Applied Computer Science 2025, vol. 21, no. 1, pp. 18–30
https://doi.org/10.35784/acs_6867

Submitted: 2024-11-28 | Revised: 2025-02-22 | Accepted: 2025-03-01

Keywords: spatial identification, shape recognition, bionic limb prosthetics, artificial neural networks, microcontroller module

Yurii LOBUR [1*], *Kostiantyn VONSEVYCH* [1], *Natalia BEZUGLA* [1]
[1] Igor Sikorsky Kyiv Polytechnic Institute, Ukraine, juderyesp@gmail.com
[*] Corresponding author: juderyesp@gmail.com

Spatial identification of manipulable objects for a bionic hand prosthesis

Abstract

This article presents a method for the spatial identification of objects for bionic upper limb prostheses, utilizing the analysis of digital images captured by an optoelectronic module based on the ESP32-CAM and classified using neural network algorithms, specifically FOMO (MobileNetV2). Modern bionic prostheses that imitate natural limb functions, as well as their advantages and significance for restoring the functionality of the human body, are analysed. An algorithm for a grip-type recognition system is proposed, integrating spatial identification of object shapes with the analysis of myographic signals to enable accurate selection and execution of appropriate manipulations. The neural network was trained on a set of images of basic shapes (spherical, rectangular, cylindrical), which achieved an average identification accuracy of over 89% with a processing time of one image of 2 ms. Due to its compactness and low cost, the developed system is suitable for integration into low-cost prostheses, ensuring adaptation of the movements of the artificial limb to the shape of the objects of manipulation and minimizing the risk of slipping objects. The proposed approach helps to increase the accuracy of movement execution and reduce dependence on expensive and complex technologies. The system has potential for further improvement, as it can operate with objects of complex shapes and handle scenarios involving multiple objects within the camera's field of view simultaneously.

1. INTRODUCTION

The modern development of additive technologies, artificial intelligence, and the general availability of information opens up new ways for researchers to solve complex applied problems, particularly in robotics, bionics, and restoration of human body functions. Research aimed at creating biomimetic analogues and robotic devices, including reproducing human natural capabilities, is becoming increasingly popular. A typical example is the development of devices for restoring the sense of smell (Hurot et al., 2020; Kim et al., 2021) or vision (Kim et al., 2022; Chinnery et al., 2016), the operation of which is based on the principles of functioning of the relevant systems of the human body. Equally high-tech and essential is the field of limb prosthetics, where modern bionic arm (Ortiz-Catalan et al., 2023; Wijk et al., 2024) and leg prostheses (Azocar et al., 2020; Tran et al., 2022) not only simulate the visual presence of the lost limb but also replicate a specific range of movements. For example, Azocar et al. (2020) describe the development and clinical implementation of a lower limb prosthesis for patients with transfemoral amputation. In practical applications, this prosthesis not only replicates the supporting functions of a healthy leg but also enables complex tasks such as dynamic movement on flat surfaces, ramps, and ascending and descending stairs.

In addition, when designing such devices, their total weight, autonomy of operation, trajectory of the main drive mechanisms, etc., are considered. In the case of bionic hands, beyond achieving aesthetic similarity to natural limbs, it is crucial to replicate fine motor skills and elements of natural sensory feedback. This presents developers with a set of highly ambitious challenges, particularly those related to the biomechanics of prosthetic movement. Specifically, the ability to reproduce precise movements of artificial joints, perform a set of natural grasps or gestures, manipulate individual fingers of the palm, etc. The robotic arm system usually includes several fundamental components that are necessary for the implementation of such tasks. For example, a biological signal measurement module, a central processing unit (CPU), drive mechanisms and their control circuit, a set of sensors, a power supply system, etc. (Vonsevych, 2024). Precision control of the movement of the artificial limb is ensured using biological signals from the patient's body, which are further analyzed as part of the CPU using machine learning methods (Chen et al., 2023; Zbinden et al., 2024) and directly affect the further performance of manipulations according to the reaction of the user's body.

Machine learning (ML) methods are widely utilized for the analytical processing of large datasets, particularly in solving nonlinear and nonstationary problems, identifying interdependencies and correlations. Specific ML mathematical algorithms can be applied across various domains, such as analysing purely technical parameters of industrial equipment, adjusting actuator parameters (Józwik et al., 2024), or addressing scenarios characterized by high uncertainty of outcomes, including medical applications, the analysis of biological signals, images, disease diagnosis, etc. (Machrowska et al., 2024).

Among the wide range of machine learning algorithms notable examples include Artificial Neural Networks, Decision Trees, Support Vector Machines, Elastic Net, Random Forest and others. Each of these algorithms offers unique advantages, whether in terms of computational speed or the accuracy of decisions, depending on the type of task for which the specific method is employed. For instance, algorithms such as Multilayer Perceptrons (MLPs) and Radial Basis Function (RBF) networks are frequently used for classification tasks but exhibit limitations in areas such as image processing. In contrast, convolutional neural networks (CNNs) are actively employed in computer vision technologies and various image analysis processes. Due to their use of convolutional layers, CNNs enable the identification of hidden dependencies and the automatic extraction of spatial patterns within images (e.g., lines, textures, contours) which can be particularly effective in the medical field.

Neural networks trained using Bayesian Regularization algorithms are especially beneficial for forecasting tasks involving noisy data under conditions of limited input availability. These ML algorithms are particularly valuable in medical diagnostics, where data may contain significant artefacts but require reliable predictions. They are utilized in the analysis of electrocardiograms, X-ray or MRI images, as well as in bionic prosthetics and robotics for predicting user movements based on electromyography signals.

Measurement of control biological signals can be carried out by various methods, among which the most used in the field of prosthetics are brain-computer interface (BCI) methods (EEG, ECoG, ENG), electromyography (EMG), measurement of mechanical muscle movement (FMG, MMG), etc. (Esposito et al., 2021; Jiang et al., 2023; Mereu et al., 2021). At the same time, using a combination of sEMG signals with the technology of artificial neural networks is one of the leading combinations used in the practice of non-invasive control of bionic limbs and recognition of natural gestures (Said et al., 2020; Lee et al., 2021; Vonsevych et al., 2019).

Additional modules and sensors that are part of the bionic system provide the performance of a specific set of gestures and the further adaptation of limb movements, touch strength, and grip strength. For this purpose, a feedback link with the patient's body can also be introduced into the general system of the prosthesis. The practical implementation of such a link can be performed both with further impact on the user's body (human body interaction, HBI) and in the form of a module for autonomous object identification (self-automatic control, SAC) (Vonsevych et al., 2019). Usually, the implementation of HBI systems is carried out by transferring an external stimulus to the patient's stump after the fingers of the prosthesis touch the object or object of manipulation. At the same time, the effect on the body can be carried out by proportional transmission of vibration, superficial electrical stimulation of afferent fibers, temperature change at some point on the surface of the patient's stump, etc. (Sensinger et al., 2020; Lan et al., 2023).

