Predicting Grasp Success with a Soft Sensing Skin and Shape-Memory Actuated Gripper

Julian Zimmer[1], Tess Hellebrekers[2], Tamim Asfour[1], Carmel Majidi[2,3], and Oliver Kroemer[2]

***Abstract*— Tactile sensors have been increasingly used to support rigid robot grippers in object grasping and manipulation. However, rigid grippers are often limited in their ability to handle compliant, delicate, or irregularly shaped objects. In recent years, grippers made from soft and flexible materials have become increasingly popular for certain manipulation tasks, e.g., grasping, due to their ability to conform to the object shape without the need for precise control. Although promising, such soft robot grippers currently suffer from the lack of available sensing modalities. In this work, we introduce a soft and stretchable sensing skin and incorporate it into the two fingers of a shape-memory actuated soft gripper. The on-board sensing skin includes a 9-axis inertial measurement unit (IMU) and five discrete pressure sensors per finger. We use this sensorized soft gripper to study grasp success and stability of over 2585 grasps with various objects using several machine learning methods. Our experiments show that LSTMs were the most accurate predictors of grasp success and stability, compared to SVMs, FFNNs, and ST-HMP. We also evaluated the effects on performance of each sensor's data, and the success rates for individual objects. The results show that the accelerometer data of the IMUs has the largest contribution to the overall grasp prediction, which we attribute to its ability to detect precise movements of the gripper during grasping.**

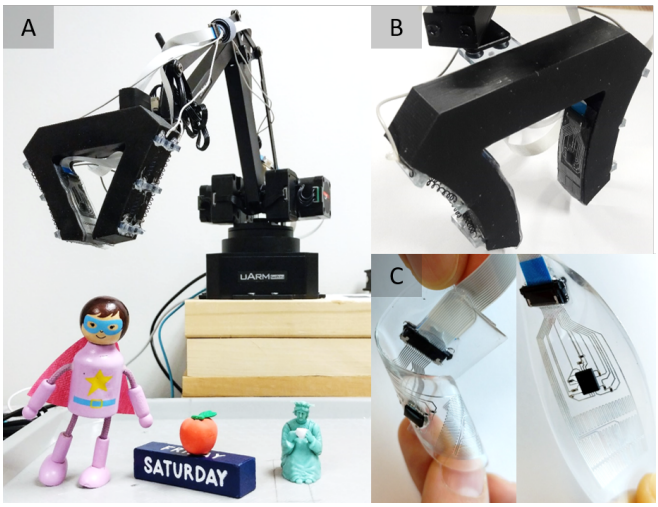

Fig. 1. A) Overview of soft, sensorized gripper and most common objects used for grasping data set. B) Gripper activated by shape-memory coils in open position with shape-memory coils heated. C) Overview of flexible and stretchable hybrid liquid-metal circuit

I. INTRODUCTION

Soft robots represent a paradigm shift in the way that they are engineered to operate safely and effectively in unstructured environments. Unlike traditional robots, soft robots are built from compliant materials such as gels, elastomers, or fluids that can undergo large deformations. Soft grippers have garnered particular interest for manipulation tasks due to their compliant behavior that can delicately and securely conform to irregular object surfaces. While they provide an inherent mechanical advantage for safety, soft grippers often lack the diverse range of sensing modes and off-the-shelf electronics available to conventional robot grippers. This is a critical disadvantage, as sensing is an important aspect of robotic manipulation.

The introduction of tactile and proprioceptive sensing has been shown to enable a variety of high-level tasks including slip detection [1], grasp stability estimation, object recognition and tactile servoing [2]. Predicting the success of a grasp before it is fully executed can enable safer and more efficient grasping. In response, soft and stretchable sensors are being developed to enable proprioceptive and tactile sensing in soft robots. Soft sensors must provide reliable and fast information without adversely affecting the underlying mechanics of the soft robot. Current research in soft sensing has largely focused on tactile sensing. Pressure and strain sensors are commonly made from liquid-metals, optics, or ionic fluids [3], [4], [5]. The integration of these pressure and strain sensors has been shown to capture responses to both internal and external deformations. However, integrated systems of soft actuators and sensors are challenging to accurately model using non-linear viscoelastic models. This has motivated an alternative data-driven approach to isolate and approximate the soft body deformation from sensor responses. The resulting models can be used for open-loop and closed-loop control [6].

We present a study on grasp success and stability estimation using a shape-memory soft gripper and integrated hybrid microelectronic sensing skin (Fig. 1A). Our shape-memory actuation method (Fig. 1B) lends itself to straight-forward integration with a robotic arm, and the sensor skin fabrication provides both a 9-axis IMU and five pressure sensors per finger (Fig. 1C). We use these aspects of our hardware

*This work was in part supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE 1252522, the Office of Naval Research under Grant No. N00014-18-1-2775, and German Academic Exchange Service (DAAD) within the Continuous Learning in Collaborative Studies project (CLICS). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation, ONR, or the German Academic Exchange Service.

