Brain Tumour Detection Using Dempster Shafer Combination

Agustin CARTAYA
University of Burgundy – ImViA Lab
Internship Report, September 2023

Introduction

Segmentation of brain tumors from magnetic resonance imaging (MRI) is a challenging and essential task with various applications in the field of medical analysis. Each brain imaging modality offers distinct and fundamental information regarding different aspects of the tumor. Therefore, in this project, the proposed method is based on integrating evidence (masses) [1] obtained from these modalities (sources) using the Dempster-Shafer theory [2] to achieve a 3D segmentation. Figure 1 illustrates the key steps of this combination process.

Sources:

The sources are 3D images of patients with glioma from the UCSF-PDGM database [3]. Each source comes from an MRI modality. The modalities used in this work were T1, T2, FLAIR and ADC.

The masses:

Masses are 3D images where each voxel varies from 0 to 1 indicating the belief that a voxel may (1) or may not be (0) part of a tumor. These are obtained from the sources.

The combination:

Once the masses are obtained, they are combined using Dempster's rule of combination to obtain a resulting image with greater certainty that a voxel belongs or not to a tumor.

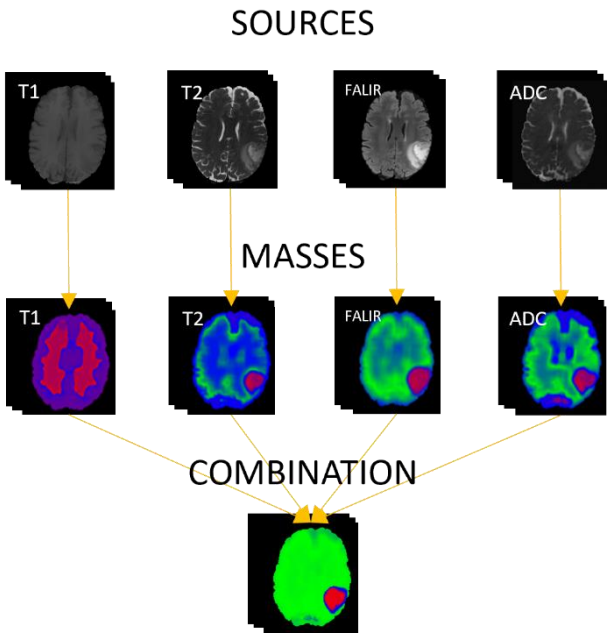

Figure 1: Main steps to get the combined image. Sources: the 4 original MRI modalities. Masses: the mass images of each modality. green voxels (high probability of belonging to a healthy area), blue (areas with uncertainty) and red (high probability of belonging to a tumor). Combination: Image obtained from the combination of mass images; the color of the voxels is the same as for mass images. Patient shown: 4 Layer 109.

Proposed method

Step 0 – Obtaining texture images:

To add more information to the combination of normal images (intensity images from MR), texture images were created. These were obtained as follows [4]:

$$img\_texture[i,j,k] = img[i,j,k] - img[i + desp_y, j + desp_x, k]$$

where, img is one of the sources. i, j and k are the position of the voxel in the texture image. $desp_y$ and $desp_x$ are the necessary shift to reach a neighboring voxel. These are calculated from a direction (°) and a distance. In this work, a direction of 0° and a distance of 3 were used.

Figure 2: Texture images of each MRI modality. Patient shown: 4 Layer 109.

Step 1 – Obtaining distance images:

Distance images are images in which each voxel represents the distance between two normalized histograms [4]. The first histogram is constant and was obtained from a healthy area (gt healthy histogram) or with tumor (gt tumor histogram). The second was obtained by centering a 2D or 3D window of size N x M (N x M x K in the case of 3D windows) in each voxel of the source and calculating the histogram with the values within that window.

In the following steps the size of the window used was 15x15 and the number of bins was $\sqrt{15\ x\ 15} = 15$ for 2D and 15x15x15 and $\sqrt{15\ x\ 15\ x\ 15} = 58$, for 3D.

The range of each histogram varied with respect to each source image, as shown in Table 1. These ranges were obtained from the standard deviation of the voxel value of the first 30 patients in the dataset:

Modality	Normal range	Texture range
T1	0 – 2000	-1000 – 1000
T2	0 – 2400	-1200 – 1200
FLAIR	0 – 3200	-1600 – 1600
ADC	0 – 0.004	-0.002 – 0.002

Table 1: Range of groundtruth histograms for each modality and image type

Step 1.1 – Obtaining gt histograms:

To obtain the gt histograms, the segmentation obtained from the dataset was used. Only the areas of the growing tumor (1) and the area of the necrotic tumor (2) were used for the gt tumor histogram, since the edema areas (4) had a lot of similarity with the healthy area of the brain in many patients. 3 types of gt histograms were obtained to compare the results:

Gt histogram Patient 4 layer 109:

These gt histograms were obtained only from layer 109 of patient 4 of the dataset with a 15x15 window. For the gt tumor histogram, the window was focused on the centroid of the largest region between the junction of zones 1 and 2 of the tumor. For the gt healthy histogram the window was centered horizontally aligned with the previous one. Figure 3 shows the position of the windows and histograms.

