Deep Convolutional Feature Extraction and Brain Morphometry Intelligence: AI-Powered Computational Solutions for Enhanced Neuroimaging Analysis and Interpretation

Dr. Javad Salehi, Professor of Electrical Engineering, University of Tehran, Iran

1. Introduction to Neuroimaging and Machine Learning

Neuroimaging and machine learning are two cutting-edge, exciting fields that have the ability to reshape our world over the next decades. A conflux of these two fields is particularly appealing because of the immense need and potential of tech-enabled neuroimaging applications. Neuroimaging refers to a series of noninvasive neurophysiological techniques used in medicine to visualize the structure, function, and the tiniest level of detail of the brain. It forms the foundation of neuroimaging research, which in turn guides advances in psychology, psychiatry, neurology, and beyond. Precise image analysis is crucial to the success of any neuroimaging project. This has been done manually or semi-automatically for decades. Machine learning techniques can provide a better alternative for far greater accuracy, less error, and greater speed than manual intervention.

The increasing use of AI-powered neuroimaging solutions reflects the growing importance of this technology in the industry. A well-known example is an AI for accurately interpreting eye scans. The artificial intelligence arm has developed an AI algorithm 25 times more accurate at interpreting optical coherence tomography scans of the eye than fully trained ophthalmologists. The algorithm, which can help identify patients that need life-changing treatment before irreversible deterioration, detected 50 different eye diseases as well as ophthalmologists. A second example is that of a project commissioned to provide insight into the progression of visual and neurodegenerative disorders, with potential applications for glaucoma, age-related macular degeneration, and Alzheimer's. As a subset of AI, deep learning is a particularly powerful tool for extracting useful patterns from neuroimaging data. However, the methods must be sensitive to neuroimaging data-specific particulars and challenges, and the interpretation of deep learning models should also be carefully considered.

1.1. Overview of Neuroimaging Techniques

Introduction Images of the living flesh of our brain are often seen on TV or on the web. This conjures an aura of beauty, amazing features, and is often related to specific abilities, disabilities, or even geniality. But what can neuroimaging really tell us in a scientifically sound way? Before discussing the principles of machine learning in the context of neuroimaging, let's provide a general overview of the so-called neuroimaging techniques.

1.1. Overview of Neuroimaging Techniques Two contemporary techniques for obtaining 3D images of the brain are MRI and CT, which are particularly used to investigate lymphoma detection or cerebrovascular pathology. This type of scan makes use of X-rays: they provide a rich contrast in tumors because pathological tissue lacks healthy neurons and white matter, which attenuate X-rays at much slower rates. The CT is a quick technique, and its repetition is allowed even some days later. The images obtained with the MRI scanner are of higher quality, and it is the most used system for neuroimaging. Many techniques are available to improve contrast within a given image, which makes MRI a very flexible tool for different purposes. Depending on the image properties used for imaging, at least four classic types of MRI techniques are known to date: relaxation-weighted MRI, flow-weighted MRI, nuclear spectroscopy, and diffusion-weighted MRI.

Positron emission tomography can be used to label radioactive isotopes or tracers, which over time are selectively retained from different cellular or molecular compartments into the blood. Signals from the labeled emission are then accumulated post-acquisition to build a tomographic image of the acquired radioactive traits, either intracellular or extracellular, naturally occurring or stimulated by antecedent interventions such as brain activation, also by using, for example, mental tasks and cognitive stimuli. All different imaging tools have their strengths and some limitations; thus, a possible strategy could be ionizing radiation in clinical practice and the use of the best system for a given purpose in the neuroscience field. A small note has to be made regarding cognitive applications of brain dysfunctions. All of the above-described neuroimaging techniques allow one to reveal how the brain looks at a given moment. With a time resolution sufficient to explore the mind's non-stop changes over time, using advanced methods of the same type, it was possible to reveal the brain structure related to functions, possibly using refined MRI imaging systems and machine learning algorithms to create detailed graphs of the structural network defining the different brain regions. Diffusion tensor imaging was developed to extend the fMRI studies, which were derived from MRI, with the ability to image white tissue tracts in vivo.

