

Article

# CADUCEO: A Platform to Support Federated Healthcare Facilities through Artificial Intelligence

Danilo Menegatti , Alessandro Giuseppe, Francesco Delli Priscoli, Antonio Pietrabissa, Alessandro Di Giorgio , Federico Baldisseri \*, Mattia Mattioni, Salvatore Monaco, Leonardo Lanari  and Martina Panfilì and Vincenzo Suraci 

Department of Computer, Control and Management Engineering “Antonio Ruberti”, Sapienza University of Rome, Via Ariosto 25, 00185 Rome, Italy; menegatti@diag.uniroma1.it (D.M.); giuseppi@diag.uniroma1.it (A.G.); dellipriscoli@diag.uniroma1.it (F.D.P.); pietrabissa@diag.uniroma1.it (A.P.); digiorgio@diag.uniroma1.it (A.D.G.); mattioni@diag.uniroma1.it (M.M.); monaco@diag.uniroma1.it (S.M.); lanari@diag.uniroma1.it (L.L.); panfilì@diag.uniroma1.it (M.P.); suraci@diag.uniroma1.it (V.S.)

\* Correspondence: baldisseri@diag.uniroma1.it

**Abstract:** Data-driven algorithms have proven to be effective for a variety of medical tasks, including disease categorization and prediction, personalized medicine design, and imaging diagnostics. Although their performance is frequently on par with that of clinicians, their widespread use is constrained by a number of obstacles, including the requirement for high-quality data that are typical of the population, the difficulty of explaining how they operate, and ethical and regulatory concerns. The use of data augmentation and synthetic data generation methodologies, such as federated learning and explainable artificial intelligence ones, could provide a viable solution to the current issues, facilitating the widespread application of artificial intelligence algorithms in the clinical application domain and reducing the time needed for prevention, diagnosis, and prognosis by up to 70%. To this end, a novel AI-based functional framework is conceived and presented in this paper.

**Keywords:** deep learning; artificial intelligence; e-health



**Citation:** Menegatti, D.; Giuseppe, A.; Delli Priscoli, F.; Pietrabissa, A.; Di Giorgio, A.; Baldisseri, F.; Mattioni, M.; Monaco, S.; Lanari, L.; Panfilì, M.; et al. CADUCEO: A Platform to Support Federated Healthcare Facilities through Artificial Intelligence. *Healthcare* **2023**, *11*, 2199. <https://doi.org/10.3390/healthcare11152199>

Academic Editor: Joaquim Carreras

Received: 13 July 2023

Revised: 26 July 2023

Accepted: 31 July 2023

Published: 4 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Deep learning (DL) is a branch of artificial intelligence (AI) that relies on artificial neural networks (ANN) capable of learning from large amounts of data and extracting information relevant to various tasks. DL has found numerous applications in the medical field, particularly in imaging diagnosis [1], clinical and drug research [2], disease classification and prediction [3], personalized therapy design [4], and public health monitoring [5]. It can offer significant advantages over traditional data analysis methods, both in terms of performances and automation [6]. Moreover, it can also complement and enrich the knowledge and skills of healthcare professionals, providing them with evidence-based clinical decision support tools. However, DL also presents challenges and limitations related to the following:

- The need for quality and representative data;
- The difficulty of interpreting and explaining the results given as output by ANNs, i.e., explaining their inner workings;
- The legal and ethical implications arising from the use of intelligent systems in clinical practice.

Having domain-specific representative data has always been the bottleneck for data-driven methods; if ANNs are trained over data which are not statistically informative with respect to the problem being addressed, then, even if they exhibit pleasing behavior over training data, they may fail to generalize when applied in a real-world scenario.

This problem, although common, becomes crucial in the medical field where most clinical data are protected by privacy laws, such as the General Data Protection Regulation (GDPR) [7], and cannot be used without the patient's consent. Furthermore, even if the patients give consent, their data need to be anonymized before being used to train ANNs, so that it is not possible in any way to track it back to the single patient. Additionally, even if data are maintained by hospitals or other clinical institutes and have previously been anonymized, ethical approval must be obtained before using it.

To this end, federated learning (FL) [8] offers a valid strategy to overcome most of the above-mentioned problems. FL is a solution for distributed machine learning (ML) problems which allows a federation of clients sharing the same machine learning model to learn a model without any exchange of data, in a privacy-preserving and GDPR-compliant way. It can allow exploitation of the vast amount and variety of medical data available in different sources, increasing the statistical power and generalizability of ML models while also providing a solution to overcome privacy, security, and data governance issues that prevent the sharing of medical data between various healthcare institutions. FL has been successfully applied to various clinical fields, such as imaging diagnosis, drug research, and genomics.

Although FL makes it possible to overcome data sharing issues, and to implement strategies such as Continuous Learning (CL) [9], the lack of explainability of the inner workings of ML models, such as ANNs, is a major limitation. Explainable AI (XAI) solutions [10] offer valid tools to explain the functioning of ML algorithms in a way that humans can easily understand and interpret. Most of the time, the specific XAI solution depends on the particular type of data being processed, such as Grad-CAM [11] for convolutional neural networks (CNNs) for image classification; however, data type-independent solutions, such as LIME [12] or NAMs [13], can also be considered.

Making the inner workings of ML models interpretable to humans is the first step toward Trustworthy AI which focuses on creating AI systems which are reliable and respect human rights. To this end, XAI can help ensure that AI systems are robust, reliable, and ethically sound.

This work proposes a novel functional architecture for the introduction and use of AI algorithms in clinical practice. In particular, we will focus on the project "Cloud pLatform for intelligent prevention and Diagnosis sUpported by artifiCial intelligEnce solutiOns" (CADUCEO) which aims to design and develop a cloud platform for the realization of medical applications for the prevention, diagnosis, and prognosis of diseases, integrating advanced AI technologies and solutions with reference to three diseases of the digestive system: eosinophilic esophagitis [14], chronic inflammatory bowel diseases [15], and portal hypertension [16].

The main contribution of this paper is in the discussion of the functional architecture designed in the project CADUCEO. The project proposes a federated platform which runs based on FL algorithms, which allows the cohort medical centers to foster the performance of their machine learning-based algorithms by sharing the knowledge they locally generate, without having to share privacy-sensitive medical information. In addition, this paper introduces the main machine learning-based algorithms developed in the CADUCEO project for providing decision support to medical operators in some operations, including automated image processing (e.g., for image inpainting and cell counting), data augmentation for improving the classifiers' detection accuracy, and support to diagnosis through automated image classification. In conclusion, this paper provides a broad overview of the role that the latest AI algorithms can play in improving the efficiency of medical operations, in terms of agility and improved quality of treatment.

The remainder of this paper is organized as follows. Section 2 reports used materials and methods. Section 3 details the results of the system design. Section 4 discusses the functionalities and the considered use cases. Finally, Section 5 draws the conclusion and future works.

## 2. Materials and Methods

### 2.1. Data Augmentation

In the machine learning literature, data augmentation techniques are commonly classified into two main categories [17]. The first category involves methodologies that directly manipulate images, such as algebraic–geometric transformations [18], color modifications [19], noise addition [20], image combinations [21], and random deletions [22]. However, these methods are known to have some drawbacks and often result in marginal performance improvements that may not justify their usage [23].

On the other hand, the second category employs deep learning techniques for data augmentation, which includes increasing the feature space [24], adversarial training [25], and generative adversarial networks (GANs). GANs have gained significant popularity in the clinical field since their introduction in 2014 [26]. These networks have the ability to generate new training data, leading to substantial performance enhancements for classifiers compared to training on the original, limited dataset. GANs have been successfully applied to various medical applications, including the study of electroencephalograms [27], brain tumor detection [28], and analysis of dermatological lesions [29]. Moreover, diversified GAN architectures have been utilized in [30] for risk stratification and disease classification in rare clinical cases, namely, primary Sjögren’s syndrome and hypertrophic cardiomyopathy. Meanwhile, traditional algebraic transformation methods continue to find applications in the medical field, such as in [31], where five different direct image manipulation methods are analyzed for prostate cancer detection.