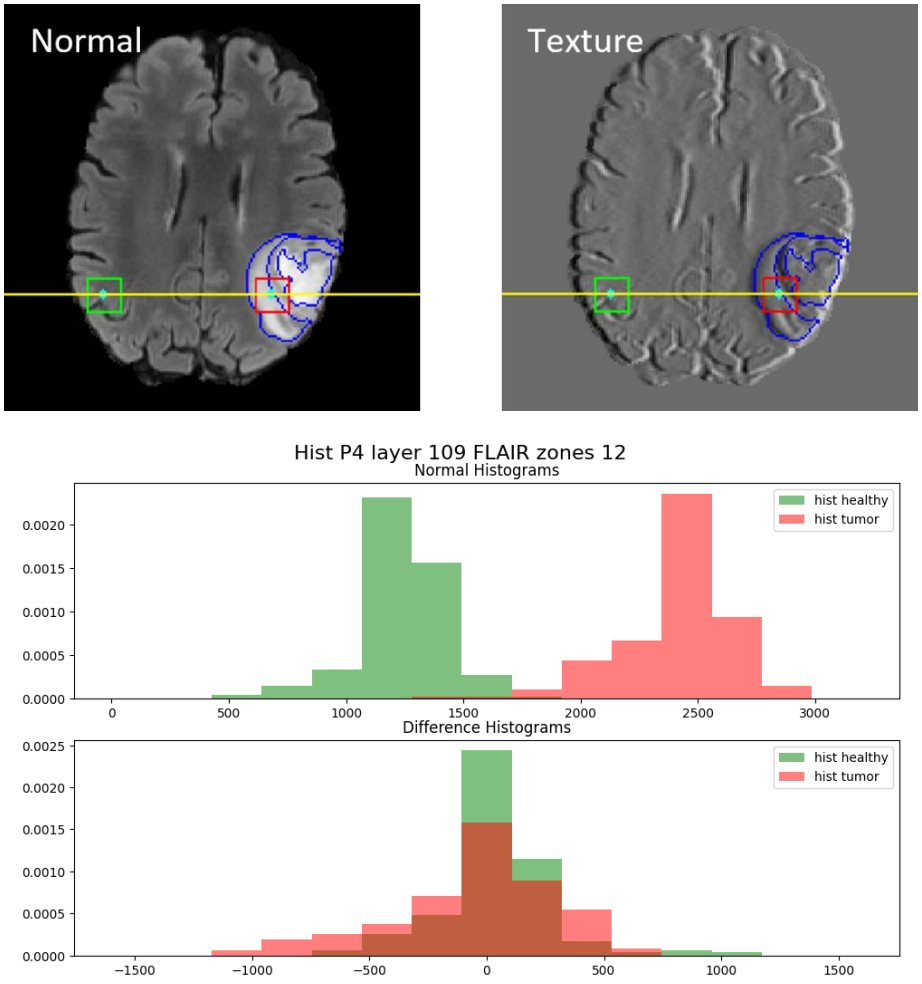

Figure 3: Gt histogram patient 4 layer 109 for FLAIR source. At the top it can be seen the windows where the gt histograms were taken, the green windows correspond to the gt healthy histogram and the red ones to the gt tumor histogram. At the bottom it can be seen the histograms obtained for the normal image (top) and for the texture image (bottom).

Gt histogram patient 4:

These gt histograms were obtained from all the voxels of patient 4 (excluding those from the background). For the gt tumor histogram, zones 1 and 2 of the tumor were used in all layers. For the gt healthy histogram, all healthy areas of the brain were used. Figure 4 shows the histograms.

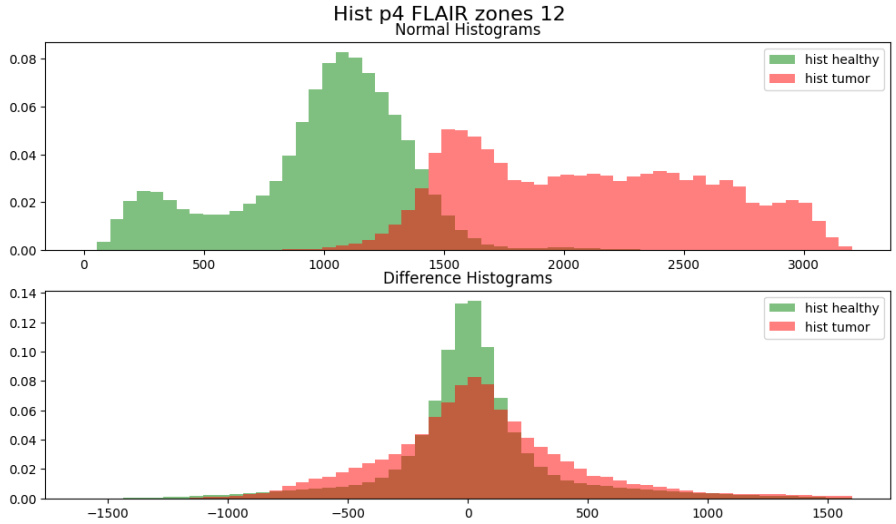

Figure 4: Gt histogram patient 4 for FLAIR source. Histograms obtained for the normal image (top) and for the texture image (bottom).

Gt histogram global:

These gt histograms were obtained from all voxels (excluding those in the background) of patients 4 to 35 of the dataset. For the gt tumor histogram, zones 1 and 2 of the tumor were used in all layers of each patient. For the gt healthy histogram all healthy areas of the brain of all patients were used. Figure 5 shows the histograms.

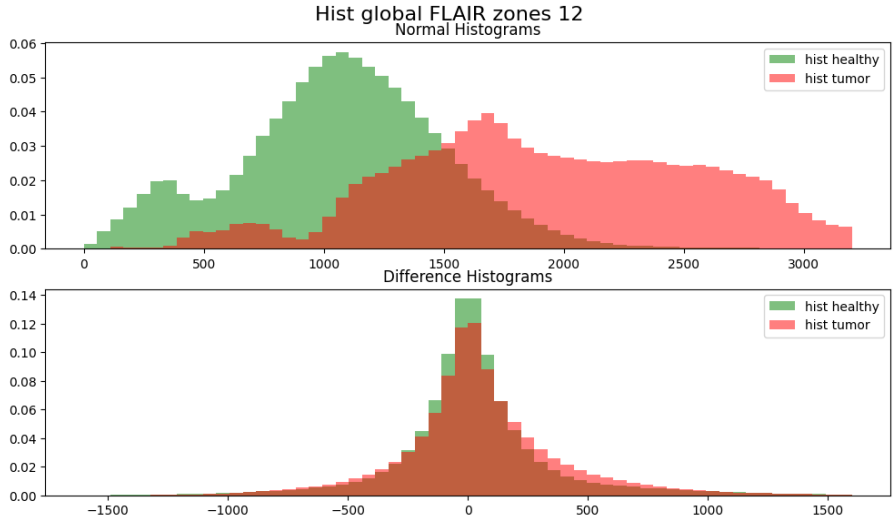

Figure 5: Gt histogram global for FLAIR source. Histograms obtained for the normal image (top) and for the texture image (bottom).

Step 1.2 – Obtaining histogram of each voxel:

To obtain the histogram of each voxel, a 2D or 3D window centered on the voxel was used. All these histograms were stored in a 3D image where each voxel contains its histogram (image histogram).

Step 1.3 – Obtaining distance image:

For each source 2 distance images were created, one between each element of the image histogram and the gt tumor histogram (image distance tumor) and another between each element of the image histogram and the gt healthy histogram (image distance healthy).

