

---

## Research Article

---

Open Access, Volume 3

# Eye biometry prediction from ultrasound images using convolutional neural networks

John Lipor<sup>1</sup>; Hadi Khazaei<sup>2\*</sup>; Danesh Khazaei<sup>3</sup>; Andrew Franck<sup>3</sup>; John Ng<sup>4</sup>; Faryar Etesami<sup>5</sup>

<sup>1</sup>Associate Professor, Electrical & Computer Engineering, Portland State University, USA.

<sup>2</sup>Research Scholar, OHSU Portland, OR, USA.

<sup>3</sup>Student, Portland State University, USA.

<sup>4</sup>Professor and Head, Division of Oculofacial Plastic, Orbital, and Reconstructive Surgery, Oregon Health and Science University, Portland, Oregon, USA.

<sup>5</sup>Emeritus Professor, Department of Mechanical Engineering, Portland State University, USA.

**\*Corresponding Author: Hadi Khazaei**

Research Scholar, OHSU Portland, OR, USA.

Email: khazaei@pdx.edu

Received: Sep 12, 2023

Accepted: Oct 10, 2023

Published: Oct 17, 2023

Archived: www.jclinmedimages.org

Copyright: © Khazaei H (2023).

### Abstract

Medical imaging is important in clinical diagnosis and individualized treatment of eye diseases. Ultrasound imaging is one of the most prominent technologies to evaluate the orientations, anomalies and anatomical features of the eye and orbit. However, the interpretation of the data obtained from such studies is best left to expert physicians and technicians who are trained and well-versed in analyzing such images. This technology can provide high-resolution information regarding anatomic and functional changes. In recent years, imaging techniques have developed rapidly, together with therapeutic advances. However, with the increasing sophistication of imaging technology, comprehension and management of eye disease has become more complex due to the large numbers of images and findings that can be recorded for individual patients, as well as the hypotheses supported by these data. Thus, each patient has become a “big data” challenge. Conventional diagnostic methods greatly depend on physicians’ professional experience and knowledge, which can lead to a high rate of misdiagnosis and wastage of medical data [6]. The new era of clinical diagnostics and therapeutics urgently requires intelligent tools to manage medical data safely and efficiently.

### Introduction

There are different ways for ultrasound-based diagnostic procedures. Depending on the application, the sonographer acquires either a single image or an image series. The second approach is better when a further automated image-processing step is introduced. Simultaneous analysis of multiple data provides reliable results, less prone to artifacts and outliers. At the same time, the analysis of the whole recording might be disturbed by strongly distorted data or the artifacts influencing the geometry of visualized structures, appearing on the part of frames. Consequently, it leads to misclassification, false-

positive detections, and finally, inaccurate results of measurements. Therefore, the overall goal of this study was to develop and evaluate the classification framework, which enables robust and fast POCUS series analysis. Ocular ultrasonography in the ambulatory and critical care setting has become an invaluable diagnostic tool for patients presenting with traumatic or atraumatic vision and ocular complaints. Sonographic bedside evaluation is intuitive and easy to perform and can accurately diagnose a variety of pathologies. These include detachment or hemorrhage of the retina or vitreous, lens dislocation, retrobulbar hematoma or air, as well as ocular foreign bodies, infections, tumors, and increased optic nerve sheath diameter that can be

assessed in the setting of suspected increased intracranial pressure. The ocular anatomy is easy to visualize with sonography, as the eye is a superficial structure filled with fluid. Over the last two decades, many scientific publications have documented that ocular ultrasound in emergent or critical care settings is an accurate diagnostic tool and expands and improves emergency diagnosis and management.

### Research methodology

**Significance:** There is an abundance of ultrasound datasets for various use-cases, which can be used to generate DNN-based models for classification and segmentation. For instance, the breast ultrasound image dataset presented by Al-Dhabyani et al. (Al-Dhabyani et al., 2020), which is composed of normal, benign, and malignant images that can be used to train to a model to act as a classifier. Similarly, the POCUS dataset, presented by Born et al. (Born et al., 2020), and the COVIDX-US dataset, by Ebadi et al. (Ebadi et al., 2021), are openly accessible for building DNN-based clinical assistants that can aid in the analytics and diagnosis of COVID19. Leclerc et al. (Leclerc et al., 2019) presented a cardiac ultrasound electrocardiography dataset containing image sequences with two and four-chamber views of the heart of 500 patients. Likewise, there are a wide number of ultrasound datasets for diagnosing and analyzing several internal bodies' organs.

