Classification of Humans Social Relations Within Urban Areas

Oscar Castro[1], Ely Repiso[12], Anaís Garrell[1], and Alberto Sanfeliu[1]

[1] Institut de Robòtica i Informàtica Industrial (CSIC-UPC).
Llorens Artigas 4-6, 08028 Barcelona, Spain.
[2] LAAS-CNRS, Université de Toulouse, CNRS, Toulouse, France.
7, avenue du Colonel Roche BP 54200 31031 Toulouse cedex 4, France.
{ocastro,agarrell,sanfeliu}@iri.upc.edu;ely.repiso-polo@laas.fr

Abstract. This paper presents the design of deep learning architectures which allow to classify the social relationship existing between two people who are walking in a side-by-side formation into four possible categories –colleagues, couple, family or friendship. The models are developed using Neural Networks or Recurrent Neural Networks to achieve the classification and are trained and evaluated using a database obtained from humans walking together in an urban environment. The best achieved model accomplishes a good accuracy in the classification problem and its results enhance the outcomes from a previous study [1]. In addition, we have developed several models to classify the social interactions in two categories –"intimate" and "acquaintances", where the best model achieves a very good performance, and for a real robot this classification is enough to be able to customize its behavior to its users. Furthermore, the proposed models show their future potential to improve its efficiency and to be implemented in a real robot.

Keywords: Human Behavior Classification, Human-Human Accompaniment, Social Relation, Pedestrian Groups

1 Introduction

The world of robotics is experiencing an unprecedented growth towards artificial intelligence and big data. Also, the field of Human-Robot Interaction (HRI) is not an exception, more and more researchers include these theories to allow the development of robots capable of executing more natural, safe, social and comfortable tasks for humans who interact with them to perform everyday tasks [2].

Furthermore, the interaction between robots and humans is presented as one of the greatest challenges for robotics that must be faced. For this reason, we must equip robots with more human and social skills, like the accompaniment task which is present in many situations, such as shopping [3], Universities [4], or visiting museums [5]. In this research, we are interested in the development

This work has been supported by the Artificial Intelligence for Human-Robot Interaction (AI4HRI) project ANR-20-IADJ-0006. Also, this Work has been supported under the Spanish State Research Agency through the ROCOTRANSP project (PID2019-106702RB-C21 / AEI / 10.13039/501100011033)) and the EU project CANOPIES (H2020- ICT-2020-2-101016906).

Fig. 1. Examples of human-human accompaniment relations. *Left*: a couple, *center*: friends, and *right*: colleagues. Notice: These images are from an own data set of our institute included in [6]. The provided one from [7] only includes the numeric data about the people tracks.

of social robots capable of accompanying pedestrians where the robot should be adapted to allow more human, predictable, and comfortable accompaniment behaviors. To do so, the classification of these human social relationships while they walk together is crucial to be able to include this human behavior, in the future, in the accompaniment behavior that the robot will carry out with its companions.

Then, this work presents the first steps in which a new framework is capable of classifying the relationship between humans, which it can later be used between robots and humans. Fig. 1 shows a set examples of different relationships between humans navigating in a side-by-side formation.

In the remainder of the paper, Sec. 2 includes the related work. Sec. 3 describes the used database and exposes the implemented architectures of Neural Networks (NNs) and Recurrent Neural Networks (RNNs). Sec. 4 shows our results. Sec. 5 includes some discussions. Finally, conclusions are given in Sec. 6.

2 Related Work

In general, there are several studies focused on knowing the social relationship between people inside several images by finding the context of these images [8–10], to design machines and robots capable of interacting socially with humans. But few works try to find these social relations between humans using geometrical properties that a social robot can easily detect to classify them, and customize its relation with its human users while their are walking together.

We have found a similar study in Yucel et al. [1] that classifies the relationship of pedestrian couples into four categories: colleagues, couple, family, or friends. They use two methods to obtain relatively good results in classifying the couples' relationships and distinguishing between the four categories above. We do similar work to these authors, obtaining similar results in the classification in four categories. Still, we obtain better results in classification in two new categories that should be sufficient to customize the robot's behavior of accompaniment. Also, we use deep learning techniques that have more potential to be improved, and we focus on this classification to be able to use it in the future in a social robot accompaniment implemented in several previous works of us [11, 12].

