

Chapter 1

Combining Radiology Images and Meta-data for Multimodal Medical Case-based Retrieval

Oscar Alfonso Jiménez-del-Toro ^{1,2}, Pol Cirujeda ³ and Henning Müller ^{1,2}

Abstract As part of their daily workload, clinicians examine patient cases in the process of formulating a diagnosis. These large multimodal patient data sets stored in hospitals could help in retrieving relevant information for a differential diagnosis, but are currently not fully exploited. The VISCERAL Retrieval benchmark organized a medical case-based retrieval algorithm evaluation using multimodal (text and visual) data from radiology reports. The common data set contained patient CT (Computed Tomography) or MRI (Magnetic Resonance Imaging) scans and RadLex term anatomy-pathology lists from the radiology reports. A content-based retrieval method for medical cases that uses both textual and visual features is presented. It defines a weighting scheme that combines the anatomical and clinical correlations of the RadLex terms with local texture features obtained from the region of interest in the query cases. The visual features are computed using a 3D Riesz wavelet texture analysis performed on a common spatial domain to compare the images in the analogous anatomical regions of interest in the data set images. The proposed method obtained the best mean average precision in 6 out 10 topics and the highest number of relevant cases retrieved in the benchmark. Obtaining robust results for various pathologies, it could further be developed to perform medical case-based retrieval on large multimodal clinical data sets.

Oscar Alfonso Jiménez-del-Toro

Henning Müller

¹Institute of Information Systems, University of Applied Sciences Western Switzerland Sierre (HES-SO), Sierre, Switzerland,

²University and University Hospitals of Geneva, Geneva, Switzerland, e-mail: oscar.jimenez@hevs.ch

³Pol Cirujeda

Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain, e-mail: pol.cirujeda@upf.edu

1.1 Introduction

As part of their daily workload, clinicians have to visualize and interpret a large number of medical images and radiology reports [15]. In recent years, the volume of images in medical records has increased due to the continuous development of imaging modalities and storage capabilities in hospitals [17]. Going through these large amounts of data is time consuming and not scalable with the current trend of big data analysis [14]. Therefore, the challenge to make efficient use of these large data sets and to provide useful information for the diagnostic decisions of clinicians is of high relevance [16]. It is of particular significance to effectively combine the information contained both in the patients medical imaging and the clinical meta-data from their reports [12].

It is now common for research groups to test their retrieval algorithms on private data sets, impeding the repeatability of their results and comparison to other algorithms [4]. The Visual Concept Extraction Challenge in Radiology (VISCERAL) project was developed as a cloud-based infrastructure for the evaluation of medical image analysis techniques on large data sets [14]. Through evaluation campaigns, challenges, benchmarks and competitions, tasks of general interest can be selected to compare the algorithms on a large scale. One of these tasks is the Retrieval Benchmark, which aims to find cases with similar anomalies based on query cases [9].

In this paper a multimodal (text and visual) approach for medical case-based retrieval is presented. It uses the RadLex [13] terminology and 3D texture features extracted from medical images to objectively compare and rank the relevance of medical cases for a differential diagnosis. Via an estimation of anatomical regions of interest in the spatial domain delineated in medical images, it exploits the visual similarities in the 3D patient scans to improve the baseline text rankings [8]. The implementation of the method, set up and results in the VISCERAL Retrieval Benchmark and lessons learned are explained in the following sections.

1.2 Materials and Methods

The proposed approach to retrieve relevant medical cases was based on a weighting score scheme that combined RadLex terms anatomical and pathological correlations with local visual information.

The VISCERAL Retrieval data set on which this method was implemented and tested is initially addressed. The clinical meta-data weighting scheme is then explained. Afterwards, the various image processing techniques used for the visual feature extraction approach are shown. Finally the fusion of both data, RadLex term lists and 3D patient scans, is explained in the multimodal fusion section.

Data Set

The Retrieval benchmark data set was composed of patient scans (3D volumes) and RadLex Anatomy–Pathology term lists. The 2311 images in the data set were obtained during clinical routine from two data providers¹. The data set had a heterogeneous collection of images including computed tomography (CT) and magnetic resonance imaging (MRI) T1 and T2 weighted imaging, enhanced and unenhanced, in various fields of view (e.g. abdomen, whole body, unknown). For 1813 cases, RadLex anatomy–pathology term lists were generated automatically from the radiology reports of the images. They included the affected anatomical structures and their RadLex term ID, the pathologies and their Radlex term ID, and whether the findings were negated or not in the report. The number of findings and anatomical structures involved varied from case to case.

