UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

Brain Tumor MRI Classification

Franceschin Sarah, Grisi Caterina, Pozza Giacomo, Tonello Alessio, Viberti Andrea

DEEP LEARNING APPLIED TO NEUROSCIENCE AND REHABILITATION
AA 2021/22

Index

1 – Understanding the project

- Aim
- Methods
- Results
- Inside the CNN

2 – Highlighted problems

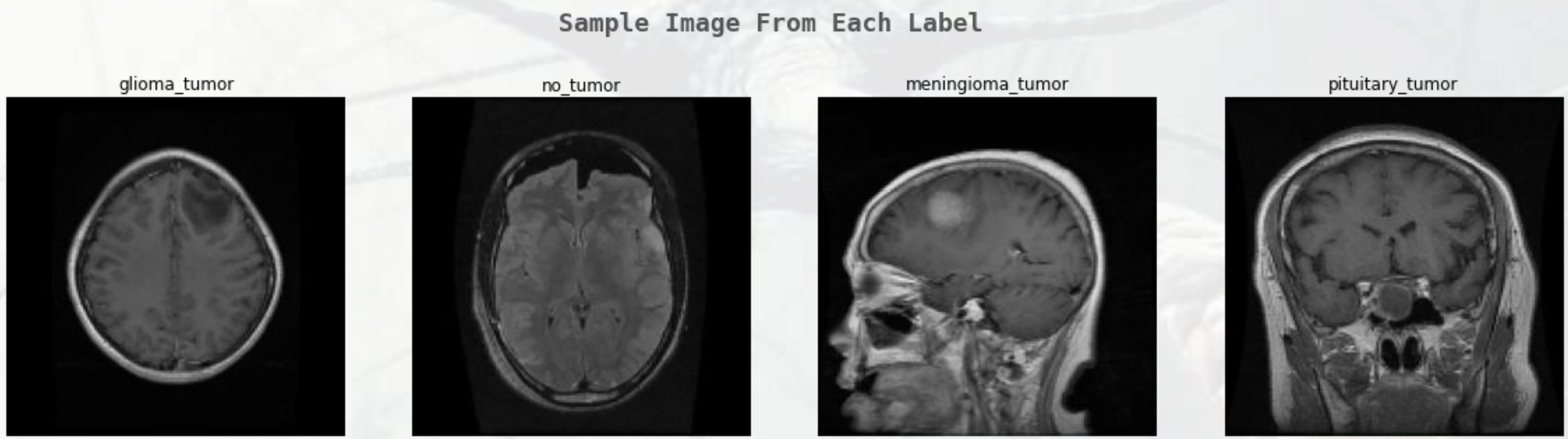
- Unbalanced data
- Why efficient net?

3 – Aim variation

- Tumor VS No Tumor

Understanding the project - AIM

Build a CNN able to classify between 4 different type of Tumor from images as input:



926 **500** **937** **901**

Understanding the project - METHODS

DATA PREPARATION:

- Rescaling of the images
- One hot encoding for labels
- Splitting into training set (90%) and test set (10%)

Transfer learning technique with EfficientNetB0:

Weights = from 'Imagenet'

Layers = Not included the last one and add few layers in order to adapt the model for the 4-classification

Global Average Pooling 2D → Dropout (rate=0.5) → Dense (activation = softmax)

 # Epochs = 12

 # Batches = 83

Validation Split = 10%

Metrics evaluated = Accuracy and Loss both of the training set and the test set

Optimizer = Adam

CV?

Understanding the project - METHODS

THE CROSS VALIDATION TECHINQUE:

PRO

Gives an overall view and comprehension about the accuracy of the model

↓

LESS SUSCETTIBLE FROM THE TYPE OF IMAGES TRAINED

CONS

Very expensive

↓

NOT DONE ORIGINALLY

↓

LET’S DO IT!

CV method: 5 fold cross validation

Understanding the project- METHODS

Other analysis performed by us:

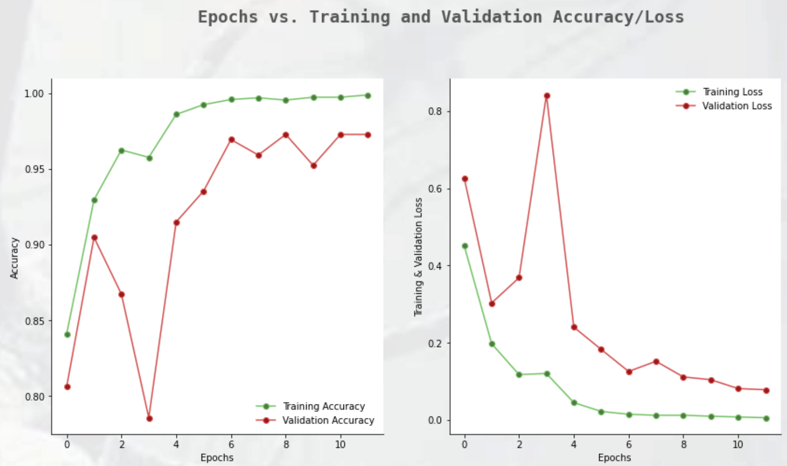
- Precision’s monitoring (Average)
- Recall’s monitoring (Average)
- Training with the SGD optimizer to see differences

Understanding the project – ORIGINAL RESULTS – TRAINING & TESTING

TRAINING

High Accuracy
- Training set = 0.9993
- Validation set = 0.9728

Low Loss
- Training set = 0.0059
- Validation set = 0.0785



TESTING

	precision	recall	f1-score	support
0	0.98	0.96	0.97	93
1	0.96	1.00	0.98	51
2	0.98	0.98	0.98	96
3	1.00	1.00	1.00	87
accuracy			0.98	327
macro avg	0.98	0.98	0.98	327
weighted avg	0.98	0.98	0.98	327