**Innovation:** The required ultrasound data was generated and collected using ocular phantom tissue, which has been typically used to train sonographers to assess the development and condition of the eye. The eye model includes anatomical structure and key organic features that can be observed and used to train the sonographer to assess the eye anatomy (like vertical and horizontal meridians, circumferences, area and volume) and internal orbit structures (like lacrimal gland, optic nerve, muscles and blood vasculature, etc.). The biometric parameters of the eye can also be measured/learned using an ultrasound of the eye phantom at the appropriate positions or the correct diagnostic planes. We use the butterfly IQ plus, which is interfaced with 3D facilities to collect and process the data and generate the final ultrasound image. The Anatomically Intelligent Ultrasound (AIUS) imaging technology deploys, an Ultrasound Phantom Dataset with DNN evaluations advanced organ modeling and imaging techniques to generate a three-dimensional image of the ocular phantom using the default settings for the "Ophthalmology" imaging option. The imaging depth was set to 5 cm and captured at a 10 Hz frame rate. Sufficient ultrasound gel is applied on the phantom model to ensure acoustic coupling with the probe, thereby reducing acoustic impedance, and enabling clear imaging. We executed two protocols to collect the images used in this dataset [1]. Protocol-I the probe is placed on the phantom surface and navigated to the correct diagnostic planes that can be used for the measurement of the three primary biometric parameters of the eye, namely the horizontal plane, which is used to obtain the transverse diameter), vertical standard plane, which is used to estimate the eye circumference, and the sagittal standard plane, which is used to estimate the optic nerve diameter. The correct diagnostic planes were identified using the proposed research protocol. To further enrich the dataset, after the acquisition of several frames at the correct diagnostic plane, we tilt, rotate, or traverse the ultra-

sound probe in random directions to collect more information [2]. Protocol-II: The focus of this protocol is to obtain images capturing the anatomies of the eye phantom in the generated images. We do this by navigating the probe to obtain the lacrimal gland, optic nerve, muscles and orbital fats, individually or combined, in the picture and move the probe in different directions to obtain a heterogeneous set of images capturing the orbital anatomies. Furthermore, the phantom tissue model was also rotated and placed in the four possible orientations [head up or down, view front or back, when collecting the ultrasound data, to potentially mimic the real-life behavior of anatomical orientation and presentation. Additionally, the probe orientation was also changed between horizontal and vertical, with respect to the midline, when the data was collected to enhance the dataset with more information.



### Approach

**Background and justification:** The data streams obtained by the butterfly ultrasound system are converted to PNG image sequences, of dimension 664x388, using custom in-house software for easier labeling, annotating, and processing. The stored PNG files are annotated using a customized version of the computer vision annotation tool -Desmos which is an online interactive graphing calculator utilized in this study to measure the desired dimensions of the eyeball. The salient features of this online program include the ability to pre-populate a plot with the necessary formulas to calculate the diameters, circumference, and area of the eyeball, which was treated like a simple ellipse. A template program was created, and the ultrasound images were each uploaded to the program and the coefficients were manipulated to yield manual readings of the dimensions. The ultrasound images were superimposed onto the graph to scale.

Each acquired ultrasound frame was subsequently annotated by scientists with experience in eye ultrasound imaging. The sequences obtained using Protocol-I are labeled as a correct diagnostic plane for one of the three biometric parameters or as a non-diagnostic plane. Since the number of data samples obtained for each of the diagnostic planes is quite smaller than the non-diagnostic plane output class, the samples were augmented in each of the other three output classes to ensure

equal representation of data across all classes. The data obtained using Protocol-II is, first, labeled with eye orientations, as discussed earlier, based on the position of the phantom midline when the scans are made. Next, the images are tagged with the anatomies present in the image, such as lacrimal gland, optic nerve, muscles and orbital fat. The images are subsequently exhaustively annotated. Images that do not contain any vital and/or relevant information regarding the orbital structures are not labeled or annotated. These finalized labels and annotations, for each valid image in the dataset, can be used to train various deep learning models as required.

**Preliminary studies and feasibility:** We analyzed a dataset consisting of 107 ultrasound images of eyes captured with the Butterfly iQ+ ultrasound probe. Of these images, 77 are taken from typical/healthy patients, while 30 are taken from eyes with known abnormalities. The goal is to utilize a convolutional neural network (CNN) to predict both the vertical and horizontal diameter of each eye, removing the need for the costly procedure of expert analysis of images.