VISCERAL Retrieval Benchmark Evaluation Setup

Ten query topics, not included in the data set, were distributed to the participants for the evaluation of their retrieval algorithms. The goal of the benchmark was to detect and rank relevant cases in the data set that could potentially aid in the process of diagnosing the query cases. Each query topic was composed of the following data:

- List of RadLex anatomy–pathology terms from the radiology report
- 3D patient scan (CT or MRT1/MRT2)
- Manually annotated 3D mask of the main organ affected
- Manually annotated 3D region of interest (ROI) from the radiologist’s perspective

Participants submitted their rankings and medical experts performed relevance judgments on the submitted cases to determine if they were relevant for the diagnosis of each of the query topics.

Multimodal Medical Case Retrieval

Text Retrieval

Given a set of N medical cases $C = \langle \mathbf{R}_1, \dots, \mathbf{R}_N; \mathbf{V}_1, \dots, \mathbf{V}_N; \mathbf{M}_1, \dots, \mathbf{M}_N; \mathbf{F}_1, \dots, \mathbf{F}_N \rangle$ where the textual information from a radiological report $\mathbf{R}_n$ contains a list L of the anatomies A and pathologies P present in the medical case C_n . The visual information $(\mathbf{V}_n, \mathbf{M}_n, \mathbf{F}_n)$ includes a triple of 3D volumes, containing the patient volume

¹ <http://www.visceral.eu/benchmarks/retrieval-benchmark/>, as of 20th October 2015

$\mathbf{V}_n$, binary label organ mask (annotation) $\mathbf{M}_n$ and binary label region-of-interest (annotation) $\mathbf{F}_n$.

The aim is to create a ranking S of relevant cases $\mathcal{S} = \langle \mathbf{C}^1, \dots, \mathbf{C}^S \rangle$ useful for a differential diagnosis with the target case $\mathbf{C}_T$. Each case $\mathbf{C}_n$ is evaluated according to its radiology report $\mathbf{R}_n$ and visual information $(\mathbf{V}_n, \mathbf{M}_n, \mathbf{F}_n)$ and a final score A ranks the set of cases according to their relevance weight w .

$$\mathbf{S}'_T = (\mathbf{C}_1(w_1), \mathbf{C}_2(w_2), \dots, \mathbf{C}_n(w_n)) \quad (1.1)$$

The correlations were computed with the RadLex term lists provided from the radiology reports. Each similarity feature had a different weight in the final decision for the differential diagnosis and retrieval of cases. The textual similarity between two cases was computed according to the following correlations and their correspondent weighting score (in brackets):

1. Same anatomy with same pathology [0.6]
2. Same anatomy with same pathology negated [0.55]
3. Same anatomy present multiple times [0.2]
4. Same anatomy mentioned once [0.1]
5. Same pathology with different anatomy [0.05]
6. Similar anatomies [0.05]
7. Same imaging modality [0.02]

The weights were defined using a heuristic approach after a review of a subset of the RadLex term lists in the data set from the medical expert. The aim of the weightings is to identify and highlight clinical features that could be relevant for a differential diagnosis and incorporate *a priori* knowledge of the types of image scans contained in the data set. The ranking was performed with the sum of all the weights from the different similarity features with an independent score generated for each case in the retrieval data set. To define similar anatomies, a list of correlating RadLex terms (e.g. lung, superior lobe, pleura...) was manually generated by a medical expert from the standard RadLex term hierarchy ². These lists were generated for each of the query topics in the benchmark.

Helping Multimodal Retrieval with Visual Texture Features

Multimodal retrieval can be influenced by common image processing techniques used in template matching or visual likelihood metrics for content-based image retrieval. Computer vision research areas as image classification and pattern recognition from visible features such as color, contours or texture have been present in recent approaches for the retrieval of medical cases with likely affected organs, image modalities or diagnosis [7].

This section defines a methodology for content-based image retrieval via a similarity measurement from texture-based visual cues. First, a region-of-interest from

² www.RadLex.org, as of 20th may 2015

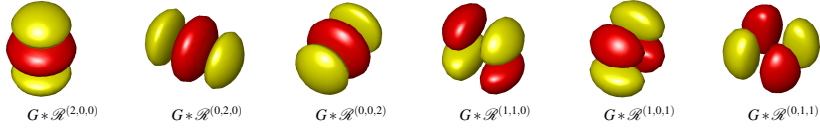


Fig. 1.1 Second-order Riesz kernels $\mathcal{R}^{(n_1, n_2, n_3)}$ convolved with isotropic Gaussian kernels $G(x)$.

a query image is characterized thanks to its computed 3D Riesz wavelet coefficients. In order to deal with 3D structure and also to provide a more compact representation, these features are translated into a particular descriptor space which arises from modelling the covariance matrices of the coefficient observations within a volumetric region, instead of keeping the whole set of feature values. This compact data representation is of crucial interest as it allows to translate both learning image templates and unknown testing image candidates to a common space which can be used in a dictionary-seeking fashion for visual-based retrieval.

