

Binary and nonbinary description of hypointensity for search and retrieval of brain MR images

Devrim Unay^a, Xiaojing Chen^b, Aytul Ercil^c, Mujdat Cetin^c, Radu Jasinschi^a, Mark A. van Buchem^d, Ahmet Ekin^a

^a Philips Research, Video Processing Group, HTC 36, 5656 AE, Eindhoven, The Netherlands
^b LIACS, Leiden University, Leiden, Netherlands

^c Faculty of Engineering and Natural Sciences, Sabanci University, Istanbul, Turkey

^d Department of Radiology, Leiden University Medical Center, Leiden, Netherlands

ABSTRACT

Diagnosis accuracy in the medical field, is mainly affected by either lack of sufficient understanding of some diseases or the inter/intra-observer variability of the diagnoses. We believe that mining of large medical databases can help improve the current status of disease understanding and decision making. In a previous study based on binary description of hypointensity in the brain, it was shown that brain iron accumulation shape provides additional information to the shape-insensitive features, such as the total brain iron load, that are commonly used in clinics. This paper proposes a novel, nonbinary description of hypointensity in the brain based on principal component analysis. We compare the complementary and redundant information provided by the two descriptions using Kendall's rank correlation coefficient in order to better understand the individual descriptions of iron accumulation in the brain and obtain a more robust and accurate search and retrieval system.

Keywords: search and retrieval in medical databases, Kendall's rank correlation coefficient, brain MR image analysis, brain iron deposition, hypointense features, principal component analysis, shape-based brain structure detection, segmentation, neurodegenerative diseases

1. INTRODUCTION

Population aging - explained as a shift in the distribution of a country's population towards greater ages - is a process common to the whole world, while more developed countries experience it with more severity. Although aging is associated with living longer and healthier lives, it comes with the burden of increase in age-related chronic diseases, such as cardiovascular problems, neurological diseases and cancer.

In certain neurological diseases, such as Alzheimer's disease, neurodegeneration or progressive loss of neurons occurs and aging is considered to be one of the primary risk factors of developing such a disease. For example, in the U.S. alone, 5.2 million people of all ages have Alzheimer's and more than 400,000 new cases are diagnosed per year. More importantly, this number is expected to more than double and increase to 959,000 new cases a year by 2050.¹ Unfortunately, for most of the neurodegenerative diseases, neither their detection nor their diagnosis is a fully solved problem. Accordingly, researchers from various disciplines increasingly focus on the (early) detection, understanding, and diagnosis of neurodegenerative diseases.

Iron is required to maintain the brain's optimal functioning, and therefore is essential for life.² Its accumulation in the brain is a normal process that starts at the early ages and is observed in every individual. However, in those developing neurodegenerative diseases there is strong evidence that basal ganglia and other regions of the brain accumulate abnormal (larger) amounts of iron.^{3,4}

Magnetic resonance (MR) imaging is a commonly used modality for visualizing brain, thanks to its safety (as it requires no ionizing radiation), its excellent image contrast to discern various tissue types and its ability to obtain different images emphasizing various physiological features. These different images are called contrasts (e.g. T1, T2 and PD) and they are often used in combination by the radiologists in decision-making.

Send correspondence to Devrim Unay at unaydevrim@hotmail.com

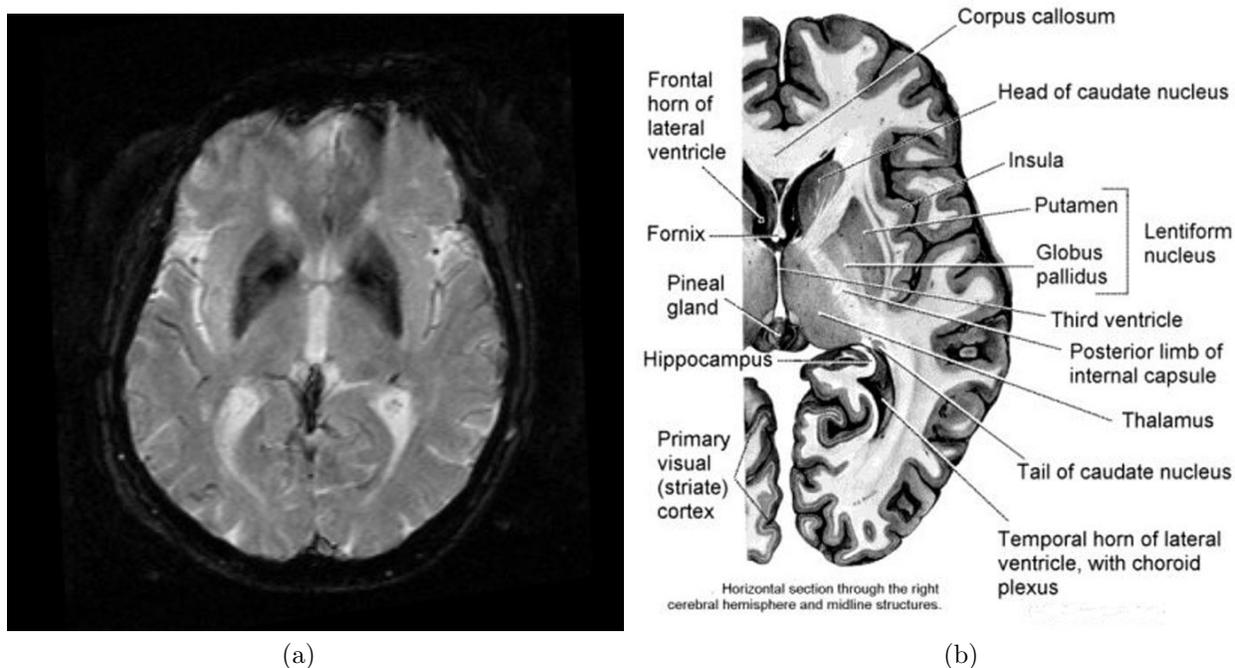


Figure 1. (a) SWI image of a patient with large amount of iron deposition in the basal ganglia, (b) the anatomical information for the approximate location.