To calculate the distance between the histograms, different distance functions proposed in [5] were tested. It was decided to use Euclidean distance of Cumulative Spectrum (ECS) since it provided the greatest distance between the gt tumor histogram and the gt healthy histogram, and fulfilled all the properties of a distance function:

- Reflexive: $d(x,x) = 0$
- Nonnegative: $d(x,y) \geq 0$
- Symmetric: $d(x,y) = d(y,x)$
- Identity of indiscernibles: $d(x,y) = 0 \Leftrightarrow x = y$
- Triangular inequality: $d(x,y) \leq d(x,z) + d(z,y)$

Table 2 shows the results of different distances between the gt tumor histogram and the gt healthy histogram obtained from patient 4 layer 109, where the largest distance is that obtained using ECS. In this table it can also be seen that the distance functions meet the following properties. Reflexive and Identity of indiscernibles: column 2; Symmetric: column 3 and 4.

Distance	nor HvsH	nor HvsT	nor TvsH	dif HvsT
ECS	0	2.126382	2.126382	0.245009
Hellinger	0	0.967237	0.967237	0.267993
Combined EMD	0	0.181302	0.174253	0.328317
Earth mover	0	0.017778	0.017778	0.031407

Table 2: Comparison of distance functions between histograms. The histograms used were those from gt histogram patient 4 layer 109. Column 1: distance used; column 2: distance between gt healthy histogram and itself obtained from the normal image; Column 3: distance between gt healthy histogram and gt tumor histogram obtained from normal image; column 4: distance between gt tumor histogram and gt healthy histogram obtained from normal image; Column 5: Distance between gt healthy histogram and gt tumor histogram obtained from the texture image.

In total for each patient, 48 distance images were obtained, 16 for each type of gt histogram (gt histogram Patient 4 layer 109, gt histogram Patient 4, gt histogram Global), in which 4 for each source, 2 with respect to the gt healthy histogram and 2 with respect to the gt tumor histogram. Figure 6 shows the distance images of the FLAIR source for each type of gt histogram.

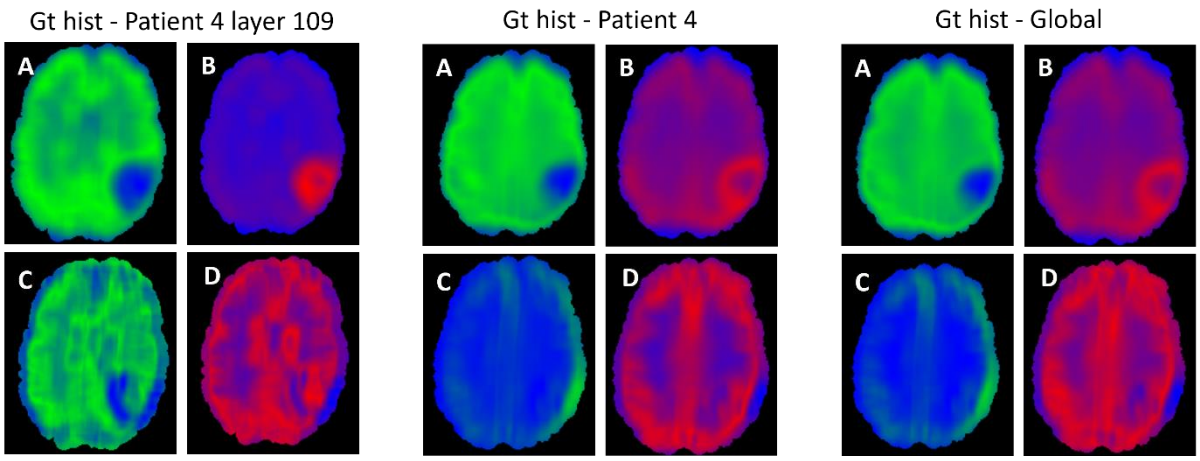

Figure 6: Distance images using different gt histograms for the FLAIR source. On the left it can be seen the distance images obtained using the gt histogram patient 4 layer 109. In the center those obtained using gt histogram patient 4 and on the right, those obtained using gt global histogram. A: distance image from normal source with respect to gt healthy histogram, B: distance image from normal source with respect to gt tumor histogram. C: distance image from texture source with respect to gt healthy histogram, D: distance image from texture source with respect to gt tumor histogram. For images A and C, the voxels closest to the gt healthy histogram are green and the farthest blue. For images B and D, the voxels closest to the gt tumor histogram are red and the farthest blue. Patient shown: 4 Layer 109.

Step 2 – Obtaining the angle images:

After obtaining the distance images, they were joined following [4] to obtain the angle images. Each voxel of the angle images corresponds to the angle (in radians) of the vector formed by the corresponding voxel in the image distance healthy (y axis) and the image distance tumor (x axis). Therefore, values close to 0 have a high probability of belonging to a healthy part of the brain and values close to Pi/2, to a tumor. Figure 7 shows an example.

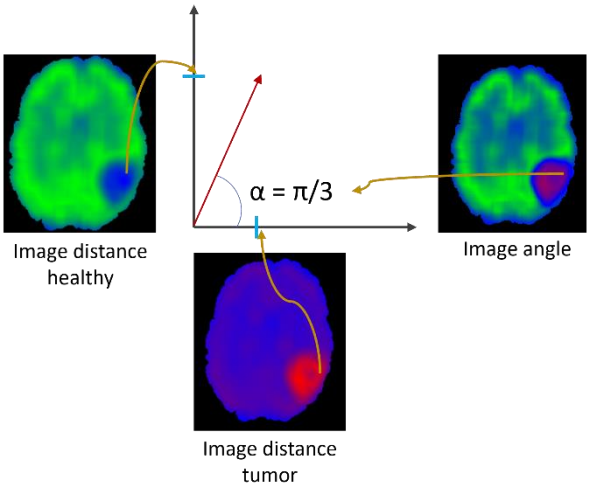

Figure 7: Graphical explanation of how to obtain an angle image.

In total, 24 angle images were obtained for each patient. 8 for each type of gt histogram, in which 2 for each source, one from the normal sources (image angle normal), and one from the texture sources (image angle difference). Figure 8 shows the angle images of the FLAIR source for each type of gt histogram of patient 4 (layer 109).