ONE VS ALL APPROACH

PRECISION

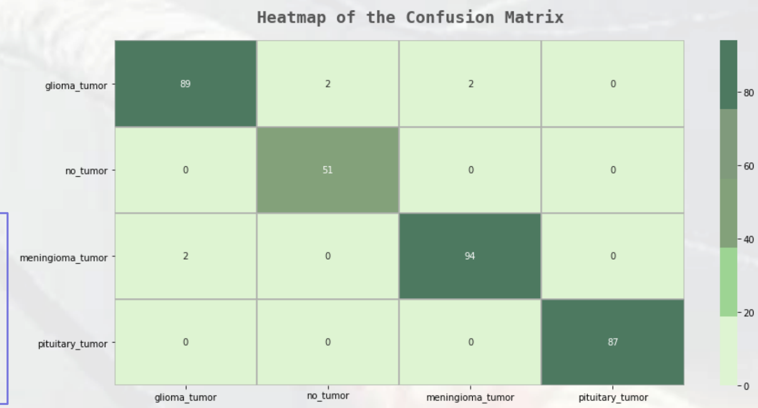
$$\frac{TP}{TP+FP}$$

RECALL

$$\frac{TP}{TP+FN}$$

F1-SCORE

$$\frac{TP}{TP+\frac{FN+FP}{2}}$$



Understanding the project – OUR RESULTS – TRAINING

High Accuracy (Mean values)

ADAM

- Training set = 0.9962
- Validation set = 0.9707

SGD

- Training set = 0.9576
- Validation set = 0.9339

Low Loss (Mean Values)

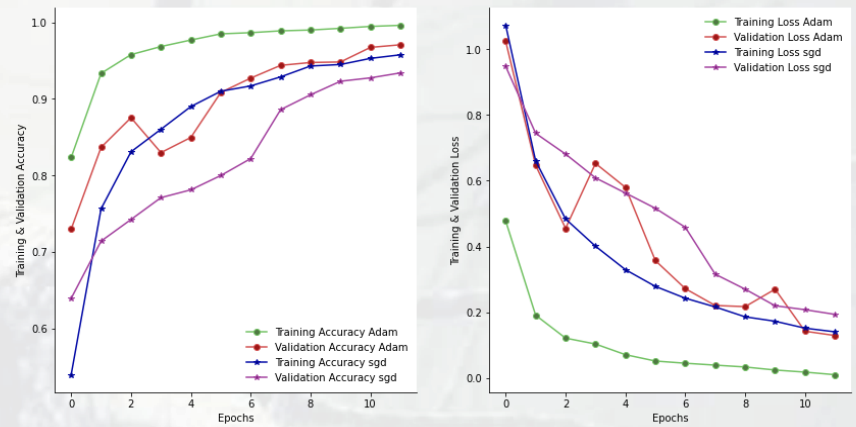
ADAM

- Training set = 0.0099
- Validation set = 0.1292

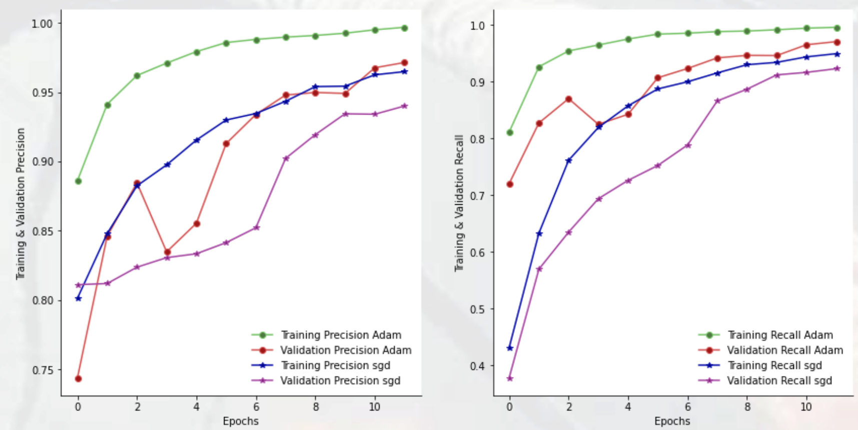
SGD

- Training set = 0.1404
- Validation set = 0.1935

Epochs vs. Training and Validation Accuracy/Loss



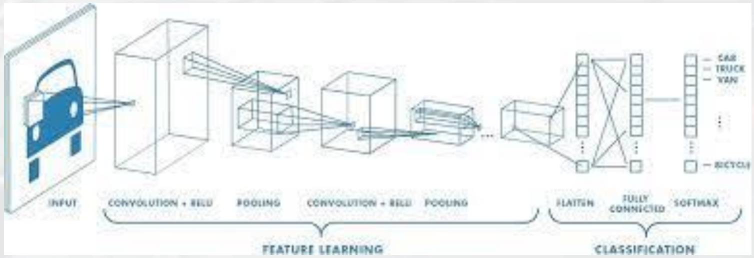
Epochs vs. Training and Validation Precision/Recall



Inside the CNN

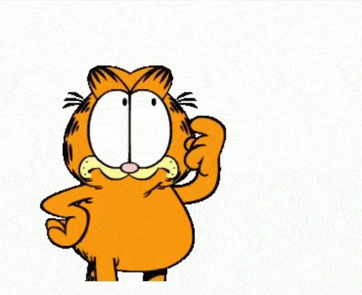
Problems of CNNs applied to Computer Vision:

- Lack of intuitiveness and understandability
- CNNs are hard to interpret
- They are treated as a black box



Goal: make CNNs more transparent

- Why the models predict what they predict?
- What CNNs look in the input image to classify it?

