Analogy-based post-treatment of CNN image segmentations

Justine Duck, Romain Schaller, Frédéric Auber, Yann Chaussy, Julien Henriet, Jean Lieber, Emmanuel Nauer, Henri Prade

Analogy-based post-treatment of CNN image segmentations

Justine Duck[1,2], Romain Schaller[2], Frédéric Auber[3], Yann Chaussy[3], Julien Henriet[1], Jean Lieber[2], Emmanuel Nauer[2], and Henri Prade[4]

[1] FEMTO-ST DISC, Univ. Bourgogne-Franche-Comté, F-25000 Besançon, France
[2] Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy, France
[3] CHU / LNIT UFC, F-25000 Besançon, France
[4] IRIT, CNRS, Université de Toulouse, France

Abstract. Convolutional neural networks (CNNs) have proven to be efficient tools for image segmentation when a large number of segmented images are available. However, when the number of segmented images is not so large, the CNN segmentations are less accurate. It is the case for nephroblastoma (kidney cancer) in particular. When a new patient arrives, the expert can only manually segment a sample of scanned images since manual segmentation is a time-consuming process. As a consequence, the question of how to compute accurate segmentations using both the trained CNN and such a sample is raised. A CBR approach based on proportional analogy is proposed in this paper. For a source image segmented by the expert, let a be the CNN segmentation of this image, b be its expert segmentation and c be the CNN segmentation of a target image close to the source image. The proposed approach aims at solving the analogical equation "a is to b as c is to d" with unknown d: the solution d of this equation is proposed as a segmentation of the target image. This approach and some of its improvements are evaluated and show an accuracy increase of the segmentation with respect to the CNN segmentation.

Keywords: analogical extrapolation, case-based reasoning, convolutional neural networks, medical image segmentation, kidney cancer

1 Introduction

Convolutional neural networks (CNNs) constitute powerful tools for many tasks, such as image segmentation [9]. However, as many techniques of deep learning, they are demanding in terms of the computing time they require for the learning phase. This is particularly true when the dataset is continuously enriched, as it leads to re-run regularly the learning process. By contrast, case-based reasoning (CBR [14]) is usually less demanding in terms of computing time and also for the volume of resources (including cases) and it is "naturally" fitted to the continuous enrichment of the case base.

This article studies the issue of using CBR to improve the results of a CNN when few additional data are available, without having to re-train the CNN. This

issue is considered for an application of image segmentation of kidney cancer scans. More specifically, a CNN has been run on images segmented by an expert for 14 patients, with about 100 slices per patient. Given new patient's images, the CNN can propose some segmentations, but with an insufficient precision. To increase this precision, the expert manually segments about 10% of these images and then, the goal is to exploit this 10% to improve the precision of the CNN segmentations of the 90% remaining images. For this purpose, an approach called OV^{2}ASSION based on re-running the CNN with the additional examples (the 10%), has been studied in previous works. In this paper, an alternative approach to OV^{2}ASSION based on CBR methodology and on analogical proportions is studied.

After some necessary preliminaries (Section 2), a general approach is presented (Section 3) explaining how the CNN image segmentation can be modified by retrieval and adaptation of an image manually segmented. The result is a partial segmentation, meaning that some pixels are *undecided*. Section 4 presents some improvements to the approach, in particular, to make decisions for undecided pixels. Section 5 concludes and describes future directions of work.

2 Preliminaries

This section introduces the notions and notations that are useful for the article. Section 2.1 presents the issue of image segmentation, in the framework of kidney cancer images. Section 2.2 presents two deep-learning approaches for addressing this issue: one based on a CNN and one which is an improvement of this approach in the context of the study, called OV^{2}ASSION. This article proposes another improvement of the approach of the outcome of the CNN that is based on CBR and analogical proportions introduced in Sections 2.3 and 2.4.

2.1 Image segmentation and its application to kidney cancer management

Nephroblastoma is one of the most frequently abdominal tumor observed in children (generally 1 to 5-year-old boys and girls). This cancer represents 5 to 14% of malignant paediatric tumors. This tumor is developped in the kidney. Most of the time, its initial diagnosis is based on imaging. Generally, ultrasounds are planned first in order to confirm its existence and approximate its position. Then, a CT-scan provides its position, and the healthy tissues and organs are reached with a higher accuracy. Radiologists and surgeons need 3D representation of the tumor and the border organs in order to plan the surgery (e.g. anticipate vascular risks, choose between a total or partial nephrectomy), and also to inform the family.

Image segmentation is a topic of image processing. It consists in associating a single label to each pixel of an image. In the field of health, segmenting a scan consists in defining the anatomical structure to which each pixel belongs. Segmentation is one of the key steps of the construction of 3D representations

of all the organs, veins, arteries, cavities and other anatomical elements (such as a tumor). Formally, for an image of $m \times n$ pixels, let $\mathcal{D}om = \{0, 1, \ldots, m-1\} \times \{0, 1, \ldots, n-1\}$. An image is a mapping associating to each $(i, j) \in \mathcal{D}om$ the code for the color of the pixel. In this paper, only binary segmentations are considered: there are 2 labels, denoted by the Boolean values 0 and 1. More precisely, the segmentation distinguishes pixels associated to the tumor (label 1) from other pixels (label 0). Formally, a segmentation of a tumor is a mapping $\mathtt{S} : \mathcal{D}om \to \{0, 1\}$. For $(i, j) \in \mathcal{D}om$, $\mathtt{S}(i, j)$ is denoted by $\mathtt{S}_{ij}$.

Automatic segmentation is one of the actual key challenges of image processing since most of the time, and in the particular case of the segmentation of nephroblastoma in children, surgeons and radiologists must lead the segmentation process manually. The manual segmentation of the images of a kidney is time consuming (it requires about 6 to 8 hours of medical expert time [2]), hence the usefulness of tools for assisting these experts. The quality of an automatic or semi-automatic segmentation is evaluated comparing it to a reference one given by an expert. In the field of image processing applied to healthcare, the Dice coefficient is usually employed [4]. It gives a similarity value (on $[0, 1]$) between two segmentations $\mathtt{S}^1$ and $\mathtt{S}^2$ defined by

$$\mathtt{DICE}(\mathtt{S}^1, \mathtt{S}^2) = \frac{2 \times \text{number of } (i,j) \in \mathcal{D}om \text{ such that } \mathtt{S}^1_{ij} = \mathtt{S}^2_{ij} = 1}{\#\mathtt{S}^1 + \#\mathtt{S}^2} \quad (1)$$

where $\#\mathtt{S}$ is the number of $(i, j) \in \mathcal{D}om$ such that $\mathtt{S}_{ij} = 1$.