At the same time, SAC identification does not necessarily imply an impact on the limb. Still, it provides the CPU with additional information about the object and, as a result, about the need to adjust the current position of the actuators. Such information may include the type of object, the hardness or roughness of its surface, its overall dimensions, the current position of the limb relative to the object, and other relevant parameters. For example, the papers (Vonsevych et al., 2019; Shi et al., 2023) present the implementation of modules for identifying the type of contact surface of the objects manipulating a bionic prosthesis. In studies (Lin et al., 2024; Huang & Wu, 2021), an algorithm and measuring tools for determining the roughness and texture of objects with which an artificial limb can interact. Thus, the use of both HBI and SAC feedback links can serve as an additional channel of helpful information for the CPU of the prosthesis, which in turn affects the overall functionality of the device, the ability to adjust grip strength, determine the optimal position of the actuators of the artificial limb, etc.

One of the critical aspects of enhancing the functionality of such devices is their ability to recognize the shape of objects being manipulated, thereby preventing slippage from plastic finger phalanges, including those manufactured using 3D printing technologies. The detection of object slippage from the fingers of a prosthesis can be achieved through various algorithms and technical solutions, including the integration of multiple sensors within a single limb. For example, Zeng et al., (2022) describes a system for controlling the slippage of objects from the grip of an artificial limb using tactile sensors placed on the executive surfaces of the bionic arm. Similarly, the study by James et al. (2018) presents a prototype of an upper limb prosthesis capable of detecting object slippage through a combination of optical and tactile sensors. Initial object identification and subsequent shape determination are typically performed using optical scanning or by integrating more advanced technologies and computer vision principles into the prosthesis.

For example, Fejér et al. (2022), it is proposed to recognize the shape of objects by using a hybrid hardware-software visual control system. Computer vision technology is used here, which makes it possible to distinguish the composition of the general scene presented to the prosthesis user from the location of a particular object and its approximate dimensions. Although this technology allows for machine recognition of objects and increases the possibilities of using an artificial limb, it requires the mandatory use of additional devices that are not an actual part of the prosthetic arm (for example, glasses with a set of special sensors). As a result, this can complicate the further user training process and increase the product's final cost. Even more technologically complex is the recognition of detailed contours of objects of manipulation. Thus, the study by Wei and Chen (2020) presents an example of determining complex object shapes for robotic limbs, requiring the execution of several computationally intensive operations. These include background leveling, shape descriptor extraction, superpixel segmentation, image smoothing, edge detection, intermediate refinements of reference lines, and contour reference points, among others. Although such a shape recognition algorithm allows for high-precision detection of the shape of objects with which the user will interact in the future, they require powerful computing equipment and computer resources, which is a rather tricky task in the practice of bionic prosthetics. More promising for practical application in the composition of myoprotheses seems to be the approach proposed in the paper (Vásquez & Perdereau, 2017), where the recognition of the shape of objects is carried out by combining several types of proprioceptive sensors, which are located in different parts of the fingers of the robotic limb and generate a set of signals on which in the future, after machine classification, a decision is formed on a generalized proprioceptive signature of the shape, which does not change concerning the size of the objects. Given the complexity of processing large datasets of images in computer vision technologies, the extensive mathematical transformations required for digital object recognition, and the consequent need for bulky equipment, the practical application of computer vision in bionic prosthetics remains limited.

2. THE PROPOSED METHOD

The analysis of myoelectric signals from the human upper limb is a complex task, particularly when interpreting the measured indicators of residual muscle activity in relation to the set of movements or grasps that the robotic prosthesis should replicate. The natural biomechanics of the limbs implement a wide range of supporting movements and precise manipulations using the fingers. For example, to reproduce elements of sign language (Llop-Harillo et al., 2019) or to perform various grasps used in a person's daily life. According to Cutkosky's taxonomy and studies of hand grasps presented in papers (Llop-Harillo et al., 2019), the main types used for most manipulations can be divided into three general categories. Depending on the shape and how large the object the palm should cover, there are power, intermediate, and precision types of grasps. At the same time, depending on the position of the thumb used in implementing these movements, each category contains subcategories and certain types of movements that can be used to manipulate a specific object (for example, hook, lateral pinch, tripod, etc.). As a result, implementing a measurement system and high-precision recognition of myosignals for such a significant number of manipulations as part of an artificial limb is a challenging task that requires significant computing power and a large number of measuring channels. At the same time, if you pay attention to the shape of the objects or their separate part, about which the grip and interaction with the fingers of the palm occur, then you can trace a specific relationship between different types of grasps and the shape of the object of manipulation for which these grasps are used. For example, an extension grip is used for large rectangular objects, a tripod - for medium and small objects of spherical and cylindrical shape, etc. The authors of this article propose a method for identifying the basic shape of the object of manipulation as an element of the general algorithm for controlling a bionic hand prosthesis, which is designed to increase the efficiency of recognizing the patient's motor activity in the process of forming a decision on the type of grasp that an artificial limb should reproduce. The proposed method's general principle can be presented as a diagram in Fig. 1.

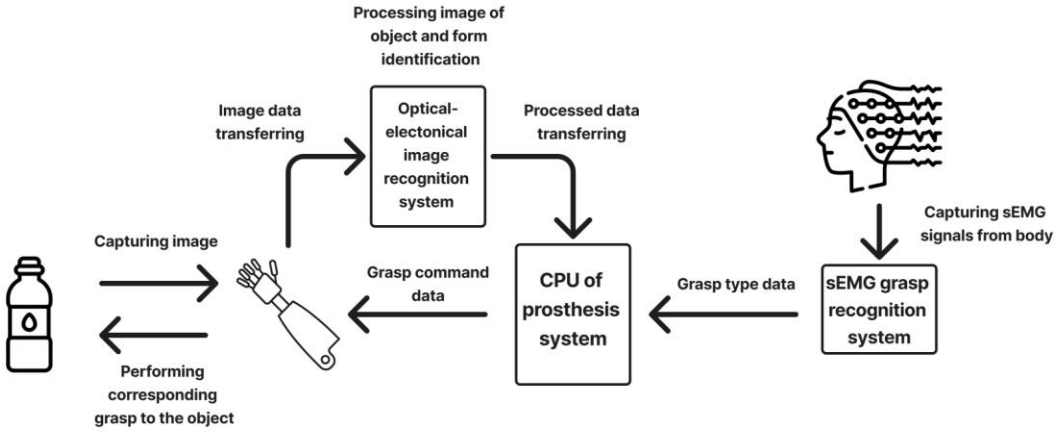


Fig. 1. The general principle of the proposed method for grasp type recognition

As can be seen in the diagram, the result of spatial identification of the object shape with which the artificial limb will interact should be correlated with a predetermined type of grasp, which is recognized by the myographic system of the prosthesis in the process of analyzing signals of the patient's stump muscles. Suppose the kind of grasp chosen by the CPU of the prosthesis based on the myogram corresponds to the identified shape of the object of manipulation to which it can be applied. In that case, the final decision is made to allow the activation of the fingers and the choice of their movement parameters. For example, a power grip is for large-sized spherical objects, a hook is for cylindrical objects, a tripod is for small spherical objects, etc. At the same time, the parameters of finger movement include the actuator's final position and speed, the phalanges' maximum contact force, etc. At the same time, it should be noted that the method of measuring the myographic signal in this case is unimportant and, as a result, is not limited to EMG measurements only.

