

Reasoning the Trust of Humans in Robots through Physiological Biometrics in Human-Robot Collaborative Contexts

Tiffany Guo
Department of Electrical and
Computer Engineering
Cornell University
Ithaca, New York
tg382@cornell.edu

Omar Obidat
Department of Computer
Science
Montclair State University
Montclair, USA
obidato1@montclair.edu

Laury Rodriguez
Department of Computer
Science
Montclair State University
Montclair, USA
rodriguez116@montclair.edu

Jesse Parron
Department of Computer
Science
Montclair State University
Montclair, USA
parronj1@montclair.edu

Weitian Wang*
Department of Computer
Science
Montclair State University
Montclair, USA
wangw@montclair.edu

Abstract—With the rapid recent growth of automation and artificial intelligence, human-robot collaboration (HRC) is playing a significant role across a variety of fields. Trust between humans and robots is an important element to enable the efficiency and success of HRC. The lack of trust of humans in robots can have critical consequences, especially in real-world applications in which humans must adapt to unfamiliar situations. In this work, we develop a novel and effective approach for robots to actively reason and respond to dynamic human emotions and trust levels during shared tasks. We implement a real-world validation experiment in the context of human-robot object hand-over, which shows the robot's ability to correctly identify and predict the human's trust levels in real-time and assist the human accordingly in human-robot collaborative tasks. Future work on how to improve the performance of the proposed approach is also discussed.

Keywords—Robotics, human-robot interaction, trust, computer vision, Extreme Learning Machine

I. INTRODUCTION

In a world facing many human struggles, robots who are stronger, speedier, and smarter hold the key to breaking through many centuries-old roadblocks. With the rapid recent growth of automation and artificial intelligence, human-robot collaboration (HRC) is playing a large role across a variety of fields [1-3]. In the present day, robots have made their way into a variety of vastly differing but equally significant fields. For instance, their superior strength and durability are utilized in the areas of labor-intensive manufacturing and product assembly. Human-robot hand overs, which are used to optimize the robot's strength and precision and the human's flexibility and knowledge, play a key role in this collaboration [4]. Other fields involving robots require them to be more intelligent or have increased computing power. These areas often involve high-risk situations, such as advanced surgeries, search and rescue missions, military defense, and assistant care for the injured, disabled, and elderly [5-7].

In many of these areas, robots are still seen as tools to be manipulated rather than teammates for humans to work alongside [8]. However, advancements are being made to allow robots to take more of an autonomous, teammate-like role in HRC situations [9, 10]. For example, in the 2015 disaster relief project TRADR, robots utilized semi-autonomous navigation along with data gathering abilities, making use of their strengths to contribute independently to the mission [6]. Moreover, another research work about teaching-learning-collaboration in intelligent manufacturing allowed robots to learn from human

demonstrations and natural language inputs similar to a human collaborator [11].

However, despite various advances in robotics and technology, effective HRC will not occur simply by inserting robots into human teams. In order to facilitate highly functioning human-robot partnerships, trust must be forged between each teammate. The establishment of the trust is especially important in high-risk tasks, such as rescue missions and medical procedures, where a human's lack of trust in the robot may result in a life-threatening mistake [12-14].

Several studies have been conducted to model the trust of humans in robots in recent years. Researchers have defined various metrics with which to quantify trust and developed models to classify trust [15-17]. However, while these studies identify factors that affect trust between humans and robots, methods to assess human-robot trust with higher accuracy and approaches to increase human trust in robots are still being developed. This should be addressed in order to increase the effectiveness of HRC as well as to have quantitative measures with which to compare different approaches.