If $\mathtt{S}^1$ is the segmentation automatically computed, and $\mathtt{S}^2$ is the desired segmentation (the ground truth, given by the expert), the closest to 1 is $\mathtt{DICE}(\mathtt{S}^1, \mathtt{S}^2)$, the better the segmentation is and the least the expert has additional work to do.

Now, for N images of the same kidney (e.g. N slices as in Figure 1) providing a 3D representation, a segmentation of this collection of images can be defined as $\mathbf{S} : k \in \{1, 2, \ldots, N\} \mapsto \mathbf{S}(k)$, where $\mathbf{S}(k)$ is a segmentation of the image number k. Now, given $\mathbf{S}^1$ and $\mathbf{S}^2$ two segmentations of the same collection of images, the Dice coefficient of $(\mathbf{S}^1, \mathbf{S}^2)$ can be computed with a definition similar to (1).[1] This is called the *3D Dice coefficient* in the following of the paper and provides a global assessment of a similarity between two collections of segmentations (e.g., for a given patient, the predicted collection of segmentations and the expert collection of segmentations).

The final 3D representations are used by surgeons and radiologists in order to plan the surgery, to have a clear vision of the healthy tissues and also to communicate with the children's families [2].

[1] $(i, j) \in \mathcal{D}om$ is substituted by $((i, j), k) \in \mathcal{D}om \times \{1, 2, \ldots, N\}$, $\mathtt{S}^1$ and $\mathtt{S}^2$ are substituted by $\mathbf{S}_1(k)$ and $\mathbf{S}_2(k)$ in the numerator and $\#\mathtt{S}^1 + \#\mathtt{S}^2$ is substituted with $\#\mathbf{S}_1 + \#\mathbf{S}_2$, where $\#\mathbf{S}$ is the number of $((i, j), k)$ such that $\mathbf{S}_{ij}(k)$.

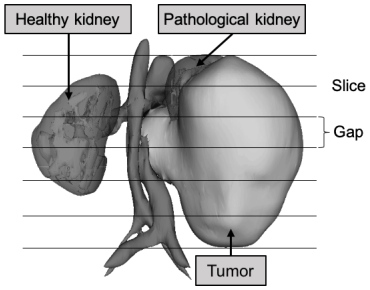

Fig. 1: 3D representation of a kidney with nephroblastoma, with the horizontal slices of the scanner imaging.

2.2 Using a CNN for kidney tumor segmentation

Inside the deep learning paradigm, CNN principles [6] have been implemented in Fully Convolutional Network (FCN) [10], SegNet [1], DeconvNet [12], DeepLab [3], and appear to be high-performance tools for image segmentation.

In medical and biology applications, Thong et al. [16] used CNN to perform segmentation of healthy kidneys. U-Net [15] performed segmentation of cells in microscopy images. Currently, CNNs obtain accurate results on the recognition of the shape of a healthy kidney, because the shapes and areas are more or less the same from one subject to another. However, when the form of the considered structure is complex and varies a lot from one case to another, the segmentation is still a challenge. Neural networks for segmentation also need many heterogeneous data in order to be able to transcribe reliable results. In particular, pathological kidneys deformed by nephroblastoma have very different forms from one patient to another, with unpredictable shapes and situations. This then has led to find another method to segment more complicated structures with limited data.

Having a sufficiently large volume of data representative of all possible data is essential for training a deep neural network. As manual segmentation is expert time-consuming, at the scale of a hospital, the learning set composed of the entire segmented abdomens of patients may be composed of tens of cases only. This may not be large enough for conventional learning since each tumor and pathological kidney is unique and varies greatly from one patient to another. This is the reason why we have designed a new method for training on a small dataset: the OV^{2}ASSION (Overlearning Vector for Valid Sparse SegmentatIONs) [11]. As shown in Figure 1, the OV^{2}ASSION method is based on the overlearning of some manually segmented slices of the patient, separated by a gap in order to calculate the segmentation of the entire set of unsegmented slices of this patient automatically. Each black line in this figure represents the selected slice for the training of the neural network. The gap between the chosen slices is the same in order to recover information homogeneously at different levels. This method is used in order to train a CNN based on the U-Net architecture.

2.3 Case-based reasoning

Let $\mathcal{P}$ (resp. $\mathcal{S}$) be a set. An element $\mathtt{x}$ of $\mathcal{P}$ (resp. $\mathtt{y}$ of $\mathcal{S}$) is called a *problem* (resp. a *solution*). A relation on $\mathcal{P} \times \mathcal{S}$ is assumed to exist, with the meaning "has for solution". A *case* is a pair $(\mathtt{x}, \mathtt{y}) \in \mathcal{P} \times \mathcal{S}$ such that $\mathtt{x}$ has for solution $\mathtt{y}$. This "has for solution" relation is not completely known to the system. However, it is assumed that some cases $(\mathtt{x}^s, \mathtt{y}^s) \in \mathcal{P} \times \mathcal{S}$ are known: they are the *source cases* and constitute the *case base* $\mathtt{CB}$.

CBR aims at solving a problem, called the *target problem* and denoted by $\mathtt{x}^{\mathtt{tgt}}$. A classical process model of CBR consists in (1) selecting k source cases relevant to solve $\mathtt{x}^{\mathtt{tgt}}$ (retrieval step), (2) inferring a solution $\mathtt{y}^{\mathtt{tgt}}$ of $\mathtt{x}^{\mathtt{tgt}}$ by reusing these source cases (adaptation step). If $k = 1$, step (2) is qualified as *single case adaptation*. Other steps, not considered here, follow adaptation.

