RESEARCH ARTICLE

Statistics in Medicine WILEY

Spatially informed Bayesian neural network for neurodegenerative diseases classification

David Payares-Garcia[1,2] | Jorge Mateu[3] | Wiebke Schick[2]

[1]ITC Faculty Geo-Information Science and Earth Observation, University of Twente, Enschede

[2]Institute for Geoinformatics, University of Münster, Münster, Germany

[3]Department of Mathematics, University Jaume I, Castellón de la Plana, Castellón, Spain

Correspondence
David Payares Garcia, ITC Faculty Geo-Information Science and Earth Observation, University of Twente, Hengelosestraat 99, 7514 AE Enschede, The Netherlands.
Email: d.e.payaresgarcia@utwente.nl

Magnetic resonance imaging (MRI) plays an increasingly important role in the diagnosis and prognosis of neurodegenerative diseases. One field of extensive clinical use of MRI is the accurate and automated classification of degenerative disorders. Most of current classification studies either do not mirror medical practice where patients may exhibit early stages of the disease, comorbidities, or atypical variants, or they are not able to produce probabilistic predictions nor account for uncertainty. Also, the spatial heterogeneity of the brain alterations caused by neurodegenerative processes is not usually considered, despite the spatial configuration of the neuronal loss is a characteristic hallmark for each disorder. In this article, we propose a classification technique that incorporates uncertainty and spatial information for distinguishing between healthy subjects and patients from four distinct neurodegenerative diseases: Alzheimer's disease, mild cognitive impairment, Parkinson's disease, and Multiple Sclerosis. We introduce a spatially informed Bayesian neural network (SBNN) that combines a three-dimensional neural network to extract neurodegeneration features from MRI, Bayesian inference to account for uncertainty in diagnosis, and a spatially informed MRI image using hidden Markov random fields to encode cerebral spatial information. The SBNN model demonstrates that classification accuracy increases up to 25% by including a spatially informed MRI scan. Furthermore, the SBNN provides a robust probabilistic diagnosis that resembles clinical decision-making and can account for the heterogeneous medical presentations of neurodegenerative disorders.

KEYWORDS
Bayesian inference, classification, deep learning, magnetic resonance imaging, neurodegenerative diseases, spatial information

1 | INTRODUCTION

Neurodegenerative diseases comprise a heterogeneous group of chronic and irreversible conditions characterized by the progressive degeneration of the structure and function of the central nervous system.[1] Neurodegenerative diseases attack the brain neurons by causing abnormal functioning and degeneration which eventually leads to their demise. The World Health Organization (WHO) predicts that by 2040, as many developed countries' aging population keeps

growing, neurodegenerative disorders will become the second leading cause of death after cardiovascular disease.[2] Thus, identifying and classifying these disorders is not only crucial in clinical practice but also beneficial for developing oriented treatments and enriching clinical research.

Magnetic resonance imaging (MRI) has become a reliable tool for the diagnosis and assessment of neurodegenerative diseases. In MRI data, neuronal damage caused by neurodegeneration is represented as atrophies or lesions whose intensities differ from those in surrounding apparently healthy tissue.[3] Measures of abnormalities derived from structural MRI data are markedly useful in the characterization of neurodegenerative chronic disorders such as Parkinson's, Alzheimer's, and Multiple Sclerosis.[1]

Due to the availability of MRI data, there has been an inherent interest in applying analytical methods to provide a differential diagnosis of neurodegenerative disorders. Many studies have focused on measuring the volume of regions of interest,[4,5] the volume of cortical thickness[6] and the concentration of gray matter (voxel-based morphometry)[7] as distinctive indicators of neurodegenerative conditions. However, these methods measure group differences and are not extendable to individual diagnoses. Other studies based on independent component analysis (ICA)[8] and principal component analysis (PCA)[9] have uncovered neurodegeneration hallmarks and relate them to a neurodegenerative disorder. Although these techniques have sounded in neurodegenerative diseases research, they are bounded by a modest number of patients and intolerable classification accuracy from a medical diagnosis scope.

In recent years, advances in machine learning and pattern recognition methods capable of handling high-dimensional data have enabled the development of new diagnostic tools derived from morphological analyses of MRI. One prominent approach is imaging feature extraction for classification and prediction using deep learning. Multiple deep learning studies have been successful in classifying and predicting neurodegenerative diseases by using supported vector machine (SVM)[10] and convolutional neural networks (CNNs).[11,12]

These approaches have achieved high accuracy in binary (over 93%) and multi-class (over 79%) disease classifications. Nonetheless, most of the studies focus on the distinction of different stages within the same disorder or the differentiation between healthy and pathological subjects. Such distinction does not mirror clinical practice, where brain abnormalities that define neurodegenerative diseases can be present before the onset of clinical features, and more than one neurodegeneration process can be found in an individual.[1] Studies in the classification of multiple neurodegenerative diseases are scarce since it entails combining and standardizing diverse MRI sources. Furthermore, predictions in deep learning methods are only partly suited for clinical diagnosis since they cannot produce probabilistic predictions nor account for uncertainty. This is particularly alarming in medical diagnostic applications, where silent failure can lead to dramatic results.[13] There is thus a need for classification methods that resemble clinical decision-making and account for atypical, multiple, and early presentations of neurodegenerative disorders.

Additionally, classification MRI-based studies ignore the spatial heterogeneity of the brain alterations caused by neurodegenerative processes. The spatial configuration of the neuronal loss is a characteristic hallmark for each disorder. In a number of studies mainly using statistical methods,[14-16] it has been shown that coupling the MRI data with spatial information significantly increases both the detection of and discrimination of neurodegenerative diseases. Novel approaches have proved that incorporating spatial components, such as geographic coordinates and region labels, into deep learning techniques for surface interpolation,[17] image segmentation and image recognition[18] can enhance model stability and classification accuracy.

This article aims to develop a classification technique that incorporates pattern recognition, uncertainty, and spatial information for distinguishing between healthy subjects and patients from four distinct neurodegenerative diseases: Alzheimer's disease, mild cognitive impairment (MCI), Parkinson's disease, and Multiple Sclerosis. In this work, a spatially informed Bayesian neural network (SBNN) is introduced to produce automated, probabilistic, and accurate predictions that resemble clinical diagnosis. A CNN is developed to detect a neurodegenerative process based on T_1-weighted MRI scans from patients and healthy controls, then Bayesian inference is incorporated into the CNN parameters to quantify uncertainty and produce probabilistic predictions and finally, a spatially informed MRI scan using hidden Markov random fields (HMRFs) is included as a CNN input to improve feature detection and classification accuracy. To illustrate the proposed method, this study uses 330 T1-weighted MRI scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database, the Parkinson's Progression Markers Initiative (PPMI) database and the Multiple Sclerosis database of the University Medical Center Ljubljana.

–WILEY

2 | MATERIALS AND DATA

2.1 | Study population

The dataset for this study was obtained from three different sources: the ADNI database, the PPMI database and the University Medical Center Ljubljana (UMCL).[19] The study comprises 330 participants divided into 150 healthy subjects and 180 neurodegenerative patients. A subset of 60 patients with Alzheimer's, 30 suffering from MCI and 75 cognitively normal subjects were obtained from the ADNI 3 cohort of the ADNI study. 60 patients diagnosed with Parkinson's and 75 healthy subjects were acquired from the PPMI database. The UMCL provided the data of 30 Multiple Sclerosis patients. A summary statistics of participants' demographic characteristics is reported in Table 1.

2.2 | Image sequences

Subjects were scanned on either 1.5 T or 3.0 T MRI devices according to the three studies' protocols. All scans include a 3-dimensional MP RANGE T1-weighted gradient-echo sequence. The voxel size of the T1-images oscillates between $0.42 \times 0.42 \times 3.3$ mm^3 and $1.0 \times 1.0 \times 1.2$ mm^3. 135 and 195 patients were imaged using 1.5 T and 3.0 T devices, respectively.

