Vision-Based Soft Mobile Robot Inspired by Silkworm Body and Movement Behavior

Ali A. Abed 1,*, Alaa Al-Ibadi 2, Issa A. Abed 3

1, 2 Department of Computer Engineering, University of Basrah, Basrah, Iraq
3 Basrah Engineering Technical College, Southern Technical University, Basrah, Iraq

Email: 1 ali.abed@uobasrah.edu.iq, 2 alaa.abdulhassan@uobasrah.edu.iq, 3 issaahmedabed@stu.edu.iq

*Corresponding Author

Abstract—Designing an inexpensive, low-noise, safe for individual, mobile robot with an efficient vision system represents a challenge. This paper proposes a soft mobile robot inspired by the silkworm body structure and moving behavior. Two identical pneumatic artificial muscles (PAM) have been used to design the body of the robot by sewing the PAMs longitudinally. The proposed robot moves forward, left, and right in steps depending on the relative contraction ratio of the actuators. The connection between the two artificial muscles gives the steering performance at different air pressures of each PAM. A camera (eye) integrated into the proposed soft robot helps it to control its motion and direction. The silkworm soft robot detects a specific object and tracks it continuously. The proposed vision system is used to help with automatic tracking based on deep learning platforms with real-time live IR camera. The object detection platform, named, YOLOv3 is used effectively to solve the challenge of detecting high-speed tiny objects like Tennis balls. The model is trained with a dataset consisting of images of Tennis balls. The work is simulated with Google Colab and then tested in real-time on an embedded device mated with a fast GPU called Jetson Nano development kit. The presented object follower robot is cheap, fast-tracking, and friendly to the environment. The system reaches a 99% accuracy rate during training and testing. Validation results are obtained and recorded to prove the effectiveness of this novel silkworm soft robot. The research contribution is designing and implementing a soft mobile robot with an effective vision system.

Keywords—Silkworm; Pneumatic Artificial Muscle; Jetson Nano; Camera Pi, Object Detection; YOLOv3; Google Colab.

I. INTRODUCTION

A new generation of robotic systems has been developed recently by utilizing soft materials to create the actuators and/ or the body of the machines. The term “Soft robotic” expresses that fact. Soft robots are either manufactured entirely of soft materials such as rubber, silicon, and polyethylene [1], [2] or, by a combination of rigid parts and soft textures. Numerous types of soft robots can be designed using these compliance materials including robot arms, soft grippers, humanoid robots, and mobile robots [3].

Pneumatic artificial muscle (PAM) represents the main part of the many recently designed soft robots. The PAM shows up as a contraction-like muscle [4], extension muscle [5], the contraction-based bending actuator [6], the extension-based bending actuator [7], twisting actuator, and circular actuator [8]. A single type or combination of these actuators can be used to design the robot body and the required movements. Numerous prototypes have been designed newly either for industry or medical areas due to the advantages of these systems such the lightweight, compliance, low cost, high force, high capacity of deformation, safe to wear by humans, and environment beneficial [7], [9]–[11].

The mobile robot is one of the most important types of robotic systems because of its wide use, especially during the Covid19 pandemic and the isolation requirements [12], [13]. These robots can be seen at health centers, restaurants, factories, or during disasters [14], [15].

Mostly, the mobile robot movement and shape is constructed by using wheels and legs, or it is inspired by spherical, snake, and worm. Hybrid design of two or more inspiration is exist [16]–[20]. Each type of this robot moving technique has several advantages and limitations. The main feature of the wheeled mobile robots is the high velocity [21]–[23], while both snake and worm robots consider more suitable for narrow paths [24]. A soft floor is the primary work area for sphere-shaped machine [25], [26]. A hybrid mobile robot obviously gives at least two various positives. On the other hand, each type shows limitations either because of the way it moves or the environment the robot serves. Moreover, most transporting robots fall off from the restriction of power supply [27].

