VISCERAL Silver Corpus Label Fusion Framework

Markus Krenn
Medical University Of Vienna

1 Silver corpus merging framework

The *VISCERAL* silver corpus merging framework provides the functionality to generate silver corpus annotations of anatomical structures in medical images by fusing data that is typically available from challenges or benchmarks. This is (1) a small set of manually performed expert annotations of high quality (gold corpus) and (2) large sets of possibly less accurate segmentation estimates obtained from algorithms participanting in a benchmark.

Label fusion approaches are used to fuse both segmentation sources to compute segmentation estimates of higher quality (than relying on one of both sources) in a large set of medical images resulting in a so called *silver corpus*.

The framework established during the VISCERAL project `http://www.visceral.eu/` is publicly available for the research community at `https://github.com/Visceral-Project/silverCorpusFramework`. The present document describes the frameworks components, the underlying database and provides installation guidelines. Figure 1 illustrates the frameworks components and aim.

1.1 Components

The framework is based on three components, which are described in the following paragraphs

Gold- and Silver Corpus Volumes The framework distinguishes between two sets of medical images:

1. **Gold corpus volumes:** A set of volumes that carry manually performed expert annotations of all anatomical structures for which silver corpus segmentations should be generated. Gold corpus annotations are brought to novel images by medical image registration and label propagation to compute segmentation estimates of the target structures in volumes of the silver corpus.

2. **Silver corpus volumes:** A from the gold corpus disjunct set of volumes that are not manually annotated. Instead the framework computes annotations of all target structures in volumes of the silver corpus by fusing segmentation estimates originated from atlases (GC volumes) and algorithmic segmentations.

Participant segmentations The framework uses segmentation estimates of in a benchmark participating algorithms to compute silver corpus segmentations. Participant segmentations are available for gold- and silver corpus volumes. The performance of each participant is determined on each structure by comparision to the gold corpus ground truth annotations.

Expert annotations Mannually performed expert annotations of all target structures annotated in each volume of the gold corpus. Those annotations are brought to volumes of the silver corpus by medical image registration and label propagation. The performance of each atlas (medical image and corresponding annotations) on each target structure is evaluated by cross validation on the gold corpus volumes.

1.2 Implemented label fusion approaches

The *VISCERAL Silver Corpus Framework* provides implementations of three label fusion approaches to fuse atlas and participant segmentations. Let $\mathcal{L} = \langle \mathbf{L}_1, \ldots, \mathbf{L}_M \rangle$ denote as a set of M binary label images and $\mathbf{u} = u_1, \ldots u_M$ as a vector of corresponding segmentation weights. $\mathbf{L}'(\mathbf{x})$ denotes the silver corpus segmentation estimate of a target image $\mathbf{I}_T$.

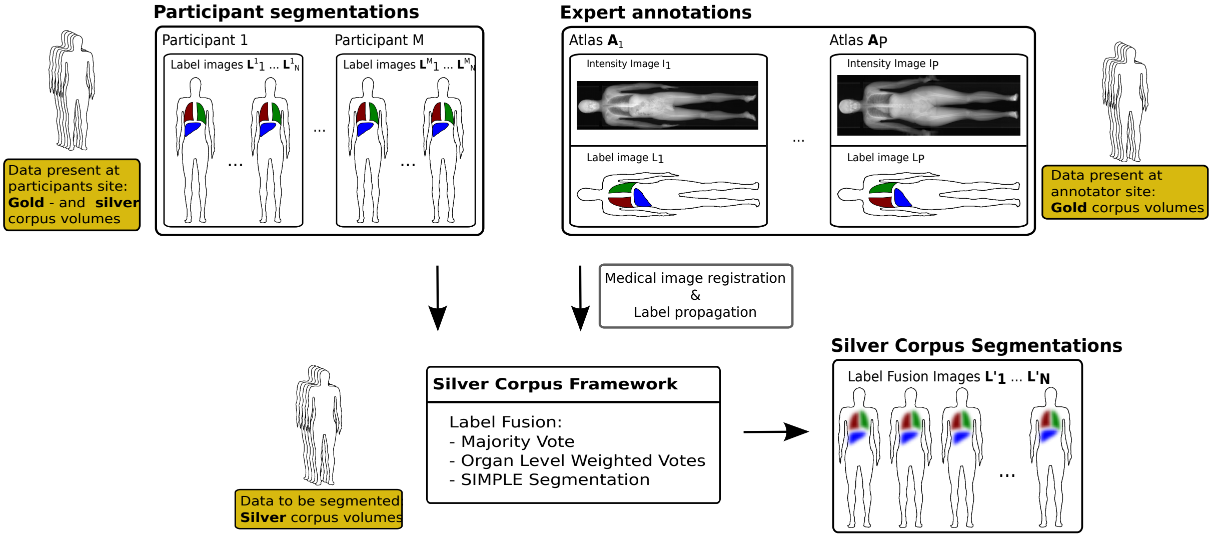

Figure 1: Overview of all components of the VISCERAL Silver Corpus Merging Framework

