Ground Truth Force Distribution for Learning-Based Tactile Sensing: A Finite Element Approach

CARMELO SFERRAZZA[1]**, (Student Member, IEEE), ADAM WAHLSTEN**[2]**, CAMILL TRUEEB**[1]**, AND RAFFAELLO D'ANDREA**[1]**, (Fellow, IEEE)**

[1]Institute for Dynamic Systems and Control, ETH Zurich, 8092 Zurich, Switzerland
[2]Institute for Mechanical Systems, ETH Zurich, 8092 Zurich, Switzerland

Corresponding author: Carmelo Sferrazza (csferrazza@ethz.ch)

ABSTRACT Skin-like tactile sensors provide robots with rich feedback related to the force distribution applied to their soft surface. The complexity of interpreting raw tactile information has driven the use of machine learning algorithms to convert the sensory feedback to the quantities of interest. However, the lack of ground truth sources for the entire contact force distribution has mainly limited these techniques to the sole estimation of the total contact force and the contact center on the sensor's surface. The method presented in this article uses a finite element model to obtain ground truth data for the three-dimensional force distribution. The model is obtained with state-of-the-art material characterization methods and is evaluated in an indentation setup, where it shows high agreement with the measurements retrieved from a commercial force-torque sensor. The proposed technique is applied to a vision-based tactile sensor, which aims to reconstruct the contact force distribution purely from images. Thousands of images are matched to ground truth data and are used to train a neural network architecture, which is suitable for real-time predictions.

INDEX TERMS Computer vision, finite element analysis, machine learning, soft robotics, tactile sensors.

I. INTRODUCTION

A growing number of applications require robots to interact with the environment [1] and with humans [2]. The use of soft materials for robotics applications [3] introduces intrinsic safety during interactive tasks [4]. In addition, precise estimation of contact forces is crucial for effective operation without damaging the robot's surroundings, e.g., for manipulation of fragile objects [5].

Modeling the interaction of soft materials with generic objects is highly complex. As a consequence, several tactile sensing strategies leverage the use of machine learning algorithms to map sensory feedback to the corresponding quantities of interest, e.g., contact forces, shape and materials, see [6]–[8]. These maps are generally retrieved by means of supervised learning techniques, which fit a model to a large amount of labeled data, i.e., sensory data paired with the corresponding ground truth.

The associate editor coordinating the review of this manuscript and approving it for publication was Chenguang Yang.

However, the estimation of the full contact force distribution purely from data is limited by the lack of a ground truth source that does not alter the interaction between the soft material and the objects in contact. This article aims to provide a systematic way of labeling data with ground truth for the three-dimensional force distribution, which is obtained in simulation through the finite element method (FEM).

The approach is evaluated on a vision-based tactile sensor, originally presented in [9], which uses a camera to track spherical particles within a transparent gel. Hyperelastic models of the sensor's materials are retrieved from state-of-the-art material characterization tests, which are fully independent of the evaluation experiments. A label vector representing the ground truth force distribution is assigned to each image collected during an automatic indentation procedure. The total contact force also retrieved from the FEM simulations shows a satisfactory agreement with the measurements obtained from a commercial force-torque (F/T) sensor.
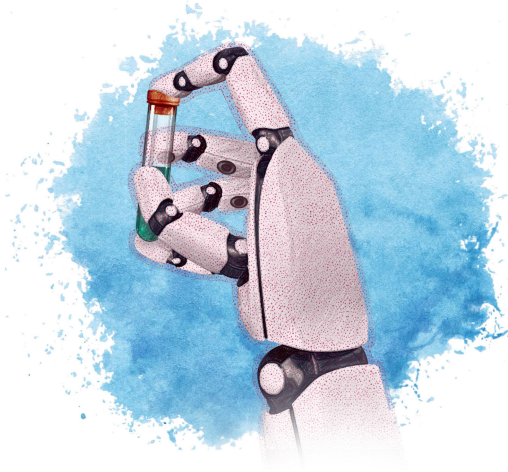
FIGURE 1. The tactile sensing technique presented in [9] is suitable to cover large surfaces of arbitrary shape and dimension. A concept of a robotic hand covered with cameras and a transparent gel embedding a spread of red particles is shown in this figure.

The dataset generated with the strategy proposed here is then used to train a deep neural network (DNN) architecture [10], which maps optical flow features to the contact force distribution. The evaluation of this strategy is carried out in a specific indentation setup, with the resulting pipeline running in real-time on the CPU of a standard laptop computer (dual-core, 2.80 GHz) at 40 Hz.

A. RELATED WORK

In recent decades, tactile sensing research has shown the potential of providing robots with the sense of touch, exploited both singly [11] or in combination with vision [12].

Among the various categories of tactile sensors, see [13], [14] for a survey, vision-based (or optical) tactile sensors are based on optical devices that monitor properties related to the contact between the sensor's surface (generally soft) and the environment. Among the advantages of this type of tactile sensors are high resolution, low cost, ease of manufacture and the preservation of the surface softness.

One category of optical tactile sensors uses a camera to track sparse markers within a soft, transparent gel, which deforms when subject to external forces, see for example [15], [16]. Other optical devices are able to provide information about the contact with the environment, as shown with the use of dynamic vision sensors [17] and depth cameras [18]. The sensor used here for the evaluation of the proposed approach is based on an RGB camera (which retains a small size) that tracks a dense spread of particles within a soft gel, see for example Fig. 1. This design is presented in [9] and shows performance advantages over sparse marker tracking, and ease of manufacture, without any assumptions about the surface shape.

Converting tactile information to quantities, such as the force distribution, which are of high relevance for many robotics tasks (e.g., grasping or manipulation), is not trivial. In fact, the main complexity is introduced by the absence of a generally valid closed-form model, which maps the deformation of a hyperelastic material to the external forces applied to it. The use of data-driven techniques aims to overcome this problem, approximating this map with a model learned from a collection of past data. In [19], an optical tactile sensor that exploits photometric stereo and markers painted on its soft surface is used to reconstruct the total contact force by means of a neural network architecture. In [20], an array of light emitters and receivers is placed below a soft gel to create tactile information, which is then provided to machine learning algorithms that reconstruct the location and the depth of an indentation, as well as the type of the employed indenter. Although these techniques generally require large datasets, transfer learning techniques can reuse information extracted across different sensors, as shown in [21].

The FEM [22] is a powerful numerical technique that provides approximate solutions of boundary value problems arising in engineering, by subdividing a large system into many smaller parts (called elements). One of the widespread applications of this technique is the analysis of the behavior of soft materials under various loading conditions. In [23], the silicone gel pad of an optical tactile sensor is modeled as a linear elastic material, and the FEM is used to compute the stiffness matrix that approximates the relation between external forces and displacements of the sensor's material. Based on reconstructed surface displacements, this matrix is then used to compute an estimate of the force distribution applied to the sensor. FEM simulations of a flexible 3D shape sensor are used in [24] to optimize design parameters. Furthermore, these simulations show the uniqueness of a strain-to-shape mapping for the case considered. In [25], theoretical justifications based on FEM equations are provided to reconstruct the external forces applied to soft bodies using information about their deformation.

The strategy followed in this article exploits FEM simulations to obtain ground truth data for the contact force distribution applied to the soft surface of a tactile sensor. The lack of ground truth data has so far prevented the development of learning-based tactile sensors that predict the full force distribution, limiting them to the estimation of simpler quantities, e.g., the resultant force and its location, and the depth of a contact. The hyperelastic models identified capture the full material behavior, including nonlinearities, rendering highly accurate simulations. Images collected in experiments on a vision-based tactile sensor are matched to the ground truth and used to train a DNN that reconstructs the force distribution with high accuracy and in real-time.

