



Multi-Task Deep Learning for Predicting Depression Severity and Treatment Response Using Synthetic Neuroimaging: A Fully Controlled Simulation Framework

Rocco de Filippis^{1*}, Abdullah Al Foysal²

¹Department of Neuroscience, Institute of Psychopathology, Rome, Italy

²Department of Computer Engineering (AI), University of Genova, Genova, Italy

Email: *roccodefilippis@istitutodipsicopatologia.it, niloyhasanfoysal440@gmail.com

How to cite this paper: de Filippis, R. and Al Foysal, A. (2025) Multi-Task Deep Learning for Predicting Depression Severity and Treatment Response Using Synthetic Neuroimaging: A Fully Controlled Simulation Framework. *Open Access Library Journal*, **12**: e13955.

<https://doi.org/10.4236/oalib.1113955>

Received: July 15, 2025

Accepted: August 19, 2025

Published: August 22, 2025

Copyright © 2025 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Major Depressive Disorder (MDD) remains a complex and debilitating psychiatric condition characterized by heterogeneous symptom profiles and substantial variability in treatment response. Despite extensive clinical research, the process of selecting effective pharmacological interventions for individual patients continues to rely heavily on trial-and-error, leading to prolonged suffering and increased healthcare costs. Neuroimaging-based predictive models have shown potential to support precision psychiatry; however, their progress is limited by the scarcity of large, annotated clinical datasets, high imaging costs, and strict patient privacy regulations. In this study, we propose a comprehensive synthetic neuroimaging and clinical data simulation framework that enables the development and testing of deep learning models in a fully controlled environment. Specifically, we designed a multi-task Convolutional Neural Network (CNN) that simultaneously predicts depression severity as a continuous outcome and treatment response as a binary classification task. The synthetic data were generated with biologically plausible brain activation patterns, targeting clinically relevant regions such as the prefrontal cortex, cingulate cortex, amygdala, and hippocampus. Clinical variables, including age, gender, baseline severity, and treatment group, were integrated into the simulation to reflect real-world heterogeneity. The proposed model achieved a Mean Absolute Error (MAE) of 2.84 for severity prediction and an Area Under the Curve (AUC) of 0.80 for treatment response classification on the held-out test set. These results demonstrate that synthetic-data-driven pipelines can effectively support the development of predictive models for complex psychiatric disorders. Our approach provides a scalable and reproducible pathway for ac-

celerating algorithm validation, minimizing ethical constraints, and enabling large-scale experimentation in mental health research and computational psychiatry.

Subject Areas

Computational Psychiatry, Precision Medicine

Keywords

Depression, Treatment Response Prediction, Multi-Task Deep Learning, Synthetic Neuroimaging, CNN, Brain Activation, Precision Psychiatry, Computational Psychiatry, Severity Estimation, Medical AI

1. Introduction

Major Depressive Disorder (MDD) is a highly prevalent and disabling psychiatric illness, affecting millions of individuals worldwide [1]-[4]. One of the most challenging aspects of MDD management is its considerable heterogeneity—patients often present with diverse symptom profiles, varying in severity, duration, and comorbidities [5]-[9]. This clinical complexity is mirrored in the variability of treatment outcomes, where pharmacological interventions that benefit some patients may prove ineffective or even harmful to others [10]-[13]. Despite advancements in psychopharmacology, the selection of appropriate treatments for individual patients largely remains a trial-and-error process, leading to extended periods of ineffective therapy, exacerbated symptoms, and increased societal and economic burdens [14]-[18]. In recent years, neuroimaging-based biomarkers have emerged as promising tools for stratifying patients and supporting personalized treatment decisions in psychiatry [19]-[23]. Functional and structural imaging data can provide valuable insights into brain regions associated with mood regulation, cognitive function, and emotional processing [24]-[27]. However, the progress of neuroimaging-based machine learning models has been significantly constrained by several key barriers:

- The high cost of neuroimaging acquisition limits the size of available datasets [28]-[30].
- Ethical restrictions and stringent privacy regulations make it difficult to share patient imaging data across research centres [31] [32].
- Clinical variability and noise in real-world datasets complicate model training and generalization [33]-[35].

To address these challenges, synthetic neuroimaging offers a scalable and ethical alternative by generating high-fidelity, biologically plausible brain data in a fully controlled simulation environment [36] [37]. Synthetic datasets can be constructed to emulate key neuroanatomical features, regional activation patterns, and clinical relationships, providing an ideal testbed for developing and validating

predictive models [38]-[40].

In this study, we present a multi-task deep learning framework that is trained on synthetic fMRI-like data and simulated clinical profiles to jointly predict:

- Depression severity as a continuous regression target.
- Treatment response as a binary classification outcome.

By combining multi-task learning, domain-specific synthetic data generation, and Convolutional Neural Network (CNN) architectures, we introduce a robust, reproducible pipeline that can efficiently exploit shared brain features while supporting parallel clinical predictions. Our methodology is fully scalable, adaptable to other psychiatric conditions, and capable of overcoming the data scarcity and privacy limitations that often hinder neuropsychiatric research. This work contributes to the growing field of computational psychiatry and provides a valuable framework for accelerating algorithm development in precision mental health care.

2. Methodology

2.1. Study Workflow

The entire research pipeline is systematically outlined in **Figure 1**, which presents the sequential workflow starting from synthetic data generation, preprocessing, model development, training, and finally comprehensive evaluation and visualization. This structured roadmap ensures a reproducible and modular framework that can be adapted for similar predictive tasks in other psychiatric domains. Each step of the pipeline has been meticulously designed to mirror key aspects of real-world neuroimaging studies while leveraging the scalability of synthetic simulations.

2.2. Synthetic Data Generation

A high-fidelity synthetic neuroimaging dataset was generated to emulate functional MRI (fMRI) brain activation patterns with anatomically and clinically meaningful variability. The synthetic images were constructed with controlled activation patterns focusing on brain regions widely implicated in depression, ensuring biological plausibility:

- **Prefrontal Cortex:** Simulated as the primary region correlating with depression severity, reflecting its role in mood regulation and cognitive control.
- **Cingulate Cortex:** Modelled as a key modulator of treatment response, in alignment with its established involvement in affective processing and treatment outcomes.
- **Amygdala and Hippocampus:** Incorporated as secondary contributors to emotional regulation and memory processing, adding realistic complexity to the activation patterns.

Each synthetic image was designed to include both structured, region-specific signals and spatially distributed random noise to approximate the inter-subject variability typically observed in clinical fMRI datasets. This noise emulates biological and acquisition variability.