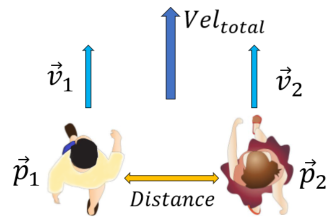

Fig. 2. Representation of the variables of the accompaniment process obtained from the examples in the database.

3 Classification of Humans' Relationship

In order to make robots capable of classifying social relationships between humans, we describe the database that we use in Sec. 3.1, and the deep learning architectures created to achieve the mentioned classifier in Sec.3.2 and Sec.3.3. We have chosen these methods because these ones are suitable for time series, which are the type of data we have, whereas other methods like Bayesian and SVN are not suitable for time series.

3.1 Database Description

The used database[3] in this work was provided by Dr. Francesco Zanlungo of the Intelligent Robotics and Communication Laboratory of the Advanced Telecommunications Research Institute (ATR) in Kyoto, author of several studies in the field of HRI and social robots [7]. How this database was obtained is included in [13].

This database includes 867 examples of groups of two people performing an accompaniment process, walking through an urban environment, distributed into four categories –267 colleagues, 96 couple, 218 family or 286 friendship. This database of the previous website contains directly the readings of 13 different variables and the labels corresponding to the groups' social relationships. These variables are: detection time; $\overrightarrow{p}_c = (p_X,\ p_Y,\ p_Z)_c$ where $c \in \{1,2\}$ meaning position of pedestrian 1 and 2; $\overrightarrow{V}_c = (V_X, V_y)$ velocity of pedestrian 1 and 2; and ${V^T}_c$ total velocity of pedestrian 1 and 2. Most of these variables can be seen graphically in the Fig. 2. In addition, to the aforementioned 13 parameters, three new variables are calculated using these previous parameters due to the importance they prove to have in the processes of social accompaniment and navigation [7,14]. The addition of these three new parameters allow to increase the accuracy. These new three variables are: the distance between pedestrians, the relative pedestrian-to-pedestrian velocity, and total velocity of the couple in Eq. 1. Then, finally we have 16 parameters. Further details about the database can be obtained in [1].

$$Dist = ||\overrightarrow{p}_2 - \overrightarrow{p}_1||;\ Vel_{Relative} = |V_1^T - V_2^T|;\ Vel_{Total} = \frac{V_1^T + V_2^T}{2}, \quad (1)$$

[3] `https://dil.atr.jp/sets/groups/`

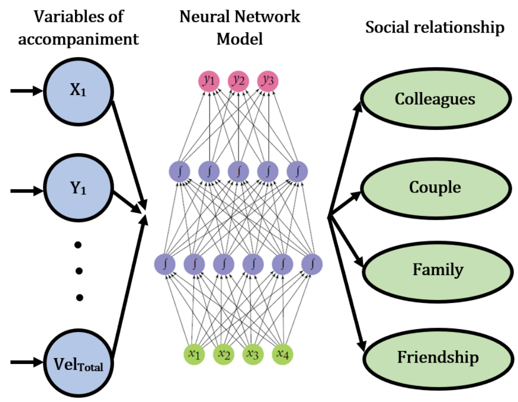

Fig. 3. General outline for all the developed classification methods.

where the position of pedestrian 1 and 2 are respectively $\overrightarrow{p}_1$ and $\overrightarrow{p}_2$. The total velocity of pedestrian 1 and 2 are respectively V_1^T and V_2^T.

A group accompaniment is a dynamic process over time, then the readings of variables at a given instant may not accurately represent the real relationship. To represent it better, we took the readings from each experiment and averaged each variable; except for the time variable, which is obtained by subtracting the final and initial times. To train the different NNs designed, the existing database is divided into training set (90 % a total of 780 experiments) and test set (10 % a total of 87 experiments); both randomly chosen from all the data of the database of 780 experiments.

3.2 Neural Networks (NN)

Different NN designs are used to test which one offers the best accuracy. In each design implemented, several hyperparameters of the network are varied: the number of hidden layers, the number of neurons in the hidden layers, the learning rate and the number of epochs in the training process. In addition, the L2 regularisation and dropout methods are used to try to solve the overfitting problem. There are several design conditions that are met for all NNs developed. These conditions are: The input layer consists of 16 neurons, where each neuron corresponds to one of the input parameters explained in the Sub.Sec. 3.1; the hidden layers are designed with the $ReLU$ activation function; the output layer is designed with the *Softmax* activation function; the Mini-Batch Gradient Descent is used as gradient descent algorithm; and the croos-entropy function is used as the loss function. An outline of the developed classification method is in Fig. 3. Where, the variables extracted from the database feed the designed deep learning models; and the classification of the group in one of the defined social relationships is obtained at the NN output.

