






Article

A Meal-Assistance Robot System for Asian Food and Its Food Acquisition Point Estimation and User Interface Based on Face Recognition

Iksu Choi ^{1,2} , KwangEun Ko ¹ , Hajun Song ³, Byung-Jin Jung ³ , Jung-Hoon Hwang ³, Hyungpil Moon ^{2,*} 
and Gi-Hun Yang ^{1,*} 

¹ AI Robot R&D Department, Korea Institute of Industrial Technology (KITECH), Ansan 15588, Republic of Korea

² Mechanical Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea

³ Korea Electronics Technology Institute (KETI), Bucheon 14502, Republic of Korea

* Correspondence: hyungpil@g.skku.edu (H.M.); yanggh@kitech.re.kr (G.-H.Y.); Tel.: +82-31-290-4842 (H.M.); +82-31-8040-6389 (G.-H.Y.)

Abstract: Various meal-assistance robot (MAR) systems are being studied, and several products have already been commercialized to alleviate the imbalance between the rising demand and diminishing supply of meal care services. However, several challenges remain. First, most of these services can serve limited types of western food using a predefined route. Additionally, their spoon or fork sometimes makes it difficult to acquire Asian food that is easy to handle with chopsticks. In addition, their limited user interface, requiring physical contact, makes it difficult for people with severe disabilities to use MARs alone. This paper proposes an MAR system that is suitable for the diet of Asians who use chopsticks. This system uses Mask R-CNN to recognize the food area on the plate and estimates the acquisition points for each side dish. The points become target points for robot motion planning. Depending on which food the user selects, the robot uses chopsticks or a spoon to obtain the food. In addition, a non-contact user interface based on face recognition was developed for users with difficulty physically manipulating the interface. This interface can be operated on the user's Android OS tablet without the need for a separate dedicated display. A series of experiments verified the proposed system's effectiveness and feasibility.

Keywords: meal-assistance robot; deep learning; instance segmentation; human–robot interaction; face recognition



Citation: Choi, I.; Ko, K.; Song, H.; Jung, B.-J.; Hwang, J.-H.; Moon, H.; Yang, G.-H. A Meal-Assistance Robot System for Asian Food and Its Food Acquisition Point Estimation and User Interface Based on Face Recognition. *Appl. Sci.* **2023**, *13*, 3216. <https://doi.org/10.3390/app13053216>

Academic Editor: Carlos A. Jara and Juan Antonio Corrales Ramon

Received: 20 January 2023

Revised: 16 February 2023

Accepted: 21 February 2023

Published: 2 March 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

According to the latest statistics from the World Health Organization, more than 1 billion people, about 15% of the world's population, currently experience a disability, and this number is steadily growing due to the aging population, the rapid spread of chronic diseases, and the increasing prevalence of non-communicable diseases. Moreover, it is reported that nearly 1.9 million people, or 3.8% of the world's population, suffer from significant movement difficulties due to a physical disability or a brain lesion disorder [1]. Because they have difficulties performing activities of daily living (ADLs) alone, such as eating, dressing, toileting, and bathing, they often need assistance from a caregiver or family member [2]. However, while the need for care is increasing, the working-age population is decreasing, so there is a shortage of human resources to provide care. This problem is expected to worsen over time. In addition, family care was common in the past, but the use of a professional workforce is becoming more common. Therefore, expanding the care supply is attracting attention as an important task.

Among the ADLs, feeding is the most important activity to maintain life. This should be performed every day and at every meal. However, for the caregiver, this is a burdensome

task that requires a lot of time and effort [3]. This situation is similar for those requiring care, who often feel sorry when they receive meal assistance from their family members or caregivers. The fact that they cannot feed themselves can make cause them to feel depressed and not independent. Therefore, if the disabled and the elderly can eat independently, without the help of others, they would maintain their self-esteem and improve their quality of life [4,5]. In this situation, robots offer alternative solutions to meet the demands of daily problems in various fields [6,7]. In particular, meal-assistance robots (MARs) have emerged as a solution to replace human caregivers for patients in the field of meal assistance, leading to the development of numerous commercial products and related research.

In this work, we would like to introduce a new portable MAR system to alleviate these issues (see Figure 1). This MAR system was invented so that the patient can eat on their own with the help of a robot, only requiring the caregivers to help with the initial setting and finishing of the meal. The MAR has 7-degrees-of-freedom (DOF), which includes a robot arm and controller, food tray, and tablet user interface. For ease of portability, we fabricated the hardware with a total weight of 4.90 kg and designed the system to be suitable for Korean-style meals in hospitals or sanatoriums. We developed software to obtain food acquisition points based on deep learning instance segmentation. Furthermore, we implemented a tablet PC user interface based on face recognition for the severely disabled. Utilizing this MAR system is expected to help ameliorate the lack of caregivers, reduce their burden regarding meal care, and improve caretakers' self-esteem and quality of life by providing better assistance.

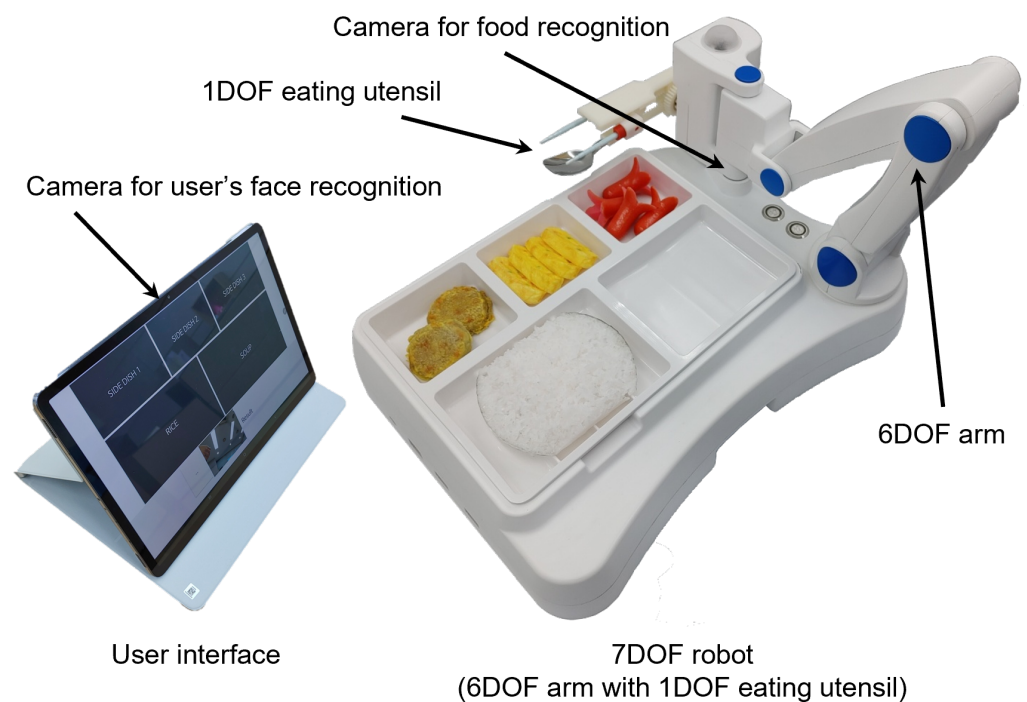


Figure 1. The proposed meal-assistance robot system.

The various types of MARs that have been presented to date are often operated in a limited laboratory environment or are non-portable. In addition, even commercialized products cannot provide high-quality services due to their functionally low performance and high prices. Furthermore, they do not consider various types of user disabilities and provide only limited user interfaces. These previous studies and their limitations will be dealt with in more detail in the next section, and our proposed MAR system has several points of difference. The major contributions of this article are summarized as follows:

- Development of a compact and portable meal assistance robot system capable of both food recognition and robot control without external computing.

- Development of a meal assistance robot system that can serve realistic Korean diets, which can easily be used in a cafeteria rather than a simple lab environment.
- Development of deep-learning-based food acquisition point estimators to utilize chopstick-type dining tools.
- Development of a non-contact user interface based on face (gaze/nod/mouth shape) recognition for critically ill patients who have difficulty using a touch interface.

The remaining parts of this paper are organized as follows. Section 2 introduces various recent meal assistance robots and related research trends. Next, the proposed MAR system is illustrated in Section 3. Section 4 presents the experimental results using the developed system, and finally, the conclusion is presented in Section 5.