B. OUTLINE

The sensing principle and the hardware used for the evaluation are presented in Section II, while the material characterization is discussed in Section III. In Section IV, the dataset generation is described, from the collection of images to the approach proposed for assigning ground truth labels. The learning algorithm and the results are presented in Section V. Section VI concludes the article with a brief discussion.
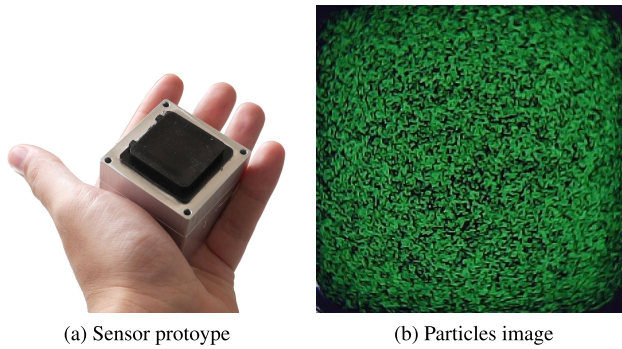
(a) Sensor protoype   (b) Particles image

FIGURE 2. The prototype of the sensor (designed for desktop testing) is shown in (a). The dense spread of green particles is captured by the camera placed inside the aluminum part. The resulting RGB image at a state of zero force is shown in (b).
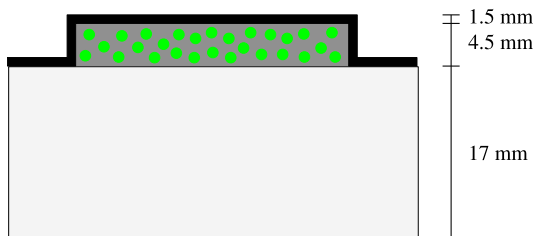

1.5 mm
4.5 mm

17 mm

FIGURE 3. A scheme of the three-layer structure that composes the soft material. The thickness of the different layers is shown in the figure above. This structure yields a top surface of 32 × 32 mm.

II. HARDWARE

The approach discussed in this article for generating ground truth labels is evaluated on a vision-based tactile sensor. The tactile sensing strategy is presented in [9], and is based on tracking the movement of spherical particles, which are randomly spread within a soft gel placed in front of a camera. The prototype used for the experiments and the camera image at a state of zero force are shown in Fig. 2.

The soft material is produced in a three-layer structure, as depicted in Fig. 3. From the bottom (which touches the camera lens) to the top surface, the following materials are used: 1) a stiffer layer (ELASTOSIL® RT 601 RTV-2, mixing ratio 7:1, shore hardness 45A); 2) the soft gel (Ecoflex™ GEL, mixing ratio 1:1, shore hardness 000-35) embedding the particles, which comprise 1.96 % of the layer volume; 3) a black surface layer (ELASTOSIL® RT 601 RTV-2, mixing ratio 25:1, shore hardness 10A). After curing, the complete sensor is placed in an oven at 60 °C for 8 hours. This step has the effect of reducing the aging of the materials, which is discussed in further detail in Section III.

III. MATERIAL CHARACTERIZATION

Finite element analysis (FEA) of arbitrary contact interactions with the sensor's soft surface requires material models that account for geometrical and material nonlinearities. Soft elastomers are often modeled as hyperelastic materials [26], and finding a suitable model formulation and corresponding parameters generally necessitates experimental

data from both uniaxial and biaxial stress states [27], [28]. To this end, a large-strain multiaxial characterization of the two most compliant materials, the Ecoflex GEL and the Elastosil 25:1, is performed. Samples of both materials are tested in uniaxial tension (UA), pure shear (PS), and equibiaxial tension (EB) based on previously described protocols [28]. The bottom layer of Elastosil with the mixing ratio 7:1 is considerably stiffer than the soft adjacent Ecoflex GEL, see Section III-E, and is therefore modeled as rigid in the subsequent FEA.

A. SAMPLE PREPARATION

Thin material sheets of each elastomer are prepared as described in Section II, and cast to a nominal thickness of 0.5 mm. Test pieces are cut to obtain gauge dimensions (length × width) of 40 mm × 10 mm for UA, 10 mm × 60 mm for PS, and a diameter of 30 mm for membrane inflation tests (EB). The test pieces of Ecoflex GEL used for the mechanical characterization do not contain the spherical particles, i.e. the pure material behavior is tested. An ink pattern is applied to the sample surface to facilitate optical strain analysis [28]. After each experiment, the sample thickness h_0 is measured on cross-sections cut from the central region using a confocal microscope (LSM 5 Pascal, Carl Zeiss AG) with a 10× objective in brightfield mode.

B. MECHANICAL TESTING

UA and PS tests are performed on a tensile testing set-up (MTS Systems) consisting of horizontal hydraulic actuators, 50 N force sensors, and a CCD-camera (Pike F-100B, Allied Vision Technologies GmbH) equipped with a 0.25× telecentric lens (NT55-349, Edmund Optics Ltd.) that captures top-view images of the deforming test piece. Displacement-controlled monotonic tests are performed up to a specified nominal strain (Ecoflex GEL: 200 %; Elastosil 25:1: 100 %) at a nominal strain rate of 0.3 %/s. The strain-rate dependence is analyzed in an additional UA test, where the sample is loaded cyclically with strain rates increasing from 0.1 %/s up to 10 %/s.

An EB state of tension is realized in a pressure-controlled membrane inflation test (see [28] for details). Briefly, a thin, circular sample is clamped on top of a hollow cylinder and inflated by means of a syringe pump (PhD Ultra, Harvard Apparatus), while a pressure sensor (LEX 1, Keller AG) measures the inflation pressure p. Top and side-view images are recorded with CCD cameras (GRAS-14S5C-C, Point Grey Research), and the image sequences are used for evaluating the in-plane deformation at the apex and the apex radius of curvature r, respectively.

All experiments are performed at room temperature and on the same day as completed curing. The mechanical properties of soft elastomers are known to change with aging [28], [29], a process attributed to additional, thermally activated cross-linking [29]. To assess the influence of aging, additional UA test pieces of the same sheets were kept at room temperature and tested several weeks after fabrication.

Material	μ_1 [kPa]	α_1 [-]	μ_2 [kPa]	α_2 [-]
Ecoflex GEL	7.9652	1.2769	0.3093	3.5676
Elastosil 25:1	85.1168	2.8991	−0.0020	−8.2915

C. EXPERIMENTAL DATA ANALYSIS

Since nominal strains computed from the clamp displacements are prone to errors due to sample slippage [30], the local in-plane principal stretches in the center of the test-piece, λ_1 and λ_2, are computed from the top-view image sequences using a custom optical flow-tracking algorithm [28]. The principal stretch in thickness direction is calculated by assuming material incompressibility, i.e., $\lambda_3 = 1/(\lambda_1\lambda_2)$. In the UA and PS configurations, the Cauchy stress in loading direction is evaluated as $\sigma = F\lambda/(w_0 h_0)$, where F are the measured force values, $\lambda := \lambda_1$ is the principal stretch in loading direction, and w_0 is the reference width of the test piece. For inflation tests, the measured inflation pressure and apex radius of curvature can be used to approximate the equibiaxial Cauchy stress at the apex as $\sigma = pr/(2h_0\lambda_3)$, which holds for $h_0 \ll r$ [31].

