

# Learning to estimate the proximity of slip using high-resolution tactile sensing

For obtaining the degree of Master of Science in Mechanical Engineering  
Thursday 3<sup>rd</sup> February, 2022

Dirk-Jan Boonstra (4309359), Delft University of Technology

**Abstract**— Tactile sensing provides crucial information about the stability of a grasped object by a robotic gripper. Tactile feedback can be used to predict slip, allowing for timely response to perturbations and to avoid dropping objects. Tactile sensors, included in robotic grippers, measure vibrations, strain or shearing forces which are produced by the movement of the grasped object. With sufficient spatial resolution, tactile sensors can even classify slip or estimate the 3d force displacement field. However, current tactile sensors fail to preemptively detect slippage, requiring fast reaction times during applications in real-time control. Here we show a perception framework that can predict slippage before it occurs by estimating the frictional safety margin. The safety margin indicates the margin to the frictional strength of a grasp, which decreases for reduced friction or increased load force. An accurate safety margin estimate allows for more efficient robot grip force control while providing robustness against object uncertainty and frictional conditions. We developed a high resolution tactile sensor, on which we trained a convolutional neural network to learn the relationship between tactile images and the safety margin. The network’s performance is evaluated on unseen test data, showing robustness to variations in environmental conditions. The results demonstrate that the tactile images contain the information needed to produce accurate safety margin estimates. These estimates can be used for control up to 20% of the minimum required grip force, mimicking human grasping behavior. This approach can drive new grasp control methods and enable robotic grasping of fragile objects in highly dynamic environments. Applications can be found in harvesting, parcel sorting, or improving human-robot interaction.

**Index Terms**—robotic grasping, grip force control, friction, safety margin estimation

## I. INTRODUCTION

Robots have an increasing appearance in our daily lives. Our food is sorted by automated transporting systems for increased greenhouse efficiency [1], parcels are sorted by robotic sorting systems [2], and we even let robots take care of our family in elderly homes [3]. Yet, all these robotic systems are limited in their applications by a lack of stable grasping in unknown settings, making grasping one of the most fundamental problems in robotic manipulation [4]. Tactile feedback can provide the necessary information to improve robotic grasping behavior with timely feedback to unforeseen conditions. This thesis work improves on the current field of robotic tactile sensing by presenting a new vision-based tactile sensor. Data obtained from this sensor were used to train a convolutional neural network, resulting in a tactile perception framework that is able to estimate the proximity of slip.

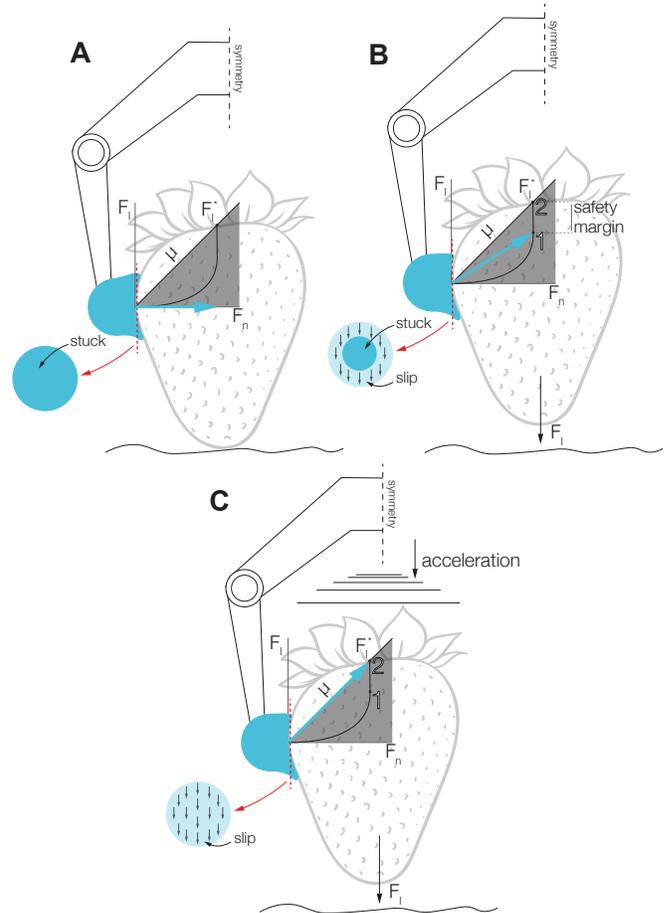


Fig. 1. This figure illustrates the contact mechanics and resultant forces present when grasping a fragile object with a robotic pinch grasp in three dynamical scenarios. In **A**, the object is at rest. The gripper fingers apply a certain grasping (normal) force ( $F_n$ ), and in this static case, no lateral force ( $F_l$ ) is present at the fingertip. The robotic finger is completely stuck against the object, and the blue arrow indicates we are in the lower end of the friction cone. In **B**, the object is lifted off the ground. This results in  $F_l$  at the fingertip, causing the blue resultant force to rise in the friction cone towards point 1. The critical lateral force ( $F_l^*$ ) at point 2 indicates the point of full slippage. The frictional safety margin is given by the distance between points 1 and 2. The margin decreases together with the declining apparent contact area at the fingertip that is still stuck to the object. In **C**,  $F_l$  is increased until the resulting blue arrow reaches outside the friction cone, past  $F_l^*$  at point 2. There is 0% safety margin remaining and the object starts to slip.

### A. Mechanical analysis of the grasping process

Robotic manipulation often makes use of the pinch grasp to manipulate objects. In a pinch grasp the object is enclosed between multiple (gripper) fingers which constrain the free configuration space around the object. The constraint is realized by applying sufficient grip force ( $F_n$ ) to overcome the lateral force ( $F_l$ ) caused by the object's weight. The frictional relationship between these two forces is given by the friction coefficient ( $\mu$ ) as follows:

$$\begin{aligned} \text{Static friction: } & F_l \leq \mu_s F_n \\ \text{Kinetic friction: } & F_l^* = \mu_k F_n \end{aligned} \quad (1)$$

In these equations, a distinction is made between  $\mu_s$  for the static case before the onset of slip, and  $\mu_k$  for the kinetic scenario in the case of full slippage. As our focus is to prevent the grasp from failing, this work will focus on the static case only and we will refer to the applicable friction coefficient as  $\mu$  from now on.  $F_n$  and  $F_l$  are the forces that can be applied through a frictional contact. They are related by  $\mu$  to form the friction cone, indicated with the gray area in Figure 1. The friction cone describes the equilibrium range of the pinch grasp. When the grasped object is in equilibrium, you are within the gray area as  $F_n$  is large enough to overcome  $F_l/\mu$  and the static part in Equation 1 holds. All points within the friction cone from Figure 1 will therefore result in a stable grasp. The edge case of stability is reached on the line  $\mu = F_l^*/F_n$ , where  $F_n$  and the maximum lateral force ( $F_l^*$ ) caused by the weight of the object are in exact equilibrium at the current frictional state. If the edge case is not yet reached, we can describe the distance from the line  $\mu$  with the frictional safety margin, on which we will elaborate in Section I-D.

We can increase the safety margin of a grasp by increasing  $F_n$ . However, forces can be limited by the grasping hardware, and it might not always be desirable to exert large forces on fragile or deformable objects. This limitation results in a friction cone of finite size, which can only evaluate grasp equilibrium when exact values for  $F_l$  and  $\mu$  are known. However, estimates of  $F_l$  fluctuate during manipulation tasks where dynamic movements come into play. Furthermore, the friction cone relies on a constant  $\mu$  throughout the grasping scenario. Equation 1 already showed that the friction coefficient can change between static and dynamic cases. A small overshoot in  $F_l$  can cause the object to slip indefinitely if the dynamic  $\mu$  is lower than the static  $\mu$ . Even when limiting ourselves to the static situation only,  $\mu$  can differ as a result of a variety of aspects. Variations in e.g. apparent contact area (by varying  $F_n$ ), temperature, humidity [5] or rapid dynamic movements [6] will result in variability of  $\mu$ . These variations will cause alterations in both magnitude and direction of the resulting force in the robotic fingertips, as displayed in Figure 1. To react accordingly, we need to measure the full spatial force distribution field at the entire robotic fingertip surface.

### B. Related work in robotic tactile sensing

Tactile sensors can measure the contact mechanics at the robotic fingertip surface needed to quantify the entire force distribution field at the surface. Next to the spatial force distribution, the measured surface can contain information about the object's shape, local texture or other material properties. Early research into robotic tactile sensing shows the use of strain gauges, acceleration, and pressure sensors to measure static and dynamic deformations up to 1000 Hz at the robotic fingertip skin [7]–[10]. The static deformations can be used to quantify the touch of an object, while the highly dynamic signals can aid in indicating the onset of slip, as movement in the local micro texture results in vibrations before the object completely slips.

A disadvantage in these works is the lack of spatial information. Using strain gauges to properly replicate the 2d receptive area of the human hand would require a plethora of sensors and requires high effort for robust integration and wiring. An indication of the onset of slip by an acceleration or pressure sensor is useful only if the magnitude of vibrations until full slip is known. Furthermore, these sensors can experience noise coming from other vibration sources such as the motors of the robot.

More recent works in artificial tactile sensing deploy a miniature camera that can track deformations in the soft skin of the robotic fingertip. These vision-based tactile sensors can achieve high spatial resolutions at moderately high sampling rates, with minimal integration effort. By equipping the inside of the fingertip surface with reflective colors and markers, the internal camera can detect object movements along the robotic fingertip surface. Subsequently, a model can be developed to map the obtained pixel values to perception metrics on slip and shear in multiple directions.

