Application of Deep Learning in Medical Image Analysis

Maria Pia Bellini

June 2024

1 Abstract

Image processing and analysis are fundamental to disease detection, diagnosis, and treatment in several medical fields. In the last few decades, severe limitations in traditional techniques have been highlighted due to the reliance on expert radiologists and physicians. These issues led to the necessity of developing new techniques, in particular employing Deep Learning (DL), whose automation can overcome these boundaries. This essay explores the implementation of DL techniques in medical image processing in various fields, such as early detection of cardiovascular diseases, and central nervous system tumours, and their role during the COVID-19 pandemic. Notable DL models and studies across these cases will be looked at in detail and discussed.

2 Introduction

Medical image interpretation and processing such as computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), mammography, ultrasound, and X-rays are essential for the early detection, diagnosis, and appropriate treatment of diseases [13]. Computer-assisted analysis has gained popularity across medical fields since, traditionally, image interpretation was entrusted exclusively to expert radiologists and physicians. This human reliance raises issues such as the significant variability in pathological presentation and human fatigue.

2.1 The limitations of traditional techniques

Machine learning for image processing consists of recognising relevant patterns in data according to some relevant features. Those features were traditionally designed by experts in the sector based on their domain knowledge, which makes it difficult for non-experts to understand and apply them. A solution developed to overcome this issue is **Sparse representation** using predefined dictionaries. This approach derives dictionaries from training samples and is based on the principle of parsimony, meaning that simpler structures are preferred over complex ones. The implementation of sparse representation, along with sparsity-inducing penalisation and dictionary learning, have been used effectively in medical image analysis, but still operate with superficial architectures, failing to capture complex patterns. Deep Learning overcomes this problem by shifting from manual to **Automated Feature Engineering**, which allows non-experts to develop their research and applications in medical image analysis. DL manages to do so by integrating the feature engineering process into the learning phase, meaning that instead of manually extracting features, it only requires a dataset with minimal preprocessing [11].

2.2 Deep Learning Methods

Deep Learning is characterized by the possibility of creating networks with multiple layers, leading to advancements in high-performance central processing units (CPUs) and graphics processing units (GPUs), the availability of vast amounts of data (big data), and innovations in learning algorithms [7]. That implies deep neural networks go beyond traditional artificial neural networks, being able to learn hierarchical feature representation, which means they can autonomously learn higher-level features from lower-level ones [26].

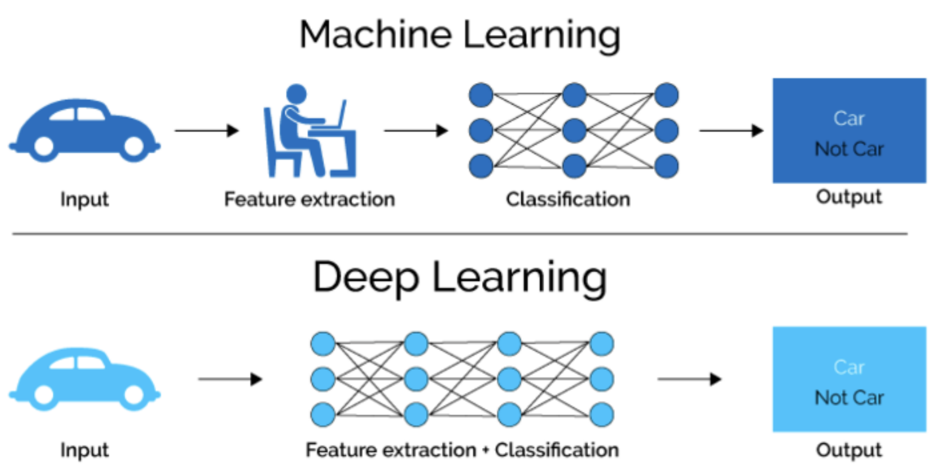

Figure 1: Machine Learning vs Deep Learning [1]

The application of Deep Learning is implemented in several medical image analysis tasks, including segmentation, registration, fusion, and annotation of the image; computer-aided diagnosis (CADx) and prognosis; lesion/landmark detection; and microscopic image analysis.

Deep Learning methods have one important limitation: they require a large number of samples during the training phase, and millions of images need to be involved. Medical applications often provide much smaller datasets, typically less than 1,000 images, leading to overfitting DL models.

