Eurographics Workshop on Visual Computing for Biology and Medicine (2015)
K. Bühler, L. Linsen, and N. W. John (Editors)

Uncertainty Estimation and Visualization for Multi-modal Image Segmentation

Ahmed Al-Taie[1,3], Horst K. Hahn[1,2], and Lars Linsen[1]

[1]Jacobs University, Bremen, Germany
[2]Fraunhofer MEVIS, Bremen, Germany
[3] Computer Science Department, College of Science for Women, Baghdad University, Baghdad, Iraq

Abstract

Multi-modal imaging allows for the integration of complementary information from multiple medical imaging modalities for an improved analysis. The multiple information channels may lead to a reduction of the uncertainty in the analysis and decision-making process. Recently, efforts have been made to estimate the uncertainty in uni-modal image segmentation decisions and visually convey this information to the medical experts that examine the image segmentation results. We propose an approach to extend uncertainty estimation and visualization methods to multi-modal image segmentations. We combine probabilistic uni-modal image segmentation results using the concept of ensemble of classifiers. The uncertainty is computed using a measure that is based on the Kullback-Leibler divergence. We apply our approach for an improved segmentation of Multiple Sclerosis (MS) lesions from multiple MR brain imaging modalities. Moreover, we demonstrate how our approach can be used to estimate and visualize the growth of a brain tumor area for imaging data taken at multiple points in time. Both the MS lesion and the area of tumor growth are detected as areas of high uncertainty due to different characteristics in different imaging modalities and changes over time, respectively.

1. Introduction

Different medical imaging techniques are applied in clinical settings. Each imaging modality has its very own strengths and weaknesses. Hence, it has become a common procedure to use multiple imaging modalities. The complimentary information that is captured by the different modalities allow for an improved analysis and decision making. A core component of the analysis process is image segmentation. Such algorithms can try to operate directly on the multi-modal images or they can be applied separately to the individual image modalities followed by a combination of the segmentation results.

In the context of multi-modal medical image analysis, the process of medical image fusion where a set of images from single or multiple imaging modalities are registered and combined to improve the imaging quality by reducing imaging errors and artifacts and by merging complementary information is very common. The main goal of such a process is to increase the clinical applicability of medical images for diagnosis and assessment of medical problems. Throughout this paper, we assume that the different modalities are registered. Recently, James and Dasarathy [JD14] in their review of over 300 approaches in this field conclude that multi-modal medical image fusion algorithms and devices have shown notable achievements in improving clinical accuracy of decisions based on medical images. Moreover, the use of multi-modal image fusion methods offer a greater diversity of the features used for the medical analysis applications. This diversity often leads to a more robust information analysis that reveals information that can otherwise not be observed. The extra information obtained from the fused images can be used for more precise localization of abnormalities. The survey also listed a wide range of medical applications that make use of these techniques to support the medical analysis. There are two possible scenarios for medical image fusion. In the first scenario, the different modalities are combined to produce a single image that is assumed to be better for the analysis than each uni-modal image. In the second scenario, the different modalities are segmented before they are combined to produce an aggregated segmentation that is assumed to be better than the uni-modal image segmentations. Our approach follows the latter concept, as the individual segmentations add meaningful in-

DOI: 10.2312/vcbm.20151205

formation that could be helpful for improving the accuracy of the fusion process. Multi-modal medical image segmentation is considered a subfield of medical image fusion which is called decisions fusion and can be achieved through an ensemble of classifiers model.

The concept of combining the results of multiple segmentation or classifiers for more reliable and accurate results when compared to the results of individual classifiers is known as committee machine, mixture of experts, or ensemble of classifiers. This concept has been confirmed by several studies in pattern recognition and machine learning community. An important aspect of such an ensemble of classifiers is the diversity, i.e., that the complementary information of the individual classifiers can improve the final result when combining them. Multi-modal medical data represent a good example of such kind of diversity.

Recently, many approaches have been presented to tackle uncertainty estimation and visualization including a few techniques in the context of medical image segmentation, but they typically address the uncertainty associated with a single segmentation approach. Only recently, Al-Taie et al. [ATHL14a] generalize their methods for single segmentation to an ensemble of classifier segmentation using multiple unsupervised segmentations of the same input image as individual classifiers. These studies show the importance of uncertainty-aware medical visualization in supporting the analysis and decision-making process.

To our knowledge, there exists no method to estimate and visualize the uncertainty associated with multi-modal image segmentation fusion, although integrating the complementary information of multi-modal images may reduce the aggregated uncertainty and, consequently, improve the analysis. In this paper, first we apply a multi-modal ensemble of classifier segmentation that makes use of the complementary information captured by the different modalities. Second, we extend the uncertainty estimation and visualization methods for uni-modal imaging data proposed in [ATHL14b] to a multi-modal imaging data. We have developed an interactive analysis tool that incorporates uncertainty visualization methods. Finally, we combine our methods and apply them to some case studies. In particular, we show how multi-modal brain MR image segmentations can be fused to detect and visualize Multiple Sclerosis (MS) lesions. Due to the different classifications of the lesion for different imaging modalities, the lesion shows up as a high-uncertainty region. Also, we apply our method to the fusion of brain MR images taken at different points in time. Here, the uncertain regions deliver the changes over time, i.e., the growth of the tumor. Hence, we can detect and visualize the areas of tumor growth.

The main contributions of this paper can be summarized as: (1) Uncertainty estimation for multi-modal medical image segmentation from ensemble of classifiers with their visualization. (2) Interactive uncertainty visualization tool for visual analysis. (3) Application of our approach for an improved segmentation result for multi-modal medical imaging data.

2. Related Work

2.1. Ensemble of classifiers segmentation

In recent years, combining ensemble of classifiers in order to improve their performance have witnessed a great attention by researchers across different fields to solve different classification problems. Kittler et al. have reviewed the combining rules and introduced a common theoretical framework of these rules [KHDM98]. Dietterich has reviewed the ensemble methods algorithms and explained from a statistical, computational, and representational point of views why ensembles can often performs better than any individual classifier [Die00]. Mignotte introduced the probabilistic Rand index (PRI) as combining strategy in a label field fusion Bayesian model for image segmentation [Mig10]. Fred et al. have explored the idea of evidence accumulation for combining the results of multiple clusterings using different ways of producing data partitions in order to achieve the diversity for more improvement [FJ05]. Recently, Paci et al. proposed an ensemble-based texture classification system [PNS13].