**Research design:** The dataset is split into a training and validation set as follows. For the typical images, we perform an 80/20 split, so that 64 images are used for training and 13 are used for validation. For the atypical images, we allot 20 to the training set and 10 to the test set, in order to ensure that there sufficiently many atypical images in the validation set. The training and validation datasets are much smaller than those typically used when training a CNN. As a result, we investigated the use of various forms of image augmentation. However, we did not find any augmentation to be beneficial to the training process. We note that, in contrast to images of typical objects (e.g., dogs, cars), ultrasound images have distinct spatial structure, with the main image having a specific orientation and an objective scale on the right side. This fact may account for the lack of benefit seen, and further investigation with larger labeled datasets is a topic of future research. Our model is based on the Network-in-Network (NiN) structure [1], which utilizes  $1 \times 1$  convolutions to aggregate information across image channels without destroying spatial structure. While we experimented with other network types, we found that any network employing linear layers ultimately learned the median value of the dataset. However, we note that the choice of network structure may vary as more images are considered in the training set. Each NiN block consists of [1] a convolutional layer of arbitrary kernel size followed by a rectified linear unit (ReLU) (2) A  $1 \times 1$  convolutional layer followed by a ReLU, and (3) A second  $1 \times 1$  convolutional layer followed by batch normalization and a ReLU. The first convolutional layer of each NiN block learns filters of user-specified size to detect the salient features in the image, while the subsequent two  $1 \times 1$  convolutional layers act as nonlinear predictions for each pixel across all image channels. In this way, the NiN block performs convolution and a nonlinear transformation without destroying spatial structure. For our model, we use four NiN blocks with 96 output channels (filters) each. The first layer uses a kernel size of  $11 \times 11$ , while the next three use  $5 \times 5$  kernels. A final NiN block consists of two output channels using kernel size  $3 \times 3$ , and the final prediction is performed by a global average pooling over each channel. The first channel corresponds to prediction of vertical diameters, while the second corresponds to horizontal diameters. We train the above network using the mean absolute error (MAE), optimized using Adam [2] with a learning rate of 0.01 over 1000 epochs.

### Expected outcomes

Figure 1 shows the training and validation error versus epoch. Interestingly, even for a very large number of epochs, the validation loss does not appear to diverge greatly from the training loss. This is likely due to the homogeneity of the training and validation sets, in which the images are very similar. We note that a more careful consideration of overfitting may be necessary when larger training and validation sets become available. The final validation MAE is 0.0925, corresponding to a relative error of 3.96%. For comparison, a simple algorithm that always predicts the median diameters of the training set results in a validation error of 8.72%, indicating that our trained network learns nontrivial predictions corresponding to the actual images, even from this small dataset.

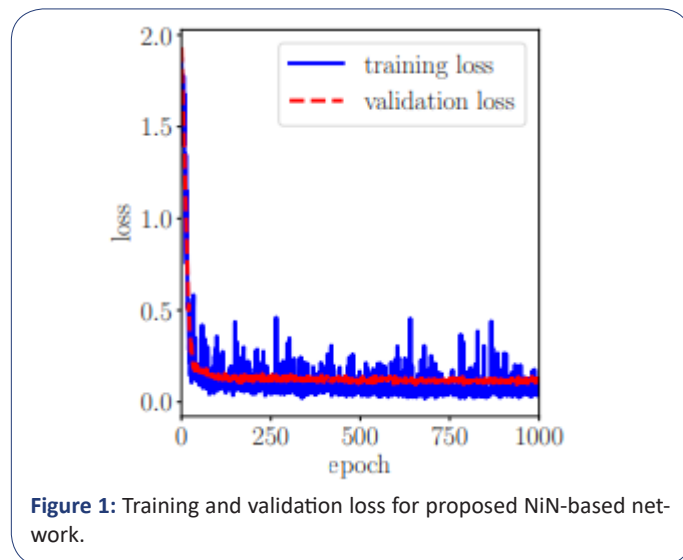
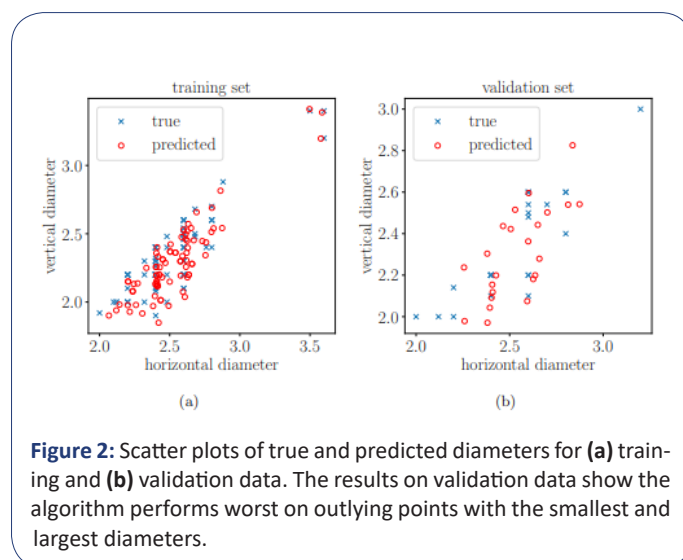
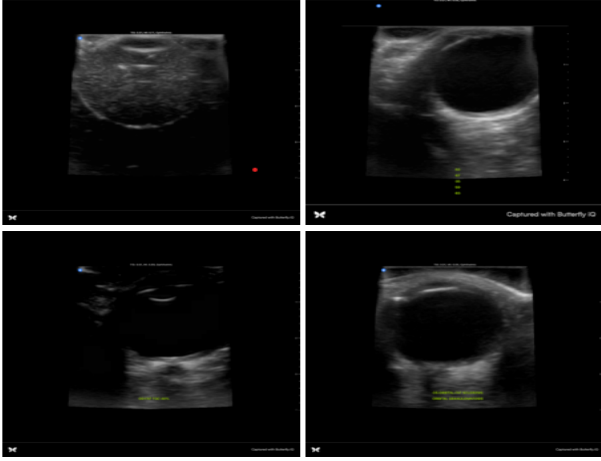


