Computers in Biology and Medicine 185 (2025) 109518

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

MRI classification of progressive supranuclear palsy, Parkinson disease and controls using deep learning and machine learning algorithms for the identification of regions and tracts of interest as potential biomarkers

Heiko Volkmann [a], Günter U. Höglinger [b,c,d], Georg Grön [e], Lavinia A. Bârlescu [a], the DESCRIBE-PSP study group, Hans-Peter Müller [a,*,1], Jan Kassubek [a,1]

[a] Department of Neurology, University of Ulm, Ulm, Germany
[b] Department of Neurology, LMU University Hospital, Ludwig-Maximilians-Universität (LMU), Munich, Germany
[c] German Center for Neurodegenerative Diseases (DZNE), Site, Munich, Germany
[d] Munich Cluster for Systems Neurology (SyNergy), Munich, Germany
[e] Section for Neuropsychology and Functional Imaging, Dept. of Psychiatry III, University of Ulm, Germany

ARTICLE INFO

Keywords:
Progressive supranuclear palsy
Machine learning
Deep learning
Diffusion tensor imaging (DTI)
Magnetic resonance imaging (MRI)
Neuropathology
tau protein

ABSTRACT

Background: Quantitative magnetic resonance imaging (MRI) analysis has shown promise in differentiating neurodegenerative Parkinsonian syndromes and has significantly advanced our understanding of diseases like progressive supranuclear palsy (PSP) in recent years.
Objective: The aim of this study was to develop, implement and compare MRI analysis algorithms based on artificial intelligence (AI) that can differentiate PSP not only from healthy controls but also from Parkinson disease (PD), by analyzing changes in brain structure and microstructure. Specifically, this study focused on identifying regions of interest (ROIs) and tracts of interest (TOIs) that are crucial for the algorithms to provide clinically relevant performance indices for the distinction between disease variants.
Methods: MR data comprised diffusion tensor imaging (DTI – tractwise fractional anisotropy statistics (TFAS)) and T1-weighted (T1-w) data (texture analysis of the corpus callosum (CC)). One subject sample with 74 PSP patients and 63 controls was recorded at 3.0T at multiple sites. The other sample came from a single site, consisting of 66 PSP patients, 66 PD patients, and 44 controls, recorded at 1.5T. Four different machine learning algorithms (ML) and a deep learning (DL) neural network approach using Tensor Flow were implemented for the study. The training of the algorithms was performed on 80 % of the data, which included the entire single-site data and parts of the multiple-site data. The validation process was conducted on the remaining data, thereby consistently separating training and validation data.
Results: A random forest algorithm and a DL neural network classified PSP and healthy controls with accuracies of 92 % and 95 %, respectively. Particularly, DTI derived measures for the pons, midbrain tegmentum, superior cerebral peduncle, putamen, and CC contributed to high accuracies. Furthermore, DL neural network classification of PSP and PD with 86 % accuracy showed the importance of 19 structures. The four most important features were DTI derived measures for prefrontal white matter, the fasciculus frontooccipitalis, the midbrain tegmentum, and the CC area II. This DL network achieved a sensitivity of 88 % and specificity of 85 %, resulting in a Youden-index of 0.72.
Conclusion: The primary goal of the present study was to compare multiple ML-methods and a DL approach to identify the least necessary set of brain structures to classify PSP vs. controls and PSP vs. PD by ranking them in a hierarchical order of importance. That way, this study demonstrated the potential of AI approaches to MRI as possible diagnostic and scientific tools to differentiate variants of neurodegenerative Parkinsonism.

* Corresponding author. Dept. of Neurology, University of Ulm, Oberer Eselsberg 45, 89081, Ulm, Germany.
E-mail addresses: heiko.volkmann@uni-ulm.de (H. Volkmann), guenter.hoeglinger@med.uni-muenchen.de (G.U. Höglinger), georg.groen@uni-ulm.de (G. Grön), lavinia.barlescu@uni-ulm.de (L.A. Bârlescu), hans-peter.mueller@uni-ulm.de (H.-P. Müller), jan.kassubek@uni-ulm.de (J. Kassubek).
[1] shared senior authorship.

https://doi.org/10.1016/j.compbiomed.2024.109518
Received 6 August 2024; Received in revised form 2 December 2024; Accepted 2 December 2024
Available online 10 December 2024

1. Introduction

Progressive supranuclear palsy (PSP) is a 4R–tauopathy with pathology in neurons, astrocytes, and oligodendroglia [1] associated with different clinical phenotypes such as PSP-Richardson's syndrome (PSP-RS) and PSP–parkinsonism (PSP-P) [2,3]. To capture the multifaceted phenotypical presentations of PSP, four functional domains (ocular motor dysfunction, postural instability, akinesia, and cognitive dysfunction) have been identified as clinical predictors of PSP in the current MDS diagnostic criteria according to [2,4]). Still, PSP patients might not be diagnosed or might be misdiagnosed with Parkinson's disease (PD) especially in the early stages after the first symptoms have been recognized [5]. Therefore, there is the need to use technical approaches beyond clinical symptoms to identify early PSP. In recent years, neuroimaging has significantly improved our understanding and diagnosis of PSP [6].

Quantitative MRI analysis has been shown to be useful in differentiating atypical Parkinsonian syndromes from PD [7]. Several diagnostic methods and potential biomarkers using MRI have been described in the literature, including frontal brain atrophy, alterations of the frontal part of the corpus callosum (CC), and changes in the midbrain–pons area (e.g., superior cerebellar peduncle) as possible biomarkers for PSP–RS [8–12]. Methods such as MRPI (Magnetic resonance parkinsonism index), MRPI 2.0 [13], and the mesencephalon/pons ratio [14] have achieved promising results. Possible correlations have already been reported between the level of glial cell line-derived neurotrophic factor (GDNF) and neuroimaging parameters. GDNF in the CSF was negatively correlated with the midbrain-to-pons (M/P) ratio, and positively correlated with MRPI and MRPI 2.0 in PSP-RS [15]. Brain charts for the human lifespan have been recently proposed to build dynamic models of brain anatomy in normal aging and various neurological conditions. Planche and colleagues observed six major consecutive stages of atrophy progression in PSP-RS matching the neuropathological staging of tauopathy progression [16].

From a computational perspective, artificial intelligence (AI)-based approaches [17] have been reported to enable earlier detection of PSP patients. Especially a current Deep Learning (DL) approach [18] combining an automated measurement and segmentation of specific brain regions has been suggested as a tool for the early diagnosis of PSP.

In addition to those previous works, the primary goal of the present study was to compare multiple machine learning (ML) algorithms and a DL neural network approach to identify the least necessary set of brain structures to classify PSP vs. controls, and PSP vs. PD. Thereby pathoanatomically informing brain structures were ranked in a hierarchical order of importance needed to provide clinically relevant performance indices (accuracy, sensitivity, specificity, and Youden index). Importance of a parameter was defined relatively. This means that an "important" parameter can classify many data sets correctly and a parameter that is less important can only classify a few data sets correctly. Different parameters from T1-weighted (T1-w) and DTI data were derived from predefined brain structures.

2. Methods

2.1. Subjects and patient characteristics

This study consisted of two cohorts of PSP patients at different clinical stages who met the MDS diagnostic criteria for PSP [2] and the Multiple Allocations Extinction (MAX) rules [2,19] which were applied to all PSP subjects in both cohorts. The MDS-PSP criteria have been established in 2017 by the Movement Disorder Society-endorsed PSP Study Group as the international standard for the operationalization of clinical features for the diagnosis of PSP, stratified by predominance type and diagnostic certainty [2]. Since 2019, the application of the MAX rules - also established by the Movement Disorder Society-endorsed PSP Study Group - has been used as an amendment to the MDS diagnostic criteria to reduce the number of patients with multiple diagnostic allocations, i.e., to simplify and standardize the use of the MDS-PSP criteria for both research and clinical care [19].

Cohort A included initially 78 PSP patients (mean age 70.0 ± 7.4 years, male/female = 38/40); of those, 21 PSP patients underwent a follow–up scan after 17.5 months on average (range 11.3–34.8 months), and 63 controls (mean age 68.6 ± 7.7 years, male/female = 30/33, with 17 controls undergoing a follow–up scan after 17.0 months on average (range 10.0–37.4 months). Data were multi-centrically recorded with 3.0T MRI, as part of the DESCRIBE and DANCER studies initiated by the DZNE (German Center for Neurodegenerative Diseases) across 9 sites. Cohort B comprised a single-site sample of 66 PSP patients (mean age 70.5 ± 9.1 years, male/female = 38/28) 66 PD (mean age 71 ± 10 years, male/female = 41/25) and 44 controls (mean age 68.5 ± 5.3 years, male/female = 25/19), recorded at 1.5T at the Department of Neurology, University of Ulm, Germany [8]. These patients were characterized with respect to disease severity (PSP Rating Scale (PSPRS), [20]), clinical stage (according to the staging system by [21], and phenotype according to the MDS Diagnostic Criteria of PSP and the MAX rules [2,19]. In cohort A, 78 patients with a diagnosis of PSP (PSP-RS: 55; PSP-P: 23) were initially included. Cohort B included 66 PSP patients, with 46 of the phenotype PSP–RS and 20 of the phenotype PSP–P. In both cohorts, patients had the diagnosis of probable PSP. Only PSP patients who underwent MRI scans including MPRAGE and DTI without relevant artefacts and imaging abnormalities compromising the accurate assessment of the scans (e.g., extended vascular lesions) were considered for the study. All controls were well-matched volunteers without a history of neurological or psychiatric disease or other relevant medical conditions. Table 1 summarizes the demographic and clinical data of the participants.

This experimental study was conducted in compliance with the declaration of Helsinki. All subjects provided written informed consent according to institutional guidelines approved by the DZNE (for Cohort A: "Klinische Register–Studie neurodegenerativer Erkrankungen (DESCRIBE)" and "Vertiefte Phänotypisierung der Progressiven Supranukleären Parese (DESCRIBE–PSP)", No. 311/14), and the Ethics Committee of Ulm University, Germany (for Cohort B and PD, No. 279/19, No. 284/22).