To address this issue, researchers have developed various strategies, including:

- **Patch-Based Input:** using two-dimensional (2D) or three-dimensional (3D) image patches instead of full-sized images to reduce input dimensionality and the number of model parameters;
- **Data Augmentation:** augmenting datasets by artificially generating samples through affine transformations and training networks from scratch with this augmented data;
- **Feature Extraction:** employing deep models trained on extensive datasets of natural images as feature extractors, then training the final classifier or output layer with the specific medical samples;
- **Transfer Learning:** initializing model parameters with those from pre-trained models on non-medical or natural images and fine-tuning the network with the medical task-related samples;
- **Model Adaptation:** Adapting models trained with small-sized inputs to handle arbitrarily sized inputs by converting weights in fully connected layers into convolutional kernels. [11]

2.2.1 DL methods for Medical Images

The Deep Learning process for ultrasound medical imaging can be divided into three phases: data acquisition and processing, network selection, and trading and evaluation of network performance [6].

1. **Data Acquisition and Preprocessing:** When acquired, medical images are typically in a particular format called DICOM; hence, conversion to a standard image format, such as JPG or PNG, is required before processing. Various tools have been developed for direct conversions, such as the dicomread function in MATLAB. Preprocessing means taking the images, now in the right format, and performing some steps, such as filtering out the low-quality images, reducing noise, and adding tags. After that, processed images can be utilized for subsequent analysis [6].

Table 1: Common pre-processing techniques

Preprocessing technique	Description	Modality
Thresholding	The process of constraining the pixel values of an image to be between predefined values.	CT, MRI
Normalisation and whitening	The process of transforming the distribution of image pixels to some distribution which is standardised across images.	CT, MRI, X-ray
Cropping	The process of removing unwanted outer pixels or voxels of an image prior to being inputted to the network. Cropping is commonly used to reduce computational cost and/or eliminate the influence of background voxels.	CT, MRI, X-ray, PET
Denoising	The process of removing noise from images in order to improve their quality.	CT, MRI

2. **Network selection:** In recent years, Deep Learning technology has enhanced image recognition. That led to different networks, each with their strengths and weaknesses and their environment of applications. Different networks can be chosen according to the different tasks that have to be performed and the data to analyze. In 2012, Hinton et al. [2] proposed the deep convolutional neural network AlexNet, which implemented stochastic gradient descent and dropout optimization techniques. Hinton won the ILSVRC championship and achieved a top-5 error rate of 15.3This led to explosive research, with many successful DL models from 2014, such as Google LeNet by Szeged C. et al. [5] and VGGNet by Simonyan K. et al. [19]. In 2015, He K. et al. designed 152 layers of ResNet [18], with an error rate of 3.6%, which shocked the academic community by surpassing the human level [6].

3. **Training and evaluation:** After choosing an adequate network, it needs to be trained and evaluated. That means adjusting the network parameters and layer settings and subsequentially validating the network on a test set [6].
   The principal categories of validation techniques are reported in the figure below:

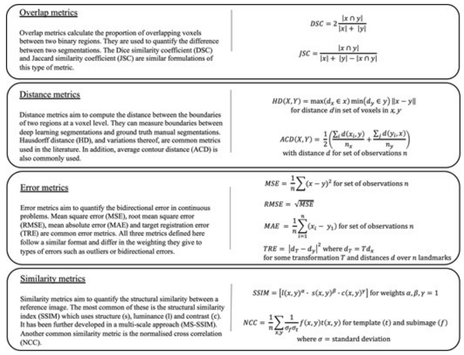

Figure 2: Overview of four key categories of evaluation metrics (overlap, distance, error and similarity) [3]

3 Related work

3.1 Cardiovascular Diseases

Cardiovascular disease diagnosis and prevention have become more and more relevant in recent years. Despite the initial decline over the past decade, global incidence and mortality rates of coronary heart disease have been on the upswing [25]. Structural and functional changes in the myocardium often lead to serious complications, such as ischemic cardiomyopathy, malignant arrhythmias, systemic circulation embolism, and multi-organ damage. The 5-year survival rate for ischemic cardiomyopathy ranges from 26% to 52%.

Clinical management of cardiovascular diseases gives rise to several challenges, including optimization of prevention and treatment costs, and minimizing excessive interventions and inadequate patient care. The research on DL techniques for this scope gained popularity to efficiently face these problems and reduce high readmission rates and mortality [6].