D. CONSTITUTIVE MODELS

The experimental data are used to fit the parameters of a hyperelastic, incompressible Ogden model [31], for which the strain-energy density per unit reference volume reads

$$W = \sum_{k=1}^{K} \frac{\mu_k}{\alpha_k}\left(\lambda_1^{\alpha_k} + \lambda_2^{\alpha_k} + \lambda_3^{\alpha_k} - 3\right), \quad \lambda_1\lambda_2\lambda_3 = 1. \quad (1)$$

The material parameters μ_k, α_k must satisfy the constraint $\mu_k\alpha_k > 0$, $k = 1, 2, \ldots, K$, and can be used to calculate the corresponding Young's modulus as $E = (1 + \nu)\sum_{k=1}^{K}\mu_k\alpha_k$, with $\nu = 0.5$ being the Poisson's ratio of an isotropic incompressible material. The principal Cauchy stresses immediately follow from (1) as (see [31], p. 571)

$$\sigma_i = \sum_{k=1}^{K}\mu_k\lambda_i^{\alpha_k} - q, \quad i = 1, 2, 3, \quad (2)$$

where q is an arbitrary hydrostatic pressure arising due to the incompressibility constraint, whose value depends on the boundary conditions. By specializing (2) to the three experimentally considered load cases (see [31]), the analytical formulas were used to minimize the squared error between the model and the experiments using the minimization routine `fmincon` available in MATLAB (R2018b, The MathWorks, Inc.). Ogden models of order $K = 2$ were found to provide the best description of the data for both materials compared to neo-Hookean or Mooney–Rivlin formulations; the resulting parameter sets are reported in Table 1.

E. RESULTS

The individual stress-stretch curves for each sample of the two elastomers tested are reported in Fig. 4, together with the sample averages and the model predictions. Both models identified provide an excellent description of the mechanical behavior over the whole range of deformation for all three load cases. The additional UA tests suggest a negligible influence of both strain rate and shelf time (over 5 weeks) on the mechanical behavior of the Elastosil 25:1 for the rates and times tested. However, the Ecoflex GEL shows a dependence on both strain rate and aging, see Fig. 12, in the Appendix. These dependencies, as well as potential softening phenomena upon cyclic loading (Mullins effect), are neglected in the hyperelastic model.

Corresponding Young's moduli (calculated using the Ogden model coefficients) of the Ecoflex GEL and the Elastosil 25:1 are 16.9 kPa and 370.2 kPa, respectively. For comparison, the Young's modulus of the stiffer Elastosil 7:1, determined by microindentation tests (FT-MTA02, FemtoTools AG), is found to be 0.97 MPa, i.e. more than 50 times stiffer than the Ecoflex GEL.

IV. GENERATING A DATASET

The task of mapping the information extracted from the images to the applied contact force distribution is formulated here as a supervised learning problem. This requires a training dataset composed of input features (here retrieved from the images) and the respective ground truth labels (here obtained from finite element simulations, using the material models derived in Section III). These labels represent the quantities of interest in the inference process (e.g., the contact force distribution). The following subsections describe in detail each of the components of the dataset.

A. FEATURES

In order to perform a large number of indentations within a feasible time, an automatic milling and drilling machine (Fehlmann PICOMAX 56 TOP) is used to press an indenter against the soft surface of the tactile sensor at different locations and depths. The machine is equipped with fast and precise motion control (up to 10^{-3} mm). In the experiments presented here, a stainless steel spherical-ended cylindrical indenter is used. The indenter has a diameter of 10 mm and is attached to the spindle of the milling machine, together with a six-axis F/T sensor (ATI Mini 27 Titanium). The experimental data collection setup is shown in Fig. 5.

A total of 13,448 vertical indentations were performed, on an horizontal grid with a regular spacing of 0.55 mm, at various depths (with a maximum depth of 2 mm). The RGB images of the particle spread are captured once the indentation has reached the commanded position. Five images were collected for each indentation in order to improve the robustness to image noise. The F/T sensor's measurements of the total vertical and horizontal (over two perpendicular axes) contact force were recorded.
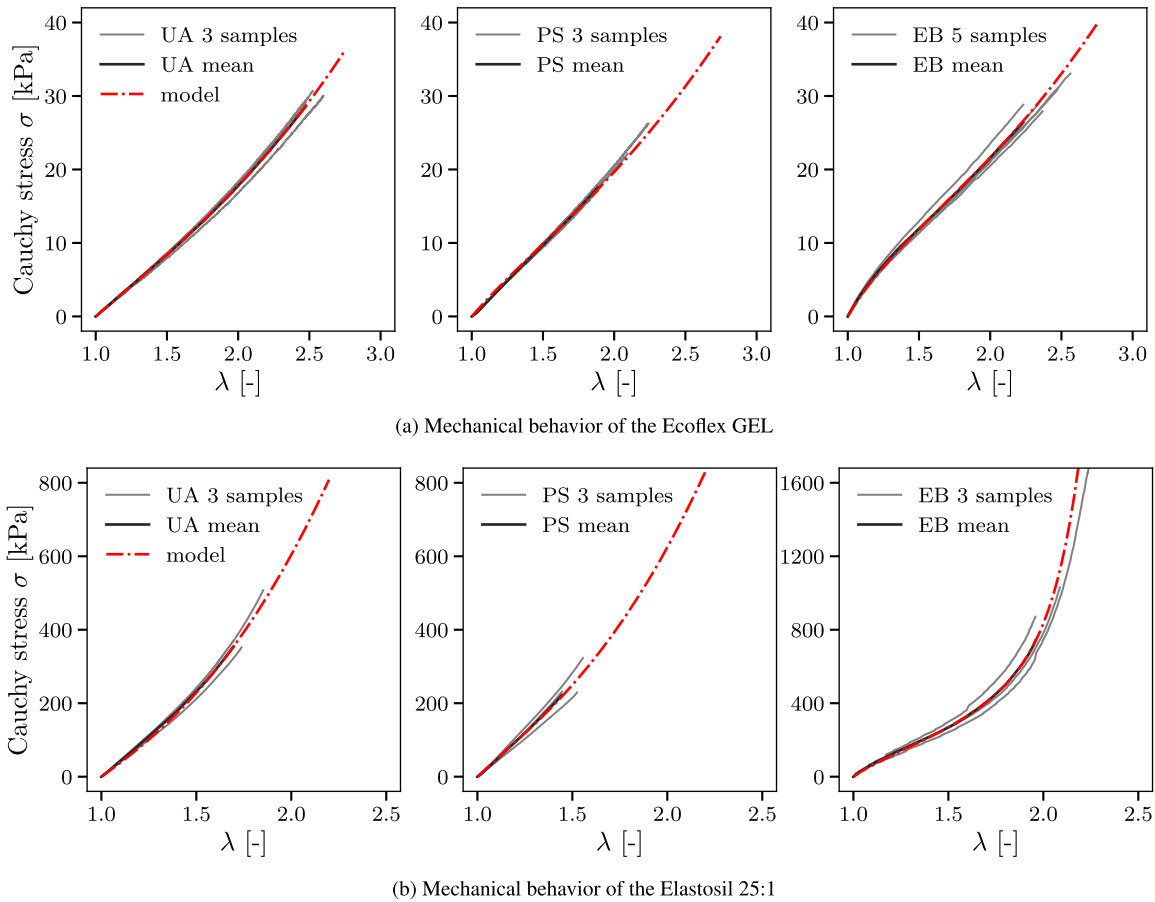
(a) Mechanical behavior of the Ecoflex GEL

(b) Mechanical behavior of the Elastosil 25:1

FIGURE 4. Stress-stretch response of the Ecoflex GEL (a) and the Elastosil 25:1 (b) in, from left to right, uniaxial tension (UA), pure shear (PS), and equibiaxial tension (EB), together with corresponding hyperelastic model predictions. Note the different scales in (a) and (b), in particular the significantly stiffer equibiaxial response of the Elastosil 25:1.
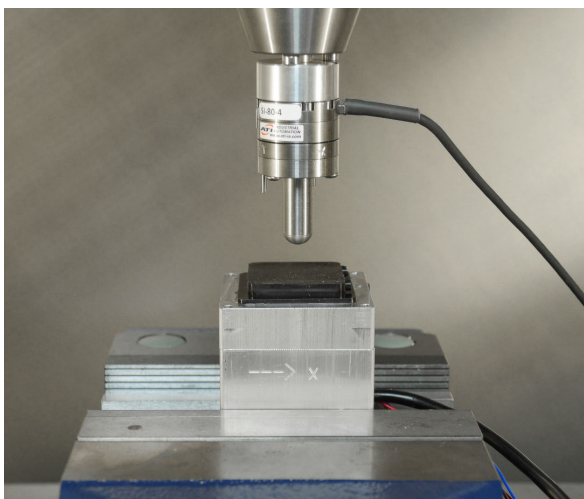

FIGURE 5. The experimental data collection setup is shown above. The indenter and the F/T sensor (connected through the cable on the top right) are attached to the spindle of an automatic milling machine.