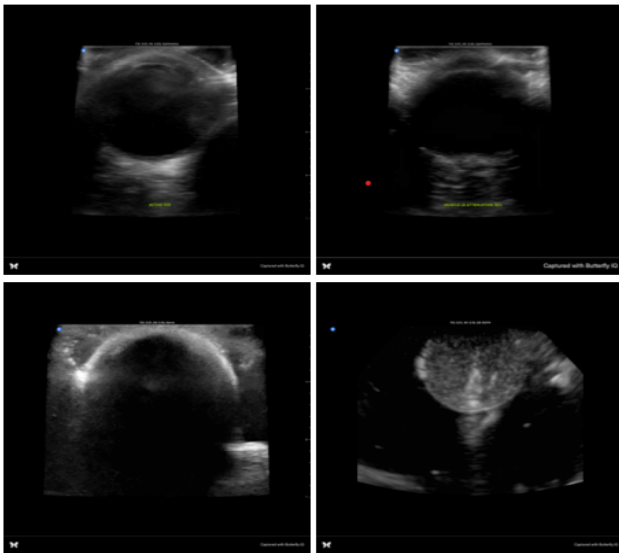
Figure 2 shows a scatter plot of the true and predicted diameters for (a) The training dataset and (b) The validation dataset. The figure shows that the network can accurately capture the training data while still generalizing to the validation set. For the validation data, we see that the network overpredicts the smallest horizontal diameters (bottom left points in Figure 2(b)).



**Figure 2:** Scatter plots of true and predicted diameters for (a) training and (b) validation data. The results on validation data show the algorithm performs worst on outlying points with the smallest and largest diameters.



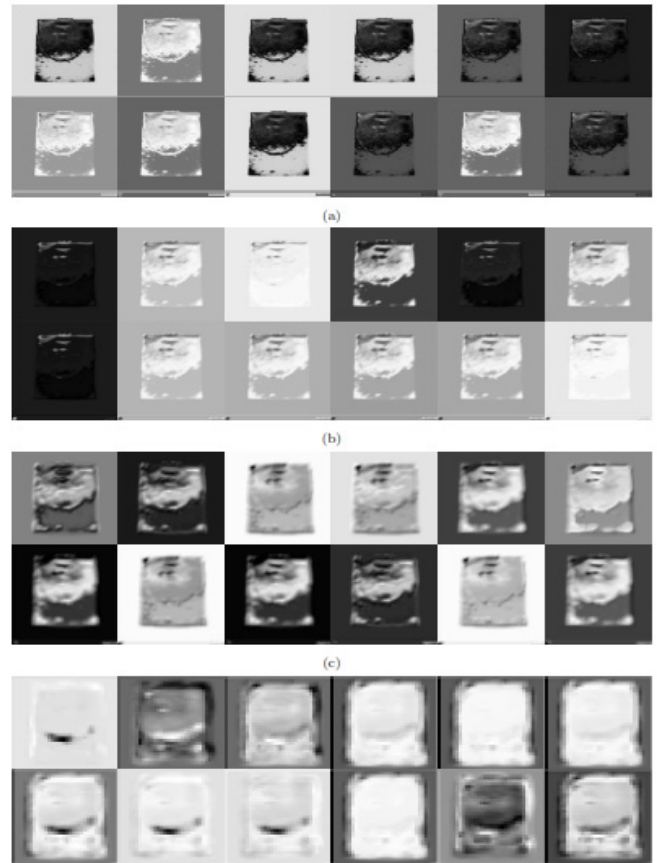
**Figure 3:** Ultrasound images corresponding to four most accurate predictions in validation set. Relative errors range from 0.12%-0.71%. The images all contain eye images with clearly identified boundaries.