It is well-known that presence of iron leads to changes in MR signal, both magnitude and phase. Accordingly, tissues with high iron concentration appear hypo-intense (darker than usual) in MR contrasts, such as T2 and T2*, which focus mainly on the signal magnitude. Susceptibility-weighted imaging (SWI), on the other hand, is a novel MR contrast that exploits magnitude and phase information collectively.^{5,6} Due to the presence of a phase difference between tissues with iron and those without, SWI exhibits enhanced signal difference, and therefore is promising in non-invasive quantification of brain iron by computer-aided tools.

Figure 1(a) demonstrates the iron accumulation effect on a susceptibility-weighted image (SWI) with significantly darker regions in the globus pallidus and the putamen. A corresponding atlas⁷ for approximately the same slice is also shown on the right. The structures relevant for iron accumulation include putamen, globus pallidus, and caudate nucleus, which form the basal ganglia responsible for motor control, cognition, emotions and learning. Any abnormality in this region may lead to devastating disorders, including Parkinson's disease and Huntington's disease.

Our general purpose is to search large medical databases of brain scans by the similarity of brain iron accumulation in the basal ganglia. Such a system will be useful in linking the iron accumulation similarity with neurodegenerative diseases and revealing its relation to various complications, such as high cholesterol or blood pressure. It will also allow the medical experts to compare multiple patients before making a diagnosis, which in turn will reduce the inter/intra-observer variability in the diagnosis. In a previous study, a search and retrieval system for brain MR databases was introduced, where iron accumulation was quantified using a binary hypointensity description approach and it was shown that brain iron accumulation shape provides additional information compared to the shape-insensitive features.⁸ In this paper, we propose a novel nonbinary description of hypointensity based on principal component analysis and study the complementary and similar information provided by the two descriptions. To do this, we compare the values of Kendall's rank correlation coefficient from the results of queries with the hypointensity features of both descriptions.

The following section presents the algorithms for hypointensity feature extraction. In Section 3, we describe the image-based search and retrieval algorithm to find the patients with similar brain iron accumulation to

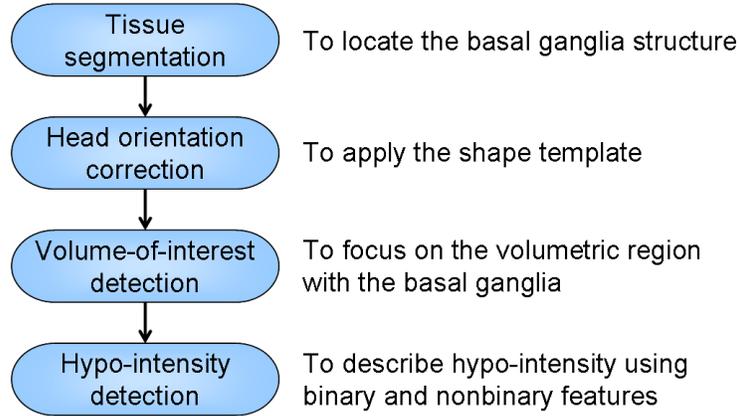


Figure 2. The image processing steps involved.

the current patient and to understand the correlation between the hypointensity features. We provide the experimental results in Section 4. Finally, in Section 5 we conclude the paper.

2. FEATURE EXTRACTION

In this section, we explain semi-automatic detection of hypo-intense voxels in the basal ganglia structure of the brain. Our input is volumetric brain MR data consisting of three contrasts, T2, PD and SWI. T2 and PD were acquired in the same brain scan, while SWI in a separate acquisition, so we performed their spatial alignment to each other using a registration algorithm. Our output is hypo-intense voxel mask in the brain basal ganglia. This involves several image processing steps for various purposes, as displayed in Figure 2. In the following, we explain each of these steps briefly.

2.1. Tissue Segmentation

We use a clustering algorithm to segment the brain MR data into three classes: 1) cerebrospinal fluid (CSF) 2) white matter (WM) and gray matter (GM), and 3) background. Initially, the intensity histogram of the brain tissue from the T2 and PD-weighted data is built. From this histogram, the significant histogram modes are found using the mean-shift algorithm.^{9,10} The most salient three modes are determined and employed for the initialization of the cluster centers.

In clustering, segmentation performance of the K-harmonic means (KHM)¹¹ and the well-known K-means algorithms were compared previously.⁸ It was observed that KHM provided slightly better results at the expense of more computation time. As the segmentation accuracy of K-means is sufficient for our purpose, in this work we favor the use of K-means. Figure 3 displays an exemplary segmentation result obtained by K-means where the CSF region is highlighted in white. Eqn. 1 shows the objective function of K-means, where n is the number of data points, K is the number of clusters, x_i is the feature value of the i^{th} data point, and c_j is the feature value of the j^{th} cluster.

$$KM(X, C) = \sum_{i=1}^n \min \|x_i - c_j\|^2 \quad j \in 1, 2, \dots, K \quad (1)$$

2.2. Head orientation correction

Healthy brain exhibits a rough bilateral symmetry with respect to the interhemispheric fissure, which is commonly known as the anatomical mid-sagittal plane (MSP). The location of the MSP on a 2-D slice is pinpointed as a white line in Figure 4.