[1]Institute for Anthropomatics and Robotics, Karlsruhe Institute of Technology, Karlsruhe, Germany

[2]Robotics Institute, School of Computer Science, Carnegie Mellon University, Pittsburgh PA 15123, USA

[3]Department of Mechanical Engineering, Carnegie Mellon University, Pittsburgh PA 15213

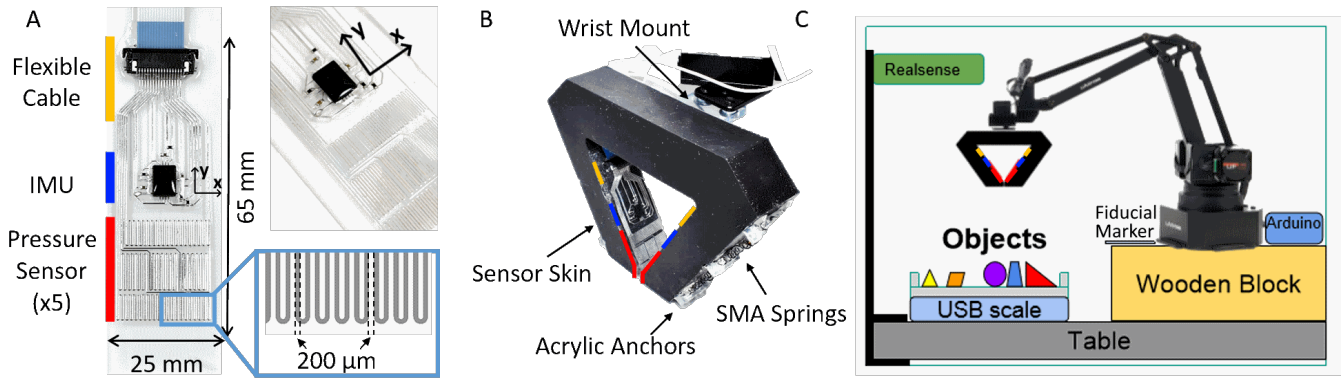

Fig. 2. A) Close-up of sensor skin with flexible cable, IMU, and pressure sensors visible. The pressure sensors have a channel width and spacing of 200μm. B) Detailed overview of the sensorized gripper with sensor skins, shape-memory coils, and acrylic mounts. C) Detailed overview of experimental set-up with the uArm Swift Pro, soft gripper, Arduino Due, RealSense camera, ARTag, and USB scale for ground truth.

platform to collect the first grasping data set for a soft gripper with both proprioceptive and tactile data. In sections IV and V, we will discuss the hardware configuration and data collection process, as well as the algorithms used. In section VII, we then present the results of our experiments on predicting grasp success and stability.

II. RELATED WORK

Tactile sensing is known to be a key component in human grasping [7] and has recently shown increased activity in robotics applications such as detection of object properties [8], object recognition [9], and friction estimation [10]. Similarly, there has been an increasing focus towards integrated sensing for soft grippers [11]. Soft robot grippers typically employ pneumatic/hydraulic, dielectric, or granular jamming forces to achieve actuation [12]. Therefore, the on-board sensors must be soft and elastic such that they can continue to operate while undergoing large deformations and not impair the natural mechanics of the host system. Researchers have taken several different approaches to stretchable sensing, including optics, deterministic architectures, and nanomaterials [13]. Liquid-metal strain and pressure sensing is most commonly used for soft sensing due to the ability to maintain electrical conductivity over large strains [14]. With the advancement of stretchable sensing technologies, opportunities for control and high-level tasks has also progressed [15], [16], [17].

The earliest works that use tactile sensing to estimate grasp success employed hidden-markov-models (HMMs) and support vector machines (SVMs) to classify the tactile data from a 3-finger rigid gripper [18], [19], [20]. Bekiroglu et al. [20] also included joint configurations and object information into their feature vector. This inspired similar approaches in the following years that focused on SVMs [21], [22] and bag-of-words models [23]. Researchers have also proposed algorithms to adapt grasps when failure is predicted [24], [25], [26]. A stream of temporal and spatial tactile data has been shown to learn features for improved grasp success prediction and object recognition with SVMs [27]. In recent years, the usage of neural networks to process tactile data has become popular. Calandra et al. [28] learn grasp success in an end-to-end manner by employing convolutional neural networks on image data and optical tactile sensor data. Qin et al. [29] concatenated tactile time series data to a temporal tactile image which was processed with a CNN.

Techniques from the artificial intelligence and machine learning communities has been also employed to model the complex non-linear behaviors of soft robot deformation [30]. It has been shown that such techniques are helpful for characterizing liquid-metal based sensors [31], [32], track soft body deformation [33], [34], [35], [36], [37], and classify objects [38], [39].

In contrast to these approaches, we study the problem of grasp success for a soft gripper, considering not only tactile but also proprioceptive data. We compare multiple machine learning algorithms, including SVMs and long short-term memory networks (LSTMs), to predict grasp outcomes, i.e. grasp success and stability. If the gripper successfully lifts an object, we consider it a *successful* grasp success. If it additionally withstands a rapid shaking motion, we consider it a *stable* grasp. Similar to our approach, Thuruthel et al. [33] also employ recurrent networks to process soft sensor data. While their goal is to learn the forward kinematics of a single soft finger based on strain sensor data and control input, we aim to predict grasp stability and success. The non-linear mechanics of the gripper and unpredictable properties of objects make this a good opportunity for data-driven techniques.

Our previous work demonstrates a two-finger soft gripper implementation with shape-memory alloy (SMA) and a soft, stretchable sensor skin with liquid-metal traces and embedded microelectronic components [40]. In this work, we have scaled up the design of the soft gripper and redesigned the skin for manipulation by incorporating a 9-axis IMU and five liquid-metal pressure sensors [41] per finger. Additionally, we have integrated the sensorized soft gripper into a four degree-of-freedom (DOF) robotic arm and collected data over 2500 grasp attempts.

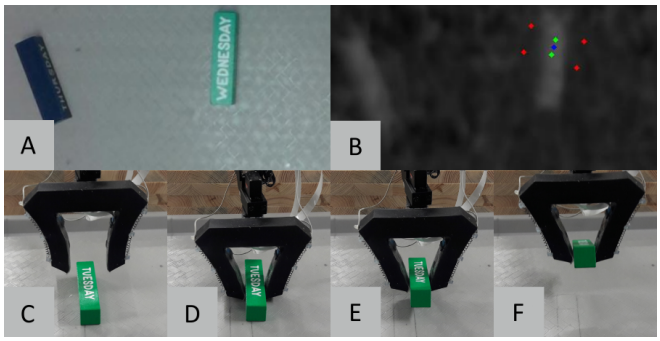

Fig. 3. Height map and different grasp stages. A) Exemplary object scene. B) Corresponding height map calculated from point cloud. Randomly generated pose has been accepted based on the values at the colored points (the two green points must be higher than a threshold, while the four red points must be lower than it). C) The gripper opened and moves down onto the object tray. D) The SMA is deactivated and the gripper closes onto the object (*CG*). E) The small lift is performed, lifting the gripper 10 mm vertically. Note that the object is still in contact with the plane, as it slightly spun in the grasp (*SL*). F) The object is lifted up and is further moving in the grasp. A stability test is executed afterwards.

