

REVIEW

Open Access



The machine learning horizon in cardiac hybrid imaging

Luis Eduardo Juarez-Orozco^{1,2*} , Octavio Martinez-Manzanera³, Sergey V. Nesterov^{1,4}, Sami Kajander¹ and Juhani Knuuti¹

* Correspondence:

l.e.juarez.orozco@gmail.com

¹Turku PET Centre, Turku University Hospital and University of Turku, Kiinamylynkatu 4-8, 20520 Turku, Finland

²Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, Hanzplein 1, 9700RB, Groningen, The Netherlands

Full list of author information is available at the end of the article

Abstract

Background: Machine learning (ML) represents a family of algorithms that has rapidly developed within the last years in a wide variety of knowledge areas. ML is able to elucidate and grasp complex patterns from data in order to approach prediction and classification problems. The present narrative review summarizes fundamental notions in ML as well as the evidence of its application in standard cardiac imaging and the potential for implementation in cardiac hybrid imaging.

Results: ML, and in particular Deep Learning, has begun to revolutionize medical imaging through the optimization of diagnostic and prognostic estimations at the individual-patient level. On the other hand, the spread and availability of high quality non-invasive imaging has provided growing amounts of data in the characterization of suspected cardiovascular diseases. At the same time, modern combined imaging equipment has set the ground for the concept of hybrid imaging to develop. Cardiac hybrid imaging refers to the combination of diagnostic images and offers the possibility to comprehensively characterize the heart and great vessels when a pathology is suspected or clinically known. Analysis and integration of large amounts of cardiac hybrid imaging data (and corresponding clinical profiles) constitutes a highly complex process and ML will likely be able to enhance it in the near future.

Conclusion: ML conveys novel and powerful approaches in the processing of large and complex datasets that may include images as well as imaging-derived data. Given the growing amount of data in the realm of cardiac hybrid imaging and the rapid development of ML, it is highly desirable to implement and test ML in the optimization of our multimodality imaging diagnostic and prognostic evaluations in cardiovascular disease.

Keywords: Machine learning, Cardiac hybrid imaging, Deep learning

Background

Machine learning (ML) has rapidly developed within the last years and its extension into medical sciences offers the potential to revolutionize the way in which complex diagnostic and prognostic estimations at the level of the individual patient are performed. At the same time, the spread and availability of high quality (non-invasive) imaging has provided growing amounts of data in the characterization of suspected cardiovascular diseases. Analysis and integration of such data constitutes a highly complex process for which ML seems to provide a novel suitable approach. Presently, the combination of informative data obtained through medical imaging delineates the realm

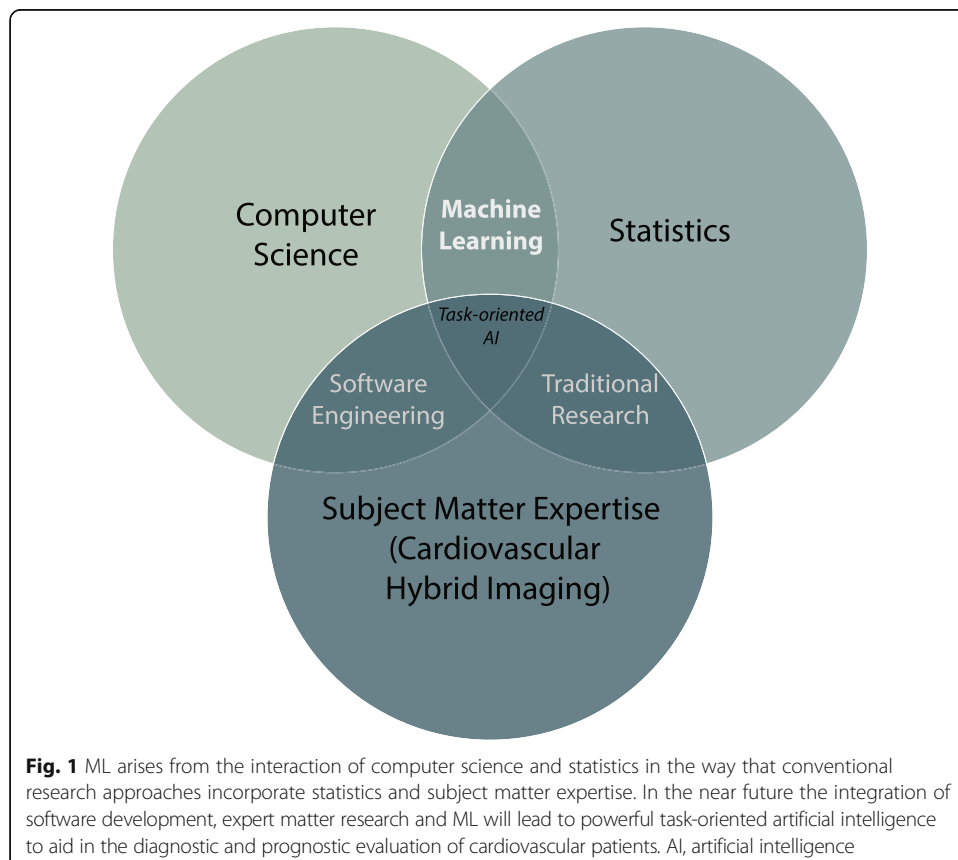
of *hybrid imaging*. Hence, it is only germane to outline the concepts, evidence and potential of ML within the context of the emerging demands of cardiac hybrid imaging.

The present narrative review will summarize the fundamental notions in ML as well as the evidence of its application in standard cardiac imaging. Thereon, the potential for implementation in order to address particular needs conveyed by cardiac *hybrid imaging* will be discussed.