2. Related Work

2.1. Commercial Products

Several MARs have already been commercialized, mainly in the US and Western countries. The most commercially successful product is Obi [8], a 6-DOF robotic arm developed in the United States. This product uses a bowl divided into four compartments as a tray, and whenever the user presses a button interface, the bowl rotates so that the arm can take another food. Once a food acquisition command is given, the robot arm scoops the food with a spoon along predefined paths. Afterward, it brings the food to the user's mouth, the location of which is taught by the caregiver. However, despite these well-made products, only limited types of food can be served. Another US product, Meal Buddy [9], serves food using a 3-DOF robot arm with three independent bowls and spoons. Mealtime Partner [10] is also a US product, consisting of three bowls on the rotational plate and 2-DOF eating utensils. This product is placed right in front of the user's mouth and provides food to the user through a very simple movement of the spoon. Neater Eater [11], developed in the UK, consists of a rotating dish and a 3-DOF robotic arm and acquires food with a spoon. This product is characterized by a tablet interface that allows for the easy setting of various detailed robot functions. Bestic [12], from Sweden, consists of one rotational bowl and a 4-DOF robotic arm. It assists with simple meals using a spoon based on preset parameters.

MARs developed in the Western world, mainly in the US, do not consider how to handle food such as rice or soup placed on a tray with separate compartments because it was developed to suit the food culture of putting food on a plate and eating it with a spoon or fork. On the other hand, the cases developed in Asia are slightly different. My Spoon [13], developed in Japan, consists of a rectangular tray divided into four compartments and a 5-DOF robot arm. It can handle various types of food using a fork and a spoon simultaneously as tongs. CareMeal [14,15], from Korea, includes a tray with multiple compartments and two robotic arms. Each arm picks up food with tongs and serves it to the user's mouth with a spoon. These two products were developed to fit the Asian diet and have one thing in common: they use tools that can act as tongs or chopsticks to handle rice or vegetable side dishes. Although no device has commercialized at present, several studies have been aimed to create meal assistance robots that use chopsticks [16–19]. Using chopsticks as tableware is not only more emotionally satisfying in Asian communities but also has the advantage of better handling Asian food.

The commercialized MARs introduced above have common limitations. First, because they move along predefined paths, they cannot actively acquire food by recognizing the location of the actual food. In addition, since most are focused on foods that are easy to consume with a spoon, it is difficult to serve a variety of food. Finally, users that struggle with physical manipulation may have difficulty using buttons or joystick-type user interfaces.

2.2. Food Perception

Food perception technology is important in the construction of autonomous MARs because food area recognition and picking points are vital in the acquisition of food. Os-

hima et al. conducted a study using a laser range finder [20], but most other studies used a vision-based recognition method. With the recent development of deep learning technology, visual recognition technology has been significantly improved. Researchers suggested the use of ResNet [21], DenseNet [22], and the Feature Pyramid Network (FPN) [23] for object classification. Several networks, including Yolo [24], SSD [25], RetinaNet [26], and DetNet [27], have been presented for object detection. Moreover, SegNet [28], two-way DenseNet [29], Mask R-CNN [30], and BlitzNet [31] have been proposed for image segmentation. These networks have been utilized in various food recognition applications [32–34]. The deterministic annealing neural network approach [35,36] is also being actively studied. It has the potential to be applied to food recognition for robotic meal assistant tasks.

In the case of an MAR system that recognizes food through artificial intelligence technology, a series of research conducted at the University of Washington can be regarded as state-of-the-art. In their early research [37], they implemented a two-stage food detector using RetinaNet, which is more accurate than single-shot algorithms such as SSD or YOLO and faster than other two-stage detectors such as Faster RCNN [38]. In addition, they proposed SPANet, which estimates the success rates for six acquisition actions by receiving detection results from RetinaNet and the results for classifying the environmental conditions of each object into three types [39]. The following study presented a method for obtaining unseen food items by receiving success/failure feedback as a reward after attempting to obtain food with SPANet [40]. In addition, an improved method utilizing additional feedback on haptic context information was presented [41].

However, these studies have the limitation that they can only be applied to foods comprising pieces without sauce or water. In other words, they are unsuitable for some Asian cuisine, such as kimchi and seasoned vegetables, which are easy to pick up with chopsticks as many objects are intertwined. In addition, since a fork is used as an eating utensil, they struggle to satisfy the sentiment of Asians who prefer chopsticks to forks.

2.3. User Interface

Various user interfaces for robot operation have already been developed and are still being developed [42,43]. More specifically, voice [44,45] interfaces and augmented/virtual reality (AR/VR) [46,47] interfaces have been attempted in various fields. Brain-computer interfaces (BCIs) [48] using an electroencephalogram (EEG) [49,50] or an electromyogram (EMG) [51,52] are also being studied. However, the user interface types are inflexible if the scope is limited to commercialized MARs. Most adopt button or joystick interfaces because they are easy and intuitive [8–10,12–14]. However, this may limit the range of available patients, since these interfaces require physical contact from the user. Nonetheless, the development a fully automated MAR that rarely requires user commands is also not a perfect solution; users prefer things that can be operated by themselves, rather than autonomous robots [53].

This study adopted a face-recognition-based interface that does not demand physical contact, so that even severely handicapped users with spinal cord injuries can operate it alone. Some more closely related works are introduced as follows. Perera et al. tracked the mouth's position by recognizing the mouth's shape [50], and other studies recognized mouth-opening [54] or eye-closing [55] as a user command. Liu et al. developed an interface that uses deep learning to recognize the opening/closing of the mouth and eyes [56]. Some studies also utilized facial gesture [57] or head movement [58] recognition in the interface. Lopes et al. implemented an interface that selects the desired food by tracking the user's eye gaze [59].

Most of the existing gaze recognition research has the drawback of requiring a calibration process and expensive sensors to accurately detect the user's gaze. Meanwhile, in a study using a universal webcam, it was verified that real-time recognition is possible by extracting a low-resolution (320×240) eye image from a user's facial video and recognizing the gaze point in 200 ms for each image [60]. However, in a study that recognized the user's gaze based on the user's geometric characteristics in the mobile environment, only

the approximate gaze area was predicted, despite a calibration process longer than one minute [61,62]. Studies were conducted to recognize the approximate direction of the gaze, not the exact gazing point, using a CNN structure and an RGB camera without a separate correction process. The proposed method was confirmed to be feasible by classifying the gaze direction of the user's pupil, estimated by two CNNs with 97% accuracy for the three classes of 'center,' 'left,' and 'right' [63,64].

3. Materials and Methods

This section describes the MAR system and its components in detail.

3.1. MAR System Overview

The proposed system includes a 7-degrees-of-freedom (DOF) robot arm with spoon–chopstick utensil [65], a food tray, and vision SWs for food recognition and user input. These components can be divided into three groups, as shown in Figure 2. The first group is robot HW, such as the robot arm, its controller, and food tray. The second is the acquisition point estimator, and a tablet user interface based on the user's face (eye gaze/nod/mouth shape) recognition is the last. Each group is introduced in turn in the following subsections. The operation procedure of the whole system is as follows (see Figure 3):

1. The caregiver holds the robot arm and directly teaches the position that the user's mouth can easily reach.
2. The caregiver presses the start button after setting a meal on the plate. (After this, the user starts to eat alone using the user interface.)
3. The robot moves to a predefined ready position and takes an image of the food tray using a camera attached to the robot arm.
4. From the image, the acquisition point estimator perceives the food area and extracts picking points for each dish through post-processing. Then, it sends the robot controller the points.
5. Once a user command selecting a particular food is transmitted from the user interface to the controller, the robotic arm moves to the coordinates of the acquisition point for that food.
6. The robot performs a food-acquisition motion and then delivers food to the mouth position, as taught directly by the caregiver in step 1.
7. The robot waits for the user to eat food for a preset amount of time.
8. Steps 3–7 are repeated until the end of the meal.