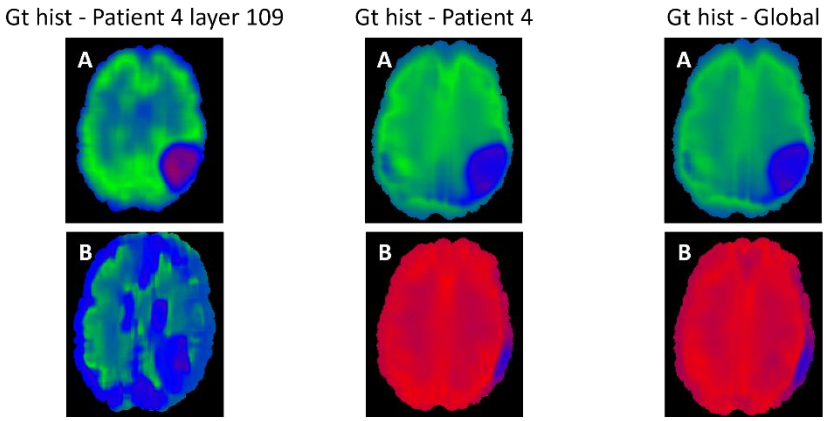

Figure 8: Angle images using different gt histograms for the FLAIR source. On the left it can be seen the angle images obtained using the gt histogram patient 4 layer 109. In the center those obtained using gt histogram patient 4 and on the right, those obtained using gt global histogram. A: Angle image of the normal source. B: Angle image of the texture source. Green voxels: high probability of belonging to a healthy area, blue: areas with uncertainty and red: high probability of belonging to a tumor. Patient shown: 4 Layer 109.

Step 3 – Selection of the best gt histogram:

After calculating the angle images for the first 19 patients (from 4 to 23), the selection of the best gt histogram was done. For this, the segmentation provided by the angle images was obtained and the global f1 score based on voxels was calculated. f1 scores were used due to class imbalance.

To achieve a segmentation with the angle images, a threshold between 0 and pi/2 was chosen and the voxels greater than or equal to that threshold were taken as areas with tumor. The choice of threshold was made in 2 ways:

- Individual thresholds: A unique threshold was calculated for each source within each patient based on the segmentation provided by the dataset, aiming to maximize the F1 score. Subsequently, a global F1 score was calculated for each source by averaging across all the patients.

- Global threshold: a global threshold for each source was chosen using all patients. For this, a 5-fold cross validation was used, in each fold the threshold that maximized the f1 score of the train set was obtained and saved together with the f1 score of the validation set. Then the average of the f1 score of the validation set obtained in each set was taken.

Figure 9 shows the overall f1 score for each source. Both methods for calculating threshold are shown. In both cases the gt histogram Patient 4 layer 109 obtained a better balance between sources. This balance is expected to enhance the performance of the final combination. Consequently, the angle images to be used in future steps will be those that were obtained with the gt histogram patient 4 layer 109. It's worth noting that texture sources were not considered in this selection process, as the difference in histogram distances compared to normal images was negligible.

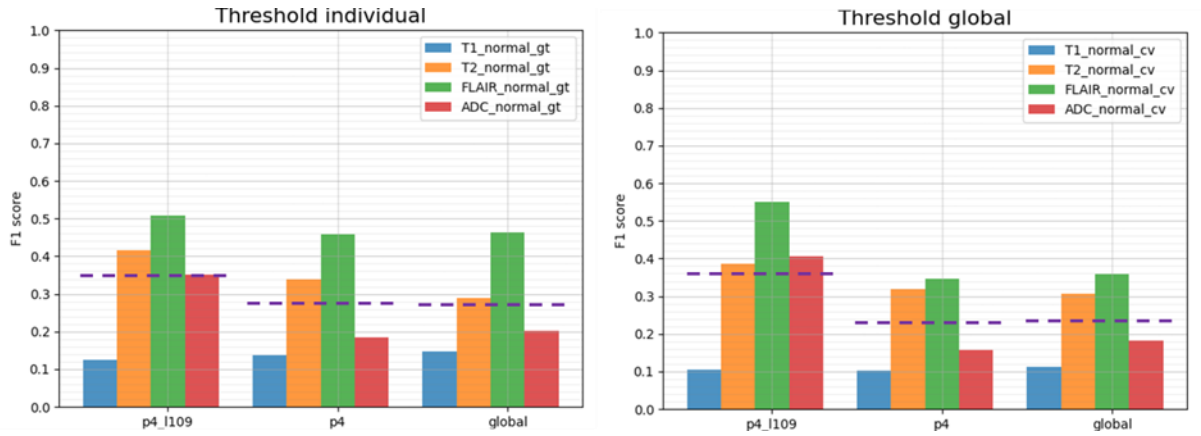

Figure 9: F1 score for each source of the angle images using different gt histograms. On the left it can be seen the results calculated from the individual threshold and on the right from the global threshold. The dotted purple lines indicate the average of the 4 modalities.

Step 4 – Obtaining mass images:

After obtaining the angle images, they must be converted into mass images to combine them. Since segmentation is binary, each voxel must have a degree of belonging (also known as mass) to one of the 2 classes ("healthy" or "tumor").

In this work, a threshold and a confidence level (weight) were sought for each source in order to convert the angle images into mass images.

Threshold

Establishing a threshold in the angle image helped to obtain the certainty of each voxel of belonging to one class or another. In this way, voxels very far from the threshold will have greater certainty of belonging to one of the classes and voxels close to the threshold will have greater uncertainty.

Confidence

Assigning greater weight to the more reliable angle images helped to decrease uncertainty in these and to influence much more in the combination. On the other hand, assigning less weight to less reliable angle images helped to increase uncertainty and influence less in the combination.

In a first attempt it was proposed to find a global threshold for each source and base its confidence on the f1 score obtained. The drawback is that for each patient the best threshold and f1 score of each source may vary. As can be seen in Figure 10, the highest overall confidence is in the FLAIR source but for certain patients this is not the predominant source, so it would affect the combination. Therefore, in the following steps, a method to find an individual threshold and weight for each source for each patient is proposed.

