

Conflict Management in the Fusion of Complementary Segmentations of Deformed Kidneys and Nephroblastoma

Lisa Corbat^{a,*}, Julien Henriet^a, Jean-Christophe Lapayre^a

^a*FEMTO-ST Institute, DISC, CNRS, Univ. Bourgogne-Franche-Comté, 16 route de Gray, 25030 Besançon, France*

Abstract

The fusion of multiple segmentations aims to improve their accuracy in order to make them exploitable. However, conflicts may appear. In this paper, two conflict-management models are proposed for the fusion of complementary segmentations. This conflict-management and fusion procedure, integrated into the SAIAD project, carries out the fusion of deformed kidneys and nephroblastoma using the combination of six independent methods. These methods are based on different criteria, like the adjacent segmented slices, the variation of information, the Dice, the neighbouring labels, the pixel intensity by scanner images, and the fully connected CRFs. The performances of our fusion models was evaluated on 139 scans for three patients with nephroblastoma, and the results demonstrate its effectiveness and the improvement of the resulting segmentations.

Keywords: Fusion, Conflict management, Segmentation, Cancer tumour

1. Introduction

Nephroblastoma, also called Wilms tumour, is the abdominal tumour the most frequently observed in children (generally 1- to 5-years-old boys and girls). This cancer disease represents 5% to 14% of malignant paediatric tumours. As

*Corresponding author, email adress: lisa.corbat@univ-fcomte.fr
Email addresses: lisa.corbat@univ-fcomte.fr (Lisa Corbat),
julien.henriet@univ-fcomte.fr (Julien Henriet),
jean-christophe.lapayre@univ-fcomte.fr (Jean-Christophe Lapayre)

5 indicated by its name, this type of tumour is situated in the kidney. Most
of the time, its initial diagnosis is based on imaging. Generally, ultrasound
observations are planned first in order to confirm its existence and approximate
its position. Then, a medical scanner provides its position, and the healthy
tissues and organs are reached with higher accuracy. Radiologists and surgeons
10 need 3-dimensional (3D) representation of the tumour and the border organs
in order to establish the diagnosis, plan surgery (estimated quantity of blood,
specialized equipment required, estimation of the duration of the surgery, etc.),
and also guide the actions of the surgeon during the surgery.

This 3D representation is currently done through manual segmentations,
15 which is a time-consuming task. The French-Swiss border project SAIAD (Au-
tomated Segmentation of Medical Images Using Distributed Artificial Intelli-
gence) aims at obtaining automatic segmentations of the kidney tumours and
nephroblastoma through artificial intelligence methods. There is a lot of re-
search in publications on automatic segmentation using AI like Deep Learning
20 with Convolutional Neural Networks (CNNs) (Long et al. (2015); Ronneberger
et al. (2015)), Random Markov Fields (Kato et al. (2012)), and also the Fully
Connected Conditional Random Fields (CRFs) Krähenbühl & Koltun (2011)
coupled with CNNs to refine segmentations Chen et al. (2018); Noh et al. (2015);
Kamnitsas et al. (2017), genetic algorithm (Tosta et al. (2017); Khan & Jaffar
25 (2015)), and Case Based Reasoning (CBR) (Marie et al. (2018); Kausar et al.
(2016); Frucci et al. (2008)). But a single method is not efficient enough to
achieve a correct segmentation of all structures on an image. Each structure
can be calculated using one optimal method of its own in order to obtain its
best segmentation. Thus, in the SAIAD project, a method based on CBR and
30 region growing is used in order to segment the kidney deformed by nephrob-
lastoma on each of the patients' 2D images (Marie et al. (2018)), and the OV^2
ASSION method, based on CNN, is used to segment nephroblastoma (Marie
et al. (2019)).

The next step is the fusion of the nephroblastoma and kidney segmenta-
35 tions in order to obtain the final segmentation of a scanner image. This fusion

can cause conflicts between pixels, and a conflict-management protocol must determine the true labels. Most of the conflicts are found at the intersection of the pathological kidney and the nephroblastoma. These border zones are the most difficult to label because there are no clear and/or visible boundaries
40 on the scanned images. Only advanced algorithms or an experienced expert (radiologist or surgeon) could succeed in determining these boundaries. In our approach, we consider all the segmentations of deformed kidneys and nephroblastoma calculated for one patient in order to resolve the conflicts using a fusion and conflict-management model based on several combined methods.

45 After the study of works related to segmentation fusion in Section 2, this paper presents, in Section 3, the fusion and conflict-management model we have designed for kidney and nephroblastoma segments. Section 4 explains the experiments for our methods and the discussions around the results.

2. Related work

50 Information fusion has emerged in order to manage data from multiple sources. The fusion of information can also be applied to the fusion of images. Image fusion is applied in a wide scope of applications, including remote sensing (Ghassemian (2016)), surveillance (Jin et al. (2017)), photography (Ma et al. (2015)), and medicine (James & Dasarathy (2014); Liu et al. (2014)). Several fusion techniques are used, such as the wavelet transform (Li et al. (1995)),
55 image pyramids decomposition (Mertens et al. (2007)), etc. These methods are categorized according to different fusion levels : pixel-level fusion, feature-level fusion, and decision-level fusion (Pohl & Van Genderen (1998)). However, there are also other methods specific to the merging of multiple segmentations.

60 The main strategy for segmentation consists in determining the best algorithm with optimal parameters, but these algorithms can be costly and complex. Another strategy aims at combining several segmentations into a single, consensual one. The best strategy or the best criterion for fusion is then applied to the consensual segmentation. Multiple segmentations are then calculated applying

65 different segmentation methods or using the same method but with different
parameters. Many techniques use the over-segmentation method in order to
obtain a large number of regions and a large amount of information about each
segmentation (Huang et al. (2016); Li et al. (2012)). Other fusion methods are
applied to the multi-atlas based segmentation (Aljabar et al. (2009); Nguyen
70 et al. (2015)). The atlas, which contains a set of images and corresponding
segmentations, is used to determine the images that are the most similar to
the target image and fuse the segmentations of the images chosen to obtain the
segmentation of the target image.

It is possible to group most of these fusion strategies into different ap-
75 proaches. The first and most intuitive approach is the democratic approach
with fusion voting techniques like majority vote, global weighted vote, local
weighted vote, and so on (Artaechevarria et al. (2009)). Another approach is
the morphological approach, which uses contours and shapes of the elements
in the segmentations. The Shape-Based Averaging (SBA) method (Rohlfing &
80 Maurer (2007)) uses the signed Euclidean distance maps in order to determine
the segmentation contours of each possible label. F-Measure Martin et al. (2004)
is a criterion used in Mignotte & H elou (2014) which evaluates the quality of
the resulting contours and uses the notions of precision and recall.

Many other methods are based on probabilistic approaches, like the STAPLE
85 method (Warfield et al. (2004)), which estimates the optimal combination of seg-
mentations to fuse, weighting each segmentation according to the estimated per-
formance level using the Expectation-Maximization algorithm (Dempster et al.
(1977)). Mignotte (2010) used a method based on the Probabilistic Rand Index
(PRI) criterion (based on pairwise relationships), as well as the method based
90 on the Variation of Information (VoI) criterion (Mignotte (2014)). This method
uses mutual information metrics and entropies in order to measure the amount
of information lost or gained when turning one label into another. Another
VoI-based fusion procedure is used in Nguyen et al. (2018) in order to merge
multiple atlases of X-ray images. This method is successfully applied to complex
95 bone regions like the patella, talus, and pelvis.

