

Physics-Informed Explainable Machine Learning for Predicting Wavelength- Dependent Optical Modulation of Neuronal Activity

Md. Rasel Islam

Islamic University

Shazid Mahmud Labim

Islamic University

Nazmul Hasan

Islamic University

Md. Sadiqur Rahman Zihad

Islamic University

Md. Rabiul Islam

Islamic University

Md. Shariful Islam

afmsi1976@gmail.com

Islamic University

Research Article

Keywords: Physics-informed machine learning, Explainable artificial intelligence, Neuronal optical modulation, Wavelength-dependent photo-modulation, Optical neuromodulation; Uncertainty quantification

Posted Date: October 26th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11345215/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Physics-Informed Explainable Machine Learning for Predicting Wavelength-Dependent Optical Modulation of Neuronal Activity

¹Md. Rasel Islam, ²Shazid Mahmud Labim, ²Nazmul Hasan, ²Md. Sadiqur Rahman Zihad, ²Md. Rabiul Islam, ^{*}2Md. Shariful Islam

¹Department of Biomedical Engineering, Islamic University, Kushtia-7003, Bangladesh

²Department of Information and Communication Technology, Islamic University, Kushtia-7003, Bangladesh

Corresponding author: Md. Shariful Islam, afmsi1976@gmail.com or afmsi76@ice.iu.ac.bd

ORCID: <https://orcid.org/0000-0003-1641-8981>

Abstract

Wavelength-dependent optical stimulation may also be a potential method for non-invasive modulation of neural activity, but the effects of this stimulation are controlled by a series of factors that include the energy of the photons, the optical properties of the tissue, the optical dose delivered, wavelength sensitivity, and neuronal characteristics. This work introduces a physics-informed machine learning framework to explore the responses of neurons across different wavelengths and to measure predictive uncertainty, in an explainable manner. Two optical features derived from experimental measurements and physics modelling were combined with the biological sensitivity descriptors and tested on a data set of 100,000 samples across two subsets of neuronal phenotypes and across a 400-700 nm wavelength range using multiple regression models. Wavelength grouped validation and locked test set were used in model evaluation. Using the full mechanistic feature set increased the locked-test coefficient of determination R^2 from 0.012 for the data-only feature set to 0.035, and the RMSE and MAE for primary endpoint: percentage change in firing rate were 8.00 percentage points and 6.35 percentage points, respectively. The modest increase in predictive value from the mechanistic features suggests incremental value, but the modest R^2 suggests that there was still significant variation in firing-rate modulation that wasn't predicted. Explainability analysis revealed that the most important predictive features were photo-modulation gain, biological optical dose, wavelength sensitivity, tissue scattering, and peak wavelength. Predictive performance was endpoint dependent; the modulation of membrane-potential was the best with R^2 0.388. Split-conformal calibration resulted in an empirical coverage of 90.2% for nominal 90% prediction intervals, whereas increasing the input noise resulted in a reduced performance of the model. In summary, the framework allows for an interpretable computational method to investigate neuronal responses across all wavelengths, to characterize endpoint-specific predictability and to assess uncertainty. It needs to be validated by independent experiments to determine its generalizability to real neuronal systems.

Keywords: Physics-informed machine learning, Explainable artificial intelligence, Neuronal optical modulation, Wavelength-dependent photo-modulation, Optical neuromodulation; Uncertainty quantification

1. Introduction

Photobiomodulation (PBM) is a new, low-intensity, non-invasive method to modulate brain and neuronal function through the use of optical radiation. PBM may offer a safer, more flexible approach to modulate neural physiology and is not as invasive as conventional neuromodulation methods, which would require implanting electrodes or directly penetrating tissue. Experimental and clinical studies are accumulating and demonstrate that transcranial PBM can modulate brain activity, cerebral hemodynamics, metabolism, cognitive function, and neurological outcomes [1-4]. Systematic studies have also shown that measurable changes in human brain activity can be induced by PBM, but that the changes depend on the experimental protocol and the biological context in a highly variable way [1, 5]. These observations have identified optical neuromodulation as a growing field of research and identified an underlying problem: there is still

a lack of understanding about the quantitative link between the stimulus parameters and the neuronal response.

PBM's therapeutic and neuromodulatory properties are very sensitive to the optical parameters of the stimulus. The quantity of optical energy delivered to neural tissue can depend on several factors including the wavelength, the irradiance, the time of exposure, the fluence, the structure of the pulse, the depth of the tissue to which the light is delivered and the optical delivery configuration. Experimental evidence shows these parameters are not interchangeable. For instance, extending transcranial stimulation to 670 nm has been shown to induce measurable changes in ATP-related metabolic processes in older subjects [6] and comparative experiments [11] have shown differences in brain metabolic activity depending on the wavelength of transcranial stimulation. There are other studies examining 808-nm stimulation and/or exposure to visible light, which also suggest that different spectral regimes may have different effects on the brain [13]. Furthermore, clinical and experimental studies have found that PBM can have a positive impact in a variety of neurological settings, including cognitive impairment, traumatic brain injury, neurodegenerative diseases, and other contexts of clinical use. [4,7,12,14] Together, these results indicate that the wavelength should not be considered just as a treatment label, but as a basic physical parameter in a multi-dimensional dose–response relationship with light. The substantial variability in experimental PBM protocols and devices further underscores the importance of optical parameters. The systematic review of brain-PBM devices was extensive and showed significant differences in the wavelength used, power, irradiance, treatment duration, delivery geometry and stimulation configuration reported in published studies [17]. This heterogeneity makes it hard to compare experiments and can lead to apparently conflicting biological results. In fact, while the majority of investigations have shown positive effects with PBM, certain studies have also shown null or even limited effects under specific conditions [10]. This range of variation makes it impossible to make confident inferences of neuronal response from the nominal use of PBM. Rather, the biological reaction might be a nonlinear function of wavelength, optical dose, nature of the tissue, exposure conditions, and state of the target neural system. The molecular and cellular processes that link optical stimulation to modulation of the nervous system are inherently multi-scale. When photons enter biological tissue they are absorbed and scattered at specific wavelengths, and deposition of the photon energy within cells may affect mitochondrial energy production, redox signaling, synaptic function, membrane excitability and intracellular signaling pathways. PBM has been linked to modulation of microglial activity and drainage processes in the brain [8], β -amyloid clearance mechanisms [9] and neurochemical alterations in experimental models of Parkinsonian pathology [10]. PBM has also been found to affect the organization of the synapses and function of the mitochondria in depressive-like behavior models [15] and to improve neurological function after repeated head injury [16]. The results indicate that optical stimulation may induce a cascade of tissue–light interactions that can result in complex neural responses at the system and cellular levels. One pathway that is becoming more significant is modulation of neuronal ion channels. Evidence that PBM can alter ion channels within the nervous system over a wide range of optical wavelengths has been identified in a systematic review, which is a link between optical energy and neuronal excitability [18]. Complementary experimental data have demonstrated that exposure to a near infra-red laser in a cultured primary sensory neuron depolarized the cell and induced a change in the biophysical properties of human Nav1.5 sodium channels [24]. These observations support the idea that optical stimulation might have direct effects on electrophysiological behavior at the cellular level in addition to indirect metabolic or vascular effects. More recently, it has been shown that PBM could modulate neuronal injury via AMPK/ULK1 signaling pathway [28]; and wavelength-dependent PBM has been correlated to increase of synaptic plasticity and decrease of cortical hyperexcitability in neurological dysfunction models [25]. These findings together strongly suggest that optical stimulation of the brain is capable of inducing measurable and mechanistically significant changes in neuronal activity which depend on wavelength. At the systems level, PBM has also been linked to measurable changes in the physiology of the cortex and functional organization of the brain. Repeated PBM is effective in decreasing the bilateral

cortical hemodynamic activation during working memory tasks in healthy older adults [5] and transcranial PBM has been shown to alter prefrontal cortical networks organization by directed physiological network analysis [19]. Recently, transcranial PBM has been shown to improve functional brain networks and working memory functioning in healthy older adults using functional near-infrared spectroscopy (fNIRS) measurements [21]. PBM has also been studied for its effects on cerebrospinal-fluid dynamics [26] suggesting that optical stimulation can have effects on a variety of physiological pathways other than local neuronal excitability. It was concluded from these findings that optical stimulation is capable of inducing measurable changes at different levels, molecular, cellular, regional and whole brain. Even with these advances, the majority of the current PBM investigations are still experimentally limited. Several studies have investigated one or a few wavelengths, one or a few exposure times, or a small range of irradiance and fluence. Thus, the literature is replete with evidence that PBM can affect neural physiology, but has relatively little capacity to predict the neuronal response to combinations of optical parameters that have not yet been tested. The results of the experiments also demonstrate a non-linear, context-dependent optical-to-neural relationship because experimental results both positive and null exist [10]. This is a great motivation to develop computational models that can learn complex response surfaces as a function of wavelength, intensity, exposure duration, tissue depth, and other physical parameters. One promising approach to this challenge is through machine learning (ML) that is able to combine multiple optical, physical and biological properties in one nonlinear learning algorithm. Importantly, the start of prediction using machine learning for PBM is beginning to emerge. Lee et al. have created a machine-learning prediction model for the effects of PBM in older adults with cognitive decline based on fNIRS [20]. Subsequently, Gonzalez-Lima et al. showed that ML based prediction of the cognitive responses to transcranial prefrontal PBM can be achieved in people with BD using fNIRS measures [22]. The studies confirm that machine learning is able to uncover predictive relationships between complex PBM-related neurophysiological data. They are mostly focused on forecasting the observed physiological or cognitive response based on experimentation rather than explicitly modelling the physical processes between the optical stimulation parameters and the neuronal response. One of the key challenges of purely data-driven ML is then the ability to ensure that predictions are not only accurate but also physically plausible. A model can be highly accurate if it learns statistical correlations of the dataset available that reflect the processes of the model, but not the physical mechanisms that control photon energy, tissue attenuation, optical penetration or the deposition of energy at different wavelengths. This is especially relevant in optical neuromodulation where the delivered amount of optical energy at the light source may be markedly different from the actual optical dose to neural tissue. Therefore, a good black-box model might not be able to predict good results for new wavelengths, exposure, tissue, and/or stimulation. One way to overcome this limitation is to leverage the known physics in the learning process, which is known as physics-informed machine learning (PIML). While the relationship between input variables and the response of a neuron could simply be a statistical map, physics-informed approaches can introduce variables or equations known from physics, mechanisms, or constraints in the map. Recent studies have shown physics-based deep learning for imaging neuronal activity with a two-photon and light-field microscope [32] which made a direct connection between physics-based optical modeling and reconstruction of neuronal activity. In the same manner, model-based explainable deep learning has been introduced in light-field microscopy, showcasing the seamless combination of physical knowledge and explainability in the optical computational model [33]. Additionally, this methodological approach is supported by the broader development of physics-guided learning in photonics. Physics-informed neural networks have been used for wavefront correction [34] and physics-aware analytic-gradient methods have been developed for photonic neural networks [35]. Physics-guided machine learning has also been explored on optical-spectrum inversion [36] and building a digital twin of an optically turbid medium [37]. The development of physics-informed AI, from Maxwell equations to nanophotonics and electromagnetics has been summarized recently [38]. Physics-informed learning has also been used for imaging scattering media [39] and optical phased-array calibration [40] as well as nonlinear optical wave propagation [41] and learning of general dynamical systems [42]. The developments show that physical constraints can enhance the scientific relevance and generalizability of machine-learning models in complex optical systems. Explicit

