Active Vision Control Policies for Face Recognition using Deep Reinforcement Learning

Pavlos Tosidis, Nikolaos Passalis, and Anastasios Tefas
Computational Intelligence and Deep Learning Group, AIIA Lab
School of Informatics, Faculty of Sciences
Aristotle University of Thessaloniki
Thessaloniki, Greece
{ptosidis, passalis, tefas}@csd.auth.gr

***Abstract*—Robotic systems are capable of interacting with their environment in order to better sense their surroundings. This key ability of robotic systems is often ignored when developing Deep Learning models, since the later are usually trained using static datasets. This limits the ability of robots to perceive the environment in challenging scenarios. On the other hand, integrating perception and action in tightly coupled systems while operating on-the-edge, holds the credentials for deploying DL-enabled robots in such scenarios; Thus leading to more robust agents that can solve challenging tasks more accurately. In this work, we investigate whether active perception approaches can be employed and integrated into robotic systems in order to improve face recognition accuracy, as well as, study the effect of such an approach on the computational requirements for edge applications. To this end, we propose a DRL-based control approach for training agents that are able to identify task-relevant objects, as well as, issue the appropriate control commands to acquire better results. Through the conducted experimental evaluation, we demonstrate that the proposed method leads to significant improvements in face recognition over the rest of the evaluated approaches by providing accurate control commands.**

***Index Terms*—Active Perception, Active Vision, Deep Reinforcement Learning**

I. Introduction

Recent advances in Deep Learning (DL) led to a number of spectacular applications, ranging from self-driving cars and robots that outperform humans in various tasks [1]. Despite the enormous success in these areas, DL methods operate in a static fashion, i.e., they do not typically provide means for interacting with the environment in order to better perceive it. This is in contrast, with the way many organisms, including humans, perceive their environment since perception and action usually form a tightly coupled system at various levels. For example, eyes can adjust to various illumination conditions while we tend to examine an object from different angles and/or distances when we are uncertain about it. This process is called *active perception* [2], and it is thought to be a critical component of robotic agents that can work in challenging real-world scenarios.

There have been several recent attempts to integrate active perception principles into DL models [3], [4]. Most of them focused on robotics tasks, where they attempt to appropriately manipulate a camera and/or a robot in order to improve the accuracy of the models. However, training DL models for such tasks is not trivial, since most datasets used for training DL models do not provide the appropriate data and/or annotations that can be exploited in active perception scenarios. Indeed, active perception requires an agent that can *interact* with its environment and acquire an improved view of the world. To overcome this limitation, existing methods either employ simple handcrafted rules for implementing active perception feedback [3], or use multi-view datasets to simulate some of the effects of active perception feedback [4]. However, due to the lack of appropriate datasets, such methods are still usually trained with simplistic rules, e.g., to predict if moving left/right will increase/decrease the confidence on correctly recognizing a person [4]. Another closely related line of work employs Deep Reinforcement Learning (DRL) algorithms to perform a specific control task [5], [6], [7], e.g., acquire a frontal view of a person [8]. Despite the effectiveness of DRL approaches in these robotics tasks, applying them on challenging computer vision tasks typically require realistic simulation environments and/or appropriate training methods, e.g., *sim2real* approaches [9]. At the same time, the lengthy training time of DRL methods further limits their applications in robotics. As a result, despite their enormous potential for developing active perception approaches their application faces significant obstacles.

The main contribution of this work is to propose a DRL-based active perception approach integrated with state-of-the-art DL-based face recognition models. More specifically, our goal is to investigate whether active perception approaches can be employed and integrated into robotic systems, in order to improve face recognition results, as well as, study the effect of such an approach on the computational requirements. To this end, we propose a DRL-based control approach for training agents that are able to identify and focus on task-relevant objects, i.e., humans, as well as issue appropriate control commands accordingly to acquire better results. To train and evaluate the proposed method, we developed a simulation environment using the Webots simulator [10] and generated several 3D human models using the MakeHuman software [11]. The proposed method aims to control a drone, equipped with a camera, in order to improve face recognition results over existing baseline and rule-based active perception approaches. Indeed, as the experimental results demonstrate, the proposed method managed to lead to significant improvements in face recognition over the rest of the evaluated approaches by issuing the appropriate control commands. Indeed, the trained agents showed an emergent behavior that can resemble those of humans, e.g., move closer or around a person in order to more confidently identify it. At the same time, it is demonstrated that the proposed method can also lead to computational savings under certain conditions.

