Porcelli Gianluca

Prof. GUERRIERO Andrea

IMAGE PROCESSING AND ARTIFICIAL VISION

GBM saliency detection

The work presented explains how to segment the brain tumour area in absence of interaction with user basing his technique on a saliency map constructed from three different resonance techniques.

1. Overview
2. Requested material
3. Pseudo-color technique
4. Saliency map
5. Modification of saliency map
6. Main 3D object detection
7. Automated segmentation
8. Conclusions and further developments

1.Overview

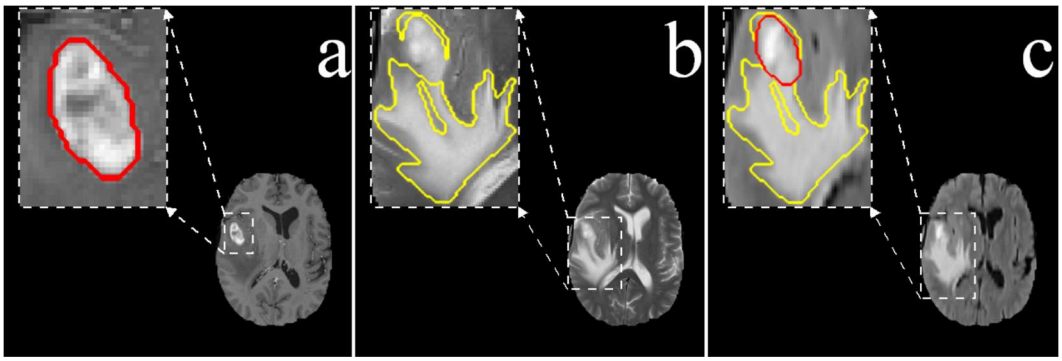
The computerized automatic detection of ROI (regions of interest) is an important step towards a progressive streamlining of a doctor's workload. A new algorithm, based on visual salience, is analysed here for the identification of tumour regions. The GBM salience detection model is designed starting from the concept of visual salience. A visually salient region is generally rare in an image and contains much discriminating information that draws attention to it. Exploiting images, such as magnetic resonance imaging (MRI) is important in order to improve tumour management through non-invasive detection. Currently it is necessary to integrate different images in order to provide a specific clinical prediction of patients. Glioblastoma multiforme (GBM) is the most common brain tumour and for this reason it is the most studied. Repeating tumour biopsies can be harmful for brain tumour patients, therefore, non-invasive techniques such as imaging are playing an important role in the treatment of GBM. Magnetic resonance imaging provides detection capabilities concerning abnormalities at brain level in terms of both shape and volume. Reconstruction and visualization of detailed 3D images is possible using MRI. We'll use different pulse sequences and modified images because MRI depends on biologically variable parameters such as longitudinal relaxation time (T1) and transverse relaxation time (T2). None of these sequences is known to be capable of depict the full extent of a malignant tumour. Segmentation and detection, of different regions of interest (ROI) in medical images, are usually performed manually by experts for treatment planning and diagnosis.

Tumours may have different biological characteristics in different patients although they have developed in the same organ. Furthermore, variations within a single tumour can cause marked differences in its biological characteristics mainly due by changes in blood flow; this variations are present in different ways in Flair, T1c and T2 sequences. This topic highlights the usefulness of overlapping multiple MR imaging channels in the identification and extraction of heterogeneous tumour regions.

It was observed that radiologists generally delineate the nucleus of the gross tumour from the T1C plates because the tumour border becomes more visible for the accentuated contrast between grey and white matter. The T2 channel, which provides a better contrast between brain tissues and cerebrospinal fluid (CSF), is

preferred to delineate the edema region. Although the border of the edema becomes blurred in FLAIR, it contains both the tumour and the edema.

Each pixel in the tumour is, therefore, defined by its Lab image given by the overlapping of the T1c, T2, Flair sequences.



(a) T1c focuses on tumour structure and border (Red)

(b) T2 associated with edema (Yellow)

(c) Flair demonstrates both the edema (Yellow) and solid core (Red).

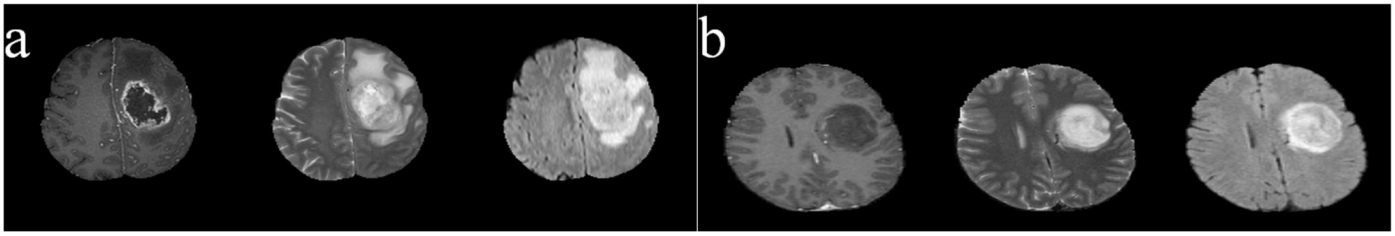
When we look at an image our eyes are immediately attracted by salient areas. Attention is a high-level visual task that requires the search for a specific object. Having an image as input, the computational salience generates a topographic map used to determine what is salient; the salience algorithm draws attention to an object by simulating the movement of the human eye.

