IEEE *Access*
Multidisciplinary : Rapid Review : Open Access Journal

Longitudinal Brain MRI Retrieval for Alzheimer's Disease Using Different Temporal Information

KATARINA TROJACHANEC, (Member, IEEE), **IVAN KITANOVSKI, IVICA DIMITROVSKI, AND SUZANA LOSHKOVSKA, (Senior Member, IEEE), for the Alzheimer's Disease Neuroimaging Initiative***

Faculty of Computer Science and Engineering, Saints Cyril and Methodius University of Skopje, 1000 Skopje, Macedonia

Corresponding author: Katarina Trojachanec (katarina.trojacanec@finki.ukim.mk)

ABSTRACT This paper describes the research made toward improving medical case retrieval for Alzheimer's Disease (AD). Our approach considers using Magnetic Resonance Images as an input for the search. To improve the retrieval process, we used longitudinal information extracted from the different sets of scans acquired at different time points and automatically extracted descriptors to represent input images. All experiments were performed with and without quality control (QC) to determine the influence of the errors caused by the automated processing to the results relevance. For the experiments, a total of 267 subjects from the AD Neuroimaging Initiative database with available scans at baseline, the 6-month, 12-month, and 24-month follow-ups were selected. The obtained results showed that the selection of the time points for extraction of the longitudinal information influences the retrieval performance. Results also showed that not all automatically generated descriptors lead to improvement of the results. Longitudinal volume changes provide the most relevant representation. Adding QC phase in the experiments leads to improvements in all examined scenarios. The results showed that the most frequent automatically selected features are common semantic markers for AD.

INDEX TERMS Longitudinal, Alzheimer's disease, image retrieval, feature extraction, brain, magnetic resonance imaging.

I. INTRODUCTION

Alzheimer'S disease (AD) is one of the most frequent neurodegenerative diseases in older adults nowadays. It leads to a nerve cells degeneration and tissue loss as the disease progresses. Until now, no cure has been found for the disease. However, finding relevant biomarkers, reaction to a certain therapy, monitoring the progression of the disease, early disease detection and prognosis are active research fields [1].

Advances in medicine and genetics, as well as the rapid evolution of technology and medical imaging techniques increase the amount of generated data for the AD medical cases. For each medical case, variety of data types are produced, including medical images of the brain, genetic markers, cognitive tests results, blood and cerebrospinal fluid biomarkers. They all need to be efficiently organized, stored, and represented in order to provide easier and more appropri-

ate access to the medical cases and to enable use of those data in the research for knowledge discovery. This could improve the diagnostic and therapeutic processes [2], [3].

A great part of the generated data for AD cases consists of medical images. Considering that the information extracted from the brain image provide precise and consistent markers for diagnosis and monitoring the development of the disease, our research is focused on medical case retrieval by using medical images as input queries [4]–[10].

The image retrieval process requires representing the image query with a descriptor and then comparing it with the descriptors of all of the images stored in the medical database. The system returns a sorted list of the database images with the most similar at the top [11].

Most of the current systems apply the approach of describing the image using their visual characteristics [12]. This strategy is facing with three major limitations. First, such representation does not have semantic relevance for the image, which results in imprecise retrieval that can be irrelevant and impractical in the clinical and research centers [2], [13]. Second, the dimension of the generated feature vector might be very high [14]–[17], reducing the efficiency of the retrieval.

Thus, finding the most appropriate and efficient image representation is crucial for providing semantically relevant result. Third, having in mind that AD is characterized by progressive neurodegeneration, the brain changes over time are even more descriptive than the brain condition at a single occasion. This can provide additional information for the retrieval process.

The paper presents the results of the research done to overcome these limitations and hence, to improve medical case retrieval for AD. Therefore, we used the domain knowledge to add the semantical meaning of the visual image content. Additionally, to provide more precise and clinically more relevant answer from the system, we decided to use the longitudinal rather than cross-sectional representation.

The objectives of the study are to: (1) test the hypothesis that the features generated using an appropriate set of time points to extract the longitudinal information will improve the overall relevance of the retrieval response; (2) identify the most relevant and efficient longitudinal changes (semantic descriptors) to represent the subjects/cases (e.g. rate of change, percent change, etc.); and (3) test the hypothesis that the errors arising from the image processing influence the retrieval performance.

For that purpose, we examined image representations comprised of longitudinal changes, such as rate of change (RC), percent change (PC), and symmetrized percent change (SPC) of volume and cortical thickness of the brain structures. They are meant to reflect the severity of the disease and the advance of the degeneration [18], unlike the raw volume or cortical thickness measurements representing the static state at a specific time point of degeneration [19]. To calculate these longitudinal changes, we created an unbiased within-subject template space and image [20] using robust, inverse consistent registration [21] with automated longitudinal processing stream [22] in FreeSurfer. Additionally, we built the template with different sets of time points in separate scenarios and evaluated the aforementioned type of feature vectors derived from the different templates separately.

Automated processing might result in failures, possibly affecting the retrieval process. To examine to which extent they influence the results, we performed all experiments with and without quality control (QC).

To evaluate the strategy, we selected subjects from AD and Normal Controls group with scans at four available time points from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu).

The benefits from this research are multifold. First, the retrieval process is based on rich longitudinal information rather than information extracted from one specific occasion. Results showed that different sets of time points contribute differently to the relevance of the retrieval response. Second, we recommend the most appropriate subject representation for case retrieval based on imaging markers from the evaluated pool of longitudinal measurements. According to the results, we identified volume percent change and volume symmetrized percentage as most promising for

representation, yet with comparable descriptor dimension with the other cases. In this way, the retrieval is directed towards answering the questions of the type "find all subjects/cases that have similar longitudinal changes in the brain anatomy to the query". From the medical point of view, this is very beneficial because the retrieved top similar cases might empower physicians with rich information covering the similar disease progression, treatment reaction, and consistency of the imaging markers as the disease advances. Third, the influence of the possible errors due to the automated processing was evaluated. The obtained results showed significant improvement by involving the QC phase. All these findings are toward semantically more relevant retrieval result. Providing this is very important from the clinical, as well as from the research and educational perspective.

We provide the related work in Section II. The methods used in this study are presented in Section III, whereas the experimental results are shown in Section IV. We discuss the results in Section V and present the concluding remarks in Section VI.