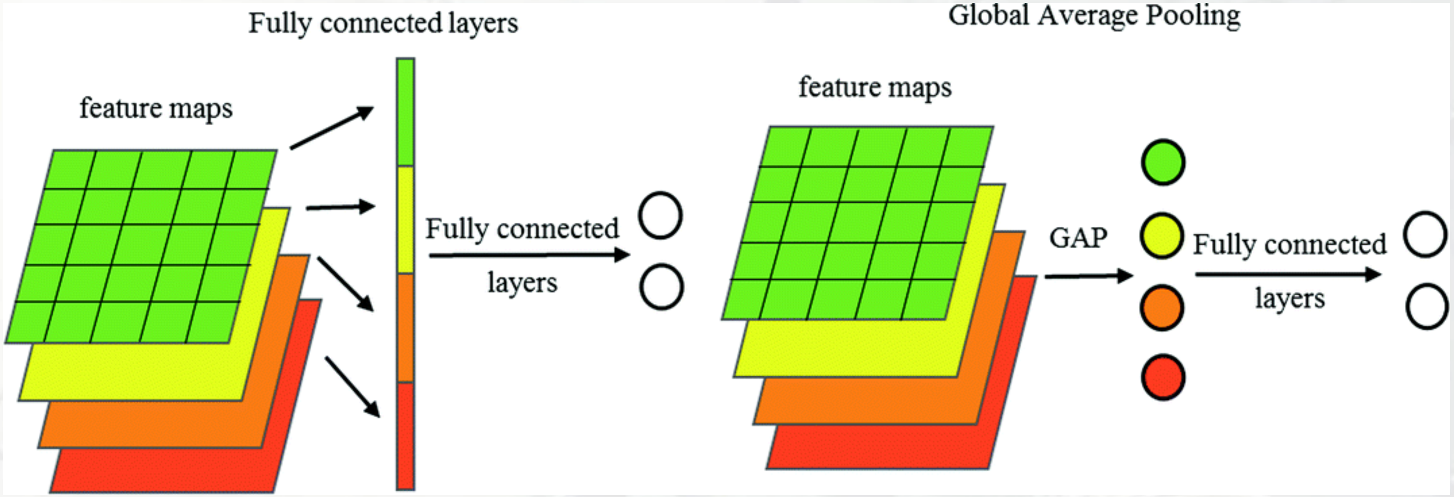


Inside the CNN - What CNNs look in the input image to classify it?

Answer: **Grad-CAM**

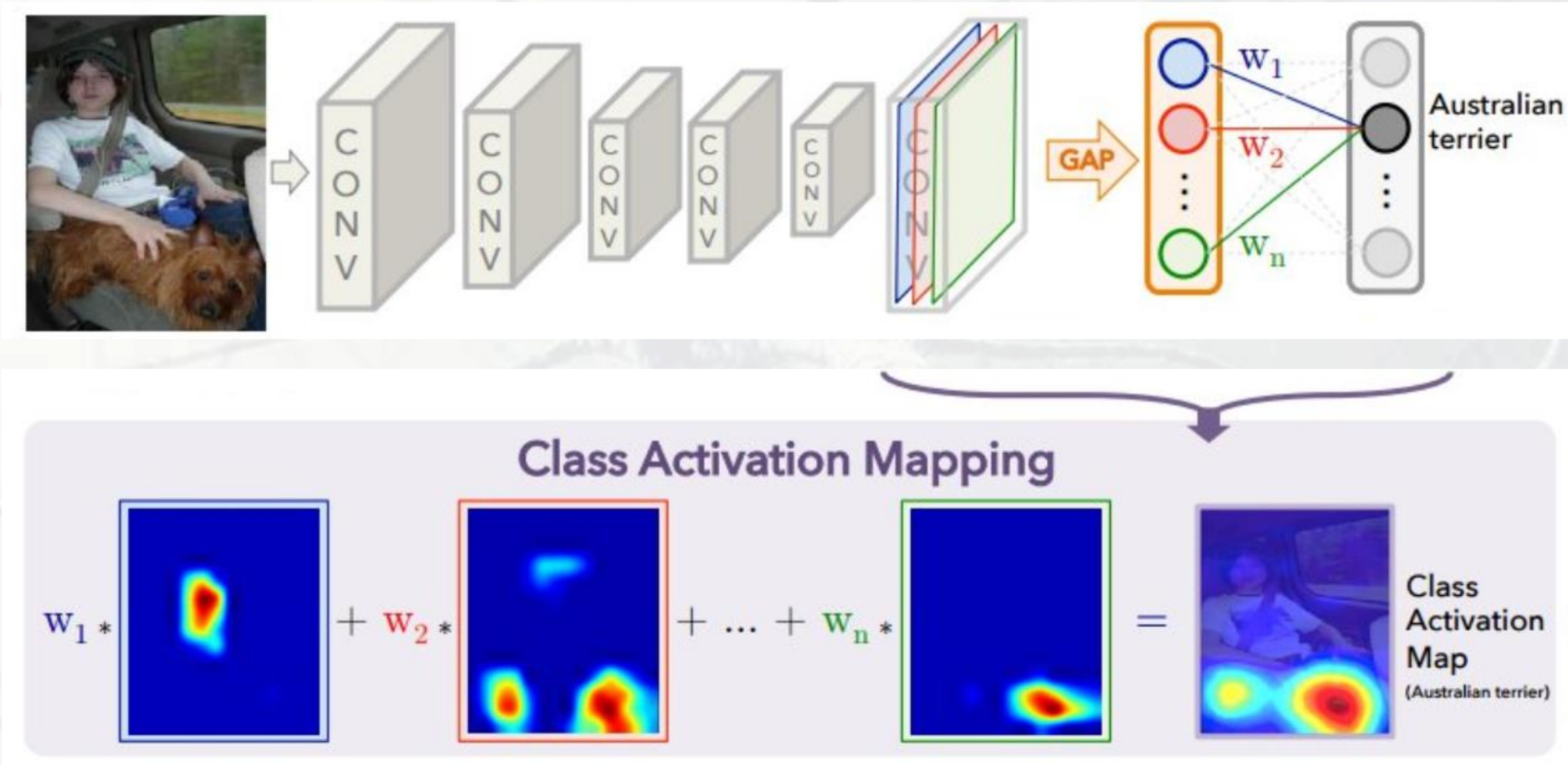
Gradient-weighted Class Activation Mapping (Grad-CAM) is a technique for producing visual explanations of what the model is looking in the image to perform classification

Inside the CNN - CAM & Grad-CAM



Grad-CAM is a generalization of a Class Activation Mapping (CAM) technique, that don't require the GAP layer.

