gait recognition to identify a human. Some standard statistical tools, i.e., PCA was used to implement the technique. Later, a gait recognition technique was proposed without the subject's cooperation [18]. This invention lifts the ideas of contemporary gait research. In an unknown circumstance, this idea was more well-performing than others. A dynamics normalization [19] was used in research to identify a person in gait recognition. Comprehensive implementation of the Hidden Markov Model was done to improve the performance. Static (body shape) and dynamic (arm and leg movements) body features are the key influencer of automatic gait recognition. Person identification [20] was performed on gait recognition based on these dynamic body features.

Human walking speed may vary in different situations. It also influences gait analysis. A traditional gait recognition system will underfit these speed transition situations. A gait identification approach was proposed to cope with these situations that deal with this speed change situation [21]. Sometimes gait is chaos with body structures that always makes confusion. Two-Point Gait [22] dissociates these two issues, which considerably process the limb motion instead of body structure. Only dynamic parameters can effectively implement a gait recognition system where body shape may be ignored in the real-time calculation. This research showed a way of using ankle level sensors in our study. Deep Convolutional Networks (CNN) [28] are recent learning techniques that improve the recognition's efficacy. One of the very first works of CNN deals with similarity learning for gait-based identification [23]. Here experimental results outperform the existing evolutions in a good score. The Koopman operator theory has recently been applied in cross-view gait recognition [24]. Universal deep linear embedding was applied

here with a large public dataset, and the performance was impressive. Rather than using the whole body, a part-based model GaitPart [25], corresponds with the different body parts individually to boost performance. All research mentioned above was based on RGB/RGB-D camera. In the next section, we describe one type of research performed using LiDAR data. A 4D visualization and tracking system was proposed there.

2.3: Person recognition using LiDAR data

As so much research has been done to identify a person in different parameters and biometric features, in most of the cases, video cameras and images were their main analyzing inputs. An alternative to the camera (RGB, depth, or Infrared), the LiDAR sensor was rarely used for recognition. We found insufficient research on the topic and thought many more must be put on it. A gait-based person identification with a 3D LiDAR sensor [3] using an LSTM network was performed by Yamada, H. et al. very recently. The average accuracy (60%) and processing of high-volume data (3D LiDAR) also influence us to diversify the research domain. Benedek et al.'s [4] research were one of the first initiatives with LiDAR sensors to recognize a person giving out with gait analysis. They used a Rotating Multi-Beam (RMB) LiDAR sensor, which eventually increased the processing complexity, and correctness was not much passable, which has a vast space to improve. We have tried to come up with a timely solution to these problems. That's why we used a 2D LiDAR sensor for data compilation so that the processing time is minimal and the cost becomes affordable. Parallely the accuracy is impressive to rely on the system.

We found a very initial initiative with Rotating Multi-Beam (RMB) LiDAR [29] to facilitate gait analysis. A 2D-LiDAR based gait analysis [30] system was proposed very recently. The study showed only walking direction with motion captured by the sensor, where individual tracking and identification were not performed. 3D LiDAR is a modern trend used in human detection [31]. Point cloud clusters and classifications were applied for this purpose. With the variation of the distance from the sensor, its accuracy was not so impressive. 3D LiDAR sensors also analyzed person behavior measurement [32]. For a service robot application, this system allows it to interact with people more accurately. To the best of our knowledge, our proposed 2D LiDAR-based person tracking and identification is the initial initiative to do such. We believe this system has a broad scope in real applications with rigorous outcomes.

3. Proposed Method

The fundamental concept of this research is to introduce 2D LiDAR sensors as an alternative to 3D Lidar sensors that comprehensively minimize the intrinsic and computational cost also enhance the system integrity on a big scale. Our previous proposed methods were a step-by-step enhancement of LiDAR-based individual following and their property estimation. We were sensible to use video cameras, often liable for compromise privacy issues. Moreover, some natural and environmental shortcomings depleted the performance of RGB/RGB-D cameras.

3.1: Integrated System overview

In Fig. 2, an overall system diagram is presented here. 2D LiDAR sensor is placed at ankle level to acquire the data. All time-series data are plotted on blank images with 40 frames per second rate. We named these images 'Motion History Image (MHI)'. The main benefit of MHI is to encode a range of time data in a single frame. Thus human gestures and movements can be represented by MHI spans [33]. An update function $\mu(x, y, t_i)$ could calculate the MHI $M_{\mathcal{F}}(x, y, t_i)$:

$$M_{\mathcal{F}}(x, y, t_i) = \begin{cases} \mathcal{E}, & \text{if } \mu(x, y, t_i) = 1 \\ \max(0, M_{\mathcal{F}}(x, y, t_i - 1) - \phi), & \text{otherwise} \end{cases} \quad (1)$$

Here, (x, y) is the position and t_i is the time; $\mu(x, y, t_i)$ shows ankle position or motion in the current frame. The temporal extent of the movement is decided by the duration \mathcal{E} , where ϕ indicates decay in the images.

These MHI are considered as input of our system. Besides traditional clustering approaches, we used modified density-based clustering techniques to determine an ankle of a person. Similarly, the same clustering approaches were used to determine a person based on closely moving two ankles on the plane. This heuristic approach was analyzed repeatedly until it came to tolerable accuracy. The second part of image 2 is clustering which elaborated scenario is shown in Fig. 3. We developed in our earlier article two enhanced algorithms (EDBSCAN and EOPTICS) to thoroughly evaluate the clusters and define the person's position on the plane. We used the Kalman filter for tracking the persons moving in front of the sensor. Our previously published article [5] introduced person tracking based on LiDAR. We enhanced density-based clustering algorithms in our nest article [6] and showed a significant improvement over the traditional density-based approach. Further, we developed a well-structured dataset considering all international audiences that were perfectly diversified with height, age, gender, and ethnicity.

Our next attempt was person property estimation (age and height) from 2D LiDAR data [7]. Though these abstract properties are very susceptible to actual calculation, we created a model. Our experimented results were very competitive to others' estimates with a loose contact of the subject. Figure 4 is a pictorial representation of the property estimation technique. We used a pre-trained Deep Residual Network for

training the system. Here, we got a fantastic outcome with a sensible dataset with rigorous training, validation, and property height and age test data. Our designed system categorized the properties in only two classes (Height: tall, short; Age: Young, Elder), but for the test, we demonstrated more categories also. Next section, we describe our current interest, gait-based identification. Further, we elaborate our study for person recognition based on 2D LiDAR only. Such types of research were not initiated previously with a 2D sensor only.