3.3 Recurrent Neural Network (RNN)

A RNN is a type of NN capable of working with temporal data sequences. The ability to have memory makes RNNs a suitable tool for machine learning tasks

	# Hidden layers	# Neurons in hidden layers	# Epochs	Learning rate	L2 Reg.	Dropout
NN2-3	2	1500-600	2500	0,00011	Yes	No
NN2-4	4	1800-2500-1600-600	2500	0,00011	Yes	No
NN2-5	2	1500-600	2500	0,00011	Yes	Yes

Table 1. Features of standard NN models implemented.

	Training set accuracy	Test set accuracy
NN2-2	94,10%	33,33%
NN2-3	99,10%	40,23%
NN2-4	95,90%	36,78%
NN2-5	63,08%	33,33%

Table 2. Accuracy of standard NN models implemented.

involving sequential data; by using relevant information from past input data in the training process, they can make more accurate predictions [15].

Once the standard NN models had been implemented, several RNNs were designed to test their efficiency. Having readings of the database samples over a certain period of time, it is decided to implement RNNs due to their ability to work with temporal sequences of data and extract possible dependencies between them. Specifically, Long Short-Term Memory (LSTM) networks are an extension of RNNs that use special hidden units, called memory cells, whose goal is to increase the network's memory so that it can remember important information over time [16]. In this way, LSTMs can better capture temporal dependencies of long-term input data due to the ability of LSTM units to register possible temporal behavioral patterns. In addition, they can avoid the phenomenon of vanishing or exploding gradients.

All our developed RNNs follow the same design conditions as NNs, in subsec. 3.2. In addition, the first hidden layer of the RNNs is a LSTM layer, which allows to capture the temporal dependencies of the input data in the long term.

4 Results

In this section, we present the results obtained from the different neural networks implemented. First, the NNs and the RNNs models that are used to classify the people behavior in four categories in Sec. 4.1 and second, the same methods are used to classify the behaviors in only two categories, because the categories of couple, family, and friendship can be merged in one "intimate", due to the fact that they have a similar degree of intimacy, very different than the one of the colleagues category.

4.1 Results of NNs models

The characteristics of the NNs implemented are detailed in Tab. 1. By running the networks defined in Tab. 1, the results are obtained as shown in Tab. 2. To try to avoid the phenomenon of overfitting, in some models L2 regularisation and dropout methods have been implemented. We only show the confusion matrix of the best model selected in Tab. 3.

		Predicted Value			
		Colleagues	**Couple**	**Family**	**Friends**
Real Value	**Colleagues**	**37,50**	4,17	29,17	29,17
	Couple	28,57	**21,43**	28,57	**21,43**
	Family	19,05	0	**42,86**	38,10
	Friends	32,14	10,71	7,14	**50**

Table 3. Confusion matrix of the NN2-3 model (in %).

	# Hidden layers	**# Neurons in hidden layers**	**# Epochs**	**Learning rate**	**L2 Reg.**	**Dropout**
RNN2-1	2	25-12	1500	0,00015	Yes	No
RNN2-2	4	2500-1800-1200-600	10	0,00011	Yes	No
RNN2-3	2	1500-600	25	0,00011	Yes	No

Table 4. Features of RNN models implemented.

Tab. 3 shows that the examples in the family and friendship categories are correctly classified with accuracies of $42,86$ % and $50,00$ %, respectively, when using the NN2-3 model. Also, it may classify part of the other examples of the two other categories, couples and colleagues.

4.2 Results of RNNs models

The characteristics of the RNNs implemented are in Tab. 4. The results obtained by executing the networks defined in Tab. 4 are in Tab. 5. In order to try to avoid the phenomenon of overfitting, the L2 regularisation method is implemented in all models. The confusion of the best obtained method is shown in Tab. 6.

Tab. 6 highlights the accuracies of $82,61$ % and $30,30$ % obtained by correctly classifying the examples in the categories of colleagues and friendship, respectively, when using the RNN2-1 model. Also, we can see here that this method tries to classify in the friendship category the categories of family and couple due to the similarity of these three categories. The fact that supports that after, we try to test our classifiers in a new database that join these three types of relation in one.