III. HAND DESIGN AND FABRICATION

The sensor skin is composed of a thin sheet of soft silicone rubber (polydimethylsiloxane; Sylgard 184; Dow Corning) embedded with liquid-metal circuit wiring and microelectronic components for sensing and digital interfacing. The skin is fabricated using steps previously reported in [42]. As shown in Fig. 2A, the circuit includes a 3-axis magnetometer, 3-axis accelerometer, and 3-axis gyroscope within a 9-axis IMU integrated chip (BNO055; Bosch Sensortec). There are also five liquid-metal pressure gauges with 200 μm spacing and channel width. Contact forces acting on the pressure gauges causes the liquid-metal channel to collapse, resulting in an increase in electrical resistance proportional to the applied pressure [41]. The IMU was chosen to respond to position changes in the fingers as they interact with the object, while the pressure sensors were placed to respond to contact forces. The flexible cable provides power and carries five analog outputs and I2C data to the off-board microcontroller.

The shape-memory actuated soft gripper was fabricated using a 3D-printed mold (Objet24; Stratasys), elastomer (Dragonskin 30; Smooth-On), and black silicone paint (SilcPig; Smooth-On). Further details can be found in [40]. The shape-memory coils are adhered along each finger and connected in series to a 12 V power source. The coils are activated via Joule heating through the digital power switch, pulling the gripper into the 'open' configuration. When the shape-memory alloy cools, the fingers relax and passively conform around an object. The resulting two-finger soft gripper behaves similarly to a binary parallel-plate gripper and was designed for simple grasping mechanics. The sensor skin is cut to size and adhered (Silpoxy; Smooth-On) to the inside surface of each finger to create the final sensorized soft gripper (Fig. 2B).

IV. ROBOT GRASPING

To collect the data, the soft gripper is integrated into a setup with a small robotic arm. The idea for the design of the setup was to implement an automatic grasping routine that allows collecting sensor data as well as ground truth of grasp success and stability.

The robot platform includes a uArm Swift Pro, which is a small, lightweight 4-degree-of-freedom robot. The soft gripper is connected to the wrist axis by screwing the acrylic plate of the gripper to the wrist axis of the uArm. In front of the robotic arm, a plastic tray holding the objects is placed on top of a USB scale, which provides ground truth labels for lifting or dropping objects. Next to the object tray, a RealSense D435 depth camera is mounted 60 cm above the tray to observe the scene. The robot also has an Arduino to process the gripper data and a fiducial marker to visually localize the robot (Fig. 2C).

A grasping attempt is based on a depth image where the resulting point cloud is transformed to the robot's coordinate system by combining the transformations from the camera to the fiducial marker and from the marker to the robot. The former is calculated using the Aruco library (see [43], [44]) with the RGB image of the RealSense camera. The transformation between marker and robot has been calculated previously based on measurements. After the transformation,the point cloud is cropped to a 15×30 cm region within the tray. The cropped point cloud is used to calculate a 200×400 height map of the scene (Fig. 3A, B). A height threshold is used to segment out the tray. To generate grasp candidates, we uniformly sample x-y positions and z axis orientations. Candidates are rejected if they do not follow a specific height pattern, as shown in Fig. 3B, and executed otherwise. This approach significantly reduces the number of empty grasps.

When a grasp is executed, the Arduino communicates with the sensors and controls the SMA actuation. The different steps of a grasp attempt can be seen in Fig. 3C-F. First, the gripper is moved over the desired grasp location in a safety distance, based on the height map. The SMA is activated, while the gripper moves down towards the object. The robot moves the gripper downwards, until the USB scale detects additional weight. The SMA springs are then deactivated and the gripper will close. Once the gripper is fully closed, a small lift displacement of 10 mm is performed. The gripper is then fully lifted and the value of the USB scale is read. If the scale indicates less weight than before the grasp, the object was successfully lifted and the stability test can be performed. To do this, the wrist is moved rapidly back and forth twice by 10 degrees. The scale value is again used to check if the object was dropped.

To get the most relevant information about the grasp, the data from the sensor skins is recorded during two important parts of the grasping routine. The first one is the closing of the gripper (Fig. 3D), which includes the moment when the gripper initially comes into contact with the object. The second part is the small lift (Fig. 3E), executed when the gripper has grasped the object but only attempted to

lift it 10 mm, which may provide some additional object weight information. We will be referencing these phases with *SL* for small lift and *CG* for close gripper in the following. The accelerometer, magnetometer, gyroscope and pressure data from both sensor skins, a total of 28 values, are sampled at 30 Hz continuously during these two phases. The grasp is marked as successful (object in hand after lift) and stable (object in hand after shake), depending on the weight registered by the scale after the lifting and shaking actions.

V. LEARNING TO PREDICT GRASP OUTCOMES

Learning grasp success or stability requires the employment of appropriate machine learning methods. Since grasping is a dynamic process, we record data across multiple time steps of the grasp execution. Consequently, the employed algorithms need to handle high dimensional time series data.

Preprocessing: Before passing the sensor data into the machine learning algorithms, we normalize the data per time step across different grasps. To this end, the data from each time step is concatenated to form a single feature vector per recording. In this way, we preserve temporal information between the features.