The GelSight/GelSlim family of tactile sensors [11]–[13] comprise of a spherical membrane. A set of RGB lights illuminates parts of the membrane, each reflecting on one-third of a special reflective coating on the inside of the sensing skin. Deformations in the tactile surface result in a changing color pattern which is converted to geometric height maps representing the object in contact. The addition of single-colored markers provide information about shear and slip. By calculating the entropy over the shear marker distribution, the authors are able to give an estimate of the stick-slip ratio. The stick-slip ratio is where part of the robotic finger, often the middle, is still stuck on the object, while other parts of the robotic finger have already slipped. The respective contact areas of the robotic fingertips representing the stick and slip cases can be divided to obtain the stick-slip ratio. Estimating this ratio allows for classifying the onset of slip by comparing against a predefined threshold. However, these works have to manually provide a threshold to classify the onset of slip, which can vary amongst different gripper-object interfaces. Furthermore, no measure of accuracy in slip detection is mentioned in any of these works, and they state to have

difficulties with smooth objects and small variations in surface geometries. Lastly, their algorithm needs to be simplified for improved prediction speed for it to be used in real-time control tasks. More specifically, one of these GelSight works used a machine learning approach [11]. Here, a convolutional neural network (VGG-16 net) is trained to learn the relationship between tactile images, and forces sensed by a separate sensor. For some basic shapes (sphere, cylinder, flat) the neural net learned from tactile image inputs the in-plane normal and lateral force, and torque around the axis perpendicular to the sensor. However when evaluating, the network shows moderate performance on similar objects and is not able to generalize to new objects or contact conditions.

A similar work using another vision-based tactile sensor also train a convolutional neural network (ShuffleNetV2) to reproduce the 3d force displacement field [14]. They developed a tactile sensor with single-colored fluorescent markers, randomly distributed over a tactile sensing dome [15]. When training the tactile images and their respective polar coordinates against ground-truth force data, they are able to discriminate linear from rotational slip. When evaluating on unseen objects with different geometric properties, they are able to manipulate the amount of rotational slip and perform successful pendulum swing-up manipulation tasks.

Another vision-based tactile sensor is the TacTip [16], that increases single-color marker movements by placement on pins, thereby making the sensor more sensitive to small deformations. The authors also adopt machine learning approaches to classify their tactile data for ‘slipped’ and ‘static’ objects during robotic grasping. In a recent TacTip work [17], a support vector machine is trained on pin velocity data. The pin velocities are obtained by extracting marker locations and taking the difference between two consecutive images. The train data for the support vector machine, which is a binary classifier, were manually labeled with the labels ‘slip’ and ‘static’ by investigation of the velocity displacement field. During evaluation, the authors obtained a slip detection classification accuracy of 99.88%. In a following work [18], a convolutional neural network is trained on raw input images to perform edge detection. The addition of machine learning improved the results compared to the probabilistic model developed in their previous work [19].

Despite these impressive results, all these works focus on a binary classification of slip or static situations. This classification in two groups will only result in feedback once the object starts sliding, which will put strong requirements on the real-time controller to reinstate stability. Furthermore, numerous cases can be thought of where the object should not even move in the first place. We can conclude that when performing highly dynamic manipulation tasks, as for robotic pick-and-place, these classification systems do not provide the crucial information for preventing dropping objects.

### C. Human grasp force regulation

Humans have, as opposed to the previously presented works in robotics, a more distinct skill set to regulate grasp force during manipulation. Humans use multiple sensory inputs during manipulation tasks, such as visual, auditory and proprioceptive feedback [20], [21]. Although we know that inferring close-contact information from the object is predominantly done by the four mechanoreceptor types forming the human sense of touch [21]–[23]. These mechanoreceptors are excited by different dynamic events at which neural spikes will be sent to the central nervous system for processing. Our brain translates the neural activity within 15 ms to a proper response for our motor control system [24]. Recent work [25] explains that the neural activity is likely to be efficiently encoded, resulting in fast response times of 100 - 150 ms from the tactile sense to motor control [26]. When discussing the quality of a human grasp, the frictional safety margin is often used to describe the distance between the frictional strength of the grasp and the current external load forces acting on the object [27], [28]. The safety margin describes how far the grasp is from failing. While maintaining grip, humans are able to keep these margins within 10 - 40% of the minimum required grip force, dependent on the desired slip probability and the estimated environmental uncertainty [21], [29].

### D. Safety margin in grasp force control

The frictional safety margin can be evaluated by comparing only the current lateral force to the maximum lateral force reached at full sliding. Using the margin in grasp force control provides advantages over evaluation of the friction cone as it is more resistant against varying friction conditions. From a contact mechanics point-of-view, the safety margin provides us with an understanding about the amount of stiction at the sensor surface. From a motor control point-of-view, the safety margin gives us an estimate how far we are from full slippage, i.e. the point of minimum required grip force to overcome the lateral force at the current frictional state. The safety margin is a relative estimation between the current lateral force  $F_l$  and the critical lateral force  $F_l^*$  at which slippage occurs, and is calculated as shown in Equation 2.

$$SM(t) = \frac{F_l^* - F_l(t)}{F_l^*} \quad (2)$$

The frictional safety margin relates to the friction cone as depicted in Figure 1. It ranges from 100% when normal force is present but zero lateral force is pulling on the object, to 0% when  $F_l \geq \mu \cdot F_n$  and the object is at the onset of slip. Equation 2 only includes components in the lateral force, making the safety margin resistant to variations in (unpredictable) contact mechanics. Therefore, estimating the safety margin using tactile sensing is the logical step in improving grasping robustness.

### E. New ChromaTouch tactile sensor

This thesis work advances on the state of the art in vision-based tactile sensing, by presenting a new iteration in the ChromaTouch family, initially developed and improved in [30], [31]. As part of this thesis work, a new version of the tactile sensor is produced and presented in [32]. Section II will elaborate on the design and manufacturing of the new tactile sensor.

### F. Contributions

This work develops a perception framework capable of estimating grasp equilibrium using the frictional safety margin. The acquired estimates provide enough accuracy to allow for grip force manipulation at 20% of the minimum required grip force, mimicking the efficiency of a human grasp which usually keeps the safety margin below 10 - 40%. To achieve this goal, we developed our own vision-based tactile sensor according to the design principles first presented in [30]. This sensor produces high-resolution tactile images describing the tactile surface displacement field. A perception framework based on a convolutional neural network is trained to estimate the frictional safety margin for the current grasp by only using these tactile images as input. The output result is a value ranging from 100 - 0% indicating the margin to an unstable grasp. Using this metric, we are able to estimate the onset of slip, invariant to changes in friction caused by changing environments. To best of the author's knowledge, this is the first work presented where the frictional safety margin is used to improve the quality of a robotic grasp.

### G. Outline

The following section presents the design principles and manufacturing process of the newly developed tactile sensor. Section III discusses the experimental setup and data collection to establish the perception framework. The results will be presented in Section IV. Finally, in Section V the results will be discussed and this work is concluded by comparing the performance and limitations with closely related literature.

## II. HARDWARE

The first part of this thesis work is focused on developing the tactile sensor. The design is a new iteration on the ChromaTouch tactile sensor, where advancements have been made in resolution, robustness, and manufacturability of the sensor.

### A. Improved high-resolution tactile sensing

State-of-the-art robotic tactile sensors exploit the high resolution of miniature cameras to perceive deformations from the sensor surface skin. In earlier work from our lab [30], a new type in this class of vision-based tactile sensors is presented. Compared to the related works, our tactile sensor has increased resolution for detecting partial slip by measuring the full three-dimensional deformation field. The sensor dome encompasses two layers of overlapping markers, translucent

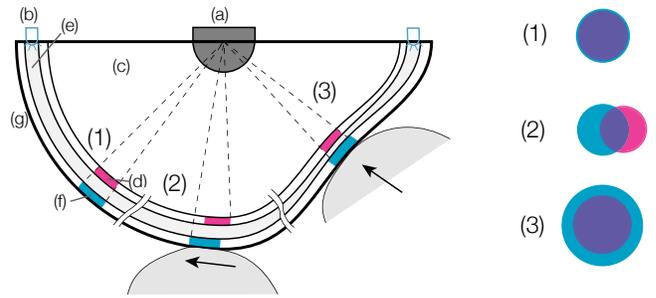


Fig. 2. The subtractive color mixing principle by means of two shifting marker layers. **1**) shows the position of the markers when in rest. The marker placement provides sensitivity to **2**) shearing force, and **3**) normal indentation. On the top, from the inside out: **a**) the fish-eye camera maps the hemispherical dome with markers to the flat surface, **b**) LED ring, **c**) empty, open area allows for large deformations of the sensing skin, **d**) inner layer with magenta markers, **e**) soft silicone intermediate layer, **f**) outer layer with cyan markers, **g**) outside white silicone layer diffuses internal light and blocks external influences. Adapted with permission from the author from [32].

and opaque, which can be tracked simultaneously by the camera. This method fully exploits all three RGB color channels from the camera at all marker locations, thereby increasing the sensing resolution compared to the other related works.

This first spherical version of the ChromaTouch is presented in [31]. Fabrication involved a complex process where the elastomeric skin is molded in the 2d surface, and then folded to form a 3d sphere. Apart from the residual stress in the skin causing decreased resolution, this process creates lines where the four quadrants are glued together. The authors argue that these lines decrease resolution as no markers can be present here. Furthermore, sensor robustness is decreased as these are possible points of failure.

During the time span of this thesis work, a new iteration of the ChromaTouch vision-based tactile sensor is developed, which is presented here and described in more detail in [32]. The main contribution of that work is a simplification of the manufacturing process, where the spherical dome can be printed in its 3d shape directly. Apart from ease of manufacturing, this offers more flexibility in marker placement and alignment. As the dome can be printed in one piece, there are no break lines with possible cause for failure. In our related work [32], we present that up to 400 markers can be placed on the 42 mm outside diameter dome. In the current work, the 200 marker version is used.