3D Riesz Transform for Texture Features

Riesz filterbanks are used in order to characterize the 3D texture of regions of interest in CT images. In previous work, 3D Riesz wavelets have demonstrated successful performance in the modelling task of subtle local 3D texture properties with high reproducibility compared to other methods [20, 19].

The N -th order Riesz transform $\mathcal{R}^{(N)}$ of a three-dimensional signal $f(x)$ is defined in the Fourier frequency domain as:

$$\widehat{\mathcal{R}^{(n_1, n_2, n_3)} f}(\omega) = \sqrt{\frac{n_1 + n_2 + n_3}{n_1! n_2! n_3!}} \frac{(-j\omega_1)^{n_1} (-j\omega_2)^{n_2} (-j\omega_3)^{n_3}}{||\omega||^{n_1 + n_2 + n_3}} \hat{f}(\omega), \quad (1.2)$$

for all combinations of (n_1, n_2, n_3) with $n_1 + n_2 + n_3 = N$ and $n_{1,2,3} \in \mathbb{N}$. Eq. 1.2 yields $\binom{N+2}{2}$ templates $\mathcal{R}^{(n_1, n_2, n_3)}$ and forms multiscale filterbanks when coupled with a multi-resolution framework.

In order to achieve three-dimensional representation, the second-order Riesz filterbank (depicted in Fig. 1.1) is used and rotation-invariance is obtained by locally aligning the Riesz components $\mathcal{R}^{(n_1, n_2, n_3)}$ of all scales based on the locally prevailing orientation as presented in [3].

Invariant Representation via 3D Covariance Descriptors

The choice of a particular set of features for an accurate texture description is as important as a representation that is able to yield invariance to scale, rotation or other spatial changes of the described region of interest. Riesz features are used in conjunction with a representation that takes into account their statistical distribution,

leading to a compact and discriminative notation with several benefits for pattern recognition.

First, a spatial homogenization baseline is achieved by an indirect 3D spatial registration, where a reference image is used to register all the images from the data set and generate a common space domain for visual comparison. The reference image is obtained from a control case of a complete patient scan in order to provide a complete alignment frame. Once a new image is provided as a query, it is first registered to the reference image and included in this rough alignment of the data set images. Then, a set of derived regions of interest is determined for each of the images in the data set by directly transforming the same coordinates from the ROI in the query image. See Fig. 1.2 for a scheme of this workflow.

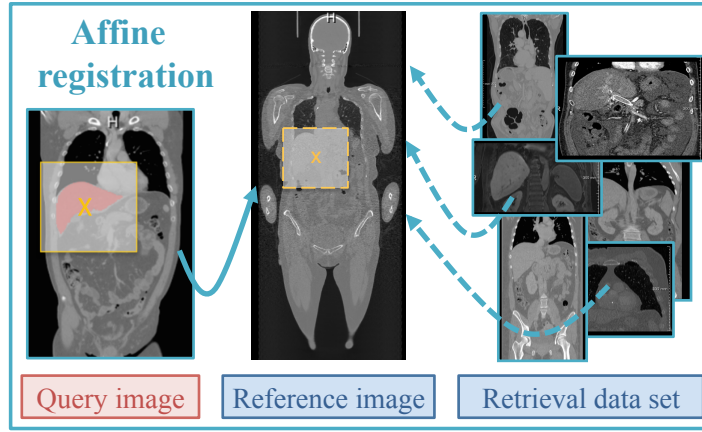


Fig. 1.2 Finding the region of interest (ROI) from the query image in the data set. The image with the biggest size from the data set was selected as the reference image. In order to have a common spatial domain to compare the images, all the images from the data set were registered in advance to this reference image using affine registration (dashed blue arrows). With a new query, the query images were also registered to the reference image and the provided binary mask for the ROI (yellow borders) was transformed using the coordinate transformation from the affine registration of the query image. This procedure defined an indirect ROI (dashed yellow borders) in each of the data set images to compare the visual similarities with the query image.

The required registrations for this step were computed using the image registration implementation from the Elastix software [11]³. The quality of the registration is iteratively evaluated in each optimization of a cost function that aims to minimize the normalized cross correlation from the voxel intensities of the transformed moving image to the fixed target image. Using affine registration the 3D volumes are globally aligned through an iterative stochastic gradient descent optimizer with a multi-resolution approach [10].