While there are not many studies systematically comparing various data augmentation techniques, a study [32] demonstrates that the most effective methods include geometric transformations (rotations, flips, and cropping) and GANs.

In recent years, a promising class of methodologies known as meta-learning has gained popularity [33]. These methods rely on the principle of neural architecture search [34], wherein neural networks optimize other neural networks. Specifically, optimization mechanisms are employed to tune the hyperparameters of a neural network, maximizing classification accuracy. Meta-learning and neural network optimization for data augmentation are addressed in [35,36], where a double-cascade neural network architecture is proposed, with the second network trained to optimize the parameters of the first one.

### 2.2. Federated Learning

Federated learning (FL), which is distributed and protects privacy, is a great option for healthcare applications. FL has a wide range of applications in situations where cooperation among stakeholders, including hospitals, laboratories, and governmental organizations, would considerably enhance the management of particular issues. The management of electronic health records (EHRs), remote health monitoring, and medical imaging are a few examples of such applications [37].

EHRs digitally archive patient health data, such as diagnosis, treatments, and analyses [38]. They serve as an important source of information that could help with disease diagnosis and evaluation. Sharing EHRs among various institutions can, however, provide regulatory difficulties. The authors of [39] suggest a collaborative learning protocol (PRCL) based on the FL framework that is privacy aware and resource efficient.

In order to avoid data breaches, PRCL employs a model-splitting technique in which the most computationally intensive portion of the learning process is delegated to a cloud server.

The FL-based method SplitNN is introduced in [40] which also presents a model-splitting strategy. The plan is to divide the model into various components, each of which will be trained by a different customer, avoiding a direct data exchange. Reference [41] uses data from patient EHRs to predict hospitalizations for cardiac events using an FL setup based on the soft-margin l1-regularized sparse Support Vector Machine (sSVM) classifier and an iterative cluster Primal–Dual Splitting (sPDS) method.

In order to solve privacy concerns, ref. [42] introduces an FL framework built on a differential privacy mechanism that has been successfully used with real-world EHRs from one million patients. In [43], where synthetic sounds are applied to the parameters at the clients' side prior to aggregation, differential privacy is also used. Federated-autonomous deep learning (FADL), an FL technique, is utilized in [44] to forecast patients' death using their EHRs. The innovative feature is that some parts of the model are trained using distributed distribution of data from all sources, while other parts are trained using specific sources.

With the introduction of the stochastic FL technique known as stochastic channel-based federated learning (SCBF), the prediction of patients' mortality as a binary classification problem is investigated in [45]. By stochastically choosing the clients with the biggest local gradients, it is possible to update the server model's weight while maintaining privacy.

### 2.3. Explainable AI

Explainable artificial intelligence (XAI) is a set of methods and processes that allow users to understand and trust the results produced by machine learning algorithms [46]. Generally, the inner workings of AI algorithms are inaccessible and/or incomprehensible to humans. However, in certain applications, it is necessary to provide clear explanations for the reasons behind a specific output. Moreover, the interpretability of an algorithm's output enables users to provide feedback that can be used to improve its accuracy and performance [47].

The main limitation of XAI lies in the empirical trade-off relationship between prediction accuracy and interpretability: as the model becomes more performant, it becomes less interpretable [48].

There are three categories of XAI methods [49]:

- **Intrinsically Interpretable Methods:** These methods use inherently interpretable models, such as Decision Trees and Linear Regression.
- **Sample-Based Methods:** Specific samples from the dataset are used to describe the model's behavior. An example is the K-Nearest Neighbors method.
- **Model-Agnostic Methods:** The explanation method is separate from the model being explained. Examples include Global Surrogate and Local Surrogate.

Methods in the first category generally achieve lower performance because, as mentioned earlier, high interpretability is traded-off with prediction accuracy. Possible solutions to this problem include using specific sample sets to describe the model (Category 2) or decoupling the explanation method from the model (Category 3), which will be discussed further.

Perceptive Interpretability methods generate elements assumed to be immediately interpretable [50]. For example, the Saliency Method explains the algorithm's decision by assigning values that reflect the importance of input components in contributing to the decision. This can be achieved through probabilities or heatmaps [47].

Signal Methods monitor the activation of specific neurons or neuron groups, allowing the activated values to be manipulated or transformed into interpretable forms. For example, the activation of neurons in a given layer can be used to reconstruct an image similar to the input. Signal Methods are considered interpretable as they connect the generation of a certain output to a reduced set of influential features. Deconvolution is often used to make the output more interpretable [50].

Verbal Interpretability methods aim to provide fragments of text that are naturally understandable to humans. These methods seek to create sentences that indicate logical causality relationships, like "If A happens, then B happens", or suggest actions based on certain conditions, like "If condition A occurs, then execute action B". The output is represented as a binary sequence, which follows a probability distribution [50].

Interpretability via Mathematical Structure methods generate outputs that require additional cognitive processing by humans compared to the immediately interpretable outputs of Perceptive Interpretability methods.

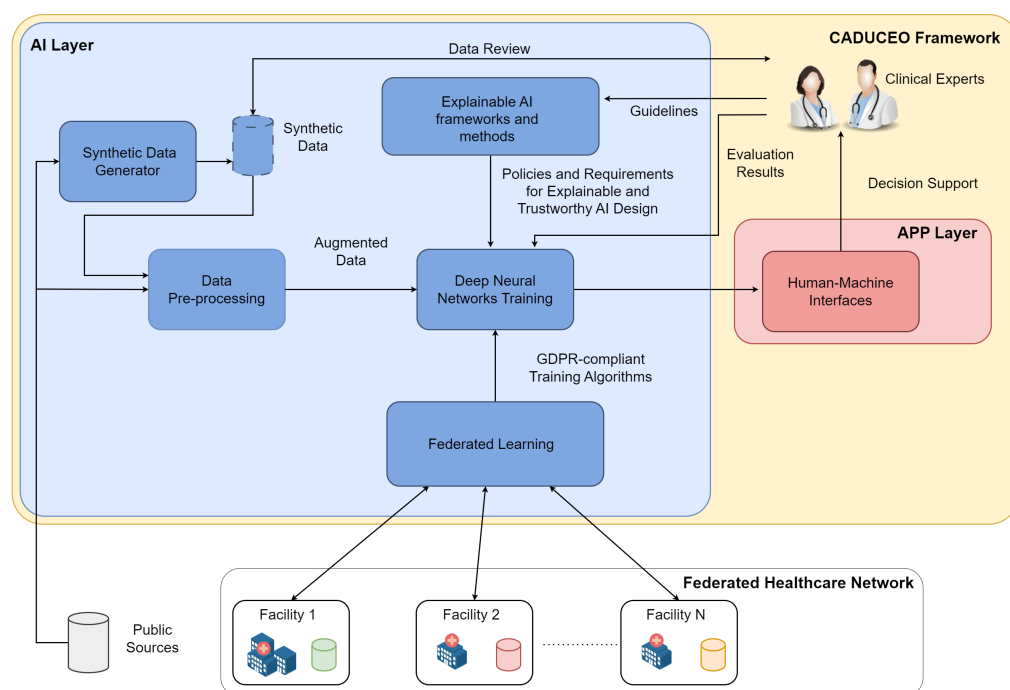
Predefined Models allow the study of systems with unknown behavior by defining parametric models with relevant and easily interpretable terms. The simplest interpretable

predefined model is the linear combination of variables. The Generalized Additive Model (GAM) is a generalization of linear models that includes easily interpretable terms [50].

### 3. Results

In light of the previous discussion, in this section, we will detail the overall functional architecture that we propose for the design of a software platform able to support healthcare operators in their decisions.