### B. Sensor dome

The tactile sensing dome is the most vital part of the assembly, essential for transforming tactile information to deformations which can be captured by the camera. The colored markers are embedded within an inner and outer layer, separated by a distance  $\delta$ . The marker layers are modeled in Grasshopper<sup>1</sup>, a parametric CAD design tool, where the markers are equally distributed along the sensor

<sup>1</sup><https://www.rhino3d.com/6/new/grasshopper/>

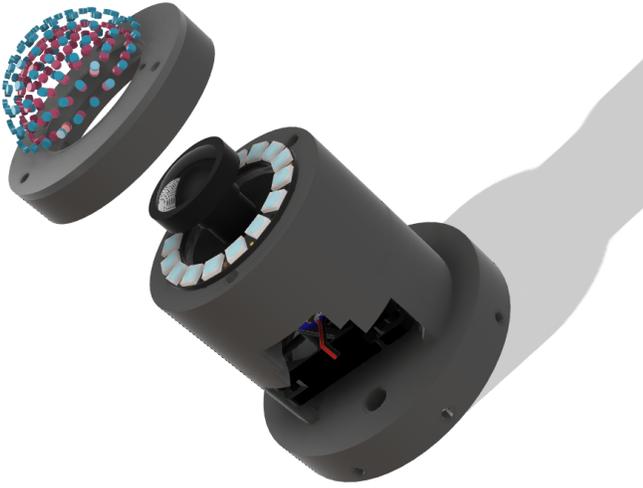


Fig. 3. An exploded view of the sensor assembly comprising the ChromaTouch tactile dome on top (white outside layer omitted for visibility), and the assembly holding the camera and fish-eye lens, LED ring, and a cooling fan for the camera body during long operations.

surface according to [33]. The resulting 3d CAD file is printed in two separate parts on a Stratasys J735 polyjet additive manufacturing printer<sup>2</sup>, which allows to create an all-in-one design including the various materials needed for the layers, embedded markers and a rigid frame for mounting. The layers supporting the colored markers consist of the AgilusClear30 elastomer with a shore hardness of 30A. The high softness of the material allows for large deformations in the tactile sensing dome, increasing the resolution when tracking subtle contacts. However, the soft material comes with the disadvantage of having strong viscoelastic effects. This is partly counteracted by casting elastic silicone (Smooth-On SORTA-Clear 12) in between the marker layers, and on the outside of the dome, thereby improving the sensor’s time response. The outside silicone layer is mixed with white pigment, to diffuse the internal light along the inside of the dome. This provides equally distributed lighting conditions at all marker locations while minimizing internal reflections. Furthermore, the white outside layer acts as a barrier against perturbations in external light. Figure 2 illustrates the principle of subtractive color mixing. More details on the design principles and manufacturing choices of the current version of ChromaTouch can be found in our previous work [32].

### C. Electronics and housing

For easy and robust use of the tactile sensor, a handheld device is designed that accommodates the sensor dome, camera and all electronics needed for operation. The housing is modular, allowing for easy repairs or modifications, and it can be used in-hand or mounted on a robot. The sensor dome on top can be replaced easily if we experience effects of wear and tear. The integrated USB-camera (Basler Dart daA1600-

<sup>2</sup><https://www.stratasys.com/-/media/files/printer-spec-sheets/j735-j750-3d-printers-spec-sheet.pdf>

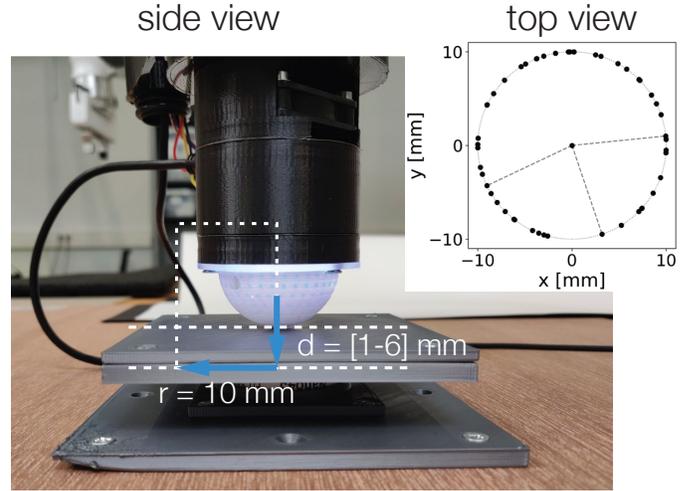


Fig. 4. Side and top view of the experimental setup for data collection. From the side, the sensor moves an indentation depth  $d = [1-6]$  mm into the surface, after which a 10 mm lateral sliding movement is executed, reaching full slippage of the sensor. Subsequently, the sensor moves back up and resets to the origin position. Data gathered at the solid lines is included in the dataset. The top view shows the target positions, randomly sampled on a circle with radius  $r=10$  mm. The sensor indents to the desired depth at the center point, after which the robot will perform a linear motion to any of the 50 sampled points, holding the indentation depth constant. 3 sliding trials are indicated with the dashed lines. The indicated points are sampled for the 4 mm indented train dataset.

60uc with a Basler Evetar M118B029520IR fish-eye lens) is used for marker tracking. The camera has a maximum spatial resolution of 1600 x 1200 px at a sampling frequency of 60 Hz. The lens comes with a 178 degree viewing angle, which maps the markers from the hemisphere onto the camera sensor. Furthermore, the focal point is placed somewhat inside the hemispherical surface, to optimize focus at several indentation depths. These two design choices minimize the need for digital preprocessing when we are integrating the tactile perception framework. Lighting is provided by an Adafruit NeoPixel Ring 16. The 16 RGB LEDs provide enough light to overcome external influences, and allow the camera to be used at short exposure times. The integrated LED drivers allow for straightforward control using a microprocessor capable of providing high frequency Pulse-width Modulation signals, such as the popular Arduino Uno. The complete sensor assembly with the ChromaTouch tactile dome, fish-eye lens, camera, lighting, and housing is shown in Figure 3.

## III. METHODS

### A. Experimental setup

An experimental setup is developed to obtain tactile image data relevant for training our tactile perception framework on the frictional safety margin. For ease of implementation, the tactile sensor is mounted directly on the wrist of a Universal Robots UR5 6 degrees of freedom robotic manipulator. Our initial experiment was to have an object sliding between two tactile gripper fingers. We can now translate this experiment to the horizontal plane, where the tactile sensor is sliding

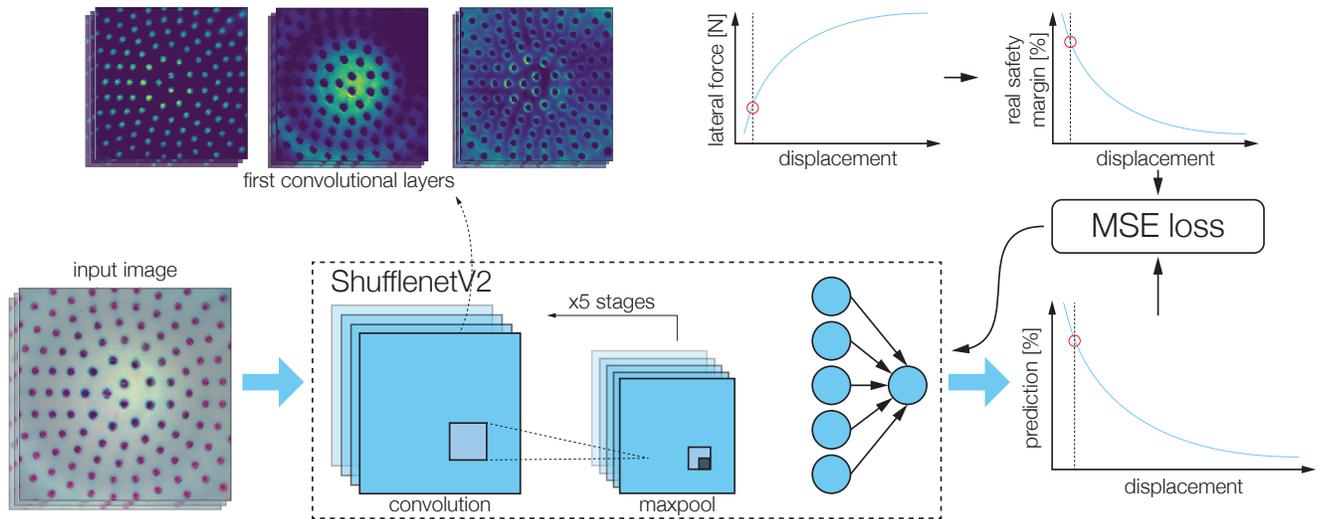


Fig. 5. This figure illustrates the complete pipeline of the tactile perception framework. During training, input images are fed one-by-one into the convolutional neural network. In this network based on the ShufflenetV2 architecture, 5 stages of convolutional and pooling layers are completed. Some of the activation functions of the first convolutional layer are shown in the top-left, indicating that the model has trained on several marker properties and on the image background. This information is condensed in the final fully connected layer to a single output, the safety margin prediction. The prediction is compared with the target (real) safety margin value using a Mean-Squared-Error loss function. Backpropagation is then used to optimize the network’s weights and biases according to the computed loss until sufficient performance is reached.

across the workspace surface. A side view of the experimental setup is displayed in Figure 4. The training phase of the convolutional neural network requires real safety margin values as target data. Therefore, we measure lateral sliding force using an ATI Nano43 6 degree of freedom force-torque sensor. A 3d-printed flat surface is mounted on the force sensor, on which the tactile sensor has ample space for sliding. The force sensor is connected through a National Instruments USB-6215 DAQ to a Windows laptop running MATLAB. Robot control is done with a separate Linux machine, which also collects the tactile images from our sensor. Both computers are synchronized using the network time protocol as described in Section III-D.