Data resulting from a study showed that often attention is immediately attracted by a salient element, despite the existence of many other objects (or distractors), without scanning the image. The role of the visual attention mechanism has been used in medical images and the goal was to model visual search. Nodine and Kundel introduced a study on perception and they collected tracking data on the eyes of radiologists while observing radiographic images; the study aimed to develop a model used to predict the event sequence that is useful for tumour segmentation.

Generally, the salience mechanism uses the characteristics of colour images, which is why a pseudo-colouring strategy is planned; for this purpose are combined three different sequences (Flair, T1c, T2) useful for generating the saliency maps. Finally, a three-dimensional model is generated by stacking the individual maps. Starting from the three-dimensional model it will be possible to identify the most relevant object that will be finally segmented automatically.

2.Requested material

The figure shows example images of glioblastoma multiforme with the three columns corresponding to the sequences T1C, T2 and FLAIR. The concept of salience is used to allow rapid algorithm focusing on ROI.



a) High-grade (HG)
b) Low-grade (LG)

The wanted tumour is often classified as "low grade", which means that the tumour cells seem to divide more slowly under the microscope, or "high grade", which means that the cells appear more aggressive under a microscope. Low-grade tumours can cause problems even when they are not malignant (cancerous) by pressing on normal brain structures and causing symptoms. Glial cells support and isolate neurons.

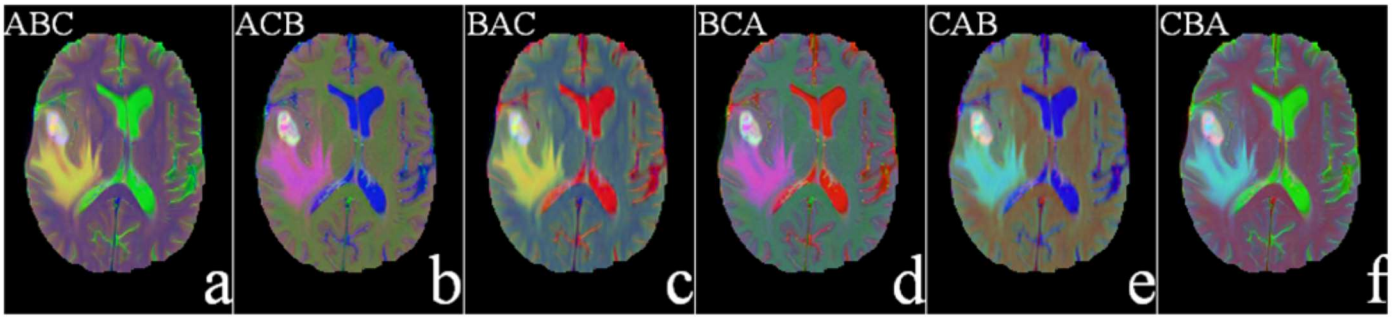
For example, about 60% of brain tumours in children occur in the infratentorial brain. The infratentorial brain is the lower part of the brain near the centre of the back of the head and includes the cerebellum and the brain stem. This is the part of the brain that controls movement and balance.

In this example we will treat high grade tumours as they are easier to spot.

3.Pseudo-color technique

Colour digital images are often constructed from three overlapping colour channels, ie. red, green and blue (RGB). These kind of images can be composed by three grayscale images in six different combinations varying the channel associated with the individual sequence type.

The colours visible in the images are "false" colours (or pseudo-colours) and do not match the colour of the fabrics. However, such pseudo-colouring produces an MR image containing about 65,536 times more information than a single-channel gray scale; the tumour area appears white in all combinations of sequences. This happens because tumour is equally bright in all three RGB channels. The edema, on the other hand, appears yellow in figures (a) and (c) due to its brightness along the R and G channels.