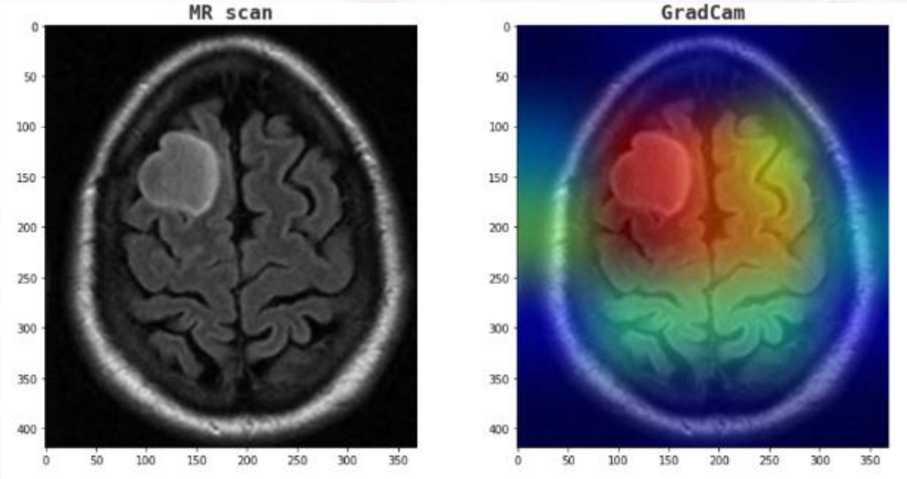
Inside the CNN - How CAM Works



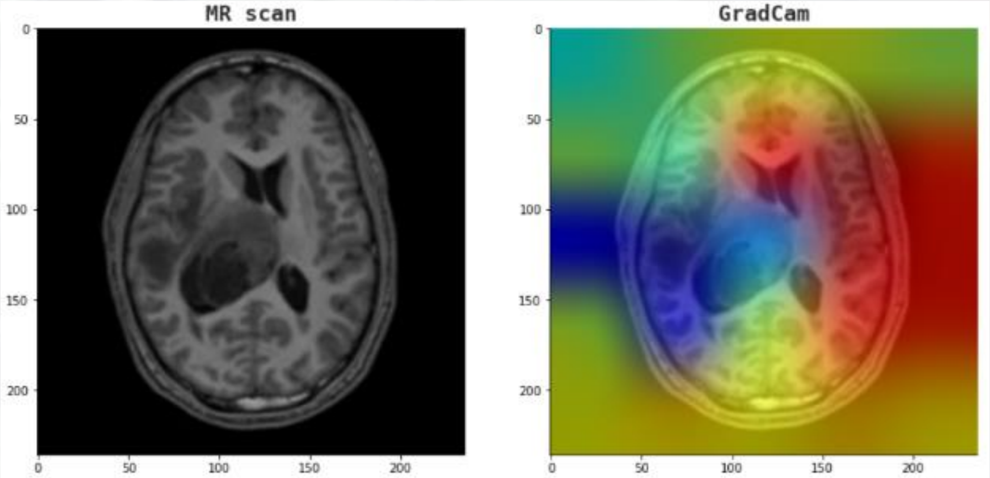
Inside the CNN - Result of Grad-CAM

Correct classification

glioma_tumor

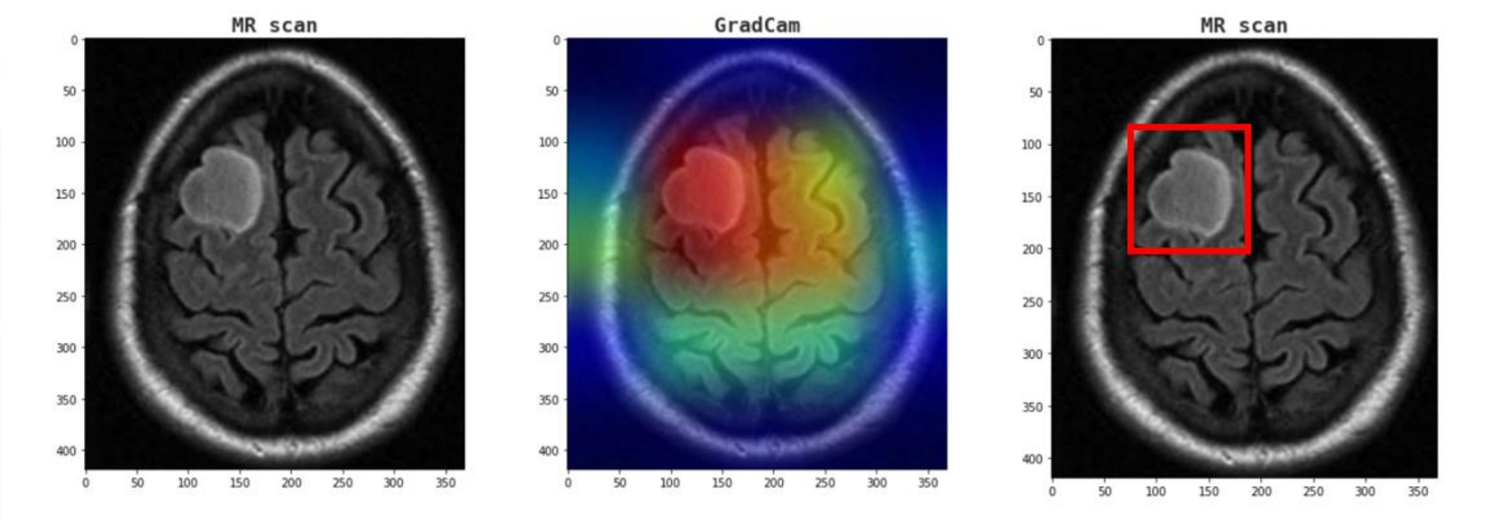


no_tumor



Other finding and future improvements

Grad-CAM underlines that CNNs have remarkable localization ability despite being trained on image-level labels.

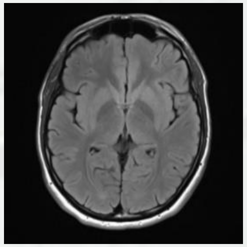


Highlighted problems – UNBALANCED DATA

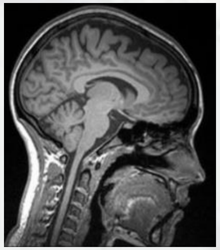
Unbalance for the class no-tumor

Glioma tumor	926
No tumor	500
Meningioma tumor	937
Pituitary tumor	901

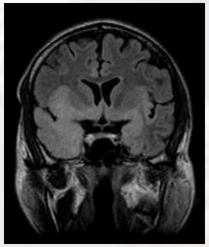
Within the no-tumor class, the majority of the images is assial



Assial



Sagittal



Coronal

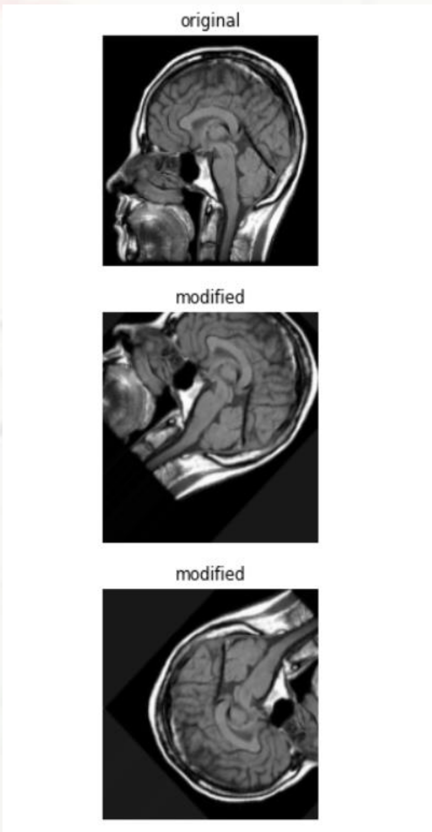
? Does the model learn that no-tumor images are all assial?

