A Combined Simulation & Machine Learning Approach for Image-based Force Classification during Robotized Intravitreal Injections

Andrea Mendizabal, Tatiana Fountoukidou, Jan Hermann, Raphael Sznitman, Stéphane Cotin

▶ **To cite this version:**

Andrea Mendizabal, Tatiana Fountoukidou, Jan Hermann, Raphael Sznitman, Stéphane Cotin. A Combined Simulation & Machine Learning Approach for Image-based Force Classification during Robotized Intravitreal Injections. Medical Image Computing and Computer Assisted Interventions Conference, Sep 2018, Granada, Spain. hal-01878682v2

HAL Id: hal-01878682
https://hal.science/hal-01878682v2

Submitted on 21 May 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Combined Simulation & Machine Learning Approach for Image-based Force Classification during Robotized Intravitreal Injections

Andrea Mendizabal[1,2], Tatiana Fountoukidou[3], Jan Hermann[3], Raphael Sznitman[3] and Stephane Cotin[1]

[1] Inria, Strasbourg, France
[2] University of Strasbourg, ICube, Strasbourg, France
[3] ARTORG Center, University of Bern, Switzerland

Abstract. Intravitreal injection is one of the most common treatment strategies for chronic ophthalmic diseases. The last decade has seen the number of intravitreal injections dramatically increase, and with it, adverse effects and limitations. To overcome these issues, medical assistive devices for robotized injections have been proposed and are projected to improve delivery mechanisms for new generation of pharmacological solutions. In our work, we propose a method aimed at improving the safety features of such envisioned robotic systems. Our vision-based method uses a combination of 2D OCT data, numerical simulation and machine learning to estimate the range of the force applied by an injection needle on the sclera. We build a Neural Network (NN) to predict force ranges from Optical Coherence Tomography (OCT) images of the sclera directly. To avoid the need of large training data sets, the NN is trained on images of simulated deformed sclera. We validate our approach on real OCT images collected on five *ex vivo* porcine eyes using a robotically-controlled needle. Results show that the applied force range can be predicted with 94% accuracy. Being real-time, this solution can be integrated in the control loop of the system, allowing for in-time withdrawal of the needle.

1 Introduction

Intravitreal injections are one of the most frequent surgical interventions in opthalmology with more than 4 million injections in 2014 alone [1]. This procedure is used, for instance, in the treatment of age-related macular degeneration for injecting vascular endothelial growth factor inhibitors. Similarly, intravitreal therapy is also used in the treatment of diabetic maculopathy and retinopathy. With the increasing prevalence of diabetic patients and aging demographics, the demand for such intravitreal therapy is growing significantly.

At the same time, robotic assistance in ophthalmology offers the ability to improve manipulation dexterity, along with shorter and safer surgeries [2]. Ullrich *et al.* [1] has also proposed a robotized intravitreal injection device capable of assisting injections into the vitreous cavity, whereby faster injections in safer conditions are projected. Designing such robotic systems involves significant hurdles and challenges however. Among others, accurate force sensing represents an

important topic, whereby the force required to puncture the sclera with a needle is very small and measuring this force is essential for the safety of the patient.

For many devices, force sensing plays a central role in the control loop and safety system of surgical robots [3]. The use of force sensors and force-based control algorithms allows for higher quality human-machine interaction, more realistic sensory feedback and telepresence. Beyond this, it can facilitate the deployment of important safety features [4]. Considerable amount of work relying on force sensors has previously been done, focusing on the development of miniaturized sensors to ease their integration with actual systems. In addition, they must be water resistant, sterilizable and insensitive to changes in temperature [5]. The major limitation of standard force sensors is thus the associated cost, since most surgical tools are disposable [4]. Alternatively, qualitative estimation of forces based on images has been proposed in the past [6, 5]. Mura *et al.* presented a vision-based haptic feedback system to assist the movement of an endoscopic device using 3D maps. Haouchine *et al.* estimates force feedback in robot assisted surgery using a biomechanical model of the organ in addition to the 3D maps. In contrast, including detailed material characteristics allows precise quantitative assessment of forces but calculating the properties of the materials still remains highly complex [7]. Recent research have proposed the use of neural networks for force estimation, such as in [8] where interaction forces using recurrent neural networks are estimated from camera acquisitions, kinematic variables and the deformation mappings. In a follow up, [7] used a neuro-vision based approach for estimating applied forces. In this approach tissue surface deformation is reconstructed by minimizing an energy functional.