In cases of mismatch between the categories determined by the myogram and by classifying the shape of the objects, it is envisaged that re-identification should be decided, or another possible type of movement from the categories available in the prosthesis control system should be chosen.

3. MATERIALS AND METHODS

Identification of the shape of the manipulation objects was carried out based on the classification of a set of images obtained from the camera of the optical-electronic module using an artificial neural network (ANN) based on the FOMO architecture from Edge-impulse (Sandler et al., 2018), which is called MobileNetV2. The general structure of the ANN model used is shown in Fig. 2 and consists of a convolutional layer, deep layers of a split convolution, a layer of inverted residues with linear constrictions, and a block of standard ANN layers. The peculiarity of the architecture of the FOMO model is that the number of convolutional channels here is reduced and is set by a particular width factor. The main blocks of architecture at the level of deep divided convolutions consist of two parts: deep and point convolution—moreover, inverted remnants with linear narrowing containing expansion point, depth and projection convolutions. After passing through each block with a particular step sf, linear activation functions used for storage of helpful information, ReLU6 activation function ensures nonlinearity after expansion and projection, and Batch normalization is performed after each convolution.

This approach reduces the overall complexity of working with the computational algorithm. It makes it relatively easy to integrate the FOMO model into portable controllers of low power, such as the ESP-32. Compared to similar algorithms for detecting objects such as YOLOv5, YOLOX, or Deep SORT, which can be used in industrial robotic arms (Durve et al., 2023; Luo et al., 2024), the FOMO model integrated into the ESP microcontroller does not require the use of large-sized strapping or additional computing power, which is an undeniable advantage when using it as an integral part of a portable bionic prosthesis system.

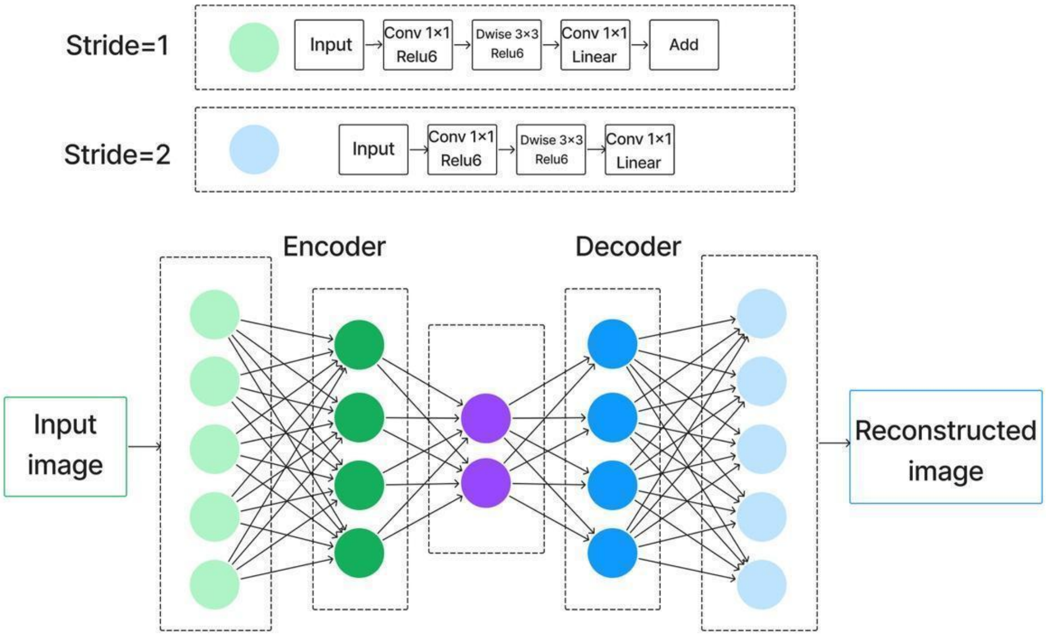


Fig. 2. The general architecture diagram and ANN graphs for used MobileNetV2 FOMO model

After preliminary testing of various mathematical models, MobileNetV2 was chosen for practical testing of ANN with a set of parameters shown in Table 1 and the width factor k_w = 0.35, which reduces the number of filters at each stage of neural network training by proportion (1):

$$f_k = f_i \cdot k_w \tag{1}$$

where: f_k– the number of filters in the block after applying the width factor,
f_i – the initial number of filters, according to the convolutional configuration of the model,
k_w– width factor value.

A microprocessor optoelectronic module for forming a set of images for a neural network and subsequent spatial identification of the shape was implemented as one of the components of the feedback link for the general system of the bionic arm (Fig. 3). This link also includes sensors for the strength and spatial position of the limb. Such a set of sensitive elements is designed to determine the initial position of the palm relative to the object of manipulation and as a result, to further adjust parameters of the movement and control the grip force on each prosthetic finger.

However, within the framework of this article, only the optoelectronic module is tested without its use in conjunction with other sensors used to adjust the grip parameters for specific objects.

Tab. 1. Set of parameters used for ANN model

Layer name	**Block Number**	**Steps Number, s_f**	**Filters Number, f_k**
Initial convolution			
Convolution	-	3	6
Inverted Residual Block			
Residual Block	1	1	6
Residual Block	2	2	8
Residual Block	3	2	11
Residual Block	4	2	22
Residual Block	5	1	34
Residual Block	6	2	56
Residual Block	7	1	112
Final Layers	-	1448	

The module's constituent components were selected based on the possibility of its further practical implementation into the bionic prosthesis and certain principles of implementation of computer vision technologies.

