Estimation of Contact Force on an Inflatable Structure Using Vision Sensors

António Franco
antonio.franco@tecnico.ulisboa.pt

Instituto Superior Técnico, Universidade de Lisboa, Lisboa, Portugal

November 2019

Abstract

This work presents a study of force estimation on an inflatable structure using computational vision tools. The structure is intended to be used in medical applications to complement a robotic manipulator with passive safety features. The inflatable structure has a truncated conical shape, and was designed to be easily deformed in order to enhance its sensibility to external forces. Its lateral surface was built using polyamide fabric coated with neoprene, a synthetic rubber, in order to make the structure airtight. Both the top and bottom surfaces were made of ethylene-vinyl acetate foam for soft contact surfaces. An image processing algorithm was developed to infer the structure deformation from internal markings captured by a camera placed inside the prototype. The results were then used as input in an artificial neural network that yielded the estimated force being exerted on the structure. The neural network was firstly trained by using data where both the input and output were known and then tested by comparing its performance against a exterior force sensor. By analysing of the performance of the prototype it was concluded that employing it under certain conditions displays satisfactory results, achieving a mean error of only 0.051 N, thus presenting a viable option for measuring force.

Keywords: Soft Robotics, Tactile Sensor, Computational Vision, Artificial Neural Networks

1. Introduction

In the scope of robotics, the acquisition of tactile sensing information is frequently required in order to successfully carry out object manipulation operations. For this requirement there is a extensive arrangement of solutions but only a few respond to this need using flexible structures.

With this necessity in mind this work focuses on the design and manufacture of a compliant robotic link with the capability of estimating the force. For this end computer vision techniques are integrated with artificial neural networks to calculate the contact force applied on the structure.

Soft robotics is a field in robotics that focus on the design and production of robots made from flexible materials. Taking into account their compliant properties, they constitute versatile solutions to accomplish tasks and provide safety when working around humans [1]. Although the topic is relatively new, its firstly application dates back to 1963 with the development of pneumatic actuators to aid polio patients [2]. Despite the resulting prototypes were used as base for following robot designs, this subject only started to grow later on, being highly recognized nowadays by the robotic community. The research in this field is widely spread, from covering the application of this concept using a broad range of compliant materials to the implementation of different types of actuation systems, being pneumatic actuation the most popular one [3]. Different configurations of structures are also analysed, mainly from bio-mimicry, taking inspiration as per example the tentacles of octopuses[4], the trunks of elephants [5] and even caterpillars [6].

Regarding the measurement of the force in the field of soft robotics, one of the possible solutions for this matter is the utilization of tactile sensors placed on the robot, with the cost of adding complexity to the structure.

Tactile sensing technology relates to solutions and equipment capable of measuring information produced from physical interaction with the environment. It is a broad concept that covers the computation of force, torque, and/or surface shape by processing information generated from sensing mechanisms. These can be: detection of the displacement of an elastic component, measurement of pneumatic pressure, piezoelectric effect, piezoresistive effect, capacitive-based sensing, etc [7]. Its first uses dates back to 1973, where a tactile sensor was mounted on an anthropomorphic hand to recognize shapes from the obtained data [8]. Follow-

ing sensing devices were designed for industrial automation but went no further than prototype stages given the wear that they would be subjected to and the fact that computational vision tools constitute a much more simpler solution [9]. Considering this tactile sensing technology took a turn from its initial intended purpose towards service robotics, food processing and medical applications with a strong emphasis on the latter [10].

Recently developed tactile sensors take inspiration from biological structures, namely the morphology of skin mechanoreceptors that convert compressive strain into neural signals [11]. From this, in 2009, a sensor was built with the intent of inferring force from the displacement of internal markers placed on a compliant contact surface [12]. The prototype developed in the scope of this work employs this concept in soft robotics, merging the tactile sensor function into a robotic link.