The optical flow field is extracted from the images (converted to grayscale) through an algorithm based on Dense Inverse Search [32]. The magnitude and the direction of the

field are then averaged in m image regions of equal area, as described in [9]. The tuples of magnitude and direction for each of these regions yield a set of $2 \times m$ features for each data point.

The readings from the F/T sensor are used to assess the quality of the ground truth labels, as described in the next subsection. The range of forces recorded in this procedure spans up to 1.7 N in the vertical direction and 0.15 N in each of the horizontal axes. Note that the large difference in magnitude between the vertical and horizontal forces is mainly due to the symmetry of the indentations, which leads to the cancellation of the symmetric contributions to the total horizontal force, with the exception of the regions close to the edges of the surface.

B. LABELS

Although the F/T sensor provides the total contact force in each direction, it does not provide any information about the force distribution over the contact surface. The force distribution renders a compact representation of various contact aspects for generic indentations. In fact, it encodes information about the contact area and forces applied to the surface, even in the case of interactions with objects of
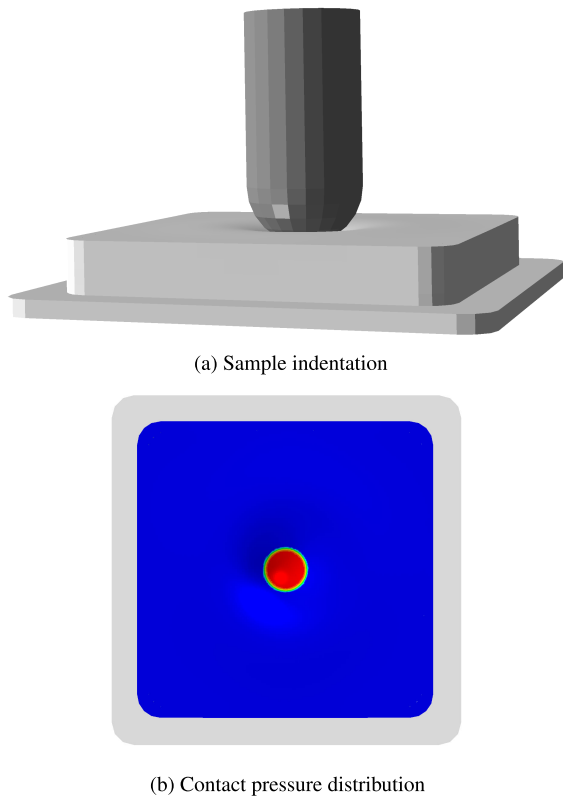
(a) Sample indentation

(b) Contact pressure distribution

FIGURE 6. The result of a sample FEM indentation in Abaqus is shown in this figure. The indenter and the gel are modeled to reflect their actual material and geometric properties, see (a). An example of the resulting contact pressure distribution (top view) is shown in (b), where the colors are mapped to the pressure magnitude (from zero, in blue, to the maximum, in red).

FIGURE 7. The procedure to discretize the force distribution is sketched in this figure. The sensor's surface is discretized into *n* bins of equal size, with the boundaries shown in red. The surface mesh used for the FEA is shown in black in the undeformed state. Each node is assigned to the bin that contains it, as is the case for the green nodes contained in the bin shown in solid red. For each indentation, the resulting forces at the nodes assigned to the same bin are summed along each axis to determine the three label vector components for the corresponding bin.

complex geometries or when multiple and distinct contact points are present. An example application, in which both the contact area and the total contact force are necessary, is presented in [33], where the interaction with a specific object is exploited. The contact force distribution is obtained in this article through FEA, which essentially simulates the indentation experiments performed with the milling machine, see for example Fig. 6.

The FEM simulations are carried out in Abaqus/ Standard [34]. The geometry of the two top layers of the sensor is modeled as shown in Fig. 3, and material properties are assigned to each layer as described in Section III, using the implementation of Ogden's model provided by Abaqus. Note that this neglects the influence of the spherical particles, as the model was derived from tests on the pure Ecoflex GEL. Assuming rigid particles, the ratio between the Ecoflex– particle composite modulus E_c and the Young's modulus E of the pure Ecoflex GEL can be estimated using Eshelby inclusion theory [35], $E_c/E = 1/(1 - 5\phi/2) \approx 1.05$ for a particle volume fraction $\phi = 0.0196$. The spherical-ended indenter is modeled as an analytical rigid shell. The finite element mesh is composed of linear tetrahedral elements (C3D4H) and hexahedral elements with reduced integration (C3D8RH). Both element types are used with hybrid formulation as appropriate for incompressible materials. A local mesh
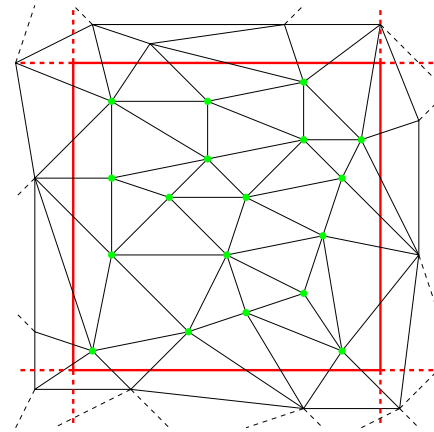
refinement is performed at the contact and at the interface of the materials, with a characteristic element size of 0.3 mm. Tie constraints are applied at the material interface to enforce the same displacement of the nodes in contact. The bottom nodes are fixed, reflecting the interface with the much stiffer bottom layer (Elastosil 7:1).

The contact between the top surface and the indenter is modeled as a hard contact and discretized with a surface-to-surface method. The friction coefficient between the indenter and the top layer is estimated by letting a block of Elastosil 25:1 rest on an inclined stainless steel plate. The maximum tilt angle θ before the block begins to slide is recorded with an external camera, and the static friction coefficient μ_0 is calculated from static equilibrium as $\mu_0 = \tan\theta$. This procedure yields a friction coefficient of 0.45. The friction coefficient was assumed constant, a possible dependence on the contact pressure was neglected.

The FEM simulations generate the normal and shear components of the contact force distribution resulting from each indentation. Note that both the normal and shear forces acting at each node are generally 3D vectors. As an example, the normal force that stems from a pure vertical indentation is not necessarily vertical, as a consequence of the material deformation (although the vertical component generally has the largest magnitude).

The normal and shear force distributions are discretized by summing the respective nodal forces inside *n* surface bins, as shown in Fig. 7. The resulting 3D force for each of these bins is used as a ground truth label with $3 \times n$ components for each data point.

This procedure is applied to assign ground truth labels to the 13,448 indentations described in the previous subsection. Since there are no readily available sensors that measure the full contact force distribution with high spatial resolution and without altering the sensor's soft surface, the quality

TABLE 2. Total force agreement (FEA vs F/T sensor).

Axis	$RMSE_{GT}$	F/T resolution
x	0.02 N	0.03 N
y	0.02 N	0.03 N
z	0.06 N	0.06 N

of the labels is evaluated by comparing the components of the total force resulting from the FEM simulations with the ones measured by the F/T sensor. Note that the total force components can be obtained from the FEM simulations by summing the contact force values of all the n bins, or simply summing the force values at all nodes of the surface mesh used for the FEA. The resulting root-mean-square error on the ground truth ($RMSE_{GT}$) for the entire dataset is reported in Table 2 for each component. x and y are the horizontal axes, and z is the vertical axis, which is positive pointing from the camera towards the top surface. The resulting errors are comparable to the F/T sensor's resolution, shown in the table as a reference. In Fig. 8, the plots show the agreement on the z component of the total force between the F/T sensor's readings and the results from the FEA for two of the indentation locations. The good agreement between the F/T measurements and the FEM simulations further justifies the simplifying assumptions taken in the material characterization and the FEM modeling. Additionally, the same plots show that using a linear elastic material model and neglecting geometric nonlinearities (i.e., NLgeom flag in Abaqus) lead to a considerable performance loss for large deformations.