To cope with the data paucity, the complexity of the ML model training and to ensure GDPR compliance and trustworthiness of the overall system, we have identified five key functionalities, gathered in the so-called AI Layer, whose logical interaction is depicted in Figure 1.



**Figure 1.** General architecture of the proposed distributed AI system.

- The data pre-processing system is a functional block that contains all the software solutions for data and clinical image enhancement. These solutions may employ both classical computer vision techniques, such as de-noising [51] and feature selection [52], and deep learning techniques, such as deep image inpainting [53], depending on the nature and characteristics of the data being processed. Data transformation and curation is one of the fundamental steps behind the design and implementation of any data-driven system and has to be designed in such a way that the integrity and correctness of the data to be analyzed are preserved.
- The synthetic data generator block gathers the functionalities related to the generation of synthetic, or artificial, data, whose function is to support and improve the training process of the AI functionalities. The role of synthetic data is of central importance in all applications where the availability (and quality) of real data is limited, as is common in the medical domain. Starting from publicly available data, as well as anonymized data samples coming from clinical facilities and deep learning solutions, including generative adversarial networks [54] and deep reinforcement learning [55], appropriately specialized on the specific clinical use cases, synthetic data will be generated. And it will undergo a review process by clinical experts to assess its quality to increase the trustworthiness of the overall trained system.
- Federated learning (FL) has been identified as an enabling technology for the distributed and decentralized training of the ML models underlying the platform's AI

capabilities. As the solutions of the CADUCEO framework are developed for clinical use cases, it is crucial to ensure their full compatibility with the GDPR and to observe in their design all appropriate precautions related to the protection of patient privacy. In this respect, FL makes it possible to ensure collaboration between different institutions, including prospective customers interested in improving certain functionalities, without the need for any exchange of clinical data. Integrating an FL system into the AI Layer thus ensures its maximum compatibility in terms of adoption and further developments.

- The explainable AI framework and methods functional block contains all the software features and guidelines to direct the design and development of AI solutions from the perspective of the explainability and interpretability of their results. Indeed, it is critical to the goals of the CADUCEO framework to offer clinical experts not only a set of recommendations and suggestions but to supplement them with adequate information critical to their correct interpretation.
- The training of machine learning models block, representing from a functional point of view the whole design process of the machine learning models (such as deep learning solutions), deals with the process of training them on the basis of available data (from project sources, public sources, and synthetic ones). The training of the models will take place under the constant supervision of clinical experts so that aspects related to the interpretability of the results and the quality of the proposed suggestions can be improved.

The figure details how these functional blocks are interconnected, the main functional interfaces with the Platform Layer (mainly related to the fruition of data by machine learning models) and to the APP Layer (mainly related to the interface of AI Layer functionalities with the clinical staff).

The following sections provide a survey that covers the main results and techniques available to reduce the amount of data required for the ML training. Such solutions involve data pre-processing, augmentation, and federated learning.

## 4. Discussion

### 4.1. Data Pre-Processing

In clinical settings, the first major problem faced when performing statistics or using AI systems is that of managing large amounts of data. The latter are often raw, heterogeneous, and come from different hardware technologies, aspects that make it difficult for a human operator to process them manually. The need to process huge amounts of data (Big Data) has quickly led to the development of a specific branch of mathematics and computer science known as data mining [56], whose main focus is to extract patterns from large and unstructured datasets. Thus, an attempt is made to bring order to the Big Data environment to create ordered datasets ready to be exploited for traditional statistical analysis or supervised and unsupervised learning algorithms [57]. The performance and quality of knowledge extracted by a data mining method in any framework depends not only on the method itself but also on the nature of the data itself. Factors such as noise, missing values, inconsistent and redundant data, and duplicates and multi-repeats strongly affect the performance of the algorithms used to learn and extract knowledge. Indeed, it has been widely demonstrated that low-quality data lead to low-quality knowledge and poor performance regarding artificial intelligence systems [58]. Thus, the importance of data pre-processing seems evident: this is a preliminary stage prior to data mining, the primary goal of which is to obtain final datasets that can be considered correct and useful for data mining algorithms and, subsequently, ML [57]. When few datasets are available, it is simple and easy to manually pre-process them, eliminating inconsistencies and possible measurement and/or transcription errors. Pre-processing Big Data constitutes a challenging task, as manual clean-up is not possible. Therefore, it is necessary to design data pre-processing strategies that can appropriately and sequentially filter out incorrect data.

Various data pre-processing techniques can be used to address a multitude of raw data issues, such as the ones listed below:

- **Missing data.** It is often the case that, during data acquisition, some features could not be recorded for several records. In this case, we speak of missing data, which therefore cannot be used as is for data mining. There are several approaches to deal with this problem, first of all eliminating all instances that are stumped [59]. However, this method results in a substantial thinning of the dataset itself, resulting in worse performance when (for example) training a neural network. Other more refined methods, such as the ones employed in this work, make use of statistical procedures such as maximum likelihood to model each feature, as if it were a random variable, and then include an estimate of the missing numerical values. In this case, the drawback lies in the “goodness-of-fit” of the random model, which may even deviate greatly from the true value of the stumped data [60].
- **Noise in the data.** Even when there are no missing data, those that are present are often characterized by imperfections and inaccuracies, mainly due to measurement errors committed by hardware instruments or, especially in the clinical setting, by human operators. In these cases, it is possible to intervene by constructing noise filtering or data polishing algorithms [61], with the purpose of automatic noise detection and the consequent choice of intervention, which may reside in the modification of the data or, in extreme cases, its elimination.
- **Feature selection.** In every dataset, each record is characterized by countless features, which often turn out to be useless for the purpose and/or redundant. Feature selection algorithms allow features with high correlation to be eliminated, an aspect that results in lowering the risk of overfitting and speeding up data mining processes [62].
- **Record selection or generation.** Shifting the discussion from features to records, high correlation can also occur between data [63]. In the medical domain, for example, it may be the case that two different patients have very similar medical records, and as a result, for some machine learning or data mining applications, knowledge of only one clinical status may suffice. Conversely, it may be the case that a data acquisition campaign has collected too few or too little variety, so that manual augmentation of the dataset itself is necessary. Such augmentation techniques will be discussed at length in the next section.
- **Class imbalance.** In many supervised learning applications whose goal is classification, there is a significant difference among the absolute number of records of different classes in a classification problem. This situation is known as class imbalance [64]. The presence of class imbalance leads to a net imbalance toward the majority class and, as a result, there is a higher rate of misclassification for all other classes, which are in the minority. In these cases, pre-processing techniques can be used to resample the data to balance the class distribution. Within resampling, two main groups can be distinguished [65]: (i) undersampling, through which majority instances are removed to create a balance, and (ii) oversampling, which creates a superset of the original dataset by replicating some instances or creating new instances from existing ones. The application of data pre-processing techniques in the medical context improves clinical interpretation, exempts healthcare personnel from manual data cleaning, reduces computational costs, and improves the performance of predictors [66]. The study conducted in [67] found that the use of pre-processing is most common in critical disciplines of medicine, such as cardiology, endocrinology, and oncology. Regardless of the relevant clinical field, the most commonly used techniques are data reduction due to noise and/or the absence of data, and feature selection.

#### 4.2. Data Augmentation

Deep convolutional neural networks constitute a widely used methodology for image classification in a variety of domains, including the clinical domain [68]. However, these networks are highly dependent on the presence of large amounts of data, which are neces-

sary to train the networks to correctly classify images in order to avoid the phenomenon of overfitting [69], which is the tendency of a neural network to fit the training data far too well, losing the generalization capabilities. In the medical field, very often, it is not possible to have access to a sufficient amount of data for the proper training of neural networks. In this respect, data augmentation practices may be crucial when dealing with clinical analysis [23,70].