2. Sensing Methodology

The methodology devised in this works starts with the processing of images acquired from a webcam placed inside the prototype and the resulting data is then used as input in an artificial neural network that infers the applied force.

2.1. Computational Vision

Several computational vision tools and methodologies were employed with the goal of collecting information regarding the deformation of the prototype from images retrieved from a video camera device.

Firstly by placing the camera in a central position relative to the axisymmetrical soft structure, the obtained images compose orthographic projections of the prototype's interior surface. This way, although it is not possible to directly measure the displacement, it allows the acquisition of relevant vision data about the prototype's deformation.

The image processing is initiated by converting the retrieved images from RGB color space to greyscale, where the intensity of each pixel is calculated from a weighted sum. In order to distinguish the painted markers from the remaining structure surface a binarization process was implemented, where a static threshold value was used since the conditions of the region captured by the camera will remain constant.

With the image divided in foreground and background it is possible to identify each connected region as a blob and proceed with a blob analysis to retrieve data associated with the prototype's deformation. In view of this, the centroid of each marker was assessed (1)(2), given that from this parameter the displacement can be calculated, along with the measurement of the eccentricity value (3), to evaluate the structure warping.

$$\text{Centroid X component: } \boldsymbol{C_x} = \frac{\sum_{i=1}^{n} X_i}{n} \tag{1}$$

$$\text{Centroid Y component: } \boldsymbol{C_y} = \frac{\sum_{i=1}^{n} Y_i}{n} \tag{2}$$

Where $\mathbf{n}$ is the number of pixels in the blob, and $\mathbf{X_i}$ and $\mathbf{Y_i}$ are the X and Y components of pixel i, respectively.

$$\mathbf{e} = \sqrt{1 - \frac{b^2}{a^2}} \tag{3}$$

Where $\mathbf{a}$ is the blob's semimajor axle length and $\mathbf{b}$ its semiminor axle dimension.

2.2. Artificial Neural Networks

The resulting data from the image processing is then used as input of an artificial neural network (ANN) that outputs the estimated force value. To achieve this it is necessary to firstly train the net using vision data with the expected force value in order to calculate the weights of each node.

In terms of structure, given the purpose of the task, a feedforward network was designed with 6 nodes in the input layer and 1 node in the output layer. The number of nodes of the inlet layer derive from the information of the 2 captured markers, where each one has 2 values for the centroid (x and y component) and 1 value for the eccentricity.

As mentioned previously, the ANN was trained using sets of inputs with the corresponding desired output. This learning approach is designated as supervised learning and from its implementation the error can be quickly calculated to use as feedback on the quality of obtained solutions. A downside of this method is that the excessive use of training using the same data leads to overfitting and thus reducing the net's accuracy. The computed error is then also utilized to adjust the net's weights if the stopping criteria has not been reached yet.

The modification of weights was done recurring to the Levenberg-Marquardt algorithm considering that it is one of the fastest methods and that it is suitable for the type of network implemented. This process starts with the computation of the jacobian matrix of the loss function (4), a matrix that expresses the variation of the ANN network error from changing the net's weights.

$$J_{i,j} = \frac{\partial e_i}{\partial \omega_j} \tag{4}$$

Where $\mathbf{e_i}$ is the error generated from data point i and $\boldsymbol{\omega}_\mathrm{i}$ the node's input j weight.

From this it is then possible to calculate the weights for the next iteration using the following equation:

$$\omega^{k+1} = \omega^k - [J^T J + \mu J]^{-1} J^T e \tag{5}$$

Where $\boldsymbol{\mu}$ is a damping factor that is decreased after each successful iteration to increase the convergence speed and accuracy.

3. Experimental Setup

3.1. Soft Link

The soft link consists in the central part of the prototype, it is formed by the flexible sleeve and the foam peripherals that will be in contact with the environment. It's also responsible for the harboring of the image acquisition device.