2.2. Scan protocols

MRI scanning was performed using two different protocols in the two cohorts. The MRI protocol for Cohort A, as part of the multicentric DZNE studies DESCRIBE and DANCER, used the following sequences: high resolution T1-w scans consisting of 192 sagittal slices with a resolution of 256 × 256 pixels and a slice thickness of 1.0 mm and an in-plane pixel size of 1.0 mm × 1.0 mm. Echo time (TE) was 4.3 ms and repetition time (TR) was 2500 ms. DTI included 70 gradient directions (GD) with a b-value of 1000 s/mm^2, and including 10 acquisitions with a b-value of b = 0 s/mm^2. Each DTI volume consisted of 72 slices, a resolution of 2.0x2.0x2.0 mm^3 (matrix 120x120x72), TE was 88 ms, and TR was 12100 ms. All participants of cohort A had the identical acquisition protocol; for details of the center distribution please refer to **Supplementary Information I**. The protocol for Cohort B at 1.5 T (Magnetom Symphony; Siemens Medical, Erlangen, Germany) used the following sequences: high resolution T1-w scans consisted of 144 sagittal slices with a resolution of 256 × 256 pixels and a slice thickness of 1.2 mm and an in-plane pixel size of 1.0 mm × 1.0 mm; TE was 4.2 ms and TR was 1640 ms. DTI used 52 GD with a b-value of 1000 s/mm^2, and including 4 acquisitions with a b-value of 0 s/mm^2). 64 slices were acquired with resolution of 2.0x2.0x2.8 mm^3 (matrix 120x120x64), TE was 85 ms, and TR was 7600 ms.

2.3. Data analysis

Pre– and post–processing (Fig. 1) of the data was performed using

Table 1
Subjects′ characteristics. PSP – progressive supranuclear palsy, PSP-P – progressive supranuclear palsy with predominant parkinsonism, PSP-RS – progressive supranuclear palsy with Richardson's syndrome, PSPRS – Progressive supranuclear palsy rating scale, PD – Parkinson disease. (*a) [20]; (*b) [21]; (*c) [2]. Values of continuous variables are given in mean ± standard deviation and range.

		N (m/f) (baseline)	N (m/f) (follow-up)	Time to follow-up/ months	Age/years (baseline)	Age/years (follow-up)	Disease duration/ years (baseline)	Disease duration/ years (follow-up)
Cohort A (3.0T)	PSP-RS patients	55 (24/31)	13 (5/8)	17.6 ± 8.1 (10.4–34.3)	71.0 ± 7.4 (50.6–86.2)	72.3 ± 7.7 (59.4–81.4)	4.3 ± 2.8 (0.6–14.4)	5.2 ± 2.2 (2.4–9.8)
	PSP-P patients	23 (14/9)	8 (5/3)	16.0 ± 5.8 (11.2–25.4)	68.2 ± 7.6 (52.0–83.9)	70.0 ± 5.9 (62.8–77.8)	3.8 ± 1.8 (0.5–6.9)	5.0 ± 1.6 (2.6–7.2)
	p (*t*-test, PSP-RS vs. PSP-P)	–	–	–	n.s.	n.s.	n.s.	n.s.
	all PSP patients	78 (38/40)	21 (10/11)	17.8 ± 7.2 (11.3–34.8)	70.0 ± 7.4 (50.6–86.2)	70.8 ± 6.5 (59.4–81.4)	4.2 ± 2.7 (0.5–14.4)	5.2 ± 2.7 (2.4–9.8)
	Controls	63 (30/33)	17 (6/11)	17.0 ± 7.5 (10.0–37.4)	68.6 ± 7.7 (51.2–89.4)	71.1 ± 8.7 (53.1–82.2)	–	–
	p (*t*-test, controls vs PSP)				n.s.	n.s.		
Cohort B (1.5T)	PSP-RS patients	46 (24/22)	–	–	70.0 ± 9.2 (49.0–84.3)	–	2.9 ± 1.6 (0.5–6.9)	–
	PSP-P patients	20 (14/6)	–	–	70.9 ± 8.9 (50.2–91.3)	–	3.4 ± 1.6 (1.1–8.7)	–
	p (*t*-test, PSP-RS vs. PSP-P)	–			n.s.		n.s.	
	all PSP patients	66 (38/28)	–	–	70.5 ± 9.1 (49.0–91.3)	–	3.1 ± 1.6 (0.5–8.7)	–
	Controls	44 (25/19)	–	–	68.5 ± 5.3 (57.2–81.9)	–	–	–
	PD patients	66 (41/25)	–	–	70.4 ± 10.4 (52.0–93.5)	–	3.6 ± 2.6 (0.8–9.7)	–
	p (controls vs PSP vs PD)	–			n.s. (Oone-way ANOVA)		n.s. (*t*-test)	

		PSPRS (*a) (baseline)	**PSPRS (*a) (follow-up)**	**Golbe (*b) stage (baseline)**	**Golbe (*b) stage (follow-up)**	**PSP-RS/PSP-P (*c) (baseline)**	**PSP-RS/PSP-P (*c) (follow-up)**
Cohort A (3.0T)	PSP patients	36 ± 8 (21–52)	43 ± 17 (14–70)	2 ± 2 (1–4)	2 ± 1 (1–3)	55/23	13/8
Cohort B (1.5T)	PSP patients	35 ± 11 (15–61)	–	2 ± 1 (1–4)	–	46/20	–

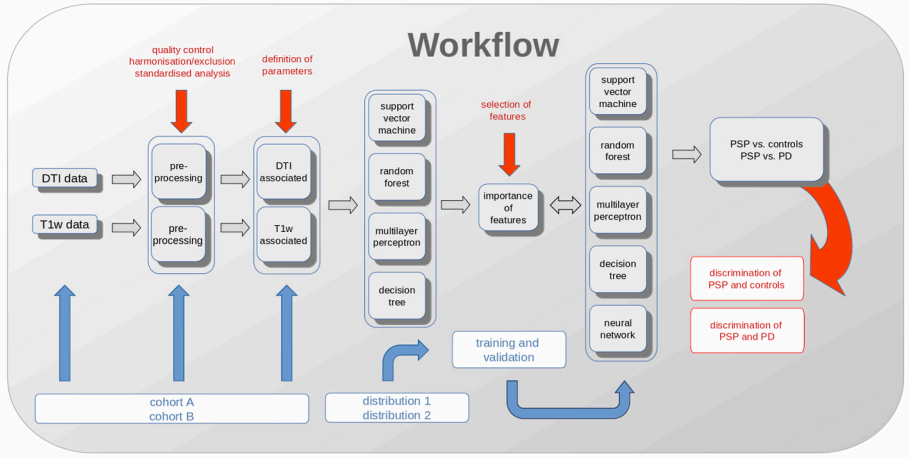

Fig. 1. Workflow of the categorization cascade. Initially, DTI and T1-weighted data underwent preprocessing, harmonization (or exclusion of data), and analysis as detailed in section 2.3. After that, parameters ROIs and TOIs were defined as described in section 2.3.3. Diverse training and validation datasets were then curated (3.1.1), initially engaging a random forest approach which subsequently led to the prioritization of specific parameters features for in-depth examination, as outlined in sections 3.1. These selected parameters were dynamically adjusted throughout training and validation, employing different AI algorithms.

the *Tensor Imaging and Fiber Tracking* software (TIFT; [22]).

2.3.1. *DTI data – fractional anisotropy*

All DTI data were assessed for completeness and, following an established quality control protocol [23], DTI data with corrupted GD or relevant motion artefacts were excluded before correcting for eddy current induced geometric distortions. DTI data could not be corrected for susceptibility artefacts because acquisition was performed only in A-P direction. However, the susceptibility-induced field will (to first approximation) be constant for all the acquired images, which means that the set of images will be internally consistent [24]. DTI data were then transferred to a 1 mm iso-grid for subsequent analyses [25] – all referenced voxel sizes refer to 1 mm^3 voxels. Non–linear spatial normalization to the Montreal Neurological Institute (MNI) stereotaxic standard space [26] was performed using study–specific templates (b0 template as well as FA template) to ensure the preservation of directional information [22,25]. This process applied to both baseline and follow–up data, incorporating a prior intra–subject alignment [27].

Fractional anisotropy (FA) maps were generated from the MNI–normalized DTI data, and a Gaussian smoothing filter of 8 mm full width at half maximum [28] was applied to these individual FA maps to optimize the balance between sensitivity and specificity. The final FA maps were corrected for age [29] but not for sex, since an effect of sex on FA has not yet been reported for PD and PSP patients – a detailed analysis showing the absence of a sex effect in the present data is provided in **Supplementary Information V**. Data from protocol A were tested for harmony and small center contributions (independent of the controls/patient ratio) were eliminated from the study. For details of the handling of center-effects please refer to **Supplementary Information II**. It is generally well known that especially DTI metrics could be influenced by the scanner protocol and by the field strength (although they are physical metrics with the intrinsic property of being independent of the measurement device – despite device specific aspects). The residual contribution to variation, in e.g. FA, is the acquisition voxel size. In this study, we assumed the voxel size differences (2.0x2.0x2.0 mm^3 in protocol A, 2.0x2.0x2.8 mm^3 in protocol B) to be negligible compared to disease-specific effects. Nevertheless, a protocol harmonization of FA maps according to the acquisition protocols of cohorts A and B was performed according to a previously published algorithm [30]. Note: PSP–RS and PSP–P generally showed similar sequential disease progression (see disease duration in Table 1) so that, for statistical sample size considerations, PSP–RS and PSP–P were analyzed together.

2.3.2. Texture analysis

Details of the T1-w texture analysis cascade have been described previously [8,31]. In short, T1-w data were assessed for completeness and motion artefacts and, after isometric and affine alignment to the anterior commissure/posterior commissure line and adjustment of the intensity threshold to automatically segment the CC, a subdivision of the CC into areas I–V according to the Hofer and Frahm scheme [32] was performed. Finally, calculation of area sizes and texture parameters [33] was applied. In the current study, the parameters entropy and homogeneity were analyzed.

2.3.3. Region of Interest (ROI) and Tract of Interest (TOI) analysis

Region of interest (ROI) analysis was performed by arithmetically averaging FA values within a given ROI for each subject, considering only voxels with an FA value greater than 0.2 [34]. An averaged DTI data was computed from the control group's data by arithmetically averaging the MNI–transformed data. This averaged control DTI data was used to identify specific tracts and pathways through FT using a seed–to–target approach. The tracts of interest (TOIs) were defined as all potential tracts originating in the start region and terminating in the target region. A modified deterministic streamline tracking technique for FT was employed, which accounts for the directional information of adjacent tracts [35].

Multiple brain structures known to be affected in PSP were analyzed, as evidenced by imaging and neuropathological studies. Clinical MRI classically reveals midbrain atrophy [11,36], a finding supported by neuropathological studies [1,37]. In addition, reduced volumetric measurements of the putamen and globus pallidus have been reported [11,38], with the globus pallidus being one of the earliest structures affected [37]. Regarding the putamen and globus pallidus, we decided to analyze them as a single entity due to their close proximity on DTI imaging so that we found it very difficult to separate these two structures accurately in order to analyze each of them entirely and exclusively. We realize that these two structures have different PSP pathology qualitatively and quantitatively, however, we found that by analyzing them in the same ROI, the possible erroneous results would be mitigated. The pons, although less severely impacted, also shows involvement [11]. The cerebral peduncle, nigrostriatal tract, and subthalamopallidal tract which interconnect these early affected structures have been studied in recent PSP imaging research [39–41]. Additionally, the medial lemniscus was included to the study due to its course through the brainstem and observed involvement in PSP, despite its unclear role in clinical presentation [42].