The main motivation behind the ensembles of classifiers concept is the ensemble's diversity. Diversity can be achieved in multiple ways: One way is to apply different algorithms or the same algorithm with different settings to the same image. Another way is to apply the same algorithm to different representations of the image such as using different intensity mappings or color spaces. Multi-modal medical data belong to the later class of diversity, where we are dealing with different image representations of the same subject.

In the context of combining the members of the ensemble of classifiers, there are several combining rules proposed in the literature. Examples of these rules are majority vote, weighted majority vote, or probability rules such as product, sum, maximum, minimum, median, etc. In general, the concept of ensemble of classifiers was mostly used in machine learning applications for supervised classification.

As image segmentation plays an essential role in any medical visualization system, medical image segmentation is the most addressed problem to be solved using the ensembles of classifiers concept in the biomedical field. Several researchers exploited the concept of ensemble methods to tackle the drawbacks of the individual segmentation approaches or to estimate the accuracy of individual approaches . Rohlfing et al. proposed a multi-classifier framework for atlas-based image segmentation. Images from several subjects have been segmented using multiple individual atlases, or using one atlas registered with different parameter settings for different subjects. Then, the combining rules are used to produce the final segmentation [RM05].

Warfield et al. presented the STAPLE algorithm for the validation of image segmentation using a collection of segmentations produced by human raters or by automated segmentation algorithms [WZW04]. The algorithm uses an expectation-maximization approach in an iterative way to estimate a probabilistic ground truth. The estimated ground truth is then used for performance assessment of an automated image segmentation algorithm or for performance comparison of human raters and the automated algorithms. Langerak et al. have proposed the SIMPLE algorithm as an improvement to the STAPLE algorithm. Artaechevarria et al. [AMBdS09] followed up on the idea by Rohlfing et al. in combining multi-atlas-based image segmentations. Rohlfing et al. [RM05] pointed out that producing multiple atlases (also human-rater segmentations for STAPLE or SIMPLE) is time-consuming and tedious, such that atlases are, in practice, not always available. Langerak et al. [LvdHK*10] referred to the shortcoming of atlas-based segmentation as being equivalent to the segmentation by human expert. These drawbacks may lead to the fact that the ensemble methods using atlas-based segmentations become impractical.

In this paper, we combine the result of several unsupervised classification-based segmentations of the multi-modal input images using similar or different segmentation approaches with acceptable accuracies. We achieve the required diversity and remove the above-mentioned drawbacks, i.e., the requirement for producing atlas-based or human-rater segmentations.

2.2. Uncertainty estimation and visualization in segmentation results

Recently, several approaches presented robust methods to estimate and visualize the uncertainties associated with probabilistic segmentations. These studies show how the methods can be useful for post-segmentation visual analysis and for decision-making support. Saad et al. [SMH10] introduce two-way and three-way interactive tools, which measure the difference between the first and second largest and between the second and third largest probabilities, respectively. These tools are used to highlight the uncertainty regions in the segmentation results. Praßni et al. [PRH10] use the probabilistic segmentation result of a random walker algorithm. After classifying the pixels into being certain or uncertain based on some selected probability thresholds, they use the gradient of the maximum probability of the uncertainty information to estimate the uncertain area at the boundary of segments. The approaches by Potter et al. [PGA13] and Al-Taie et al. [ATHL14b] use concepts from information theory to estimate and visualize the uncertainty of a probabilistic segmentation result. In [ATHL14a], Al--Taie et al. generalize their methods for single segmentation in the latter work to estimate the uncertainty associated with multiple segmentations using the ensemble-based framwork. Ristovski et al. [RPHL14] present a taxonomy to a wide range of uncertainty sources that encountered in the medical visualization pipeline.

2.3. Multi-modal image segmentation

James and Dasarathy contracted a survey of more than 300 approaches on medical image fusion that encompasses a broad range of techniques from image fusion to address medical issues captured by human body, organ, and cell images. The techniques include several multi-modal medical image segmentation methods using a decision fusion model. This review shows a growing interest and application of the imaging technologies in the areas of medical diagnostics, analysis, and historical documentation. The survey article lists a wide range of applications that use multi-modal fusion in diagnosis and assessments of medical conditions effecting brain, breast, lungs, liver, bone marrow, stomach, mouth, teeth, intestines, soft tissues and bones. They also provide a collective view of the applicability and progress of information fusion techniques in medical imaging for clinical studies and document that the three major areas of studies in medical image fusion comprise: (a) identification, improvement, and development of imaging modalities useful for medical image fusion, (b) development of different techniques for medical image fusion, and (c) application of medical image fusion for studying human organs of interest in assessments of medical conditions (for more details see [JD14]).

Although there exists this wide range of methods, we are not aware of any approach that estimates the uncertainty associated with multi-modal image segmentation. In this paper, we extend the recently developed uncertainty measure by Al-Taie et al. [ATHL14b] and their generalization in [ATHL14a] for multiple segmentations to estimate the uncertainty associated with ensemble-based multi-modal image segmentations suitable for several combining rules. The proposed method does not rely on ground truth. Moreover, we apply the proposed methods for certain multi-modal medical image analysis applications. An interactive tool supports the visual analysis of the estimated uncertainty.

3. Combining segmentation ensembles

In the context of probabilistic segmentation, the output associated with each voxel x is the probability vector $P(x)$ where the i^{th} entry $P_i(x)$ of the vector denotes the probability that voxel x belongs to the segment (or class) i out of C segments (classes) such that $\sum_{i=1}^{C} P_i(x) = 1$ (i.e., $P_i(x)$ is the a posteriori probability for class i). Traditionally, the maximum a posteriori (MAP) Bayesian principle is applied to obtain a hard classification from this "soft" output.

In the framework of combining the results of L classifiers, some combining rules depend on the soft output (the a posteriori probabilities) of the individual classifiers such as the product, sum, max, min, and median rules, while other rules

depend on the label field (i.e., on the hard classification output) such as the majority voting or the weighted majority voting (see [KHDM98]). In this paper, we use the majority voting (as it represents the rule with best performance from an uncertainty point of view according to [ATHL14a]) and its variants the weighted majority voting rules.