2.4 Analogical proportions

Following [13], an *analogical proportion* on a set $\mathcal{U}$ is a quaternary relation on $\mathcal{U}$ denoted, for $(a, b, c, d) \in \mathcal{U}^4$, by $a\!:\!b\!::\!c\!:\!d$ and read "a is to b as c is to d", that satisfies the following postulates:

- $a\!:\!b\!::\!a\!:\!b$ (*reflexivity*);
- if $a\!:\!b\!::\!c\!:\!d$ then $c\!:\!d\!::\!a\!:\!b$ (*symmetry*);
- if $a\!:\!b\!::\!c\!:\!d$ then $a\!:\!c\!::\!b\!:\!d$ (*central permutation*).

On the set $\mathcal{U} = \{0, 1\}$ of the Booleans, the minimal analogical proportion (according to the inclusion of quaternary relations) can be defined by

$$a\!:\!b\!::\!c\!:\!d \quad \text{if} \quad b - a = d - c$$

where the differences take their values in $\{-1, 0, 1\}$. Another analogical proportion, called in the following Sheldon Klein's proportion [5], can be defined by $|b - a| = |d - c|$.

An *analogical equation* is an expression of the form $a\!:\!b\!::\!c\!:\!\mathtt{?}$ where $\mathtt{?}$ is the unknown. For the minimal analogical proportion on Booleans, solving this equation consists in computing $b - a + c$: if this value belongs to $\{0, 1\}$ then it is the unique solution. Otherwise, $b - a + c \in \{-1, 2\}$ and the equation has no solution: the triples (a, b, c) such that $a\!:\!b\!::\!c\!:\!\mathtt{?}$ has no solution are $(0, 1, 1)$ and $(1, 0, 0)$. By contrast, an equation $a\!:\!b\!::\!c\!:\!\mathtt{?}$ based on Sheldon Klein proportional analogy always has a solution. If $(a, b, c) = (0, 1, 1)$ then the unique solution is 0. If $(a, b, c) = (1, 0, 0)$ then the unique solution is 1. Otherwise, the solution is the same as for the minimal analogical proportion, i.e., $b - a + c$.

Analogical proportions and analogical equations have been used in CBR in (at least) two ways. First, when the problem and solution spaces are not the same ($\mathcal{P} \neq \mathcal{S}$), a process called analogical extrapolation is applied (see e.g. [8]): it consists in retrieving $k = 3$ source cases whose problem parts are in analogy with the target problem and then in solving the analogical equation given by their solution parts. Second, when the problem and solution spaces are the

same ($\mathcal{P} = \mathcal{S}$), a single case adaptation can be performed by solving an equation $\mathtt{x}^s : \mathtt{y}^s :: \mathtt{x}^{\mathtt{tgt}} : ?$ (see [7]). In the following, analogical proportions are used according to this second way.

3 General approach

This section presents the general approach to propose an alternative to OV^{2}ASSION based on CBR. This approach outputs a *partial segmentation*, that is, a segmentation with some "undecided" pixels, as described in Section 3.1. Then, the problem setting is formulated in CBR terms (§3.2) and a CBR approach is proposed with a straightforward retrieval step (§3.3) and an adaptation step based on analogical proportions (§3.4). Finally, the approach is evaluated wrt OV^{2}ASSION, from accuracy and computation time viewpoints (§3.5).

3.1 Partial Segmentations

A *partial segmentation* is a mapping $\mathtt{PS} : \mathcal{D}om \to \{0, \mathtt{u}, 1\}$ where $\mathtt{PS}_{ij} = 0$ and $\mathtt{PS}_{ij} = 1$ have the same meaning as for segmentations and $\mathtt{PS}_{ij} = \mathtt{u}$ means that $\mathtt{PS}$ is *undecided* about the status of the pixel. For some computations in this paper, $\mathtt{u}$ is replaced by the value 0.5. A partial segmentation $\mathtt{PS}'$ *extends* a partial segmentation $\mathtt{PS}$ means that, for every $(i, j) \in \mathcal{D}om$, if $\mathtt{PS}_{ij} \neq \mathtt{u}$ then $\mathtt{PS}'_{ij} = \mathtt{PS}_{ij}$.

Now, in order to evaluate a partial segmentation $\mathtt{PS}^1$ proposed by the system in comparison to a reference segmentation $\mathtt{S}^2$, the idea is to consider the values $\mathtt{DICE}(\mathtt{S}^1, \mathtt{S}^2)$ for $\mathtt{S}^1$ in the set of segmentations that extend $\mathtt{PS}^1$. Following this idea, the definition of pessimistic, optimistic and average Dice coefficients (the latter being the average of the two formers) can be defined:

$$\mathtt{DICE}^{\mathtt{pessim}}(\mathtt{PS}^1, \mathtt{S}^2) = \min\left\{\mathtt{DICE}(\mathtt{S}^1, \mathtt{S}^2) \mid \mathtt{S}^1 \text{ extends } \mathtt{PS}^1\right\} \quad (2)$$

$$\mathtt{DICE}^{\mathtt{optim}}(\mathtt{PS}^1, \mathtt{S}^2) = \max\left\{\mathtt{DICE}(\mathtt{S}^1, \mathtt{S}^2) \mid \mathtt{S}^1 \text{ extends } \mathtt{PS}^1\right\} \quad (3)$$

$$\mathtt{DICE}^{\mathtt{avg}}(\mathtt{PS}^1, \mathtt{S}^2) = \left(\mathtt{DICE}^{\mathtt{pessim}}(\mathtt{PS}^1, \mathtt{S}^2) + \mathtt{DICE}^{\mathtt{optim}}(\mathtt{PS}^1, \mathtt{S}^2)\right)/2 \quad (4)$$

Pessimistic and optimistic Dice coefficients can be computed as follows. Let $\mathtt{S}^{1,\mathtt{pessim}}$ and $\mathtt{S}^{1,\mathtt{optim}}$ be the segmentations extending $\mathtt{PS}^1$ such that, for $(i, j) \in \mathcal{D}om$:

$$\mathtt{S}^{1,\mathtt{pessim}}_{ij} = \begin{cases} 1 - \mathtt{S}^2_{ij} & \text{if } \mathtt{PS}^1_{ij} = \mathtt{u} \\ \mathtt{PS}^1_{ij} & \text{otherwise} \end{cases} \qquad \mathtt{S}^{1,\mathtt{optim}}_{ij} = \begin{cases} \mathtt{S}^2_{ij} & \text{if } \mathtt{PS}^1_{ij} = \mathtt{u} \\ \mathtt{PS}^1_{ij} & \text{otherwise} \end{cases}$$