Owing to anatomical connections and the results of previous DTI study findings in PSP [43], the thalamic anterior and posterior radiations were also included. Additionally, the caudate nucleus was selected which shows microscopic changes and imaging correlates of degeneration [1,11,37,44,45]. There is ample evidence for the involvement of frontal white matter (WM) in several imaging and histological studies [1,11,37,46–49], leading to our subdivision of the frontal WM into four ROIs (fronto–orbital, prefrontal, premotor, and precentral WM). Furthermore, several frontal lobe WM tracts which showed changes in DTI studies were included [45,50], such as the fronto-occipital fascicle, uncinate fascicle, and superior longitudinal fascicle. The CC, particularly its anterior part, has shown involvement in PSP in numerous studies [8,51–54], leading to the inclusion of its anterior three segments (areas I, II, and III according to the definition by Hofer and Frahm [32]). In addition, the anterior limb of the internal capsule [55,56] and the corticostriatal tract were included due to their connections with the thalamus, striatum, and frontal lobe. The cerebellar WM and the dentate nucleus, both noted for PSP involvement in imaging and histological studies [1,11,44,48,57], were also included. The superior and middle cerebellar peduncles were considered, with the former being a WM bundle important for PSP diagnosis [11,44,48,49,58]. Additionally, the CC in its different segments was examined focusing on homogeneity and entropy in T1-w data (**Supplementary Information III**).

2.3.4. Evaluation of center-effects

Before starting the classification analyses, the extent to which the PSP data of cohort A were homogeneously distributed across those five centers contributing at least four PSP cases was examined. To check for center effects, non-parametric Kruskal-Wallis analyses of variance were calculated, with center as a five-fold independent factor. The dependent variables were the FA values averaged for the predefined ROIs. Significant effects of the factor *center* were found for two of the 49 ROIs: midbrain tegmentum (test statistic (h) = 19.0119; p-value = 0.0008) and substantia nigra (h = 18.6678; p-value = 0.0009). Pairwise comparisons between centers using Dunn's tests confirmed that one center in particular, which contributed four PSP cases to the study, was responsible for the abnormalities (Supplementary Information, Tables II–1). These four MRI data were therefore excluded from the study since the small sample was not suitable for numerical harmonization. The final sample size of PSP patients of cohort A was therefore 74.

3. Theory and calculation

3.1. Application of ML algorithms

The application of ML algorithms consisted of two steps. The initial first step involved the training of well-established and rather robust AI methods, such as support vector machines (SVM) [59], multilayer perceptrons (MLP) [60], decision trees [61], and random forests [62] (Fig. 1). The aim was twofold: first, to identify the most effective ML-method, and to determine the optimal set of parameters for the second step of the ML application using Gini importance information [62]. This process was integral in synthesizing the accumulated insights from the initial first step, specifically the selected classification methodology and the refined parameter list. With this prioritized parameter list, sensitivity and specificity of the ML-classification algorithms were optimized. Only after achieving results with highest accuracy by one of these ML algorithms, a DL approach was implemented which specifically aimed at further improving accuracy and validating pre–selected parameters. Not only the parameter list but also the information about the structures of the associated classification method was used to build up different DL approaches aiming at the highest accuracy and Youden-index (sensitivity + specificity −1).

3.1.1. Data partitioning into training and validation

For the application of the AI algorithms, all data were randomly split into 80 % for training and 20 % for validation (Fig. 2). Two different distributions were set up. In distribution 1 (D1), the training dataset comprised 197 participants' data sets consisting of all PSP patients' and controls' data from cohort B (N = 110), and additionally 45 randomly selected PSP patients and 42 randomly selected controls from cohort A (N = 87). Validation was performed on a set of 50 participants' data consisting of further randomly selected 21 controls and 29 PSP patients from cohort A. Distribution 2 (D2) consisted again of all data from cohort B and in addition another randomly selected 23 PSP patients and 41 controls from cohort A. Thus, in D2, 174 participants' data were used for training, and 44 participants' data from cohort A (22 PSP and 22 controls) were used for validation. There was no further preselection of participants' data for training or validation. Follow-up scans were pseudo-randomly allocated to either validation or training groups (see Section 3.2). For the classification of PSP vs PD, only 1.5 T MRI data from cohort B, consisting of 66 PSP and 66 PD patients, respectively, could be used (no PD data in cohort A).

3.1.2. SVM

An SVM [59] was included in the study because of a good performance with larger data samples and the ability to handle non-linear data through kernel transformation [63]. The SVM was trained and validated on all dataset distributions, using scikit-learn library [64] with the following configuration: kernel = 'rbf', C = 1.0, degree = 3, gamma = 'scale', coeff0 = 0.0, shrinking = true, probability = false, tol = 0.001, cache_size = 200 MB, class_weight = none, verbose = false, max_iter = −1, decision_function_shape = 'ovr'.

3.1.3. Decision Trees

This method was chosen because of its simple visualization, understandability, and the well-controllable adjustments of parameters [61]. The scikit-learn library was used [64], and the decision tree was trained and validated on all dataset distributions with the following configuration: criterion set to 'gini', splitter strategy as 'best', no limit on maximum depth, a minimum of two samples required to split a node, at least one sample required at each leaf node. The tree had no maximum leaf node restriction, no minimum impurity decrease was required for splits, and no class weights or complexity pruning was applied.

3.1.4. Random forest models

The random forest model was included in the study because of its straightforward implementation, its versatility in being applied to various data structures, its ability to output results in probabilities, and its capability to perform variable feature importance measurement [62].

The random forest model was configured with the following parameters using the scikit-learn library [64]: The number of trees was set to 100. No limit for the maximum depth of each tree was set, minimal samples split = 2, or until all leaves were pure. Minimal samples leaf = 1. Max features was set to 'auto', which means that the square root of the total number of features was used. Bootstrap sampling was used to build the trees.

The Gini importance coefficient was chosen as the criterion for measuring the quality of a split, favoring the selection of the most discriminating features at each node. These parameters features were carefully selected to optimize the performance of the random forest model in our analysis.

3.1.5. Multilayer perceptron (MLP)

MLPs are neural networks for a large range of applications, using the backpropagation algorithm for training [60]. This method was included in the study because of its strengths in several key areas: it can be applied to complex non-linear problems, works well on large datasets, and provides fast predictions after training [65].

In this study, MLP experiments were carried out on Distributions D1 and D2 using the scikit-learn library [64]. The standard configuration for the MLPClassifier included a single hidden layer with 100 neurons, activation function = 'relu'. The 'adam' solver was used for weight optimization, batch sizes = less than 200 samples or the total number of samples. Learning rate = 0.001, iteration number = 200. Regularization term (alpha) = 0.0001. The optimization would stop if the improvement were less than a tolerance of 0.0001. No specific random state was set.

3.1.6. DL neural networks

This classification method is built up from 'neurons' which in that case are mathematical functions often similar to XOR or OR functions and have a weight that can be interpreted as a threshold potential to be crossed by the input values [66]. The DL neural network enables complex, multidimensional classification. However, it was applied only in the second step of the study, to first identify important features using conventional ML techniques. Specifically, various configurations of DL neural networks were applied using TensorFlow [67], optimized on accuracy (Youden-index) (Table 2, Fig. 3).

The method of permutation importance [68] was used to evaluate the neural network designed to differentiate between PD and controls and for the neural network implemented to differentiate between PSP and PD. This technique involves systematically changing the order of the features (in the dataset) and observing the resulting effect on the accuracy of the network. In detail, each feature is randomly shuffled while the other features are kept constant, and the change in the models accuracy is recorded. A significant drop in accuracy when a particular feature is shuffled indicates the high importance of that feature for the models accuracy. This approach is particularly effective in assessing the relative contribution of each feature in complex neural network architectures, providing valuable insight into the model's decision–making process and guiding further refinements in feature selection and model optimization.

3.2. Incorporation of longitudinal data

As we decided to also include longitudinal data (available in cohort

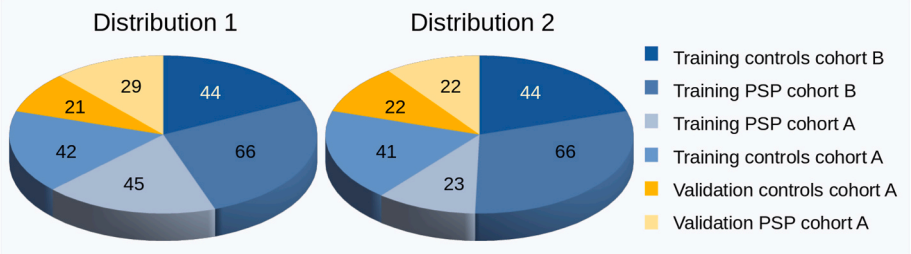

Fig. 2. Different distributions of the training and validation set. The illustrations depict the data distributions employed for the classification techniques. Distribution 1 consisted of 247 data, i.e., 197 data for training (blueish background) and 50 data for validation (brownish background). Distribution 2 included a dataset of 218 data, i.e., 174 data for training (blueish background) and 44 data for validation (brownish background).

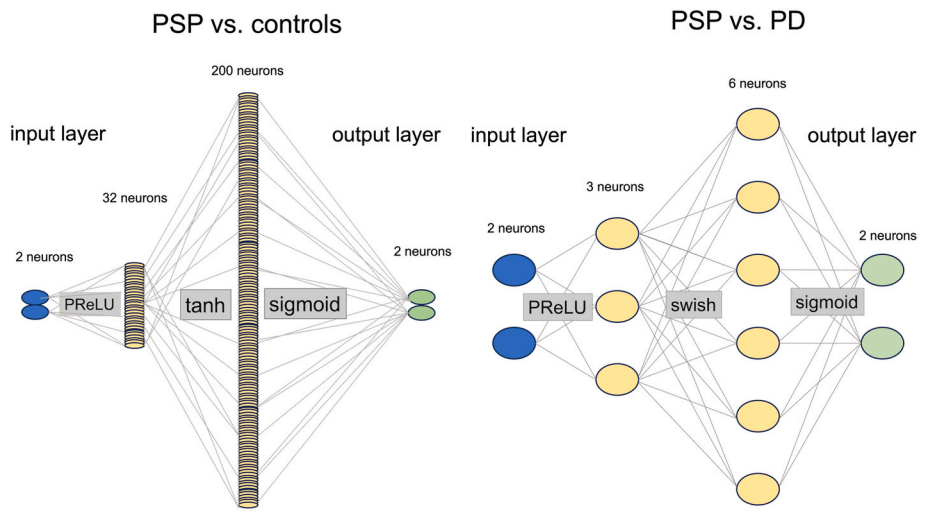

Fig. 3. Layer structure of the neural networks. This figure shows the two different layer structures of the used deep learning (DL) neural networks. **Left:** Network used for the classification of progressive supranuclear palsy (PSP) vs (healthy) controls. For the input layer, a parametric rectified linear unit (PReLU) function was used, while the output layer was implemented with a sigmoid function **Right:** the layer architecture of the network used to discriminate Parkinson disease (PD) vs PSP. The input layer was implemented with a PReLU function and the output layer calculated with a sigmoid function.

