

Machine learning and deep learning applied in ultrasound

„Machine“ und „Deep Learning“ im Ultraschall angewandt



Lea Marie Pehrson



Carsten Lauridsen



Michael Bachmann Nielsen

Correspondence

Michael Bachmann Nielsen
 Department of Radiology, Section of Ultrasound,
 Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen,
 Denmark
mbn@dadlnet.dk

Bibliography

DOI <https://doi.org/10.1055/a-0642-9545>
 Published online: 2018
 Ultraschall in Med 2018; 39: 379–381
 © Georg Thieme Verlag KG, Stuttgart · New York
 ISSN 0172-4614

We live in an exciting time, which many call the Information Age, signifying the fact that the amount of information in the world is growing exponentially. The use of medical imaging is vastly increasing [1]. From 1996 until 2010 the number of ultrasound exams has approximately doubled, while the number of CT and MRI examinations has almost tripled and quadrupled, respectively [2]. The large amount of data and the need for the early detection of pathologies are very demanding [3]. Due to the increasing workload, data handling needs to be faster and more precise. This editorial will showcase some of the basic principles of machine learning and deep learning within ultrasound imaging.

Machine learning is a method established from artificial intelligence, where the computer captures patterns in data sets and uses these patterns extensively in decision making. Machine learning offers different abilities with regards to medical imaging [4, 5]. The main purpose of the different types of algorithms is to elevate the diagnostic accuracy and the consistency of image interpretation. These observations and predictions are constructed on the basis of the data presented to the algorithm.

Deep learning has attracted a lot of attention over the past two years. It emerged from machine learning and automatically learns hierarchical features. The algorithm consists of multiple layers composed of simple and nonlinear modules. The data is transformed into representations which is important for the algorithm to discriminate the data [6]. Deep learning algorithms can learn from former mistakes, whereas traditional machine learning algorithms are not able to. Machine learning algorithms are constructed using hand-engineered features. These features are not

able to be adjusted after configuration of the machine learning algorithm.

As the name of the algorithms suggest, machine and deep learning algorithms can learn to detect. This is usually done by either supervised or unsupervised learning. Supervised learning consists of labeled data presented to the algorithm. The images are presented with an appropriate classification outcome. The classifier can be trained to output a value of 1 for input images with malignancy occurring on the image. For images with no pathology or benign lesions, the algorithm is taught to give an output with the value 0 [7]. Once the training of the algorithm has been completed, the algorithm can begin its classification of unseen images.

The algorithm must be trained correctly to allow classification. From a supervised learning point of view, this requires large amounts of well annotated images or scans [8]. This can be challenging to obtain. A way of optimizing this is to use weakly supervised learning. This method reduces the amount of information that is annotated and extracted [7]. Decreasing the number of details that are annotated by the expert simplifies the process and annotation effort. An example of weakly labeled data could be an image of a tumor which is annotated, while the precise location or boundaries are not. The disadvantage of this method is that the number of annotated images needs to be substantially larger for the algorithm to learn to the same degree.

Unsupervised learning consists of unlabeled data presented to the algorithm. The algorithms search and analyze the data to detect clusterings or tendencies. These clusterings and tenden-

cies can be used to determine different features. Features can be applied to distinguish between the different classes. This is especially useful when working with content-based retrieval. A feature selection algorithm is often applied to reduce the number of features into a smaller composite set [7].

The workload that goes into handcrafting features within machine learning has prompted researchers to look at algorithms that are able to acquire features from data without human intervention, such as deep learning. In ultrasound, acoustic patterns are not obvious nor easily engineered. Given the ability to extract non-linear features from data, a deep learning algorithm is an especially good choice when working with ultrasound.

Chan et al. reported the first evidence that the ROC curve for radiologists' detection of clustered micro-calcifications was improved significantly when a computer output was available [9]. Brattain et al. has published an in-depth review of machine learning discussing the opportunities within medical ultrasound, which methods are applied and the status of the research published in February 2018 [7]. This review showcases some of the opportunities within medical ultrasound with regard to machine learning and deep learning. The author surveyed 56 papers with the aim of providing insight into the progress and the best approaches within ultrasound imaging. In particular, approaches using deep learning are highlighted and compared to approaches using handcrafted features [7]. Litjens et al. has published a review of the literature within different anatomical regions, modalities and architectures, which included over 300 contributions and goes in depth explaining the relevant concepts concerning medical image analysis [8].

Becker et al. has demonstrated the use of deep learning in a study applying generic software for industrial image analysis to diagnose breast cancer on breast ultrasound images. These results show high accuracy, comparable to human readers [10]. The algorithm was able to allow real-time analysis during the ultrasound examination [10]. These results show that the algorithm optimizes potential detection during the examination. The deep learning algorithm learns faster and better than a human reader with no prior experience given the same amount of training data [10].

The development of machine and deep learning from idea to a clinical product will require close collaboration between medical and data sciences and begins with the definition of needs. The concept of deep learning can potentially be applied to all imaging modalities and examinations allowing new standards for image interpretation systems. The first results are beginning to make their way into ultrasound conferences and we expect them to appear in journals like ours in the coming years.