The steerability property [21] of Riesz features asserts that voxel intensity values are projected to the direction of maximum variability within the region of inter-

³ <http://elastix.isi.uu.nl>, as of 20th October 2015.

est, thus providing a common reference space for all the observable tissue patterns. Therefore features are guaranteed to be directionality-invariant which, added to the rotation-invariant representation explained below, adds an additional robustness to spatial changes in the proposed covariance descriptor framework.

By their construction, covariance descriptors are suitable for unstructured, abstract texture characterization inside a region, regardless of spatial rigid transformations such as rotation, scale or translations [2]. This is due to a statistics-based representation in which covariance is used as a measure of how several random variables change together (3D Riesz texture features in this case) and used as a discriminative signature of a region of interest. This notion translates the absolute feature space, which is sparse and high dimensional, to a meaningful lower dimensional space of feature covariances where regions with similar texture variabilities lie clustered and differentiated. Furthermore, the construction of covariance descriptors in their natural shape as symmetric positive definite matrices adds an inherent analytical methodology: these matrices form a manifold which can be analysed by its own defined Riemannian metrics [1] for the comparison of descriptor samples.

In order to formally define the 3D Riesz-covariance descriptors, a feature selection function $\Phi(ct, v)$ is denoted for a given 3D CT volume v (in this approach, a single $96 \times 96 \times 96$ block generated using the center of the bounding box surrounding the manually annotated mask of the main organ affected in each the query topics) as:

$$\Phi(v) = \left\{ \mathcal{R}_{x,y,z}^{(n_1, n_2, n_3)}, \forall x, y, z \in v \right\}, \quad (1.3)$$

which denotes the set of 6-dimensional Riesz feature vectors, as defined in Eq. 1.2, obtained at each one of the coordinates $\{x, y, z\}$ contained in the volume cube v .

Then, for a given region v of the CT image, the associated covariance descriptor can be obtained as:

$$Cov(\Phi(v)) = \frac{1}{N-1} \sum_{i=1}^N (\Phi - \mu)(\Phi - \mu)^T, \quad (1.4)$$

where μ is the vector mean of the set of feature vectors $\{\Phi_{x,y,z}\}$ within the volumetric neighbourhood made of $N = 96^3$ samples. Figure 1.3 shows the construction of a sample 3D Riesz-covariance descriptor.

Pattern Matching in the Sym_d^+ Manifold

The resulting 6×6 covariance descriptors are symmetric matrices in which the diagonal elements represent the variance of each Riesz feature, and the non-diagonal elements represent their pairwise covariance. As previously stated, these descriptors are used as discriminative signatures of the texture patterns found in the block v . 3D Riesz-based covariance descriptors do not only provide a representative entity: they also lie in the Riemannian manifold of symmetric definite positive matrices Sym_d^+ . The spatial distribution of the descriptor space is geometrically meaningful

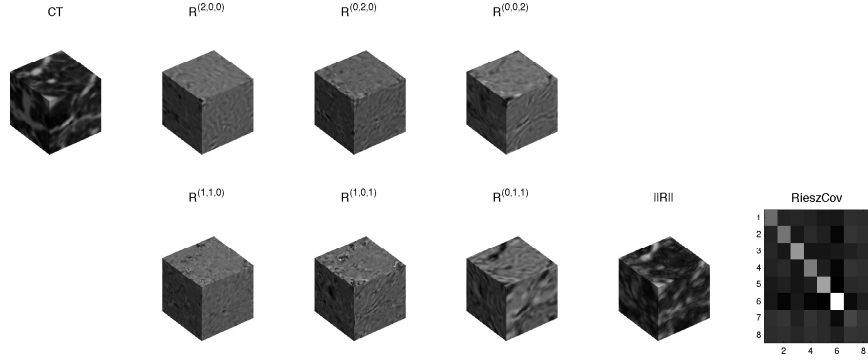


Fig. 1.3 Cues involved in the descriptor calculation for a given CT cubic region. The initial cube depicts the values within a $96 \times 96 \times 96$ pixel volume with its CT intensities; the 6 central cubes depict the 2^{nd} order 3D-Riesz wavelet responses and the Riesz norm is included as well. The matrix in the right sub-figure depicts the resulting covariance descriptor, encoding the different correlations between the distributions of the observed cues.

as 3D regions sharing similar texture characteristics remain clustered when descriptor similarity is computed by means of the Riemannian metrics defined for this