4.3 Results analysis

At first glance, when analysing the results presented in Tab. 2 and 5, it is observed that the best accuracies in the training set are obtained in the models based on standard NNs. On the contrary, the best accuracies in the test set are obtained in the models based on RNNs. Specifically, the NN2-3 model, which has an accuracy of $99,10$ % in the training set and $40,23$ % in the test set, and the RNN2-3 model, which has an accuracy of $66,52$ % in the training set and $42,31$ % in the test set, stand out. Although the accuracies of the methods in the training set and test set are basic evaluation metrics, an analysis of the confusion matrices obtained is necessary to understand how the classification of the examples into the four categories studied is being performed.

When analysing the confusion matrices of the models based on standard NNs that show better accuracy, different phenomena are observed. In Tab. 3,

	Training set accuracy	**Test set accuracy**
RNN2-1	45,17%	41,03%
RNN2-2	62,05%	30,77%
RNN2-3	66,52%	42,31%

Table 5. Accuracy of RNN models implemented.

		Predicted value			
		Colleagues	**Couple**	**Family**	**Friendship**
Real Val.	**Colleagues**	**82,61**	0	8,70	8,70
	Couple	33,33	0	0	**66,67**
	Family	**43,75**	0	18,75	37,50
	Friendship	**60,61**	0	9,09	30,30

Table 6. Confusion matrix of the RNN2-1 model (in %).

corresponding to the model with the highest accuracy in both, the training set and the test set, it is observed that the NN2-3 model recognises the correct categories more accurately than the others, even reaching $50,00$ % accuracy in the friendship category, and with the exception of the examples in the couples category, which are assigned in a similar proportion among the four categories. This same pattern is repeated in other models, as for example the NN2-4 model, which is not included in the tables due to space constrains, although the accuracy of the classification of the friendship and couple categories is worse, it is improved in the family category. In general, they are not able to classify the examples in the correct categories with a higher accuracy than the others and, moreover, the effectiveness of the classification is in all cases surpassed by the one of the NN2-3 model. Thus, the NN2-3 model is postulated as the model based on a standard NN capable of performing a better classification of human relationships in the four categories studied.

When analysing the confusion matrices generated from the models based on RNNs, different phenomena are observed. The model RNN2-1, in Tab. 6, has a preference when performing the classification to assign the examples in the categories of colleagues or friendship. This results in very high accuracy for the category of colleagues $82,61$ %, but null for the category of a couple and very low for the category of family. Due to its similarity, the method classifies all in the category of friendship. This pattern is repeated for all the RNN models, as all the models designed have zero accuracy in correctly classifying the examples into the category of couple. Taking this into consideration, the rest of the models present a similar situation by having more facilities to classify the data in certain categories than in others. For example, while the RNN2-3 model is highly accurate in correctly classifying the examples in the colleague category, the RNN2-2 model is highly accurate in classifying the examples in the family category.

In addition, models based on RNNs have, in general, considerably good accuracy in classifying examples of the colleagues category compared to models based on standard NNs. After all, the RNN2-3 model seems to emerge as the model based on a RNN that can perform a better classification of human relationships by showing, in general, a relatively higher efficiency than the other models in classifying the examples into the four correct categories. It has in its confusion matrix a 69,57% of accuracy to classify correctly colleagues, a 0% for couple, a 37,50% for family and 33,33% for friendship that includes most of the couple cases classified as friendship.

Therefore, it is considered that the model that can best perform the classification of the social relationship in human couples is the NN2-3 model, as it presents the best accuracy of all the models implemented in the training set, as well as the best accuracies in the test set, and the best efficiencies as a whole when classifying the data correctly in the four categories, not only in two.

4.4 Sighting extracted from the implementation of the methods

The differences between the obtained accuracies in the designed models are due to the different configurations of hyperparameters that have been implemented in them. The networks present higher or lower accuracies in the training and test sets depending on these configurations. Even so, when analysing the confusion matrices of the models designed, it is possible to draw several conclusions.