(a) FLAIR—T2—T1C

(b) FLAIR—T1C—T2

(c) T2—FLAIR—T1C

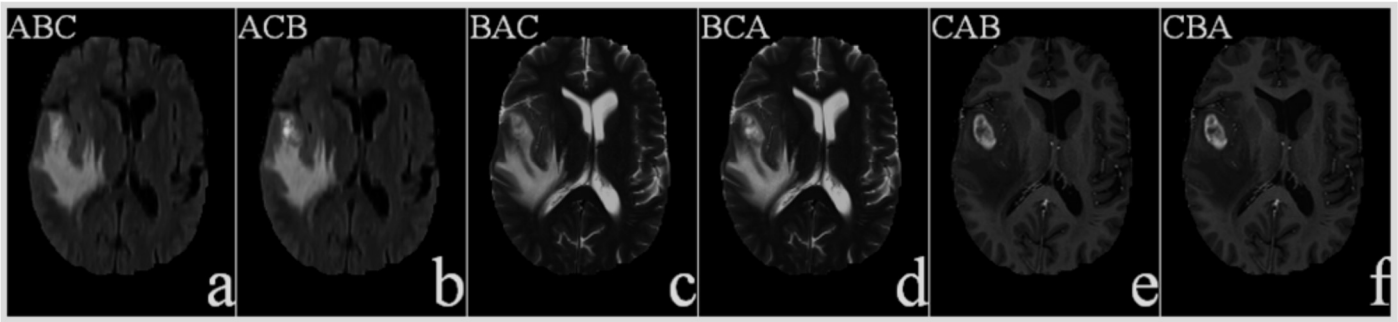
(d) T2—T1C—FLAIR

(e) T1C—FLAIR—T2

(f) T1C—T2—FLAIR

Since the salience detection algorithm depends on the colour difference between centre and contour of a region with its neighbours, it would be desirable to decorrelate the luminance from the chrominance information. The RGB system is considered not suitable for delineating tumour regions. "Brightness" is designed to approximate human vision, which is very sensitive to green, but less to blue. If you change the current RGB colour space into a space based on luminance, the result often seems more correct for the eye. For this purpose the institution (CIE) recommended the use of CIE - Lab to represent colour difference. Lab is a

conversion of the same information to a component of brightness L * and two components of colour - a * and b *. The brightness is kept separate from the colour that are not real but due to a pseudo-colouring. In the conversion of the images from RGB to Lab it has been proved that in the first two sequences (ABC and ACB) only the tumour regions and edema are highlighted while all the other regions are suppressed. Therefore we can use one of these two sequences for next passes. In this study the ABC sequence was chosen for the subsequent detection of salience.



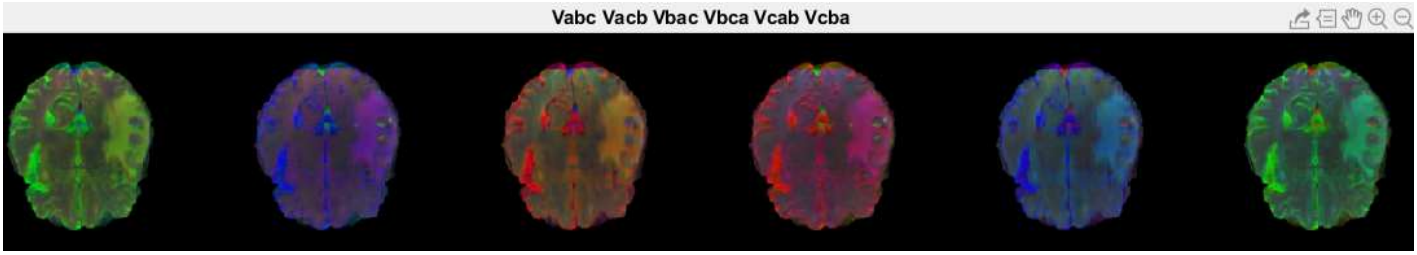
The algorithm for the pseudo-colouring scheme consists of two steps.

Step 1: Create a pseudo-coloured RGB image from the FLAIR, T2, T1C sequences.

Step 2: Switch from RGB colour space to CIE - Lab colour space to improve local contrast.

This step will be followed by the generation of contrast-based saliency maps for the detection of tumour regions, as described in the following section.

