LITERATURE REVIEW

Reviewing Brain Tumour Diagnoses Using Deep Convolutional Neural Network

Anne Odeh[1*†] and Delano M Thomas[1,2]

Abstract

This paper summarizes and examines the accuracy of the news piece "A.I. Comes to the Operating Room" [4] and its reference journal [5] paying particular attention to the machine learning method – Convolution Neural Network (CNN), carried out in the research. It further showed how the technique in the journal is compared to other state-of-the-art neural networks applied in the field of medicine. This paper will describe improvements in CCN concerning other state-of-the-art methods in brain tumour diagnosis.

Keywords: brain tumour; convolutional neural network; deep learning; MRI

Introduction

Accurate diagnosis of diseases depends equally on their inputs and analytical procedure. This report will express this concept through image acquisition and interpretation. X-Ray, Computerized Tomography (C.T.), and Magnetic Resonance Imaging (MRI) scans have improved significantly over the years through the use of much higher resolution [1]. However, the benefits of automated image interpretation are on the rise as well. Due to the differences between patient data, the non-automated methods mentioned previously are not reliable. As a result, machine learning techniques are now employed to use pattern recognition on datasets to retrieve information from medical images for the identification of diseases and their treatments [2,3].

Summary of News Article

The news article "A.I. Comes to the Operating Room" [4], published on January 6, 2020, by the New York Times, reflects on how Artificial Intelligence (A.I.) and new imaging techniques can expedite the process of diagnosing brain tumours during surgery. The newfound method proved to be superior by detecting brain tumours within two and a half minutes compared to the conventional process, which takes up to 20 to 30 minutes. Another plus point for the new technique against the old is that the former can detect the spread of tumours along nerve fibres, all while preserving the tissue for reuse.

The study involves the collection of sample tissue from 278 participants, which was split evenly between the A.I. system and a group of neuropathologists – the conventional procedure for brain tumour diagnosis. Overall, both analyses predicted equally well at 93.9% for the neuropathologists and 94.6% for the A.I. However, the A.I. procedure misdiagnosed about 14 cases that the neuropathologists got right. The A.I. excellently passed another 17 instances that the neuropathologists missed.

This A.I. approach employed Stimulated Raman Histology (SRH) for image generation, which was captured and processed by a computer. A total of 415 multiple images from brain surgery patients acted as trainers for the algorithm using a Deep Neural Network methodology to classify ten common types of tumours. Neurosurgeon and Senior author of the report, Dr. Daniel A. Orringer, indicated that he had already started using the new technology during his surgeries to guide him when the operations are complete. Other surgeries can benefit from the revolutionary A.I. technique as well, including head and neck, breast, skin, and gynaecology surgery; even lung cancer detection using C.T. could benefit from this technology.

Summary of Journal article

The authors of the journal [5] aim to develop an alternative to the present state of intra-operative diagnosis during cancer surgery – a frozen section and smear preparation technique on haematoxylin and eosin tissue (H&E). The current method relies on an unbalanced contracting pathology workforce, which is time, resource, and labour-intensive. The proposed solution relies on SRH and CNN where both were used in cascade to automatically predict the diagnosis of a brain tumour in real-time.

*Correspondence: aaodeh@mun.ca
[1]Department of Computer Science, St John's, Elizabeth Avenue, A1C 5S7 Newfoundland and Labrador, Canada
Full list of author information is available at the end of the article
†Equal contributor

According to the journal, the conventional process first involves transporting the tissue to the laboratory and processing the specimen, then technicians will prepare the slides, and finally, the pathologist steps in to interpret the slides/images.

Unlike the conventional method, the A.I. solution's implementation involves a three-step interoperative tissue diagnosis pipeline: image acquisition, image processing, and diagnostic prediction using CNN. First, the sample tissue gets taken from the surgical field, and a small portion is compressed into the microscopic slide and inserted into the SRH imager to acquire the images. Secondly, these images are processed through an algorithm to generate high-resolution and high magnification patches for CNN training. At the prediction phase, the patches pass through the Inception-ResNet-v2 network for convolution neural network image classification.

The conventional method takes between 20-30 minutes for diagnosis, while the Artificial Intelligence procedure is under 150 seconds. The A.I. solution was trained to diagnose brain tumours using 2.5 million SRH images from 415 patients, and it learned to classify 13 significant classes of brain tumours.

In a multicentre clinical trial with 278 cases, the A.I. procedure performed equally well to the pathologist-based diagnosis. Out of all the tumours diagnosed, 93.9% (261/278) accuracy was achieved using the conventional method, while SRH with CCN accomplished 94.6% (264/278) accuracy. Each diagnosis type was conducted on the same sample split intraoperatively. Notably, nine of the fourteen A.I. errors were glial tumours, while ten of the seventeen misdiagnoses from the conventional approach were malignant gliomas. Interestingly, all incorrectly diagnosed cases from one test were correctly predicted by the other analysis method, indicating that CNN could assist pathologists in classifying challenging specimens.

The new system was trained to detect rare brain tumours using a Mahalanobis distance-based within a confidence scoring algorithm; this assisted with addressing the scarcity of training data for rare tumour diagnosis.

News and Journal Articles Assessment

Comparing both the news and the journal articles, the news material was incorrect in suggesting that the new procedure could be used in other surgeries like head and neck, breast, and gynaecology surgeries because it is not yet a possibility but a hope for future use. Denise Grady misrepresented the facts when she stated [4], "The researchers used images from tissue from 415 brain surgery patients to train an artificial intelligence system to identify the ten most common types of brain tumour". The artificial intelligence system was able to classify tissues into 13 categories. Aside from these points, the article was quite precise in bringing the information from the material to its readers.