The Global Consistency Error (GCE) criterion is used in Khelifi & Mignotte (2017c) for the fusion of multiple segmentations. This method measures the extent to which one segmentation can be viewed as a refinement of another segmentation, based on segmentation regions. The K-Mean method (Mignotte (2008); Harrabi & Braiek (2012)) aggregates pixels with similar characteristics. A method using weakly supervised trace-norm multi-task learning (Liang & Huang (2018)) consists of considering the segmentation fusion problem as a weakly supervised learning problem, in order to use the information from superpixels. In this method, the multiples segmentations are treated as multiple closely related tasks and utilise multi-task learning methods to evaluate the reliability of the segmentations. The segmentation maps obtained are then aggregated according to a fusion strategy. Finally, another approach uses the spatial and intensity information of an image (i.e. using the 2D information) for the fusion of the MR-T2 brain images segmentation map (Feng et al. (2017)).

Furthermore, a recent multi-objective methods combines several criteria like the combination of the VoI criterion with the F-Measure criterion and the GCE criterion with the F-Measure criterion (Khelifi & Mignotte (2017a,b)), and then a new classification is proposed: mono-objective methods and multi-objective methods.

However, most of these approaches have not been implemented on medical images, and they fuse concurrent segmentations of the same initial image. Only Nguyen et al. (2018) and Feng et al. (2017) use medical images and merges the corresponding segmentations. Our approach aims at carrying out fusion of complementary segmentations, i.e. the fusion of kidney segmentations and tumour segmentations together, and resolving the pixels in conflict (labelled as belonging to the kidney and the tumour in the fused segmentation) by retrieving the patient's 3D information (i.e. all other segmented slices of the patient). Two conflict-management models are defined and presented in the next section of this paper and are based on the set of combined methods using the informations from the Adjacent Segmented Slices (ADS2) method, a method based on VoI (Mignotte (2014)), on Dice, on the spatial and intensity information of one

image (renamed here NandI for Neighbours and Intensity) (Feng et al. (2017)), on a 3D-NandI method which is an improvement of the NandI method and on the Fully Connected CRF (Krähenbühl & Koltun (2011)).

130 **3. Proposed method**

The general architecture of the system designed in the SAIAD project is described in Figure 1. It is composed of three layers. The first one is the data layer, which contains a database for each segmentations system. Each database has access to all the scanned images and corresponding manual segmentations
135 and expert knowledge like the information from scanner images in Dicom files but they have their own case base. The definition of a case can be different from one segmentation system to another. The second layer is the segmentation layer, where images are segmented by artificial intelligence systems. For example, the deformed kidney can be segmented by a CBR system coupled with region
140 growing (Marie et al. (2018)), and the tumour can be segmented by a Deep Learning system (Marie et al. (2019)). At the end of the segmentation process, the system gives two complementary segmentations for each image. Finally, the fusion layer combines the complementary segmentations and resolves the conflicts.

145 This part of the paper presents our conflict-management models and our different methods used within our models.

3.1. Fusion of complementary segmentations

Since nephroblastoma and kidneys are segmented individually, it is necessary to fuse them in order to obtain segmentation with both of these structures. In
150 order to manage the possible conflicting pixels, we need to obtain fused segmentation for all methods used except the method based on the Fully Connected CRF.

As shown in Figure 2, during this first step of fusion, each structure is merged and a new temporary label is assigned to the pixels in conflict (i.e. the pixels
155 labelled as belonging to different structures the complementary segmentations).

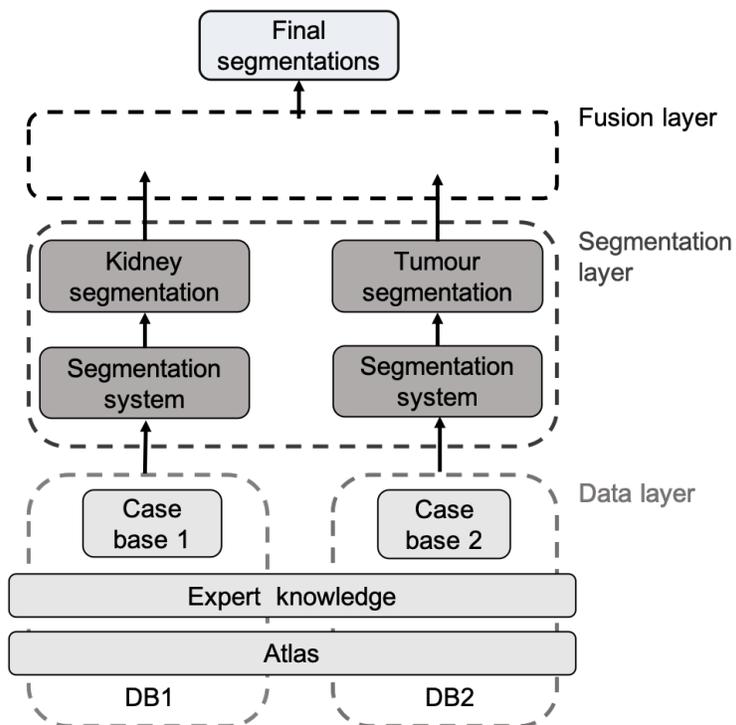


Figure 1: General schema of the SAIAD project.

3.2. Adjacent segmented slices information

The segmentations of adjacent slices may give information about a pixel in conflict. Indeed, we assumed that the segmentations of adjacent slices are relatively similar, since the structures presented have more or less the same shapes and positions. For that reason, we have designed the Adjacent Segmented Slices (ADS2) method for the resolution of conflicts. This method examines the labels of each conflicting pixel in the adjacent slices.

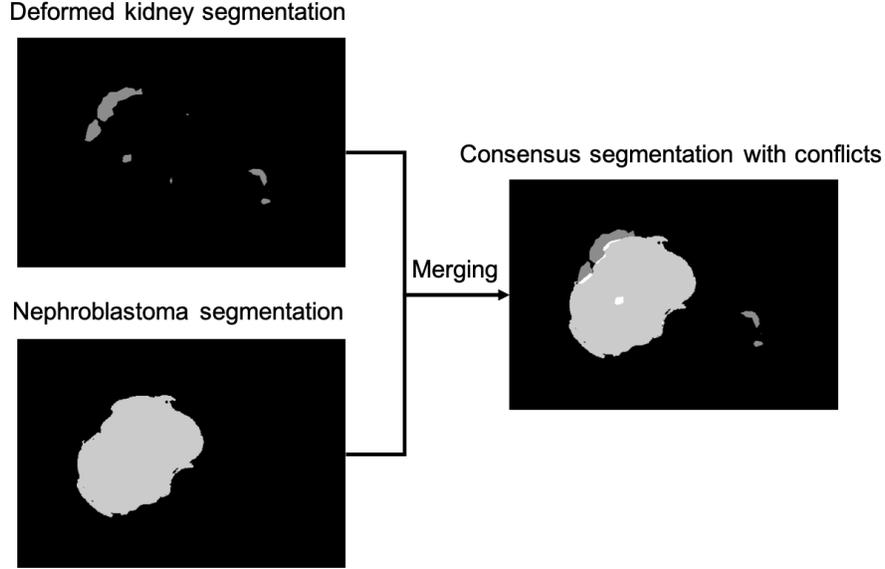


Figure 2: Merging of deformed kidney and nephroblastoma segmentations.

Algorithm 1: ADS2-based conflict resolution method.

```

for each segmented slice  $s$  do
  for each conflicted pixel  $p$  with label  $l_s^p$  do
    if  $l_{s-1}^p$  and  $l_{s+1}^p$  do not conflict then
      if  $l_{s-1}^p == l_{s+1}^p$  then
         $l_s^p = l_{s-1}^p = l_{s+1}^p$ ;
      end
    end
  end
end

```

The pseudo-code of the ADS2 method is shown in Algorithm 1. Considering
 165 a conflicting pixel p labelled with l_s^p in slice s , if the labels l_{s-1}^p and l_{s+1}^p of p in
 the adjacent slices of s are equal to l (i.e $l_{s-1}^p = l_{s+1}^p = l$), then we assume that
 l_s^p is equal to l .