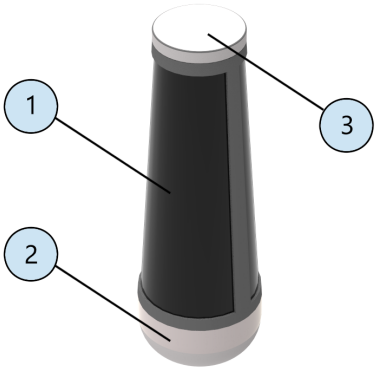


Figure 1: Soft Link, 1 - Flexible Sleeve; 2 - Foam Proximal Extremity; 3 - Foam Distal Extremity

The extremities of the soft link needed to be made from a flexible and smooth material. They can't provide stiffness to the structure, since the softer it is, the higher the sensitivity will be, allowing the measurement of smaller forces. The material must also be airtight to keep the interior of the soft link pressurized. Considering these conditions the best pick is a foam, so ethylene-vinyl acetate foam (EVA foam) was selected.

The distal extremity has a simple design where the main purpose is to have a shape that can fit in the top of the flexible sleeve and provide a good lateral surface that will be in contact with the surroundings. The whole extremity is made from two concentric discs joined together with appropriate glue.

The proximal extremity has the crucial function of serving as a central link, to assemble all different sections of the prototype. It receives the air intake duct, holds the image acquisition device (and its housing) and fits into the flexible sleeve, closing out the flexible link. It is shaped by heating the EVA foam and then pressing it against a flattened semi spherical mold, solidifying after cooling down.

The flexible sleeve, placed in the middle of the soft link defines most of the structure's stiffness and contains the computer vision markers, key elements to the acquisition of data. In respect to the size, by increasing the length the more compliant it will be so an hollow truncated cone with 230 [mm] of length was designed. The material selected was neoprene rubber padded with polyester given that its rubber side compromises as a good painting surface and can be extended without losing its air impermeability. During the manufacturing process, three annuli were painted, equally separated from each other, but with different widths since the farther away from the camera, the smaller the area it will capture. The markers were painted using white acrylic paint on the account that when it dries it has a dull finish and is flexible. In terms of design, they were placed as close to the top of the cone as possible, with the premise that the top section of the sleeve will be subjected to higher deformation therefore easier to detect the application of force.

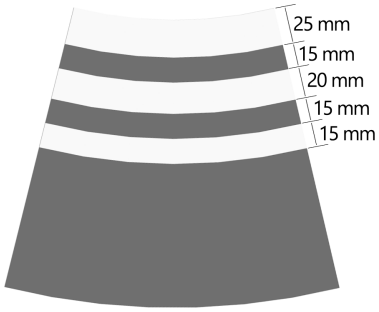


Figure 2: Painted Cone Planification

Finally the lateral edges are joined together using glue and reinforced with a strip of the same fabric, closing out any gaps between the two surfaces.

A Trust Spotlight webcam was used to capture images considering that it is small enough to fit inside the soft link and that it has a light source incorporated, a necessary feature given that otherwise the prototype's interior will be completely dark. It also has a wide field of view, enabling it to fully capture the markers when they are in a central position (null force being exerted on the estimator) and when they are diverted from this position (non null force being exerted).

Source: Trust[1]

Figure 3: Trust Spotlight Webcam

[1] https://www.trust.com/en/product/16428

With the purpose of securing the webcam to the bottom of the base foam extremity while also preventing it from rotating, an housing for it was designed. Due to its complex geometry the best option was to manufacture this component recurring to 3D printing technology, where PLA (Polyactic Acid) was selected as printing material.

(a) (b)

Figure 4: Webcamera Cover, isometric view (a) and bottom view (b)

3.2. Prototype Handle

This component provides a solid surface to help on the manipulation of the prototype, by having a small squared cross section that can be easily grabbed by a robotic arm gripper or by a human hand. It also acts as foundation to the soft link, proving a rigid base to where it will be fastened using bolts. In similarity to the camera's housing, given its format this constituent was also 3D printed, using the same technique and material.