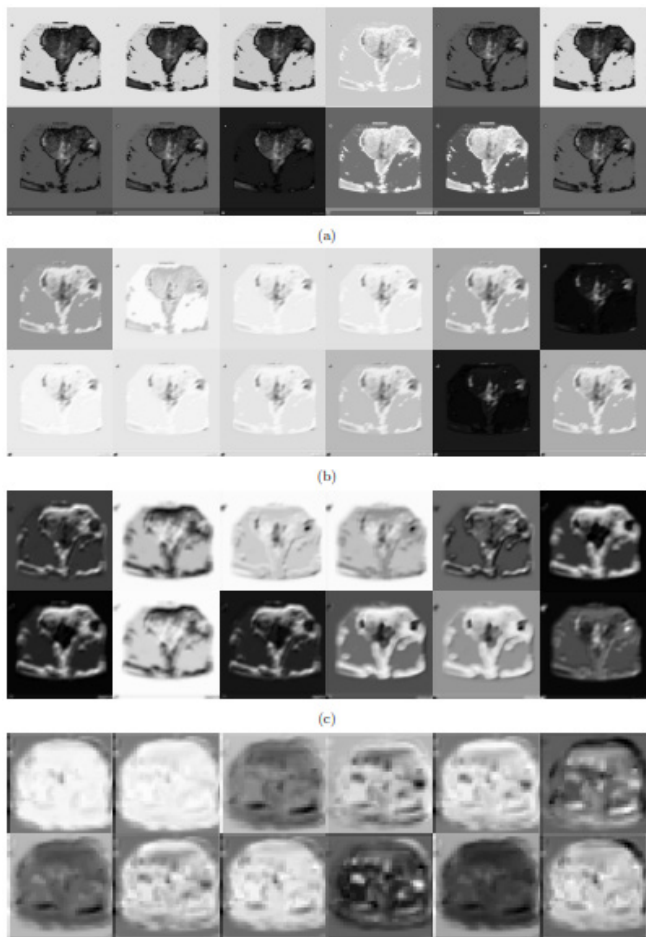
**Figure 4:** Ultrasound images corresponding to the four least accurate predictions in validation set. Relative errors are 5.03%, 6.90%, 8.59%, and 30.09% (left to right, top to bottom). Failed predictions are likely due to lack of clear visual boundaries (first three images) or a different ultrasound field (fourth image).

These points likely correspond to images of atypical scenarios, for which there is less training data. Figure 3 shows the images corresponding to the four most accurate predictions. In all cases, the entire eyeball is within the frame and largely surrounded by contrasting tissue, with the lens clearly visible. Figure 4 shows the images corresponding to the worst four predictions. In this case, we see that images either have cloudiness within the eyeball region or dark regions that make it difficult to clearly delineate the boundary of the eyeball. We note that the worst two predictions correspond to the largest and smallest examples in the dataset, indicating that the algorithm is biased toward typical eyeball sizes. In the most extreme example (Figure 4, right image), we see that the ultrasound scan covers much more of the image than other examples in the dataset, that the model had a difficult time determining the absolute scale of the eye.

