Analysis of Brain Tumor Using MRI Images

Lewi L. Uberg
Applied Data Science
Noroff University College
Kristiansand, Norway
lewi@uberg.me

Seifedine Kadry
Applied Data Science
Noroff University College
Kristiansand, Norway
seifedine.kadry@noroff.no

Abstract—**The increasing rates of deadly brain tumors in humans correspondingly increase the need for highly experienced medical personnel for diagnosis and treatment. Therefore, to reduce the workload and the time from suspicion of disease to diagnosis, then plan for suitable treatment, there is a need to automate the initial part of the process by implementing a Computer-Aided-Disease-Diagnosis (CADD) system for brain tumor classification. The aim of this research is to develop and deploy all the needed components to design and implement a working Computer-Aided-Disease-Diagnosis (CADD) system for brain tumor classification. First, by understanding the brain tumors themselves and the convention of their classification. Second, the convolution neural network (CNN) structure and functionality and how to manipulate it. Third, find different CNN architectures with promising results in similar tasks. Fourth, find and evaluate the applicability of available data sources. Fifth, implement the most promising solutions and explore the applicability of using transfer learning for the given task to take advantage of previously gained knowledge. Finally, evaluate the results of the experiments, where we show that the DenseNet121 architecture, either fully trained or using transfer learning, likely is the most appropriate candidate for the CADD system in development.**

Index Terms—**Brain Tumor Classification, Medical Imaging, CADD, Convolutional Neural Networks, Machine Learning, Deep learning.**

I. INTRODUCTION

THE CELL of origin and features found when examining the cells tissue define central nervous system tumors and predict their behavior [1]. After meningiomas, the most common primary brain tumor in adults is gliomas, with a rate of 5 to 6 persons per 100,000 annually [2].

The World Health Organization (WHO) tissue classification system categorizes gliomas, with grade 1 as the lowest and grade 4 as the highest. Thus, low-grade gliomas (LGG) consist of grade I and II tumors [3], while high-grade gliomas (HGG) consist of grade III and IV [2]. Grade I are the least malignant or benign tumors. Grade II is relatively slow-growing but may recur as a higher grade. Grade III is malignant and tends to recur as higher grades. Finally, grade IV is the most malignant, aggressive, necrosis and recurrence prone [1].

Histopathologic examination to study tissue morphology, diagnose, and grade brain tumors is the gold standard [3]. However, surgical resection for diagnosing a brain tumor is invasive and risky. Nevertheless, non-invasive diagnostic methods exist, like neuroimaging with Magnetic Resonance Imaging (MRI) [4].

Conventional MRI is the current imaging procedure of choice and identifies tumor size and associated Peritumoral Edema (PTE), one of the main features of malignant glioma [5]. To diagnose the brain tumor without surgical intervention, researchers have developed more advanced imaging methods such as texture analysis, mechanic modeling, and machine learning (ML) that form a predictive multi-parametric image-based model. The application of ML models is an emerging field in radiogenomics and represents a data-driven approach to identifying meaningful patterns and correlations from often complex data sources. ML models train by feeding the ML algorithm a substantial amount of pre-classified image data as input, such as MRI images, to learn which patterns belong to the different classes.

The convolutional neural network (CNN) is a concept introduced by [6] as the Neocognitron, as a model of the brain's visual cortex; an improvement of [7] previous model for visual pattern recognition. Furthermore, [8] significantly improved the Neocognitron to one of the most successful pattern recognition models, dramatically influencing the field of computer vision and pattern detection in images.