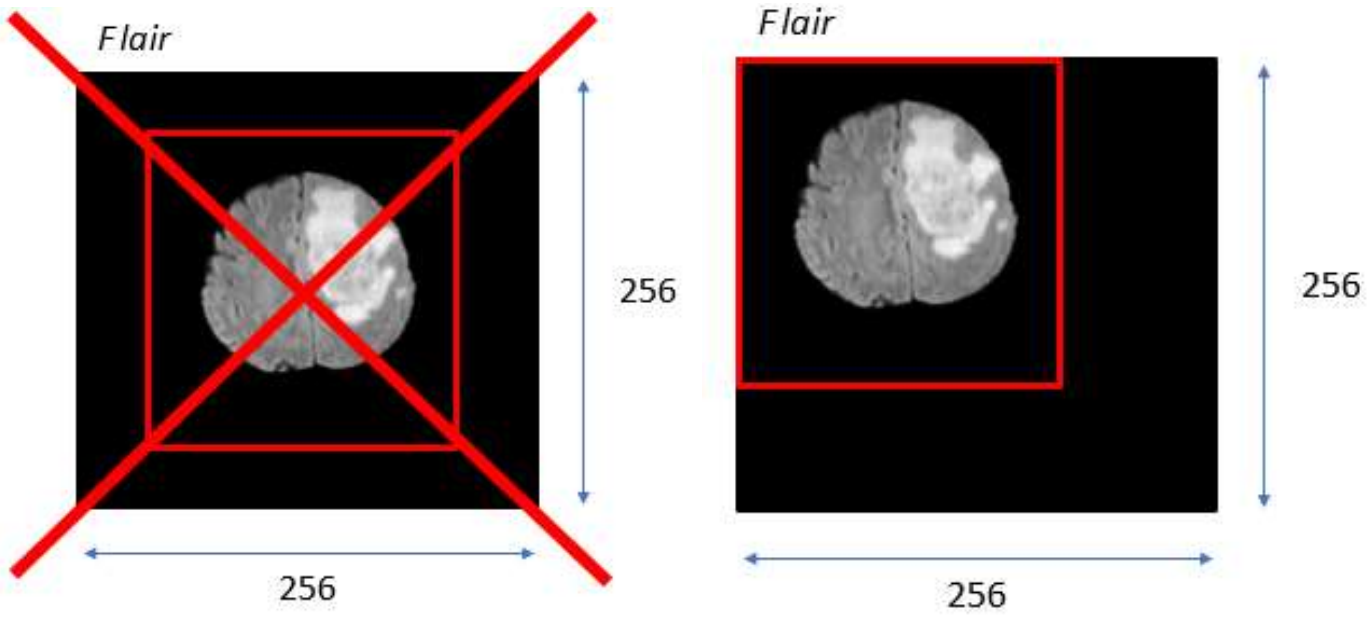
The figure shows the real pseudo-coloration of the third clinical case:



4.Saliency map

In an image a salient region is formed by one or more important objects because they are distinguished from their surroundings. A Lab image (of size M × N) is preliminarily traced to a square one for reasons understandable later. Since the database can contain images of different sizes, these must be converted into a uniform size (approximately with width = 256).

it is not specified where additional pixels useful for algorithm operation should be entered. Studying salience building I thought of positioning the additional pixels not around the initial image but in this way.



This choice means that in terms of salience the additional pixels give less contribution and this is useful as they are not part of the original image.

The initial image is then decomposed into several non-overlapping blocks Ri (or patch) with dimensions k × k pixels (where w is a multiple of k). The number of patches (w / k × w / k) corresponds to the number of pixels in the salience map. The i-th block will be represented by its average L, a, b.

$$\bar{R_i^{L^*}} = \frac{\sum I(R_i^{L^*})}{k \times k}, \bar{R_i^{a^*}} = \frac{\sum I(R_i^{a^*})}{k \times k}, \bar{R_i^{b^*}} = \frac{\sum I(R_i^{b^*})}{k \times k}.$$

Then the salience of each patch is calculated with respect to all the other patches in the image. Colour is considered the most important feature in the approach that uses the colour difference between regions. This approach provides an effective way to highlight the salient regions compared to non-salient patches.

The colour difference between a pair of patches is defined as "Euclidean distance" between the corresponding mean chromatic values of L a b. Therefore, for the patch Ri, the salience Sc (Ri) is calculated as the sum of the colour difference.

$$S_c(R_i) = \sum_{j,j\neq i} \sqrt{(\bar{R_i^{L^*}} - \bar{R_j^{L^*}})^2 + (\bar{R_i^{a^*}} - \bar{R_j^{a^*}})^2 + (\bar{R_i^{b^*}} - \bar{R_j^{b^*}})^2}$$

$$\forall i, j \in \{1, \ldots, (w/k \times w/k)\}.$$

The salience is calculated as the sum of the colour differences between the patch under examination and all the other ones. If this sum is big, the patch under examination is considered salient. Generally, while the most salient patches are concentrated around spatially adjacent areas, the non-salient ones can be distributed throughout the entire image. If a region is relevant, then the probability is large for its surrounding regions to be salient, while the probability of finding salient patches far from the ROI is very low. Therefore the influence of adjacent regions is important when calculating the salience of a region. With this in mind the spatial distance between the patches has been incorporated as another factor for calculating the saliency map. In the process we consider the difference of the Lab colour values between any two blocks and the spatial distance between them.