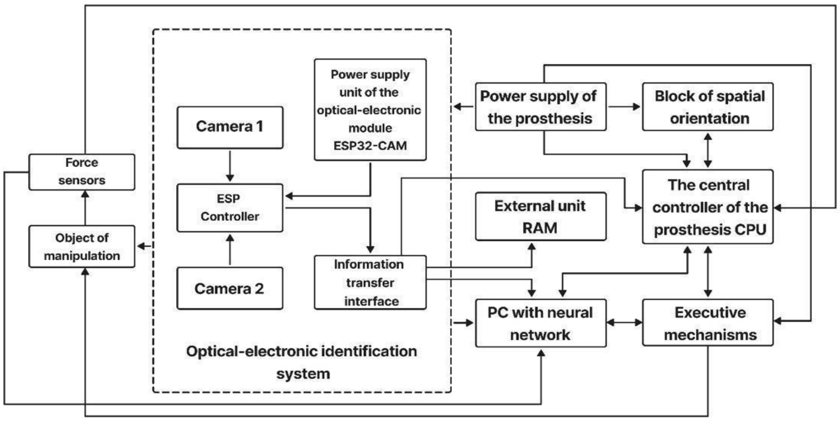
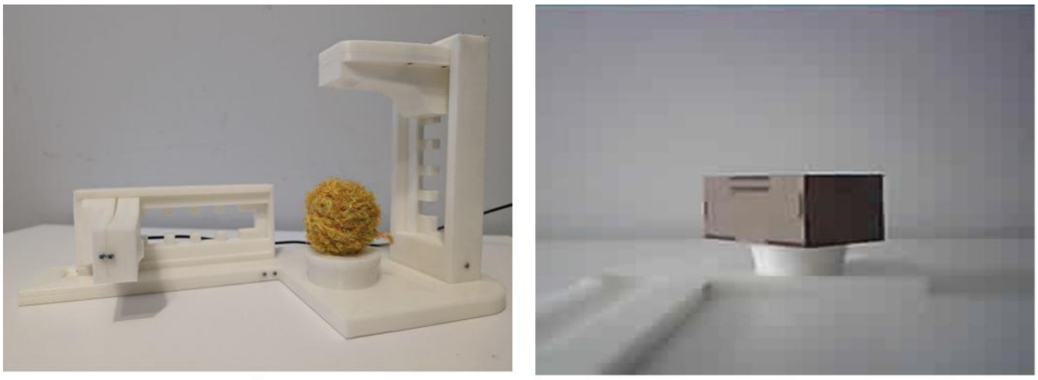


Fig. 3. Functional diagram of an optoelectronic identification module based on ESP32-CAM as part of a bionic prosthesis

Specifically, the authors determined that the module should include multiple optical sensors for capturing images from different planes of the object space, a microcontroller independent of the prosthesis's CPU with the capability to connect to a PC, external memory, a power supply, and an information transmission system. Two built-in cameras of the ESP32-CAM module were chosen as measuring elements for the optical channel. The resolution of the cameras is 2MP, and the module contains a hardware controller and built-in Bluetooth, USB, and Wi-Fi interfaces for transmitting information. A separate battery powers the proposed module, and raw images can also be stored on an external RAM unit connected to a computer or the CPU of the bionic prosthesis.

For the convenience of the experiment, before the start of the measurement procedure, the recording elements of the module were placed on a specially designed measuring device, the image of which is presented in Fig.4a. The installation was designed in such a way that the distance from the entry point of the optical channel to the platform on which the object of study was placed could be changed in a specific step (1 cm). The final set of images could be recorded at a clearly defined distance from the object. The ANN training of the optoelectronic system was carried out on a set of 235 color images of 240x240 pixels in size, obtained for different objects at two points of the object space (X and Y planes).



a **b**

Fig. 4. Placement of the module relative to the object of manipulation:
a) measuring setup, b) example of the image from the X plane

According to the terms of the study, the detailed contour of each test item was subject to approximation to a generalized shape corresponding to one of the basic categories of simple shapes: spherical, rectangular, cylindrical in vertical, and cylindrical in horizontal arrangement. This choice of categories was made based on the AHA protocol (Llop-Harillo et al., 2019) and the analysis of the taxonomy of the main grasps of the hand presented in the papers (Sharma et al., 2019; Cini et al., 2019). Here, a spherical category of objects is defined as one to which grasps such as tripod, disk, pulp can be applied, rectangular - palmar, pinch, cylindrical - wrap, hook, ring, etc. When training the network, each category was given a set of images for two different objects of similar shapes, and the image of a third item of the same category was used only to test the system. In addition, the training set also used the background image without placing any other components of the overall scene in the camera's field of view. Such information in the training set is necessary to isolate the background as a noise component in the detection algorithm, which can affect the reliability of the ANN model. The percentage distribution of images was 80% for testing and 20% for network training, the number of training epochs e = 30, and the learning rate (error) = 0.005. The efficiency of object classification was assessed by constructing an error matrix (Vonsevych et al., 2019) and calculating the F1-score parameter (Yacouby & Axman, 2020).

4. RESULTS AND DISCUSSION

To determine the optimal parameters of the FOMO model for spatial identification of objects, several additional stages of testing the optoelectronic module on static images of objects of different sizes corresponding to the previously defined set of basic shapes of manipulation objects were carried out (Fig. 5). The model was tested on images of objects that had not previously been used in the training dataset, placed in the X plane. The object of study was at a working distance of l_1 = 30 cm from the camera lens.