For quick reference, we rewrite these rules here (the majority vote as defined in [KHDM98]). To each pixel x, we assign the class that maximizes the value of the argument of the corresponding rule. Hence, we assign the following classes:

- Majority Vote Rule: Applying the MAP Bayesian principle to the a posteriori probabilities P_{kj} produces a binary-valued function Δ_{kj} as

$$\Delta_{kj} = \begin{cases} 1 & \text{if } P_{kj}(x) = \max\limits_{i=1}^{C} P_{ij}(x) \\ 0 & \text{otherwise.} \end{cases}$$

  Then, under the assumption of equal priors, the majority vote rule simply counts the votes received for each class from the individual classifiers and selects as final decision the class with the largest number of votes:

$$\arg\max_{k=1}^{C} \sum_{j=1}^{L} \Delta_{kj}$$

- Weighted Majority Vote Rule: Based on some assumptions, the individual classifiers are assigned different weights (e.g., the accuracy level of the individual classifiers). In this case, the majority vote rule becomes a weighted majority vote rule

$$\arg\max_{k=1}^{C} \sum_{j=1}^{L} \omega_j \Delta_{kj},$$

  where ω_j is the weight assigned to classifier j.

4. Uncertainty Estimation for Uni-modal Image Segmentation

Al-Taie et al. [ATHL14b] proposed several forms of the normalized Kullback-Leibler divergence as measures to estimate the uncertainty associated with the probabilistic segmentation result for uni-modal images. We give their second form here, which we will extend in the next section: The uncertainty for voxel v using the second form of Al-Taie et al. [ATHL14b] is defined by

$$U_{\overline{KL1}}(v) = 1 - \frac{D_{KL}(\mathbf{P}_v||\mathbf{P}_{max})}{D_{KL}(\mathbf{P}_{min}||\mathbf{P}_{max})}.$$

D_{KL} denotes the Kullback-Leibler divergence. For two probability distribution P and Q, the Kullback-Leibler divergence is defined as: $D_{KL}(\mathbf{P}||\mathbf{Q}) = \sum_i P_i log_2(P_i/Q_i)$. $\mathbf{P}_{min}$ represents the minimum (i.e., no) uncertainty which is obtained when one entry of the probability vector is 1 and all the others 0 (e.g., $\mathbf{P}_{min} = (1,0,\ldots,0)$), and $\mathbf{P}_{max}$ represents the maximum uncertainty which is obtained when a pixel is equally likely to belong to all segments, i.e., $\mathbf{P}_{max} = (1/c,\ldots,1/c)$. The vector $\mathbf{P}_v$ represents the segmentation probability vector for voxel v. The normalization term $D_{KL}(\mathbf{P}_{min}||\mathbf{P}_{max}))$ represents the maximum amount of randomness, which amounts to $log_2(c)$ for c segments. In case of no uncertainty, i.e., $\mathbf{P}_v = \mathbf{P}_{min}$, we obtain that $U(v) = 0$. Likewise, in case of maximum uncertainty, i.e., $\mathbf{P}_v = \mathbf{P}_{max}$, we obtain that $U(v) = 1$. If one wants to distinguish uncertain from certain voxels, one defines an uncertainty threshold δ. All voxels with uncertainty larger than δ are then considered uncertain, the others certain.

5. Uncertainty Estimation for Multi-modal Image Segmentation

We now extend the methods for uncertainty estimation and visualization from uni-modal to multi-modal image segmentation. First, we apply to each modality of the multi-modal imaging data a probabilistic segmentation algorithm to produce a soft classification. The combination of the individual segmentations form an ensemble of classifiers. For the combination, one could use the (weighted) majority vote rule to produce a hard classification. However, we would like to estimate and visualize the uncertainty associated with multi-modal image segmentation using the uncertainty measure presented in the previous section. Hence, we need a soft classification of the multi-modal images. We achieve this by converting the hard segmentation of the ensemble-based model to a probabilistic ensemble-based multi-modal image segmentation. The probabilistic ensemble-based multi-modal segmentation can be computed by using the respective combining rule without the application of the final maximum operator and normalizing the vectors. The probability vectors of the probabilistic version of the combining rule can be computed by assigning:

$$x \mapsto P_i(x), i = 1,\ldots,C$$

where $P_i(x)$ in the probabilistic combining rules represents the probability that voxel x belongs to class i according to the corresponding combining rule. I.e., $P_i(x)$ represents the i^{th} entry of the probability vector of the probabilistic ensemble segmentation result at each voxel x. Extending the concept of the (weighted) majority vote rule, we can compute the probabilities as follows:

- Probabilistic Majority Vote Rule:

$$P_i(x) = \frac{\sum_{j=1}^{L} \Delta_{ij}}{\sum_{k=1}^{C}\sum_{j=1}^{L} \Delta_{kj}}$$

- Probabilistic Weighted Majority Vote Rule:

$$P_i(x) = \frac{\sum_{j=1}^{L} \omega_j \Delta_{ij}}{\sum_{k=1}^{C}\sum_{j=1}^{L} \omega_j \Delta_{kj}}.$$

In a similar way, the concept can be extended for the other combining rules such as the sum, maximum, minimum, and

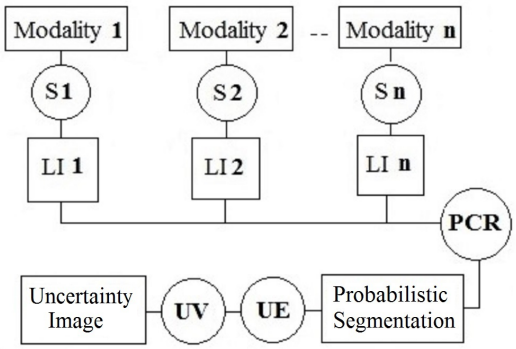
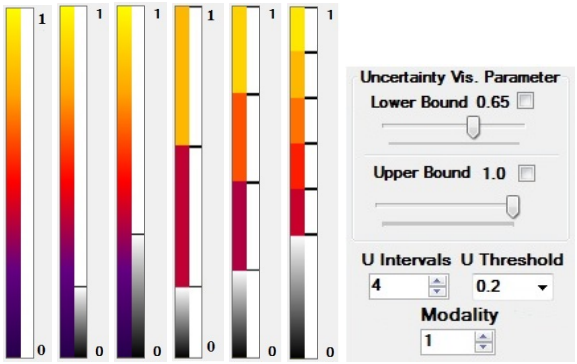