A set of universal facial expressions mapping to emotions are established and recognized over some studies. These findings have been reproduced in numerous studies originating from different cultures around the world, yielding strong evidence for seven universal facial expressions: happiness, sadness, anger, surprise, fear, contempt, and disgust [18]. Using human facial features to detect or predict emotion, however, is a more difficult task due to the variations present from one face to another as well as the ability to accurately detect facial features in relatively low-quality images. Several studies have set out to accomplish this task using a variety of different approaches and algorithms. Said et al. trained a face-sensitive convolutional neural network in 2020 which reached a facial detection accuracy of around 93% and emotion recognition accuracy of around 95% [19]. Liu et al. applied a convolutional neural network to emotion color transfer, their model reaching an average classification accuracy of 73.17% [20].

However, throughout the various recent studies that attempt to use facial expression to predict and classify emotion, most are done through traditional machine learning methods such as convolutional neural networks and do not exhibit very high accuracy, especially in the case of prediction using non-

standardized input. To this end, we develop an Extreme Learning Machine (ELM)-based model, which has many advantages compared with traditional machine learning approaches, for robots to actively reason and respond to changing human emotions and trust levels during shared tasks.

II. METHODOLOGY

A. Data Collection

1800 sample images of human facial expressions were collected for training and testing of the personalized machine learning model. The collection of these physiological biometrics samples was conducted in Python through the OpenCV package [21, 22], which was used to capture 1800 photos of the subject's face. The photos were distributed evenly among 6 subsets of emotions: neutral, happy, angry, fear, surprise, and sadness. Photos of contempt, disgust, and anger were collectively grouped as "anger," as there was no noticeable difference in the photos for all three emotions.

After collecting these samples, the photos were processed using the DLib Facial Landmark Point Detection library [23, 24]. 68 facial landmarks were labeled in each photo and indexed from 0 through 67, as shown in Fig. 5. These labels consist of the (x,y) coordinates of each landmark and are used in the facial feature calculations in the following section.

B. Data Processing

10 features were defined as inputs to the ELM model. Each feature is derived from the facial landmark map and describes a distance, ratio, or angle between different landmarks. In addition to the 4 features defined by previous studies [25, 26], 6 more features were defined to improve the accuracy of the current machine learning model. The features are defined as follows:

- Eye Aspect Ratio (EAR):

$$EAR = \frac{|P37.y - P41.y| + |P38.y - P40.y|}{2 * |P36.x - P39.x|} \quad (1)$$

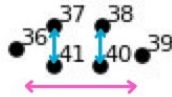


Fig. 1. Eye aspect ratio.

As shown in Fig. 1, EAR represents the aspect ratio of the eye, which is calculated using the average length of the eye divided by the width. This value is greater in expressions of surprise and fear and smaller in anger or sadness.

- Mouth Aspect Ratio (MAR):

$$MAR = \frac{P66.y - P62.y}{P54.x - P48.x} \quad (2)$$

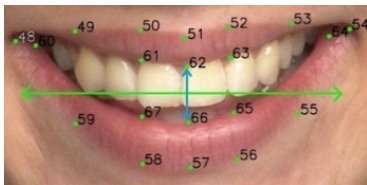


Fig. 2. Mouth aspect ratio.

As shown in Fig. 2, MAR represents the aspect ratio of the mouth, which is calculated by dividing the length of the mouth by the width. This value is higher in expressions of surprise or happiness and smaller in anger and sadness.

- Pupil circularity:

$$Circularity = \frac{4 * \pi * Area}{Perimeter^2} \quad (3)$$

where $Perimeter = Dist(P36, P37) + Dist(P37, P38) + Dist(P38, P39) + Dist(P39, P40) + Dist(P40, P41) + Dist(P41, P36)$

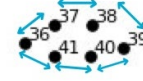


Fig. 3. Pupil circularity.

As presented in Fig. 3, pupil circularity measures the size of the pupil. Like EAR, this value is higher when the eyes are wider, such as in surprise or fear.

- MAR over EAR (MOE):

$$MOE = \frac{MAR}{EAR} \quad (4)$$

MOE is the ratio of MAR to EAR. As the MAR of an individual increases, EAR is expected to decrease. Thus, this ratio captures small changes in these two values.