Unfortunately, this method cannot resolve all the conflicts. Indeed, ADS2 can only be applied in specific situations where adjacent pixels are not in conflict. For this reason, we have also designed and implemented other methods.

3.3. Method based on the variation of information criterion

The Variation of Information (VoI) criterion (Meilă (2003)) is based on an information theory that compares two samples in a set of data. It is also used in order to compute the similarity between two segmentations and thus to quantify the quality of a segmentation regarding its ground truth given by an expert (Sathya & Manavalan (2011); Mobahi et al. (2011)). Its formula is based on a probabilistic approach using the entropy and the mutual information of the segmentations to be compared. The VoI score is between 0 and 1, and the lowest scores indicate similar segmentations. The highest scores indicate mostly different segmentations.

In the case of the comparison of two segmentations S and S' , the VoI is defined as:

$$VoI(S, S') = H(S) + H(S') - 2 * I(S, S') \quad (1)$$

where $H(S)$ and $H(S')$ represent the entropy of S and S' , and $I(S, S')$ represents the mutual information between the two segmentations.

The entropy is defined as:

$$H(S) = - \sum_{i=1}^R P(i) \log P(i) = - \sum_{i=1}^R \frac{n_i}{n} \log \frac{n_i}{n} \quad (2)$$

where R is the number of regions of S , and $P(i)$ is the probability that a pixel belongs to the class i based on the total number n of pixels in the segmentation. In imagery, the entropy can be calculated by the normalized histogram of the considered image.

190 The mutual information is also defined as:

$$I(S, S') = \sum_{i=1}^R \sum_{j=1}^{R'} P(i, j) \log\left(\frac{P(i, j)}{P(i)P(j)}\right) = \sum_{i=1}^R \sum_{j=1}^{R'} \frac{n_{ij}}{n} \log\left(\frac{\frac{n_{ij}}{n}}{\frac{n_i}{n} \frac{n_j}{n}}\right) \quad (3)$$

where $P(i, j)$ is the probability that a pixel belongs to the class i of S and to the class j of S' .

In this method, considering a set of L segmented slices $\{S_k\}_{k \leq L}$, the comparison of S to $\{S_k\}_{k \leq L}$ is based on the computation of the mean VoI:

$$\overline{VoI}(S, \{S_k\}_{k \leq L}) = \frac{1}{L} \sum_{k=1}^L VoI(S, S_k) \quad (4)$$

195 The theoretical consensus segmentation is simply the segmentation having the weakest \overline{VoI} among all the possible labels of each pixel:

$$\hat{S}_{\overline{VoI}} = \underset{S \in S_n}{\operatorname{argmin}} \overline{VoI}(S, \{S_k\}_{k \leq L}) \quad (5)$$

We could use the mean VoI metric (see Eq. 4) for conflict management but the computation time would be excessive. Consequently, we use its local expression derived $\Delta \overline{VoI}$ in Eq. 6 used in Mignotte (2014) and Nguyen et al. 200 (2018) where $s : m \rightarrow x$ is the pixel at position s with label m replaced by the label x in S , \mathcal{L}_s^l is the pixel label at position s in the segmentation $S_l \in \{S_k\}_{k \leq L}$, and $n_{m\mathcal{L}_s^l}$ is the number of pixels which have the same label m in S and the label \mathcal{L}_s^l at position s in S_l . In contrast to the mean VoI metric, the maximum

value for the derived mean VoI is searched for.

$$\begin{aligned}
\Delta \overline{VoI}(S, \{S_k\}_{k \leq L})_{s:m \rightarrow x} = & L \cdot \left\{ -\frac{n_m}{n} \log\left(\frac{n_m}{n}\right) - \frac{n_x}{n} \log\left(\frac{n_x}{n}\right) \right. \\
& + \frac{n_m - 1}{n} \log\left(\frac{n_m - 1}{n}\right) + \frac{n_x + 1}{n} \log\left(\frac{n_x + 1}{n}\right) \left. \right\} \\
& - 2 \cdot \sum_{l=1}^L \left\{ \frac{n_m \mathcal{L}_s^l}{n} \log\left(\frac{n_m \mathcal{L}_s^l}{n} \cdot \frac{n}{n_m} \cdot \frac{n}{n_{\mathcal{L}_s^l}}\right) \right. \\
& + \frac{n_x \mathcal{L}_s^l}{n} \log\left(\frac{n_x \mathcal{L}_s^l}{n} \cdot \frac{n}{n_x} \cdot \frac{n}{n_{\mathcal{L}_s^l}}\right) \\
& - \frac{(n_m \mathcal{L}_s^l - 1)}{n} \log\left(\frac{(n_m \mathcal{L}_s^l - 1)}{n} \cdot \frac{n}{n_m - 1} \cdot \frac{n}{n_{\mathcal{L}_s^l}}\right) \\
& \left. - \frac{(n_x \mathcal{L}_s^l + 1)}{n} \log\left(\frac{(n_x \mathcal{L}_s^l + 1)}{n} \cdot \frac{n}{n_x + 1} \cdot \frac{n}{n_{\mathcal{L}_s^l}}\right) \right\} \quad (6)
\end{aligned}$$

205

The maximization of the Eq. 6 can be realized by an ICM algorithm Besag (1986) (Iterative Conditional Modes) an iterative steepest local energy descent. Like the pseudo-code of the VoI-based method shown in Algorithm 2, for each conflicting pixel, the derived mean VoI is calculated for each possible label. The final label will be the one with the highest value. As the adjacent slices are similar, this method only takes into account the segmentations of the n_n adjacent slices of S . For example, for $n_n = 2$ and $S = 4$, segmentations $\{2, 3, 4, 5, 6\}$ are taken into account. This step is repeated until $p < T_{max}$ and for each segmented slice.

210

Algorithm 2: VoI-based conflict resolution method.

Data: $\{S_k\}_{k \leq L}$: The set of L segmented slices n_n : The number of adjacent slices taken into account ε : The set of possible labels for a pixel T_{max} : The maximal number of iterations

```
215 for each segmented slice  $S$  in  $\{S_k\}_{k \leq L}$  do
     $\{S_k\}_{k \in \{k-n_n, \dots, k+n_n\}}$ : Set of neighbor fused segmentations to  $S$ 
    segmentation of one patient with conflicting pixels
    while  $i < T_{max}$  do
        for each conflicted pixel  $p$  with label  $l_s^p$  of segmentation  $S$  do
            Compute
             $VoIMax = \arg \max_{x \in \varepsilon} \Delta \overline{VoI}(S, \{S_k\}_{k \in \{k-n_n, \dots, k+n_n\}})_{l_s^p \rightarrow x}$ 
            Replace label  $l_s^p$  by label  $x$  of  $VoIMax$  at the segmentation  $S$ 
        end
         $i = i + 1$ 
    end
end
```

3.4. Method based on the dice criterion

The Dice-based method is inspired by the VoI-based one: the VoI criterion is replaced by the Dice. This metric is commonly used in the medical field. The F-Measure (Martin et al. (2004)) is also based on the Dice for the segmentation fusion (Mignotte & Hérou (2014)) at the region level. In the present approach, 220 the Dice criterion is used for the segmentation fusion at the pixel level. The Dice is defined as:

$$Dice(S, S') = \frac{2 * TP_{S, S'}}{2 * TP_{S, S'} + FP_{S, S'} + FN_{S, S'}} \quad (7)$$

where $TP_{S, S'}$ is the number of true positive pixels between S and S' (pixels labelled true in both of the segmentations). FP is the number of false positive

225 pixels (pixels labelled true in S and false in S') and FN the number of false
negative pixels (pixels labelled false in S and true in S').