SVMs: The first evaluated method for this task were Support Vector Machines (SVMs) which are commonly used for binary classification and have been often used for classifying tactile data [20], [21], [22] although they are not specifically designed to handle time series data. The scikit-learn implementation was used.

ST-HMP: Spatio-Temporal Hierarchical Matching Pursuit (ST-HMP) is a feature learning approach for temporal tactile data [27]. Its core idea is to create a dictionary of the most explanatory vectors in the feature space and describe the data through elements from the dictionary. Spatial and temporal pooling are then employed to reduce the dimension and incorporate invariances for additional robustness. In [27] the ST-HMP features were successfully used as input to SVMs to predict grasp success and perform object instance recognition. Given that we only have five pressure sensors per finger, we do not use spatial pooling. We used a dictionary size of 50 and a $1-2-4-8$ pattern for the temporal pooling (pooling over the entire data in the first level and pooling over each eighth of the data in the last levels).

FFNNs: Given their ability to learn features for challenging problems, we trained feedforward neural networks to classify the data. Our networks were implemented in PyTorch and used 3 fully connected hidden layers with 200/40/10 units and rectified linear unit activation functions. The inputs to the network were the concatenated sensor recordings. Given our limited number of pressure sensors, we did not explore convolutional layers.

LSTMs: Recurrent Networks have been designed specifically for time series data. *Long short-term memory* networks (LSTMs) are the most frequently used recurrent nets due to their ability to retain information over many time steps. For the LSTMs, which were also implemented in PyTorch, the input was not concatenated, but the different time steps were processed sequentially. The LSTMs used a single layer of 20 hidden units.

VI. GRASPING DATASET

The main dataset consists of grasps on 38 objects. The objects were chosen to provide a large variety while still being able to be grasped. 75 % of the grasps have been recorded on 10 objects that cover different materials (wood, plastic, grape), shapes (sphere, cylindrical, rectangular prism) and sizes (maximum dimension between 20 and 80 mm). These objects are visualized with the object specific results in Fig. 5. The resulting dataset contains 2585 grasps with a grasping success rate of $50.\overline{90}\%$. Out of the 1316 successful grasps 872 were stable, giving a stability rate of 66.26%. The machine learning models have been evaluated on a 5-fold cross validation. For each calculation, 4 folds were used for training, while the last fold was split in half to create the validation and test data. For the neural networks, early stopping based on the validation loss was employed.

VII. RESULTS

In this section we explain the results of our experiments. We evaluated the effects of using data from different sensors and time periods during the evaluations. We also evaluated using different types of learning algorithms for performing the classification.

A. Comparing Close Gripper and Small Lift Data

To compare the data recorded during different stages of a grasp attempt, the data was processed using the same method. The main difference between them is the length of each recording. While the small lift (*SL*) is a very short action (1.5 seconds), the passive closing of the gripper (*CG*) takes 16.5 seconds. The feature dimension increases accordingly. A clear difference is noticeable for the grasp success predictions (Fig. 4), with the *SL* data achieving about 6% higher accuracy. For both target predictions, the LSTM performs best on the *SL* data. On *CG* data, a FFNN and a polynomial SVM reached the highest classification scores. The grasp stability prediction performances were very similar between the two datasets. As expected, the *SL* data performs better, although this is not true for grasp stability predictions. The *SL* recordings seem to be capturing meaningful temporal data, as for both grasp success and grasp stability predictions, an LSTM yields the best performance.

TABLE I
GRASP SUCCESS LEARNING RESULTS OVERVIEW

ML Method	*CG* Data	*SL* Data
LSTM	$82.88 \pm 1.7\%$	$91.71 \pm 1.0\%$
FFNN	$85.67 \pm 1.5\%$	$90.78 \pm 1.2\%$
SVM RBF	$83.11 \pm 1.7\%$	$89.53 \pm 1.4\%$
SVM Linear	$82.18 \pm 2.2\%$	$78.68 \pm 2.0\%$
ST-HMP	$75.29 \pm 3.1\%$	$80.54 \pm 1.4\%$

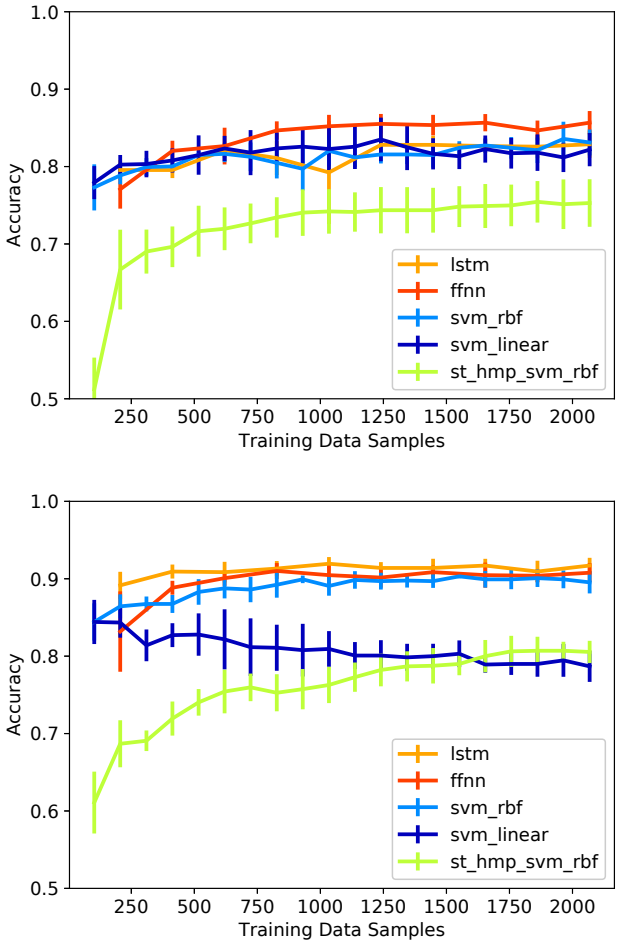

Fig. 4. Success prediction based on the close gripper data (top) and success prediction based on the small lift data (bottom). Grasp success accuracy for different amounts of training data. Showing LSTM, and feed forward neural networks, linear and radial basis function SVMs and ST-HMP results.