consideration of the propagation of light through a medium is especially critical for neuronal modulation of optical effects. Optical energy received by a neural target is influenced by the tissue optical properties, depth of penetration, absorption, scattering, and geometry, as well as the wavelength and power of the incident optical energy. Recently, simulation-based dosimetry has shown that the wavelengths, power density, and skin properties of the transcranial and intranasal PBM [23] can be important factors in determining the optical dose delivered to the human brain during such procedures. Thus, the nominal irradiation parameters could not represent the effective stimulus to the neuronal tissue. The inclusion of physical parameters like photon energy, frequency, tissue depth, optical attenuation, exposure dose and other optical parameters, therefore, can represent the biological stimulation process more meaningfully. A second limitation of current predictive models is interpretability: they are difficult to understand. Interpretability is another limitation of the existing predictive models—they are hard to understand. Modern nonlinear ML algorithms can model the complicated interactions between wavelength, response and exposure time correctly but these nonlinearities can obscure the resulting predictions. Predictive accuracy is not enough in biomedical applications if the purpose is to comprehend the effect of physical variables on a biological outcome. This limitation can be overcome by using explainable artificial intelligence (XAI), which quantifies the contribution of individual predictors and uncovers some nonlinear relationships between model inputs and predicted neuronal responses. Recent advances show the usefulness of XAI in neural and biomedical data. The applicability of feature-attribution methods to electrophysiological measurements has been shown by using SHAP-based ERP analysis to enhance the sensitivity and interpretation of EEG signals [43]. In terms of explainable AI in healthcare, existing research has focused on the requirement of useful explanations to give meaningful information about the reasoning for health-related predictions and not simply on the contents of the model output [44]. This is especially appealing for optical neuromodulation due to the potential to be able to explain what the model predicts, and why the model predicts it, by combining physics-informed learning and explainability. The current trend of integrating AI with optical neural technologies further emphasizes the importance of such a comprehensive approach. In recent years, there have been exciting developments involving the integration of AI and optogenetics, highlighting the increasing convergence of computational intelligence and optical neuromodulation [22]. Optogenetic systems and PBM, however, are two completely different paradigms of optical/neuronal stimulation, and a lack of common computational methods clearly accounting for the wavelength dependence of the response of neuronal systems to non-invasive optical stimulation remains. Previous work has been mainly conducted on experimental PBM success [1–19], on the prediction of physiological or cognitive outcomes using ML [20,21] or on the physics-informed optical reconstruction and imaging [32–42]. The streams of research have been very little brought together in one framework for predicting neuronal modulation as a function of wavelength. Another challenge is the conversion of optical parameters to biologically meaningful response variables. Photon energy is determined by the optical wavelength, and absorption and scattering in tissue determines the propagation and spatial distribution of photons. The delivered optical dose depends on exposure duration and intensity, while changes in excitability and activity depend on the mechanisms of cells and neurons. Therefore, the relationship of this process can be summarized as a hierarchy: optical stimulus – photon/tissue interaction – effective optical dose – cellular mechanisms – neuronal activity. A model that only includes one step of this cascade can only give partial mechanistic understanding. A physics-informed approach, however, can explicitly represent intermediate physical descriptors, enabling the ML model to learn nonlinear relationships between these physical properties and neuronal responses. This method not only lessens reliance on empirical correlations but also allows for predictions over a wider range of optical parameters. Considering this, there is a significant gap in the research regarding the relationship between optical neuromodulation via wavelength, physics-informed machine learning, and explainable artificial intelligence. Previous experiments demonstrate that PBM can affect neural metabolism, hemodynamics, ion-channel activity, synaptic function and behavioral outcomes [1-19, 24-31]. Past ML research has shown that neurophysiological and cognitive responses associated with PBMs can be predicted by measured data [20, 21]. As for physics-informed optical ML, it has shown that physical knowledge can be incorporated into complicated optical systems [32, 33, 39, 40, 42]. For XAI, a number of approaches have been proposed that

provide interpretable analysis of neural signals [43, 44]. However, none of these advances is yet fully included in a consistent predictive framework for the evolution of neuronal activity as a function of wavelength and other physically meaningful parameters of the optical stimulation. Given this, this study investigates a Physics-Informed Explainable Machine Learning (PIEML) framework to predict wavelength dependent optical modulation of neuronal activity. The proposed framework couples together wavelength, tissue intensity, exposure time, tissue depth and other stimulation parameters and physics-derived descriptors of photon and tissue-light interactions. A machine-learning model is used to learn nonlinear relationships between the resulting physical feature space and neuronal activity, and explainable ML is used to quantify the contributions of the features, identify nonlinear parameter interaction, and characterize wavelength-sensitive response regimes. The framework aims to be predictive, physically plausible and mechanistic interpretable. The main idea is that the neuronal optical modulation can be represented with better accuracy when the predictive model is based on physical properties of the optical image and not only on the nominal values of the experimental parameters. Here wavelength is not just assigned as an experimental parameter but is defined as a fundamental parameter of the photon energy and interaction with tissue. Thus the resulting model can give a quantitative representation of the relationship between optical stimulation and neuronal response, and the physical variables which have the greatest influence on predicted activity. There are a number of implications to the proposed approach. First, it can give a systematic representation of the neuronal response as a function of wavelength over a multidimensional stimulation space, where the response is computed in a computational way. Second, physics-informed feature construction can help make learned relationships more interpretable to physics and help potentially promote transferability of the model beyond the context of the training set. Third, the explainability analysis can be used to detect the most important optical and biological predictors and to uncover nonlinear interactions that are not apparent in traditional statistical methods. Lastly, the resulting framework could serve as a template for computational optimization of optical stimulation protocols for the development of more reproducible, interpretable and potentially personalized, non-invasive neuromodulation strategies.

The current research aims to fill the gap between three different domains of research that are traditionally divided: optical neuromodulation, physics-informed artificial intelligence, and explainable machine learning within a single predictive framework. The proposed approach is aimed at going beyond descriptive characterization of PBM to predict, physically inform and interpret wavelength-dependent neuronal activity by explicitly connecting optical wavelength, stimulation parameters and the physics-derived descriptors and neuronal response.

2. Materials and Methods

2.1 Study Design and Computational Framework

A physics-informed explainable machine-learning framework was developed to predict wavelength-dependent optical modulation of neuronal activity. The overall workflow integrated optical excitation parameters, tissue optical properties, wavelength-dependent biological sensitivity, neuronal baseline characteristics, and neuronal response variables within a leakage-controlled supervised learning framework. The computational workflow comprised six major stages: (i) dataset preparation and quality control; (ii) physical consistency verification; (iii) physics- and biology-informed feature construction; (iv) leakage-controlled model development and wavelength-based validation; (v) predictive benchmarking and physics-informed residual learning; and (vi) explainability, uncertainty quantification, robustness assessment, and wavelength-response analysis.

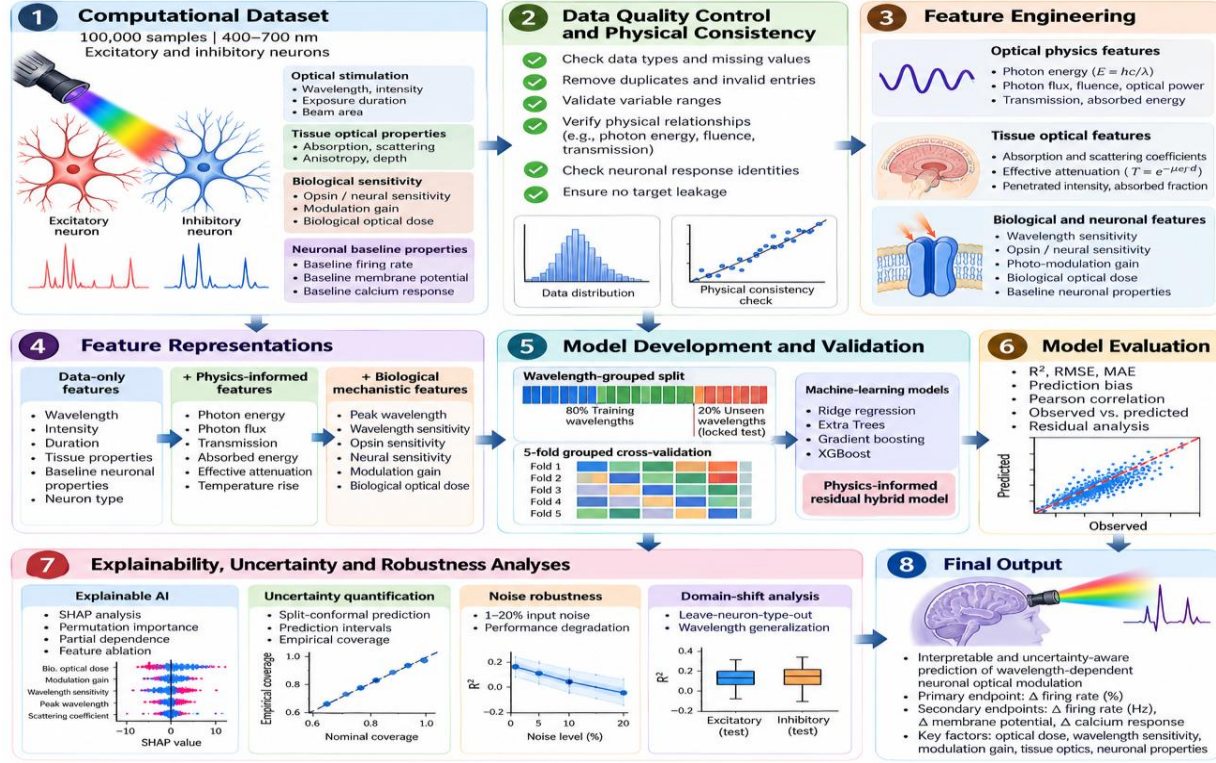


Figure 1. Overall Workflow of the Proposed Physics-Informed Explainable Machine-Learning Framework for Wavelength-Dependent Neuronal Optical Modulation (AI generated)

Figure 1. Overall process of the proposed physics-informed explainable machine learning approach for wavelength-dependent neuronal optical modulation. The workflow combines computational optical-neuronal data with physical consistency checking, leakage-controlled feature engineering and increasingly complete data-only, physics-informed and biological-mechanistic feature representations. The models are developed based on the wavelength-grouped training and cross-validation, and a subset of the wavelengths is locked as a test domain to evaluate the wavelength-domain generalization. Complementary regression metrics are used to assess multiple machine-learning algorithms alongside a physics-informed residual hybrid model. SHAP analysis, permutation importance, partial-dependence analysis, feature ablation, conformal prediction, input-noise perturbation and leave-neuron-type-out testing are then used to further evaluate model interpretation and reliability. The final framework offers an interpretable and uncertainty-aware method to identify optical and biological factors correlated with the wavelength-dependent neuronal responses.

2.2 Computational Dataset

The analysis used a computationally generated dataset containing 100,000 observations and 44 original variables describing optical stimulation, tissue optical interactions, biological wavelength sensitivity, neuronal baseline properties, and neuronal responses. The wavelength range covered 400-700 nm and included excitatory and inhibitory neuronal phenotypes. The optical stimulation variables included wavelength, incident intensity, exposure duration, beam area, optical power, fluence, photon energy, photon flux, and total photon number. Tissue-interaction variables included tissue depth, absorption coefficient, scattering coefficient, anisotropy, effective attenuation, transmission, penetrated intensity, absorbed fraction, and absorbed optical energy. Biological-response descriptors included peak wavelength, wavelength sensitivity, opsin sensitivity, neural sensitivity, photo-modulation gain, depth modulation factor, effective optical dose, and biological optical dose. Neuronal baseline characteristics included baseline firing rate, baseline calcium response, baseline membrane potential, membrane resistance,

membrane time constant, and neuronal phenotype. The complete dataset characteristics and Statistics of Principal Optical, Biological, and Neuronal Variables are summarized in Table 1 and Table 2.

Table 1. Dataset Characteristics

Characteristic	Value
Total observations	100,000
Original variables	44
Unique wavelengths	56
Wavelength range	400–700 nm
Neuronal phenotypes	2
Neuronal types	Excitatory, inhibitory
Primary endpoint	Δ firing rate (%)
Secondary endpoints	Δ firing rate (Hz), Δ membrane potential (mV), Δ calcium
Optical stimulation variables	Wavelength, intensity, exposure, beam area, optical power, fluence
Tissue optical variables	Absorption, scattering, anisotropy, attenuation, transmission
Biological variables	Wavelength sensitivity, opsin sensitivity, neural sensitivity, modulation gain
Neuronal baseline variables	Baseline firing rate, membrane potential, calcium, membrane resistance, time constant
Validation strategy	Wavelength-grouped training/testing + 5-fold grouped CV
Locked test domain	Unseen wavelengths
Primary ML model	XGBoost
Explainability methods	SHAP, permutation importance, partial dependence
Uncertainty method	Split-conformal prediction
Robustness analysis	1–20% input-noise perturbation
Domain-shift analysis	Leave-neuron-type-out validation

Table 2. Descriptive Statistics of Principal Optical, Biological, and Neuronal Variables

Variable	Mean	SD	Min	Q1	Median	Q3	Max
Wavelength (nm)	538.948	82.994	400.000	470.000	535.000	605.000	700.000
Intensity (mW/mm ²)	1.855	1.340	0.083	0.970	1.502	2.327	20.000
Exposure time (s)	301.107	173.108	1.001	150.877	301.548	450.937	599.988
Tissue depth (mm)	2.553	1.412	0.100	1.333	2.563	3.770	5.000
Photon energy (eV)	2.357	0.370	1.771	2.049	2.318	2.638	3.100
Transmission	0.382	0.250	0.006	0.173	0.330	0.562	0.980
Biological optical dose	0.0115	0.0302	0.0000	0.0000	0.0005	0.0090	1.1470
Wavelength sensitivity	0.326	0.358	0.000	0.003	0.158	0.644	1.000
Opsin sensitivity	0.326	0.368	0.000	0.003	0.154	0.620	1.299
Neural sensitivity	1.033	0.263	0.331	0.845	1.002	1.185	3.167
Δ firing rate (%)	0.407	8.175	−68.442	−5.040	0.360	5.777	80.486
Δ firing rate (Hz)	0.024	0.762	−6.125	−0.434	0.030	0.490	4.823
Δ membrane potential (mV)	0.055	0.503	−8.103	−0.247	0.035	0.323	9.760
Δ calcium response	0.0019	0.0354	−0.1573	−0.0219	0.0018	0.0254	0.2420

2.3 Data Quality Control

Before model development, the dataset underwent systematic quality-control procedures. Data types, missing values, infinite values, duplicated observations, variable ranges, and categorical distributions were examined. Continuous variables were inspected for physically implausible values and numerical inconsistencies. The number of unique wavelengths and the distribution of neuronal phenotypes were also assessed to ensure that the dataset represented the intended experimental parameter space. No target-derived variables were allowed to enter the predictor matrix. In particular, stimulated neuronal response variables and mathematically derived response variables were excluded from model predictors. This leakage-control procedure was considered essential because several variables in the dataset are deterministic transformations of other variables.