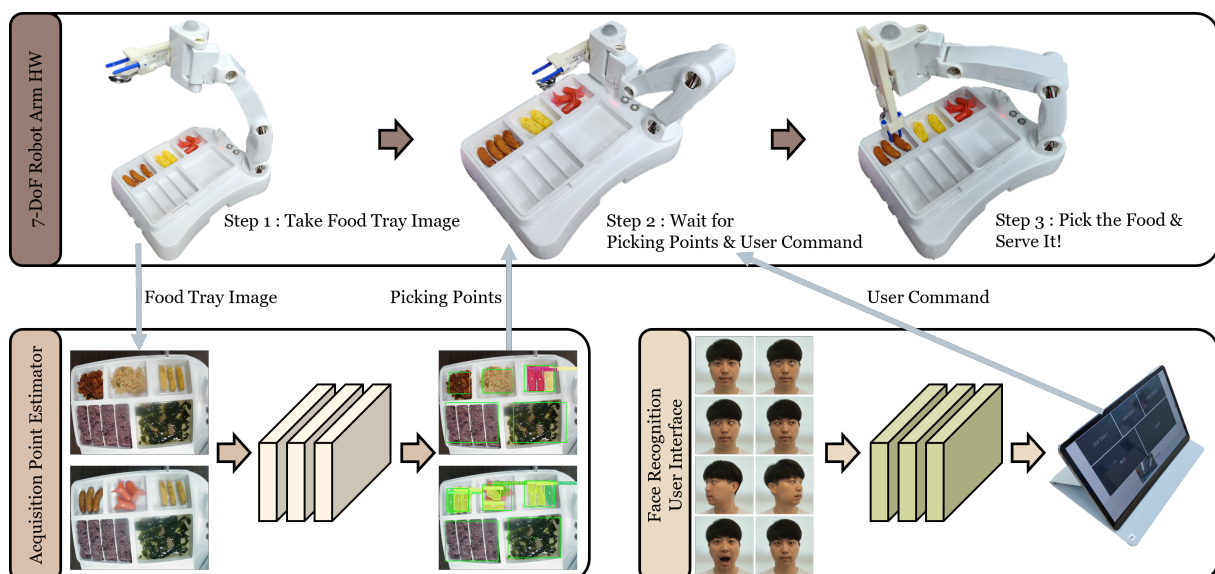


Figure 2. The overall flowchart of the meal-assistance process proposed in this work.

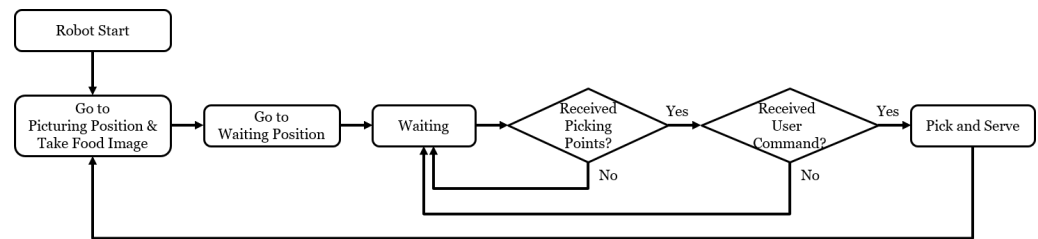


Figure 3. The simple flowchart for the whole meal-assistance task.

3.2. Robot HW and Control

The 7-DOF manipulator is designed to perform food acquisition motions at various positions on the food tray. It was fabricated using the modules of seven ROBOTIS Dynamixel. The manufacture and installation of a controller and an encoder board capable of EtherCAT communication secured the position sensing precision and control cycle time necessary to operate the control algorithm [66,67]. The last of the seven modules was used to drive the eating utensil part, which consists of spoons and chopsticks, which are familiar to Asian users. The dining tool part has a structure that can switch between the two tools [65]. A camera is mounted near the tool part to obtain the plate image required for food recognition. On the upper part of the robot base, where the manipulator workspace exists, a food tray (300 mm × 210 mm) is placed containing a meal consisting of rice, soup, and three side dishes commonly seen in cafeterias in Korea. The motion controller, mini PC (Latte Panda Alpha), and 24 V battery are located at the lower part of the robot's base to drive the robot HW. The total weight of the MAR is around 5 kg for easy installation and portability.

The performance requirements for the robot HW are as follows: First, it must have a payload of 0.3 kg for the stable handling of eating utensils and acquired food. Second, it must have a maximum speed of 250 mm/s at the tip for the user's safety and to deliver the acquired food within 2 s. According to KS B ISO 9283, the manufactured HW was subjected to a performance evaluation under a load weight of 0.315 kg and rated speed of 250 mm/s, and the results are as follows [67]:

- Position repeat accuracy: less than 30 mm;
- Distance accuracy: 4.39 mm;
- Distance repeatability: 0.26 mm;
- Path speed accuracy: 4.81% (@target speed 250 mm/s).

The SW structure for the meal assistance robot is shown in Figure 4. The motion controller acts as the main controller, which directly drives the robot. This was implemented in the Xenomai environment, where real-time computing is possible, with an ARM Cortex-A9-based embedded system. EtherCAT was used to control the joints of the manipulator within a control period of 1 ms. To ensure stable acquisition, accurate delivery, and safety in case of external contact, the control algorithm was designed to adjust the impedance in the workspace, as shown in Figure 5. Latte Panda, a sub-controller that connects to the main controller and tablet interface via TCP/IP, performs additional tasks, except directly driving the robot. In the ROS environment, it manages the acquisition point estimation task and the overall signal flow of RAS: it transmits the coordinates of the acquisition points and user commands to the main controller. The non-contact user interface on the Android tablet sends the user commands to the sub-controller according to the face recognition results.

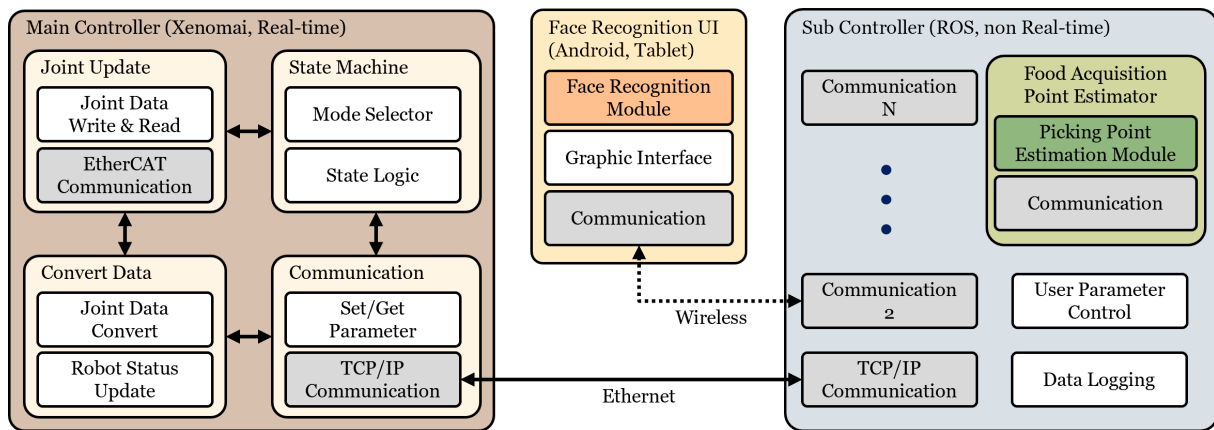


Figure 4. Control system diagram.

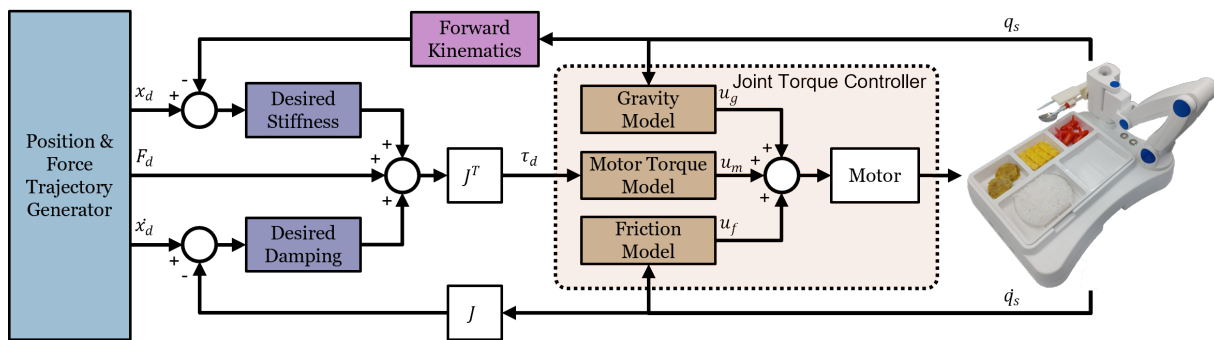


Figure 5. The control schematic for the MAR's motion control.

3.3. Food Acquisition Point Estimator

Instance segmentation in computer vision for an image is the pixel-by-pixel representation in which instances belonging to a specific class are located in a given image. Unlike semantic segmentation, instance segmentation considers instances of the same class to be different. As a deep learning model for instance segmentation, Mask R-CNN, which adds a process in which the mask of an object is predicted for the Faster R-CNN used for object detection, is widely used. Mask R-CNN, a network that combines various architectures, consists of a convolutional backbone that extracts features from images and a network head that estimates bounding boxes, classifies classes and predicts masks. This network head can be divided into a classification branch that predicts the classes of objects for each localized Region of Interest (RoI), a bbox regression branch for bounding box regression, and a mask branch that predicts segmentation masks in parallel.