### B. Convolutional neural network training

In recent years, machine learning models such as the convolutional neural network have found increasing applications in vision for robotics [34]. Section I-B provided several examples of vision-based tactile sensors, which all incorporate convolutional neural networks to perceive contact mechanics from tactile images. The to us most closely related work [14] used ShufflenetV2, an implementation in the field of convolutional neural networks. ShufflenetV2 is a new iteration in a class of lightweight, mobile convolutional neural networks [35]. The updated version offers increased speed and accuracy compared to other lightweight networks such as MobilenetV2. We have resolved to using the same architecture. Due to the practical advantages of being able to train on a simple machine, this makes our work more suitable for widespread use. We trained all of our models on a mobile NVIDIA Quadro P1000 4GB GPU with 640 CUDA cores. For the full model, we trained the ShufflenetV2 architecture

to 30 epochs (iterations) in just under 3 hours. For simplicity reasons, our predictions are evaluated on six cores of a mobile Intel i7-8750H CPU @ 2.20/4.10 GHz.

The dataset is split in 70% train, 10% test and 20% evaluation data, and loaded into the network in batches with *batch size* = 32. The standard ShufflenetV2 architecture is used, with one adaptation to have the final fully connected layer connected to a single output node, which in our case represents the frictional safety margin value. During the training phase, the ShufflenetV2 weights and biases are adjusted to better fit the train dataset. For intermediate evaluation of the network, the mean-squared-error loss is calculated on the test dataset. Subsequently, the network’s weights and biases are updated during the backpropagation step. Here, we used the AdamW optimizer with the following standard parameters: *learning rate* =  $1^{-3}$ ,  $\beta$ 's = (0.9, 0.999),  $\epsilon$  =  $1^{-8}$ , *weight decay* =  $1^{-2}$ . These form the hyperparameters, together with the *batch size*, which is the amount of images used for training at the same time. Using the Optuna<sup>3</sup> framework, the evaluation dataset is used to tune the hyperparameters of our network. However, these newly trained parameters resulted in models which showed strong overfitting behavior. We want our model to generalize between various datasets, frictional contacts and indentation depths, so we reverted to the original parameters which showed better generalization performance.

### C. Data collection

As our goal is to make an estimator for the frictional safety margin, the convolutional neural network will be trained

<sup>3</sup><https://optuna.org/>

to model the relation between the tactile input images and target values containing the safety margin. This dataset is obtained with the following experimental procedure: The robot compresses the tactile image sensor into the center of the force sensor, to indentation depths in the discrete range [1-6] mm. When looking from the top, a circle can be drawn with the indentation point at the center, and a radius  $r=10$  mm. Points are sampled randomly on this circle, after which the robot moves the sensor to the desired points while keeping the indentation depth constant. To calculate the safety margin according to Equation 2, it is necessary that every movement continues until full sliding. The selected radius should be large enough to overcome deformations of the tactile sensing surface, and thus full slippage should be present between the sliding tactile sensor and the surface. When measuring lateral force over displacement and verifying the critical points, we obtained with some trial-and-error a desired sliding distance of  $r=10$  mm. The points are randomly sampled in the  $(x,y)$ -plane to allow the network to generalize to movements in both lateral directions. After reaching the sampled points, the sensor retracts from the surface, allowing for residual forces and sensor displacements to be released, and resets to the start position. Finally, we record the sliding trials at various indentation depths, enabling the model to generalize for varying grip forces. All movements are executed in a quasi-static manner, at 5% of maximum robot velocity ( $3.14$  m/s) and acceleration ( $3.14$  m/s<sup>2</sup>), to limit unwanted dynamic events in the tactile images. Figure 4 shows a top and side view of the data collection procedure. One full sliding trial, as indicated with the blue arrows in this figure, has a trial time of approximately 1 second.

The frictional safety margin can be calculated from the lateral force  $F_l$  as described by Equation 2. Therefore, the raw experiment data have to be shaped into safety margin target values. The force sensor provides 3-dimensional force data. We can compute the magnitude of the x and y dimensions to obtain the lateral force ( $F_l$ ) in Newton. By setting the gradient to  $\frac{dF_l}{ds_l} < 0.05$ , we can find the first critical point for zero safety margin. To limit the influence of noise at the boundaries of the gradient calculation, a feasible search range in lateral displacement ( $s_l$ ) of [2-6] mm has been defined on a trial-and-error basis.

During the experiments, tactile image data are collected from the sensor at 60 Hz at a resolution of 800 x 600 px. This resolution is more than twice the input size of the convolutional neural network, providing room for later alterations, if needed. The acquisition speed is high enough to provide continuous measurements during the sliding trials. Furthermore, position data from the robot's joints are collected at the maximum robot's communication speed of 125 Hz. This position data are used for plotting results and aids in defining the search range for the safety margin critical points. The force sensor provides the force data needed for training at a sampling frequency of 1000 Hz. Figure 5 shows the complete pipeline for training of the friction perception network.

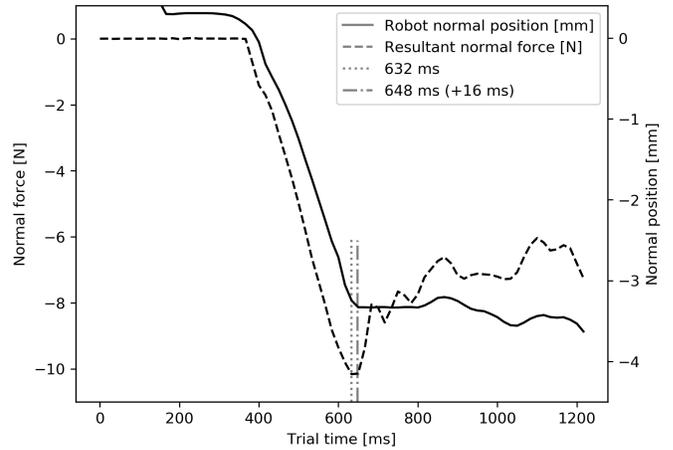


Fig. 6. After every collected dataset, a timing plot has been made to verify the alignment of the datasets. This figure shows two events over time: the robot position moving into the force plate (solid line) and the resultant normal force (dashed). Two vertical lines indicate where the indentation procedure stops. For this trial the time delay is measured to be 16 ms, equal to half the resolution of the slowest (position) sampling rate at 125 Hz.

#### D. Data synchronization

Data collection is done on a Windows and a Linux machine simultaneously. It is important that these data are synchronized, as otherwise essential information can get lost. For example, we determine the safety margin critical points in a range restricted by the position data. Furthermore, the tactile images are collected by a different computer than the force data. When these do not align perfectly, the model relates the wrong safety margin target value to the indentation image. When working with the high sampling rates mentioned above, it can be difficult to synchronize data at a later stage. Therefore, we should make sure that the data are already in sync during collection. This is achieved by connecting both computers on the same time Ethernet network. Using the network time protocol framework<sup>4</sup>, we can host a time server on the Linux computer. The network time protocol client on the Windows computer subscribes to this server and syncs its time to the nanosecond with the host. Figure 6 shows some data of both computers plotted together to verify that they are in sync. As the robot's normal indentation movement stops, there should be no further increase in normal force measured by the force sensor. The figure shows that both data types are synchronous within 16 ms, e.g. half the resolution of the (lowest) position sampling rate at 125 Hz.

<sup>4</sup><http://www.ntp.org/>

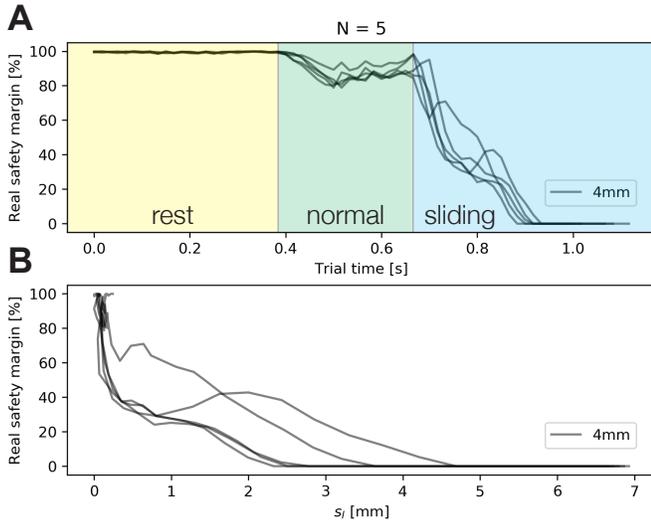


Fig. 7. This figure shows the typical trend of the safety margin, in **A** plotted over trial time, and in **B** over lateral displacement  $s_l$ . The raw data of 5 sliding trials are plotted in gray with the mean of the trials in black. It can be seen that a sliding trial consists of three phases. In the resting phase, both the tactile sensor and force plate are not yet in contact, resulting in a 100% safety margin. During the normal phase, the robot moves the tactile sensor perpendicular into the force plate, resulting in some noisy activation of the safety margin estimate due to non-linear effects in the force sensor. Finally, during the sliding phase, the robot moves the tactile sensor in the horizontal plane, causing the safety margin to decrease to 0% at full slippage.

#### IV. RESULTS

After all preparations have been done, the convolutional neural network presented in the previous section is trained. In this section, the behavior of the trained network is analyzed, and the resulting performance is shown.

##### A. Safety margin trend

Figure 7 shows several typical trends for the real safety margin value as measured by the force sensor in the experimental setup. In subplot **7A**, safety margin is plotted over the trial time duration. The plot can be divided in three stages. The first stage is the resting phase where the tactile sensor and force sensor are not in contact yet. The second stage is where the robot moves the tactile sensor perpendicular into the force plate. As can be seen, this results in a decrease of the measured safety margin, as the force sensor does not have a perfect response to shearing forces caused by imperfect orthogonal alignment. The third stage is where the robot starts moving in the lateral plane, which decreases the safety margin towards the critical point of 0%. Figure **7B** shows the same real safety margin measurement plotted against the lateral displacement  $s_l$  [mm]. This plot shows that the tactile sensor maintains some stiction to the force plate before full slippage occurs between [2-5] mm of lateral displacement. The noise around the 100% mark of the plot is caused by the nonlinear activation of the force sensor during the normal indentation phase.