1.2. Fundamentals of Machine Learning in Neuroimaging

Neuroimaging has become an essential biotechnology for the assessment of the healthy and pathological brain in human and animal populations. Analyses of neuroimaging datasets demand complex computations that are often addressed using artificial intelligence in the form of machine learning. Some machine learning algorithms include supervised learning, where algorithms or models are trained from labeled data, and unsupervised learning, covering clustering, dimensionality reduction, density estimation methods, and algorithms. In the literature, many terms have been employed to describe machine learning tasks in neuroimaging. Building a comprehensive guide for convention will be beneficial. Neuroimaging holds unique challenges compared to other datasets, and it is well recognized for the importance of robust data representation. In the image processing tasks of machine learning, features typically refer to patches or regions resulting from image segmentation pipelines that have high discriminative patterns. Automatic and unsupervised techniques, such as in convolutional neural networks, often concurrently learn this on the fly. Accurate identification of true patterns, crucial for clinical applications, is challenged by the potential for mass irregularities in non-imaged data.

The concept of overfitting, applicable to many machine learning and statistical models, can be tailored to the domain of neuroimaging in some models. The tight tuning of model hyperparameters and structure for the noise of distinct scans may result in our model being more suitable to the struck-through noise of the images analyzed than is ideal for adequate generalizability. For example, overfitting of ICT versus non-ICT brain pathology may not be perfectly generalizable to a completely novel scanner type. For compatibility with established statistical frameworks, the identification of overfitting and the incorporation of regularized estimation is a recommended practice. As computational approaches to diagnostics have become more complex, collaborations to integrate diverse fields have become more common. Many computational tools involved in deep learning applications are shared across different software programs. Numerous software programs specifically designed to handle neuroimaging data analytics and machine learning, whether as individual programs or contained within scripting languages, have been developed. In-depth coverage of developing these scripts and APIs is beyond the scope of this paper. Standardization of tools forms a desirable endpoint when developing a systematic approach to accelerate diagnostic discovery in psychiatry.

2. Challenges in Neuroimaging Analysis

The analysis of neuroimaging data, such as MRI and positron emission tomography, suffers from various challenges that make the preprocessing particularly complex. The aim of preprocessing is to transform the raw data to reveal specific features or to enhance the quality of the data to increase the quality of the outcome. One of the major challenges in preprocessing is the complexity of MRI data, where images need to be corrected for different types of artifacts, such as echo-planar imaging distortions. Moreover, there are other types of artifacts like random noise or scanner-derived artifacts, and differences in scanning parameters across different sites that need to be addressed as part of the preprocessing. The difficulty of the data preprocessing would, in turn, imply that the disentangled variation, also known as weak signals from the data, may also represent data samples as noise, leading to the variation of interest being destroyed during the preprocessing.

Moreover, the various sources of noise and artifacts make the preprocessing even more complicated. Multiplicative noises and artifacts would be introduced, and some of the weak signals may also be regarded as errors, which would be removed from the data. Modern machine learning models have the potential to learn from heterogeneous multimodal data and offer more accurate diagnosis, prognostication, and treatment indication based on neuroimaging data. These models can help inform diagnosis and personalized treatment plans. Machine learning has the potential to offer more accurate and faster analysis. However, the main challenges of interpretability and explainability have been well documented. A model may be offering great insights, but its internal workings can be highly intricate and impossible to understand in any meaningful way. Thus, unexplainable AI solutions might be a no-go in sensitive medical contexts, including the analysis of neuroimaging data. To date, the majority of AI models have been evaluated on site and population-specific training data, which constrains the models' generalizability. This has led to ethical considerations around the clinical utility of neuroimaging AI solutions.