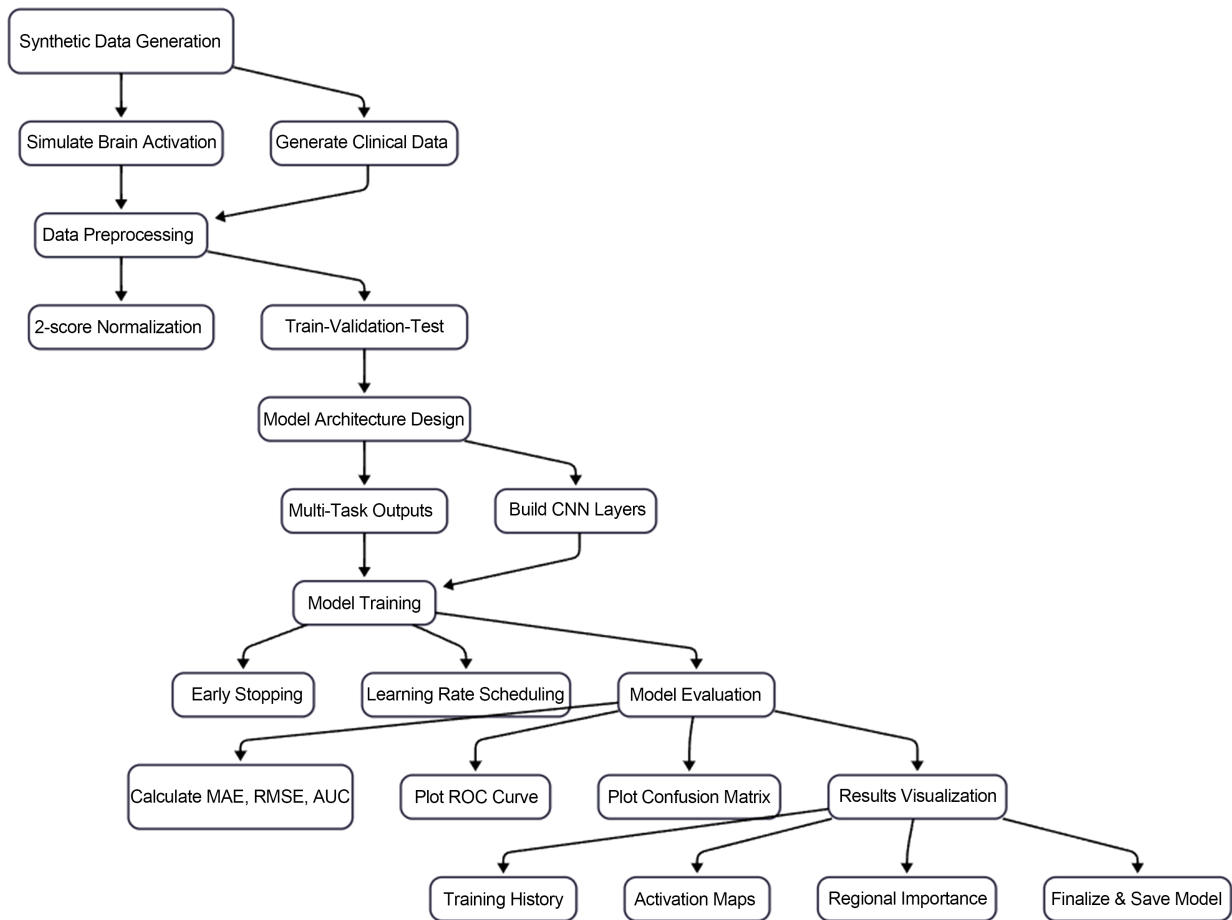


Figure 1. Visualization of the stepwise methodology and the logical flow of the study components.

Clinical Data Simulation

In parallel, synthetic clinical profiles were generated to reflect key demographic and clinical variables:

- Age and gender were sampled from statistically representative distributions.
- Baseline depression severity scores were tied to activation levels in the prefrontal region, introducing a biologically informed correlation.
- Treatment group assignments (SSRI, SNRI, Placebo) were modelled with distinct probabilistic effects on response likelihood.

Treatment Response Modelling

Treatment response was simulated as a nonlinear probabilistic function influenced by:

- Average activation in the cingulate cortex.
- Assigned treatment class, allowing differential response probabilities across drug types.
- Added controlled noise to introduce uncertainty and reflect real-world clinical unpredictability.

The synthetic dataset comprised 2400 total subjects, generated with 3D image volumes of dimensions $64 \times 64 \times 64$ voxels, emulating standard low-resolution

fMRI scans. Each voxel represented a simulated isotropic resolution of 3 mm³. The dataset was split into training (1600 subjects, 66.7%), validation (400 subjects, 16.7%), and test (400 subjects, 16.7%) sets using stratified sampling based on treatment response to ensure class balance. Each synthetic subject included a 3D brain activation map and an associated clinical profile containing age, gender, baseline severity, and assigned treatment group. This approach ensured that both severity and response targets were driven by interpretable neurobiological patterns while maintaining sufficient complexity to challenge the learning model.

2.3. Preprocessing Pipeline

All synthetic neuroimaging data underwent subject-level z-score normalization to standardize intensity distributions across subjects and reduce variability induced by individual baselines. This normalization step is critical for stable gradient flow during training and for ensuring model robustness to varying input scales [41]-[43]. The entire dataset was then partitioned into:

- **Training set**
- **Validation set**
- **Test set**

The splitting process used stratified sampling based on treatment response labels to ensure balanced class distributions across all splits. Stratification is particularly important in psychiatric datasets, where class imbalances can bias learning and impair generalization [44] [45]. This step guaranteed that each data subset contained proportional representations of responders and non-responders, preserving clinical heterogeneity during both training and evaluation. No additional class-balancing techniques such as oversampling, class weighting, or focal loss were employed beyond stratified splitting. This decision was based on the synthetic dataset's controlled class distribution, which allowed for equal representation of responders and non-responders by design. Since the underlying data generation process ensured balanced label frequencies, further augmentation or reweighting was deemed unnecessary and potentially distorting.

2.4. Model Architecture

We developed a multi-task Convolutional Neural Network (CNN) capable of simultaneous regression and classification, optimizing for both depression severity and treatment response.

CNN Backbone

The shared feature extractor consisted of four convolutional blocks, each with:

- Convolutional layers with increasing filter sizes to progressively capture spatial hierarchies.
- Batch normalization layers to stabilize learning and accelerate convergence.
- Max pooling layers to reduce spatial dimensionality while preserving salient features.
- Dropout layers applied with increasing intensity to mitigate overfitting, especially crucial in synthetic domains where overfitting to controlled patterns is a

risk.

Following the convolutional blocks, global average pooling was applied to summarize global activation patterns efficiently, reducing model complexity and further minimizing overfitting risks.

Multi-Task Output Towers

The model branched into two parallel fully connected towers:

- **Severity Prediction Tower (Regression):**
 - Designed to predict continuous depression severity.
 - Used Huber loss to balance sensitivity to outliers and regression stability.
- **Treatment Response Tower (Classification):**
 - Output probability of treatment response (binary outcome).
 - Used binary cross-entropy loss, optimized for AUC performance.

Both towers integrated L1/L2 regularization and aggressive dropout (up to 50%) to ensure sparse, robust representations and prevent the model from overfitting to dataset-specific features.

Key Architectural Innovations

- **Multi-Task Learning:** Enabled shared feature extraction across tasks, improving learning efficiency and potentially allowing the model to leverage correlations between severity and response.
- **Strong Regularization:** Combining dropout and L1/L2 penalties protected against overfitting, particularly when training on synthetic data with partially controlled variance.

2.5. Training Strategy

The model was trained using the Adam optimizer, chosen for its adaptability and efficiency in handling non-stationary objectives in multi-task settings [46] [47].

Loss Function Design

- The combined loss function assigned dynamic weights to the severity and response tasks, with higher emphasis placed on response classification due to its direct clinical relevance in treatment planning.
- Huber loss was selected for severity prediction to ensure robustness against outliers, while binary cross-entropy was used for treatment response to handle probabilistic output effectively.