1. **Majority Vote (MV):** The computed segmentation performance estimates are not considered during fusion, and all contributing segmentations are weighted equally. Each voxel $\mathbf{x}$ of $\mathbf{L}'_T$ is assigned with the label that is most frequent in corresponding voxels of all contributing segmentation estimates [1]:

$$\mathbf{L}'_T(\mathbf{x}) = \left\{ \begin{array}{ll} 1, & \left(\sum_{i=1}^{M} \mathbf{L}_i(\mathbf{x})\right) \geq \frac{M}{2} \\ 0, & \left(\sum_{i=1}^{M} \mathbf{L}_i(\mathbf{x})\right) < \frac{M}{2} \end{array} \right\} \tag{1}$$

2. **Organ Level Weighted Voting:** $\mathbf{L}'_T$ is derived by a majority vote where the impact of each $\mathbf{L}_i \in \mathcal{L}$ is weighted by u_i. Even though the method is initially called Global Weighted Voting [1], we refer to this approach as Organ Level Weighted Voting (OLWV), since weights are determined for each organ independently.

$$\mathbf{L}'_T(\mathbf{x}) = \left\{ \begin{array}{ll} 1, & \left(\sum_{i=1}^{M} \mathbf{L}_i(\mathbf{x}) \cdot u_i\right) \geq \frac{\sum u}{2} \\ 0, & \left(\sum_{i=1}^{M} \mathbf{L}_i(\mathbf{x}) \cdot u_i\right) < \frac{\sum u}{2} \end{array} \right\} \tag{2}$$

3. **SIMPLE:** Selective and Iterative Method for Performance Level Estimation (SIMPLE) proposed by Langerak et. al. [2] is based on an iterative strategy that alternates on

   (a) Estimating the hidden ground truth segmentation of the target image

   (b) Estimating the performances of the contributing segmentations

   Until both, the estimated of the ground truth and the performance estimates of contributing segmentation converges. Additionally SIMPLE excludes badly performing segmentations from the fusino process in each iteration.

   Next to $\mathcal{L}$, the algorithm is parametrized by k and α. Where α influences the performance level that contributing segmentations must exceed to remain in the set of contributing segmentation and k defines the number of iterations in which segmentations are kept for fusion even though the performance threshold is not reached. Optional, SIMPLE takes initial segmentation weights into account.

2 Data base

The present section describes the underlying database design of the silver corpus merging framework. The framework stores all information in the following data base tables:

- **Volume** A volume is defined by its VolumeID and PatientID. Additionally, *Modality Bodyregion* and the volumes *Filename* is stored. The fields *GoldCorpus* and *SilverCorpus* indicate whether a volume is part of the gold or the silver corpus.
- **Structure** Stores anatomical target structures which intended to be segmented by the framework. A structure is identified by its StructureID. Additionally the structures *Name* is stored.
- **ExpertAnnotation** Contains manual expert annotations of anatomical structures in gold corpus volumes. An annotation is defined by its VolumeID, PatientID and StructureID. The field *Filename* holds the path to the annotation volume.
- **ParticipantSegmentation** This table holds segmentation estimates originated by different participants of a benchmark (algorithmic segmentations). A paricitpant segmentation is identified by its VolumeID, PatientID, StructureID and ParticipantID. *Filename* points to the segmentation volume, segmentation quality (if ground truth is available) is stored in the field *Performance*. These values are used to compute structure specific performance estimates of participants and furthermore to generate segmentation weights during the label fusion process.
- **RegistrationResults** This table holds computed registration results of two volumes. An entry is defined by its SourcePatientID,SourceVolumeID,TargetPatientID and TargetVolumeID. The fields *Affine* and *NonRigid* poit to files containing the results of affine and non-rigid image alignments.
- **AtlasSegmentation** This table stores computed atlas segmentations, i.e. segmentation estimates generated by the propagation of an expert annotation to another volume. An entry is defined by its SourcePatientID,SourceVolumeID,TargetPatientID, TargetVolumeID and the StructureID. *Filename* holds the path to the segmentation volume, segmentation quality (if ground truth is available) is stored in the field *Performance*.
- **SilverCorpusSegmentation** This table stores the computed silver corpus segmentations. An entry is identified by its VolumeID, PatientID, StructureID and LabelFusionTypeID. Additionally the path to the resulting segmentation volume is stored in *Filename*, segmentation quality (if ground truth is available) is stored in the field *Performance*.

Figure 2 shows the enhanced entity-relationship model of the data base.

3 Installation guidelines

The final section of this documents provides installation guidelines to set up the VISCERAL Silver Corpus Merging framework.