Table 2
Architecture of the implemented models. This table shows the specific build-up for each neural network used in the study. The model for the PSP vs. controls discrimination stopped at 3000 epochs. For the more complex classification to discriminate between PD vs. PSP, the number of epochs was increased by factor 10, i.e. a maximum of 30,000 epochs was permitted.

	Model for PSP vs. controls	Model for PD vs. PSP
Network Architecture	Sequential	Sequential
Layers	- Dense: Output Shape (None, 32), Parameters: 320 - Dense: Output Shape (None, 200), Parameters: 6600 - Dense: Output Shape (None, 2), Parameters: 402	- Dense: Output Shape (None, 3), Parameters: 63 - Dense: Output Shape (None, 6), Parameters: 24 - Dense: Output Shape (None, 2), Parameters: 14
Total Parameters	7322	101
Trainable Parameters	7322	101
Non-trainable Parameters	0	0
Activation Functions	- Dense Layer: PReLU - Dense_1 Layer: tanh - Dense_2 Layer: sigmoid	- Dense Layer: PReLU - Dense_1 Layer: swish - Dense_2 Layer: sigmoid
Loss Function	Binary Crossentropy	Binary Crossentropy
Optimizer	Adam	Adam
Batch Size	2	4
Number of Epochs	3000	30,000

A) into the two different distributions D1 und D2, special care was taken to ascertain different longitudinal data to be included in either distribution. The pseudorandomized allocation procedure resulted in the following selections: Twenty-seven longitudinal follow-up scans were included in D1, with 22 for training (5 healthy controls and 17 PSP), and 5 for validation (4 healthy controls and 1 PSP). D2 comprised 26 follow-up scans, 17 (8 healthy controls and 9 PSP) for training, and 9 (1 healthy control and 8 PSP) for validation.

3.3. Association of MRI-based results to clinical stages

To align the DL neural network's (PSP vs. controls) categorizations with the clinical assessments, the DL network's confidence in its classifications was stratified into five distinct levels, with each level representing a 20 % probability interval, providing a gradation of the network's certainty in its diagnostic categorization. The primary objective was to compare these probabilistically determined intervals with the (clinically determined) Golbe stages [20,21] (for details refer to **Supplementary Information IV**).

3.4. T-SNE representation (t-distributed stochastic neighbor embedding)

The t-SNE algorithm reduces the dimensionality by capturing the underlying structure and relationships between data points. The X-axis (T-SNE (X)) and Y-axis (T-SNE (Y)) depict the transformed coordinates in a two-dimensional space, where similar data points are positioned closer together, and dissimilar points are further apart [69], so that it is possible to depict multi-dimensional data on a two-dimensional coordinate system.

4. Results

4.1. Selection of features for PSP vs controls and performance of ML-algorithms

For the selection of parameters to classify PSP vs controls, a random forest method was implemented including all ROIs and TOIs of the study and was then trained and validated with distribution D1 (Fig. 4). This approach already resulted in an accuracy of 0.84 with sensitivity of 0.83, specificity of 0.86, and a Youden index of 0.69. The most important parameters according to the Gini importance were chosen for further implementations. Every change in the parameters used for the random forest led to different hierarchical outcomes considering the ROIs and TOIs. After excluding the least important parameters in every iteration, the random forest achieved its best accuracy with the following specific ROIs and TOIs: CC (T1-w -homogeneity index), pons, midbrain tegmentum, cerebral peduncles, and putamen; for the latter regions, read-out was FA.

During the parameter feature selection process the validation accuracy was similar between random forest applications using either T1-w based data of the CC, or DTI based data of the CC. However, during training the random forest algorithms including parameters based on T1-w were slightly more prone to overfitting relative to DTI data. Therefore, only DTI data were selected for use in the following computations.

A random forest trial implemented on D1 yielded the most promising outcome, finally achieving a Youden index of 0.87, accuracy of 0.94, and

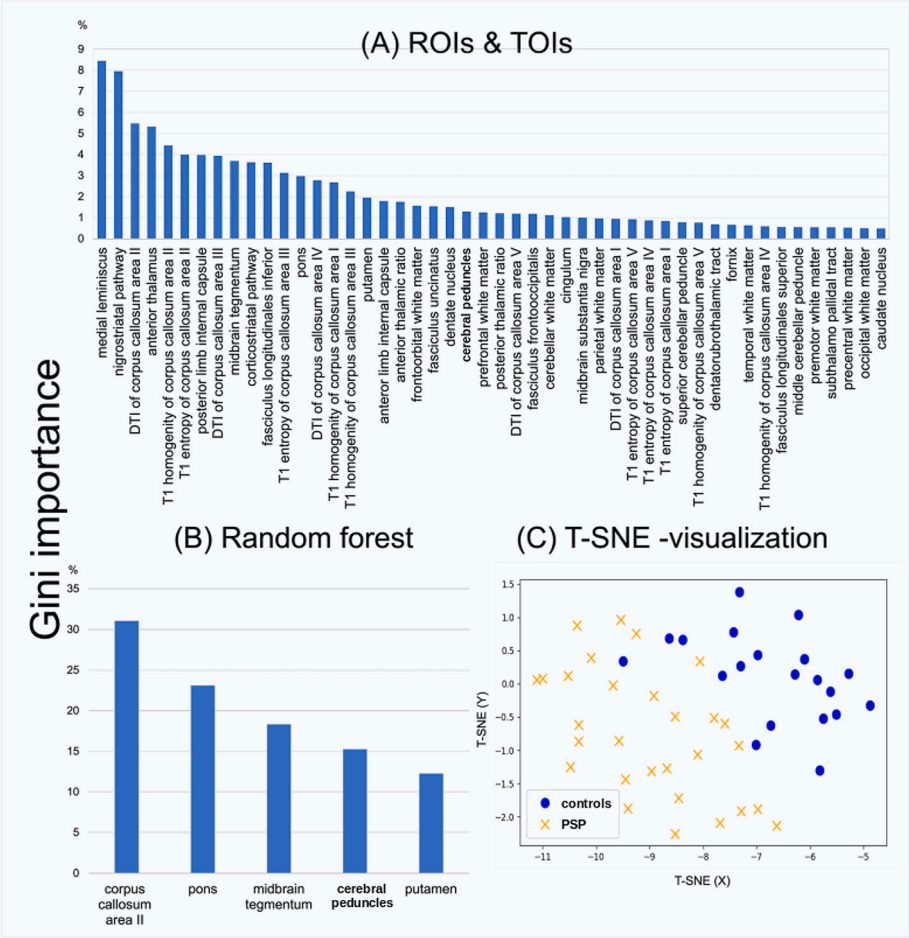

Fig. 4. Random forest approach PSP vs controls. (A): This illustration depicts the Gini importance of the random forest approach using all ROIs and TOIs of this study, with highest Gini importance value on the left and with the lowest Gini importance values on the right. **(B)** Gini importance of the parameters of the best performing random forest approach. **(C)** Exemplary T -SNE representation of the validation datasets (D1).

sensitivity and specificity of 0.97 and 0.90, respectively (Table 3A). This parameter feature selection was also used to implement a multilayer perceptron, a support vector machine and a decision tree. These results are summarized in Table 3A.

4.2. Selection of features for PSP vs PD and performance of ML-algorithms

For the selection of the features to differentiate PSP vs PD, an initial implementation of a random forest model on cohort B using all ROIs and TOIs yielded a Youden-index of 0.10. Repetitively pruning away less important features (Gini importance), this implementation yielded an hierarchically organized list of features which included: prefrontal WM, fasciculus frontooccipitalis, CC area II, midbrain tegmentum, subthalamopallidal tract, nigrostriatal pathway, caudate nucleus, medial lemniscus, precentral WM, dentatorubrothalamic tract, putamen, nucleus dentatus, anterior thalamus, cerebellar WM, premotor WM, frontoorbital WM, anterior thalamic radiation, substantia nigra, and anterior limb of internal capsule. With this set of features in use, performance of the different ML algorithms are summarized in Table 3B.

4.3. DL neural networks in PSP vs controls and PSP vs PD

For the classification of PSP vs controls, a neural network was used to further improve the accuracy of the random forest algorithm, and finally leading to an accuracy of 95 %, sensitivity of 95 %, specificity of 97 %, and a Youden-index of 92 % (Table 3A). In order to arrive at that effective classification performance, features had again to be selected iteratively using the permutation method. The final set of features included four of the five 5 selected features from the random forest approach above, and three additional features: substantia nigra, nigrostriatal pathway, medial leminiscus, midbrain tegmentum, pons, cerebral peduncles, putamen, subthalamopallidal tract. The hierarchical ordering of those features' importance is summarized in Fig. 5A.

For the differentiation of PSP vs PD, the final DL neural network model was using the selected parameters of the previous random forest model above (Table 3B). Here, the performance indices showed that PSP could be differentiated from PD with accuracy of 86 %, sensitivity of 88 %, specificity of 85 %, and a Youden-index of 0.72. The hierarchically organized importance was again determined by using the permutation method, and is summarized in Fig. 5B.

For the association of Golbe stages and AI results refer to **Supplementary Information IV.**

5. Discussion

5.1. Application of AI approaches to MRI for the diagnostic classification of PSP

In this study, different traditional ML and less traditional DL neural network approaches were evaluated for their applicability in diagnostically relevant classification of PSP versus controls and PSP versus PD, respectively, based on FA indices derived from DTI data, and callosal texture parameters from T1-w data. The classification importance of

Table 3
The best performance parameters for each of the different AI approaches for different distributions and classifications. (A) AI approaches for PSP vs controls (B) AI approaches for PSP vs. PD.

(A) PSP vs controls					
Distribution	AI approach	accuracy	sensitivity	specificity	Youden index
D1	Random forest	0.94	0.97	0.90	0.87
D2	Random forest	0.86	0.95	0.77	0.73
D1	Multilayer-perceptron	0.88	1.0	0.71	0.71
D2	Multilayer-perceptron	0.84	0.95	0.73	0.68
D1	Support vector machine	0.88	1.0	0.71	0.71
D2	Support vector machine	0.86	0.95	0.77	0.73
D1	Decision tree	0.84	0.93	0.71	0.64
D2	Decision tree	0.73	0.77	0.68	0.45
D1	Neural network	0.88	0.90	0.86	0.75
D2	Neural network	0.95	1.0	0.91	0.91
(B) PSP vs PD					
	Random forest	0.59	0.46	0.69	0.15
	Multilayer-perceptron	0.48	0.84	0.19	0.03
	Support vector machine	0.59	0.62	0.56	0.18
	Neural network	0.86	0.85	0.88	0.72

different features to obtain the best classification performance indices for each of the different algorithms applied was hierarchically defined. This study used a rather large multi-center setup, including 140 patients with PSP, 66 patients with PD, and 107 heathy controls. The differentiation of PSP versus controls was achieved with an accuracy of 92 % using a random forest model, and with 95 % accuracy using a DL neural network. The differentiation of PSP versus PD achieved an accuracy of 86 % using a DL neural network and clearly outperformed each of the different ML approaches (see Table 3). The number of relevant features for the differentiation between PSP and controls using ML algorithms, and between PSP and PD using DL neural networks, respectively, were different. While this issue did not allow for direct comparisons between ML- and DL-based categorization performances, it allowed us to identify the least necessary set of features in order to achieve between-methods comparability (ML algorithms and DL) at diagnostically relevant higher levels of accuracy, specificity, sensitivity, and the Youden-index.