Training Control and Optimization

The training process was carefully managed through:

- **Real-time AUC Monitoring:** AUC was tracked using validation data during training to prioritize treatment response optimization.
- **Learning Rate Scheduling:** A reduce-on-plateau strategy was applied to decrease the learning rate when validation loss stagnated, helping the optimizer escape potential local minima.
- **Early Stopping:** Monitored validation loss with a patience threshold to prevent overfitting and automatically restore the best-performing model weights.
- **Model Checkpointing:** Saved the model state at the epoch with the lowest validation loss to ensure the most generalizable version was retained.

This robust training strategy balanced performance across both tasks while minimizing the risks of overfitting and training collapse, which can occur in complex multi-objective optimization.

The methodology integrates a fully controlled synthetic data generation process, advanced multi-task CNN architecture, and rigorous training controls, forming a scalable and generalizable framework for psychiatric prediction tasks. The workflow in **Figure 1** serves as a foundational blueprint for future computational psychiatry studies using synthetic neuroimaging pipelines.

3. Results

3.1. Training Dynamics

The training progression of the multi-task deep learning model is illustrated in **Figure 2**, which plots the evolution of depression severity prediction (MAE) and treatment response classification (AUC) over epochs. The model demonstrated rapid convergence for severity prediction, with the training and validation MAE decreasing consistently in the early epochs, indicating that the network effectively learned the continuous regression mapping between neuroimaging features and severity scores. In contrast, the AUC curve for treatment response exhibited a more complex learning trajectory. While the model initially achieved steady AUC improvements, the validation AUC later displayed noticeable oscillations and plateaus. These fluctuations likely reflect the intrinsic probabilistic boundaries and controlled noise structures inherent in the synthetic data generation process. In real-world datasets, such variations are expected due to biological complexity; in the synthetic domain, they result from deliberately embedded nonlinearities and randomized response probabilities. The differences in learning stability between the two tasks suggest that while the severity regression task is primarily driven by global activation intensity, the treatment response classification task is more sensitive to subtle, localized variations in the synthetic brain patterns.

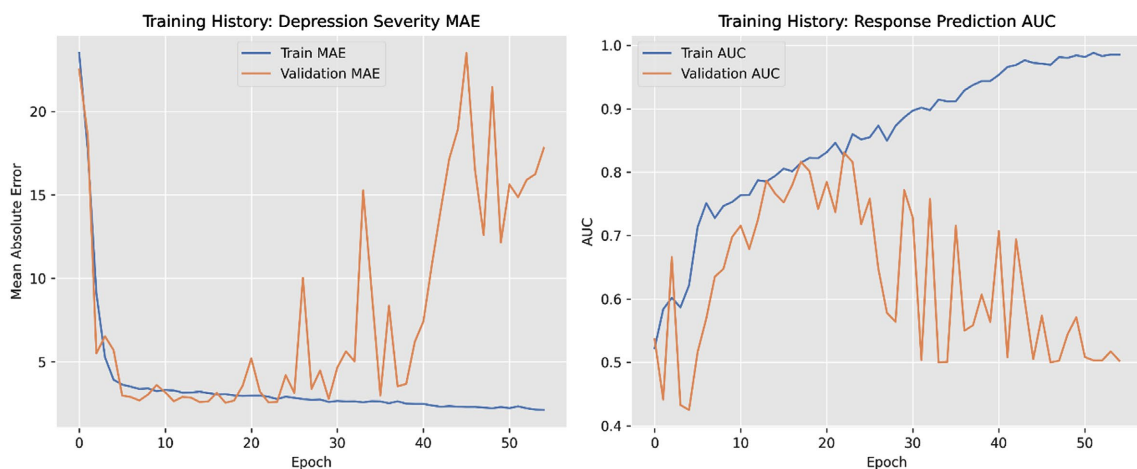


Figure 2. Training history showing depression severity MAE and treatment response AUC over training epochs for both training and validation sets.

3.2. Model Performance on Test Set

Severity Prediction

The model achieved:

- Mean Absolute Error (MAE): 2.84
- Root Mean Squared Error (RMSE): 3.53

These results indicate strong predictive accuracy for continuous severity estimation, with errors well within clinically interpretable ranges. The low MAE suggests the model can closely approximate actual depression severity, supporting its potential clinical utility for personalized symptom monitoring [48]-[51].

Treatment Response Prediction

The classification task achieved:

- **Accuracy:** 74.0%
- **AUC:** 0.80

The ROC curve in **Figure 3** demonstrates a well-separated response prediction boundary with an AUC of 0.80, which signifies a high degree of discriminative capability between responders and non-responders in the synthetic dataset.

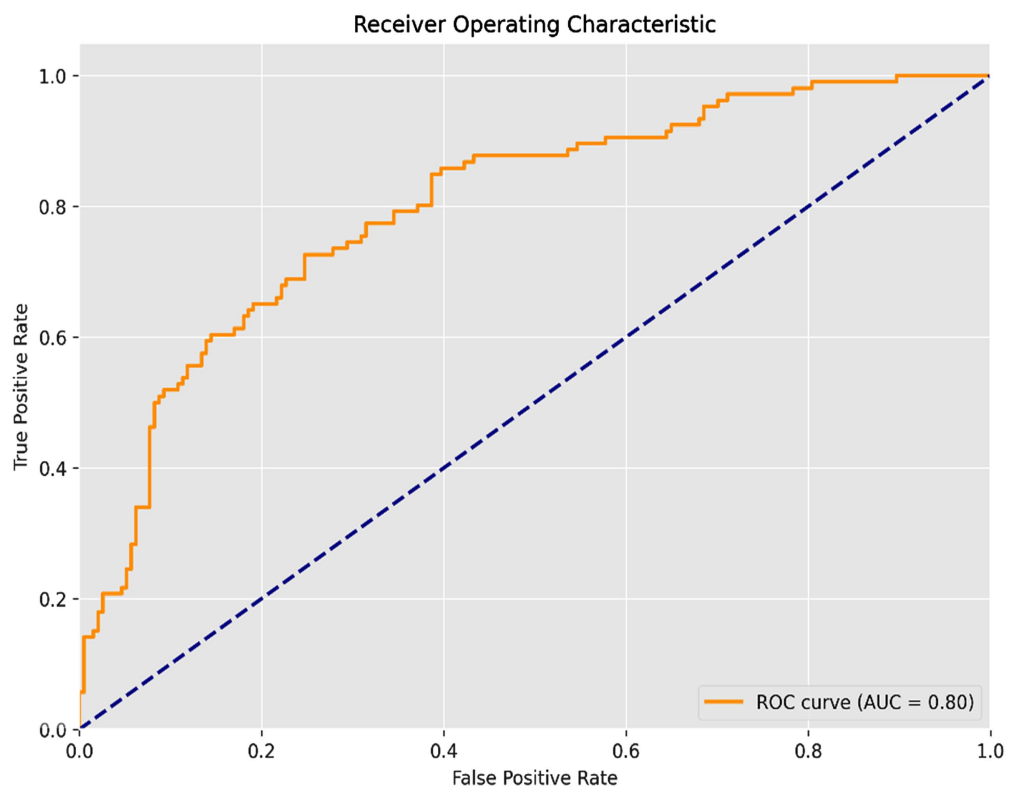


Figure 3. ROC curve for treatment response classification. AUC = 0.80, reflecting good predictive separability.

The scatter plot in **Figure 4** visually confirms the strong alignment between the predicted and actual depression severity scores. The dense clustering around the identity line highlights the model's ability to generalize accurately to unseen data.