CermgApp (Cardiac Electro-Mechanics Research Group Application) is a platform for applying statistical, ML, and simulation approaches to the diagnosis and treatment of the cardiovascular system, developed by Razeghi O. et al. [**?**]. The platform provides an integrated environment, with a user-friendly interface that makes it easy to use and approachable for non-experts. A simple scheme of the CermgApp working mechanism:

1. Loading and visualizing CMR images.
2. Converting the scans from the DICOM6 format into the anonymized NIfTI7 format.
3. Segmentation step, which can be performed automatically by the convolutional neural network [10].

4. Alignment of the segmentation with the LGE scan using the MIRTK rigid registration option [23].

5. , Automatic localization and vessel truncation tool to isolate the LA cavity from the segmentation.

6. Creation of a surface mesh from the truncated segmentation. Normals are taken, 3 mm externally and 1 mm internally, to the nodes of the mesh, and a maximum intensity projection technique is used to interrogate the LGE scan and identify fibrotic regions.

7. Finally, global scar burdens are calculated using predefined thresholds [22].

CermgApp speeds up clinical translation due to its capacity to handle large datasets and generate automated and reproducible results. That also reduces variability in studies, leading to more cost-effective and quicker clinical trials.

To date, successful studies employed by clinical researchers for various applications proved the efficiency of utilizing CermgApp and increased confidence in outcomes, proving that the results are not dependent on the user [22].

These studies include improving co-registration of ex-vivo and in-vivo CMR images [17], assessing the reproducibility of atrial fibrosis quantification [15], and guiding lead placement in cardiac resynchronization therapy (CRT) [23].

3.2 Central Nervous System Tumors

In oncology, the interpretation of medical imaging raises peculiar challenges, such as the need for detailed clinical and disease-specific histories, and the reliance on human interpretations, which often involve incompletely liked data sources.

Artificial Intelligence (AI)-driven methods have been implemented to facilitate imaging applications in oncology treatment, including the generation of radiation therapy plans, assessment of treatment responses, and prognosis prediction. For instance, Computer-Aided Detection (CADe) and Computer-Aided Diagnosis (CADx) systems, have been successfully implemented for the detection of tuberculosis, lung cancer, and brain metastases [21]. AI methods have also been explored for the early detection of various cancer types, including lung, breast, prostate, central nervous system cancers, and others [14].

Deep Learning methods have been implemented and combined with traditional methods, leading to hybrid models that result in excellent outcomes.

3.2.1 DL-based methods

Molecular subtyping is a significant task in central nervous system tumours, especially in diffuse gliomas. Li et al. used preoperative multiparametric MRI in 1016 diffuse glioma patients, dividing them into training (n = 780) and validation (n = 236) sets. They developed predictive models based on radiomics and deep convolutional neural networks (DCNNs), finding that while both radiomics and DCNN models could predict molecular subtypes of diffuse gliomas preoperatively, DL models generally outperformed, with AUCs ranging from 0.85 to 0.89. However, they also noted a low correlation between radiomics and DCNN features [20].

3.2.2 Hybrid approaches

Hybrid models combine Machine Learning and Deep Learning approaches [12], integrating preprocessing and segmentation, and leveraging both manual techniques and ML/DL methods. For instance, Fu et al. employed a 3-dimensional CNN trained on manually annotated MRI images of glioma patients to automatically segment tumour contours [16]. Subsequently, both handcrafted and DL-based radiomic features were extracted from these auto-segmented contours using specialized algorithms and pre-trained CNNs. Cox regression models then utilized these features to create prognostic signatures, demonstrating that the DL-based approach achieved higher performance (mean Dice coefficient of 0.67) compared to handcrafted methods (0.64). This method successfully stratified patients into distinct prognostic groups, showcasing the efficacy of DL in improving clinical outcomes [20].

3.3 Covid-19 Pandemic

The COVID-19 pandemic and the associated containment measures have caused a global health crisis, affecting all aspects of human life. Initially, when the number of infected individuals was low, the disease did not seem to pose a significant threat, with most cases resolving on their own. However, as time went on, the World Health Organization (WHO) declared COVID-19 an outbreak with a very high risk of impacting millions of lives, particularly in countries with weaker healthcare systems. The virus is particularly dangerous for two main reasons: it is new and lacks a vaccine, and it spreads easily through direct or indirect contact with infected individuals. The daily increase in new cases has been rapid, forcing governments and authorities worldwide to implement strict lockdowns to enforce social distancing and contain the disease [4].