Robots that have been made of rigid materials are heavy, risky for human being, costly, require a wide workspace, and they cannot adapt to the nature [23], [28], [29]. Applications can be selected for each type of moving robot depending on the construction, materials, and environment. The mobile robot is created for investigation by [30] and [31] by implanting numerous sensors. While the investigation and monitoring are extensively applied applications [31]–[34], several tasks challenge the researchers such as obstacle prevention [35], [36], path planning [37]–[39] localization approaches [28][40], and the requires of adaptation at the limited paths.

Several mobile robots have been made with soft materials or hybrid structures including both rigid and soft. A tripod rigid mobile robot is designed by [41] with a soft membrane in the bottom. A fluidic elastomer actuator is used to construct a mobile robot with bending behavior to establish forward moving [42]. Numerous soft mobile robots have been inspired by biology are presented by researchers, such as the Caterpillar-inspired robot by [43], the quadruped legs robot [44], Worm locomotion inspiration [45]–[47], Snake-inspired soft robots [48]–[51]. On the other hand, many
robots are made by using shape memory alloys and dielectric actuators such as [52]–[55].

Nowadays, computer vision is becoming more important and has been widely applied in several industrial and service sectors, especially in the field of robot vision. One of the critical and beneficial tasks for vision-based systems is to navigate mobile robots safely in different types of environments. Navigation should avoid obstacles and collisions in small environments using sensors or cameras for the sake of path finding [56]–[58]. However, the sensors may be inaccurate due to interruption of the transmission signal caused by some disturbance from the surrounding environments resulting in incorrect data reception and leading to a false decision [59], [60]. Therefore, camera is a high-performance alternative to replace the classical sensors used in mobile robots [61]–[64].

To involve the vision system with the soft robot motion system, a deep learning platform called YOLOv3 is adapted to create an algorithm for computing the value of motion parameters. Object detection and tracking are considered in relation to the information extracted from the interpreted data of the image. The motion parameter values are required to send a signal command to the pneumatic actuators of the robot for its corresponding movement to follow the target object. The proposed vision-guided navigation system based on YOLOv3 provides more accuracy and precision for the soft mobile robot to do its tasks in an indoor environment. Additionally, obstacle avoidance is internally included to make the soft robot moves smoothly in any dynamic environment with the capability of collision avoidance. Sports ball detection is one of the most challenging tasks because of its high speed and tiny size, which is an essential part of a sports ball robot. In this paper, the detection and tracking of Tennis balls are achieved despite that the Tennis ball is tiny and moving very fast.

Transfer learning indicates a model that has been developed for certain tasks and use again once adapted as the initial point for a prototype on a associated principle [65]. This technique is applied in this project. It is a wide use strategy in deep learning where pre-trained models are implemented as the early point in visualization projects. This leads to decreasing the time and computational sources required to develop neural network models from scratch.

Several past works are presented based on YOLOv3 AI (Artificial Intelligent) deep learning platforms but for other applications such as face recognition, face mask detection, and surveillance systems [66]–[71]. To the knowledge of the authors, this is the first time to combine YOLOv3 deep learning algorithm as a vision system with soft mobile robots. The main contributions of this article are the design inspiration of the mobile robot and the vision system that has been combined.

The sections of the paper are organized as follows. Section 2 contains the description of the natural silkworm and its main parts. Section 3 details the main components of the robot design. Section 4 explains the principle of operation for the proposed silkworm soft robot. In section 5, the vision system designed for the proposed soft robot is detailed. Experimental simulation and real-time results are displayed. In section 6, the main conclusions have been discussed and some future works are suggested.

II. SILKWORM

Silkworm is a slow and weak walking insect with a segmented body and three normal legs in addition to rear leg-like parts [54]–[57]. Fig.1 shows a photo of the silkworm in movement posture.

![Fig. 1. A photo of Silkworm at movement posture](image)

The segmented body of the silkworm gives the ability to steer to the right or left. The worm uses the front three pairs of its legs to move forward with the help of the rear underbody parts. The structure and the movement parts give the inspiration to design a mobile robot with an efficient vision system by using two contraction PAM.