It can be shown that the pessimistic and optimistic Dice coefficients can be computed according to the following equalities:

$$\mathtt{DICE}^{\mathtt{pessim}}(\mathtt{PS}^1, \mathtt{S}^2) = \mathtt{DICE}(\mathtt{S}^{1,\mathtt{pessim}}, \mathtt{S}^2) \quad \mathtt{DICE}^{\mathtt{optim}}(\mathtt{PS}^1, \mathtt{S}^2) = \mathtt{DICE}(\mathtt{S}^{1,\mathtt{optim}}, \mathtt{S}^2)$$

3.2 Problem setting in CBR terms

It is assumed that a CNN has been trained on the images of a set of patients and that it can be used on a new patient to propose segmentations. This CNN is used in this approach as a black box:

$\texttt{CNN}$: scanner image of a slice $\mapsto$ proposed segmentation of this image

Now, let $\{\texttt{image}^1, \texttt{image}2, \ldots, \texttt{image}^N\}$ be the set of scanner images of slices for the new patient, at the level of the kidneys. The image indexes are ordered from the top to the bottom: $\texttt{image}^i$ is above $\texttt{image}^{i+1}$ ($i \in \{1, 2, \ldots, N-1\}$). In this application, for the new patient, $N = 136$ (there is a variation of the value of N depending on the size of the kidney which depends in particular on the age of the patient). Only a sample $\{\texttt{image}^s \mid s \in \texttt{IIMS}\}$ of images are manually segmented, to lower the required expert time ($\texttt{IIMS}$ is the set of indexes of the images that are manually segmented). In this application, $\texttt{IIMS} = \{1 + 11k \mid k \in \{0, 1, \ldots, 12\}\}$, so the required manual segmentation is about 30 to 45 minutes of expert time. For $s \in \texttt{IIMS}$, two segmentations are available:

- The segmentation computed by the CNN and denoted in the following $\texttt{x}^s$: $\texttt{x}^s = \texttt{CNN}(\texttt{image}^s)$;
- The manual segmentation of $\texttt{image}^s$ given by the expert and denoted in the following $\texttt{y}^s$.

The pair $(\texttt{x}^s, \texttt{y}^s)$ is considered as a source case that encodes the experience of *correcting* the CNN segmentation. So, the case base is $\texttt{CB} = \{(\texttt{x}^s, \texttt{y}^s) \mid s \in \texttt{IIMS}\}$.

Now, let $\texttt{IINMS} = \{1, 2, \ldots, N\} \setminus \texttt{IIMS}$, the set of indexes of images for which no manual segmentation is available. Let $\texttt{tgt} \in \texttt{IINMS}$ and $\texttt{x}^{\texttt{tgt}} = \texttt{CNN}(\texttt{image}^{\texttt{tgt}})$. The objective is to find a segmentation $\texttt{y}^{\texttt{tgt}}$ of $\texttt{image}^{\texttt{tgt}}$ by reusing $\texttt{CB}$, the available experience on CNN to expert corrections in the context of the new patient. For this purpose, a classical CBR approach with the retrieval of a single source case and an adaptation of this case is considered.

3.3 Retrieval

In this application, the implemented case retrieval is straightforward: for a given target problem $\texttt{x}^{\texttt{tgt}}$ with $\texttt{tgt} \in \texttt{IINMS}$, the source case $(\texttt{x}^s, \texttt{y}^s)$ which minimizes $|s - \texttt{tgt}|$ is chosen. This choice is justified by the fact that the CNN segmentation of the tumor varies rather smoothly according to the slice vertical position.

3.4 Adaptation using an analogical proportion

Given the target problem $\texttt{x}^{\texttt{tgt}}$ and the retrieved case $(\texttt{x}^s, \texttt{y}^s)$, the adaptation consists in considering that the proposed solution $\texttt{y}^{\texttt{tgt}}$ to $\texttt{x}^{\texttt{tgt}}$ is such that $\texttt{x}^s$ is to $\texttt{y}^s$ as $\texttt{x}^{\texttt{tgt}}$ is to $\texttt{y}^{\texttt{tgt}}$. In other words, the expert "correction" from $\texttt{x}^s$ to $\texttt{y}^s$ is applied on $\texttt{x}^{\texttt{tgt}}$. The implemented approach works at the pixel level, meaning that it is based on the analogical equations $\texttt{x}^s_{ij} : \texttt{y}^s_{ij} :: \texttt{x}^{\texttt{tgt}}_{ij} : \texttt{?}$ for each $(i, j) \in \mathcal{D}om$.

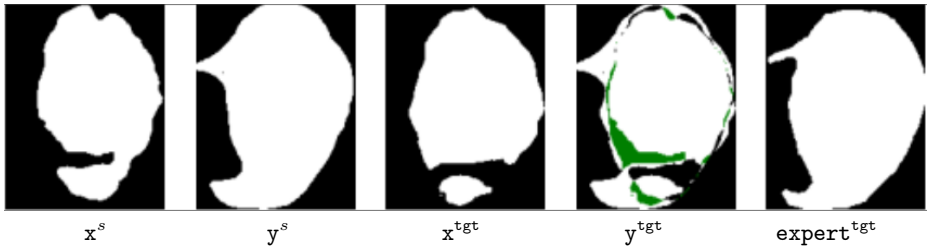

Fig. 2: Example of adaptation based on the proposed approach, with segmentations $\mathtt{x}^s$, $\mathtt{y}^s$ and $\mathtt{x}^{\mathtt{tgt}}$, and partial segmentation $\mathtt{y}^{\mathtt{tgt}}$. In a (partial) segmentation $\mathtt{S}$, $\mathtt{S}_{ij} = 1$ (resp. $\mathtt{S}_{ij} = 0$) is represented by a white (resp. black) pixel. For the partial segmentation $\mathtt{PS} = \mathtt{y}^{\mathtt{tgt}}$, $\mathtt{PS}_{ij} = \mathtt{u}$ is represented by a green pixel. $\mathtt{x}^s$ and $\mathtt{y}^s$ correspond to slice number $s = 90$, $\mathtt{x}^{\mathtt{tgt}}$, $\mathtt{y}^{\mathtt{tgt}}$ and $\mathtt{expert}^{\mathtt{tgt}}$ correspond to slice number $\mathtt{tgt} = 85$, respectively corresponding to what the CNN proposes, how the CBR approach corrects it and what the expert gives.