Like the VoI-based method, the mean Dice of S_i is computed considering a
set of L segmented slices $\{S_k\}_{k \leq L}$. In addition, a weight is associated to each
slice in order to give a bigger Dice to the nearest slices. The equation then
230 becomes:

$$\overline{Dice}(S_i, \{S_k\}_{k \leq L}) = \frac{1}{L} \sum_{k=1}^L Dice(S_i, S_k) * r_{i,k} \quad (8)$$

with weight $r_{i,k}$:

$$r_{i,k} = 1 - \frac{|i - k|}{n_n * 2 + 1} \quad (9)$$

where n_n is the number of adjacent slices taken into account at each side of S_i .

With this method, The fused segmentation is the one which obtains the
biggest \overline{Dice} :

$$\hat{S}_{i \overline{Dice}} = \underset{S \in S_n}{argmax} \overline{Dice}(S_i, \{S_k\}_{k \leq L}) \quad (10)$$

235 Finally, the pseudo-code of the Dice-based method is the same as Algorithm
2 where $\Delta \overline{VoI}$ is replaced by \overline{Dice} .

3.5. 2D and 3D method based on neighbouring information

3.5.1. NandI method

The method based on neighbouring information, like the pixel's grey level
240 and the neighbour's labels, is used in Feng et al. (2017) for the fusion of brain-
images segmentation. This method, known as NandI (Neighbours and Inten-
sity), calculates a similarity index between two pixels p and q of the same slice,
defined as:

$$Sim(p, q) = exp \left(-\frac{d^2(p, q)}{2\alpha^2} - \frac{dif^2(p, q)}{2\beta^2} \right) \quad (11)$$

where $d(p, q) = \|p - q\|$ is the Euclidean distance between p and q , and $dif(p, q) =$
245 $|I(p) - I(q)|$ is the difference of intensity between them. α and β are two weights.

A set of coherent neighbours must then be determined for each conflicting pixel p in order to determine their final label. Let E_p^r be the set of an effective neighbourhood of p . A pixel q must satisfy the next three conditions in order to become an effective neighbour of p :

- 250 1. $d(p, q) < r$: The Euclidean distance between p and q must be smaller than radius r ,
2. $l^q \in S = l^q \in S'$: q is not a conflicting pixel,
3. $l^q \in \varepsilon$: The label of q belongs to the set of possible labels, i.e the tumour label or the kidney label.

255 The new label of p is also obtained by:

$$l^p = \begin{cases} \textit{tumour} & \text{if } \sum_{q \in E_p^r, l^q = \textit{tumour}} \textit{Sim}(p, q) > \sum_{q \in E_p^r, l^q = \textit{kidney}} \textit{Sim}(p, q), \\ \textit{kidney} & \text{else} \end{cases} \quad (12)$$

Finally, the pseudo-code of the NandI method for the conflict resolution of the set of segmented slices for one patient is shown in Algorithm 3.

Algorithm 3: The NandI-method algorithm.

Data:

$\{S_k\}_{k \leq L}$: The set of L segmented slices

r : The radius

for each segmented slice S in $\{S_k\}_{k \leq L}$ **do**

for each conflicted pixel p of segmentation S **do**

Compute E_p^r

Compute l_s^p

end

end

3.5.2. 3D-NandI method

260 In order to manage the conflicts of a slice, the NandI method only uses the information from this slice. However, as we are working on a set of segmented slices corresponding to a patient's abdomen, 3D information can also be taken

into account. The 3D-NandI method is an evolution of the NandI method, in which a 3D effective neighbourhood and a 3D Euclidean distance are calculated.

265 3.6. Fully connected CRF based method

The fully connected CRF (Krähenbühl & Koltun (2011)) is mainly used as CNN post-processing, from the generated probability maps, in order to refine the calculated segmentations (Chen et al. (2018); Noh et al. (2015); Kamnitsas et al. (2017)). Indeed, they refine the edges of the structures while taking into
270 account the possible dependencies through fast and accurate inferences. Unlike the widespread use of Fully Connected CRF to refine segmentation results, we use them here to resolve the conflicts generated by the fusion of complementary segmentations.

Figure 3 shows the process of this method. At first, the probability maps of
275 the CNN segmentations of each structure of the image are used and modified. These values, between 0 and 1 and corresponding to the percentage of membership in each class, are adjusted to 0 or 1 for non-conflicting pixels. On the contrary, the values of the conflicting pixels stay in this interval. These modified probability maps are given as input to the Fully Connected CRF (in addition to
280 the corresponding scanner images), which will determine the labels of the pixels in conflict by outputting consensus segmentation without conflicts.

3.7. Conflict-management model

Figures 4 and 5 show our two versions of the conflict-management process based on the combination of all the methods described above. The set of calcu-
285 lated complementary segmentations and the corresponding scans are the inputs for the entire process. In the first version in Figure 4, the conflicts are first positioned and treated with all the methods presented above in parallel.

The results of all the methods can then be aggregated in two different ways. The first is based on a majority vote for the association of the final labels of
290 each pixel in conflict. However, with the majority vote, there may be a tie for pixels that can be resolved by the ADS2 method because there is then an even

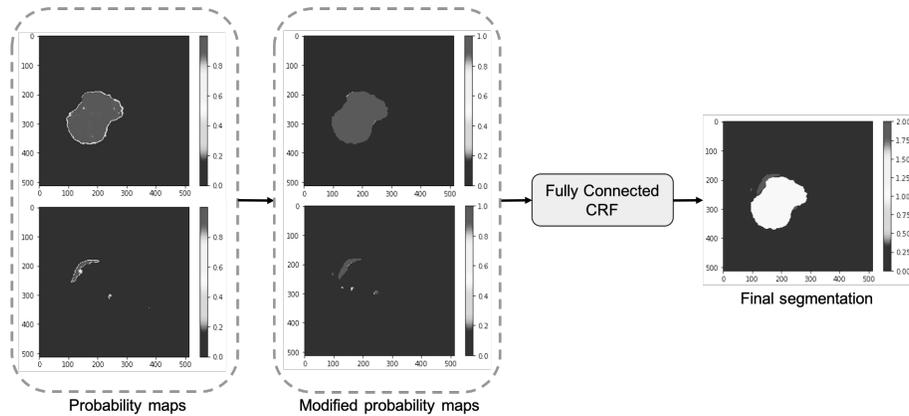


Figure 3: Fusion and conflict management of segmentations with the Fully Connected CRF-based method.

number of results to be processed. In this case, several choices are possible, such as choosing the final label determined by a certain method, or deciding that the pixels in the tie be labelled as belonging to the tumour or the kidney.

295 The second method resulting in segmentation fusion uses the CRF. During the CRF algorithm, probability maps are generated. For each conflicting pixel, the value of this pixel in the probability maps is the number of times that this label is chosen divided by the total number of results. For example, if there are two possible labels (kidney and tumour), four results in favour of the kidney
 300 label and two in favour of the tumour label, then the value of this pixel in the probability map of the kidney label will be $4/6$ and $2/6$ in the probability map of the tumour label.

The second version of the conflict-management process is described in Figure 5. The ADS2 method performs a pre-treatment in order to obtain an odd result.
 305 Then, the other methods are applied, and a majority vote or a decision based on the CRF probability map is launched in order to obtain the final segmentations.