2.1. Data Preprocessing and Quality Control

Changes in neuroimaging data representation of brain function and activity have a wide range of applications, including diagnosis, prediction, and meta-analysis. However, the lack of standardized data processing workflows can have a negative impact on the readouts of machine learning algorithms. This subsection presents a range of data preprocessing and quality control steps, showcasing that despite these challenges, there are common procedures for the preprocessing of neuroimaging data across different fields. Data Preprocessing: Key Steps Noise Reduction and Artifact Correction Functional neuroimaging data composition can be vastly different from other data types. As such, before even beginning to process the acquired readings, there must be a brief consideration of the type of data likely to be given. Secondly, how is that data stored? And finally, how is this data presented? Data Annotation An expert in the field relating to the data at hand should be consulted when considering initial annotations for the data. The key to effectively combining datasets or creating storable outputs is to use metadata. In the medical data industry, such data is normally regulated and stored securely in a patient's record on the server. This metadata provides the train of data necessary to produce exactly the same output if reprocessing is needed. Data Normalization or Scaling

2.2. Interpretability and Explainability of AI Models

In many clinical scenarios, any diagnostic decision taken based on AI predictions may affect the quality of patient care, requiring a certain level of mental and professional responsibility. In these settings, undesirable model behavior and predictions cannot be justified solely by stating good model performance, but should be explained. Model interpretability, an active field of AI research, addresses the extent to which a human can understand how a model makes its predictions. Neural network models trained on neuroimaging data often operate as black boxes, leaving users, especially clinicians, unable to provide clinical rationale for predictions. Technically, if AI models are directly applied to patient cohort analysis or treatment recommendations, it is recommended that they capture relevant features of the underlying neurobiology and can be interpreted to support the respective clinical decision. The need for interpretable models has been recognized in many clinical and biomedical applications to promote trust and acceptance among practitioners.

Methods that address model interpretation focus on normal and challenging prediction cases and indicate the input features relevant for the model's decision. For instance, while some imaging biomarkers or phenotypes need further studies for a reliable association with pathophysiological changes, they may help to support clinicians' decisions as explainable AI features. One such method is the framework called Local Interpretable Model-agnostic Explanations that fits a linear model to the output of the last layer before the final prediction for artificially perturbed samples of original data; another relies on Shapley additive explanations to weight evidence of related states regarding the target variable of interest. When evaluating deep learning models in clinical applications, it is recommended to balance accuracy, interpretability, and the level of agreement between them as well. As AI technologies progress from clinical trials to real-world settings, it is expected that the demand for transparent AI models will also increase.

3. AI Applications in Neuroimaging

Neuroimaging or image-related data that reflect brain structures or functions are used in various research fields to observe differences arising from diverse contexts, including the presence of neurological diseases. Anatomically, the brain appears as a complex structure of gray and white matter, made up of a large number of neurons and neuroglia, and surrounded by complicated structures as well as the meninges. This complex structure is supported by a rich vascular system, which provides adequate blood flow and energy support. A large number of synapses in the brain contribute to nuanced responses to sensory, cognitive, or emotional events, and such complex systems generate electrophysiologic activities that can be recorded and chemically released neurotransmitters that can affect other parts of the brain. It is challenging but meaningful to find commonalities or rules from such complex phenomena, and related clinical questions and predictions are matters of paramount importance.

An AI technique can be applied to analyze such complex data and find any possible rules or commonalities. A typical application is found in a technique to automatically segment the brain regions from neuroimaging data, which is called brain structure segmentation. Once the brain regions are extracted, one can find harmony in each brain system, e.g., between the gray matter density of the amygdala and subjects' anxiety levels. Such a technique also provided clinical improvements in diagnosing and tracking neurological diseases. For an in-depth trial, one could find substantially more possible applications. For example, another AI technique can be helpful to predict disease progression from brain cancer for all time points with poor prognosis or predict monthly future cancer from initial fMRI using a deep neural network. In other words, AI and neuroimaging aim to link the complex brain data that can be measured via imaging with a broad range of issues. They aim to identify links between brain imaging data, cognitive function, specific behaviors, and more, across large samples of individuals or samples from precise demographic groups.