II. RELATED WORK

Several studies address image retrieval applied to AD data. However, they challenge different aspects of the process. Most of them follow the traditional feature extraction techniques, deriving features from the visual cues in the images [14], [16], [17], [23]. Intensity histograms, local binary pattern, and gradient magnitude histograms are used in [23] to generate feature vector for the middle slice. Global and local texture features, such as Discrete Cosine Transform (DCT), Daubechie's Wavelet Transform (DWT), and Local Binary Patterns (LBP), are applied on a selected subset of slices in [14], [16]. Laguerre Circular Harmonic Functions expansions are used to capture the local image patch structure in [17]. Then, they apply the Bag-of-Visual-Words to a specific region of interest, the hippocampus. The feature extraction in this research is performed in a slice-by-slice manner. The main drawback encountered in these studies is the retrieval semantics. Even more, because the research is performed on 2D bases on one or multiple slices, possibly significant spatial information might be excluded. Authors in [24] present attempts to improve the retrieval semantics by explicitly using the domain knowledge. They represent the images with subcortical volumes and cortical thickness of the brain regions improving results with fewer features.

The main common limitation of the previous studies is performing the retrieval on the bases of the brain structural condition at a specific time point, i.e. cross-sectionally. On the opposite, the neurodegenerative diseases, among which is AD, are characterized with a specific pattern of atrophy that can be monitored and reflected in a longitudinal manner [25]–[28]. Existing research based only on two time points showed that data obtained by these time points might not be sufficient and reliable enough to represent brain change due to the disease progression [29]. Additionally, extract-

ing the information from longitudinal data with independently cross-sectional processing of the images may provide longitudinally inconsistent estimates leading to unreliable results [30]. To overcome this limitation, research is going toward specialized solutions using multiple time points [19]. They use structural brain information, such as volume of subcortical structures or cortical thickness, at multiple time points or a combination of them. Taking into consideration that such features have a sort of a static nature reflecting the brain condition at each available time point, there is still possibility to investigate and improve results obtained by this strategy.

The evolution of all these strategies appears in the input question for the retrieval system. They start from the question "find all subjects/cases with similar visual features to the query" when using cross-sectional traditional feature extraction methods. Then, they go through "find all subjects/cases with similar structural characteristics to the query" in the case of cross-sectional studies, using the domain knowledge for feature extraction. They finally end with "find all subjects/cases that have similar anatomical changes to the query" in the research based on longitudinal structural features.

This paper focuses on the last strategy for the input question aiming to improve medical case retrieval.

III. MATERIALS AND METHODS

A. LONGITUDINAL DATA AND PROCESSING

We based our study on longitudinal data for patients that underwent examination for Alzheimer's disease. Longitudinal data is obtained by measuring the outcome variables repeatedly on the same cohort of individuals at multiple time points, ordered in time. This enables identifying the influence of the changes in the measurements over time to the examined clinical, biological, or experimental factor. Considering the examined disease, which is a non-stationary continuous process, those changes might reflect the diseases progress and/or reaction to the treatment. Additionally, longitudinal studies provide direct assessment of within-subject changes across different time points, free of any between-subject variability, more precisely and without confounding cohort effects [31], [32].

Reliable estimations of volume and thickness and their changes over time are extracted by automatically processing all the images with the FreeSurfer' longitudinal stream [22].

At first, we processed cross-sectionally all the available time points. This processing includes motion correction, removal of non-brain tissue using a hybrid watershed/surface deformation procedure, automated Talairach transformation, segmentation of the subcortical white matter and deep grey matter volumetric structures, intensity normalization, tessellation of the grey matter - white matter boundary, automated topology correction, and surface deformation following intensity gradients. After computing the cortical models, registration to a spherical atlas was applied. It utilizes individual cortical folding patterns to match cortical geometry across subjects, ending with the parcellation, labeling, and statistics.

Then, we generated an unbiased within-subject template space and image [20] using robust, inverse consistent registration [21] for previously independently processed time points. The longitudinal scheme was designed to be unbiased with respect to any time point (TP). There was no initialization with information from a specific time point. A template volume was created and run through FreeSurfer treating all time points the same. Then, we longitudinally processed each time point. This step used information from the subject-template and from the individual runs to initialize several algorithms. These included skull stripping, Talairach transforms, atlas registration, as well as spherical surface maps and parcellations. In this way, reliability and statistical power were significantly increased. After automated segmentation and quantification, several types of measurement were available for analysis, including volume, cortical thickness, surface area, etc. [22].

To evaluate whether the selected time points included in the template generation influence the appropriateness and reliability of the derived measurements in the retrieval, we created the template with different sets of time points. In this examination, we did not consider templates with two time points, basically because the template was generated as a voxel-wise median, which is equal to mean in the case of two time points. This results in blurry edges, instead of crispy in the case of more time points included in the template. Using intensity median instead of mean enables to remove the influence of outliers, such as scans with strong motion artifacts. Thus, more time points lead to more stable templates meaning that more reliable and accurate measurements are expected [22]. We considered sets of three and four time points and examined the following four scenarios:

- T123 – template generated using TP1, TP2, and TP3
- T134 – template generated using TP1, TP3, and TP4
- T234 – template generated using TP2, TP3, and TP4
- T1234 – template generated using TP1, TP2, TP3, and TP4.

TP1 denotes the first time point available for the patient -namely, the scan acquired at the first hospital visit (the baseline). The others represent the follow-ups, where TP2 is 6 months later, TP3 – one year later, and TP4 is two years later with respect to the baseline. The templates T123 and T134 are equally spaced with six months difference between the time points in the first one and 12 months difference in the second one. Although the T234 and T1234 are differently spaced, the time spacing is (approximately) the same across the subjects in all the scenarios.

After the fully automated longitudinal processing, ROI measurements, such as volume, cortical thickness, surface area, etc. are available. Among them, volumes and cortical thickness estimates are found to be valuable imaging markers [33], [34]. On the other hand, research showed that the surface area has significantly low performance [34] and it is not considered for further analysis in this paper.

Plain subcortical volume measurements and cortical thickness estimates were used in the similar setup in the previous

work [19]. However, they reflect the static state at specific time point of degeneration, whereas longitudinal changes such as rate of change and percent change represent the speed and severity of degeneration. Considering this, we reduced the temporal data within each subject to a single statistic (e.g. annualized atrophy rate or percent change) of the subcortical and cortical volumes and cortical thickness, and subsequently used in the experiments.

B. FEATURE VECTOR GENERATION AND FEATURE SELECTION

The image retrieval uses feature vector that represents the image given as a query, and then compares it with other feature vectors for all the images stored in the image base.