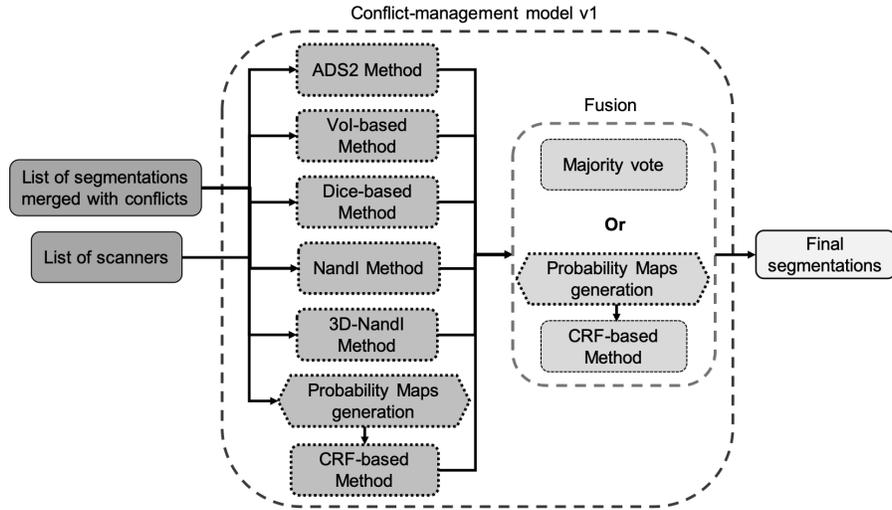


Figure 4: First version of the conflict-management model, for the fusion of tumorous kidney segmentations.

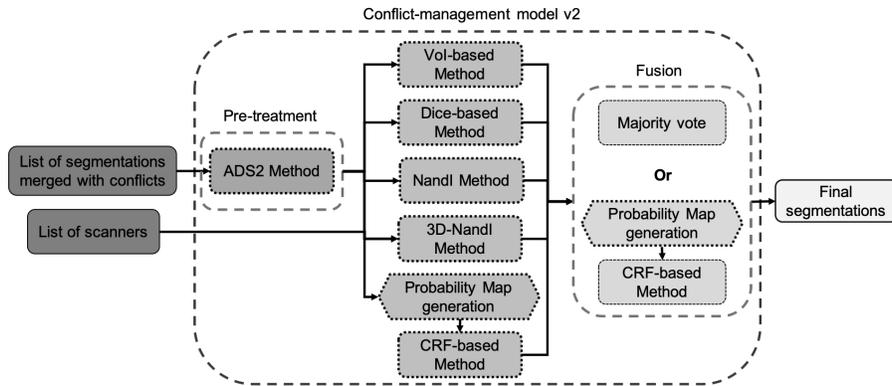


Figure 5: Second version of the conflict-management model, with a pre-treatment, for the fusion of tumorous kidney segmentations.

4. Experiments

4.1. Generation of segmentation set

We have tested the performance of our models over a set of 139 segmented
310 scans of three patients. Patients 1, 2, and 3 have 47, 62, and 30 scans, respec-
tively. All of the scans and calculated segmentations have the same size: 512
x 512. The pathological kidney and the nephroblastoma of each patient were
segmented by Deep Learning (the convolutional neural network FCN-8s first
trained with the PASCAL VOC 2012 database in Everingham et al. (2015))
315 enhanced with the OV²ASSION (Marie et al. (2019)) with 10000 iterations for
training. The learning protocol OV²ASSION is adapted to small learning sets.

We also have the ground truth segmentations of our three patients, carried
out by experts (surgeons and radiologists) at the *Centre hospitalier Régional
Universitaire de Besançon*. These ground truths are used in order to verify the
320 reliability of our processes and the resulting fused segmentations.

4.2. Results

4.2.1. Optimal parameters

We first determined the optimal parameters for each fusion method. In
particular, we determined the optimal number of adjacent slices for the VoI-
325 based and the Dice-based methods. Figure 6 shows the results obtained for the
VoI and Dice-based methods regarding the number of adjacent slices taken into
account. For all the patients, the best average performance of the VoI-based
method and Dice-based method is obtained with two adjacent slices at each side
of the considered segmentations for $T_{max} = 3$. Three loops were sufficient to
330 become stable because the initial image is very close to the expected image, and
there are only a few pixels in conflict (compared to the total number of image
pixels).

The performance of the NandI and the 3D-NandI methods have been evalu-
ated according to the size of the chosen radius. The results of both methods are
335 shown in Figure 7. The average performance of the 3D-NandI method stagnates

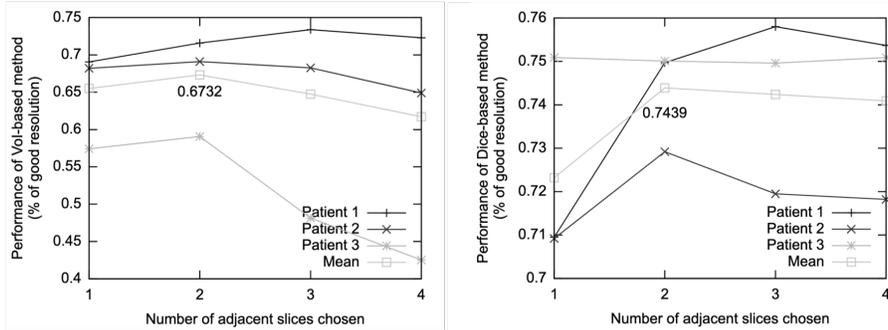


Figure 6: Evolution of the performance of the VoI-based method and the Dice-based method regarding the number of adjacent slices at each side of the considered segmentations.

after a radius of 5. Consequently, a radius of 3 has been retained for the NandI method and a radius of 5 for the 3D-NandI one.

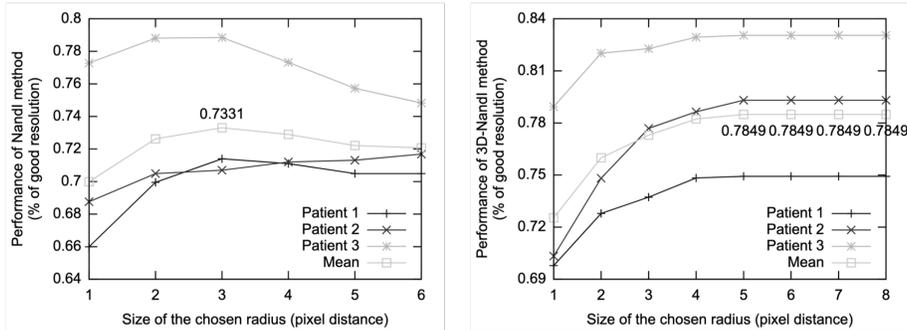


Figure 7: Evolution of the performance of the NandI method and the 3D-NandI method regarding the size of the chosen radius.

4.2.2. Results of both of the versions

Table 1 shows the performance of the ADS2 method for three patients. This table presents the number of conflicting pixels that have been correctly classified and, in between brackets, the number of conflicting pixels that were correctly classified by the total number of conflicting pixels that could be resolved by this method. We obtained 88.87% of good resolution for Patient 1 (1.749 correctly

classified pixels for 1.968 pixels that could be resolved), 92.11% for Patient 2,
 345 and 86.82% for Patient 3. An average of 89.56% of good resolution was obtained,
 which makes ADS2 a good method for a little less than a quarter of the pixels
 in conflict. However, this method cannot be applied to all conflicting pixels.
 Indeed, we must be in the favourable case where the pixels of adjacent slices
 have the same non-conflicting label for the conflicting pixel. Table 2 shows the
 350 percentage of conflicting pixels that can be resolved by the ADS2 method, which
 represents on average of 23.30%.

	Patient 1	Patient 2	Patient 3	Total Mean
ADS2	0.8887 (1749/1968)	0.9211 (1004/1090)	0.8682 (448/516)	0.8956

Table 1: Average performance on different patients and for the ADS2 method.

	Patient 1	Patient 2	Patient 3	Mean
ADS2	31.2%	26.3%	12.5%	23.3%

Table 2: Percentage of conflicting pixels that can be resolved by the ADS2 method for three patients.

