Precision medicine and quantitative imaging in glioblastoma

ELMED219 (Artificial Intelligence and Computational Medicine), 4-29 January 2021

https://github.com/MMIV-ML/ELMED219-2021

Team #3

1 Research plan

1.1 A brief background to the field

Recognized as the most common and aggressive form of primary brain tumors in adults, but can present at any age [1], glioblastomas (GBM) are lethal with a poor 5-year survival rate of only 5.0% [2]. Initial diagnosis can be achieved through medical imaging techniques such as computed tomography (CT) or magnetic resonance imaging (MRI) to provide contrast enhanced images which help identify important tumor and tissue characteristics both pre- and post-operative [3]. Standard treatment after diagnosis include surgical resection, although due to the invasive nature of tumor cells, prolonged survival after surgical intervention remains a debate [4]. Although surgical intervention remains one of the most common methods of treatment, it is frequently used in combination with other therapeutic options such as radiotherapy, chemotherapy, immunotherapy, and basic supportive care. Due to the inherent dangers glioblastomas present to the patient, treatment options must be considered based on tumor size, location, edema, age, presence of necrosis, molecular characteristics and the overall quality of life of the patient [6] which can be attributed to overall survival.

One of the best methods to get useful information about the tumor is magnetic resonance imaging. MRI takes advantage of how different bound protons react differently to external magnetic fields. Radiowaves are sent in to the tissue, and exites the proton. Depending on the tissue surrounding the proton, it reemits the radiowave at different times. MRI are one of the best imaging modalities when it comes to soft-tissue contrast. This is very helpful to detect and localize tumors. Different weightings (T1, T2, Flair, etc) for the pictures helps to separate tumor and cancerous tissue from healthy tissue. It is also possible to combine MR-imaging with contrast medium to inspect vascular properties. As you can see under (In figure 1), the different tissues have different colors with different weightings:

Tissue	T1-Weighted	T2-Weighted	Flair
CSF	Dark	Bright	Dark
White Matter	Light	Dark Gray	Dark Gray
Cortex	Gray	Light Gray	Light Gray
Fat (within bone marrow)	Bright	Light	Light
Inflammation (infection, demyelination)	Dark	Bright	Bright

Figure 1: A table showing the colors of different tissues at various MRI-weightings, gathered from [7]

In addition to more general medical imaging techniques such as MRI, histological tissue sections analyzed by imaging mass cytometry (IMC) have the potential to bring forward many more biomarkers and phenotypic characteristics of tumor cells. Heavy-metal conjugated antibodies are used to stain tissue sections which are then subjected to laser ablation and measured by time-of-flight mass spectrometry [8]. Each ablation spot and the corresponding measurements of the metal isotope contribute to 1 pixel in the resulting image, and can viewed as a collection of different isotope biomarker-targeting channels. There is already established several supervised machine learning methods used to analyze IMC data [9, 10] and has been proposed that IMC would be an ideal candidate for training deep learning models where thousands of cells and their labels are already embedded in a single image [11].

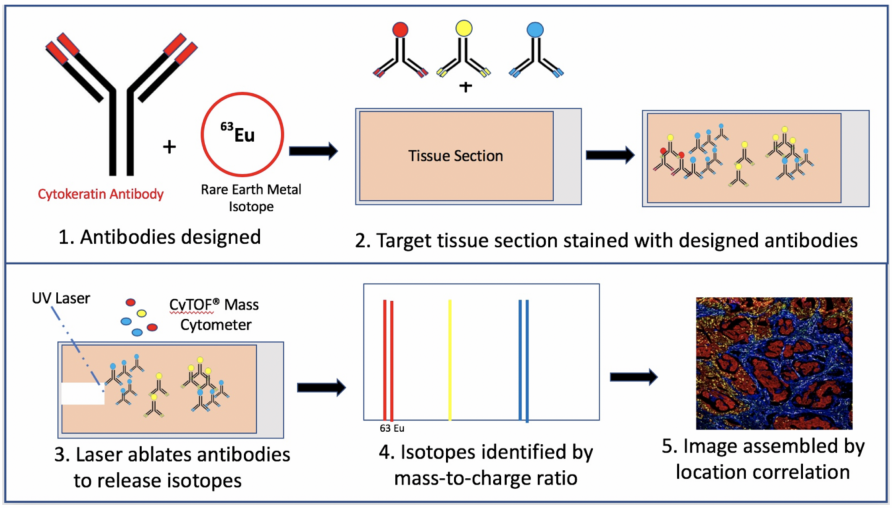

Figure 2: Simplified overview of IMC workflow (Image reproduced from [12])