We created the feature vector using the longitudinal information extracted from the images. We considered the following statistics calculated using FreeSurfer:

- **Rate of change (RC)** - the difference per time unit. In fact, the slope of a linear fit across time within each subject is computed. Depending on the time variable this will yield the volume loss in mm^3/time or thinning in mm/time for each brain region. In this study, we measured the time variable in years.
- **Percent change (PC1/fit)** – percent thinning/volume loss per year. It is calculated as the rate normalized by the measure at the first time point times 100 (e.g. percent tinning per year). It represents the amount of percent thinning/volume loss at a given region. In fact, in this research we used the PC1/fit. It was calculated with respect to the value obtained from the linear fit at baseline, which is more reliable and noiseless than the baseline value directly.
- **Symmetrized percent change (SPC)** - the rate with respect to the temporal average (instead of taking it from the first time point). The temporal average is computed from the linear fit at the middle of the time interval. Symmetrized percent change is more robust measure than PC1. The reason is that the thickness at time point 1 is noisier than the average or an outlier. Moreover, SPC is symmetric. Considering this, SPC is recommended to be used in the analyses [22].

On the basis of this, we generated six types of feature vectors:

Type 1: volume RC (VolumesRC) – a statistic derived from the volumetric temporal information per case for each region. We calculated RC for 55 measures obtained from the subcortical segmentation and 68 from the cortical parcellation (34 from each hemisphere, left and right), 123 measures in total for each case;

Type 2: cortical thickness RC (CTRC) – a statistic derived from the cortical thickness temporal information per subject for each region, 70 measures in total (35 for the regions of each hemisphere, left and right, together with the mean thickness for each hemisphere);

Type 3: volume PC1fit (VolumesPCfit) – a statistic derived from the volumetric temporal information per subject for each subcortical region, 123 measures in total for each case;

Type 4: cortical thickness PC1fit (CTPCfit) – a statistic derived from the cortical thickness temporal information per subject for each region, 70 measures in total;

Type 5: volume SPC (VolumesSPC) – a statistic derived from the volumetric temporal information per subject for each subcortical region, 123 measurements in total for each case;

Type 6: cortical thickness SPC (CTSPC) – a statistic derived from the cortical thickness temporal information per subject for each region, 70 measures in total.

It should be noted that regarding the available volumetric measurements, we excluded the volume of the fifth ventricle, left and right non-white matter hypointensities, because they were not available for most or for all the subjects. We considered the regions for which the estimates were available for the left and right hemisphere separately without averaging or summing them.

To analyze the power and semantical appropriateness of the longitudinal subject representation, we derived all the six feature vector types based on the four listed templates. However, to provide fair and clear evaluation, each experiment was based on the same conditions (included time points/feature vector type) for all the subjects.

We also applied feature subset selection to reduce the feature vector dimensionality and to select the most relevant features. In fact, we used the Correlation-based Feature Selection (CFS) method [35], due to its powerful influence on the retrieval in the previous studies [19], [24]. It is based on evaluation of the subsets of features, considering the usefulness of the individual features for predicting the class along the degree of intercorrelation among them. This means that good feature subsets should contain features highly correlated with the class, yet uncorrelated to each other [35]. Considering the application domain, we expected features sensitive to the disease to be selected.

C. IMAGE RETRIEVAL

When a query is given to the image retrieval system, it compares the query to all other subjects in the database. As a result, it generates a sorted list of cases based on the subjects' similarity to the query, with the most similar one at the top.

We represented the subjects with features vectors generated using the strategy described in the previous subsections. We performed the feature selection independently of the query subject information to provide an unbiased result. Hence, we obtained a specific feature subset for each query subject using the information of all other subjects in the database. To examine the stability of the selected features, we recorded the inclusion rate, i.e. how frequent each feature was selected.

Considering the small number of subjects included in this study, we used leave-one-out strategy. This means that each subject representation was used as a query against all other representations stored in the database. The similarity measurement between the feature vectors was performed using Manhattan distance. We selected Manhattan distance on the experimental bases, because it provided

TABLE 1. Longitudinal ADNI sample demographics.

Variable	Diagnostic group	
	AD	NL
Subjects (no.)	99	168
Gender (male/female no.)	52/47	85/83
Education (years)	14.91 (2.95)	15.98 (2.85)
TP1 (Baseline)		
Age (years)	75.32 (7.46)	76.03 (5.10)
MMSE Score	23.17 (1.94)	29.19 (0.95)
Global CDR	0.77 (0.25)	0.00 (0.00)
NPI-Q Total Score	3.42 (3.25)	0.30 (0.80)
FAQ Total Score	13.40 (7.00)	0.15 (0.64)
TP2 (month 6)		
Age (years)	75.88 (7.46)	76.60 (5.11)
MMSE Score	22.32 (3.23)	29.08 (1.00)
Global CDR[a]	0.89 (0.27)	0.02 (0.09)
NPI-Q Total Score	3.71 (3.78)	0.57 (1.26)
FAQ Total Score	15.65 (7.49)	0.23 (0.79)
TP3 (year 1)		
Age (years)	76.38 (7.46)	77.10 (5.11)
MMSE Score	20.97 (4.27)	29.18 (1.15)
Global CDR[b]	0.98 (0.47)	0.03 (0.13)
NPI-Q Total Score	5.04 (4.86)	0.63 (1.48)
FAQ Total Score[c]	17.62 (6.88)	0.21 (0.98)
TP4 (year 2)		
Age (years)	77.41 (7.46)	78.12 (5.12)
MMSE Score[d]	18.88 (5.62)	29.10 (1.14)
Global CDR[e]	1.28 (0.62)	0.05 (0.16)
NPI-Q Total Score	4.87 (4.79)	0.71 (1.57)
FAQ Total Score[f]	20.34 (6.78)	0.38 (1.26)

Abbreviations: AD = Alzheimer's disease group, NL = Normal Controls group, TP = Time Point, MMSE = Mini-Mental State Examination, CDR = Clinical Dementia Rating, NPI-Q = Neuropsychiatric Inventory Questionnaire, FAQ = Functional Activities Questionnaires.

The education, age, MMSE Score, Global CDR, NPI-Q Total Score and FAQ Total Score are given in mean (standard deviation).

Due to missing data, sample size for the labeled variables is reduced to:
[a] 167 for NL
[b] 165 for NL
[c] 167 for NL
[d] 97 for AD
[e] 96 for AD and 166 for NL
[f] 98 for AD and 165 for NL.

TABLE 2. Timing of scans per time point by clinical group.