The concept and rise of ML

ML constitutes a swiftly growing field located at the intersection of computer science, statistics and subject-matter expertise (in our case, cardiac hybrid imaging) (Fig. 1). ML is not a single method but rather a family of algorithms with a particular feature, i.e. the ability to learn complex patterns from data through an iterative process (training). The objective in ML is viewed as an optimization problem with aims to elucidate, re-fine and apply the learned patterns in order to predict or classify unseen data (testing). Adequate generalization to independent data is assumed when prediction or classification performance tracks the one reported for the original training and testing phase.

Many of the algorithms considered in ML are not formally new. For instance, the base computational models that paved the way for artificial neural networks date back to the 1940's. The reason why ML has just recently grown into a major topic within other areas of scientific research is that currently, we find ourselves in the cross-roads



of three necessary components for its implementation, namely: a complex problem, large amounts of data and sufficient computational power.

Biological problems, such as cardiovascular diseases, are adequate examples of complex problems. The patent and suspected interrelations at every level of human (patho) physiological processes are considerably intricate. This notion is contained in the concept of *precision medicine*. Ultimately, precision medicine presumes that the individual patient is a conjunction of all the particular patterns that distinguish the subject and his/her manifestation of a disease from the rest of the people, and that this profile can and should be harvested in order to individualize treatments and risk estimations. Such patterns can be found in all interactions in the gene-protein-anatomy-function continuum.

Currently, larger amounts of data are available to researchers as we have witnessed exponential gains in storage capacity in the last decades. But subtle and complex patterns within data can only be deduced in as much as large sets of representative data can be utilized. In this sense, organized efforts to overcome accessibility and standardization problems are being undertaken.

Finally, the third component is sufficient computational power. The amount of calculations performed when comprehensively evaluating all possible interrelations within datasets and the amount of available data mutually enhance the computational time required for such operations. The advent of new graphics-processing units has allowed us to boost processing speed, making it possible to operate within the limits of practicality.

Types of learning

One can classify ML approaches in terms of the type of learning process they utilize. Overall, there are four types of learning: supervised, unsupervised, semi-supervised and reinforcement.

Supervised learning works through the utilization of labeled data (i.e. data which has been effectively classified or for which the outcome value is already known). This labeling may have been done by an expert physician (e.g. a radiologist that diagnosed the presence of a cardiac mass, a nuclear physician who annotates a perfusion defect in the anterior wall of the left ventricle or a cardiologist who identifies cardiac involvement in a patient with sarcoidosis) or by a data manager who has confirmed vital status and/or adverse outcomes in the electronic health record of a given subject. Learning then takes place by allowing the algorithm to make a classification or prediction, then comparing it to the known label in order to *support* or *penalize* the result and adjusting the model accordingly in order to refine its performance. While supervised learning has obtained remarkable results (i.e. studies in different medical fields achieve performances comparable to the ones of expert clinicians [diabetic retinopathy (Gulshan et al., 2016) and skin cancer (Esteva et al., 2017) labelling the data can be remarkably time consuming (e.g. segmentation).

Unsupervised learning, conversely, does not utilize available labels or known outcomes in order to elucidate patterns within the data. It can be useful to identify novel clusters among, for example, patients originally considered to present with the same disease without the bias of our established disease phenotypes. As novel as such new clusters may be, it is yet up to the human observers to decide whether such

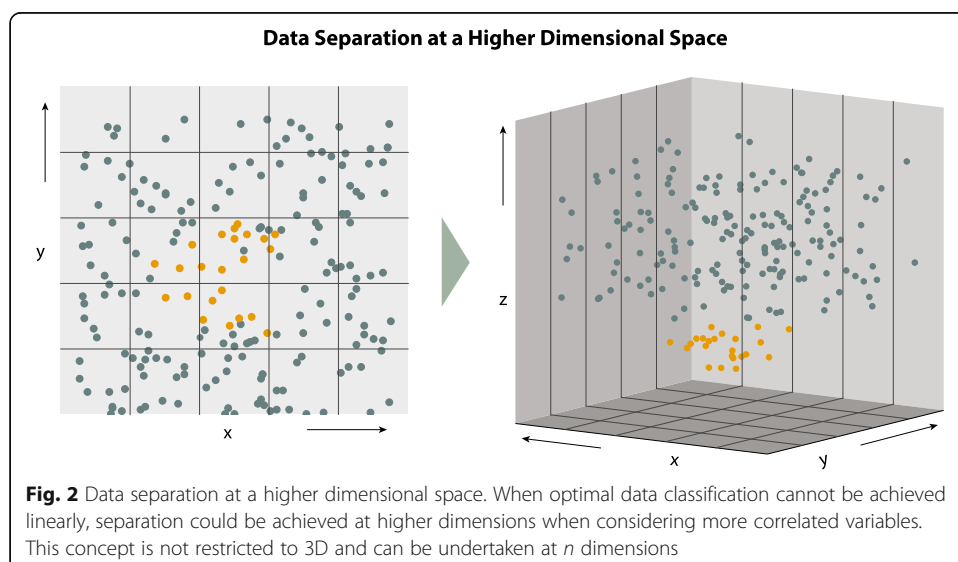
differentiations are of value in order to improve treatments and outcomes. In other words, to make sure that what we obtain are not distinctions without differences.

Semi-supervised learning falls somewhere in between the two aforementioned types of learning. It utilizes labeled and unlabeled data (mostly a smaller quantity of the former) in order to optimize the efficiency of classification. Semi-supervised learning is mainly useful when the human-power necessary to provide full labeling of all working data exceeds practical possibilities. However, it also conveys the advantage of keeping some distance from intrinsic biases provided by human operators, which probably exist even when such operators are considered experts.

Finally, reinforcement learning is considered when there is a conjunction of other types of learning with the underpinning of constant trial and error. Reinforcement learning progresses without an initial notion of the objective and relies on constant interaction with new conditions and input in order to maximize the reward or benefit (performance), even if not immediately. Most notable examples of reinforcement learning can be found outside the medical field such as self-driving cars and advance systems for game implementations (e.g. AlphaGo Zero (Silver et al., 2017)).