The minimal analogical equation is chosen, thus such an equation may have no solution. Therefore, the result of this adaptation is only a partial segmentation.

Formally, adaptation is computed this way:

- For each $(i, j) \in \mathcal{D}om$,
  - Solve the analogical equation $\mathtt{x}^s_{ij} : \mathtt{y}^s_{ij} :: \mathtt{x}^{\mathtt{tgt}}_{ij} : ?$.
  - If it has no solution, then $\mathtt{y}^{\mathtt{tgt}}_{ij} \leftarrow \mathtt{u}$.
  - Otherwise, let d be the solution of this equation.[2] Then $\mathtt{y}^{\mathtt{tgt}}_{ij} \leftarrow d$.
- The partial segmentation $\mathtt{y}^{\mathtt{tgt}}$ is proposed as a (partial) solution to $\mathtt{x}^{\mathtt{tgt}}$.

Figure 2 presents an example of the outcome of such a process.

3.5 Evaluation

The data for this evaluation are constituted by the scanner images of 14 patients with about 100 images per patient and an expert segmentation for each image. A CNN was trained on the images and segmentations of 13 patients. For the 14^{th} patient, the expert segmentations of 1/11 images where kept, with a constant gap between them, which has constituted the case base. The approach was tested on the 10/11 images remaining.

Figure 3 and Table 1 present the results with pessimistic, average and optimistic 3D Dice coefficients, compared to the 3D Dice coefficients for the segmentations provided by the CNN and by OV^{2}ASSION. The figure presents these results slice by slice, for slices number 40 to 100 (for the other slices, the number of pixels for the predicted and expert segmentations that belong to the tumor is very low, making the Dice coefficient of little relevance). The table presents these results with 3D Dice coefficients, comparing predicted and expert segmentations on the tumor as a whole.

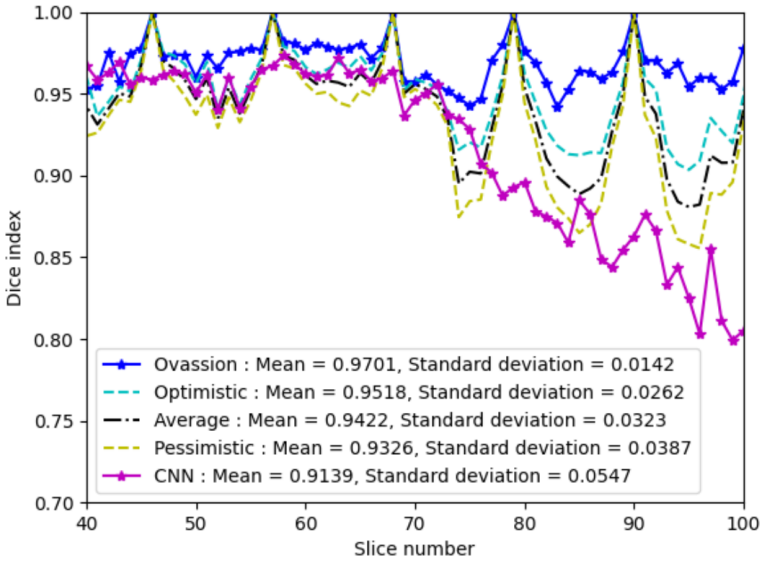

Fig. 3: Evaluation of the approach, slice by slice.

Approach	$\texttt{DICE}^{\texttt{pessim}}$	$\texttt{DICE}^{\texttt{avg}}$	$\texttt{DICE}^{\texttt{optim}}$	Calculation time (s)
OV^{2}ASSION	—	0.9489	—	$762 + 209 = 971$
proposed approach	0.8806	0.8974	0.9143	$299 + 473 = 772$
CNN	—	0.8648	—	299

Table 1: Dice evaluation and calculation time of the general approach. For approaches without undecided pixels (CNN and OV^{2}ASSION), the values $\texttt{DICE}^{\texttt{avg}}$ and DICE are equal. The computation time does not include the CNN training.

The initial CNN training (with 200 epochs, and with 16 as batch size) took 2 hours and 42 minutes and the segmentation 5 minutes (this time is common to the three approaches). As indicated in Table 1, the computation on the images of the new patient take 299 seconds with the trained CNN, 772 seconds with our proposed approach (written in Python), and 971 with OV^{2}ASSION.[3]

This shows that the simple case-based approach presented in this section actually improves the output of the CNN thanks to a sample of additional segmentations, but that it is still overcome by the OV^{2}ASSION approach using the same sample (and using a little bit more time). It also shows that the choice of values for undecided pixels could improve the approach, which is studied in the next section.

[2] With the chosen analogical proportion, the solution, when it exists is unique.

[3] All these computations have been made on the Mesocenter of computation of Franche-Comté, equiped of processor Intel(R) Xeon(R) Gold 6126 CPU @ 2.60GHz and Nvidia Volta V100 GPU.

4 Improvements of the approach based on post-processing techniques

This section presents some ways to improve the approach of Section 3 and ends with an evaluation.

4.1 Two local strategies for choosing a value for undecided pixels

An analogical equation $a:b::c:?$ on Booleans with minimal analogical proportion has no solution d for $(a,b,c) \in \{(0,1,1),(1,0,0)\}$. Two ways of proposing a solution are considered.

First, considering the Sheldon Klein proportion gives the solution $d = 1 - c$.