In this paper, we implemented a food acquisition point estimator that detects the areas of side dishes and extracts the acquisition points of each side dish by utilizing Mask R-CNN-based instance segmentation [30]. It recognizes side dishes, excluding rice and soup, and distinguishes between lump food and food made up of pieces. Then, it uses different strategies for different types of food to calculate adequate acquisition points. Figure 6 shows the whole process.

Figure 7 shows the instance segmentation process in our food acquisition point estimator. We used ResNet50-FPN pre-trained with the ImageNet dataset [68] as the backbone. A Fully Convolutional Network (FCN) was applied to the branch at which mask segmentation takes place, considering both the performance and speed of the model for food made up of pieces. The feature map extracted from the input image through the backbone network is sent to the Region Proposal Network (RPN), and food instances are detected through RoIAlign [30]. The feature map obtained from this is passed back to the network head and used as input for class classification and bounding box creation. It goes through the FCN, consisting of several convolutional layers, and generates an instance mask for each class.

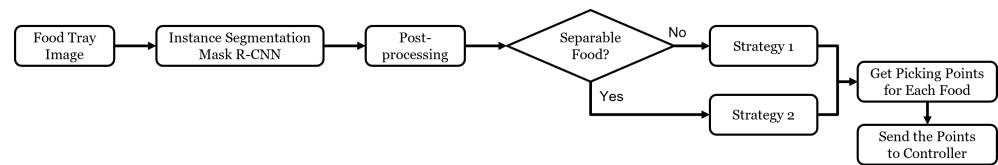


Figure 6. The flowchart for food acquisition point estimator.

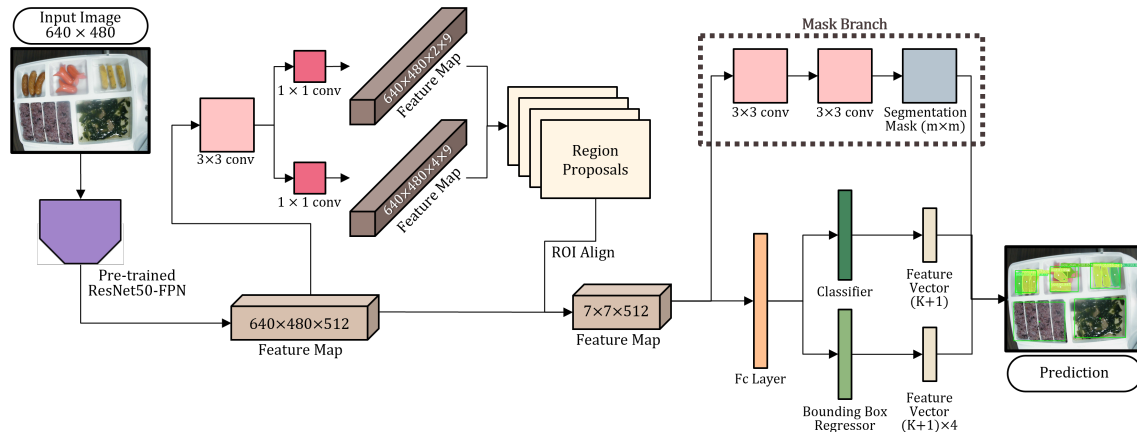


Figure 7. Mask R-CNN instance segmentation pipeline for the proposed MAR system.

The segmentation results of the 640×480 RGB image by Mask R-CNN undergo post-processing to extract acquisition points. First, to prevent the calculation acquisition points by recognizing some objects outside the plate, the estimator algorithm deletes instance masks created outside the pre-specified food tray area. Next, acquisition points are obtained through different strategies according to the types of side dishes that are recognized. Specifically, in the case of common side dishes in the form of a lump, the acquisition points are generated at the upper right corner of the instance mask considering the width of the tool. In the case of food made up of pieces, the center of the detected region becomes an acquisition point. Among these points, the point obtained from the instance with the highest score is selected as an acquisition point. In the end, only one point is extracted for each side dish. Points obtained on the image coordinate system are not converted to robot base frame coordinates; instead, they are calculated as relative coordinates in pixels based on the center of each cell of the food tray, as shown in Figure 8. Therefore, it is necessary to calibrate the position at which the centerpoint of each cell appears on the image coordinate system and the scale difference between the two coordinate systems. That is, errors that may occur during coordinate conversion between two frames due to lens distortion can be minimized by receiving relative coordinates for each cell center point rather than the robot reference point. The robot performs a meal assistance task by obtaining the acquisition points necessary for motion planning before each acquisition motion using the above process.

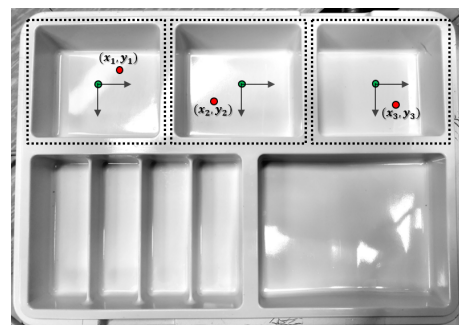


Figure 8. Relative coordinates for each cell centerpoint.

3.4. Face Recognition User Interface

Most MASs adopt a user interface that requires physical input, such as a button or a joystick. In this study, however, we propose a face-recognition-based HRI interface that operates on mobile devices and does not require physical contact for critically ill patients with poor upper-extremity manipulation. The proposed interface utilizes a deep learning model that estimates facial movement from the user's facial image input; using the estimated results, it recognizes the user's intention and transmits the corresponding command to the robots (see Figure 9). First, in order to implement this, the user's facial movements are defined into eight classes:

- Four types of eye-gaze directions ("left up," "up," "right up," "center");
- Two types of head directions ("head_left," "head_right");
- Two types of mouth shapes ("open," "close").

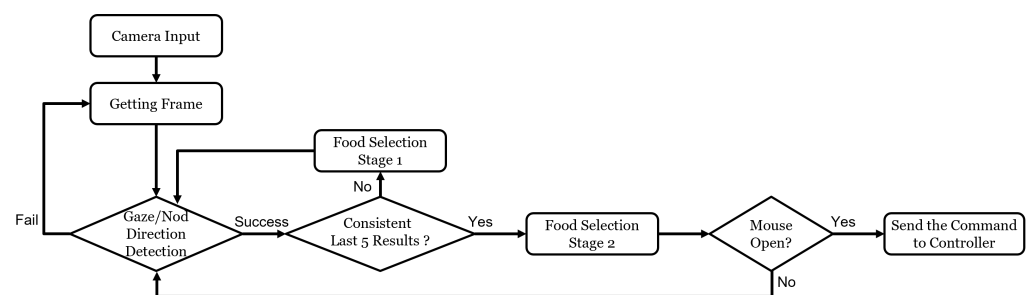


Figure 9. The flowchart for the user interface based on face recognition. The real-time recognition result from the user's face motion input is highlighted in blue on the tablet interface screen at food selection stage 1. If the same results are obtained more than five times consecutively, it goes to food selection stage 2 and changes to a red highlight. When the food displayed in red matches what the user wants, the user can confirm the selection using an "open" command. In that case, the color changes to green, and an acquisition command is sent to the robot to serve the food.

This improves on our previous study [69] by adding two head-direction and two mouth-shape classes, instead of similar-gaze classes, which have a high probability of confusion.

Next, a deep learning model for classifying these eight classes was designed, based on the use of a one-stage detector to operate in real-time and on mobile AP environments. As shown in Figure 10, a CNN model consisting of nine convolutional layers was constructed based on YOLOv4-tiny [70], a lightweight, one-step detector structure. Facial images were collected from the front camera of android tablets (Samsung Galaxy Tab S6, S7+). Then, a benchmark dataset, Eye-Chimera [71,72], made up of males/females in their 20s and 50s, was added: a dataset consisting of four eye-gaze classes, two head-direction classes, and two mouth classes was established.