- Eyebrow Angle (EBA):

$$EBA = Average \left(\frac{Dist(P19, P20)}{P20.x - P19.x}, \frac{Dist(P23, P24)}{P24.x - P23.x} \right) \quad (5)$$

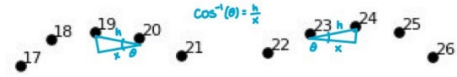


Fig. 4. Eyebrow angle.

As shown in Fig. 4, EBA measures the secant of the angle of the front of the eyebrows. A person's eyebrows are typically more angled when angry, and they are more flat when happy or relaxed.

- Chin Aspect Ratio (CAR):

$$CAR = \frac{Dist(P0, P8)}{Dist(P0, P16)} \quad (6)$$

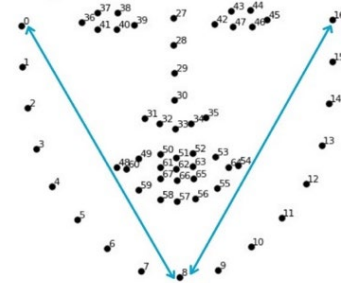


Fig. 5. Chin aspect ratio.

As presented in Fig. 5, CAR measures the distance between the upper cheek and the chin. This value typically increases when expressing surprise.

- Eyebrow to Nose (ETN):

$$ETN = (|P21 - P27| + |P22 - P27|)/2 \quad (7)$$

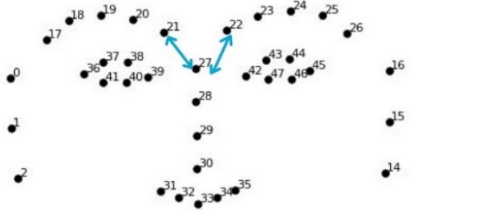


Fig. 6. Eyebrow to nose.

As shown in Fig. 6, ETN is the distance from the front of the eyebrow to the nose bridge. This distance increases when the eyebrows are lifted, such as when a person is happy, and it decreases when the eyebrows are furrowed from anger.

- Left Mouth Corner Ratio (LMCR):

$$LMCR = -1 * \frac{(P60.y - P48.y)}{(P60.x - P48.x)} \quad (8)$$

- Right Mouth Corner Ratio (RMCR):

$$RMCR = -1 * \frac{(P54.y - P64.y)}{(P54.x - P64.x)} \quad (9)$$

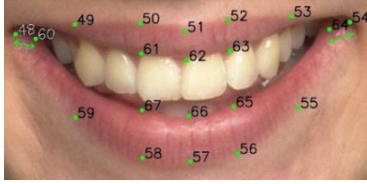


Fig. 7. Left and right mouth corner ratio.

As presented in Fig. 7, LMCR and RMCR measure the angle of the corners of the mouth. This feature is angled upwards when an individual is happy and downward when they are sad. A -1 is applied to each quotient because the y-axis is inverted, so lower y-axis values represent higher points on the face.

- Tip of Eyebrow (TOE):

$$TOE = Average((P21.y - P17.y), (P26.y - P22.y)) \quad (10)$$

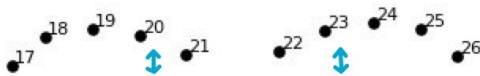


Fig. 8. Tip of eyebrow.

As shown in Fig. 8, TOE measures the y-coordinate of the front of the eyebrow versus the back. If the front is higher than usual, this shows happiness, and if it is lower, it shows anger, similar to ETN.

After the 10 features were calculated for each sample, the data were then normalized so that each value falls between 0 and 1. Normalization was achieved using Eq. (11):

$$Normalized = \frac{Sample - Max}{Min - Max} \quad (11)$$

The data was then split using an 80:20 ratio into a training subset and a testing subset.

