



DISCRIMINATION BETWEEN ALZHEIMER'S DISEASE AND AMYOTROPHIC LATERAL SCLEROSIS VIA AFFINE INVARIANT SPHERICAL HARMONICS ANALYSIS OF SPECT IMAGES

*K. Horaisova**, *J. Dudasova*[†], *J. Kukal*[‡], *R. Rusina*[‡], *R. Matej*[‡], *M. Buncova*[§]

Abstract: Alzheimer's Disease (AD) is the most frequent form of degenerative dementia and its early diagnosis is essential for effective treatment. Functional imaging modalities including Single Photon Emission Computed Tomography (SPECT) are often used with such an aim. However, conventional evaluation of SPECT images relies on manual reorientation and visual evaluation of tomographic slices which is time consuming, subjective and therefore prone to error. Our aim is to show an automatic Computer-Aided Diagnosis (CAD) system for improving the early detection of the AD. For this purpose, affine invariant descriptors of 3D SPECT image can be useful. The method consists of four steps: evaluation of invariant descriptors obtained using spherical harmonic analysis, statistical testing of their significance, application of regularized binary index models, and model verification via leave-one-out cross-validation scheme. The second approach is based on Support Vector Machine (SVM) classifier and visualization with use of self-organizing maps. Our approaches were tested on SPECT data from 11 adult patients with definite Alzheimer's disease and 10 adult patients with Amyotrophic Lateral Sclerosis (ALS) who were used as controls. A significant difference between SPECT spherical cuts of AD group and ALS group was both visually and numerically evaluated.

Key words: *Alzheimer's disease, amyotrophic lateral sclerosis, image processing, SPECT, affine transformation, spherical harmonics, classification, statistical analysis, logit model, probit model, Cauchy model, support vector machine, self-organizing neural network*

Received: June 1, 2015

DOI: 10.14311/NNW.2018.28.002

Revised and accepted: February 19, 2018

*Katerina Horaisova – Corresponding author; Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Software Engineering, Trojanova 13, Prague, Czech Republic, E-mail: katerina.horaisova@jfji.cvut.cz

[†]Julie Dudasova; Jaromir Kukal; Czech Technical University in Prague, Faculty of Nuclear Sciences and Physical Engineering, Department of Software Engineering, Trojanova 13, Prague, Czech Republic, E-mail: julie.vavrycukova@gmail.com jaromir.kukal@jfji.cvut.cz

[‡]Robert Rusina; Radoslav Matej; Thomayer Hospital, Videnska 800, Prague, Czech Republic, E-mail: robert.rusina@lf1.cuni.cz radoslav.matej@ftn.cz

[§]Marie Buncova; Institute for Clinical and Experimental Medicine, Videnska 1958/9, Prague, Czech Republic, E-mail: marie.buncova@ikem.cz

1. Introduction

Alzheimer's Disease (AD) is the most frequent form of degenerative dementia and early diagnosis is essential for effective treatment [8]. Despite significant progress in the field of neurology and neurosciences during the last years, accurate diagnosis of AD is still challenging. Current diagnostic criteria for AD remain based on clinical appreciation, such as Diagnostic and Statistical Manual of Mental Disorders, fourth edition: DSM-IV-TR [1]; and the National Institute of Neurological Disorders and Stroke-Alzheimer Disease and Related Disorders working group: NINCDS-ADRDA [19]. Its sensitivity and specificity reach about 80-87% [9]. Neuroimaging may be particularly helpful in increase the diagnostic precision. For this reason, many research groups have recently focused on additional tools based on various image processing principles in order to refine diagnostic results: (i) hippocampal shape analysis of magnetic resonance images [12, 13, 25, 28]; (ii) hippocampal volume analysis of magnetic resonance images [18]; (iii) cerebral texture analysis of positron emission tomography sinograms [24].

Single Photon Emission Computed Tomography (SPECT) is a non-invasive and widely accessible imaging method which maps regional blood flow in different cerebral areas. Selective hypoperfusion in temporal and parietal regions has been attributed to AD with an overall specificity of 60-70% and sensitivity of about 80% [20, 26]. When using professional SPECT images obtained by experienced physicians in specialized centers, both sensitivity and specificity increase. However, the sensitivity and specificity proved not high enough for routine use of SPECT to be recommended for diagnostic purposes [6]. On the contrary, if the sensitivity and specificity further increase by more relevant data processing, SPECT would be ideal for routine use because of its reasonable financial costs and the possibility to use SPECT repeatedly for monitoring of disease's progression and dynamic patterns. Moreover, SPECT imaging could reliably differentiate local hypoperfusion in frontal and cingulate areas in depressive pseudodementia (memory clinics in depression, second most frequent condition in patients with memory complaints) from the typical temporo-parietal pattern in AD. As Magnetic Resonance Imaging (MRI) is unable to distinguish between AD and depression, SPECT could be very helpful in this respect.

The aim of our work was to develop a more adapted SPECT analysis protocol as a supportive method for discerning early symptoms of Alzheimer's disease. The principal approach takes advantage of the medical finding that affected brain is usually characterized by a different structure of gray and white matter and is based on mathematical and computer analysis of 3D brain scans. Having 3D intensity image of any structure, various characteristics can be defined and measured. We focus on characteristics, which are invariant with respect to any spatial affine transformation. In other words and more specifically, via 3D integration of SPECT brain scans with spherical harmonic functions we obtain characteristics (descriptors) which are not sensitive to body translation, rotation and scaling. Such characteristics are further processed, their dimension is reduced and their classification ability is determined and verified using statistical binary response models [27], particularly logistic regression [14], and the datamining techniques such as support vector machine or self-organizing neural network. Comparison of

various classification approaches including regularized regression is included in [4]. Self-organizing neural network was successfully used as classifier for 2D image invariant descriptors in [15].

2. Affine invariant descriptors

Our approach is based on affine invariant descriptors, which are independent on translation, rotation, and coordinate-by-coordinate scaling of 3D SPECT image. It is useful for elimination of image registration techniques from data preprocessing.

2.1 Affine transform of intensity image in \mathbb{R}^n

The first aim of *affine invariance* can be introduced via affine transformation description in \mathbb{R}^n related to n -dimensional intensity image investigated in continuous coordinates. Lately, the dimension will be set to $n = 3$ and space discretization will be permitted. The *original intensity image* can be declared as a density of n -dimensional random variable X in mathematical statistics

$$f: \mathbb{R}^n \rightarrow \mathbb{R}_0^+ \quad (1)$$

satisfying

$$\int_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \, d\mathbf{x} = 1, \quad (2)$$

where

$$d\mathbf{x} = \prod_{k=1}^n dx_k. \quad (3)$$

When the image is constrained in size and intensity, all image moments are finite. Then we can define the *image momentum vector* as mean value

$$\mathbf{x}_0 = EX = \int_{\mathbf{x} \in \mathbb{R}^n} \mathbf{x} f(\mathbf{x}) \, d\mathbf{x}. \quad (4)$$

Corresponding covariance matrix \mathbf{C} of original image f is supposed to be positive definite, which is valid for non-degenerated image of full dimension. Adequate formula is then

$$\mathbf{C} = E(X - \mathbf{x}_0)(X - \mathbf{x}_0)^T = \int_{\mathbf{x} \in \mathbb{R}^n} (\mathbf{x} - \mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)^T f(\mathbf{x}) \, d\mathbf{x}. \quad (5)$$

Let Y be a random vector satisfying the *whitening conditions*

$$EY = \mathbf{0} \in \mathbb{R}^n, \quad EYY^T = \mathbf{I} \in \mathbb{R}^{n \times n}. \quad (6)$$