In 2021 a fully automatic hybrid solution for brain tumor classification comprised of several steps was proposed [9]. First, pre-process the brain MRI images by cropping, resizing, and augmenting. Second, use pre-trained CNN models for feature extraction with better generalization. Third, select the top three performing features using fined-tuned ML classifiers and concatenate these features. Finally, use the concatenated feature as input for the ML classifiers to predict the final output for the brain tumor MRI. The proposed scheme uses a novel feature evaluation and selection mechanism, an ensemble of 13 pre-trained CNNs, to extract robust and discriminative features from brain MRI images without human supervision. The CNN ensemble, is comprised of ResNet-50, ResNet-101, DenseNet-121, DenseNet-169, VGG-16, VGG-19, AlexNet, Inception V3, ResNext-50, ResNext-101, ShuffleNet, MobileNet, and MnasNet. Since the researchers use fairly small datasets for training, they take a transfer learning-based approach by using the fixed weights on the bottleneck layers of each CNN model pre-trained on the ImageNet dataset.

II. DATA GATHERING & PRE-PROCESSING

The data-gathering part of the research starts with outlining some criteria. The data should contain non-tumorous, LGG,

and HGG samples and be well balanced between the labels. The latter is essential since too much data augmentation on medical images often does not generalize well. A well-suited candidate was found at The Cancer Imaging Archive (TCIA). The REMBRANDT (REpository for Molecular BRAin Neoplasia DaTa) Dataset [10] seemed to have all this research's characteristics. The dataset is one of the most trusted publicly available datasets, comprised of MRI scans from 130 subjects of three classes, non-tumorous, LGG, and HGG. Furthermore, the LGG and HGG classes have subclasses that opened the opportunity to make the classification outcome more extensive. The dataset is a combination of metadata from various text and spreadsheet files and the 110,020 images in the DICOM format, which also includes a vast quantity of metadata. After preprocessing, the dataset comprises 123 patients and 105,265 slides, distributed as shown in Table I. At this point, it was discovered that one key bit of information was missing; how to separate the MRI slides that contained the tumorous cells from the ones that did not contain them. All the source data were reexamined, but the answer was not found. While searching online for how to find the needed key, one paper stood out [11]. The paper used that same dataset for a similar application. While being reduced to 4069 slides, it seemed that the authors of this paper had found a way to filter the data further. A meeting with the paper's main author was arranged to understand how to reproduce the dataset. His research team had employed help from neurologists to go through each slide manually and label them correctly. Fortunately, the main author offered to share a smaller dataset version. After further processing, the dataset has 735 samples distributed, as shown in Table II, and is now ready for training CNN models.

TABLE I
DATASET SAMPLE DISTRIBUTION

Disease	Grade	Label	Unique Samples	Count
Astrocytoma	II	LGG	30	25286
Astrocytoma	III	HGG	17	16038
GBM	IV	HGG	43	32837
Non-Tumorous	n/a	n/a	15	17041
Oligodendroglioma	II	LGG	11	9335
Oligodendroglioma	III	HGG	7	4728
Total			**123**	**105265**

TABLE II
MANUALLY LABELED DATASET SAMPLE DISTRIBUTION

Label	Sample Count
Normal	168
LGG	287
HGG	280
Total	735

III. METHODOLOGY

Finding the CNN architecture with or without transfer learning that yields the best results for the classification task is the primary goal of the current research. Therefore designing a