In contrast, our method consists in estimating the force, based only on an OCT image of the deformation of the eye sclera. The method relies on a deep learning classifier for estimating quantiles of forces during robotized intravitreal injections. Our contribution hence consists of two critical stages. We first build a biomechanical model of the sclera in order to generate virtually infinite force-OCT image pairs for training any supervised machine learning method. In particular, we will show that this allows us to avoid the need for large data sets of real OCT images. Second, we train end-to-end, a two-layer image classifier algorithm with the synthetic images and the corresponding forces. This allows a very simple estimation process to take place, without the need of a specific image feature extraction method beforehand [7, 8]. The consequence of this two-stage process to produce synthetic data for training a NN is the ability to classify force ranges with 94% accuracy. We validate this claim on real OCT images collected on five *ex vivo* porcine eyes using a robotically-controlled needle.

2 Method

Deep learning has already been proposed to improve existent features of robotic assisted surgeries such as instrument detection and segmentation [9]. A key requirement for deep learning methods to work is the high amount of data to train on. Since for the moment, intravitreal injections are manually performed, there

is no available information on the force applied by the needle. In this paper, we build a numerical model of the relevant part of the eye to generate images of the deformed sclera under needle-induced forces. The simulations are parametrized to match experimental results and compensate for the lack of real data.

2.1 Numerical simulation for fast data generation

Biomechanical model: The deformation of the sclera under needle-induced forces can be approximated by modeling the eye as an elastic material, shaped like a half-sphere subject to specific boundary conditions. Since applied forces and resulting deformations remain small, we choose a linear relationship between strain ϵ and stress σ, known as Hooke's law (1):

$$\sigma = 2\mu\epsilon + \lambda tr(\epsilon)I \tag{1}$$

where λ and μ are the Lamé coefficients that can be determined from the Young's modulus $\mathbf{E}$ and Poisson's ratio $\boldsymbol{\nu}$ of the material. The linearity of Hooke's law leads us to a simple relation $\sigma = \mathbf{C}\epsilon$ where $\mathbf{C}$ is the linear elasticity constitutive matrix of homogeneous and isotropic materials. Boundary conditions are then added. Dirichlet boundary conditions are used to prevent rigid body motion of the sclera, while a constant pressure is applied on the inner domain boundary to simulate the intraocular pressure (IOP). The IOP plays an important role in the apparent stiffness of the eye and its variability is well studied, as high eye pressure can be an indication for glaucoma. The external force due to IOP is simply given by $F = S \times P$ where S is the surface area of the eye given in m^2 and P the IOP in Pa. The external forces of the system are formed by the IOP, the force induced by the needle and gravity. Scleral thickness also plays an important role in the deformation of the sclera so it has to be taken into account.

Finite Element simulation: We solve the equation for the constitutive model using a finite element method. To discretize the eyeball, which we consider as nearly spherical, we generated a quadrilateral surface mesh of a half sphere of radius 12 mm, discretized using the Catmull-Clark subdivision method. This quadrilateral surface is then extruded according to the scleral thickness to generate almost regular hexahedra. The deformation is specified by the nodal displacements $\mathbf{u}$ and the nodal forces $\mathbf{f}$, according to the following equation:

$$\mathbf{Ku} = \mathbf{f_p} + \mathbf{f_g} + \mathbf{f_n} \tag{2}$$

The matrix $\mathbf{K}$ is the stiffness matrix, and can be computed thanks to the elastic parameters of the material $\mathbf{E}$ and $\boldsymbol{\nu}$. $\mathbf{E}$ is Young's modulus and is a measure of the stiffness of the material while $\boldsymbol{\nu}$ is Poisson's ratio and estimates the compressibility of the material. According to the values in the literature [10] we set $\mathbf{E} = 0.25\ MPa$ and $\boldsymbol{\nu} = 0.45$. To model the needle pushing on the sclera, we apply a local force $\mathbf{f_n}$ on a subset of nodes within a small region of interest near the virtual needle tip. $\mathbf{f_g}$ is the force due to gravity, and $\mathbf{f_p}$ the force normal to the surface due to the IOP. To compute accurately the deformation of the