III. ROBOT DESIGN

Two identical contraction muscles are sewn in parallel to construct the body. The dimension and the specification of each actuator are listed in Table I.

<table>
<thead>
<tr>
<th>Initial Length $L_a$ (cm)</th>
<th>Initial Diameter $D_a$ (cm)</th>
<th>Stiffness (N/m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 cm</td>
<td>3.5 cm</td>
<td>545</td>
</tr>
</tbody>
</table>

The contraction ratio and the expansion of the diameter depending on the specifications of the expandable braided mesh. Table II shows these specifications.

<table>
<thead>
<tr>
<th>Initial Mesh Length (cm)</th>
<th>Initial Diameter $D_a$ (cm)</th>
<th>Maximum Diameter (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 cm</td>
<td>3.5 cm</td>
<td>5.2</td>
</tr>
</tbody>
</table>

Fig.2 illustrates the structure of the contraction PAM. The basic behavior of this actuator is the shrinking at pressurized condition up to 30% of its initial length [4], [19]. Fig. 3 gives a block diagram of the design procedure.

![Fig. 2. The structure of the contraction PAM](image)
braided angle $\theta$ starts at a value less than the threshold of 54.7° [11], the PAM stop contracting either because the angle reaches the threshold or the maximum mesh diameter is achieved [58]. Fig. 4 shows the constructed air muscle of 30 cm in length.

![Diagram](image)

**Fig. 3.** The flow chart of the design procedure

Fig. 3 shows that the first step is looking to living creatures and select one that gives an idea of moving and turning with a suitable vision system. While the second step to design and construct the closest shape, then test the behavior of the proposed robot. Select the proper design and material leads to finish the process or repeat step two again.

In this article we decided to use the PAM to design the robot for their advantages in spite of its slow performance, therefore, the authors searched for a slow motion animals or insects.

![Image](image)

**Fig. 4.** The constructed pneumatic artificial muscle

To specify the performance of the actuators, an experiment is done to show the length and tensile force by applying an air pressure through a solenoid valve (Matrix 3/3 by BIBUS Shanghai Mechanical, Ltd, China). The air pressure and the length of the actuator are measured by air pressure sensor and ultrasound sensors respectively. On the other hand, a digital weight scale is attached to the free end of the PAM to show the tensile force. Fig. 5 illustrates the experiment rig.

Fig. 6 (a) shows the length of the air muscle against the air pressure and Fig. 6 (b) gives the contraction force. The air pressure is applied from 0 to 500 kPa to show the behavior, and the 500 kPa is considered as a maximum safe working pressure.

![Fig. 5. The experiment rigs](image)

![Fig. 6. The performance of the PAM. (a) the length of the actuator. (b) the tensile force](image)

Instead of using segments to design the proposed robot, the two contraction actuators give the exact behavior of silkworm. Fig. 7 illustrates the design and the actual robot.
Fig. 7. The proposed silkworm robot. (a) The design schemes. (b) The photograph

Fig. 7 (a) shows that two identical contraction actuators are used in parallel to design the robot. A high-tension thread is used to connect the PAMs longitudinally to establish a bending to the direction of the actuator that carry higher air pressure. On the other hand, Fig. 7 (b) gives a photograph of the designed robot.

IV. THE OPERATION OF THE SILKWORM ROBOT

The basic operation of the contraction PAM is the shortening in its length and expanding in diameter. The contraction ratio $\varepsilon$ reaches its maximum value at maximum air pressure and/or the braided angle approaches to its threshold. The contraction ratio can be formulated as (1).

$$\varepsilon = \frac{L_0 - L}{L_0}$$  \hspace{1cm} (1)

$L_0$ and $L$ represent the initial and actual length respectively. The minimum length of the actuator is as (2).

$$L = L_0(1 - \varepsilon)$$  \hspace{1cm} (2)

While the maximum contraction ratio considers to be 30%, the minimum length of the air muscle is 70% of its initial length. As a result, the maximum length shrinking of the proposed robot is 9 cm. For safe operation pressure, a 250 kPa is assigned as a forward moving air pressure. This gives about 7 cm step at full cycle of filling and venting processes.