The rest of the paper is structured as followed. First, the related work is briefly introduced in Section II. Then, the proposed method is introduced in Section III. The experimental evaluation is provided in Section IV, while conclusions are drawn in Section V.

II. Related Work

Face recognition research in the past years has made tremendous leaps. From traditional approaches that represent faces with hand-crafted features extracted from an image [12], to modern deep learning approaches that automatically learn the distinctive features of a face, when trained on massive datasets [13], [14], [15]. The face recognition pipeline of such approaches typically consists of four stages: a) face detection and cropping, b) (optionally) face alignment, c) feature extraction, and d) classification/verification. The two first stages are often considered as preprocessing stages. A face recognition model requires an input image that is carefully cropped and aligned. This preprocessed image is then fed into a DL model which extracts a discriminative feature vector. Finally, this vector is compared to a set of feature vectors of people of interest [13], [14], [15], performing the final classification or verification task. The method proposed in this paper is orthogonal to these approaches, since it can be readily combined with any face recognition model and further increase its accuracy. Indeed, as demonstrated in Section IV, the proposed method can be readily combined with a state-of-the-art DL-based face recognition system and increase its accuracy by integrating it into an active vision pipeline.

This work is also closely related to active perception approaches. According to Bajscy [16], [17], an actively perceiving agent is one which can, among others, appropriately control its mechanical components in order to enable the best sensing of its surroundings, as well as, select the best viewpoint to achieve the task in hand. However, there are only a few recent approaches to active face recognition using DL [3], [4]. An active face recognition system that employs a DL model to extract the facial features and a controller module to act based on the results of the DL model was proposed in [3]. The controller module works as a rule-based controller that selects the most appropriate action according to the face recognition confidence and predefined thresholds for each action. A fully end-to-end trainable DL-based approach was also proposed in [4], where a DL model was trained to output both the face feature embeddings, as well as, a suggested action. The network was trained on a small dataset containing facial images at various pans and tilts, providing a proof-of-concept demonstration for a DL-based pipeline for active face recognition. Moreover, this approach cannot fully exploit the potential of active perception, since it only considers 1-step actions for training the control branch of the DL model.

The proposed method goes beyond these approaches by employing a powerful RL-based formulation that is both end-to-end trainable and does not make any assumption regarding the control policy. In this way, more advanced policies can be discovered without introducing any strong prior, using handcrafted rules either for training or inference. However, the proposed method requires a realistic simulation environment for training, since the control module cannot be trained using the existing static datasets. To overcome this limitation, in this work, we employed the realistic Webots simulator, along with 3D human models generated using the MakeHuman software. A sample of the generated human models can be seen in Fig. 2. Furthermore, both aforementioned works require the use of a face detector to appropriately crop the face image before feeding it to the face recognition module. On the other hand, the proposed method allows for significantly reducing the computational requirements by working independently of the face recognition model. In this way, a lightweight DL model is used for performing control and the heavy face recognition pipeline (face detection and recognition) is only employed when deemed appropriate.

III. Proposed method

Let $\mathbf{x} \in \mathbb{R}^{W \times H \times C}$ be an image that contains a face to be recognized, where W, H and C are the width, height, and number of channels of the corresponding image. As described before, face recognition algorithms require to first employ a face detection model to detect and crop the bounding box that encloses each face. Therefore, let

$$\mathbf{x}_p = f_p(\mathbf{x}) \in \mathbb{R}^{W_p \times H_p \times C_p} \tag{1}$$

be the cropped face image, where the notation $f_p(\cdot)$ is used to refer to the face detector and preprocessing pipeline employed to crop the image and W_p, H_p, and C_p are the width, height and number of channels of the cropped image. Most recent deep face recognition methods, e.g., [15], aim at learning an appropriate model $\mathbf{y} = f_r(\mathbf{x}_p) \in \mathbb{R}^D$ that will extract a discriminative identify-oriented representation from each face image, where D denotes the dimensionality of the embedding space used for representing the input face images.