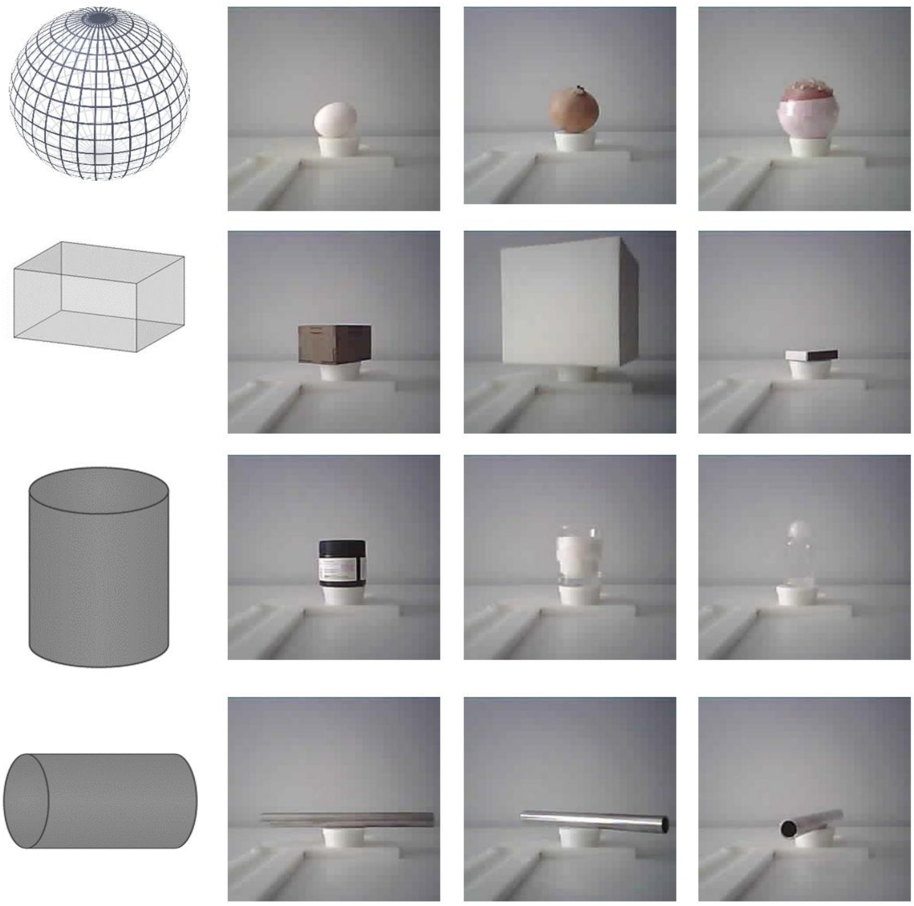


Fig. 5. Examples of images of objects obtained from the optoelectronic module: spherical, rectangular, cylindrical vertical and cylindrical horizontal shapes

According to the F1-score indicator, the average identification reliability obtained at different stages of testing was 89.2%, where the system's solutions were wrong in most cases only for the category of large-sized rectangular objects. This peculiarity can be explained by the fact that depending on the angle of the obtained image, the size of the object, and the distance from it to the entrance to the optical channel, one of its sides can be recognized by the system as an element of the general background. To further analyse classification errors, the authors plan to expand the training dataset with images of rectangular-shaped objects in future studies. Additionally, the system will undergo further testing under varying conditions, including changes in background, lighting intensity, camera angles and the positions of the rectangular object. The confusion matrix for the practical test is presented in Table 2, and the corresponding ANN parameters were used to load into an optoelectronic module based on ESP32-CAM to further assess the reliability of spatial identification of objects in real-time.

Tab. 2. Confusion matrix identification for the ANN FOMO model

	Background	**Cylinder horizontal**	**Cylinder vertical**	**Rectangle**	**Sphere**
Background	99.4%	0%	0.6%	0%	0%
Cylinder horizontal	0%	100%	0%	0%	0%
Cylinder vertical	0%	0%	100%	0%	0%
Rectangle	0%	0%	0%	100%	0%
Sphere	0%	0%	0%	0%	100%
F1 Score	1.00	1.00	0.60	1.00	1.00

Real-time identification was carried out on a set of 12 items that the user may encounter in everyday life, which were not included in the set of images used for training or practical testing of ANN. Namely spherical eggs, rectangular boxes of various sizes, cylindrical vials and tubes, etc. (Fig. 6).

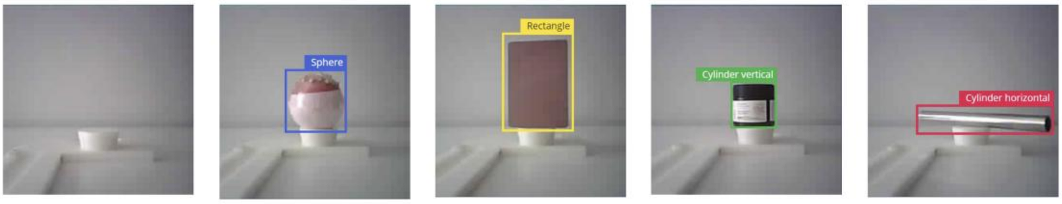
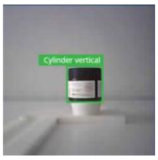
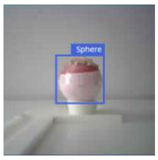


Fig. 6. Examples of images of objects obtained from the optoelectronic module: spherical, rectangular, cylindrical vertical and cylindrical horizontal shapes

This item selection was done to create relevant images, an example of what the prosthesis user may encounter when performing everyday manipulations. At the same time, for the convenience of further integration of the proposed optoelectronic module into the control system of the bionic prosthesis, information about the result of spatial identification was displayed by logging in the form of a data package. The package included: an obtained result of classification accuracy for a specific category of the form, the coordinates of the object's location in the camera's field of view, and its dimensions in the X and Y planes. Moreover, the total time of DSP calculations during the classification for one essential form category was 2 ms.

As can be seen from Table 3, the results of logging the identification process demonstrate the absence of any information about the object's size or the reliability of determining the category of its shape in the presence of only the general background within the studied scene. However, immediately after appearing in the field of view against the general background of any object, its relative dimensions relative to the dimensional frame and the coordinates of the location of the object of study in the X and Y planes are determined. At the same time, it can be seen that the designed network as part of the optoelectronic module requires some time to stabilize its algorithm, which in most cases takes about 1000 ms and reduces the initial reliability of identification (the first calculated value of F1-score after the appearance of the item).

Tab. 3. Real-time object identification results

Real-time images					
Object Shape	Background	Cylinder vertical	Rectangle	Cylinder horizontal	Sphere
Authenticity of identification according to F1-score	-	0.81328	0.99218	0.89843	0.80781
	-	0.98046	0.99218	0.92187	0.98828
	-	0.98828	0.99218	0.89843	0.98046
Coordinates of the object in the X and Y planes	-	x: 24, y: 8	x: 16, y: 24	x: 16, y: 24	x: 16, y: 16
	-	x: 24, y: 8	x: 0, y: 24	x: 16, y: 24	x: 24, y: 16
	-	x: 16, y: 0	x: 8, y: 8	x: 16, y: 24	x: 24, y: 16
Relative dimensions of an object in the X and Y planes	-	width: 8, height: 8	width: 16, height: 8	width: 16, height: 8	width: 16, height: 16
	-	width: 8, height: 16	width: 8, height: 24	width: 8, height: 8	width: 8, height: 16
	-	width: 16, height: 24	width: 8, height: 8	width: 8, height: 8	width: 8, height: 16

For example, in the case of a vertically placed cylindrical object (Table 3), the reliability of initial identification is approximately 81%, according to the F1-score indicator. However, immediately after the module is stabilized, the confidence score increases to 98% and remains consistently high throughout all subsequent iterations. Similarly, high performance, even at the first appearance of the object as part of the general scene, is demonstrated by the spherical and rectangular shape of the manipulation object of both small and medium sizes (Table 3). However, although indicators of identification reliability at 99.2% were obtained during the approbation of rectangular objects, such a class of objects in the training set of images also demonstrates a negative impact on the classification of horizontally placed cylindrical objects. Here the accuracy identification value is reduced to values within the range of 89.8% to 92.2% and is sometimes recognized by the network precisely as a cylindrical horizontal shape.