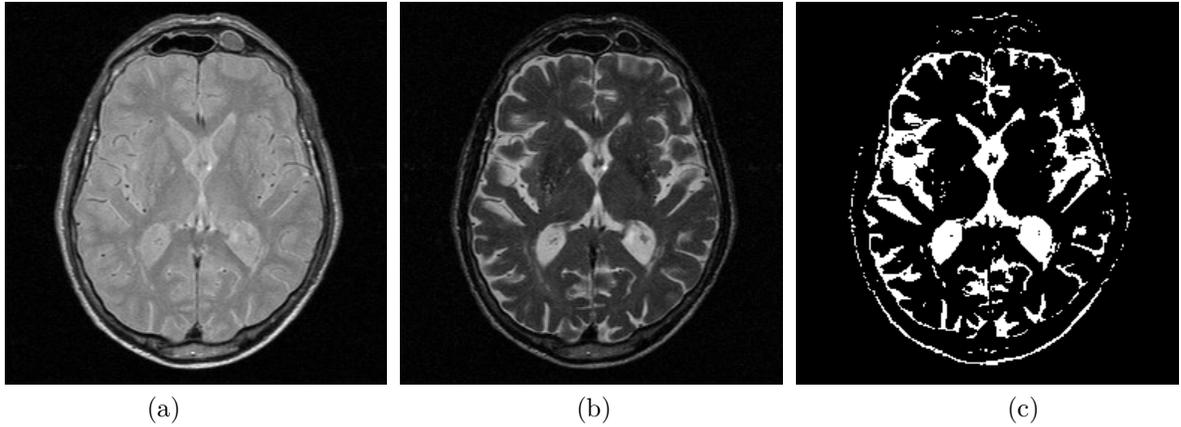


Figure 3. MR dual echo images (PD on the left, T2 in the middle) and the extracted CSF mask (right), T2 has hypo-intense regions due to the abnormally high iron deposition.

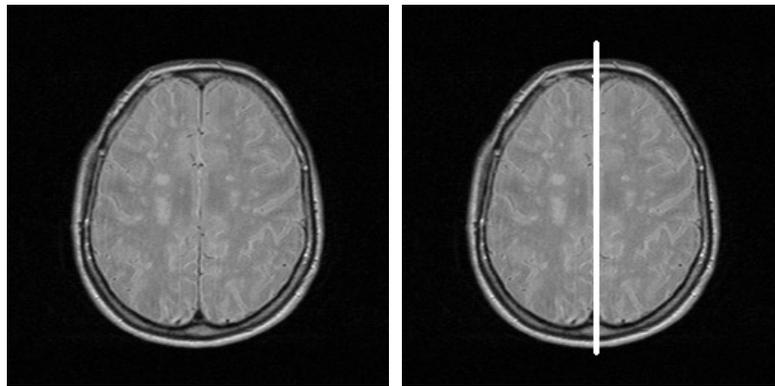


Figure 4. The mid-sagittal plane location for the brain MR slice on the left is shown as a white line on the right.

As the following volume-of-interest detection step in our application assumes a standard head orientation in the images, we detect MSP using a fast feature-based algorithm and correct for the head orientation. The algorithm extracts feature points from each slice, fits a line to these points by RANSAC, and fuses the lines to detect the MSP. A detailed explanation of the algorithm can be found elsewhere.¹²

2.3. VOI detection by landmark shape characteristics

The organs we are interested in (e.g. putamen, globus pallidus, and caudate nucleus) are visible in three or four MR slices at the T2 and PD contrasts. Therefore, we define these slices as the VOI (volume-of-interest) whose length is automatically computed as a function of slice thickness and inter-slice gap. Observing that the shape of the lateral ventricle (filled with CSF) in the frontal lobe is consistent across a large database of individuals, we automatically detect the VOI using a shape-based algorithm, where the lateral ventricle shape is approximated by the V-shaped region between the triangles in Figure 5.

Prior to using the proposed template, we correct the head orientation (by rotation and translation using bilinear interpolation) with respect to the estimated mid-sagittal plane. Afterwards, the size of the template is tailored to the head size and it is overlaid onto each slice (Figure 6). The similarity between the template and each slice is defined as the total number of CSF voxels in the V-shaped region divided by those outside the region. Finally, the VOI consists of a consecutive number of slices whose total similarity value to the template is the maximum.

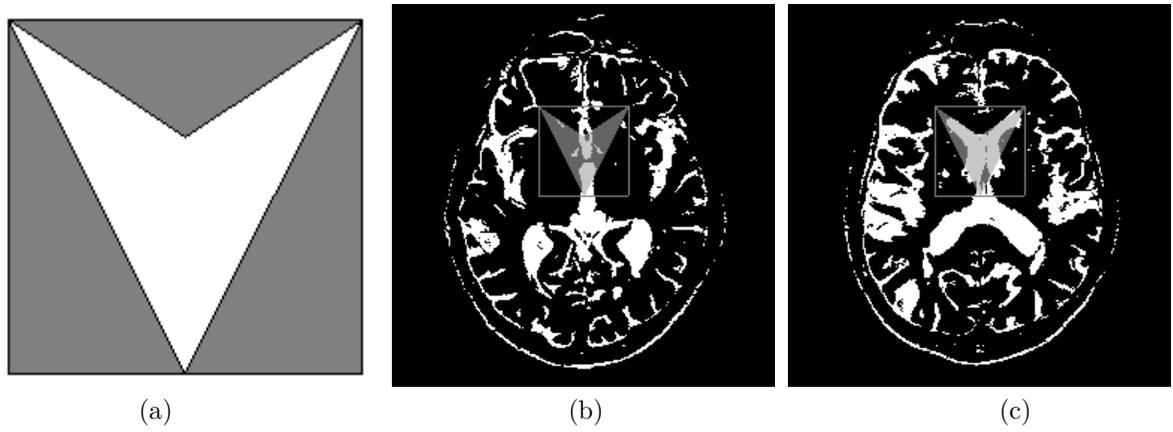


Figure 5. (a) The approximate template of the CSF shape for the frontal lobe of the lateral ventricle, (b) shape model overlaid onto a slice outside the VOI, and (c) on a slice in the VOI.