Figure 5: Prototype Handle

3.3. Pneumatic Circuit

In order to pressurize the soft link an elementary pneumatic circuit was devised. This assemblage was not only to inflate but also to compensate for leaks by continuously pumping air into the designed soft structure. This circuit is composed of a gauge pressure sensor (a), DC air compressor (b) and the soft link(c - represented by an air reservoir), all connected with 4mm tubing.

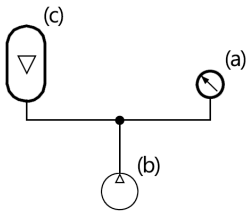
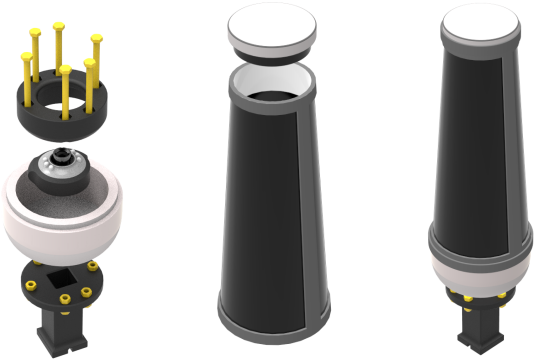


Figure 6: Pneumatic Circuit Representation

3.4. Prototype Assembly

The assembly started by fitting the webcam into its housing and then fastening this set to the foam proximal extremity using M5 bolts (7a). Following this, the foam distal extremity was mounted into the flexible sleeve, using contact glue to bind the two components (7b). Subsequently this previously assembled set was then fitted into the foam proximal extremity, closing out the soft link (7c). To finish the pneumatic circuit was connected and all edges between the different components were covered with silicone to serve as a sealant against air leaks.

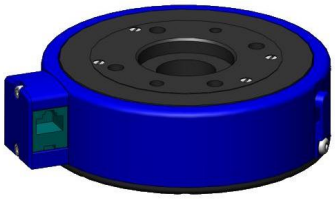


(a) (b) (c)

Figure 7: Prototype Assembly

3.5. Force Sensor

With the intention of obtaining force data to subsequently train the neural network it was necessary to set a force sensor. The employed sensor was a JR3 multi-axis load cell (8), capable of measuring both forces and torques along three orthogonal axes (x, y and z) through strain gauge signals. The origin point used to calculate the torques is located at the center of the sensor and the z axis is aligned with its central axis.



Source: JR3[2]

Figure 8: JR3 Force Sensor CAD

Its important that the force data and vision data are processed at the same time to make sure that they are synchronized. This is crucial to make sure that the input and output data for the neural network match correctly otherwise it will be trained with erroneous values and hinder its accuracy.

[2]https://www.jr3.com/products/force-torque-sensors

4. Results

4.1. Experimental Data

Experimental data was collected using the MATLAB® tool Simulink®. This software was running in two separate computers connected thought a LAN connection, were one retrieved force data (target computer) and would then send it to the computer responsible for collecting the vision data (host computer).

After having both computers configured, the experimental setup was then prepared to acquire the force data collected from the JR3 sensor.