Since the two actuators are connected together and maintain forward movement, an equal air pressure needs to apply simultaneously at both PAMs. However, the contraction occurs towards the center of the actuator if there are no restrictions at both ends. To establish the forward moving single direction wheels are used. Bicycles freewheel of 5 cm in diameter is used as shown in Fig. 8.

The bike structure is significantly increasing because of various purposes not only for transports. Numerous GYM kits are types of bicycle that needs freewheel element, and numerous health care machines consist of bicycle components too [59], [60].

The freewheel designed to turn easily for single path. This performance has been developed in this article to maintain the forward movement.

V. VISION SYSTEM

A. YOLOv3

YOLOv3 (You Only Look Once) [79] object detection model suggested in this article is created on YOLOv1 recognition model. Various adjustments to the loss task are done ahead for further strong feature extractor network producing a multi-object finding procedure. Consequently, this platform could recognize large types of big or tiny objects, ranging in number from 1 to 10. Additionally, YOLOv3 is fast and enables short inference time with high FPS (Frame Per Second) on GPU (Graphical Processing Unit) edge devices [80]. Hence, image classification network developed further developed as compared to regular deep stacks of layers of the earlier forms of YOLO. Consequently, the present extractor of this design has effectively been expanded from 19 levels (in YOLOv1) to 53 levels (in YOLOv3).

B. Dataset and Training

The first step in obtaining an object follower soft robot is the objects’ images collection. A Tennis ball custom dataset [81] is adopted. This dataset is a category from “Google Open Images” that has about 9M images marked with image-level tags, object bounding frames, item segmentation masks, and graphical connections. It includes a total of 16M bounding boxes for 600 object classes on 1.9M images with object location annotations. In this work, the selected class has more than 700 images trained using Darknet-53 with YOLOv3 for 4000 epochs. Training/testing is performed with 75/25% of the total data respectively. Google Colab, a powerful tool and free Jupyter notebook environment running in the cloud, is firstly used for training. It does not require a setup and supports most popular machine learning libraries need to be loaded in the notebook. It is used instead of our local computer to decrease training time. Model storage for future testing and validation on Jetson Nano fast GPU is performed. Mathematical equations for evaluating and measuring training metrics such as Accuracy, Precision, Recall, and F1-Score are found in [82].
The precision and the cost of the suggested model are calculated 20 times through training on Google Colab, although 10 values (for short) have been indicated in Table III. This table is rather adequate to demonstrate that precision raises while loss declines until achieving steady state. Table IV shows the evaluation outcomes of the suggested model throughout the second phase.

### TABLE III. LOSS/ACCURACY VS. EPOCHS

<table>
<thead>
<tr>
<th>Epoch</th>
<th>Train loss</th>
<th>Val loss</th>
<th>Train accuracy</th>
<th>Val accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.34</td>
<td>0.14</td>
<td>0.87</td>
<td>0.97</td>
</tr>
<tr>
<td>3</td>
<td>0.12</td>
<td>0.07</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td>5</td>
<td>0.07</td>
<td>0.07</td>
<td>0.97</td>
<td>0.98</td>
</tr>
<tr>
<td>7</td>
<td>0.07</td>
<td>0.06</td>
<td>0.97</td>
<td>0.99</td>
</tr>
<tr>
<td>9</td>
<td>0.06</td>
<td>0.05</td>
<td>0.98</td>
<td>0.99</td>
</tr>
<tr>
<td>11</td>
<td>0.05</td>
<td>0.05</td>
<td>0.98</td>
<td>1.0</td>
</tr>
<tr>
<td>13</td>
<td>0.04</td>
<td>0.05</td>
<td>0.98</td>
<td>1.0</td>
</tr>
<tr>
<td>15</td>
<td>0.04</td>
<td>0.04</td>
<td>0.99</td>
<td>1.0</td>
</tr>
<tr>
<td>17</td>
<td>0.04</td>
<td>0.04</td>
<td>0.99</td>
<td>1.0</td>
</tr>
<tr>
<td>19</td>
<td>0.03</td>
<td>0.03</td>
<td>0.99</td>
<td>1.0</td>
</tr>
</tbody>
</table>