pipeline where the CNN architecture can easily be substituted is essential. Doing so required finding the generally most suitable division of the Test, Validation, and Test subsets, the most suitable baseline hyperparameter settings, and whether image augmentation was to be used. For this task, VGG16 [12] was selected, as it had been used for the same application with promising results [13] [14]. After training a large number of VGG16 models, a baseline for training other CNN architectures was established. A total of 60 images, 20 from the three classes, were randomly picked for the Test set. The remaining images are divided into 30% and 70% for the Validation and Training sets. The Training set is also augmented with rotation, width and height shift, shearing, zoom and brightness adjustment, and horizontal flipping. Filters with minimal adjustments to the original gave better results than fewer with more extensive adjustments. The Normal label has a little over half of the samples as LGG and HGG, and is therefore producing a higher number of augmented images. The Normal label in the Train set is increased from 103 to 309, LGG from 186 to 372, and HGG from 182 to 364. With a batch size of 24, the baseline uses 10 epochs for training with a learning rate of 0.001, Adam for optimization with categorical cross-entropy as the loss function. In addition to the VGG16 architecture, five other CNN architectures were used to train models, including AlexNet, MobileNet [15], ResNext [16], DenseNet121 [17], and a custom-designed architecture. The custom CNN architecture is designed to accept N amount of 224x224 image matrices with 3 color channels; it consists of three convolutional layers of 32, 64, and 64 filters, of filter size 3x3, with zero-padding, and ReLU as the activation function. The first convolutional layer is followed by a max-pooling layer of pool size 4x4, and the last two convolutional layers have a pool size of 2x2, each followed by a dropout layer with a dropout rate of 0.15. Next, a flattening layer is added to transform matrix output to vector inputs to be accepted by the first dense layer, which is comprised of 512 ReLU activated neurons, followed by a dropout layer with a 0.5 dropout rate. The last hidden layer is a dense layer of 256 ReLU activated neurons. The model's final layer, its output, is a 3 neuron softmax activated dense layer for classification. The general architecture of this model is shown in Fig. 1, and the folowchart is provided in Fig. 2.

IV. RESULTS

All the CNN architectures implemented to this point show outstanding results, as shown in Table III. While multiple runs were made on each architecture to obtain the given results, and there is an indication that the Test set in particular needs expansion, the results indicate that DenseNet121 is the best candidate for the given classification task. Surprisingly enough, with the custom architecture as the runner-up.

V. RESEARCH CONTRIBUTIONS

The research conducted has been a good start for further developing a usable CADD system. A pipeline for data preprocessing that makes preparing new data samples both easy and
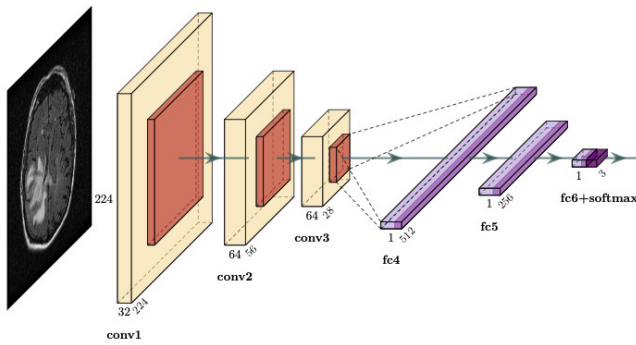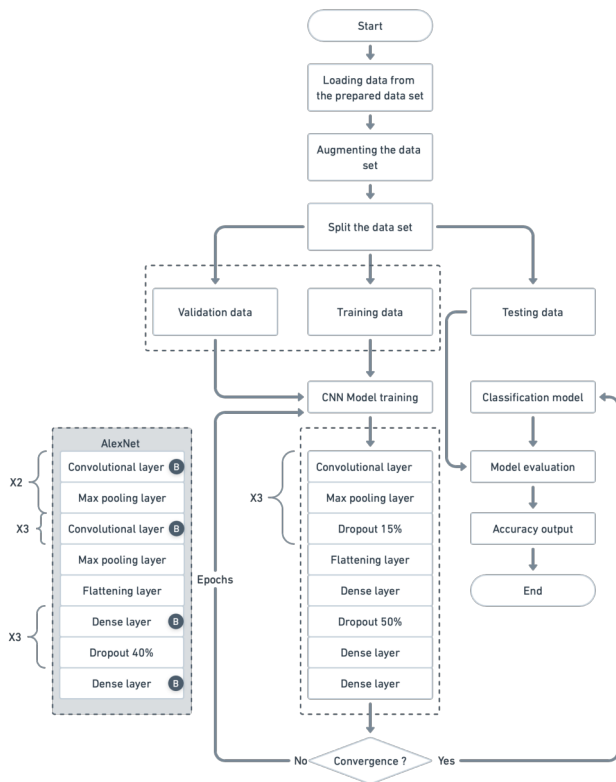
Fig. 1. Custom CNN Architecture.