3.1. Segmentation of Brain Structures

Accurate segmentation is key to neuroimaging, as it allows the delineation of the various anatomical regions that compose the brain and whose definition is provided in volumetric images. Consequently, it is at the heart of many imaging studies. Manual segmentation is common in preclinical and clinical imaging research, as it is considered the ground truth. Manual delineation is, however, time-consuming and rater-dependent segmentation; therefore, it is not a feasible strategy for segmenting the entire brain, particularly in large-scale, multi-centric studies. AI-driven image segmentation has gained momentum in recent years due to its improved efficiency and accuracy. Convolutional neural networks are the most commonly used architecture and provide state-of-the-art results. Despite the strong performance of AI-driven methods, some challenges are still limiting the field, including the fact that brain anatomies are highly variable across patients.

Further, the design of AI methods typically requires very large annotated datasets of patient images as inputs. While traditional methods have the considerable overhead of visualization of machine learning-based prediction errors, AI-driven methods are praised for allowing the generation of continuous visual output that aims at luring the eye of radiologists. Their aim is to help radiologists confirm their findings and aid the reader in understanding the diagnosis. Successful AI implementations in the clinical workflow have been reported, such as for volumetric hippocampus segmentation.

Segmentation is an essential step, as it provides a 3D representation of the anatomical abnormalities, allowing for a direct visualization of the volume of the lesions, potentially improving diagnostic accuracy and treatment planning. The overall purpose of CNN-based automated segmentation is to introduce wider practical applications to the imaging research field, improving anatomical visualization.

3.2. Functional Connectivity Analysis

Functional connectivity analysis. Functional connectivity is broadly defined as the temporal correlation between spatially remote brain regions. The objective of FC analysis is to uncover intrinsic temporal correlations between distinct neuroanatomical regions and identify putative functional brain networks. FC quantification is instrumental in generating and interpreting graph-based models of the human connectome and provides valuable insights into sensorimotor integration patterns across multiple brain regions. Additionally, resting-state fMRI is a sensitive tool for probing large-scale brain networks and has been shown to be a promising potential diagnostic test in several pathological conditions. Aberrant FC estimates might contribute to evidence-based disease diagnosis or overshadow standard behavioral measures. However, despite the significant growth in methods and applications, there are unresolved issues in choosing a specific feature selection strategy for FC data and for interpreting the complex network effects.

A considerable amount of work has been dedicated to providing novel bioinformatics and medical informatics criteria for AI-driven FC features and to highlight the role that FC plays in the pathophysiology and molecular underpinning of several neurological and psychiatric conditions. Several investigations on rs-fMRI data have focused on incorporating machine learning algorithms that reveal the optimal arrangement of regions and edge weights that can better discriminate effectively between groups. FC-based research has mainly exploited graph characteristics, such as topological thresholding procedures to convert correlation matrices or adjacency matrices into a binarized version or perform embedding procedures to reduce the dimensionality-driven challenging effects into a lower representation of weighted connectivity data. However, it is important to emphasize that no consensus can be reached on how to define the functional connectivity properties across subjects when using complex statistical approaches. Therefore, most studies across different disease states have treated the mathematical features derived from FC as purely data-driven, but this can be interpreted differently based on the results of related works.

4. Developing Machine Learning Models

Developing powerful machine learning models for neuroimaging analysis is a multi-step process that requires a range of neuroimaging and machine learning expertise. The first indispensable step is to perform data collection, which involves collecting and then anonymizing a sufficient number of samples representative of the targeted clinical application, ensuring that the quality of the neuroimaging data is good enough to describe the targeted brain property. With a few exceptions, data used for training such models requires annotation. This involves an exhaustive review of the data by a skilled clinician or neuroscientist, in order to guide the learning algorithm with human expertise and to quantitatively define a ground truth.