3.2: Gait-based Person Identification

Person identification deals with much research over so many years. Other biometric features-based identification made this research more authentic and accurate. Various types of cameras and sensors were used to make this perfect. With the invention of the deep neural network, this research got a new dimension. Our focus on this research is to contribute differently to the sense of sensor and calculation.

3.2.1: Experimental Setup

Figure 5 shows the experimental setup for the research. We use four LiDAR sensors to collect the data at different angles and heights. Two LiDAR stands are placed at a two-meter distance, and the angled gap is 90 degrees. We put four various LiDAR sensors in the two stands so that everyone's data can be collected easily. We named two LiDARs in the same stand as multi-layer, and LiDARs in separate stands are named multi-angle. For the analysis, we used different sensors' data in different ways. For this experiment, we used HOKUYO UTM-30LX 2D LiDAR sensors. It has a 30 meter and 270-degree scanning range. This sensor is lightweight and very suitable to use outdoor. We will discuss these in the result discussion section.

3.2.2: Dataset Preparation

In this research work, we used our developed dataset. No such research was done with this sensor setup, so no public dataset is available. 2D LiDAR sensors only provide distance values of objects in front of it, making our image dataset challenging. We created different datasets based on enormous parameters and named these KoLaSU (Kobayashi Laboratory of Saitama University). Twenty-nine independent observers have participated in our study, and all of them were unbiasedly allowed to walk in front of the LiDAR sensors. We created one aggregated dataset considering all four LiDAR sensors. Then individual four sensors' data. After that, we created a multi-layer and multi-angle dataset considering all possible conditions, i.e., multi-layer-13, multi-layer-24, multi-angle-12, multi-angle-14, etc. Here 1,2,3, and 4 indicate LiDAR position as per Fig. 5. Different colors in the MHI indicate various LiDAR sensors' data in the images. Figure 6 shows a brief description of our prepared dataset. Here seven chronological movements of two participants are shown with our created images.

3.2.3: ResNet based identification

Figure 7, showing detail of gait-based identification. As mentioned earlier, we created motion history images (MHI) from LiDAR data given as an input of the neural network. We categorized different ways and trained our model for the experiments. A pre-trained residual neural network (ResNet) was applied to validate and test the system. Identity mapping vectors and residual learning extract convolutional features by training a ResNet. In a ResNet architecture, residual can be defined by

$$Y = f(x) + x$$

Here, x is the input vector, Y is the output vector, and $f(x)$ is the residual mapping function. In this article, we used 18 layers and 35 layer ResNet to extract the convolutional features of KoLaSU datasets. The second layer of the network is Max pooling comes after the first convolutional layer is used to minimize the data overfitting problem. A Fully Connected (FC) layer combined with Average Pooling and Softmax layers provides features of person identification based on gait data. Without raising the training error rate, a substantial number of layers can be trained by ResNet easily. ResNet is also competent to nullify the vanishing gradient problem, giving it an extra advantage over other traditional neural networks. Finally, a consistent gait-based classification was done with meaningful accuracy.

4. Experiments And Discussion

LiDAR-based identification is a recent yet challenging initiation considering various circumstances and precisions. Few comprehensive studies have begun with the parallel of conventional video-based tracking with the LiDAR sensor. But all these studies considered 3D LiDAR as their alternative. We thought differently and used a 2D LiDAR sensor to do so. Considerable outcomes, less processing time, inexpensive installation, and wide practical applications made this research worthy.

4.1: Gait based person identification

We thought of the different data combinations during acquisition to make the application-wide in range and reliability. A homogeneous and heterogeneous LiDAR setup was considered for the collection of data. Though a single LiDAR is sufficient to get a pedestrian's desired data, critically distributed and overlapped data are always dilute the system performance. Ankle level LiDAR setup was our primary focus to track and identify a pedestrian. Besides, the uppermost setup of the sensors does not enhance the system performance when the measurements consider a two-dimensional scenario. We placed two LiDAR sensors on a single tripod; parallelly, another two sensors were placed on the second tripod. These two tripods were kept 2 meters distance in a ninety-degree angle experiment setup. We considered different ages and heights pedestrians. They were unbiased in gender, clothes, shoes, and region also. We substantially found numerous walking styles based on their properties. To plot LiDAR data on an image was our primary challenge, but we cautiously did the experiments and found excellent accuracy there. Our proposed KoLaSU dataset consists of fourteen outdoor sequences where twenty-nine participants were attended with five to ten minutes walking with the experiment setup. We considered a standard forty fps rate to write the data. For cross-validation, we considered the 100 fps rate also. In all our experiments, we kept our datasets into three groups. Sixty percent of the data were considered train data, where the remaining 40 percent were split equally into test and validation sets.

Table 1 shows an overall experimental result of gait-based recognition. As discussed, we considered 14 different conditions data for our KoLaSU person tracking dataset; here, we placed 9 of these. We split out all four LiDAR's data individually as we kept four top rows in table 1. Similarly, in the table, we merged different LiDAR data as the number assigned (i.e., LiDAR 12, merged LiDAR 1 and LiDAR 2's data in MHI with 40 fps). Among 29 participants, we keep 26 pedestrians for our experiment after the filter. Here 60 percent of the total data of a set were kept for training the system, and the remaining 40 percent were divided equally for validation and test dataset. For this research, we used the GIGABYTE BRIX GPU machine for the analysis of our data. The batch size was considered as 38, and the number of epochs varied from 25 to 50 periodically. We used a deep neural network to train our model. Here a pre-trained ResNet 18 network was used to train the dataset. We used ResNet 34 and ResNet 50 also for cross-checking the system performance. We placed some results of these in the next section. We randomly selected our train, test, and validation datasets among all machine-generated data, where all segments were utterly disjoint. Initially, we kept every person's data in three parts: train, test, and validation group. Further, we enhanced our study for unknown test data set. Without prior information, man or machine cannot identify a new individual; thus, the system reacts. From table 1, we see that accuracy in three different segments is very impressive and near about 99 percent correctly identified. Though some data are not accurately captured due to congestion problems (i.e., KoLaSU LiDAR 3 and KoLaSU LiDAR 4), their performance in the test case did not go below 93 percent. Moreover, their combined dataset (KoLaSU LiDAR 34) performed significantly well with 99 percent precession.