Fig. 2. Flowchart of the custom framework.

TABLE III
CNN ARCHITECTURE RESULTS

Architecture	Epochs	Initial learning rate	Loss	Accuracy & F1
VGG16	10	0.001	0.3218	0.933
ResNext	12	0.001	0.1244	0.933
MobileNet	28	0.001	0.1141	0.950
Custom	22	0.000316	0.0989	**1.00**
AlexNet	28	0.001	0.0582	0.983
DenseNet121	12	0.001	**0.0254**	**1.00**

fast has been developed. Similarly, a pipeline for augmenting the training data and a pipeline for defining CNNs, training them, and providing a classification model as the output has been developed successfully. Also, to demonstrate the potential of the research, a user interface has been developed and deployed to a web server where the general public can accessed at tumorclass.info. All the source code used in this project can be found in the GitHub repository.

VI. DEPLOYMENT

A service called ngrok was selected for the deployment. Ngrok is a free service that allows the exposure of localhost to the internet in a secure manner. With ngrok installed, the only steps needed were to run the FastAPI in uvicorn on the local machine and then run ngrok in a separate terminal to expose the local host to the internet. To make it a little more elegant, a domain name was purchased, and the DNS was set to give access to ngrok. A monthly subscription on ngrok was also purchased to connect the autogenerated ngrok domain name to the custom domain name. In figures Fig. 3, Fig. 4, Fig. 5, and Fig. 6 the implemented user interface is shown.
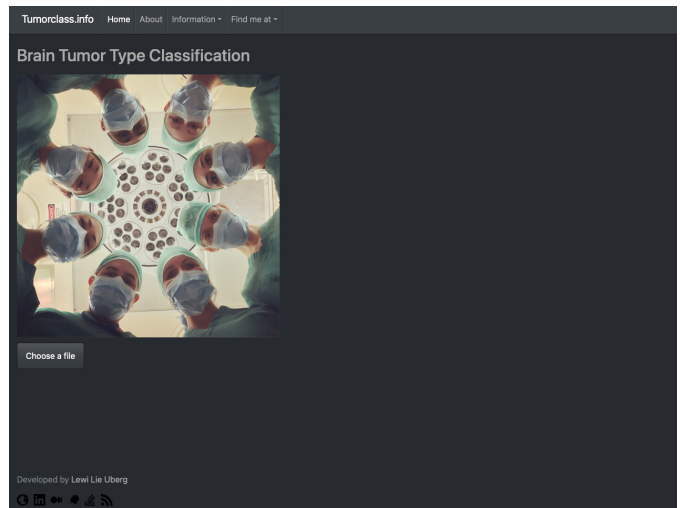

Fig. 3. Interface of our platform.

VII. CONCLUSION & FUTURE WORK

The next step after our CADD system is to expand the data set. The amount of suitable data available is limited. However, the REMBRANDT dataset has a vast number of images suitable for manual labeling performed by radiologists. Expanding the data set will make it more transparent, which architectures generalize well. As such, these architectures can be the focus of further development. These days, MRI scans provide 3D images, so further development to facilitate 3D classification is also a viable option.