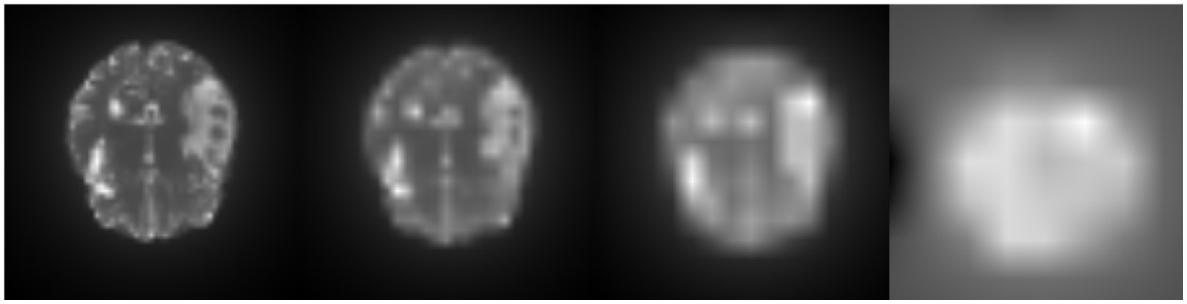
$$S(R_i) = \sum_{j,j\neq i} \frac{1}{1 + d(R_i, R_j)} \times S_c(R_i).$$

$$d(R_i, R_j) = \sqrt{(\bar{x}_{R_i} - \bar{x}_{R_j})^2 + (\bar{y}_{R_i} - \bar{y}_{R_j})^2},$$ *Spatial distance between patches*

Thinking of how an observer looks at a distant scene, we imagine that his attention would focus on the whole salient region; approaching the scene, he would pay more attention to the details within that region of interest. This characteristic of human visual attention has been exploited in the model through the evaluation of salience on multiple scales. By partitioning an image into smaller patches, we can clearly highlight the salient object along with its details, but the salience map for a larger patch can help to better locate a salient object. Finally, a rescaling was applied to the different maps using the bilinear interpolation method. After rescaling operations a cut of the additional pixels is executed.

The block size k refers to the resolution of the saliency map and consequently its patch size used for the construction.

The figure shows the maps corresponding to the third real case under examination:



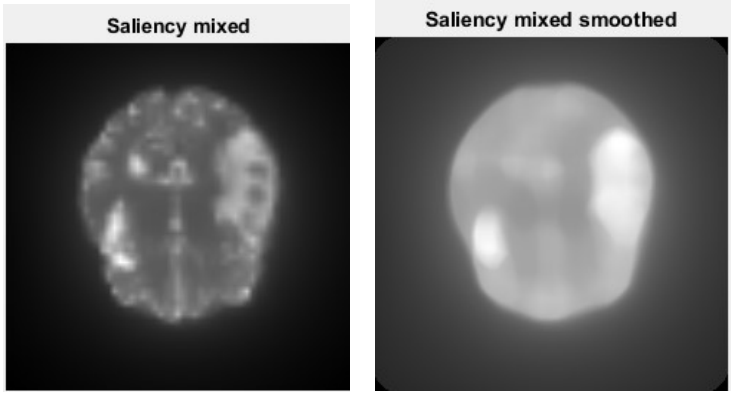
(a) 4 × 4

(b) 8 × 8

(c) 16 × 16

(d) 32 × 32

The saliency algorithm is used simultaneously on several scales to capture the salient regions in the MR image at different resolution levels. It is assumed that those regions consistently highlighted on different resolutions (possibly on all) are the salient ones. Therefore we superimpose these salience maps to calculate the final map. The figure shows the final map and its filtered version of the third real case under examination:



a)Final map

b)Smoothed final map

The final map is composed by a weighted sum of saliency maps:

$$S = \sum_{k=4,8,16,32} r^k \times \hat{S}^k$$

The first term is the weight corresponding to the k-th saliency map. In the case shown, a weight of 1/4 was chosen for all maps.

In the future we could evaluate a different weighing approach by studying which maps are more significant.

Multiple scale saliency maps are used to generate a final map using a fusion strategy. The assumption underlying spatial coincidence identifies a region as salient only if it is constantly salient on several scales.

Finally the filtered map is obtained by applying a median filter of size 25 × 25 to uniform the saliency map S; this filtering is useful to focus attention on the main region (ROI). The last map acts as a reference for the subsequent segmentation.

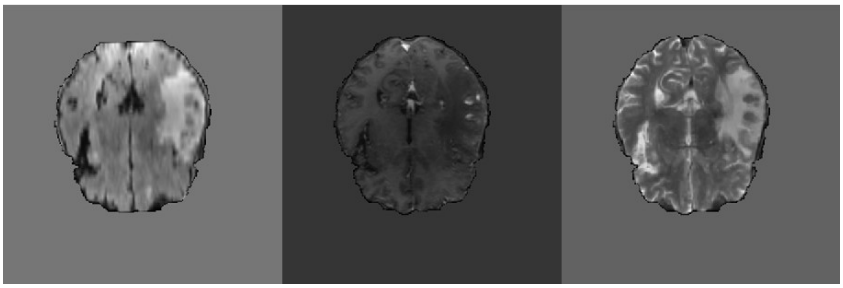
5.Modification of saliency map

In the sections that have a foreground too small compared to the background, this salience detection algorithm can produce false positive results by confusing the brain section with the tumour area. This is evident from the MR images located in initial and final sections of the 3D stack. This is because the guiding equation used calculates the salience based on the average colour difference between the patches.