Previous studies based on atlas-based volumetry [4] of T1-w data or based on automated segmentation and measurement of specific brain regions (MRPI 2.0 [13,18], have already been successful in aiding for diagnosis of PSP, or quantitatively detecting disease progression in PSP. The FA-based microstructure analysis of this study could further increase the set of features and, with it, accuracy in the diagnosis of PSP.

5.2. Comparison with previous studies

Previous studies focused on the discrimination between PSP and controls as well as on the discrimination between PSP and other Parkinsonian syndromes. Using data from atlas-based volumetry (ABV), Huppertz and colleagues (2016) found the majority of binary SVM classifications between PD, PSP-RS, and multiple system atrophy (MSA) with balanced accuracies of >80 %; SVM classifications between PD, PSP, and MSA achieved sensitivities from 79 % to 87 % [70]. classified MSA and PD against healthy controls with 95 % accuracy. More recent studies [71] differentiated between PD and atypical neurodegenerative Parkinsonism (MSA and PSP) reaching 95 % accuracy with DTI data. Chougar and colleagues (2020) reviewed various imaging approaches to discriminate between PD and PSP with high accuracies; highest accuracies (up to 100 %) were reported by combination of volumetry and DTI metrics using SVM [72].

Concerning MR imaging modalities, previous studies [73,74] showed that a combination of DTI and T1-w imaging achieved high accuracy in finding PSP-associated alterations that support classification and diagnostic accuracy. In contrast, another study [44] showed that diffusion parameters did not significantly contribute to the classification, perhaps due to scanner variability. It should be noted, however, that the present study was not intended as a systematic head-to-head comparison of both MRI modalities.

It should also be noted that while this is not the first study to address the differentiation of PSP versus controls and PD, previous studies often relied on smaller samples. In the study by [74], the authors did also not separate training and validation data completely [74], but used a leave-one-out approach which increases the risk of over-fitting [75]. Another previous study [58] performed comparisons among MRPI, MRPI 2.0, volumetric/thickness data and used traditional ML methods for the differentiation of PSP-RS and PSP-P. That study suggested that ML models using a combination of MRPI, and volumetric/thickness data achieve the best classification performance in distinguishing between these two PSP phenotypes. However, DTI data and DL-neural networks were not employed.

The present study highlights numerous brain regions, including those predominantly composed of white matter, which is specifically targeted by DTI, but also incorporated several grey matter structures. The most important structures for differentiating PSP from PD were prefrontal white matter, the fasciculus occipitalis, the CC area II (FA), and the midbrain tegmentum. This result is in agreement with the findings by [76] who reported reductions in FA within the prefrontal white matter of PD patients [76]. The outcomes from the random forest analysis revealed that particularly the grey matter regions midbrain tegmentum, pons, putamen and globus pallidus, previously identified as pathological by [1], are important for AI methods to differentiate between PSP and controls at high performance indices. Furthermore, the contribution of the CC to the differentiation of PSP versus controls was demonstrated in the present study, confirming previous findings [8].

5.3. Limitations

The first limitation of our study was the difference of imaging protocols between the cohorts. Cohort A consisted exclusively of 3.0 T MRI scans, whereas cohort B included 1.5 T scans only. To address the potential impact of these differences in image quality, we included 1.5 T images only at the training level and did not use them for validation. However, the inclusion of more 3.0 T scans might likely further improve the overall results. The AI models for differentiating between PSP and PD were trained and validated exclusively on 1.5 T data since no 3.0 T data were available for PD patients.

Even though this study used one of the largest DTI data samples for investigating classification of PSP vs controls and PD, the limitation to 140 PSP data may still cause a significant challenge for the implementation of AI methods. Nonetheless, it was ensured that the data used in the validation process were randomly selected, and one validation even used data from a different site that was entirely not involved in delivering training data.

Furthermore, longitudinal data were not excluded. While that data represented less than one third of baseline data, the ratios between healthy controls and PSP patients were rather alike, i.e., 0.85 for baseline and 0.81 for the follow up. Nevertheless, inclusion of longitudinal data might lead to overfitting in the ML methods However, all ML methods were validated across both distributions D1 and D2 and due to

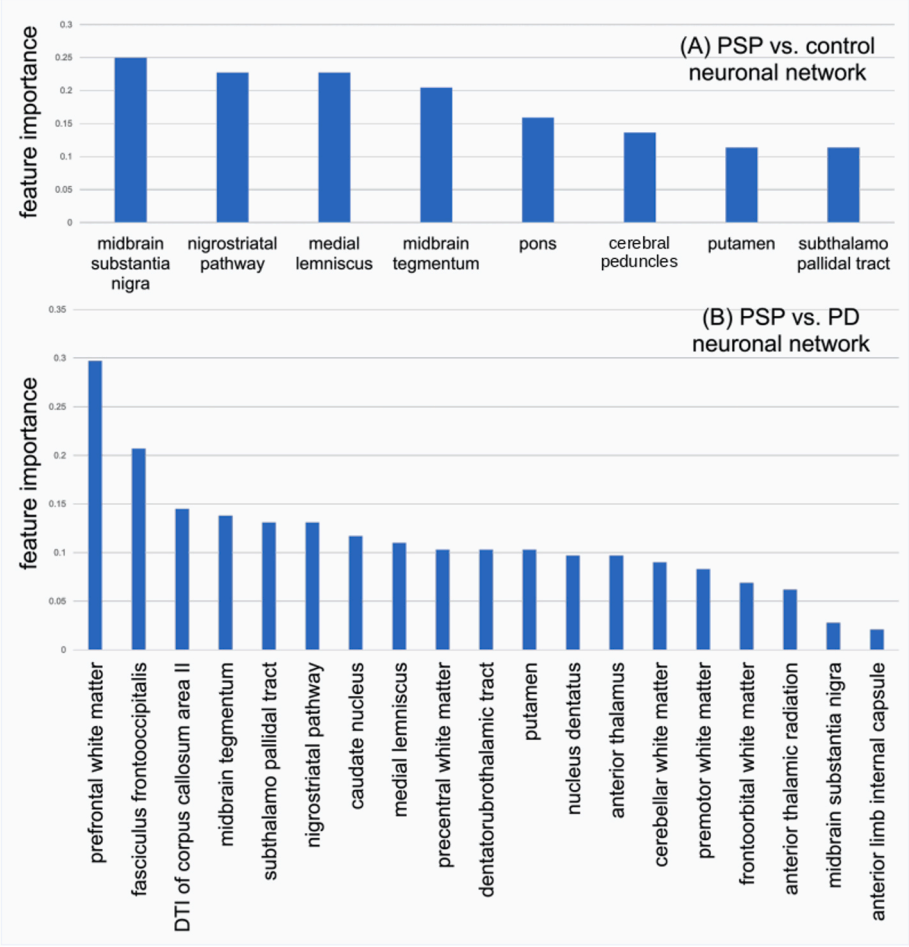

Fig. 5. Descending feature importance of the PSP vs controls neural network analysis. (A) This figure displays descending feature importance of the PSP vs controls neural network analysis with the most important parameter on the left and the least important parameter on the right. **(B)** This figure shows descending feature importance of Parameters of the PSP vs PD neural network analysis with the most important parameter on the left and the least important parameter on the right.

the random selection process, it happened, that D1 compromised just one PSP follow-up. The accuracy of the random forest algorithm of 0.94 for D1 indicated a low risk of overfitting due to follow-up scans.

Furthermore, in the current study, we specifically focused on microstructural alterations. Structural alterations in PSP, e.g. e vacuo expansion of the third ventricle [4,18] were not addressed but can in combination with volumetric techniques, principally further improve accuracy. Similarly, a recent study investigated PSP patients and controls showing interconnections among different brain areas by multiparametric MRI, providing a new perspective on the coordinated changes in structure and function in PSP pathogenesis and their relationships with the microstructural cellular structure hierarchy [77].

An important challenge in clinical routine is also to differentiate PD, MSA and PSP [70,71,73,78], but this study addressed the task to differentiate PSP and PD patients. A future task is to extend the developed methodological approach to other neurodegenerative Parkinsonian syndromes such as MSA.

Finally, as in many other current studies, the PSP diagnoses in our study were not verified post–mortem.

6. Conclusion

In conclusion, we demonstrated that AI applications to DTI data can be used to differentiate PSP from healthy controls and PD, respectively. These results highlight that the patho-anatomical complexity of PSP can be deduced by ML and DL neural network techniques. ML techniques appear already rather useful for classifying patients with PSP against controls. However, to achieve reliability at 100 % accuracy for automated diagnostics while also acknowledging computational feasibility in clinical routines, further evaluation of DTI metrics other than FA in combination with informing brain structures seems to be a worthwhile task. More complex AI methods, such as DL neural networks are apparently necessary for distinguishing PSP from PD. Further research is needed to determine the most important combinations of affected brain structures and DTI metrics, providing the least necessary set of features that would inform DL neural networks to reliably solve this challenge at 100 % accuracy.

CRediT authorship contribution statement

Heiko Volkmann: Writing – original draft, Formal analysis. **Günter U. Höglinger:** Writing – review & editing, Data curation. **Georg Grön:** Writing – review & editing, Validation, Methodology. **Lavinia A. Bârlescu:** Writing – review & editing, Formal analysis. **Hans-Peter Müller:** Writing – original draft, Software, Methodology, Formal analysis, Conceptualization. **Jan Kassubek:** Writing – review & editing, Supervision, Project administration, Data curation, Conceptualization.

Ethics statement

This experimental study was conducted in compliance with the declaration of Helsinki. All subjects provided written informed consent

according to institutional guidelines approved by the DZNE (for Cohort A: "Klinische Register–Studie neurodegenerativer Erkrankungen (DESCRIBE)" and "Vertiefte Phänotypisierung der Progressiven Supranukleären Parese (DESCRIBE–PSP)", No. 311/14), and the Ethics Committee of Ulm University, Germany (for Cohort B and PD, No. 279/19, No. 284/22).