2.4 Physical Consistency Verification

The physical validity of the computational dataset was evaluated using independent analytical relationships. Photon energy was calculated from wavelength according to $E_{ph} = \frac{hc}{\lambda}$, where E_{ph} is photon energy, h is Planck's constant, c is the speed of light, and λ is wavelength. For wavelength expressed in nanometers, photon energy was additionally represented in electron-volts as $E_{ph} (eV) = \frac{1239.84198}{\lambda(nm)}$. Optical fluence was independently verified from incident intensity and exposure duration: $F = \frac{It}{1000}$, where I is incident intensity in $mWmm^{-2}$, t is exposure duration in seconds, and the factor 1000 converts milliwatts to watts. Optical power was checked using $P = IA$, where A represents the illuminated beam area. For homogeneous effective attenuation, optical transmission was independently verified using a Beer–Lambert-type relationship: $T = \exp(-\mu_{eff}d)$, where μ_{eff} is the effective attenuation coefficient and d is tissue depth. The neuronal response variables were also checked for algebraic consistency. Specifically, the stimulated response was verified according to $R_{stim} = R_{baseline} + \Delta R$, where R_{stim} is the stimulated neuronal response, $R_{baseline}$ is the baseline response, and ΔR represents the absolute change induced by optical stimulation. The corresponding relative change was calculated as $\Delta R_{\%} = 100 \frac{\Delta R}{R_{baseline}}$. Maximum numerical deviations from these analytical relationships were calculated and reported in Table 3.

Table 3. Physical Consistency Verification of the Computational Dataset

Physical/Mathematical Relationship	Verification Equation	Maximum Absolute Error	Interpretation
Photon energy–wavelength relationship	$E_{ph} = \frac{1239.841984}{\lambda_{nm}}$	$\approx 4.44 \times 10^{-16} \text{eV}$	Physically consistent
Fluence–intensity–exposure relationship	$F = \frac{It}{1000}$	$3.55 \times 10^{-15} \text{J/mm}^2$	Physically consistent
Transmission–attenuation–depth relationship	$T = e^{-\mu_{eff}d}$	$\approx 3.33 \times 10^{-16}$	Physically consistent
Stimulated firing-rate identity	$R_{stim} = R_{baseline} + \Delta R$	$\approx 8.88 \times 10^{-16} \text{Hz}$	Algebraically consistent
Percentage firing-rate change	$\Delta R_{\%} = 100 \frac{\Delta R}{R_{baseline}}$	$\approx 8.88 \times 10^{-16} \text{Hz}$	Algebraically consistent

2.5 Target Definition

The primary endpoint was the percentage change in neuronal firing rate:

$$\Delta \text{Firing}_{\%} = 100 \times \frac{\text{Firing}_{stim} - \text{Firing}_{baseline}}{\text{Firing}}$$

This endpoint represents the relative change in neuronal firing following optical stimulation. Three secondary endpoints were additionally investigated:

$$Y_1 = \Delta \text{Firing}_{Hz}, \text{ and } Y_3 = \Delta \text{Calcium}.$$

$$Y_2 = \Delta V_m,$$

The changes in firing rate, membrane potential, and calcium response were defined as:

$$\Delta \text{Firing}_{Hz} = \text{Firing}_{stim} - \text{Firing}_{baseline},$$

$$\Delta V_m = V_{m, stim} - V_{m, baseline},$$

and $\Delta \text{Calcium} = \text{Calcium}_{Response} - \text{Calcium}_{Baseline}$. The use of changes rather than stimulated absolute values for the secondary endpoints was intended to reduce dependence on baseline magnitude and provide a more physiologically interpretable measure of optical modulation.

2.6 Leakage-Controlled Feature Engineering

To prevent target leakage and distinguish the contribution of raw experimental variables from physics-informed information, predictors were organized into three feature representations.

2.6.1 Data-only feature representation

The data-only representation contained variables that can be regarded as direct optical-stimulation, tissue, and neuronal baseline descriptors: wavelength, incident intensity, exposure duration, beam area, tissue depth, absorption coefficient, scattering coefficient, anisotropy, baseline firing rate, baseline calcium, baseline membrane potential, membrane resistance, membrane time constant, and neuron type. Neuron type was numerically encoded for machine-learning models.

2.6.2 Physics-informed feature representation

The second representation augmented the data-only variables with optical-physics descriptors, including: photon energy, photon energy in joules, optical frequency, photon flux, total photon number, reduced scattering coefficient, effective attenuation, transmission, penetrated intensity, absorbed fraction, absorbed energy, and temperature rise. These variables were used to provide the machine-learning models with physically meaningful representations of wavelength-dependent optical propagation and energy deposition.

2.6.3 Full mechanistic representation

The full mechanistic feature set further incorporated biological optical-response descriptors: peak wavelength, sensitivity width, wavelength sensitivity, opsin sensitivity, neural sensitivity, photo-modulation gain, depth modulation factor, effective optical dose, and biological optical dose. The three representations enabled direct assessment of the incremental predictive contribution of physics-derived and biological mechanistic variables. Variables directly representing stimulated neuronal responses were strictly excluded from all predictor matrices.

2.7 Exploratory Data Analysis

Exploratory analysis was conducted to characterize distributions, relationships, and wavelength-dependent response patterns before predictive modeling. Univariate distributions were examined for the primary and secondary endpoints. Pearson correlation analysis was used to quantify linear associations among optical, biological, and neuronal variables. Wavelength-response profiles were calculated separately for neuronal phenotypes using wavelength-specific means and standard errors. These analyses were used to identify systematic wavelength-dependent changes in neuronal firing and membrane potential.

2.8 Statistical Effect-Size Analysis

To quantify the contribution of wavelength and neuronal phenotype to firing modulation, a two-way analysis of variance was performed using neuronal firing modulation as the dependent variable and wavelength, neuron type, and their interaction as explanatory factors. Wavelength was treated as a categorical factor to evaluate differences among the discrete wavelength levels represented in the dataset.

The corresponding F-statistics and significance values were calculated. Partial eta-squared was used as an effect-size measure: $\eta^2 = \frac{SS_{effect}}{SS_{effect} + SS_{error}}$. Here, η^2 is eta squared (effect size), SS_{effect} is the sum of squares attributable to the effect, and SS_{error} is the residual/error sum of squares.

2.9 Leakage-Controlled Training and Test Strategy

A major methodological objective was to evaluate whether models could generalize to wavelengths not observed during training. Rather than relying exclusively on a conventional random split, a wavelength-grouped train/test strategy was adopted. Wavelength groups were divided into approximately 80% training and 20% locked testing groups using group-based random splitting. All observations associated with a wavelength were assigned to the same partition. Consequently, the locked test set contained wavelengths that were not used for model fitting. This strategy provides a substantially more stringent test of wavelength-domain generalization than random row-level splitting because neighboring observations from the same wavelength cannot simultaneously occur in training and testing subsets. A conventional random 80/20 split was also evaluated as a comparator to quantify the potential optimism introduced by non-grouped validation.

2.10 Wavelength-Grouped Cross-Validation

Five-fold grouped cross-validation was subsequently performed within the development dataset. Wavelength was used as the grouping variable, ensuring that observations belonging to the same wavelength remained within a single fold. For each fold, the model was trained on four wavelength groups and evaluated on the remaining wavelength group. Mean and standard deviation of R^2 , RMSE, and MAE were calculated across folds. The locked wavelength test set was not used during cross-validation or model selection. This nested separation of development, grouped cross-validation, and locked testing was designed to reduce overestimation of generalization performance.

2.11 Machine-Learning Models

Four regression algorithms were benchmarked: Ridge regression, Extra Trees regression, histogram-based gradient boosting, and extreme gradient boosting (XGBoost). Ridge regression was included as a regularized linear baseline. Extra Trees and gradient boosting models were included to capture nonlinear relationships and feature interactions. XGBoost was selected as the principal nonlinear ensemble model because of its ability to represent complex nonlinear relationships between wavelength, optical dose, tissue properties, and neuronal characteristics. Model hyperparameters were fixed using the predefined computational pipeline and evaluated consistently across feature representations.

2.12 Prediction Performance Metrics

Model performance was quantified using the coefficient of determination R^2 , root mean squared error (RMSE), mean absolute error (MAE), prediction bias, and Pearson correlation. The coefficient of determination was calculated as $R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$. RMSE was calculated as $RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$. MAE was calculated as $MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i|$. Pearson correlation stated as: $r = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}}$. Here, y_i denotes the observed value, $\hat{y}_i$ denotes the model-predicted value, $\bar{y}$ is the

mean observed value, $\bar{\hat{y}}$ is the mean predicted value, and n is the number of observations. These complementary metrics were used because R^2 alone does not adequately describe absolute prediction error, correlation strength, or systematic over- and under-prediction.

2.13 Physics-Informed Residual Learning

To explicitly integrate mechanistic information into the predictive framework, a physics-informed residual model was developed. A transparent mechanistic prior was constructed from the biological optical dose and photo-modulation gain: $P_{\text{mech}} = D_{\text{bio}} G_{\text{photo}}$, where D_{bio} represents the biological optical dose and G_{photo} represents the photo-modulation gain. A linear model was first fitted between this mechanistic prior and the primary neuronal response:

$$\hat{Y}_{\text{mech}} = f(P_{\text{mech}}),$$

where $\hat{Y}_{\text{mech}}$ denotes the response predicted by the mechanistic model. The residual response was then defined as $r_i = Y_i - \hat{Y}_{\text{mech},i}$, where Y_i is the observed neuronal response and $\hat{Y}_{\text{mech},i}$ is the corresponding mechanistic prediction. XGBoost was subsequently trained to predict this residual using the complete mechanistic feature representation: $\hat{r}_{\text{ML},i} = f_{\text{XGB}}(\mathbf{X}_i)$, where $\mathbf{X}_i$ represents the complete feature vector for observation i . The final prediction was obtained by combining the mechanistic prediction with the machine-learning residual correction: $\hat{Y}_i = \hat{Y}_{\text{mech},i} + \hat{r}_{\text{ML},i}$. This formulation allows the machine-learning model to learn nonlinear corrections to a transparent mechanistic prior rather than replacing the physical/biological representation entirely.

2.14 Feature Ablation Analysis

Ablation experiments were performed to quantify the incremental contribution of physics-informed and biological mechanistic variables. Three feature configurations were compared: data-only, data-only + physics-derived features, and data-only + physics-derived + biological mechanistic features. All models were evaluated using the same locked wavelength test set and XGBoost architecture. The improvement in R^2 , RMSE, and MAE across feature configurations was used to quantify the contribution of progressively richer physical and biological representations.

2.15 Explainable Artificial Intelligence

Model interpretability was investigated using complementary global explanation methods.

2.15.1 SHAP analysis

SHapley Additive exPlanations (SHAP) were calculated using a tree-based explainer for the final XGBoost model. A maximum of 3,000 randomly selected test observations was used for computational efficiency. For each feature, the mean absolute SHAP value was calculated as $I_j^{\text{SHAP}} = \frac{1}{N} \sum_{i=1}^N |\phi_{ij}|$, where ϕ_{ij} is the SHAP contribution of feature j for observation i .

2.15.2 Permutation importance

Permutation importance was independently calculated on the locked test set using repeated feature permutation. $I_j^{\text{perm}} = R_{\text{baseline}}^2 - R_{\text{permuted},j}^2$, where R_{baseline}^2 is the coefficient of determination using the original test set and $R_{\text{permuted},j}^2$ is the coefficient of determination after permuting feature j . The decrease in predictive R^2 following feature permutation was used as a model-agnostic measure of feature importance.

2.15.3 Partial dependence

Partial-dependence analysis was performed for major optical and mechanistic variables, including wavelength, intensity, exposure duration, tissue depth, and biological optical dose and can be calculated as

$PD_j(x_j) = \frac{1}{N} \sum_{i=1}^N f(x_j, X_{i,-j})$ where $f(\cdot)$ is the trained prediction function and $X_{i,-j}$ represents all features other than x_j .

2.16 Prediction Uncertainty and Conformal Calibration

Predictive uncertainty was quantified using split-conformal prediction. The non-test development dataset was divided into a model-fitting subset and an independent calibration subset using wavelength-grouped splitting. Absolute calibration residuals were calculated as Calibration residual: $s_i = |y_i - \hat{y}_i|$ 90% conformal residual threshold: $q_{0.90} = Q_{0.90}(s_1, s_2, \dots, s_N)$, Prediction interval: $C(x) = [\hat{y}(x) - q_{0.90}, \hat{y}(x) + q_{0.90}]$, Empirical coverage: $Coverage = \frac{1}{N} \sum_{i=1}^N 1[y_i \in C(x_i)]$, Mean prediction-interval width: $MPIW = \frac{1}{N} \sum_{i=1}^N (U_i - L_i)$, where U_i and L_i are the upper and lower prediction bounds, respectively.