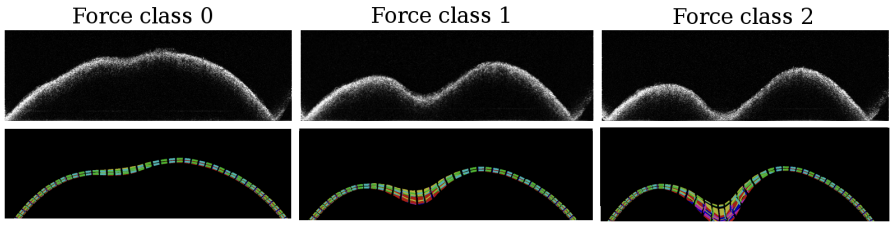


Fig. 1. Real OCT images of the sclera and their corresponding simulated images for the three different force ranges.

sclera, the finite element mesh needs to be sufficiently fine (14,643 hexahedral elements in our case), resulting in about 5 seconds of computation to obtain the solution of the deformation. Since we need to repeat thousands of times this type of computation in order to generate the training data set, we take advantage of the linearity of the model and pre-compute the inverse of $\mathbf{K}$. This significantly speeds up the generation of the training data. In Fig. 1 are shown different examples of the output of the simulation (bottom) matching the OCT images (top). The simulated images correspond to a 2D cross-section of the entire 3D mesh.

2.2 Neural network image classification

Our goal is to create an artificial NN model to predict the force based solely on a single OCT image of the deformation. Using cross sectional OCT images allows one to visualize millimeters beneath the surface of the sclera and therefore provides information about the induced deformation. Using 2D OCT B-scans over 3D scans is preferred here as can be acquired at high-frame rates using low cost hardware [11]. Instead of estimating the force as a scalar value, we opt to estimate an interval of forces in order to be less sensitive to exact force values. As such, we set up our inference problem as a classification task where a probability for a class label (i.e. a force interval) is desired.

Given our above simulation model, we can now train our prediction model with virtually an *infinite* amount of data. In our case, the forces applied in the experiments varied between 0.0 N to 0.08325 N. To ensure quality of the simulations (i.e. small displacements) forces going from 0.0 N to 0.06 N were applied on the Finite Element mesh of the sclera. The objective being the in-time withdrawal of the needle to limit scleral damage, we decided to label the forces only using three classes. The classes are such that the first one translates no danger at damaging the sclera, the second means that a considerable force is being applied and the last class triggers the alarm to withdraw the needle. If we set the alarm threshold to 0.03 N, we have the following class ranges (see Fig. 1 for the deformation corresponding to each class):

- Class 0: force values smaller than 0.005 N
- Class 1: force values from 0.005 N to 0.03 N
- Class 2: force values bigger than 0.03 N

Our NN then consists of four layers: an input layer having 5600 neurons, two fully connected hidden layers having 600 neurons each and an output layer of 3 neurons. The output layer is softmax activated. ReLu activation was used as well as a 0.8 dropout factor between layers. For training, we used Gradient descent optimization and the cross-entropy loss function. The network was implemented using Tensorflow and the Keras Python library and was trained from scratch. Validation was performed on unseen simulated images.

3 Results

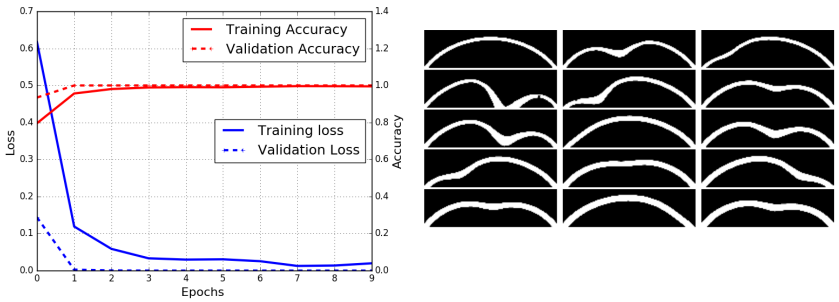
In this section we first report the real data acquisition on *ex vivo* porcine eyes. Then, the virtually generated data set and the following training are presented. And last, we validate the NN on unseen real data.