1.2 Objectives and expected impact

Our goal with this project is to use machine learning to predict how long the patients will survive following initial and/or recurrent diagnosis of glioblastoma, classifying the predicted outcomes as either short ($0-5$ months), medium ($6-12$ months) or long term (>12 months) survival. Using MRI features like tumor size, location, surrounding tissue, edema, necrotic core, and patient information (age, gender, previous treatments), our model can predict the probability of survival for a given patient. This model could be used by clinicians as a guiding tool to evaluate the different treatments options for the patient.

There are many things to consider while choosing the treatment method to use. Efficiency, early and late side effects for the patients, age, etc., is important variables to consider. If a patient has a low survival probability, it is possible to review palliative treatment, which has a goal to relieve the patients from pain rather than cure the cancer. In this case, we can ignore the late side effects for the treatment, and for example deliver more radiation to the tumor.

Using this model will be faster than evaluating the data manually, which could be crucial to cure the cancer. The faster the treatment starts, the better. This will also relieve crucial time for health workers, which can use the saved time on other important tasks. It has to be stated that right now, machine learning can not replace the knowledge of clinicians, but can be used as a guiding tool to save time.

In addition we will analyze tissue sections through IMC, which will be able to give a much more in depth description of how the tumors are built up, the cellular interactions, and possibly identify sub-populations within the tumor microenvironment which will help contribute to the overall understanding of why and how

glioblastomas function they way they do. These analyses are not meant to be used as a part of a predictive model, but will be able to be studied and support in the development of precision therapeutic treatments in the future.

1.3 Material and methods

The study we are planning to do will use data from the Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM) [1] and BraTS[2]. The data includes MR- images in a DICOM-format, information about age, gender, therapy, treatments and survival. We will extract features from the MRI using deep learning. The general information about the patients can be used as additional features for our supervised learning model that in the end will predict how long our patients will live. The features the model extracts, along with the features already included about the patients will contribute to be useful for the doctor in determining which therapy to use. In addition our study will address an alternative method with histological tissue sections and analyze them by using imaging mass cytometry. Tissue samples from biopsies of patients with varying survival length will be provided by biobanks.

1.3.1 MRI data

For patients with progressive glioblastoma multiforme, the best MRI weightings to use are found to be T1 + GBCA (contrast agent) and FLAIR [13]. These weightings are the best to extract useful information about the glioblastoma. Data like size, location, surrounding organs at risk, necrotic core, edema, etc are important features to map for the model to predict the survival probability.

In addition to the MRI pictures, patient data will also be used in training of the model. Data like age, gender, and comorbidities may be useful to predict survival rate and/or choose future treatment of a patient.

1.3.2 Convolutional Neural Network

MRI-images cannot be put directly in a predicting, supervised ML-model which we are planning to use. The images are too information rich, and are not possible for our planned model to evaluate. We need to find a way to quantify the important features of the MRI-images. Using Convolutional Neural Network (CNN) is a good method to gather MRI-features from the images. Studies find that features like amount of necrosis and level of tumor enhancement are important in predicting survival of the patient [18].

To find these features, among others, we can use a multilayer, multiscale CNN model like [19]. Here, Perkuhn evaluated a CNN-model based on DeepMedic, which is trained on a BRATS dataset. The model segmented edema, contrast-enhancing tumor, necrosis and nonenhancing tumor. The model consists of a 3D network with 11 layers.

In our project, features like this are extremely helpful to have available. Once these features are gathered, they can be put into a supervised ML-model to predict the survival of the patient, which we will describe under.