Let $\mathbf{A} \in \mathbb{R}^{n \times n}$, $\mathbf{b} \in \mathbb{R}^n$ be unknown regular matrix and shifting vector of general affine transform from \mathbb{R}^n to \mathbb{R}^n as

$$X = \mathbf{A}Y + \mathbf{b}. \quad (7)$$

According to (4) and linearity principle we obtained the shifting vector

$$\mathbf{x}_0 = \mathbf{A}(\mathbf{E}Y) + \mathbf{b} = \mathbf{b}. \quad (8)$$

According to (5) and linearity principle we obtained the relationship between covariance matrix \mathbf{C} and matrix \mathbf{A} as

$$\begin{aligned} \mathbf{C} &= \mathbf{E}(X - \mathbf{x}_0)(X - \mathbf{x}_0)^T = \mathbf{E}(X - \mathbf{b})(X - \mathbf{b})^T = \\ &= \mathbf{E}A Y Y^T \mathbf{A}^T = \mathbf{A}(\mathbf{E}Y Y^T) \mathbf{A}^T = \mathbf{A} \mathbf{I} \mathbf{A}^T = \mathbf{A} \mathbf{A}^T > 0, \end{aligned} \quad (9)$$

which is equation for unknown affine transformation matrix \mathbf{A} with infinite number of solutions, which can be generated by rotation of a particular solution. The combination of *Singular Value Decomposition* (SVD) of matrix \mathbf{A} and *Eigen Value Decomposition* (EVD) of matrix \mathbf{C} is the trick for the solution of (9).

Let \mathbf{U} , \mathbf{S} , \mathbf{V} , \mathbf{Q} , $\mathbf{E} \in \mathbb{R}^{n \times n}$ be unknown matrices satisfying conditions SVD and EVD as

$$\mathbf{A} = \mathbf{U} \mathbf{S} \mathbf{V}^T, \quad \mathbf{C} = \mathbf{Q} \mathbf{E} \mathbf{Q}^T, \quad (10)$$

$$\mathbf{U}^T \mathbf{U} = \mathbf{V}^T \mathbf{V} = \mathbf{Q}^T \mathbf{Q} = \mathbf{I} \in \mathbb{R}^{n \times n}. \quad (11)$$

Finally, \mathbf{S} , \mathbf{E} are diagonal positive definite matrices.

Performing EVD of covariance matrix \mathbf{C} , we obtain diagonal matrix \mathbf{E} of eigenvalues and orthogonal matrix \mathbf{Q} of eigenvectors. The solution of (9) can be obtained by substitution of \mathbf{A} via SVD as

$$\mathbf{C} = \mathbf{A} \mathbf{A}^T = \mathbf{U} \mathbf{S} \mathbf{V}^T (\mathbf{U} \mathbf{S} \mathbf{V}^T)^T = \mathbf{U} \mathbf{S} \mathbf{V}^T \mathbf{V} \mathbf{S} \mathbf{U}^T = \mathbf{U} \mathbf{S}^2 \mathbf{U}^T = \mathbf{Q} \mathbf{E} \mathbf{Q}^T. \quad (12)$$

Finally, we directly obtained the matrices

$$\mathbf{U} = \mathbf{Q}, \quad \mathbf{S} = \mathbf{E}^{\frac{1}{2}} \quad (13)$$

and the matrix \mathbf{V} of the *second rotation* is undetermined, then free. The solution of (9) is any matrix of the form

$$\mathbf{A} = \mathbf{Q} \mathbf{E}^{\frac{1}{2}} \mathbf{V}^T. \quad (14)$$

The inverse transformation of (7) is very useful for the future affine invariant system. It is defined as

$$Y = \mathbf{A}^{-1} (X - \mathbf{x}_0) = \mathbf{V} \mathbf{E}^{-\frac{1}{2}} \mathbf{Q}^T (X - \mathbf{x}_0). \quad (15)$$

It can be decomposed using unique *particular transform*

$$Y_p = \mathbf{E}^{-\frac{1}{2}} \mathbf{Q}^T (X - \mathbf{x}_0). \quad (16)$$

And then perform unknown rotation around origin as

$$Y = \mathbf{V} Y_p. \quad (17)$$

Adequate density of stochastic variable Y_p is

$$g(Y_p) = f(\mathbf{x}_0 + \mathbf{Q} \mathbf{E}^{\frac{1}{2}} Y_p) \det \left(\mathbf{E}^{\frac{1}{2}} \right). \quad (18)$$

We plan a processing of Y_p which will be invariant to any rotation around origin. So it will bring the same values of characteristics as the processing of Y . Then the whole system will be affine invariant.

2.2 Spherical harmonic functions

Before defining spherical harmonics [21], let us assume that *spherical functions* are square integrable mappings $(\theta, \varphi) \rightarrow \mathbb{C}$ lying in the Hilbert space $L^2(S^2)$, where S^2 is two-dimensional sphere in \mathbb{R}^3 and variables $0 \leq \theta < \pi, 0 \leq \varphi < 2\pi$ come from spherical coordinate system. Then *spherical harmonics* are a set of functions $Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi)$ solving the Laplace equation $\nabla^2 F = 0$ expressed in the spherical coordinates. Assuming that $P_l^m(x)$ is an *associated Legendre polynomial* [22] over interval $[-1; 1]$ with integers $l \in \mathbb{N}_0$ and $m \in \{0, \dots, l\}$, the solution of such Laplace differential equation is

$$Y_l^m(\theta, \varphi) = N_l^{|m|} P_l^{|m|}(\cos \theta) e^{im\varphi}. \quad (19)$$

Thus, $Y_l^m(\theta, \varphi)$ is a single *spherical harmonic function* with degree l and order m , where $-l \leq m \leq l$ and N_l^m stands for a normalization coefficient. This coefficient can be derived from the equation which proves orthogonal property of spherical harmonic system as

$$\int_0^{2\pi} \int_0^\pi Y_l^m(\theta, \varphi) \overline{Y_{l'}^{m'}}(\theta, \varphi) \sin \theta \, d\theta \, d\varphi = \delta_{m,m'} \delta_{l,l'}, \quad (20)$$

where $\delta_{i,j}$ is the Kronecker delta. Then the coefficient is

$$N_l^m = \sqrt{\frac{2l+1}{4\pi} \cdot \frac{(l-m)!}{(l+m)!}}. \quad (21)$$

The orthogonal property of $Y_l^m(\theta, \varphi)$ enables us to constitute an orthonormal basis of the $L^2(S^2)$ space for the projection of other spherical functions similarly as is done for the Fourier transform. The projection of a real-valued spherical function $F(\theta, \varphi) \in L^2(S^2)$ into spectral coefficients can be computed as

$$c_{l,m} = \int_0^{2\pi} \int_0^\pi F(\theta, \varphi) \overline{Y_l^m}(\theta, \varphi) \sin \theta \, d\theta \, d\varphi. \quad (22)$$

There is a special spectral characteristic

$$s_l = \left(\sum_{m=-l}^{+l} |c_{l,m}|^2 \right)^{\frac{1}{2}}, \quad (23)$$

which is invariant to any rotation of $SO(3)$ group [11] for given $l \in \mathbb{N}_0$. It means that any rotation of function f within sphere S^2 does not change the value of s_l which is useful in classification applications.