Time Point	Time from baseline (years)	
	AD	NL
TP1 (Baseline)	0	0
TP2 (month 6)	0.57 (0.05)	0.58 (0.06)
TP3 (year 1)	1.07 (0.05)	1.08 (0.07)
TP4 (year 2)	2.09 (0.14)	2.1 (0.11)

Abbreviations: TP = Time Point.
The time from baseline is given in mean (standard deviation).

retrieval systems that return more relevant cases at the top of the list;

- Precision at first 1 (P1) – precision of the first (top) returned case;
- Precision at first 5 (P5) – precision of the first (top) five returned cases;
- Precision at first 10 (P10) – precision of the first (top) 10 returned cases;
- Precision at first 20 (P20) - precision of the first (top) 20 returned cases;
- Precision at first 30 (P30) - precision of the first (top) 30 returned cases;
- R-precision (RP) – precision at first (top) R returned subjects, where R is the total number of relevant cases.

The retrieved case is assumed to be relevant if the patient has the same diagnosis as the query one. The higher the relevant cases in the retrieved list, the higher the value of the precision is.

D. PARTICIPANTS AND INCLUSION CRITERIA

Data used in this study were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The data contains serial magnetic resonance images (MRI), positron emission tomography (PET) images, other biological markers, such as cerebrospinal fluid (CSF) markers, APOE status and full-genome genotyping via blood sample, as well as clinical and neuropsychological assessments. This information can be combined to measure the progression of mild cognitive impairment and early Alzheimer's disease that has been the primary goal of ADNI. The database contains data about cognitively normal individuals, adults with early or late Mild Cognitive Impairment, and people with early AD with different follow-up duration of each group, specified in the protocols for ADNI-1, ADNI-2, and ADNI-GO (see http://www.adni-info.org).

To evaluate the presented strategy, we selected a total of 267 subjects from ADNI-1 standardized lists by using the following criteria: (1) each subject belongs to AD or NL group; (2) for each subject, scan at baseline (TP1) and the 6-month (TP2), 12-month (TP3), and 24-month (TP4) follow-ups are available. Table 1 provides subject demographics for

better retrieval results among several other similarity measures including Canberra, Euclidean, Chebyshev distance, and Cosine similarity.

To evaluate the examined strategies, we used standardized evaluation metrics for quantitative measurement of the retrieval performance:

- Mean Average Precision (MAP) – the mean of the average precision scores for each query, evaluation metric for the general retrieval performance. It is meant to favor

TABLE 3. Value of MAP for the retrieval based on the measurements extracted using the template T123.

Feature type	Without QC		With QC	
	MAP	RP	MAP	RP
T123_VolumesRC	0.69	0.65	0.75	0.72
T123_VolumesPCfit	0.70	0.64	0.75	0.71
T123_VolumesSPC	0.70	0.64	0.76	0.71
T123_CTRC	0.62	0.59	0.69	0.66
T123_CTPCfit	0.65	0.61	0.70	0.66
T123_CTSPC	0.65	0.61	0.70	0.67

Abbreviations: MAP = mean average precision; RP = precision at first (top) R returned subjects, where R is the total number of relevant subjects; QC = quality control; RC = rate of change; PCfit = percent change with respect to the value obtained from the linear fit at baseline; SPC = symmetrized percent change; CT = cortical thickness.

Feature type is denoted as T123_X, where T123 stands for the within-subject template generated using the longitudinal processing using the scans at time point 1, 2, and 3; X is the feature vector based on RC/PCfit/SPC of volumetric measurements and cortical thickness respectively.

TABLE 4. Value of MAP for the retrieval based on the measurements extracted using the template T134.

Feature type	Without QC		With QC	
	MAP	RP	MAP	RP
T134_VolumesRC	0.74	0.70	0.81	0.77
T134_VolumesPCfit	0.78	0.72	0.83	0.78
T134_VolumesSPC	0.78	0.72	0.83	0.78
T134_CTRC	0.69	0.64	0.74	0.70
T134_CTPCfit	0.72	0.66	0.76	0.72
T134_CTSPC	0.72	0.67	0.77	0.72

Abbreviations: MAP = mean average precision; RP = precision at first (top) R returned subjects, where R is the total number of relevant subjects; QC = quality control; RC = rate of change; PCfit = percent change with respect to the value obtained from the linear fit at baseline; SPC = symmetrized percent change; CT = cortical thickness.

Feature type is denoted as T134_X, where T134 stands for the within-subject template generated using the longitudinal processing using the scans at time point 1, 3 and 4; X is the feature vector based on RC / PCfit / SPC of volumetric measurements and cortical thickness respectively.

TABLE 5. Value of MAP for the retrieval based on the measurements extracted using the template T234.

Feature type	Without QC		With QC	
	MAP	RP	MAP	RP
T234_VolumesRC	0.72	0.68	0.78	0.75
T234_VolumesPCfit	0.74	0.69	0.80	0.76
T234_VolumesSPC	0.74	0.69	0.80	0.76
T234_CTRC	0.65	0.61	0.73	0.70
T234_CTPCfit	0.68	0.63	0.74	0.71
T234_CTSPC	0.67	0.63	0.74	0.71

Abbreviations: MAP = mean average precision; RP = precision at first (top) R returned subjects, where R is the total number of relevant subjects; QC = quality control; RC = rate of change; PCfit = percent change with respect to the value obtained from the linear fit at baseline; SPC = symmetrized percent change; CT = cortical thickness.

Feature type is denoted as T234_X, where T234 stands for the within-subject template generated using the longitudinal processing using the scans at time point 2, 3, and 4; X is the feature vector based on RC/PCfit/SPC of volumetric measurements and cortical thickness respectively.

TABLE 6. Value of MAP for the retrieval based on the measurements extracted using the template T1234.

Feature type	Without QC		With QC	
	MAP	RP	MAP	RP
T1234_VolumesRC	0.76	0.71	0.82	0.78
T1234_VolumesPCfit	0.79	0.73	0.84	0.80
T1234_VolumesSPC	0.78	0.73	0.84	0.79
T1234_CTRC	0.69	0.66	0.77	0.74
T1234_CTPCfit	0.73	0.67	0.79	0.75
T1234_CTSPC	0.73	0.68	0.79	0.75

Abbreviations: MAP = mean average precision; RP = precision at first (top) R returned subjects, where R is the total number of relevant subjects; QC = quality control; RC = rate of change; PCfit = percent change with respect to the value obtained from the linear fit at baseline; SPC = symmetrized percent change; CT = cortical thickness.