2.3 | Image preprocessing

The N4 inhomogeneity correction (improved N3 bias correction) algorithm[20] was applied to each MRI scan. As this study involves a population-level analysis, every image must have the same conditions in resolution, size, and spatial distribution. A non-linear spatial registration is performed to the N4-corrected T1-weighted scans to transform them into an unbiased high-resolution image with 2 mm isotropic voxels. The symmetric diffeomorphic image registration algorithm[21] implemented in the Advanced Normalization Tools (ANTs)[22] is used for this procedure. The use of a non-linear registration method guarantees the alignment of corresponding brain regions across subjects and allows the comparison of neural loss that characterizes each degenerative disorder. To remove non-cerebral tissue such as the skull and neck from the N4-corrected and registered images, a brain mask of the atlas template was applied. The Removal of Artificial Voxel Effect by Linear regression (RAVEL) method[23] was employed to normalize the image voxel intensities since these were collected on different scanners across different sites as part of different studies. RAVEL removes present unwanted variation by adopting control voxels unassociated with disease status and clinical covariates. 3-class tissue segmentation is performed using the FSL FAST segmentation algorithm[24] to obtain the control voxels within the cerebrospinal fluid (CSF) structure. Age and sex are used as clinical covariates. Finally, to correct the scanner and protocol effects and to preserve biological associations on the MRI scans, the ComBat image harmonization algorithm[25] was used. The cortical thicknesses of the 330 N4-corrected, registered, skull stripped and normalized images were harmonized. The steps mentioned above were integrated into an automatic pipeline to automate MRI preprocessing. We developed `Neurorom`, an R package to prepare structural MRI scans for analysis. `Neurorom` is an easily accessible package that allows users to preprocess multiple MRI scans in an automated fashion using minimal user input. The output of `Neurorom` is preprocessed MRI ready to use in our approach (see methodology). The reader is referred to Section 7 for further details on accessing the package and the raw code.

TABLE 1 Summary statistics of participants

	Healthy	AD	MCI	PD	MS
n	150	60	30	60	30
% Female	42	33	38	70	60
Median age	66	76	75	64	39
$[Q_1, Q_3]$	[58.0, 72.0]	[70.5, 80.2]	[65.5, 81.5]	[56.0, 69.2]	[33.0, 42.0]

Note: Q_1 and Q_3 stand for the first and third quartile of the patient's age, respectively.
Abbreviations: AD, Alzheimer's disease; MCI, mild cognitive impairment; MS, Multiple Sclerosis; PD, Parkinson's disease.
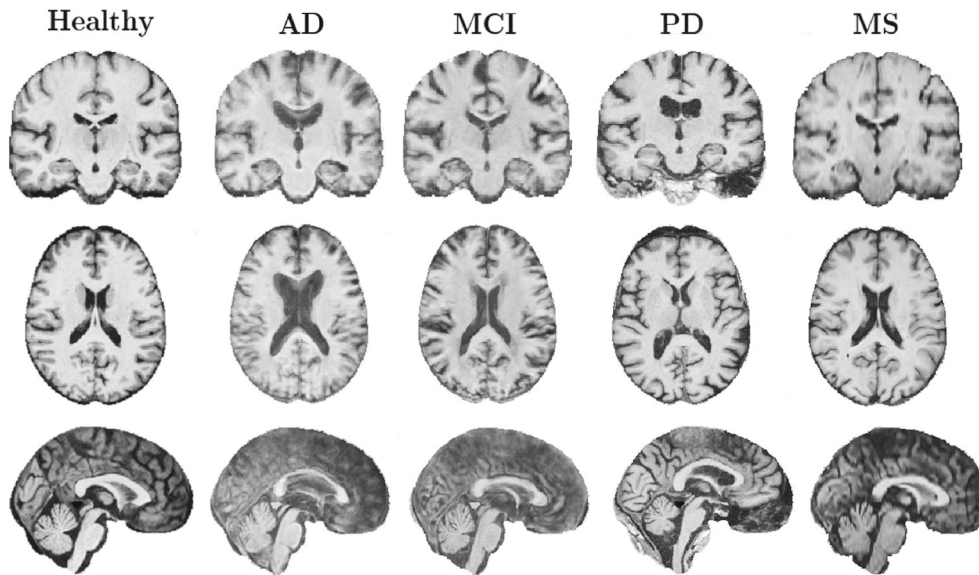
FIGURE 1 Sagittal, coronal, and axial planes of preprocessed MRI images for patients with advanced stage of a neurodegenerative disease (AD, MCI, PD, and MS) and a healthy subject

See Figure 1 as an example of preprocessed MRI images used. Figure 1 illustrates how preprocessing allows comparison and how spatially distributed features in the brain anatomy inform of a particular neurodegenerative disease. The differences are visually evident in Figure 1 as we deliberately selected MRI scans for patients in advanced stages.

3 | METHODOLOGY

3.1 | Spatially informed Bayesian neural network

Bayesian neural networks (BNN) combine neural network and stochastic models with the latter forming the core of this integration.[26] BNN introduces stochastic components in a neural network to simulate multiple possible models with an associated probability distribution. This allows explicit uncertainty inference in the underlying processes of the neural network.[13] BNN can produce uncertainties on the predictions and draw the distributions of the parameters learned by the model.

The SBNN is an extension of a BNN. The SBNN combines a CNN, a stochastic variational model, and a spatially informed input to produce classification predictions. Figure 2 displays the workflow to build the SBNN.

Since one of the main goals of the SBNN is to prove that the inclusion of an explicit spatial input in a BNN model improves classification accuracy, the SBNN model is fed not only by standard MRI data but also by spatially informed data. The standard data corresponds to the preprocessed MRI scans (x_λ). The spatially informed data (x_γ) is obtained after applying a HMRF model proposed by Zhange et al[24] to the preprocessed MRI scans. The HMRF model informs about the spatial structure using a neighboring context which allows forming mutually exclusive classes that depend on the spatial similarities of the entities. In brain imaging context, the HMRF segments MRI scans into classes that encode anatomical spatial relations. The labeling of each MRI scan voxel depends on the probability of that voxel's intensity belonging to a spatial neighborhood of similar voxel intensities. The HMRF model transforms the voxels' intensity response into a measure of spatial correlation. For further details, please refer to the aforementioned paper. In this study, three classes ($\ell = 3$) corresponding to the white matter, grey matter, and cerebrospinal fluid are selected. These structures are chosen as their spatial distribution and heterogeneity over the brain and volumes are highly correlated to the brain damage caused by neurodegenerative diseases such as Alzheimer's, Parkinson's, Multiple Sclerosis, and MCI.[27] The output spatially informed MRI scans summarize implicitly the neighborhood structures among the voxel intensities and, evidence abnormal changes—neuronal damage or death—in the three selected classes.