Table 1
Gait-based person identification on different parameters

| Data | Experiments Type | Batch Size | Epoch Size | GPU | DNN Model | Train Accuracy | Train Loss | Test Accuracy | Test Loss | Validation Accuracy | Validation Loss |
|-------------------------|---|---------------|---------------|-----|--------------|-------------------|---------------|------------------|--------------|------------------------|--------------------|
| KoLaSU LiDAR 1 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.99421 | 0.02 | 0.9843 | 0.05 | 0.9851 | 0.04718 |
| KoLaSU LiDAR 2 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.99324 | 0.02329 | 0.9846 | 0.0494 | 0.98502 | 0.04515 |
| KoLaSU LiDAR 3 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.97721 | 0.00388 | 0.9336 | 0.2144 | 0.93478 | 0.2078 |
| KoLaSU LiDAR 4 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.98038 | 0.0627 | 0.9479 | 0.1689 | 0.94711 | 0.16857 |
| KoLaSU LiDAR 13 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.9982 | 0.0076 | 0.9962 | 0.0132 | 0.996 | 0.0145 |
| KoLaSU LiDAR 24 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.99831 | 0.00717 | 0.99622 | 0.01315 | 0.99611 | 0.013 |
| KoLaSU LiDAR 12 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.99719 | 0.00033 | 0.9934 | 0.0208 | 0.99305 | 0.0208 |
| KoLaSU LiDAR 34 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.99821 | 0.00772 | 0.99343 | 0.0209 | 0.9942 | 0.0198 |
| KoLaSU LiDAR 1234 | 26 Persons Individual (60%,20%,20%) | 38 | 25 | Yes | Resnet 18 | 0.99869 | 0.00561 | 0.99658 | 0.01118 | 0.99713 | 0.0095 |

4.2 Comparison to different data types

Figure 8 shows three different input datasets' accuracy with a different dataset, also showing designed network is accurately trained with all kinds of data. To reduce the overfitting, we validated the network, and here its accuracy is impressive, and none of the datasets goes below 93 percent. The test accuracy also goes through with validation accuracy and follows its footstep. Accuracies and losses are inversely proportional in a system. Our system is also showing this trend. Most of the loss calculations remain on the borderline of the scale in Fig. 9. The symmetry in the accuracy and loss curves of the three parameters substantially shows the network credibility. This designed network and its performances instinctively develop a logical ground of using two-dimensional LiDAR sensors for person identification in broad.

We performed rigorous testing with different datasets to test the system performance. All fourteen datasets were considered for this cross-testing. We analyzed the results and found an essential symmetry in the performance analysis phase, and those wholly aligned with our theoretical expectations—figure 10 shows these results in detail. Suppose the top four datasets on the figure are KoLaSU LiDAR 24, 13 respectively. We trained and validated the system with the same dataset but changed the test data only in four cases. LiDAR 24 and LiDAR 13 are created with merging sensors 2 and 4 and sensors 1 and 3, respectively. For testing, we used only LiDAR 4 and 2 and LiDAR 3 and LiDAR 1 separately. The bar chart shows that training and validation accuracy is nearly absolute, with a test accuracy below 20 percent. We considered unbiased disjoint data for all cases. As the figure shows, if a person is accurately trained by the neural network and tested differently, the system performance will degrade. The same performance persists for all the cases except the combined dataset LiDAR 1234 is tried with test data 24 and 13. Here the performance reached up to 38 percent but was not impressive. So here, we can conclude that to achieve the best performance of the system; it should be trained and tested with the same types of data; any other biasedness is not necessary there.

Table 2
Performance test with different DNN model

| Data | KoLaSU LiDAR 1234 TestCross24 | KoLaSU LiDAR 1234 TestCross24 |
|----------------------------|-------------------------------------|-------------------------------------|
| Experiments Type | 26 Persons Individual (60%,20%,20%) | 26 Persons Individual (60%,20%,20%) |
| Batch Size | 38 | 38 |
| Epoch | 25 | 40 |
| GPU | Yes | Yes |
| Model | Resnet 18 | Resnet 50_2 |
| Train Accuracy | 0.99864 | 0.99999 |
| Train Loss | 0.00589 | 0.000354 |
| Test Accuracy | 0.379 | 0.4007 |
| Test Loss | 4.2741 | 3.5852 |
| Validation Accuracy | 0.99721 | 0.99956 |
| Validation Loss | 0.00589 | 0.001578 |

Besides ResNet 18, we analyzed different neural networks to test our data's system performance and effectiveness. We show one of such types of analysis in Table 2. For the same dataset KoLaSU LiDAR 1234, we trained and validated it by ResNet 18 and ResNet 50. The rest of the parameters remained the same except epoch size. We tested the system with different datasets KoLaSU LiDAR 24 and tried to analyze the performance of two separate networks. ResNet 18 gave almost 38 percent accuracy, whereas ResNet 50 performed with 40 percent accuracy. But network size of ResNet 50 is abruptly huge than ResNet 18, and computation time is highly excessive. Thus, we decided to consider ResNet 18 rather than ResNet 50 even though its performance is little improved.

To test the system performance in different ways, we combined different datasets and trained and validated our system. Our achieved accuracies were impressive in all cases. In Fig. 11, we placed some experiments based on combined datasets. Here we put system accuracy and loss together. Six datasets were connected with their aligned ones. Suppose LiDAR 24 data was combined with LiDAR 2 and 4 data for training and validation of the system. Further, we tested the system individually with LiDAR 24, 2, and 4 data. The same scenarios were performed with LiDAR 1234, 13, and 24 also. In all cases, though the system was trained with multiple groups of data and tested with individual one, it performed as a regular system performed previously. From Fig. 11, we see that all six combinations' train, validation, and test accuracies lay around 99 percent, where its loss remains more minor as below 20 percent. Here accuracy and loss curves follow a symmetry as well, which is an indication of network performance.