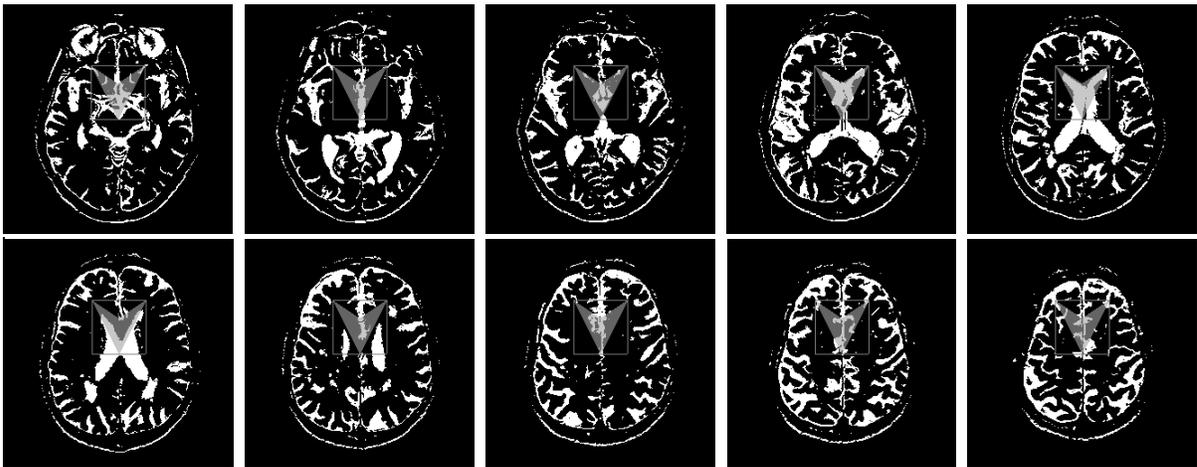


Figure 6. The shape template overlaid on the CSF mask of the MR axial slices.

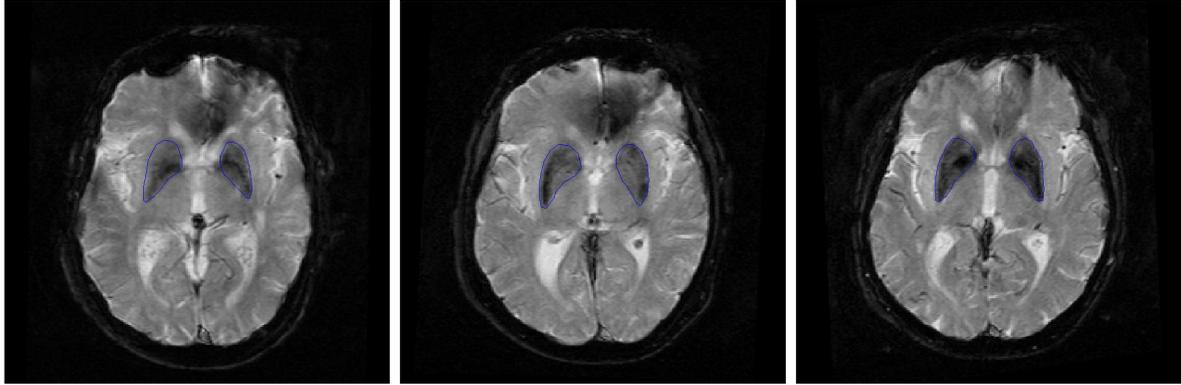


Figure 7. Manual delineation of the basal ganglia structure (SWI images).

2.4. Structure localization and hypointensity detection

In this section, we describe localization of the basal ganglia structure and the detection of hypointense voxels in the VOI. For the localization of basal ganglia, previously we used a Talairach brain atlas based approach that tessellates the minimum bounding rectangle of the brain tissue in each slice into rectangular regions.^{8,13} Accordingly, certain grid locations of the atlas indicated the position of basal ganglia structures. Our observations, however, showed that with this atlas-based approach the boundary regions introduced inconsistency to the detection, most probably due to the imperfect structure localization. Therefore, in this work we opted for manual delineations of the basal ganglia by an expert. These delineations, of which some examples are shown in Figure 7, are used as region-of-interests (ROIs) for the following hypointensity detection step. Please note that, the proposed method can be fully-automated by replacing manual delineation by an automatic and accurate structure localization technique. In fact, there exists some recent preliminary work on shape-based automatic segmentation of basal ganglia structures (see e.g.¹⁴). We intend to employ such techniques in our framework in the future.

2.4.1. Binary Hypointensity Detection

In MRI the absolute intensity value does not have any quantitative meaning unlike some of the other medical imaging modalities, such as X-ray and CT (computer tomography). One way to deal with non-standard values is to normalize the intensity. Accordingly, we normalized the ROIs by subtracting the mean from intensity data I and then dividing the result by the mean, as shown in Eqn. 2.

$$I_{norm} = \frac{I - \text{mean}(I)}{\text{mean}(I)} \quad (2)$$

From this normalization, we obtain a variation range of normalized intensity values as well as the distribution of the values. In the next step, an adaptive threshold is generated for hypointensity quantification, as explained below.

First we select several values in this variation range, considering them as thresholds and apply them to all MR images. In our application, pixels which have values less than the threshold are regarded as hypointense (darker), and hypointensity is defined as percentage of hypointense pixels in an ROI. In this way, each threshold gives one hypointensity result for all MR images, and then we sort these results in descending order to get a ranking. Next, a comparison between the rankings and an expert's clustering result is carried out. The expert has grouped the database into three clusters: dark, medium and light, according to the darkness of the basal ganglia. In order to compare, every ranking result is also divided into three clusters, the ones at the top of the ranking are categorized as light while the ones at the bottom as dark, and for each cluster the number of included subjects is the same as in the expert's corresponding cluster. Given these two clustering results, fractions of true