Feature type is denoted as T1234_X, where 1T234 stands for the within-subject template generated using the longitudinal processing using the scans at time point 1, 2, 3 and 4; X is the feature vector based on RC / PCfit / SPC of volumetric measurements and cortical thickness respectively.

the analyzed sample. The timing of scans per time point by clinical group is reported in Table 2.

As explained, we processed all subjects with FreeSurfer image analysis suite. After inspecting the processed data, we encountered global or regional failures in some of them. The subjects with such failures in at le ast one time point were detected. According to the previous research performed on the same dataset [36], the quality control was recommended on the automatically processed data. Due to the lack of a medical expert and to ensure the fully automated processing, we excluded these cases and performed the examination on the selected cohort that consists of 153 subjects, 41 AD and 112 NL.

IV. EXPERIMENTAL RESULTS

In this section, we summarize the results from the evaluation of the presented strategy. The values of the evaluation metrics MAP and RP are grouped by template scenarios. Table 3 contains the results in the case of template T123. The evaluation scores considering the template T134 are presented in Table 4. Table 5 lists MAP and RP values for the template T234, whereas the results regarding the template T1234 are summarized in Table 6. These tables contain results in both cases, with and without QC.

The measurements extracted from the longitudinal processing using TP1, TP3, and TP4 provide the best results for the cases that include temporal information from three time points. Both cases that include the last two time points, the case based on T134 and T234, gave better results than the case that uses T123. The same conclusions were derived in both situations, with and without QC (Table 3-5).

We noticed improvement of the MAP and RP values when we based the retrieval process on the information extracted

TABLE 7. Precision at specific point in the case with QC.

FeatureType	P1	P5	P10	P20	P30
T123_VolumesRC	0.78	0.78	0.77	0.77	0.76
T123_VolumesPCfit	0.76	0.79	0.78	0.78	0.77
T123_VolumesSPC	0.79	0.79	0.79	0.78	0.76
T123_CTRC	0.76	0.72	0.71	0.70	0.69
T123_CTPCfit	0.74	0.74	0.72	0.71	0.70
T123_CTSPC	0.74	0.74	0.73	0.71	0.71
T134_VolumesRC	0.84	0.84	0.85	0.83	0.81
T134_VolumesPCfit	0.89	0.87	0.86	0.85	0.83
T134_VolumesSPC	0.87	0.87	0.86	0.84	0.83
T134_CTRC	0.78	0.76	0.76	0.76	0.75
T134_CTPCfit	0.80	0.80	0.79	0.78	0.77
T134_CTSPC	0.78	0.79	0.79	0.79	0.77
T234_VolumesRC	0.85	0.80	0.80	0.79	0.78
T234_VolumesPCfit	0.85	0.82	0.81	0.81	0.80
T234_VolumesSPC	0.84	0.82	0.81	0.81	0.80
T234_CTRC	0.79	0.74	0.75	0.73	0.73
T234_CTPCfit	0.80	0.75	0.76	0.75	0.74
T234_CTSPC	0.80	0.75	0.76	0.75	0.74
T1234_VolumesRC	0.84	0.85	0.84	0.83	0.82
T1234_VolumesPCfit	0.90	0.88	0.87	0.86	0.84
T1234_VolumesSPC	0.90	0.87	0.86	0.85	0.84
T1234_CTRC	0.82	0.80	0.79	0.77	0.77
T1234_CTPCfit	0.82	0.82	0.81	0.80	0.79
T1234_CTSPC	0.85	0.82	0.81	0.80	0.79

Abbreviations: P1 = precision of the first (top) returned subject, P5 = precision of the first (top) 5 returned subjects, P10 = precision of the first (top) 10 returned subjects, P20 = precision of the first (top) 20 returned subjects, P30 = precision of the first (top) 30 returned subjects, QC = quality control

TABLE 8. Precision at specific point in the case without QC.

FeatureType	P1	P5	P10	P20	P30
T123_VolumesRC	0.73	0.74	0.73	0.72	0.72
T123_VolumesPCfit	0.81	0.77	0.76	0.74	0.74
T123_VolumesSPC	0.81	0.76	0.75	0.74	0.73
T123_CTRC	0.63	0.66	0.65	0.64	0.63
T123_CTPCfit	0.63	0.68	0.68	0.68	0.67
T123_CTSPC	0.64	0.68	0.68	0.68	0.67
T134_VolumesRC	0.78	0.79	0.79	0.78	0.77
T134_VolumesPCfit	0.86	0.84	0.83	0.83	0.82
T134_VolumesSPC	0.85	0.83	0.83	0.82	0.82
T134_CTRC	0.78	0.73	0.73	0.73	0.72
T134_CTPCfit	0.75	0.76	0.76	0.75	0.75
T134_CTSPC	0.76	0.75	0.75	0.75	0.75
T234_VolumesRC	0.73	0.77	0.76	0.75	0.75
T234_VolumesPCfit	0.81	0.78	0.78	0.78	0.78
T234_VolumesSPC	0.78	0.77	0.78	0.77	0.77
T234_CTRC	0.67	0.68	0.68	0.67	0.66
T234_CTPCfit	0.72	0.71	0.71	0.70	0.70
T234_CTSPC	0.73	0.70	0.70	0.70	0.69
T1234_VolumesRC	0.81	0.81	0.81	0.80	0.80
T1234_VolumesPCfit	0.85	0.84	0.83	0.83	0.82
T1234_VolumesSPC	0.87	0.84	0.83	0.83	0.82
T1234_CTRC	0.73	0.73	0.73	0.72	0.72
T1234_CTPCfit	0.77	0.77	0.77	0.76	0.76
T1234_CTSPC	0.78	0.78	0.77	0.76	0.76

Abbreviations: P1 = precision of the first (top) returned subject, P5 = precision of the first (top) 5 returned subjects, P10 = precision of the first (top) 10 returned subjects, P20 = precision of the first (top) 20 returned subjects, P30 = precision of the first (top) 30 returned subjects, QC = quality control

from four time points (Table 6). Particularly, there was a general improvement over the scenarios based on T123 and T234 in all of the cases on the bases of MAP. This stands for all experiments, with and without QC. However, comparing the cases based on T134 and T1234, the MAP value was the same when the subjects were represented by the volumetric SPC (0.78) or by cortical thickness rate of change (0.69) in the cases without QC. On the other hand, in the experiments with QC, the measurements extracted from four time points provided higher MAP values than those with three time points in all of the cases. The representations based on four time points led to higher RP in all of the cases in general.