ML algorithms and the special case of deep learning

The range of supervised ML algorithms is wide, their foundations, weaknesses and strong points are different but their objectives are constant. Every approach aims to classify (separate) or predict data as effectively as possible. K-nearest neighbors (k-NN), for instance, is a simple algorithm that classifies data points according to the most common class represented in the “k” (i.e. an adjustable parameter) nearest neighboring data points within the data space. Support vector machine (SVM) is another useful algorithm which aims to construct classification models by transforming a subset of predictors in a multidimensional space where it then can find a non-linear boundary with the widest (and therefore the best) separation between classes (Smola & Schölkopf, 2004). A simple graphic depiction of how data separation can, in many instances, be achieved at higher dimensional spaces is shown in Fig. 2. Decision trees (or regression trees, depending on the nature of the objective) represent another approach in ML in which individual



predictors are aggregated or used sequentially to separate a subset of data in order to determine its class value (outcome label). When several decision trees are built and aggregated the resulting method is called random forest (Random Forests, 2001), and when they are used sequentially as ensembles they are called extreme gradient boosting. Yet another example is that of the naïve Bayes algorithm. This approach aims to classify data instances by a simple probabilistic approach that uses the pre-probability (prevalence) and the likelihood of belonging to a specific class (Miranda et al., 2016).

Overall, every method conveys both advantages and disadvantages, and knowledge of the algorithms is paramount in order to select the better suited, based on the characteristics of a specific problem and dataset. A quick view of common ML algorithms along with their main advantages is shown in Table 1 (the list is not meant to be exhaustive).

Then there is the case of artificial neural networks (ANNs). These algorithms were originally inspired in the architecture and function of biological neurons. ANNs are formed by layers of processing units (neurons) through which the learning process is conducted in an iterative two-step sequence. First, for each subsequent layers, information is carried forward to each neuron of the later layer by applying weights to the values of the neurons of earlier layers. The resulting values are then added and an activation function is applied. Once the processing delivers a result in the output layer, supervised learning can be used to compare the classification results to that of the known labels. In the second step, the errors in classification are carried backwards in the network in order to adjust the weights applied in the connections between processing units. In this way, the estimations are iteratively improved and the accuracy of the classification is optimized. The basic architecture of an ANN is depicted in Fig. 3.

The reason why only until recently the interest in ANNs has reshaped ML is two-fold: first, the high computational power currently at our disposal, and second, the superior performance of recent network architectures (e.g. Inception (Szegedy et al., 2015), Resnet (He et al., 2015), U-Net) and refinements in the operations performed in the processing units (batch normalization (Ioffe & Szegedy, 2015) and rectified (Krizhevsky et al., 2012) or exponential non-linear activations (Clevert et al., 2015)). These optimizations have allowed researchers to unleash the power of deep convolutional neural networks (CNNs) in tasks that use data with very high dimensionality such as images. This collection of applications is what is now commonly known as *Deep Learning* and it has started to demonstrate contributions in medical imaging.

In general, ML modelling is performed by data parcellation into a training and a testing dataset. Traditionally, a 60:40 or 70:30 split have been used, but a number of concerns based on possible unbalance of positive and negative cases have promoted the incorporation of resampling methods such as cross-validation in order to be able to utilize all available data for both purposes. This delivers robust performance estimations that provide a confident overview of the capacities of the ML models generated. A depiction of the steps involved in data handling in ML modelling is shown in Fig. 4.

ML applications in cardiac imaging tasks

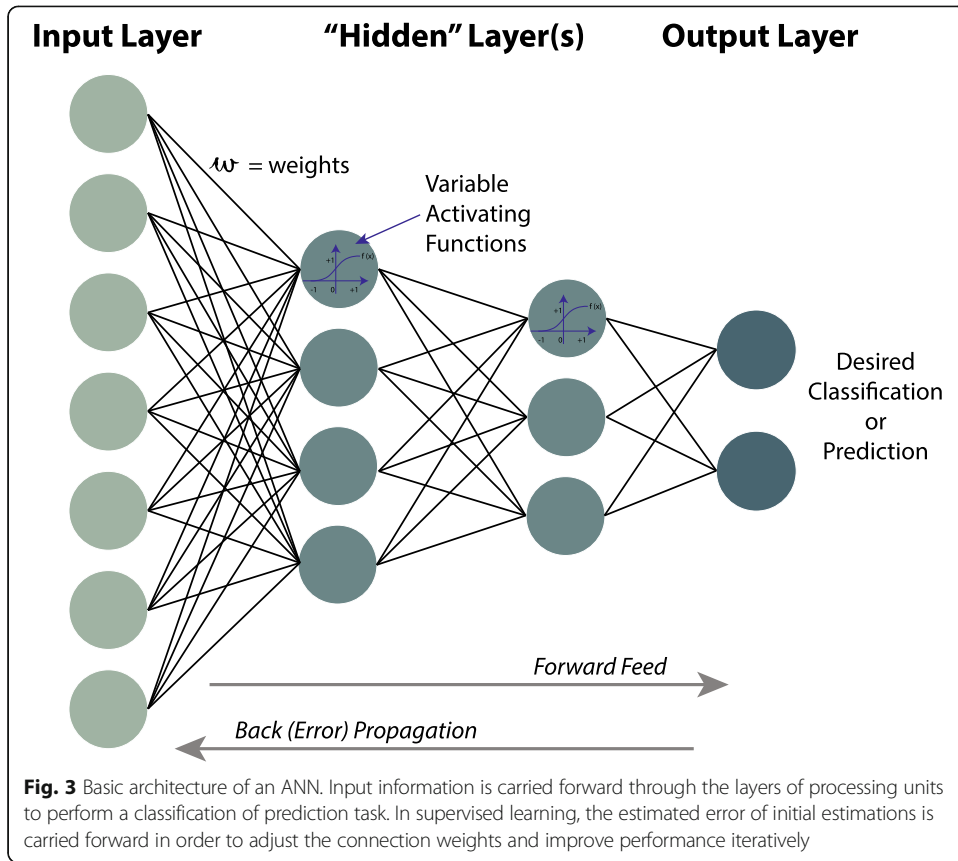
Feature selection and feature engineering