One of the significant challenges in COVID-19 research is the lack of reliable and sufficient data. Due to the limited number of tests conducted, many deaths and cases of the virus go unreported. **Information fusion** is defined as the process of combining and associating information from one or multiple sources to provide useful data for detecting, identifying, and characterizing a specific entity. In machine learning (ML) and deep learning (DL) applications, the availability of large-scale, high-quality datasets is crucial for achieving accurate results. Information fusion facilitates the integration of multiple datasets, which can be used in DL models to improve prediction accuracy.
In the study by Apostolopoulos and Mpesiana [8], X-ray image datasets from GitHub, Cohen, the Radiology Society of North America (RSNA), and the Italian Society of Medical and Interventional Radiology (SIRM) were combined and

utilized to train a convolutional neural network (CNN) for COVID-19 detection.

The phenotype of COVID-19 often begins with mild or no symptoms but can quickly progress to severe, critical conditions, sometimes resulting in multi-organ failure and fatal outcomes. The primary objective is to **prevent such rapid deteriorations** by detecting the disease as early as possible, despite the limited knowledge and resources available. Traditional lab-based RT-PCR (real-time polymerase chain reaction) tests using nose and throat swabs have limited sensitivity and are time-consuming. When patient numbers surge, shortages of RT-PCR reagents and specialized lab resources for COVID screening become inevitable. Thus, tools that can augment these resources are crucial.
Machine learning (ML) and artificial intelligence (AI) are techniques with the potential to speed up decision-making and improve patient outcomes. Various studies have developed ML models that use minimal resources while achieving accuracy (82–86%) and sensitivity (92–95%) comparable to the gold standard RT-PCR test [9].
Integrating AI and ML models with radiological images can lead to more accurate early detection of the disease, ensuring timely treatment availability. For instance, models like DarkNet can detect and classify COVID-19 cases as pneumonia, proving particularly useful in areas experiencing a shortage of radiologists due to an overwhelming number of patients [24].

4 Conclusions

Deep learning (DL) has revolutionized the field of medical image processing, offering numerous advantages over traditional methods. These advantages have been instrumental in advancing diagnostics, treatment planning, and patient outcomes across various medical domains.

- **Automated Feature Extraction:** DL techniques automate the feature extraction process, eliminating the need for manual feature engineering and enabling the discovery of complex patterns in medical images.

- **Improved Accuracy:** DL models have demonstrated higher accuracy in image classification, segmentation, and detection tasks compared to traditional methods, often rivalling or exceeding the performance of human experts.

- **Scalability:** Once trained, DL models can process large volumes of medical images rapidly and consistently, making them suitable for high-throughput applications.

- **Adaptability:** DL models can be fine-tuned for various medical imaging modalities and applications, providing flexibility in addressing diverse clinical challenges.

- **Enhanced Sensitivity and Specificity:** DL techniques have shown high sensitivity and specificity in detecting diseases, contributing to early diagnosis and improved patient outcomes.
- **Reduction of Human Bias:** By relying on data-driven approaches, DL models minimize the influence of human bias in image interpretation and analysis.
- **Integration with Clinical Workflows:** DL tools can be seamlessly integrated into existing clinical workflows, enhancing the efficiency and effectiveness of diagnostic processes.

Table 2: summary of DL pros for Specific Medical Application

Application	**Advantages of DL Techniques**	**Key Benefits**	**Notable Models/ Studies**
Cardiovascular Diseases	Enhanced image segmentation	Improved detection and characterization of heart conditions	CemrgApp
	Accurate disease classification	Better management of ischemic cardiomyopathy	Automated atrial segmentation
	Prognostic predictions	Reduction in intra- and interobserver variability	
Central Nervous System Tumors	Detailed tumor segmentation	Enhanced diagnostic precision	Radiomics and DCNN models for gliomas
	Molecular subtyping	Personalized treatment plans	3D CNNs for tumor contouring
	Treatment response prediction	Early detection of tumor recurrence	
COVID-19	Rapid diagnosis	Early detection of COVID-19 from radiological images	CNN for X-ray detection
	Efficient resource utilization	Alleviating healthcare resource constraints	DarkNet model for classification
	High sensitivity and specificity	Support in areas with radiologist shortages	

References

[1] Log analytics with generative ai. https://www.xenonstack.com/blog/log-analytics-generative-ai.

[2] Krizhevsky A., Sutskever I., and Hinton G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 2012.

[3] Joshua R. Astley, Jim M. Wild, and Bilal A. Tahir. Deep learning in structural and functional lung image analysis. *British Journal of Radiology*, 2022.