Different loss functions have been proposed to train the face recognition model f_r, to extract discriminative embeddings. In this work, we employ the *Additive Angular Margin Loss* [13], which is minimized when embeddings that belong to the same identity are close, while the representations of face images that do not belong to the same person are far. After training the model $\mathbf{y} = f_r(\mathbf{x}_p)$, the identity of a person depicted in an image $\mathbf{x}_p$ can be obtained simply by calculating the Euclidean distance between the feature vector of that image and the feature vectors on a database that contains images $\mathbf{x}_i$ of known identities, i.e., $\mathcal{X}_d = \{(\mathbf{x}_i, l_i)\}$, where l_i is the identity of the person depicted in the i-th image. Therefore, during inference

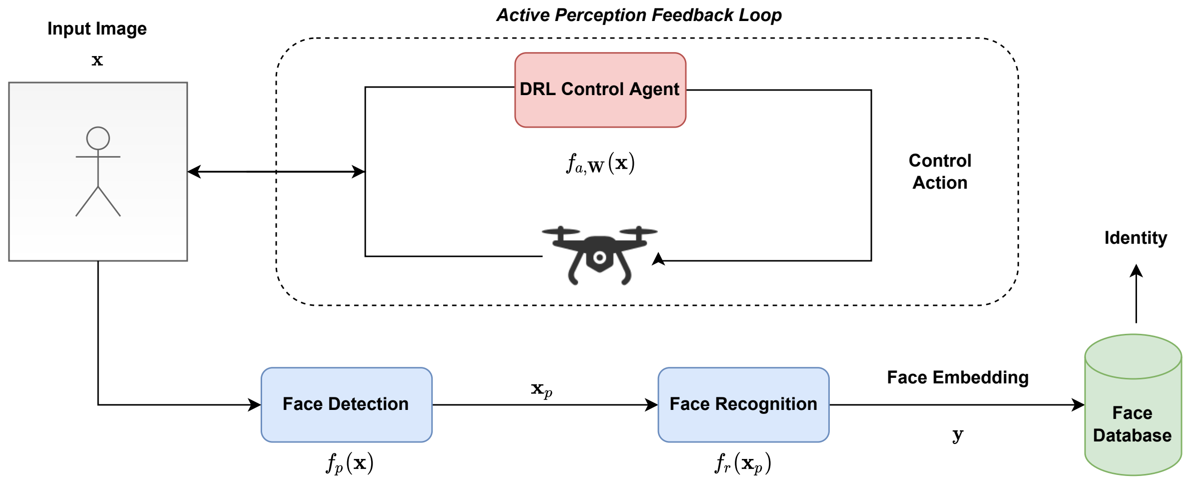

Fig. 1. Proposed Active Perception Approach: A DRL agent is employed to issue control commands in order to acquire the most appropriate view for improve face recognition accuracy.

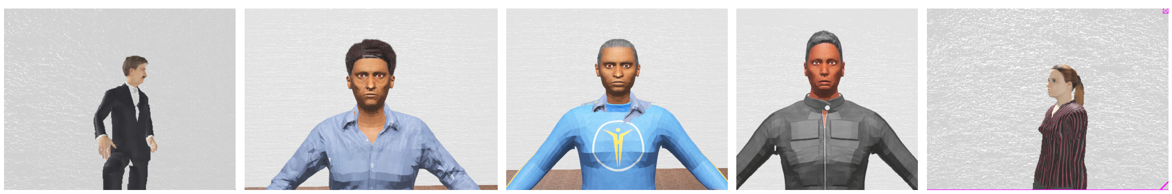

Fig. 2. Realistic human models generated using MakeHuman.

the identity l of a person appearing in a novel image $\mathbf{x}$ is obtained as $l = l_k$, where

$$k = \arg\min_i ||f(\mathbf{x}_i) - f(\mathbf{x})||_2 \ (\forall(\mathbf{x}_i, l_i) \in \mathcal{X}_d). \quad (2)$$