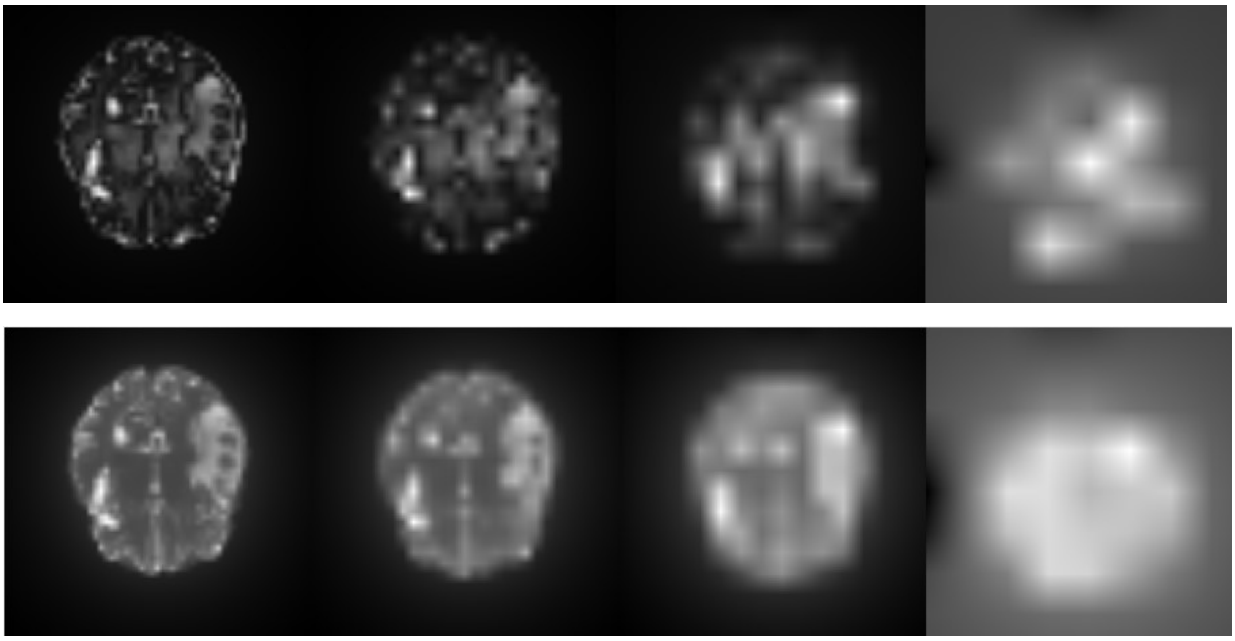
The background in general is not part of what needs to be analysed and therefore it is a problem; you have more complications if the object is much smaller than the background because the salience is calculated with respect to all the pixels of the image including the background ones. In these cases the whole object is considered salient but for this purpose it is not correct. In cases where the object is instead of a size comparable with the background, the phenomenon will still be present but will have a slight impact on the saliency detection.

To remedy this problem, all the background pixels, in each of the three sequences, are replaced by the average value of gray image intensity excluding the background pixels. Acting in this way, the salience detection algorithm effectively ignores such distorted situations.

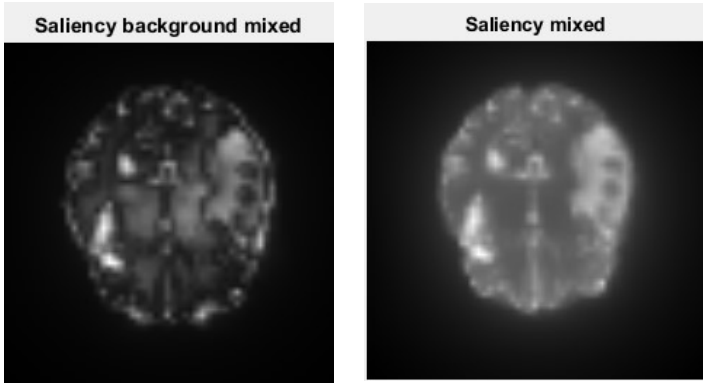
The starting sequences from which the 4 saliency maps are constructed will be similar to these:



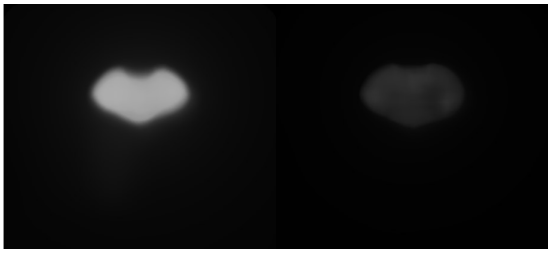
The modified saliency maps (up) compared with the previous saliency maps (down):



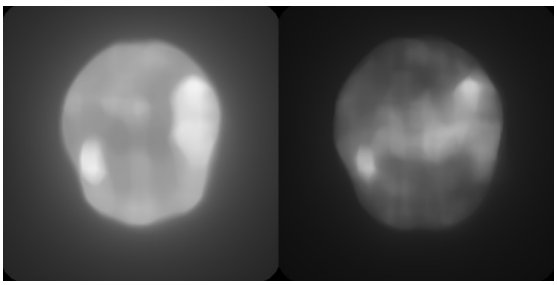
The saliency map with background modification (left) and the previous one (right):



The following figure shows the twentieth slice of the real case discussed (case 3); the section on the right shows the modified map:



On the other hand, analysing one of the central sequence slices (70) there are no obvious differences between the highlighted areas. The same salient regions are highlighted in both images but in the left one there is a globally brighter image due to the contrast between object and background.