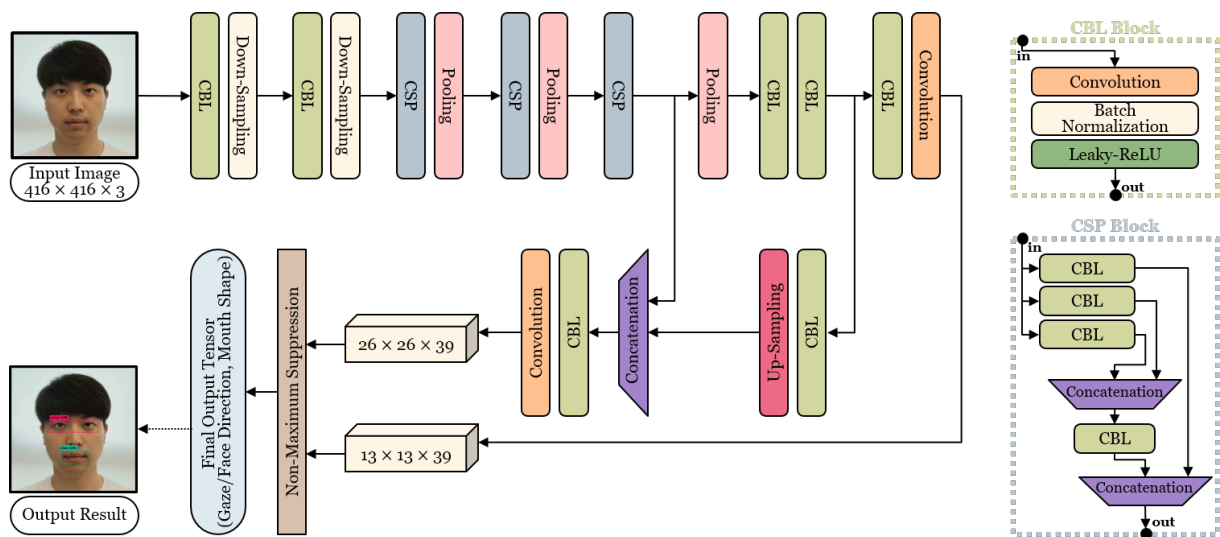


Figure 10. A CNN model architecture for face recognition.

Meanwhile, several processes were added to reduce the probability of misrecognition [69]. The eye-gaze “center” is not directly used for food selection but was added as a neutral command. This is because eye-gaze classes are prone to being misrecognized due to the difference in the position of the front cameras between the Galaxy Tab S6 and S7. Since the training data for mouth classes were images of the mouth opening and closing while gazing at the center of the tablet, most were misidentified as the “center” class with model overfitting. To solve this problem, we added a preprocessing process that recognizes the face area by dividing it into two parts: the upper face and lower face. A small change in lighting can also significantly influence recognition performance. Thus, four types of lighting filters and artificial brightness changes were applied to the dataset to overcome this phenomenon.

The trained model based on YOLOv4-tiny, TensorFlow Lite [73], Java, and Kotlin languages were used to implement a face recognition interface that operates in real-time on the Android system. The operation flow of the interface is shown in Figure 9. The results recognizing the user’s facial images from the tablet’s front camera were mapped to three side dishes, rice, and soup, respectively, as shown in Figure 11. Whenever the real-time recognition result is provided, the food mapped to this is highlighted in blue on the tablet interface. The food turns red if the same results are repeated five or more times in a row. If another recognition result is derived at this time, the red highlight disappears, and the most recent recognition result is displayed in blue again. After that, if the user wants to acquire the food currently highlighted in red, he can give the “open” command by opening his mouth. In this case, when “open” is properly recognized, the red food turns green, and a command to acquire the food is transmitted to the robot (see Figure 12).

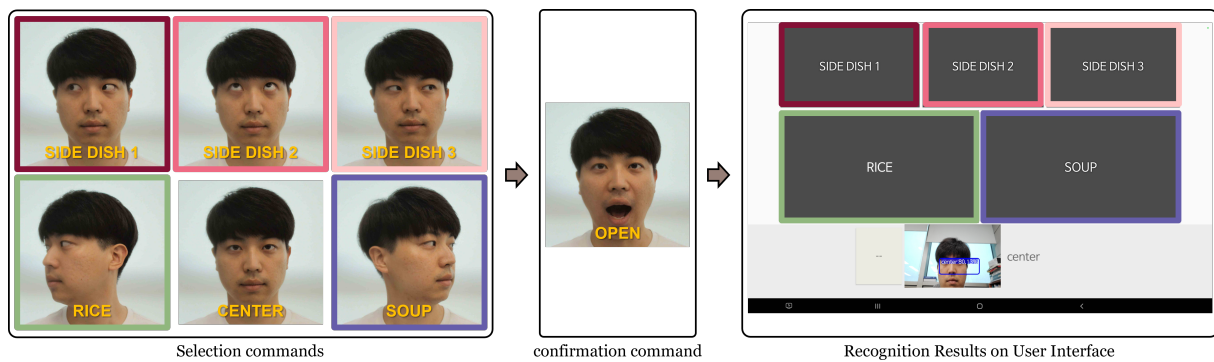


Figure 11. Mapping between facial recognition results and food selection. The color-boxed user facial movements on the left are mapped to the same color-boxed food in the tablet interface on the right. The user can select the food they want by using their facial movements as input commands. Once the desired food is selected, the user can send an acquisition command to the robot using the “open” command. In the case of ‘center’ without a colored box, this is not mapped with a specific dish and is used as a neutral command to reduce the misselection of side dishes.

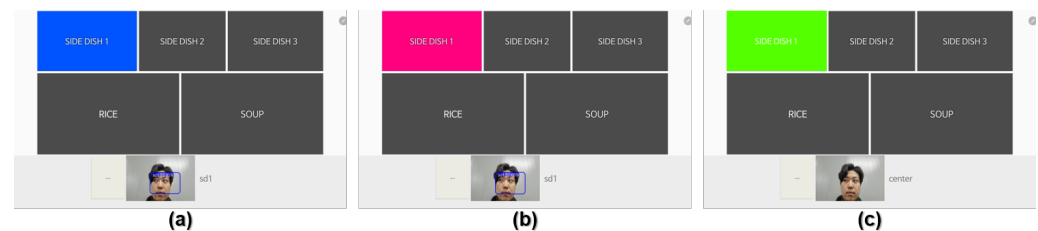


Figure 12. Tablet UI screens showing each step of face recognition. (a) A screen of food selection stage 1. This shows the real-time recognition result before entering food selection stage 2. The detected food is highlighted in blue. (b) Upon entering food selection stage 2, the result turns red. (c) The state in which the choice in food selection step 2 is finally confirmed. The food is highlighted in green, and the corresponding command is sent to the robot controller.

4. Results

4.1. Experiment 1—Acquisition Point Estimator

The training data for the Mask R-CNN-based model, used to detect food areas, were collected by taking pictures of the food tray with a camera mounted on the robot arm end-effector of the MAR system. Each food area in the acquired images was masked. In particular, some side dishes composed of pieces such as radish kimchi and sausage were masked so that each piece was recognized as a separate area. The trained model was converted to ONNX format and applied to the MAR system, which was used for the acquisition point estimator, as shown in Figure 6.

The performance evaluation of the overall acquisition point estimator was performed in the developed MAR system. The estimator SW was run on Latte Panda, a sub-controller located under the tray of MAR, in Ubuntu 18.04 and ROS Melodic environment. The model operated through CPU calculation, utilizing the ONNX runtime. Side dishes not used for model training were prepared on the MAR plate, and the acquisition point estimation process was tested 100 times, changing the type and amount of food each time (see Figure 13).

For food made up of pieces, success depends on whether the final selected acquisition point is in the area of that object. If all three acquisition points individually estimated three side dishes located in each food image area, this was classified as a success. In other words, the recognition was considered a failure if even one of the three food acquisition points was extracted outside the food area. As a result, 99 out of 100 times were successful, and a 99% success rate was obtained.

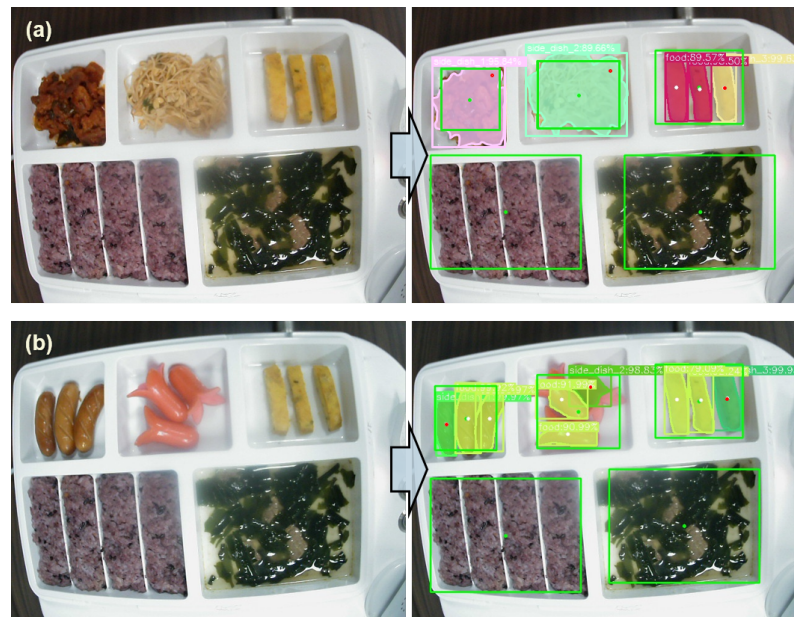


Figure 13. Food recognition results in the MAR system. (a,b) The pictures on the left are food images taken by the robot, which are input to the acquisition point estimator. The figures on the right are the result of the acquisition point estimator. The red dots in the right pictures indicate the points transmitted to the robot controller as acquisition points for each side dish.