Highlighted problems – UNBALANCED DATA

Solution:

New data and data augmentation

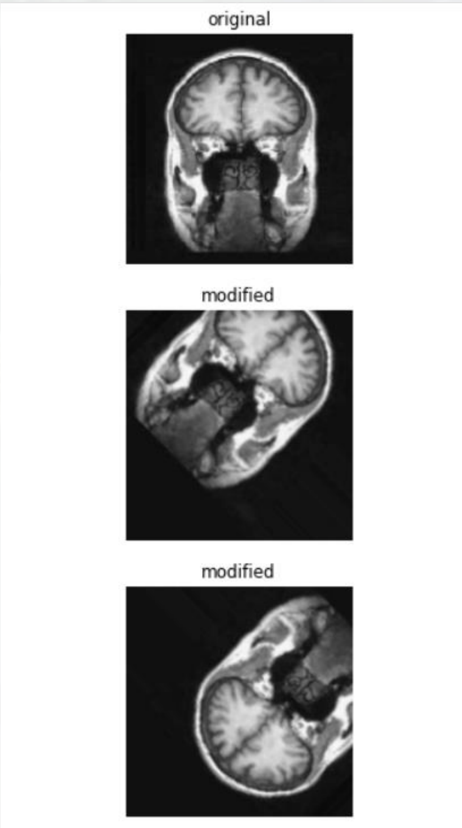
No tumor → 890



New sagittal and coronal images (130)

X translation: 30
Rotation: 45

Y translation: -30
Rotation: 225

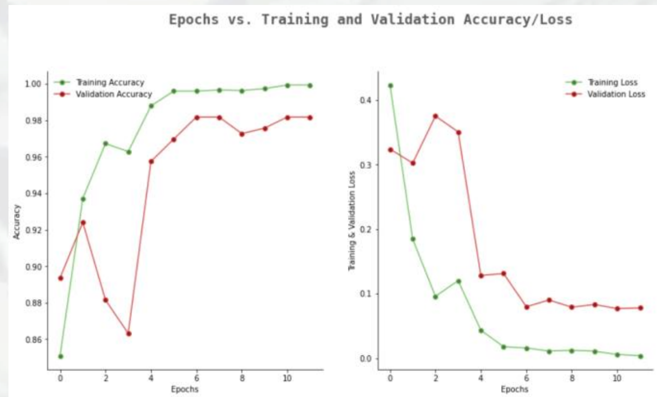


Highlighted problems – UNBALANCED DATA

Results:

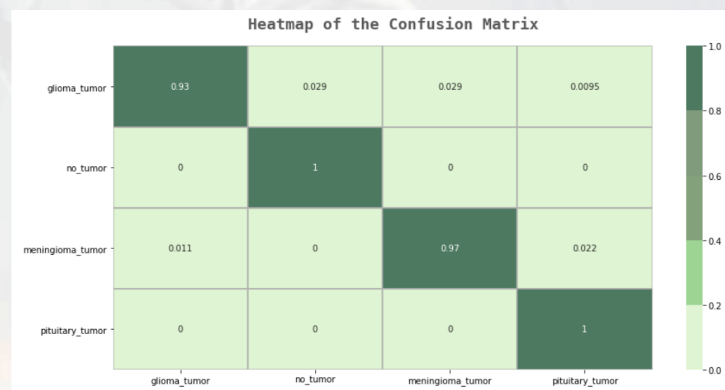
Performances didn't get worse → the original model wasn't affected by the two problems

TRAINING



TESTING

	precision	recall	f1-score	support
0	0.99	0.93	0.96	105
1	0.97	1.00	0.98	86
2	0.97	0.97	0.97	92
3	0.97	1.00	0.98	83
accuracy			0.97	366
macro avg	0.97	0.98	0.97	366
weighted avg	0.97	0.97	0.97	366

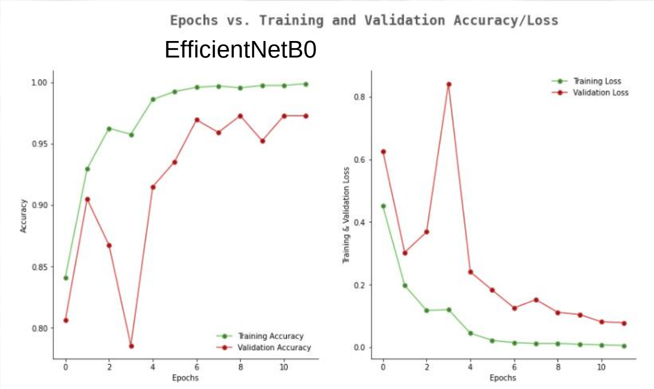


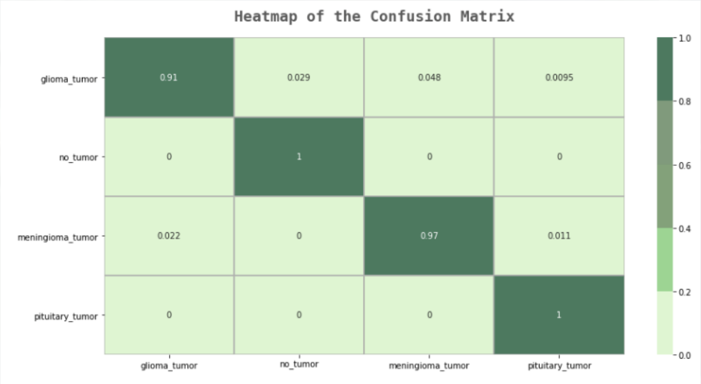
Highlighted problems – WHY EFFICIENT NET?