Based on the obtained indicators of the reliability of spatial identification of the shape of manipulation objects and the results obtained in similar studies by other authors, it can be argued that the proposed combination of the optoelectronic module and the FOMO model of ANN is quite effective and can be used for further testing as part of a bionic upper limb prosthesis. For example, in the work (Cai et al., 2016), which investigates the relationship between the types of grasp and the attributes of the object, the obtained indicators of reliability of the classification of 4 forms of objects vary from 80.2% to 94.0%, depending on the algorithm used and the corresponding shape of the object under study. A similar approach to recognizing the shape of objects is also presented in papers (Zhang et al., 2020; Wang et al., 2021), where information from the optical channel and heat maps is used for further capture of objects by robotic manipulators. At the same time, for objects of different basic shapes, the optimal location of the rectangular gripping zone for the gripper of the manipulator's actuators is determined here. An example of the practical integration of optical identification technologies into the general system of the bionic limb is presented in (Castro et al., 2022). Where the identification of objects is also carried out, the shape of which is similar to the one discussed in this article. However, here, the practical approbation is performed for four types of grippers that correlate with the shape of objects, and a laser designator is also used as an additional information channel.

The proposed solution based on the FOMO classifier while requiring further validation on more complex input datasets, demonstrates quite relevant results compared to other neural networks models and machine learning algorithms for spatial identification of manipulation objects. According to the analysis presented in the study (Bai et al., 2020) an example of using a machine classifier for recognizing four categories of objects (cylindrical, spherical, cubic and conical) with similar classification accuracy is the CNN model of the VGG16 type. This method provides an object recognition accuracy at the level of 93% in real-time. At the same time, the PointNetGPD end-to-end grasping evaluation model, similar to the proposed FOMO model, is lightweight and categorizes objects according to their general shape. This algorithm uses point cloud images as input data, enabling the analysis of objects with complex shapes. However, it introduces a strong dependency on the quality of 3D reconstruction of the manipulation object, based on RGB-D image sets. This factor directly impacts the reliability of subsequent object classification and consequently the selection of grips and grasping strategies.

In the works (He et al., 2019; He et al., 2020), object detection models trained on five everyday items (a cup, bottle, spray bottle, ball and stapler) are presented, utilizing YOLO algorithms. However, in the study (He et al., 2019), YOLO used with a Pascal Titan X5 graphics processor demonstrates classification accuracy exceeding 80%. In the study (He et al., 2020) an adapted model is presented, which unlike the classic YOLO algorithm, includes 13 convolutional layers and achieves recognition accuracy of over 90% for the five objects. However, this requires two high-performance GPU computing nodes based on GeForce GTX Titan X and GeForce GTX 970 that consume 2.2 [GB] of GPU memory during image processing.

In further works by the authors, it will be advisable to check the proposed methodology and technologies for spatial identification of objects of a composite shape. An example of objects with a composite shape can be items, the individual parts of which have different categories of basic shapes considered in this article. For example, a flask, a comb, scissors, etc. It will also be advisable to approbate a scenario in which several objects of different shapes are simultaneously in the optoelectronic module camera's field of view, for example, as shown in a paper (Roy et al., 2018).

Considering that the current study focuses specifically on analysing static images of objects, the authors acknowledge that challenges related to the practical implementation of the proposed technology in a dynamic environment remain insufficiently explored and require additional attention before its direct integration into a bionic hand. Future research should include dynamic testing, such as recognizing moving objects of various sizes and testing the system under changing lighting conditions to determine its operational limits. Providing future tests on more diverse datasets will enhance the practical utility of the results and the relevance of the proposed method to real-world conditions in which a prosthetic user may operate.

Another issue requiring further attention is integrating the bionic limb with proposed solution into clinical practice. While technical factors such as the compact size of the optoelectronic module, its computational capabilities for image analysis and the overall cost of electronics are not critical (through the combination of a compact low-cost system based on the ESP-32CAM and the FOMO model), the training and adaptation of prosthetic users, integration of the module with existing medical systems for recognizing myoelectric signals, compliance with regulatory standards and preparation for approval by regulatory authorities should be further considered.

Integrating computer vision technologies into prosthetics introduces a new user interface, which may complicate the adaptation process for patients using a bionic hand, potentially increasing the time required for their training. The operation of the proposed module must synchronize seamlessly not only with the custom-designed general control interface for the artificial hand but also adapt to existing systems and technologies for myosignal recognition, such as COAPT, Myo Plus by Ottobock, or I-Limb Quantum by Össur. Additionally, the final cost of implementing computer vision technologies into bionic prosthesis before their actual clinical application, may also be influenced by the certification process required by regulatory authorities such as CE in Europe or FDA in the United States.

5. CONCLUSIONS

In this study, the authors proposed an approach to spatial identification of objects of manipulation of bionic upper limb prostheses by using the optical channel for obtaining information and determining the basic shapes of objects to further adapt the movements of the robotic limb to a specific type of grasping. The basis of the technical component in the proposed approach is the further use of microcontroller optoelectronic module

ESP-32 CAM and an artificial neural network based on the FOMO (MobileNetV2) algorithm as additional parts in the bionic prosthesis system.

The study results showed that the proposed approach provides high reliability of identification of objects of various basic shapes (spherical, cylindrical, rectangular), which is essential for accurate execution of grasps and further minimizing the risk of objects slipping from the fingers of the prosthesis.

In particular, the average identification reliability indicator (estimated by the F1-score criterion) was over 89%, an acceptable value for practical use in bionic limb. A significant advantage of the proposed approach and its technical component is the overall compactness of the measuring equipment and the ability to work in real-time on limited computing power. This allows it to be integrated into low-cost prostheses without further using more complex and expensive equipment. Using the FOMO neural network model also provides the ability to quickly process images and classify objects with the processing time of one image within 2 milliseconds.

However, it is worth noting that the proposed solution also demonstrates specific difficulties when working with large rectangular objects due to the possibility of their partial loss in the camera's field of view, which, as a result, can lead to erroneous results. The authors also note that one of the critical areas of further research is adapting the proposed solution to work with objects of complex shapes and expanding the identification capabilities in situations when several objects of different shapes are simultaneously in the camera's field of view. It is also necessary to provide dynamic tests with moving objects under varying lighting conditions and backgrounds. This can significantly increase the system's overall application effectiveness in patients' daily lives, where manipulations often require simultaneous interaction with several static or moving objects. The authors also identify potential challenges that may arise during further integrating the artificial limb with the proposed technologies into clinical practice.