Both the preprocessed (x_λ) and the spatially informed MRI scans (or spatially informed input) (x_γ) are split into training, validation, and test datasets. Suppose we obtain observations from n subjects and let ($x_{\lambda i}, x_{\gamma i}$) denotes the preprocessed
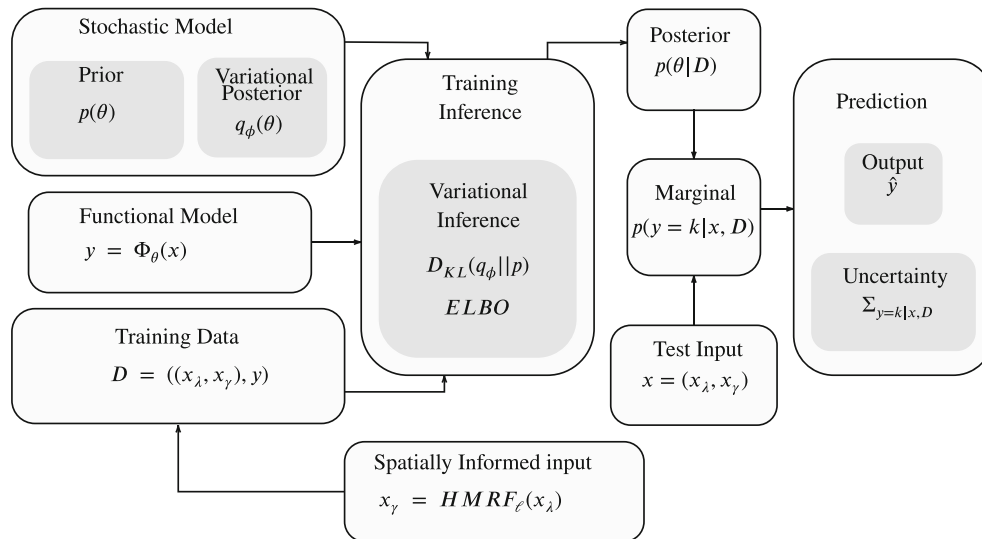
FIGURE 2 Pipeline to build, train, and test the spatially informed Bayesian neural network. Further details of the formalities are given in the text

and spatially informed MRI scans, and y_i denotes the label of the true ground disease k for the ith subject; k refers to one of the K classes, that is, one of the four neurodegenerative diseases or the healthy state. For training and validation purposes, each pair of MRI scans $(x_{\lambda i}, x_{\gamma i})$ is accompanied by its corresponding true ground disease (y_i). In the case of model testing, true classes y are only used to evaluate the predicted classes $\hat{y}$ of the test dataset. The training dataset (D)—composed by x_λ and x_γ—is the input of the SBNN model. A neural network model (Φ_θ) is chosen, in particular, we use a CNN. This study selected 4 brain classification state-of-the-art neural network architectures and, we additionally proposed a new one. The 5 neural networks are trained to detect the architecture that best represents and generalizes the MRI data and to evaluate the effect of the spatially informed input, that is, the spatially informed MRI scans. The next step is the choice of a stochastic model that will provide a prior distribution $p(\theta)$ for the bias parameters over the CNN weights and also for the variational posterior distribution $q_\Phi(\theta)$ to approximate the exact posterior probability distribution. The selected CNN is then trained again considering the introduced stochastic parameters and probabilistic loss function (see Section 3.4). The output is based on the posterior distribution over the model coefficients $p(\theta|D)$ needed to further obtain $p(y = k|x, D)$ which provides, given MRI scans x and the trained dataset D, the distribution of $\hat{y}$, the classification predictions. The predictions are summarized to find the final prediction—the highest averaged probability—and associated uncertainty.

Data augmentation for the input x was performed by rotating, translating, and flipping the MRI scans. However, it did not affect the models' performance and prediction.

3.2 | CNN architecture

The aim of a neural network model Φ_θ is to produce predictions by representing an arbitrary function.[13] Neural networks' architecture is built using a sequence of input layers, hidden layers, and one output layer composed of a linear transformation and an activation function. Each layer depends on a parameter θ which represents the weights W and bias b of the network. In a Φ_θ, θ is optimized to minimize the differences between the predicted and real values given the training data D.

A 3D CNN is a special kind of neural network applied to analyze 3D visual imagery. It exploits the image data topology and reduces it to its key features which are used as classification criteria. The neural network Φ_θ used in this article is a 3D CNN with two inputs and two branches. The first branch is designed to extract high-level features from the preprocessed MRI scans. The second branch extracts key characteristics from the spatially informed images. Both branches have identical architectures to maximize feature detection and guarantee common output dimensions. A fusion layer combines the output semantic features of both CNN branches which are passed on to a set of fully connected layers (FC). The FC layers and a SoftMax activation at the last layer are trained to assign a predefined class to the input 3D images, that is, to predict the neurodegenerative disease for each patient.

TABLE 2 State-of-the-art CNN-based methods for neuroimaging multi-class classification

Author	Disease	n classes	MRI scans	Testing accuracy (%)
Ker et al[28]	CT brain hemorrhages	5	399	72.68
Payan and Montana[29]	Alzheimer's stages	3	2275	89.47
Mzoughi et al[30]	Brain glioma tumors	2	284	87.2
Parmar et al[31]	Alzheimer's stages	5	120	93.10

TABLE 3 Description of the convolutional neural network (CNN) architecture proposed in this article

Operation layer	Filters	Filter size	Stride value	Units
Convolution (ReLU)	64	$3 \times 3 \times 3$	$1 \times 1 \times 1$	-
Max pooling	1	$2 \times 2 \times 2$	$1 \times 1 \times 1$	-
Batch normalization	-	-	-	-
Convolution (ReLU)	64	$5 \times 5 \times 5$	$1 \times 1 \times 1$	-
Max pooling	1	$2 \times 2 \times 2$	$2 \times 2 \times 2$	-
Batch normalization	-	-	-	-
Convolution (ReLU)	64	$7 \times 7 \times 7$	$1 \times 1 \times 1$	-
Max pooling	1	$2 \times 2 \times 2$	$3 \times 3 \times 3$	-
Batch normalization	-	-	-	-
Convolution (ReLU)	64	$9 \times 9 \times 9$	$1 \times 1 \times 1$	-
Max pooling	1	$2 \times 2 \times 2$	$4 \times 4 \times 4$	-
Batch normalization	-	-	-	-
Fully connected (ReLU)	-	-	-	512
Fully connected (ReLU)	-	-	-	128
Fully connected (SoftMax)	-	-	-	5

Five 3D CNN architectures are used for the neural network model branches. We use four 3D CNN architectures based on related state-of-the-art brain multi-class classification and we additionally propose a new one.

The first state-of-the-art 3D CNN architecture[28] uses three simple convolutional layers and two fully connected layers to classify computed tomography brain scans into healthy scans and abnormal scans containing subarachnoid hemorrhage, intraparenchymal hemorrhage, acute subdural hemorrhage, and brain polytrauma hemorrhage. The second architecture, proposed by Payan and Montana,[29] consists of three convolutional layers and three FC layers. This architecture differentiates between Alzheimer's disease, MCI, and healthy subjects. In a third architecture, Mzoughi et al[30] propose a complex CNN architecture to distinguish between low-grade and high-grade gliomas: eight convolutional layers with three fully connected layers. Finally, another multi-class classification approach is suggested by Parmar et al.[31] In their study, they classify four Alzheimer's disease stages using fMRI scans and a set of five convolutional layers and three FC ones. Table 2 shows the details of the used state-of-the-art 3D CNN architectures for brain imaging classification. Further details of the models can be found in the mentioned studies.

The architecture we propose has four convolutional layer groups and three fully-connected layers. The detailed configuration of the network is shown in Table 3. The input to the CNN is a 3D preprocessed and spatially informed MRI scan with $91 \times 109 \times 91$ dimensions. The convolutional layers compute their outputs by applying the convolutional operations and ReLU activations. The results are followed by max-pooling and batch normalization. The last three layers in the CNN are fully connected. These FC layers include the neurons connected to all outputs of their precedent layers. The last FC layer with a SoftMax activation has 5 neurons, which represent the probabilities of a patient to belong to one of the five classes (healthy, AD, MCI, PD, MS).

We proposed the architecture in Table 3 responding to the aspect and dimensions of the brain changes caused by neurodegenerative diseases. Initially, our network uses smaller filters to collect as much local information as possible,

for example, the transition from one bran tissue to another. Then, we gradually increase the filters' width to reduce the generated feature space to represent more general, high-level and representative information such as brain lesions, abrupt tissue changes, and anatomical distortions.

The first convolutional layer easily captures minute details. Once the layer deepens, we accumulate and aggregate the features by increasing the layer window size. We also increase the max-pool accordingly to reduce redundancies in the features. We gradually increased our filter size by two strides in each successive layer so that the detected feature will be extracted at a low, intermediate, and high level with a bigger window for the successive layers. We call this a hierarchical network because the size of the kernel increases as the stride does. Nonetheless, we keep constant the number of filters in each as we seek to retain details in more expansive windows, that is, account for brain changes in extensive regions of the brain.