Additionally, Table 7 (for the case with QC) and Table 8 (regarding the case without QC) contain values of the precision at a specific point. In each table, the results are grouped by the template type. In the case with QC (Table 7), the precision at specific point calculated to evaluate the retrieval based on four time points, outperforms all of the cases with three time points based on T123 and T234, except one case in the

scenario with QC. That is the P1 provided by the feature type VolumesRC based on TP2, TP3, and TP4. It is 1% higher than in the case based on the template T1234.

Regarding the comparison between the templates T1234 and T134, the results are comparable, even though the template T1234 provides higher precision in most of the cases. In fact, regarding the cases based on volume rates derived from the TP1, TP2, TP3, and TP4 from one side, and TP1, TP3, and TP4 from the other side, there were situations in which the three time points led to the same precision (for VolumesRC and VolumesSPC) and 1% better for P10 on the bases of T134_VolumesRC. This stands for the cases with QC (Table 7). In the experiments without QC, the situation is similar in the cases of feature types VolumesPCfit, VolumesSPC, and CTRC, for which we noticed the same or 1% higher precision value in the cases that included TP1, TP3, and TP4 than in the cases based on four time points. For all the other situations, the longitudinal measurements extracted by using temporal information from four time points provided

better retrieval performance on the bases of precision at a specific level.

Considering the QC, it is evident that in all the cases in general it leads to significantly improved retrieval performance on the bases of MAP and RP. Regarding the precision at specific point, the precision at level 1 is higher in the experiments without QC, in the case when volumetric PCfit and SPC are derived from the time points 1, 2, and 3, or the same in the case based on CTRC generated from TP1, TP3, and TP4. But, it is significantly faster dropping through the other levels than in the experiments with QC. In all the other cases, we noticed higher precision at a specific point in the scenarios with QC than in those without QC.

When analyzing the feature vector type, we came to the conclusion that the volume changes provide better results than the cortical thickness changes, on the bases of the general retrieval performance evaluation metric MAP, as well as the R precision (Table 3-6). The best MAP value in general (0.84) was calculated in two cases, based on T1234_VolumesPCfit and T1234_VolumesSPC in the experiments with QC. The T1234_VolumesPCfit also led to the best RP value (0.80). In the experiments without QC, we also calculated the highest MAP when we used VolumesPCfit features, whereas we obtained the highest RP in the cases of VolumesPCfit and VolumesSPC, all derived from the four time points.

We gained the best P1 precision (0.9) when the QC phase was included in the scenario based on volume symmetrized percent change and volume percent change, whereas in the case without QC - in the scenario based on VolumesSPC. This was obtained when we used the longitudinal information from all four time points.

In the cases where we used cortical thickness changes to generate the feature vector, the scenarios based on T1234_CTPCfit and T1234_CTSPC led to the highest MAP value of 0.79. These scenarios also led to the highest RP from all scenarios based on cortical thickness changes. However, these results are worse than those obtained when using volumetric based measures. We also came to this conclusion considering the detailed precision at a specific level.

After the feature selection, a comparable number of features were selected as the most relevant between the examined feature vector types, yet significantly smaller in comparison to the original feature vector dimension. In fact, 19–21 features were selected in most of the queries in the volumes-based scenarios (feature type 1, 3, and 5) that provide the best results. Considering the CT-based scenarios (feature type 2, 4, and 6), 12–14 features were selected in most of the queries. Even though the number of the selected features in most of the cases based on volumetric rates is slightly higher than in the cases of cortical thickness rates, it is still minor increase considering the level of improvement of the retrieval results.

V. DISCUSSION

Our examination provides a good insight of the appropriateness and efficiency of longitudinal information used in the retrieval process. Regarding the information extracted

during the longitudinal processing based on three time points, the cases that incorporate later time points provide better retrieval results. The rationale behind this is that the last scans are acquired when the disease is in more advanced stage and the brain changes are prominent. Retrieval based on four time points led to general improvements comparing to the cases based on three time points. This is reasonable because the information derived from the within-subject template using more time points is more stable and reliable [22]. However, we noticed comparable results with the scenarios based on the template T134. Thus, according to the obtained results, we recommend using as much time points as possible to provide more robust and precise measurements. We believe that the equal time spacing between the time points, together with the presence of the later scans, make the information based on TP1, TP3, and TP4 more suitable than the other three time points based information and comparable to the four time points based scenarios. Thus, in the situation when the processing time or the number of scans are limited, the processing based on TP1, TP3, and TP4 is advisable.

When choosing whether to use volume or thickness changes, the experiments showed that volume-based changes are preferable. They globally provide better retrieval results, with few more features than the cases based on cortical thickness changes. Particularly, on the basis of the obtained results, we recommend volume percent change and symmetrized percent change due to their powerful capability to represent the disease development. Additionally, the feature subset selection provides à significant reduction in the feature vector dimension, making the retrieval process more efficient. It is important to emphasize that among the most frequently automatically selected features using the feature selection, we identified the changes in the brain structures which, according to the literature, are highly affected by the disease. These include volume change of lateral ventricles, inferior lateral ventricles, third ventricle, hippocampus, left part of the amygdala, entorhinal cortex, fusiform, parahippocampal, as well as changes in entorhinal thickness, thickness of posterior cortical areas, or insula cortex.

The obtained results in this study are better than those achieved in [19] on the same cohort. However, they are not directly comparable, because the volume and cortical thickness measurements in that study were generated using the information from at least four time points. According to our study, the inclusion of more time points might influence the results, so the comparison might not be reasonable and fair. Due to the diversity of the datasets or the selected subjects that are used for evaluation, we also cannot make and report objective comparison with the other studies directed to image retrieval for AD [14], [17], [24], [29], even though we can notice improved retrieval performance using our strategy.

The main findings considering the general evaluation metrics are:

1) Using more and latter time points improves the retrieval results. But, in the situation when the processing time or the number of scans are limited, then latter and,

if possible, equally spaced time points might be more appropriate.

2) Volume changes with applied feature selection provide more powerful and semantically relevant subject representation that leads to better retrieval performance. Specifically, percent change and symmetrized percent change cortical and subcortical volumes are recommended. Also, these features might be seen as promising imaging markers for AD.

3) Quality control improves the results in all experiments, thus we recommend to be used whenever possible.

Our future work is directed toward finding a proper combination of the longitudinal changes, plain volume and plain cortical thickness measurements to further improve the semantics derived from the medical cases. Moreover, we are going to extend this research to a bigger cohort to provide more reliable evaluation.