Thus, the proposed approach and its technical implementation can be essential to the general algorithm for controlling a bionic hand prosthesis. This establishes an additional information channel that aids in identifying the shape of the manipulation object, thereby enabling potential improvements in the efficiency of grasping by the robotic limb. This is particularly beneficial in scenarios where myosignal classification is challenging or limited.

Conflicts of Interest

The authors declare no conflict of interest.

REFERENCES

Azocar, A. F., Mooney, L. M., Duval, J. F., Simon, A. M., Hargrove, L. J., & Rouse, E. J. (2020). Design and clinical implementation of an open-source bionic leg. *Nature biomedical engineering, 4,* 941-953. https://doi.org/10.1038/s41551-020-00619-3

Bai, Q., Li, S., Yang, J., Song, Q., Li, Z., & Zhang, X. (2020). Object detection recognition and robot grasping based on machine learning: A survey. *IEEE Access*, *8*, 181855-181879. https://doi.org/10.1109/ACCESS.2020.3028740

Cai, M., Kitani, K. M., & Sato, Y. (2016). Understanding hand-object manipulation with grasp types and object attributes. *ArXiv, abs/1807.08254*. http://dx.doi.org/10.48550/arXiv.1807.08254

Castro, M. C. F., Pinheiro, W. C., & Rigolin, G. (2022). A hybrid 3D printed hand prosthesis prototype based on sEMG and a fully embedded computer vision system. *Frontiers in Neurorobotics*, *15*, 751282. http://dx.doi.org/10.3389/fnbot.2021.751282

Chen, Z., Min, H., Wang, D., Xia, Z., Sun, F., & Fang, B. (2023). A review of myoelectric control for prosthetic hand manipulation. *Biomimetics*, *8*(3), 328. https://doi.org/10.3390/biomimetics8030328

Chinnery, H., Thompson, S., Noroozi, S., & Dyer, B. T. (2016). Scoping review of the development of artificial eyes throughout the years. *Edorium Journal of Disability and Rehabilitation*, *3*, 1-10. http://dx.doi.org/10.5348/D05-2017-25-RA-1

Cini, F., Ortenzi, V., Corke, P., & Controzzi, M. J. S. R. (2019). On the choice of grasp type and location when handing over an object. *Science Robotics*, *4*(27), eaau9757. https://doi.org/10.1126/scirobotics.aau9757

Durve, M., Orsini, S., Tiribocchi, A., Montessori, A., Tucny, J. M., Lauricella, M., Camposeo, A., Pisignano, D., & Succi, S. (2023). Benchmarking YOLOv5 and YOLOv7 models with DeepSORT for droplet tracking applications. *The European Physical Journal E, 46*, 32. https://doi.org/10.1140/epje/s10189-023-00290-x

Esposito, D., Centracchio, J., Andreozzi, E., Gargiulo, G. D., Naik, G. R., & Bifulco, P. (2021). Biosignal-based human–machine interfaces for assistance and rehabilitation: A survey. *Sensors, 21*(20), 6863. https://doi.org/10.3390/s21206863

Fejér, A., Nagy, Z., Benois-Pineau, J., Szolgay, P., de Rugy, A., & Domenger, J. P. (2022). Hybrid FPGA–CPU-Based architecture for object recognition in visual servoing of arm prosthesis. *Journal of Imaging*, *8*(2), 44. https://doi.org/10.3390/jimaging8020044

He, Y., Fukuda, O., Yamaguchi, N., Okumura, H., & Arai, K. (2020). Novel control scheme for prosthetic hands through spatial understanding. *International Journal of Advanced Computer Science and Applications*, *11*(10). https://dx.doi.org/10.14569/IJACSA.2020.0111088

He, Y., Shima, R., Fukuda, O., Bu, N., Yamaguchi, N., & Okumura, H. (2019). Development of distributed control system for vision-based myoelectric prosthetic hand. *IEEE Access*, *7*, 54542-54549. https://doi.org/10.1109/ACCESS.2019.2911968

Huang, S., & Wu, H. (2021). Texture recognition based on perception data from a bionic tactile sensor. *Sensors*, *21*(15), 5224. https://doi.org/10.3390/s21155224

Hurot, C., Scaramozzino, N., Buhot, A., & Hou, Y. (2020). Bio-inspired strategies for improving the selectivity and sensitivity of artificial noses: A review. *Sensors*, *20*(6), 1803. https://doi.org/10.3390/s20061803

James, J. W., Pestell, N., & Lepora, N. F. (2018). Slip detection with a biomimetic tactile sensor. *IEEE Robotics and Automation Letters*, *3*(4), 3340-3346. http://dx.doi.org/10.1109/LRA.2018.2852797

Jiang, N., Chen, C., He, J., Meng, J., Pan, L., Su, S., & Zhu, X. (2023). Bio-robotics research for non-invasive myoelectric neural interfaces for upper-limb prosthetic control: A 10-year perspective review. *National Science Review*, *10*(5), nwad048. https://doi.org/10.1093/nsr/nwad048

Józwik, J., Zawada-Michałowska, M., Kulisz, M., Tomiło, P., Barszcz, M., Pieśko, P., Leleń, M., & Cybul, K. (2024). Modeling the optimal measurement time with a probe on the machine tool using machine learning methods. *Applied Computer Science, 20*(2), 43–59. https://doi.org/10.35784/acs-2024-15

Kim, S., Brady, J., Al-Badani, F., Yu, S., Hart, J., Jung, S., Tran, T.-T., & Myung, N. V. (2021). Nanoengineering approaches toward artificial nose. *Frontiers in Chemistry, 9*, 629329. https://doi.org/10.3389/fchem.2021.629329

Kim, S., Choi, Y. Y., Kim, T., Kim, Y. M., Ho, D. H., Choi, Y. J., Roe, D. G., Lee, J.-H., Park, J., Choi, J.-W., Kim, J. W., Park, J.-H. Jo, S. B., Moon, H. C., Jeong, S., & Cho, J. H. (2022). A biomimetic ocular prosthesis system: emulating autonomic pupil and corneal reflections. *Nature communications*, *13*(1), 6760. https://doi.org/10.1038/s41467-022-34448-6

Lan, N., Zhang, J., Zhang, Z., Chou, C. H., Rymer, W. Z., Niu, C. M., & Fang, P. (2023). Biorealistic hand prosthesis with compliance control and noninvasive somatotopic sensory feedback. *Progress in Biomedical Engineering*, *5*, 023001. https://doi.org/10.1088/2516-1091/acc625

Lee, K. H., Min, J. Y., & Byun, S. (2021). Electromyogram-based classification of hand and finger gestures using artificial neural networks. *Sensors*, *22*(1), 225. https://doi.org/10.3390/s22010225