3.3 | Bayesian inference

The Bayesian paradigm in statistics is based on two simple ideas. The first is that probability is a measure of belief in the occurrence of events, and the second idea is that prior beliefs influence posterior beliefs. Bayes' theorem summarizes this interpretation, stating that

$$P(H|D) = \frac{P(D|H)p(H)}{P(D)} = \frac{P(D, H)}{\int_H P(D, H')dH'}. \tag{1}$$

Here, $P(D|H)$ is the likelihood whereas $P(H)$ is the prior and $P(D)$ the evidence. Using Bayes' formula to train a predictor can be understood as learning from the data D. BNNs are stochastic neural networks trained using a Bayesian approach. The goal of artificial neural networks (ANNs) is to represent an arbitrary function $y = \Phi(x)$. ANNS are built using one input layer, a succession of hidden layers and one output layer, where each layer is represented as a linear transformation, followed by a nonlinear operation (activation function). In this process, we have a set of parameters of the network θ formed by the weights W of the network connections and the biases b of the network layers. Deep learning is the process of regressing the parameter vector θ from the training data D, with D composed of a series of input x and their corresponding labels y. In addition, the stochastic model assumes that θ is drawn from a probability distribution. The main goal in using this stochastic framework is to obtain a better idea of the uncertainty associated to the underlying processes. This can be done by comparing the predictions of multiple sampled model parametrizations θ. This process can be formalized as $\theta \sim p(\theta)$, and $y = \Phi_\theta(x) + \epsilon$, with ϵ a random noise. Thus, we need a prior distribution over the possible model parametrization $p(\theta)$, and a prior confidence in the predictive power of the model $p(y = k|x, \theta)$, where k is one possible label of the response y.

We opt here for a variational inference method, following.[32] This method performs a Monte Carlo approximation of the posterior distribution by integrating over the kernel and bias. In this case, a surrogate posterior $q_\phi(\theta)$ (parametrized by ϕ to match the exact posterior), and divergence for both the weights and bias distributions are to be specified.

The prior $p(\theta)$ and variational posterior $q_\phi(\theta)$ are taken as Gaussian distributions with initialized mean $\mu = 0$ and a small variance σ^2. The choice of using a normal prior is not arbitrary. Assuming a known normal prior is a common practice to not over-parametrized neural networks.[32] Besides the ideal mathematical properties of a Gaussian process and the simple formulation of its log, which is used in most of the learning algorithms, a multivariate normal surrogate posterior $q_\phi(\theta)$ dramatically simplifies the probabilistic loss function (ELBO), leading to a generalization of the expectation-maximization algorithm.[13] In Vladimirova et al,[33] the author proved that Gaussian priors also avoid overfitting and improve generalization. Malinin et al[34] also showed that Gaussian priors in multi-layered architectures perform comparatively well against other distributions such as the (non-informative) Uniform and the Weibull. Furthermore, as the number of hidden layers increases in a neural network, the prior will converge to a Gaussian process regardless of the defined initial distribution.[35]

Assuming the training set is D, we denote by D_x the training inputs, and by D_y the training labels. Thus, applying Bayes' theorem, the resulting posterior distribution of θ, given the training data D, can be written as

$$p(\theta|D) = \frac{p(D_y|D_x, \theta)p(\theta)}{\int_\theta p(D_y|D_x, \theta')p(\theta')d\theta'} \propto p(D_y|D_x, \theta)p(\theta). \tag{2}$$

Note that the marginal probability distribution $p(y|x, D)$ quantifies the model's uncertainty on its prediction, and it is our quantity of interest. Given $p(\theta|D)$, it can be calculated for each $k = 1, \ldots, K$ as

$$p(y = k|x, D) = \int_{\theta} p\left(y = k|x, \theta'\right) p\left(\theta'|D\right) d\theta'. \tag{3}$$

In practice, a large number of θ_j are sampled from $p(\theta|D)$, and used to compute the y_j from the equation $y = \Phi_{\theta}(x) + \epsilon$, taking samples from $p(y = k|x, D)$. The output becomes $Y = \{y_j | j \in [1, N]\}$ a set of samples from $p(y|x, D)$, and $\Theta = \{\theta_j | j \in [1, N]\}$, where N represents the Monte Carlo samples.

For classification purposes, the average model prediction will give the relative probability of each class k, considered a measure of uncertainty,[13] and it is given by

$$\hat{p_k} = \frac{1}{N} \sum_{j=1}^{N} I\{\Phi_{\theta_j}(x) = k\}, \quad \text{for } k = 1, \ldots, K, \tag{4}$$

where I is the indicator function, and $N = |\Theta|$. The final prediction is taken as the most likely class of the set $\hat{\boldsymbol{p}} = \{\hat{p_1}, \hat{p_2}, \ldots, \hat{p_K}\}$

$$\hat{y} = \arg \max_{k} \hat{p_k}. \tag{5}$$

3.4 | Training, inference, and prediction

The SBNN is trained using variational inference given the training dataset D. The variational inference samples the distribution $q_{\phi}(\theta)$ parametrized by parameters ϕ. The SBNN learns the values of ϕ such that $q_{\phi}(\theta)$ resembles closely the exact posterior, and the Kullback-Leibler divergence measures this matching.

In terms of ϕ, the Kullback-Leibler divergence is defined as

$$D_{KL}\left(q_{\phi}\|p\right) = \int_{\theta} q_{\phi}\left(\theta'\right) \log\left(\frac{q_{\phi}\left(\theta'\right)}{p\left(\theta'|D\right)}\right) d\theta'. \tag{6}$$

In order to compute $D_{KL}\left(q_{\phi}\|p\right)$, $p(\theta'|D)$ has also to be computed. To avoid this problem, the evidence lower bound ($ELBO$) is defined as

$$ELBO = \int_{\theta} q_{\phi}\left(\theta'\right) \log\left(\frac{p\left(\theta', D\right)}{q_{\phi}\left(\theta'\right)}\right) d\theta' = \log(p(D)) - D_{KL}\left(q_{\phi}\|p\right). \tag{7}$$

Minimizing $D_{KL}\left(q_{\phi}\|p\right)$ is equivalent to maximizing the $ELBO$ given that $\log(p(D))$ only depends on the prior. The $ELBO$, as any other loss function in CNN, can be optimized by applying stochastic gradient descent to variational inference.[36]

The trained SBNN is used to produce predictions over the test dataset. Monte-Carlo inference is applied to each input of the test dataset to get the corresponding probability distribution $p(y|x, D)$ depending on the estimated output posterior $P(\theta|D)$. This is performed by executing the SBNN n times (or iterations) for a test dataset input x.

A vector of relative probabilities (uncertainties) per disease is obtained by averaging the n predictions, that is, the mean value of the probability distribution. The final predicted class $\hat{y}$ is taken as the most likely class.

4 | EXPERIMENTAL SETUP

4.1 | Evaluation metrics

The SBNN performance is measured in terms of accuracy, sensitivity, specificity, and F_1-score statistics assessed with true positive (TP), true negative (TN), false positive (FP), and false negative (FN) values. The expressions of the statistical

measures are given in Equations (8)–(11)

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}, \tag{8}$$

$$Sensitivity = \frac{TP}{TP + FN}, \tag{9}$$

$$Specificity = \frac{TN}{TN + FP}, \tag{10}$$

$$F_1\text{-score} = \frac{2TP}{2TP + FP + FN}. \tag{11}$$

Accuracy measures the correctly identified MRI scans over the total classification. Sensitivity describes the portion of MRI scans correctly classified in the correct class, whereas specificity refers to the ratio of MRI scans incorrectly classified within the incorrect class. F_1-score measures the similarity between the predicted $\hat{y}$ classes and the target y classes.