In the literature, two main classes of methods for performing data augmentation are distinguished [17]. The first one belongs to methodologies based on direct image manipulation: in this family, we find methods such as algebraic–geometric transformations, color modifications, noise addition, image combination, and random deletions [17]. Although these methods are known to suffer from some drawbacks and often low performance gains that do not justify their use in some domains, traditional algebraic transformation methods also remain widespread in the medical field: in [31], the authors analyze five different methods belonging to the family of direct image manipulation for prostate cancer detection. The second class makes use of deep learning to perform data augmentation: examples include the feature space augmentation [24], adversarial training [25], and generative adversarial networks (GAN) [26] methods. The latter has been one of the most widely used approaches in clinical settings since its introduction in 2014. Such networks have the ability to generate new training data that allow, for the same classifier, significantly better classification performance than using the original dataset. The method of GANs has been applied to the study of electroencephalograms [27], the detection of brain tumors [28], and dermatological lesions [29]. GANs diverse in network structure are employed by the authors in [30] for risk stratification and disease classification in two rare clinical cases, namely, primary Sjögren’s syndrome and hypertrophic cardiomyopathy.

Although there are not many studies that perform a systematic comparison of the various data augmentation techniques, in [32] the authors show that the most effective methods are those based on geometric transformations (rotations, flips, and cutoffs) and GANs.

In recent years, very promising new methodologies, called meta-learning [33], are gaining popularity. These methods are based on the principle of neural architecture search [34], which uses neural networks to optimize other neural networks. Specifically, optimization mechanisms are employed to perform tuning of the hyperparameters of a neural network so as to maximize the accuracy of classifications. Data augmentation by meta-learning and optimization of neural networks are addressed in [35,36]. In these two papers, the use of a dual neural network cascade can be seen: the second is trained to optimize the parameters of the first.

In other cases, however, the optimization of network parameters is sometimes carried out by making use of a class of algorithms belonging to one of the three branches of machine learning, namely, reinforcement learning [71]. For example, AutoAugment [72] is a reinforcement learning algorithm that seeks an optimal data augmentation “policy” among a constrained set of geometric transformations with various levels of distortion. In reinforcement learning algorithms, a policy is analogous to the strategy of the learning algorithm. This policy determines what actions to take in given states to achieve a goal. The AutoAugment approach learns a policy that consists of many sub-policies, each consisting of a specific image transformation, namely, to associate a geometric transformation to the given data as input, i.e., the horizontal flip of an input image. Reinforcement learning is then used as a discrete search algorithm to perform data augmentation optimally. In another paper [73], the authors show how it is possible to improve the results of meta-learning by reinforcement learning by selecting transformations using, instead of the entire dataset, only small parts of it.

Meta-learning by reinforcement learning is a technique recently proposed in the literature and still little explored [74]. One of the major challenges concerns the design of meta-learning schemes dedicated to the specific application being considered: certain architectures may be valid for some clinical cases and prove unsuitable for others. Another limitation lies in the problem of gradient nulling [75], a very frequent phenomenon that



prevents the training of the network. Future developments, which are of interest to the CADUCEO project within the scope of the case studies defined therein, lie in the development of deep reinforcement learning algorithms [76] for the purpose of both solving optimization problems aimed at parameter tuning and selecting the data augmentation method most suitable for the specific clinical case study.

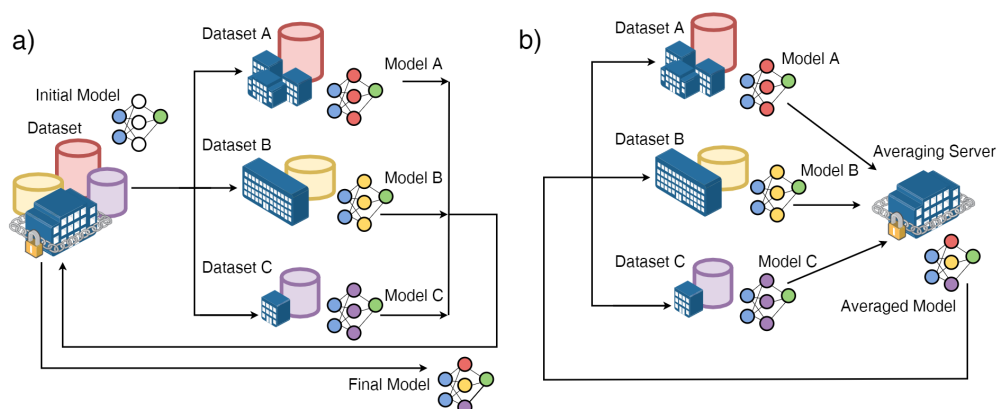
#### 4.3. Federated Learning

This section details the concept and architecture of federated learning schemes, emphasizing its importance in the medical field and the characteristics, in terms of GDPR compliance and Privacy Assurance, that make it an ideal tool in applications using clinical data—or data which are confidential in nature.

Federated learning (FL) constitutes a distributed machine learning solution employed for solving machine learning problems in the absence of a data exchange. First introduced in 2016 [59] as an alternative to conventional approaches for training deep models based on data from mobile devices, it was readily applied in the field of mobile keyboard prediction for GBoard [77].

The federated reference architecture for FL consists of a set of devices, also called clients, in communication with a server, that drives the entire learning process, which relies on local client data. In contrast to a typical distributed optimization scenario, where a central entity distributes data to peripheral ones and governs the training process, in the federated one there is no exchange of data, and training is based solely on the clients' data "as it is" (Figure 2); because the latter are strictly dependent on the use made by the device owner, they will be as follows:

- Non-IID (not independent and identically distributed), i.e., not representative of the statistical properties of the population identified for the problem under consideration;
- Unbalanced, from the point of view of numerosity, because the users do not behave in the same way;
- Highly distributed, because the number of devices participating in the training process may be greater than the total number of data available;
- Communication-limited, because they may not be accessible during training.



**Figure 2.** (a) Distributed optimization scenario, in which a centralized entity distributes data to peripheral entities and governs the training process; (b) federated optimization (FL) scenario, in which there is no data exchange. Image derived from [78].

Under the assumption that the server and the devices share the same neural network model (in terms of architecture), the FederatedAveraging (FedAvg) algorithm first chooses a subset of the clients that will participate in the training process, and then sends the model, with the given initial parameters, to the clients. The clients perform the training phase of the model based on their own local data—the training step—and send the resulting local model back to the server. The server updates its model by performing the weighted

average of the parameters of the local models received from the clients—the averaging step—and send the resulting model back to the clients. By iterating the two described steps, it can be shown that the model converges to a neural network with similar performances to a neural network centrally trained on the whole dataset.

The distributed nature and privacy protection, strengthened by the absence of a data exchange, of FedAvg make FL an ideal choice for medical applications, given the limitations imposed by the GDPR [79]. To date, FL is employed not only for predictive purposes based on information in electronic patient records, predicting hospitalizations for cardiac events [41], patient mortality [44], and ICU length of stay [80], but also in clinical research activities, such as the investigation of brain structural relationships [78], analysis of brain tissue from MRI images [81], and activities supporting clinical diagnostics, such as detection of COVID-19 from CT scans [82,83] or chest X-ray [84].

In addition to FL solutions developed ad hoc for the particular application under consideration, such as the previous ones, theoretical solutions have been proposed in the last period, aimed at enhancing FedAvg. Among them, FedProx [85] allows devices with optimization characteristics strongly different from others to be included in the optimization process, AdaFed [86] replaces the averaging mechanism of clients with a dynamic and adaptive heuristic weighting mechanism based on client performance, FOCUS [87] replaces the averaging mechanism with a credibility mechanism based on Shapley values [88], DecFedAvg [89] and CFA-GE [90] represent the first attempts to decentralize FedAvg by allowing each client to communicate with its neighbors, and FedLCon [91], the first consensus-based FL algorithm.