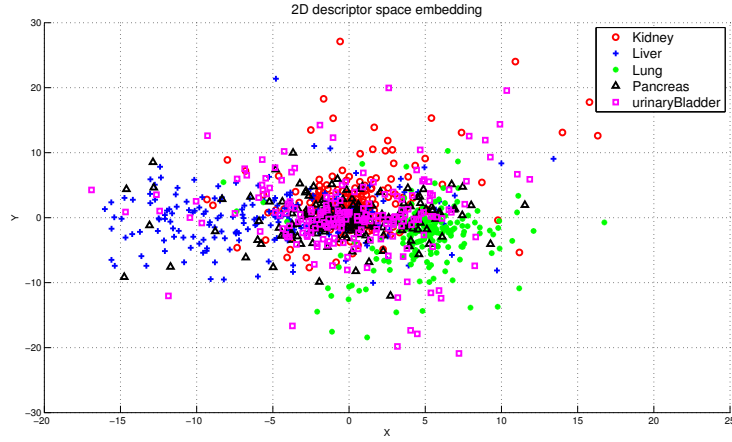


Fig. 1.4 Set of image descriptors obtained from 5 organ textures, belonging to 200 different cubic samples from different patient CT scans. The different descriptors for each class are plotted in different colors in the embedded two-dimensional space, via the multidimensional scaling dimensionality reduction technique, according to the descriptor similarity metric defined in Eq. 1.8. This plot demonstrates geometrical coherence as the class distribution is correlated in the descriptor space: areas with different texture features, such as liver, lung or urinary bladder appear clustered in the descriptor space. Some areas that share texture features, such as pancreas appear more overlapped to other regions. In any case, this descriptor space can be used in linear or non-linear machine learning classification methods for further texture modelling.

non-Euclidean spatial distribution, as defined below. This is depicted in Fig. 1.4, where multidimensional scaling is used for projecting the descriptor space into a two-dimensional plot for visualization. The same notion can be used for feature selection or dimensionality reduction in the non-linear descriptor space.

According to [1], the Sym_d^+ Riemannian manifold constituting the covariance descriptor space can be approximated in close neighbourhoods by the Euclidean metric in its tangent space, T_Y , where the symmetric matrix Y is a reference projection point in the manifold. T_Y is formed by a vector space of $d \times d$ symmetric matrices, and the tangent mapping of a manifold element X to $x \in T_Y$ is made by the point-dependent $\log_Y$ operation:

$$x = \log_Y(X) = Y^{\frac{1}{2}} \log \left(Y^{-\frac{1}{2}} X Y^{-\frac{1}{2}} \right) Y^{\frac{1}{2}}. \quad (1.5)$$

As a computational approximation in certain classification problems, the projection point can be established in a common point such as the Identity matrix, and therefore the tangent mapping becomes:

$$\log(X) = U \log(D) U', \quad (1.6)$$

where U and D are the elements of the single value decomposition (SVD) of $X \in Sym_d^+$.

One property of the projected symmetric matrices in the tangent space T_Y is that they contain only $d(d+1)/2$ independent coefficients, in their upper or lower triangular parts. Therefore it is possible to apply the vectorization operation in order to obtain a linear orthonormal space for the independent coefficients:

$$\hat{x} = \text{vect}(x) = (x_{1,1}, x_{1,2}, \dots, x_{1,d}, x_{2,2}, x_{2,3}, \dots, x_{d,d}), \quad (1.7)$$

where x is the mapping of $X \in Sym_d^+$ to the tangent space, resulting from Eq. 1.5. The obtained vector $\hat{x}$ lies in the Euclidean space $\mathbb{R}^m$, where $m = d(d+1)/2$. This can be used for efficient template storage in cases with big data volumes.

This set of operations is useful for data visualization, feature selection, and for developing machine Learning and classification techniques on top of the particular geometric space of the proposed covariance descriptors. The tangent mapping operator can be taken into account leading to the following Riemannian metric, which expresses the geodesic distance between two points X_1 and X_2 on Sym_d^+ [1]:

$$\delta(X_1, X_2) = \sqrt{\text{Trace} \left(\log \left(X_1^{-\frac{1}{2}} X_2 X_1^{-\frac{1}{2}} \right)^2 \right)}, \quad (1.8)$$

or more simply $\delta(X_1, X_2) = \sqrt{\sum_{i=1}^d \log(\lambda_i)^2}$, where λ_i are the positive eigenvalues of $X_1^{-\frac{1}{2}} X_2 X_1^{-\frac{1}{2}}$.

Therefore, in a similarity retrieval application in which a query region obtained covariance descriptor Q has to be matched against a set of template region descrip-

tors $\{T_i\}$ belonging to different classes, this distance can be used as a supporting metric for a weighted scoring system for multimodal retrieval:

$$\text{class}(Q) = \underset{i}{\operatorname{argmin}} \{ \delta(Q, T_i) \mid \forall i \in T \}, \quad (1.9)$$

since the dimensionality of the proposed descriptors is very compact, this scoring function is computationally feasible for datasets of reasonable sizes.