B. ML Methods Comparison - Grasp Success

We compared the different machine learning methods on the grasp success prediction based on *CG* and *SL* data. Fig. 4 show the results for increasing amounts of training data. Table I additionally shows the values when the models were trained with the full training data (rightmost points in the figures). In the table, the accuracy and standard deviation for different machine learning models predicting grasp success are given.

All models, except for the linear SVM, performed better on the *SL* data. The larger variance in error indicate that the predictions from *CG* data are less consistent than the ones from *SL* data. The *CG* results still show an improvement over the baseline of the dataset's 50.91% successful grasps. Looking at the specific methods, the linear SVM and the ST-HMP approach have the lowest grasp prediction accuracy and greatest noise. It should be noted that ST-HMP may be more suitable for higher-resolution data as in the original paper [27]. In summary, LSTMs perform best for *CG* data ($85.67\% \pm 1.5\%$) and FFNNs perform best for *SL* data ($91.94\% \pm 1.0\%$), both in accuracy and noise. Looking at the failed classifications shows that false positives are more frequent than false negatives, i.e. the grasp success likelihood is slightly overestimated. Still, most failed grasps are classified correctly. Overall, the neural architectures are the most suitable for grasp success classification.

C. ML Methods Comparison - Grasp Stability

Predicting grasp stability is a significantly harder learning task than predicting grasp success. Since only successful grasps can be tested for stability, our available dataset is reduced to 1316 grasps. Out of these successful grasps, 66.26% turned out to be stable. In addition to the smaller dataset, the stability prediction problem is also more challenging due to the increased dependence of the shaking outcome on object properties like friction, weight, and size. These properties are very difficult to estimate from a single grasp given our limited set of sensors.

The increased difficulty is apparent in the results. The best achieved accuracy was $71.23\% \pm 4.3\%$ by a polynomial SVM using *CG* data. The best *SL* result was $70.43\% \pm 3.6\%$ by LSTMs. Most other methods ended with accuracies in the range of $65 - 68\%$ only slightly outperform the baseline of 66.26%.

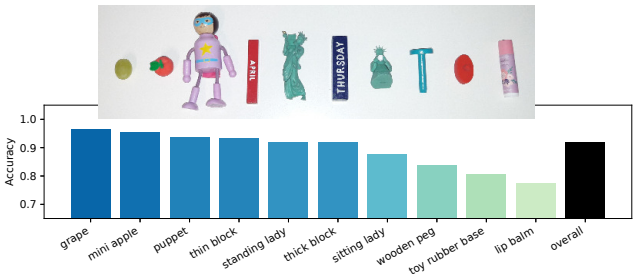

Fig. 5. LSTM prediction accuracy per object for small lift data. The objects in the top of the figure are aligned with the bars indicating the accuracy of their grasp success predictions.

D. Grasping Specific Objects

To check whether grasping specific objects has a strong effect on predictability, i.e., if the robot does not need to generalize across objects, we analysed the LSTM grasp success results per object. Fig. 5 shows the 10 objects with the most recorded grasps on top and their prediction scores in the test sets in the bar chart below. Most of the objects have prediction scores in the range of $90 - 96\%$. The four objects with the worst predictions have several physical properties that might have resulted in less reliable predictions. The *sitting figurine*, *wooden peg* and *toy rubber base* objects all have unusual shapes and rather thin features that can easily slip out of a grasp. The *lip balm* on the other hand has a relatively slippery surface. That being said, they still reached $77 - 87\%$ accuracy.

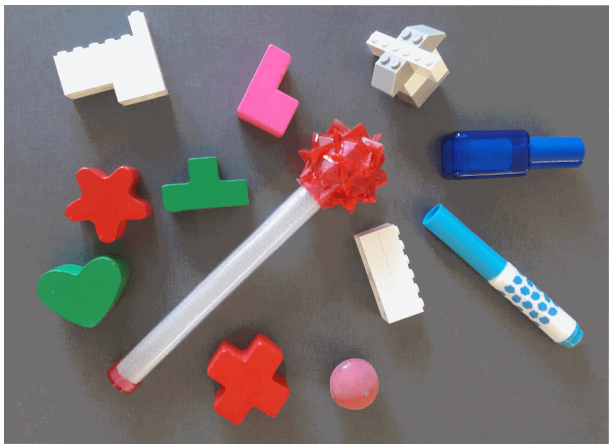

Fig. 6. These objects were used to test the generalization of the grasp success prediction and were not part of the training dataset.

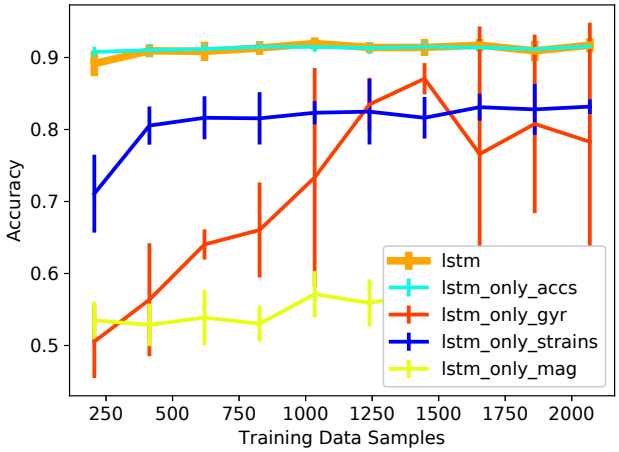

Fig. 8. Comparison of LSTM performance on *SL* data when the input is limited to one sensor type.