2.17 Noise-Robustness Analysis

The robustness of the final model to measurement uncertainty was assessed by perturbing physical input variables with multiplicative Gaussian noise. Noise levels of 1%, 5%, 10%, and 20% were evaluated. The perturbed features included incident intensity, exposure duration, tissue depth, absorption coefficient, scattering coefficient, and anisotropy. The trained model was evaluated without retraining under each perturbation level. Changes in R^2 , RMSE, MAE, and bias were recorded. Perturbed physical feature: $x^* = x(1 + \epsilon)$, where the multiplicative noise term is $\epsilon \sim \mathcal{N}(0, \sigma^2)$ and σ represents the specified relative noise level. This analysis was intended to approximate the sensitivity of predictions to imperfect optical and tissue-property measurements.

2.18 Wavelength-Response Landscape Analysis

To investigate the learned wavelength dependence, test-set observations were aggregated by wavelength. For each wavelength, the mean observed firing modulation and mean model prediction were calculated. Median biological optical dose, photo-modulation gain, and transmission were additionally summarized at each wavelength to provide mechanistic context. The resulting observed-versus-predicted wavelength-response landscape was used to identify wavelength regions associated with stronger or weaker predicted modulation. Mean observed response at wavelength λ_k : $\bar{Y}_{obs}(\lambda_k) = \frac{1}{N_k} \sum_{i:\lambda_i=\lambda_k} Y_i$, Mean predicted response: $\bar{Y}_{pred}(\lambda_k) = \frac{1}{N_k} \sum_{i:\lambda_i=\lambda_k} \hat{Y}_i$, Median biological optical dose: $\widetilde{D}_{bio}(\lambda_k) = median[D_{bio,i} : \lambda_i = \lambda_k]$, Median photo-modulation gain: $\widetilde{G}_{photo}(\lambda_k) = median[G_{photo,i} : \lambda_i = \lambda_k]$, Median transmission: $\widetilde{T}(\lambda_k) = median[T_i : \lambda_i = \lambda_k]$.

2.19 Secondary Endpoint Analysis

To assess whether the framework generalized across different neuronal response measures, independent XGBoost models were trained for: absolute firing-rate change, membrane-potential change, and calcium-response change. The same wavelength-grouped locked-test strategy and mechanistic feature representation were used. Performance was evaluated using R^2 , RMSE, MAE, bias, and Pearson correlation. For the three secondary endpoints: $\hat{Y}_k = f_{XGB,k}(X)$, $k \in \{1,2,3\}$ where, $Y_1 = \Delta Firing_{Hz}$, $Y_2 = \Delta V_m$, $Y_3 = \Delta Calcium$

2.20 Leave-Neuron-Type-Out Generalization

A domain-shift experiment was performed to evaluate whether the learned relationships generalized between neuronal phenotypes. In each experiment, one neuronal type was completely excluded from model training and used as the held-out evaluation domain. The model was trained using the remaining neuronal type and subsequently evaluated on the excluded type. This analysis provides a stringent test of phenotype transferability and complements the wavelength-domain generalization experiment. Training dataset excluding neuronal type c : $\mathcal{D}_{train}^{(-c)} = \{(x_i, y_i): c_i \neq c\}$, Held-out evaluation dataset: $\mathcal{D}_{test}^{(c)} = \{(x_i, y_i): c_i = c\}$. The model is therefore trained on $\mathcal{D}_{train}^{(-c)}$ and evaluated exclusively on $\mathcal{D}_{test}^{(c)}$.

2.21 Reproducibility and Computational Implementation

All analyses were implemented in Python using NumPy, pandas, SciPy, scikit-learn, XGBoost, SHAP, statsmodels, Matplotlib, and Seaborn. A fixed random seed was used throughout the analysis to improve reproducibility. All data partitions, model training procedures, feature definitions, evaluation metrics, figures, tables, predictions, and trained models were generated programmatically. The complete computational workflow was organized so that the locked wavelength test set remained isolated from model development and cross-validation. Final model predictions, uncertainty estimates, feature-importance results, and manuscript tables were exported as machine-readable CSV files, while publication figures were exported at high resolution.

2.22 Statistical and Scientific Reporting Considerations

All predictive analyses were interpreted according to their corresponding validation design. In particular, performance obtained from conventional random splitting was not considered equivalent to performance obtained from wavelength-grouped validation. The principal evidence for generalization was therefore based on the locked unseen-wavelength test set and wavelength-grouped cross-validation. Feature importance was interpreted as predictive association rather than causal evidence. Similarly, high performance for any endpoint was not interpreted as experimental confirmation of neuronal optical modulation because the underlying dataset was computationally generated. The combination of physical consistency testing, leakage-controlled validation, feature ablation, explainability, uncertainty calibration, robustness analysis, and domain-shift evaluation were used to provide a multidimensional assessment of the proposed physics-informed machine-learning framework.

3. Results and Discussion

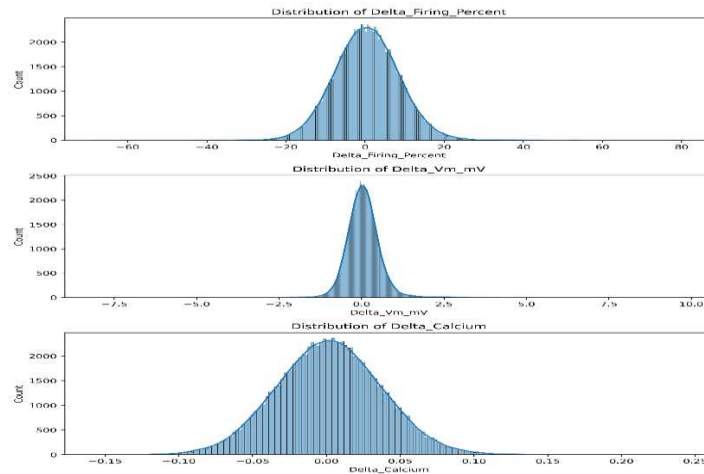


Figure 1. Distribution of neuronal response endpoints in the computational dataset.

Figure 1 shows the distributions of the three main neuronal response endpoints: percentage change in firing rate, change in membrane potential and change in intracellular calcium. The distribution of the primary endpoint, ΔFiring (%), was fairly wide and showed little sign of either a maximum or a minimum value, with a mean of 0.407% and a standard deviation of 8.175%. The range observed was from -68.44% to 80.49%, suggesting a large spread of neuronal responses in the model. The variance around a zero mean value is relatively large indicating that optical stimulation does not have a uniformly excitatory or inhibitory response in the simulated parameter space. The distributions of the secondary endpoints were much more tightly clustered. Changes in firing rate, membrane potential, and calcium concentration were 0.024 Hz, 0.055 mV, and 0.0019, respectively, on average. Positive responses along with negative responses further support bidirectional modulation instead of a monotonic increase in neuronal activity. This heterogeneity is an important reason why to use distribution-aware metrics and external/generalization analyses as evaluation methods for the model, in addition to the global correlation or R^2 . Notably, the wide range of responses also sets the challenge of the prediction problem. Conventional regression models may be more likely to induce training that reflects the more dominant portion of the observations that have relatively small response values, which may result in less representation of the more extreme responses. Therefore, Figure 1 serves as an important reference point to interpret the residual structure and modest predictive power in subsequent analyses.

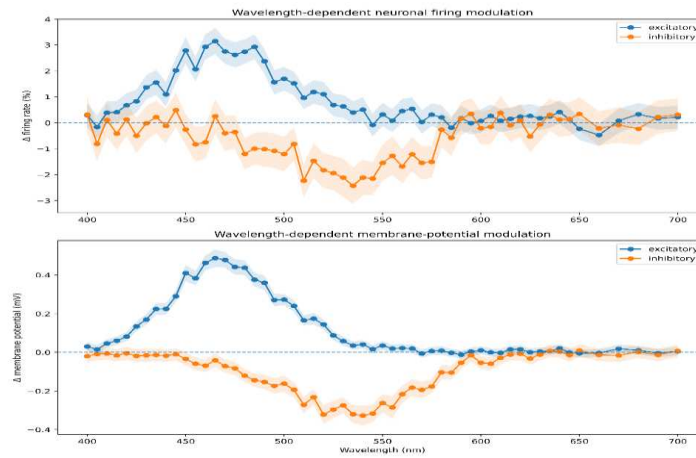


Figure 2. Wavelength-dependent neuronal optical modulation.

Figure 2 demonstrates a pronounced wavelength dependence of the predicted neuronal response. The response was not monotonic across the 400-700 nm spectral range, but instead exhibited a distinct spectral optimum in the blue-green region. The strongest positive modulation occurred approximately between 460 and 495 nm, with the observed response reaching approximately 2.28% at 465 nm. In contrast, responses around 530-650 nm were substantially smaller and approached zero. This behavior is consistent with the incorporation of wavelength-dependent optical and biological sensitivity into the computational framework. The corresponding biological optical dose was also substantially higher in the shorter-wavelength region. For example, the calculated biological optical dose was approximately 1.44×10^{-2} at 465 nm, compared with approximately 6.77×10^{-12} at 650 nm. Thus, the spectral response cannot be interpreted solely as a direct effect of wavelength; rather, wavelength modifies the interaction between photon energy, tissue optics, transmission, wavelength sensitivity, and biological modulation. Interestingly, the excitatory and inhibitory response profiles exhibited different magnitudes and spectral patterns, indicating that the same optical stimulus can produce neuron-type-dependent responses. This finding supports the inclusion of neuron identity and biological sensitivity variables in the predictive framework. The observed non-monotonic spectral structure is particularly important because the primary validation strategy deliberately used wavelength-grouped splitting, meaning that the model was required to generalize to wavelengths not represented in the training data. Therefore, the spectral landscape provides a more stringent test of model extrapolation than a conventional random split.

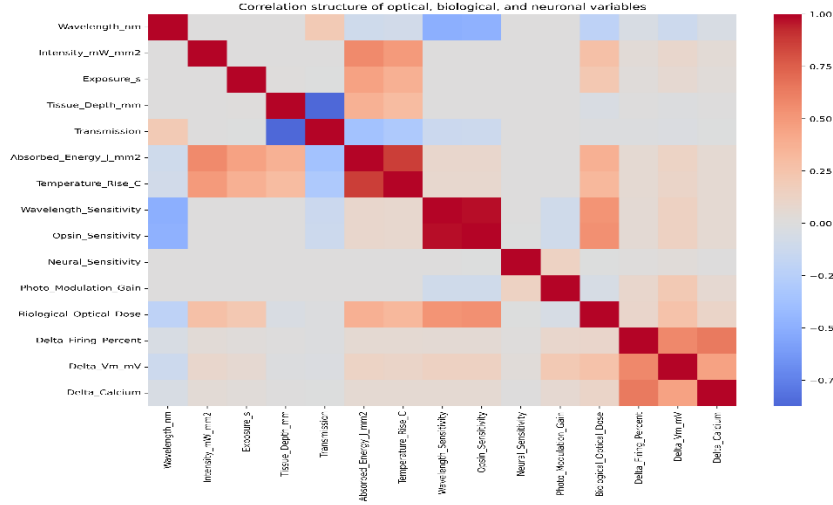


Figure 3. Correlation structure among optical, physical, and neuronal response variables.

The correlation heatmap in Figure 3 reveals a strongly structured relationship among wavelength, optical transport variables, biological sensitivity, and neuronal endpoints. Several optical variables exhibit expected interdependencies because they are mathematically or physically related. For example, wavelength is intrinsically connected to photon energy, whereas transmission is associated with attenuation and tissue depth. The primary neuronal response does not exhibit a single dominant linear association with wavelength alone. Instead, its relationships are distributed across several optical and biological variables. This observation is important because it suggests that neuronal photomodulation is governed by a combination of stimulus intensity, optical propagation, biological sensitivity, and tissue-dependent attenuation rather than by wavelength as an isolated predictor. The correlation structure also highlights the potential for multicollinearity among engineered mechanistic features. Variables such as optical dose, photon flux, absorbed energy, transmission, attenuation, and wavelength sensitivity are not statistically independent. Consequently, interpretation based exclusively on marginal correlation could be misleading. This motivated the use of complementary explainability approaches, including SHAP and permutation importance, to identify variables that retain predictive relevance within the multivariate model.

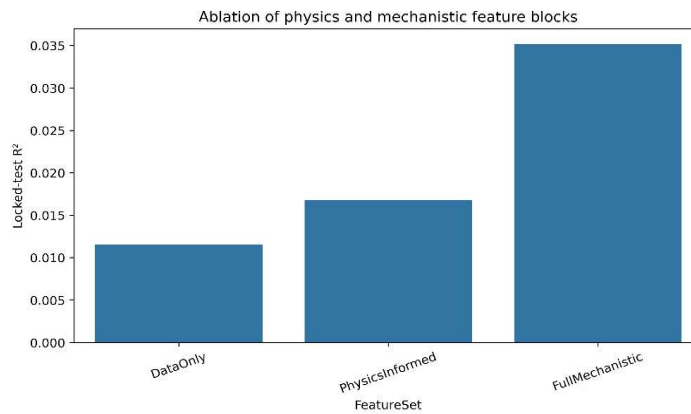


Figure 4. Ablation analysis demonstrating the incremental contribution of mechanistic and physics-informed feature representations.

