Editorial: Computational modelling of cardiovascular hemodynamics and machine learning

Christos Bourantas[1,2], Ryo Torii[3], Sergey Karabasov[4] and Rob Krams[4]*

[1]Department of Cardiology, Bart's Heart Centre, Barts Health NHS Trust, London, United Kingdom, [2]Device and Innovation Centre, William Harvey Research Institute, Queen Mary University, London, United Kingdom, [3]Department of Mechanical Engineering, University College, London, United Kingdom, [4]School for Science and Engineering, Queen Mary University, London, United Kingdom

Editorial on the Research Topic
Computational modelling of cardiovascular hemodynamics and machine learning

Artificial Intelligence (A.I.) holds promises in many fields, especially in the health sector. Here, pattern recognition of complex problems—a major strength of A.I. is what makes A.I. so useful. Despite its promises, applying A.I. to the health sector comes with specific challenges, of which the papers in the current issue aim to provide solutions (1–3). This editorial will offer specific background theory on A.I. to better understand the solutions offered in this issue.

The general requirements for high precision, generalizable A.I. models are the data set on which the model is trained/validated is homogeneous, well-labelled (for supervised learning), the data sets are of sufficient size, and the data set contains more measurements repeats than features (−10:1) (4, 5). The size of the data set is dependent on the complexity of the model. In general, the rule of thumb is to have 10 times as much data as the number of parameters in your model. As a deep neural network (NN) contains easily more than 100,000 parameters, data sets of 1M points (e.g., 1M patients) are often required, which is quite an unrealistic assumption. Hence, several engineering methods have been proposed to overcome this challenge, like data augmentation, transfer learning or data synthesis. Some of these methods have been proposed in the current issue.

The neural network predictions are best described by conditional Bayesian statistics, which depend not only on the statistical distributions (variance, mean, outliers, missing values) of the data sets (e.g., feature space), but also on the modelling complexity (e.g., weight distribution) (6–8). There is a well-known trade-off between variance and bias which is determined by the modelling complexity and size of the feature space (6–8). Increasing complexity of the model at constant input, reduces the variance and improves training predictions but tend to increase the bias thereby reducing the generalizability of the model to new data sets. A recent study shed light on how model generalizability is affected. The two main factors are the curvature of the loss function,

and the variability of the weights within each layer of the network (9). Well-known procedures to improve generalizability (tuning parameters, regularisation, drop out) affected these two parameters as expected (10, 11). It is therefore essential to not only validate the model by splitting the data set in training and validation data, but also to test the model on new, independent data. Prediction accuracies need to be high in both the validation and testing studies.

This issue of Frontiers had as subject Computational modelling, Cardiovascular Hemodynamic and Artificial Intelligence and contained 9 papers. Two (2) papers in this issue focussed on rapid flow modelling with A.I. In the first paper (Pavlo Yevtushenko et al.), a RNN (1,100 nodes) was trained on 244 patients and was validated on 23 patients using stationary flow data generated with Computational Fluid Dynamics (CFD) applied on a systolic geometry of aortic valves and the proximal aorta obtained with a combination of echocardiography and computed tomography (CT) imaging. The NN predicted the pressure and wall shear stress distribution over the centreline geometry with >90% accuracy and within seconds. The limitations of this approach are: (i) the simplicity of the ground truth (stationary flow, stiff walls, absence of moving valves), (ii) relatively small data set and (iii) the absence of testing on an independent data set. The second study (Morgan et al.) used the CFD derived ground truth but predicted the velocity field instead of pressure and shear stress. This creates a higher flexibility to predict more sophisticated biomechanical factors derived from the velocity field. The training set was created using synthetic data generation algorithm, a novel method to overcome the "curse of dimensionality" of A.I. In addition, it was optimised for speed, by using state-of-the art feature reduction methods (t-SNE) and Proper Orthogonal Decomposition of the velocity field (cPOD). This method could easily be extended to include time-dependence, elastic walls, and moving boundaries, making this method applicable to a variety of health problems.

Non-invasive technologies (Computer Tomography (CT), magnetic resonance imaging (MRI)) have been increasingly used for screening in Cardiology. These techniques are getting more sophisticated in their speed, and analysis, but have a lower spatial resolution (−200 micron) than invasive techniques (−10–50 micron). In addition, the radiation dose of CT is high, preventing the creation of time series. In two review papers, the authors (Liao et al. and Baeßler et al.) describe A.I. studies aimed at reducing imaging noise when radiation dose in CT is reduced. It is in this application, that CNN models proved to be superior to more classical methods (regression trees, support vector machine, k-NN). The same studies indicated that A.I. methods are useful for vessel segmentation, and stenosis detection. While these studies indicate a high sensitivity of DNN classifiers, the specificity of these Deep Neural Network models appeared rather low. This is probably due to the effect of noise in the feature space on classification, as indicated above. AI for calcification classification and plaque composition detection with CT radiometry shows improvement over the more classical methods but needs further development because of low accuracies. In their review papers, Liao and Baessler discussed coronary artery stenosis detection and fractional flow reserve (FFR). Their review indicated that both can be detected with high accuracy.