3.1 Experimental set-up

To validate the ability of the NN on real data, we collected pairs of force-OCT images from *ex vivo* porcine eyes. Five porcine eyes were obtained from the local abattoir and transported in ice. Experiments began within 3 hours of death and were completed between 4 and 10 hours postmortem. During the experiments the eyes were moisturized with water. For the experiments, the eyes were fixed with super glue on a 3D-printed holder to ensure fixed boundary conditions on the lower half of the eyeball. The IOP was measured with a tonometer. For all the eyes the IOP was close to 2 $mmHg$ (i.e. 266 Pa). The IOP is low since it decreases dramatically after death [12] and no fluid injection was made during the experiments. A medical robot [3] was used to guide a needle while measuring the applied force. A 22G Fine-Ject needle (0.7×30 mm) was mounted at the tip of the end effector of the robot. The needle was placed as close as possible to the B-scan but without intersecting it to avoid shadows on OCT images. The margin between the needle and the B-scan was taken into account in the simulation. The robot moved towards the center of the eye (i.e. normally to the sclera) and forces in the direction of movement were continuously recorded with a sensor. To collect OCT images an imaging system was used to record B-scans over time. The OCT device used 840 $nm \pm 40$ nm wavelength light source, with an A-scan rate of 50 kHz and 12 bits per pixel. The resulting 2D image had a resolution of 512×512 pixels, corresponding to roughly 15×4 mm. As the images were acquired at a lower rate than the forces, the corresponding forces were averaged over the imaging time for each frame (i.e. two seconds after the starting moment of the line move). Overall, 54 trials were performed. For each eye, one position near the corneal-scleral limbus was chosen and several line moves were performed in the same direction.

3.2 Data generation for neural network training

Measurements on ten porcine eyes are made to estimate the thickness of the sclera at the locations where the forces are applied. The values range from



(a) (b)

Fig. 2. (a) Loss and Accuracy curves for training and validation sets. (b) Fragment of the training data set generated by our numerical simulation.

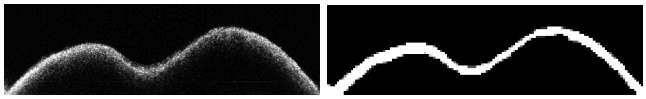
460 μm to 650 μm with a mean value of 593 μm. From the literature [13], the thickness near corneal-scleral limbus is estimated to be between 630 μm and 1030 μm. Hence, we simulated scleras of five different thicknesses : 400 μm, 500 μm, 600 μm, 700 μm and 800 μm. The IOP is the same for all the eyes in the experiments so it is fixed at 266 Pa for all simulations. For the smallest thickness, we generated 3200 images of deformed scleras undergoing the stated IOP and forces going from 0.0 N to 0.045 N. For each of the other four thicknesses, we generated 4000 images of deformed scleras where the forces vary from 0.0 N to 0.06 N at different random locations. For each thickness, the simulation took approximatively half an hour. Overall, a data set of 19200 synthetic images was generated within 2 hours and a half (see Fig. 2(b)). The generated images look like images in Fig. 1 below. This images are post-processed with a contour detection algorithm using OpenCV functions and binarized to obtain the images used to train the NN (see Fig. 2(b)).

To train the NN, the data set is split such that 90% of the images are for training and the remaining 10% for validating. Hence, the NN is trained on 17280 images and validated on the other 1920 images. In Fig. 2(a) the accuracy and loss of the model are shown on both training and validation data sets over each epoch. The validation accuracy curve shows 100% accuracy when classifying unseen synthetic images. This curve is above the training accuracy curve probably because of the high dropout applied during the training.

3.3 Tests on unseen data

The aim of our work is to classify force ranges from real OCT images of the deformed scleras. All the OCT images obtained during the experiments (see Fig. 3(a)) are blurred and thresholded to obtain similar images to the synthetic ones (see Fig. 3(b) and Fig. 2(b)).

Now that the OCT images look like the simulated ones they can be given as input to the NN for predictions. For each OCT acquisition, the force measured



(a) (b)

Fig. 3. (a) Unprocessed OCT image. (b) Blurred and thresholded OCT image.

by the robot is converted into a class label (going from 0 to 2) and is considered as the ground truth (target class). The performance of the classification is given in the confusion matrix in table 1. Each raw of the table gives the instances in a target class and each column the instances in a predicted class. For each class, the correct decisions are highlighted in red. Overall the accuracy of the classifier is 94%.

A force class with label 0 is understood as a very small force meaning that there is no risk of damaging the sclera. On the contrary, a force class with label 1 means the sclera is being considerably deformed (forces in this range go from 0.005 N to 0.03 N) and a force class with label 2 means that the sclera is potentially being damaged and a withdrawal of the needle is advised. With the eyes used during the experiments, the lowest precision of the NN was found for target class 0 (71%). For target class 2, the precision is of 100%. It is not dramatic if forces of range 0 and 1 are misclassified since the risk of damaging the sclera with forces so close to the alarm threshold is almost null (assuming the threshold is correctly set). On the other hand it is essential that the forces of range 2 are predicted correctly.