4.2. Experiment 2—Face Recognition User Interface

The deep learning model used for the face recognition user interface was trained using a combination of the Eye-ChIMERA dataset and our own dataset. As a result of the five-fold cross-validation of the model, an average F1 score of 0.95 and an average accuracy of 0.95 were achieved. The trained model was utilized as one of the SW components of that recognizes the user's intention through the process shown in Figure 9.

A performance evaluation of the entire face recognition user interface was also conducted. Five people, two glasses-wearers and three non-glasses-wearers, participated in the experiment using a Samsung Galaxy tab S7+ with Android 12. The evaluation process was as follows. First, the experiment instructor commanded the participant to select a specific food, and the participant made eye or head movements to select the food. Recognition status appeared on the user interface running on the tablet, as shown in Figure 12. Real-time recognition results were highlighted in blue, and when the same result occurred consecutively, the result turns red. The case highlighted in red is the food selection stage 2, not the final selection result. Even if unintended food is selected in that stage, the user can continue to select the designated food. If the instructed food is correctly selected in red, the participant may make a final decision about the food using mouth motions. When the mouth command was successfully recognized, the red turned into a green highlight, and the food was finally selected. At this time, success or failure was judged according to whether the green matched the food ordered by the experiment instructor. However, if the food was not highlighted in green within 10 s after the instructor's direction, this was considered a failure.

A total of 200 recognition tests were conducted, with each person trying to select five types of food eight times, and a 100% success rate was obtained, with all 200 trials ending in success. However, there were six cases of misrecognition of foods marked in red during the food selection stage, stage 2. If the success condition is more strictly defined, considering the misrecognitions in selection stage 2 as a failure, the success rate would be calculated as 97.0%.

4.3. Experiment 3—Whole MAR System Demonstration

As shown in Figure 14, a real meal was served on a tray, and the user ate by himself with the assistance of the MAR system, without using his own arms. The acquisition points of each side dish were estimated using the food image taken from the camera attached to the robot arm, and then the coordinates were sent to the robot controller. Next, whenever the user selected food using gaze/head/mouth commands, the robot moves to the acquisition point of the chosen food, obtains the food, and serves it to the designated user's mouth position.

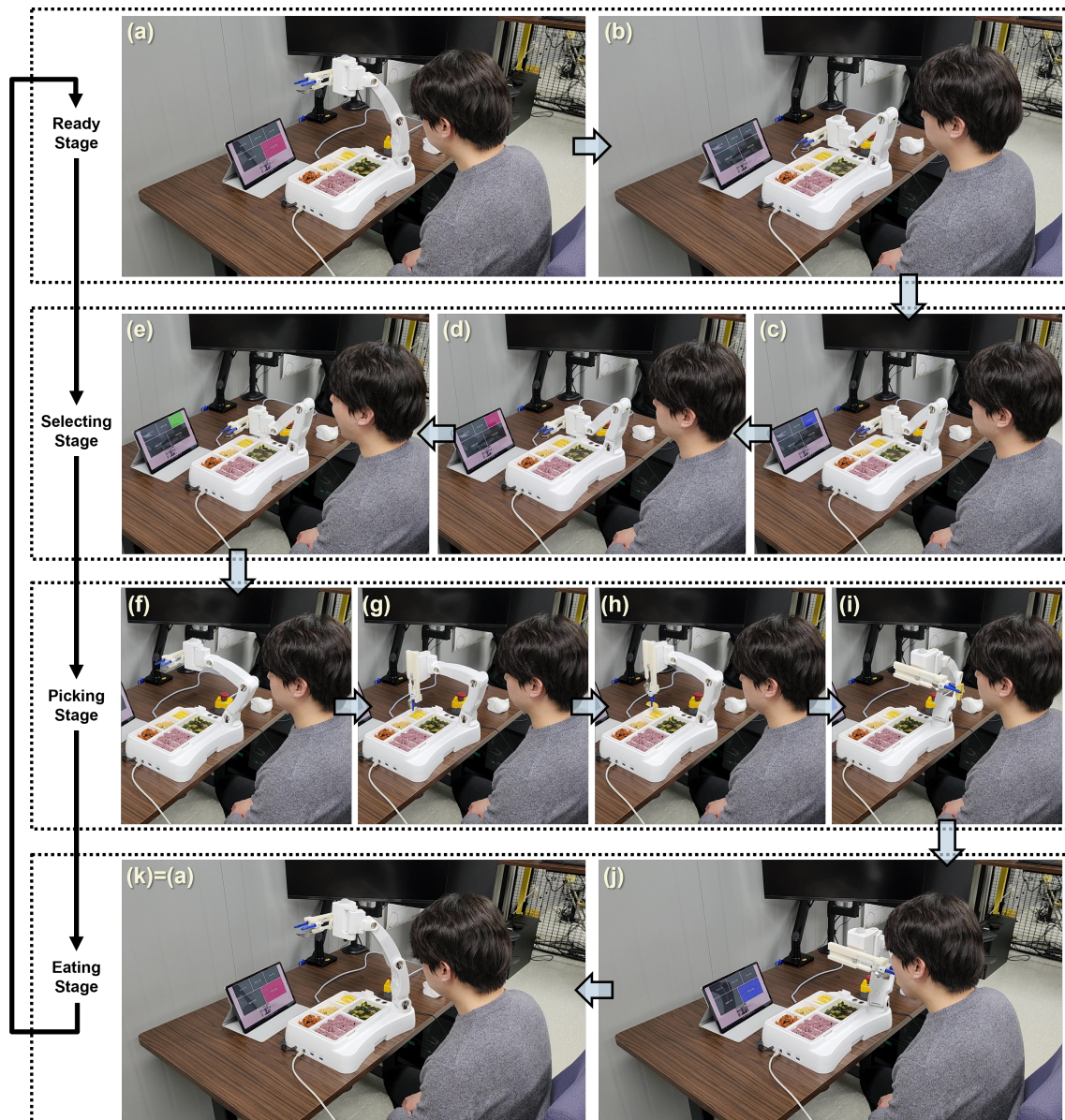


Figure 14. Demonstration of the actual use of the proposed MAR system. (a) The robot arm returns to the ready position and takes an image of the food tray. (b) After taking the image, the robot moves to the standby position, transfers acquisition points for each side dish to the robot controller and waits for the user's command. (c) Food selection stage 1. (d) Food selection stage 2. (e) Once the user finalizes the result of food selection step 2, a corresponding signal is sent to the robot controller. (f–i) The robot arm moves to the acquisition point of the food received from the UI, acquires the food and delivers it in front of the user's mouth. (j) The user eats the food brought by the robot. (k) The robot arm returns to the same posture as in (a), takes an image of the tray again, and repeats the above process.

5. Conclusions

In this work, we proposed a new MAR system that could overcome the limitations of previously studied or commercialized MAR systems. This was developed so that both spoons and chopsticks can be used to accommodate the Asian diet, which typically includes soups and side dishes such as clumped greens. Unlike commercial products, which require dedicated rotating trays to obtain various dishes while moving in one fixed position, this system uses a common tray that is easily found in Korean cafeterias. The MAR system includes an acquisition point estimator that derives acquisition points from a tray image and a user interface based on facial recognition, considering the use of the severely disabled. In addition, with a portable weight of less than 5 kg, it can perform meal assistance tasks by computation alone.

Through a test conducted by setting a menu that is easily seen in Korean hospitals, a 99% success rate was obtained by the acquisition point estimator. Side dishes composed of pieces, as well as the acquisition points for lump side dishes, were well extracted. The accuracy of the deep learning model when used alone for facial recognition was only about 97%. However, the entire user interface experiment, programmed to select user intent through additional algorithms, recorded a 100% success rate. Further, we confirmed the actual usability of the system by demonstrating the eating of a meal using the entire proposed MAR system.

Future research should be devoted to advancing food acquisition strategies using this system. It is necessary to measure the actual acquisition success rate of food using the acquisition point estimator and to increase this. It is expected that this will be applied to a realistic environment such as a hospital and will greatly help the elderly and the disabled to eat on their own. In this way, it will not only reduce the caregiver workload but also improve the users' quality of life by instilling independence.