A successful ML workflow can be generated by considering two sequential processes, namely, feature selection and modelling. In feature selection, the variables that contribute

Table 1 Overview of advantages and disadvantages of commonly utilized machine learning algorithm clustered by utilized type of learning

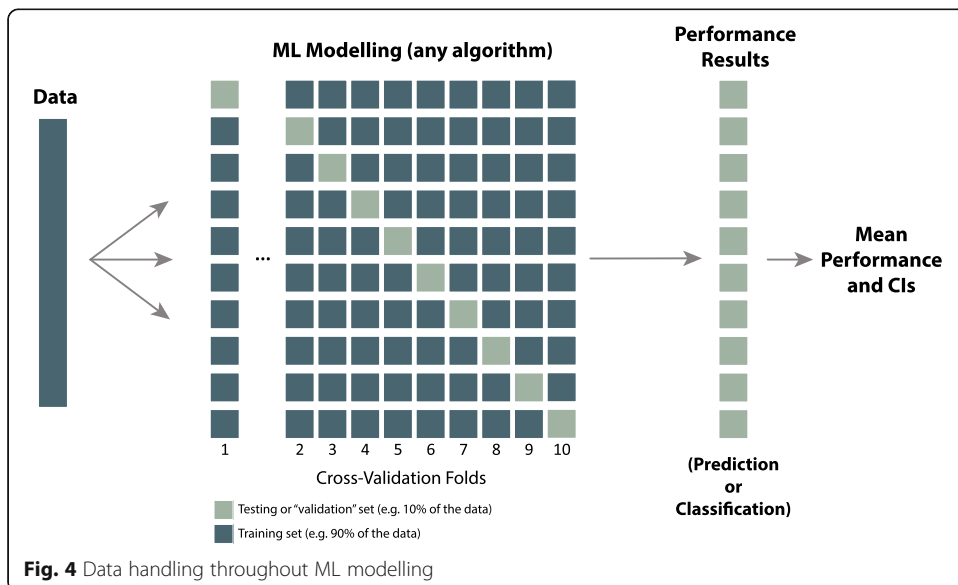
Learning Type	Purpose	Algorithm	Advantages	Disadvantages
Unsupervised	Clustering	K-means	<ul style="list-style-type: none"> Intuitive algorithm with linear relative computational complexity 	<ul style="list-style-type: none"> Number of clusters or classes must be defined by user Poor performance with clusters of irregular shapes
		Mean Shift	<ul style="list-style-type: none"> No need to select number of clusters or classes 	<ul style="list-style-type: none"> Utilized window size must be defined by user
		DBSCAN	<ul style="list-style-type: none"> No need to select number of clusters or classes It can identify points as noise and cluster with irregular shapes 	<ul style="list-style-type: none"> Windows size must be defined by user Reduced performance when clusters have different densities
		Expectation-Maximization	<ul style="list-style-type: none"> It can identify clusters with ellipsoidal shapes It assigns membership probabilities to each point 	<ul style="list-style-type: none"> Number of clusters or classes must be defined by user
Supervised	Classification	Logistic Regression	<ul style="list-style-type: none"> Good performance with small datasets Its output can be interpreted as a probability 	<ul style="list-style-type: none"> Data assumptions are needed to be complied It can only provide linear solutions
		K-Nearest Neighbors	<ul style="list-style-type: none"> Intuitive algorithm 	<ul style="list-style-type: none"> Number of neighbors must be defined by user High relative computational complexity
		Naive Bayes	<ul style="list-style-type: none"> Performs well in small datasets if conditional independent assumption holds 	<ul style="list-style-type: none"> Assumption of independence between features
		Support Vector Machines	<ul style="list-style-type: none"> It can provide non-linear solutions 	<ul style="list-style-type: none"> To achieve good performance, they require knowledge about the kernel employed
		Decision trees ensembles	<ul style="list-style-type: none"> They can handle categorical features Few parameters to tune They perform well in datasets with large number of features 	<ul style="list-style-type: none"> Interpretability of ensemble can be questioned
		Neural Networks	<ul style="list-style-type: none"> State-of-the-art results Direct complex image processing 	<ul style="list-style-type: none"> Many parameters to fine-tune Large number of samples are required to achieve good performance
	Regression	Linear (LASSO, Ridge)	<ul style="list-style-type: none"> Good performance with small datasets 	<ul style="list-style-type: none"> Data assumptions are needed to be complied Can only provide linear solutions
		Decision trees	<ul style="list-style-type: none"> They can provide non-linear solutions 	<ul style="list-style-type: none"> Interpretability of ensemble can be questioned
	Segmentation	U-Net	<ul style="list-style-type: none"> One stage algorithm with good performance and variants 	
		Mask-RCNN	<ul style="list-style-type: none"> State-of-the-art performance 	<ul style="list-style-type: none"> Two-stage algorithm

the most with the task at hand according to a specific model are selected. Although this may somehow resemble the way in which predictors are selected based on their univariate correlations with the outcome variable in traditional linear statistics, ML feature selection commonly explores the relevance of the predictors and their interactions in non-linear ways and at a higher dimensional level. This means that relevant variables can potentially



be selected more efficiently and without compromising possibly relevant interactions that are relevant in distinguishing groups of subjects or predicting continuous outcomes.

Deep learning in this case, is also noteworthy. Unlike other models that require features to be created in a preliminary process, CNNs have the ability to generate optimal features based on the raw data (i.e. images). Therefore, they not only inherently



employ the most relevant features, but they also add a component that is not present on other models.