Although the FEM simulations can be time consuming to carry out, depending on the accuracy required, most of the operations are highly parallelizable, as for example, the several indentations. This makes it possible to exploit cluster computers or GPUs to reduce the time consumption. The simulations presented here are carried out on the Euler cluster of ETH Zurich.

Note that the strategy presented above provides the ground truth for the full contact force distribution under no assumptions on the specific tactile sensing technique. It is therefore not limited to use on vision-based devices, but more generally on data-driven approaches to the force reconstruction task.

V. NEURAL NETWORK TRAINING
A. LEARNING ARCHITECTURE
A feedforward DNN architecture (see Fig. 9) is used to address the supervised learning task of reconstructing the full contact force distribution from the features extracted from the images. An input layer with $2 \times m$ units represents the image features described in Section IV-A (a tuple of averaged optical flow magnitude and direction for each of the chosen image regions). Similarly, an output layer with $3 \times n$ units represents the discretized force distribution applied to the surface of the sensor (a three-dimensional force vector for each of the discrete surface bins).

Three fully connected hidden layers with a sigmoid activation function are used to model the map between the inputs
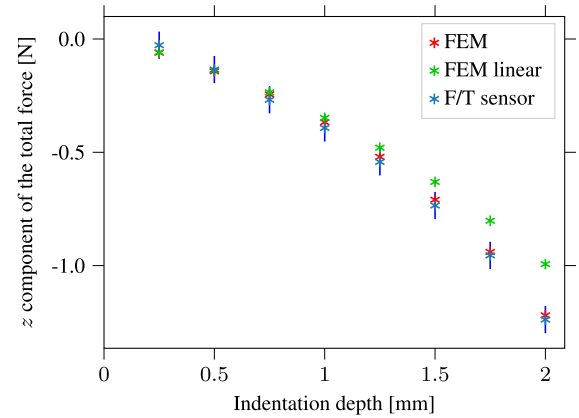
(a) Center indentation
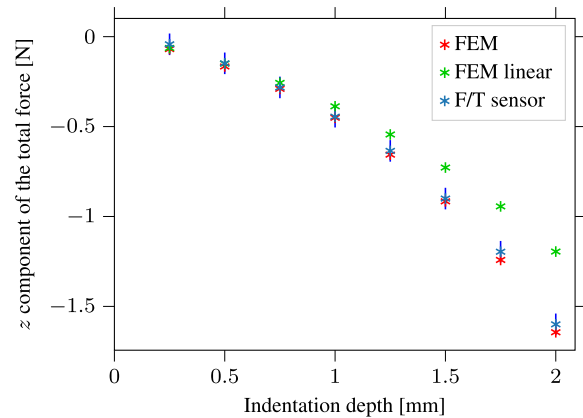
(b) Corner indentation

FIGURE 8. The plots above show the agreement on the total vertical contact force between the measurements obtained from the F/T sensor (in blue) and the FEM simulations (in red). The results from the simulations are accurate for indentations at the center of the surface (a) and close to the corners (b) (5 mm from each of the edges) for different indentation depths. The F/T sensor readings are shown with ±0.06 N bars, representing the resolution of the F/T sensor. In green, the results obtained using a linear elastic model (as opposed to the hyperelastic model described in Section III) and neglecting geometric nonlinearities are shown.
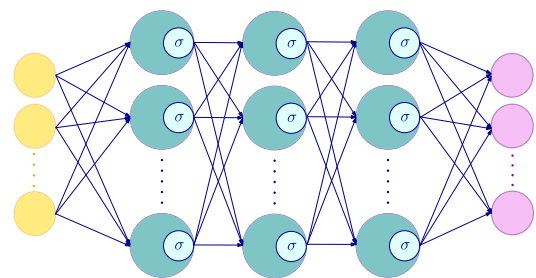

FIGURE 9. A diagram of the learning architecture used to predict the 3D contact force distribution. In yellow the input layer, representing the image features, in cyan the hidden layers, and in magenta the output layer, representing the discretized force distribution.

and the outputs. Dropout layers are used after each of the hidden layers during the training phase. Twenty percent of the dataset is used as a test set, while the remaining data are
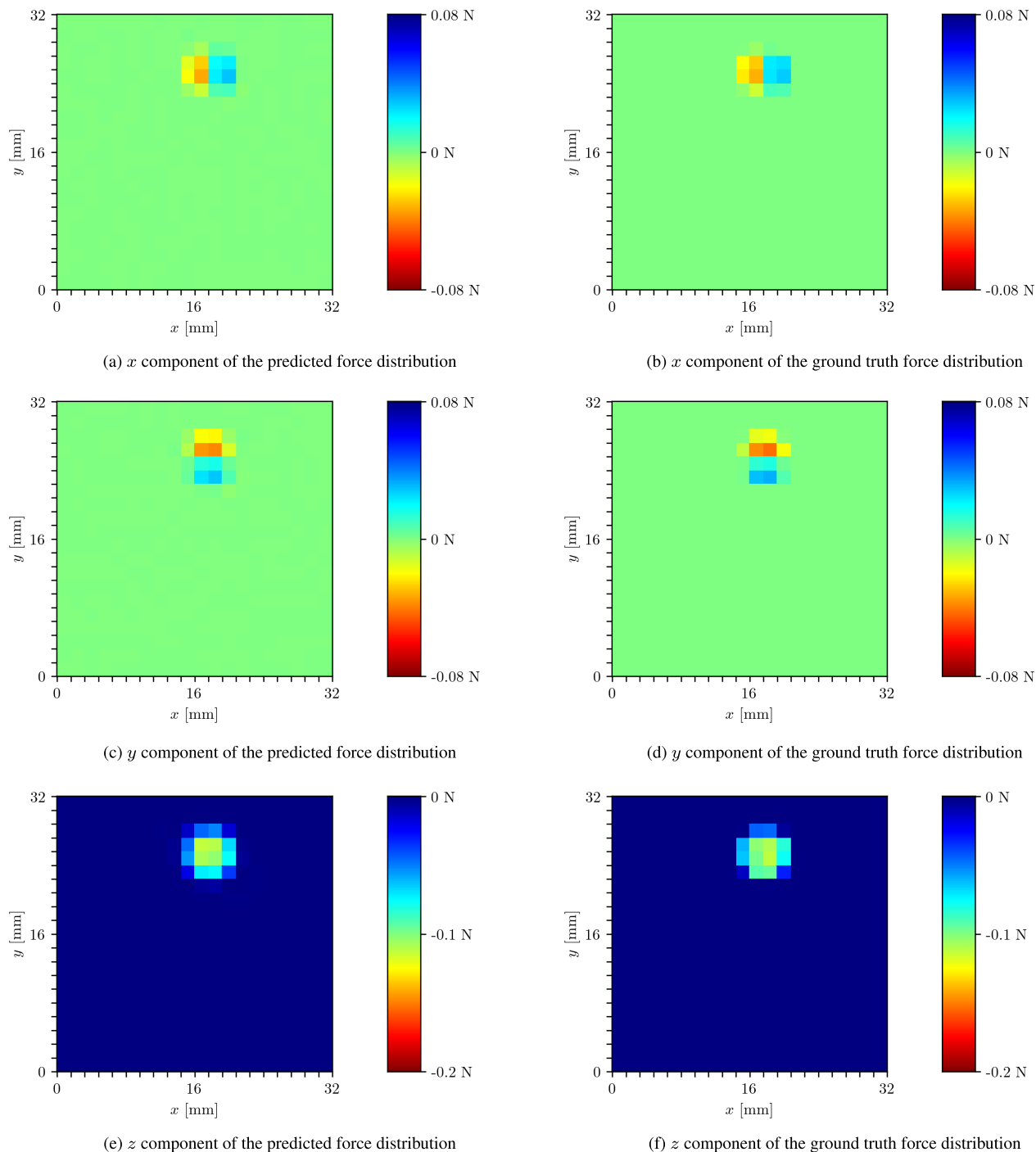
(a) *x* component of the predicted force distribution