As part of the CADUCEO project, the interconnection of the participating hospitals in a federated manner is being evaluated, in such a way so as to increase the amount of data available, in full compliance with the privacy of patients, and then with current regulations, as well as the statistical representativeness of the data considered in relation to the use cases under consideration.

#### 4.4. Use Cases

This section describes the main use cases investigated for the design of the first iteration of the software tools developed in the scope of the CADUCEO project, highlighting their clinical features and sketching the ad hoc AI functionalities under development.

##### 4.4.1. Eosinophilic Esophagitis (EoE)

Eosinophilic esophagitis is a chronic inflammatory disease located in the esophagus, characterized by an elevated number of a particular type of white blood cells, called eosinophils, that cause its dysfunction [14]. It is one of the most frequent esophageal diseases in children and young adults, with an incidence of 20 new cases per 100,000 inhabitants per year.

The diagnosis and monitoring of this disease is based on the identification, by means of esophageal biopsies, of the concentration of eosinophils. The presence of eosinophils is detected by staining with haematoxylin-eosin.

Specifically, the diagnosis of EoE requires manual microscopic inspection of biopsies with a diagnostic threshold of at least 15 eosinophils in at least one high-magnification field. The protocol requires identifying the area of tissue with the greatest infiltration of eosinophils and quantifying the peak eosinophil count (PEC) to compare it with the mentioned diagnostic threshold, taking care to distinguish intact eosinophils, i.e., those with clearly visible cytoplasmic granules and nucleus, which contribute to the PEC, from non-intact eosinophils, which do not.

In addition, EoE monitoring is based on the inspection of endoscopic images looking for one or more macroscopic features, i.e., exudates, rings, edema, longitudinal sulci, stenosis, for example, in such a way so as to associate each of them with a value of the Endoscopic Reference Score (EREFS) [92].

Automatically counting the eosinophils together with staging the pathology on the basis of the EREFS score would obviously significantly speed up the diagnosis and treat-

ment processes. In both cases, GAN-based deep image inpainting techniques, such as [93] as well as Conditional GANs (cGANs) [94], are employed for data augmentation purposes; in the first case, the automated count fits in the general object detection branch of computer vision, with streamlined solutions such as YOLO [95] offering great opportunities in terms of performance and ease of training, while the EREFS-based classification is carried out by means of deep reinforcement learning algorithms, such as [96], in addition with XAI solutions, such as Grad-CAM [11]. Finally, FL methodologies, such as [91], are employed to increase the numerosity of training data allowing for the distributed training of the federation of hospitals, as well as for continual learning, namely, to prevent performance degradation over time.

#### 4.4.2. Inflammatory Bowel Disease (IBD)

Chronic inflammatory intestinal diseases are characterized by chronic inflammation of the intestine, due to an improper interaction between the intestinal microbiota, i.e., the set of symbiotic microorganisms that coexist with the human being without damaging them, and the mucosal immune system, i.e., the set of organs and cells specialized in defending the organism from external agents localized at the level of the intestinal mucosa. Depending on their characteristics and the intestinal tract affected, they can be divided into [15] the following:

- Ulcerative Rectocolitis (UCR), characterized by an inflammatory process localized exclusively at the level of the colon, which starts in the rectum and may extend continuously to the cecum;
- Crohn's disease (MC), characterized by inflammation that can affect the entire intestinal tract and involve all layers of the intestinal wall, leaving several healthy sections between the inflamed areas;
- Unclassified chronic inflammatory bowel disease, a term reserved for those cases in which anatomopathological features common to both UC and MC are present.

The main risks of such diseases are essentially twofold: on the one hand, we find that of resective surgery, i.e., the reasoned removal of a more or less long segment of the intestine, and, on the other hand, we find the development, after approximately ten years of disease, of dysplastic lesions (cancer) in the intestinal and colic mucosa.

To date, the only diagnostic tool is the evaluation of inflammation indices and the performance of instrumental examinations in a cyclic manner. Building a predictive model able to classify autonomously the staging of the disease in terms of the Geboes score [97], the Robarts histopathology index [98] on the basis of histopathology data, and in terms of the Mayo score [99], Ulcerative Colitis Endoscopic Index of Severity (UCEIS) [100], Harvey Bradshaw Index [101], and Simple Endoscopic Score for Crohn's Disease (SES-CD) [102] on the basis of endoscopic data would represent a step forward in terms of speeding up treatment and prognosis processes. To this end, GAN-based deep image inpainting methodologies, such as [103], deep reinforcement learning for classification, and FL are employed.

#### 4.4.3. Portal Hypertension (PH)

Portal hypertension (PH) is a pathological condition resulting from the evolution of chronic liver disease that results in an excessive increase in portal venous pressure [16], i.e., an increase in the pressure gradient between the portal vein and the inferior vena cava above the nominal value of 1–5 mmHg.

This increase leads to the formation of new collateral venous circles, called esophageal varices, in the esophagus and stomach, with a consequent increased risk of hemorrhage. While in pediatric age the most frequent form is pre-hepatic, caused by portal vein thrombosis, in adulthood the most frequent form is intrahepatic, commonly called cirrhosis.

Building an ML model able to estimate PH values on the basis of instrumental data, thus making a prediction about the prognosis of the disease, would be beneficial in terms of diagnosis and treatment processes. To this end, deep regression models, such as [104], will be employed, together with FL techniques for training and performance keeping purposes.

## 5. Conclusions

This paper has introduced the e-health project CADUCEO “Cloud pLatform for intelligent prevention and Diagnosis sUpported by artifiCial intelligEnce solutiOns”, which aims to design and develop novel machine learning based tools for assisting medical operators in improving the efficiency of today’s medical practice. The paper has presented the functional architecture designed by the project partners, which is centered around the idea of the federation of cooperating medical centers which, through federated learning algorithms, can share knowledge and improve the respective processes, without having to share sensitive medical data. Four key AI-based algorithms are developed and made available through the CADUCEO platform. They include algorithms for support to diagnosis through the automated analysis of medical images and other data; algorithms for automated counting of cells (as a relevant example of the contribution that AI can give to speed up the current, very time consuming tasks of medical practice); algorithms for image processing, focusing on the removal of undesired artifacts that might be present in the image, to make them usable; and algorithms for data augmentation, which allow to expand artificially the amount of data available for training the AI-based algorithms.

Finally, the three use cases that are the focus of the CADUCEO project have been presented, discussing the impact that the above AI-based tools can bring.