Automation of detection and segmentation

ML has been used in the automation of detection and delineation tasks. For instance, SVM was used by Kang and colleagues in order to automatically identify atherosclerotic lesions on coronary computed tomography angiography (CCTA) images with an accuracy well above 90% (Kang et al., 2015). An engaging early report by Išgum and colleagues described automatic calcium scoring (CaSc) in low-dose CT scans using and merging the results of two ML algorithms (k-NN and SVM) (Išgum et al., 2012). Later, deep learning has been utilized with a similar intent demonstrating moderate to good results (Wolterink et al., 2016; Lessmann et al., 2017).

In echocardiographic imaging, cardiac magnetic resonance (CMR) and computed tomography, segmentation of (dynamic) views of the left ventricle, boundary detection and measurement derivation constitute interesting targets for ML. Knackstedt et al. documented the utilization and superiority of an original ML-based software in the automatic measurement of left ventricular ejection fraction and longitudinal strain (Knackstedt et al., 2015). Ngo and colleagues merged deep learning and level set for ventricular segmentation in cine CMR images in a discrete sample advancing the steps towards full task automation (Ngo et al., 2017).

Cardiovascular diagnosis

ML has been utilized to try to optimize detection of coronary artery disease (CAD). Mohammadpour and colleagues, for instance, defined important risk factors for CAD and then identified patients who would not need an invasive evaluation of CAD with high accuracy (93%) (Mohammadpour et al., 2015). Further, Xiong et al. used an ensembles approach to assess myocardial perfusion CCTA images for the detection of significant CAD (Xiong et al., 2015). Similar work was performed more recently by Zreik et al. They reported the use of a sequential approach that included CNN-based segmentation, convolutional *autoencoding* and SVM classification of suspected myocardial ischemia in rest CCTA (Zreik et al., 2018). Another report also studied the utilization of ML in the form of boosted ensembles in order to predict early CAD revascularization indications (Arsanjani et al., 2015).

Finally, Dey and colleagues have reported on the utility of ML and boosted ensembles for the detection of plaque-specific ischemia through CCTA (AUC = 0.84) using invasive FFR as reference (Dey et al., 2018). Their report concentrated on the integration of plaque features and their integration into a strong classifying score, an advantage offered by iterative ML algorithms.

Cardiovascular prognosis evaluation

Another objective that has been addressed with the implementation of ML algorithms is the improvement in the prognostic value of cardiac imaging. In this sense, Berchiolla and colleagues were among the firsts to describe the use of ML in the form of a Bayesian Network to evaluate predictors from stress echocardiography and CCTA in the prediction of myocardial infarction or death. Moreover, they described the incremental performance of ML against traditional logistic regression (Berchiolla et al., 2012). Thereon, Motwani and colleagues utilized CCTA measurements characterizing

atherosclerotic plaques, demonstrating an improvement in the 5-year prediction of mortality (Motwani et al., 2016). Lately, the same group was able to demonstrate an improvement in the 3-year predictive capacity for the occurrence of major adverse cardiovascular events (MACE) when implementing a ML workflow that integrated single photon emission computed tomography (SPECT) perfusion data and electronic health record clinical variables (Betancur et al., 2017). Given that quantitative positron emission tomography (PET) imaging has also demonstrated robust prognostic value (Juárez-Orozco et al., 2017), ML implementations in PET prognostic data for the classification of patients who will experience particular MACE are warranted.

Cardiac disease profiles exploration

A remarkable implementation has been the use of unsupervised ML for the identification of cardiac disease profiles. In particular, ML has been used in echocardiography in order to characterize healthy and diseased hypertrophic states (Narula et al., 2016), while ML-based *phenomapping* has now been used to evaluate heart failure with preserved ejection fraction for new and clinically relevant disease phenotypes (Shah et al., 2015).

The nature of hybrid imaging

Available non-invasive techniques can be robustly divided by the kind of information they convey and reports of their diagnostic and prognostic performance have generally taken place independently, as *stand-alone* modalities. There are techniques such as CCTA that are mainly deemed as anatomical (although contemporary developments are extending the application of CCTA beyond anatomy and into ventricular function, myocardial perfusion and even FFR evaluation [FFR_{CT}]), while others can evaluate the pattern and distribution of specific physiological processes, such as perfusion, viability, innervation and angiogenesis, within the myocardium (i.e. SPECT and positron emission tomography [PET]). Furthermore, echocardiography can easily provide prompt assessment of ventricular motion and dynamic insights of the heart's function. Finally, CMR can offer both anatomical and functional information, such as motion evaluation of the ventricles, myocardial mass and tissue content or composition based on numerous developed acquisition sequences.

The clinical benefit (beyond the research value) of the data provided by the aforementioned range of approaches follows especially when such information is complementary in nature (e.g. anatomical and functional), rather than overlapping (e.g. anatomical and anatomical). Consequently, the concept of *hybrid imaging* has rapidly gained popularity.

A multimodal approach, ideally boosted by ML-based optimizations, can deliver insight into different aspects of the status of the cardiovascular system could improve identification and characterization of CAD, heart failure, etc. The expected benefit, therefore, boils down to the additive value of the strengths from stand-alone techniques, and at the same time, the overcoming of their independent pitfalls. However, it is possible that the speed of adoption of the concept has obfuscated the general principle that researchers and clinicians assign to hybrid imaging. Already by 2009 the European Society of Cardiology put forward efforts to define and substantiate the

concept of *hybrid imaging*. It was suggested that the term should be applied when “images are fused combining two data sets, whereby both modalities are equally important in contributing to image information...” (Knuuti & Kaufmann, 2009). However, there is a variety of situations where the concept may demonstrate some fluidity.