REFERENCES

[1] D. N. Louis, H. Ohgaki, O. D. Wiestler, W. K. Cavenee, P. C. Burger, A. Jouvet, B. W. Scheithauer, and P. Kleihues, "The 2007 WHO classification of tumours of the central nervous system," *Acta neuropathologica*, vol. 114, no. 2, pp. 97,109, Aug. 2007.
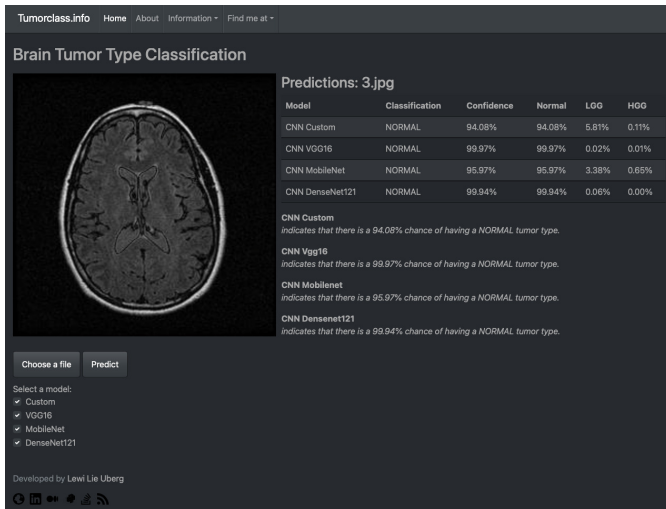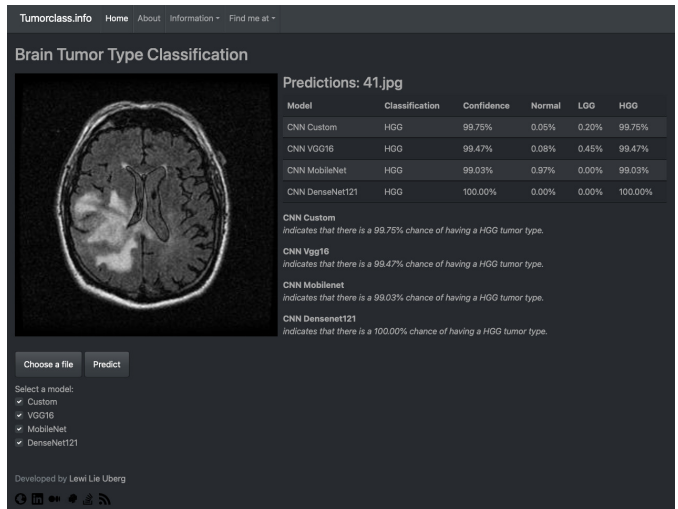
Fig. 4. Classification demo for normal MRI.

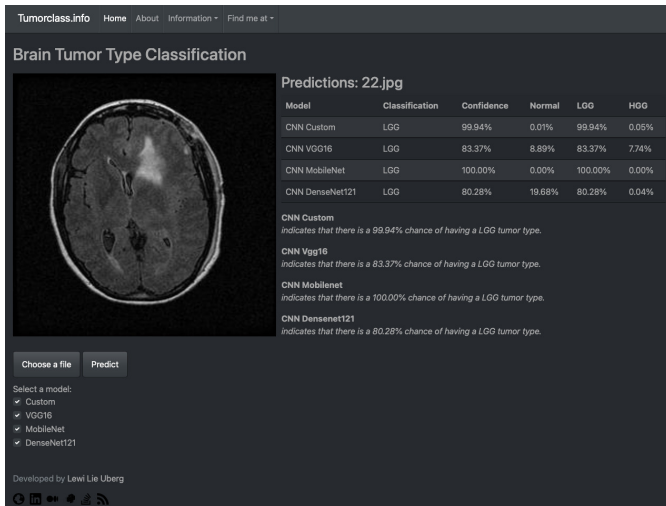Fig. 6. Classification demo for HGG MRI.

Fig. 5. Classification demo for LGG MRI.