2.3 Affine invariant characteristics in \mathbb{R}^3

The affine invariant characteristics for the classification of 3D image will be designed in \mathbb{R}^3 . The combination of spherical harmonic function with any radial symmetric

function will produce affine invariant weighted moments. The problem will be studied in *spherical polar coordinates* [22] of \mathbf{y}_p as

$$\begin{aligned} y_{1p} &= r \sin \theta \cos \varphi, \\ y_{2p} &= r \sin \theta \sin \varphi, \\ y_{3p} &= r \cos \theta. \end{aligned} \tag{24}$$

Let $w : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ be weight function. The recognition system is based on *generalized spectral coefficients*

$$q_{l,m} = \int_0^\pi \int_0^{2\pi} \int_0^\infty g(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) w(r) \bar{Y}_l^m(\theta, \varphi) r^2 \sin \theta dr d\varphi d\theta. \tag{25}$$

We can define *radial projection of image intensity* (density) h as

$$h(\theta, \varphi) = \int_0^\infty r^2 w(r) g(r \sin \theta \cos \varphi, r \sin \theta \sin \varphi, r \cos \theta) dr. \tag{26}$$

After the direct application of Fubini theorem [22] to (25) and using (26) we derived

$$q_{l,m} = \int_0^{2\pi} \int_0^\pi h(\theta, \varphi) \bar{Y}_l^m(\theta, \varphi) \sin \theta d\theta d\varphi. \tag{27}$$

There is a direct analogy between (22) and (27), where $q_{l,m}$ are spectral coefficients of radial projection h . Finally, according to (23), we obtained new *affine invariant characteristics*

$$I_l = \left(\sum_{m=-l}^{+l} |q_{l,m}|^2 \right)^{\frac{1}{2}}. \tag{28}$$

Both radial projection function (26) and the spectral coefficients (27) are linear with respect to weight function $w(r)$, which means the world of spectral coefficients is not too rich. But the final invariant characteristics (28) are not linear with respect to the weight function. So, the world of invariant characteristics has many useful properties and special shapes of $w(r)$ can bring an advantage in the recognition task. The practical recognition is based on the discrete 3D image in \mathbb{Z}^3 with many additional problems. They are caused by space discretization, intensity quantization and noise. Resulting recognition system is not perfectly affine invariant as we would wish to be. This negative effect can be decreased by the higher resolution in the space coordinates and intensity axes together with noise filtering.

There are many possibilities how to design the weight function $w(r)$. The analyzed image after the transform $g(\mathbf{y}_p)$ is normalized and thus the radius r is dimensionless. We introduce scaling factor $\rho > 0$ in radius, which enables to analyze the normalized image in given resolution. First we define *constrained weight function* by conditions

$$w(\rho) > 0, \quad \forall r > \rho : w(r) = 0. \tag{29}$$

As a case of constrained weight function which is based on the tradition in 3D image processing was chosen *spherical cut* defined as

$$w(r) = \rho^{-2} \delta(r - \rho), \quad (30)$$

where $\delta(x)$ is Dirac function. After the integration in (26) we recognize that the radial projection is

$$h(\theta, \varphi) = g(\rho \sin \theta \cos \varphi, \rho \sin \theta \sin \varphi, \rho \cos \theta). \quad (31)$$

Therefore, we are prepared to calculate individual invariant descriptors of 3D SPECT image of human brain and for hypotheses testing about their significance in Alzheimer's disease diagnostics.

3. Binary response index model

For better classification power, we can combine individual descriptors of various orders and radii. Binary response index model (logit, probit) is a traditional choice in the case of classification into two classes.

We suppose a model [27] with m real inputs \mathbf{x} and single binary output y in the form

$$y = h(\mathbf{x}^T \beta + e), \quad (32)$$

where

$$h(z) = \begin{cases} 1, & z > 0 \\ 0, & z \leq 0 \end{cases} \quad (33)$$

is the Heaviside's unit step function, $\mathbf{x}, \beta \in \mathbb{R}^{m+1}$, $x_0 = 1$, e is the continuous random variable with positive and symmetric probability density function $g(z)$ around zero. Its cumulative distribution function is

$$G(z) = \int_{-\infty}^z g(u) du. \quad (34)$$

So, the output y is also of stochastic nature and it can be described via probability

$$\begin{aligned} p(\mathbf{x}) &= \text{prob}(y = 1 | \mathbf{x}) = \text{prob}(\mathbf{x}^T \beta + e > 0) = \\ &= \text{prob}(e > -\mathbf{x}^T \beta) = 1 - G(-\mathbf{x}^T \beta) = G(\mathbf{x}^T \beta), \end{aligned} \quad (35)$$

which is well known formula [14, 27] for *binary model* related to *logistic regression*. It is necessary to satisfy $0 < G(z) < 1$ for all real arguments z . Fortunately, it implies from $g(z) > 0$ everywhere, which was declared above.

3.1 Special cases of binary model

Binary model was first published [3] as a *probit model* with nonlinearity

$$G(z) = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^z \exp\left(-\frac{u^2}{2}\right) du \quad (36)$$

and corresponding density function of standard normal distribution as

$$g(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right). \quad (37)$$

Then, the *logit model* (logistic regression) with nonlinearity

$$G(z) = \frac{1}{1 + \exp(-z)} \quad (38)$$

and corresponding density function of logistic distribution

$$g(z) = \frac{\exp(-z)}{(1 + \exp(-z))^2} \quad (39)$$

was published [2].

This model is frequently used in many applications. The logit model approaches its asymptotes less rapidly than the probit one. The other models are also possible to use. The density of Cauchy (t_1) distribution

$$g(z) = \frac{1}{\pi(1 + z^2)} \quad (40)$$

generates the nonlinearity

$$G(z) = \frac{1}{2} + \frac{1}{\pi} \arctan z \quad (41)$$

with several amazing properties corresponding to heavy tail effect of Cauchy distribution. The variety of nonlinear characteristics can help us to choose a binary model and its parameters with the best possible quality of fitting.

3.2 Parameter estimation

The estimation of model parameters is frequently performed via maximization of likelihood function or its logarithm. Resulting point estimate (if exists) has a very good asymptotic properties [14, 27]. Let N be number of observations and $(\mathbf{x}_k, \mathbf{y}_k)$ be individual observation for $k = 1, \dots, N$. Thus, the density of y for individual \mathbf{x}_k is

$$\varphi(y|\mathbf{x}_k, \beta) = [G(\mathbf{x}_k^T \beta)]^y [1 - G(\mathbf{x}_k^T \beta)]^{1-y}. \quad (42)$$

The *logarithmic likelihood function* over all observations is defined as

$$L(\beta) = \sum_{k=1}^N (y_k \log G(\mathbf{x}_k^T \beta) + (1 - y_k) \log (1 - G(\mathbf{x}_k^T \beta))). \quad (43)$$

Maximization of $L(\beta)$ on \mathbb{R}^n is impossible in the case of linear separable pattern sets. Using Bayesian approach, we suppose prior parameter distribution $\beta \sim N(0, \sigma^2 \mathbf{I})$ as n -dimensional Gaussian one with parameter $\sigma = 0$. Adequate *quasi-log-likelihood* function is

$$B(\beta) = L(\beta) - \frac{1}{2\sigma^2} \|\beta\|^2. \quad (44)$$

This approach is called *Bayesian regularization*.

The point estimate \mathbf{b} of parameter vector is obtainable via maximization of objective function L on \mathbb{R}^n as

$$\mathbf{b} = \hat{\beta} \in \operatorname{argmax}_{\beta \in \mathbb{R}^n} B(\beta). \quad (45)$$

The existence but not uniqueness of the estimate \mathbf{b} is guaranteed in this case. The analysis of asymptomatic variance begins with matrix

$$\mathbf{U} = \sum_{k=1}^N \frac{g(\mathbf{x}_k^T \mathbf{b}) \mathbf{x}_k \mathbf{x}_k^T}{G(\mathbf{x}_k^T \mathbf{b}) (1 - G(\mathbf{x}_k^T \mathbf{b}))} + \sigma^2 \mathbf{I}. \quad (46)$$

When the matrix \mathbf{U} is regular, it is also positive definite and the asymptotic variance of estimate \mathbf{b} is

$$\operatorname{Avar}(\mathbf{b}) = \mathbf{V} = \mathbf{U}^{-1}. \quad (47)$$

The asymptotic standard error of estimate \mathbf{b} is

$$\operatorname{Astd}(\mathbf{b}) = \mathbf{s} = \operatorname{diag}(\mathbf{V})^{\frac{1}{2}} \quad (48)$$

and corresponding approximate 95% confidence interval (CI) is

$$\beta \in [\mathbf{b} - 1.96\mathbf{s}, \mathbf{b} + 1.96\mathbf{s}]. \quad (49)$$

The confidence interval is more useful for final report than for significance testing due to its sensitivity to singularity of matrix \mathbf{U} .