Hybrid imaging may refer to simultaneous acquisitions in integrated protocols using equipment that feature two imaging techniques. This is the case for modern PET/CT, SPECT/CT and PET/MR studies. “Simultaneous” is understood in these cases as “within the same imaging session”. Furthermore, the large array of radiotracers utilized in the nuclear component of these hybrid equipment opens the possibility to image more than one physiological aspect of the coronary-myocardium continuum. Although *dual tracer* imaging within the same imaging session has not been established in the cardiovascular area, performing separate PET scans using a perfusion tracer (^{82}Rb , ^{13}N -ammonia or ^{15}O - H_2O) plus CCTA and a glucose-metabolism tracer (^{18}F -FDG) for myocardial viability demonstrated utility to a certain extent. Interestingly, although current cardiac PET/CT scanning makes obligatory use of the CT information for attenuation correction, the *hybrid imaging* label is not applied whenever “full” CCTA is not performed even though such correction is may be seen as inherently important for providing adequate images.

One can also consider hybrid imaging the simultaneous interpretation of individually obtained scans. In this case, “simultaneous” conveys a parallel or side-by-side visualization of images. Moreover, it can also imply the purposeful fusion of images in order to track a suspected anatomical origin (e.g. atherosclerotic plaque) of a functionally patent abnormality (regional myocardial ischemia). However, although there are dedicated software extensions for such purpose, such *fusion* is ultimately a process that clinicians perform in the presence of complementary imaging information independently from the refinements of the software they work with. Notably, there are commonly instances when clinicians ponder complementary information from independent tests (e.g. echocardiography and CCTA). This is not per se considered as *hybrid imaging* although the modalities may provide equally important data for diagnostic and prognostic purposes.

Finally, another possibility is when sequential and selective application of complementary techniques is performed as an established workflow in patients within certain limits of pre-test probability of disease. In this situation additive (hybrid) information is obtained with a confirmatory objective when the initial results of an individual technique are positive for the suspected disease.

A niche for ML in cardiac hybrid imaging

The potential for ML implementation in cardiac *hybrid* imaging follows from the notion that the combination of data provided by complementary imaging methods can enrich and improve our confidence on the presence, origin and clinical significance of a suspected cardiovascular condition. Given that ML algorithms have increasingly been applied in the described range of tasks pertaining to non-invasive imaging modalities, as well as in the integration of image-derived data and clinical data, it is likely that improved results will translate an additive utility in accelerating processes and refining the characterization of patients who undergo cardiac hybrid imaging.

With the spread of powerful non-invasive cardiovascular diagnostic approaches, clinicians are now provided with an increasing amount of information regarding the (patho) physiological status of the heart and the great vessels of a given patient. ML is likely suitable in combining data provided by individual techniques and in addressing the complexity of such data, which ranges from standard numerical variables to entire 2-, 3- and even 4-dimensional images in standardized or unstandardized formats. Furthermore, it is understood that imaging data should be integrated with the outlook offered by electronic health records, which contain demographic, clinical, and biomarker measurements. This coupling is highly desired as it may allow us to further understand the nature and development of cardiovascular disease, but it has been deemed as challenging. It is within this scenario that the recent explosion of ML is expected to impact the way in which we analyze data provided by a variety of sources and use meaningful information to optimize the diagnostic process and prognostic assessment of individual patients. We believe this may further expand the conceptual framework understood as *hybrid imaging* by itself.

It should be noted that ML in cardiovascular clinical research has predominantly been utilized for the analysis of numerical data obtained from stand-alone imaging techniques. In this scenario, numerical predictors translate particular characteristics of a pathophysiological process but it is being demonstrated that larger benefits may be obtained from the analysis of hybrid images directly irrespectively from the particular definition applied.

Deep learning has revolutionized the way in which complex sets of data can be analyzed and classified. Wide and deep neural networks can define image characteristics that optimize the classification of complex images. The seminal paper by Esteva et al. has already demonstrated how a complex classification of benign and malignant skin lesions can be reliably performed by deep learning (Esteva et al., 2017). *Stand-alone* cardiac imaging has already seen the benefit of such analyses and cardiovascular hybrid imaging will likely see implementations of such principles when the challenges of image complexity, integration and amount of available data are overcome.

On the brink

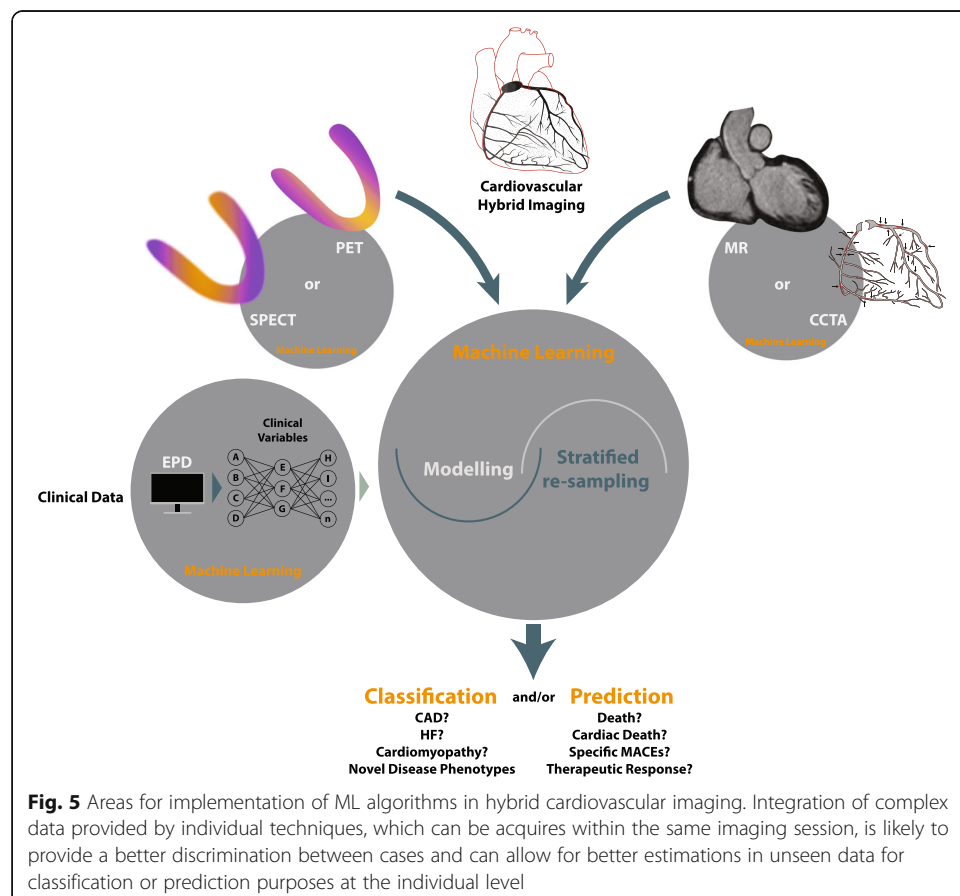
The potential of ML in the setting of hybrid imaging is considerable. Hybrid imaging conveys the availability of a full set of data (in several forms). Although no major report has addressed the utilization of ML on hybrid imaging data per se, we believe that the path is becoming clearer.