Author Contributions: Conceptualization, I.C., K.K., B.-J.J., J.-H.H., G.-H.Y.; methodology, I.C., K.K., G.-H.Y.; software, I.C., K.K., H.S., B.-J.J.; investigation, I.C.; resources, H.S., B.-J.J.; data curation, I.C., H.S., B.-J.J., J.-H.H.; writing—original draft preparation, I.C., H.S., B.-J.J.; writing—review and editing, K.K., J.-H.H., G.-H.Y., H.M.; visualization, I.C. and B.-J.J.; supervision, H.M. and G.-H.Y.; project administration, G.-H.Y.; funding acquisition, G.-H.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Industrial Strategic Technology Development Program (No. 20005096, Development of Intelligent Meal Assistant Robot with Easy Installation for the elderly and disabled) funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea). Also, it was financially supported by the Korea Institute of Industrial Technology (KITECH) through the In-House Research Program (Development of Core Technologies for a Working Partner Robot in the Manufacturing Field, Grant Number: EO210004).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank our research consortium including CyMechs Co., Korea Electronics Technology Institute, Chonnam National University, Gwangju Institute of Science and Technology, and Seoul National University Bundang Hospital, for their assistance for this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. World Health Organization. Available online: <https://www.who.int/news-room/fact-sheets/detail/disability-and-health/> (accessed on 8 December 2022).
2. Naotunna, I.; Perera, C.J.; Sandaruwan, C.; Gopura, R.; Lalitharatne, T.D. Meal assistance robots: A review on current status, challenges and future directions. In Proceedings of the 2015 IEEE/SICE International Symposium on System Integration (SII), Nagoya, Japan, 11–13 December 2015; IEEE: New York, NY, USA, 2015; pp. 211–216.

3. Bhattacharjee, T.; Lee, G.; Song, H.; Srinivasa, S.S. Towards robotic feeding: Role of haptics in fork-based food manipulation. *IEEE Robot. Autom. Lett.* **2019**, *4*, 1485–1492. [CrossRef]
4. Prior, S.D. An electric wheelchair mounted robotic arm—A survey of potential users. *J. Med. Eng. Technol.* **1990**, *14*, 143–154. [CrossRef] [PubMed]
5. Stanger, C.A.; Anglin, C.; Harwin, W.S.; Romilly, D.P. Devices for assisting manipulation: A summary of user task priorities. *IEEE Trans. Rehabil. Eng.* **1994**, *2*, 256–265. [CrossRef]
6. Pico, N.; Jung, H.R.; Medrano, J.; Abayebas, M.; Kim, D.Y. Climbing control of autonomous mobile robot with estimation of wheel slip and wheel-ground contact angle. *J. Mech. Sci. Technol.* **2022**, *36*, 1–10. [CrossRef]
7. Pico, N.; Park, S.H.; Yi, J.S.; Moon, H. Six-Wheel Robot Design Methodology and Emergency Control to Prevent the Robot from Falling down the Stairs. *Appl. Sci.* **2022**, *12*, 4403. [CrossRef]
8. Obi. Available online: <https://meetobi.com/meet-obi/> (accessed on 8 December 2022).
9. Meal Buddy. Available online: <https://www.performancehealth.com/meal-buddy-systems> (accessed on 8 December 2022).
10. Specializing in Assistive Eating and Assistive Drinking Equipment for Individuals with Disabilities. Available online: <https://mealtimepartners.com/> (accessed on 8 December 2022).
11. Neater. Available online: <https://www.neater.co.uk/> (accessed on 8 December 2022).
12. Bestic Eating Assistive Device. Available online: <https://at-aust.org/items/13566> (accessed on 8 December 2022).
13. Automation My Spoon through Image Processing. Available online: <https://www.secom.co.jp/isl/e2/research/mw/report04/> (accessed on 8 December 2022).
14. CareMeal. Available online: <http://www.ntrobot.net/myboard/product> (accessed on 8 December 2022).
15. Song, W.K.; Kim, J. Novel assistive robot for self-feeding. In *Robotic Systems-Applications, Control and Programming*; Intech Open: London, UK, 2012; pp. 43–60.
16. Song, K.; Cha, Y. Chopstick Robot Driven by X-shaped Soft Actuator. *Actuators* **2020**, *9*, 32. [CrossRef]
17. Oka, T.; Solis, J.; Lindborg, A.L.; Matsuura, D.; Sugahara, Y.; Takeda, Y. Kineto-Elasto-Static design of underactuated chopstick-type gripper mechanism for meal-assistance robot. *Robotics* **2020**, *9*, 50. [CrossRef]
18. Koshizaki, T.; Masuda, R. Control of a meal assistance robot capable of using chopsticks. In Proceedings of the ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics), Munich, Germany, 7–9 June 2010; VDE: Berlin, Germany, 2010; pp. 1–6.
19. Yamazaki, A.; Masuda, R. Autonomous foods handling by chopsticks for meal assistant robot. In Proceedings of the ROBOTIK 2012, 7th German Conference on Robotics, Munich, Germany, 21–22 May 2012; VDE: Berlin, Germany, 2012; pp. 1–6.
20. Ohshima, Y.; Kobayashi, Y.; Kaneko, T.; Yamashita, A.; Asama, H. Meal support system with spoon using laser range finder and manipulator. In Proceedings of the 2013 IEEE Workshop on Robot Vision (WORV), Clearwater Beach, FL, USA, 15–17 January 2013; IEEE: New York, NY, USA, 2013; pp. 82–87.
21. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
22. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
23. Lin, T.Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature pyramid networks for object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 2117–2125.
24. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 779–788.
25. Liu, W.; Anguelov, D.; Erhan, D.; Szegedy, C.; Reed, S.; Fu, C.Y.; Berg, A.C. Ssd: Single shot multibox detector. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 8–16 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 21–37.
26. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
27. Li, Z.; Peng, C.; Yu, G.; Zhang, X.; Deng, Y.; Sun, J. Detnet: A backbone network for object detection. *arXiv* **2018**, arXiv:1804.06215.
28. Badrinarayanan, V.; Kendall, A.; Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef] [PubMed]
29. Jégou, S.; Drozdal, M.; Vazquez, D.; Romero, A.; Bengio, Y. The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 11–19.
30. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2961–2969.
31. Dvornik, N.; Shmelkov, K.; Mairal, J.; Schmid, C. Blitznet: A real-time deep network for scene understanding. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 4154–4162.
32. Yanai, K.; Kawano, Y. Food image recognition using deep convolutional network with pre-training and fine-tuning. In Proceedings of the 2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Turin, Italy, 18–22 July 2022; IEEE: New York, NY, USA, 2015; pp. 1–6.