VI. CONCLUSION

The current results showed that the time points used to extract the longitudinal information and their number influence the retrieval performance for the selected cohort. Additionally, we evaluated and recommended the most suitable longitudinal changes that lead to improved and semantically more reliable retrieval. In fact, volume percent change and volume symmetrized percentage change extracted from all four time points were found to be the most representative for the disease, outperforming all the other descriptors. The feature selection provides significantly lower dimension of the feature vectors, making the retrieval more efficient. We also discovered that errors encountered from the automated processing negatively influence the retrieval performance. Thus, we recommend to use the quality control.

REFERENCES

[1] M. W. Weiner *et al.*, "The Alzheimer's disease neuroimaging initiative: A review of papers published since its inception," *Alzheimer's Dementia*, vol. 8, no. 5, pp. e111–e194, 2012.
[2] A. Mourão, F. Martins, and J. Magalhães, "Multimodal medical information retrieval with unsupervised rank fusion," *Comput. Med. Imag. Graph.*, vol. 39, pp. 35–45, Jan. 2015.
[3] E. Pinho, T. Godinho, F. Valente, and C. Costa, "A multimodal search engine for medical imaging studies," *J. Digit. Imag.*, vol. 30, no. 1, pp. 39–48, 2016.
[4] J. Ye, T. Wu, J. Li, and K. Chen, "Machine learning approaches for the neuroimaging study of Alzheimer's disease," *Computer*, vol. 44, no. 4, pp. 99–101, 2011.
[5] L. Velayudhan *et al.*, "Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease," *Alzheimer's Disease*, vol. 33, no. 3, pp. 755–766, 2013.
[6] D. Dai, H. He, J. T. Vogelstein, and Z. Hou, "Accurate prediction of AD patients using cortical thickness networks," *Mach. Vis. Appl.*, vol. 24, no. 7, pp. 1445–1457, 2012.
[7] K. Nho *et al.*, "Voxel and surface-based topography of memory and executive deficits in mild cognitive impairment and Alzheimer's disease," *Brain Imag. Behavior*, vol. 6, no. 4, pp. 551–567, 2012.
[8] J. Lötjönen *et al.*, "Fast and robust extraction of hippocampus from MR images for diagnostics of Alzheimer's disease," *NeuroImage*, vol. 56, no. 1, pp. 185–196, 2011.
[9] M. R. Sabuncu *et al.*, "The dynamics of cortical and hippocampal atrophy in Alzheimer disease," *Arch. Neurol.*, vol. 68, no. 8, pp. 1040–1048, Jan. 2011.
[10] E. Gerardin *et al.*, "Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging," *NeuroImage*, vol. 47, no. 4, pp. 1476–1486, 2009.
[11] L. Piras and G. Giacinto, "Information fusion in content based image retrieval: A comprehensive overview," *Inf. Fusion*, vol. 37, pp. 50–60, Sep. 2017.
[12] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler, "A review of content-based image retrieval systems in medical applications—Clinical benefits and future directions," *Int. J. Med. Inform.*, vol. 73, no. 1, pp. 1–23, 2004.
[13] A. V. Faria, K. Oishi, S. Yoshida, A. Hillis, M. I. Miller, and S. Mori, "Content-based image retrieval for brain MRI: An image-searching engine and population-based analysis to utilize past clinical data for future diagnosis," *NeuroImage, Clin.*, vol. 7, pp. 367–376, Jan. 2015.
[14] M. Agarwal and J. Mostafa, "Image retrieval for Alzheimer's disease detection," in *Medical Content-Based Retrieval for Clinical Decision Support* (Lecture Notes in Computer Science). Berlin, Germany: Springer-Verlag, 2010, pp. 49–60.
[15] W. Cai, S. Liu, L. Wen, S. Eberl, M. J. Fulham, and D. Feng, "3D neurological image retrieval with localized pathology-centric CMRGlc patterns," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 2010, pp. 3201–3204.
[16] M. Agarwal and J. Mostafa, "Content-based image retrieval for Alzheimer's disease detection," in *Proc. 9th Int. Workshop Content-Based Multimedia Indexing (CBMI)*, Jun. 2011, pp. 13–18.
[17] M. Mizotin, J. Benois-Pineau, M. Allard, and G. Catheline, "Feature-based brain MRI retrieval for Alzheimer disease diagnosis," in *Proc. 19th IEEE Int. Conf. Image Process.*, Sep./Oct. 2012, pp. 1241–1244.
[18] Y. Li *et al.*, "Discriminant analysis of longitudinal cortical thickness changes in Alzheimer's disease using dynamic and network features," *Neurobiol. Aging*, vol. 33, no. 2, pp. 427.e15–427.e30, 2012.
[19] K. Trojacanec, I. Kitanovski, I. Dimitrovski, and S. Loshkovska, "Medical image retrieval for Alzheimer's disease using data from multiple time points," in *ICT Innovations* (Advances in Intelligent Systems and Computing). Cham, Switzerland: Springer, 2016, pp. 215–224.

[20] M. Reuter and B. Fischl, "Avoiding asymmetry-induced bias in longitudinal image processing," *NeuroImage*, vol. 57, no. 1, pp. 19–21, 2011.

[21] M. Reuter, H. D. Rosas, and B. Fischl, "Highly accurate inverse consistent registration: A robust approach," *NeuroImage*, vol. 53, no. 4, pp. 1181–1196, 2010.

[22] M. Reuter, N. J. Schmansky, H. D. Rosas, and B. Fischl, "Within-subject template estimation for unbiased longitudinal image analysis," *NeuroImage*, vol. 61, no. 4, pp. 1402–1418, 2012.

[23] C. B. Akgül, D. Ünay, and A. Ekin, "Automated diagnosis of Alzheimer's disease using image similarity and user feedback," in *Proc. ACM Int. Conf. Image Video Retr. (CIVR)*, 2009, Art. no. 34.

[24] K. Trojacanec, I. Kitanovski, I. Dimitrovski, and S. Loshkovska, "Medical image retrieval for Alzheimer's disease using structural MRI measures," in *Biomedical Engineering Systems and Technologies* (Communications in Computer and Information Science). Cham, Switzerland: Springer, 2015, pp. 126–141.

[25] L. Huang, Y. Jin, Y. Gao, K.-H. Thung, and D. Shen, "Longitudinal clinical score prediction in Alzheimer's disease with soft-split sparse regression based random forest," *Neurobiol. Aging*, vol. 46, pp. 180–191, Oct. 2016.