Figure 1: The proposed uncertainty estimation and visualization model for an ensemble-based multi-modal image segmentation (Si: Segmentation i, LI i: Labeled Image i, PCR: Probabilistic Combining Rule, UE: Uncertainty Estimation, UV: Uncertainty Visualization).

median rules. Throughout this paper, we focus on the majority vote rule, as it fits the applications we address. Once we have the probabilistic multi-modal segmentation result, we can directly apply the uncertainty measure presented in the previous section to estimate and visualize the uncertainty associated with each voxel. The estimated uncertainty using majority vote rule has zero uncertainty when all individual segmentations of the different modalities agree in their decision (label). The uncertainty increases as the agreement among the segmentations decreases and reaches the maximum uncertainty (i.e., 1) when each modality has a different decision (maximum diversity). Figure 1 shows the proposed system to estimate and visualize the uncertainty associated with an ensemble-based multi-modal image segmentation.

6. Interactive Visual Analysis

The visualization of the uncertainty in the multi-modal segmentation result is performed using a color-coding of the original 2D images. For the color-coding, we follow the ideas presented by Al-Taie et al.'s [ATHL14b], but we integrate them into an interactive visual analysis tool. The color-coding uses a color map with multiple hues but increasing luminance. The hues vary from purple over red and orange to yellow, see Figure 2(a). So, the first option is to just apply this color map to the uncertainty estimated for each pixel of the 2D image. Alternatively, we can combine the color mapping with the uncertainty threshold δ. The threshold is defined by the user. All pixel with uncertainty values above the threshold use the color map as described above to encode the amount of uncertainty of the pixel. All pixels with uncertainty values below the threshold use a greyscale color mapping of the intensity value of the pixel. Figures 2(b) and (c) show examples of the applied color mapping. Finally, the chosen color map can be discretized to a banded color map. The amount of bands is, again, defined by the



(a) (b) (c) (d) (e) (f) (g)

Figure 2: Uncertainty visualization interacting tools: uncertainty color-mapping legends for continuous mapping (a), continuous mapping with uncertainty threshold = 0.2 (b) and 0.35 (c), banded mapping with uncertainty threshold 0.2 and 2 bands (d), threshold 0.25 and 3 bands (e), and threshold 0.35 and 5 bands; (g) interactive tool panel for parameter setting of uncertainty threshold, number of bands, lower and upper bound, and modality.

user. The banded color map can also be combined with the thresholding we just explained for the continuous color map. Figures (d)-(f) show respective banded color maps with different number of bands and different thresholds δ.

The user can interactively define the different visualization parameters that are shown in Figure 2(g). Obviously, the user can choose the uncertainty threshold δ (none if $\delta = 0$) and the number of bands (U Intervals) in case the banded color maps are used. The color map is, by default, applied to the interval $[0,1]$ for the first visualization option and $[\delta,1]$ for the second and third visualization option (i.e., when the continuous or the banded color map is combined with the original image intensities through thresholding), but can be applied to a subinterval of $[0,1]$ defined by a lower and upper bound. A modality parameter allows for switching between the different image modalities to be combined with the continuous and banded color maps. These simple interacting tools allow for different visual interpretation, better analysis, and highlighting the different categories of uncertain areas.

7. Uncertainty-based Multi-modal Image Segmentation

In addition to estimating the uncertainty and visually analyzing it the main goal of the proposed system is to enhance the visual analysis through the combined information integrated from the different modalities which is not possible using only a single modality.

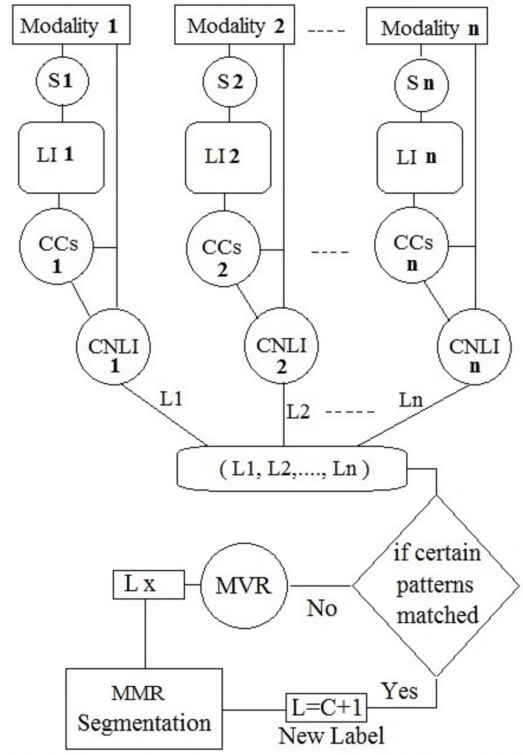
The main reason for using multi-modal imaging is that a single modality is insufficient to capture all the necessary information. Often, there are certain structures that cannot

be distinguished in one modality, but can very well be distinguished in another modality, where the other modality then has other difficulties. Consequently, there may be regions that have conflicting classifications in different modalities. Those would show up as areas of high uncertainties. Hence, we could easily identify those regions using our approach. Figure 3(a) shows the proposed ensemble-based multi-modal segmentation approach that incorporates the complementary information integrated from the different modalities to segment suspicious or special regions. We refer to this segmentation result obtained by our model as multi-modal rule (MMR) segmentation. Figure 3(b) shows how the proposed MMR segmentation can be combined with the proposed uncertainty estimation system to segment and highlight the suspicious or special regions that cannot be detected using only a single modality. In the application scenarios presented in the subsequent two section, we show how we can make use of our concept to detect MS tumors from multi-modal MR brain imaging data and tumor growth in brain images taken at different points in time.