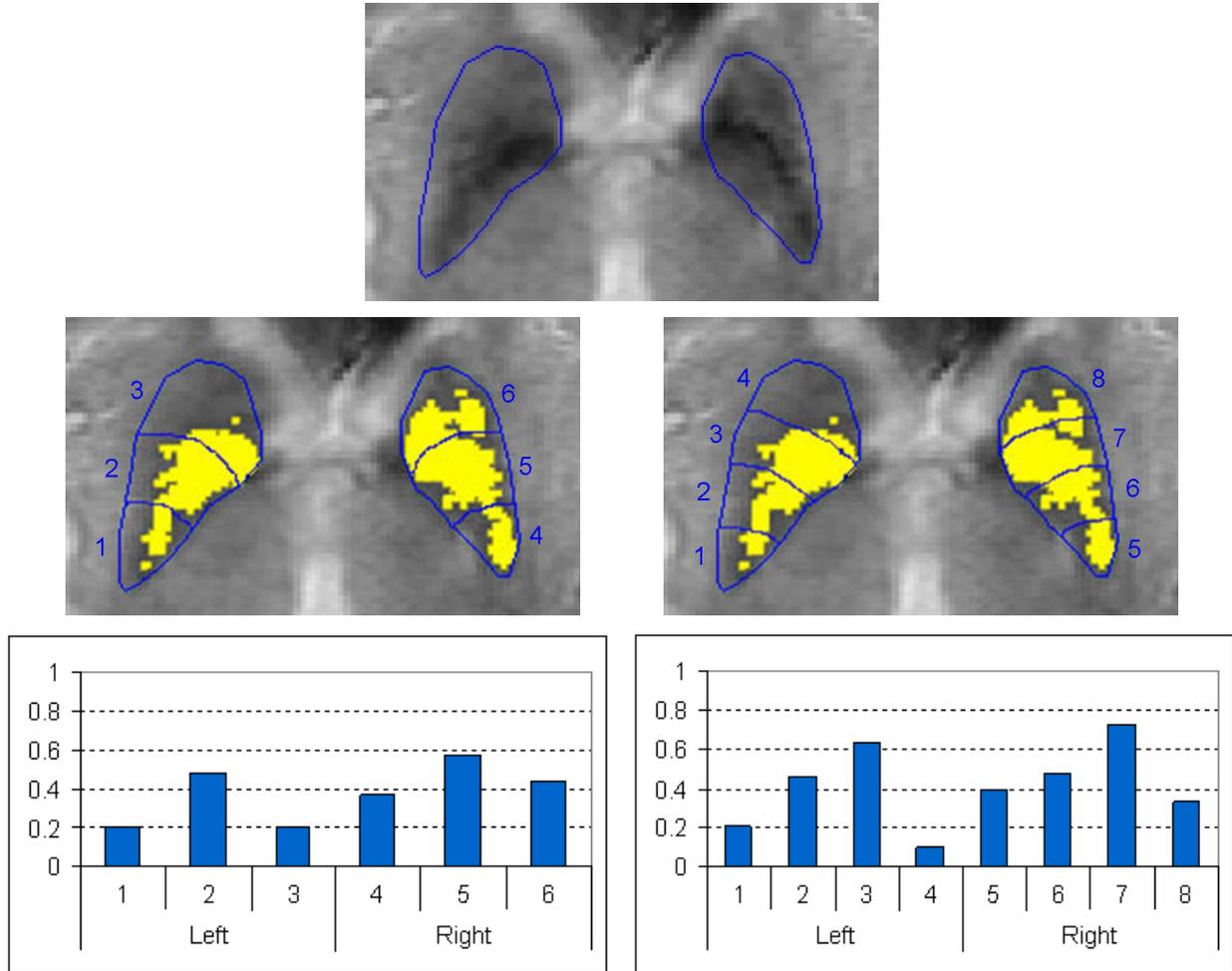


Figure 8. Binary hypointensity detection example. Top: Zoomed SWI image of one basal ganglia delineation. Middle: Same basal ganglia tessellated into radially equidistant subregions (left: $N=3$, right: $N=4$) and hypointense voxels highlighted. Bottom: Corresponding subregion hypointensities displayed in a graph.

positives (TPR) and false positives (FPR) are computed to estimate the agreement between each ranking and expert's clustering. In the end, the threshold which has the highest TPR and the lowest FRP is chosen and the hypointensities defined by this threshold are assigned as the final binary hypointensity features.

Furthermore, in order to capture the spatial distribution of iron in basal ganglia we tessellate the ROIs as shown in Fig.8. Each hemispheric part of the ROI is divided into N radially equidistant subregions with the outermost posterior voxel selected as the central point. Please note that, we define radially equidistant subregions by concentric circles with $(r_{i+1} - r_i) = \Delta r$, where Δr is constant and i is a positive integer. Exemplary tessellations for $N = 3$ and $N = 4$, and the corresponding subregion hypointensities are displayed in Fig. 8.

2.4.2. Nonbinary Hypointensity Detection

The binary hypointensity detection approach explained above has the following two drawbacks: 1) optimum tessellation is unknown, and 2) the features proposed are correlated with each other. In order to overcome these issues, we propose to use principal component analysis (PCA), also known as Karhunen-Loève transform, which is a vector space transformation often used to reduce multidimensional datasets to lower dimensions for analysis.¹⁵

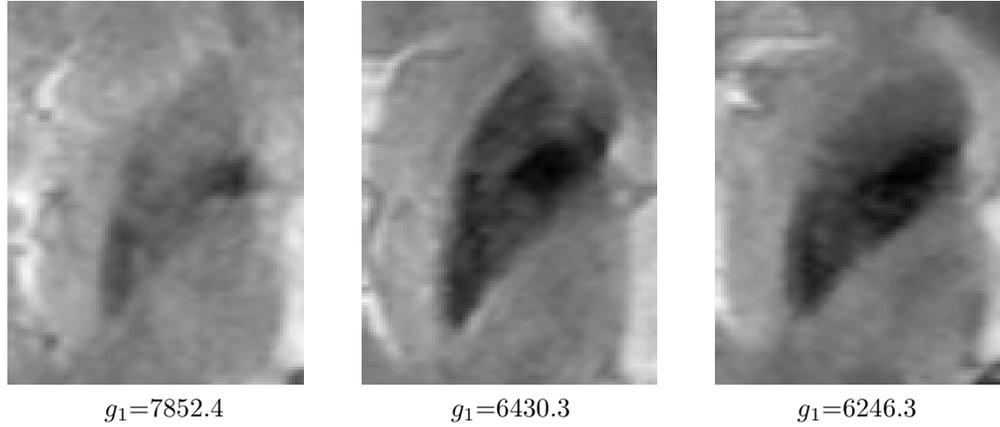


Figure 9. Exemplary SWI images with varying hypointensities in the basal ganglia and the corresponding eigenspace projection values, g , from the first principal component.