„Machine“ und „Deep Learning“ im Ultraschall angewandt

Wir leben in einer aufregenden Zeit, um das exponentielle Informationswachstum zu betonen, von vielen als Informationszeitalter bezeichnet. Der Einsatz von medizinischen Bildgebungsverfahren nimmt stark zu [1]. So hat sich von 1996 bis 2010 die Anzahl der Ultraschalluntersuchungen etwa verdoppelt und die CT- und

MRT-Untersuchungen haben sich nahezu verdoppelt bzw. vervierfacht [2]. Die große Menge an Daten sowie die Notwendigkeit der Früherkennung von Pathologien sind eine Herausforderung [3]. Aufgrund der zunehmenden Arbeitsbelastung muss die Datenverarbeitung schneller und präziser sein. In diesem Editorial werden einige Grundprinzipien des maschinellen Lernens (Machine Learning) und des tiefgehenden Lernens (Deep Learning) in der Ultraschallbildgebung vorgestellt.

„Machine Learning“ ist eine Methode aus dem Bereich der künstlichen Intelligenz, bei der der Computer Muster in Datensätzen erfasst und diese Muster zur Entscheidungsfindung verwendet. „Machine Learning“ bietet verschiedene Vorteile in Bezug auf medizinische Bildgebung [4, 5]. Der Hauptzweck der verschiedenen Algorithmen besteht darin, die diagnostische Genauigkeit und die Konsistenz der Bildinterpretation zu erhöhen. Die Beobachtungen und Vorhersagen werden auf Grundlage von Daten erstellt, die dem Algorithmus zur Verfügung gestellt werden.

„Deep Learning“ als Teilgebiet des „Machine Learning“ hat in den letzten zwei Jahren besonders viel Aufmerksamkeit auf sich gezogen. Ein „Deep Learning“-Algorithmus lernt automatisch hierarchische Merkmale. Der Algorithmus besteht aus mehreren Schichten, die aus einfachen und nichtlinearen Modulen bestehen. Die Daten werden in Darstellungen umgewandelt, sodass der Algorithmus in der Lage ist die Daten zu unterscheiden [6]. „Deep Learning“-Algorithmen können aus früheren Fehlern lernen, wozu die traditionellen „Machine Learning“-Algorithmen nicht in der Lage sind, da diese mit manuell erstellten Merkmalen ausgestattet werden. Diese Merkmale sind nach der Konfiguration des „Machine Learning“-Algorithmus nicht anpassungsfähig.

Wie der Name der Algorithmen nahelegt, können „Machine“ und „Deep Learning“-Algorithmen lernen etwas zu erkennen. Dies geschieht normalerweise entweder durch überwachtes oder unüberwachtes Lernen. Beim überwachten Lernen werden dem Algorithmus markierte Daten präsentiert. Die Bilder sind mit einem geeigneten Klassifizierungsergebnis verknüpft. Der Algorithmus kann dann trainiert werden, für Eingabebilder bei denen eine Malignität auftritt, einen Wert von 1 auszugeben. Für Bilder ohne Pathologie oder benigne Läsionen wird dem Algorithmus beigebracht, eine Ausgabe mit dem Wert 0 [7] zu geben. Sobald das Training des Algorithmus abgeschlossen ist, kann der Algorithmus seine Klassifizierung von ungesehenen Bildern beginnen.

Damit der Algorithmus korrekt klassifizieren kann, muss er korrekt trainiert werden. Beim überwachten Lernen erfordert dies große Mengen gut annotierter Bilder oder Scans [8]. Dies stellt eine beträchtliche Herausforderung dar. Das semi-überwachte Lernen ist hierfür eine mögliche Lösung, denn diese Methode braucht weniger annotierte Informationen [7]. Indem die Anzahl vom Experten kommentierter Details verringert wird, wird der Prozess- und Annotationsaufwand vereinfacht. Ein Beispiel für schwach markierte Daten könnte ein Bild eines Tumors sein, der annotiert ist, aber seine genaue Position oder Grenzen nicht. Der Nachteil dieses Verfahrens besteht darin, dass die Menge der annotierten Bilder wesentlich größer sein muss, damit der Algorithmus in gleichem Maße lernt.

Unüberwachtes Lernen besteht aus unmarkierten Daten, die dem Algorithmus präsentiert werden. Der Algorithmus durchsucht und analysiert die Daten, um Cluster oder Tendenzen zu

erkennen. Diese Cluster und Tendenzen können verwendet werden, um verschiedene Merkmale zu bestimmen. Verschiedene Merkmale können angewendet werden, um zwischen den verschiedenen Klassen zu unterscheiden. Dies ist besonders nützlich, wenn mit inhaltsbasierter Suche gearbeitet wird. Häufig wird ein Merkmal-Auswahlalgorithmus verwendet, um die Anzahl der verwendeten Merkmale zu begrenzen [7].