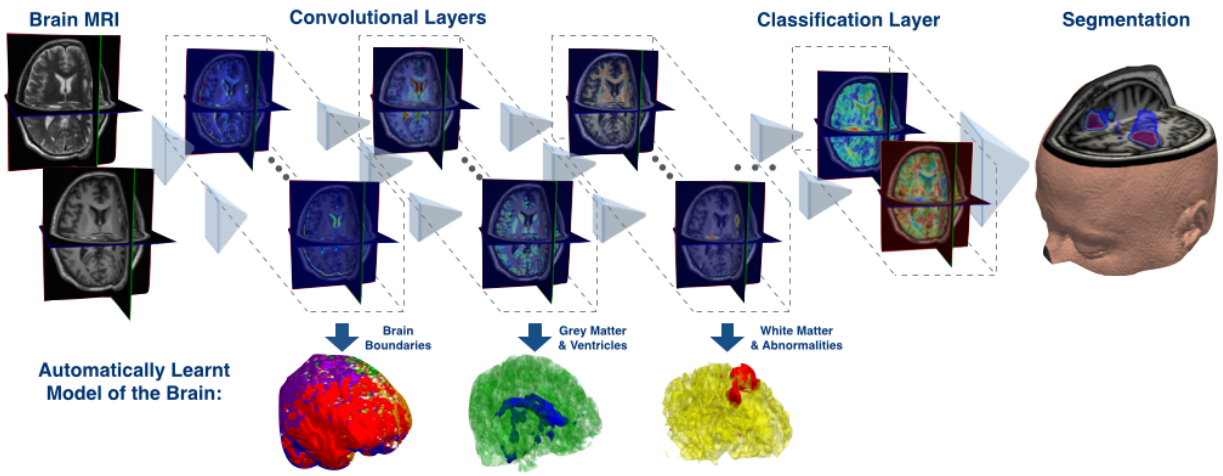

Figure 3: Architecture of the DeepMedic software, gathered from [20]

1.3.3 Supervised ML-model

With the CNN-model above, we can now add the features gathered with other patient data. The MRI-features (edema, contrast-enhancing tumor, necrosis and nonenhancing tumor) along with patient age, gender, previous treatments, etc, will be the input for the new supervised model. Each patient will be labeled with a survival outcome as explained in 1.2. There are three possible labels: short term survival (0-5 months), medium (6-12 months) and long term survival ($>$ 12 months). We will train a supervised ML-model to predict the survival of the patient.

The model we will use is Random Forest Classifier. The RFC-model is made of a large number of uncorrelated decision trees, where each tree makes a class prediction [21]. The model are known to be one of the most accurate classification models. The trees spilts the data points according to the features they are put in with. The class who get the most 'votes' from the classification trees becomes the whole model's prediction.

1.3.4 Tissue staining and IMC image acquisition

Histological tissue sections of 15 patients (with varying survival status: short, medium, and long term) obtained from glioblastoma biopsies will be provided for analysis by imaging mass cytometry. The antibody panel will consist of heavy-metal isotope conjugated antibodies directed towards biomarkers of immune cells, glioblastoma stem cells, DNA, cells undergoing apoptosis/necrosis, markers associated with hypoxia, cell membrane markers, and vascular endothelial growth factor. Tissue staining will be performed as outlined in the Imaging Mass Cytometry Staining Protocol supplied by Fluidigm, and the IMC image stacks and corresponding data will be acquired through analysis by use of the Hyperion Imaging System (Fluidigm).

1.3.5 Supervised segmentation and classification of the IMC stack

Cell segmentation and pixel feature classification will be performed on the IMC image stacks by using the supervised machine learning tool ilastik, which is adapted for (bio)image analysis. The different classes of our pixel features will be grouped into user-defined categories such as 'nucleus', 'membrane', and 'background' to define cell borders, as well as classified into groups of 'immune cells', 'glioblastoma stem cells' and 'other' based on the expression of relevant markers. Ilastik also provides it's own feature selection by color/intensity, edge, and texture. Classification is performed using a built in Random Forest classifier, and

trained directly through the addition of labels. By drawing directly on the image in batch mode we can label these classes and get immediate predictive feedback after giving additional annotations. The final segmentation masks for each class of cell should then be able to be turned on and the IMC channels can be visualized as a spatially organized image depicting different cell phenotypes, as well as confirming the up-regulated presence of specific biomarkers associated with aggressive GBM.