4.3 Comparison with contemporary studies

To the best of our knowledge, there was no such research performed with 2D LiDAR sensors to identify a person based on gait analysis. Even a few studies were conducted with a 3D LiDAR sensor. Though they used different sensor setups and designed their model, we compared the overall system precisions in Table 3. Benedek et al. [4] initiated the research for lidar-based gait analysis. They prepared their dataset, SZTAK-LGA, with 28 participants. They used CNN (convolutional neural network) and MLP (multi-layer perceptron) for training and testing the system. They used a different number of people in their experiments; an increased number of people degraded the system performance from 92 percent (for five people) to 75 percent (for 28 people). Yamada et al. [3] performed a thorough experiment on lidar-based gait analysis. This research was also conducted with a 3D LiDAR sensor. They also prepared their dataset, PCG (point cloud gait), with 30 participants. A CNN and LSTM (long short-term memory) neural network model was applied for training and testing the system. Though they get different accuracies in different input patterns, here $l = 1:8$ gave a maximum of 72 percent in general. Compared with rest two, we used our own dataset KoLaSU, with 29 participants, and a two-dimensional LiDAR sensor only. We used a residual deep neural network (ResNet) for training, validation, and testing the system. We randomly used utterly unbiased datasets categorized into three classes (train, test, and validation). The average system performance is greater than 98 percent, emphasizing a wide use of 2D sensors in different applications.

Table 3
Performance comparison with state-of-the-art technologies

| Method | Dataset | Sensor | Model | Accuracies |
|--------------------|------------------------|----------|------------|---------------|
| Benedek et al. [4] | SZTAKI-LGA (28 People) | 3D LiDAR | CNN + MLP | 80% (approx.) |
| Yamada et al. [3] | PCG (30 People) | 3D LiDAR | CNN + LSTM | 72% (approx.) |
| Our | KoLaSU (29 People) | 2D LiDAR | ResNet | 98% (approx.) |

5. Conclusion

This paper presents a way of person identification with LiDAR sensor. Very little research has been done in this arena those were used only this sensor setup. We tried to elaborate the use of 2D LiDAR sensors with real-time calculations and enhanced accuracies. Creating an expanded dataset KoLaSU makes the study well suited in this research. A pre-trained deep neural network, ResNet effectively fit the data and found its losses for gait recognition. The overall precision of the system is awe-inspiring. A combined application of person tracking and identification with this sensor could be applied in autonomous robot movements.

Declarations

Availability of data and material

We created our dataset, which can be provided on demand of any research.

Competing interests

We know of no conflicts of interest associated with this publication, and there has been no significant financial support for this work that could have influenced its outcome.

Funding

There is no external funding for this research. This is going through with traditional research to fulfill Ph.D. requirements partially.

Authors' contributions

MH has written and conducted all experiments of this research. JH and RG contributed to dataset preparations and experiment setups. HF and YK guided and discussed the topic explicitly. Y Kobayashi is the main supervisor of the research. He guided and finetuned the manuscript thoroughly.

Corresponding Author: Correspondence to Mahmudul Hasan

Acknowledgments

We are very grateful to all our lab members of Kobayashi Laboratory for their active cooperation during the dataset preparation. We thank all volunteers from Bangladesh, Sri Lanka, Nepal, the Philippines, and Japan for their participation in collecting data. We also thank Saitama University and MEXT for providing us necessary support for this research.

References

1. Bolle RM, Connell J, Pankanti S, Ratha NK, Senior AW (2013) Guide to Biometrics. Springer Science & Business Media
2. Wan C, Wang L, Phoha VV (2018 Aug) A survey on gait recognition. ACM Computing Surveys. 51:89. <https://doi.org/10.1145/3230633>. 5
3. Yamada H, Ahn J, Mozos OM, Iwashita Y, Kurazume R (2020) Gait-based person identification using 3D LiDAR and long short-term memory deep networks. Adv Robot 1–11. <https://doi.org/10.1080/01691864.2020.1793812>
4. Benedek C, Gálai B, Nagy B et al Lidar-based gait analysis and activity recognition in a 4d surveillance system.IEEE Trans Circuits Syst Video Technol. 2018Jan; 28(1):101–113. doi: 10.1109/TCSVT.2016.2595331
5. Hasan M, Hanawa J, Goto R, Fukuda H, Kuno Y, Kobayashi Y (2021) Tracking People Using Ankle-Level 2D LiDAR for Gait Analysis. In: Ahram T. (eds) Advances in Artificial Intelligence, Software and Systems Engineering. AHFE 2020. Advances in Intelligent Systems and Computing, ; vol 1213. Springer, Cham. https://doi.org/10.1007/978-3-030-51328-3_7
6. Hasan M, Hanawa J, Goto R, Fukuda H, Kuno Y, Kobayashi Y (2021) Person Tracking Using Ankle-Level LiDAR Based on Enhanced DBSCAN and OPTICS. IEEJ Trans Elec Electron Eng 16:778–786. <https://doi.org/10.1002/tee.23358>