**Author Contributions:** Conceptualization, F.D.P., V.S., A.P. and A.D.G.; methodology, A.G., D.M., S.M. and F.B.; software, D.M. and M.M.; resources, F.D.P. and A.P.; data curation, M.P. and M.M.; writing—original draft preparation, A.G., A.D.G., L.L., V.S., S.M. and M.P.; supervision, F.D.P., S.M. and L.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has been partially carried out in the framework of the CADUCEO project (No. F/180025/01-05/X43), supported by the Italian Ministry of Enterprises and Made in Italy.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors acknowledge all the colleagues involved in the CADUCEO project, and in particular its coordinator Eng. Francesco Garofano.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Pantanowitz, L.; Quiroga-Garza, G.M.; Bien, L.; Heled, R.; Laifenfeld, D.; Linhart, C.; Sandbank, J.; Shach, A.A.; Shalev, V.; Vecsler, M.; et al. An artificial intelligence algorithm for prostate cancer diagnosis in whole slide images of core needle biopsies: A blinded clinical validation and deployment study. *Lancet Digit. Health* **2020**, *2*, e407–e416. [[CrossRef](#)]
- Smalley, E. AI-powered drug discovery captures pharma interest. *Nat. Biotechnol.* **2017**, *35*, 604–606. [[PubMed](#)]
- Gunduz, H. Deep learning-based Parkinson’s disease classification using vocal feature sets. *IEEE Access* **2019**, *7*, 115540–115551.
- Wen, Z.; Wang, S.; Yang, D.M.; Xie, Y.; Chen, M.; Bishop, J.; Xiao, G. Deep learning in digital pathology for personalized treatment plans of cancer patients. *Semin. Diagn. Pathol.* **2023**, *40*, 109–119.
- Wang, Y.; Nazir, S.; Shafiq, M. An overview on analyzing deep learning and transfer learning approaches for health monitoring. *Comput. Math. Methods Med.* **2021**, *2021*, 5552743 .
- Budd, S.; Robinson, E.C.; Kainz, B. A survey on active learning and human-in-the-loop deep learning for medical image analysis. *Med. Image Anal.* **2021**, *71*, 102062. [[CrossRef](#)]
- European Parliament; Council of the European Union. Regulation (EU) 2016/679 of the European Parliament and of the Council. *Regulation* **2016**, *679*, 2016 .
- McMahan, B.; Moore, E.; Ramage, D.; Hampson, S.; y Arcas, B.A. Communication-efficient learning of deep networks from decentralized data. *Artif. Intell. Stat.* **2017**, *54*, 1273–1282.
- Wang, L.; Zhang, X.; Su, H.; Zhu, J. A comprehensive survey of continual learning: Theory, method and application. *arXiv* **2023**, arXiv:2302.00487.
- Xu, F.; Uszkoreit, H.; Du, Y.; Fan, W.; Zhao, D.; Zhu, J. Explainable AI: A brief survey on history, research areas, approaches and challenges. In Proceedings of the Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019, Dunhuang, China, 9–14 October 2019; Proceedings, Part II 8; Springer: Berlin/Heidelberg, Germany, 2019; pp. 563–574.

11. Selvaraju, R.R.; Das, A.; Vedantam, R.; Cogswell, M.; Parikh, D.; Batra, D. Grad-CAM: Why did you say that? *arXiv* **2016**, arXiv:1611.07450.
12. Ribeiro, M.T.; Singh, S.; Guestrin, C. Why Should I Trust You?. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; ACM: New York, NY, USA, 2016. [[CrossRef](#)]
13. Agarwal, R.; Melnick, L.; Frosst, N.; Zhang, X.; Lengerich, B.; Caruana, R.; Hinton, G.E. Neural additive models: Interpretable machine learning with neural nets. *Adv. Neural Inf. Process. Syst.* **2021**, *34*, 4699–4711.
14. Daniel, N.; Larey, A.; Akin, E.; Osswald, G.A.; Caldwell, J.M.; Rochman, M.; Collins, M.H.; Yang, G.Y.; Arva, N.C.; Capocelli, K.E.; et al. A Deep Multi-Label Segmentation Network For Eosinophilic Esophagitis Whole Slide Biopsy Diagnostics. In Proceedings of the 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Glasgow, Scotland, UK, 11–15 July 2022. [[CrossRef](#)]
15. Ghimire, K.; Lai, W.; Omar, Y.; Schwebke, T.; Vo, J. Machine Learning Approach to Distinguish Ulcerative Colitis and Crohn’s Disease Using SMOTE (Synthetic Minority Oversampling Technique) Methods. *SMU Data Sci. Rev.* **2021**, *5*, 9.
16. Reiniš, J.; Petrenko, O.; Simbrunner, B.; Hofer, B.S.; Schepis, F.; Scoppettuolo, M.; Saltini, D.; Indulti, F.; Guasconi, T.; Albillos, A.; et al. Assessment of portal hypertension severity using machine learning models in patients with compensated cirrhosis. *J. Hepatol.* **2023**, *78*, 390–400. [[CrossRef](#)]
17. Shorten, C.; Khoshgoftaar, T.M. A survey on image data augmentation for deep learning. *J. Big Data* **2019**, *6*, 1–48. [[CrossRef](#)]
18. Bagherinezhad, H.; Horton, M.; Rastegari, M.; Farhadi, A. Label refinery: Improving imagenet classification through label progression. *arXiv* **2018**, arXiv:1805.02641.
19. Jurio, A.; Pagola, M.; Galar, M.; Lopez-Molina, C.; Paternain, D. A comparison study of different color spaces in clustering based image segmentation. In Proceedings of the International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Dortmund, Germany, 28 June–2 July 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 532–541.
20. Moreno-Barea, F.J.; Strazzer, F.; Jerez, J.M.; Urda, D.; Franco, L. Forward noise adjustment scheme for data augmentation. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 18–21 November 2018; pp. 728–734.
21. Inoue, H. Data augmentation by pairing samples for images classification. *arXiv* **2018**, arXiv:1801.02929.
22. Zhong, Z.; Zheng, L.; Kang, G.; Li, S.; Yang, Y. Random erasing data augmentation. *Proc. Aaai Conf. Artif. Intell.* **2020**, *34*, 13001–13008.
23. Abdollahi, B.; Tomita, N.; Hassanpour, S. Data augmentation in training deep learning models for medical image analysis. In *Deep Learners and Deep Learner Descriptors for Medical Applications*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 167–180.
24. Konno, T.; Iwazume, M. Icing on the cake: An easy and quick post-learnig method you can try after deep learning. *arXiv* **2018**, arXiv:1807.06540.
25. Zajac, M.; Zołna, K.; Rostamzadeh, N.; Pinheiro, P.O. Adversarial framing for image and video classification. *Proc. Aaai Conf. Artif. Intell.* **2019**, *33*, 10077–10078. [[CrossRef](#)]
26. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [[CrossRef](#)]
27. Lashgari, E.; Liang, D.; Maoz, U. Data augmentation for deep-learning-based electroencephalography. *J. Neurosci. Methods* **2020**, *346*, 108885.
28. Nalepa, J.; Marcinkiewicz, M.; Kawulok, M. Data augmentation for brain-tumor segmentation: A review. *Front. Comput. Neurosci.* **2019**, *13*, 83. [[CrossRef](#)] [[PubMed](#)]
29. Bissoto, A.; Valle, E.; Avila, S. Gan-based data augmentation and anonymization for skin-lesion analysis: A critical review. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 1847–1856.
30. Pezoulas, V.C.; Grigoriadis, G.I.; Gkois, G.; Tachos, N.S.; Smole, T.; Bosnić, Z.; Pičulin, M.; Olivotto, I.; Barlocco, F.; Robnik-Šikonja, M.; et al. A computational pipeline for data augmentation towards the improvement of disease classification and risk stratification models: A case study in two clinical domains. *Comput. Biol. Med.* **2021**, *134*, 104520.
31. Hao, R.; Namdar, K.; Liu, L.; Haider, M.A.; Khalvati, F. A comprehensive study of data augmentation strategies for prostate cancer detection in diffusion-weighted MRI using convolutional neural networks. *J. Digit. Imaging* **2021**, *34*, 862–876. [[CrossRef](#)]
32. Shijie, J.; Ping, W.; Peiyi, J.; Siping, H. Research on data augmentation for image classification based on convolution neural networks. In Proceedings of the 2017 Chinese Automation Congress (CAC), Jinan, China, 20–22 October 2017; pp. 4165–4170.
33. Hospedales, T.; Antoniou, A.; Micaelli, P.; Storkey, A. Meta-learning in neural networks: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *44*, 5149–5169. [[CrossRef](#)] [[PubMed](#)]
34. Zoph, B.; Le, Q.V. Neural architecture search with reinforcement learning. *arXiv* **2016**, arXiv:1611.01578.
35. Lemley, J.; Bazrafkan, S.; Corcoran, P. Smart augmentation learning an optimal data augmentation strategy. *IEEE Access* **2017**, *5*, 5858–5869. [[CrossRef](#)]
36. Wang, J.; Perez, L. The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Netw. Vis. Recognit* **2017**, *11*, 1–8.