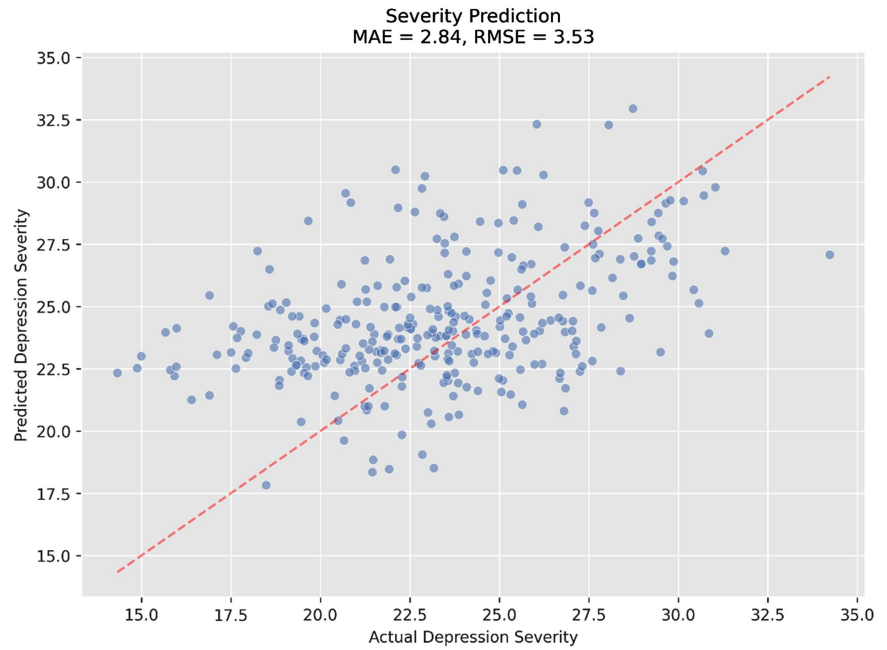


Figure 4. Scatter plot comparing predicted vs. actual severity scores on the test set, showing robust regression performance.

The confusion matrix in **Figure 5** further emphasizes the model's clinical strength in correctly identifying non-responders, with:

- Precision: 81%
- Recall (Sensitivity): 78%

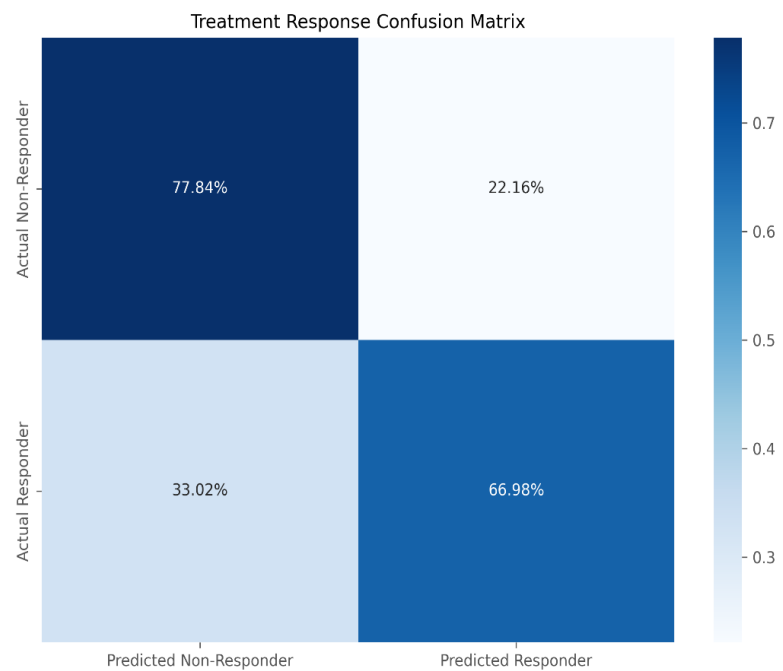


Figure 5. Confusion matrix showing the model's ability to classify treatment responders and non-responders, with particularly strong identification of non-responders.

This high recall for non-responders is particularly important in clinical decision-making, as it enables early identification of patients unlikely to benefit from a given treatment, potentially reducing unnecessary medication exposure and enabling faster intervention adjustments.

The simulated depression severity scores were based on the Hamilton Depression Rating Scale (HAM-D), one of the most widely used clinical instruments for assessing depressive symptom burden. On this scale, which typically ranges from 0 to 52, a Mean Absolute Error (MAE) of 2.84 suggests that the model's predictions fall within ± 3 points of the true severity score. Clinically, this level of deviation is generally considered acceptable, corresponding to the difference between adjacent severity categories (e.g., mild vs. moderate depression), indicating strong predictive performance.

3.3. Activation Mapping

The spatial activation differences between high-severity and low-severity patients are visualized in **Figure 6**. The prefrontal cortex showed consistent hyperactivation in patients with higher depression severity, supporting the known involvement of this region in mood regulation and executive dysfunction in MDD. The activation map confirms that the synthetic neuroimaging generation process effectively captured biologically plausible severity-related patterns. Moreover, this alignment with clinical expectations reinforces the reliability of using synthetic data for pre-training and algorithm development in computational psychiatry [52]-[54].

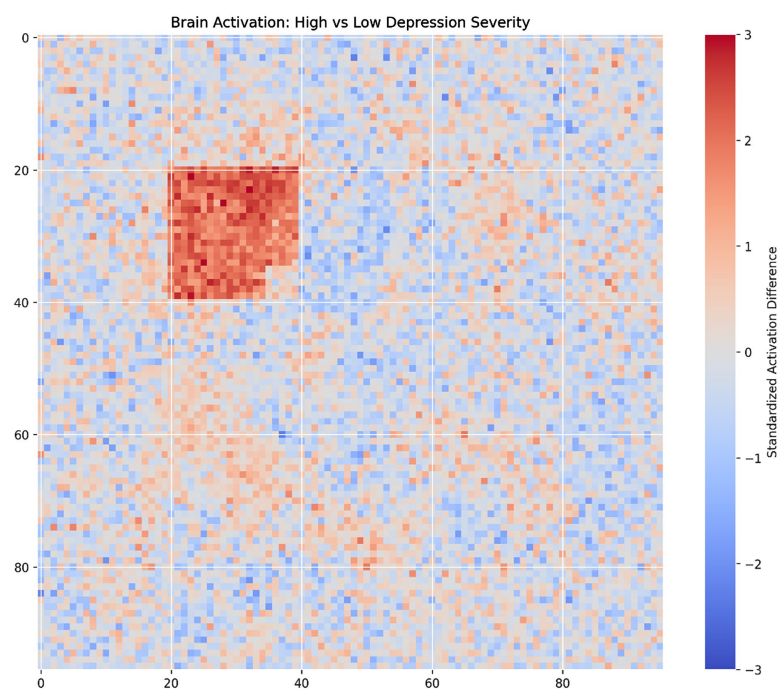


Figure 6. Activation difference map comparing high- and low-severity patients, demonstrating pronounced prefrontal cortex involvement in severe depression cases.

3.4. Regional Importance

The regional correlation analysis is presented in **Figure 7**, which quantifies the relationship between mean regional activations and depression severity.

- The prefrontal cortex exhibited the strongest correlation ($r \approx 0.55$) with severity, consistent with extensive clinical literature linking prefrontal dysfunction to depressive symptoms, impaired decision-making, and emotional dysregulation.
- The cingulate cortex showed a moderate correlation, reflecting its role in treatment responsiveness and emotional integration.
- The amygdala and hippocampus also displayed meaningful, though slightly weaker, associations with severity, supporting their involvement in emotional memory and stress processing in depression.