The development of machine learning models consists of several main phases: i) data curation; ii) model building; iii) performance estimation; iv) model interpretation. Each phase is generally iterative, and many decisions may have to be revised over time. First, data is usually divided into a training set, a validation set, and a test set. These phases may lead to the modification of data curation or data collection to enlarge datasets or to ensure the representativity of samples. While the massive development of data-driven models has led to an explosion in the number of machine learning models over recent years, many challenges have to be tackled to maximize the effectiveness of such models. Model selection is still a challenge in the field, as no consensus on the best approach to be used has yet emerged. Nonetheless, model selection will strongly depend on the targeted application in neuroimaging as well as on the characteristics of the dataset. Some tasks and some datasets will favor non-supervised models, like neural networks with a pre-trained backbone, while others will require contrasting several non-supervised models. Since most classifiers rely on at least one hyperparameter, searching for the best model setup becomes critical to ensure satisfying classification performance. Model optimization or hyperparameter tuning comprises various techniques such as grid search, random search, or Bayesian optimization. To avoid reporting unfair performance, the chosen hyperparameters of the selected model should be based on the validation set. Including the test set in an iterative model optimization and evaluation will lead to biased and over-optimistic results. Reporting all model parameters, as well as the model evaluation on unseen data, enables reproducibility and limits bias. Reproducibility may still be further enhanced by publicly sharing data to enable the community to reproduce, retrain, and re-evaluate the proposed model. Finally, it is crucial to ensure the generalizability of the model across diverse and large clinics, scanners, or acquisition protocols for future use in clinical practice. In order to fill the gap between research and clinical applications, many works are currently aiming at ensuring the cross-site generalization of deep learning-based model performance.

4.1. Data Collection and Annotation

Data collection and curation is a time-consuming and expensive process, but also the backbone of any machine learning system. It is essential to collect a high-quality and diverse dataset that accurately reflects the target population. Electronic Health Record systems, as well as various clinical databases, can be mined to collect sufficient data. Nevertheless, it is important to underline that the use, storage, and handling of personal data for the purposes of scientific investigation must comply with national and international legal provisions. Notably, the regulations outline the principles and requirements for processing personal data. Although there is room for member state-specific complementations, it is key to consider that the regulations are clear that personal data processing activities require a legal basis, one of which can be in the public interest. Thus, research ethics committees should be contacted for necessary ethical approval in order to be certain of any needed administrative licenses. Furthermore, other constraints like consent of the participants, sociodemographics of the patient population, and data privacy laws might affect data collection and analysis.

For the data itself, the best data collection procedures would be prospective, patient-based, and multicentric in order to maximize its generalizability. Since manual or electronic patient charts are vulnerable to errors and bias due to misinterpretation or incorrect documentation of clinical findings, the creation of a ground truth by annotation is necessary to increase the reliability of downstream analyses. Data annotation typically involves labeling regions of interest or landmarks. Participants are trained in the use of standardized protocols and a pre-defined lexicon to minimize discrepancies. After observation, a second person validates the tracking procedure, or the output is independently reviewed by another person. The goal is to ensure that the annotation is correct with respect to the ground truth and adequate references. The inter-observer variability and the single observer variability are calculated as measures to assess the quality of the annotations.

4.2. Model Selection and Optimization

Model selection refers to the task of choosing which algorithm to use, given the problem and data characteristics at hand. Due to the complex nature of neurological data, non-exhaustive evaluation of various models could result in the misuse of the algorithm despite the availability of a more beneficial one. Therefore, practitioners should test the performance of multiple algorithms and select a trade-off between model predictive accuracy and interpretability. Techniques such as cross-validation, inverse validations, or holdout subsets are often used to evaluate the usefulness of a given model. The improved availability of computational resources has also made optimization over model hyperparameters more feasible. One-dimensional hyperparameter space search methods such as grid search are easy and do not require the need for any matheuristic optimizations to fine-tune model performance.

By choosing the optimal model automatically, we accept that the resulting suggestions, insights, or conclusions, including the possibilities of interpreting the data, are provided automatically. But to obtain a useful induced model, sometimes a model that offers reliable, complete, and insightful advice has to be traded off. When training a machine learning algorithm, the number of options one has concerning how well the model can tune to the actual noise of the data is limitless. Therefore, it is important to be aware of possible overfitting and underfitting during the interpretation process. The occasional practical trade-off that leads to the generation of more interpretable, less accurate models because of reliability should not be ignored.

Selecting an appropriate metric should come down to practical matters, but it is important to remember that sometimes the equilibrium could lean more towards accuracy and sometimes towards interpretability. Additionally, to successfully introduce currently available AI algorithms and models to the field of neuroimaging, these models have to be evaluated, including checking robustness and the necessity of preprocessing, and that a minimum of 10-50 subjects' datasets can be used to train the model.