4. Case study: Alzheimer's disease diagnostics

Both affine invariant descriptors (28) and three binary index models (36), (38), (41) can be directly applied to 3D SPECT of human brain related to diagnostics of Alzheimer's disease.

4.1 Patients and control groups

SPECT data from 11 adult patients (6 males, 5 females) with definite Alzheimer's disease confirmed by post mortem brain autopsy were used for analysis. Hippocampal regions and neocortical areas observations using a silver stain impregnation method was confirmed through immunohistochemical testing using specific monoclonal antibodies against hyperphosphorylated tau protein and amyloid beta peptide. The neocortical stage of Alzheimer's disease was consistent with the final diagnostic features. When SPECT images were obtained (between 2003 and 2005), the patients fulfilled diagnostic clinical criteria for mild to moderate probable AD. All of these patients were demented.

SPECT data from 10 adult patients (7 males, 3 females) with Amyotrophic Lateral Sclerosis (ALS) were used as controls. Besides SPECT, these patients underwent detailed cognitive evaluation focusing on the general cognitive status as well as the executive, language and visuospatial functions. Using such protocol it

is safe to claim, that these patients were definitely not demented at the moment of SPECT data acquisition. Also the post mortem neuropathological investigation did not show significant signs of AD related pathology.

SPECT study was started after the injection of hexamethyl propylene amine oxime labeled with ^{99m}Tc Technetium. Dual-head gamma camera DST-XL SOPHA with LEHR collimator [26] was used. Software Multi Dim Cerebral SPECT for semiquantitative analysis regional uptake was applied. Data from SPECT regional cerebral perfusion in selected brain regions were analyzed for both hemispheres separately, preferentially in orbito-frontal, dorsolateral prefrontal, anterior temporal and parietal regions.

4.2 Methodology of image processing and testing

Individual 3D scans were represented as $128 \times 128 \times 128$ matrices of non-negative intensities (see Fig. 1). Applying normalization conditions (2), we can continue in evaluation of \mathbf{x}_0 (4), \mathbf{C} (5), and transformed density $g(\mathbf{y}_p)$ (18) for any dimensionless vector \mathbf{y}_p . Using spherical cut (30) for radii $\rho = 0.3, 0.4, 0.5, \dots, 3.0$ (see Fig. 2, 3) we directly calculate affine invariant descriptors I_l (28) for $l = 0, 1, \dots, 20$ via integration with spherical harmonic functions (Fig. 4). Therefore, we obtain 588 numerical descriptors of various orders and radii for each patient.

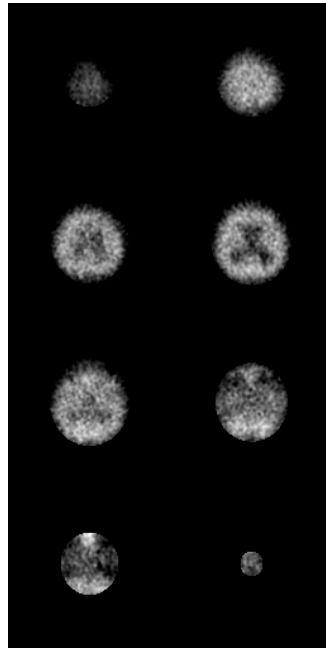


Fig. 1 Original SPECT image data of AD patient (8 equidistant slices).

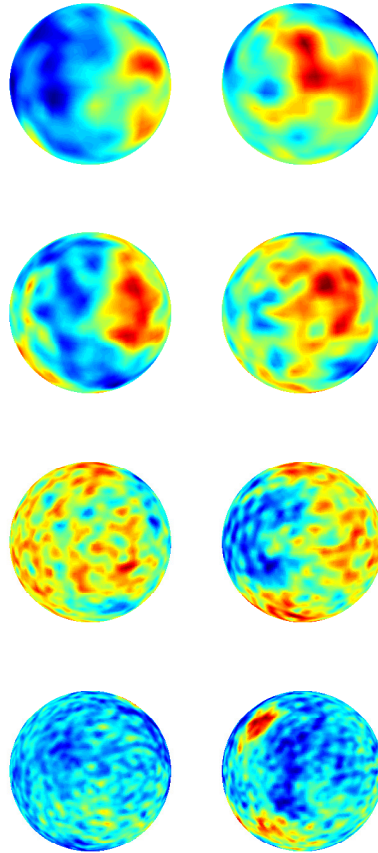


Fig. 2 Spherical cuts (30) of typical AD patient for $\rho = 0.6, 1.5, 1.9, 2.6$ in front (left) and back (right) view.

4.2.1 Traditional analysis with logistic regression

In the first experiment, we test hypothesis $H_0 : M_k(\text{AD}) = M_k(\text{ALS})$ against alternative $H_A : M_k(\text{AD}) \neq M_k(\text{ALS})$ using Wilcoxon-Mann-Whitney (WMW) test on level $\alpha = 0.05$ where M_k is median of k^{th} descriptor in given group of patients.

Probit (36), logit (38), and Cauchy (41) models with regularization are used in the third step. The main aim of binary response model application consists in selection of statistically significant orders l and radii ρ . It is useful for dimensionality reduction, biomedical interpretation, and good cross-validation properties.

Leave-one-out scheme of cross-validation is used in the last step. This approach is useful mainly in the case of small number of patterns. Binary models (36), (38), (41) are applied to complete pattern set except i^{th} pattern. This removed pattern is then used for response verification. Repeating this procedure for all patients, we can easily evaluate classification error err as a fraction of misclassified patterns.

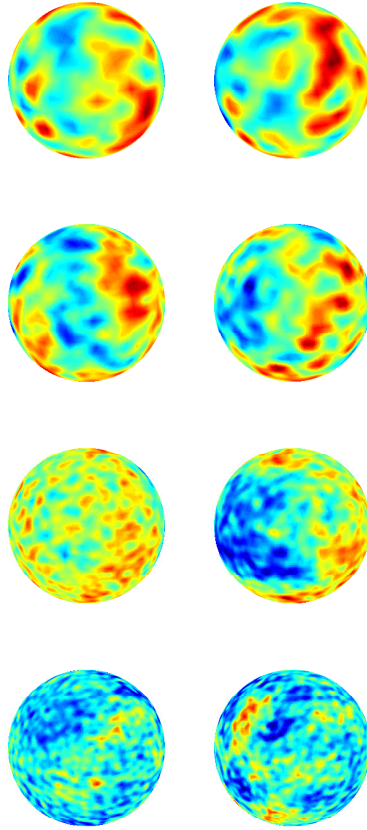


Fig. 3 Spherical cuts (30) of typical ALS patient for $\rho = 0.6, 1.5, 1.9, 2.6$ in front (left) and back (right) view.

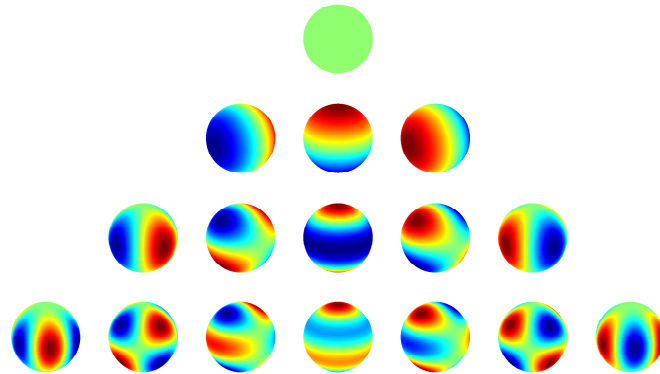


Fig. 4 Real part of spherical harmonic functions (19) for $l = 0, 1, 2, 3$ (rows) and $-l \leq m \leq l$ (columns).

4.2.2 Advanced analysis via SVM and SOM

Another experiment was based on use of Support Vector Machine (SVM) classifier [7]. We use standard library LIBSVM [5] which implements among others also linear SVM. All 588 invariant features are used for classification by SVM. Our aim is to find groups of three, four, and five features, which separate AD and ALS patients by linear hyperplane with the smallest classification error.