The ablation analysis provides evidence that incorporating mechanistic information improves predictive performance under the locked wavelength-based evaluation. The data-only representation achieved an R^2 of 0.0116, whereas addition of physics-informed variables increased R^2 to 0.0168. The full mechanistic

representation produced the strongest performance, reaching an R^2 of 0.0352, with RMSE decreasing from 8.100 for the data-only representation to 8.002. Thus, the transition from data-only to physics-informed and finally to full mechanistic representation produced a progressive improvement in performance. The approximately 0.10 percentage-point reduction in RMSE may appear modest in absolute terms, but the consistent ordering across representations supports the hypothesis that physically meaningful variables provide additional information beyond generic predictors. The results also demonstrate an important limitation: physics-informed features improve the model but do not transform the task into a highly accurate deterministic prediction problem. This distinction strengthens the scientific interpretation of the framework because the objective is better characterized as extracting biologically and physically meaningful structure from a heterogeneous response landscape rather than achieving artificially high predictive accuracy.

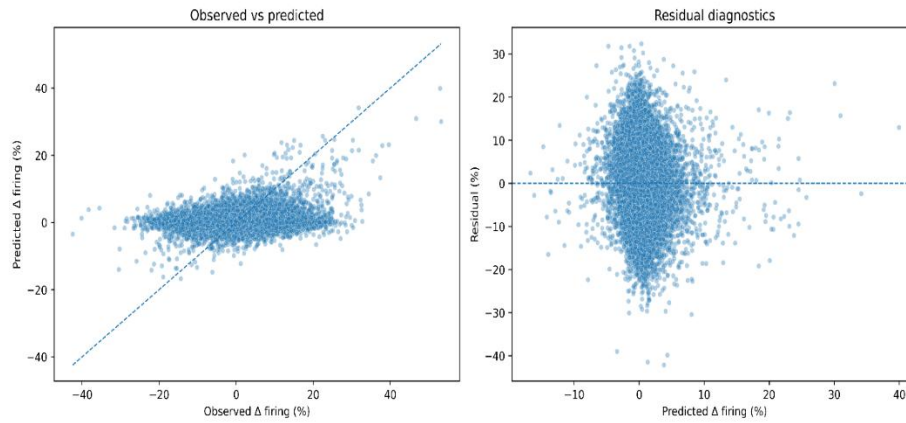


Figure 5. Observed-versus-predicted responses and residual diagnostics for the primary model.

Figure 5 illustrates the agreement between observed and predicted neuronal modulation together with the corresponding residual structure. The predicted values are concentrated around the central response range, whereas observations extend considerably farther from zero. Consequently, the model captures the dominant central tendency but has limited ability to reproduce the largest positive and negative responses. The residual plot shows a relatively dense central residual distribution accompanied by increased dispersion at extreme response values. This pattern indicates that prediction error is not completely homogeneous across the response range. In particular, the model tends to regress extreme responses toward the population mean, a common behavior in heterogeneous regression problems where extreme outcomes are relatively sparse. The relatively low global R^2 therefore should not be interpreted as evidence that the model contains no useful information. Instead, the modest positive Pearson correlation and the structured residual distribution indicate that the model identifies part of the underlying response variation while failing to capture a substantial fraction of individual-level variability. This result is particularly relevant for computational photomodulation modeling because neuronal responses can be affected by nonlinear interactions among optical dose, wavelength sensitivity, neuron phenotype, tissue attenuation, and baseline electrophysiological state.

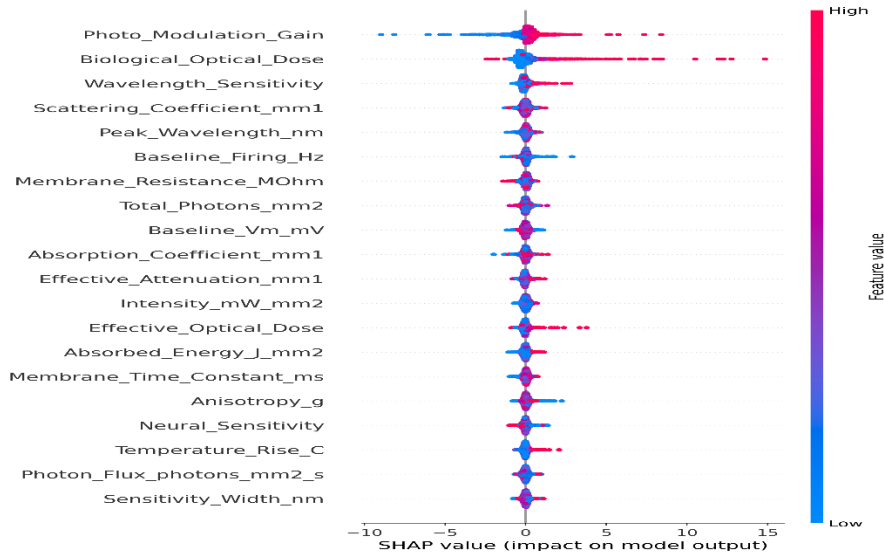


Figure 6. Global SHAP summary of the mechanistic predictive model.

The SHAP analysis identifies a clear hierarchy of influential predictors. Photo_Modulation_Gain was the most influential feature, with a mean absolute SHAP value of 0.484, followed closely by Biological_Optical_Dose with a value of 0.473. These were substantially more influential than the third-ranked Wavelength_Sensitivity feature 0.193. This ranking is biologically meaningful because photomodulation gain represents the model's integrated response sensitivity to optical stimulation, whereas biological optical dose captures the effective optical exposure after accounting for the relevant biological and propagation processes. Their high SHAP contributions therefore indicate that the model primarily relies on biologically interpretable stimulus-response variables rather than wavelength alone. Other important predictors included scattering coefficient, peak wavelength, baseline firing rate, membrane resistance, total photon number, baseline membrane potential, and absorption coefficient. The simultaneous importance of optical transport and neuronal state variables indicates that neuronal photomodulation is represented as a coupled optical-biophysical process. Notably, raw wavelength ranked only around the middle of the feature hierarchy, despite wavelength producing the most visually apparent spectral structure. This finding suggests that wavelength acts partly through derived mechanistic variables such as wavelength sensitivity, photon energy, absorption, and optical dose rather than functioning as an isolated determinant.

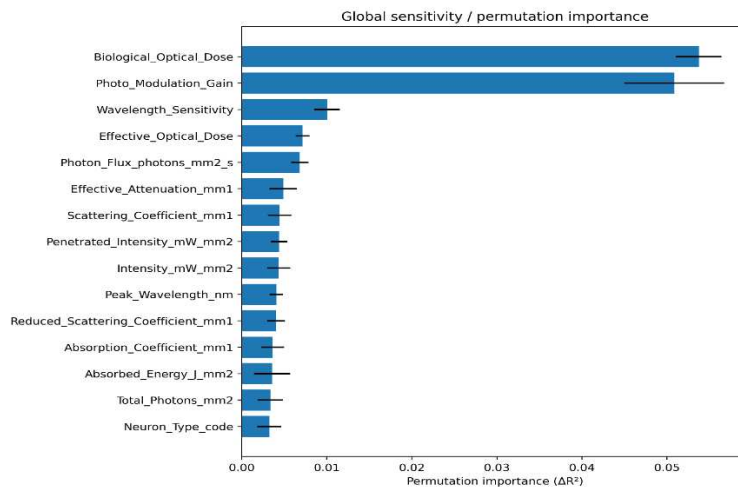


Figure 7. Global permutation importance of model predictors.

Permutation importance independently confirmed the dominant role of biologically and physically meaningful predictors. Biological_Optical_Dose ranked first with an importance of 0.0537, followed by Photo_Modulation_Gain 0.0508 and Wavelength_Sensitivity 0.0100. The agreement between permutation importance and SHAP analysis is particularly important because the two methods quantify feature contribution differently. The convergence of both approaches on biological optical dose and photo-modulation gain therefore provides stronger evidence that these variables represent central information channels within the predictive framework. Secondary contributions were observed for effective optical dose, photon flux, attenuation, scattering, penetrated intensity, and incident intensity. In contrast, photon energy, frequency, membrane resistance, baseline calcium, and baseline membrane potential had limited or slightly negative permutation importance under this evaluation. Taken together, these results suggest that effective optical delivery and biological responsiveness are more informative than simple spectral descriptors alone. This supports the mechanistic premise of the framework while also demonstrating that explainability should be interpreted at the level of interacting feature groups rather than isolated variables.

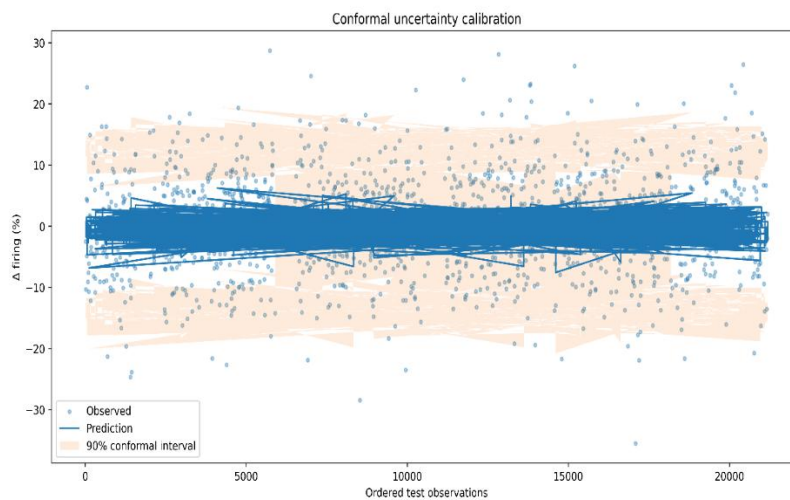


Figure 8. Conformal prediction intervals for uncertainty-aware neuronal response prediction.

Figure 8 demonstrates the uncertainty behavior of the predictive framework using conformal prediction intervals. The nominal 90% prediction interval achieved an empirical coverage of 90.24%, closely matching the target coverage. The corresponding mean interval width was approximately 26.43 percentage points. The close agreement between nominal and empirical coverage indicates that the uncertainty quantification procedure provides reasonably calibrated predictive intervals under the evaluated test distribution. This is an important complement to the relatively modest point-prediction performance because uncertainty estimates allow the model to distinguish between prediction and confidence in prediction. However, the relatively broad interval width indicates substantial irreducible or unmodeled variability in the primary endpoint. Therefore, the uncertainty analysis reinforces rather than contradicts the moderate predictive performance observed elsewhere. From a translational perspective, such calibrated uncertainty is preferable to overconfident predictions, particularly when the model is used to explore optical stimulation conditions outside the dominant training distribution.

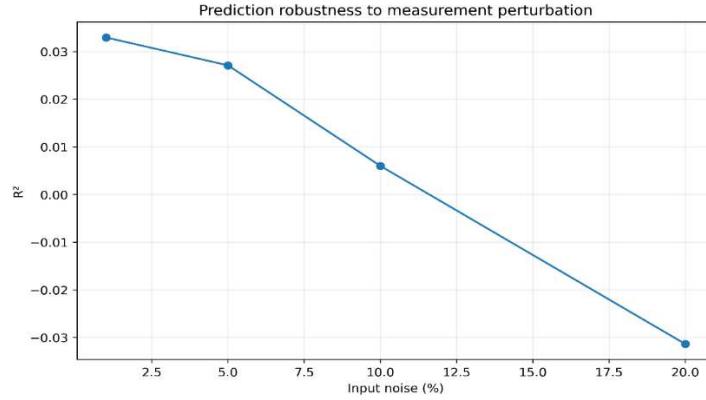


Figure 9. Prediction robustness under increasing input measurement noise.

The noise robustness experiment demonstrates a progressive deterioration in predictive performance as input perturbation increases. At a noise level of 1%, the model retained an R^2 of 0.0330 and RMSE of 8.012. At 5% noise, R^2 decreased slightly to 0.0271, whereas at 10% noise it declined to 0.0060. At the highest evaluated perturbation level of 20%, the model exhibited a negative R^2 of -0.0314, with RMSE increasing to 8.274. Pearson correlation similarly declined from 0.202 at 1% noise to 0.134 at 20% noise. The monotonic decline demonstrates that the predictive framework is sensitive to substantial measurement perturbation, particularly beyond approximately 10% noise. Nevertheless, the gradual rather than abrupt degradation at lower noise levels suggests that the model retains limited structural information under moderate perturbations. These findings emphasize that accurate acquisition of optical and biological input variables remains important for reliable neuronal response prediction.

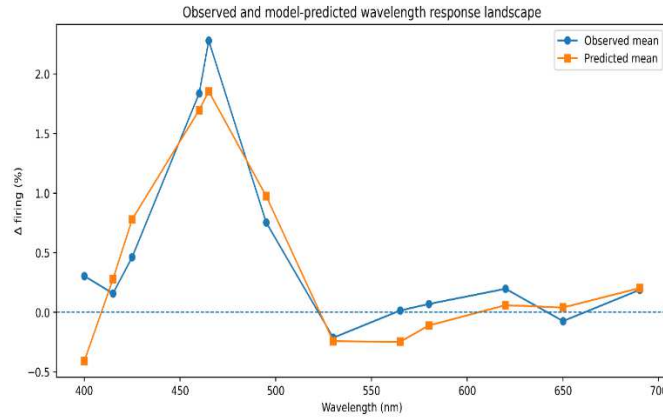


Figure 10. Observed and predicted wavelength-response landscape across the evaluated spectral range.