Second, considering that the patterns for minimal analogical proportions are (u,u,v,v) and (u,v,u,v) $(u,v \in \{0,1\})$, the solution requiring the least modifications to these patterns are $d = 1$ for $(a,b,c) = (0,1,1)$ and $d = 0$ for $(a,b,c) = (1,0,0)$, i.e., $d = c$ for both cases (where the considered modifications are flips of Boolean values). This strategy is called *conservative approach* in the following.

Therefore, these two strategies give the opposite proposition for changing a value $\mathtt{y}_{ij}^{\mathtt{tgt}} = \mathtt{u}$ to $\mathtt{y}_{ij}^{\mathtt{tgt}} \in \{0,1\}$. These two strategies have been compared, following the same evaluation protocol as in Section 3.5, and gives a significant preference for the second one (cf. the evaluation, at the end of this section).

4.2 Use of convolution masks

The approaches presented and compared in Section 4.1 are limited by the fact that they do not take into account the context, i.e., the segmentations of the pixels that are around an undecided pixel. For example, it seems natural to consider that an undecided pixel surrounded by pixels that are all 0s or all 1s is likely to be associated with the same common value. More generally, the strategy consists in choosing the "main tendancy" of these pixels. For this purpose, a convolution mask is used. By contrast to the CNN approach, the weights of such a mask are not learned but are chosen parameters.

For the sake of simplicity of the presentation, in this section, for a partial segmentation $\mathtt{PS}$, if (i,j) is an "out of bound index" pair (i.e., $(i,j) \notin \mathcal{D}om$), then $\mathtt{PS}_{ij} = \mathtt{u}$ (the undecided value $\mathtt{u}$ is assimilated to $\mathtt{u} = 0.5$). 3 types of convolution masks were tested, with different values of their weights, knowing that the results (described in§4.4) are poorly sensitive to the choice of the weights.

Use of a 3×3 *convolution.* A 3×3 convolution mask is a 3×3 matrix $\mathtt{M}$ of real numbers, assumed to be indexed on $\{-1,0,1\}$ (i.e., $\mathtt{M}_{kl}$ is defined for $-1 \leq k,\ell \leq 1$). Following a symmetry principle, the masks considered here are of the form $\begin{bmatrix} w_C & w_B & w_C \\ w_B & w_M & w_B \\ w_C & w_B & w_C \end{bmatrix}$, thus are parametrized by w_M (the weight at

the middle), w_B (the weight at a border) and w_C (the weight at a corner). Moreover, the mask is assumed to be normalized: the sum of its elements is $w_M + 4w_B + 4w_C = 1$.

Given a partial segmentation $\mathtt{PS}$ and a mask $\mathtt{M}$, a new partial segmentation $\mathtt{PS}'$ is defined as follows, for $(i,j) \in \mathcal{D}om$:

$$\text{with } \sigma = \sum_{(k,\ell)\in\{-1,0,1\}^2} \mathtt{M}_{k\ell} \times \mathtt{PS}_{i+k,j+\ell}, \qquad \mathtt{PS}'_{ij} = \begin{cases} 0 & \text{if } \sigma < 0.5 \\ \mathtt{u} & \text{if } \sigma = 0.5 \\ 1 & \text{otherwise} \end{cases}$$

This partial segmentation $\mathtt{PS}'$ is then proposed as a result.

Different values of the parameter triple (w_M, w_B, w_C) were considered. In order to have $\mathtt{PS}'$ extending $\mathtt{PS}$, the choice $w_M > 0.5$ was made: when $\mathtt{PS}_{ij} = \mathtt{u}$, this weight is of no consequence on the decision made and when $\mathtt{PS}_{ij} \neq \mathtt{u}$ then $\mathtt{PS}'_{ij} = \mathtt{PS}_{ij}$. As a consequence, $w_M > 4w_B + 4w_C$: the weight of the middle is greater than the weight of all the other pixels (borders and corners). Moreover, the relation $w_B > w_C$ was made, since the non-corner borders of the mask being closest to the middle that the corners, they have more importance. With these constraints, various values of the parameter triple were chosen, given similar result. The evaluation was made with the triple $(0.6, 0.075, 0.025)$.

Use of a 5×5 *convolution.* A 5×5 convolution mask was used in a similar way. The convolution mask is based on 6-tuple of weights (because of the symmetry, there are only 6 different values and they are ordered in the tuple by decreasing distance to the center). The evaluation was carried out with the weight tuple $(0.6, 0.05, 0.025, 0.025, 0, 0)$.

Use of a $3 \times 3 \times 3$ *convolution.* A $3 \times 3 \times 3$ mask is the fusion of three 3×3 masks obtained on the target slice and the slices preceding and following this target slice. This construction supposes beforehand to have obtained the partial segmentations of all the target slices. The $3D$ convolution mask is based on a 4-tuple of weights (because of the symmetry, there are only 4 different values and they are ordered in the tuple by decreasing distance to the center). In the experiment, the chosen tuple was $(0.6, 0.0333, 0.0083, 0.0125)$.

4.3 Use of closures

The idea has emerged to use convex closures for improving the result and then, more generally, to use other closure functions.

Convex closure and other closures. Given a set X of points of the plane, the *convex closure* (also called convex hull) of X, is the minimal set of points $\mathcal{CC}(X)$ such that for every pair $(P,Q) \in X^2$, the segment $\mathtt{[}P,Q\mathtt{]}$ is included in X. Formally:

$$\mathcal{CC}(X) = \bigcup \{ \mathtt{[}P,Q\mathtt{]} \mid (P,Q) \in X^2 \}$$

Fig. 4: Example of post-treatment using the rectangular closure. Left: $\mathtt{y}^{\mathtt{tgt}}$ provided by the approach of §3. Center: correction of $\mathtt{y}^{\mathtt{tgt}}$ using rectangular closure. Right: expert segmentation.

The function $\mathcal{CC}$ belongs to the family of closure functions. A *closure function* $\mathcal{C}$ on sets of points X is a function that is extensive ($X \subseteq \mathcal{C}(X)$), non-decreasing (if $X \subseteq Y$ then $\mathcal{C}(X) \subseteq \mathcal{C}(Y)$), and idempotent ($\mathcal{C}(\mathcal{C}(X)) = \mathcal{C}(X)$).