Exported as tiff images along with their accompanying csv files further anaylsis can be performed in a Jupyter notebook. Working in python we can transform this data into a pandas dataframe and calculate pairwise association between the metal tags, potentially revealing any biological relationships between the biomarkers of interest. By importing scikit-learn and other libraries into our jupyter notebook, this will allow us to perform further data analysis of the tissue micro-environment which can then be studied and utilized by physicians when analyzing tumor heterogeneity and considering treatment options.

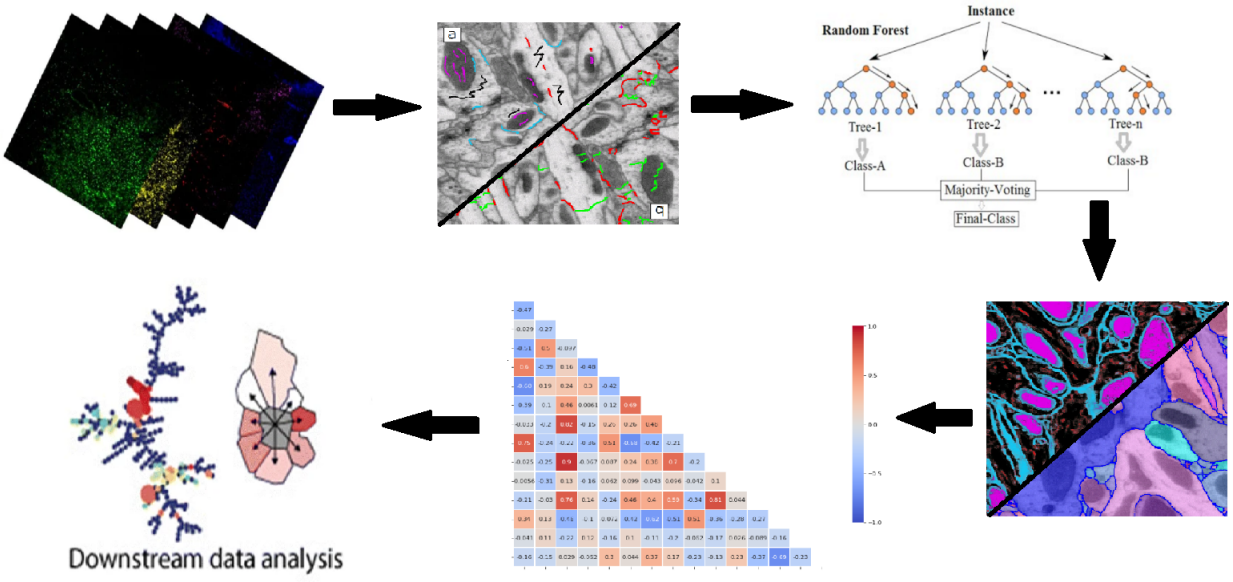

Figure 4: Supervised segmentation and classification of the IMC image stack using Ilastik and possible downstream data analysis working in jupyter notebook (Individual images reproduced and adapted from [10, 15, 16, 17, 14])

1.4 Evaluation

After our model has been trained on the training data, a confusion matrix will be used to evaluate the model. This will give a representation of the model's predicted outcomes compared to the "ground truth". The predicted outcomes (y) are measured from test data (obtained MRI images), and will indicate short, medium or long term survival rate. The "ground truth" is the actual survival rate of a patient, which are pre-labeled to the test data. The accuracy is the number of correct predictions divided by the total number of predictions. Accuracy can also be performed on test data, after the model is trained. It will result in a value between 0 and 1, where 1 represents 100 percent correct predictions. The accuracy is strongly related to the amount of training data that is applied to the model, where more data will give a higher accuracy.

In ilastik, evaluation properties of pixel classification and eventual segmentation are already built in. Users can view an uncertainty map which indicates areas where the classifier is unsure about the prediction results, and manually add additional annotations to these regions. With interactive predictive results, users can correct mistakes directly to improve the classifier, and in order to provide feedback as to which features

were important for a specific classification task, ilastik outputs a variable importance.