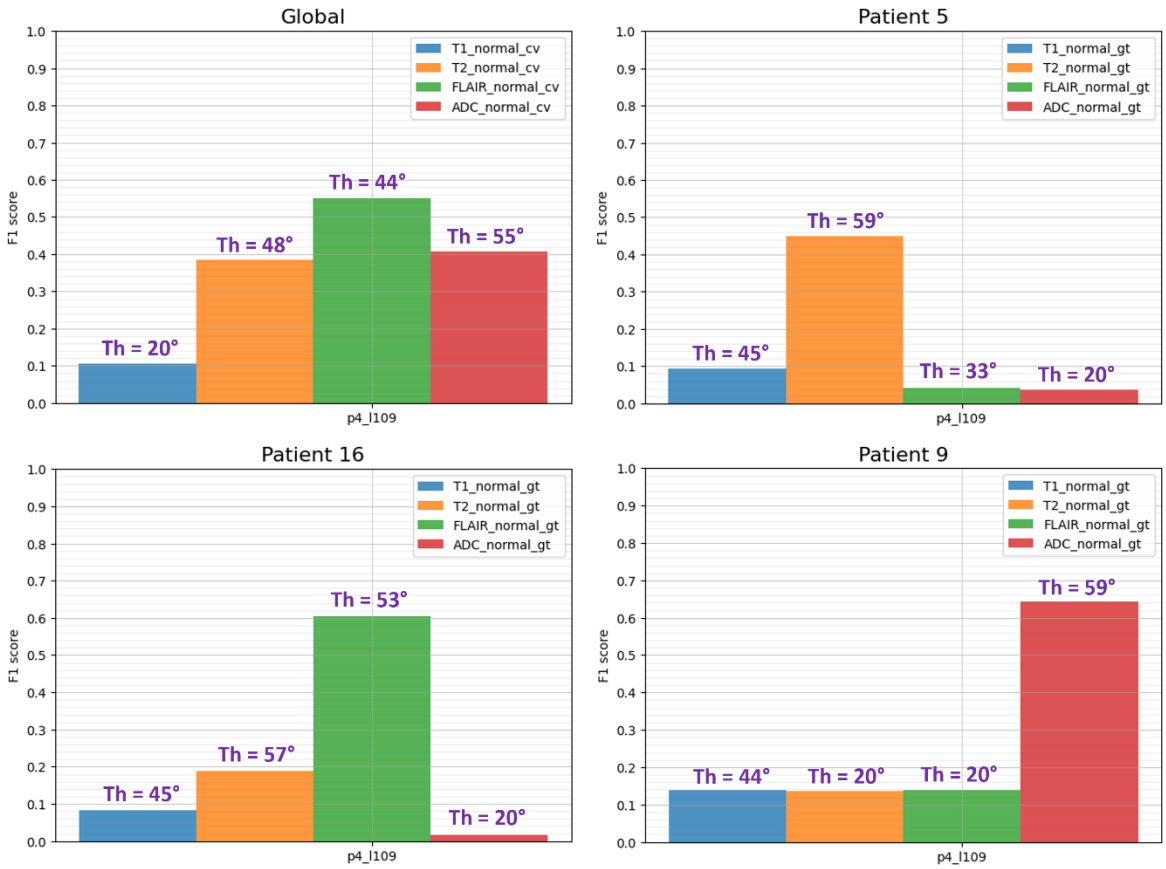

Figure 10: F1 score for each source using different thresholds and different patients.

Step 4.1 – Obtaining weight for each source

Basing the confidence of each source on its f1 score would be the best alternative, However, without the real segmentation of the tumor, this metric cannot be obtained. To solve this problem, a method based on the normalized histogram of the source is proposed in this project in which it is sought to obtain the uncertainty of the source and the certainty to predict voxels belonging to healthy areas or tumors:

Uncertainty:

The voxels of the angle image contain values ranging from 0 to pi/2. If the certainty of the image is 100% then there should be no doubtful values (close to pi/4). Therefore, to obtain uncertainty, the sum of histogram values under a Gaussian curve centered on pi/4 was obtained.

Certainty for voxels belonging to healthy areas and tumors:

To obtain these two values, the histogram was divided into 2 equal parts and their values were normalized. To obtain the certainty of the voxels belonging to healthy areas, the first part of the histogram was taken, and the values under a Gaussian curve centered on 0 were added. Then, to obtain the certainty of the voxels belonging to tumors, the second part of the histogram was taken, and the values under a Gaussian curve centered on pi / 2 were added.

Figure 11 graphically shows how the weights were obtained for two types of sources from the same patient.

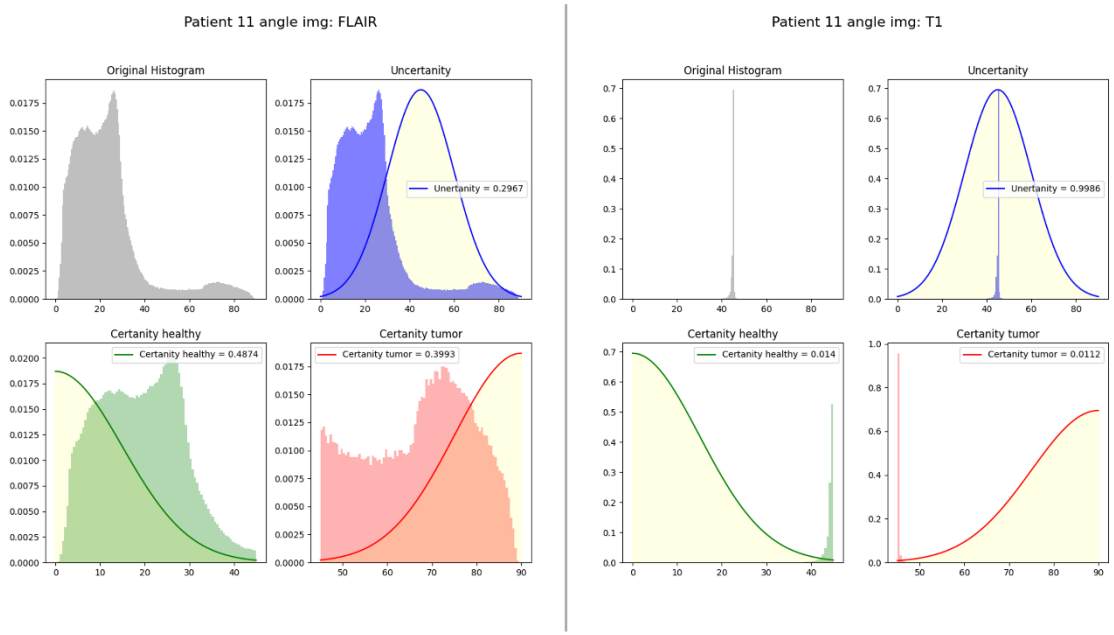

Figure 11: Graphic explanation for obtaining weights: Left: FLAIR source. Right: T1 source. In each set of 4 images. Top left corner: histogram of the source. Top right corner: Values used to obtain uncertainty. Lower left corner: values used to obtain the certainty of voxels in healthy areas. Lower right corner: values used to obtain the certainty of voxels in tumor. Patient shown: 11.