Figure 10 provides a direct comparison between the observed and predicted wavelength-dependent response landscape. Both profiles reproduce the major spectral pattern, including the prominent response maximum around 460-495 nm and the subsequent reduction toward longer wavelengths. The maximum observed response in the wavelength landscape was approximately 2.28% at 465 nm, while the corresponding model prediction was approximately 1.85%. At 530 nm, the observed response approached slightly negative modulation -0.21%, whereas the predicted value was approximately -0.24%. At longer wavelengths, both observed and predicted responses remained relatively close to zero. The agreement in the broad spectral topology is notable because the model was evaluated using wavelength-grouped generalization. Thus, the model was not simply memorizing individual wavelength labels; rather, it learned part of the underlying relationship between wavelength-associated optical and biological variables.

Nevertheless, discrepancies at individual wavelengths demonstrate that the model does not fully reconstruct the fine-scale spectral response. Therefore, Figure 10 should be interpreted as evidence for recovery of the dominant spectral structure rather than precise wavelength-specific quantitative prediction.

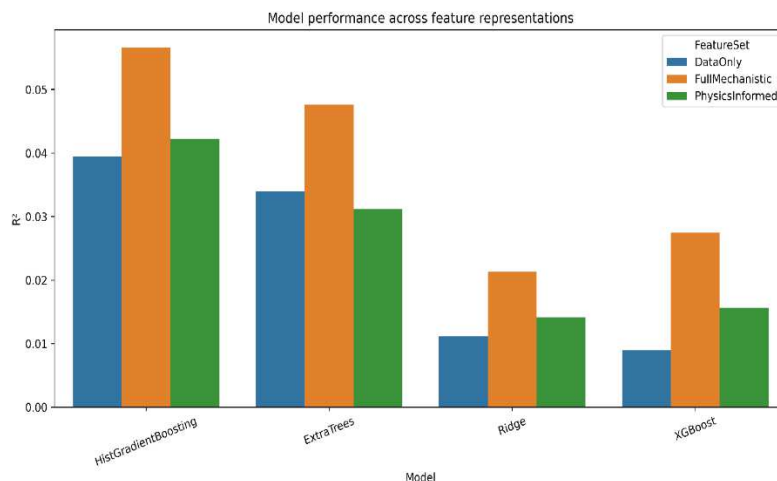


Figure 11. Comparison of predictive performance across data-only, full-mechanistic, and physics-informed representations.

Figure 11 compares multiple machine-learning algorithms across three feature representations. Among the evaluated models, HistGradientBoosting with the full mechanistic representation produced the strongest performance, achieving $R^2=0.0566$, RMSE=7.913, and MAE=6.281. The corresponding data-only HistGradientBoosting model achieved $R^2=0.0395$, while the physics-informed version achieved $R^2=0.0422$. A similar pattern was observed for ExtraTrees, where the full mechanistic representation outperformed both data-only and physics-informed representations. The consistency of the improvement across several algorithms indicates that the gain is not solely attributable to the choice of learning algorithm. Instead, it suggests that richer mechanistic representations provide additional predictive information. At the same time, the relatively low absolute R^2 values demonstrate that feature engineering alone cannot resolve the full complexity of neuronal response variability. The result therefore supports a mechanism-aware but uncertainty-conscious modeling strategy, rather than a claim of deterministic prediction. The key statistical and methodological characteristics of the analyzed dataset are summarized in Table 4, providing an overview of the variables and their corresponding descriptive statistics.

Table 4. Comparative Performance of Machine-Learning Models Under Locked Wavelength-Domain Testing

Feature representation	Model	R ²	RMSE (%)	MAE (%)	Bias (%)	Pearson r
Data-only	Ridge	0.0111	8.1016	6.3996	−0.0472	0.1057
Data-only	ExtraTrees	0.0340	8.0076	6.3482	−0.0673	0.1867
Data-only	HistGradientBoosting	0.0395	7.9847	6.3294	−0.0769	0.1999
Data-only	XGBoost	0.0090	8.1105	6.4245	−0.0684	0.1506
Physics-informed	Ridge	0.0141	8.0893	6.3939	0.0014	0.1189
Physics-informed	ExtraTrees	0.0312	8.0191	6.3573	−0.0799	0.1858
Physics-informed	HistGradientBoosting	0.0422	7.9734	6.3243	−0.0795	0.2075
Physics-informed	XGBoost	0.0156	8.0832	6.4052	−0.0597	0.1646
Full mechanistic	Ridge	0.0213	8.0597	6.3770	−0.0007	0.1462
Full mechanistic	ExtraTrees	0.0476	7.9508	6.3111	−0.0558	0.2215
Full mechanistic	HistGradientBoosting	0.0566	7.9133	6.2809	−0.0483	0.2389
Full mechanistic	XGBoost	0.0274	8.0345	6.3784	−0.0544	0.1936

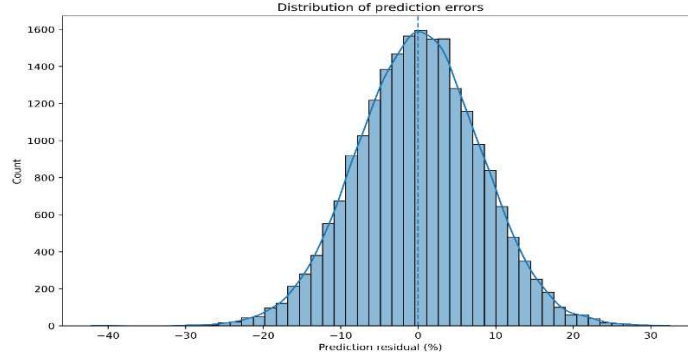


Figure 12. Distribution of prediction errors for the primary endpoint.

Figure 12 shows a concentrated, approximately unimodal distribution of prediction residuals centered near zero. The concentration around the central region indicates that large systematic bias is not dominant across the majority of observations. The relatively symmetric residual profile is favorable from a calibration perspective, although the tails indicate that substantial errors remain for a subset of observations. These larger residuals correspond to the extreme response values that were also apparent in the observed-versus-predicted analysis. The residual distribution therefore supports two complementary conclusions. First, the model does not exhibit a strong global directional bias. Second, prediction uncertainty increases for atypical response conditions, indicating that the remaining error is likely associated with nonlinear interactions or response heterogeneity rather than a simple systematic offset.

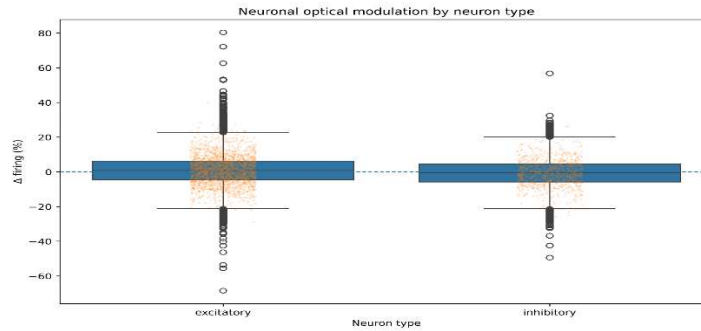


Figure 13. Neuron-type-specific distribution of optical modulation responses.

Figure 13 demonstrates clear heterogeneity between excitatory and inhibitory neuronal populations. Both groups exhibited broad distributions centered near relatively small changes in firing, but the distributions contained substantial outliers extending toward both positive and negative modulation. The existence of neuron-type-specific response distributions supports the biological relevance of including *Neuron_Type* as an explicit model feature. Nevertheless, the leave-one-neuron-type-out experiment revealed poor cross-type generalization: when excitatory neurons were held out, R^2 was -0.201, whereas holding out inhibitory neurons resulted in R^2 = -0.167. This finding is particularly informative. It indicates that neuron type is not merely a minor categorical variable but is associated with distributional differences that cannot be reliably transferred from one neuronal population to another using the available feature representation. Consequently, the model should not be interpreted as demonstrating universal neuron-independent optical response prediction. Instead, the results suggest that cell-type-specific calibration or training data may be necessary for reliable cross-population deployment.

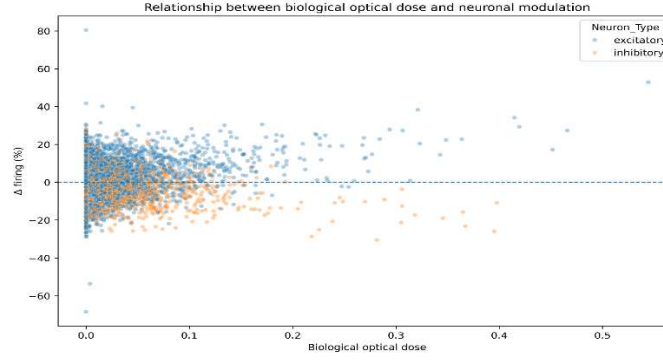


Figure 14. Relationship between biological optical dose and neuronal modulation.

Figure 14 illustrates the relationship between biological optical dose and neuronal response. The response magnitude generally increased in the region of higher biological optical dose, but substantial dispersion remained throughout the dose range. This pattern is consistent with the SHAP and permutation analyses, where biological optical dose emerged as the most influential predictor. However, the broad vertical dispersion at comparable dose levels demonstrates that optical dose alone is insufficient to uniquely determine neuronal modulation. The observed heterogeneity likely reflects the contribution of wavelength sensitivity, photo-modulation gain, neuron type, baseline electrophysiological state, tissue optical properties, and other biological variables. Therefore, biological optical dose should be regarded as a major exposure descriptor rather than a complete dose-response determinant. The absence of a simple one-to-one dose-response relationship is also relevant for experimental design. Equivalent optical doses may not necessarily produce equivalent neuronal responses if delivered at different wavelengths or under different biological conditions.

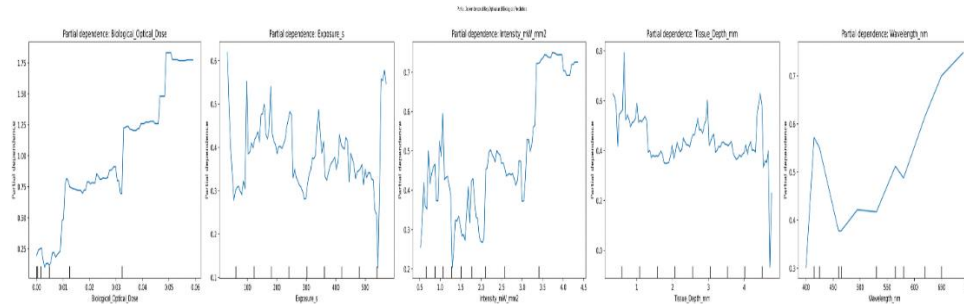


Figure 15. Partial Dependence of Key Optical and Exposure Features on Predicted Neuronal Photomodulation

The partial dependence profiles for the five important predictors Biological Optical Dose, Exposure, Optical Intensity, Tissue Depth, and Wavelength of the model-predicted neuronal photomodulation response are shown in figure 15. The profiles show that the model learnt mostly nonlinear, feature specific relations instead of simple monotonic relations. Biological Optical Dose has the most evident positive dependence with the predicted response when presented in the form of partial dependence. A progression of response is predicted from low optical-dose values, with higher predicted responses at the 0.01 and 0.03-0.05 dose levels, before settling at a relatively high maximum level at higher doses. This pattern suggests that higher BEE exposure tends to lead to better predictions of photomodulatory responses, but not in a linear fashion between the two: there are clear nonlinear transitions, not a linear dose-response relationship. The partial dependence profile is much more varied for Exposure. The response predicted varies throughout the range of exposure, having a number of local peaks and valleys, and no clear monotonic response. This indicates that the predicted neuronal response is not a function of exposure time, but is probably modulated by interactions with optical intensity, wavelength-dependent sensitivity and tissue transmittance. An improvement of the positive trend is seen for Optical Intensity. The response is relatively low for intensities less than about 3.2 mW/mm², but rises rapidly beyond this point and enters a relatively high-response regime at higher intensities. This is a non-linear intensity response relationship, as modelled, because at high enough optical intensities, a significantly larger

modulation is predicted, indicating this non-linear relationship. The Tissue Depth profile shows a fairly consistent response throughout most of the depth range studied with some moderate variations and a number of local deviations. A lack of a monotonic decline suggests that the depth of the tissue itself is not the sole factor controlling the predicted response in the modeled range. Rather, it is likely to impact through optical attenuation and interaction with intensity, absorption, scattering, and biological optical dose. These abrupt transitions at the extreme depth ranges should be treated with care as they may represent limited coverage in the tails of the feature distribution. Finally, Wavelength is definitely nonlinear. The predicted response shows a sudden rise from around 400 nm towards the blue region, a local maximum at 420-430 nm, a rapid drop through 450-500 nm, and a gradual increase towards the red region, with the greatest partial dependence around 700 nm. This is a non-monotonic spectral pattern, indicating that wavelength plays an indirect role in the model prediction, in a non-linear fashion. Figure 15 overall gives important evidence on the model level that the modulation of neurons is a nonlinear dependence of optical dose, intensity, exposure, tissue depth, and wavelength. The strong wavelength dependence of the features is in line with the particularly strong dependence that was observed on the biological dose and intensity, and the irregular wavelength profile indicates the presence of wavelength-dependent biological and optical interactions. More importantly, the partial dependence is the average marginal model response and it is not a causal dose-response relationship. The nonlinearities observed are mostly a reflection of how the trained model is using the variables to make predictions in the computational feature space.