7. Hasan M, Goto R, Hanawa J, Fukuda H, Kuno Y, Kobayashi Y (2021) In: Huang DS, Jo KH, Li J, Gribova V, Bevilacqua V (eds) Intelligent Computing Theories and Application. ICIC 2021. Lecture Notes in Computer Science, vol 12836. Springer, Cham.
https://doi.org/10.1007/978-3-030-84522-3_62 Person Property Estimation Based on 2D LiDAR Data Using Deep Neural Network
8. Read JD (1995) The availability heuristic in person identification: The sometimes misleading consequences of enhanced contextual information. *Appl Cognit Psychol* 9:91–121. <https://doi.org/10.1002/acp.2350090202>
9. Choudhury T, Proceedings et al (1999) International Conference on Audio-and Video-Based Person Authentication. ; 1999; Washington DC
10. Matta F, Jean-Luc D (2009) Person recognition using facial video information: A state of the art. *J Visual Lang Comput* 20(3):180–187
11. Huang Z, Wang R, Shan S, Chen X, Recognition P (2015) (CVPR); 2015; pp. 140–149, doi: 10.1109/CVPR.2015.7298609
12. Liu H et al (2017) Neural Person Search Machines. IEEE International Conference on Computer Vision (ICCV); 2017; pp. 493–501, doi: 10.1109/ICCV.2017.61
13. Zhu X, Lei Z, Yan J, Yi D, Li SZ, Recognition P (2015) (CVPR); 2015; pp. 787–796, doi: 10.1109/CVPR.2015.7298679
14. Taigman Y, Yang M, Ranzato M, DeepFace WL (2014) : Closing the Gap to Human-Level Performance in Face Verification. IEEE Conference on Computer Vision and Pattern Recognition; 2014; pp. 1701–1708, doi: 10.1109/CVPR.2014.220
15. Schroff F, Kalenichenko D, Philbin J, FaceNet, Recognition P (2015) (CVPR); 2015; pp. 815–823, doi: 10.1109/CVPR.2015.7298682
16. Meng Q, Zhao S, Huang Z, Zhou F, MagFace:, Recognition P (2021) (CVPR); ; pp. 14225–14234
17. Wang L, Tan T, Ning H, Hu W (2003) Silhouette analysis-based gait recognition for human identification. *IEEE Trans Pattern Anal Mach Intell* 25(12):1505–1518. doi: 10.1109/TPAMI.2003.1251144
18. Bashir K, Xiang T, Gong S (2010) Gait recognition without subject cooperation. *Pattern Recognition Letters*; ; Volume 31, Issue 13, Pages 2052–2060.
19. Liu Z, Sarkar S (2006) Improved gait recognition by gait dynamics normalization. *IEEE Trans Pattern Anal Mach Intell* 28(6):863–876. doi: 10.1109/TPAMI.2006.122
20. Singh JP, Jain S (2010) Person identification based on gait using dynamic body parameters. *Trendz in Information Sciences* 248–252. doi: 10.1109/TISC.2010.5714649. & Computing (TISC2010)
21. Mansur A, Makihara Y, Aqmar R, Yagi Y, Recognition P (2014) (CVPR); ; pp. 2521–2528, doi: 10.1109/CVPR.2014.323
22. Lombardi S, Nishino K, Makihara Y, Yagi Y (2013) Two-Point Gait: Decoupling Gait from Body Shape. IEEE International Conference on Computer Vision; 2013; pp. 1041–1048, doi: 10.1109/ICCV.2013.133
23. Wu Z, Huang Y, Wang L, Wang X, Tan TA (2017) Comprehensive Study on Cross-View Gait Based Human Identification with Deep CNNs. *IEEE Trans Pattern Anal Mach Intell* 39(2):209–226. doi: 10.1109/TPAMI.2016.2545669
24. Zhang S, Wang Y, Li A, Recognition P (2021) (CVPR); ; pp. 9095–9104
25. Fan C et al (2020) GaitPart: Temporal Part-Based Model for Gait Recognition. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); ; pp. 14213–14221, doi: 10.1109/CVPR42600.2020.01423
26. Wang M, Deng W (2021) Deep face recognition: A survey. *Neurocomputing*; ; Volume 29, 2021, Pages 215–244, doi: 10.1016/j.neucom.2020.10.081
27. Dargan S, Kumar M (2020) A comprehensive survey on the biometric recognition systems based on physiological and behavioural modalities. *Expert Syst Appl* 143. doi: 10.1016/j.eswa.2019.113114
28. Sepas-Moghaddam A, Etemad A (2021) Deep Gait Recognition: A Survey. *ArXiv abs/2102.09546*;
29. Benedek C, Nagy B, Gálai B, Jankó Z (2015) Lidar-based gait analysis in people tracking and 4D visualization. 2015 23rd European Signal Processing Conference (EUSIPCO); ; pp. 1138–1142, doi: 10.1109/EUSIPCO.2015.7362561
30. Yoon S, Jung H-W, Jung H, Kim K, Hong S-K, Roh H, Oh B-M (2021) Development and validation of 2D-LiDAR-Based Gait Analysis Instrument and Algorithm. *Sensors* 21(2):414. doi: 10.3390/s21020414
31. Yan Z, Duckett T, Bellotto N (2020) Online learning for 3D LiDAR-based human detection: experimental analysis of point cloud clustering and classification methods. *Auton Robot* 44:147–164. doi: 10.1007/s10514-019-09883-y
32. Koide K, Miura J, Menegatti E (2019) A portable three-dimensional LIDAR-based system for long-term and wide-area people behaviour measurement. *Int J Adv Rob Syst*. doi: 10.1177/1729881419841532
33. Bradski G, Davis J (2000) Motion segmentation and pose recognition with motion history gradients. IEEE Workshop on Applications of Computer Vision, pp. 174–184,