C. Trust Reasoning via Extreme Learning Machine

The Extreme Learning Machine is the machine learning model employed to predict human emotion and trust levels using the 10 input features. ELM is a single hidden layer feedforward neural network that is lightweight and has a very fast computing time, which is advantageous for real-time training and prediction [27-29]. Other advantages of ELM compared to traditional machine learning models is that it has greater generalization performance and is relatively simple to implement [30].

The structure of ELM is shown in Fig. 9:

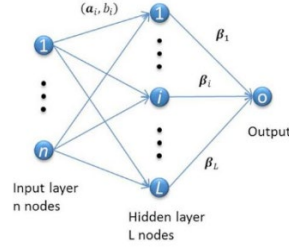


Fig. 9. Diagram of the Extreme Learning Machine.

In our work, the input layer consists of 10 nodes representing the 10 features calculated in the previous section. In the initialization phase, the ELM randomly assigns weights to each input layer node, generating the hidden layer nodes. Each node in the hidden layer is a combination of the input layer nodes, with the weight of each node stochastically generated using a continuous probability distribution. The mapping from the n-dimensional input layer to the L-dimensional hidden layer H is defined by Eq. (12):

$$h(x) = [g(a_1, b_1, x), g(a_2, b_2, x), \dots, g(a_L, b_L, x)] \quad (12)$$

where $\{(a_i, b_i)\}_{i=1}^L$ represent the randomly generated weights with a representing the input weight vector and b the hidden node bias. Increasing the number of hidden layer nodes generally increases the accuracy of the ELM.

The goal of the ELM is to calculate the correct weights to map the hidden layer nodes to the final output, pictured on the right of Figure 2. In other words, these weights serve as the activation function mappings of the input facial features to their respective emotion labels. In the training phase, these weights are calculated and stored in the β matrix.

The output T of the ELM is described in Eq. (13):

$$T = \sum_{i=1}^L \beta_i * g_i(a_i * x_k + b_i) = h(x) * \beta \quad (13)$$

Each hidden layer node, $g_i(a_i * x_k + b_i)$, is multiplied by its corresponding calculated weight from the β matrix to calculate the output T. In the context of the current project, T takes a discrete value between 0 and 5 denoting the 6 labels of human emotion.

In order to choose the prediction that is most likely to be accurate, the ELM minimizes the approximation error and the norm of the output weights using the minimization function in Eq. (14):

$$\min(\beta) \|\beta\|_{\mu}^{\sigma_1} + C\|H\beta - T\|_v^{\sigma_2} \quad (14)$$

where $H\beta$ is the output of the ELM, H is the hidden layer matrix

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} g(a_1, b_1, x_1) & \cdots & g(a_L, b_L, x_1) \\ \vdots & \vdots & \vdots \\ g(a_1, b_1, x_N) & \cdots & g(a_L, b_L, x_N) \end{bmatrix} \quad (15)$$

where N is the number of input features, and T is the learning target matrix

$$T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix} = \begin{bmatrix} t_{11} & \cdots & t_{1m} \\ \vdots & \vdots & \vdots \\ t_{N1} & \cdots & t_{Nm} \end{bmatrix} \quad (16)$$

The optimal β matrix is calculated using Eq. (17):

$$\beta = (\frac{I}{C} + H^T H)^{-1} H^T T \quad (17)$$

where I is the L -dimensional identity matrix.

The resulting β matrix is used in the testing phase to select the output with the highest likelihood based on the input features. In other words, the β matrix uses the input facial landmark data to select the emotion which is most likely being expressed. Based on the predicted emotion, the outputs are then grouped into three levels of trust: trust, low trust, and no trust. This emotion-trust mapping is depicted in Table I.