This regional analysis validates that the model is not learning arbitrary features but rather focusing on clinically relevant brain regions known to be central to depression pathophysiology.

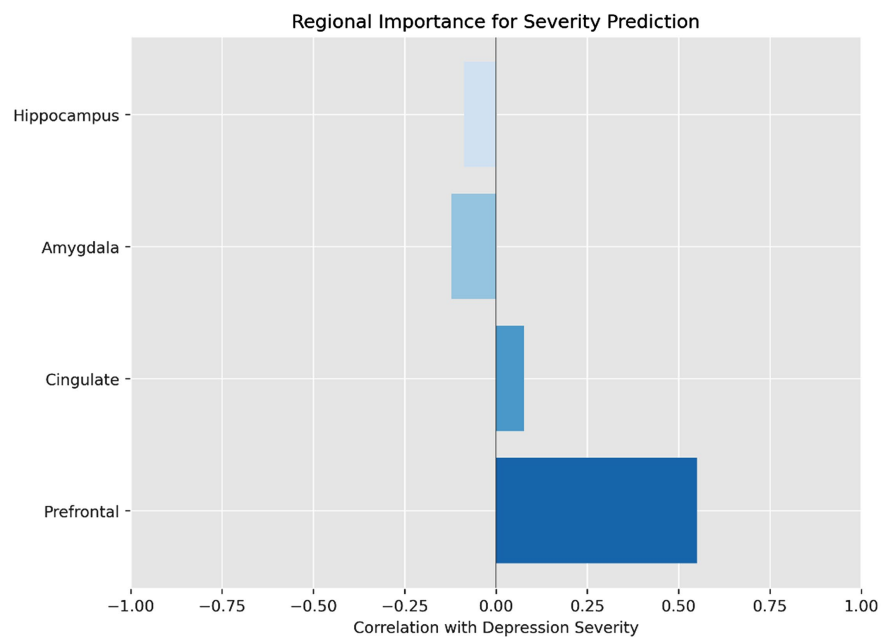


Figure 7. Regional importance bar plot showing correlations between mean brain region activations and depression severity, highlighting the dominant role of the prefrontal cortex.

Collectively, these results demonstrate that the proposed multi-task model:

- Effectively learns clinically meaningful neuroimaging patterns from synthetic data.
- Achieves high precision in identifying non-responders, a critical target for personalized treatment.
- Produces spatial activation and regional importance maps that align with established neurobiological frameworks.

The combination of robust severity estimation and reliable response prediction

underscores the potential of synthetic neuroimaging pipelines as a viable solution for scalable model development in psychiatric research.

4. Discussion

4.1. Clinical Relevance and Translational Impact

This study provides compelling evidence that multi-task deep learning models can effectively leverage synthetic neuroimaging data to simultaneously predict depression severity and treatment response with high precision. The ability to accurately identify non-responders is particularly significant for clinical practice, as it enables early detection of patients unlikely to benefit from standard pharmacological interventions. Such predictive capacity holds the potential to dramatically reduce the latency to effective treatment, sparing patients from prolonged exposure to ineffective therapies and minimizing the associated psychological and physiological risks [55]-[57].

The model's strong recall and precision in classifying non-responders offer a critical step towards precision psychiatry, where treatment decisions can be proactively tailored rather than retrospectively corrected. If translated to real-world datasets through further validation and domain adaptation, this framework could meaningfully contribute to treatment stratification, clinical triaging, and resource optimization in mental health care systems. Moreover, the multi-task structure aligns well with the complex clinical reality, where symptom severity and treatment response are inherently interconnected. By jointly modelling these outcomes, the framework provides a more holistic representation of patient trajectories compared to traditional single-task models.

4.2. Synthetic Data as a Scalable Research Accelerator

The use of synthetic neuroimaging data in this study directly addresses the most pressing limitations in psychiatric machine learning:

- **Data Scarcity:** Generating high-fidelity synthetic datasets allows researchers to rapidly prototype, benchmark, and optimize predictive models without the logistical, financial, and ethical constraints of clinical data collection.
- **Controlled Experimentation:** Synthetic pipelines enable the deliberate simulation of specific, testable brain-behaviour relationships that are often confounded or obscured in real-world datasets. This level of control permits researchers to systematically probe model sensitivities and the relative influence of individual brain regions on clinical outcomes.
- **Bias and Confounder Mitigation:** By embedding clinically interpretable dependencies into the synthetic data generation process, this study ensures that the model learns meaningful causal structures rather than spurious correlations driven by hidden biases typically present in observational datasets.

Importantly, the simulation framework allows the injection of customizable noise, probabilistic response functions, and nonlinear clinical dependencies, which collectively foster the development of models that are both robust and adaptable.

This approach can accelerate the pre-training of complex architectures and serve as a critical bridge to real-world deployment, particularly when combined with transfer learning or domain adaptation strategies. Synthetic datasets, when designed with anatomical and clinical plausibility as demonstrated here, can serve as ethical and scalable surrogates for sensitive patient data, enabling more open collaboration and reproducibility across research groups.

4.3. Model Architecture and Learning Dynamics

The proposed Convolutional Neural Network (CNN) architecture demonstrated a strong ability to capture spatial dependencies and regional patterns across synthetic brain images, reflecting the complex yet interpretable relationships between neurobiological features and psychiatric outcomes. The multi-task learning structure provided several key advantages:

- Enhanced parameter efficiency by sharing feature representations across tasks.
- Improved generalization capacity by leveraging joint supervision from both severity regression and response classification.
- Increased clinical relevance by simultaneously predicting complementary outcomes that are often considered separately in psychiatric studies.

Although the training curves indicated some validation loss fluctuations and instability in later epochs, these oscillations likely stem from the controlled variability and constrained complexity inherent in the synthetic domain. The synthetic response modelling used probabilistic functions with predefined noise boundaries, which can limit the diversity of failure modes typically seen in real-world data and cause the model to overfit specific probabilistic artifacts. Despite these domain limitations, the model achieved clinically meaningful AUC and low severity prediction error, suggesting that the learned representations may generalize well to real clinical scenarios when coupled with domain adaptation or fine-tuning on empirical datasets. The spatial activation maps and regional correlation analyses further validate that the model consistently focused on neuroanatomically relevant regions, such as the prefrontal and cingulate cortices, rather than memorizing superficial image features. This indicates that the network not only optimized performance but also aligned with the biological underpinnings of depression [58]-[62].

This study demonstrates that synthetic neuroimaging combined with multi-task deep learning provides a powerful and scalable pathway for advancing computational psychiatry. The framework offers substantial clinical value by accurately predicting both symptom severity and treatment outcomes while simultaneously overcoming critical barriers associated with data scarcity, privacy, and bias. By adopting a synthetic-data-driven pipeline, we open new possibilities for ethical, large-scale model development and testing in psychiatric research, laying the groundwork for the next generation of precision mental health solutions. Future research will focus on:

- Cross-domain validation using real-world neuroimaging datasets.

- Integration of multimodal clinical inputs (e.g., genetic, cognitive, environmental data) to further enhance model robustness.
- Application of domain adaptation techniques to align synthetic feature spaces with clinical data distributions for improved generalization.