Given data X consisting of N images, in PCA we first perform data normalization by subtracting the mean vector m from the data. Then the covariance matrix Σ of the normalized data $(X - m)$ is computed.

$$m = \frac{1}{N} \sum_{i=1}^N X_i \quad (3)$$

$$\Sigma = (X - m)(X - m)^T \quad (4)$$

Afterwards, the basis functions are obtained by solving the algebraic eigenvalue problem

$$\Lambda = \Phi^T \Sigma \Phi \quad (5)$$

where Φ is the eigenvector matrix of Σ , and Λ is the corresponding diagonal matrix of eigenvalues. Consequently, projection of a new data Y to the eigenspace is achieved by

$$G = \Sigma^T (Y - m) \quad (6)$$

where $G = (g_1, g_2, \dots, g_N)$ and $\Sigma = (e_1, e_2, \dots, e_N)$ with g_i being a scalar value representing the degree-of-match between the image and the eigenvector e_i . Finally, data (image) X_j can be reconstructed exactly by weighted sum of the eigenvectors.

$$X_j = \sum_{j=1}^N g_j e_j + m \quad (7)$$

Our initial observations showed that images of basal ganglia with varying hypointensities have distinctive eigenspace projections, as shown in Fig. 9. Therefore, we define nonbinary hypointensity features of an image as the corresponding eigenvectors from each principal component.

3. FEATURE CORRELATION ANALYSIS WITH SEARCH AND RETRIEVAL

In this paper we aim to determine the correlation between the proposed binary and nonbinary hypointensity features. Accordingly, we propose to use Kendall's rank correlation coefficient, also known as Kendall's tau, which measures the correlation between two rankings¹⁶ as a non-parametric statistical measure. In the domain of search and retrieval, the returns to various queries are often represented as the (similarity) ranks of the items in the database, and therefore Kendall's tau is a very convenient way to compare the performance of search and retrieval methods. Higher values of Kendall's tau show that the two ranks are correlated; hence, in our case, the features that have higher correlation value between each other possess similar information.

Kendall's tau between feature A and B is computed as follows:

For each patient in the database used as query

1. Select one of the features as feature A and another one as feature B
2. Find the ranked list (Rank A) of the returns to the query with feature A
3. Find the ranked list (Rank B) of the returns to the query with feature B
4. Match Rank B with Rank A to find the respective position of same database items. For example,

Patients	a	b	c	d	e	f	g	h
Rank by Feature A	1	2	3	4	5	6	7	8
Rank by Feature B	3	4	1	2	5	7	8	6
5. For each item in the Rank B row, compute the number of entries on its right that have higher rank and sum these numbers to find the P value
6. For each item in the Rank B row, compute the number of entries on its right that have lower rank and sum these numbers to find the Q value
7. Having N as the total number of entries (eight in the above example), compute tau as:

$$\tau = \frac{P - Q}{\frac{1}{2} \times N \times (N - 1)} \quad (8)$$

Finally, for each feature pair we repeat the above steps and assign the average of the individual values computed as the corresponding Kendall's correlation value.

4. RESULTS

4.1. Feature Extraction

We have tested the proposed algorithms on a brain MRI dataset of more than 600 patients. The proposed set of feature extraction algorithms have so far generated excellent results that have been verified visually. For quantification, we extracted a subset of this dataset and verified the performance of the proposed algorithms. In this set, we included 37 subjects with atherosclerotic risk factors. MRI was performed on a Philips Intera 1.5T whole body scanner at Leiden University Medical Center. Resulting volumetric data was composed of 48 slices with 256×256 pixels for T2 and PD, and 22 slices with 256×256 pixels for SWI.

The performance of the CSF segmentation, mid-sagittal plane detection and VOI detection stages were previously evaluated on the same dataset, where the algorithm reached 100% detection accuracy.⁸

4.2. Feature Correlation Analysis

Previously we have defined the binary hypointensity features as the percentages of hypointense pixels in each basal ganglia subregion, and the nonbinary hypointensity features as the eigenvectors of principal component analysis. Given this, the reader may notice that in both cases the number of features depend on the tessellation size: 1) In the binary case, tessellation is done spatially and results in hypointensities from different number of subregions. 2) In the nonbinary case, tessellation is performed at the eigenspace and defines the number of eigenvectors kept. Hence, we repeated the experiments with varying tessellation sizes and computed Kendall's correlation coefficient between each pair of the corresponding binary and nonbinary hypointensity features. The following subsections present the results and our analysis.

4.2.1. Single Tessellation

In order to analyze the relationships between individual features we first introduce the results of a single tessellation. Fig. 10 visualizes the resulting correlation measurements for the binary, nonbinary and binary versus nonbinary features that are computed from the tessellation size of 10 using heat maps - a data visualization technique based on color. Regarding the binary features, we observe medium-strength (in the range of [0.4-0.7]) correlation between neighboring features (or subregions), and this correlation weakens as the proximity between features increases. Nonbinary feature results show weak (below 0.4) correlations between eigenvectors, which is consistent with the fact that PCA finds a transformation that uncorrelates the variables. When we compare binary features with the nonbinary ones we see that the first nonbinary feature, which is the eigenvector with the largest variance, exhibits medium correlation with binary features, while the rest have weak figures.