Leave-one-out scheme of cross-validation is used because we have small number of patterns. SVM is applied to complete pattern set except i^{th} pattern. This removed pattern is then used for response verification. Repeating this procedure for all patients, we can easily evaluate classification error err as a fraction of misclassified patterns. As the last step we use Kohonen Self-Organizing Maps (SOM) [16] neural network for cluster forming and relationship visualization. Our aim is categorization of patients with minimum errors. This technique will be applied only to the promising feature combinations.

4.3 Results

4.3.1 Traditional analysis with logistic regression

After the first two steps we obtained statistically significant affine invariants. Their list is included in Tab. I for $p_{\text{value}} < 0.05$. Dimensionless radius $\rho = 0.6$ frequently

invariant	ρ	median(AD)	median(ALS)	p_{value}
0	0.6	0.0606	0.0668	0.0257
1	2.6	0.0006	0.0030	0.0036
1	0.5	0.0135	0.0072	0.0058
1	2.5	0.0026	0.0055	0.0073
1	0.6	0.0131	0.0057	0.0091
1	0.4	0.0137	0.0099	0.0113
1	0.7	0.0125	0.0078	0.0113
1	1.9	0.0051	0.0081	0.0211
3	2.6	0.0012	0.0021	0.0211
5	0.9	0.0061	0.0038	0.0058
8	1.7	0.0038	0.0033	0.0257
10	2.2	0.0026	0.0018	0.0036
11	2.3	0.0022	0.0018	0.0257
12	0.5	0.0016	0.0021	0.0091
13	1.7	0.0021	0.0024	0.0211
14	1.2	0.0023	0.0028	0.0173
16	0.5	0.0004	0.0007	0.0257
17	0.3	0.0003	0.0003	0.0257
18	0.7	0.0011	0.0014	0.0211
18	2.6	0.0003	0.0006	0.0211
20	2.3	0.0014	0.0011	0.0173

Tab. I Individual invariants with $p_{\text{value}} < 0.05$ (according to WMW test).

occurs in Tab. I. Therefore, all invariants of this radius are collected in Tab. II. The best invariant from this table is I_1 . Its dependency on dimensionless radius ρ is demonstrated in Tab. III. There are two interesting ranges: $\rho \in [0.4, 0.7]$ and $\rho \in [2.5, 2.7]$, where I_1 reaches very low p_{value} 's.

invariant	median(AD)	median(ALS)	p_{value}
0	0.0606	0.0668	0.0257
1	0.0131	0.0057	0.0091
2	0.0070	0.0062	0.3447
3	0.0083	0.0054	0.0539
4	0.0051	0.0044	0.0640
5	0.0034	0.0041	0.4727
6	0.0039	0.0037	1,0000
7	0.0036	0.0040	0.7913
8	0.0037	0.0037	0.6232
9	0.0035	0.0034	0.3847
10	0.0025	0.0033	0.0757
11	0.0026	0.0033	0.1620
12	0.0020	0.0028	0.0890
13	0.0019	0.0024	0.0640
14	0.0016	0.0019	0.1041
15	0.0013	0.0016	0.1212
16	0.0009	0.0013	0.0539
17	0.0007	0.0009	0.0452
18	0.0006	0.0009	0.0539
19	0.0004	0.0006	0.0640
20	0.0004	0.0005	0.0890

Tab. II *Individual invariants for $\rho = 0.6$.*

General aim of cross-validation is to find the best selection of invariants, the best selection of radii, and the best regularization parameter σ , which have the lowest possible error err . Final results are collected in Tab. IV. The best choice is to use I_1 , I_3 , and I_5 for radii $\rho = 0.6$ and $\rho = 2.6$ in combination with logit or Cauchy model, which comes to 5% error in leave-one-out cross-validation. The probit model was less successful because of absence of heavy-tails.

4.3.2 Advanced analysis via SVM and SOM

Linear SVM classifier was used for all combinations of three, four, and five features separately to find the optimal combinations, which were able to separate 11 AD patients and 10 ALS patients with linear hyperplane without error. Our aim is to create an automatic CAD system for detection of the AD, therefore other combinations are not useful. The value of margin M between feature vectors of AD and ALS patients is measured. Consequently, the leave-one-out cross-validation was performed for these successful combinations. The minimum value of all 21 margins

ρ	median(AD)	median(ALS)	p_{value}
0.3	0.0145	0.0104	0.0757
0.4	0.0137	0.0099	0.0113
0.5	0.0135	0.0072	0.0058
0.6	0.0131	0.0057	0.0091
0.7	0.0125	0.0078	0.0113
0.8	0.0114	0.0077	0.0376
0.9	0.0109	0.0080	0.2413
1.0	0.0090	0.0087	0.6776
1.1	0.0082	0.0101	0.5708
1.2	0.0086	0.0117	0.5708
1.3	0.0097	0.0128	0.6232
1.4	0.0123	0.0142	0.5708
1.5	0.0141	0.0153	0.5205
1.6	0.0123	0.0158	0.4727
1.7	0.0124	0.0146	0.1620
1.8	0.0094	0.0116	0.0452
1.9	0.0051	0.0081	0.0211
2.0	0.0038	0.0048	0.0757
2.1	0.0038	0.0037	0.5205
2.2	0.0072	0.0055	0.1212
2.3	0.0082	0.0066	0.2413
2.4	0.0055	0.0066	0.6232
2.5	0.0026	0.0055	0.0073
2.6	0.0006	0.0030	0.0036
2.7	0.0001	0.0012	0.0312
2.8	0.0000	0.0002	0.3038
2.9	0.0000	0.0000	0.353
3.0	0.0000	0.0000	0.7798

Tab. III Individual radii for the first invariant I_1 .

type	σ	invariants	ρ 's	err	p_{value}
probit	1885	1; 3; 5	0.6; 2.6	0.10	3.61×10^{-3}
logit	17510	1; 3; 5	0.6; 2.6	0.05	1.83×10^{-4}
Cauchy	2361	1; 3; 5	0.6; 2.6	0.05	1.83×10^{-4}

Tab. IV Cross-validation of regularized logistic regression.

obtained from all leave-one-out tests is measured and denoted as M_{loa} . Results are summarized in Tab. V for five best groups of three features, in Tab. VI for five best groups of four features, and in Tabs. VII, VIII for five best groups of five features.

invariants (ρ 's)	all samples	training set		testing set
	M	M_{loa}	err	err
1 (2.6); 3 (1.6); 19 (0.9)	0.1471	0.0205	0.0000	0.1052
1 (2.6); 3 (2.6); 19 (0.9)	0.1249	0.0117	0.0000	0.1052
1 (2.6); 8 (2.0); 19 (0.9)	0.0999	0.0108	0.0000	0.1667
1 (0.7); 1 (2.6); 19 (0.9)	0.0794	0.0078	0.0000	0.2353
1 (0.6); 1 (2.6); 19 (0.9)	0.0768	0.0075	0.0000	0.2353

Tab. V Cross-validation for combinations of three features.

We can easily find out from Tab. V, that no combination of three features is able to separate the patients without error by the leave-one-out cross validation. First column of the table includes invariants and their appropriate values of ρ , second and third column include the value of margins M and M_{loa} , fourth column includes the mean value of errors measured on 21 training sets, and the last column includes the error measured consecutively on 21 testing sets. Reached results are ordered by descend margin M (we obtained the same order for M_{loa}) and the best is the combination of I_1 for $\rho = 2.6$, I_3 for $\rho = 1.6$, and I_{19} for $\rho = 0.9$, where we reached the margin $M = 0.15$ and the classification error $err = 0.11$. Many combinations of four features were able to distinguish between AD and ALS patients without error. Five combinations with the top margin M (we obtained the same order for M_{loa}) can be seen in Tab. VI. The best is the combination of I_1 for $\rho = 0.6$ and $\rho = 2.6$, I_{11} for $\rho = 0.8$, and I_{19} for $\rho = 0.9$, where we reached the margin $M = 0.56$.