Step 4.2 – Obtaining threshold for each source:

When deciding the best gt histograms, an individual threshold was used for each modality of each patient, the problem is that the threshold was also chosen based on the actual segmentation of the dataset. To solve this problem, a method based on the characteristics of the histograms of the angle images was proposed for the choice of the best threshold.

After analysing the best thresholds for different angle images in various patients, it was noted that:

Almost all are in a valley or area where the height difference (slope) of the histogram is less than a value m_{max}

Once all the possible candidates were obtained, they were filtered based on several criteria.

1. The size of a tumor must be within a range $[vol_{min}, vol_{max}]$ of the total volume of the brain.
2. In a valley there can only be one threshold and it is the lowest among them.
3. The threshold must be within a range $[th_{min}, th_{max}]$
4. When the distance between 2 thresholds is less than a $dist_{min}$ value, its mean is calculated.

Once these filters were applied, if there were still multiple thresholds, the one with the lowest height between them was taken.

Sometimes no threshold may be found, so the search for these was done in groups of multiples sources from the same patient as follows:

1. Find a threshold for each image of the previous way.
2. If there are one or more images whose thresholds were not found (image without th), the uncertainty of all the images is calculated, and an image with the most similar uncertainty in which a threshold could have been found (reference image) is sought.
3. From the reference image the percentage of brain occupied by the tumor is obtained
4. Then, a threshold is sought for the image without th so that it allows segmenting the tumor with a percentage of brain similar to that of the reference image.

In image 12 it can be seen the thresholds found by the algorithm and those found with the actual segmentation of the tumor. It can be noted that when the f1 score of the source is not so low, both thresholds tend to be very close.

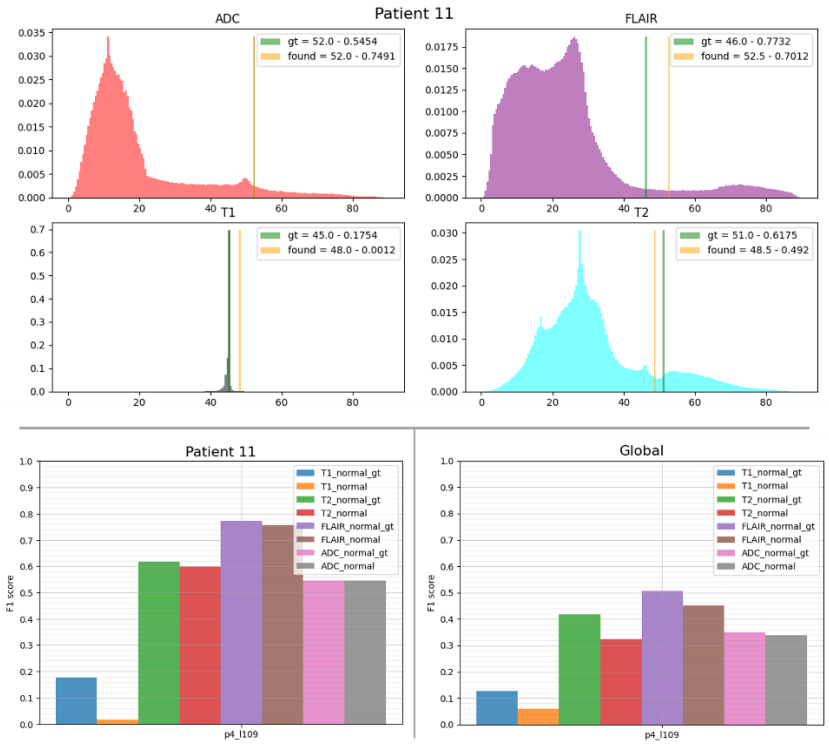

Figure 12: Thresholds difference: in the top image, it is seen the threshold obtained (yellow) and the best one (green) for each source. In the bottom, it is seen the f1 score using the threshold obtained and the best threshold of each source (suffix _gt at the end) for patient 11 (left) and the mean for all the patients (right).

Step 4.3 – Creating the mass functions:

In order to convert an angle image to a mass image, a mass function was required. In this work, 2 mass functions were proposed that meet the following characteristics:

- They receive values in a range $[0, \frac{\pi}{2}]$ and bring it to the range (0, 1).
- Values near the threshold at the input are close to 0.5 (uncertainty) at the output.
- If the weight of the source is equal to 0 (untrusted source), no matter what the value of the input, the value of the output will be very close to 0.5 at the output.
- They can assign importance to the weight of the source.

The fact that the output values do not contain the extremes of the interval was done to avoid problems caused by conflicts when combining masses using Dempster's rule of combination.

Sigmoid mass function:

$$\text{mass(x)} = \frac{1}{1 + \exp\big(-\exp(\text{h}) \cdot \pi \cdot \text{w} \cdot ((x\ -\ th) \pm \text{desp})\big)}$$

where:

- h: importance of source weight
- w: source weight
- th: threshold
- desp: source certainty for voxels in healthy areas (-desp) or with tumor (+desp)

Linear mass function:

$$mass(x) = \begin{cases} x - th \geq 0, & 0.5 + \min\left(0.49, \tan\left(\frac{\pi}{2} * w_{pos} * \exp(h)\right) * \frac{x - th}{\pi - 2 * th}\right. \\ x - th < 0, & 0.5 + \min\left(-0.49, \tan\left(\frac{\pi}{2} * w_{neg} * \exp(h)\right) * \frac{x - th}{2 * th}\right. \end{cases}$$

Where:

- h: importance of source weight
- w_{pos}, w_{neg}: source certainty for voxels in healthy or tumor areas respectively.
- th: threshold

The conversion of angle images to mass images for different patients can be seen in Figure 13.