Another example of closures on sets of points of the Euclidian plane is $\mathcal{C}_{\Delta}$, where Δ is a line of the plan: for a set of points X of the plane, $\mathcal{C}_{\Delta}(X)$ is defined in a similar way as $\mathcal{CC}(X)$, except that the segment $[P, Q]$ has to be parallel to Δ. In particular, given an affine coordinate system Oxy, horizontal closures (resp. vertical closure, rising diagonal closure, and downward diagonal closure) is the closure $\mathcal{C}_{\Delta}$ where Δ is defined by the equation $y = 0$ (resp. $x = 0$, $y = x$ and $y = -x$). A final example of closure is the rectangular closure, i.e., the smallest rectangle whose edges are parallel to the axis that contains all the values of a set of points X.

How to use closures for the post-treatment. The notions of closures are defined on sets of points of the plane, but are applied on images, making the approximation that a pixel corresponds to a point. Moreover, a segmentation $\mathtt{S}$ is assimilated to the set of pixels (or points) $(i, j) \in \mathcal{D}om$ such that $\mathtt{S}_{ij} = 1$: $\mathtt{S} = \big\{(i, j) \in \mathcal{D}om \mid \mathtt{S}_{ij} = 1\big\}$, hence the notion of closure $\mathcal{C}(\mathtt{S})$ of a segmentation $\mathtt{S}$ (given a closure function).

Now, given three segmentations $\mathtt{x}^s$, $\mathtt{y}^s$ and $\mathtt{x}^{\mathtt{tgt}}$ (corresponding to a source case $(\mathtt{x}^s, \mathtt{y}^s)$ and a target problem $\mathtt{x}^{\mathtt{tgt}}$), the approach of section 3 (possibly with some post-treatments of the current section) provides a partial segmentation $\mathtt{y}^{\mathtt{tgt}}$. For each of these three segmentations, a closure can be computed with a closure function $\mathcal{C}$, hence the three segmentations $\mathcal{C}(\mathtt{x}^s)$, $\mathcal{C}(\mathtt{y}^s)$ and $\mathcal{C}(\mathtt{x}^{\mathtt{tgt}})$. Then, a partial segmentation $\mathtt{PS}$ can be found by applying the section 3 approach on these three segmentations: $\mathtt{PS}_{ij}$ is obtained by solving the analogical equations $\mathcal{C}(\mathtt{x}^s)_{ij} : \mathcal{C}(\mathtt{y}^s)_{ij} :: \mathcal{C}(\mathtt{x}^{\mathtt{tgt}})_{ij} : ?$ (for any $(i, j) \in \mathcal{D}om$). This partial segmentation $\mathtt{PS}$ is used as a prediction of the closure of the desired segmentation. Thus, for $(i, j) \in \mathcal{D}om$, it is expected that $\mathtt{y}^{\mathtt{tgt}}_{ij} \leq \mathtt{PS}_{ij}$. Therefore, the situation $\mathtt{y}^{\mathtt{tgt}}_{ij} > \mathtt{PS}_{ij}$ is considered to be abnormal and, when it occurs, $\mathtt{y}^{\mathtt{tgt}}$ is changed into $\mathtt{PS}_{ij}$. Figure 4 shows that a closure can remove distant noisy pixels. This approach has been tested for horizontal, vertical, rising diagonal, downward diagonal, and rectangular closures.

Method	DICE$^{\text{pessim}}$	DICE$^{\text{avg}}$	DICE$^{\text{optim}}$	Computation time (s)
OV^{2}ASSION	—	0.9489	—	$762 + 209$
(b) conservative approach	—	0.9129	—	$299 + 473 + 471$
(c) 5×5 convolution	0.9078	0.9078	0.9078	$299 + 473 + 643$
(c) 3×3 convolution	0.8984	0.9056	0.9128	$299 + 473 + 496$
(c) $3 \times 3 \times 3$ convolution	0.9010	0.9028	0.9045	$299 + 473 + 711$
(a) general approach	0.8806	0.8974	0.9143	$299 + 473$
(d) rectangular closure	0.8713	0.8904	0.9095	$299 + 473 + 1644$
(d) horizontal closure	0.8635	0.8829	0.9024	$299 + 473 + 1626$
(d) vertical closure	0.8630	0.8828	0.9026	$299 + 473 + 1632$
(b) Sheldon Klein	—	0.8820	—	$299 + 473 + 487$
(d) rising diagonal closure	0.8621	0.8815	0.9010	$299 + 473 + 1573$
(d) downward diagonal closure	0.8612	0.8808	0.9005	$299 + 473 + 1385$
CNN	—	0.8648	—	299

(a) General approach, providing a partial segmentation (cf. §3).
(b) Using two local strategies for choosing a value for undecided pixels (cf. §4.1).
(c) Using convolution masks (cf. §4.2).
(d) Using horizontal, vertical, diagonal and rectangular closures (cf. §4.3).

Table 2: Evaluation of various improvement strategies.

4.4 Evaluation

Table 2 presents an evaluation of the different segmentation approaches: CNN, OV^{2}ASSION, the case-based approach of Section 3, and the improvements of this approach proposed in the current section.

The lines are ordered by decreasing 3D DICE$^{\text{avg}}$. For these data, the best result, according to DICE$^{\text{avg}}$ is obtained by applying the general approach of Section 3 followed by the conservative choice of pixels described in Section 4.1. It is noteworthy that (1) for other segmentation problems (e.g. kidneys) this order may be different and (2) some of the other approaches may be improved.

For this first version of our approach, the results are just over the middle between CNN and OV^{2}ASSION ones, but there is room for improvement, as is detailed in the conclusion of the paper. Every approach considered here uses the CNN computation (that requires 2 hours and 42 minutes of computing time).

The required computing times of our best approach and OV^{2}ASSION are similar (1243 s for the former and 971 s for the latter, knowing that the former has been programmed in Python and the second uses a highly optimized library). Now, if the expert just want an improvement of the CNN segmentation of one slice (at at time), as a starting point for manual segmentation, our approach needs only about $1243/100 \simeq 12\,s$.