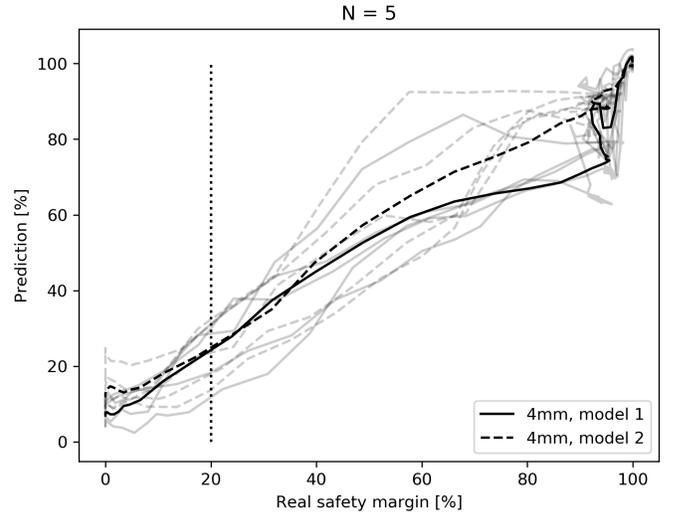


Fig. 8. This figure shows for two differently trained models the network prediction accuracy when interpolating at an unseen indentation depth. The means of the datasets are plotted in black lines, while the raw data are plotted in gray. Model 1 is trained at indentation depths 1, 2, 3, and 5 mm, whilst model 2 is trained at 2, 3, 5, and 6 mm. Evaluating both models at the unseen 4 mm indentation depth shows that both models have comparable interpolation performance, and they follow the real safety margin within a small margin of error.

##### B. Interpolation and extrapolation behavior

It is important that the trained network works at varying indentation depths which will occur during a normal robotic grasp. Our train and test datasets are all discretized to the 1 mm precision of the robot. Therefore, we can investigate the interpolation capabilities by removing an indentation depth from the train set, and subsequently evaluating the performance of our test set on that indentation depth. We have trained model 1 at indentation depths 1, 2, 3, and 5 mm, and model 2 at 2, 3, 5, and 6 mm. We have chosen to exclude 4 mm from the train dataset as it is roughly in the middle of the dataset to interpolate upon. Figure 8 shows an evaluation of both models at the unknown indentation depth, showing the interpolation behavior of the models. It can be seen that, apart from noise around the 100% mark, both models interpolate fairly well. They both show the same trend, and cross the 20% real safety margin line within respectable margin.

Next to interpolation, extrapolation behavior is important to maintain grasping performance at unseen ranges in indentation depth. Figure 9 shows the extrapolation capabilities for the two trained models from the interpolation section. Model 1 did not see any data during training at 6 mm, while for model 2 the 1 mm data were excluded. Figure 9 shows the performance of the two models at both indentation depths. The black lines again show the mean of the datasets, with the raw data from 5 sliding trials per dataset plotted in gray. We can see that the extrapolation performance of model 2 is worse than the performance of model 1 which has seen

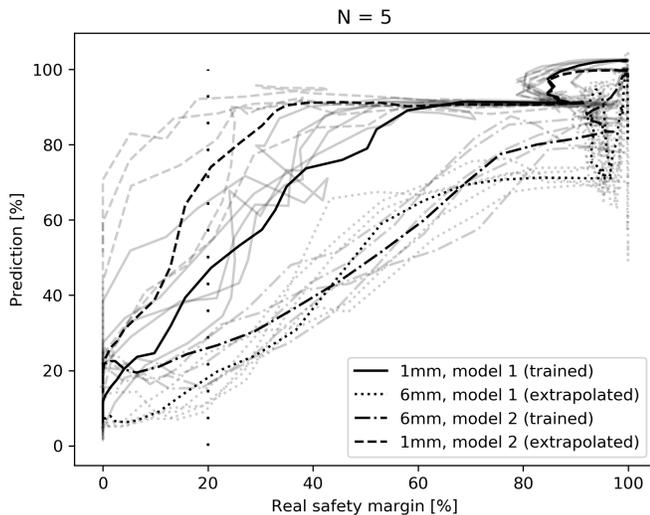


Fig. 9. This figure shows the extrapolation behavior of the two models from the interpolation section. The extrapolation capabilities are explored in the low and high indentation range. The figure shows that extrapolation to the low (1 mm) indentation range results in late activation and a general underestimation of the real safety margin. Extrapolation to the high (6 mm) indentation range results in generally better performance, but comes with more noise effects during the normal indentation phase.

data at this indentation depth. The point of first activation for the extrapolated model is where the real safety margin already decreased to 40%. We see that model 1 which has preliminary information at this indentation depth, is able to make more accurate estimates although the predicted safety margin of 40% is still an underestimation of the real margin of 20%. However, when performing the same type of evaluation for the extrapolation behavior at 6 mm indentation depth, we see that the mean performance of both models is much more similar compared to the 1 mm case.

These different results can be explained by comparing the signal strength resulting from both indentation depths, displayed in Table I. The table shows that the signal strength increases with indentation depth. Especially the low signal strength at 1 mm explains the models' slower activation, as there is a small amount of moving pixels which can be used to perform the safety margin estimation. For the 6 mm indentation depth, we can observe the opposite behavior. Due to the high signal strength, a lot of pixels are already moving when the sensor is indented into the force plate. This causes more pixels to activate and results in a decreasing safety margin before there is any lateral movement.

### C. Generalization behavior

One aim of this work is to show that the perception framework is robust against minor changes in frictional conditions. Therefore, train and evaluation data are gathered at multiple days. As explained in the introduction, changes in e.g. humidity will result in variability in the friction coefficient, as can be seen in Figure 10. It shows that there is a large variation in frictional coefficient ( $\mu$ ) between our evaluation dataset,

TABLE I

SIGNAL STRENGTH AT ALL SIX INDENTATION DEPTHS, MEAN AND STANDARD DEVIATION OVER ALL 5 TRIALS FROM THE EVALUATION DATASET. THE SIGNAL STRENGTH IS A METRIC FOR THE AMOUNT OF MOVING PIXELS IN THE IMAGE ON WHICH THE MODEL CAN ESTIMATE THE SAFETY MARGIN. IT IS CALCULATED AS FOLLOWS:  $\text{SUM}(I_l - I_n) / (\text{TOTAL PIXELS})$ , WITH  $I_l$  THE IMAGE DURING COMPLETE SLIDING, AND  $I_n$  DURING THE NORMAL INDENTATION PHASE.

	1 mm	2 mm	3 mm	4 mm	5 mm	6 mm
<b>Sig. strength [-]</b>	5.40	12.36	13.18	16.35	16.04	17.52
<b>std [-]</b>	0.36	1.37	0.25	0.51	0.19	0.56

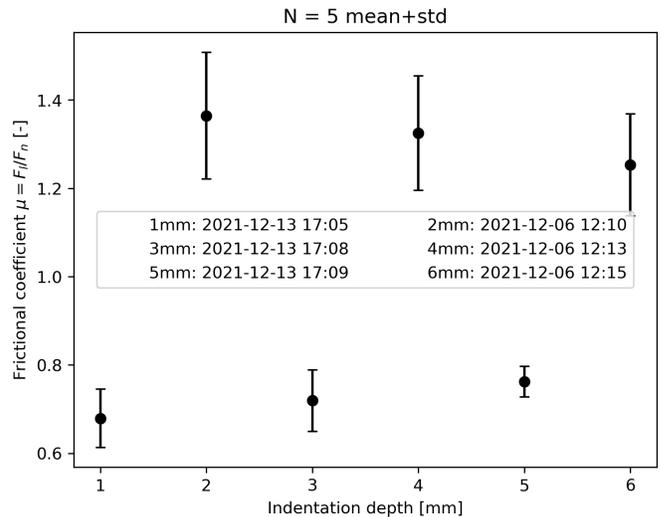


Fig. 10. The calculated friction coefficient  $\mu$  shows large variability between datasets calculated on different days, due to e.g. changes in environmental humidity. It can be seen that the (2,4,6) dataset follows the theory for decreasing  $\mu$  when increasing indentation depth, while the other dataset shows more unexpected behavior.

gathered at two different days. Furthermore,  $\mu$  in the 2, 4 & 6 mm dataset decreases with increasing apparent contact area, while the opposite is true for the 1, 3 & 5 mm dataset. These differences underline the importance of robustness against friction variability, as we show in this results section.

### D. Full model

The complete model is trained at all six indentations depths. Our total dataset is gathered during 300 sliding trials divided over the six indentation depths, resulting in 66.500 images or 100 GB of data. Figure 11 shows the obtained results evaluated at 1, 5 and 6 mm indentation. These three indentations are displayed here as 1 and 6 mm show considerably lower performance than the rest, and the most consistent result is obtained at 5 mm. The mean and standard deviation of all errors can be found in Table II. Figures of all results can be found in Appendix A. Figure 11 shows that 1 mm and 6 mm are still under- and overestimating. This is comparable to behavior found in the extrapolation models from Section IV-B. The results for the four other indentation depths all show to be close to the optimal diagonal between prediction and real safety margin.