E. Generalization to New Objects

To show the generalization potential of the approach, an additional generalization dataset has been recorded, where the grasp success prediction was performed live during the grasping routine. We performed 10 grasps each for 13 novel objects (see Fig. 6). After the small lift is executed, two previously calculated LSTM models predict if the grasp is successful and stable. The grasp attempts had a lower success rate of 35% on these objects than the main dataset, but they showed a similar stability rate of 70%. The accuracies for the predictions are similar to the ones on the test data of the main dataset, scoring 90% and 71% for grasp success and stability prediction respectively. Thus, the predictions made on a dataset of novel objects confirm the results on the main dataset.

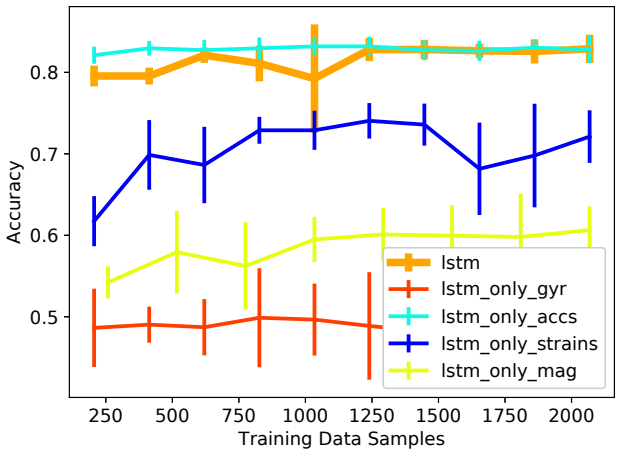

Fig. 7. Comparison of LSTM performance on *CG* data when the input is limited to one sensor type.

F. Sensor Relevance

The sensorized skin provides data from four types of sensors. To determine the relevance of each sensor type, we trained four LSTMs using only one type of sensor's data as input at a time: accelerometer, magnetometer, gyroscope or pressure data. As Fig. 7 and 8 show, for both cases the magnetometer data is not able to produce meaningful results, ending with low accuracy. The gyroscope data recorded during the gripper closing cannot predict the grasp success. While the respective *SL* data achieves a higher accuracy, it is still not close to the performance when considering all sensor inputs. Using the pressure sensor data only, the prediction is significantly improved compared to the magnetometer or gyroscope data. The pressure data scores are about 10% lower than the LSTM results with all sensor inputs. The accelerometer stands out as the sensor type with the best prediction capabilities. For both grasping stages, the LSTM model based on the accelerometer data is on par with the model which processes all different sensor types. This analyses indicates that measuring the accelerations is the mose informative for learning to monitor soft robots. Intuitively, an IMU that is placed on a soft robot with deformable parts is able to partially make up for the lack of classical proprioceptive sensors. Pressure sensors seem to be able to sense the movement of the gripper as well, but in a more limited fashion.

G. Grasp Location Sensing

As an additional demonstration of the accelerometer's sensing capabilities, we evaluated its ability to estimate where a specific elongated object was grasped along its length. This data could then be used to regrasp the object. A wooden block was grasped at five different locations along its axis, as shown in Fig. 9. For each of these grasp locations, 50 grasp attempts were recorded. The resulting average x-axis accelerometer recordings for the five locations can be seen in Fig. 10.

Given the limited amount of training data, we used an SVM for classification. We calculated the five class classification 10 times on random 80/20 train and test data splits

Fig. 9. The grasp location analysis was based on grasps at 5 different locations. These locations (left, semi-left, middle, semi-right, right) are shown in this figure, corresponding to figure 10.

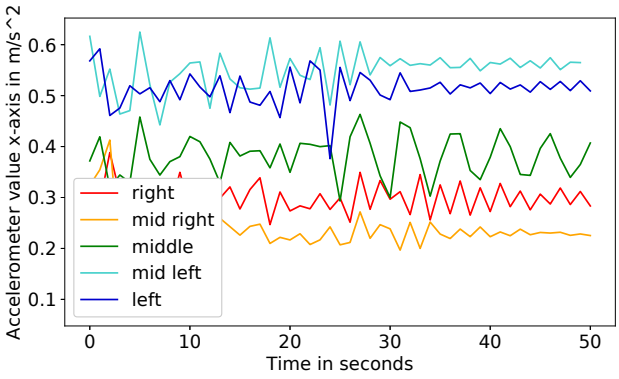

Fig. 10. Average x-axis acceleration when a small lift is performed on different grasp locations. 50 grasps have been performed on each of the five locations shown in figure 9. This plot shows the average accelerometer value of the x-axis of the right finger during the 10 mm small lift. The classes can be distinguished relatively clearly.

and averaged the results. For the classification, a linear SVM kernel performed the best, achieving an average accuracy of 79.6% with a standard deviation of 4.4%. Fig. 11 shows the respective confusion matrix. The values clearly indicate that the accelerometer *SL* data can sense where the object is grasped in most cases.

VIII. DISCUSSION

The sensorized skin was able to gather valuable information of the state of the soft gripper during grasping. In many cases, recordings taken while closing the gripper are sufficient to predict if a grasp will be successful. In comparison, the data obtained during the small lift performs even better. When appropriate sensing (like an IMU or pressure sensor) is available, the induced movement can be perceived and improve the estimate of the system state. The fact that LSTMs performed best on this data indicates a valuable temporal component in the data. The results have been validated on an additional generalization data set. The grasp predictions were similar over different objects.