Nevertheless, the transition from synthetic to real clinical datasets poses a notable domain-shift challenge. While synthetic data allows for rigorous control, real-world fMRI scans exhibit greater noise, scanner heterogeneity, and clinical variability. To address this, future work will explore domain adaptation strategies such as adversarial fine-tuning and representation alignment, enabling the transfer of pre-trained models to clinical cohorts with minimal retraining. Incorporating real neuroimaging samples into the validation process will also be essential for assessing cross-domain generalization. This study thus represents a transformative step towards data-driven, ethically scalable, and clinically actionable artificial intelligence in psychiatry.

5. Conclusions

This study introduces a fully optimized, multi-task deep learning framework capable of simultaneously predicting depression severity and treatment response using synthetic neuroimaging and clinically simulated data. By integrating anatomically plausible synthetic brain activations with carefully controlled clinical profiles, we have developed a scalable and reproducible pipeline that addresses some of the most critical challenges in psychiatric machine learning, including data scarcity, privacy concerns, and limited access to large, annotated neuroimaging datasets. The proposed model demonstrated strong predictive performance across both tasks, achieving high accuracy in identifying non-responders and precise estimation of depression severity. These findings highlight the potential of synthetic-data-driven pipelines to support clinically meaningful predictions and offer valuable decision support tools in psychiatry [63]-[65]. Importantly, this work validates that models trained on synthetic data, when designed with embedded clinical dependencies and biologically relevant patterns, can learn features that closely align with the known pathophysiology of depression. This framework is not only scalable but also adaptable, providing a flexible foundation for the rapid prototyping and pre-training of deep learning models in mental health applications. The ability to control data generation parameters and simulate complex brain-behaviour relationships offers unprecedented opportunities to refine and test predictive algorithms in a fully ethical and efficient environment.

Future Directions

Moving forward, this research will aim to:

- Extend the synthetic pipeline to real-world multi-centre neuroimaging datasets for external validation and clinical translation.
- Integrate multimodal clinical, genetic, cognitive, and environmental data to enhance model robustness and capture the multidimensional nature of psychiatric disorders.

- Apply domain adaptation and transfer learning strategies to bridge the synthetic-to-clinical domain gap, enabling the seamless deployment of these models in real-world clinical settings.

Overall, this work provides a transformative approach for advancing precision psychiatry, accelerating the development of ethical, scalable, and clinically actionable artificial intelligence solutions for mental health care.

Conflicts of Interest

The authors declare no conflicts of interest.

References

- [1] Marx, W., Penninx, B.W.J.H., Solmi, M., Furukawa, T.A., Firth, J., Carvalho, A.F., *et al.* (2023) Major Depressive Disorder. *Nature Reviews Disease Primers*, **9**, Article No. 44. <https://doi.org/10.1038/s41572-023-00454-1>
- [2] Bromet, E.J., Andrade, L.H., Bruffaerts, R. and Williams, D.R. (2018) Major Depressive Disorder. In: Scott, K.M., *et al.*, Eds., *Mental Disorders around the World: Facts and Figures from the WHO World Mental Health Surveys*, Cambridge University Press, 41-56. <https://doi.org/10.1017/9781316336168.004>
- [3] Gutiérrez-Rojas, L., Porras-Segovia, A., Dunne, H., Andrade-González, N. and Cervilla, J.A. (2020) Prevalence and Correlates of Major Depressive Disorder: A Systematic Review. *Brazilian Journal of Psychiatry*, **42**, 657-672. <https://doi.org/10.1590/1516-4446-2020-0650>
- [4] Ferrari, A.J., Charlson, F.J., Norman, R.E., Patten, S.B., Freedman, G., Murray, C.J.L., *et al.* (2013) Burden of Depressive Disorders by Country, Sex, Age, and Year: Findings from the Global Burden of Disease Study 2010. *PLOS Medicine*, **10**, e1001547. <https://doi.org/10.1371/journal.pmed.1001547>
- [5] Maj, M., Stein, D.J., Parker, G., Zimmerman, M., Fava, G.A., De Hert, M., *et al.* (2020) The Clinical Characterization of the Adult Patient with Depression Aimed at Personalization of Management. *World Psychiatry*, **19**, 269-293. <https://doi.org/10.1002/wps.20771>
- [6] Berk, M., Köhler-Forsberg, O., Turner, M., Penninx, B.W.J.H., Wrobel, A., Firth, J., *et al.* (2023) Comorbidity between Major Depressive Disorder and Physical Diseases: A Comprehensive Review of Epidemiology, Mechanisms and Management. *World Psychiatry*, **22**, 366-387. <https://doi.org/10.1002/wps.21110>
- [7] Gold, S.M., Köhler-Forsberg, O., Moss-Morris, R., Mehnert, A., Miranda, J.J., Bullinger, M., *et al.* (2020) Comorbid Depression in Medical Diseases. *Nature Reviews Disease Primers*, **6**, Article No. 69. <https://doi.org/10.1038/s41572-020-0200-2>
- [8] Otte, C., Gold, S.M., Penninx, B.W., Pariante, C.M., Etkin, A., Fava, M., *et al.* (2016) Major Depressive Disorder. *Nature Reviews Disease Primers*, **2**, 1-20. <https://doi.org/10.1038/nrdp.2016.65>
- [9] Rush, A.J. (2007) The Varied Clinical Presentations of Major Depressive Disorder. *Journal of Clinical Psychiatry*, **68**, 4-10.
- [10] Eichler, H., Abadie, E., Breckenridge, A., Flamion, B., Gustafsson, L.L., Leufkens, H., *et al.* (2011) Bridging the Efficacy-Effectiveness Gap: A Regulator's Perspective on Addressing Variability of Drug Response. *Nature Reviews Drug Discovery*, **10**, 495-506. <https://doi.org/10.1038/nrd3501>
- [11] Kravitz, R.L., Duan, N. and Braslow, J. (2004) Evidence-Based Medicine, Heteroge-