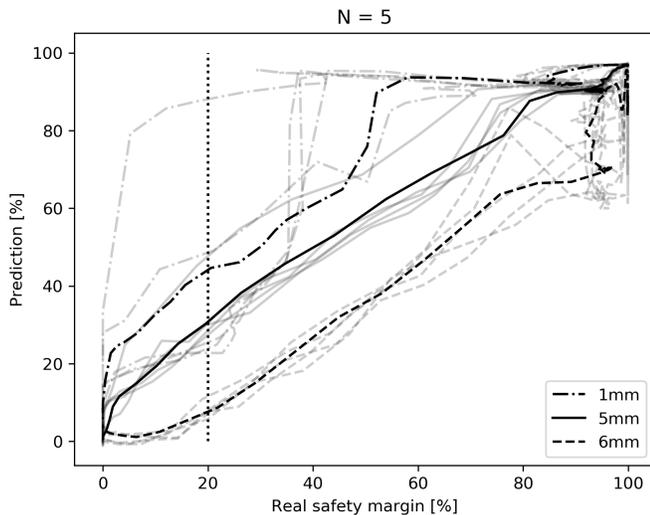


Fig. 11. Results for the complete model trained at all six indentation depths. It can be seen that both boundary scenarios (1 and 6 mm) are under performing compared to the rest of the model. This can be explained by fewer data available for the model to generalize here. The four other indentations showed comparable good performance, with the most consistent result at 5 mm showed here as well.

TABLE II

ERROR AND STANDARD DEVIATION OF THE COMPLETE TRAINED MODEL IN PERCENTAGE POINT [PP]. IT CAN BE SEEN THAT THE BEST RESULTS ARE OBTAINED AT 2, 3, 4, AND 5 MM INDENTATION DEPTH, WHILE THE TWO EDGE CASES SHOW LESS ACCURATE PERFORMANCE.

	1 mm	2 mm	3 mm	4 mm	5 mm	6 mm
<b>Error [pp]</b>	9.28	4.99	4.25	3.78	3.83	11.51
<b>std [pp]</b>	7.10	2.74	2.04	2.83	1.69	6.06

### E. Model speed

Next to having accurate results, the model is only useful when the prediction speed is fast enough to run in a real-time controller. On our evaluation set, the average prediction time on 360 images is  $19.07 \text{ ms} \pm 0.22 \text{ ms}$ . This allows for a controller to run at 50 Hz, which is close to our camera’s maximum acquisition rate of 60 Hz. If more speed is desired, predictions can be made on the above mentioned GPU instead of the CPU, but this comes at the cost of copying data to the video memory, possibly increasing the total prediction time.

## V. DISCUSSION & CONCLUSION

The goal of this thesis is to evaluate the stability of a robotic grasp by estimating the frictional safety margin from the gripper-object contact. Unfortunately, there are no current works in robotics which use the safety margin for evaluating grasp stability. It is therefore difficult to compare our performance and accuracy scores directly to other related works. The more closely related works in robotic literature are those which use vision-based tactile sensors to e.g. perform slip detection. The works on the GelSight/GelSlim [11]–[13] sensors predict slip based on the stick/slip ratio. However, as explained in the introduction, their predictions come with moderate performance. More

important: as they only *classify* slip, the point of activation is generally too late to be used for real-time control. The to us most closely related work is from Bi et al. [14], who use tactile images to obtain the contact force field in continuous time. They do not estimate the onset of slip directly, but by comparing the predicted forces to ground-truth values, they are able to make an accurate estimate on the onset of slip.

We have brought slip prediction a step further by directly estimating the frictional safety margin from raw RGB images. We have improved upon tactile sensing results obtained with the ChromaTouch tactile sensor compared to earlier works [30]–[32], where amongst others the sensor’s capability of exploring object shapes and curvatures were demonstrated. As part of the current work, we have improved the vision-based tactile sensor by developing a new manufacturing method, aiding in sensor versatility and robustness. Furthermore, this increased tactile resolution which is properly captured by a new vision system. Using the improved hardware, we have trained a tactile perception framework able to mimic human grasping behavior by estimating the frictional safety margin [25]. We have shown that we can estimate the safety margin to a high enough accuracy which allows the grasp to be controlled at 20% of the minimum required grip force. This is done at varying indentation depths, as will be the case in real-life grasping tasks with a controller on grip force. It is shown that the network architecture is capable of interpolating for these indentation depths in the whole mm space, leading us to the conclusion that it would most likely be able to perform smaller interpolations in the sub-mm space with comparable accuracy. We have found that the model shows decreased performance on the boundaries of the train set because the model has fewer data points to generalize to. It became apparent that the responsiveness of the model at low indentation depths is slow because of a lower signal strength in the tactile images. The best performance was found in the range of [2-5] mm indentation depth, with safety margin prediction errors  $< 5 \pm 3 \text{ pp}$ . Furthermore, these safety margin estimations are at a high enough refresh rate to be used in real-time control tasks. Lastly, we have shown that estimating the friction coefficient on our evaluation data gives highly variable results. This is in part due to the viscous effects of our dome, and can be eliminated by putting more strict quasi-static constraints on our data collection process. This high variability in friction indicates that grip force control based on the friction cone directly will likely be unsuccessful, whilst our safety margin estimator is robust against these variances.

The developed perception framework also comes with its limitations. Although the safety margin gives a clear indication of the amount of incipient slippage between gripper and object, it does not contain spatial information. Furthermore, despite it has been shown that the trained convolutional neural network generalizes well between parts of the training set, it shows

moderate performance at low indentation depths and at the boundary conditions. This can be partly solved by adapting the train dataset to better incorporate these lower indentations, but the relatively low signal strength in this range makes the safety margin estimate noisy. This can be improved by using a softer silicone during manufacturing, thereby lowering the force range to which the sensor is sensitive and increasing deformability. However, it should be noted that softer domes can show more viscous behavior, increasing response times. We have proven that fabrication of such sensors is possible, but these new prototypes are not yet tested in estimating the safety margin. Furthermore, changing sensor domes for varying tasks could be a solution, but no analysis is done about the perception framework’s robustness to these changes.

The dataset is collected using a specifically designed experimental procedure. If we are able to develop a training process using data more regularly encountered in robotic manipulation, we could train and improve the network online, as is done in other robotic works [36]. However, the real safety margin can only be computed when the real-life objects are being dropped, as we need to measure the critical point in lateral force. Even with the specific training setup, the safety margin estimation suffers from noise during normal indentation of the sensor. This is in part due to non-linear measurements from the force sensor, but it also has to do with the network’s incapability to capture dynamical phenomena, as it is trained on single images. Capabilities to capture dynamical phenomena can be implemented by training on sequences of input images, by e.g. using a Recurrent Neural Network, a Long Short-Term Memory network, or the now rapidly evolving Transformers [37], [38]. Furthermore, minimizing oscillations should improve the current quasi-static performance, as the tactile sensing surface is relying on its (fast) elastic response. Instead of actuating an entire arm, a more precise gripper can be used.

The last limitations to this work are on sensor robustness and manufacturing, and are not necessary limitations but more small drawbacks inherent to new soft robotic designs. After a several hundred sliding trials on dry surfaces, the tactile sensor starts to show some wear and tear. Especially the separately casted white diffusive outer layer shows signs of detachment, which influences the deformation image perceived by the camera. These effects are visible in some of our train and test data. This wear did not seem to have a noticeable influence yet, but might worsen the signal strength over time. Last, the addition of additive manufacturing greatly simplified the sensor construction with respect to [31], but naturally there is still room for improvement, as some manual labor remains with respect to casting silicone between the marker layers.

For future work, it would be interesting to move to a more realistic manipulation scenario with variations in speed, shape, and contact conditions. Quantifying and controlling rotational slip would be a topic of interest, as it is also controlled during human manipulation and is demonstrated

for other vision-based tactile sensors in the literature. Another compelling path would be to explore the indentation sensitivity for softer, more compliant, ChromaTouch sensors. A last recommendation for future research would be to explore the possibilities for a complete robotic gripper with ChromaTouch tactile sensors equipped on all fingertips.

That brings us to the conclusion of this thesis work. Our main contribution is the introduction of the frictional safety margin to robotic applications. We have shown that we can correctly estimate these margins, leading to improved robotic manipulation behavior. To the best of the author’s knowledge, no other works in robotics are available where the safety margin is estimated to process information from tactile sensors. We showed that the trained perception network is capable of estimating these safety margins with a high enough accuracy to allow for control at 20% of the minimum required grip force, thereby mimicking the efficiency of a human grasp. The obtained results are robust to fluctuations in frictional conditions within-trial (e.g. change in force changes the apparent contact area) and environmental conditions (humidity). Furthermore, because of this invariance to small fluctuations in friction, the perception algorithm should be able to generalize to measurements of new objects with comparable geometry and hardness. Although the perception network does not achieve 100% accuracy, we have shown that the safety margin estimates show accurate trends over time. Combined with a safety margin of 20% makes this work well suited for controlling the grip force during robotic grasping.

## VI. ACKNOWLEDGMENTS

I would like to thank Jens Kober and Michaël Wiertelwski from Delft University of Technology for defining this interesting thesis topic, and for providing me with continuous feedback throughout the entire graduation process. Furthermore, I would like to thank Michaël Wiertelwski, Padmaja Kulkarni and Jelle Luijkx for their daily supervision on the different thesis-related aspects, and providing me with relevant literature. We have had a lot of interesting discussions, which both helped broadening my horizon, as well as helped me converge to the relevant details of my thesis. I would like to thank Laurence Willemet for her expert knowledge on the field of friction safety margin estimation. Special thanks go out to Jelle Luijkx, who was eager to step into the project when I needed a new daily supervisor, helped me extensively with the lab experiments, was always in for a coffee break and has provided detailed feedback during the writing of this thesis. Furthermore, I would like to thank my fellow students from the Human-Robot Interaction department for the interesting meetings and feedback sessions, where we got to know each others projects extensively. I would like to thank Thomas Frateur and Taeke de Haan for their constructive feedback on the writing part of the thesis. Also, I would like to thank the IO Advanced Prototyping minor students Tiemen Kos, Mathijs de Heer, Mattia Strocchi, Luca Elbracht, and Mark Kruijthoff for their hard work into the development of early prototypes

of 3D printed tactile sensing domes. The last spotlight is for my parents and Anna van Soest, who helped me through the difficult parts of the thesis and helped to enjoy the more memorable events.

