Retraining strategies for improved generalization of deep neural networks in ultrasound imaging

David Pasdeloup[1], Andreas Østvik[1, 2], Erik Smistad[1, 2], Lasse Lovstakken[1]; [1]CIUS, Department of Circulation and Medical Imaging, Faculty of Medicine and Health Sciences, Norwegian University of Science and Technology, Trondheim, Norway; [2]SINTEF Medical Technology, Trondheim, Norway

Background, Motivation and Objective

Deep convolutional neural networks have shown a high potential for improving image analysis and ultrasound imaging. These networks have shown excellent generalization ability when training and inference frames are acquired with the same ultrasound system. However, performance may drop when the system is different between training and inference. This work investigates strategies to adapt existing networks to data acquired by a different system than the one used to acquire the training data.

Statement of Contribution/Methods

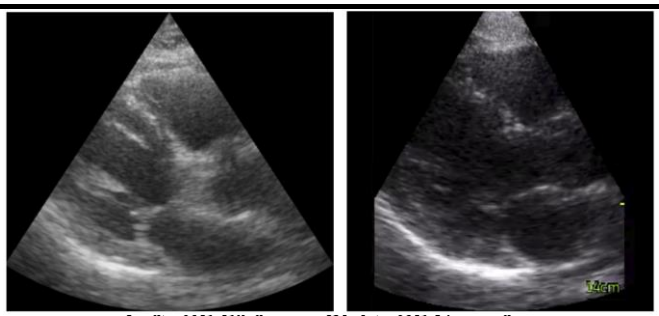
Two cardiac view classification networks, CVCNet [1] and MobileNet were investigated. While accuracy was 99.68% and 99.30% on test data from the training system (GE Vivid E9), only 77.65% and 77.07% was achieved on data from a different system (GE Vscan). Three alternatives were investigated to reduce this drop in accuracy. The first two was to fine-tune either the input or output network layers with a limited amount of training data from the new system. As a third alternative, the full network was retrained from scratch. For each of these alternatives, 4-fold validation and test was used to estimate the accuracy of the new network.

Results/Discussion

Between the three alternatives, input layer fine-tuning achieved the best accuracy, improving accuracy from 78 to 92%. Output layer fine-tuning did not bring significant improvements, and retraining was not successful due to the limited data available. This suggests that the local features learned by input layers are device dependent, while global features can be learned regardless from the data origin. Input fine-tuning thus appears as a good solution to minimize the drop in accuracy for data from different systems, where only a smaller set of annotated data is required from the unknown system.

Accuracy results

		CVCnet	MobileNet	Test method
Out of the box model	Known scanner – GE Vivid E9	**99,7%**	**99,3%**	1 inference pass over 230711 frames
Out of the box model	Unknown scanner – GE Vscan	77,7%	77,1%	1 inference pass over 8234 frames
Retrained from scratch		82,1±5,6%	64,9±5,6%	4-fold validation and test over 8234 frames (train 0.5, validation 0.25, test 0.25)
Output layers fine-tuned		79,6±8,6%	78,6±4,6%	
Input layers fine-tuned		**92,4±2,0%**	**83,8±2,9%**	



Left: GE E9 frame, Right: GE Vscan frame

[1] A. Østvik, E. Smistad, S.A. Aase, B.O. Haugen, and L. Lovstakken, « Real-time Standard View Classification in Transthoracic Echocardiography using Convolutional Neural Networks », *Ultrasound in Medicine and Biology,* 2018.