Repository content The source code is available at `https://github.com/Visceral-Project/silverCorpusFramework` and consists of two components:

- **MySQL database:** Create statements for all data base tables and views requred are provided at /database/ of the repository. Those scripts also fill the DB tables with example data.
- **Matlab backend:** The framework is implemented in Matlab, the source code is available at /matlabBackend/ of the repository. The sources include a tutorial script to show how DB getters and setters can be used and provides a guideline to build silver corpus segmentations as described within this document.

Required software packages The following software packages are required to set up the VISCERAL silver corpus merging framework:

- **Nifty Reg Toolbox:** The nifty-reg toolbox is used for the purpose of medical image registration and label propagation. Binaries can be downloaded from `http://cmictig.cs.ucl.ac.uk/wiki/index.php/NiftyReg`

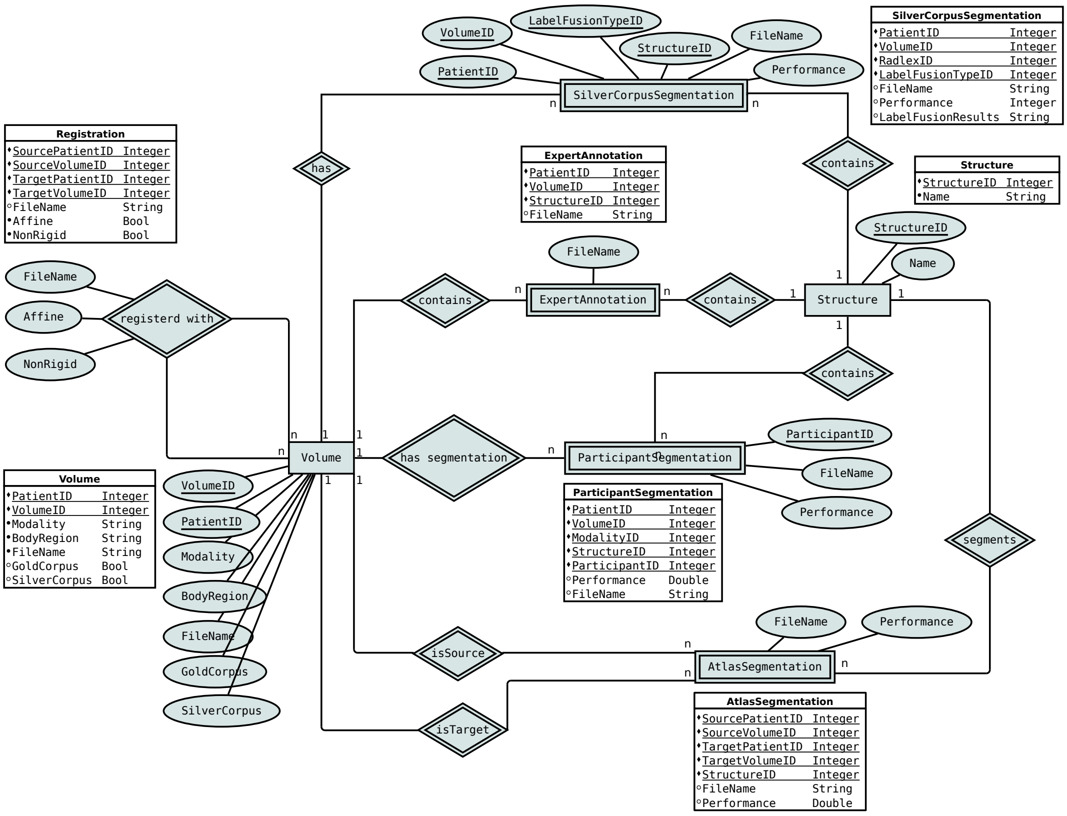

Figure 2: Entity Relationship Model of the Silver Corpus Framework data base

Stepwise installation

1. **Creating the DB:** All create statemetns are provided at */database/*
2. **Set DB access information:** Enter the DB access information (server-adress, username, password) in */matlabBackend/db_functions/dbOpenConnection.m*
3. **Setup Matlab environment:** Navigate to the folder */matlabBackend/* and edit the data directories for storing volumes, registrations, annotations and segmentations in *initializeFramework.m*
4. **Run the tutorial:** Run the script *tutorial.m*, which provides a stepwise description of all components of the VISCERAL silver corpus merging framework.

References

[1] X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de Solorzano. Combination strategies in multi-atlas image segmentation: Application to brain mr data. *Medical Imaging, IEEE Transactions on*, 28(8):1266–1277, 2009.

[2] T.R. Langerak, U.A. Van der Heide, A. N T J Kotte, M.A. Viergever, M. Van Vulpen, and J. P W Pluim. Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (simple). *Medical Imaging, IEEE Transactions on*, 29(12):2000–2008, 2010.