TABLE I. MAPPING OF EMOTION TO TRUST LEVELS

Emotion	Trust level
Neutral	Trust
Happy	Trust
Surprised	Low trust
Sad	Low trust
Angry	No trust
Fearful	No trust

III. RESULTS AND ANALYSIS

A. Experiment Setup

After training and tuning the ELM model to achieve high accuracy, the model was tested in a real-time experiment. The experiment setup, shown in Fig. 10, included a webcam to continuously capture frames of the participant's face, the Franka Emika Panda collaborative robot with 7 degrees of freedom [31], a ThinkStation P520, and the human-robot collaborative task. In our experiment, the task was for the human to hand an object to the robot, which would receive it in varying ways depending on the human's facial expression and trust levels. The facial expressions captured by the webcam were used by the ELM to predict the trust level of the participant in each frame, and the collaborative robot moved in response to this prediction. The trust level prediction and robotic arm movement planning were all computed on the ThinkStation. Depending on the trust level predicted, the robot performed one of three actions, which are shown in Table II. If the human shows an expression of trust, the robot will continue to work at a relatively efficient pace. However, if the human shows low trust, such as a surprised expression, it will slow down and try to gently gain the human's trust. Lastly, if the human shows signs of no trust, the robot will

back away completely until the human shows trust again. Each scenario was tested multiple times, and the robot's responses were recorded.

TABLE II. MAPPING OF TRUST LEVELS TO ROBOT MOVEMENTS

Trust level	Robot movement
Trust	Move to the human and pick the object at a normal pace.
Low trust	Move to the human and pick the object at a slower pace.
No trust	Back away and do not attempt to pick the object from the human.



Fig. 10. Experimental setup.

B. ELM Parameter Tuning

The ELM was trained using $n = 1,440$ samples, $L = 500$ hidden units, rectified linear activation function, and a regularization factor $C = 1$. The regularization parameter C determines how much the ELM should generalize from the training to the testing subsets. A very high value for C will cause the ELM to accurately learn and predict the correct output for the training set, but it will fail to generalize the results to the testing set. A very low value of C will overgeneralize the predictions, and the accuracy of the ELM will be low overall. Thus, as shown in Fig. 11, comparing the training and testing accuracy (Y-axis) for a range of C values (X-axis), a value of 1 was chosen for the ELM. This value makes sure that the ELM is able to achieve high accuracy but does not overgeneralize when making predictions.

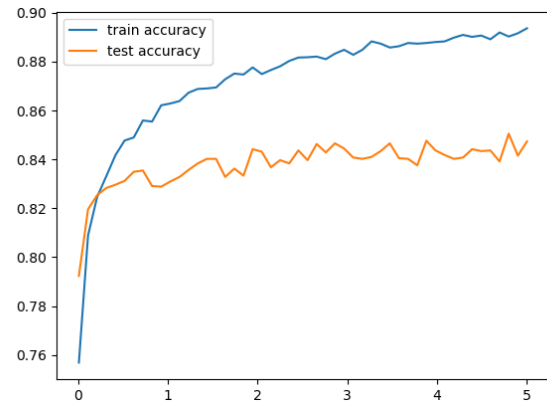


Fig. 11. Training and testing accuracies vs regularization factor C .

C. Offline Testing

The overall testing accuracy of the ELM in predicting human trust levels is approximately 91%. The average running time was 0.6 seconds, exhibiting the ELM's fast learning pace. The ELM

is able to predict the correct corresponding trust level when given a human facial expression input with high accuracy. In the training phase, the ELM consistently reached an accuracy of approximately 86% when predicting human emotion from the six established labels. However, since multiple emotions map to the same level of trust as defined in Table I, the overall accuracy for trust prediction is higher. This is partially because similarities exist between the emotions which map onto the same trust level, so they are more likely to be mislabeled as the corresponding emotion. For example, the emotions happy and neutral are more likely to be mislabeled as each other while they are rarely labeled as anger or fear since they share few similarities. By grouping similar emotions together, the ELM output is less precise but more accurate, and thus it is able to meet the goal of this project.