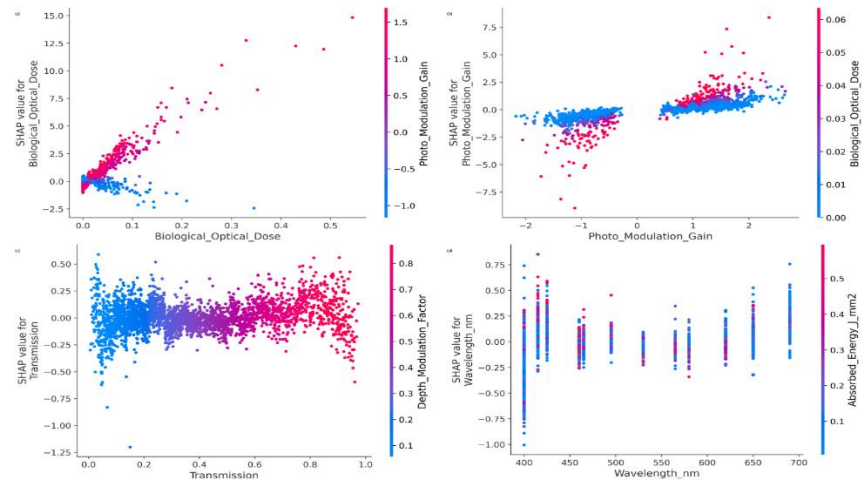


Figure 16. SHAP Dependence Analysis Revealing Nonlinear and Interaction Effects of Key Optical Features on Neuronal Photomodulation

Figure 16 shows four of the most influential predictor Biological Optical Dose, Photo-Modulation Gain, Transmission, and Wavelength with the all panels colored by a biologically or physically related feature to see how feature interactions might be present. The results show that there is no independent additive effect required to use the predictive model rather, the contributions of several predictors to the predicted neuronal photomodulation are highly nonlinear and dependent on context. Biological Optical Dose and Photo-Modulation Gain shown in Figure 16(a). Biological Optical Dose has a strong positive and non-linear SHAP dependence profile with the model output. Note that the contributions for SHAP are small for the low optical-dose values while they increase with increasing biological optical dose. When the dose is raised (above c. 0.1-0.2) there is a significant spread in the SHAP values, with the highest positive values at the highest dose. Importantly, the colour scale shows that there is strong interaction with Photo-Modulation Gain. Observations with higher modulation gain are more often found to have positive SHAP values, while observations with lower modulation gain are more often found to have negative or weak values of the SHAP, especially at intermediate to high optical dose. This pattern implies that the model prediction is not completely determined by biological optical dose, the effects of which are strongly modulated by the intrinsic modulation gain of the modeled system. The growing dispersion of SHAP values at higher doses also indicate that the effect of optical dose is becoming more variable, depending on the biological response. In this way, the model seems to be an important driver that is modulated by biological sensitivity. The data is shown as the photo-modulation gain and biological optical dose. As shown in the reciprocal analysis of Figure 16(b), Photo-Modulation Gain is important. The SHAP values tend to be more positive with higher modulation gain, and more negative with lower modulation gain. Additional evidence of interaction with Biological Optical Dose is provided by the color encoding. Warm colors indicate high dose observations and cover a greater range of SHAP values; these SHAP values are also the biggest

positive contribution. By contrast, these low dose observations are focused around the main SHAP area. The reciprocal pattern is relevant because it shows that it is not just the fact that one feature is important in isolation that results in the relationship between biological optical dose and photo-modulation gain. Rather, the two variables seem to be involved in a coupled response structure. This finding is supported by the convergence of this result with the global SHAP and permutation importance analyses, which led to the conclusion that the main predictive dimensions for this model are effective optical exposure and biological modulation sensitivity. Figure 16(c) shows that there is a much more complicated relationship between Transmission and the neural response predicted. In the region of low transmissions, the contributions to SHAP can be widely spread ranging from negative to positive values. The distribution shifts to be more peaked around zero as the transmission value increases, and hence has a relatively small and more diverse contribution to the distribution. For higher values of the transmission (higher than around 0.6–0.9), positive SHAP contributions are more pronounced, indicating that at higher levels of optical transmission, the predicted response may be greater under a favorable biological scenario. However, the effect is not strictly monotonic; there are significant negative effects at the highest values of the transmission. The Depth-Modulation Factor is shown as a color gradient, which also shows interaction between transmission and tissue-dependent modulation. A higher depth modulation has a wider spectrum of SHAP effects, especially in the high transmission region. This implies that the modulation due to the effect of optical transmission is not an independent variable but is dependent on the modulation due to tissue-depth. Together, this pattern allows for a mechanistic interpretation where the amount of optical energy that reaches the biological compartment interacts with modulation processes that are specific to the tissue. The large spread also suggests that variability in the response of neurons cannot be entirely accounted for by differences in the transmission. The wavelength dependence as depicted in Figure 16(d) is definitely nonlinear. The contribution by SHAP is mostly negative at 400nm while the proportion of positive contributions is higher at 420–430nm. The contribution then moves toward zero and shows about 460–650 nm with some intra-wavelength variations. Positive SHAP values are more prominent at 650–700 nm. The color distribution showing the amount of absorbed optical energy suggests that the wavelength effects are also influenced by the amount of absorbed energy. However, there is no straightforward relationship between SHAP magnitude and absorbed energy of sunlight across all wavelengths. This is significant as it implies that the model is not only describing the effects of the wavelength as energy absorption. Rather, wavelength probably affects a variety of optical and biological parameters such as tissue transmission, absorption, biological sensitivity and photo-modulation parameters. The separate vertical clusters of observations at one wavelength also illustrate that SHAP contributions from observations at the same nominal wavelength can vary significantly. This intra-wavelength variation is another indication that the predicted neuronal response cannot be determined based on wavelength alone. Overall, Figure 16 shows that the effective optical dose, the gain of photomodulation, the photo-modulation gain, the tissue transmission, the depth-dependent modulation, and the wavelength exhibit nonlinear, context-dependent interactions to determine the predicted neuronal photomodulation, which can be explained mechanistically beyond the traditional global feature-importance ranking.

4. Limitations

The results of this study are subject to a few limitations. Firstly, the analysis was carried out on a generated database of data from computational simulations of neurons, instead of data obtained from experimental recordings of neurons. The data set includes both physically and biologically interpretable optical variables, but the synthetic data has a limited ability to mimic the biological heterogeneity, measurement noise, non-linear cell responses, and experimental confounds found in real neuronal systems. Thus, the claimed predictive performance should be considered as a computational test of the proposed framework instead of the direct experimental test of neuronal modulation by light. Second, the primary endpoint, percentage change in neuronal firing rate was relatively poorly predictable in the strict unseen-wavelength scheme. That suggests that the modulation of neuronal firing might be influenced by other biological factors that are not captured in the current feature space. More specifically, the intracellular signaling pathways, the dynamics of the ion channels, the morphology of the cells, the variability of the cell membranes, and the temporal adaptation of the cells were not explicitly modeled. Third, a number of optical and biological variables in the data set are derived from other variables mechanistically. While we excluded response related variables from the predictor sets to prevent target leakage, there is the possibility of the physics derived features being correlated, which can result in redundancy and apparent stability in the model. Thus, the importance of features is not to be understood as independent causal effects, but rather as predictive association. Last, the wavelength-grouped validation strategy is a rigorous measure of the

interpolation/extrapolation of different wavelengths, but does not represent external validation across independent laboratories, experimental protocols, species, brain region, or neuronal population. The leave-neuron-type-out analysis is also a restricted form of domain-shift evaluation, and not a full biological generalization. Fifth, simple optical and biological representations are used in the physics-informed part. The present formulation does not include the wavelength dependent scattering, heterogeneous absorption, anisotropic propagation, complex geometry, and potentially depth dependent photobiological responses of real tissue. Physical fidelity could be increased by more detailed light-transport models, such as Monte Carlo or finite-element models, or models calibrated from experiments. Lastly, computational evaluation of uncertainty calibration was performed based on the split-conformal prediction and measurement-noise perturbation. These analyses offer good estimates of predictive reliability, but cannot quantify all biological uncertainties, such as inter-animal variability, batch effects, differences in experimental instrument, and unobserved physiological states. The framework should be further tested using independent experimental datasets and extended to include multimodal neuronal measurements, longitudinal responses and mechanistic cellular models in future studies. In sum, these restrictions are the features of the present study that primarily identify it as a physics-informed computational framework and hypothesis generation tool. The next step to validate its usefulness for quantitative prediction and optimization of wavelength dependent neuronal optical modulation is independent experimental validation.

5. Conclusions

In this study, an explainable machine learning model that incorporates optical physics, tissue interaction parameters, biological sensitivity descriptors and neuronal baseline parameters was developed for prediction of the wavelength-dependent optical modulation of neuronal activity. The proposed framework systematically assessed the predictive capability, physics-informed contributions of features, model interpretability, uncertainty, and robustness, using a computational dataset of 100,000 samples, with both types of neurons and across the range of wavelengths between 400-700 nm. Results showed that these physics-derived and mechanistic features did offer measurable advantages over data-only representations in predicting the main endpoint of percent change in firing rate, but with a relatively low performance under the stringent unseen-wavelength validation, it suggests that firing modulation is still difficult to predict from the currently available variables alone. The learnability of neuronal optical responses was, in contrast, highly reproducible for the modulation of the membrane potential, indicating important endpoint-specific differences in the learnability of neuronal optical responses. Optical dose, wavelength sensitivity, photo-modulation gain, tissue optical properties and wavelength-related descriptors were identified as the main predictive factors across all the explainability analyses. Ablation analysis also showed the benefit of utilizing physics-based information, and the physics-informed residual approach gave a clear pathway for combining mechanical knowledge with nonlinear machine learning. Importantly different was the evidence on generalization and predictive uncertainty, which was provided by wavelength-grouped validation, leave-neuron-type-out testing, noise-robustness analysis, and split-conformal prediction, which complimented each other and did not depend exclusively on the traditional random train–test performance. These findings provide a clear computational model for probing and understanding the wavelength-specific modulation of neuronal response and for the development of hypotheses of optical and biological control of neuronal response that can be tested. The framework should be viewed, however, as a methodological and hypothesis generating tool, rather than an experimental validation, due to the fact that the data used in the present analysis is computer-generated. Further future studies should focus on independent experimental confirmation, electrophysiology and calcium imaging data in the dynamic component, more realistic models of light transport in a heterogeneous array of cells, and data sets across species and laboratories. The extensions would allow for experimentally validated and uncertainty-aware predictive models for the rational optimization of optical neuromodulation strategies.

Declarations

Ethics approval and consent to participate: This study did not involve human participants, human-derived biological samples, or animal experiments. Therefore, ethical approval and consent to participate were not required.

Consent for publication: Not Applicable

Data and materials availability: The datasets and computational resources used in this study are described in the manuscript. Where permitted by the relevant data sources, the data and/or processed datasets can be made available from the corresponding author upon reasonable request.

Competing interests: The authors declare that they have no competing interests.

Funding: This work was not supported by any specific grant.

Authors' contributions:

Md. Rasel Islam: Conceptualization, methodology, software, data curation, formal analysis, investigation, visualization, validation, and writing-original draft.

Shazid Mahmud Labim: Investigation, data curation, validation, formal analysis, and writing-review and editing.

Nazmul Hasan: Investigation, data curation, validation

Md. Sadiqur Rahman Zihad: Formal analysis, and writing-review and editing

Md. Rabiul Islam: Investigation, formal analysis, and validation

Md. Shariful Islam: Conceptualization, methodology, supervision, project administration, validation, and writing-review and editing.

Acknowledgements

The authors gratefully acknowledge the Emerging Research Team for their continuous assistance, constructive cooperation, and support during the completion of this study. The authors also extend their appreciation to the Department of Information and Communication Technology and the Department of Biomedical Engineering, Islamic University, Kushtia-7003, Bangladesh, for providing institutional support and encouragement.

Conflict of Interest

The authors affirm that they have no relevant financial or non-financial interests that influenced the research, analysis, or preparation of this manuscript.

Funding Declaration

The authors received no financial support or funding for the research, authorship, or publication of this article.