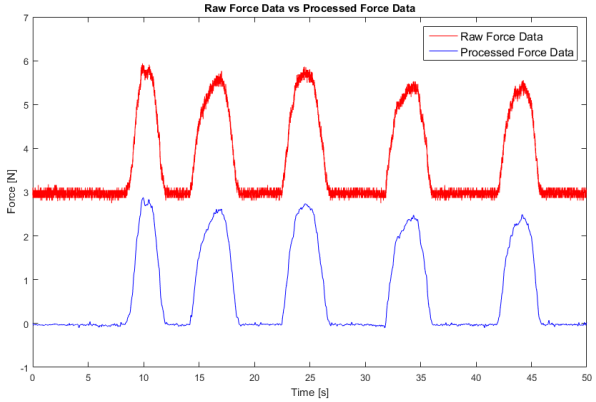


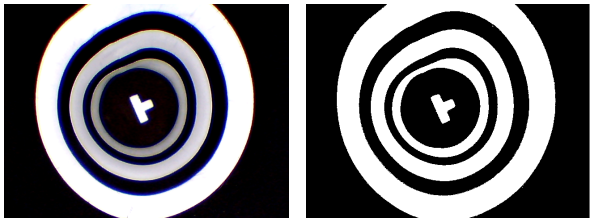
Figure 9: Raw vs Processed Force Data

Given the high frequency noise present in the force values, as seen from figure 9, a lowpass FIR filter was applied. Considering that filtering in real-time introduces a delay in the signal, to avoid desynchronization of the force data in comparison with the vision data, this filter was applied only after the force values were collected. To finalize this process, the bias was removed from the signal by firstly estimating its value and then subtracting it to the filtered data.

With respect to the acquisition of vision data it was crucial that the prototype was appropriately prepared by having the internal pressure stable after the air compressor was switched on. For the performed trials the compressor was running with 3.5V and the pressure stabilized at 4kPa. At the start of each trial, vision data was retrieved with null force applied to the prototype. This data was posteriorly used as reference to calculate the variation in the centroid and eccentricity values.

The prototype's distal extremity was then slowly pressed against the force sensor while maintaining a constant downwards inclination of 30° with the horizontal plane. It was important to keep this slope stable to prevent interference with the uniformity of the trials conditions. Furthermore the speed at which the experiment was conducted is not important in view of the fact that the results don't account for dynamical effects, this is because each data point was processed by the neural network individually.

Figure 10 presents the original camera's output and after thresholding.



(a) (b)

Figure 10: Camera's output, original (a) and binarized (b)

The binarized image was then processed using blob analysis to retrieve the statistical information of the connected regions, namely the centroids and the eccentricities. Only the blobs of the two smaller annuli were analysed considering that the third one isn't captured entirely by the camera, avoiding the usage of inaccurate data.

Finally the reference data was subtracted to the obtained values in order to calculate the statistical variation triggered by the deformation of the prototype.

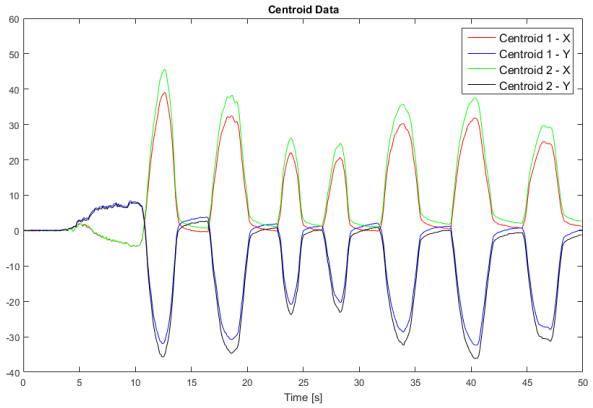


Figure 11: Centroid Data

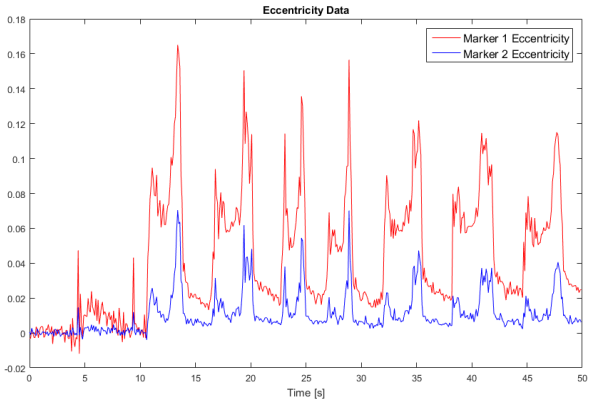


Figure 12: Eccentricity Data

4.2. Artificial Neural Networks Performance

Given the prototype's structure, its behaviour will not be axisymmetrical so its necessary to explore two scenarios: with and without the rotation of the prototype when collecting the data.