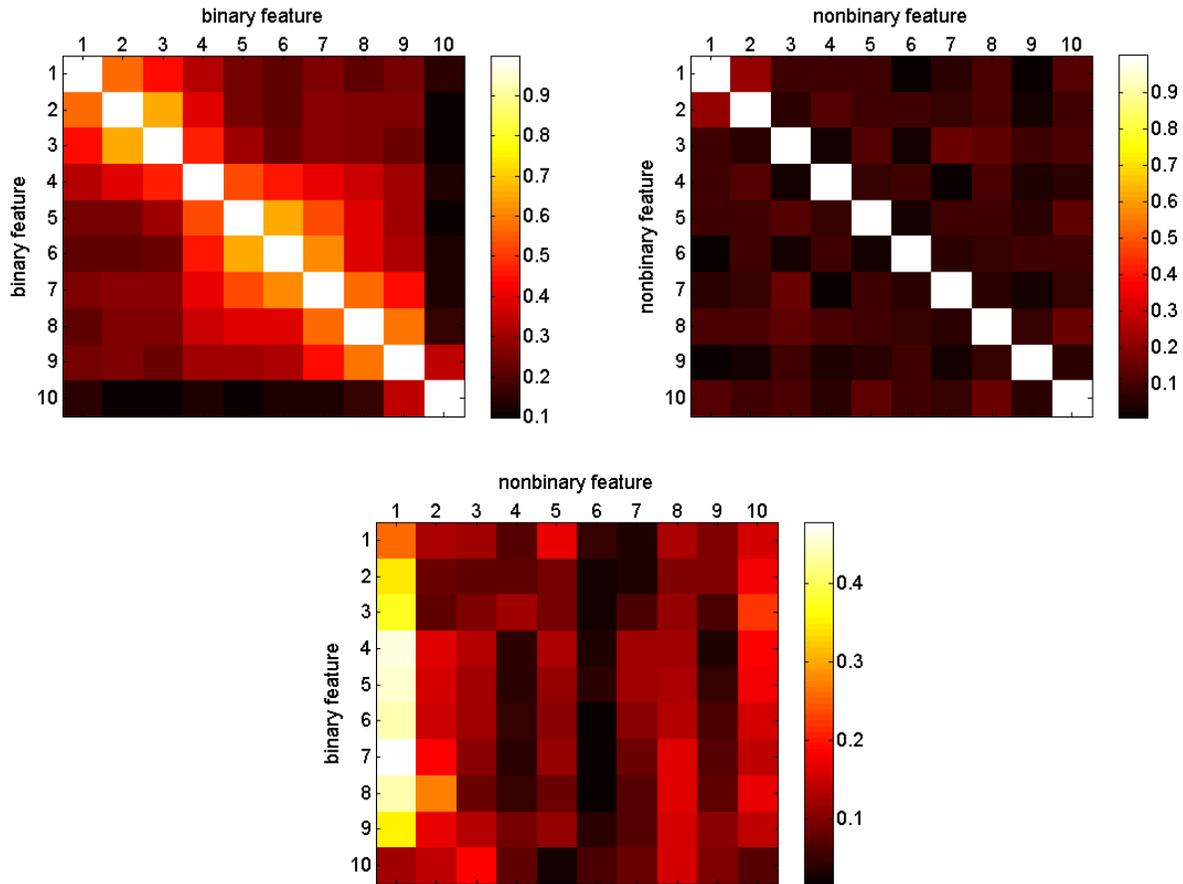


Figure 10. Heat maps of Kendall’s correlation values for binary (top-left), nonbinary (top-right) and binary versus nonbinary (bottom) features. Tessellation size is 10.

4.2.2. Multiple Tessellations

“Does the tessellation size affect correlations?” To answer this question, in this subsection we present the results of multiple tessellations. Contrary to the previous subsection, here we use combination of features to describe hypointensity and refer to them as descriptions (e.g. binary description=4 refers to a 4-dimensional vector of hypointensities computed from 4 basal ganglia subregions, or nonbinary description=7 means 7 eigenvectors with highest variance are kept).

Fig. 11 visualizes the corresponding averaged correlation values over multiple runs for the binary-only, nonbinary-only and binary versus nonbinary descriptions, respectively. The analysis reveals that within-description (binary-only and nonbinary-only) correlations are strong (above 0.7), while between-description (binary versus nonbinary) correlations have medium-strength. Furthermore, in the latter, we observe that tessellation size has no considerable effect on the results, meaning spatial and eigenspace tessellations interrelate with each other. We conclude that combination of the two descriptions may provide a better representation of hypointensity. Hence further evaluation is needed to find the optimal tessellation and combination of descriptions.

5. CONCLUSION

In this work we presented novel methods to describe hypointensity in the basal ganglia regions of brain MR images and evaluate the brain iron deposition as well as a search and retrieval mechanism over a large database. In