## REFERENCES

- [1] T. Zhang, W. Zhou, F. Meng, and Z. Li, "Efficiency analysis and improvement of an intelligent transportation system for the application in greenhouse," *Electronics (Switzerland)*, vol. 8, no. 9, 2019.
- [2] B. Zou, R. De Koster, Y. Gong, X. Xu, and G. Shen, "Robotic Sorting Systems: Performance Estimation and Operating Policies Analysis," *Transportation Science*, vol. 55, no. 6, pp. 1430–1455, 2021.
- [3] M. Hans, B. Graf, and R. D. Schraft, "Robotic home assistant care-O-bot: Past-present-future," in *Proceedings - IEEE International Workshop on Robot and Human Interactive Communication*, 2002, pp. 380–385.
- [4] A. Bicchi, "Hands for dexterous manipulation and robust grasping: A difficult road toward simplicity," *IEEE Transactions on Robotics and Automation*, vol. 16, no. 6, pp. 652–662, 2000.
- [5] M. J. Adams, S. A. Johnson, P. Lefèvre, V. Lévesque, V. Hayward, T. André, and J. L. Thonnard, "Finger pad friction and its role in grip and touch," *Journal of the Royal Society Interface*, vol. 10, no. 80, 2013.
- [6] S. M. Rubinstein, G. Cohen, and J. Fineberg, "Detachment fronts and the onset of dynamic friction," *Nature*, vol. 430, no. 7003, pp. 1005–1009, 2004.
- [7] M. R. Cutkosky and R. D. Howe, "Dynamic tactile sensing," in *Romansy 88: Seventh CISM-IFTOMM Symposium on the Theory and Practice of Robots and Manipulators*, September 1988.
- [8] C. H. Chuang, M. S. Wang, Y. C. Yu, C. L. Mu, K. F. Lu, and C. T. Lin, "Flexible tactile sensor for the grasping control of robot fingers," *2013 International Conference on Advanced Robotics and Intelligent Systems, ARIS 2013 - Conference Proceedings*, pp. 141–146, 2013.
- [9] R. D. Howe and M. R. Cutkosky, "Howe, Cutkosky - Sensing skin acceleration for slip and texture perception," *IEEE*, pp. 145–150, 1989.
- [10] N. Wettels, J. A. Fishel, and G. E. Loeb, "Multimodal tactile sensor," *Springer Tracts in Advanced Robotics*, vol. 95, no. 0912260, pp. 405–405, 2014.
- [11] W. Yuan, S. Dong, and E. H. Adelson, "GelSight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors (Switzerland)*, vol. 17, no. 12, 2017.
- [12] S. Dong, W. Yuan, and E. H. Adelson, "Improved GelSight tactile sensor for measuring geometry and slip," *IEEE International Conference on Intelligent Robots and Systems*, vol. 2017-Sept, pp. 137–144, 2017.
- [13] I. Taylor, S. Dong, and A. Rodriguez, "GelSlim3.0: High-Resolution Measurement of Shape, Force and Slip in a Compact Tactile-Sensing Finger," 2021. [Online]. Available: <http://arxiv.org/abs/2103.12269>
- [14] T. Bi, C. Sferrazza, and R. D'Andrea, "Zero-shot sim-to-real transfer of tactile control policies for aggressive swing-up manipulation," *arXiv cs.RO (preprint)*, 2021. [Online]. Available: <http://arxiv.org/abs/2101.02680>
- [15] C. Sferrazza and R. D'Andrea, "Design, motivation and evaluation of a full-resolution optical tactile sensor," *Sensors (Switzerland)*, vol. 19, no. 4, 2019.
- [16] N. F. Lepora, "Biomimetic Active Touch with Fingertips and Whiskers," *IEEE Transactions on Haptics*, vol. 9, no. 2, pp. 170–183, 2016.
- [17] J. W. James, N. Pestell, and N. F. Lepora, "Slip detection with a biomimetic tactile sensor," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 3340–3346, 2018.
- [18] N. F. Lepora, A. Church, C. de Kerckhove, R. Hadsell, and J. Lloyd, "From pixels to percepts: Highly robust edge perception and contour following using deep learning and an optical biomimetic tactile sensor," *arXiv*, vol. 4, no. 2, pp. 2101–2107, 2018.
- [19] N. F. Lepora, K. Aquilina, and L. Cramphorn, "Exploratory Tactile Servoing with Active Touch," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1156–1163, 2017.
- [20] A. Amedi, K. Von Kriegstein, N. M. Van Atteveldt, M. S. Beauchamp, and M. J. Naumer, "Functional imaging of human crossmodal identification and object recognition," *Experimental Brain Research*, vol. 166, no. 3–4, pp. 559–571, 2005.
- [21] 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, pp. 345–359, 2009. [Online]. Available: [www.nature.com/reviews/neuro](http://www.nature.com/reviews/neuro)
- [22] A. S. Augurelle, A. M. Smith, T. Lejeune, and J. L. Thonnard, "Importance of cutaneous feedback in maintaining a secure grip during manipulation of hand-held objects," *Journal of Neurophysiology*, vol. 89, no. 2, pp. 665–671, 2003.
- [23] A. G. Witney, A. Wing, J. L. Thonnard, and A. M. Smith, "The cutaneous contribution to adaptive precision grip," *Trends in Neurosciences*, vol. 27, no. 10, pp. 637–643, 2004.
- [24] R. S. Johansson and I. Birznieks, "First spikes in ensembles of human tactile afferents code complex spatial fingertip events," *Nature Neuroscience*, vol. 7, no. 2, pp. 170–177, 2004.
- [25] L. Willemet, N. Huloux, and M. Wiertlewski, "Efficient tactile encoding of object slippage."
- [26] K. J. Cole and J. H. Abbs, "Grip force adjustments evoked by load force perturbations of a grasped object," *Journal of Neurophysiology*, vol. 60, no. 4, pp. 1513–1522, 1988.
- [27] K. J. Cole and R. S. Johansson, "Transformation for Grip Responses To Pulling Loads," *Experiment*, vol. 95, no. 1993, pp. 523–532, 1993.
- [28] T. Zackrisson, B. Eriksson, N. Hosseini, B. Johnels, and A. L. Krogstad, "Patients with hyperhidrosis have changed grip force, coefficient of friction and safety margin," *Acta Neurologica Scandinavica*, vol. 117, no. 4, pp. 279–284, 2008.
- [29] A. M. Hadjiosif and M. A. Smith, "Flexible control of safety margins for action based on environmental variability," *Journal of Neuroscience*, vol. 35, no. 24, pp. 9106–9121, 2015.
- [30] X. Lin and M. Wiertlewski, "Sensing the Frictional State of a Robotic Skin via Subtractive Color Mixing," *IEEE Robotics and Automation Letters*, vol. 4, no. 3, pp. 2386–2392, 2019.
- [31] X. Lin, L. Willemet, A. Bailleul, and M. Wiertlewski, "Curvature sensing with a spherical tactile sensor using the color-interference of a marker array," *IEEE International Conference on Robotics and Automation (ICRA)*, no. c, 2020.
- [32] R. B. N. Scharff, D.-J. Boonstra, L. Willemet, X. Lin, and M. Wiertlewski, "Rapid manufacturing of color-based hemispherical soft tactile fingertips," 2021. [Online]. Available: <http://arxiv.org/abs/2112.05536>
- [33] M. Deserno, "How to generate equidistributed points on the surface of a sphere," p. 55128, 2004.
- [34] S. Luo, J. Bimbo, R. Dahiya, and H. Liu, "Robotic tactile perception of object properties: A review," pp. 54–67, dec 2017. [Online]. Available: [www.elsevier.com/locate/mechatronics](http://www.elsevier.com/locate/mechatronics)
- [35] N. Ma, X. Zhang, H. T. Zheng, and J. Sun, "Shufflenet V2: Practical guidelines for efficient cnn architecture design," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11218 LNCS, pp. 122–138, 2018.
- [36] R. Zenha, B. Denoun, C. Coppola, and L. Jamone, "Tactile Slip Detection in the Wild Leveraging Distributed Sensing of both Normal and Shear Forces," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, sep 2021, pp. 2708–2713. [Online]. Available: <https://ieeexplore.ieee.org/document/9636602/>
- [37] Y. K. Chia, S. Witteveen, and M. Andrews, "Transformer to CNN: Label-scarce distillation for efficient text classification," no. Nips, 2019. [Online]. Available: <http://arxiv.org/abs/1909.03508>
- [38] C. Ge, Y. Liang, Y. Song, J. Jiao, J. Wang, and P. Luo, "Revitalizing CNN Attention via Transformers in Self-Supervised Visual Representation Learning," no. NeurIPS, pp. 1–14, 2021. [Online]. Available: <http://arxiv.org/abs/2110.05340>

APPENDIX

A. Final results, presented per indentation depth

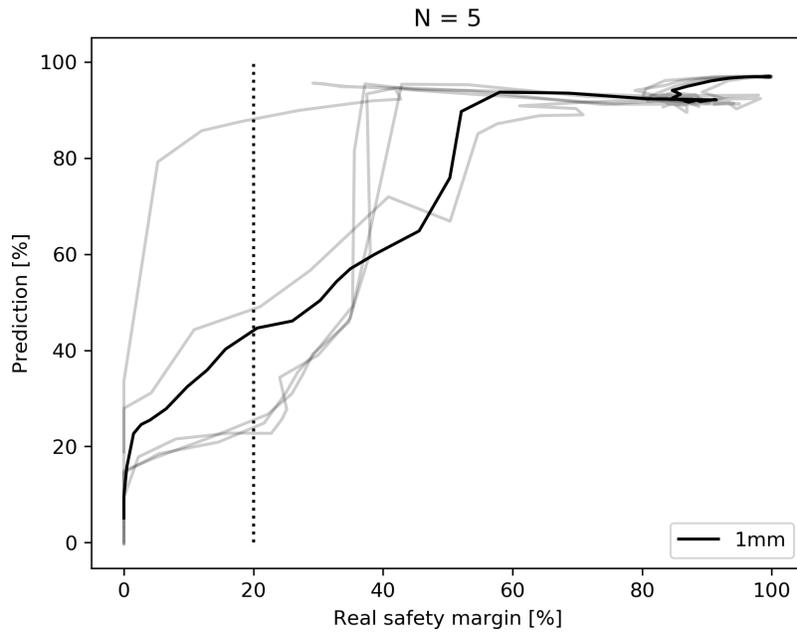


Fig. 12. Prediction accuracy of the safety margin at 1 mm indentation depth. The 5 trials in grey are plotted with the mean in the dark black line.