4.2 | Experimental scheme

To train, validate, and test the SBNN model, the 330 MRI scans were split using a simple random sampling technique. The training dataset, used by the SBNN model to learn the features of each disease, consists of 231 (70%) T_1-weighted preprocessed scans. The validation dataset and test dataset are composed of 51 (15.4%) and 49 (14.6%) images, respectively. The full dataset was split per disease to guarantee a balanced distribution in each subset dataset.

The CNN architectures are trained for a maximum of 100 iterations (epochs) using the Adam optimizer with an exponential decay function of initial rate $\delta = 1 \times 10^{-5}$. For fine-tuning the CNNs, we include L2 regularizations and dropouts of 0.5 after each dense layer. We also added batch normalization layers. The batch size is set to a default value of 32 training samples. The input training and validation sets are shuffled upon generation. The CNN architectures were trained using six random dataset splits to assess their robustness in the presence of different data partitions.

The 3D CNN with the highest accuracy over the test dataset is chosen and transformed into a SBNN using variational inference. Then, the SBNN is trained, validated and tested using the same parameters mentioned above, except the loss function, which in a BNN can be defined as in expression (5). The bias and weights defining the parameter θ are given a Gaussian distribution for both the prior and variational posterior. The mean μ and σ^2 variance are initially set to 0 and 0.1 respectively. The Kullback-Leibler divergence function is weighted by 231, the number of training samples.

The Monte-Carlo number of iterations is set to $n = 1000$. Probability distributions per disease are obtained for each of the 49 test samples. For one sample, the mean value of every disease probabilities is chosen to obtain the predicted relative probabilities. The disease with the highest relative probability is taken as the final prediction.

5 | RESULTS

In this section, the performance of the proposed SBNN for the classification of neurodegenerative disorders is evaluated. First, the overall classification accuracy of the five CNN architectures with and without the explicit spatially informed input is measured to quantify the spatial information incidence and to select a suitable architecture for the current application. The trained SBNN is used to predict the unknown testing dataset with 49 preprocessed and spatially informed MRI scans. The performance of the SBNN is discussed based on classification evaluation metrics. Since uncertainty is a relevant aspect in clinical diagnosis, confidence intervals are conducted for each neurodegenerative disorder given the test dataset. Finally, the predictions of three MRI scans from three different AD patients are constrained with the patients' clinical evaluation to assess the SBNN predictive properties. The proposed SBNN is implemented with the high-level neural networks API Keras using Python 3.7. The networks are trained on the Google Colab platform equipped with 12 GB NVIDIA Tesla K80 GPU and Xeon Processors CPU@2.3 GHz.

TABLE 4 Testing accuracy for the four state-of-the-art architectures and our proposed one

	Testing accuracy	
Architecture	**MRI**	**SiMRI**
Ker et al28, [28]	0.49	0.74
Payan and Montana[29]	0.60	0.68
Mzoughi et al[30]	0.61	0.76
Parmar et al[31]	0.56	0.68
Our architecture	0.64	0.83

Abbreviations: MRI, one branch, only preprocessed MRI scans; SiMRI, two branches, preprocessed and spatially informed MRI scans.

5.1 | Spatially informed input

The five CNN architectures were trained using the same model hyper-parameters, training, validation, and test datasets. The accuracy for the testing dataset defined in Equation (8) is used to measure the performance of each CNN architecture and the effect of the spatially informed input on them.

Table 4 shows the accuracy results of the five networks for one random split of the dataset including and excluding the spatially informed input. For the five architectures, omitting the spatially informed input, the accuracy oscillates between 49% to 64%. Among the CNN architectures, the lowest classification accuracy (49%) was obtained from the model of Reference 28. The highest accuracy for this group corresponds to our proposed architecture with a value of 64%. Although it represents a relatively high value compared to other architectures, it is considered a low value for classification purposes.

Furthermore, the accuracy for the five CNNs architectures with the spatially informed input ranges from 68% to 83%. Our architecture outperforms the state-of-the-art ones in terms of the accuracy of the testing dataset (83%). The architecture presented by Mzoughi et al[30] shows the second-highest accuracy among the five models. The CNN configuration from Reference 31 as well as the one from Reference 29 exhibit the lowest accuracy with a same value of 68%. A similar scenario occurred with other five random partitions of the dataset, that were considered to check for robustness in the results against the selected starting partition. Our architecture displays the highest testing accuracy in all splits ranging from 76.5% to 87.4%.

In each architecture, the accuracy for the testing dataset improved by adding the spatially informed input. Overall, the accuracies enhanced between 8% and 25%, four out of the five architectures presented an increment of more than 10%.

To visualize how the spatially informed input is improving the architectures accuracy, slices of a feature map for the first convolutional layer of our proposed architecture are shown in Figure 3. The feature map corresponds to a MRI scan of a MS patient. Typically, the areas colored in yellow and red in the feature map are called activated regions which represent the extracted key features from the input data. In the brain imaging context of neurodegenerative conditions, the activated regions express characteristic features associated with a neurodegenerative process, that is, brain regions affected by the neuronal damage.

It can be seen that the activated regions in the preprocessed MRI scan (Figure 3A) are mainly located in the outer parts of the occipital and parietal lobes, which are unassociated manifestations in the brain produced by MS.[1] Contrarily, the activated regions in the spatially informed scan (Figure 3B) are distributed over the brain. Some activations are found in the front and parietal lobes as small polygons which represent the white matter lesions caused by the disease. Likewise, other activated regions emphasize the corpus callosum and the cerebellum. One manifestation of MS is the slimming of these structures. Other activated areas might reflect possible cerebral atrophy in most parts of the lobes.

The combination of the preprocessed and spatially informed MRI scan boosts the CNN detection of physiological changes in the brain produced by a neurodegenerative disease. Note that abnormalities identified by the CNN only inform about the current type of degenerative disorder but do not provide any insight about disease staging or prognosis. Furthermore, although activation regions in the feature maps matched some known neurodegenerative disorders biomarkers, these cannot be directly considered informative neuroimaging features.
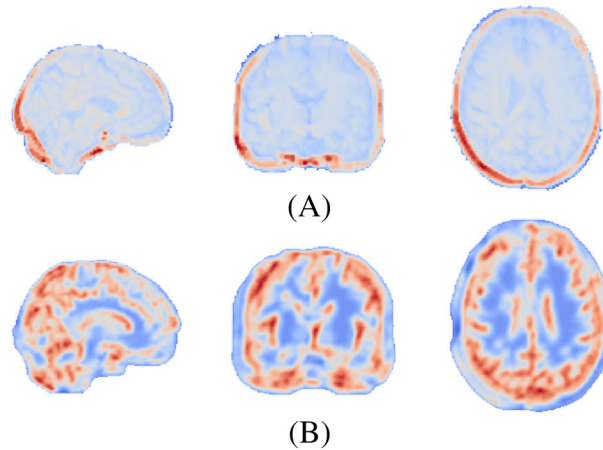
FIGURE 3 Sagittal, coronal, and axial slices from the feature map of the first convolutional layer for a MS patient; (A) preprocessed scan MRI scan, (B) spatially informed MRI scan. The color of each voxel represents the activation value

TABLE 5 Evaluation metrics for the SBNN model and the studied classes

Disease	Sensitivity	Specificity	F_1-score
AD	0.76	0.94	0.70
MCI	0.25	1.00	0.40
PD	1.00	0.92	0.85
MS	1.00	1.00	1.00
Healthy	0.88	0.86	0.88
Accuracy	-	-	0.81
Macro avg	0.77	0.77	0.75
Weighted avg	0.85	0.81	0.82

5.2 | Model evaluation

As our proposed CNN architecture outperformed the four state-of-the-art ones, it was transformed using variational inference to include uncertainty (our SBNN approach).