In our notebook we adopted the transfer learning approach:

- it is the exploitation of knowledge achived from solving a pa rticular task to resolve a new one
- we took the EfficientNetB0 Neural Network as backbone
- we removed the top of the Network
- we added new top layers targeting the new classification task

```
Trainable params: 4,012,672
```





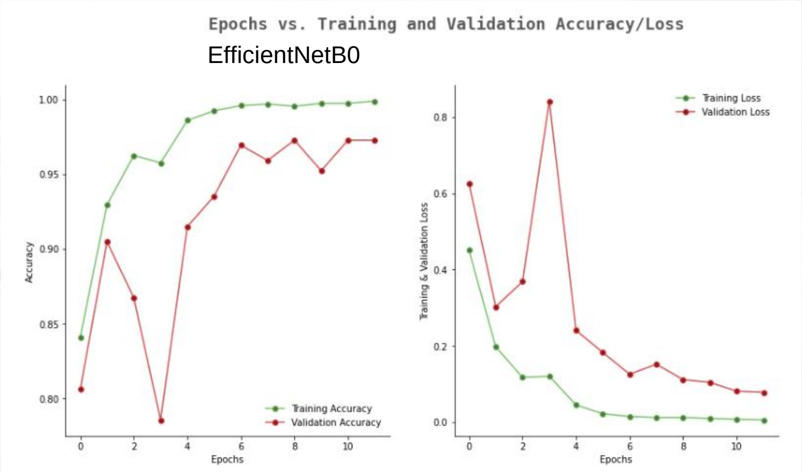
We can see our original model is characterized by: 4M trainable parameters and very good performance

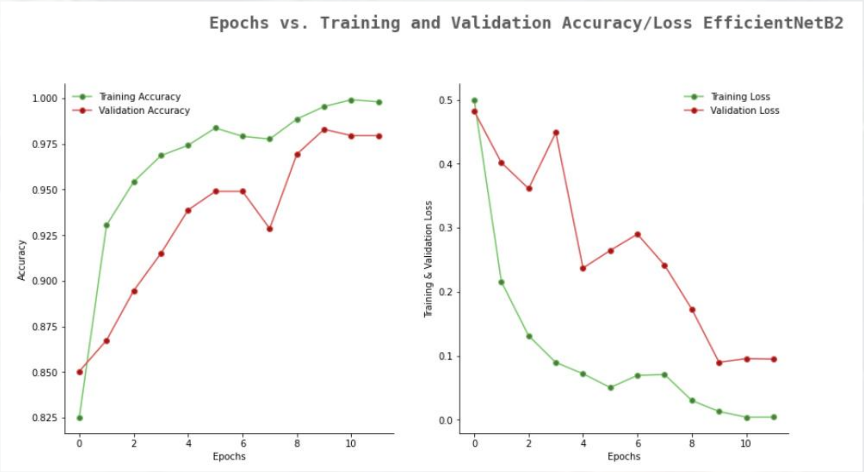
Highlighted problems – WHY EFFICIENT NET?

EfficientNetB2

At this point, we tried to improve the original model performance by exploiting different types of pre set Deep Neural Networks:

- first of all, we took into account the EfficientNetB2
- it is of course similar to EfficientNetB0, but the higher number of parameters could lead to better performance





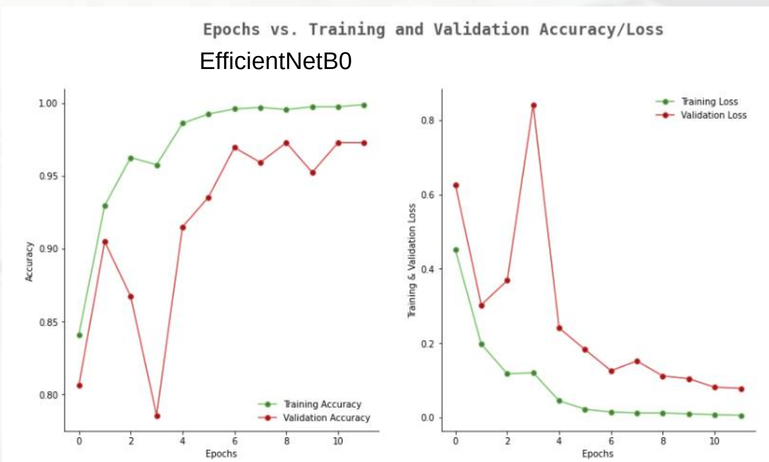
- Despite the number of learnable parameters is heavily increased from 4M to 7.7M, we can not appreciate remarkable improvements
- This leads us to prefer the original Network to accomplish this task.

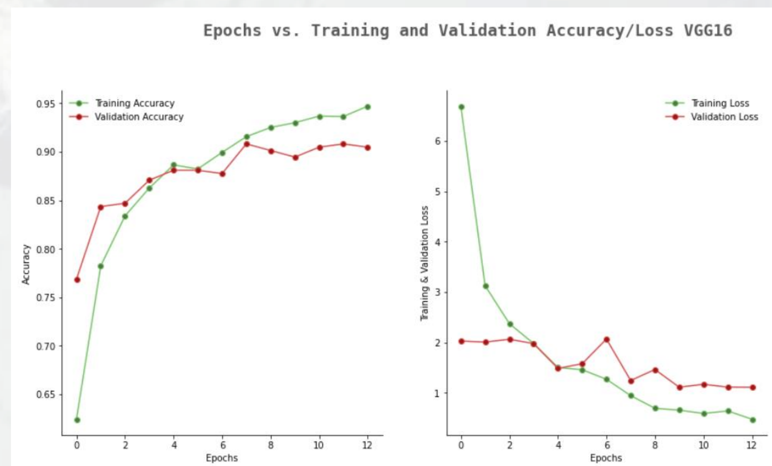
Highlighted problems – WHY EFFICIENT NET?

VGG16

At this point, we tried to improve the original model performances by exploiting different types of pre set Deep Neural Networks:

- in this second attempt, a VGG16 pre-trained model has been adopted as backbone
- we wanted to introduce a new CNN, different from EfficientNet
- we preferred to freeze the VGG16 base model parameters





- as depicted from the graphics, we do not achieve comparable results to the original pre-trained Network.
- this is probably due to the decreased number of trainable parameters of VGG16 (whose pre-trained weigths have been frozen)
- given these facts, also in this case we prefer EfficientNetB0 to accomplish this tumor classification task.