The first tackled issue was the force estimation when the direction of the applied force is kept constant ensuring that the relation between displacement-force doesn't vary. This corresponded to trials where the prototype was pressed against the force sensor without rotation. Figure 13 portrays this operation, where **F** represents the force acting on the prototype.

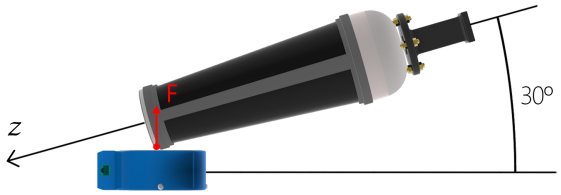


Figure 13: Prototype Rotation

From the perspective of the resulting vision data this can be reduced to a one-dimensional problem, where the absolute value of the displacement of the centroids would be enough to calculate the force being applied, although with less accurary. Considering this, a one-layered feedforward neural network is enough to solve this problem if designed with enough neurons.

The number of neurons in the ANN was calculated by creating batches of net's with the same structure, computing their error and finally, evaluating which design had the best performance. Taking into account that the measured performance is actually the mean squared error, then the lower the performance value, the better the net is.

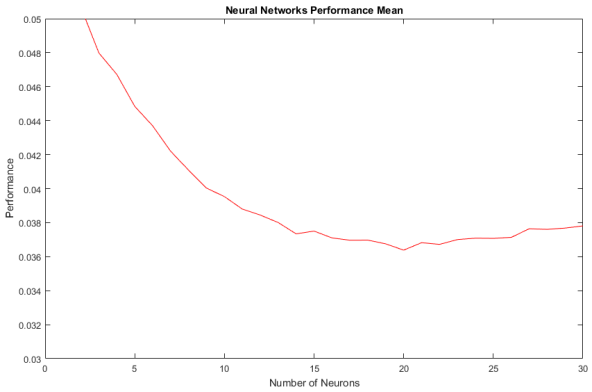


Figure 14: ANN Performance Mean

From 14 it can concluded that the net's with best performances have between **15-25** neurons, with the minimum at **20** neurons. The next figure presents the results from implementation of the best obtained net:

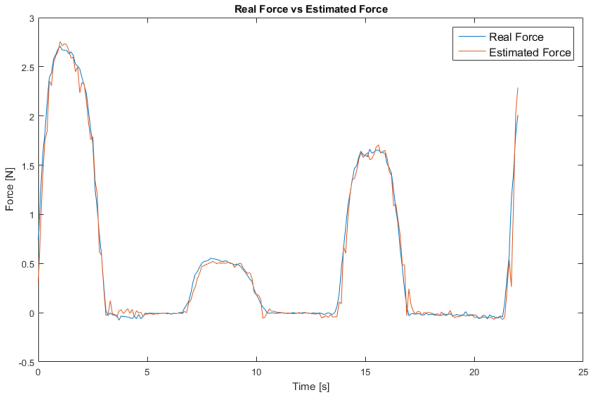


Figure 15: ANN Testing

The force estimation when there is variation of the force direction composes a more complex problem and it was simulated by rotating the prototype when pressing it against the force sensor, as illustrated by figure 16.