(b) *x* component of the ground truth force distribution

(c) *y* component of the predicted force distribution

(d) *y* component of the ground truth force distribution

(e) *z* component of the predicted force distribution

(f) *z* component of the ground truth force distribution

FIGURE 10. The plots above show the predicted (left) and ground truth (right) 3D contact force distribution applied to the top surface of the tactile sensor for an indentation in the test set. Note that the axes are defined as in Section IV, that is, with two perpendicular horizontal axes *x* and *y*, aligned with two of the top surface edges, and a vertical axis *z*, which is positive pointing from the camera towards the top surface.

used for training. The architecture is trained with PyTorch[1] by minimizing the mean squared error (MSE) through the Adam optimizer, see [36]. The remaining parameters chosen for the optimization, as well as the size of each layer, are summarized

in Table 3. Note that the spatial resolution of the tactile sensor is determined by the size of the surface bins, which have a side of 1.6 mm, comparable to the spatial resolution of the human fingertip [37]. However, a finer resolution may yield additional discrimination capabilities, i.e., for sensing an object's roughness or texture.

[1]www.pytorch.org

(a) *x* component of the predicted force distribution

(b) *x* component of the ground truth force distribution

(c) *y* component of the predicted force distribution

(d) *y* component of the ground truth force distribution

(e) *z* component of the predicted force distribution

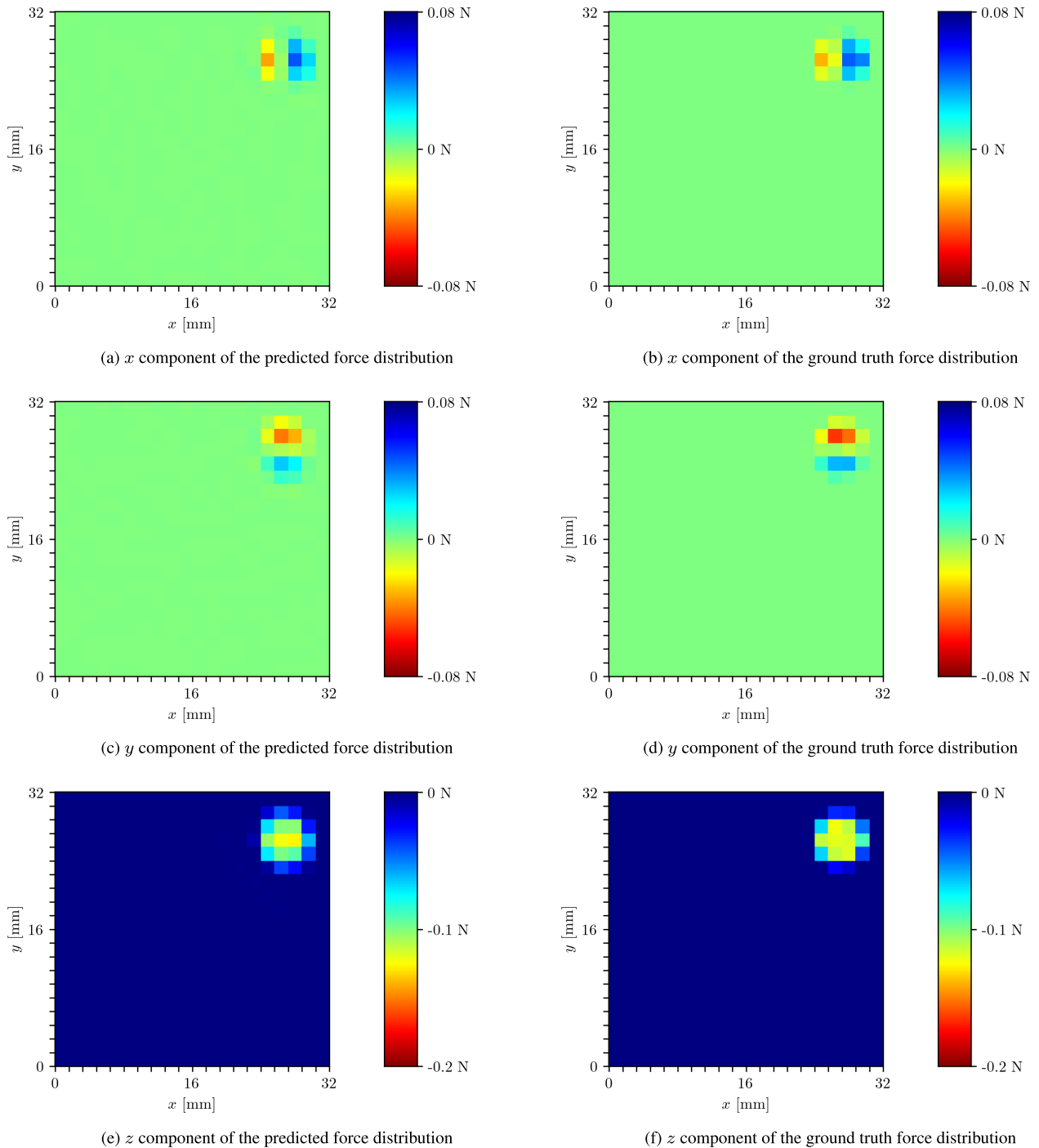(f) *z* component of the ground truth force distribution

FIGURE 11. Similarly to Figure 10, the plots above show the predicted (left) and ground truth (right) 3D contact force distribution applied to the top surface of the tactile sensor for an indentation in the test set. In this case, the indentation was centered at 5 mm from the top and right edges, with the indenter's radius being 5 mm. It can be noted a larger asymmetry in the force distribution, affected by the stiffer edges of the gel.

In contrast to architectures that directly learn from the pixels (e.g., convolutional neural networks), the extraction of the optical flow features described in Section IV-A yields a relatively shallow model, reducing the training times and the data requirements. Additionally, the use of these features makes it possible to efficiently transfer the model across different sensors, as shown in previous work [21]. However,

these features do not exploit the information about particles placed at different distances from the camera.

B. RESULTS
After training, the quality of the DNN predictions is evaluated on the test set. Additionally to the root-mean-square error (RMSE) on the entire test set, the sparse RMSE on the

TABLE 3. DNN parameters.

Symbol	Value	Description
m	1600	# of averaging image regions
n	400	# of discrete surface bins
-	(800, 600, 400)	hidden layers' size
-	1E-4	learning rate
-	400	training batch size
-	0.1	dropout rate

TABLE 4. Resulting errors on force distribution and total force.

Metric	x	y	z
RMSE	0.001 N	0.001 N	0.003 N
RMSES	0.007 N	0.006 N	0.016 N
RMSET$_{\text{FEM}}$	0.004 N	0.004 N	0.045 N
RMSET$_{\text{F/T}}$	0.025 N	0.021 N	0.082 N

non-zero values of the FEM ground truth is also computed as

$$\text{RMSES} := \sqrt{\frac{1}{|\mathcal{I}|} \sum_{(i,l) \in \mathcal{I}} \left(f_i^{(l)} - \hat{f}_i^{(l)} \right)^2},$$

where $f_i^{(l)}$ and $\hat{f}_i^{(l)}$ are the i-th components of the ground truth and the predicted label, respectively, for the l-th sample in the test set, and,

$$\mathcal{I} := \left\{ (i,l) \in \{0, \ldots, 3n-1\} \times \{0, \ldots, N_{\text{set}} - 1\} \mid f_i^{(l)} \neq 0 \right\},$$

with N_{set} the number of samples in the test set. This metric emphasizes the prediction performance in the location where the contact is expected.