It can be seen that the tumour area previously visible actually has significantly reduced dimensions. This background modification process drastically reduces the false positive rating.

6.Main 3D object detection

According to the object-based attention theory we want to identify different objects within an image. The importance value is subsequently assigned to each discovered object and it's called salience score. The object wanted is the one with the highest salience score. The evaluation metric for assigning the salience score is the intensity of gray in the saliency map.

The function used to delimit the three-dimensional objects is superpixels3; it calculates the 3D superpixels of the selected image in as many objects as they are specified in the express parameter. The function returns a matrix of 3D labels and the actual number of superpixels found. The superpixels will be managed as a minimum processing unit and each of these will be coloured with the average intensity value of the entire occupied area. The superpixel with maximum salience score is then searched; first the maximum intensity value is found and then a new volume is created containing only the superpixel with that intensity value.

The algorithm used by the superpixel function is called SLIC (Simple Linear Iterative Clustering). In this case I used a version of this algorithm with an automatic manage of compactness parameter (defined later) called SLICO.

The algorithm creates as many clusters (they cannot be overlapped) as the required superpixels and modifies their form iteratively. For each iteration, the distance from each cluster is calculated for each pixel. At the end of the calculation of the distance from all the clusters it will be added to the nearest cluster modifying its shape.

The distance from the cluster is calculated as a parameterized sum of the spatial distance and the intensity distance. The spatial distance will be calculated from a single pixel to the centre of all clusters. The spatial distance will be calculated for each pixel from the spatial centres of all clusters. X_j,Y_j,Z_j are the spatial coordinate of each cluster centres. Clusters that are represented by their mean intensity value and the centre cluster location.

$$d_{\text{int ensity}} = \sqrt{(l_i - l_j)^2}$$

$$d_{spatial} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2}$$

$$D_s = d_{lab} + \frac{m}{S} d_{xy}$$

"m" is the term of compactness. By decreasing this parameter, more irregular shapes will be composed and the clusters would better represent the object to be segmented.

The parameter required by superpixel function are:

N: number of image pixels taken from the image.

K: amount of superpixels provided by parameter.

m: compactness

Derived parameters:

N/K: Average area of superpixels

$S=\sqrt{N/K}$: Distance between cluster centres

The algorithm starts from the subdivision of the starting image into N equal cubes having the size N / K. At each iteration all the pixels are added to the nearest cluster. The algorithm ends when convergence is reached and every pixel is already present in the cluster that is most suitable for it or if the maximum number of iterations has been reached.

Algorithm:

1. Initialize cluster centres that are represented by their mean intensity value and the centre location

2. All the pixels are associated with the nearest cluster. The admissible area is a square 2Sx2S around the cluster centre.

3. Update cluster information

4. Return to point 2 while there is a condition

The figure shows a section of the superpixels generated during this procedure and the object that has the maximum intensity value found:

The image represented is a horizontal cut of the segmented volume. The initial volume has been divided into K objects. The K objects constructed will have a gray intensity equal to the average of the pixels that they represent. The object with higher intensity will be the one composed by pixels with higher average intensity.

The object that have higher intensity (image with light blue background) will be subsequently segmented to extract the final tumour area.

7. Automated segmentation

The last step of the proposed algorithm involves the automated segmentation of tumour region. Any segmentation method can be used. In this project I choose the Chan-Vese algorithm; it is implemented by the activecountor function as a default segmentation method.

Segmentation is the process used for splitting a digital image into multiple segments (set of pixels). The Chan-Vese model for active contours is a powerful and flexible method that is able to segment many types of images. This model is widely used in the medical imaging industry, mainly for brain segmentation.

The activecontour function will have to be informed about which is the initial image and which is the object to be segmented; it then reconstructs the object using the initial image. Note that a good detection of objects with superpixel 3 requires few segmentation iterations to achieve an acceptable result.

The algorithm aims to minimize the following function.

$$F(c_1,c_2,c)=\iiint_{\Omega 1=\omega} (u_0(x,y,z)-c_1)^2 dxdydz+\iiint_{\Omega 2=\Omega-\omega} (u_0(x,y,z)-c_2)^2 dxdydz$$

- u_0 = initial image
- ω = initial map
- C = $\delta\omega$ *(contorno di ω)*
- C_i =mean(u_0) in Ω_i

The first term is the force to shrink the contour.