Aim variation- TUMOR VS NO TUMOR

We used the technique of Transfer Learning with the same network (EfficientNet B0) for a simpler task: classification of tumor vs. no tumor images. We did this to better understand the efficiency of this method and the dependence from the proposed dataset.

2 datasets used:

- Original dataset (*brain-tumor-classification-mri*): https://www.kaggle.com/sartajbhuvaji/brain-tumor-classification-mri
- Additional dataset (*new-dataset-yes-no*): https://www.kaggle.com/navoneel/brain-mri-images-for-brain-tumor-detection → 155 tumor images, 98 no tumor images

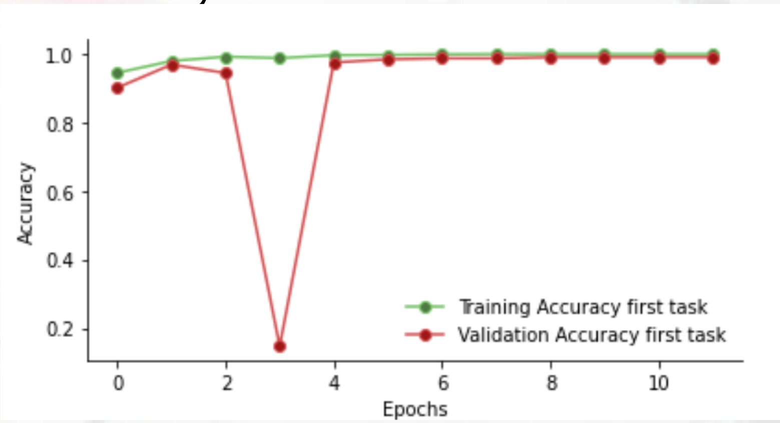
Binary classification problem:

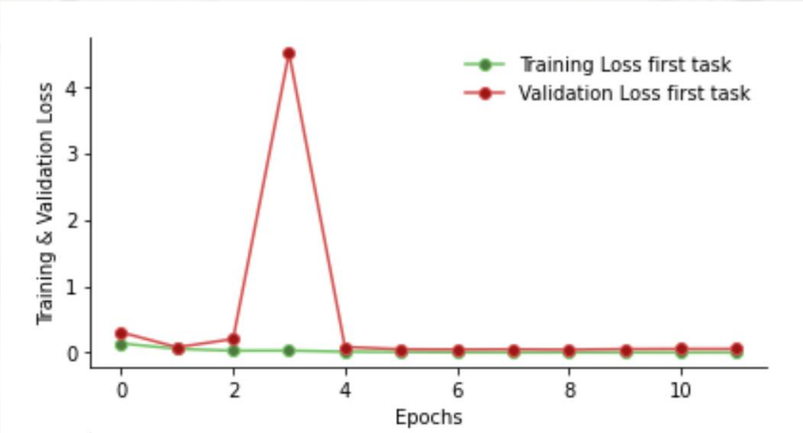
- Sigmoidal activation function in Dense Layer;
- Binary cross-entropy Loss Function;

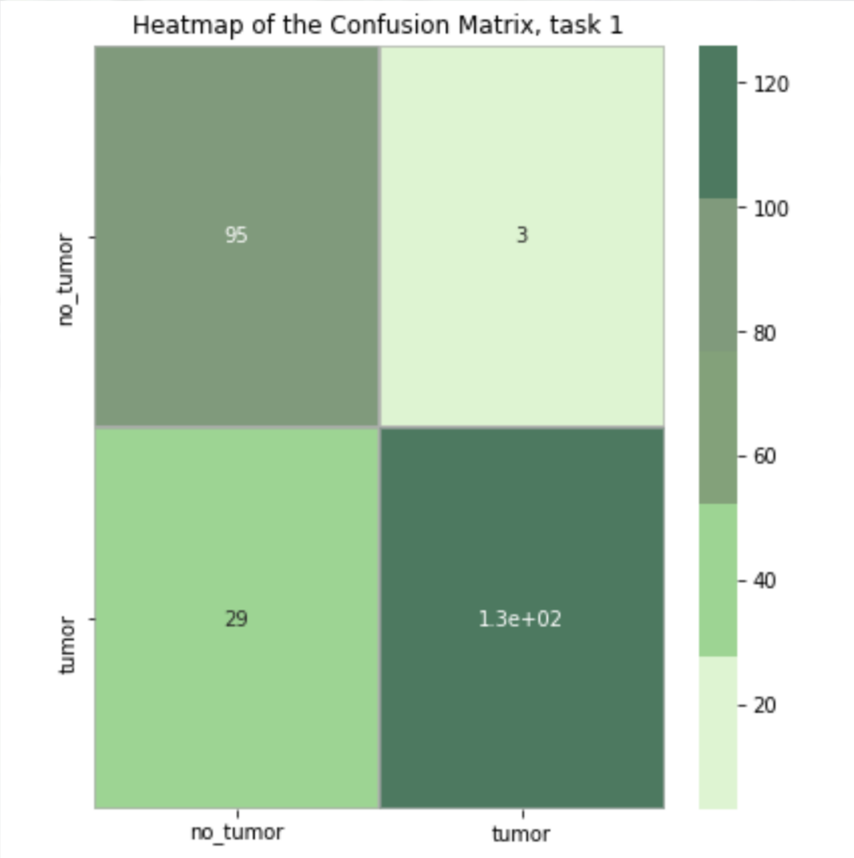
Aim variation- TUMOR VS NO TUMOR

We proposed two tasks:

- **Task 1:** use original dataset (*brain-tumor-classification-mri)* as training set, and additional dataset with few images (*new-dataset-yes-no*) as test set. In this case we reached 87% accuracy on the test set.