Two studies investigated the origins of errors in parameters derived from flow measurements in models of idealised coronary stenosis (Nguyen Ho et al.), and cardiac output derived from CT coronary aniography (Leiknes et al.). These studies indicate that the major problem is in classification of the anatomical parameters. In order to overcome this problem, Nguyen et al. applied ensemble averaging. The principle is based on repeated analysis of slightly different model parameters, thereby reducing Bayesian error propagation in neural networks. Indeed, this technique showed an improvement of classification accuracy over classical backpropagation methods and feature averaging methods.

In summary, A.I. models are increasingly applied to problems in Cardiology. These models show excellent results in detecting the complicated patterns underlying cardiovascular disease. The requirement to acquire a well-validated, generalizable predictor are demanding and are briefly discussed in the introduction of this paper. As most models presented here use rather small data sets, their accuracy and generalizability may be improved. Besides increasing the feature space (number of cases), several solutions for working with small data sets have been introduced in this journal, like ensemble averaging, and synthetic data generation.

Author contributions

RK: Conceptualization, Formal Analysis, Writing – review & editing. CB: Writing – review & editing. RT: Writing – review & editing. SK: Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

References

1. Wang M, Yang Y, Zhang W, Zhou F, Zhang X, Zhang J, et al. Risk factors for cerebrovascular events in moyamoya angiopathy using 4D flow MRI: a pilot study. *J Magn Reson Imaging*. (2023) 58(1):61–8. doi: 10.1002/jmri.28522

2. Skandha SS, Gupta SK, Saba L, Koppula VK, Johri AM, Khanna NN, et al. 3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: atheromatic 2.0. *Comput Biol Med*. (2020) 125:103958. doi: 10.1016/j.compbiomed.2020.103958

3. Belal SY, Taktak AF, Nevill AJ, Spencer SA, Roden D, Bevan S. Automatic detection of distorted plethysmogram pulses in neonates and paediatric patients using an adaptive-network-based fuzzy inference system. *Artif Intell Med*. (2002) 24:149–65. doi: S0933365701000999

4. Way TW, Sahiner B, Hadjiiski LM, Chan HP. Effect of finite sample size on feature selection and classification: a simulation study. *Med Phys*. (2010) 37:907–20. doi: 10.1118/1.3284974

5. Ding Y, Tang S, Liao SG, Jia J, Oesterreich S, Lin Y, et al. Bias correction for selecting the minimal-error classifier from many machine learning models. *Bioinformatics*. (2014) 30:3152–8. doi: 10.1093/bioinformatics/btu520

6. Ghahremanpour MM, Saar A, Tirado-Rives J, Jorgensen WL. Ensemble geometric deep learning of aqueous solubility. *J Chem Inf Model*. (2023) 63:7338–49. doi: 10.1021/acs.jcim.3c01536

7. Gumasing MJJ, Ong AKS, Sy M, Prasetyo YT, Persada SF. A machine learning ensemble approach to predicting factors affecting the intention and usage behavior towards online groceries applications in the Philippines. *Heliyon*. (2023) 9:e20644. doi: 10.1016/j.heliyon.2023.e20644

8. Wisesa P, Andolina CM, Saidi WA. Machine-learning accelerated first-principles accurate modeling of the solid-liquid phase transition in MgO under mantle conditions. *J Phys Chem Lett*. (2023) 14:8741–8. doi: 10.1021/acs.jpclett.3c02424

9. Genkin M, Engel TA. Moving beyond generalization to accurate interpretation of flexible models. *Nat Mach Intell*. (2020) 2:674–83. doi: 10.1038/s42256-020-00242-6

10. Pai TY, Wang SC, Chiang CF, Su HC, Yu LF, Sung PJ, et al. Improving neural network prediction of effluent from biological wastewater treatment plant of industrial park using fuzzy learning approach. *Bioprocess Biosyst Eng*. (2009) 32:781–90. doi: 10.1007/s00449-009-0304-2

11. ElGokhy SM, ElHefnawi M, Shoukry A. Ensemble-based classification approach for micro-RNA mining applied on diverse metagenomic sequences. *BMC Res Notes*. (2014) 7:286. doi: 10.1186/1756-0500-7-286