Multimodal Fusion

It is known from previous medical case-based retrieval benchmarks that the text queries obtain much better results than visual queries [6, 5]. This has been attributed to the currently much more consistent representation of clinical signs in medical images by text labels than by their visual features that are not always very specific. Therefore, it is of high interest to the retrieval information community to find robust visual features that can be combined with semantic terms [12]. To include the information obtained from the visual ranking of the cases into the semantic text weighting scheme, we give an additional weighting if the visual similarity score is high. The additional weight [0.05] is added to the total sum from the textual score of the case if it is in the top 20% of the ranking obtained from the similarity score of the covariance descriptor. These parameters were manually optimized using a small subset of the data set. A medical expert provided a list of correlation-based similarities that are of interest for finding relevant cases in the data set. A single main combination of anatomy and pathology RadLex terms was manually selected from each of the query topic RadLex term lists. This decision was based on the region of interest and organ mask provided in the benchmark to the participants.

Results

Only one run was submitted for the VISCERAL Retrieval benchmark 2015, which combined both the RadlexID weighting score scheme and the visual texture features. This run contained a ranking of 300 cases for each of the ten query topics in the benchmark. The cases were ranked in descending order according to the computed similarity to the query topic.

The proposed approach obtained the highest mean average precision (MAP) scores in 6 out of the 10 query topics [9]. The topic with highest MAP was topic 08—kidney cyst with 0.5131, and the lowest was topic 10—rib fracture with 0.0467. The mean MAP was 0.2367 which was the second best MAP score of the benchmark.

Although the precision from 10 (P₁₀) to 30 (P₃₀) documents retrieved is lower than the method by Spanier et al. [18], our method presented a more stable decline of precision scores obtaining the benchmark top scores for 100 (P₁₀₀) and 200 (P₂₀₀) documents retrieved (see Fig.1.5). Moreover, the proposed method ob-

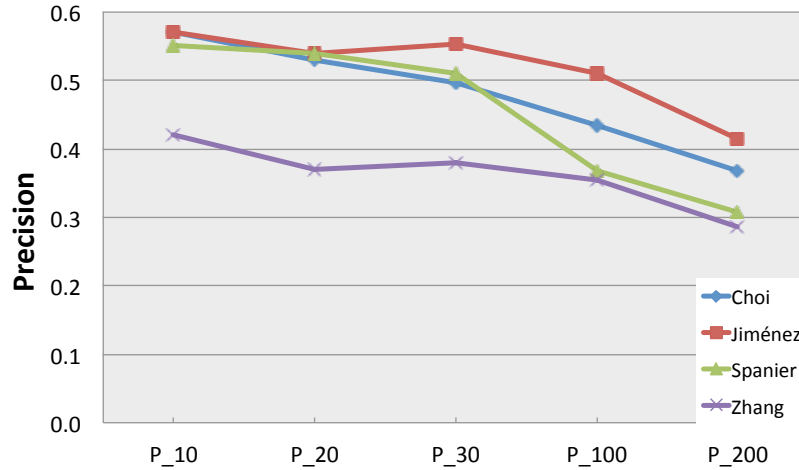


Fig. 1.5 Line graph showing the mean precision scores over all the topics at varying number of cases retrieved: 10 to 200. The best run was selected per participant considering all possible techniques: only text, only visual or mixed. A maximum of 300 cases could be included in each of the submitted rankings per topic.

tained the highest total of relevant documents retrieved (num_rel_ret): 1077 out of a maximum of 2462. In 7 of the 10 query topics it obtained the top number of total documents retrieved out of all the different 21 runs in the benchmark (see Fig.1.6). The mean average precision (MAP), precision after query relevant cases retrieved (Rprec), binary preference (bpref), precision after 10 cases retrieved (P_10) and precision after 30 cases retrieved (P_30) from our method are shown per query topic in Table 1.1. These results take special interest when compared against the other retrieval methods proposed in order to identify which components receive a particular benefit when a multimodal based scoring is introduced. There is a clear advantage

Table 1.1 Results for the 10 query topics in the VISCERAL Retrieval benchmark using the proposed multimodal (text and visual) retrieval approach.