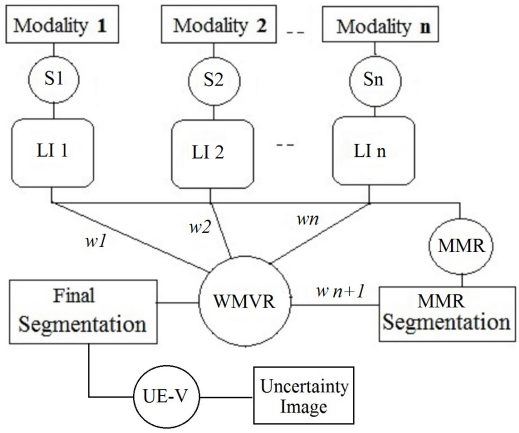
8. Application Scenario to Multi-modal Brain MRI for MS Lesion Segmentation

The characteristics of an MS lesion tissue in brain MR images is similar to healthy brain tissues, which makes the task of segmenting it as a separate tissue difficult when using state-of-the-art segmentation approaches. Luckily, the MS lesion has different characteristics for different imaging modalities. For example, the MS lesion in PD- , T1-, and T2-weighted MR imaging could have intensities similar to cerebrospinal fluid (CSF), gray matter, and CSF, respectively, or CSF, white matter, and CSF, respectively, or sometimes gray matter, white matter, and CSF, respectively, or even as CSF, white matter, and grey matter, respectively (see the individual segmentations in Figure 6). Hence, when using uni-modal image segmentations, one cannot separate the lesion from other brain tissues. However, when combining the conflicting information from the uni-modal image segmentation results in the sense of an ensemble of classifiers, one can use the complementary information to draw correct conclusions. To segment and highlight the MS lesion area using the majority voting combining rule, we perform the following steps:

First, we produce a new segmentation result in addition to the uni-modal image segmentation results. We call the additional segmentation result multi-modal rule (MMR) segmentation (see Figure 3(a)). It is based on the segmentations of the three modalities using the following procedure: 1) Generate a probabilistic segmentation using fuzzy c-means (FCM) or one of its variants to segment each of the three modalities assuming only normal brain tissues, i.e., cerebrospinal fluid (CSF), gray matter (GM), white matter (WM), and background (Bg). Hence, we generate a segmentation of each modality using exactly four classes. 2) Calcu-



(a) MMR segmentation



(b) The combined system

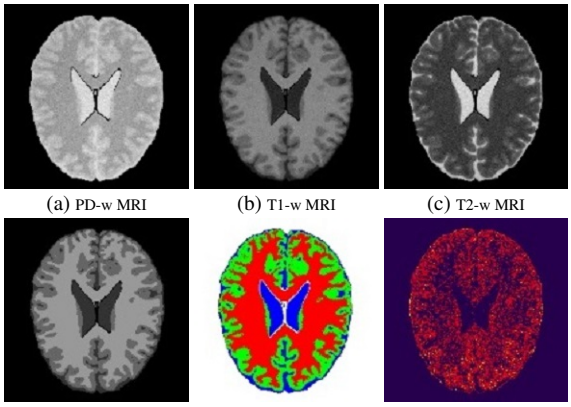
Figure 3: The proposed multi-modal image segmentation rule (MMR) (a), and the combined MMR with MM uncertainty estimation system (b) (Si: Segmentation i, LI i: Labeled Image i, CCs: Compute Centroids, CNLI: Compute New Labeled Image, (W)MVR: (Weighted) Majority Vote Rule, MMR: Multi-Modal Rule, UE-V: Uncertainty Estimation and Visualization).

late the centroids of the corresponding tissue classes in the segmentation results for each of the three modalities. 3) For each modality, assign to each voxel the tissue whose centroid is closest to the voxel's intensity according to the corresponding uni-modal segmentation result. 4) Apply to each voxel the ensemble voting procedure for the three modalities to assign the winning class as the label of the new MMR segmentation. However, if the uni-modal segmentation results deliver the pattern for MS lesions (i.e., one of the combinations listed above), the MMR segmentation assigns a new (fifth) class label "MS lesion". 5) Finally, we correct outlier voxels, i.e., voxels whose label differs from all surrounding neighboring voxels, in a post-segmentation step.

Second, we use our uncertainty visualization methods to highlight the MS lesion area. To assure that the uncertainty is sufficiently high for the MS lesion area and sufficiently low for other tissue types, we use a weighted majority vote combining rule, where the MMR segmentation result gets double the weight (weight 2) as the other three uni-modal segmentation results (weight 1) (see Figure 3(b)). Thus, we reduce the uncertainty for the normal tissue types, as we have a higher vote for the already winning tissue, and we increase the uncertainty of the MS lesion area, as we raise the voting probability for the newly inserted class leading to an approximately similar voting probability as other (two or three) classes.

In the following, we present results that document that this scenario succeeds in segmenting the MS lesion with acceptable accuracy with only a few false negatives and a few false positives voxels around the MS lesion area (i.e., inside the safety margin of lesion treatment).

In the first experiment, we apply the proposed ensemble-based multi-modal image segmentation with its uncertainty estimation and visualization to a simulated MR images of PD-, T1-, and T2-weighted modalities for a healthy brain without MS lesions [MNI97]. As individual classifiers for the three modalities, we use the standard fuzzy c-means (FCM) algorithm introduced by Bezdek [Bez81]. In our proposed system, it is possible to use the same or different approaches of probabilistic or FCM segmentation as individual classifiers. Also, the ensemble segmentation and its uncertainty estimation can be done using the other known combining rules such as median and sum rules (through this paper we use only the majority vote rule). Figure 4 shows the segmentation result in (e), and the uncertainty visualization in (f) of the proposed methods for an ensemble-based multi-modal image segmentation with three simulated brain MRI modalities (PD-weighted in (a), T1-weighted in (b), and T2-weighted in (c)). Only the result when using the majority voting rule is shown. Figure 4(d) shows the ground truth segmentation. We can observe that pixels in the background area have zero uncertainty (shown in dark purple) as all three modalities agree for this area. In the brain area, a large portion of the pixels have low uncertainty of around 0.4 (shown



(a) PD-w MRI (b) T1-w MRI (c) T2-w MRI

(d) Ground Truth (e) MV Rule (f) Uncertainty

Figure 4: Uncertainty visualization of an ensemble-based multi-modal segmentation result for simulated of (a) PD-weighted, (b) T1-weighted, and (c) T2-weighted MR images: (d) ground truth, (e) segmentation result, and (f) uncertainty visualization using the proposed methods with the majority voting as combining rule.

in red), a small portion have zero uncertainty, and very few pixels have high uncertainty of around 1.0 (shown in yellow). The concentration of pixels with low and zero uncertainty differs for the different tissues as the distribution of the aggregated uncertainty (such as noise or partial volume effect) differs for them. The red color can be interpreted that one of the three modalities makes an error, while yellow color (high uncertainty where the user should be careful about) can be interpreted that the three modalities have three different decisions.