Die Arbeitslast, die in die manuell erstellten Funktionen des „Machine Learning“ fließt, hat Forscher dazu bewogen Algorithmen zu untersuchen, welche ohne menschliches Eingreifen Merkmale aus Daten erfassen können, wie die des „Deep Learning“. Im Ultraschall sind akustische Muster weder offensichtlich noch leicht zu konstruieren. Angesichts der Fähigkeit, nichtlineare Merkmale aus Daten zu extrahieren, ist ein „Deep-Learning“-Algorithmus besonders gut für die Arbeit mit Ultraschall geeignet.

Chan et al. zeigt den ersten Nachweis, dass die ROC-Kurve für die Erkennung von Mikroverkalkungen durch Radiologen signifikant verbessert wurde, wenn eine Computerausgabe verfügbar war [9]. Brattain et al. gibt eine ausführliche Übersicht über „Machine Learning“ und die Möglichkeiten im medizinischen Ultraschall; welche Methoden angewendet werden und den Stand der Forschung, Februar 2018 [7]. Diese Übersicht zeigt einige der Möglichkeiten, die „Machine Learning“ und „Deep Learning“ im Bereich des medizinischen Ultraschalls bieten. Der Autor untersuchte 56 Arbeiten um einen Einblick in den Fortschritt und die besten Ansätze im Bereich der Ultraschallbildgebung zu geben. Ansätze, die „Deep Learning“ anwenden, werden besonders hervorgehoben und mit Herangehensweisen verglichen, die manuell erstellte Merkmale verwenden [7]. Litjens et al. gibt eine Literaturübersicht zu verschiedenen anatomischen Regionen, Modalitäten und Architekturen, wobei mehr als 300 Beiträge untersucht und die relevanten Konzepte zur medizinischen Bildanalyse vertieft werden [8].

Becker et al. hat in einer Studie, in der generische Software zur industriellen Bildanalyse zur Diagnose von Brustkrebs in Brust-Ultraschallbildern eingesetzt wird, die Anwendung von „Deep Learning“ demonstriert. Diese Ergebnisse zeigen eine hohe Genauigkeit, vergleichbar mit den Ergebnissen menschlicher Beurteiler [10]. Der Algorithmus war zu einer Echtzeitanalyse während der Ultraschalluntersuchung in der Lage [10]. Diese Ergebnisse zeigen, dass der Algorithmus die Diagnose während der Untersuchung verbessert. Der „Deep-Learning“-Algorithmus lernt schneller und besser als ein menschlicher Beurteiler ohne vorherige Erfahrung bei gleicher Menge an Trainingsdaten [10].

Die Entwicklung von „Machine Learning“ und „Deep Learning“ von der Idee bis zum klinischen Produkt erfordern eine enge Zusammenarbeit zwischen Medizin- und Datenwissenschaften und beginnt mit der Definition der Bedürfnisse. Das Konzept des „Deep Learning“ kann potenziell auf alle Bildgebungsmodalitäten und -untersuchungen angewendet werden, was neue Standards für Bildinterpretationssysteme ermöglicht. Die ersten Ergebnisse beginnen in Ultraschallkonferenzen Einzug zu halten, und wir gehen davon aus, dass sie in den nächsten Jahren in Zeitschriften wie der unseren publiziert werden.

References

- [1] Doi K. Diagnostic imaging over the last 50 years: Research and development in medical imaging science and technology. *Phys Med Biol* 2006; 51: R5–R27
- [2] Smith-Bindman R, Miglioretti DL, Johnson E et al. Use of Diagnostic Imaging Studies and Associated Radiation Exposure for Patients Enrolled in Large Integrated Health Systems, 1996-2010. *JAMA* 2012; 307: 2400–2409
- [3] Cancer Research UK, “Spot Cancer Early : Why is early diagnosis important?,”. 2015. [Online]. Available: <http://www.cancerresearchuk.org/about-cancer/cancer-symptoms/why-is-early-diagnosis-important>. [Accessed: 08-May-2018]
- [4] Azizi S, Imani F, Ghavidel S et al. Detection of prostate cancer using temporal sequences of ultrasound data: a large clinical feasibility study. *Int J Comput Assist Radiol Surg* 2016; 11: 947–956
- [5] Chi J, Walia E, Babyn P et al. Thyroid Nodule Classification in Ultrasound Images by Fine-Tuning Deep Convolutional Neural Network. *J Digit Imaging* 2017; 30: 477–486
- [6] Choi H. Deep Learning in Nuclear Medicine and Molecular Imaging: Current Perspectives and Future Directions. *Nucl Med Mol Imaging* 2018; 52: 109–118
- [7] Brattain LJ, Telfer BA, Dhyani M et al. Machine learning for medical ultrasound: status, methods, and future opportunities. *Abdom Radiol* 2018; 43: 786–799
- [8] Litjens G, Kooi T, Bejnordi BE et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017; 42: 60–88
- [9] Chan HP, Koi K, Vyborny CJ et al. Improvement in Radiologists Detection of Clustered Microcalcifications on Mammograms. The potential of computer-aided diagnosis. *Invest Radiol* 1990; 25: 1102–1110
- [10] Becker AS, Mueller M, Stoffel E et al. Classification of breast cancer from ultrasound imaging using a generic deep learning analysis software: a pilot study. *Br J Radiol* 2018; 91: 20170576