We attribute reduced performance on grasp stability to increased object dependence. A possible way to overcome this could be the incorporation of more sensors that detect

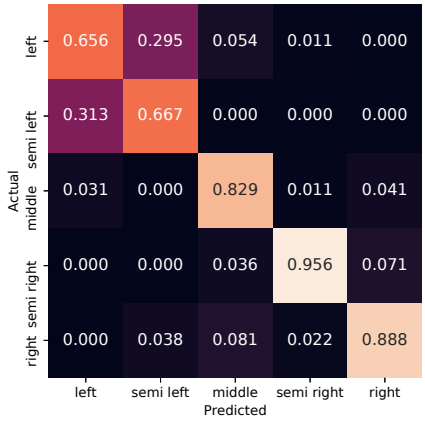

Fig. 11. Confusion matrix of the classification results of the linear SVM for the five different grasp locations in the grasp location sensing.

object properties or object classes. Our experiments also showed that our sensor setup has the potential to estimate further interesting properties, e.g., the center of mass location relative to the fingers. Based on this data, a regrasping approach would be possible.

To investigate the change in sensor readings over time, we compared data collected two months apart. We observe slightly decreased idle values for the pressure sensors. We attribute this change to general wear of the elastomer surrounding the liquid-metal traces. The IMU data is more consistent between the two recordings, due to its epoxy packaging, which is not affected by elastomer wear. The slight effects of wear on the pressure sensors might also be a reason for the overall better performance of the accelerometer data.

IX. CONCLUSIONS AND FUTURE WORK

In summary, a sensorized soft gripper was integrated with a 4-DOF robotic arm. The soft gripper behaves similar to a binary parallel-plate gripper, actuated by shape-memory springs fixed along each finger. Each finger had a stretchable sensing skin with a 9-axis IMU and five discrete pressure sensors. This system was used to collect data of over 2585 grasp attempts. LSTMs showed the best performance for grasp success prediction based on both the close gripper (CG) and small lift (SL) datasets.

In the future, we plan to incorporate more microelectronics, such as time-of-flight and microphones, into the sensing skin. Adding a denser pattern of pressure sensors to the skin or covering more parts of the gripper with the skin would also increase the tactile sensing capabilities. Improved sensor output and spatial resolution could improve the overall understanding of the object interaction.

REFERENCES

[1] R. D. Howe and M. R. Cutkosky, "Sensing skin acceleration for slip and texture perception," in *Proceedings, 1989 International Conference on Robotics and Automation*. IEEE, 1989, pp. 145–150.

[2] Z. Kappassov, J.-A. Corrales, and V. Perdereau, "Tactile sensing in dexterous robot hands," *Robotics and Autonomous Systems*, vol. 74, pp. 195–220, 2015.

[3] M. D. Dickey, "Stretchable and soft electronics using liquid metals," *Advanced materials*, vol. 29, no. 27, p. 1606425, 2017.

[4] M. Amjadi, K.-U. Kyung, I. Park, and M. Sitti, "Stretchable, skin-mountable, and wearable strain sensors and their potential applications: a review," *Advanced Functional Materials*, vol. 26, no. 11, pp. 1678–1698, 2016.

[5] H. Zhao, K. OBrien, S. Li, and R. F. Shepherd, "Optoelectronically innervated soft prosthetic hand via stretchable optical waveguides," *Sci. Robot.*, vol. 1, no. 1, p. eaai7529, 2016.

[6] D. Rus and M. T. Tolley, "Design, fabrication and control of soft robots," *Nature*, vol. 521, no. 7553, p. 467, 2015.

[7] R. S. Johansson and J. R. Flanagan, "Coding and use of tactile signals from the fingertips in object manipulation tasks," *Nature Reviews Neuroscience*, vol. 10, no. 5, p. 345, 2009.

[8] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, "Robotic tactile perception of object properties: A review," *Mechatronics*, vol. 48, pp. 54–67, 2017.

[9] H. Liu, Y. Wu, F. Sun, and D. Guo, "Recent progress on tactile object recognition," *International Journal of Advanced Robotic Systems*, vol. 14, no. 4, p. 1729881417717056, 2017.

[10] W. Chen, H. Khamis, I. Birznieks, N. F. Lepora, and S. J. Redmond, "Tactile sensors for friction estimation and incipient slip detection-toward dexterous robotic manipulation: A review," *IEEE Sensors Journal*, vol. 18, no. 22, pp. 9049–9064, 2018.

[11] H. Wang, M. Totaro, and L. Beccai, "Toward perceptive soft robots: Progress and challenges," *Advanced Science*, vol. 5, no. 9, p. 1800541, 2018.

[12] J. Hughes, U. Culha, F. Giardina, F. Guenther, A. Rosendo, and F. Iida, "Soft manipulators and grippers: a review," *Frontiers in Robotics and AI*, vol. 3, p. 69, 2016.

[13] N. Lu and D.-H. Kim, "Flexible and stretchable electronics paving the way for soft robotics," *Soft Robotics*, vol. 1, no. 1, pp. 53–62, 2014.

[14] X. Wang, R. Guo, and J. Liu, "Liquid metal based soft robotics: Materials, designs, and applications," *Advanced Materials Technologies*, p. 1800549, 2018.

[15] F. Largilliere, V. Verona, E. Coevoet, M. Sanz-Lopez, J. Dequidt, and C. Duriez, "Real-time control of soft-robots using asynchronous finite element modeling," in *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2015, pp. 2550–2555.

[16] R. F. Shepherd, F. Ilievski, W. Choi, S. A. Morin, A. A. Stokes, A. D. Mazzeo, X. Chen, M. Wang, and G. M. Whitesides, "Multigait soft robot," *Proceedings of the national academy of sciences*, vol. 108, no. 51, pp. 20 400–20 403, 2011.

[17] R. K. Katzschmann, A. D. Marchese, and D. Rus, "Hydraulic autonomous soft robotic fish for 3d swimming," in *Experimental Robotics*. Springer, 2016, pp. 405–420.