Funding

No external funding.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Günter Höglinger has ongoing research collaborations with Roche, UCB, Abbvie; serves as a consultant for Abbvie, Alzprotect, Amylyx, Aprinoia, Asceneuron, Bayer, Bial, Biogen, Biohaven, Epidarex, Ferrer, Kyowa Kirin, Lundbeck, Novartis, Retrotope, Roche, Sanofi, Servier, Takeda, Teva, UCB; received honoraria for scientific presentations from Abbvie, Bayer, Bial, Biogen, Bristol Myers Squibb, Esteve, Kyowa Kirin, Pfizer, Roche, Teva, UCB, Zambon. holds a patent "Treatment of Synucleinopathies" United States Patent No.: US 10,918,628 B2/European Patent Patent No.: EP 17 787 904.6–1109/3 525 788; received publication royalties from Academic Press, Kohlhammer, and Thieme. Günter Höglinger was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy within the framework of the Munich Cluster for Systems Neurology (EXC 2145 SyNergy – ID 390857198); European Joint Programme on Rare Diseases (Improve-PSP); Petermax-Müller Foundation (Etiology and Therapy of Synucleinopathies and Tauopathies).

Jan Kassubek has received honoraria or consultation fees from AbbVie, Bial, Biogen, Desitin, Esteve, Licher MT, Medtronic, NeuroDerm, Novartis, STADA, UCB Pharma, and Zambon; in addition, he is Specialty Chief Editor for Frontiers in Neurology (section Applied Neuroimaging) and Associate Editor (Neurology) for Therapeutic Advances in Chronic Disease.

Acknowledgement

Data were provided by the study groups of the DESCRIBE studies of the Clinical Research of the German Center for Neurogenerative Diseases (DZNE).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2024.109518.

Appendix– Members of the *DESCRIBE-PSP* study group

Moritz Brandt (Department of Neurology, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany), Katharina Buerger (Institute for Stroke and Dementia Research, University Hospital, LMU Munich, Munich, Germany), Emrah Düzel (Institute of Cognitive Neurology and Dementia Research, Otto-von-Guericke University, Magdeburg, Germany), Björn Falkenburger (Department of Neurology, University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany), Agnes Flöel (Department of Neurology, University Medicine Greifswald, Greifswald, Germany), Wenzel Glanz (Clinic for Neurology, Medical Faculty, University Hospital Magdeburg, Magdeburg, Germany), Günter U. Höglinger (Department of Neurology, University Hospital, LMU Munich, Munich, Germany), Daniel Janowitz (Institute for Stroke and Dementia Research, University Hospital, LMU Munich, Munich, Germany), Sabrina Katzdobler (Department of Neurology, University Hospital of Munich, Ludwig-Maximilians-Universität (LMU) Munich, Munich, Germany), Ingo Kilimann (Department of Psychosomatic Medicine, Rostock University Medical Center, Rostock, Germany), Okka Kimmich (Department of Neurology, University of Bonn, Bonn, Germany), Johannes Levin (Department of Neurology, University Hospital of Munich, Ludwig-Maximilians-Universität (LMU) Munich, Munich, Germany), Oliver Peters (Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin-Institute of Psychiatry and Psychotherapy), Josef Priller (Department of Psychiatry and Psychotherapy, Charité, Berlin, Germany), Johannes Prudlo (Department of Neurology, University Medical Centre, Rostock, Germany), Luisa-Sophie Schneider (Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin-Institute of Psychiatry and Psychotherapy), Annika Spottke (Department of Neurology, University of Bonn, Bonn, Germany), Eike Jakob Spruth (Department of Psychiatry and Psychotherapy, Charité, Berlin, Germany), Matthis Synofzik (Department of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and Center of Neurology, University of Tübingen, Tübingen, Germany), Stefan Teipel (Department of Psychosomatic Medicine, Rostock University Medical Center, Rostock, Germany), Carlo Wilke (Department of Neurodegenerative Diseases, Hertie Institute for Clinical Brain Research and Center of Neurology, University of Tübingen, Tübingen, Germany).

References

[1] G.G. Kovacs, M.J. Lukic, D.J. Irwin, T. Arzberger, G. Respondek, E.B. Lee, D. Coughlin, A. Giese, M. Grossman, C. Kurz, C.T. McMillan, E. Gelpi, Y. Compta, J. C. Van Swieten, L.D. Laat, C. Troakes, S. Al-Sarraj, J.L. Robinson, S. Roeber, G. U. Höglinger, Distribution patterns of tau pathology in progressive supranuclear palsy, Acta Neuropathol. 140 (2) (2020) 99–119, https://doi.org/10.1007/s00401-020-02158-2.

[2] G.U. Höglinger, G. Respondek, M. Stamelou, C. Kurz, K.A. Josephs, A.E. Lang, B. Mollenhauer, U. Müller, C. Nilsson, J.L. Whitwell, T. Arzberger, E. Englund, E. Gelpi, A. Giese, D.J. Irwin, W.G. Meissner, A. Pantelyat, A. Rajput, J.C. Van Swieten, for the Movement Disorder Society-endorsed PSP Study Group, Clinical diagnosis of progressive supranuclear palsy: the movement disorder society criteria: MDS Clinical Diagnostic Criteria for PSP, Mov. Disord. 32 (6) (2017) 853–864, https://doi.org/10.1002/mds.26987.

[3] G. Respondek, G.U. Höglinger, The phenotypic spectrum of progressive supranuclear palsy, Park. Relat. Disord. 22 (2016) S34–S36, https://doi.org/10.1016/j.parkreldis.2015.09.041.

[4] G.U. Höglinger, J. Schöpe, M. Stamelou, J. Kassubek, T. Del Ser, A.L. Boxer, S. Wagenpfeil, H.-J. Huppertz, for the AL-108-231 Investigators, the Tauros MRI Investigators, & the Movement Disorder Society-Endorsed PSP Study Group, Longitudinal magnetic resonance imaging in progressive supranuclear palsy: a new combined score for clinical trials: PSP MRI Atrophy Score, Mov. Disord. 32 (6) (2017) 842–852, https://doi.org/10.1002/mds.26973.

[5] I.T.S. Coyle-Gilchrist, K.M. Dick, K. Patterson, P. Vázquez Rodríquez, E. Wehmann, A. Wilcox, C.J. Lansdall, K.E. Dawson, J. Wiggins, S. Mead, C. Brayne, J.B. Rowe, Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes, Neurology 86 (18) (2016) 1736–1743, https://doi.org/10.1212/WNL.0000000000002638.

[6] C. Peralta, A.P. Strafella, T. Van Eimeren, R. Ceravolo, K. Seppi, V. Kaasinen, J. E. Arena, S. Lehericy, the International Parkinson Movement Disorders Society-Neuroimaging Study Group, Pragmatic approach on neuroimaging techniques for the differential diagnosis of parkinsonisms, Movement Disorders Clinical Practice 9 (1) (2022) 6–19, https://doi.org/10.1002/mdc3.13354.

[7] J. Kassubek, MRI-based neuroimaging: atypical parkinsonisms and other movement disorders, Curr. Opin. Neurol. 31 (4) (2018) 425–430, https://doi.org/10.1097/WCO.0000000000000578.

[8] L.A. Bârlescu, H.-P. Müller, I. Uttner, A.C. Ludolph, E.H. Pinkhardt, H.J. Huppertz, J. Kassubek, Segmental alterations of the corpus callosum in progressive supranuclear palsy: a multiparametric magnetic resonance imaging study, Front. Aging Neurosci. 13 (2021) 720634, https://doi.org/10.3389/fnagi.2021.720634.

[9] M. Cosottini, R. Ceravolo, L. Faggioni, G. Lazzarotti, M.C. Michelassi, U. Bonuccelli, L. Murri, C. Bartolozzi, Assessment of midbrain atrophy in patients with progressive supranuclear palsy with routine magnetic resonance imaging, Acta Neurol. Scand. 116 (1) (2007) 37–42, https://doi.org/10.1111/j.1600-0404.2006.00767.x.

[10] L. Möller, J. Kassubek, M. Südmeyer, R. Hilker, E. Hattingen, K. Egger, F. Amtage, E.H. Pinkhardt, G. Respondek, M. Stamelou, F. Möller, A. Schnitzler, W.H. Oertel, S. Knake, H. Huppertz, G.U. Höglinger, Manual MRI morphometry in Parkinsonian syndromes, Mov. Disord. 32 (5) (2017) 778–782, https://doi.org/10.1002/mds.26921.

[11] J.L. Whitwell, G.U. Höglinger, A. Antonini, Y. Bordelon, A.L. Boxer, C. Colosimo, T. Van Eimeren, L.I. Golbe, J. Kassubek, C. Kurz, I. Litvan, A. Pantelyat, G. Rabinovici, G. Respondek, A. Rominger, J.B. Rowe, M. Stamelou, K.A. Josephs,