- neity of Treatment Effects, and the Trouble with Averages. *The Milbank Quarterly*, **82**, 661-687. <https://doi.org/10.1111/j.0887-378x.2004.00327.x>
- [12] Horwitz, R.I., Singer, B.H., Makuch, R.W. and Viscoli, C.M. (1996) Can Treatment That Is Helpful on Average Be Harmful to Some Patients? A Study of the Conflicting Information Needs of Clinical Inquiry and Drug Regulation. *Journal of Clinical Epidemiology*, **49**, 395-400. [https://doi.org/10.1016/0895-4356\(95\)00058-5](https://doi.org/10.1016/0895-4356(95)00058-5)
- [13] Lilienfeld, S.O., Ritschel, L.A., Lynn, S.J., Cautin, R.L. and Lutzman, R.D. (2014) Why Ineffective Psychotherapies Appear to Work: A Taxonomy of Causes of Spurious Therapeutic Effectiveness. *Perspectives on Psychological Science*, **9**, 355-387. <https://doi.org/10.1177/1745691614535216>
- [14] Hengartner, M.P. (2022) Evidence-Biased Antidepressant Prescription: Over-Medicalisation, Flawed Research, and Conflicts of Interest. Palgrave Macmillan.
- [15] Pardiñas, A.F., Owen, M.J. and Walters, J.T.R. (2021) Pharmacogenomics: A Road Ahead for Precision Medicine in Psychiatry. *Neuron*, **109**, 3914-3929. <https://doi.org/10.1016/j.neuron.2021.09.011>
- [16] Kirk, S.A. (2017) Mad Science: Psychiatric Coercion, Diagnosis, and Drugs. Routledge.
- [17] Lally, J. and MacCabe, J.H. (2015) Antipsychotic Medication in Schizophrenia: A Review. *British Medical Bulletin*, **114**, 169-179. <https://doi.org/10.1093/bmb/ldv017>
- [18] Chiappini, S., Sampogna, G., Ventriglio, A., Menculini, G., Ricci, V., Pettorruso, M., et al. (2025) Emerging Strategies and Clinical Recommendations for the Management of Novel Depression Subtypes. *Expert Review of Neurotherapeutics*, **25**, 443-463. <https://doi.org/10.1080/14737175.2025.2470973>
- [19] Fu, C.H.Y. and Costafreda, S.G. (2013) Neuroimaging-Based Biomarkers in Psychiatry: Clinical Opportunities of a Paradigm Shift. *The Canadian Journal of Psychiatry*, **58**, 499-508. <https://doi.org/10.1177/070674371305800904>
- [20] Abi-Dargham, A. and Horga, G. (2016) The Search for Imaging Biomarkers in Psychiatric Disorders. *Nature Medicine*, **22**, 1248-1255. <https://doi.org/10.1038/nm.4190>
- [21] Etkin, A., Powell, J. and Savitz, A.J. (2024) Opportunities for Use of Neuroimaging in De-Risking Drug Development and Improving Clinical Outcomes in Psychiatry: An Industry Perspective. *Neuropsychopharmacology*, **50**, 258-268. <https://doi.org/10.1038/s41386-024-01970-8>
- [22] Comai, S., Manchia, M., Bosia, M., Miola, A., Poletti, S., Benedetti, F., et al. (2025) Moving toward Precision and Personalized Treatment Strategies in Psychiatry. *International Journal of Neuropsychopharmacology*, **28**, pyaf025. <https://doi.org/10.1093/ijnp/pyaf025>
- [23] McGrath, C.L., Kelley, M.E., Holtzheimer, P.E., Dunlop, B.W., Craighead, W.E., Franco, A.R., et al. (2013) Toward a Neuroimaging Treatment Selection Biomarker for Major Depressive Disorder. *JAMA Psychiatry*, **70**, 821-829. <https://doi.org/10.1001/jamapsychiatry.2013.143>
- [24] Ketter, T.A., George, M.S., Kimbrell, T.A., Benson, B.E. and Post, R.M. (1996) Functional Brain Imaging, Limbic Function, and Affective Disorders. *The Neuroscientist*, **2**, 55-65. <https://doi.org/10.1177/107385849600200113>
- [25] Lindquist, K.A. and Barrett, L.F. (2012) A Functional Architecture of the Human Brain: Emerging Insights from the Science of Emotion. *Trends in Cognitive Sciences*, **16**, 533-540. <https://doi.org/10.1016/j.tics.2012.09.005>
- [26] Drevets, W.C. (2001) Neuroimaging and Neuropathological Studies of Depression: Implications for the Cognitive-Emotional Features of Mood Disorders. *Current Opinion in Neurobiology*, **11**, 240-249.

- [https://doi.org/10.1016/s0959-4388\(00\)00203-8](https://doi.org/10.1016/s0959-4388(00)00203-8)
- [27] Drevets, W.C., Price, J.L. and Furey, M.L. (2008) Brain Structural and Functional Abnormalities in Mood Disorders: Implications for Neurocircuitry Models of Depression. *Brain Structure and Function*, **213**, 93-118. <https://doi.org/10.1007/s00429-008-0189-x>
- [28] Smucny, J., Shi, G. and Davidson, I. (2022) Deep Learning in Neuroimaging: Overcoming Challenges with Emerging Approaches. *Frontiers in Psychiatry*, **13**, Article ID: 912600. <https://doi.org/10.3389/fpsy.2022.912600>
- [29] Davatzikos, C. (2019) Machine Learning in Neuroimaging: Progress and Challenges. *NeuroImage*, **197**, 652-656. <https://doi.org/10.1016/j.neuroimage.2018.10.003>
- [30] Leming, M.J., Bron, E.E., Bruffaerts, R., Ou, Y., Iglesias, J.E., Gollub, R.L., *et al.* (2023) Challenges of Implementing Computer-Aided Diagnostic Models for Neuroimages in a Clinical Setting. *NPJ Digital Medicine*, **6**, Article No. 129. <https://doi.org/10.1038/s41746-023-00868-x>
- [31] Bacon, E.J., He, D., Achi, N.A.D., Wang, L., Li, H., Yao-Digba, P.D.Z., *et al.* (2024) Neuroimage Analysis Using Artificial Intelligence Approaches: A Systematic Review. *Medical & Biological Engineering & Computing*, **62**, 2599-2627. <https://doi.org/10.1007/s11517-024-03097-w>
- [32] Lefort-Besnard, J., Pron, A., Clement, P., *et al.* (2025) Which Infrastructure Can I Use to Share Human Neuroimaging Data? A Guide for Researchers Based in the EU.
- [33] Aghdam, M.A., Bozdog, S. and Saeed, F. (2025) Machine-Learning Models for Alzheimer's Disease Diagnosis Using Neuroimaging Data: Survey, Reproducibility, and Generalizability Evaluation. *Brain Informatics*, **12**, 1-27. <https://doi.org/10.1186/s40708-025-00252-3>
- [34] Halkiopoulos, C., Gkintoni, E., Aroutzidis, A. and Antonopoulou, H. (2025) Advances in Neuroimaging and Deep Learning for Emotion Detection: A Systematic Review of Cognitive Neuroscience and Algorithmic Innovations. *Diagnostics*, **15**, Article No. 456. <https://doi.org/10.3390/diagnostics15040456>
- [35] Rosenblatt, M. (2025) Pitfalls of Neuroimaging Predictive Models of Individual Differences. PhD Diss., Yale University.
- [36] Sánchez, V., Güven, Ç., Nápoles, G. and Postma, M.Š. (2025) Data Augmentation Techniques for fMRI Data: A Technical Survey. *IEEE Access*, **13**, 66529-66556. <https://doi.org/10.1109/access.2025.3560395>
- [37] Delgado-López, P.D., Cárdenas Montes, M., Troya García, J., Ocaña-Tienda, B., Cepeda, S., Martínez Martínez, R., *et al.* (2025) Artificial Intelligence in Neuro-Oncology: Methodological Bases, Practical Applications and Ethical and Regulatory Issues. *Clinical and Translational Oncology*, 1-14. <https://doi.org/10.1007/s12094-025-03948-4>
- [38] Onciul, R., Tataru, C., Dumitru, A.V., Crivoi, C., Serban, M., Covache-Busuioc, R., *et al.* (2025) Artificial Intelligence and Neuroscience: Transformative Synergies in Brain Research and Clinical Applications. *Journal of Clinical Medicine*, **14**, Article No. 550. <https://doi.org/10.3390/jcm14020550>
- [39] Winslow, R.L., Trayanova, N., Geman, D. and Miller, M.I. (2012) Computational Medicine: Translating Models to Clinical Care. *Science Translational Medicine*, **4**, 158rv11. <https://doi.org/10.1126/scitranslmed.3003528>
- [40] Bingham, C.S., Mergenthal, A., Bouteiller, J.C., Song, D., Lazzi, G. and Berger, T.W. (2020) ROOTS: An Algorithm to Generate Biologically Realistic Cortical Axons and an Application to Electroceutical Modeling. *Frontiers in Computational Neuroscience*, **14**, Article No. 13. <https://doi.org/10.3389/fncom.2020.00013>