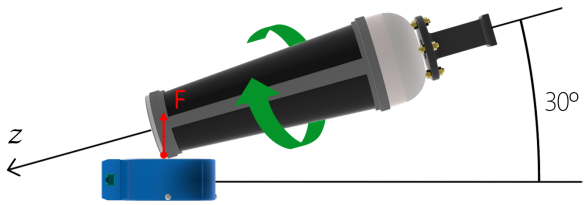


Figure 16: Procedure With Prototype Rotation

This problem presents higher complexity given that the relation between the centroid displacement and the force is no longer linear and depends on the direction of the displacement. This factor can be inferred by the combination of the centroid components along with their respective signs. Even if more difficult, the same procedure was used to compute the number of neurons for the ANN structure. Using only one hidden layer the following mean performances were obtained:

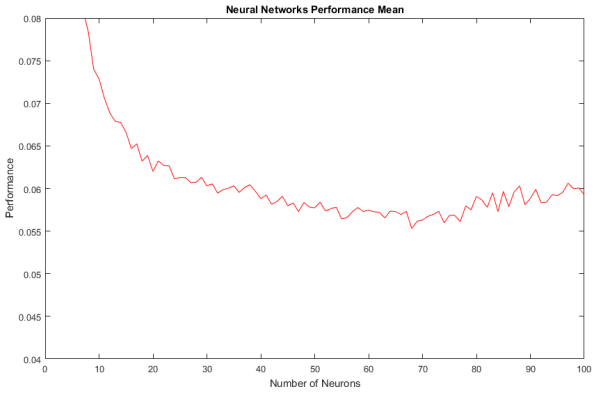


Figure 17: ANN Performance Mean

As can be noted, the ANN's with better performances have between **60-80** neurons, with the mean minimum at **68** neurons but individually the best result was reached by a net with **74** neurons and a performance of **0.0487**. From implementation of the this net the following behaviour was achieved:

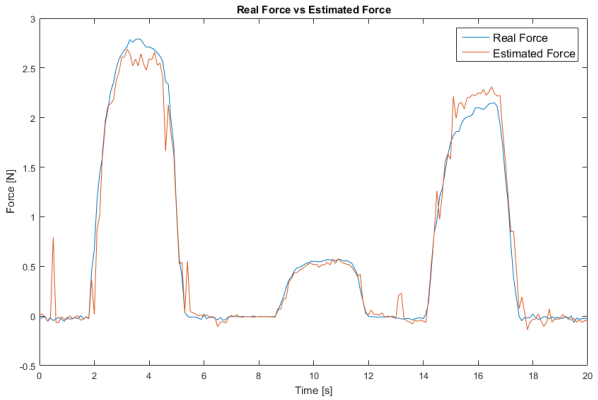


Figure 18: 1 Layer ANN Testing

With the purpose of trying to improve the problem was also approached using a neural network with two hidden layers. To evaluate the results a surface plot was employed where the X and Y axis relate to the number of neurons in the first and second layers respectively and the Z axis as the average performance value (19).

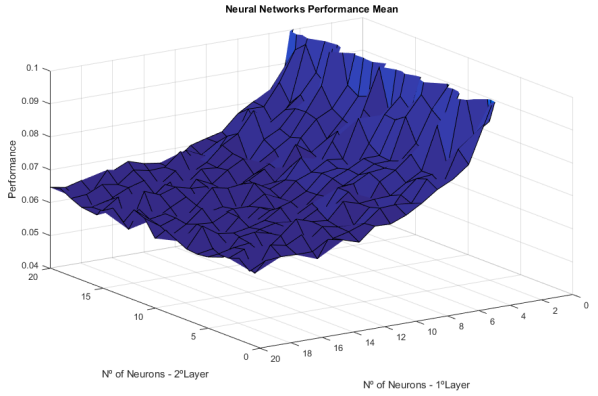


Figure 19: ANN Performance Mean

From reviewing the results it was concluded that ANN's with two hidden layers performed worse with respect to net's with a single layer. Furthermore, it was noted that the net's performance is independent of the dimension of the second layer, only changing when there is variation of the number of neurons in the first layer. In conformity with these results the best obtained ANN had a 16-1 structure (16 neurons in the first hidden layer and only 1 in the second). This means that the net behaved as if it only had a single hidden layer with 16 neurons, where the second layer behaved simply as an output layer that only added a scaling factor to the final result. In conclusion the second hidden layer incremented unnecessarily complexity to the neural network.