Metric	01	02	03	04	05	06	07	08	09	10	All
MAP	0.2293	0.2227	0.2227	0.2497	0.1949	0.3883	0.1780	0.5131	0.1212	0.0467	0.2367
Rprec	0.4576	0.3575	0.3575	0.3106	0.3508	0.4985	0.3483	0.6399	0.1667	0.0851	0.3572
bpref	0.5035	0.3466	0.3466	0.4047	0.3542	0.4912	0.3444	0.6307	0.1580	0.0837	0.3664
P10	0.2000	0.6000	0.6000	0.8000	0.7000	0.9000	0.6000	0.8000	0.3000	0.2000	0.5700
P30	0.5000	0.5333	0.5333	0.8000	0.5667	0.8667	0.7000	0.8000	0.1333	0.1000	0.5533

of the method by Spanier et al. in query topic 10 when compared to our method. This topic, with radiological diagnosis of rib fracture, had only 47 relevant cases selected by the relevance judges making it the second topic with few relevant cases

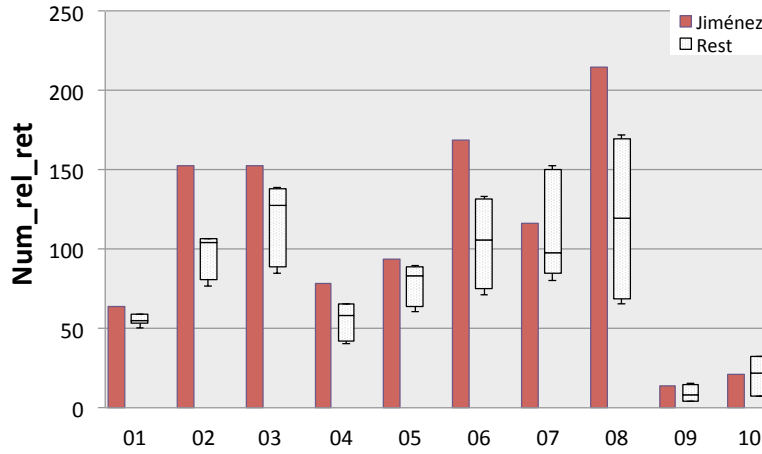


Fig. 1.6 Box plot chart with the total number of relevant cases retrieved (num_rel_ret) per topic in the VISCERAL Retrieval benchmark. The method proposed in this paper is represented with red solid bars. The results of the other participants, including text, visual and mixed runs are shown as white boxes. The horizontal lines inside the boxes mark the median number of relevant cases retrieved. Each box extends from the first to the third quantile of the run results.

in the benchmark. On the other hand, our method was the only run with a mixed technique that produced a ranking for all of the query topics available.

Lessons learned

Having a common data set is fundamental to make objective comparisons between different retrieval methods. Since both multimodal techniques in the benchmark obtained overall the best scores, analysing the topics where other text only techniques worked better (e.g. 07 and 09) is important to improve the features used per topic.

An advantage of scanning a large data set of patient cases is that, like in a real clinical scenario, the distribution of diseases is not uniform. This requires a robust selection of relevant features for a successful retrieval, particularly for those diseases with few cases in the data set.

Visual retrieval is still a complementary technique that is best used with a strong baseline of text related similarities between medical cases. Further research is needed to detect the most relevant region of interest in the images as well as the best visual features per topic. Manual annotation of the medical images in the data set can be useful to improve even further this technique, although it implies a significant increase in the workload of the clinicians when handling this data.

1.3 Conclusions

An innovative multimodal (using text and visual information) medical case-based retrieval approach is presented. A rule-based weighting of the anatomical and clinical RadLex term correlations from radiology reports is used as a baseline to find relevance features from the cases. The results of the RadLex IDs are further improved with state-of-the-art techniques (Riesz wavelets, image registration and covariance descriptors) to compute a visual similarity score between the medical images in the cases. The method was implemented and tested in the VISCERAL Retrieval Benchmark 2015, with overall promising results for the retrieval of relevant cases for differential medical diagnosis.

1.3.0.1 Acknowledgments

This work was supported by the EU in FP7 through VISCERAL (318068), Khresmoi (257528) and the Swiss National National Foundation (SNF grant 205320-141300/1).