33. Singla, A.; Yuan, L.; Ebrahimi, T. Food/non-food image classification and food categorization using pre-trained googlenet model. In Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management, Amsterdam, The Netherlands, 16 October 2016; pp. 3–11.
34. Hassannejad, H.; Matrella, G.; Ciampolini, P.; De Munari, I.; Mordonini, M.; Cagnoni, S. Food image recognition using very deep convolutional networks. In Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management, Amsterdam, The Netherlands, 16 October 2016; pp. 41–49.
35. Wu, Z.; Gao, Q.; Jiang, B.; Karimi, H.R. Solving the production transportation problem via a deterministic annealing neural network method. *Appl. Math. Comput.* **2021**, *411*, 126518. [\[CrossRef\]](#)
36. Wu, Z.; Karimi, H.R.; Dang, C. An approximation algorithm for graph partitioning via deterministic annealing neural network. *Neural Netw.* **2019**, *117*, 191–200. [\[CrossRef\]](#) [\[PubMed\]](#)
37. Gallenberger, D.; Bhattacharjee, T.; Kim, Y.; Srinivasa, S.S. Transfer depends on acquisition: Analyzing manipulation strategies for robotic feeding. In Proceedings of the 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Daegu, Korea, 11–14 March 2019; IEEE: New York, NY, USA, 2019; pp. 267–276.
38. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Adv. Neural Inf. Process. Syst.* **2015**, *28*. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Feng, R.; Kim, Y.; Lee, G.; Gordon, E.K.; Schmitt, M.; Kumar, S.; Bhattacharjee, T.; Srinivasa, S.S. Robot-Assisted Feeding: Generalizing Skewing Strategies Across Food Items on a Plate. In Proceedings of the International Symposium of Robotics Research, Hanoi, Vietnam, 6–10 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 427–442.
40. Gordon, E.K.; Meng, X.; Bhattacharjee, T.; Barnes, M.; Srinivasa, S.S. Adaptive robot-assisted feeding: An online learning framework for acquiring previously unseen food items. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October–24 January 2020; IEEE: New York, NY, USA, 2020; pp. 9659–9666.
41. Gordon, E.K.; Roychowdhury, S.; Bhattacharjee, T.; Jamieson, K.; Srinivasa, S.S. Leveraging Post Hoc Context for Faster Learning in Bandit Settings with Applications in Robot-Assisted Feeding. In Proceedings of the 2021 IEEE International Conference on Robotics and Automation (ICRA), Xi'an, China, 30 May–June 5 2021; IEEE: New York, NY, USA, 2021; pp. 10528–10535.
42. Berg, J.; Lu, S. Review of interfaces for industrial human-robot interaction. *Curr. Robot. Rep.* **2020**, *1*, 27–34. [\[CrossRef\]](#)
43. Mahmud, S.; Lin, X.; Kim, J.H. Interface for human machine interaction for assistant devices: A review. In Proceedings of the 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 6–8 January 2020; IEEE: New York, NY, USA, 2020; pp. 768–773.
44. Porcheron, M.; Fischer, J.E.; Reeves, S.; Sharples, S. Voice interfaces in everyday life. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, New York, NY, USA, 21–26 April 2018; pp. 1–12.
45. Suzuki, R.; Ogino, K.; Nobuaki, K.; Kogure, K.; Tanaka, K. Development of meal support system with voice input interface for upper limb disabilities. In Proceedings of the 2013 IEEE 8th Conference on Industrial Electronics and Applications (ICIEA), Melbourne, Australia, 19–21 June 2013; IEEE: New York, NY, USA, 2013; pp. 714–718.
46. De Pace, F.; Manuri, F.; Sanna, A.; Fornaro, C. A systematic review of Augmented Reality interfaces for collaborative industrial robots. *Comput. Ind. Eng.* **2020**, *149*, 106806. [\[CrossRef\]](#)
47. Dianatfar, M.; Latokartano, J.; Lanz, M. Review on existing VR/AR solutions in human–robot collaboration. *Procedia CIRP* **2021**, *97*, 407–411. [\[CrossRef\]](#)
48. Chamola, V.; Vineet, A.; Nayyar, A.; Hossain, E. Brain-computer interface-based humanoid control: A review. *Sensors* **2020**, *20*, 3620. [\[CrossRef\]](#)
49. Perera, C.J.; Naotunna, I.; Sadaruwan, C.; Gopura, R.A.R.C.; Lalitharatne, T.D. SSVEP based BMI for a meal assistance robot. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; IEEE: New York, NY, USA, 2016; pp. 2295–2300.
50. Perera, C.J.; Lalitharatne, T.D.; Kiguchi, K. EEG-controlled meal assistance robot with camera-based automatic mouth position tracking and mouth open detection. In Proceedings of the 2017 IEEE international conference on robotics and automation (ICRA), Singapore, 29 May–3 June 2017; IEEE: New York, NY, USA, 2017; pp. 1760–1765.
51. Ha, J.; Park, S.; Im, C.H.; Kim, L. A hybrid brain–computer interface for real-life meal-assist robot control. *Sensors* **2021**, *21*, 4578. [\[CrossRef\]](#)
52. Zhang, X.; Wang, X.; Wang, B.; Sugi, T.; Nakamura, M. Meal assistance system operated by electromyogram (EMG) signals: Movement onset detection with adaptive threshold. *Int. J. Control Autom. Syst.* **2010**, *8*, 392–397. [\[CrossRef\]](#)
53. Bhattacharjee, T.; Gordon, E.K.; Scalise, R.; Cabrera, M.E.; Caspi, A.; Cakmak, M.; Srinivasa, S.S. Is more autonomy always better? exploring preferences of users with mobility impairments in robot-assisted feeding. In Proceedings of the 2020 15th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Cambridge, UK, 23–26 March 2020; IEEE: New York, NY, USA, 2020; pp. 181–190.
54. Jain, L. Design of Meal-Assisting Manipulator AI via Open-Mouth Target Detection. *Des. Stud. Intell. Eng.* **2022**, *347*, 208
55. Tomimoto, H.; Tanaka, K.; Haruyama, K. Meal Assistance Robot Operated by Detecting Voluntary Closing eye. *J. Inst. Ind. Appl. Eng.* **2016**, *4*, 106–111.
56. Liu, F.; Xu, P.; Yu, H. Robot-assisted feeding: A technical application that combines learning from demonstration and visual interaction. *Technol. Health Care* **2021**, *29*, 187–192. [\[CrossRef\]](#) [\[PubMed\]](#)

57. Huang, X.; Wang, L.; Fan, X.; Zhao, P.; Ji, K. Facial Gesture Controlled Low-Cost Meal Assistance Manipulator System with Real-Time Food Detection. In Proceedings of the International Conference on Intelligent Robotics and Applications, Harbin, China, 1–3 August 2022; Springer: Berlin/Heidelberg, Germany, 2022; pp. 222–231.
58. Yunardi, R.T.; Dina, N.Z.; Agustin, E.I.; Firdaus, A.A. Visual and gyroscope sensor for head movement controller system on meal-assistance application. *Majlesi J. Electr. Eng.* **2020**, *14*, 39–44. [[CrossRef](#)]
59. Lopes, P.; Lavoie, R.; Faldu, R.; Aquino, N.; Barron, J.; Kante, M.; Magfory, B. Icraft-eye-controlled robotic feeding arm technology. *Tech. Rep.* **2012**.
60. Cuong, N.H.; Hoang, H.T. Eye-gaze detection with a single WebCAM based on geometry features extraction. In Proceedings of the 2010 11th International Conference on Control Automation Robotics & Vision, Singapore, 7–10 December 2010; IEEE: New York, NY, USA, 2010; pp. 2507–2512.
61. Huang, Q.; Veeraraghavan, A.; Sabharwal, A. TabletGaze: dataset and analysis for unconstrained appearance-based gaze estimation in mobile tablets. *Mach. Vis. Appl.* **2017**, *28*, 445–461. [[CrossRef](#)]
62. WebCam Eye-Tracking Accuracy. Available online: <https://gazerecorder.com/webcam-eye-tracking-accuracy/> (accessed on 8 December 2022).
63. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
64. George, A.; Routray, A. Real-time eye gaze direction classification using convolutional neural network. In Proceedings of the 2016 International Conference on Signal Processing and Communications (SPCOM), Bangalore, India, 12–15 June 2016; IEEE: New York, NY, USA, 2016; pp. 1–5.
65. Chen, R.; Kim, T.K.; Hwang, J.H.; Ko, S.Y. A Novel Integrated Spoon-chopsticks Mechanism for a Meal Assistant Robotic System. *Int. J. Control Autom. Syst.* **2022**, *20*, 3019–3031. [[CrossRef](#)]
66. Song, H.; Jung, B.J.; Kim, T.K.; Cho, C.N.; Jeong, H.S.; Hwang, J.H. Development of Flexible Control System for the Meal Assistant Robot. In Proceedings of the KSME Conference, Daejeon, Republic of Korea, 29–31 July 2020; The Korean Society of Mechanical Engineers: Seoul, Republic of Korea, 2020; pp. 1596–1598.
67. Jung, B.J.; Kim, T.K.; Cho, C.N.; Song, H.; Kim, D.S.; Jeong, H.; Hwang, J.H. Development of Meal Assistance Robot for Generalization of Robot Care Service Using Deep Learning for the User of Meal Assistant Robot. *Korea Robot. Soc. Rev.* **2022**, *19*, 4–11. (In Korean)
68. Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; IEEE: New York, NY, USA, 2009; pp. 248–255.
69. Park, H.; Jang, I.; Ko, K. Meal Intention Recognition System based on Gaze Direction Estimation using Deep Learning for the User of Meal Assistant Robot. *J. Inst. Control Robot. Syst.* **2021**, *27*, 334–341. [[CrossRef](#)]
70. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv* **2020**, arXiv:2004.10934.
71. Vrânceanu, R.; Florea, C.; Florea, L.; Vertan, C. NLP EAC recognition by component separation in the eye region. In Proceedings of the International Conference on Computer Analysis of Images and Patterns, York, UK, 27–29 August 2013; Springer: Berlin/Heidelberg, Germany, 2013; pp. 225–232.
72. Florea, L.; Florea, C.; Vrânceanu, R.; Vertan, C. Can Your Eyes Tell Me How You Think? A Gaze Directed Estimation of the Mental Activity. In Proceedings of the BMVC, Bristol, UK, 9–13 September 2013.
73. TensorFlow. Available online: <https://www.tensorflow.org/lite/> (accessed on 8 December 2022).

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.