[4] Sweta Bhattacharyaa, Praveen Kumar Reddy Maddikunta, Quoc-Viet Pham, Thippa Reddy Gadekallu, Siva Rama Krishnan, Chiranji Lal Chowdharya, Mamoun Alazab, and Md. Jalil Piran. Deep learning and medical image processing for coronavirus (covid-19) pandemic: A survey. *Sustainable Cities and Society*, 2021.

[5] Szegedy C., Liu W., and Jia Y. et al. Going deeper with convolutions. *Computer Vision and Pattern Recognition*, 2015.

[6] Yankun Cao, Zhi Liu, Pengfei Zhang, Yushuo Zheng, Yongsheng Song, and Lizhen Cui. Deep learning methods for cardiovascular image. *Journal of Artificial Intelligence and Systems*, 1:96–109, 01 2019.

[7] Bishop CM. Neural networks for pattern recognition. *Oxford Univ. Press*, 1995.

[8] Apostolopoulos I. D. and Mpesiana. Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, 2020.

[9] Brinati D., Campagner A., Ferrari D., Locatelli M., Banfi G., and Cabitza F. Detection of covid-19 infection from routine blood exams with machine learning: A feasibility study. *medRxiv*, 2020.

[10] Rueckert D., Sonoda L., Denton E., Rankin S., Hayes C., and Leach M. et al. Comparison and evaluation of rigid and nonrigid registration of breast mr images. *Hanson KM*, 1999.

[11] Shen D., Wu G., and Suk H. Deep learning in medical image analysis. *Annu Rev Biomed*, 2017.

[12] Nassiri F., Liu J., and Patil V. et al. A clinically applicable integrative molecular classification of meningiomas. *Nature*, 2021.

[13] Brody H. *Medical Imaging*. Nature, 2013.

[14] Shimizu H. and Nakayama KI. Artificial intelligence in oncology. *Cancer Sci*, 2020.

[15] Sim I., Razeghi O., Karim R., Chubb H., Whitaker J., and O'Neill L. et al. Reproducibility of atrial fibrosis assessment using cmr imaging and an open source platform. *JACC Cardiovasc Imaging*, 2019.

[16] Fu J., Singhrao K., and Zhong X. et al. An automatic deep learning-based workflow for glioblastoma survival prediction using preoperative multimodal mr images: A feasibility study. *Adv Radiat Oncol*, 2021.

[17] Whitaker J., Neji R., Byrne N., Puyol-Antón E., Mukherjee R., and Williams S. et al. Improved co-registration of ex-vivo and in-vivo cardiovascular magnetic resonance images using heart-specific flexible 3d printed acrylic scaffold combined with non-rigid registration. *J Cardiovasc Magn Reson*, 2019.

[18] He K., Zhang X., and Ren S. et al. Deep residual learning for image recognition. *Computer Vision and Pattern Recognition*, 2016.

[19] Simonyan K. and Zisserman A. Very deep convolutional networks for large-scale image recognition. *arXiv*, 2014.

[20] AV. Krauze, Y. Zhuge, R. Zhao, E. Tasci, and K. Camphausen. Ai-driven image analysis in central nervous system tumors- traditional machine learning, deep learning and hybrid models. *J Biotechnol Biomed*, 2022.

[21] Sunwoo L., Kim YJ., and Choi SH. et al. Computer-aided detection of brain metastasis on 3d mr imaging: Observer performance study. *PLoS One 12*, 2017.

[22] Orod Razeghi, José Alonso Solís-Lemus, Angela W.C. Lee, Rashed Karim, Cesare Corrado, Caroline H. Roney, Adelaide de Vecchi, and Steven A. Niederer. Cemrgapp: An interactive medical imaging application with image processing, computer vision, and machine learning toolkits for cardiovascular research. *SoftwareX*, 2020.

[23] Banks T., Razeghi O., Ntalas I., Aziz W., Behar J., and Preston R. et al. Automated quantification of mitral valve geometry on multi-slice computed tomography in patients with dilated cardiomyopathy - implications for transcatheter mitral valve replacement. *J Cardiovasc Comput Tomogr*, 2018.

[24] Ozturk T., Talo M., Yildirim E. A., Baloglu U. B., Yildirim O., and Acharya U. R. Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in Biology and Medicine*, 2020.

[25] World Health Organization. *World Health Statistics 2018: Monitoring Health for the SDGs, Sustainable Development Goals*. World Health Organization, Geneva, 2018. Licence: CC BY-NC-SA 3.0 IGO.

[26] LeCun Y., Bengio Y., and Hinton G. Deep learning. *Nature*, 2015.