References

[1] Davis ME. Glioblastoma: Overview of Disease and Treatment. Clin J Oncol Nurs. 2016;20(5 Suppl):S2-S8. doi:10.1188/16.CJON.S1.2-8

[2] Ostrom QT, Gittleman H, Liao P, et al. CBTRUS statistical report: primary brain and central nervous system tumors diagnosed in the United States in 2007-2011. Neuro Oncol. 2014;16 Suppl 4(Suppl 4):iv1-iv63. doi:10.1093/neuonc/nou223

[3] Shukla G, Alexander GS, Bakas S, Nikam R, Talekar K, Palmer JD, Shi W. Advanced magnetic resonance imaging in glioblastoma: a review. Chin Clin Oncol 2017;6(4):40. doi:10.21037/cco.2017.06.28

[4] Delgado-López, P.D., Corrales-García, E.M. Survival in glioblastoma: a review on the impact of treatment modalities. Clin Transl Oncol 18, 1062–1071 (2016). doi.org/10.1007/s12094-016-1497-x

[5] Shboul Z.A., Vidyaratne L., Alam M., Iftekharuddin K.M. (2018) Glioblastoma and Survival Prediction. In: Crimi A., Bakas S., Kuijf H., Menze B., Reyes M. (eds) Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries. BrainLes 2017. Lecture Notes in Computer Science, vol 10670. Springer, Cham. doi.org/10.1007/978-3-319-75238-9_31

[6] Fernandes C, Costa A, Osório L, et al. Current Standards of Care in Glioblastoma Therapy. In: De Vleeschouwer S, editor. Glioblastoma [Internet]. Brisbane (AU): Codon Publications; 2017 Sep 27. Chapter 11. doi:10.15586/codon.glioblastoma.2017.ch11

[7] Preston, David C. "Magnetic Resonance Imaging (MRI) of the Brain and Spine: Basics". https://case.edu/med/neurology/NR/MRI\%20Basics.htm

[8] Gerdtsson E, Pore M, Thiele JA, et al. Multiplex protein detection on circulating tumor cells from liquid biopsies using imaging mass cytometry. Convergent Science Physical Oncology. 2018 Mar;4(1). doi:10.1088/2057-1739/aaa013

[9] Carpenter, A.E., Jones, T.R., Lamprecht, M.R. et al. CellProfiler: image analysis software for identifying and quantifying cell phenotypes. Genome Biol 7, R100 (2006). doi.org/10.1186/gb-2006-7-10-r100

[10] Berg, S., Kutra, D., Kroeger, T. et al. ilastik: interactive machine learning for (bio)image analysis. Nat Methods 16, 1226–1232 (2019). doi.org/10.1038/s41592-019-0582-9

[11] Baharlou H, Canete NP, Cunningham AL, Harman AN and Patrick E (2019) Mass Cytometry Imaging for the Study of Human Diseases—Applications and Data Analysis Strategies. Front. Immunol. 10:2657. doi:10.3389/fimmu.2019.02657

[12] de Miranda N.F., Hargraves K. Imaging Mass Cytometry: Endless possibilities to understand the microenvironment of disease. 2019 https://blog.cellsignal.com/imaging-mass-cytometry-endless-possibilities-to-understand-the-microenvironment-

[13] Thompson, Eric M et al. "Correlation of MRI sequences to assess progressive glioblastoma multiforme treated with bevacizumab", 2011. https://pubmed.ncbi.nlm.nih.gov/20848300/

[14] https://en.wikipedia.org/wiki/Random_forest

[15] Giesen, C., Wang, H., Schapiro, D. et al. Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry. Nat Methods 11, 417–422 (2014). doi.org/10.1038/nmeth.2869

[16] Kho J. Annotated Heatmaps of a Correlation Matrix in 5 Simple Steps. https://www.kdnuggets.com/2019/07/annotated-heatmaps-correlation-matrix.html

[17] Ramaglia V, Sheikh-Mohamed S, Legg K, et al. Multiplexed imaging of immune cells in staged multiple sclerosis lesions by mass cytometry. Elife. 2019;8:e48051. Published 2019 Aug 1. doi:10.7554/eLife.48051

[18] Hammoud M A, et al. 'Prognostic significance of preoperative MRI scans in gliovlastoma multiforme', 1996. https://pubmed.ncbi.nlm.nih.gov/8699228/

[19] Perkuhn, Michael. 'Clinical Evaluation of a Multiparametric Deep Learning Model for Glioblastoma Segmentation Using Heterogeneous Magnetic Resonance Imaging Data From Clinical Routine' 2018. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7598095/