Moreover, the RMSE on the total force is estimated for both the cases, in which the ground truth is provided either by the FEM simulations (RMSET$_{\text{FEM}}$) or the F/T sensor (RMSET$_{\text{F/T}}$). The resulting errors from the predictions on the test set are summarized in Table 4 for each axis. The values in the last row are affected by both the errors introduced by the FEM modeling and the DNN predictions. Note that it is only possible to compute the metrics for the force distribution (first two rows) in relation to the FEM simulations (the F/T sensor only provides total forces, without specific information about the force distribution). As a reference, the ranges of force provided by the ground truth labels are summarized in Table 5. Examples of the predicted contact force distribution are shown in Fig. 10 and Fig. 11.

The resulting DNN is deployed on the dual-core laptop computer introduced in Section I. The entire pipeline yields real-time predictions on an image stream of 40 frames per second, as shown in the experiments available in the video attached to this article. The parallelization of both the optical flow algorithm and the neural network prediction step is not exploited here, but it may be leveraged on commercially available embedded computers provided of GPUs.

TABLE 5. Range of ground truth forces.

Quantity	x	y	z
Force per bin	[-0.06,0.06] N	[-0.06,0.06] N	[-0.15,0] N
Total force	[-0.13,0.13] N	[-0.13,0.13] N	[-1.66,0] N
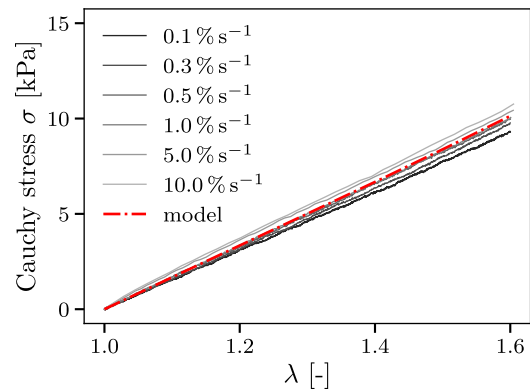
(a) Strain-rate dependence in uniaxial tension for the Ecoflex GEL
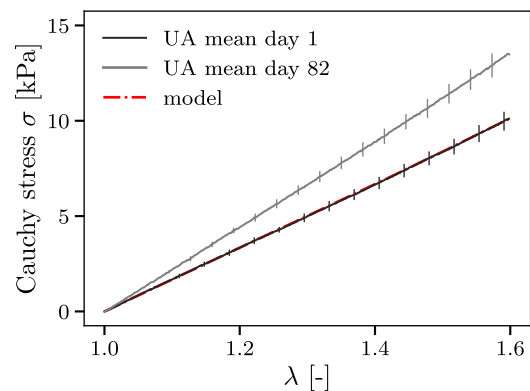
(b) Age-dependent mechanical properties of the Ecoflex GEL

FIGURE 12. Uniaxial tension (UA) tests showing (a) strain-rate and (b) shelf-time (aging) dependence on the mechanical properties of Ecoflex GEL.

VI. CONCLUSION

This article has presented a strategy to provide ground truth contact force distribution for learning-based tactile sensing. The approach has been evaluated on a vision-based tactile sensor, which is based on tracking particles spread within a soft gel. After the characterization of the hyperelastic materials, which provides accurate material models for FEA, a large number of real indentations and corresponding simulations have been performed to generate a dataset that includes image features and ground truth for the 3D contact force distribution. Although the material characterization was performed with considerably different tests and setup (i.e., UA, PS, EB tests) than the indentations considered in the evaluation, the total forces recorded in the experiments are comparable to the ones determined in simulation, showing the generalization potential of the approach proposed. Note that due to the fact that the simulation labels are assigned to real

data obtained from experimental indentations, the experimental setup needs to be carefully arranged. As an example, the alignment of the tactile sensor with the reference axes of the milling machine used for the data collection is crucial for obtaining good performance.

As shown in Section V, the dataset generated with the strategy proposed in this article can be used to train a DNN for accurate reconstruction of the force distribution applied to the surface of the tactile sensor. Although in these experiments the DNN has been trained and evaluated on a sample indenter, the techniques presented here are directly applicable to generic shapes and indentations, including multiple and distinct contact areas. However, this would likely require the collection of a dataset under various contact conditions, involving interactions with complex objects. Therefore, the generalization capabilities of this strategy will be object of future work.

APPENDIX
STRAIN-RATE AND SHELF-TIME DEPENDENT PROPERTIES OF ECOFLEX GEL

The additional experimental data on the rate-dependence and the aging effect on the mechanical behavior of the Ecoflex GEL are shown in Fig. 12. While the strain-rate dependence remains relatively low over the three decades analyzed (Fig. 12a), a significant stiffening after 9 weeks of storage at room temperature is evident due to material aging (Fig. 12b). Longer curing times and higher curing temperatures may be used to approach the final, curing-independent material properties [28], [29].

ACKNOWLEDGMENT

REFERENCES

[1] J. Mahler and K. Goldberg, "Learning deep policies for robot bin picking by simulating robust grasping sequences," in *Proc. 1st Annu. Conf. Robot Learn. (PMLR)*, vol. 78, 2017, pp. 515–524.

[2] C. D. Kidd and C. Breazeal, "Robots at home: Understanding long-term human-robot interaction," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, Sep. 2008, pp. 3230–3235.

[3] H. Shen, "Meet the soft, cuddly robots of the future," *Nature*, vol. 530, no. 7588, pp. 24–26, 2016.

[4] H. Abidi and M. Cianchetti, "On intrinsic safety of soft robots," *Frontiers Robot. AI*, vol. 4, p. 5, Feb. 2017.

[5] S. L. Gorniak, V. M. Zatsiorsky, and M. L. Latash, "Manipulation of a fragile object," *Exp. Brain Res.*, vol. 202, no. 2, pp. 413–430, 2010.

[6] S. Decherchi, P. Gastaldo, R. S. Dahiya, M. Valle, and R. Zunino, "Tactile-data classification of contact materials using computational intelligence," *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 635–639, Jun. 2011.

[7] N. Wettels and G. E. Loeb, "Haptic feature extraction from a biomimetic tactile sensor: Force, contact location and curvature," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, Dec. 2011, pp. 2471–2478.

[8] O. Kroemer, C. H. Lampert, and J. Peters, "Learning dynamic tactile sensing with robust vision-based training," *IEEE Trans. Robot.*, vol. 27, no. 3, pp. 545–557, Jun. 2011.

[9] C. Sferrazza and R. D'Andrea, "Design, motivation and evaluation of a full-resolution optical tactile sensor," *Sensors*, vol. 19, no. vol. 4, p. 928, 2019.

[10] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016.

[11] H. Dang and P. K. Allen, "Learning grasp stability," in *Proc. IEEE Int. Conf. Robot. Automat.*, May 2012, pp. 2392–2397.

[12] R. Calandra, A. Owens, D. Jayaraman, J. Lin, W. Yuan, J. Malik, E. H. Adelson, and S. Levine, "More than a feeling: Learning to grasp and regrasp using vision and touch," *IEEE Robot. Autom. Lett.*, vol. 3, no. 4, pp. 3300–3307, Oct. 2018.

[13] M. H. Lee and H. R. Nicholls, "Review article tactile sensing for mechatronics—A state of the art survey," *Mechatronics*, vol. 9, no. 1, pp. 1–31, 1999.

[14] R. S. Dahiya, G. Metta, M. Valle, and G. Sandini, "Tactile sensing—From humans to humanoids," *IEEE Trans. Robot.*, vol. 26, no. 1, pp. 1–20, Feb. 2010.

[15] A. Yamaguchi and C. G. Atkeson, "Combining finger vision and optical tactile sensing: Reducing and handling errors while cutting vegetables," in *Proc. IEEE-RAS 16th Int. Conf. Hum. Robots (Humanoids)*, Nov. 2016, pp. 1045–1051.