5. Evaluating AI Models in Neuroimaging Analysis

Empirical evaluations are extremely important to demonstrate robust and reproducible machine learning predictions. This is particularly relevant in the health domain since identifying correct predictions can have critical consequences for patients. In the realm of neuroimaging analysis, there are several effective methodologies and standard metrics often used to evaluate machine learning models. Common metrics rely upon model output in the form of performance using receiver operating characteristic analysis, confusion matrix cross-tabulations, or related statistics such as forecast/actual data, or a variety of predictive measures including accuracy, sensitivity, and specificity. Furthermore, the evaluation methodology is based on the study of a specific group. The scope of testing should then focus on that population. Several robust models will aim to assess generalizability through external testing, where new observers are used for validation. In some situations, the dataset is split into numerous sections; traditionally, cross-validation has been employed where an algorithm is trained and validated multiple times using different combinations of the existing data. The performance statistics from each of these computations are then averaged to give a more valid estimate of the model's generalizability.

One of the preferred areas for AI to show potential beyond the performance of classical neuroimaging analyses is in diagnostic and prognostic prediction. The potential value of AI-powered analysis is particularly high when accurate diagnosis and/or prediction affect treatment and patient outcomes. In the field of neuroimaging, AI has shown potential where traditional methods have had problems. Furthermore, variations in a diagnostic protocol may encourage further exploration of an AI model's clinical value. Nevertheless, although many neuroimaging studies now use some form of AI, evaluation is suboptimal and based predominantly on model accuracy, leaving uncertainty over a model's robustness. Current metrics used to evaluate AI diagnosis and prediction are not fit for purpose in a clinical setting. As a result, in the changing medical landscape, new strategies are being investigated to validate possible novel predictors.

5.1. Performance Metrics for Model Evaluation

Performance metrics quantitatively measure the AI/ML model's predictive accuracy and can help clinicians understand how often a model makes errors, along with the types and frequencies of these errors. A typical evaluation in the AI world would involve calculating indicators such as precision, recall, F1 score, accuracy, specificity, true positive, false positive, false negative, and true negative. It is crucial to choose the proper performance metrics depending on the application or task, and the importance of sensitivity versus specificity, based on the potential clinical context and consequences. Precision and recall correspond to the model's capacity to detect the "positive class" correctly – diseased in this case – and avoid false negatives: these two metrics always represent a trade-off; for example, a higher precision could be obtained at the same time with lower sensitivity and vice versa. Differently from precision and recall, specificity addresses the ability of a model to identify the truly healthy subjects. The F1 score could be used by practitioners when the false positives and false negatives carry different costs. Since neuroimaging data analysis often involves data with imbalanced class distributions, it is preferable to investigate a combination of these performance indicators to reduce the risk of drawing wrong conclusions.

Confusion matrices could be used to evaluate the AI/ML model's performance on a more detailed level: they highlight the frequencies among the four metrics, measuring the number of true positives, false positives, false negatives, and true negatives that a given model predicted. These confusion matrix values can be used to compute the following statistics: sensitivity, specificity, precision, and negative predictive value. Fundamentally, these metrics are useful to investigate the AI/ML model's performance in the differentiation between individuals with a certain disease and individuals without it. Moreover, they also provide a measure of global performance. A high specificity model will therefore rarely misclassify any controls as affected, while a high sensitivity test will correctly identify all or most of the cases; for example, in the case of a health chatbot that screens for Alzheimer's, it is preferable to have high sensitivity in order to prevent as many false negatives as possible. It is important to take these parameters into account when a diagnostic assistant for cognitive and mental diseases is in the works.

5.2. Comparison with Traditional Methods

In this section, we compare AI methods with traditional approaches in neuroimaging. AI-driven models allow for significant improvements in terms of accuracy and processing time. For example, diagnosis accuracy in ischemic stroke can be boosted from 25% for manual three-dimensional tomography analysis up to 96% using deep learning for automated analysis. MRI volumetry reduces costs and processing time—from 8 hours to 10 minutes—to calculate brain volumes and lesions in the study. Moreover, AI provides the possibility of being highly precise without human interpretation, which is increasingly needed as technology evolves. Lastly, it is important to keep in mind that the AI model's output is highly influenced by design and data selection, requiring age for a better understanding of results and their direct applicability.