Prediction / **Target**	**Class 0**	**Class 1**	**Class 2**	Precision
Class 0	5	2	0	0.71
Class 1	0	14	1	0.93
Class 2	0	0	32	1.00
Recall	1.00	0.88	0.97	

Table 1. Confusion matrix

4 Conclusion and Discussion

In this paper, we have proposed a method aimed at improving the safety features of upcoming robotized intravitreal injections. Our vision-based method combining numerical simulation and Neural Networks, predicts the range of the force applied by a needle using only 2D images of the scleral deformation with high accuracy. It indicates in real-time the need to withdraw the needle as soon as a

certain alarm threshold is reached. To avoid the need of large training data sets, the NN is trained on synthetic images from a simulated deformed sclera.

It is worth mentioning that more complex scenarios could be simulated such as different eye sizes, variable needle insertion angles and different intraocular pressures. We also propose to improve the simulated images to better match real surgical scenarios. In particular it seems important to add the shadows induced by the needle in the OCT image. From an imaging point of view, we know that the predictions of the NN are very sensitive to the image framing and scaling. To address this issue, we plan to randomly crop each simulated image, and enlarge the data set with it. We might also train a Convolutional Neural Network to reduce sensitivity to image properties and allow for more accurate predictions.

References

1. Ullrich, F., Michels, S., Lehmann, D., Pieters, R.S., Becker, M., Nelson, B.J.: Assistive device for efficient intravitreal injections. Ophthalmic Surgery, Lasers & Imaging Retina, Healio **47**(8) (June 2016) 752–762
2. Meenink, H., Hendrix, R., Naus, G., Beelen, M., Nijmeijer, H., Steinbuch, M., van Oosterhout, E., De Smet, M.: Robot-assisted vitreoretinal surgery. (October 2012) 185–209
3. Weber, S., Gavaghan, K., Wimmer, W., Williamson, T., Gerber, N., Anso, J., Bell, B., Feldmann, A., Rathgeb, C., Matulic, M., Stebinger, M., Schneider, D., Mantokoudis, G., Scheidegger, O., Wagner, F., Kompis, M., Caversaccio, M.: Instrument flight to the inner ear. (March 2017)
4. Haidegger, T., Beny, B., Kovcs, L., Beny, Z.: Force sensing and force control for surgical robots. Proceedings of the 7th IFAC Symposium on Modelling and Control in Biomedical Systems (August 2009) 401–406
5. Haouchine, N., Kuang, W., Cotin, S., Yip, M.: Vision-based Force Feedback Estimation for Robot-assisted Surgery using Instrument-constrained Biomechanical 3D Maps. IEEE Robotics and Automation Letters (February 2018)
6. Mura, M., Abu-Kheil, Y., Ciuti, G., Visentini-Scarzanella, M., Menciassi, A., Dario, P., Dias, J., Seneviratne, L.: Vision-based haptic feedback for capsule endoscopy navigation: a proof of concept. **11** (May 2016)
7. Aviles, A.I., Alsaleh, S., Sobrevilla, P., Casals, A.: Sensorless force estimation using a neuro-vision-based approach for robotic-assisted surgery. (April 2015) 86–89
8. Aviles, A.I., Marban, A., Sobrevilla, P., Fernandez, J., Casals, A.: A recurrent neural network approach for 3d vision-based force estimation. (October 2014) 1–6
9. Pakhomov, D., Premachandran, V., Allan, M., Azizian, M., Navab, N.: Deep residual learning for instrument segmentation in robotic surgery. (March 2017)
10. Asejczyk-Widlicka, M., Pierscionek, B.: The elasticity and rigidity of the outer coats of the eye. **92** (November 2008) 1415–8
11. Apostolopoulos, S., Sznitman, R.: Efficient oct volume reconstruction from slitlamp microscopes. IEEE Trans. on Biomedical Engineering **64**(10) (2017) 2403–2410
12. Gnay, Y., Basmak, H., Kenan Kocaturk, B., Sahin, A., Ozdamar, K.: The importance of measuring intraocular pressure using a tonometer in order to estimate the postmortem interval. **31** (April 2010) 151–5
13. Olsen, T., Sanderson, S., Feng, X., C Hubbard, W.: Porcine sclera: Thickness and surface area. **43** (September 2002) 2529–32