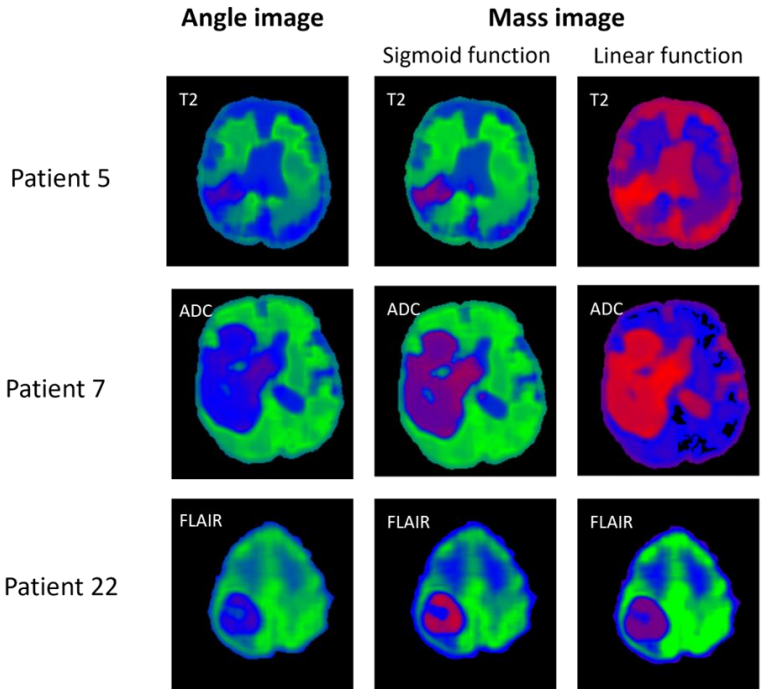

Figure 13: Conversion of angle images to mass images for different patients using the two proposed mass functions.

Step 5 – combination of mass images:

Once the mass images were obtained, they were combined using Dempster's rule of combination. For each patient, 8 mass images were combined and a single image (combined image) was obtained in which each voxel represents the belief of belonging to a healthy area (close to 0) or a tumor (close to 1). Figure 14 shows all the mass images used to obtain the combined image of a patient. The mass images were obtained using the sigmoid function.

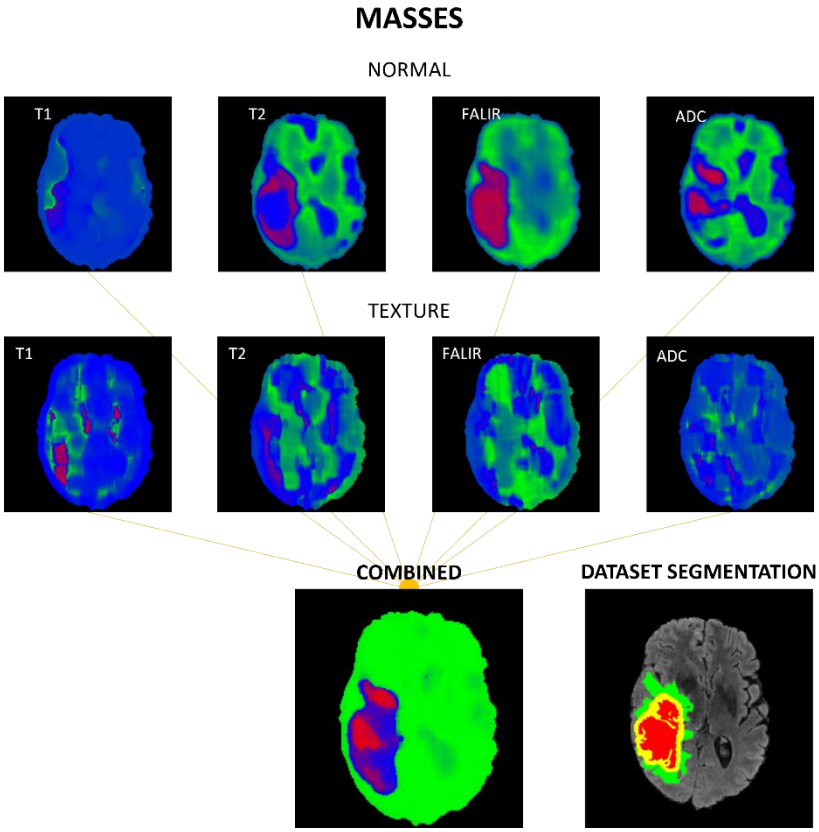

Figure 13: Images to combine. On the top it can be seen the mass images used to obtain the combined image in the bottom. On the bottom right it can be seen the actual segmentation provided by the dataset. Patient shown: 8 layer 90

Step 5.1 – Choosing the best mass function:

In order to choose the best mass function for the combination, the f1 score of the combined images was calculated. In figure 14 it can be seen the overall f1 score for the results obtained by each mass function, and there is more than a 15% difference between both results in favour of the sigmoid function. Therefore, the mass function to be used in future steps will be the sigmoid.

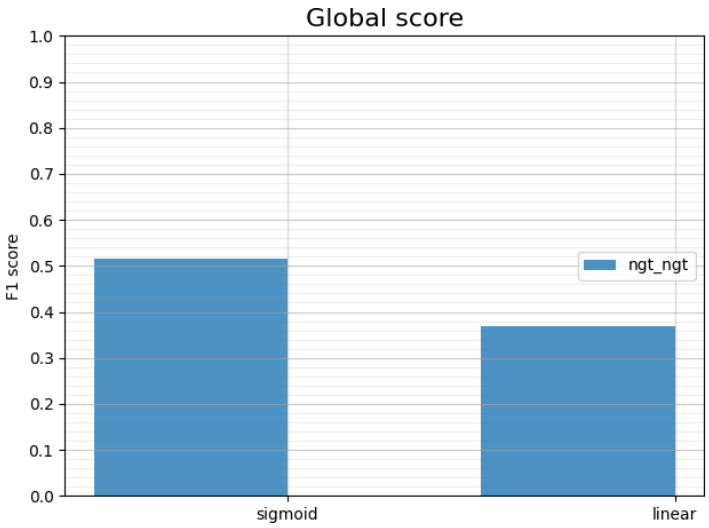

Figure 14: F1 score, of each mass function, obtained from all patients.