On the other hand, traditional methods such as visual MRI postprocessing by segmentation have several problems, as they depend on subjective interpretations by physicians. Consequently, one must know the limitations of these methods in order to use them complementarily to the AI analyses. Therefore, advantage should be taken of the best class of MRI to compare the gold standard obtained with the AI or verify new protocols, avoiding the cost of MRI. AI analysis can replace a high number of human operators in postprocessing, but they can help to equally reduce the complexity of automation. Furthermore, insight into the performance of deep learning is required to explore methodologies to decrease it.

6. Future Direction

Since neuroimaging techniques have continuously been improving, allowing for real-time, personalized, high-resolution, and high-dimensional imaging data to be acquired, it has the potential to be the future of personalized medicine. By pairing this improvement in computational power with deep learning, and by further linking it with electronic health records, in silico models, and so forth, the analysis of cognitive function can be performed in real-time in the context of where the subject is and what they are doing. Subsequently, the results could be pushed back onto a neurosurgeon's console or any clinician who is taking part in the functional recovery of subjects with brain injuries.

Nevertheless, there are several limitations and challenges that need to be addressed for the field to progress in a meaningful way. This will include the integration of neuroimaging-derived knowledge into interventional and adaptive closed-loop strategies, signing protocols to further harmonize procedures and best practices, and taking into account clinical workflows and operational workflows to guarantee that interventional protocols are feasible and competitive under clinical practice. Similarly, data protection, security, and ethical frameworks need to be well-established and comply with national and international regulations with intercultural and international adaptation. Importantly, as we progressively move towards these future directions, further interdisciplinary studies or collaborations are of great interest. It will also be necessary to have ongoing education in computational and big data workflows, including data acquisition, analytics, and artificial intelligence as a formal part of the medical school curriculum to facilitate clinical translation. We also predict that the role of AI will further embed itself within a specific imaging modality for the design of the most appropriate acquisition and/or analysis, and for the prediction of the response of an individual to a particular intervention.

7. Conclusion

This essay has demonstrated that artificial intelligence has the potential to improve the accuracy, robustness, and interpretability of neuroimaging research and has the capability to assist with screening, monitoring, and diagnosing subjects in pre-clinical research settings. There are, however, a number of technical, philosophical, and ethical challenges that will need to be addressed before the full potential of AI in this domain can be achieved: one of the biggest open challenges in this field is that there is no standardization of protocols with regard to handling of the data in the various processes from pre-processing to result analysis. So far, researchers compared models and solutions only in lab settings without making any effort to collaborate towards comparing the novel methodologies in scientific or clinical settings across the various research institutes and centers. Another open challenge is to build and deploy systems that provide explanations for decisions, insights, or discoveries in a transparent format. As neuroradiological imaging will continue to evolve and provide us with more detailed and accurate information, we need to offset the ethical and privacy implications by exploring solutions that have ethical considerations at their core, and show respect for privacy and autonomy of subjects within the neuro-imaging AI development pathway. In conclusion, neuroimaging analysis is a rapidly evolving field, and we hope to have shed some light on the potential AI and other technologies hold to radically enhance our collective understanding of how the brain functions, and the pathways implicated in neurogenerative disease. We issue a call for interdisciplinary research, fundamental collaboration across fields, and invite further research into this rapidly evolving domain. Progress in the integration of successful preclinical biomarkers for dementia or patient-specific examination by electroencephalography already gives us a glimpse of how these tools may be used in the future as part of routine healthcare and in the prevention of non-communicable diseases. The increased opportunities for combining phenotypic data from the digital domain with multiomics data on the same patients will create many more powerful tools for analysis and the delivery of personalized healthcare. Using virtual tumours to rapidly test multiple treatment combinations for nearly all patients has already become established practice, and this approach may also be useful for optimising the combinations of drugs used in resource-limited healthcare systems to improve patient outcomes.