The proposed method aims to teach an agent that can appropriately control a robot in order to re-acquire an input image $\mathbf{x}$ in which the depicted person can be more confidently identified, as shown in Fig. 1. To this end, another model $f_{a,\mathbf{W}}(\mathbf{x})$ is introduced, where $\mathbf{W}$ denotes the trainable parameters of the model. This model is responsible for controlling the position and orientation of the robot in order to recognize the human in the scene with the greatest confidence possible. Five possible actions are supported by this model:

1) *stay*, where the robot does not move and initiates the face recognition pipeline,
2) *move forward/backward*, where the robot moves forward/backward, and
3) *move left/right*, where the robot rotates and translates its position on a predefined arc either on the left or right.

All actions translate into discrete actions in the simulation environment, e.g., moving forward/backward moves the agent 0.1m to the corresponding direction. Note that the face recognition pipeline is only employed when the agent issues the *stay* command. This can significantly reduce the computational complexity of the employed pipeline, since both the face detection and recognition models run only when the control agent is confident enough that the depicted person can be indeed recognized. This is in contrast with other active vision approaches that require all models to run simultaneously, e.g., [4].

The proposed agent is trained using DRL. More specifically, the Proximal Policy Optimization (PPO) algorithm was used [6]. The reward used for training the DRL agent was defined based on the face recognition confidence of a pre-trained face recognition model. If the person was not correctly identified, the agent received a reward of 0. Therefore, after identifying the embedding of the most similar person (k) in the database according to (2), the reward at time-step t can be defined as:

$$r_t = \begin{cases} c & \text{if } ||\mathbf{y} - \mathbf{y}_k||_2 < a \\ 0 & \text{otherwise} \end{cases}, \quad (3)$$

where $\mathbf{y}_k = f(\mathbf{x}_k)$, c is the face recognition confidence, and a is a cut-off value for recognizing a person, i.e., if the Euclidean distance is larger than a, then we assume that the person has not been recognized. The face recognition confidence is calculated simply as the negative of the normalized Euclidean distance between the current embedding vector and the embedding vector of the most similar person in the database:

$$c = 1 - \frac{||\mathbf{y} - \mathbf{y}_l||_2}{a}. \quad (4)$$

Note that c is bounded between 0 and 1, since the Euclidean distance cannot exceed the value of a, due to the used cut-off threshold.

A deep convolutional neural network, receiving input images of 400×300 pixels, was used to implement the policy, i.e., $f_{a,\mathbf{W}}(\mathbf{x})$, as well as, to estimate the advantage value. A lightweight DL model was used to this end. The architecture of the model was the following: 2 convolutional layers with 16 (8×8) and 32 (4×4) filters respectively utilizing the *ReLU* activation function, one fully connected layer of 256 neurons and two output layers. The first layer was responsible for providing the policy function $f_{a,\mathbf{W}}(\mathbf{x})$ function. This layer was composed of the same number of neurons as the number of available actions and employed the *softmax* function to provide the final action probabilities. The other one was used for implementing the critic function and was composed of one output neuron providing the current advantage. To constraint the advantage values the *tanh* activation function was employed for this branch.

Each training episode lasts 1,000 steps and the agent initially starts at a random position around the human model, which also faces at different directions in each episode. The simulation world consists of a square room, a human model at the center and a drone robot controlled by the DRL agent. After each time-step the agent must decide whether or not its position and orientation must be adjusted. For training the network we employed the Adam optimization algorithm with a learning rate of 0.0003, while a total of 10,000,000 steps where performed during the training.

IV. Experimental evaluation

The proposed method was evaluated under two different setups. In the first setup, the agent was trained to select one of the first three actions ("stay", "move forward" and "move backward"). In this setup, the human was always initialized to be in front of the drone and correctly centered. The aim of this ablated setup was to evaluate the ability of the agent to control the movement in just one axis in order to increase the face recognition model's confidence. In the second setup, the agent was allowed to select any of the available control actions, evaluating the ability of the proposed method to perform more complicated sequences of actions, in order to improve face recognition accuracy.