References

1. Arsigny, V., Fillard, P., Pennec, X., Ayache, N.: Log-Euclidean metrics for fast and simple calculus on diffusion tensors. *Magnetic Resonance in Medicine* 56(2), 411–421 (2006)
2. Cirujeda, P., Mateo, X., Dicente, Y., Binefa, X.: MCOV: A covariance descriptor for fusion of texture and shape features in 3D point clouds. In: *International Conference on 3D vision (3DV)* (2014)
3. Depeursinge, A., Foncubierta-Rodríguez, A., Van De Ville, D., Müller, H.: Lung texture classification using locally-oriented Riesz components. In: Fichtinger, G., Martel, A., Peters, T. (eds.) *Medical Image Computing and Computer Assisted Intervention – MICCAI 2011*. *Lecture Notes in Computer Science*, vol. 6893, pp. 231–238. Springer Berlin / Heidelberg (September 2011)
4. Hanbury, A., Müller, H., Lings, G., Weber, M.A., Menze, B.H., Fernandez, T.S.: Bringing the algorithms to the data: cloud-based benchmarking for medical image analysis. In: *CLEF conference*. Springer *Lecture Notes in Computer Science* (2012)
5. García Seco de Herrera, A., Foncubierta-Rodríguez, A., Müller, H.: Medical case-based retrieval: Integrating query MeSH terms for query-adaptive multi-modal fusion. In: *SPIE Medical Imaging*. International Society for Optics and Photonics (2015)
6. García Seco de Herrera, A., Kalpathy-Cramer, J., Demner Fushman, D., Antani, S., Müller, H.: Overview of the ImageCLEF 2013 medical tasks. In: *Working Notes of CLEF 2013 (Cross Language Evaluation Forum)* (September 2013)
7. García Seco de Herrera, A., Müller, H., Bromuri, S.: Overview of the ImageCLEF 2015 medical classification task. In: *Working Notes of CLEF 2015 (Cross Language Evaluation Forum)* (September 2015)
8. Jiménez-del-Toro, O.A., Cirujeda, P., Dicente Cid, Y., Müller, H.: RadLex terms and local texture features for multimodal medical case retrieval. In: *Multimodal Retrieval in the Medical Domain*. *Lecture Notes in Computer Science*, vol. 9059. Springer (2015)

9. Jiménez-del-Toro, O.A., Hanbury, A., Lings, G., Foncubierta-Rodríguez, A., Müller, H.: Overview of the VISCERAL retrieval benchmark 2015. In: *Multimodal Retrieval in the Medical Domain (MRMD) 2015. Lecture Notes in Computer Science*, vol. 9059. Springer (2015)
10. Klein, S., Pluim, J.P., Staring, M., Viergever, M.A.: Adaptive stochastic gradient descent optimisation for image registration. *International Journal of Computer Vision* 81(3), 227–239 (2009)
11. Klein, S., Staring, M., Murphy, K., Viergever, M.A., Pluim, J.P.: Elastix: a toolbox for intensity-based medical image registration. *IEEE Transactions on medical imaging* 29(1), 196–205 (2010)
12. Kurtz, C., Depeursinge, A., Napel, S., Beaulieu, C.F., Rubin, D.L.: On combining visual and ontological similarities for medical image retrieval applications. *Medical Image Analysis* 18(7), 1082–1100 (2014)
13. Langlotz, C.P.: Radlex: A new method for indexing online educational materials. *Radiographics* 26(6), 1595–1597 (2006)
14. Lings, G., Hanbury, A., Menze, B., Müller, H.: Visceral: Towards large data in medical imaging — challenges and directions. In: Greenspan, H., Müller, H., Syeda-Mahmood, T. (eds.) *Medical Content-Based Retrieval for Clinical Decision Support*. pp. 92–98. *Lecture Notes in Computer Science*, Springer, Berlin, Heidelberg (2012)
15. Müller, H., Michoux, N., Bandon, D., Geissbühler, A.: A review of content-based image retrieval systems in medicine—clinical benefits and future directions. *International Journal of Medical Informatics* 73(1), 1–23 (2004)
16. Rubin, D., Napel, S.: Imaging informatics: toward capturing and processing semantic information in radiology images. *Yearbook of medical informatics* pp. 34–42 (2010)
17. Rubin, G.D.: Data explosion: the challenge of multidetector-row ct. *European Journal of Radiology* 36(2), 74–80 (2000)
18. Spanier, A.B., Joskowicz, L.: Medical Case-Based Retrieval of Patient Records Using the Radlex Hierarchical Lexicon. In: *Multimodal Retrieval in the Medical Domain. Lecture Notes in Computer Science*, vol. 9059. Springer (2015)
19. Jiménez del Toro, O.A., Foncubierta-Rodríguez, A., Depeursinge, A., Müller, H.: Texture classification of anatomical structures in CT using a context-free machine learning approach. In: *SPIE Medical Imaging 2015* (2015)
20. Jiménez del Toro, O.A., Foncubierta-Rodríguez, A., Vargas Gomez, M.I., Müller, H., Depeursinge, A.: Epileptogenic lesion quantification in MRI using contralateral 3D texture comparisons. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013. Springer Lecture Notes in Computer Science*, vol. 8150, pp. 353–360 (2013)
21. Unser, M., Van De Ville, D.: Wavelet steerability and the higher-order Riesz transform. *IEEE Transactions on Image Processing* 19(3), 636–652 (March 2010)