Table 3 shows the average performance in the first version for each of the
 methods, in parallel to the ADS2 method. We have 67.32% of correct resolu-
 tion for the VoI-based method, 74.39% for the Dice-based method, 73.31% and
 355 78.49% for the NandI and 3D-NandI methods, and 75.03% for the CRF-based
 method. The final results are shown in regards to the type of fusion applied in
 Table 4.

Applying the first method shown in Figure 4, when the fusion of results is
 realized by a majority vote (MV in the table) and, if in case of a tie, the kidney
 360 label is chosen, the conflicting pixels are correctly resolved at 78.61%. On the
 contrary, if the tumour label is chosen in case of a tie, our system gets 78.59%

	Patient 1	Patient 2	Patient 3	Total Mean
VoI	0.7159 (4946/6909)	0.6910 (3203/4635)	0.5906 (2703/4577)	0.6732
Dice	0.7497 (5180/6909)	0.7292 (3380/4635)	0.7501 (3433/4577)	0.7439
NandI	0.7140 (4933/6909)	0.7070 (3277/4635)	0.7885 (3609/4577)	0.7331
3D-NandI	0.7493 (5177/6909)	0.7931 (3676/4635)	0.8305 (3801/4577)	0.7849
CRF	0.7234 (4998/6909)	0.7458 (3457/4635)	0.7955 (3641/4577)	0.7503

Table 3: Average performance on different patients and for each method with the first version of the conflict-management model.

	Patient 1	Patient 2	Patient 3	Total Mean
MV with kidney in case of a tie	0.7746 (5352/6909)	0.7737 (3586/4635)	0.8160 (3735/4577)	0.7861
MV with tumour in case of a tie	0.7752 (5356/6909)	0.7730 (3583/4635)	0.8149 (3730/4577)	0.7859
CRF-based method	0.7658 (5291/6909)	0.7663 (3552/4635)	0.8193 (3750/4577)	0.7812

Table 4: Average performance measurements on different patients for the first model with different methods of fusion.

of correct resolution. The system also gets 78.12% of correct resolution with the CRF-based method.

For the second version of our method described in Figure 5, the performance of the ADS2 method is the same as the first model shown in Table 1. For the other methods, the performances are shown in Table 5, where the results are

	Patient 1	Patient 2	Patient 3	Total Mean
VoI	0.6446 (3185/4941)	0.6285 (2228/3545)	0.5696 (2313/4061)	0.6157
Dice	0.6871 (3395/4941)	0.6697 (2374/3545)	0.7355 (2987/4061)	0.6979
NandI	0.6912 (3415/4941)	0.6894 (2444/3545)	0.7870 (3196/4061)	0.7217
3D-NandI	0.6936 (3427/4941)	0.7540 (2673/3545)	0.8239 (3346/4061)	0.7528
CRF	0.7051 (3484/4941)	0.7255 (2572/3545)	0.7936 (3223/4061)	0.7395

Table 5: Average performance on different patients and for each method with the second version of the conflict-management model.

	Patient 1	Patient 2	Patient 3	Total Mean
MV	0.7681 (5307/6909)	0.7786 (3609/4635)	0.8184 (3746/4577)	0.7854
CRF-based method	0.7678 (5305/6909)	0.7769 (3601/4635)	0.8217 (3761/4577)	0.7857

Table 6: Average performance measurements on different patients for the second model with different methods of fusion.

lower than the first version (Mean between 61.57% and 75.28%) because the easiest conflicts have already been solved by the ADS2 method. Table 6 shows the final results of this second method. The final results of the two models are similar but remain superior to the result of each method individually (except for the fusion using the CRF-based method in the first model, which is less efficient than the 3D-NandI method alone).

Table 7 shows the average Dice of segmentations for each patient before and

	Type of Fusion	Mean Dice Patient 1	Mean Dice Patient 2	Mean Dice Patient 3
Before		0.8535	0.8491	0.8477
After V1 model	CRF-based method	0.8568	0.8503	0.8501
	MV with kidney in case of a tie	0.8569	0.8504	0.8500
	MV with tumour in case of a tie	0.8569	0.8504	0.8499
After V2 model	CRF-based method	0.8568	0.8505	0.8502
	MV	0.8567	0.8505	0.8501

Table 7: Presentation of the average Dice before and after the conflict management.

after the different conflict-management processes. The values in bold correspond
375 to the maximum values obtained for each patient. We can see that the average
Dice increased for each patient, with an average increase of 0.24%, taking into
account the maximum Dice after conflict management. We can justify the slight
improvement by the fact that the number of conflicting pixels is minimal compared
to the total number of pixels of each image. Indeed, the conflicting pixels
380 represent 0.06%, 0.03%, and 0.06% of pixels for Patients 1, 2, and 3 respectively.
Nevertheless, the conflict-management protocol improves the segmentations in
the most important areas (the edges of the different touching structures).

The result of one segmentation is presented in Figure 8. The initial seg-
mentation on the left has conflicts (light colour) at the intersection between the
385 kidney and the tumour, and inside the tumour. The latter conflict is quite easy
to solve, because they are mostly due to artefacts during the segmentation pro-
cess. These artefacts are unique to the current slice. In addition, it is possible
to simply determine the correct labels of these pixels by their grey intensity in

the scanner image. The two models of conflict management then give the same
390 result, as is the case on the right image. However, the former type of conflict is
more difficult to resolve, and this is where the propositions of the different mod-
els diverge. Our different methods, which use various information like grayscale
of the corresponding scanner image and of the neighbouring scanner images, as
well as the labels of current segmentation and of the neighbouring segmentations
395 are not enough in all cases to resolve the conflicts correctly.

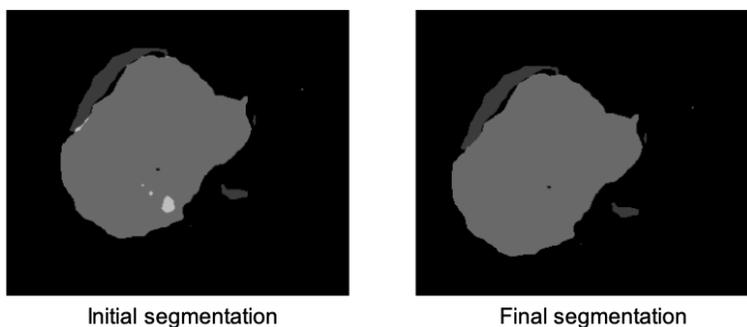


Figure 8: Example of a segmentation fusion result with the initial segmentation with conflicts (shown in light colour) on the left and the resulting segmentation on the right.

Figure 9 shows the conflict resolution between the two structures. Four pixels (shown by the arrows) were labelled "tumour" by the fusion through majority vote in both of the models and also labelled "kidney" by the fusion through the CRF-based method in both of the models.

400 4.2.3. Discussion

The results show that the combination of several fusion methods produces better results than each method individually. Indeed, the best way is to obtain complementary methods. The performance of the conflict management methods are different since they exploit different types of information (localization,
405 statistics, probabilities, similarities between slices) and a different number of images (the slices in conflicts only, adjacent slices in addition).

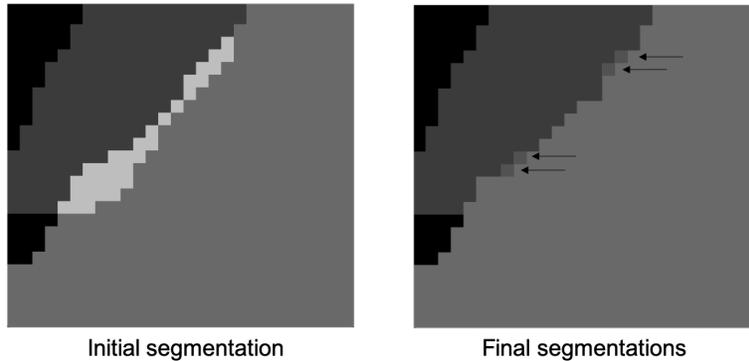


Figure 9: Focus on conflict resolution at the kidney-tumour intersection in Figure 8 with the initial segmentation with conflicts (shown in light colour) on the left and the resulting segmentations superimposed on the right.