5 Conclusion

This article has presented a first case-based approach to improve CNN-based image segmentations given a new sample of expert segmentations from the current

context, i.e., in the application, in the context of a new patient suffering from nephroblastoma. This approach improves the result of the CNN but gives still less accurate results than the OV^{2}ASSION approach that consists in re-running the CNN with the additional segmentation sample.

This work gives encouraging results, but has to be improved. Other directions of work follow this study.

First, it can be noticed that the approach presented in this paper relies only on segmentations (given by the CNN and the experts). By contrast, the CNN and the OV^{2}ASSION system use also the images taken from the scanner, with gray levels. A future work aims at using these images for improving the approach. This can be done as an additional post-processing technique for the choice of undecided values for segmentation. This can be done also by using analogical proportions directly on gray levels. More precisely, the idea is as follows. With the current approach, the adaptation consists in solving the analogical equation

$$\mathtt{x}^s : \mathtt{y}^s :: \mathtt{x}^{\mathtt{tgt}} : ?$$

pixel by pixel. Now, if the mapping of `image`, an image described by gray levels, into a segmentation `S` is denoted by $(\mathtt{image} \mapsto \mathtt{S})$ then, the adaptation would consist in solving the analogical equation

$$(\mathtt{image}^s \mapsto \mathtt{x}^s) : (\mathtt{image}^s \mapsto \mathtt{y}^s) :: (\mathtt{image}^{\mathtt{tgt}} \mapsto \mathtt{x}^{\mathtt{tgt}}) : ?$$

and obtaining a mapping $(\mathtt{image}^{\mathtt{tgt}} \mapsto \mathtt{y}^{\mathtt{tgt}})$ from which a segmentation would be obtained. How to actually realize this idea is still a challenging question.

A second work consists in studying the use of geometrical moments of a segmentation, i.e. the area of the set of tumor pixels, its centroid and its orientation. The principle is that the geometrical moments can be predicted by analogy (e.g. by solving in $\mathbb{R}$ the analogical equation $\mathtt{area}(\mathtt{x}^s) : \mathtt{area}(\mathtt{y}^s) :: \mathtt{area}(\mathtt{x}^{\mathtt{tgt}}) : ?$) and to use these predictions to modify $\mathtt{y}^{\mathtt{tgt}}$ (e.g. by making choices for undecided pixels based on the difference between the predicted area and $\mathtt{area}(\mathtt{y}^{\mathtt{tgt}})$).

Third, it is planned to use the interpolation approach presented in [8] to improve the results. This approach is based on the retrieval of two source cases $(\mathtt{x}^r, \mathtt{y}^r)$ and $(\mathtt{x}^s, \mathtt{y}^s)$ such that $\mathtt{x}^{\mathtt{tgt}}$ is *between* $\mathtt{x}^r$ and $\mathtt{x}^s$. Then, it is plausibly inferred that the expected segmentation $\mathtt{y}^{\mathtt{tgt}}$ has to be between $\mathtt{y}^r$ and $\mathtt{y}^s$ (at a pixel by pixel level or globally).

Fourth, the search for a smart combination of these approaches remains to be studied.

Finally, these approaches will be tested for other anatomical structures which still challenge the segmentation using CNN and OV^{2}ASSION: kidneys, veins, arteries, and cavities.

Acknowledgement

The authors would like to thank the European Union for financing this project as part of the SAIAD and SAIAD 2 INTERREG V programs and the SAIAD and SAIAD 2 consortiums partners. Computations have been performed on the supercomputer facilities of the *Franche-Comté Computation Mesocenter*.

References

1. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE transactions on pattern analysis and machine intelligence **39**(12) (2017) 2481–2495
2. Chaussy, Y., Vieille, L., Lacroix, E., Lenoir, M., Marie, F., Corbat, L., Henriet, J., Auber, F.: 3D reconstruction of Wilms' tumor and kidneys in children: Variability, usefulness and constraints. J. of Pediatric Urology **16**(16) (2020) 830.e1 – 830.e8
3. Chen, L.C., Papandreou, G., Kokkinos, I., Murphy, K., Yuille, A.L.: Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE transactions on pattern analysis and machine intelligence **40**(4) (2017) 834–848
4. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3) (1945) 297–302
5. Klein, S.: Analogy and mysticism and the structure of culture (and Comments & Reply). Current Anthropology **24 (2)** (1983) 151–180
6. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. nature **521**(7553) (2015) 436
7. Lepage, Y., Lieber, J., Mornard, I., Nauer, E., Romary, J., Sies, R.: The French Correction: When Retrieval is Harder to Specify than Adaptation. In Watson, I., Weber, R., eds.: ICCBR 2020 - 28th Int. Conf. on Case-Based Reasoning. (2020)
8. Lieber, J., Nauer, E., Prade, H., Richard, G.: Making the best of cases by approximation, interpolation and extrapolation. In: ICCBR 2018 - 26th Int. Conf. on Case-Based Reasoning. (2018)
9. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I.: A survey on deep learning in medical image analysis. Med Image Anal **42** (Dec 2017) 66–88
10. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition. (2015) 3431–3440
11. Marie, F., Corbat, L., Chaussy, Y., Delavelle, T., Henriet, J., Lapayre, J.C.: Segmentation of deformed kidneys and nephroblastoma using Case-Based Reasoning and Convolutional Neural Network. Expert Systems with Applications **127** (2019) 282–294
12. Noh, H., Hong, S., Han, B.: Learning deconvolution network for semantic segmentation. In: Proc. of the IEEE internat. conf. on computer vision. (2015) 1520–1528
13. Prade, H., Richard, G.: Analogical proportions: Why they are useful in AI. In: Proc. 30th Int. Joint Conf. on Artificial Intelligence (IJCAI-21), (Z.-H. Zhou, ed.) Virtual Event / Montreal, Aug. 19-27. (2021) 4568–4576
14. Riesbeck, C.K., Schank, R.C.: Inside Case-Based Reasoning. Lawrence Erlbaum Associates, Inc., Hillsdale, New Jersey (1989) Available on line.
15. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, Springer (2015) 234–241
16. Thong, W., Kadoury, S., Piché, N., Pal, C.J.: Convolutional networks for kidney segmentation in contrast-enhanced CT scans. Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization **6**(3) (2018) 277–282