In the next experiments, we test our approach to segment the brain MS lesion and highlight it in the uncertainty visualization as suspicious tissue in the multi-modal images. First, we apply the proposed methods to synthetic images that simulate normal brain tissues and MS lesion (circle) intensities in PD-, T1-, and T2-weighted MRI modalities, respectively. All the three modalities are corrupted with mixed noise (Gaussian and salt-and-pepper noise). In Figure 5, we can observe how the proposed segmentation method succeeds in segmenting the MS lesion (shown in purple) with high accuracy (we got 100% segmentation accuracy for the tumor) and how the uncertainty visualization methods highlight the lesion area successfully. Figures 5(a)-(c) show the three imaging modalities, (d) the ground truth. The MMR segmentation is shown in (e) and the majority voting output in (f) using five distinct colors for the five labels. The uncertainty according to the color map in Figure 2(a) is visualized in Figure 5(g) where the light orange represents the high uncertainty area. Figure 5(i) shows the uncertainty visualization using the color map in Figure 2(b) for threshold $\delta = 0.2$. Figure 5(h) shows the segmentation result visualization combined with the estimated uncertainty where the

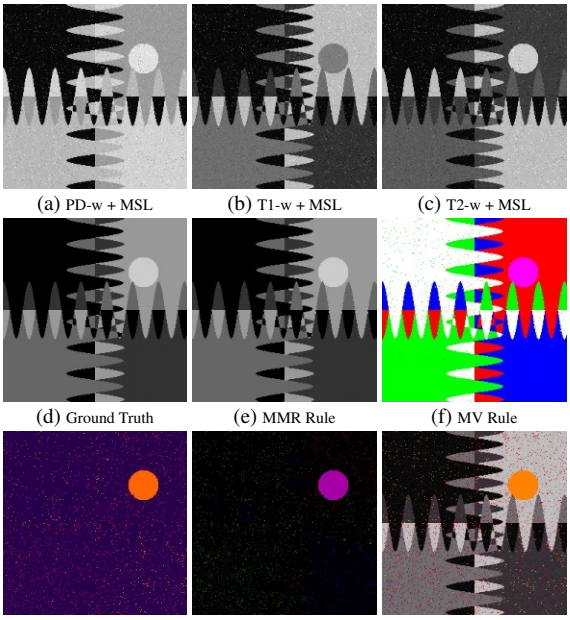


Figure 5: Segment and highlight MS lesion in an ensemble-based multi-modal segmentation result for synthetic images that simulate (a) PD-weighted, (b) T1-weighted, and (c) T2-weighted MR images: (d) ground truth, (e) MMR segmentation result, (f) majority voting segmentation result, (g) uncertainty visualization of segmentation result, (h) combined uncertainty-segmentation visualization, and (i) uncertainty overlaid with original image.

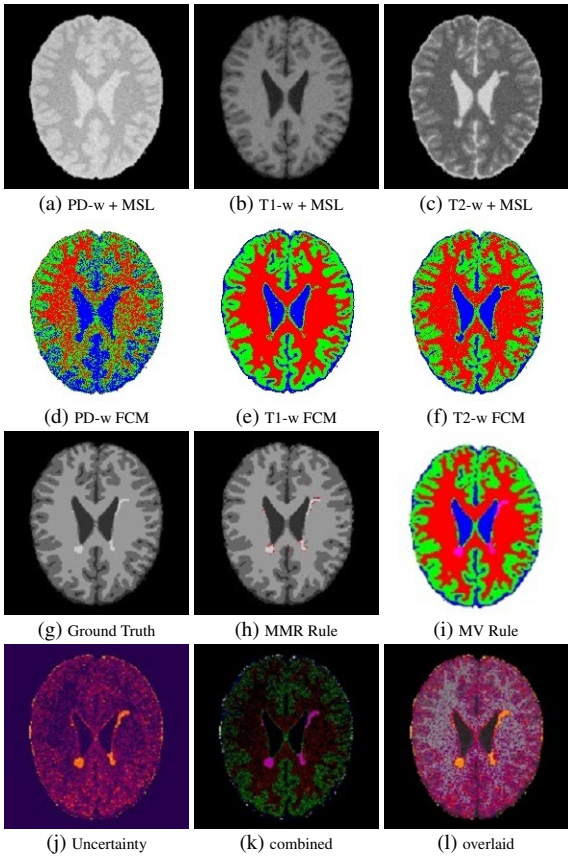


Figure 6: Segment and highlight MS lesion in an ensemble-based multi-modal segmentation result for simulated (a) PD-weighted, (b) T1-weighted, and (c) T2-weighted MR images obtained from BrainWeb: (d), (e), and (f) the FCM segmentation results for the images in (a), (b), and (c), respectively, (g) ground truth, (h) MMR segmentation result (false positives shown in red), (i) majority voting segmentation result, (j) uncertainty visualization of segmentation result, (k) combined uncertainty-segmentation visualization, and (l) uncertainty overlaid with original image.

associated uncertainty is assigned as the intensity of segmentation' color map to highlight the uncertain pixels.

Then, we apply the same experiment on simulated brain MRI PD-, T1-, and T2-weighted modalities with MS lesion obtained from BrainWeb [MNI97]. We apply the same procedure and get comparable results as shown in Figure 6. We can observe that the MS lesion is segmented (shown in purple in Figures 6 (g) and (k)) with high accuracy (see Table 1 slice No. 96) and few false positives (shown in red) surrounding the MS lesion area (see Figure 6(h)).

Table 1 provides quantitative assessments of the MS lesion segmentation accuracy and its false positive ratio with lesion size (in pixels) for 10 slices from the dataset in Figure 6. The segmentation accuracy is computed by the number of correctly classified voxels over the lesion size. The false positive ratio (FP) is computed by the number of false positives divided by the slice size.