37. Nguyen, D.C.; Pham, Q.V.; Pathirana, P.N.; Ding, M.; Seneviratne, A.; Lin, Z.; Dobre, O.A.; Hwang, W.J. Federated learning for smart healthcare: A survey. *ACM Comput. Surv. (CSUR)* **2022**, *55*, 1–37.
38. Gunter, T.D.; Terry, N.P. The emergence of national electronic health record architectures in the United States and Australia: Models, costs, and questions. *J. Med. Internet Res.* **2005**, *7*, e3. [[CrossRef](#)]
39. Hao, M.; Li, H.; Xu, G.; Liu, Z.; Chen, Z. Privacy-aware and resource-saving collaborative learning for healthcare in cloud computing. In Proceedings of the IEEE International Conference on Communications (ICC), Dublin, Ireland, 7–11 June 2020; pp. 1–6.
40. Vepakomma, P.; Gupta, O.; Swedish, T.; Raskar, R. Split learning for health: Distributed deep learning without sharing raw patient data. *arXiv* **2018**, arXiv:1812.00564.
41. Brisimi, T.S.; Chen, R.; Mela, T.; Olshevsky, A.; Paschalidis, I.C.; Shi, W. Federated learning of predictive models from federated electronic health records. *Int. J. Med. Inform.* **2018**, *112*, 59–67. [[CrossRef](#)] [[PubMed](#)]
42. Choudhury, O.; Gkoulalas-Divanis, A.; Salonidis, T.; Sylla, I.; Park, Y.; Hsu, G.; Das, A. Differential privacy-enabled federated learning for sensitive health data. *arXiv* **2019**, arXiv:1910.02578.
43. Wei, K.; Li, J.; Ding, M.; Ma, C.; Yang, H.H.; Farhad, F.; Jin, S.; Quek, T.Q.S.; Poor, H.V. Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 3454–3469.
44. Liu, D.; Miller, T.; Sayeed, R.; Mandl, K.D. FADL: Federated-autonomous deep learning for distributed electronic health record. *arXiv* **2018**, arXiv:1811.11400.
45. Shao, R.; He, H.; Liu, H.; Liu, D. Stochastic channel-based federated learning for medical data privacy preserving. *arXiv* **2019**, arXiv:1910.11160.
46. Arrieta, A.B.; Díaz-Rodríguez, N.; Ser, J.D.; Bennetot, A.; Tabik, S.; Barbado, A.; Garcia, S.; Gil-Lopez, S.; Molina, D.; Benjamins, R.; et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [[CrossRef](#)]
47. Zhang, Y.; Weng, Y.; Lund, J. Applications of Explainable Artificial Intelligence in Diagnosis and Surgery. *Diagnostics* **2022**, *12*, 237A.
48. Adadi, A.; Berrada, M. Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access* **2018**, *6*, 52138–52160. [[CrossRef](#)]
49. Islam, S.R.; Eberle, W.; Ghafoor, S.K.; Ahmed, M. Explainable artificial intelligence approaches: A survey. *arXiv* **2021**, arXiv:2101.09429.
50. Tjoa, E.; Guan, C. A Survey on Explainable Artificial Intelligence (XAI): Toward Medical XAI. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4793–4813. [[CrossRef](#)] [[PubMed](#)]
51. Kaur, R.; Kaur, R. Survey of de-noising methods using filters and fast wavelet transform. *Int. J. Adv. Res. Comput. Sci. Softw. Eng* **2013**, *3*, 133–136.
52. Miao, J.; Niu, L. A survey on feature selection. *Procedia Comput. Sci.* **2016**, *91*, 919–926.
53. Xiang, H.; Zou, Q.; Nawaz, M.A.; Huang, X.; Zhang, F.; Yu, H. Deep learning for image inpainting: A survey. *Pattern Recognit.* **2023**, *134*, 109046. [[CrossRef](#)]
54. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. *Adv. Neural Inf. Process. Syst.* **2014**, 27–32.
55. Arulkumaran, K.; Deisenroth, M.P.; Brundage, M.; Bharath, A.A. A brief survey of deep reinforcement learning. *arXiv* **2017**, arXiv:1708.05866.
56. Seifert, J.W. Data mining: An overview. *Natl. Secur.* **2004**, 201–217.
57. Garcia, S.; Ramírez-Gallego, S.; Luengo, J.; Benítez, J.M.; Herrera, F. Big data preprocessing: Methods and prospects. *Big Data Anal.* **2016**, *1*, 9. [[CrossRef](#)]
58. Pyle, D. *Data Preparation for Data Mining*; Morgan Kaufmann: Burlington, MA, USA, 1999.
59. Little, R.J.; Rubin, D.B. *Statistical Analysis with Missing Data*; John Wiley and Sons: Hoboken, NJ, USA, 2019; Volume 793.
60. Luengo, J.; García, S.; Herrera, F. On the choice of the best imputation methods for missing values considering three groups of classification methods. *Knowl. Inf. Syst.* **2011**, *32*, 77–108. [[CrossRef](#)]
61. Zhu, X.; Wu, X. Class Noise vs. Attribute Noise: A Quantitative Study. *Artif. Intell. Rev.* **2004**, *22*, 177–210. [[CrossRef](#)]
62. Hall, M.A. Correlation-Based Feature Selection for Machine Learning. Ph.D. Thesis, The University of Waikato, Hamilton, New Zealand, 1999.
63. Liu, H.; Motoda, H. On issues of instance selection. *Data Min. Knowl. Discov.* **2002**, *6*, 115.
64. López, V.; Fernández, A.; García, S.; Palade, V.; Herrera, F. An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Inf. Sci.* **2013**, *250*, 113–141. [[CrossRef](#)]
65. Mohammed, R.; Rawashdeh, J.; Abdullah, M. Machine Learning with Oversampling and Undersampling Techniques: Overview Study and Experimental Results. In Proceedings of the 2020 11th International Conference on Information and Communication Systems (ICICS), Irbid, Jordan, 7–9 April 2020. [[CrossRef](#)]
66. Mendes, D.; Paredes, S.; Rocha, T.; Carvalho, P.; Henriques, J.; Cabiddu, R.; Morais, J. Assessment of cardiovascular risk based on a data-driven knowledge discovery approach. In Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Milan, Italy, 25–29 August 2015. [[CrossRef](#)]