[20] https://biomedia.doc.ic.ac.uk/software/deepmedic/

[21] 'Understanding Random Forest'. towards data science, 2019. https://towardsdatascience.com/understanding-random-forest-58381e0602d2

2 Data management plan and ethical considerations

2.1 Description of generated data and code

The data we plan to use in this study will be directly accessed and collected from The Cancer Genome Atlas Glioblastoma Multiforme (TCGA-GBM) [1], specifically we have collected MRI images from the 256 participants with different stages of glioblastoma. The data includes MR images as well as clinical and biomedical data such as age, gender, therapy treatments and survival. In addition, we will use the dataset from BraTS [2], which includes MRI T1, T1Gd, T2 and T2-Flair images as well as overall survival data, age of patients and resection status. Although due to the inherent nature of machine learning methods and how their accuracy is largely dependent on large datasets for training, we can implement what is called data augmentation to alter the images and increase the amount of variability in the dataset. This can be done by simply flipping (mirroring horizontally or vertically), rotating, bluring, or shifting the center point of the image as is described here [3]. In this way we can artificially increase the size of the dataset in hopes of producing a model with better prediction accuracy.

For IMC data analysis, we have generated image stacks (for each donor) across all isotope measuring channels converted to TIFF files using the Bodenmiller Group preproccessing script (found at the GitHub repository here [4]), corresponding metadata provided as csv files, and panel file with information related to the specific metal-tagged antibodies used to stain the tissue along with their corresponding channels.

2.2 Sharing of data and code

Code used to build and construct the predictive neural network model will be made available on GitHub so that others can reproduce the model to use on their own datasets. The predicted outcomes of model and it's test data from this project will be uploaded to Mendely Data [5] so that it can be made publically available and citable to others. IMC generated data along with the finalized segmentation masks will also be uploaded to Mendely Data.

2.3 Ethical considerations

First and foremost this research plan will be submitted to The Norwegian National Research Ethics Committee for review and approval prior to any data analysis and experimentation. All MRI images obtained from TCGA and the open dataset from BraTS will be cited accordingly to avoid any misuse of the data. Contracts will be drawn up between the biobanks which provide the tissue samples and our research group where additional information such as patient age, overall survival data, and treatment given will be included. We will not be in direct contact with donors or their family, and they must also consent to the use of their tissue samples being used for research as outlined in a contract between the biobank and themselves.

In line with guaranteeing anonymity to the patient, treating all fairly and justly, and to protect any potential breaches of privacy, pseudo-identifiers will be used for both the MR images and tissue analysis. Our method of ground truth labeling (overall survival) is recognized as an extreme predictive outcome and we realize this model may not be at all times appropriate, although important, in the clinical setting. We therefore propose that to do the most good and avoid automation bias, the predicted outcomes be used in such a way as to guide the decision of treatment alternatives, avoiding over-treatment while keeping the patients best interest in mind. We will evaluate the trained model based on performance to ensure that before clinical application, the model has a high predictive accuracy such that we can reduce the amount of errors when being used on new and unseen data. New data, as it becomes available for research use should be used to test and train the model to ensure that it's performance does not degrade over time and working as expected. Unexpected biases may exist in the data that we are unaware of, and as we train and test the model if we

become aware of any bias we will take steps to mitigate these. If any other ethical challenges arise we aim to take contact with ethical committees that specialize in areas of artificial intelligence.

References

[1] The TCGA-GBM data collection. Accessed on 28/12/2020.
https://wiki.cancerimagingarchive.net/display/Public/TCGA-GBM

[2] http://braintumorsegmentation.org/

[3] Safdar MF, Alkobaisi SS, Zahra FT. A Comparative Analysis of Data Augmentation Approaches for Magnetic Resonance Imaging (MRI) Scan Images of Brain Tumor. Acta Inform Med. 2020;28(1):29-36.
doi:10.5455/aim.2020.28.29-36

[4] scriptsizehttps://github.com/BodenmillerGroup/imctools

[5] https://data.mendeley.com/

[6] Vollmer S, Mateen BA, Bohner G et al. Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. BMJ 2020;368:l6927.
https://www.bmj.com/content/368/bmj.l6927