[18] Y. Bekiroglu, D. Kragic, and V. Kyrki, "Learning grasp stability based on tactile data and hmms," in *19th International Symposium in Robot and Human Interactive Communication*. IEEE, 2010, pp. 132–137.

[19] J. Laaksonen, V. Kyrki, and D. Kragic, "Evaluation of feature representation and machine learning methods in grasp stability learning," in *2010 10th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 2010, pp. 112–117.

[20] Y. Bekiroglu, J. Laaksonen, J. A. Jørgensen, V. Kyrki, and D. Kragic, "Assessing grasp stability based on learning and haptic data," *IEEE Transactions on Robotics*, vol. 27, no. 3, p. 616, 2011.

[21] H. Dang, J. Weisz, and P. K. Allen, "Blind grasping: Stable robotic grasping using tactile feedback and hand kinematics." in *ICRA*, 2011, pp. 5917–5922.

[22] J. Schill, J. Laaksonen, M. Przybylski, V. Kyrki, T. Asfour, and R. Dillmann, "Learning continuous grasp stability for a humanoid robot hand based on tactile sensing," in *2012 4th IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob)*. IEEE, 2012, pp. 1901–1906.

[23] H. Dang and P. K. Allen, "Learning grasp stability," in *2012 IEEE International Conference on Robotics and Automation*. IEEE, 2012, pp. 2392–2397.

[24] H. Dang and P. Allen, "Stable grasping under pose uncertainty using tactile feedback," *Autonomous Robots*, vol. 36, no. 4, pp. 309–330, 2014.

[25] M. Li, Y. Bekiroglu, D. Kragic, and A. Billard, "Learning of grasp adaptation through experience and tactile sensing," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. Ieee, 2014, pp. 3339–3346.

[26] Y. Chebotar, K. Hausman, O. Kroemer, G. Sukhatme, and S. Schaal, "Regrasping using tactile perception and supervised policy learning," in *2017 AAAI Spring Symposium Series*, 2017.

[27] M. Madry, L. Bo, D. Kragic, and D. Fox, "St-hmp: Unsupervised spatio-temporal feature learning for tactile data," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 2262–2269.

[28] R. Calandra, A. Owens, M. Upadhyaya, W. Yuan, J. Lin, E. H. Adelson, and S. Levine, "The feeling of success: Does touch sensing help predict grasp outcomes?" *arXiv preprint arXiv:1710.05512*, 2017.

[29] J. Qin, H. Liu, G. Zhang, J. Che, and F. Sun, "Grasp stability prediction using tactile information," in *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM)*. IEEE, 2017, pp. 498–503.

[30] G.-Z. Yang, "Robot learning—beyond imitation," *Science Robotics*, vol. 4, no. 26, 2019. [Online]. Available: http://robotics.sciencemag.org/content/4/26/eaaw3520

[31] D. Kim and Y.-L. Park, "Contact localization and force estimation of soft tactile sensors using artificial intelligence," pp. 7480–7485, 2018.

[32] I. Van Meerbeek, C. De Sa, and R. Shepherd, "Soft optoelectronic sensory foams with proprioception," *Science Robotics*, vol. 3, no. 24, p. eaau2489, 2018.

[33] T. G. Thuruthel, B. Shih, C. Laschi, and M. T. Tolley, "Soft robot perception using embedded soft sensors and recurrent neural networks," *Science Robotics*, vol. 4, no. 26, p. eaav1488, 2019.

[34] V. Wall, G. Zöller, and O. Brock, "A method for sensorizing soft actuators and its application to the rbo hand 2," pp. 4965–4970, 2017.

[35] C. B. Teeple, K. P. Becker, and R. J. Wood, "Soft curvature and contact force sensors for deep-sea grasping via soft optical waveguides," pp. 1621–1627, 2018.

[36] P. Ohta, L. Valle, J. King, K. Low, J. Yi, C. G. Atkeson, and Y.-L. Park, "Design of a lightweight soft robotic arm using pneumatic artificial muscles and inflatable sleeves," *Soft robotics*, vol. 5, no. 2, pp. 204–215, 2018.

[37] G. Zöller, V. Wall, and O. Brock, "Acoustic sensing for soft pneumatic actuators," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 6986–6991.

[38] B. S. Homberg, R. K. Katzschmann, M. R. Dogar, and D. Rus, "Robust proprioceptive grasping with a soft robot hand," *Autonomous Robots*, pp. 1–16, 2018.

[39] B. S. Homberg, R. K. Katzschmann, M. R. Dogar, and D. Rus, "Haptic identification of objects using a modular soft robotic gripper," in *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Sep. 2015, pp. 1698–1705.

[40] T. Hellebrekers, K. B. Ozutemiz, J. Yin, and C. Majidi, "Liquid metal-microelectronics integration for a sensorized soft robot skin," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 5924–5929.

[41] Y.-L. Park, C. Majidi, R. Kramer, P. Bérard, and R. J. Wood, "Hyperelastic pressure sensing with a liquid-embedded elastomer," *Journal of Micromechanics and Microengineering*, vol. 20, no. 12, p. 125029, 2010.

[42] K. B. Ozutemiz, J. Wissman, O. B. Ozdoganlar, and C. Majidi, "Egain–metal interfacing for liquid metal circuitry and microelectronics integration," *Advanced Materials Interfaces*, vol. 5, no. 10, p. 1701596, 2018.

[43] S. Garrido-Jurado, R. Munoz-Salinas, F. J. Madrid-Cuevas, and R. Medina-Carnicer, "Generation of fiducial marker dictionaries using mixed integer linear programming," *Pattern Recognition*, vol. 51, pp. 481–491, 2016.

[44] F. J. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer, "Speeded up detection of squared fiducial markers," *Image and Vision Computing*, vol. 76, pp. 38–47, 2018.