Figures

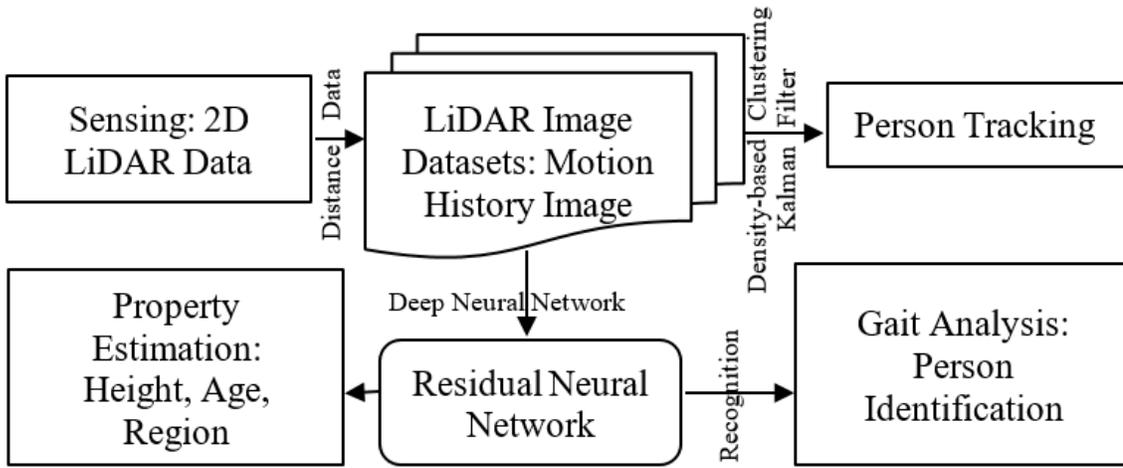


Figure 1

A system overview of 2D Lidar-based Estimation

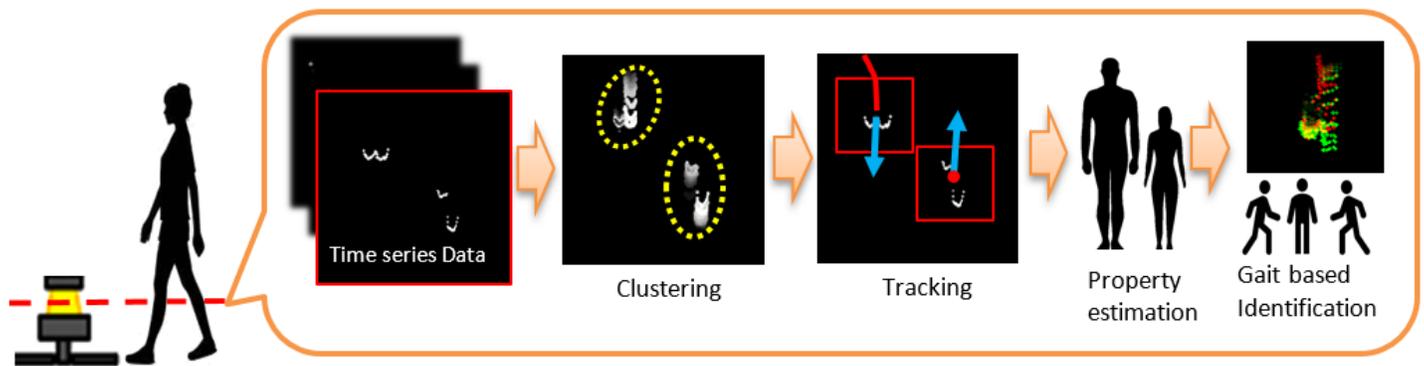


Figure 2

LiDAR-based Person Tracking, Property Estimation, and Recognition

Density based clustering (EDBSCAN and EOPTICS) was applied to cluster an ankle points to recognize a person.

Gravity point of cluster is tracked by Kalman Filter



Figure 3

Tracking using EDBSCAN and EOPTICS algorithms

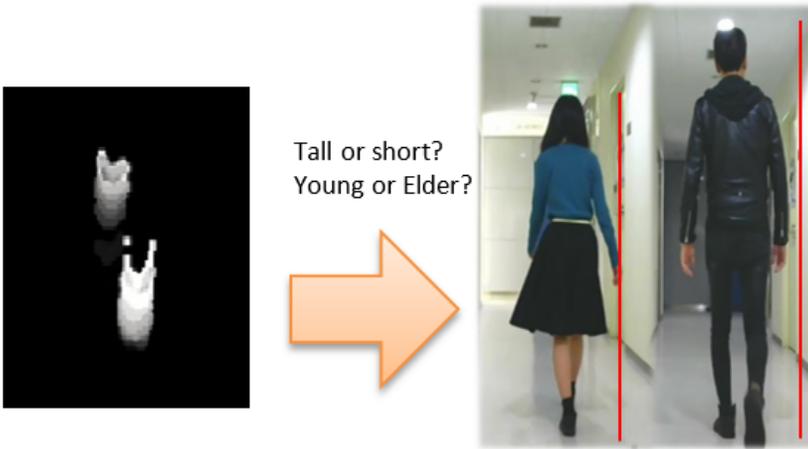


Figure 4

Person Property estimation (Height, Age, Region)