The proposed method was also compared to two other baselines. First, a face recognition pipeline was employed to evaluate the ability of existing approaches to detect and recognize humans at different distances. This setup was called "static" in the conducted experiments. Then, we also employed an active perception enabled agent that uses rules. The rule-based agent employed a face detector to detect if a face exists in the scene. If a face is found, it outputs the appropriate control commands to center it to its field of view based on the detected bounding box and then moves forward based on the face recognition model's confidence, until it reaches the maximum confidence. This method is called "rule-based" in the conducted experiments. For all methods we used ArcFace [13] for the face recognition and RetinaFace [18] for the face detection. Furthermore, the database of known identities consists of one feature vector extracted from cropped frontal face images of 5 different human models that were used for the conducted experiments.

TABLE I
Evaluation for controlling one axis (Setup 1). Face recognition confidence is reported. A value of zero is used when a person is not correctly recognized.

Distance	Static	Rule-based	Proposed
1m	0.76	0.77	**0.78**
2m	0.55	0.77	**0.78**
3m	0.43	**0.78**	0.77
4m	0.19	**0.76**	**0.76**
5m	0	**0.77**	**0.77**
6m	0	0.63	**0.75**
7m	0	0	**0.76**
10m	0	0	**0.71**
15m	0	0	**0.68**
20m	0	0	**0.48**

In this setup the drone is initialized at a distance of 20m, which decreases by 1m in every evaluation episode. We report the average face recognition confidence reached for 4 different human models at each distance.

The experimental results for the first setup are reported in Table I. In this setup, the drone was positioned at various distances in front of the human subject, ranging from 1m to 20m. Using a static setup, where the drone does not move, allows for recognizing persons only up to 4 meters. On the other hand, the rule-based approach, which allows the drone to move closer to the subject at hand, enables confident recognition up to 6 meters. This demonstrates that active perception, even when implemented using simple rules, can indeed lead to improved perception accuracy. The proposed method outperforms all the other evaluated methods, since it allows for confidently recognizing persons even up to 15m, while it can work correctly even for larger distances (up to 20m).

Similar conclusions can be also drawn for the evaluation results reported in Table II, using the second setup. Again, the proposed method can significantly improve the view invariance of face recognition, allowing not only for recognizing the persons at different distances, but also in a wide range of different angles, for some of which most face recognition pipelines typically fail. It is worth noting that at a distance of 7m only the proposed method manages to work correctly, while the provided face recognition accuracy is virtually the same with a robot that was initially placed in close distance in front of a human subject. Additionally, note that the proposed method does not need a face detector to actively perceive the surroundings, which can lead to significant performance improvements. Indeed, the proposed method runs on 180 FPS on average, while the rule-based approach runs on 62 FPS. A GPU-enabled workstation (8 GB VRAM, 9 TFLOPS) was used for measuring the performance of the evaluated agents.

TABLE II
EVALUATION FOR CONTROLLING TWO AXES (SETUP 2). FACE RECOGNITION CONFIDENCE IS REPORTED. A VALUE OF ZERO IS USED WHEN A PERSON IS NOT CORRECTLY RECOGNIZED.

Angle	Static	Rule-based	Proposed
3m			
0°	0.48	**0.77**	0.76
60°	0.24	0.32	**0.78**
120°	0	0	**0.78**
180°	0	0	**0.76**
240°	0	0	**0.79**
300°	0	0.14	**0.78**
5m			
0°	0.18	0.53	**0.79**
60°	0	0.32	**0.79**
120°	0	0	**0.79**
180°	0	0	**0.77**
240°	0	0	**0.78**
300°	0	0.13	**0.79**
7m			
0°	0	0	**0.78**
60°	0	0	**0.78**
120°	0	0	**0.77**
180°	0	0	**0.76**
240°	0	0	**0.77**
300°	0	0	**0.77**

In this setup the drone is initialized at three different distances, while for each distance we also evaluated the performance of the agents at 6 different angles around the human model. We report the average face recognition confidence reached for 4 different human models at each distance.