The Dice-based method performs better than the VoI-based method because, even if exploit all the neighbouring segmentations, the Dice-based method associates weighting to the segmentations. The NandI, 3D-NandI, and the CRF-based methods also use the scanner images themselves as inputs for their conflict management (and thus exploit more information). In addition, the 3D-NandI method is an improvement to the NandI method since it integrates all the neighbouring-slice (segmentations and scanners) information. As with the NandI method, the CRF-based method uses the information on the current slice. For that reason, the CRF-based method is outperformed by the 3D-NandI method. In any manner, the CRF-based method is the only fusion method that uses the probability map of the segmentation, which is another source of information for the segmentation fusion.

The ADS2 method is a particular method since its strength is also its weakness. Indeed, it uses the information in adjacent segmentations only if they are equal and consistent. Thus, statistically and logically, it is the most efficient method because of its precautionary principle. However, this method cannot be used on all conflicting pixels: the conflict can be resolved when, and only when, pixel labels in adjacent slices match.

425 The performance of the two proposed conflict management models are more
or less equivalent; there is no model that stands out. Two fusion methods are
proposed in the models but are questionable because they give close results.
However, using the CRF-based method makes it possible not to be blocked a
tie in the case of an even number of results (as is the case using the majority
430 vote).

Prioritizing the ADS2 method in the second model does not seem to have
a significant impact on the results. We can also imagine weighting each of the
methods used as functions of the data. For example, if the scanned images
are too noisy, it would be better to give more importance to the methods that
435 exploit only the segmentations. Conversely, with segmentations very far from
the ground truths, it would be better to give more important weighting to the
methods exploiting the scanner images.

In addition, conflicting pixels within a structure are often easy to solve, but
those within a boundary between two structures are more ambiguous. **Even**
440 **radiologists and surgeons use their experience to delimit these areas manually.**
Sometimes, the information on the scanned images and on the segmentations
does not make it possible to find a clear delimitation. More powerful methods
must be found **to provide other types of information and** improve the efficiency
of our models for these particular cases. Some artificial intelligence methods **us-**
445 **ing experience, such as the** radiologists' and surgeons' experiences, may achieve
better results on these borders.

Finally, these models are easily flexible, and our results could be enriched
by adding other methods which exploit other information, such as patient data.
The models must now be tested over a larger set of patients. Indeed, at this
450 point, we managed to have a dataset composed of 139 slices and three patients.

5. Conclusion and further work

In this paper, we have presented new conflict-management strategies for
complementary segmentations of pathological kidneys with cancerous tumours.

Our methods allow the resolution of conflicts and thus improve the accuracy
455 of segmentations. Actually, these protocols are based on the intelligent asso-
ciation of single strategies for conflict management. These strategies use very
different tools: the 3D information provided by the adjacent slices, neighbouring
labels, the pixel intensity in scanner images, use of probability maps provided
by CNN training, and criteria dedicated to the measurement of the accuracy of
460 segmentations.

These fusion strategies are also adaptable and flexible to each case, since the
different parameters can be modified according to the type of segmentation and
conflicting pixels encountered.

Finally, in our case of the fully automatic segmentation of tumorous kidneys,
465 these processes for fusion improve the Dice accuracy of the calculated segmen-
tations by 0.24% on average and thus increase the robustness of the general
system of the SAIAD project.

Further work will focus on the measurement of the robustness of the pro-
cesses. In addition, we wish to add artificial intelligence methods in our sys-
470 tems to improve the resolution of conflicting pixels. We also wish to extend the
method to other anatomical structures appearing on the scans, such as arteries,
veins, and excretory cavities.

Acknowledgements

The authors wish to thank Professor F. Auber, Dr Y. Chaussy, and Dr M.
475 Lenoir-Auber from the *BESANCON University Hospital* for their expertise in
nephroblastoma, and Loredane Vieille for performing the manual segmentations.

The authors would like to thank the European Union for financing this
project as part of the SAIAD INTERREGV program in collaboration with the
Swiss partner EPFL Lausanne.

480 The authors also wish to thank Mary Moritz (WordStyle Traduction) for her
help with the English language.

References

References

- Aljabar, P., Heckemann, R. A., Hammers, A., Hajnal, J. V., & Rueckert, D.
485 (2009). Multi-atlas based segmentation of brain images: atlas selection and its
effect on accuracy. *Neuroimage*, *46*, 726–738. doi:10.1016/j.neuroimage.
2009.02.018.
- Artaechevarria, X., Munoz-Barrutia, A., & Ortiz-de Solórzano, C. (2009).
Combination strategies in multi-atlas image segmentation: application to
490 brain mr data. *IEEE transactions on medical imaging*, *28*, 1266–1277.
doi:10.1109/TMI.2009.2014372.
- Besag, J. (1986). On the statistical analysis of dirty pictures. *Journal of the
Royal Statistical Society: Series B (Methodological)*, *48*, 259–279. doi:10.
1111/j.2517-6161.1986.tb01412.x.
- 495 Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018).
Deeplab: Semantic image segmentation with deep convolutional nets, atrous
convolution, and fully connected crfs. *IEEE transactions on pattern analysis
and machine intelligence*, *40*, 834–848. doi:10.1109/TPAMI.2017.2699184.
- Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood
500 from incomplete data via the em algorithm. *Journal of the royal statistical
society. Series B (methodological)*, (pp. 1–38). doi:10.1111/j.2517-6161.
1977.tb01600.x.
- Everingham, M., Eslami, S. A., Van Gool, L., Williams, C. K., Winn, J., &
Zisserman, A. (2015). The pascal visual object classes challenge: A ret-
505 rospective. *International journal of computer vision*, *111*, 98–136. doi:10.
1007/s11263-014-0733-5.
- Feng, Y., Shen, X., Chen, H., & Zhang, X. (2017). Segmentation fusion based
on neighboring information for mr brain images. *Multimedia Tools and Ap-
plications*, *76*, 23139–23161. doi:10.1007/s11042-016-4098-3.

- 510 Frucci, M., Perner, P., & Sanniti Di Baja, G. (2008). Case-based-reasoning for image segmentation. *International Journal of Pattern Recognition and Artificial Intelligence*, *22*, 829–842. doi:10.1142/S0218001408006491.
- Ghassemian, H. (2016). A review of remote sensing image fusion methods. *Information Fusion*, *32*, 75–89. doi:10.1016/j.inffus.2016.03.003.
- 515 Harrabi, R., & Braiek, E. B. (2012). Color image segmentation using multi-level thresholding approach and data fusion techniques: application in the breast cancer cells images. *EURASIP Journal on Image and Video Processing*, *2012*, 11. doi:10.1186/1687-5281-2012-11.
- Huang, D., Lai, J.-H., Wang, C.-D., & Yuen, P. C. (2016). Ensembling over-
520 segmentations: From weak evidence to strong segmentation. *Neurocomputing*, *207*, 416–427. doi:10.1016/j.neucom.2016.05.028.
- James, A. P., & Dasarathy, B. V. (2014). Medical image fusion: A survey of the state of the art. *Information Fusion*, *19*, 4–19. doi:10.1016/j.inffus.2013.12.002.
- 525 Jin, X., Jiang, Q., Yao, S., Zhou, D., Nie, R., Hai, J., & He, K. (2017). A survey of infrared and visual image fusion methods. *Infrared Physics & Technology*, *85*, 478–501. doi:10.1016/j.infrared.2017.07.010.
- Kamnitsas, K., Ledig, C., Newcombe, V. F., Simpson, J. P., Kane, A. D., Menon, D. K., Rueckert, D., & Glocker, B. (2017). Efficient multi-scale 3d
530 cnn with fully connected crf for accurate brain lesion segmentation. *Medical image analysis*, *36*, 61–78. doi:10.1016/j.media.2016.10.004.
- Kato, Z., Zerubia, J. et al. (2012). Markov random fields in image segmentation. *Foundations and Trends® in Signal Processing*, *5*, 1–155. doi:10.1561/20000000035.
- 535 Kausar, N., Majid, A., & Javed, S. G. (2016). Developing learning based intelligent fusion for deblurring confocal microscopic images. *Engineering Applica-*