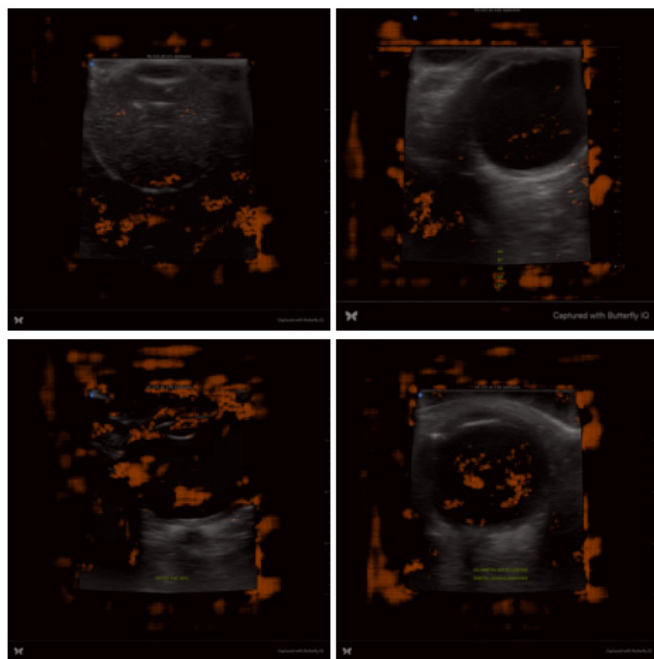
To gain insight into our network, we visualize the learned representations (feature maps) at the intermediate and output layers of the network. Figure 5 shows example feature maps at the output of the first four NiN blocks for the most accurately predicted image (Figure 3, top left). We see that the first NiN block performs broad-scale feature learning, picking out edges of various orientations and discovering contrasting regions. The second block appears to perform some smoothing, reducing the variation in locations away from the eyeball. The third block outlines both the eyeball and the ultrasound field, smoothing all other regions. Finally, the fourth block further refines and smooths these estimates. Figure 6 shows the same outputs for the lowest-accuracy prediction (Figure 4, bottom right). We note that the true vertical and horizontal diameters for this image are 20 mm and 20 mm, while the predicted diameters are 25 mm and 27 mm. The feature maps show that the model fails to determine the outline of the eyeball, and perhaps mistakes the outline of the ultrasound field for the eyeball in Figure 6(d), which would explain the upward bias in the prediction for this image. Finally, we consider a network saliency map obtained by guided backpropagation [3], which aims to determine which parts of the input image had the greatest impact on the final prediction by examining the gradients corresponding to each pixel location. For visualization purposes, we normalize all values and set values greater than 0.01 to 0.5; hence the resulting heatmaps do not indicate the strength of influence of each pixel. Figure 7 shows the saliency maps for the four most accurate images, while Figure 8 shows the maps for the worst four predictions. Red locations correspond to pixels that have greater influence over the model predictions.



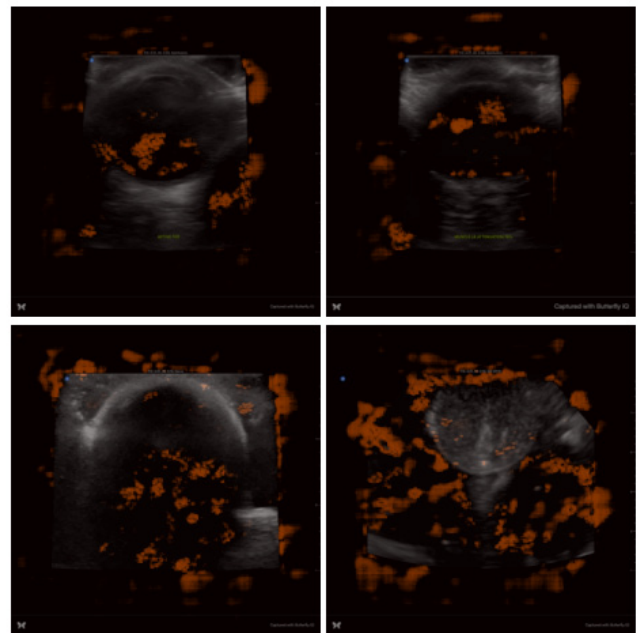
**Figure 5:** Feature maps output by each NiN block for lowest-error image. (a) First NiN block, (b) Second block, (c) Third block, (d) Fourth block. At the fourth layer, the network appears to learn the boundary of the ultrasound field, long with an outline of the eye.



**Figure 6:** Feature maps output by each NiN block for highest-error image. **(a)** First NiN block, **(b)** Second block, **(c)** Third block, **(d)** Fourth block. The third- and fourth-layer outputs indicate the model may have mistaken the ultrasound field for the eye, resulting in the high prediction for this image.



**Figure 7:** Saliency maps (via guided backpropagation) corresponding to four most accurate predictions in validation set. The model appears to determine the bounds of the ultrasound field, as well as the relative size of the eyeball within these bounds.