D. Application in Real-World Human-Robot Collaboration

The trained ELM model is validated in an online HRC task, where it predicts the trust level of its human collaborator in real-time. First, the robot assesses the human's emotional state and trust level. In the example in Fig. 12, the human is happy. The robot correctly predicts that trust is being displayed and reaches out to receive the object that the human is handing off. In a real-world situation such as smart manufacturing, this process would repeat as many times as needed. The robot would continuously assess the human's trust level and follow up with the appropriate actions. In a scenario like Fig. 12, the robot will continue to work at a relatively efficient pace, matching the human collaborator's high level of energy and trust. Online testing of the ELM showed that it is a promising method that can be extremely helpful in an online HRC situation.

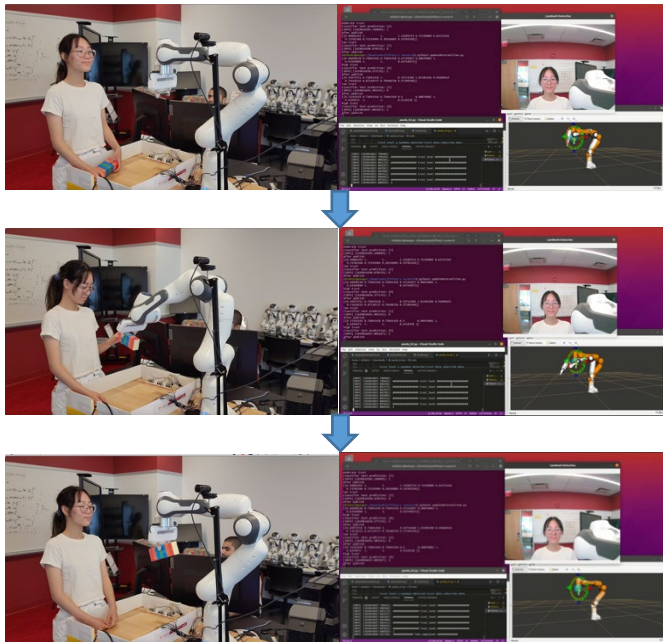


Fig. 12. Robot response to trust facial expression.

The robot can also detect and react to situations with low or no trust, as shown in Fig. 13. human emotions cannot always be positive, and the robot must learn this in order for effective

collaboration to take place. In Fig. 13, the human is angry. The robot correctly predicts that no trust is being shown, and it backs away to give the human time to calm down and recollect their emotions. Only when the human shows a low trust or trust expression, such as the one in Fig. 12, will the robot resume the collaborative task. This validation exemplifies the robot's ability to understand and react to human emotion by choosing the best course of action to accommodate their human collaborator's needs.

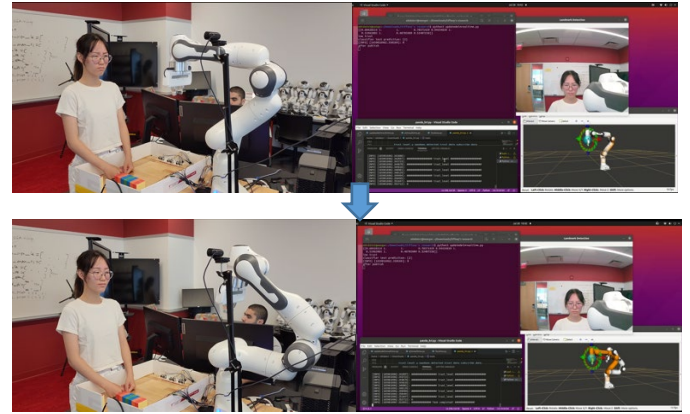


Fig. 13. Robot response to no trust facial expression.