Step 5.2 – Compare with other combination methods:

In order to compare the Dempster's Rule for combination, other combination methods were proposed:

Mean combination: Calculates the mean between the masses and obtains a combined image.

Min distance combination: The combined mass is equal to that mass with less uncertainty.

Figure 14 shows the results of the 3 types of combinations, and it can be stated that the best combination is given by Dempster's rule of combination.

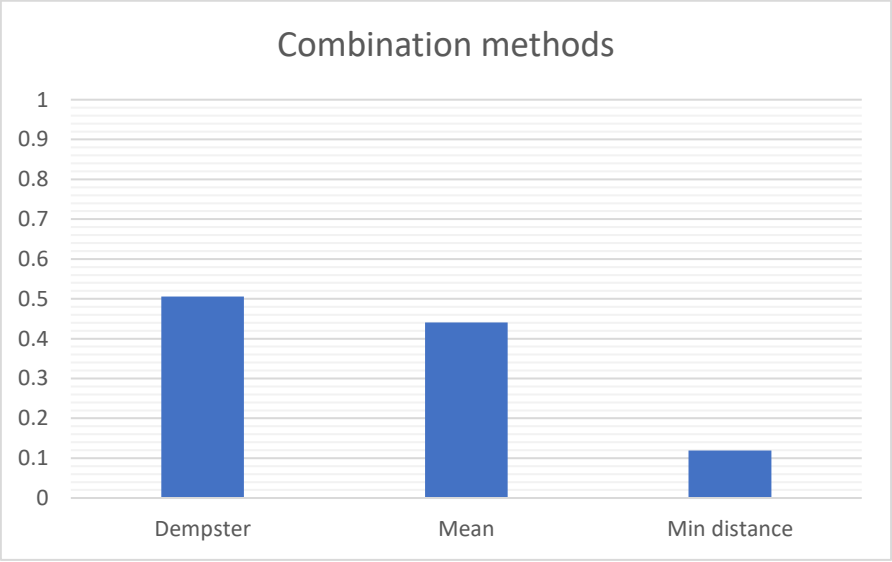

Figure 14: F1 score, for each combination method, obtained from all patients.

Results

After obtaining the combined images of the first 19 patients of the dataset, the following metrics were calculated for each mass image (normal sources) and for the combined image of each patient: Sensitivity (*sens*), Precision (*prec*), F1 score (*f1*), Balanced Accuracy (*ba*), Cross entropy (*cross_entropy*), Unceranity (*uncert*), Certainty for voxels in healthy areas (*cert_neg*), Certainty for voxels in tumors (*cert_pos*). This made it possible to compare individually and globally (mean of the individuals) the results of the proposed method.

It can be seen in the overall result that the F1 score was maximized with the combination, as well as the sensitivity, the balanced accuracy, and the certainty for voxels in healthy areas. While low uncertainty remained. See Figure 15.

In the individual results, it can be seen that the combined image for each patient most of the time is usually equal to or greater than the best source (for the metrics that are sought to maximize). See Figure 16.

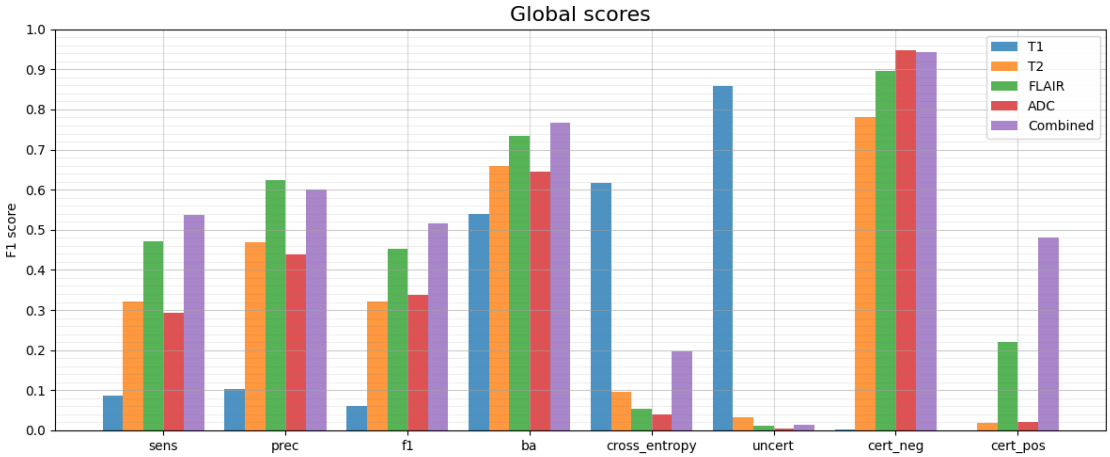

Figure 15: Global results for each source (mass image) and for the combined image. Sensitivity (sens), Precision (prec), F1 score (f1), Balanced Accuracy (ba), Cross entropy (cross_entropy), Unceranity (uncert), Certainty for voxels in healthy areas (cert_neg), Certainty for voxels in tumors (cert_pos)

Figure 15: Individual results for each source (mass image) and for the combined image of patient 4 to 23. Sensitivity (sens), Precision (prec), F1 score (f1), Balanced Accuracy (ba), Cross entropy (cross_entropy), Uncertainty (uncert), Certainty for voxels in healthy areas (cert_neg), Certainty for voxels in tumors (cert_pos).

Conclusion

Thanks to the previous results, it was possible to demonstrate that the combination helped to provide a better segmentation thanks to the information provided by the different sources.

With the help of Dempster Shafer's combination model, it was possible to reduce the conflict between sources and maintain low uncertainty.

It can also be stated that sources with little information do not affect individual patient outcomes. This was achieved thanks to the proposed mass functions and weights based on the uncertainty of each source.

Bibliography

[1] Intégration de la mesure de conflit dans la fusion de décision pour des couples (attributs, précision); Sébastien DESBOUCHAGES

[2] A mathematical theory of evidence (1976); G. Shafer

[3] Calabrese et al. The University of California San Francisco Preoperative Diffuse Glioma MRI Dataset. Radiology: Artificial Intelligence, 4(6):e220058, November 2022.

[4] Vers des attributs textures métrologiques pour les images 3D médicales multivaluées; Billaud Eliot, Et al.

[5] A Comprehensive Evaluation of Spectral Distance Functions and Metrics for Hyperspectral Image Processing; Hilda Deborah, Et al.