Initially, the CNN model without Bayesian inference was trained and tested. It showed an overall (testing) accuracy of 83%. When the stochastic model, that is, the prior and surrogate posterior over the CNN parameters were incorporated in the CNN (Bayesian CNN), the accuracy decreased to 81.5%. A loss of 1.5% occurred by introducing the stochastic model to the neural network. Although the CNN accuracy presented a reduction, confidence in the predictions can be inferred from the hybrid model. It is a fair trade-off between accuracy and inclusion of uncertainty.

Since now we have access to the probability distributions associated with the model parameters, we evaluated the model with 1000 trials and the 48 MRI scans of the test dataset to obtain an average accuracy of 81.5% with a SE of 1.5%. Evaluation metrics for each disease were calculated based on the averaged multi-class confusion matrix for the 1000 trials. Macro average and weighted average metrics are also provided to evaluate the model performance with consideration to the unbalanced number of samples per class. Table 5 shows the performance evaluation for the entire model and for each class.

As observed in Table 5, the highest sensitivity with a perfect score of 1.00 was obtained for the classification of the diseases MS and PD. Healthy and AD sensitivities are highly accurate with values of 88% and 76%, respectively. MCI presents the lowest sensitivity (25%) of the group; only one out of four MCI MRI scans is correctly classified in the MCI class. Specificity for the five diseases varies from 50% to 100%. In MS, all the images were correctly classified: the specificity and sensitivity are perfect. In the case of a healthy subject, there is only a 14% (1-0.86) chance of a MRI scan being assigned to another disease. F_1-score might be considered as an alternative precision measure since there

TABLE 6 95% confidence intervals for the mean of the predicted probabilities for each disease

Disease	Mean	Interval
AD	0.754	(0.734, 0.778)
MCI	0.336	(0.236, 0.657)
PD	0.870	(0.820, 0.917)
MS	0.976	(0.952, 0.999)
Healthy	0.868	(0.834, 0.902)

is an uneven class distribution in the test dataset. In terms of the F_1-score, most of the diseases perform well except for MCI. MS exhibits a perfect score and both PD and healthy scores are considerably high. AD F_1-score is lower than their counterparts sensitivity and specificity due to the wrong classification of MCI MRI scans into the category. Most of the classes met the criteria considered suitable to distinguish between the neurodegenerative diseases studied in this article.

Macro average resumes the evaluation metrics for all diseases assuming that the classes are balanced. Considering the unbalanced test dataset, the weighted average accuracy better summarizes these metrics. The SBNN has an overall sensitivity of 85%, a specificity of 81%, and an F_1-score of 82%. These metrics are consistent with the SBNN averaged accuracy and support the potential predictive properties of the model for classifying neurodegenerative diseases.

A confidence interval indicates a likely range of values that encompasses the expected value with a certain probability; this is a measure of uncertainty. Table 6 introduces the 95% confidence intervals for each disease. MCI probabilities are lower compared to other diseases; this can be explained by the number of samples in the test data and by the overlapping features it presents concerning other categories such as AD and healthy. As well, MCI presents a wider range which translates as diagnostic variability given that it might be a precursor of AD or PD. MS MRI range of probabilities is almost perfect which means the SBNN can easily categorize it. MS distinguishes itself from other disorders as the neural damage appears as small lesions distributed over the brain. This distinctive characteristic allows our model to easily recognize it resulting in a narrow predicted mean interval. PD and healthy classification accuracy can rise up to 90%, a relatively high accuracy value considering the false positives coming from other classes (see Table 5). AD predictions are precise given the narrow range of its confidence interval. AD predicted mean is a sensible value considering that MRI scans from AD can be misclassified as MCI and PD. Various anatomical changes of the brain produced by early stages of AD coincide with those caused by MCI and PD.[3]

5.3 | Model predictions

The SBNN was executed 1000 times in each MRI scan to obtain the probability distributions per disease. The mean and SD values were extracted from the probability distributions to represent the predicted probability and its error class-wise. The class with the highest probability is considered the predicted class for that MRI scan. If the predicted class corresponds to the ground truth class, the classification is correct. If the predicted class differs from the ground truth class, the classification is considered incorrect. If none of the average probabilities per disease exceeds a threshold probability of 0.5 (the minimum acceptable probability), the MRI scan is not assigned to any of the five classes, that is, the classification is uncertain.

This procedure was performed in the 48 images of the test dataset. Three prediction outputs for three different AD patients' MRI scans are presented in Figure 4. The probability distributions are log-transformed to improve visualization; the closer the values are to 1, the higher the probability is. This information is provided by the ADNI database. The results were compared to the diagnosis summary of the clinical evaluation of the patients.

The first MRI scan (Figure 4A) corresponds to a patient diagnosed with AD with mild dementia symptoms. Cerebral atrophy is evident in the patient's MRI. The SBNN correctly classified the MRI scan; the mean probability is 0.83 and the error is 0.13. The average probabilities for the other diseases are very low, and in most cases, are equal to 0. There is an 83% chance that the MRI scan belongs to an AD patient. The second scan (Figure 4B) belongs to an Alzheimer's
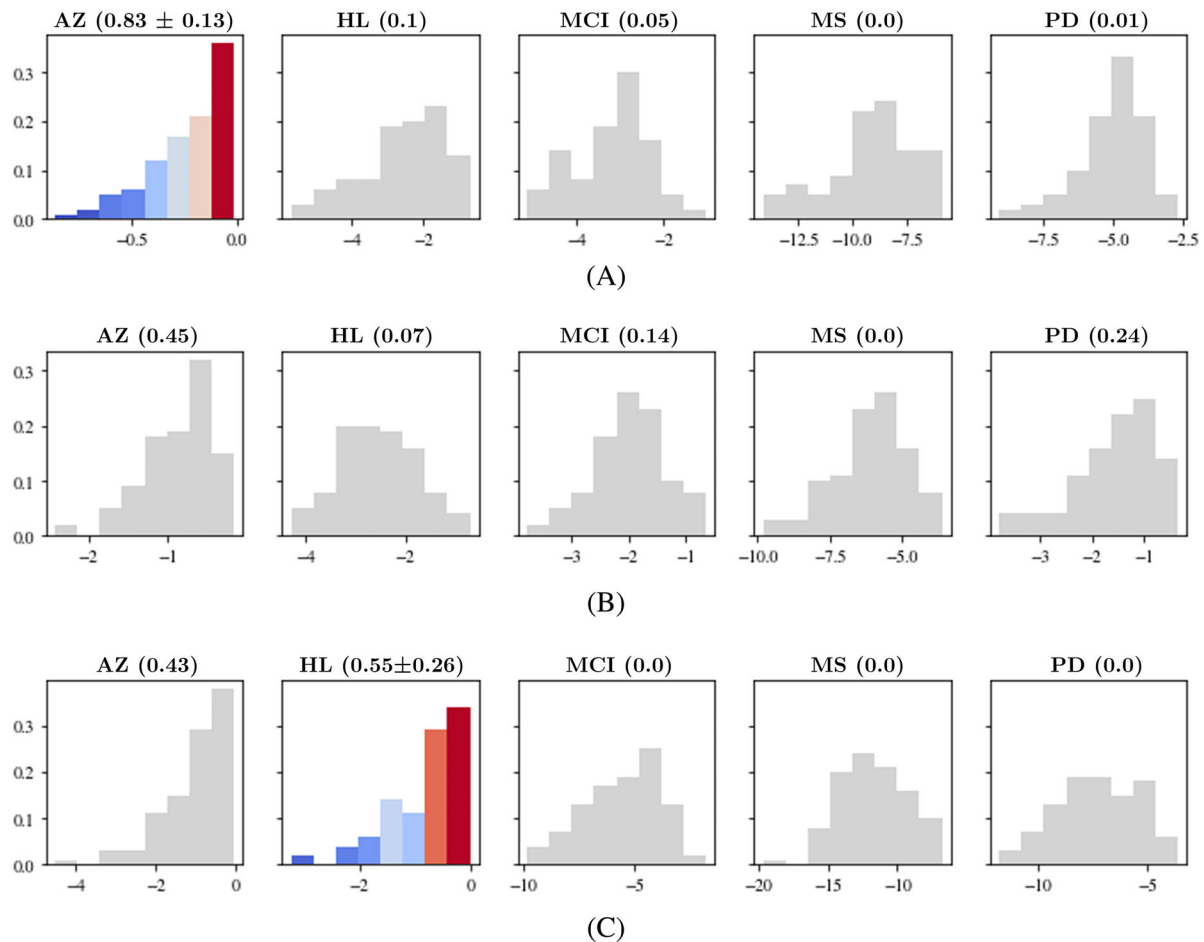
FIGURE 4 Illustration of neurodegenerative diseases classification. (A) AD patient, MRI scan correctly classified; (B) AD patient, MRI scan not classified; (C) AD patient, MRI scan incorrectly classified