Models with a higher number of neurons in their hidden layers or their epochs can train their network parameters better and obtain high accuracies in the training set. However, the accuracy slightly decreases with increasing the number of hidden layers from two to four, as can be seen when comparing the results of the two-hidden-layer NN2-3 model with the results of the four-hidden-layer NN2-4 model, Tab. 2. This phenomenon may occur as a result of the increased learning difficulty during training due to the increased number of layers. Therefore, an unnecessary increase in the number of layers should be avoided because it does not necessarily lead to an increase accuracy and, on the contrary, may lead to a reduction efficiency.

Another factor to take into account is the training time of the models. When it is increased the number of hidden layers, the number of neurons in the hidden layers, or the number of epochs, the computational time and cost of the model training process rise. In the specific case of this work, it has been decided to give priority to the accuracy of the models in the classification, but in the case of looking for a model with a shorter training period, the correct choice of these three hyperparameters must be taken into account to achieve this objective.

Moreover, the L2 regularisation method manages to slightly decrease the overfitting phenomenon and increase the accuracy of the model, as can be seen by comparing the results of the NN2-2 and NN2-3 models of Tab. 2. In addition, the dropout method not only fails to increase the accuracy of the model in the test set but also considerably decreases the accuracy in the training set, as can be seen by comparing the results of models NN2-3 and NN2-5 of Tab. 2. Finally, the choice of the learning rate is approached as a process of choosing different values and testing the effectiveness of the models for each of them until the value that offers the best possible results is found.

Consequently, the design of NNs is proposed as an iterative trial-and-error process to find the configuration of hyperparameters that offers the best efficiency of the model and, therefore, the best possible accuracy in the database classification process.

	# Hidden layers	# Neurons in hidden layers	# Epochs	Learning rate	L2 Reg.	Dropout
NN2-6	2	1500-600	2500	0,00011	No	No
NN2-7	2	1500-600	2500	0,00011	Yes	No
RNN2-4	2	25-12	500	0,00011	Yes	No

Table 7. Features of NN and RNN models implemented for the classification into two categories.

	Training set accuracy	Test set accuracy
NN2-6	96,15%	71,26%
NN2-7	95,38%	63,22%
RNN2-4	75,76%	65,38%

Table 8. Accuracy of NN and RNN models implemented for the classification into two categories.

Finally, it can be concluded that the differentiation between the four social relations is a complicated process due to the similarity of the data, even for a human being in Fig. 1. However, the best models presented perform relatively well in classifying the social relationship of people in a group into one of the four categories studied. Therefore, it can be stated that NNs-based deep learning methods are positioned as a relatively effective tool for classifying the relationship between people in a group. However, to realise their full potential, it is necessary to considerably increase the database with more examples of groups, as it is with large datasets that NNs work best. Mainly, it is essential to increase the number of couple relationships in the database, which has much fewer examples than the others, which endows in the worst accuracy in classifying this class by all the models designed. Doing that, it will increase the model's effectiveness. In addition, the process of setting up hyperparameters and testing new NN models should also be continued to see how effective they are.

4.5 Classification into two categories

Finally, in order to better analyse the results obtained, it was decided to merge the categories of couple, family and friendship into a single category called "intimate", in reference to the consideration that the three merged relationships entail a higher degree of intimacy between the members of the couple than the relationship of colleagues. In this sense, the category of colleagues is renamed as "acquaintances" for the purpose of the classification process.

In the following, we report the classification results obtained by the two types of NNs used. The characteristics of the different NN models implemented are detailed in Tab. 7. The results obtained by executing the networks defined in the Tab. 7 are listed in Tab. 8. To try to avoid the phenomenon of overfitting, L2 regularisation is implemented in some models. The confusion matrices of the different designed methods are collected in Tab. 9.

The classification into two possible categories shows a considerable increase in the accuracy in the test set of the different models designed, as can be seen

		Predicted value					
		Model NN2-6		Model NN2-7		Model RNN2-4	
		Acquaint.	**Intimate**	**Acquaint.**	**Intimate**	**Acquaint.**	**Intimate**
Real Val.	**Acq.**	45,83	**54,17**	33,33	**66,67**	17,39	**82,61**
	Intimate	19,05	**80,95**	25,40	**74,60**	14,55	**85,45**

Table 9. Confusion matrix (in %) of the NN2-6 model at left, NN2-7 model at center, and RNN2-4 model at right.