### TABLE IV. THE PROPOSED MODEL EVALUATION

<table>
<thead>
<tr>
<th>Ball</th>
<th>Recall</th>
<th>F1-score</th>
<th>Precision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tennis</td>
<td>100%</td>
<td>99%</td>
<td>99%</td>
</tr>
<tr>
<td>MA</td>
<td>98%</td>
<td>98%</td>
<td>99%</td>
</tr>
<tr>
<td>WA</td>
<td>98%</td>
<td>98%</td>
<td>99%</td>
</tr>
<tr>
<td>Accuracy</td>
<td></td>
<td>99%</td>
<td></td>
</tr>
</tbody>
</table>

The validation and testing results of the proposed deep learning model is shown in Fig.9.

Additional advantage of the improved technique is its ability of demonstrating various Tennis balls (more than 10) in a single scene although follow a specific ball depending on its prespecified color. Hence, the technique be able to be utilized simply in any busy zone to distinguish a specific-colored ball.

### C. Jetson Nano Kit and Camera Pi

Since the computational processing of the CNN (Convolutional Neural Network) network architecture of YOLOv3 is high, the hardware microcontroller has a hard constraint in that it should involve high GPU, CPU, and RAM memory. The adopted microcontroller board must support parallel programming to enable several threads to be processed concurrently. Therefore, the powerful NVIDIA Jetson Nano 4GB shown in Fig. 10 is chosen since it has 128 CUDA cores GPU and a quad-core ARM Cortex-A57 high-performance CPU.

The software implements and platforms used to assist Jetson Nano are CUDA 11.2, Cudnn 8.0, TensorRT 7.1.3. TensorRT is a deep learning inference SDK at outstanding behavior that includes an implication optimizer and a runtime for deep learning purposes at minimum expectancy and superior throughput. Additionally, Python and OpenCV are previously installed on the board. Moreover, Virtual Network Computing (VNC) server is mounted on the board to observe its Linux environment on the laptop’s monitor.

Complementary to the Jetson board, an IR Camera Pi module 2 (shown in Fig.10) is plugged-in to capture real-time live video streaming. The camera is fixed on a servo motor to get a coverage area of 180°.

The YOLOv3 platform is then applied along with the input images to the bounding box and labeling to monitor the recognized Tennis ball with their labels and recognition accuracies.

Fig. 10. NVIDIA Jetson Nano 4GB with an IR camera

### D. Final Setup

The suggested improvement kit with its fittings is applied to employ the trained model and initial real-time live video capturing and Tennis ball recognition. The categorization model will retrieve balls since video streams as required. Real ball image is positioned into the hidden CNN as an input. The output of the CNN is a “Tennis ball” decision with its bounding box and validation accuracy.

The block diagram of the whole hardware system demonstrating the association with the different hardware elements of the soft mobile robot with its vision system is shown in Fig. 11. The real image of the overall soft robot is shown in Fig. 12.

Fig. 11. Overall block diagram of the silkworm soft robot
All physical components can be carried on the back of the silkworm (as shown in Fig. 12) or may be fixed on a specified location and connected with silkworm via thin plastic hose. 

E. Real-time results

After the training process has been completed, the deep learning model is downloaded into the Jetson Nano kit to start real-time testing. The servo motor of the camera starts to turn left and right searching for the real Tennis ball used as a leader. The robot detects the ball and keep track it until reaching its goal point. The real-time detection and tracking result are realized with a confidence of 99% and FPS of 1.394 as shown in Fig. 13. Table V lists a comparison between the proposed robot and its vision system with a literature.