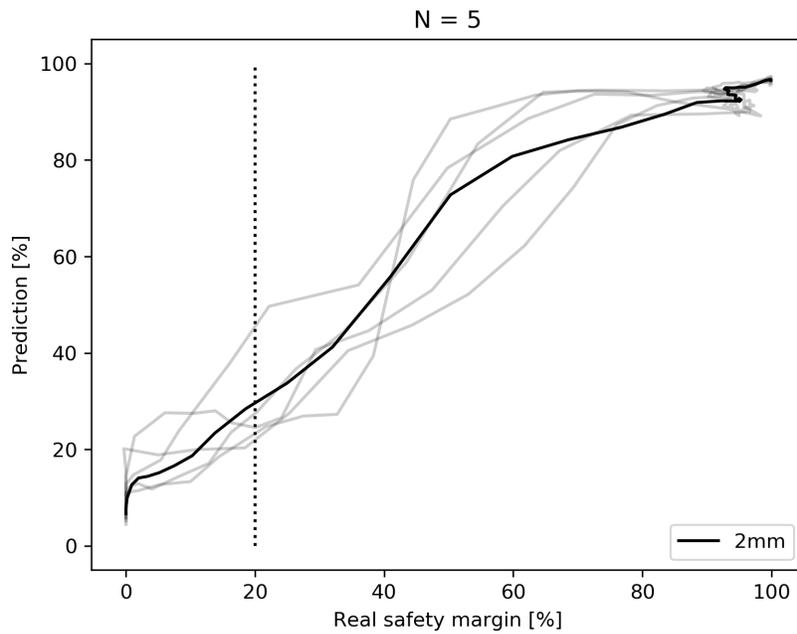


Fig. 13. Prediction accuracy of the safety margin at 2 mm indentation depth. The 5 trials in grey are plotted with the mean in the dark black line.

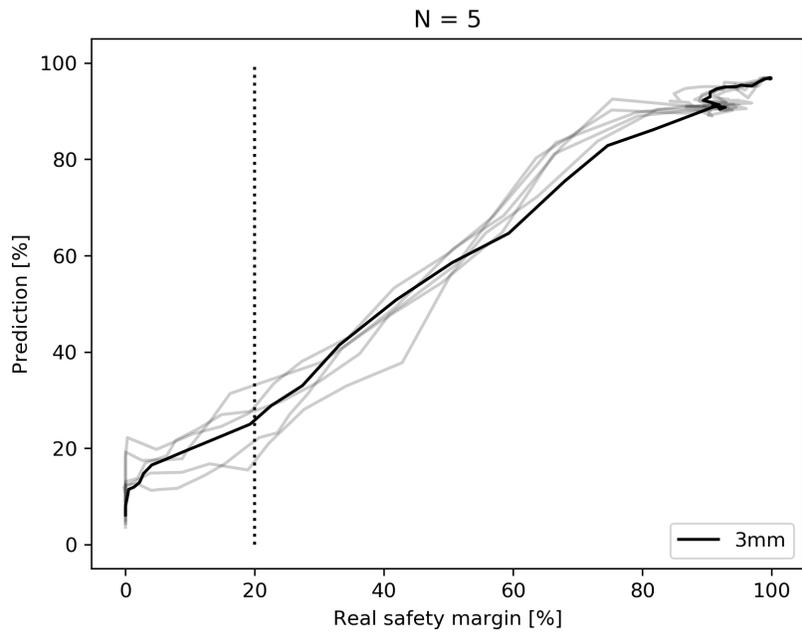


Fig. 14. Prediction accuracy of the safety margin at 3 mm indentation depth. The 5 trials in grey are plotted with the mean in the dark black line.

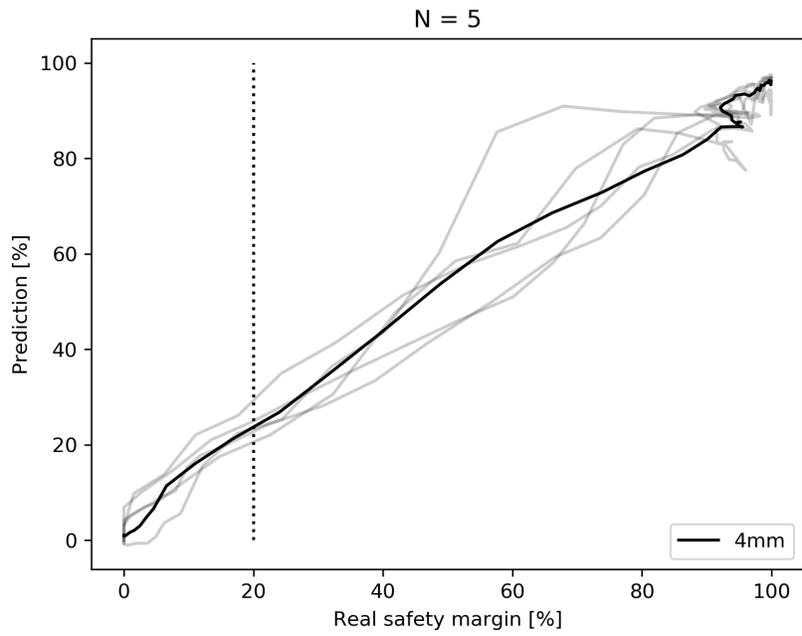


Fig. 15. Prediction accuracy of the safety margin at 4 mm indentation depth. The 5 trials in grey are plotted with the mean in the dark black line.

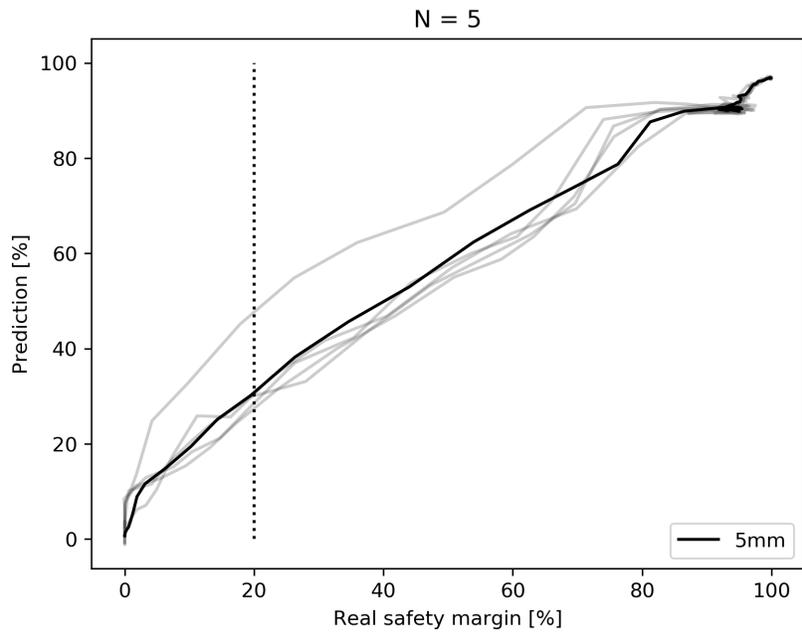


Fig. 16. Prediction accuracy of the safety margin at 5 mm indentation depth. The 5 trials in grey are plotted with the mean in the dark black line.

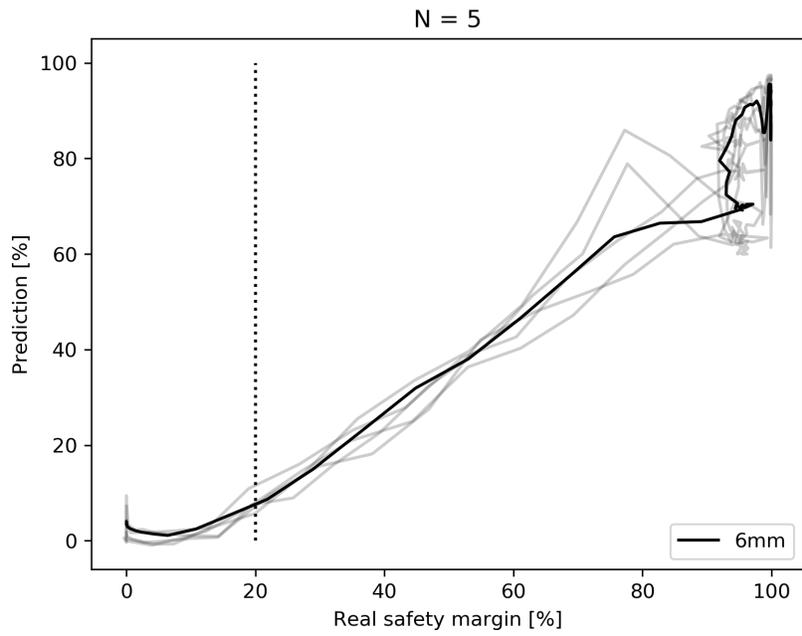


Fig. 17. Prediction accuracy of the safety margin at 6 mm indentation depth. The 5 trials in grey are plotted with the mean in the dark black line.