		Predicted value			
		Colleagues	**Couple**	**Family**	**Friendship**
Real value	**Colleagues**	**68,31**	7,29	5,37	19,03
	Couple	18,10	**38,92**	20,66	22,32
	Family	13,58	31,11	**36,57**	18,75
	Friendship	34,19	16,66	12,74	**36,41**

Table 10. Confusion matrix of one of the implemented methods in [1] (in %).

in Tab. 8. Specifically, the NN2-6 model stands out, with an accuracy of $96,15$ % in the training set and $71,26$ % in the test set. Furthermore, it can be seen in Tab. 9-*left* that the model has a high accuracy in the recognition of examples in the intimate category and by far the best accuracy of all models for the acquaintances category. In contrast, other models such as NN2-7 or RNN2-4 show a very considerable bias towards the intimate category, as can be seen in Tabs. 9-*center* and 9-*right*. Therefore, the NN2-6 model is postulated as the NN-based model capable of performing a better classification of human relationships into the two categories presented.

Thus, it can be stated that the process of classifying the social relationship into two categories is more accurate than the classification into four categories. In addition, as discussed above, it is necessary to increase the database of examples and to continue the process of setting hyperparameters in order to achieve a model with the best possible efficiency.

4.6 Comparison with state-of-art-method

Once the results of the different classification models have been obtained, they are compared with the results obtained in the work [1]. However, it is important to bear in mind that, although the objective of this project coincides with that of the comparative study, the databases used are not exactly the same. Nevertheless, it is possible to make a comparison of the results obtained by using different classification methods.

As various methods, parameters and even data are used in the study [1], the best results obtained by some of the methods presented are used to compare them with the results obtained by the models developed in this work. These different methods are based on Bayesian methods that include or not hierarchical recognition and also use or not the entire trajectory or only a single observation, more details about these methods in [1]. Specifically, we select one of the methods of the study that obtains the best classification accuracies in the four categories of relationships which also uses the whole trajectory, Tab. 10.

Therefore, we compare the results of one of the best models of the study [1], shown in Tab. 10, with the results of our developed NN2-3 model in Tab. 3, which is considered the model designed to best classify the social relationship in human groups. Thus, it is observed that while the method of the study presents a better accuracy in the categories of colleagues and couples, the developed model has a higher accuracy in the categories of family and friendship. However, the method implemented in [1] presents greater differences in the classification accuracies of the categories of colleagues and couple with respect to our method, than those presented by our method in the categories of family and friendship with respect to those of the comparative study. Consequently, it is difficult to say with certainty which model performs a better overall classification of social relationships in the four categories presented.

Even so, the potential offered by the deep learning models suggests that an increase in the database of examples of accompanying couples, and some minor adjustments in the configuration of the hyperparameters of the model designed could provide the necessary improvement to obtain greater efficiency in the classification process.

Finally, our new strategy to merge the three most similar categories in one, derived to obtain better classifications results than in the state-of-art method. And for a robot application to customize the robot's behavior to accompany people, this classification in two categories can be enough and derive in good accompaniment customization results.

5 Discussion of Contributions for HRI and its Applicability

One of the pillars that technology must always bear in mind is to make people's lives easier. Then, to have robot companions that are able to automatically adapt to people preferences can be an important improvement in HRI, here focused to the people accompaniment. To be able to adapt to people preferences, these robots should classify their social relation between them and humans; and to be able to do so, first they need to be able to recognize and classify the different human social relations while walk together between them. In addition, any accompanied person could choose, prior to the accompaniment process, what type of relationship they want that the robot use with them; and this robot only need to apply the same geometrical relations that it knows that represents this social relation.

It is important to take into account the advantage that a robotic system can adapt itself as comfortably as possible to a person while accompanying them to a destination. For example, this can be very beneficial when using robot assistants for people with special needs, sick or elderly in nursing homes. Such robots can also support care home workers, helping to reduce their heavy workload. Ultimately, providing robots with the tools that allow them to adapt their behavior in a personalised way, and behave in the most social way possible can only bring benefits to the society of the future.

6 Conclusions

We have presented several models capable of classifying the social behavior of a group of humans during their accompaniment into four possible categories (colleagues, couple, family or friendship), and into two possible categories "intimate" and "acquaintances". From the analysis of the results, the NN2-3 model based on a standard NN stands out, with which good accuracies are obtained in the process of classifying the social relationship between groups in the four categories. The confusion matrix has been used to verify the effectiveness of the classification and the results have been compared with those obtained in a similar work [1]. The process of classification into two possible categories is more accurate due to the similarity of the data of three of these categories, also for humans. Here, the model that achieves the best accuracy results is NN2-6. Also, this classification in two categories is far enough to obtain, in the future, a human-robot accompaniment more customized to its partner behavior.