5. Conclusions

The prototype's conception required thorough planning and general craftsmanship. Although it physically performed as intended there was the issue of it not being totally airtight and it could be simpler by having an axysymmetrical behaviour.

With regards to the image processing procedure designed it was shown that it was appropriate to acquire the desired data from the images obtained from the webcam. Furthermore it was computationally light enough to be able to estimate the force in real-time, enabling its application in a real-case scenario where the exerted force must be measured promptly.

From the results analysis it was shown that artificial neural networks are a suitable approach in the scope of this work, being able to infer, with some degree of precision, the force values from relevant vision data.

Without prototype rotation the prototype presented remarkable results proving to be an highly adequate solution having a mean error of only **0.051 N** and a mean percentage error of **1.25%**.

With prototype rotation, its was shown that an ANN with a single hidden layer performs better than a net with more hidden layers considering that these implementations produced mean errors of **0.185** and **0.255** respectively. Although the performance was lower when comparing to the previously mentioned application, the results were satisfactory but can be improved with further analysis.

Although this line of investigation still has a lot of capability to improve, it was demonstrated that the integration of tactile sensing with soft robotic links through the application of computer vision and artificial neural networks techniques, presents a practicable solution.

References

[1] J. Rossiter and H. Hauser. Soft robotics - the next industrial revolution? *IEEE Robotics Automation Magazine*, 23(3):17–20, 2016.

[2] Vernon L. Nickle, Jacquelin Perry, and Alice L. Garrett. Development of useful function in the severely paralyzed hand. *JBJS*, 45(5), 1963.

[3] G Robinson and John Bruce Clayfield Davies. The parallel bellows actuator. pages 195–200, 1998. Robotica 98, Robotica 98 ; Conference date: 01-11-1998.

[4] G. Immega and K. Antonelli. The ksi tentacle manipulator. In *Proceedings of 1995 IEEE International Conference on Robotics and Au-*

tomation, volume 3, pages 3149–3154 vol.3, May 1995.

[5] M. W. Hannan and I. D. Walker. The 'elephant trunk' manipulator, design and implementation. In *2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Proceedings (Cat. No.01TH8556)*, volume 1, pages 14–19 vol.1, 2001.

[6] Barry Trimmer, Ann Takesian, Brian Sweet, Chris Rogers, Daniel Hake, and Daniel Rogers. Caterpillar locomotion: A new model for soft-bodied climbing and burrowing robots. *Proceedings of the 7th International Symposium on Technology and the Mine Problem*, 04 0002.

[7] Pinyo Puangmali, Kaspar Althoefer, Lakmal Seneviratne, Declan Murphy, and Prokar Dasgupta. State-of-the-Art in Force and Tactile Sensing for Minimally Invasive Surgery. *IEEE Sensors Journal*, 8(4):371–381, 2008.

[8] Gen-ichiro Kinoshita, Shuhei Aida, and Masahiro Mori. Pattern classification of the grasped object by the artificial hand. In *Proceedings of the 3rd International Joint Conference on Artificial Intelligence. Standford, CA, USA, August 20-23, 1973*, pages 665–670, 1973.

[9] Mark H. Lee. Tactile sensing: New directions, new challenges. *I. J. Robotics Res.*, 19(7):636–643, 2000.

[10] MI Tiwana, Stephen Redmond, and Nigel Lovell. A review of tactile sensing technologies with applications in biomedical engineering. *Sensors and Actuators A: Physical*, 179:17–31, 06 2012.

[11] G. J. Gerling and G. W. Thomas. The effect of fingertip microstructures on tactile edge perception. In *First Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems. World Haptics Conference*, pages 63–72, March 2005.

[12] Goro Obinata, Ashish Dutta, Norinao Watanabe, and Nobuhiko Moriyama. *Vision Based Tactile Sensor Using Transparent Elastic Fingertip for Dexterous Handling.* 02 2007.