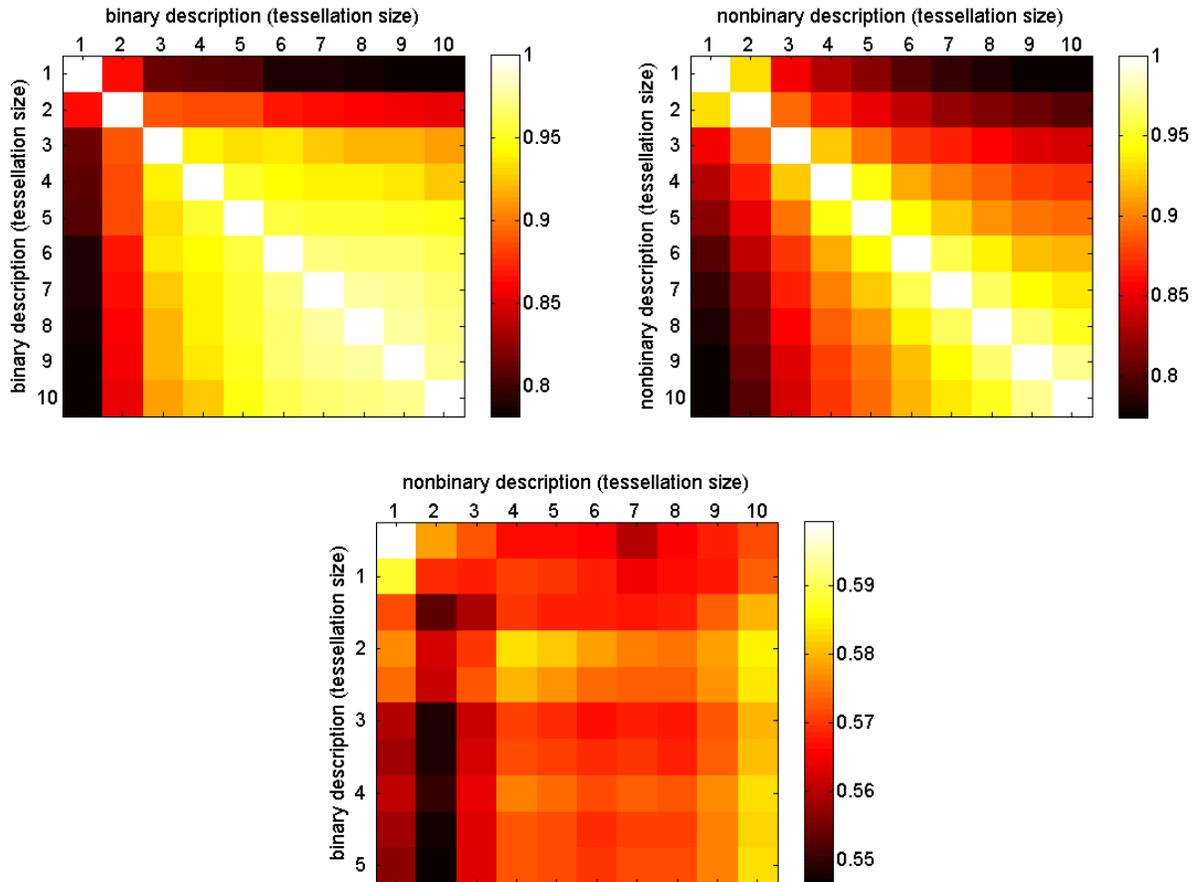


Figure 11. Heat maps of Kendall's correlation values for binary (top-left), nonbinary (top-right) and binary versus nonbinary (bottom) descriptions at varying tessellations.

a previous study based on binary description of hypointensity in the brain, it was shown that brain iron accumulation shape provides additional information to the shape-insensitive features, such as the total brain iron load, that are commonly used in clinics. In this paper we propose a novel, nonbinary description of hypointensity in the brain based on principal component analysis. We analyzed the proposed hypointensity features by computing Kendall's correlation values from the returns to large sets of queries. The analysis revealed high within-feature (binary-only and nonbinary-only) correlations, while the between-feature (binary versus nonbinary) correlations were at medium strength. Furthermore, we observed that the number of features has no considerable effect on the results, meaning binary (shape-related) and nonbinary (eigenspace-related) descriptions correlate with each other. In the future, we want to investigate the optimum tessellation for each description and evaluate their combination based on a ground truth.

REFERENCES

1. "2008 Alzheimer's disease facts and figures," 2008. Alzheimer's Association.
2. E. Madsen and J. D. Gitlin, "Copper and iron disorders of the brain," *Annual Review of Neuroscience* **30**, pp. 317–337, 2007.
3. J. F. Schenck and E. A. Zimmerman, "High-field magnetic resonance imaging of brain iron: birth of a biomarker?," *NMR in Biomedicine* **17**, pp. 433–445, 2004.

4. S. L. Harder, K. M. Hopp, H. Ward, H. Neglio, J. Gitlin, and D. Kido, "Mineralization of the deep gray matter with age: a retrospective review with susceptibility-weighted MR imaging," *American Journal of Neuroradiology* **29**, pp. 176–183, 2008.
5. E. M. Haacke, Y. Xu, Y.-C. N. Cheng, and J. R. Reichenbach, "Susceptibility weighted imaging (SWI)," *Magnetic Resonance in Medicine* **52**, pp. 612–618, 2004.
6. B. Thomas, S. Somasundaram, K. Thamburaj, C. Kesavadas, A. K. Gupta, N. K. Bodhey, and T. R. Kapilamoorthy, "Clinical applications of susceptibility weighted MR imaging of the brain - a pictorial review," *Diagnostic Neuroradiology* **50**, pp. 105–116, 2008.
7. [online] at <http://brainmind.com/BrainMaps5.html>.
8. A. Ekin, R. Jasinschi, E. Turan, R. Engbers, J. van der Grond, and M. van Buchem, "Search and retrieval of medical images for improved diagnosis of neurodegenerative diseases," 2007. Proc. SPIE Electronic Imaging.
9. K. Fukunaga, *Introduction to statistical pattern recognition*, Academic Press, Boston, 1990.
10. D. Comaniciu and P. Meer, "Robust analysis of feature spaces: color image segmentation," in *Proc. IEEE ICPR*, pp. 750–755, 1997.
11. B. Zhang, "Generalized k-harmonic means - boosting in unsupervised learning," in *Technical Report HPL-2000-137*, 2000.
12. A. Ekin, "Feature-based brain mid-sagittal plane detection by ransac," in *Proc. EURASIP EUSIPCO*, Sept. 2006.
13. R. Jasinschi, A. Ekin, A. van Es, J. van der Grond, M. van Buchem, and A. van Muiswinkel, "Automatic detection and classification of hypointensity in mr images of the brain," in *14th Scientific Meeting of Int. Society for Magnetic Resonance in Medicine (ISMRM)*, Sept. 2006.
14. G. Uzunbas, M. Cetin, G. B. Unal, and A. Ercil, "Coupled nonparametric shape priors for segmentation of multiple basal ganglia structures," in *5th IEEE Int. Symposium on Biomedical Imaging (ISBI)*, pp. 217–220, May 2008.
15. B. Moghaddam, "Principal manifolds and probabilistic subspaces for visual recognition," *IEEE Trans. PAMI* **24**(6), pp. 780–788, 2002.
16. P. Armitage, G. Berry, and J. N. S. Matthews, *Statistical Methods in Medical Research*, Blackwell Publishing, 4th ed., 2002.