Responsibility of this method could increase with growing amount of features. This is the reason, why combinations of five features were tested consequently. Values of M and M_{loa} increased as well as the amount of successful combinations. The list of five combinations with the top margin M can be seen in Tab. VII. The best is the combination of I_1 for $\rho = 0.7$, $\rho = 2.6$, and $\rho = 2.8$, I_{11} for $\rho = 0.8$, and I_{19} for $\rho = 0.9$, where we reached margins $M = 0.69$ and $M_{\text{loa}} = 0.31$. For comparison, the results ordered by M_{loa} can be seen in Tab. VIII. Here we obtained the best combination of I_1 for $\rho = 0.7$ and $\rho = 2.6$, I_3 for $\rho = 2.8$, I_{11} for $\rho = 0.8$, and I_{19} for $\rho = 0.9$, where we reached margins $M = 0.64$ and $M_{\text{loa}} = 0.41$.

invariants (ρ 's)	all samples	training set		testing set
	M	M_{loa}	err	err
1 (0.6); 1 (2.6); 11 (0.8); 19 (0.9)	0.5572	0.3081	0.0000	0.0000
1 (0.6); 1 (2.6); 11 (1.7); 19 (0.9)	0.5242	0.2959	0.0000	0.0000
1 (0.7); 1 (2.6); 11 (1.7); 19 (0.9)	0.5009	0.2306	0.0000	0.0000
1 (0.5); 1 (2.6); 11 (0.8); 19 (0.9)	0.4746	0.1950	0.0000	0.0000
1 (0.7); 1 (2.6); 11 (0.8); 19 (0.9)	0.4646	0.1716	0.0000	0.0000

Tab. VI Cross-validation for combinations of four features.

id	invariants (ρ 's)	all samples	training set		testing set
		M	M_{loa}	err	err
1	1 (0.7); 1 (2.6); 1 (2.8); 11 (0.8); 19 (0.9)	0.6930	0.3080	0.0000	0.0000
2	1 (0.6); 1 (2.6); 3 (2.6); 11 (1.7); 19 (0.9)	0.6634	0.3568	0.0000	0.0000
3	1 (0.5); 1 (2.6); 1 (2.7); 8 (1.9); 19 (0.9)	0.6614	0.3383	0.0000	0.0000
4	1 (0.7); 1 (2.6); 3 (2.6); 11 (1.7); 19 (0.9)	0.6604	0.2549	0.0000	0.0000
5	1 (0.6); 1 (2.6); 8 (1.1); 11 (1.7); 19 (0.9)	0.6594	0.3815	0.0000	0.0000

Tab. VII Cross-validation for combinations of five features (ordered by M).

id	invariants (ρ 's)	all samples	training set		testing set
		M	M_{loa}	err	err
1	1 (0.7); 1 (2.6); 3 (2.8); 11 (0.8); 19 (0.9)	0.6367	0.4123	0.0000	0.0000
2	1 (0.5); 1 (2.6); 9 (1.6); 11 (2.3); 19 (0.9)	0.5515	0.4053	0.0000	0.0000
3	1 (0.7); 1 (2.6); 1 (2.7); 11 (0.8); 19 (0.9)	0.6034	0.3860	0.0000	0.0000
4	1 (0.6); 1 (2.6); 5 (2.1); 11 (0.8); 19 (0.9)	0.5886	0.3851	0.0000	0.0000
5	1 (0.6); 1 (2.6); 8 (1.1); 11 (1.7); 19 (0.9)	0.6594	0.3815	0.0000	0.0000

Tab. VIII Cross-validation for combinations of five features (ordered by M_{loa}).

As the last step, we use Kohonen Self-Organizing Maps (SOM) neural network for cluster forming and relationship visualization with the goal to categorize the patients with minimum errors. Standard hexagonal and rectangular topologies were used for map sizes $(m - 1) \times m$ and $m \times m$ for $m \in \{3, 4, 5, 6\}$. The SOM learning began with 1000 epochs for $\alpha \in \{0.9, 0.8, 0.7, 0.6, 0.5\}$ and $R = m$, and finished with 10000 epochs for $\alpha \in \{0.5, 0.45, 0.4, \dots, 0.1, 0.05\}$ and $R = \lfloor \frac{m}{2} \rfloor$. The learning process was realized with use of the standard software [17] and the categorization quality was measured by mean value of errors reached for 21 training sets and the error for the testing set was obtained by sequel in 21 steps.

The results for five combinations from Tab. VII (for M order) are summarized in Tab. IX. For each topology and map size are specified the least errors on training and testing sets (err_{trn} and err_{tst}) and id's of combinations, that were able to achieve these best results of learning. We can see, that the first combination of features reached the minimum errors for the most of the maps. The particular

topology	map size	err_{trn}	err_{tst}	id's
hexagonal	2×3	0.0072	0.0000	1
	3×4	0.0000	0.0500	1
	4×5	0.0000	0.0500	1
	5×6	0.0000	0.0500	1, 3
	3×3	0.0024	0.0000	1
	4×4	0.0000	0.0000	1
	5×5	0.0000	0.0000	1
rectangular	6×6	0.0000	0.0000	1, 5
	2×3	0.0000	0.0500	1
	3×4	0.0000	0.0500	1
	4×5	0.0000	0.0000	1, 4
	5×6	0.0000	0.0500	1, 3
	3×3	0.0000	0.0500	1
	4×4	0.0000	0.0000	1
	5×5	0.0000	0.0500	1
	6×6	0.0000	0.0000	5

Tab. IX SOM results for M minimization.

results for this best combination are summarized in Tab. X. It is possible to see, that for hexagonal topology we reached best results with map sizes 4×4 (Fig. 5a), 5×5 , and 6×6 , for rectangular topology with map size 4×4 . Sammon's mapping [23]

topology	map size	α_1	α_2	err_{trn}	err_{tst}
hexagonal	2×3	0.60	0.15	0.0072	0.0000
	3×4	0.90	0.50	0.0000	0.0500
	4×5	0.90	0.40	0.0000	0.0500
	5×6	0.90	0.40	0.0000	0.0500
	3×3	0.90	0.30	0.0024	0.0000
	4×4	0.90	0.40	0.0000	0.0000
	5×5	0.80	0.20	0.0000	0.0000
rectangular	6×6	0.70	0.50	0.0000	0.0000
	2×3	0.90	0.15	0.0000	0.0500
	3×4	0.90	0.50	0.0000	0.0500
	4×5	0.90	0.35	0.0000	0.0000
	5×6	0.90	0.05	0.0000	0.0500
	3×3	0.80	0.30	0.0000	0.0500
	4×4	0.90	0.50	0.0000	0.0000
	5×5	0.90	0.45	0.0000	0.0500
	6×6	0.90	0.45	0.0000	0.0500

Tab. X SOM results for I_1 ($\rho = 0.7$, $\rho = 2.6$, $\rho = 2.8$), I_{11} ($\rho = 0.8$), I_{19} ($\rho = 0.9$).

of SOM codebook vectors was performed to \mathbb{R}^2 , using coordinates s_1, s_2 . The corresponding Sammon's mapping for hexagonal topology and map size 4×4 is in Fig. 5b.

The results for five combinations from Tab. VIII (for M_{loa} order) are summarized in Tab. XI. We can see, that the first combination of features reached the zero errors for the most of the maps. Thus, the minimum margin M_{loa} reached across all training sets plays the more important role than the margin M between all AD and ALS patients. The particular results for the first combination are summarized in Tab. XII. It is possible to see, that this combination of features leads to correct classification for all maps except of hexagonal topology maps with sizes $2 \times 3, 3 \times 4,$ and 3×3 . The SOM for rectangular topology map with the size 2×3 is depicted in Fig. 6a. The corresponding Sammon's mapping is in Fig. 6b.