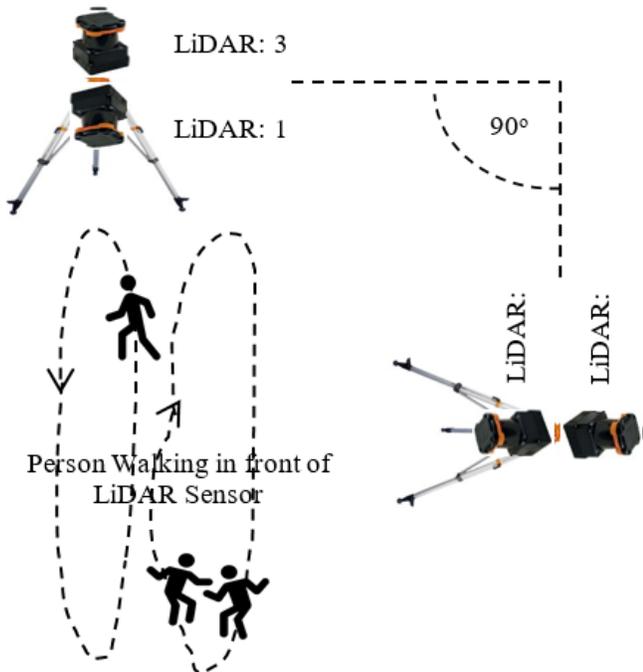


Figure 5

Experimental setup for Person Identification

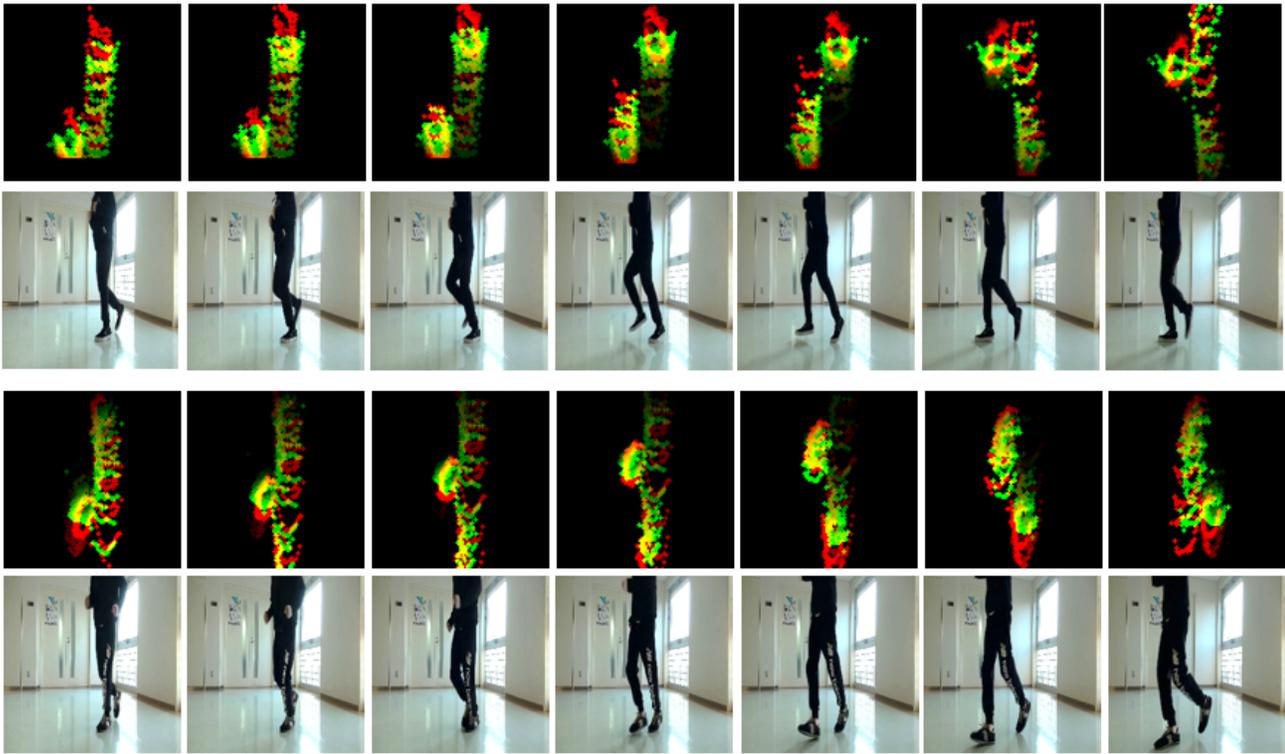


Figure 6

KoLaSU, Two Persons' data: Upper one is MHI and lower one is the corresponding pose