Figure 7 shows examples of how the segmentation result and the estimated uncertainty can be visually and interactively analyzed using different visualization strategies that support different analysis tasks. The uncertainty color mapping with MS lesion (for slice No. 101 in Table 1) is highlighted shown in Figure 7 (a) for the whole 2D image, (e)

Table 1: Examples of MS lesion segmentation accuracy (SA) and false positive ratio (FP) using the proposed methods with MS lesion size for selected slices from dataset in Figure 6.

Slice No.	SA %	FP Ratio	MSL size
94	0.8941	0.00109	85
95	0.9304	0.00127	115
96	0.9877	0.00140	163
97	0.8961	0.00096	154
98	0.8758	0.00162	153
99	0.9461	0.00193	167
100	0.9682	0.00246	189
101	0.9419	0.00208	224
102	0.9638	0.00142	221
103	0.9687	0.00229	160

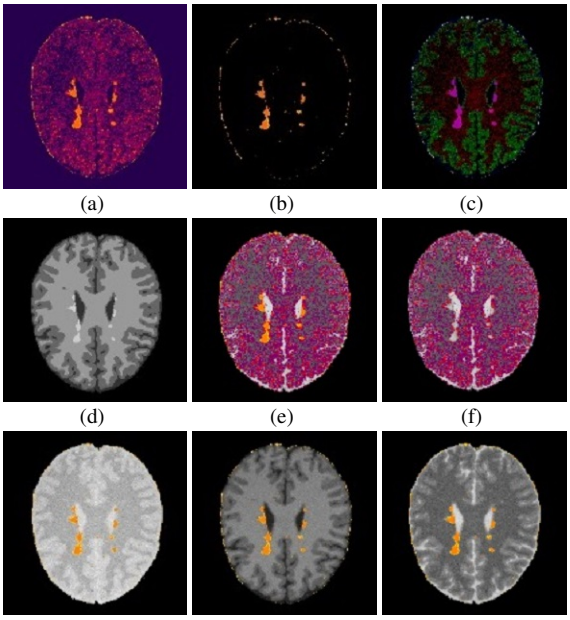


Figure 7: Interactive visualization tool provides several methods to visually highlight the segmentation results for different analysis tasks: The uncertainty color mapping with MS lesion is highlighted for the whole 2D image (a), filtered by the lower and upper bound interacting tool with values (0.6 and 0.99) (b), overlaid with T2-weighted modality (e), and combined with segmentation result visualization (MS lesion shown in purple) (c); the same visualization in (b) overlaid with the three modalities in (g), (h), and (i), ground truth (d).

overlaid with T2-weighted modality, and (c) combined with segmentation result visualization (MS lesion shown in purple). Figures 7 (b), (g), (h), and (i) show the uncertainty mapping filtered by the lower and upper bound interacting tool with values (0.6 and 0.99). We can observe how the modality interacting tool shown in Figure 2(g) is used in selecting the desired modality to overlay the uncertainty mapping with it in Figures 7(g), (h), and (i). In Figure 7 (f), different settings of the lower and upper bounds (0.26 and 0.59) are used to highlight different categories of uncertain areas for comparison. The ground truth is shown in Figure 7 (d).

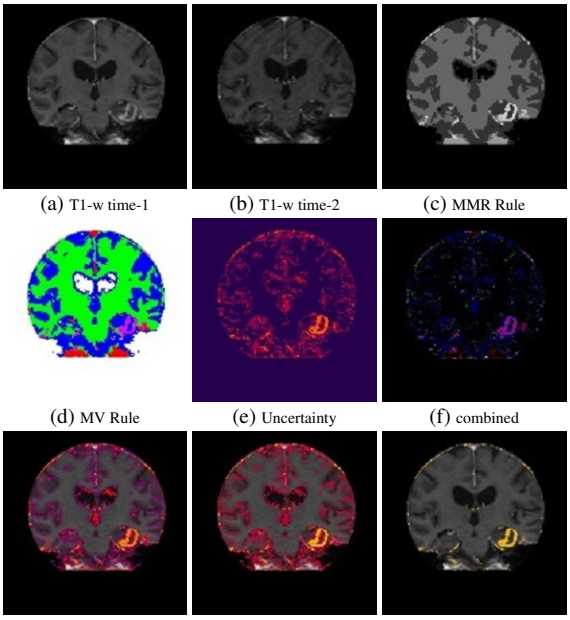
9. Application Scenario to Multiple Time Points for Tumor Growth Analysis

We use the same scenario as above to highlight the difference of the brain tumor size in real MR images of the same subject taken at two different points in time dataset taken from IBSR [IBS96]). In this application scenario, the two T1-weighted images are segmented and the tumor can be labeled as a separate class. Hence, the goal here is different than in the preceding section. Instead of combining different modalities for an improved segmentation, we are now combining two different points in time to estimate how much and where the tumor grew or shrank over time. The individual segmentation results of the two time steps would disagree in the area of growth/shrinkage. Hence, it is expected to be an area of high uncertainty and we can apply the same procedure as above for producing the desired result.

Figure 8 shows the result of the segmentation and uncertainty estimation. The tumor at time-1 (a) is not present at time-2 (b). The MMR rule detects the shrinkage area as a separate class (c). The segmentation using the majority voting rule shows the class of shrinkage area in purple (d). The uncertainty visualization conveys that the shrinkage area is the area of highest uncertainty (e). The segmentation result for uncertain pixels is shown in (f). Next, an uncertainty visualization using a banded color map with four bands (g) and two bands (h), where certain pixels show pixel intensities, is shown. Finally, we show the uncertainty visualization restricted to uncertainties out of interval $[0.6, 1]$ in Figure 8(i). The uncertainty mapping overlaid with the image at time-1 in Figures 8(g), (h), and (i). We can observe that both the segmentation and uncertainty visualization succeed in estimating and highlighting the shrinkage in tumor size accurately.