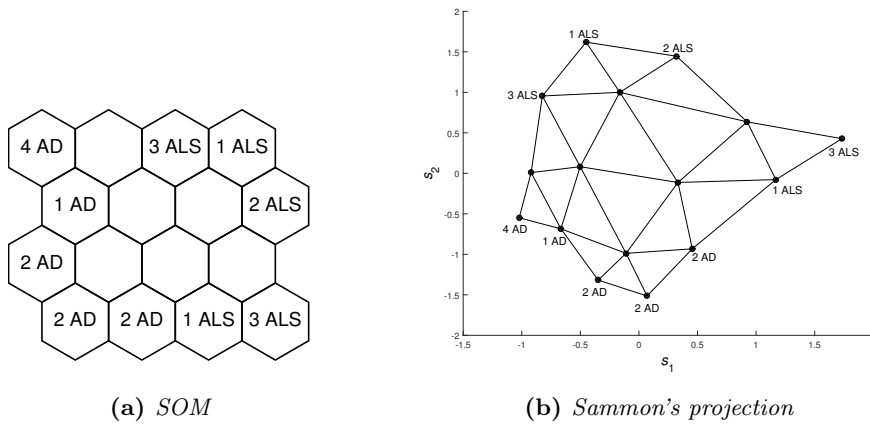


Fig. 5 Results for map of size 4×4 and hexagonal topology, features I_1 ($\rho = 0.7, \rho = 2.6, \rho = 2.8$), I_{11} ($\rho = 0.8$), I_{19} ($\rho = 0.9$).

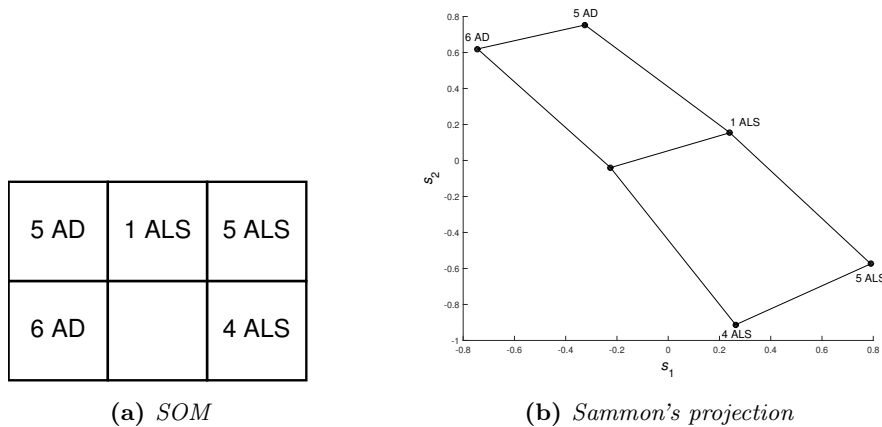


Fig. 6 Results for map of size 2×3 and rectangular topology, features I_1 ($\rho = 0.7, \rho = 2.6$), I_3 ($\rho = 2.8$), I_{11} ($\rho = 0.8$), I_{19} ($\rho = 0.9$).

topology	map size	err_{trn}	err_{tst}	id's
hexagonal	2×3	0.0000	0.0000	2
	3×4	0.0000	0.0000	2
	4×5	0.0000	0.0000	1, 2, 4
	5×6	0.0000	0.0000	1, 2, 4
	3×3	0.0000	0.1052	4
	4×4	0.0000	0.0000	1, 4
	5×5	0.0000	0.0000	1
	6×6	0.0000	0.0000	1, 4, 5
rectangular	2×3	0.0000	0.0000	1
	3×4	0.0000	0.0000	1
	4×5	0.0000	0.0000	1, 3
	5×6	0.0000	0.0000	1, 4
	3×3	0.0000	0.0000	1
	4×4	0.0000	0.0000	1
	5×5	0.0000	0.0000	1, 4
	6×6	0.0000	0.0000	1, 4

Tab. XI SOM results for M_{loa} minimization.

topology	map size	α_1	α_2	err_{trn}	err_{tst}
hexagonal	2×3	0.80	0.45	0.0145	0.0000
	3×4	0.60	0.10	0.0000	0.1052
	4×5	0.70	0.50	0.0000	0.0000
	5×6	0.70	0.35	0.0000	0.0000
	3×3	0.70	0.25	0.0048	0.0500
	4×4	0.70	0.25	0.0000	0.0000
	5×5	0.80	0.10	0.0000	0.0000
	6×6	0.90	0.50	0.0000	0.0000
rectangular	2×3	0.90	0.50	0.0000	0.0000
	3×4	0.90	0.50	0.0000	0.0000
	4×5	0.90	0.50	0.0000	0.0000
	5×6	0.90	0.50	0.0000	0.0000
	3×3	0.90	0.50	0.0000	0.0000
	4×4	0.90	0.50	0.0000	0.0000
	5×5	0.90	0.50	0.0000	0.0000
	6×6	0.90	0.50	0.0000	0.0000

Tab. XII SOM results for I_1 ($\rho = 0.7$, $\rho = 2.6$), I_3 ($\rho = 2.8$), I_{11} ($\rho = 0.8$), I_{19} ($\rho = 0.9$)

5. Conclusion

The study compares the datamining techniques (SVM, SOM) with pure statistical approach (probit, logit, and Cauchy regressions). We obtained fair comparison of mentioned approaches on given data set. The logistic and Cauchy regression lead to statistically significant classifiers meanwhile SVM guarantees linear separability and SOM also exhibit a good separation of AD and ALS patients. Invariance to translation, rotation, scaling, deformation, and radiomarker amount means elimination of measurement conditions in the case of SPECT and PET (Positron Emission Tomography) imaging and analysis. Affine invariant descriptors of 3D image were applied to SPECT scans of human brain activity. Despite of well known principle, that invariant characteristics contain less information then the original data source, we found statistically significant descriptors for the decision, whether the patient has got Alzheimer's disease or amyotrophic lateral sclerosis. Combining six descriptors of radii 0.6, 2.6 and of spherical harmonics orders 1, 3, 5 we obtained regularized logistic classifier between these two forms of dementia. Resulting classifier was tested using leave-one-out technique. Cross-validation error was only 5 % and corresponding p_{value} is 1.83×10^{-4} , which corresponds to already published results of meta-analysis [9]. In the second experiment we combined five descriptors I_1 for radii 0.7 and 2.6, I_3 for radius 2.8, I_{11} for radius 0.8, and I_{19} for radius 0.9, and consequently we obtained the classifier which distinguish AD patients from ALS patients without errors with use of leave-one-out technique. The SOM with rectangular topology and only six "neurons" was used for visualization of relationships between all AD and ALS patients. Corresponding Sammon's mapping proved, that results of learning process are correct. Such a results are promising, however in order to use our approach as a supportive tool for Alzheimer's disease diagnostics from SPECT scans, more tests with broadened data set is needed. Presented method should be applied also on SPECT scans of other brain disorders groups as well as a healthy group.

Acknowledgement

This work was supported by Czech Technical University in Prague, grant SGS17/196/OHK4/3T/14.