The future work would be oriented towards two different objectives. The first objective would be trying to improve the precision of the implemented NNs. To achieve this, it is essential to increase the database with more labeled examples of couples accompanying each other and, then, continue adjusting the parameters of the models and modifying their design to increase their effectiveness. The second objective to be carried out would be the implantation of the designed models in a real robot to classify in real-time its actual social relation with the human that is interacting with it; and adapt its accompaniment behavior to the detected relation between them; obtaining in that way a more natural, intelligent, safe, social and comfortable robot's behavior.

References

1. Z. Yucel, F. Zanlungo, C. Feliciani, A. Gregorj, and T. Kanda, "Identification of social relation within pedestrian dyads," *PloS one*, vol. 14, no. 10, p. e0223656, 2019.
2. M. Doering, D. Brščić, and T. Kanda, "Data-driven imitation learning for a shopkeeper robot with periodically changing product information," *ACM Transactions on Human-Robot Interaction (THRI)*, vol. 10, no. 4, pp. 1–20, 2021.
3. H.-M. Gross, H. Boehme, C. Schroeter, S. Müller, A. König, E. Einhorn, C. Martin, M. Merten, and A. Bley, "Toomas: interactive shopping guide robots in everyday use-final implementation and experiences from long-term field trials," in *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems.* IEEE, 2009, pp. 2005–2012.
4. E. Repiso, A. Garrell, and A. Sanfeliu, "Robot approaching and engaging people in a human-robot companion framework," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).* IEEE, 2018, pp. 8200–8205.
5. F. Faber, M. Bennewitz, C. Eppner, A. Gorog, C. Gonsior, D. Joho, M. Schreiber, and S. Behnke, "The humanoid museum tour guide robotinho," in *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication.* IEEE, 2009, pp. 891–896.
6. E. Repiso, A. Garrell, and A. Sanfeliu, "Adaptive side-by-side social robot navigation to approach and interact with people," *International Journal of Social Robotics*, vol. 12, no. 4, pp. 909–930, 2020.

7. F. Zanlungo, Z. Yücel, and T. Kanda, "Intrinsic group behaviour ii: On the dependence of triad spatial dynamics on social and personal features; and on the effect of social interaction on small group dynamics," *PloS one*, vol. 14, no. 12, p. e0225704, 2019.
8. Z. Wang, T. Chen, J. Ren, W. Yu, H. Cheng, and L. Lin, "Deep reasoning with knowledge graph for social relationship understanding," *arXiv preprint arXiv:1807.00504*, 2018.
9. Q. Sun, B. Schiele, and M. Fritz, "A domain based approach to social relation recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3481–3490.
10. J. Li, Y. Wong, Q. Zhao, and M. S. Kankanhalli, "Visual social relationship recognition," *International Journal of Computer Vision*, vol. 128, no. 6, pp. 1750–1764, 2020.
11. A. Garrell, C. Coll, R. Alquézar, and A. Sanfeliu, "Teaching a drone to accompany a person from demonstrations using non-linear asfm," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 1985–1991.
12. E. Repiso, A. Garrell, and A. Sanfeliu, "People's adaptive side-by-side model evolved to accompany groups of people by social robots," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2387–2394, 2020.
13. D. F. Glas, T. Miyashita, H. Ishiguro, and N. Hagita, "Laser-based tracking of human position and orientation using parametric shape modeling," *Advanced robotics*, vol. 23, no. 4, pp. 405–428, 2009.
14. Z. Yucel, F. Zanlungo, C. Feliciani, A. Gregorj, and T. Kanda, "Estimating social relation from trajectories," *Collective Dynamics*, vol. 5, pp. 222–229, 2020.
15. S. Li, W. Li, C. Cook, C. Zhu, and Y. Gao, "Independently recurrent neural network (indrnn): Building a longer and deeper rnn," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5457–5466.
16. Y. Yu, X. Si, C. Hu, and J. Zhang, "A review of recurrent neural networks: Lstm cells and network architectures," *Neural computation*, vol. 31, no. 7, pp. 1235–1270, 2019.