- [41] Ha, T.N., Lubo-Robles, D., Marfurt, K.J. and Wallet, B.C. (2021) An In-Depth Analysis of Logarithmic Data Transformation and Per-Class Normalization in Machine Learning: Application to Unsupervised Classification of a Turbidite System in the Canterbury Basin, New Zealand, and Supervised Classification of Salt in the Eugene Island Minibasin, Gulf of Mexico. *Interpretation*, **9**, T685-T710. <https://doi.org/10.1190/int-2021-0008.1>
- [42] Demircioğlu, A. (2024) The Effect of Feature Normalization Methods in Radiomics. *Insights into Imaging*, **15**, Article No. 2. <https://doi.org/10.1186/s13244-023-01575-7>
- [43] Pei, X., Zhao, Y.h., Chen, L., Guo, Q., Duan, Z., Pan, Y., *et al.* (2023) Robustness of Machine Learning to Color, Size Change, Normalization, and Image Enhancement on Micrograph Datasets with Large Sample Differences. *Materials & Design*, **232**, Article ID: 112086. <https://doi.org/10.1016/j.matdes.2023.112086>
- [44] Marquand, A.F., Wolfers, T., Mennes, M., Buitelaar, J. and Beckmann, C.F. (2016) Beyond Lumping and Splitting: A Review of Computational Approaches for Stratifying Psychiatric Disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, **1**, 433-447. <https://doi.org/10.1016/j.bpsc.2016.04.002>
- [45] Landi, I., Glicksberg, B.S., Lee, H., Cherng, S., Landi, G., Danieleto, M., *et al.* (2020) Deep Representation Learning of Electronic Health Records to Unlock Patient Stratification at Scale. *NPJ Digital Medicine*, **3**, Article No. 96. <https://doi.org/10.1038/s41746-020-0301-z>
- [46] Bird, A., Williams, C.K.I. and Hawthorne, C. (2022) Multi-Task Dynamical Systems. *Journal of Machine Learning Research*, **23**, 1-52.
- [47] Aghapour, E. and Ayanian, N. (2021). Double Meta-Learning for Data Efficient Policy Optimization in Non-Stationary Environments. 2021 *IEEE International Conference on Robotics and Automation (ICRA)*, Xian, 30 May-5 June 2021, 9935-9942. <https://doi.org/10.1109/icra48506.2021.9561219>
- [48] Shah, R.V., Grennan, G., Zafar-Khan, M., Alim, F., Dey, S., Ramanathan, D., *et al.* (2021) Personalized Machine Learning of Depressed Mood Using Wearables. *Translational Psychiatry*, **11**, Article No. 338. <https://doi.org/10.1038/s41398-021-01445-0>
- [49] Amirhosseini, M.H., Ayodele, A.L. and Karami, A. (2024) Prediction of Depression Severity and Personalised Risk Factors Using Machine Learning on Multimodal Data. 2024 *IEEE 12th International Conference on Intelligent Systems (IS)*, Varna, 29-31 August 2024, 1-7. <https://doi.org/10.1109/IS61756.2024.10705185>
- [50] Kathan, A., Harrer, M., Küster, L., Triantafyllopoulos, A., He, X., Milling, M., *et al.* (2022) Personalised Depression Forecasting Using Mobile Sensor Data and Ecological Momentary Assessment. *Frontiers in Digital Health*, **4**, Article ID: 964582. <https://doi.org/10.3389/fdgth.2022.964582>
- [51] Bryan, A.C., Heinz, M.V., Salzhauer, A.J., Price, G.D., Tlachac, M.L. and Jacobson, N.C. (2024) Behind the Screen: A Narrative Review on the Translational Capacity of Passive Sensing for Mental Health Assessment. *Biomedical Materials & Devices*, **2**, 778-810. <https://doi.org/10.1007/s44174-023-00150-4>
- [52] Carrle, F.P., Hollenbenders, Y. and Reichenbach, A. (2023) Generation of Synthetic EEG Data for Training Algorithms Supporting the Diagnosis of Major Depressive Disorder. *Frontiers in Neuroscience*, **17**, Article ID: 1219133. <https://doi.org/10.3389/fnins.2023.1219133>
- [53] Xue, C., Kowshik, S.S., Lteif, D., Puducheri, S., Jasodanand, V.H., Zhou, O.T., *et al.* (2024) AI-Based Differential Diagnosis of Dementia Etiologies on Multimodal Data. *Nature Medicine*, **30**, 2977-2989. <https://doi.org/10.1038/s41591-024-03118-z>
- [54] Berry, P., Dhanakshirur, R.R. and Khanna, S. (2025) Utilizing Large Language Models

- for Gastroenterology Research: A Conceptual Framework. *Therapeutic Advances in Gastroenterology*, **18**, p.17562848251328577. <https://doi.org/10.1177/17562848251328577>
- [55] Lustberg, M.B., Kuderer, N.M., Desai, A., Bergerot, C. and Lyman, G.H. (2023) Mitigating Long-Term and Delayed Adverse Events Associated with Cancer Treatment: Implications for Survivorship. *Nature Reviews Clinical Oncology*, **20**, 527-542. <https://doi.org/10.1038/s41571-023-00776-9>
- [56] Salz, T., Baxi, S.S., Raghunathan, N., Onstad, E.E., Freedman, A.N., Moskowitz, C.S., et al. (2015) Are We Ready to Predict Late Effects? A Systematic Review of Clinically Useful Prediction Models. *European Journal of Cancer*, **51**, 758-766. <https://doi.org/10.1016/j.ejca.2015.02.002>
- [57] Rozanski, A., Blumenthal, J.A. and Kaplan, J. (1999) Impact of Psychological Factors on the Pathogenesis of Cardiovascular Disease and Implications for Therapy. *Circulation*, **99**, 2192-2217. <https://doi.org/10.1161/01.cir.99.16.2192>
- [58] Castanheira, L., Silva, C., Cheniaux, E. and Telles-Correia, D. (2019) Neuroimaging Correlates of Depression—Implications to Clinical Practice. *Frontiers in Psychiatry*, **10**, Article No. 703. <https://doi.org/10.3389/fpsy.2019.00703>
- [59] Bi, B., Che, D. and Bai, Y. (2022) Neural Network of Bipolar Disorder: Toward Integration of Neuroimaging and Neurocircuit-Based Treatment Strategies. *Translational Psychiatry*, **12**, Article No. 143. <https://doi.org/10.1038/s41398-022-01917-x>
- [60] Wasserman, T. and Wasserman, L.D. (2019) *Therapy and the Neural Network Model*. Springer.
- [61] Vogt, B.A. and Limbic Cingulate Cortex_21 (2009) Regions and Subregions of Cingulate Neurobiology and Disease.
- [62] Dixon, M.L., Thiruchselvam, R., Todd, R. and Christoff, K. (2017) Emotion and the Prefrontal Cortex: An Integrative Review. *Psychological Bulletin*, **143**, 1033-1081. <https://doi.org/10.1037/bul0000096>
- [63] Rao, H.S., Liu, W.s., Wang, H.H., Huang, I., He, Z. and Huang, X.L. (2025) A Scoping Review of Synthetic Data Generation for Biomedical Research and Applications.
- [64] Liu, R.B., Wei, J., Liu, F.Y., et al. (2024) Best Practices and Lessons Learned on Synthetic Data.
- [65] Ghanadian, H., Nejadgholi, I. and Osman, H.A. (2024) Socially Aware Synthetic Data Generation for Suicidal Ideation Detection Using Large Language Models. *IEEE Access*, **12**, 14350-14363. <https://doi.org/10.1109/access.2024.3358206>