[16] B. Ward-Cherrier, N. Pestell, L. Cramphorn, B. Winstone, M. E. Giannaccini, J. Rossiter, and N. F. Lepora, "The TacTip family: Soft optical tactile sensors with 3D-printed biomimetic morphologies," *Soft Robot.*, vol. 5, no. 2, pp. 216–227, Apr. 2018.

[17] A. Rigi, F. B. Naeini, D. Makris, and Y. Zweiri, "A novel event-based incipient slip detection using dynamic active-pixel vision sensor (DAVIS)," *Sensors*, vol. 18, no. 2, p. 333, 2018.

[18] A. Alspach, K. Hashimoto, N. Kuppuswarny, and R. Tedrake, "Soft-bubble: A highly compliant dense geometry tactile sensor for robot manipulation," in *Proc. 2nd IEEE Int. Conf. Soft Robot. (RoboSoft)*, Apr. 2019, pp. 597–604.

[19] W. Yuan, S. Dong, and E. H. Adelson, "Gelsight: High-resolution robot tactile sensors for estimating geometry and force," *Sensors*, vol. 17, no. 12, p. 2762, 2017.

[20] P. Piacenza, E. Hannigan, C. Baumgart, Y. Xiao, S. Park, K. Behrman, W. Dang, J. Espinal, I. Hussain, I. Kymissis, and M. Ciocarlie, "Touch sensors with overlapping signals: Concept investigation on planar sensors with resistive or optical transduction," 2019, *arXiv:1802.08209*. [Online]. Available: https://arxiv.org/abs/1802.08209

[21] C. Sferrazza and R. D'Andrea, "Transfer learning for vision-based tactile sensing," 2019, *arXiv:1812.03163*. [Online]. Available: https://arxiv.org/abs/1812.03163

[22] D. V. Hutton, *Fundamentals of Finite Element Analysis*. New York, NY, USA: McGraw-Hill, 2004.

[23] D. Ma, E. Donlon, S. Dong, and A. Rodriguez, "Dense tactile force estimation using GelSlim and inverse FEM," in *Proc. IEEE Int. Conf. Robot. Automat. (ICRA)*, May 2019, pp. 5418–5424.

[24] T. Le Tim Lun, K. Wang, J. D. L. Ho, K.-H. Lee, K. Y. Sze, and K.-W. Kwok, "Real-time surface shape sensing for soft and flexible structures using fiber Bragg gratings," *IEEE Robot. Autom. Lett.*, vol. 4, no. 2, pp. 1454–1461, Apr. 2019.

[25] Z. Zhang, J. Dequidt, and C. Duriez, "Vision-based sensing of external forces acting on soft robots using finite element method," *IEEE Robot. Autom. Lett.*, vol. 3, no. 3, pp. 1529–1536, Jul. 2018.

[26] R. W. Ogden, *Non-Linear Elastic Deformations*. Hemel Hempstead, U.K.: Ellis Horwood, 1984.

[27] P. Steinmann, M. Hossain, and G. Possart, "Hyperelastic models for rubber-like materials: Consistent tangent operators and suitability for Treloar's data," *Arch. Appl. Mech.*, vol. 82, no. 9, pp. 1183–1217, 2012.

[28] R. Hopf, L. Bernardi, J. Menze, M. Zündel, E. Mazza, and A. E. Ehret, "Experimental and theoretical analyses of the age-dependent large-strain behavior of Sylgard 184 (10:1) silicone elastomer," *J. Mech. Behav. Biomed. Mater.*, vol. 60, pp. 425–437, Jul. 2016.

[29] V. Placet and P. Delobelle, "Mechanical properties of bulk polydimethylsiloxane for microfluidics over a large range of frequencies and aging times," *J. Micromech. Microeng.*, vol. 25, no. 3, 2015, Art. no. 035009.

[30] L. Bernardi, R. Hopf, A. Ferrari, A. E. Ehret, and E. Mazza, "On the large strain deformation behavior of silicone-based elastomers for biomedical applications," *Polym. Test.*, vol. 58, pp. 189–198, Apr. 2017.

[31] R. W. Ogden, "Large deformation isotropic elasticity—On the correlation of theory and experiment for incompressible rubberlike solids," *Proc. Roy. Soc. London A, Math. Phys. Sci.*, vol. 326, no. 1567, pp. 565–584, 1972.

[32] T. Kroeger, R. Timofte, D. Dai, and L. Van Gool, "Fast optical flow using dense inverse search," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*. Berlin, Germany: Springer, 2016, pp. 471–488.
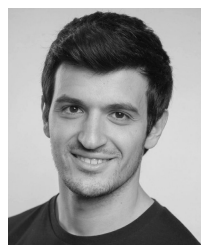
[33] K. Nozu and K. Shimonomura, "Robotic bolt insertion and tightening based on in-hand object localization and force sensing," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatronics (AIM)*, Jul. 2018, pp. 310–315.

[34] *Abaqus/Standard User's Manual, Version 6.14*, Dassault Systèmes, Vélizy-Villacoublay, France, 2014.

[35] J. D. Eshelby, "The determination of the elastic field of an ellipsoidal inclusion, and related problems," *Proc. Roy. Soc. London A, Math., Phys. Eng. Sci.*, vol. 241, no. 1226, pp. 376–396, 1957.

[36] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2015, pp. 1–15.

[37] J. Dargahi and S. Najarian, "Human tactile perception as a standard for artificial tactile sensing—A review," *Int. J. Med. Robot. Comput. Assist. Surg.*, vol. 1, no. 1, pp. 23–35, 2004.

ADAM WAHLSTEN received the M.Sc. degree in engineering physics from Lund University, Lund, Sweden, in 2017, with a focus on computational solid and fluid mechanics. He is currently pursuing the Ph.D. degree with the Institute for Mechanical Systems, ETH Zurich, Zurich, Switzerland. He also wrote his master's thesis on multiphase modeling of soft tissues at ETH Zurich. His research focuses on the mechanical characterization and biomechanical modeling of the skin and skin-like tissue-engineered materials.

CAMILL TRUEEB received the B.Sc. degree in mechanical engineering from ETH Zurich, Zurich, Switzerland, in 2017, where he is currently pursuing the M.Sc. degree, with a focus on the fields of robotics, systems, and control. During his studies, he worked as a Teaching Assistant in control systems and participated in a project to develop a stretchable volume sensor for the human bladder. After doing a semester project on ground truth acquisition for vision-based tactile sensors, he is currently writing his master's thesis on deep learning calibration approaches for tactile sensors.

CARMELO SFERRAZZA (S'17) received the B.Eng. degree in electronics and information engineering from Tongji University, Shanghai, China, the B.Sc. degree in automation engineering from Politecnico di Milano, Milan, Italy, in 2014, and the M.Sc. degree in robotics, systems, and control from ETH Zurich, Zurich, Switzerland, in 2016, where he is currently pursuing the Ph.D. degree with the Institute for Dynamic Systems and Control. His current research interests include the design and the development of vision-based, data-driven tactile sensors, and learning-based model predictive control. He received the ETEL Award for his master's thesis on model predictive control of an unmanned aerial vehicle.

RAFFAELLO D'ANDREA (F'10) received the B.Sc. degree in engineering science from the University of Toronto, in 1991, and the M.S. and Ph.D. degrees in electrical engineering from the California Institute of Technology, in 1992 and 1997, respectively. He was an Assistant Professor and then an Associate Professor with Cornell University, from 1997 to 2007. While on leave from Cornell, from 2003 to 2007, he co-founded Kiva Systems (now Amazon Robotics), where he led the systems architecture, robot design, robot navigation and coordination, and control algorithms development. He is currently a Professor of dynamic systems and control with ETH Zurich, and the Chairman of the board with Verity Studios AG.

● ● ●