[26] B. Shi, Y. Chen, K. Hobbs, C. D. Smith, and J. Liu, "Nonlinear metric learning for Alzheimer's disease diagnosis with integration of longitudinal neuroimaging features," in *Proc. Brit. Mach. Vis. Conf.*, 2015, pp. 1–13.

[27] A. Chincarini *et al.*, "Integrating longitudinal information in hippocampal volume measurements for the early detection of Alzheimer's disease," *NeuroImage*, vol. 125, pp. 834–847, Jan. 2016.

[28] S. L. Risacher *et al.*, "Longitudinal MRI atrophy biomarkers: Relationship to conversion in the ADNI cohort," *Neurobiol. Aging*, vol. 31, no. 8, pp. 1401–1418, 2010.

[29] S. Liu *et al.*, "Longitudinal brain MR retrieval with diffeomorphic demons registration: What happened to those patients with similar changes?" in *Proc. IEEE 12th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2015, pp. 588–591.

[30] G. Li, J. Nie, and D. Shen, "Consistent reconstruction of cortical surfaces from longitudinal brain MR images," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science). Berlin, Germany: Springer, 2011, pp. 671–679.

[31] G. M. Fitzmaurice, N. M. Laird, and J. H. Ware, *Applied Longitudinal Analysis*. Hoboken, NY, USA: Wiley, 2012.

[32] J. L. Bernal-Rusiel, D. N. Greve, M. Reuter, B. Fischl, and M. R. Sabuncu, "Statistical analysis of longitudinal neuroimage data with linear mixed effects models," *NeuroImage*, vol. 66, pp. 249–260, Feb. 2013.

[33] O. Voevodskaya *et al.*, "The effects of intracranial volume adjustment approaches on multiple regional MRI volumes in healthy aging and Alzheimer's disease," *Frontiers Aging Neurosci.*, vol. 6, p. 264, Jul. 2014.

[34] C. G. Schwarz *et al.*, "A large-scale comparison of cortical thickness and volume methods for measuring Alzheimer's disease," *NeuroImage, Clin.*, vol. 11, pp. 802–812, May 2016.

[35] M. A. Hall and G. Holmes, "Benchmarking attribute selection techniques for discrete class data mining," *IEEE Trans. Knowl. Data Eng.*, vol. 15, no. 6, pp. 1437–1447, Nov./Dec. 2003.

[36] K. Trojacanec, I. Kitanovski, I. Dimitrovski, and S. Loskovska, "The influence of quality control on the image retrieval: Application to longitudinal images for Alzheimer's disease," in *Proc. 14th Int. Conf. Inform. Inf. Technol.*, 2017, pp. 37–42.

KATARINA TROJACHANEC (M'06) was born in 1986, in Kumanovo, Macedonia. She received the bachelor's degree in computer science and engineering from the Faculty of Electrical Engineering and Information Technology, Saints Cyril and Methodius University of Skopje, Macedonia, in 2008, where she received the master's degree with thesis titled Content Based Image Retrieval System for Magnetic Resonance Images under the supervision of Dr. S. Loshkovska in 2010. She is currently pursuing the Ph.D. degree with the Faculty of Computer Science and Engineering, Saints Cyril and Methodius University of Skopje. She has experience with the Faculty of Computer Science, European University in Skopje, in 2008. She was with the Faculty of Electrical Engineering and Information Technologies in 2008. She is currently a Teaching Assistant and Research Assistant with the Faculty of Computer Science and Engineering, Saints Cyril and Methodius University of Skopje. Her research focuses on image processing, data mining, machine learning, neuroimaging, and image retrieval. She was awarded with an Engineering Ring by the Engineering Institution of Macedonia and with Gold Coin by the Saints Cyril and Methodius University of Skopje.

IVAN KITANOVSKI was born in 1986, in Bitola, Macedonia. He received the bachelor's degree in 2009. He is currently pursuing the Ph.D. degree. In 2005, he was enrolled with the Faculty of Electrical Engineering and Information Technologies, on the module Information and Computer Engineering.

Since 2009, he has been a Junior Research and Teaching Assistant with the Faculty of Electrical Engineering and Information Technologies in Skopje. He is currently a Teaching and Research Assistant with the Faculty of Computer Science and Engineering, Saints Cyril and Methodius University of Skopje, Macedonia. His tasks involved conducting exercises for the following courses: user interfaces, internet programming, visualization, virtual reality, object-oriented programming and structural programming.

IVICA DIMITROVSKI was born in 1981 in Kratovo, Macedonia. In 2000, he started his studies with the Faculty of Electrical Engineering and Information Technologies, Saints Cyril and Methodius University of Skopje. He received the bachelor's degree in computer science, automation and electrical engineering from the Faculty of Electrical Engineering and Information Technologies in 2005. In 2008, he received the M.S. degree and the Ph.D. degree in 2011.

In 2009, he was an Assistant with the Department of Computer Science and Engineering, Faculty of Electrical Engineering and Information Technologies, Skopje. He was a Visiting Researcher with the Jozef Stefan Institute, Slovenia, from 2009 to 2010. He is currently an Associate Professor with the Faculty of Computer Science and Engineering, Saints Cyril and Methodius University of Skopje, Macedonia. His fields of interest include image processing, medical images, machine learning, content-based retrieval of multimedia data, and web programming technologies.

SUZANA LOSHKOVSKA (SM'94) was born in 1966, in Skopje, Macedonia. She received the bachelor's and master's degrees in computer science and automation from the Faculty of Electrical Engineering, Skopje, in 1989 and 1992, respectively, and the Ph.D. degree from the Technical University of Wien, Wien, Austria, in 1995.

She has been with the Department of Computer Science and Engineering, Faculty of Electrical Engineering, since 1992. She was Visiting Researcher with the Department of System Engineering and Automation, Academy of Sciences of Lower Austria, Krems, from 1993 to 1994. During this period, she was involved in the design and development of the medical information system with Krems Hospital. Since 1994, she has been with the Faculty of Electrical Engineering. She became a Professor in 2007. She is currently a Professor with the Faculty of Computer Science and Engineering, Saints Cyril and Methodius University of Skopje. Her research interests include programming, visualization, human–computer interaction, virtual reality, medical imaging, modeling and visualization of processes in medical praxis. She has authored approximately 100 scientific papers in the field of computer science, several articles related to the implementation of the European Credit Transfer System and two books for secondary school. She was project leader and participant in a few projects. She is a member of the Editorial Board of the *International Informatica journal*, Slovenia.

● ● ●