The second term is the force to expand the contour.

The first term calculates the sum of the intensity differences between the individual pixels belonging to the object and the average intensity of the pixels in that volume region. The second term instead contains the sum of the differences between the individual pixels belonging to the background and the average intensity of the volume region itself.

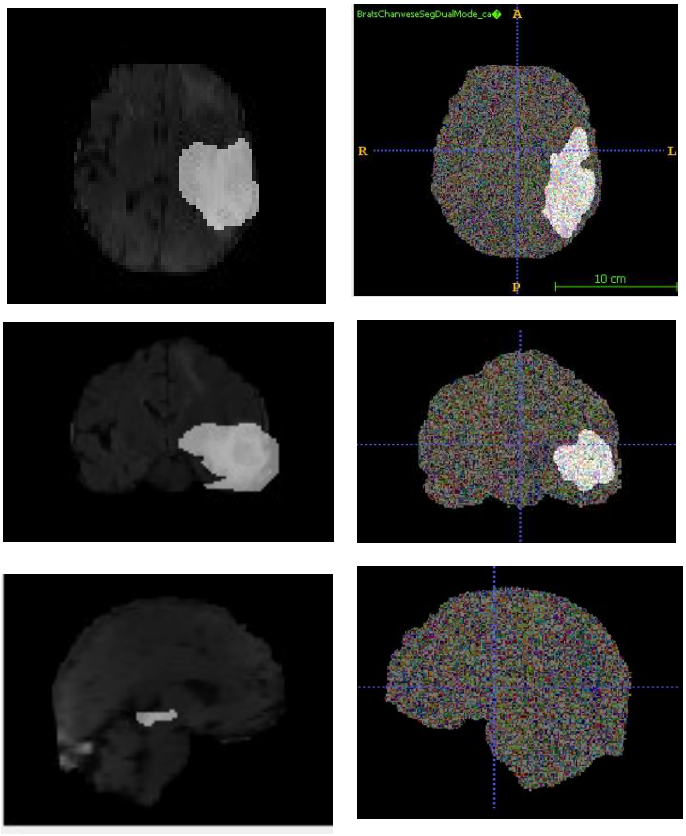
We call the first integral F_1 and the second F_2.

In the case where F_1 is greater than zero, a decrease in the region delimiting the object will be necessary. This narrowing of the region will result in a decrease of F_1 in the next iteration. Similarly a value different from zero of F_2 will involve an enlargement of the region delimiting the object.

The theoretical goal is to make F equal to zero and this would mean that the object has been cut perfectly. This perfect goal can only be reached in theory; In practice, if we wait for the function to be identical to zero, we would never reach the goal.

For this reason the minimum of F will be calculated using the gradient descent method until convergence is reached. Convergence is guided by a maximum of iterations that can be carried out or by an acceptable defined value of the function F in which the purpose will be considered achieved.

On the left is visible the segmentation realized in this project while on the right the real tumour area.



8. Conclusions and further developments

Choice number of SLICO superpixels

An incorrect choice of the superpixels number would compromise the functioning of the algorithm. In this case, too many superpixels would generate many small objects grouped by intensity. The selection of the higher intensity object may not correspond with the most relevant object; this would mean selecting a very salient object but not necessarily it will be the one sought since the algorithm should select the bigger salient object. In fact the algorithm does not consider the object size that it is selecting but only it's salience; the object size depends on the ratio between the total number image pixels and the number of superpixels and for this reason a good choice of these parameters may give better results.

A miserable number of superpixels, on the other hand, would create objects that are too large. This means that the object with maximum intensity covers a too high volume and this would lead to a disproportionate false positive index.

Rationally I think that the number of superpixels should be adjusted according to the size of the brain we are analysing. For small brains (or images much smaller than 256x256) a low number of superpixels would not be enough to discriminate a tumour, because I'm sure that it will have contained dimensions compared to a 256x256 pixel image.

Choice of the maximum number regarding the Chan-Vese iterations

The algorithm in question takes the image already segmented by the previous algorithm as input, so in general it will not require too many iterations. I estimated a maximum number of 20 iterations for a good approximation.

A number too high of iterations would require a high computational time and would lead to an exaggeratedly high false positive index.

Alternatively, a small number of iterations would not lead to significant variations with respect to the object segmented by the previous algorithm.

In conclusion, I do not believe that the last segmentation procedure is necessary since the segmentation operations are already performed in the previous step, applying a different policy, but obtaining results that are already reliable. Having two segmentation processes excessively slows execution times; simultaneously they require the management of two project parameters such as the number of superpixels and the number of iterations after which the Chan-Vese algorithm is stopped. Acting solely on the first parameter, very similar and simpler results could be obtained by having to select only the number of superpixels to be generated.