**Figure 8:** Saliency maps (via guided backpropagation) corresponding to four least accurate predictions in validation set. The model appears to either have trouble determining extremal points of the eyeball (bottom left) or bounds on the area of interest within the image (bottom right).

**Potential problems and alternative strategies:** For the top four images, as well as the best two of the worst four predictions, we see that the network focuses on bounding the ultrasound field as well as on a few extremal points within the eyeball itself. For the worst two predictions (Figure 8, bottom row), we see that the model fails to determine the boundaries of either the eyeball or of the ultrasound field. Overall, we see that the network seeks to determine bounds around the region of interest within the image, which may provide a global scale, as well as a few points within the eyeball, which provide the relative size. Together, these result in the successful predictions reported above.

#### Future directions

Relevant ocular ultrasound experience to obtain information regarding the [1] Diagnostic plane, [2] Eye orientation, [3] Ocular anatomy, and [4] their anomalies, using box annotations. The generated dataset is used to train a variety of deep learning models to illustrate the model's ability to extract vital information, which can be used to accurately distinguish among an Ultrasound Phantom Dataset with DNN evaluations the classes in different categories and detect the anomalies. Furthermore, To evaluate their deployability in portable resource-constrained devices, we evaluated the capability of a smaller DNN compressed using pruning and quantization to illustrate that smaller DNNs are equally competent at extracting relevant information from the dataset and are capable of execution on resource-constrained devices and embedded platforms.

#### References

1. Khazaei H, Khazaei D, Ashraf D, Mikkilineni S, Ng JD. Overview of Orbital Ultrasonography. *Ann Ophthalmol Vis Sci.* 2022; 5(1): 1028.
2. Joseph DP, Pieramici DJ, Beauchamp NJ. Computed tomography in the diagnosis and prognosis of open globe injuries. *Ophthalmology.* 2000; 107: 1899-906.