- tions of Artificial Intelligence*, 55, 339–352. doi:10.1016/j.engappai.2016.08.006.
- Khan, A., & Jaffar, M. A. (2015). Genetic algorithm and self organizing map based fuzzy hybrid intelligent method for color image segmentation. *Applied soft computing*, 32, 300–310. doi:10.1016/j.asoc.2015.03.029.
- Khelifi, L., & Mignotte, M. (2017a). Efa-bmfm: A multi-criteria framework for the fusion of colour image segmentation. *Information Fusion*, 38, 104–121. doi:10.1016/j.inffus.2017.03.001.
- 545 Khelifi, L., & Mignotte, M. (2017b). A multi-objective decision making approach for solving the image segmentation fusion problem. *IEEE Transactions on Image Processing*, 26, 3831–3845. doi:10.1109/TIP.2017.2699481.
- Khelifi, L., & Mignotte, M. (2017c). A novel fusion approach based on the global consistency criterion to fusing multiple segmentations. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47, 2489–2502. doi:10.1109/TSMC.2016.2531645.
- 550 Krähenbühl, P., & Koltun, V. (2011). Efficient inference in fully connected crfs with gaussian edge potentials. In *Advances in neural information processing systems* (pp. 109–117).
- 555 Li, H., Manjunath, B., & Mitra, S. K. (1995). Multisensor image fusion using the wavelet transform. *Graphical models and image processing*, 57, 235–245. doi:10.1109/ICIP.1994.413273.
- Li, Z., Wu, X.-M., & Chang, S.-F. (2012). Segmentation using superpixels: A bipartite graph partitioning approach. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (pp. 789–796). IEEE. doi:10.1109/CVPR.2012.6247750.
- 560 Liang, X., & Huang, D.-S. (2018). Image segmentation fusion using weakly supervised trace-norm multi-task learning method. *IET Image Processing*, 12, 1079–1085. doi:10.1049/iet-ipr.2017.1061.

- 565 Liu, Z., Yin, H., Chai, Y., & Yang, S. X. (2014). A novel approach for multi-modal medical image fusion. *Expert systems with applications*, *41*, 7425–7435. doi:10.1016/j.eswa.2014.05.043.
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431–3440). doi:10.1109/CVPR.2015.7298965.
- 570 Ma, K., Zeng, K., & Wang, Z. (2015). Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, *24*, 3345–3356. doi:10.1109/TIP.2015.2442920.
- 575 Marie, F., Corbat, L., Chaussy, Y., Delavelle, T., Henriët, J., & Lapayre, J.-C. (2019). Segmentation of deformed kidneys and nephroblastoma using case-based reasoning and convolutional neural network. *Expert Systems with Applications*, . doi:10.1016/j.eswa.2019.03.010.
- Marie, F., Corbat, L., Delavelle, T., Chaussy, Y., Henriët, J., & Lapayre, J.-C. (2018). Segmentation of kidneys deformed by nephroblastoma using case-based reasoning. In *ICCBR 2018, 26th International conference on Case-based reasoning* (pp. 351–365). Stockholm, Sweden. doi:10.1007/978-3-030-01081-2_16.
- 585 Martin, D. R., Fowlkes, C. C., & Malik, J. (2004). Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE transactions on pattern analysis and machine intelligence*, *26*, 530–549. doi:10.1109/TPAMI.2004.1273918.
- Meilă, M. (2003). Comparing clusterings by the variation of information. In *Learning theory and kernel machines* (pp. 173–187). Springer. doi:10.1007/978-3-540-45167-9_14.
- 590 Mertens, T., Kautz, J., & Van Reeth, F. (2007). Exposure fusion. In *Computer*

Graphics and Applications, 2007. PG'07. 15th Pacific Conference on (pp. 382–390). IEEE. doi:10.1109/PG.2007.17.

Mignotte, M. (2008). Segmentation by fusion of histogram-based k -means clusters in different color spaces. *IEEE Transactions on image processing*, 17, 595 780–787. doi:10.1109/TIP.2008.920761.

Mignotte, M. (2010). A label field fusion bayesian model and its penalized maximum rand estimator for image segmentation. *IEEE Transactions on Image Processing*, 19, 1610–1624. doi:10.1109/TIP.2010.2044965.

600 Mignotte, M. (2014). A label field fusion model with a variation of information estimator for image segmentation. *Information Fusion*, 20, 7–20. doi:10.1016/j.inffus.2013.10.012.

Mignotte, M., & Hérou, C. (2014). A precision–recall criterion based consensus model for fusing multiple segmentations. *Int J Signal Process Image Process Pattern Recognit*, 7, 61–82. doi:10.14257/ijcip.2014.7.3.07. 605

Mobahi, H., Rao, S. R., Yang, A. Y., Sastry, S. S., & Ma, Y. (2011). Segmentation of natural images by texture and boundary compression. *International journal of computer vision*, 95, 86–98. doi:10.1007/s11263-011-0444-0.

Nguyen, D., Benameur, S., Mignotte, M., & Lavoie, F. (2018). Superpixel and multi-atlas based fusion entropic model for the segmentation of x-ray images. 610 *Medical image analysis*, 48, 58–74. doi:10.1016/j.media.2018.05.006.

Nguyen, D. C. T., Benameur, S., Mignotte, M., & Lavoie, F. (2015). Superpixel and entropy-based multi-atlas fusion framework for the segmentation of x-ray images. In *International Conference on Image Analysis and Processing* (pp. 615 151–161). Springer. doi:10.1007/978-3-319-23234-8_15.

Noh, H., Hong, S., & Han, B. (2015). Learning deconvolution network for semantic segmentation. In *Proceedings of the IEEE international conference on computer vision* (pp. 1520–1528). doi:10.1109/ICCV.2015.178.

- 620 Pohl, C., & Van Genderen, J. L. (1998). Review article multisensor image fusion
in remote sensing: concepts, methods and applications. *International journal
of remote sensing*, *19*, 823–854. doi:10.1080/014311698215748.
- Rohlfing, T., & Maurer, C. R. (2007). Shape-based averaging. *IEEE Transactions on Image Processing*, *16*, 153–161. doi:10.1109/TIP.2006.884936.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks
625 for biomedical image segmentation. In *International Conference on Medical
image computing and computer-assisted intervention* (pp. 234–241). Springer.
doi:10.1007/978-3-319-24574-4_28.
- Sathya, B., & Manavalan, R. (2011). Image segmentation by clustering methods:
performance analysis. *International Journal of Computer Applications*, *29*.
630 doi:10.5120/3688-5127.
- Tosta, T. A. A., Faria, P. R., Neves, L. A., & do Nascimento, M. Z. (2017).
Computational method for unsupervised segmentation of lymphoma histolog-
ical images based on fuzzy 3-partition entropy and genetic algorithm. *Expert
Systems with Applications*, *81*, 223–243. doi:10.1016/j.eswa.2017.03.051.
- 635 Warfield, S. K., Zou, K. H., & Wells, W. M. (2004). Simultaneous truth and
performance level estimation (staple): an algorithm for the validation of image
segmentation. *IEEE transactions on medical imaging*, *23*, 903–921. doi:10.
1109/TMI.2004.828354.