Lin, Z., Zheng, H., Lu, Y., Zhang, J., Chai, G., & Zuo, G. (2024). Object surface roughness/texture recognition using machine vision enables for human-machine haptic interaction. *Frontiers in Computer Science*, *6*, 1401560. https://doi.org/10.3389/fcomp.2024.1401560

Llop-Harillo, I., Pérez-González, A., Starke, J., & Asfour, T. (2019). The anthropomorphic hand assessment protocol (AHAP). *Robotics and Autonomous Systems*, *121*, 103259. https://doi.org/10.1016/j.robot.2019.103259

Luo, Z., Bi, Y., Yang, X., Li, Y., Yu, S., Wu, M., & Ye, Q. (2024). Enhanced YOLOv5s+ DeepSORT method for highway vehicle speed detection and multi-sensor verification. *Frontiers in Physics*, *12*, 1371320. https://doi.org/10.3389/fphy.2024.1371320

Machrowska, A., Karpiński, R., Maciejewski, M., Jonak, J., & Krakowski, P. (2024). Application of EEMD-DFA algorithms and ANN classification for detection of knee osteoarthritis using vibroarthrography. *Applied Computer Science*, *20*(2), 90–108. https://doi.org/10.35784/acs-2024-18

Mereu, F., Leone, F., Gentile, C., Cordella, F., Gruppioni, E., & Zollo, L. (2021). Control strategies and performance assessment of upper-limb TMR prostheses: A review. *Sensors*, *21*(6), 1953. https://doi.org/10.3390/s21061953

Ortiz-Catalan, M., Zbinden, J., Millenaar, J., D'Accolti, D., Controzzi, M., Clemente, F., Cappello, L., Earley, E. J., Mastinu, E., Kolankowska, J., Munoz-Novoa, M., Jönsson, S., Cipriani, C., Sassu, P., & Brånemark, R. (2023). A highly integrated bionic hand with neural control and feedback for use in daily life. *Science robotics*, *8*(83), eadf7360. https://doi.org/10.1126/scirobotics.adf7360

Roy, A. C., Hossin, K., Uddin, M. P., Al Mamun, M. A., Afjal, M. I., & Nitu, A. M. (2018). Detection and classification of geometric shape objects for industrial applications. *Advancement in Image Processing and Pattern Recognition*, *1*(2), 11-19.

Said, S., Boulkaibet, I., Sheikh, M., Karar, A. S., Alkork, S., & Naït-Ali, A. (2020). Machine-learning-based muscle control of a 3D-printed bionic arm. *Sensors*, *20*(11), 3144. https://doi.org/10.3390/s20113144

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. *IEEE conference on computer vision and pattern recognition* (pp. 4510-4520). IEEE. http://dx.doi.org/10.1109/CVPR.2018.00474

Sensinger, J. W., & Dosen, S. (2020). A review of sensory feedback in upper-limb prostheses from the perspective of human motor control. *Frontiers in neuroscience*, *14*, 345. https://doi.org/10.3389/fnins.2020.00345

Sharma, A., Roo, J. S., & Steimle, J. (2019). Grasping microgestures: Eliciting single-hand microgestures for handheld objects. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (pp. 1-13). Association for Computing Machinery. https://doi.org/10.1145/3290605.3300632

Shi, X., Wang, Y., & Qin, L. (2023). Surface recognition with a Bio-inspired Tactile Fingertip. *IEEE Sensors Journal*, *23*(16), 18842–18855. http://dx.doi.org/10.1109/JSEN.2023.3291720

Tran, M., Gabert, L., Hood, S., & Lenzi, T. (2022). A lightweight robotic leg prosthesis replicating the biomechanics of the knee, ankle, and toe joint. *Science robotics*, *7*(72), eabo3996. https://doi.org/10.1126/scirobotics.abo3996

Vásquez, A., & Perdereau, V. (2017). Proprioceptive shape signatures for object manipulation and recognition purposes in a robotic hand. *Robotics and Autonomous Systems*, *98*, 135-146. https://doi.org/10.1016/j.robot.2017.06.001

Vonsevych, K. (2024). Myographic system of the bionic wrist with surface type identification. In M. Bezuglyi, N. Bouraou, V. Mykytenko, G. Tymchyk, & A. Zaporozhets (Eds.), *Advanced System Development Technologies I* (Vol. 511, pp. 193–228). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-44347-3_6

Vonsevych, K., Goethel, M.F., & Mrozowski, J., Awrejcewicz, J., & Bezuglyi, M. (2019). Fingers movements control system based on artificial neural network model. *Radioelectronics and Communications Systems*, *62*, 23–33. https://doi.org/10.3103/S0735272719010047

Vonsevych, K. P., Bezuglyi, M. A., & Prytula, O. A. (2019). Optical feedback based on photometry by ellipsoidal reflector in bionic fingers application. *KPI Science News*, *3*, 63–72. https://doi.org/10.20535/kpi-sn.2019.3.175785

Wang, J., & Li, S. (2021). Grasp detection via visual rotation object detection and point cloud spatial feature scoring. *International Journal of Advanced Robotic Systems*, *18*(6). http://dx.doi.org/10.1177/17298814211055577

Wei, A. H., & Chen, B. Y. (2020). Robotic object recognition and grasping with a natural background. *International Journal of Advanced Robotic Systems*, *17*(2). http://dx.doi.org/10.1177/1729881420921102

Wijk, U., Björkman, A., Carlsson, I., Kristjansdottir, F., Mrkonjic, A., Rosén, B., & Antfolk, C. (2024). A bionic hand vs. a replanted hand. Journal of Rehabilitation medicine. *Clinical Communications*, *7*, 24845. https://doi.org/10.2340/jrmcc.v7.24854

Yacouby, R., & Axman, D. (2020). Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. *Proceedings of the first workshop on evaluation and comparison of NLP systems* (pp. 79-91). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.eval4nlp-1.9

Zbinden, J., Molin, J., & Ortiz-Catalan, M. (2024). Deep learning for enhanced prosthetic control: Real-time motor intent decoding for simultaneous control of artificial limbs. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *32*, 1177-1186. https://doi.org/10.1109/tnsre.2024.3371896

Zeng, B., Liu, H., Song, H., Zhao, Z., Fan, S., Jiang, L., Liu, Y., Yu, Z., Zhu, X., Chen, J., & Zhang, T. (2022). Design and slip prevention control of a multi-sensory anthropomorphic prosthetic hand. *Industrial Robot: The International Journal of Robotics Research and Application*, *49(2)*, 289-300. http://dx.doi.org/10.1108/IR-07-2021-0133

Zhang, J., Li, M., Feng, Y., & Yang, C. (2020). Robotic grasp detection based on image processing and random forest. *Multimedia Tools and Applications*, *79*, 2427-2446. https://doi.org/10.1007/s11042-019-08302-9