10. Discussion and Conclusion

The imaging technologies have witnessed a growing number of advancements in recent years leads to improve imaging features and accuracies. However, every modality of imaging has its own practical limitations due to the nature of the underlying organ or tissue and no single modality can capture all details necessary for diagnosis and analysis. This observation, in addition to the diversity of features and the complementary information provided by combining different modalities, enforces the need of using multiple imaging modalities. Multi-modal approaches have proven to be more efficient and accurate when compared to uni-modal approaches. The combined information and the accumulated evidences from the different modalities can reduce the uncertainty in the decision-making and analysis process. Recently, several approaches have been proposed to estimate the uncertainty in uni-modal image segmentation decisions and visually convey this information to the medical experts that examine the image segmentation results. In this work, we presented an approach to extend uncertainty estimation and visualization methods to multi-modal image segmentations. We use the concept of ensemble of classifiers to combine probabilistic uni-modal image segmentation results. The uncertainty is estimated using a measure that is based on the Kullback-Leibler divergence. We apply our approach to the segmentation of Multiple Sclerosis (MS) lesions from multiple MR brain imaging modalities. Moreover, we demonstrate how our approach can be used to estimate and visualize the growth of a brain tumor area for imaging data taken



(a) T1-w time-1 (b) T1-w time-2 (c) MMR Rule

(d) MV Rule (e) Uncertainty (f) combined

(g) overlaid (h) overlaid (i) overlaid

Figure 8: Estimate and highlight brain tumor shrinkage in an ensemble-based multi-time image segmentation result for real T1-weighted MR images at (a) time1, and (b) time2: (c) MMR segmentation result, (d) majority voting segmentation result, (e) uncertainty visualization, (f) segmentation result for uncertain pixels, (g) uncertainty visualization with threshold $\delta = 0.2$ using banded color map with four bands, (h) uncertainty visualization with threshold $\delta = 0.2$ using banded color map with two bands, and (i) uncertainty visualization for pixels with uncertainty out of interval $[0.6, 1]$.

at multiple points in time. Both the MS lesion and the area of tumor growth are detected as areas of high uncertainty due to different characteristics in different imaging modalities and changes over time, respectively.

References

[AMBdS09] ARTAECHEVARRIA X., MUÑOZ-BARRUTIA A., DE SOLORZANO C. O.: Combination strategies in multi-atlas image segmentation: Application to brain mr data. *IEEE Transactions Medical Imaging 28*, 8 (2009), 1266–1277. 3

[ATHL14a] AL-TAIE A., HAHN H. K., LINSEN L.: Uncertainty-aware Ensemble of Classifiers for Segmenting Brain MRI Data. In *4th Eurographics Workshop on Visual Computing for Biology and Medicine 2014* (Vienna, Austria, 2014), Viola I., Buehler K., Ropinski T., (Eds.), Eurographics Association, pp. 41–50. 2, 3, 4

[ATHL14b] AL-TAIE A., HAHN H. K., LINSEN L.: Uncertainty estimation and visualization in probabilistic segmentation. *Computers & Graphics 39*, 0 (2014), 48 – 59. 2, 3, 4, 5

[Bez81] BEZDEK J.: Pattern recognition with fuzzy objective function algorithms. *Plenum, NY.* (1981). 7

[Die00] DIETTERICH T. G.: Ensemble methods in machine learning. In *Proceedings of the First International Workshop on Multiple Classifier Systems* (London, UK, UK, 2000), Springer-Verlag, pp. 1–15. 2

[FJ05] FRED A., JAIN A.: Combining multiple clusterings using evidence accumulation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 27*, 6 (Jun 2005), 835–850. 2

[IBS96] IBSR: The internet brain segmentation repository (ibsr), available since 1996. Available at http://www.cma.mgh.harvard.edu/ibsr/, access time: on October 2012, 1996. 9

[JD14] JAMES A., DASARATHY B. V.: Medical image fusion: A survey of the state of the art. *Information Fusion* (Mar 2014). 1, 3

[KHDM98] KITTLER J., HATEF M., DUIN R. P. W., MATAS J.: On combining classifiers. *Pattern Analysis and Machine Intelligence, IEEE Transactions on 20*, 3 (Mar 1998), 226–239. 2, 4

[LvdHK*10] LANGERAK R., VAN DER HEIDE U. A., KOTTE A. N. T. J., VIERGEVER M. A., VAN VULPEN M., PLUIM J. P. W.: Label fusion in atlas-based segmentation using a selective and iterative method for performance level estimation (simple). *IEEE Transactions Medical Imaging 29*, 12 (2010), 2000–2008. 3

[Mig10] MIGNOTTE M.: A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation. *IEEE Transactions on Image Processing 19*, 6 (2010), 1610–1624. 2

[MNI97] MNI: Brainweb, simulated brain database. Available at http://www.bic.mni.mcgill.ca/brainweb/, access time: on November 2012, 1997. 7, 8

[PGA13] POTTER K. C., GERBER S., ANDERSON E. W.: Visualization of uncertainty without a mean. *IEEE Computer Graphics and Applications 33*, 1 (2013), 75–79. 3

[PNS13] PACI M., NANNI L., SEVERI S.: An ensemble of classifiers based on different texture descriptors for texture classification. *Journal of King Saud University - Science 25*, 3 (2013), 235 – 244. 2

[PRH10] PRASSNI J., ROPINSKI T., HINRICHS K.: Uncertainty-aware guided volume segmentation. *IEEE Transactions on Visualization and Computer Graphics 16*, 6 (2010), 1358–1365. 3

[RM05] ROHLFING T., MAURER C. R. J.: Multi-classifier framework for atlas-based image segmentation. *Pattern Recognition Letters 26*, 13 (2005), 2070 – 2079. 2, 3

[RPHL14] RISTOVSKI G., PREUSSER T., HAHN H. K., LINSEN L.: Uncertainty in medical visualization: Towards a taxonomy. *Computers & Graphics 39*, 0 (2014), 60 – 73. 3

[SMH10] SAAD A., MÖLLER T., HAMARNEH G.: Probexplorer: Uncertainty-guided exploration and editing of probabilistic medical image segmentation. *Computer Graphics Forum 29*, 3 (2010), 1113–1122. 3

[WZW04] WARFIELD S. K., ZOU K. H., WELLS W. M.: Simultaneous truth and performance level estimation (staple): An algorithm for the validation of image segmentation. *IEEE Transactions Medical Imaging 23* (2004), 903–921. 3