for the Movement Disorder Society-endorsed PSP Study Group, Radiological biomarkers for diagnosis in PSP: where are we and where do we need to be? Mov. Disord. 32 (7) (2017) 955–971, https://doi.org/10.1002/mds.27038.
[12] S. Zanigni, G. Calandra-Buonaura, D.N. Manners, C. Testa, D. Gibertoni, S. Evangelisti, L. Sambati, M. Guarino, P. De Massis, L.L. Gramegna, C. Bianchini, P. Rucci, P. Cortelli, R. Lodi, C. Tonon, Accuracy of MR markers for differentiating progressive supranuclear palsy from Parkinson's disease, Neuroimage: Clinical 11 (2016) 736–742, https://doi.org/10.1016/j.nicl.2016.05.016.
[13] S. Nigro, A. Antonini, D.E. Vaillancourt, K. Seppi, R. Ceravolo, A.P. Strafella, A. Augimeri, A. Quattrone, M. Morelli, L. Weis, E. Fiorenzato, R. Biundo, R. G. Burciu, F. Krismer, N.R. McFarland, C. Mueller, E.R. Gizewski, M. Cosottini, E. Del Prete, A. Quattrone, Automated MRI classification in progressive supranuclear palsy: a large international cohort study, Mov. Disord. 35 (6) (2020) 976–983, https://doi.org/10.1002/mds.28007.
[14] N. Madetko, P. Alster, M. Kutyłowski, B. Migda, M. Nieciecki, D. Koziorowski, L. Królicki, Is MRPI 2.0 more useful than MRPI and M/P ratio in differential diagnosis of PSP-P with other atypical parkinsonisms? J. Clin. Med. 11 (10) (2022) 2701, https://doi.org/10.3390/jcm11102701.
[15] P. Alster, D. Otto-Ślusarczyk, S. Szlufik, K. Duszyńska-Wąs, A. Drzewińska, A. Wiercińska-Drapało, M. Struga, M. Kutyłowski, A. Friedman, N. Madetko-Alster, The significance of glial cell line-derived neurotrophic factor analysis in Progressive Supranuclear Palsy, Sci. Rep. 14 (1) (2024) 2805, https://doi.org/10.1038/s41598-024-53355-y.
[16] V. Planche, B. Mansencal, J.V. Manjon, W.G. Meissner, T. Tourdias, P. Coupé, Staging of progressive supranuclear palsy-Richardson syndrome using MRI brain charts for the human lifespan, Brain Commun. 6 (2) (2024) fcae055, https://doi.org/10.1093/braincomms/fcae055.
[17] A.S. Talai, J. Sedlacik, K. Boelmans, N.D. Forkert, Utility of multi-modal MRI for differentiating of Parkinson's disease and progressive supranuclear palsy using machine learning, Front. Neurol. 12 (2021) 648548, https://doi.org/10.3389/fneur.2021.648548.
[18] S. Nigro, M. Filardi, B. Tafuri, M. Nicolardi, R. De Blasi, A. Giugno, V. Gnoni, G. Milella, D. Urso, S. Zoccolella, G. Logroscino, Frontotemporal Lobar Degeneration Neuroimaging Initiative, 4-Repeat Tau Neuroimaging Initiative, & Alzheimer's Disease Neuroimaging Initiative, Deep learning-based approach for brainstem and ventricular MR planimetry: application in patients with progressive supranuclear palsy, Radiology. Artif. Intell. 6 (3) (2024) e230151, https://doi.org/10.1148/ryai.230151.
[19] M. Grimm, G. Respondek, M. Stamelou, T. Arzberger, L. Ferguson, E. Gelpi, A. Giese, M. Grossman, D.J. Irwin, A. Pantelyat, A. Rajput, S. Roeber, J.C. Van Swieten, C. Troakes, A. Antonini, K.P. Bhatia, C. Colosimo, T. Van Eimeren, J. Kassubek, for the Movement Disorder Society-endorsed PSP Study Group, How to apply the movement disorder society criteria for diagnosis of progressive supranuclear palsy, Mov. Disord. 34 (8) (2019) 1228–1232, https://doi.org/10.1002/mds.27666.
[20] L.I. Golbe, P.A. Ohman-Strickland, A clinical rating scale for progressive supranuclear palsy, Brain 130 (6) (2007) 1552–1565, https://doi.org/10.1093/brain/awm032.
[21] L.I. Golbe, P. Ohman-Strickland, E.B. Beisser, F.T. Elghoul, A convenient prognostic tool and staging system for progressive supranuclear palsy, Movement Disorders Clinical Practice 7 (6) (2020) 664–671, https://doi.org/10.1002/mdc3.13010.
[22] H.-P. Müller, A. Unrath, A.C. Ludolph, J. Kassubek, Preservation of diffusion tensor properties during spatial normalization by use of tensor imaging and fibre tracking on a normal brain database, Phys. Med. Biol. 52 (6) (2007) N99–N109, https://doi.org/10.1088/0031-9155/52/6/N01.
[23] H.-P. Müller, J. Kassubek, G. Grön, R. Sprengelmeyer, A.C. Ludolph, S. Klöppel, N. Z. Hobbs, R.A. Roos, A. Duerr, S.J. Tabrizi, M. Orth, S.D. Süssmuth, G. Landwehrmeyer, Impact of the control for corrupted diffusion tensor imaging data in comparisons at the group level: an application in Huntington disease, Biomed. Eng. Online 13 (1) (2014) 128, https://doi.org/10.1186/1475-925X-13-128.
[24] J.L.R. Andersson, S. Skare, J. Ashburner, How to correct susceptibility distortions in spin-echo echo-planar images: application to diffusion tensor imaging, Neuroimage 20 (2) (2003) 870–888.
[25] H.-P. Müller, G. Grön, R. Sprengelmeyer, J. Kassubek, A.C. Ludolph, N. Hobbs, J. Cole, R.A.C. Roos, A. Duerr, S.J. Tabrizi, G.B. Landwehrmeyer, S.D. Süssmuth, Evaluating multicenter DTI data in Huntington's disease on site specific effects: an ex post facto approach, Neuroimage: Clinical 2 (2013) 161–167, https://doi.org/10.1016/j.nicl.2012.12.005.
[26] M. Brett, I.S. Johnsrude, A.M. Owen, The problem of functional localization in the human brain, Nat. Rev. Neurosci. 3 (3) (2002) 243–249, https://doi.org/10.1038/nrn756.
[27] R.A.L. Menke, S. Körner, N. Filippini, G. Douaud, S. Knight, K. Talbot, M.R. Turner, Widespread grey matter pathology dominates the longitudinal cerebral MRI and clinical landscape of amyotrophic lateral sclerosis, Brain 137 (9) (2014) 2546–2555, https://doi.org/10.1093/brain/awu162.
[28] A. Unrath, H. Müller, A. Riecker, A.C. Ludolph, A. Sperfeld, J. Kassubek, Whole brain-based analysis of regional white matter tract alterations in rare motor neuron diseases by diffusion tensor imaging, Hum. Brain Mapp. 31 (11) (2010) 1727–1740, https://doi.org/10.1002/hbm.20971.
[29] A. Behler, J. Kassubek, H.-P. Müller, Age-related alterations in DTI metrics in the human brain—consequences for age correction, Front. Aging Neurosci. 13 (2021) 682109, https://doi.org/10.3389/fnagi.2021.682109.
[30] J. Rosskopf, H.-P. Müller, J. Dreyhaupt, M. Gorges, A.C. Ludolph, J. Kassubek, Ex post facto assessment of diffusion tensor imaging metrics from different MRI protocols: preparing for multicentre studies in ALS, Amyotroph Lateral Scler Frontotemporal Degener 16 (1–2) (2014) 92–101, https://doi.org/10.3109/21678421.2014.977297.
[31] H.-P. Müller, J. Dreyhaupt, F. Roselli, M. Schlecht, A.C. Ludolph, H.J. Huppertz, J. Kassubek, Focal alterations of the callosal area III in primary lateral sclerosis: an MRI planimetry and texture analysis, Neuroimage Clin. 26 (2020) 102223, https://doi.org/10.1016/j.nicl.2020.102223.
[32] S. Hofer, J. Frahm, Topography of the human corpus callosum revisited—comprehensive fiber tractography using diffusion tensor magnetic resonance imaging, Neuroimage 32 (3) (2006) 989–994, https://doi.org/10.1016/j.neuroimage.2006.05.044.
[33] G. Stockman, L.G. Shapiro, Computer Vision, Prentice Hall, Upper Saddle River, NJ, 2001. Chapter 7.
[34] A. Kunimatsu, S. Aoki, Y. Masutani, O. Abe, N. Hayashi, H. Mori, T. Masumoto, K. Ohtomo, The optimal trackability threshold of fractional anisotropy for diffusion tensor tractography of the corticospinal tract, Magn. Reson. Med. Sci. 3 (1) (2004) 11–17, https://doi.org/10.2463/mrms.3.11.
[35] H.-P. Mueller, A. Unrath, A.D. Sperfeld, A.C. Ludolph, A. Riecker, J. Kassubek, Diffusion tensor imaging and tractwise fractional anisotropy statistics: Quantitative analysis in white matter pathology, Biomed. Eng. Online 6 (1) (2007) 42, https://doi.org/10.1186/1475-925X-6-42.
[36] M. Stamelou, S. Knake, W.H. Oertel, G.U. Höglinger, Magnetic resonance imaging in progressive supranuclear palsy, J. Neurol. 258 (4) (2011) 549–558, https://doi.org/10.1007/s00415-010-5865-0.
[37] D.W. Dickson, R. Rademakers, M.L. Hutton, Progressive supranuclear palsy: pathology and genetics, Brain Pathol. 17 (1) (2007) 74–82, https://doi.org/10.1111/j.1750-3639.2007.00054.x.
[38] M.J. Armstrong, Progressive supranuclear palsy: an update, Curr. Neurol. Neurosci. Rep. 18 (3) (2018) 12, https://doi.org/10.1007/s11910-018-0819-5.
[39] F. Albrecht, S. Bisenius, J. Neumann, J. Whitwell, M.L. Schroeter, Atrophy in midbrain & cerebral/cerebellar pedunculi is characteristic for progressive supranuclear palsy – a double-validation whole-brain meta-analysis, Neuroimage: Clinical 22 (2019) 101722, https://doi.org/10.1016/j.nicl.2019.101722.
[40] D.B. Archer, T. Mitchell, R.G. Burciu, J. Yang, S. Nigro, A. Quattrone, A. Quattrone, A. Jeromin, N.R. McFarland, M.S. Okun, D.E. Vaillancourt, Magnetic resonance imaging and neurofilament light in the differentiation of parkinsonism, Mov. Disord. 35 (8) (2020) 1388–1395, https://doi.org/10.1002/mds.28060.
[41] J.-H. Lee, M.-S. Lee, Brain iron accumulation in atypical parkinsonian syndromes: in vivo MRI evidences for distinctive patterns, Front. Neurol. 10 (2019) 74, https://doi.org/10.3389/fneur.2019.00074.
[42] A. Worker, C. Blain, J. Jarosz, K.R. Chaudhuri, G.J. Barker, S.C.R. Williams, R. G. Brown, P.N. Leigh, F. Dell'Acqua, A. Simmons, Diffusion tensor imaging of Parkinson's disease, multiple system atrophy and progressive supranuclear palsy: a tract-based spatial statistics study, PLoS One 9 (11) (2014) e112638, https://doi.org/10.1371/journal.pone.0112638.
[43] M. Gorges, M.N. Maier, J. Rosskopf, O. Vintonyak, E.H. Pinkhardt, A.C. Ludolph, H.-P. Müller, J. Kassubek, Regional microstructural damage and patterns of eye movement impairment: a DTI and video-oculography study in neurodegenerative parkinsonian syndromes, J. Neurol. 264 (9) (2017) 1919–1928, https://doi.org/10.1007/s00415-017-8579-8.
[44] L. Chougar, N. Pyatigorskaya, S. Lehéricy, Update on neuroimaging for categorization of Parkinson's disease and atypical parkinsonism, Curr. Opin. Neurol. 34 (4) (2021) 514–524, https://doi.org/10.1097/WCO.0000000000000957.
[45] T. Potrusil, F. Krismer, V. Beliveau, K. Seppi, C. Müller, F. Troger, G. Göbel, R. Steiger, E.R. Gizewski, W. Poewe, C. Scherfler, Diagnostic potential of automated tractography in progressive supranuclear palsy variants, Park. Relat. Disord. 72 (2020) 65–71, https://doi.org/10.1016/j.parkreldis.2020.02.007.
[46] C. Brenneis, K. Seppi, M. Schocke, T. Benke, G.K. Wenning, W. Poewe, Voxel based morphometry reveals a distinct pattern of frontal atrophy in progressive supranuclear palsy, J. Neurol. Neurosurg. Psychiatr. 75 (2) (2004) 246–249.
[47] D.C. Paviour, S.L. Price, M. Jahanshahi, A.J. Lees, N.C. Fox, Longitudinal MRI in progressive supranuclear palsy and multiple system atrophy: rates and regions of atrophy, Brain 129 (4) (2006) 1040–1049, https://doi.org/10.1093/brain/awl021.
[48] K. Sakurai, A.M. Tokumaru, K. Shimoji, S. Murayama, K. Kanemaru, S. Morimoto, I. Aiba, M. Nakagawa, Y. Ozawa, M. Shimohira, N. Matsukawa, Y. Hashizume, Y. Shibamoto, Beyond the midbrain atrophy: wide spectrum of structural MRI finding in cases of pathologically proven progressive supranuclear palsy, Neuroradiology 59 (5) (2017) 431–443, https://doi.org/10.1007/s00234-017-1812-4.
[49] A. Shoeibi, N. Olfati, I. Litvan, Frontrunner in translation: progressive supranuclear palsy, Front. Neurol. 10 (2019) 1125, https://doi.org/10.3389/fneur.2019.01125.
[50] F. Caso, F. Agosta, M. Ječmenica-Lukić, I. Petrović, A. Meani, V.S. Kostic, M. Filippi, Progression of white matter damage in progressive supranuclear palsy with predominant parkinsonism, Park. Relat. Disord. 49 (2018) 95–99, https://doi.org/10.1016/j.parkreldis.2018.01.001.
[51] A. Lenka, S.A. Pasha, S. Mangalore, L. George, K.R. Jhunjhunwala, B.S. Bagepally, R.M. Naduthota, J. Saini, R. Yadav, P.K. Pal, Role of corpus callosum volumetry in differentiating the subtypes of progressive supranuclear palsy and early Parkinson's disease, Movement Disorders Clinical Practice 4 (4) (2017) 552–558, https://doi.org/10.1002/mdc3.12473.
[52] T.-T. Nguyen, J.-S. Cheng, Y.-L. Chen, Y.-C. Lin, C.-C. Tsai, C.-S. Lu, Y.-H. Weng, Y.-M. Wu, N.-T. Hoang, J.-J. Wang, Fixel-based analysis of white matter degeneration in patients with progressive supranuclear palsy or multiple system atrophy, as compared to Parkinson's disease, Front. Aging Neurosci. 13 (2021) 625874, https://doi.org/10.3389/fnagi.2021.625874.