V. CONCLUSIONS

Despite its potential in a wide variety of robotics systems, active perception using DL models is a field not yet explored deeply. Indeed, it is expected that a robot should be able to interact with its environment to better understand it, improve situational awareness and make informed decisions. In this work we demonstrated that active perception approaches can indeed lead to improved face recognition accuracy in a wide variety of setups, including challenging ones, e.g., when images taken from the back side of humans or faces appear too small to be detected by traditional face detection models. At the same time, it was also shown that DRL can be efficiently integrated into such active perception pipelines, given that the appropriate reward function has been defined. Also, the proposed method can lead to performance improvements, apart from more accurate agents, since it can replace part of the existing DL pipelines. This work paves the way for more advanced DRL-based active perception approaches for human-centric perception.These approaches can be trained on more complex simulation environments, employ *sim2real* approaches [9], while also consider the trade off between the expected accuracy improvement and energy expenditure for each control action.

ACKNOWLEDGMENT

This work was supported by the European Union's Horizon 2020 Research and Innovation Program (OpenDR) under Grant 871449. This publication reflects the authors' views only. The European Commission is not responsible for any use that may be made of the information it contains.

REFERENCES

[1] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436, 2015.

[2] Ruzena Bajcsy, "Active perception," 1988.

[3] Masaki Nakada, Han Wang, and Demetri Terzopoulos, "Acfr: Active face recognition using convolutional neural networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2017, pp. 35–40.

[4] Nikolaos Passalis and Anastasios Tefas, "Leveraging active perception for improving embedding-based deep face recognition," in *Proceedings of the International Workshop on Multimedia Signal Processing*, 2020, pp. 1–6.

[5] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller, "Playing atari with deep reinforcement learning," 2013.

[6] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov, "Proximal policy optimization algorithms," 2017.

[7] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin A. Riedmiller, Andreas Fidjeland, Georg Ostrovski, Stig Petersen, Charlie Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, pp. 529–533, 2015.

[8] A Tzimas, Nikolaos Passalis, and Anastasios Tefas, "Leveraging deep reinforcement learning for active shooting under open-world setting," in *Proceedings of the IEEE International Conference on Multimedia and Expo*, 2020, pp. 1–6.

[9] Andrei A Rusu, Matej Večerík, Thomas Rothörl, Nicolas Heess, Razvan Pascanu, and Raia Hadsell, "Sim-to-real robot learning from pixels with progressive nets," in *Proceedings of the Conference on Robot Learning*, 2017, pp. 262–270.

[10] Olivier Michel, "Cyberbotics ltd. webots™: professional mobile robot simulation," *International Journal of Advanced Robotic Systems*, vol. 1, no. 1, pp. 5, 2004.

[11] Manuel Bastioni, Simone Re, and Shakti Misra, "Ideas and methods for modeling 3d human figures: the principal algorithms used by makehuman and their implementation in a new approach to parametric modeling," in *Proceedings of the Bangalore Annual Compute Conference*, 2008, pp. 1–6.

[12] Jun Zhang, Yong Yan, and Martin Lades, "Face recognition: eigenface, elastic matching, and neural nets," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1423–1435, 1997.

[13] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4690–4699.

[14] Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li, and Wei Liu, "Cosface: Large margin cosine loss for deep face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5265–5274.

[15] Weiyang Liu, Yandong Wen, Zhiding Yu, Ming Li, Bhiksha Raj, and Le Song, "Sphereface: Deep hypersphere embedding for face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 212–220.

[16] R. Bajcsy, "Active perception," *Proceedings of the IEEE*, vol. 76, no. 8, pp. 966–1005, 1988.

[17] Ruzena Bajcsy, Yiannis Aloimonos, and John K Tsotsos, "Revisiting active perception," *Autonomous Robots*, vol. 42, no. 2, pp. 177–196, 2018.

[18] Jiankang Deng, Jia Guo, Evangelos Ververas, Irene Kotsia, and Stefanos Zafeiriou, "Retinaface: Single-shot multi-level face localisation in the wild," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 5203–5212.