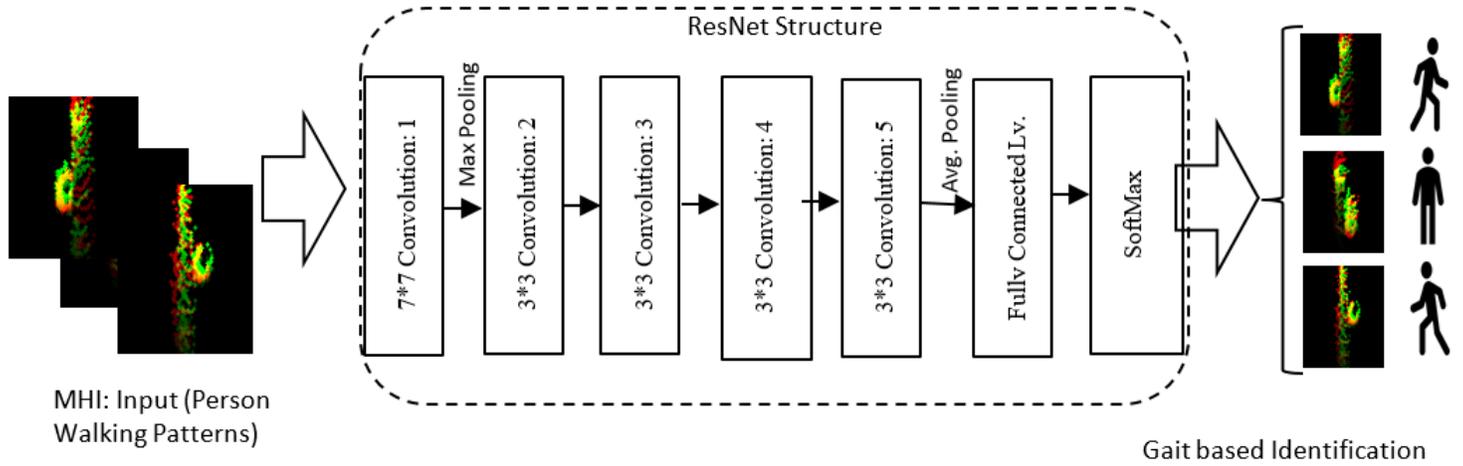


Figure 7

Gait based Person Identification

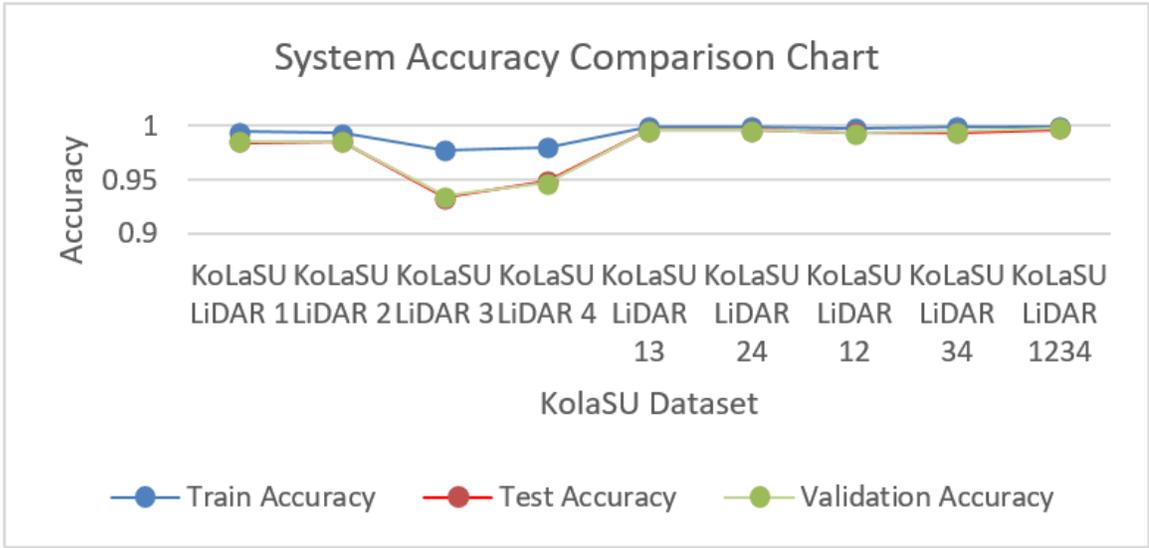


Figure 8

Gait-based Person Identification, Performance comparison

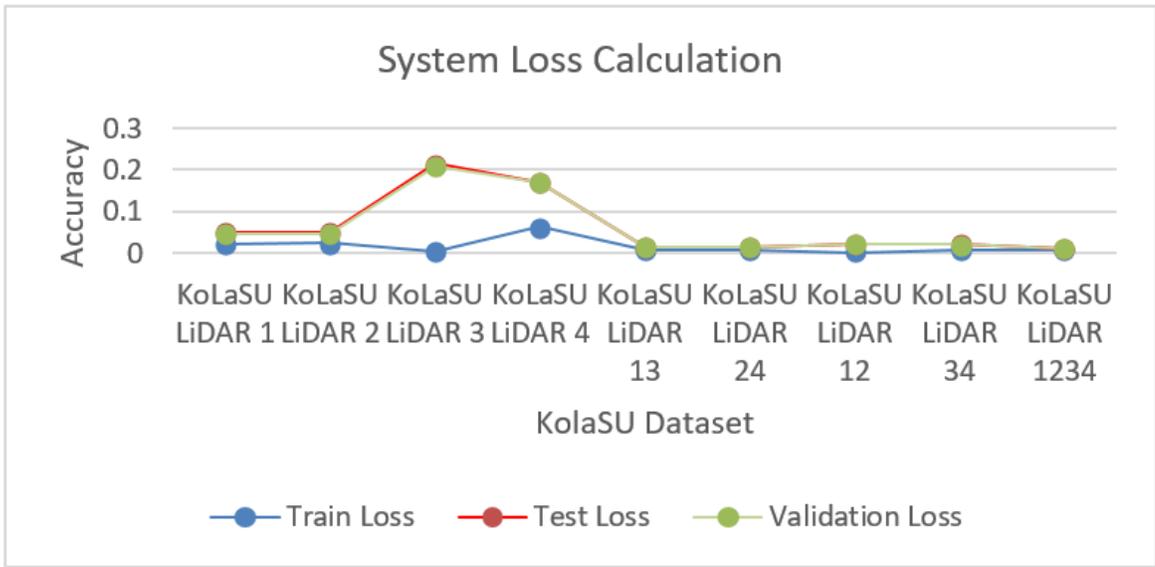


Figure 9

Gait-based Person Identification, System Loss comparison

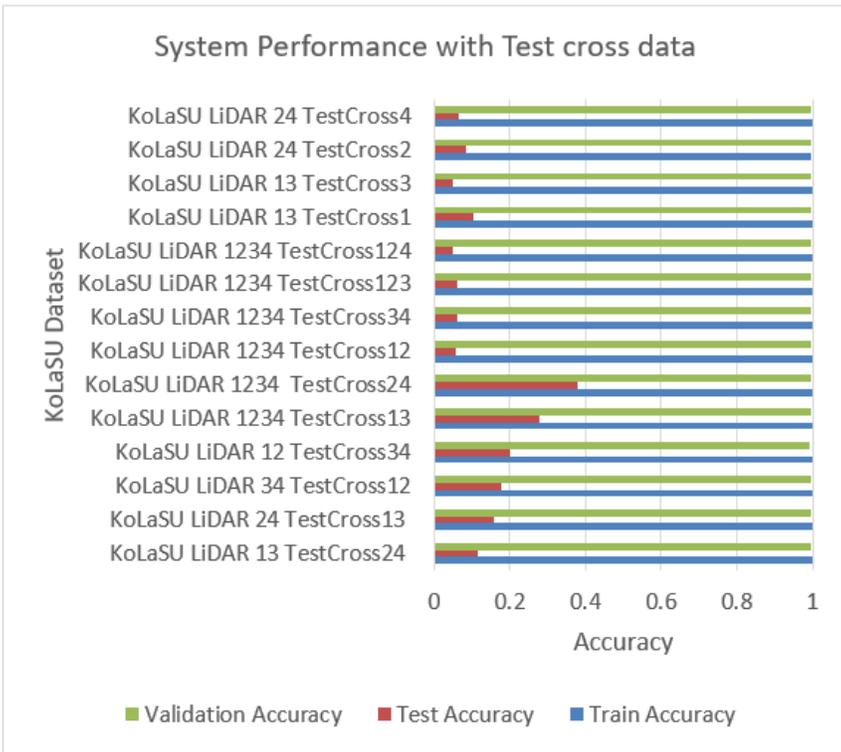


Figure 10

Gait Performance test with Cross-Data

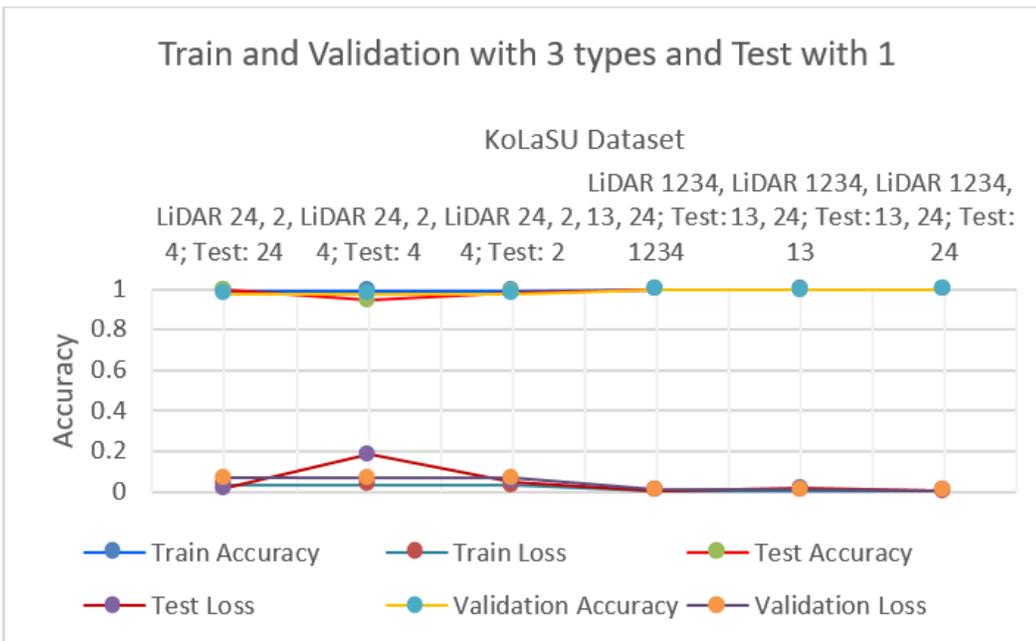


Figure 11

Combined dataset performance analysis