IV. CONCLUSIONS AND FUTURE WORK

To achieve safe and efficient human-robot collaboration, we have developed and tested a novel and effective model for robots to actively reason and respond to changing human emotions and trust levels during shared tasks. An extreme learning machine was used to enable the robot to predict human emotion and trust levels with over 90% accuracy. Using this model, the robot is able to correctly recognize and accommodate humans during collaborative tasks in real-time. We have implemented a real-world validation experiment in the context of human-robot object hand-over, which showed the robot's ability to correctly identify and predict the human's trust levels in real-time and assist the human accordingly in human-robot collaborative tasks. One area for future research is to develop methods to improve the online application of this ELM model. Although the ELM itself is very fast, the calculations of each facial landmark on the face, as well as applying the defined features to compute all the input nodes to the ELM, took numerous seconds to process one frame. Thus, there was a lag between the time of the input facial expression and the actual movement of the robot. In a real-world application, this calculation time would have to be further shortened to facilitate more effective HRC. Overall, the proposed approach to trust prediction shows an efficient and accurate solution to improve effectiveness in human-robot collaboration. Further research will be conducted to improve calculation speeds as well as further generalization of facial expressions in order to build a more accurate and precise model for HRC.

ACKNOWLEDGMENT

This work was supported in part by the National Science Foundation under Grant REU CNS-2050548 and in part by the National Science Foundation CNS-2104742. The authors would like to thank the faculty at Montclair State University's

Department of Computer Science, namely Dr. Boxiang Dong and Dr. Bharath Samanthula.