3. Arey ML, Mootha VV, Whittemore AR, Chason DP, Blomquist PH. Computed tomography in the diagnosis of occult open globe injuries. *Ophthalmology*. 2007; 114: 1448-52.
4. Turkcuer I, Atilla R, Topacoglu H, et al. Do we really need plain and soft-tissue radiographies to detect radiolucent foreign bodies in the IED. *Am J Emerg Med*. 2006; 24(7): 763-8.
5. Bray PW, Mahoney JL, Campbell JP. Sensitivity and specificity of ultrasound in the diagnosis of foreign bodies in the hand. *J Hand Surg Am*. 1995; 20(4): 661-6.
6. Oikarinen KS, Nieminen TM, Makkarainen H, Pyhtinen J. Visibility of foreign bodies in soft tissue in plain radiographs, computed tomography, magnetic resonance imaging and ultrasound. In vitro study. *Int J Oral Maxillofac Surg*. 1993; 22(2): 119-24.
7. Aras MH, Miloglu O, Barutcugil C, Kantarci M, Ozcan E, Harorli A. Comparison of the sensitivity for detecting foreign bodies among conventional plain radiography, computed tomography and ultrasonography. *Dentomaxillofac radiol*. 2010; 39(2): 72-8.
8. Blaivas M, Theodoro D, Sierzenski P. A study of bedside ocular ultrasonography in the emergency department. *Acad Emerg Med*. 2002; 9: 791-9.
9. Blaivas M: Bedside emergency department ultrasonography in the evaluation of ocular pathology. *Acad Emerg Med*. 2000; 7(8): 947-50.
10. Soldatos T, Chatzimichail K, Papathanasiou M, Gouliamos A: Optic nerve sonography. A new window for the non-invasive evaluation of intracranial pressure in brain injury. *Emerg Med J*. 2009; 26: 630-4.
11. Sargsyan AE, Dulchavsky AG, Adams J, Melton S, Hamilton DR, Dulchavsky, SA. Detection of simulated intra-ocular foreign bodies by minimally trained personnel. *Avait Space Environ Med*. 2008; 79(1): 58-61.
12. Weichel ED, Colver MH. Combat ocular trauma and systemic injury. *Curr Opin Ophthalmol*. 2008; 19(6): 519-25.
13. Colyer MH, Chun DW, Bower KS, Dick JS, Weichel ED. Perforating globe injuries during operation Iraqi Freedom. *Ophthalmology*. 2008; 115(11): 2087-93.
14. Colyer MH, Weber ED, Weichel ED, et al. Delayed intraocular foreign body removal without endophthalmitis during operations Iraqi Freedom and Enduring freedom. *Ophthalmology*. 2007; 114: 1439-47.
15. Sawyer N. Ultrasound imaging of penetrating ocular trauma. *J Emerg Med*. 2009; 36(2): 181-2.
16. Blaivas M. Bedside emergency department ultrasonography in the evaluation of ocular pathology. *Acad Emerg Med* 2000; 7:947-950.
17. Evans LV, Dodge KL, Shah TD, et al. Simulation training in central venous catheter insertion. Improved performance in clinical practice. *Acad Med*. 2010; 85:1462-1469.
18. Keenan SP. Use of ultrasound to place central lines. *J Crit Care*. 2002; 17: 126-137.
19. Sharma R, Brunette DD. *Ophthalmology*. In: Marx J, Hockberger R, Walls R (eds). *Rosen's Emergency Medicine*. 7th ed. Philadelphia, PA: Elsevier Mosby. 2010; 226-239.
20. Blaivas M. Bedside emergency department ultrasonography in the evaluation of ocular pathology. *Acad Emerg Med* 2000; 7:947-950.
21. Evans LV, Dodge KL, Shah TD, et al. Simulation training in central venous catheter insertion. Improved performance in clinical practice. *Acad Med*. 2010; 85:1462-1469.
22. Keenan SP. Use of ultrasound to place central lines. *J Crit Care*. 2002; 17: 126-137.
23. Yan H, Ahmed AS, Han J, Cui J, Yu J. A vitreous hemorrhage animal model in rabbits using force percussion injury. *Curr Eye Res*. 2009; 34: 717-726.
24. Glickman RD, Maswadi SM, Dodd SJ, Gao J. Mapping 3-dimensional temperature gradients using magnetic resonance thermography in an ocular phantom during laser-induced hyperthermia. *Invest Ophthalmol Vis Sci*. 2003; 44: 3631.
25. Kendall JL Faragher, JP. Ultrasound-guided central venous access. A homemade phantom for simulation. *CJEM*. 2007; 9: 371-373.
26. Madsen EL, Frank GR, Dong F. Liquid or solid ultrasonically tissue-mimicking materials with very low scattering. Jun Shi, Shichong Zhou, Xiao Liu, Qi Zhang, Minhua Lu, and Tianfu Wang. Stacked deep polynomial network-based representation learning for tumor classification with small ultrasound image dataset. *Neurocomputing*. 2016; 194: 87-94.
27. Zahra Sobhaninia, Shima Rafiei, Ali Emami, Nader Karimi, Kayvan Najarian, Shadrokh Samavi, and SM Reza Soroushmehr. Fetal ultrasound image segmentation for measuring biometric parameters using multi-task deep learning. In 2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC), Pages. 2019; 6545-6548. IEEE.
28. Vivienne Sze, Yu-Hsin Chen, Tien-Ju Yang, and Joel S Emer. Efficient processing of deep neural networks A tutorial and survey. *Proceedings of the IEEE*. 2017; 105(12): 2295-2329.
29. Jeya Maria Jose Valanarasu, Rajeev Yasarla, Puyang Wang, Ilker Hacihaliloglu, and Vishal M Patel. Learning to segment brain anatomy from 2d ultrasound with less data. *IEEE Journal of Selected Topics in Signal Processing*. 2020; 14(6): 1221-1234.
30. Thomas LA van den Heuvel, Dagmar de Bruijn, Chris L de Korte, And Bram van Ginneken. Automated measurement of fetal head circumference using 2d ultrasound images. *PloS one*. 2018.; 13(8): e0200412,
31. Thomas LA van den Heuvel, Hezkiel Petros, Stefano Santini, Chris L de Korte, and Bram van Ginneken. Automated fetal head detection and circumference estimation from freehand ultrasound sweeps using deep learning in resource-limited countries. *Ultrasound in medicine & biology*, 2019; 45: 773-785.
32. Tianzhe Wang, Kuan Wang, Han Cai, Ji Lin, Zhijian Liu, Hanrui Wang, Yujun Lin, and Song Han. Apq. Joint search for network architecture, Pruning and quantization policy. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Pages 2078-2087, 2019. *Ultrasound Med Biol*. 1998; 24: 535-542
33. M. Lin, Q. Chen, and S. Yan, "Network in network," *arXiv preprint arXiv*. 2013; 1312-4400.
34. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv*. 2014; 1412.6980.
35. J. T. Springenberg, A. Dosovitskiy, T. Brox, and M. Riedmiller, "Striving for simplicity: The all convolutional net," *arXiv preprint arXiv*. 2014; 1412.6806.