Future analyses will possibly aim to integrate ML-based methods in order to manage data arising at different levels of the diagnostic workup (i.e. different diagnostic techniques). Particular approaches will yield better results according to the operationalized variables. For example, pre-processed and even raw images will surely benefit from very deep CNNs, while numerical data on clinical variables may possibly benefit from boosted ensembles of simpler and faster algorithms that deliver robust estimates. An optimistic example could be the integration of CCTA anatomical data with (PET) myocardial perfusion and plaque metabolism data, which can then be enriched by comprehensive clinical and hemodynamic data in order to better comprehend the blueprint of unstable or risky atherosclerotic disease.

The capacity of ML to apply results from individual algorithms as inputs into further learning processes in order to accomplish the task at hand whatever we chose it to be, e.g. refining diagnosis with new disease phenotypes or optimizing risk stratification evaluations, must be well understood. This is possible due to the notion that ML can be implemented as a workflow with distinguishable sections, which may deal with independent sets of cardiovascular data. An outline of the components of an integrated ML workflow is presented in Fig. 5.

The black-box

There is some concern on the interpretability of the “intermediate steps” undertaken in deep learning. These take place in the so-called *hidden layers* of the CNNs, which lay between the input and the output. Efforts have been conducted to characterize the nature of such intermediate interpretations. It has been shown how, in image recognition, some initial layers start aiming to detect lines, borders, angles and shadows within the image input. This information then progresses to higher levels of integration that become increasingly abstract and therefore, difficult to interpret in any meaningful way. This *black-box* phenomenon is of relevance and opinions on how much importance it should be given vary. On the one side, it is true that blind trust in an algorithm could generate problems when the system and the generating data might suffer from known or unknown biases. But on the other hand, it is also true that the type of



appraisal that deep learning algorithms perform is constitutionally different from that applied by humans.

There will probably have to be some compromise in this sense. We will have to remain vigilant of the possible input of biases (even if human), while allowing for the systems to deploy their whole potential for performing better classifications or predictions. This can be promoted by the creation of the aforementioned step-wise workflows that implement ML in independent types of data that can be later integrated with confidence on the validity of the sequential results.

Emerging issues

With the continuous expansion of ML-applications, it will become primordial for clinicians and medical professionals to adapt their data procedures and infrastructure, as perhaps the biggest current impediment for machine learning implementation to clinical problems is the amount of available and organized data. Better communication and collaboration between clinical systems, an activity pursued by many researchers and health care administration professionals, will accelerate this process.

Notably, a previous obstacle was the infrastructure required to store and analyze very large datasets. A solution taken by many entities that can be emulated in medical sciences is to outsource the task managing data infrastructure to an expert separate entity provided with adequate security provisions. Using platforms of specialized data centers can lead to a simpler, more reliable, secure and scalable processes while the front-line clinical departments maintain control over the data and emerging analyses.

A perspective of the future

As intensive research is being performed in order to improve algorithm selection and fine-tuning, we believe that support (and even independent) systems will eventually be able to provide on-the-flight evaluation of hybrid imaging-derived information together with the clinical profile that can inform the decisions of both clinicians and scientists in their encounters with patients at the individual level in terms of therapeutic actions and risk appraisal.

ML is bound to keep revolutionizing the ways we understand cardiovascular disease through imaging due to its dynamic nature and possibility to perpetuate the learning process as more data will be inevitably available.

Conclusions

ML conveys novel and powerful approaches in the processing of large and complex datasets that may include images as well as imaging-derived data along with comprehensive and potentially relevant clinical data. As such, ML has begun to revolutionize medical imaging and currently available reports in cardiac imaging have focused on automatization and classification tasks in stand-alone techniques.

Given the growing amount of data in the realm of cardiac hybrid imaging and the rapid development of ML, it is highly desirable to implement and test ML in the optimization of our multimodality imaging diagnostic and prognostic evaluations in cardiovascular disease.

Abbreviations

ANN: Artificial neural networks; CAD: Coronary artery disease; CaSC: Calcium score; CCTA: Coronary computed tomography angiography; CMR: Cardiac magnetic resonance; CNN: Convolutional neural networks; MACE: Major adverse cardiovascular events; ML: Machine learning; PET: Positron emission tomography; SPECT: Single photon emission computed tomography; SVM: Support vector machine

Authors' contributions

LJ – Conception, design of the review, drafting the manuscript and critical revision of intellectual content. Also, giving final approval of the version to be published and agreeing to be accountable for all aspects of the work. OM – Conception, design of the review, drafting the manuscript and critical revision of intellectual content. Also, giving final approval of the version to be published and agreeing to be accountable for all aspects of the work. SN – Conception, design of the review, drafting the manuscript and critical revision of intellectual content. Also, giving final approval of the version to be published and agreeing to be accountable for all aspects of the work. SK – Conception, design of the review, drafting the manuscript and critical revision of intellectual content. Also, giving final approval of the version to be published and agreeing to be accountable for all aspects of the work. JK – Conception, design of the review, drafting the manuscript and critical revision of intellectual content. Also, giving final approval of the version to be published and agreeing to be accountable for all aspects of the work.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹Turku PET Centre, Turku University Hospital and University of Turku, Kiinamyllynkatu 4-8, 20520 Turku, Finland.