![Fig. 12. Real images of the overall silkworm soft robot](image1)

![Fig. 13. Detection and tracking results of yolov3-based soft robot using Jetson Nano](image2)

<table>
<thead>
<tr>
<th>paper</th>
<th>Body</th>
<th>Materials</th>
<th>Actuators</th>
<th>Vision system</th>
</tr>
</thead>
<tbody>
<tr>
<td>[41]</td>
<td>Rigid</td>
<td>Hybrid</td>
<td>Vibration</td>
<td>None</td>
</tr>
<tr>
<td>[42]</td>
<td>Soft</td>
<td>Elastomer Films</td>
<td>FEA</td>
<td>None</td>
</tr>
<tr>
<td>[43]</td>
<td>Soft</td>
<td>DEA</td>
<td>DEA</td>
<td>None</td>
</tr>
<tr>
<td>[44]</td>
<td>Soft</td>
<td>Silicon</td>
<td>Pneumatic</td>
<td>None</td>
</tr>
<tr>
<td>[45]</td>
<td>Rigid</td>
<td>plastic</td>
<td>Servo, Suction Cup</td>
<td>None</td>
</tr>
<tr>
<td>[46]</td>
<td>Soft</td>
<td>PAM</td>
<td>PAM</td>
<td>None</td>
</tr>
<tr>
<td>[47]</td>
<td>Soft</td>
<td>Silica Gel</td>
<td>SMA</td>
<td>None</td>
</tr>
<tr>
<td>[48]</td>
<td>Rigid</td>
<td>Metals</td>
<td>Tendons</td>
<td>None</td>
</tr>
<tr>
<td>[50]</td>
<td>Soft</td>
<td>gelatin, glycerol, and water</td>
<td>magnet</td>
<td>None</td>
</tr>
</tbody>
</table>

Table V. The Comparison between the proposed robot and the literature

A wide range of actuation systems also have been utilized. The rigid robots show vibration, servo motors, and tendons as an actuating. The soft robots use fluidic elastomer actuators (FEA), DEA, pneumatic, PAM, shape memory alloy (SMA), and magnet respectively.

The main advantage of the proposed system is the efficient vision system that have been used in comparison with no exiting vision systems. The body of the presented robot is built and actuated by using two contraction PAMs with the aid of four freewheels for keeping the forward moving. The PAMs that have utilized shows high tensile force as illustrated in Fig. 6b. Overall, the presented robot such as others has been inspired by biology, while it has an excellent vision system, actuated by air muscles like [46], and moving in steps such as many worms.

VI. Conclusion

A novel design of object following mobile soft robot in the form of silkworm is presented. The robot is designed and implemented by utilizing two identical contraction pneumatic muscle actuators. The twin air muscles are sewed together longitudinally to get three degree of freedom, forward, left, and right depending on the air pressure in each actuator. The whole system is managed with Jetson Nano Al deep learning board trained with YOLOv3 lightweight model. Consequently, this mobile soft robot might be used in crowded zones. When the system is placed in any zone, it can be configured to capture a live video stream to detect a ball object and tracking it. Our approach outperforms YOLOv3 and achieves exceptional Tennis ball detection performance. The achieved validation findings demonstrated 99.0% precision for training and testing. The obtained processing time using this developed model is about 0.71s. Another developed feature is the number of objects that can be recognized and detected in a single scene is 10. This gives the robot the ability to navigate in a complex dynamic environment without confusion. The developed robot is fast-tracking, low-cost, lower power, and small size suitable for small-sized applications such as sports ball robots.

This article presents a soft mobile robot with an efficient vision system that can be used under different situations and environments. Further research could carry out to modify the mobile robot to move backward in order to provide full movement performance.

REFERENCES

Ali A. Abed, Vision-Based Soft Mobile Robot Inspired by Silkworm Body and Movement Behavior
For a 2D vision system for mobile robotic drilling, 


With a 2D vision system for mobile robotic drilling, 


For a 2D vision system for mobile robotic drilling, 


With a 2D vision system for mobile robotic drilling, 


For a 2D vision system for mobile robotic drilling, 


With a 2D vision system for mobile robotic drilling, 


For a 2D vision system for mobile robotic drilling, 