patient with mild declining dementia impairment. In the clinical diagnosis, the patient was initially diagnosed with MCI. However, due to the frequent movement disorders reported in the cognitive assessment test, it was determined that the patient resided in the Alzheimer's disease group. According to the clinical summary, the diagnosis is provisional, and further sessions are needed to assess the patient and provide a final decision. This MRI scan is not classified into any disease even though the AD class (0.45) exhibits the highest probability among the group. The prediction also suggests that the patient presents PD features (0.24). This value matches the symptomatology reported in the cognitive test; PD is characterized by the loss of motor functions. The SBNN uncertain prediction for this scan reflects the clinical diagnosis of the patient. The third MRI scan (Figure 4C) comes from an AD patient with no evident memory disorder according to the clinical assessment test. However, the mass spectrometry measurement in the biospecimens results shows low levels of the Aβ-42 protein. Such Aβ-42 level decrease is a known biomarker in early stages of AD.[37] The prediction for this patient's scan shows an incorrect classification. Healthy is the most likely predicted class with a probability of 0.55 and an error of 0.26. Three main aspects affect the classification: (i) the mean probability surpassed the decision threshold only by 0.05, (ii) the error is half of the predicted probability, and (iii) the mean probability for AD is 0.43. For this scan, given predicted probabilities, their errors, and the level of uncertainty of the SBNN Model, a clinician can decide whether to integrate or not more information to support their diagnosis.

6 | DISCUSSION

In this work, the SBNN is presented to classify four neurodegenerative disorders, Alzheimer's disease, MCI, Parkinson's disease, and Multiple Sclerosis, and healthy subjects. To improve accuracy, the model took advantage of spatially informed

MRI data (introduced as a regional brain segmentation) and used a Bayesian approach in which the model predictions are drawn with confidence.

The results show a substantial increment (up to 25%) in the classification accuracy of 3D CNNs by incorporating a spatially informed MRI scan. Such an increment is a direct consequence of enhancing the CNN feature detection through spatial information of the brain. While the features identified by our method might represent anatomical biomarkers, we used them for no other purpose than classification. Including the explicit spatial scan results in detailed differentiation and increased separability between diseases. These results support previous findings in brain imaging statistical studies[15,16] where disease classification was improved by adding spatial information. Although the segmentation MRI scan enhances the distinction of neurodegeneration processes, there is a limitation in this approach. In highly correlated diseases, such as Alzheimer's and MCI, the captured spatial changes are insufficient to separate these conditions unequivocally.

This research also proposes a CNN architecture able to identify and classify neurodegenerative processes. The CNN yielded an accuracy of 81.5% and 83% with and without variational inference, respectively. The proposed architecture understands and generalizes the most relevant features of the diseases of this study. To the best of our knowledge, the SBNN model is the first to investigate four distinct neurodegenerative diseases using MRI data.

An advantage of the SBNN model lies in the fact that it accounts for uncertainty. The SBNN model formulation explicitly allows for probabilistic estimation of the CNN parameters which are used in the convolutional layers. Additionally, the Bayesian framework considers inference on the predictions by summarizing the posterior predictive probabilities. Concerning the prediction of a neurodegenerative disorder, the probabilistic formulation of the SBNN model naturally accounts for stochastic uncertainty (inherent random noise) and epistemic uncertainty (ignored information), and therefore, better represents the uncertainty in clinical diagnosis. SBNN predictions can not only detect the most likely disease of a patient, but they can also inform about patients with atypical variants, numerous pathologies, early disease stages, or diseases ignored in this research.

The overall classification accuracy of the SBNN model shows that approximately 8 out of 10 MRI scans are classified correctly. However, there is room for improvement regarding individual disease accuracy. PD, MS and HL accuracy surpassed the overall accuracy, but MCI and AD could not. MCI's low accuracy is not surprising since many studies consider this condition an early stage of AD and a subtle transition between normal aging old patients and AD.[7] Most of the MCI MRI scans were wrongly classified by the model into the AD class. MS class presents perfect evaluation metrics. MS is the only condition whose manifestation in the brain occurs in the white matter[5] and thus, the spatially informed input (or brain structure segmentation scan) provides richer information for the classification of this disorder.

It is worth noting that data preprocessing is instrumental in comparing and analyzing MRI data coming from diverse data sources. This study indirectly proposes a data preprocessing pipeline to integrate MRI scans from neurodegenerative patients. Although preprocessing varies depending on the research purpose, most brain imaging classification studies need to standardize and clean data to guarantee accurate predictions.[23] This work shows that the proposed preprocessing schema could be implemented in neuroimaging studies involving numerous MRI data sources.

In future studies concerned with the same problem, the accuracy may be improved by including exogenous variables, related to clinical assessments, which account for uncaptured hallmarks of the diseases. Clinical trials suggest that MRI-based information coupled with information derived from clinical evaluations and individual characteristics lead to accurate differential diagnosis,[1,2,8] for example, psychological and cognitive tests that describes the antecedent clinical history and behavioral aspects of the patient. Likewise, use datasets that balance patient's information (age and sex) since research have shown that these variables define the population affected by neurodegenerative diseases. For example, The incidence of AD is higher amongst women than men.[6] A direct extension that might be considered is the addition of other spatial information, for example, relative coordinates of voxels, distances between brain regions, and voxel intensity densities. These alternative ways to encode spatial information[17] could enrich disease detection and account for subtle differences between the voxel intensities from two similar disorders, for example, AD and MCI. Another approach unexplored by this work but frequent in clinical practice is the inclusion of other MRI modalities. Although in most of the cases T_1-weighted scans display neurodegeneration biomarkers,[3] other modalities such as T_2-weighted and FLAIR scans can be used to support the findings of the former. In a more robust model, other modalities could be added to account for disease features that are undetectable in the T_1-weighted sequence.

The findings of this study are expected to be of direct clinical interest for neurologists, physicians, and medical researchers. In a comprehensive overview, the most important contribution of this work is that it is one of the first works

that consider an explicit spatial input for disease classification, and assists diagnosis with uncertainties. The spatially informed MRI data is not only limited to neurodegenerative diseases. Given its nature, it is extendable to other brain conditions such as tumors, gliomas, and strokes, in which brain damage is also spatially located. Likewise, the SBNN model can be brought to bear on other brain imaging modalities.

7 | CONCLUSION

In conclusion, this article introduces a reliable automatic tool that can provide a fast and accurate diagnosis, and supports the diagnosis and treatment of neurodegenerative disorders. One of the main benefits of the proposed model is the combination of explicit spatial information in MRI data, multiple diseases, and the ability to generate probabilistic diagnoses.

The SBNN was conceived as a tool to facilitate automatic medical diagnosis. The `Neurorom` package transforms the raw MRI scans into prepossessed MRI scans to be the inputs of the SBNN. The R package can be installed into the R platform from https://github.com/DavidPayares/neuronorm. The SBNN implementation, including data loading, model training, and model evaluation, can be found at https://github.com/DavidPayares/SBNN. The sequential use of both tools produces an automatic pipeline to diagnose neurodegenerative diseases using MRI.