²Department of Nuclear Medicine and Molecular Imaging, University Medical Center Groningen, Hanzeplein 1, 9700RB, Groningen, The Netherlands. ³Department of Neurology, University Medical Center Groningen, University of Groningen, Hanzeplein 1, 9700RB, Groningen, The Netherlands. ⁴IM Sechenov Institute of Evolutionary Physiology and Biochemistry, RAS, Prospekt Toreza 44, 194223 St. Petersburg, Russia.

Received: 8 March 2018 Accepted: 9 May 2018

Published online: 23 July 2018

References

- Arsanjani R, Dey D, Khachatryan T, Shalev A, Hayes SW, Fish M et al (2015) Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. *J Nucl Cardiol* 22(5):877–884
- Berchiolla P, Foltran F, Bigi R, Gregori D (2012 Jun) Integrating stress-related ventricular functional and angiographic data in preventive cardiology: a unified approach implementing a Bayesian network. *J Eval Clin Pract* 18(3):637–643
- Betancur J, Otaki Y, Motwani M, Fish MB, Lemley M, Dey D et al (2017) Prognostic value of combined clinical and myocardial perfusion imaging data using machine learning. *JACC Cardiovasc Imaging*:1–10
- Clevert D-A, Unterthiner T, Hochreiter S (2015) Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs), pp 1–14 Available from: <http://arxiv.org/abs/1511.07289>
- Dey D, Gaur S, Ovrehus KA, Slomka PJ, Betancur J, Goeller M et al (2018) Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study. *Eur Radiol*:19
- Esteva A, Kuprel B, Novoa RA, Ko J, Swetter SM, Blau HM et al (2017 Jan 2) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542(7639):115–118
- Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316(22):2402–2410
- He K, Zhang X, Ren S, Sun J (2015) Deep Residual Learning for Image Recognition. *Multimed Tools Appl [Internet]*:1–17 Available from: <http://arxiv.org/abs/1512.03385>
- Ioffe S, Szegedy C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 2015; Available from: <http://arxiv.org/abs/1502.03167>
- Isgum I, Prokop M, Niemeijer M, Viergever MA, van Ginneken B (2012) Automatic coronary calcium scoring in low-dose chest computed tomography. *IEEE Trans Med Imaging* 31(12):2322–2334
- Juárez-Orozco LE, Tio RA, Alexanderson E, Dweck M, Vliegenthart R, El Moumni M, et al. Quantitative myocardial perfusion evaluation with positron emission tomography and the risk of cardiovascular events in patients with coronary artery disease: a systematic review of prognostic studies. *Eur Heart J Cardiovasc Imaging*. 2017;(January):1–9
- Kang D, Dey D, Slomka PJ, Arsanjani R, Nakazato R, Ko H et al (2015) Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography. *J med imaging (Bellingham, Wash)* 2(1):14003
- Knackstedt C, Bekkers SCAM, Schummers G, Schreckenber M, Muraru D, Badano LP et al (2015) Fully automated versus standard tracking of left ventricular ejection fraction and longitudinal strain. The FAST-EFs Multicenter Study *J Am Coll Cardiol* 66(13):1456–1466
- Knuuti J, Kaufmann PA (2009) The ESC textbook of cardiovascular imaging. In: Zamorano JL, Bax JJ, Rademakers FE, Knuuti J (eds) *The ESC textbook of cardiovascular imaging*. Springer London, London, pp 89–99

- Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*:1097–1105
- Lessmann N, van Ginneken B, Zreik M, de Jong PA, de Vos BD, Viergever MA et al (2017) Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions, pp 1–11 Available from: <http://arxiv.org/abs/1711.00349>
- Miranda E, Irwansyah E, Amelga AY, Maribondang MM, Salim M (2016 Jul) Detection of cardiovascular disease Risk's level for adults using naive Bayes classifier. *Healthc Inform Res* 22(3):196–205
- Mohammadpour RA, Abedi SM, Bagheri S, Ghaemian A (2015) Fuzzy rule-based classification system for assessing coronary artery disease. *Comput Math Methods Med* 2015:564867
- Motwani M, Dey D, Berman DSD, Germano G, Achenbach S, Al-Mallah MMH et al (2016) Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J* 52(4):468–476
- Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP (2016) Machine-learning algorithms to Automate morphological and functional assessments in 2D echocardiography. *J Am Coll Cardiol* 68(21):2287–2295
- Ngo TA, Lu Z, Carneiro G (2017) Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal* 35:159–171
- Random Forests BL (2001) *Mach Learn* 45(1):5–32
- Shah SJ, Katz DH, Selvaraj S, Burke MA, Yancy CW, Gheorghiane M et al (2015) Phenomapping for novel classification of heart failure with preserved ejection fraction. *Circulation* 131(3):269–279
- Silver D, Schrittwieser J, Simonyan K, Antonoglou I, Huang A, Guez A et al (2017) Mastering the game of go without human knowledge. *Nature* 550(7676):354–359
- Smola AJ, Schölkopf B (2004) A tutorial on support vector regression. *Stat Comput* 14(3):199–222
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D et al (2015) Going deeper with convolutions. In: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) IEEE, pp 1–9
- Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Išgum I (2016) Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Med Image Anal* 34:123–136
- Xiong G, Kola D, Heo R, Elmore K, Cho I, Min JK (2015) Myocardial perfusion analysis in cardiac computed tomography angiographic images at rest. *Med Image Anal* 24(1):77–89
- Zreik M, Lessmann N, van Hamersvelt RW, Wolterink JM, Voskuil M, Viergever MA et al (2018) Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal* 44:72–85

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