REFERENCES

- [1] E. Matheson, R. Minto, E. G. Zampieri, M. Faccio, and G. Rosati, "Human-Robot Collaboration in Manufacturing Applications: A Review," *Robotics*, vol. 8, no. 4, p. 100, 2019.
- [2] M. A. R. Garcia, E. Rauch, D. Salvalai, and D. T. Matt, "AI-Based Human-Robot Cooperation for Flexible Multi-Variant Manufacturing."
- [3] Q. Liu, Z. Liu, W. Xu, Q. Tang, Z. Zhou, and D. T. Pham, "Human-robot collaboration in disassembly for sustainable manufacturing," *IJPR*, vol. 57, no. 12, pp. 4027-4044, 2019.
- [4] W. Wang, R. Li, Y. Chen, Y. Sun, and Y. Jia, "Predicting Human Intentions in Human-Robot Hand-Over Tasks Through Multimodal Learning," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, pp. 2339-2353, 2022.
- [5] H. Su, A. Mariani, S. E. Ovrur, A. Menciassi, G. Ferrigno, and E. D. Momi, "Toward Teaching by Demonstration for Robot-Assisted Minimally Invasive Surgery," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 484-494, 2021.
- [6] I. Kruijff-Korbayová *et al.*, "TRADR Project: Long-Term Human-Robot Teaming for Robot Assisted Disaster Response," *KI - Künstliche Intelligenz*, vol. 29, no. 2, pp. 193-201, 2015.
- [7] J. Mišeikis *et al.*, "Lio-A Personal Robot Assistant for Human-Robot Interaction and Care Applications," *IEEE Robotics and Automation Letters*, vol. 5, no. 4, pp. 5339-5346, 2020.
- [8] P. Tsarouchi, S. Makris, and G. Chryssolouris, "Human-robot interaction review and challenges on task planning and programming," *International Journal of Computer Integrated Manufacturing*, vol. 29, no. 8, pp. 916-931, 2016.
- [9] W. Wang, R. Li, Z. M. Diekel, Y. Chen, Z. Zhang, and Y. Jia, "Controlling Object Hand-Over in Human-Robot Collaboration Via Natural Wearable Sensing," *IEEE Transactions on Human-Machine Systems*, vol. 49, no. 1, pp. 59-71, 2019.
- [10] B. D. Argall, S. Chernova, M. Veloso, and B. Browning, "A survey of robot learning from demonstration," *Robotics and autonomous systems*, vol. 57, no. 5, pp. 469-483, 2009.
- [11] W. Wang, R. Li, Y. Chen, Z. M. Diekel, and Y. Jia, "Facilitating Human-Robot Collaborative Tasks by Teaching-Learning-Collaboration From Human Demonstrations," *IEEE Transactions on Automation Science and Engineering*, vol. 16, no. 2, pp. 640-653, 2019, doi: 10.1109/TASE.2018.2840345.
- [12] D. R. Billings, K. E. Schaefer, J. Y. Chen, and P. A. Hancock, "Human-robot interaction: developing trust in robots," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 2012, pp. 109-110.
- [13] H. Soh, Y. Xie, M. Chen, and D. Hsu, "Multi-task trust transfer for human-robot interaction," *The International Journal of Robotics Research*, 2019, doi: 10.1177/0278364919866905.
- [14] R. H. Wortham and A. Theodorou, "Robot transparency, trust and utility," *Connection Science*, vol. 29, no. 3, pp. 242-248, 2017.
- [15] H. Khalid *et al.*, "Exploring Psycho-Physiological Correlates to Trust: Implications for Human-Robot-Human Interaction," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 60, pp. 697-701, 09/01 2016.
- [16] Z. R. Khavas, S. R. Ahmadzadeh, and P. Robinette, "Modeling Trust in Human-Robot Interaction: A Survey," in *Social Robotics*, Cham, A. R. Wagner *et al.*, Eds., 2020: Springer International Publishing, pp. 529-541.
- [17] C. Hannum, R. Li, and W. Wang, "Trust or Not?: A Computational Robot-Trusting-Human Model for Human-Robot Collaborative Tasks," in *2020 IEEE International Conference on Big Data (Big Data)*, 2020, pp. 5689-5691.
- [18] D. Matsumoto, & Hwang, H. S., "Reading facial expressions of emotion," *Psychological Science Agenda*, 2011.
- [19] Y. Said and M. Barr, "Human emotion recognition based on facial expressions via deep learning on high-resolution images," *Multimedia Tools and Applications*, vol. 80, no. 16, pp. 25241-25253, 2021.
- [20] S. Liu, H. Wang, and M. Pei, "Facial-expression-aware Emotional Color Transfer Based on Convolutional Neural Network," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 18, no. 1, p. Article 8, 2022, doi: 10.1145/3464382.
- [21] G. Bradski and A. Kaehler, *Learning OpenCV: Computer vision with the OpenCV library*. "O'Reilly Media, Inc.", 2008.
- [22] J. Howse, *OpenCV computer vision with python*. Packt Publishing Ltd, 2013.
- [23] N. Boyko, O. Basystiuk, and N. Shakhovska, "Performance evaluation and comparison of software for face recognition, based on dlib and opencv library," in *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*, 2018: IEEE, pp. 478-482.
- [24] S. Mohanty, S. V. Hegde, S. Prasad, and J. Manikandan, "Design of real-time drowsiness detection system using dlib," in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, 2019: IEEE, pp. 1-4.
- [25] R. Ghoddoosian, M. Galib, and V. Athitsos, "A realistic dataset and baseline temporal model for early drowsiness detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2019.
- [26] Drowsiness Detection with Machine Learning. [Online]. Available: <https://towardsdatascience.com/drowsiness-detection-with-machine-learning-765a16ca208a>.
- [27] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489-501, 2006.
- [28] J. Tang, C. Deng, and G.-B. Huang, "Extreme learning machine for multilayer perceptron," *IEEE transactions on neural networks and learning systems*, vol. 27, no. 4, pp. 809-821, 2016.
- [29] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513-529, 2012.
- [30] H. Guang-Bin, Z. Qin-Yu, and S. Chee-Kheong, "Extreme learning machine: a new learning scheme of feedforward neural networks," in *2004 IEEE International Joint Conference on Neural Networks (IEEE Cat. No.04CH37541)*, 2004, vol. 2, pp. 985-990 vol.2, doi: 10.1109/IJCNN.2004.1380068.
- [31] S. Bier, R. Li, and W. Wang, "A Full-Dimensional Robot Teleoperation Platform," in *2020 IEEE International Conference on Mechanical and Aerospace Engineering*, 2020: IEEE, pp. 186-191.