CONFLICT OF INTEREST
The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

DATA AVAILABILITY STATEMENT
Data sharing is not applicable to this article as no new data were collected in this study.

ORCID
David Payares-Garcia https://orcid.org/0000-0001-6130-7450
Jorge Mateu https://orcid.org/0000-0002-2868-7604

REFERENCES
1. Dugger BN, Dickson DW. Pathology of neurodegenerative diseases. *Cold Spring Harb Perspect Biol.* 2017;9(7):a028035.
2. Gammon K. Neurodegenerative disease: brain windfall. *Nature.* 2014;515(7526):299-300.

3. Stoessl AJ. Neuroimaging in the early diagnosis of neurodegenerative diseases. *Transl Neurodegener*. 2012;1(5):1-6.

4. Serrano-Pozo A, Frosch MP, Masliah E, Hyman BT. Neuropathological alterations in Alzheimer disease. *Cold Spring Harb Perspect Med*. 2011;1(1):a006189.

5. Smeets D, Ribbens A, Sima DM, et al. Reliable measurements of brain atrophy in individual patients with multiple sclerosis. *Brain Behav*. 2016;6(9):e00518.

6. Burton EJ, McKeith IG, Burn DJ, Williams ED, O'Brien JT. Cerebral atrophy in Parkinson's disease with and without dementia: a comparison with Alzheimer's disease, dementia with Lewy bodies and controls. *Brain*. 2004;127(Pt 4):791-800.

7. Schmitter D, Roche A, Maréchal B, et al. An evaluation of volume-based morphometry for prediction of mild cognitive impairment and Alzheimer's disease. *Neuroimage Clin*. 2015;7:7-17.

8. Mazón M, Vázquez Costa JF, Ten-Esteve A, Martí-Bonmatí L. Imaging biomarkers for the diagnosis and prognosis of neurodegenerative diseases. The example of amyotrophic lateral sclerosis. *Front Neurosci*. 2018;12:784.

9. Blazhenets G, Ma Y, Sörensen A, et al. Principal components analysis of brain metabolism predicts development of Alzheimer dementia. *J Nucl Med*. 2019;60(6):837-843.

10. Sharma A, Kaur S, Memon N, Jainul Fathima A, Ray S, Bhatt MW. Alzheimer's patients detection using support vector machine (SVM) with quantitative analysis. *Neurosci Inform*. 2021;1(3):100012.

11. Folle L, Vesal S, Ravikumar N, Maier A. Dilated deeply supervised networks for hippocampus segmentation in MRI. *Bildverarbeitung für die Medizin*. Informatik aktuell. Wiesbaden: Springer Vieweg; 2019:68-73.

12. Sood T, Khandnor P. Classification of Parkinson's disease using various machine learning techniques. *Advances in Computing and Data Sciences*. Singapore: Springer; 2019:296-311.

13. Valentin Jospin L, Buntine W, Boussaid F, Laga H, Bennamoun M. Hands-on Bayesian Neural Networks - A Tutorial for Deep Learning Users. *IEEE Computational Intelligence Magazine*. 2022;17(2):29-48.

14. Marschallinger R, Mühlau M, Schmidt P, et al. Using classical geostatistics to quantify the spatiotemporal dynamics of a neurodegenerative disease from brain MRI. In: Gómez-Hernández JJ, Rodrigo-Ilarri J, Rodrigo-Clavero ME, Cassiraga E, Vargas-Guzmán JA, eds. *Geostatistics Valencia 2016*. Cham: Springer International Publishing; 2017:967-974.

15. Illan IA, Górriz JM, Ramírez J, Meyer-Base A. Spatial component analysis of MRI data for Alzheimer's disease diagnosis: a Bayesian network approach. *Front Comput Neurosci*. 2014;8:156.

16. Taschler B, Ge T, Bendfeldt K, Müller-Lenke N, Johnson TD, Nichols TE. Spatial modeling of multiple sclerosis for disease subtype prediction. *Med Image Comput Comput Assist Interv*. 2014;17(Pt 2):797-804.

17. Meyer H, Reudenbach C, Hengl T, Katurji M, Nauss T. Improving performance of spatio-temporal machine learning models using forward feature selection and target-oriented validation. *Envir Modell Softw*. 2018;101:1-9.

18. Mei S, Ji J, Bi Q, Hou J, Du Q, Li W. Integrating spectral and spatial information into deep convolutional neural networks for hyperspectral classification. Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS); 2016:5067-5070.

19. Lesjak Ž, Galimzianova A, Koren A, et al. A novel public MR image dataset of multiple sclerosis patients with lesion segmentations based on multi-rater consensus. *Neuroinformatics*. 2018;16(1):51-63.

20. Tustison NJ, Avants BB, Cook PA, et al. N4ITK: improved N3 bias correction. *IEEE Trans Med Imaging*. 2010;29(6):1310-1320.

21. Avants BB, Epstein CL, Grossman M, Gee JC. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Med Image Anal*. 2008;12(1):26-41.

22. Avants BB, Tustison NJ, Stauffer M, Song G, Wu B, Gee JC. The insight toolkit image registration framework. *Front Neuroinform*. 2014;8:44.

23. Fortin JP, Sweeney EM, Muschelli J, Crainiceanu CM, Shinohara RT. Removing inter-subject technical variability in magnetic resonance imaging studies. *Neuroimage*. 2016;132:198-212.

24. Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans Med Imaging*. 2001;20(1):45-57.

25. Fortin JP, Cullen N, Sheline YI, et al. Harmonization of cortical thickness measurements across scanners and sites. *Neuroimage*. 2018;167:104-120.

26. Mullachery V, Khera A, Husain A. *Bayesian Neural Networks*. *CoRR*. 2018.

27. Gitler AD, Dhillon P, Shorter J. Neurodegenerative disease: models, mechanisms, and a new hope. *Dis Model Mech*. 2017;10(5):499-502.

28. Ker J, Singh SP, Bai Y, Rao J, Lim T, Wang L. Image thresholding improves 3-dimensional convolutional neural network diagnosis of different acute brain hemorrhages on computed tomography scans. *Sensors (Basel)*. 2019;19(9):2167.

29. Payan A, Montana G. Predicting Alzheimer's disease a neuroimaging study with 3D convolutional neural networks. Proceedings of the ICPRAM 2015 - 4th International Conference on Pattern Recognition Applications and Methods; 2015.

30. Mzoughi H, Njeh I, Wali A, et al. Deep multi-scale 3D convolutional neural network (CNN) for MRI Gliomas brain tumor classification. *J Digit Imaging*. 2020;33(4):903-915.

31. Parmar H, Nutter B, Long R, Antani S, Mitra S. Spatiotemporal feature extraction and classification of Alzheimer's disease using deep learning 3D-CNN for fMRI data. *J Med Imaging*. 2020;7(05):056001.

32. Kingma DP, Welling M. Auto-encoding variational Bayes. Proceedings of the 2nd International Conference on Learning Representations, ICLR 2014 - Conference Track; 2014.

33. Vladimirova M, Verbeek J, Mesejo P, Arbel J. Understanding priors in bayesian neural networks at the unit level. Proceedings of the 36th International Conference on Machine Learning; 2019:6458-6467; PMLR.

34. Malinin A, Gales M. Predictive Uncertainty Estimation via Prior Networks. 32nd Conference on Neural Information Processing Systems (NeurIPS 2018). *Advances in Neural Information Processing Systems 31*. Montreal, Canada: MIT Press; 2018.

35. Neal RM. Priors for infinite networks. *Bayesian Learning for Neural Networks*. New York: Springer; 1996:29-53.

36. Hoffman MD, Blei DM, Wang C, Paisley J. Stochastic variational inference. *J Mach Learn Res*. 2013;14(1):1303-1347.

37. Pesini P, Pérez-Grijalba V, Monleón I, et al. Reliable measurements of the β-amyloid pool in blood could help in the early diagnosis of AD; 2012:604141.