[2] L. S. Hu, A. Hawkins-Daarud, L. Wang, J. Li, and K. R. Swanson, "Imaging of intratumoral heterogeneity in high-grade glioma," *Cancer letters*, vol. 477, pp. 97,106, May 2020. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/32112907

[3] D. A. Forst, B. V. Nahed, J. S. Loeffler, and T. T. Batchelor, "Low-grade gliomas," *The oncologist*, vol. 19, no. 4, pp. 403,413, Apr. 2014. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/24664484

[4] Q. Luo, Y. Li, L. Luo, and W. Diao, "Comparisons of the accuracy of radiation diagnostic modalities in brain tumor: A nonrandomized, nonexperimental, cross-sectional trial," *Medicine*, vol. 97, no. 31, Aug. 2018. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/30075495

[5] C.-X. Wu, G.-S. Lin, Z.-X. Lin, J.-D. Zhang, L. Chen, S.-Y. Liu, W.-L. Tang, X.-X. Qiu, and C.-F. Zhou., "Peritumoral edema on magnetic resonance imaging predicts a poor clinical outcome in malignant glioma," *Oncology Letters*, vol. 10, no. 5, pp. 2769,2776, Aug. 2015. [Online]. Available: https://www.spandidos-publications.com/10.3892/ol.2015.3639

[6] K. Fukushima, "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position," *Biological Cybernetics*, vol. 36, no. 4, pp. 193,202, Apr. 1980. [Online]. Available: https://link.springer.com/article/10.1007/BF00344251

[7] ——, "Cognitron: A self-organizing multilayered neural network,"

[8] *Biological Cybernetics*, vol. 20, no. 3, pp. 121–136, Sep. 1975. [Online]. Available: https://doi.org/10.1007/BF00342633

[8] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278,2324, Nov. 1998. [Online]. Available: http://yann.lecun.com/exdb/publis/pdf/lecun-98.pdf

[9] J. Kang, Z. Ullah, and J. Gwak, "MRI-Based Brain Tumor Classification Using Ensemble of Deep Features and Machine Learning Classifiers," *Sensors*, vol. 21, no. 6, Mar. 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/6/2222

[10] L. Scarpace, A. E. Flanders, R. Jain, T. Mikkelsen, and D. W. Andrews, "Data From REMBRANDT [Data set]," 2019. [Online]. Available: https://wiki.cancerimagingarchive.net/display/Public/REMBRANDT#35392299515cc672b974080a1394cbe9c649c74

[11] S. Khawaldeh, U. Pervaiz, A. Rafiq, and R. S. Alkhawaldeh, "Noninvasive Grading of Glioma Tumor Using Magnetic Resonance Imaging with Convolutional Neural Networks," *Applied Sciences*, vol. 8, no. 1, 2018. [Online]. Available: https://www.mdpi.com/2076-3417/8/1/27

[12] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in *International Conference on Learning Representations*, Sep. 2015. [Online]. Available: https://arxiv.org/pdf/1409.1556.pdf

[13] O. N. Belaid and M. Loudini, "Classification of Brain Tumor by Combination of Pre-Trained VGG16 CNN," *Journal of Information Technology Management*, vol. 12, no. 2, pp. 13,25, 2020. [Online]. Available: https://jitm.ut.ac.ir/article_75788.html

[14] O. Sevli, "Performance Comparison of Different Pre-Trained Deep Learning Models in Classifying Brain MRI Images / Beyin MR Görüntülerini Sınıflandırmada Farklı Önceden Eğitilmiş Derin Öğrenme Modellerinin Performans Karşılaştırması," *Acta Infologica*, vol. 5, p. 2021, Jun. 2021. [Online]. Available: https://cdn.istanbul.edu.tr/file/JTA6CLJ8T5/99DD9C496BF14E44859851B33E49A006

[15] A. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *ArXiv*, 04 2017. [Online]. Available: https://arxiv.org/pdf/1704.04861.pdf

[16] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, Jul. 2017, pp. 5987,5995. [Online]. Available: https://ieeexplore.ieee.org/document/8100117

[17] G. Huang, Z. Liu, G. Pleiss, L. Van Der Maaten, and K. Weinberger, "Convolutional Networks with Dense Connectivity," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1,1, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8721151