[53] A. Padovani, B. Borroni, S.M. Brambati, C. Agosti, M. Broli, R. Alonso, P. Scifo, G. Bellelli, A. Alberici, R. Gasparotti, D. Perani, Diffusion tensor imaging and voxel based morphometry study in early progressive supranuclear palsy, J. Neurol. Neurosurg. Psychiatr. 77 (4) (2006) 457–463, https://doi.org/10.1136/jnnp.2005.075713.

[54] J. Rosskopf, H.-P. Müller, H.-J. Huppertz, A.C. Ludolph, E.H. Pinkhardt, J. Kassubek, Frontal corpus callosum alterations in progressive supranuclear palsy but not in Parkinson's disease, Neurodegener. Dis. 14 (4) (2014) 184–193, https://doi.org/10.1159/000367693.

[55] N. Nicastro, P.V. Rodriguez, M. Malpetti, W.R. Bevan-Jones, P. Simon Jones, L. Passamonti, F.I. Aigbirhio, J.T. O'Brien, J.B. Rowe, 18F-AV1451 PET imaging and multimodal MRI changes in progressive supranuclear palsy, J. Neurol. 267 (2) (2020) 341–349, https://doi.org/10.1007/s00415-019-09566-9.

[56] N. Spotorno, S. Hall, D.J. Irwin, T. Rumetshofer, J. Acosta-Cabronero, A.F. Deik, M. A. Spindler, E.B. Lee, J.Q. Trojanowski, D. Van Westen, M. Nilsson, M. Grossman, P.J. Nestor, C.T. McMillan, O. Hansson, Diffusion tensor MRI to distinguish progressive supranuclear palsy from α-synucleinopathies, Radiology 293 (3) (2019) 646–653, https://doi.org/10.1148/radiol.2019190406.

[57] J.L. Robinson, N. Yan, C. Caswell, S.X. Xie, E. Suh, V.M. Van Deerlin, G. Gibbons, D.J. Irwin, M. Grossman, E.B. Lee, V.M.-Y. Lee, B. Miller, J.Q. Trojanowski, Primary tau pathology, not copathology, correlates with clinical symptoms in PSP and CBD, J. Neuropathol. Exp. Neurol. 79 (3) (2020) 296–304, https://doi.org/10.1093/jnen/nlz141.

[58] A. Quattrone, M.G. Bianco, A. Antonini, D.E. Vaillancourt, K. Seppi, R. Ceravolo, A. P. Strafella, G. Tedeschi, A. Tessitore, R. Cilia, M. Morelli, S. Nigro, B. Vescio, P. P. Arcuri, R. De Micco, M. Cirillo, L. Weis, E. Fiorenzato, R. Biundo, A. Quattrone, Development and validation of automated magnetic resonance parkinsonism index 2.0 to distinguish progressive supranuclear palsy-parkinsonism from Parkinson's disease, Mov. Disord. 37 (6) (2022) 1272–1281, https://doi.org/10.1002/mds.28992.

[59] T. Evgeniou, M. Pontil, Support vector machines: theory and applications, in: G. Paliouras, V. Karkaletsis, C.D. Spyropoulos (Eds.), Machine Learning and its Applications, vol. 2049, Springer Berlin Heidelberg, 2001, pp. 249–257, https://doi.org/10.1007/3-540-44673-7_12.

[60] M.-C. Popescu, V.E. Balas, L. Perescu-Popescu, N.E. Mastorakis, Multilayer Perceptron and Neural Networks, 2009.

[61] S.B. Kotsiantis, Decision trees: a recent overview, Artif. Intell. Rev. 39 (4) (2013) 261–283, https://doi.org/10.1007/s10462-011-9272-4.

[62] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, https://doi.org/10.1023/A:1010933404324.

[63] S. Karamizadeh, S.M. Abdullah, M. Halimi, J. Shayan, M.J. Rajabi, Advantage and drawback of support vector machine functionality, in: 2014 International Conference on Computer, Communications, and Control Technology (I4CT), 2014, pp. 63–65, https://doi.org/10.1109/I4CT.2014.6914146.

[64] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, A. Müller, J. Nothman, G. Louppe, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, É. Duchesnay, Scikit-learn: Machine Learning in Python, 2012, https://doi.org/10.48550/ARXIV.1201.0490.

[65] F. Ciabuschi, B. Venkateswaran, Neural Networks with R: Smart Models Using CNN, RNN, Deep Learning, and Artificial Intelligence Principles, Packt, 2017.

[66] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444, https://doi.org/10.1038/nature14539.

[67] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G.S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, X. Zheng, TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems, 2016, https://doi.org/10.48550/ARXIV.1603.04467.

[68] A. Altmann, L. Toloşi, O. Sander, T. Lengauer, Permutation importance: a corrected feature importance measure, Bioinformatics 26 (10) (2010) 1340–1347, https://doi.org/10.1093/bioinformatics/btq134.

[69] L. van der Maaten, G. Hinton, Visualizing Data Using T-SNE, 2008, pp. 2579–2605.

[70] P. Péran, G. Barbagallo, F. Nemmi, M. Sierra, M. Galitzky, A.P. Traon, P. Payoux, W.G. Meissner, O. Rascol, MRI supervised and unsupervised classification of Parkinson's disease and multiple system atrophy, Mov. Disord. 33 (4) (2018) 600–608, https://doi.org/10.1002/mds.27307.

[71] D.B. Archer, J.T. Bricker, W.T. Chu, R.G. Burciu, J.L. Mccracken, S. Lai, S. A. Coombes, R. Fang, A. Barmpoutis, D.M. Corcos, A.S. Kurani, T. Mitchell, M. L. Black, E. Herschel, T. Simuni, T.B. Parrish, C. Comella, T. Xie, K. Seppi, N. I. Bohnen, M.L.T.M. Müller, R.L. Albin, F. Krismer, G. Du, M.M. Lewis, X. Huang, H. Li, O. Pasternak, N.R. McFarland, M.S. Okun, D.E. Vaillancourt, Development and validation of the automated imaging differentiation in parkinsonism (aid-P): a multi-site machine learning study, Lancet Digit Health 1 (5) (2019) e222–e231, https://doi.org/10.1016/s2589-7500(19)30105-0.

[72] A. Cherubini, M. Morelli, R. Nisticó, M. Salsone, G. Arabia, R. Vasta, et al., Magnetic resonance support vector machine discriminates between Parkinson disease and progressive supranuclear palsy, Mov. Disord. 29 (2014) 266–269, https://doi.org/10.1002/mds.2573799.

[73] L. Chougar, N. Pyatigorskaya, B. Degos, D. Grabli, S. Lehéricy, The role of magnetic resonance imaging for the diagnosis of atypical parkinsonism, Front. Neurol. 11 (2020) 665, https://doi.org/10.3389/fneur.2020.00665.

[74] A.S. Talai, J. Sedlacik, K. Boelmans, N.D. Forkert, Utility of multi-modal MRI for differentiating of Parkinson's disease and progressive supranuclear palsy using machine learning, Front. Neurol. 12 (2021) 648548, https://doi.org/10.3389/fneur.2021.648548.

[75] H. Li, G.K. Rajbahadur, D. Lin, C.-P. Bezemer, Z.M. Jiang, Keeping deep learning models in check: a history-based approach to mitigate overfitting, IEEE Access 12 (2024) 70676–70689, https://doi.org/10.1109/ACCESS.2024.3402543.

[76] B.R. Isaacs, A.C. Trutti, E. Pelzer, M. Tittgemeyer, Y. Temel, B.U. Forstmann, M. C. Keuken, Cortico-basal white matter alterations occurring in Parkinson's disease, PLoS One 14 (8) (2019) e0214343, https://doi.org/10.1371/journal.pone.0214343.

[77] J. Qu, R. Zhu, Y. Wu, G. Xu, D. Wang, Abnormal structural–functional coupling patterning in progressive supranuclear palsy is associated with diverse gradients and histological features, Commun. Biol. 7 (1) (2024) 1195, https://doi.org/10.1038/s42003-024-06877-0.

[78] H.J. Huppertz, L. Möller, M. Südmeyer, R. Hilker, E. Hattingen, K. Egger, F. Amtage, G. Respondek, M. Stamelou, A. Schnitzler, E.H. Pinkhardt, W.H. Oertel, S. Knake, J. Kassubek, G.U. Höglinger, Differentiation of neurodegenerative parkinsonian syndromes by volumetric magnetic resonance imaging analysis and support vector machine classification, Mov. Disord. 31 (10) (2016) 1506–1517, https://doi.org/10.1002/mds.26715.