References

- [1] AMERICAN PSYCHIATRIC ASSOCIATION. *Diagnostic and statistical manual of mental disorders (DSM-IV-TR)*, 4th ed. Washington, DC. 2000, doi: [10.1176/appi.books.9780890420249.dsm-iv-tr](https://doi.org/10.1176/appi.books.9780890420249.dsm-iv-tr)
- [2] BERKSON J. Application of the logistic function to bio-assay. *Journal of the American Statistical Association*. 1944, 39(227), pp. 357–365, doi: [10.2307/2280041](https://doi.org/10.2307/2280041)
- [3] BLISS C.I. The method of probits. *Science*. 1934, 79(2037), pp. 38–39, doi: [10.1126/science.79.2037.38](https://doi.org/10.1126/science.79.2037.38)
- [4] ČECH P., KUKAL J., ČERNÝ J., SCHNEIDER B., SVOZIL D. Automatic workflow for the classification of local DNA conformations. *BMC Bioinformatics*. 2013, 14(205), doi: [10.1186/1471-2105-14-205](https://doi.org/10.1186/1471-2105-14-205)

- [5] CHANG C.-C., LIN C.-J. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*. 2011, 2(3), pp. 27:1–27:27, doi: [10.1145/1961189.1961199](https://doi.org/10.1145/1961189.1961199)
- [6] CLAUS J.J., HASKAMP F. VAN, BRETELIER M.M.B., KRENNING E.P., KONING I. DE, CAMMEN J.M. VAN DER, HOFMAN A., HASAN D. The diagnostic value of SPECT with Tc 99m HMPAO in Alzheimer's disease: a population-based study. *Neurology*. 1994, 44(3), pp. 454–461, doi: [10.1212/wnl.44.3_part_1.454](https://doi.org/10.1212/wnl.44.3_part_1.454)
- [7] CORTES C., VAPNIK V. Support-Vector Networks. *Machine Learning*. 1995, 20(3), pp. 273–297, doi: [10.1007/bf00994018](https://doi.org/10.1007/bf00994018)
- [8] DEVOUS M.D. Functional brain imaging in the dementias: role in early detection, differential diagnosis, and longitudinal studies. *European Journal of Nuclear Medicine and Molecular Imaging*. 2002, 29(12), pp. 1685–1696, doi: [10.1007/s00259-002-0967-2](https://doi.org/10.1007/s00259-002-0967-2)
- [9] DOUGALL N., NOBILI F., EBMEIER K.P. Predicting the accuracy of a diagnosis of Alzheimer's disease with ^{99m}Tc HMPAO single photon emission computed tomography. *Psychiatry Research: Neuroimaging*. 2004, 131(2), pp. 157–168, doi: [10.1016/j.pscychresns.2003.11.001](https://doi.org/10.1016/j.pscychresns.2003.11.001)
- [10] DOUGALL N.J., BRUGGINK S., EBMEIER K.P. Systematic review of the diagnostic accuracy of ^{99m}Tc-HMPAO-SPECT in dementia. *American Journal of Geriatric Psychiatry*. 2004, 12(6), pp. 554–570, doi: [10.1097/00019442-200411000-00002](https://doi.org/10.1097/00019442-200411000-00002)
- [11] EREMENKO A., JAKOBSON D., NADIRASHVILI N. On nodal sets and nodal domains on S^2 and \mathbb{R}^2 . *Annales de l'Institut Fourier*. 2007, 57(7), pp. 2345–2360, doi: [10.5802/aif.2335](https://doi.org/10.5802/aif.2335)
- [12] GERARDIN E., CHÉTELAT G., CHUPIN M., CUINGNET R., DESGRANGES B., KIM H.-S., NIETHAMMER M., DUBOIS B., LEHÉRICY S., GARNERO L., EUSTACHE F., COLLIOT O. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *NeuroImage*. 2009, 47(4), pp. 1476–1486, doi: [10.1016/j.neuroimage.2009.05.036](https://doi.org/10.1016/j.neuroimage.2009.05.036)
- [13] GUTMAN B., WANG Y., MORRA J., TOGA A.W., THOMPSON P.M. Disease Classification With Hippocampal Shape Invariants. *Hippocampus*. 2009, 19(6), pp. 572–578, doi: [10.1002/hipo.20627](https://doi.org/10.1002/hipo.20627)
- [14] HILBE J. M. *Logistic Regression Models*. Chapman & Hall/CRC Press, 2009.
- [15] HORAISOVÁ K., KUKAL J. Leaf Classification from Binary Image via Artificial Intelligence. *Biosystems Engineering*. 2016, 142, pp. 83–100, doi: [10.1016/j.biosystemseng.2015.12.007](https://doi.org/10.1016/j.biosystemseng.2015.12.007)
- [16] KOHONEN T. Self-Organizing Maps. *Springer Series in Information Sciences*, Springer Berlin Heidelberg. 2001, 30, doi: [10.1007/978-3-642-56927-2](https://doi.org/10.1007/978-3-642-56927-2)
- [17] KOHONEN T., HYNINEN J., KANGAS J., LAAKSONEN J. *SOM_PAK: The self-organizing map program package*. Tech. Rep. A31, Helsinki University of Technology, Laboratory of Computer and Information Science. 1996.
- [18] LÖTJÖNEN J., WOLZ R., KOIKKALAINEN J., JULKUNEN V., THURFJELL L., LUNDQVIST R., WALDEMAR G., SOININEN H., RUECKERT D. Fast and robust extraction of hippocampus from MR images for diagnostics of Alzheimer's disease. *NeuroImage*. 2011, 56(1), pp. 185–196, doi: [10.1016/j.neuroimage.2011.01.062](https://doi.org/10.1016/j.neuroimage.2011.01.062)
- [19] MCKHANN G., DRACHMAN D.A., FOLSTEIN M., KATZMAN R., PRICE D.L., STADLAN E.M. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS–ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*, 1984, 34(7), pp. 939–944, doi: [10.1212/wnl.34.7.939](https://doi.org/10.1212/wnl.34.7.939)
- [20] MCNEILL R., SARE G.M., MANOHARAN M., TESTA H.J., MANN D.M.A., NEARY D., SNOWDEN J.S., VARMA A.R. Accuracy of single-photon emission computed tomography in differentiating frontotemporal dementia from Alzheimer's disease. *Journal of Neurology, Neurosurgery & Psychiatry*. 2007, 78(4), pp. 350–355, doi: [10.1136/jnnp.2006.106054](https://doi.org/10.1136/jnnp.2006.106054)
- [21] MÜLLER C. *Spherical Harmonics*. Lecture Notes in Mathematics, Springer. 1966, 17, doi: [10.1007/bfb0094775](https://doi.org/10.1007/bfb0094775)

- [22] RILEY K.F., HOBSON M.P., BENCE S.J. *Mathematical Methods for Physics and Engineering*. Cambridge University Press, Cambridge. 2006, doi: [10.1017/cco10521861535](https://doi.org/10.1017/cco10521861535)
- [23] SAMMON J.W. A nonlinear mapping for structure analysis. *IEEE Transactions and Computers*. 1969, C-18(5), pp. 401–409, doi: [10.1109/t-c.1969.222678](https://doi.org/10.1109/t-c.1969.222678)
- [24] SAYEED A., PETROU M., SPYROU N., KADYROV A., SPINKS T. Diagnostic features of Alzheimer's disease extracted from PET sinograms. *Physics in Medicine and Biology*. 2001, 47(1), pp. 137–148, doi: [10.1088/0031-9155/47/1/310](https://doi.org/10.1088/0031-9155/47/1/310)
- [25] SHI Y., THOMPSON P.M., ZUBICARAY G.I. DE, ROSE S.E., TU Z., DINOVI I., TOGA A.W. Direct mapping of hippocampal surfaces with intrinsic shape context. *NeuroImage*. 2007, 37(3), pp. 792–807, doi: [10.1016/j.neuroimage.2007.05.016](https://doi.org/10.1016/j.neuroimage.2007.05.016)
- [26] TALBOT P.R., LLOYD J.J., SNOWDEN J.S., NEARY D., TESTA H.J. A clinical role for ^{99m}Tc-HMPAO SPECT in the investigation of dementia? *Journal of Neurology, Neurosurgery & Psychiatry*. 1998, 64(3), pp. 306–313.
- [27] WOOLDRIDGE J.M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press. 2002.
- [28] ZHOU L., LIEBY P., BARNES N., RÉGLADE-MESLIN CH., WALKER J., CHERBUIN N., HARTLEY R. Hippocampal Shape Analysis for Alzheimer's Disease Using an Efficient Hypothesis Test and Regularized Discriminative Deformation. *Hippocampus*. 2009, 19(6), pp. 533–540. doi: [10.1002/hipo.20639](https://doi.org/10.1002/hipo.20639)