67. Idri, A.; Benhar, H.; Fernández-Alemán, J.; Kadi, I. A systematic map of medical data preprocessing in knowledge discovery. *Comput. Methods Programs Biomed.* **2018**, *162*, 69–85. [[CrossRef](#)] [[PubMed](#)]
68. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [[CrossRef](#)]
69. Ying, X. An Overview of Overfitting and its Solutions. *J. Phys. Conf. Ser.* **2019**, *1168*, 022022. [[CrossRef](#)]
70. Chlap, P.; Min, H.; Vandenberg, N.; Dowling, J.; Holloway, L.; Haworth, A. A review of medical image data augmentation techniques for deep learning applications. *J. Med. Imaging Radiat. Oncol.* **2021**, *65*, 545–563. [[CrossRef](#)]
71. Sutton, R.; Barto, A. Reinforcement Learning: An Introduction. *IEEE Trans. Neural Netw.* **1998**, *9*, 1054–1054. [[CrossRef](#)]
72. Cubuk, E.D.; Zoph, B.; Mane, D.; Vasudevan, V.; Le, Q.V. AutoAugment: Learning Augmentation Policies from Data. *arXiv* **2018**, arXiv:1805.09501.
73. Minh, T.N.; Sinn, M.; Lam, H.T.; Wistuba, M. Automated Image Data Preprocessing with Deep Reinforcement Learning. *arXiv* **2018**, arXiv:1806.05886.
74. Hu, M.; Zhang, J.; Matkovic, L.; Liu, T.; Yang, X. Reinforcement learning in medical image analysis: Concepts, applications, challenges, and future directions. *J. Appl. Clin. Med. Phys.* **2023**, *24*, e13898. [[CrossRef](#)]
75. Hochreiter, S. The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions. *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* **1998**, *6*, 107–116. [[CrossRef](#)]
76. Zhou, S.K.; Le, H.N.; Luu, K.; Nguyen, H.V.; Ayache, N. Deep reinforcement learning in medical imaging: A literature review. *Med. Image Anal.* **2021**, *73*, 102193. [[CrossRef](#)]
77. Hard, A.; Rao, K.; Mathews, R.; Ramaswamy, S.; Beaufays, F.; Augenstein, S.; Eichner, H.; Kiddon, C.; Ramage, D. Federated Learning for Mobile Keyboard Prediction. *arXiv* **2018**, arXiv:1811.03604.
78. Sheller, M.J.; Edwards, B.; Reina, G.A.; Martin, J.; Pati, S.; Kotrotsou, A.; Milchenko, M.; Xu, W.; Marcus, D.; Colen, R.R.; et al. Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data. *Sci. Rep.* **2020**, *10*, 12598. [[CrossRef](#)]
79. Viorescu, R. 2018 Reform Of Eu Data Protection Rules. *Eur. J. Law Public Adm.* **2017**, *4*, 27–39. [[CrossRef](#)]
80. Huang, L.; Shea, A.L.; Qian, H.; Masurkar, A.; Deng, H.; Liu, D. Patient clustering improves efficiency of federated machine learning to predict mortality and hospital stay time using distributed electronic medical records. *J. Biomed. Inform.* **2019**, *99*, 103291. [[CrossRef](#)]
81. Silva, S.; Gutman, B.A.; Romero, E.; Thompson, P.M.; Altmann, A.; Lorenzi, M. Federated Learning in Distributed Medical Databases: Meta-Analysis of Large-Scale Subcortical Brain Data. In Proceedings of the 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), Venice, Italy, 8–11 April 2019. [[CrossRef](#)]
82. Barstugan, M.; Ozkaya, U.; Ozturk, S. Coronavirus (COVID-19) Classification using CT Images by Machine Learning Methods. *arXiv* **2020**, arXiv:2003.09424.
83. Gozes, O.; Frid-Adar, M.; Greenspan, H.; Browning, P.D.; Zhang, H.; Ji, W.; Bernheim, A.; Siegel, E. Rapid ai development cycle for the coronavirus (COVID-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis. *arXiv* **2020**, arXiv:2003.05037.
84. Wang, L.; Lin, Z.Q.; Wong, A. COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci. Rep.* **2020**, *10*, 19549. [[CrossRef](#)]
85. Li, T.; Sahu, A.K.; Zaheer, M.; Sanjabi, M.; Talwalkar, A.; Smith, V. Federated optimization in heterogeneous networks. *Proc. Mach. Learn. Syst.* **2020**, *2*, 429–450.
86. Giuseppe, A.; Torre, L.D.; Menegatti, D.; Pietrabissa, A. AdaFed: Performance-based Adaptive Federated Learning. In Proceedings of the 2021 the 5th International Conference on Advances in Artificial Intelligence (ICAAI), Virtual, 20–22 November 2021; ACM: New York, NY, USA, 2021. [[CrossRef](#)]
87. Chu, W.; Xie, C.; Wang, B.; Li, L.; Yin, L.; Zhao, H.; Li, B. FOCUS: Fairness via Agent-Awareness for Federated Learning on Heterogeneous Data. *arXiv* **2022**, arXiv:2207.10265.
88. Roth, A.E. Introduction to the Shapley value. In *The Shapley Value*; Cambridge University Press: Cambridge, UK, 1988; pp. 1–28. [[CrossRef](#)]
89. Giuseppe, A.; Manfredi, S.; Menegatti, D.; Pietrabissa, A.; Poli, C. Decentralized Federated Learning for Nonintrusive Load Monitoring in Smart Energy Communities. In Proceedings of the 2022 30th Mediterranean Conference on Control and Automation (MED), Vouliagmeni, Greece, 28 June 2022–1 July 2022. [[CrossRef](#)]
90. Savazzi, S.; Nicoli, M.; Rampa, V. Federated Learning with Cooperating Devices: A Consensus Approach for Massive IoT Networks. *IEEE Internet Things J.* **2020**, *7*, 4641–4654. [[CrossRef](#)]
91. Giuseppe, A.; Manfredi, S.; Pietrabissa, A. A Weighted Average Consensus Approach for Decentralized Federated Learning. *Mach. Intell. Res.* **2022**, *19*, 319–330. [[CrossRef](#)]
92. Ahuja, N.; Weedon, J.; Schwarz, S.M.; Sklar, R.; Rabinowitz, S.S. Applying the Eosinophilic Esophagitis Endoscopic Reference Scores (EREFS) to Different Aged Children. *J. Pediatr. Gastroenterol. Nutr.* **2020**, *71*, 328–332. [[CrossRef](#)] [[PubMed](#)]
93. Jain, J.; Zhou, Y.; Yu, N.; Shi, H. Keys to Better Image Inpainting: Structure and Texture Go Hand in Hand. In Proceedings of the WACV, Waikoloa, HI, USA, 2–7 January 2023.
94. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.

95. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788.
96. Lin, E.; Chen, Q.; Qi, X. Deep reinforcement learning for imbalanced classification. *Appl. Intell.* **2020**, *50*, 2488–2502. [[CrossRef](#)]
97. Jauregui-Amezaga, A.; Geerits, A.; Das, Y.; Lemmens, B.; Sagaert, X.; Bessissow, T.; Lobatón, T.; Ferrante, M.; Assche, G.V.; Bisschops, R.; et al. A Simplified Geboes Score for Ulcerative Colitis. *J. Crohn's Colitis* **2016**, *11*, 305–313. [[CrossRef](#)]
98. Peyrin-Biroulet, L.; Arenson, E.; Rubin, D.T.; Siegel, C.A.; Lee, S.; Laroux, F.S.; Zhou, W.; Finney-Hayward, T.; Gonzalez, Y.S.; Shields, A.L. A Comparative Evaluation of the Measurement Properties of Three Histological Indices of Mucosal Healing in Ulcerative Colitis: Geboes Score, Robarts Histopathology Index and Nancy Index. *J. Crohn's Colitis* **2023**, *1*, jjad087. [[CrossRef](#)]
99. Naegeli, A.N.; Hunter, T.; Dong, Y.; Hoskin, B.; Middleton-Dalby, C.; Hetherington, J.; Stefani-Hunyady, D.; Canavan, J.B. Full, Partial, and Modified Permutations of the Mayo Score: Characterizing Clinical and Patient-Reported Outcomes in Ulcerative Colitis Patients. *Crohn's Colitis 360* **2021**, *3*, otab007. [[CrossRef](#)]
100. Ikeya, K.; Hanai, H.; Sugimoto, K.; Osawa, S.; Kawasaki, S.; Iida, T.; Maruyama, Y.; Watanabe, F. The Ulcerative Colitis Endoscopic Index of Severity More Accurately Reflects Clinical Outcomes and Long-term Prognosis than the Mayo Endoscopic Score. *J. Crohn's Colitis* **2015**, *10*, 286–295. [[CrossRef](#)]
101. Best, W.R. Predicting the Crohn's disease activity index from the harvey-bradshaw index. *Inflamm. Bowel Dis.* **2006**, *12*, 304–310. [[CrossRef](#)]
102. Koutroumpakis, E.; Katsanos, K. Implementation of the simple endoscopic activity score in crohn's disease. *Saudi J. Gastroenterol.* **2016**, *22*, 183. [[CrossRef](#)]
103. Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; Aila, T. Analyzing and improving the image quality of stylegan. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 8110–8119.
104. Lathuilière, S.; Mesejo, P.; Alameda-Pineda, X.; Horaud, R. A comprehensive analysis of deep regression. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *42*, 2065–2081. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.