References

1. Dole, M., Auboiroux, V., Langar, L., & Mitrofanis, J. (2023). A systematic review of the effects of transcranial photobiomodulation on brain activity in humans. *Reviews in the Neurosciences*, 34, 671–693. DOI: 10.1515/revneuro-2023-0003.
2. Hamblin, M. R. (2023). Transcranial photobiomodulation for the brain: a wide range of clinical applications. *Neural Regeneration Research*, 19, 483–484. DOI: 10.4103/1673-5374.380891.
3. Bowen, R. L., et al. (2023). Use of either transcranial or whole-body photobiomodulation treatments improves COVID-19 brain fog. *Journal of Biophotonics*, 16, e202200391. DOI: 10.1002/jbio.202200391.
4. Lin, YP., Ku, CH., Chang, CC. et al. Effects of intravascular photobiomodulation on cognitive impairment and crossed cerebellar diaschisis in patients with traumatic brain injury: a longitudinal study. *Lasers Med Sci* 38, 108 (2023). <https://doi.org/10.1007/s10103-023-03764-8>
5. Rodríguez-Fernández, L., Zorzo, C. & Arias, J.L. Photobiomodulation in the aging brain: a systematic review from animal models to humans. *GeroScience* 46, 6583–6623 (2024). <https://doi.org/10.1007/s11357-024-01231-y>
6. Hu, Z., Qu, X., Li, L., et al. (2023). Repeated photobiomodulation induced reduction of bilateral cortical hemodynamic activation during a working memory task in healthy older adults. *IEEE Journal of Biomedical and Health Informatics*, 27, 2876–2885. DOI: 10.1109/JBHI.2023.3259069.
7. Fear, E. J., Torkelsen, F. H., Zamboni, E., et al. (2023). Use of ³¹P magnetization transfer magnetic resonance spectroscopy to measure ATP changes after 670 nm transcranial photo biomodulation in older adults. *Aging Cell*, 22. DOI: 10.1111/accel.14005.
8. Li, S., Li, D., Yu, T., et al. (2023). Transcranial photo biomodulation improves insulin therapy in diabetic microglial reactivity and the brain drainage system. *Communications Biology*, 6, 1239. DOI: 10.1038/s42003-023-05630-3.
9. Li, D., Lin, H., Sun, S., et al. (2023). Photo stimulation of lymphatic clearance of β -amyloid from mouse brain: a new strategy for the therapy of Alzheimer's disease. *Frontiers of Optoelectronics*, 16, 45. DOI: 10.1007/s12200-023-00099-8.
10. Mohammed, H. S., Hosny, E. N., Sawie, H. G., & Khadrawy, Y. A. (2023). Transcranial photo biomodulation ameliorates midbrain and striatum neurochemical impairments and behavioral deficits in reserpine-induced parkinsonism in rats. *Photochemical & Photobiological Sciences*, 22, 2891–2904. DOI: 10.1007/s43630-023-00497-z.
11. Sipion, M., Ferreira, F. M., Scholler, J., et al. (2023). A randomized, blinded study of photo biomodulation in a mouse model of Alzheimer's disease showed no preventive effect. *Scientific Reports*, 13. DOI: 10.1038/s41598-023-47039-2.
12. Zorzo, C., Rodríguez-Fernández, L., Martínez, J. A., & Arias, J. L. (2024). Photo biomodulation increases brain metabolic activity through a combination of 810 and 660 wavelengths: a comparative study in male and female rats. *Lasers in Medical Science*, 39, 26. DOI: 10.1007/s10103-023-03966-0.
13. Razzaghi, M., Sheibani, F., Kimia, N., et al. (2024). Photo biomodulation's potential as a non-invasive therapy for Alzheimer's disease and minimal cognitive impairment: A 12-week investigation. *Photo-diagnosis and Photodynamic Therapy*, 46, 103991. DOI: 10.1016/j.pdpdt.2024.103991.
14. Xu, H., Luo, Z., Zhang, R., et al. (2024). Exploring the effect of photo biomodulation and gamma visual stimulation induced by 808 nm and visible LED in Alzheimer's disease mouse model. *Journal of Photochemistry and Photobiology B: Biology*, 250, 112816. DOI: 10.1016/j.jphotobiol.2023.112816.

15. Chen, H., Shi, X., Liu, N., et al. (2024). Photo biomodulation therapy mitigates depressive-like behaviors by remodeling synaptic links and mitochondrial function. *Journal of Photochemistry and Photobiology B: Biology*, 258, 112998. DOI: 10.1016/j.jphotobiol.2024.112998.
16. Chen, Z., Li, M., Wu, C., et al. (2024). Photo biomodulation therapy alleviates repeated closed head injury-induced anxiety-like behaviors. *Journal of Bio-photonics*, 17, e202300343. DOI: 10.1002/jbio.202300343.
17. Shahdadian, S., Wang, X., & Liu, H. (2024). Directed physiological networks in the human prefrontal cortex at rest and post transcranial photo biomodulation. *Scientific Reports*, 14, 10242. DOI: 10.1038/s41598-024-59879-7.
18. Fernandes, F., Oliveira, S., Monteiro, F., et al. (2024). Devices used for photo biomodulation of the brain—a comprehensive and systematic review. *Journal of Neuro Engineering and Rehabilitation*, 21, 53. DOI: 10.1186/s12984-024-01351-8.
19. Zhang, Z., Zhang, Z., Liu, P., et al. (2024). The Role of Photobiomodulation to Modulate Ion Channels in the Nervous System: A Systematic Review. *Cellular and Molecular Neurobiology*, 44, 79. DOI: 10.1007/s10571-024-01513-1.
20. Lee, K., Chun, M., Jung, B., et al. (2024). Machine-Learning-Based Prediction of Photobiomodulation Effects on Older Adults With Cognitive Decline Using Functional Near-Infrared Spectroscopy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 32, 3710–3718. DOI: 10.1109/TNSRE.2024.3469284.
21. Gonzalez-Lima, M. D., Barrett, D. W., Mirji, A., et al. (2025). Machine Learning Prediction of the Cognitive Responses to Transcranial Prefrontal Photobiomodulation in Individuals With Bipolar Disorder Using Functional Near-Infrared Spectroscopy. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 33, 4349–4361. DOI: 10.1109/TNSRE.2025.3625855.
22. Hyeon, B., Shin, J., Lee, J.-H., et al. (2025). Integrating artificial intelligence and optogenetics for Parkinson's disease diagnosis and therapeutics in male mice. *Nature Communications*.
23. Van Lankveld, H., Mai, A. Q., Lim, L., et al. (2025). Simulation-based dosimetry of transcranial and intranasal photobiomodulation of the human brain: the roles of wavelength, power density, and skin tone. *Biomedical Optics Express*, 16, 3295–3314. DOI: 10.1364/BOE.567345.
24. Yang, Q., Qu, X., Sheng, C., et al. (2025). Transcranial photobiomodulation improves functional brain networks and working memory in healthy older adults: An fNIRS study. *NeuroImage*, 316, 121305. DOI: 10.1016/j.neuroimage.2025.121305.
25. Hong, N., Yoon, S.-R., & Ahn, J.-C. (2025). Photobiomodulation using an 830-nm laser alleviates hippocampal reactive gliosis and cognitive dysfunction in a mouse model of adolescent chronic alcohol exposure. *Pharmacology Biochemistry and Behavior*, 248, 173956. DOI: 10.1016/j.pbb.2025.173956.
26. Florian Armășescu, Bogdan Amuzescu, Roxana-Olimpia Gheorghe, Mihail Ghenghea, Violeta Ristoiu, Jean Ciurea, Ion Gruia, Fiber-optic-guided near-infrared laser exposure induces depolarization of cultured primary sensory neurons and modifies biophysical properties of human Nav1.5 channels, *Journal of Photochemistry and Photobiology B: Biology*, Volume 269, 2025, 113191, ISSN 1011-1344, <https://doi.org/10.1016/j.jphotobiol.2025.113191>.
27. Zhang, Z., Xue, X., He, D., et al. (2026). Reversal of auditory cortical hyperexcitability and restoration of synaptic plasticity balance by GluN1-mediated photobiomodulation in noise-induced tinnitus. *Brain Research Bulletin*, 234, 111685. DOI: 10.1016/j.brainresbull.2025.111685.
28. Ariel Motsenyat, Xiaole Z. Zhong, Hannah Van Lankveld, Joanna X. Chen, Alicia Mathew, J. Jean Chen, Modulating cerebrospinal fluid dynamics using pulsed photobiomodulation, *Brain Stimulation*, Volume 19, Issue 1, 2026, 102988, ISSN 1935-861X, <https://doi.org/10.1016/j.brs.2025.102988>.

29. Ariel Motsenyat, Xiaole Z. Zhong, Hannah Van Lankveld, Joanna X. Chen, Alicia Mathew, J. Jean Chen, Modulating cerebrospinal fluid dynamics using pulsed photobiomodulation, *Brain Stimulation*, Volume 19, Issue 1, 2026, 102988, ISSN 1935-861X, <https://doi.org/10.1016/j.brs.2025.102988>.
30. Godaert, L., Dramé, M. & Lemaire, A. Transcranial photobiomodulation for the treatment of chemobrain: new perspectives from a pilot study. *Support Care Cancer* 34, 62 (2026). <https://doi.org/10.1007/s00520-025-10280-8>
31. Dirk De Ridder, Michael R. Hamblin, Sven Vanneste, Transcranial Photobiomodulation for Neuromodulation of Brain Disorders: A Perspective, *Neuromodulation: Technology at the Neural Interface*, 2026, ISSN 1094-7159, <https://doi.org/10.1016/j.neurom.2025.12.006>.
32. Guiyun Shi, Kairui Wei, Ruocun Wang, Yayu Wang, Xianyu Wang, Luyang Gao, Shuyi Li, Bo Yang, Lina Tuerxunaili, Yi Jiao, Mechanisms of action and clinical research progress of photobiomodulation, *Journal of Photochemistry and Photobiology B: Biology*, Volume 282, 2026, 113529, ISSN 1011-1344, <https://doi.org/10.1016/j.jphotobiol.2026.113529>.
33. Verinaz-Jadan, H., Howe, C. L., Song, P., et al. (2023). Physics-Based Deep Learning for Imaging Neuronal Activity via Two-Photon and Light Field Microscopy. *IEEE Transactions on Computational Imaging*, 9, 565–580. DOI: 10.1109/TCI.2023.3282052.
34. Song, P., Verinaz-Jadan, H., Howe, C. L., Foust, A. J., & Dragotti, P. L. (2024). Model-Based Explainable Deep Learning for Light-Field Microscopy Imaging. *IEEE Transactions on Image Processing*. DOI: 10.1109/TIP.2024.3387297.
35. Long, X., Gao, Y., Yuan, Z., et al. (2024). In-Situ Wavefront Correction via Physics-Informed Neural Network. *Laser & Photonics Reviews*, 18, 2470045. DOI: 10.1002/lpor.202300833.
36. Zhan, Y., Zhang, H., et al. (2024). Physics-Aware Analytic-Gradient Training of Photonic Neural Networks. *Laser & Photonics Reviews*. DOI: 10.1002/lpor.202300445.
37. Park, H., Park, J. H., & Hwang, J. (2024). An inversion problem for optical spectrum data via physics-guided machine learning. *Scientific Reports*, 14, 9042. DOI: 10.1038/s41598-024-59594-3. *Nature*
38. Kazemzadeh, M., Collard, L., Piscopo, L., De Vittorio, M., & Pisanello, F. (2025). A Physics-Informed Neural Network as a Digital Twin of Optically Turbid Media. *Advanced Intelligent Systems*, 7, 2570024. DOI: 10.1002/aisy.202400574.
39. Omar A.M. Abdelraouf, Abdulrahman M.A. Ahmed, Emadeldeen Eldele, Ahmed A. Omar, From maxwell's equations to artificial intelligence: The evolution of physics-guided AI in nanophotonics and electromagnetics, *Optics & Laser Technology*, Volume 192, Part D, 2025, 113828, ISSN 0030-3992, <https://doi.org/10.1016/j.optlastec.2025.113828>.
40. Zicheng Huang, Luofei Tu, Zhishun Guo, Mengyang Shi, Yesheng Gao, Xingzhao Liu, Scalable imaging through scattering media via physics-informed sparse optronic convolutional neural networks with knowledge distillation, *Optics Communications*, Volume 585, 2025, 131809, ISSN 0030-4018, <https://doi.org/10.1016/j.optcom.2025.131809>.
41. Zicheng Huang, Luofei Tu, Zhishun Guo, Mengyang Shi, Yesheng Gao, Xingzhao Liu, Scalable imaging through scattering media via physics-informed sparse optronic convolutional neural networks with knowledge distillation, *Optics Communications*, Volume 585, 2025, 131809, ISSN 0030-4018, <https://doi.org/10.1016/j.optcom.2025.131809>.
42. W. Gao et al., "High-Precision Large-Scale Optical Phased Array Calibration Using Physics-Informed Neural Network With Transfer Learning," in *Journal of Lightwave Technology*, vol. 43, no. 22, pp. 10203-10209, 15 Nov.15, 2025, doi: 10.1109/JLT.2025.3608167.

43. Yu R, Wang R. Learning dynamical systems from data: An introduction to physics-guided deep learning. *Proc Natl Acad Sci U S A*. 2024 Jul 2;121(27):e2311808121. doi: 10.1073/pnas.2311808121. Epub 2024 Jun 24. PMID: 38913886; PMCID: PMC11228478.
44. Chuprov, I.A., Gao, J., Efremenko, D.S. et al. Solution of the Multimode Nonlinear Schrödinger Equation Using Physics-Informed Neural Networks. *Dokl. Math.* 110 (Suppl 1), S15–S24 (2024). <https://doi.org/10.1134/S1064562424602105>
45. Sylvester, S., Sagehorn, M., Gruber, T. et al. SHAP value-based ERP analysis (SHERPA): Increasing the sensitivity of EEG signals with explainable AI methods. *Behav Res* 56, 6067–6081 (2024). <https://doi.org/10.3758/s13428-023-02335-7>
46. Kyrimi, E., McLachlan, S., Wohlgemut, J.M. et al. Explainable AI: definition and attributes of a good explanation for health AI. *AI Ethics* 5, 3883–3896 (2025). <https://doi.org/10.1007/s43681-025-00668-x>