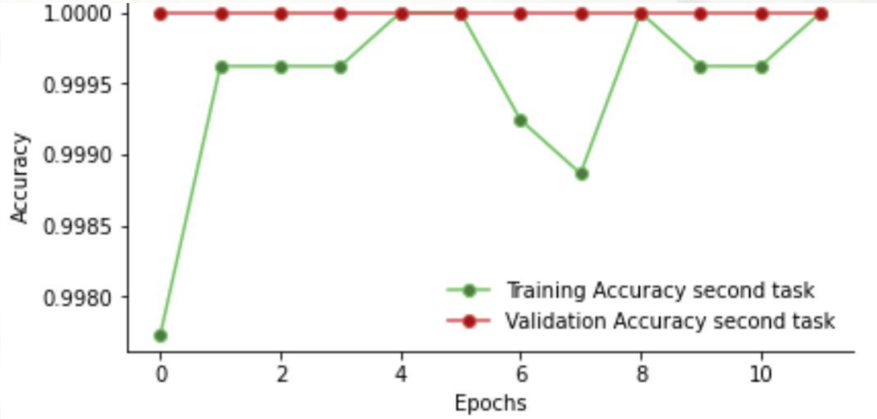


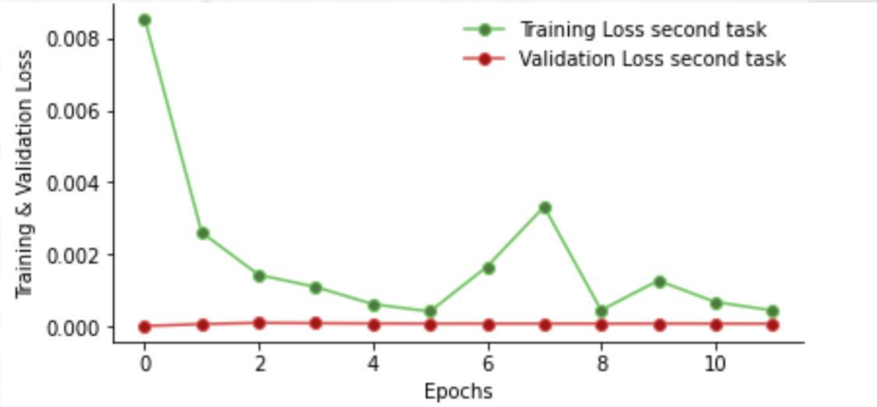


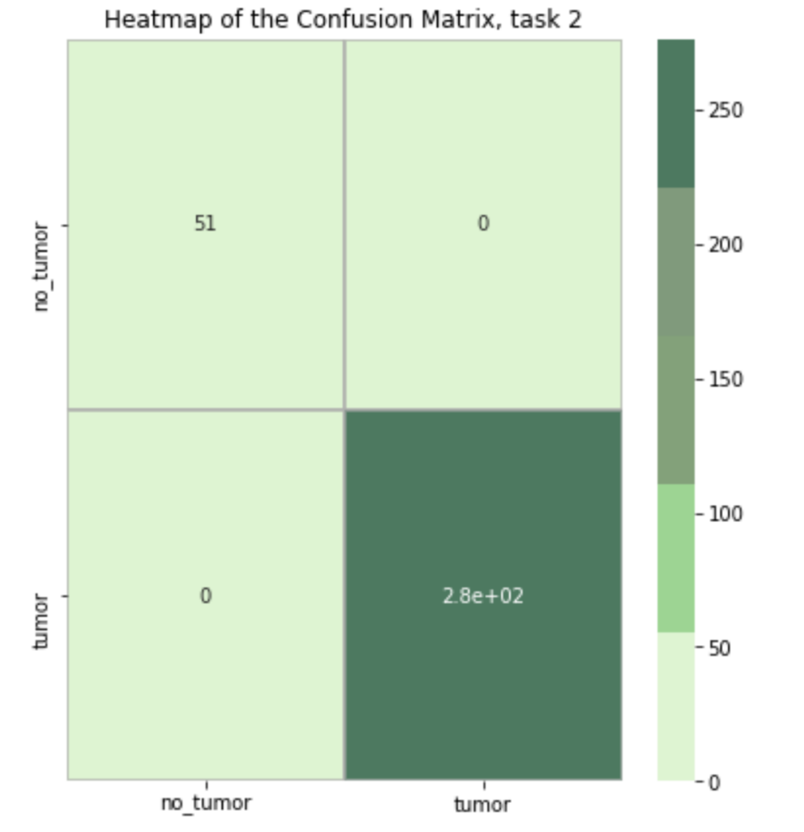


Aim variation- TUMOR VS NO TUMOR

- **Task 2:** use only original dataset, split it into training and test set. In this case we reached an accuracy of 100% on the test set, and an accuracy of 100% on the validation set reached after almost 1 epoch.



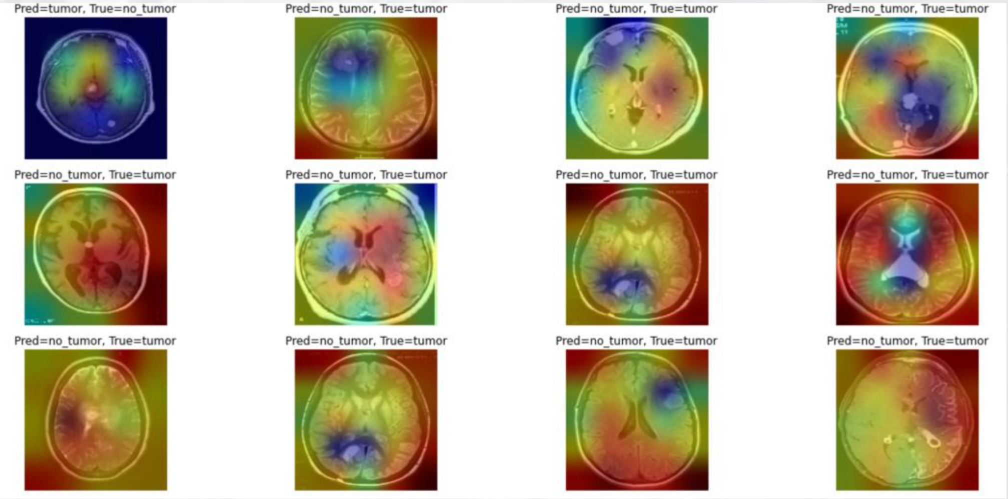




Aim variation- TUMOR VS NO TUMOR

We applied the Grad-CAM alghoritm to the misclassified images of the test set:

- **Task 1:** more or less 30 misclassified images (almost all tumor images).
- **Task 2:** 0 misclassified images.



Observations:

- We tried to use **only** the smaller dataset (*new-dataset-yes-no*) to classify, splitting it into training and test set, we reached an accuracy of almost 90%;
- We tried to use a training set composed only by axial images and a test set composed by saggital and coronal images to understand if the type of image affects the classification. We reached an accuracy of 96%.

Thank you for your attention!

Notebook link:

- FIRST PART - Group 8 | Kaggle
- SECOND PART - Group 8 | Kaggle
- THIRD PART - Group 8 | Kaggle