Method of Implementation

The study had two main objectives [5]: (1) develop an intraoperative diagnostic Computer Vision System using clinical SRH and Convolution Neural Networks that will improve interpretation of fresh surgical specimen in near real-time and (2) to perform a multi-centre clinical test to verify diagnostic accuracy. CCN training data was gathered beginning mid-2015 from patients undergoing Central Nervous System (CNS) tumour resection surgery at Michigan University, New York-Presbyterian/Columbia University, and the University of Miami. The following are the essential steps carried out to develop this A.I. solution.

Stimulated Raman Histology (SRH) Capture

A clinical stimulated Raman scattering (SRS) microscope [5] was used to capture the tissue images via beam scanning with a spatial sampling of 450nm pixels at 1000 pixels per strip for the training data. However, image capturing for the testing phase was done with a clinical fibre-lased-based SRS microscope, a virtual H&E look-up table was applied to transform the raw SRS images into SRH images.

Image Preprocessing and Data Augmentation

The images from the SRS [5] were processed to generate a three-channel image that is passed through a 300x300 pixel sliding window algorithm with a 100-pixel step and valid padding, which created high-magnification image patches necessary for input to the CNN architecture. The images were further optimized and rescaled before they were reviewed and labelled according to their diagnosis. Oversampling was used to address class imbalance for underrepresented classes. These included affine transformation, random combination of rotation, shift, and reflection.

Image Datasets

One SRH microscope [5] was used to capture images from 296 patients, and a three fibre-laser-based SRS microscope captured 339 patients' images. Training and validation made use of both SRH and SRS images from 415 patients; however, the testing used only SRS images.

CNN Training

Thirteen diagnostic classes were chosen as input for the CNN architecture [5]; they represent the most common CNS tumours from which intraoperative decisions arrive. Some of the output labels are malignant glioma, diffuse midline glioma, diffuse lower-grade glioma, ependymoma, lymphoma, and pituitary adenoma. A 'normal' label was also supplied. Google's Inception-ResNet-v2 CNN architecture, with its 55.8 million trainable parameters, was fed 2.5 million unique patches for training, and a cross-entropy loss function was applied to the algorithm. Sixteen randomly selected validation sets were used for hyperparameter tuning, while model selection relied on result accuracy. Other training features implemented included the Adam optimizer, early stopping for regularization, python-based neural APIs, Kera (v.2.2.0) with TensorFlow (v.1.8.0), were used for training, validation, and testing.

Patient-level diagnosis and inference algorithm

The output from the SoftMax algorithm [5] in the CCN is summed elementwise and renormalized to produce a probability distribution. If the probability of the normal label were greater than 90%, then the output would be normal else the normal label probability was set to zero, and the probability distribution renormalized again. The final class decision now becomes the value of the expected renormalized distribution.

Mahalanobis distance-based confidence score

The training data included mostly common tumours. A confidence score, based on Mahalanobis [5] distance and a multivariate Gaussian Distribution under Gaussian Discriminant Analysis, was assigned to the output to go along with the posterior probabilities to cater for rare tumour classification.

Literature Review

History of Convolutional Neural Networks

The evolution of CNN started with Hubei and Wiesel [6] in the 1960s. They observed the brain activities in cats and discovered a hierarchical data processing pattern in the visual cortical pathway– simple cells detect location information while complex cells process information generated by simple cells. Hubei and Wiesel's [6, 7] research breakthrough led to the first implementation of CNN in the 1980s called "neocognitron" proposed by Fukushima and Miyake [6]. Neocognitron [6] breaks down a visual model into several sub-models and then processes them on progressively connected feature planes ensuring full recognition even if the object has displacement or slight deformation. The imperfections of "neocognitron" made many researchers dig deeper into perfecting CNN. In the 1990s, Yann Lecun [8], a computer scientist, presented the very first CNN model, which was able to recognize handwritten character. Yann combined the convolutional layer with the downsampling layer to develop LeNets5 [8], which outperformed all other techniques in identifying handwritten character tasks [6]. The primary application of CNN includes image recognition, natural language, and video recognition [9, 10].

The current state-of-the-art in Convolutional Neural Networks

CNN's state-of-the-art architecture is steadily expanding, and the quality of image classification has significantly improved by using extensive and wider networks. Krizhevsky et al. [11] demonstrated in their 2012 article that CNNs can be trained on 1.2 million images [12] to classify 1000 categories with high accuracy; as a result, CNNs have been applied to medical images. CNNs have had several astonishing successes in unravelling complex problems of machine learning and are currently considered the most effective method for image processing [13].

Early recognition of brain tumours using MRI and C.T. scans [14, 15, 16] is not enough for detection. Still, it is now achievable by applying the images (MRI and C.T. scans) to deep learning methods for image processing. Medical Image Processing improves the prior diagnosis of patients who survived with a brain tumour. In this article, we present various techniques employed in brain tumours classification based on CNN.

CNN with Genetic Algorithm

Kabir et al. [17] proposed using CNN with Genetic Algorithm (G.A.) so that the various grades of glioma brain tumours, using MRI images, could be categorized as non-invasive. To evolve the architecture of CNN, G.A. was used. Instead of training and comparing more than one million different architectures, by employing G.A. and comparing less than 500 architectures, a suitable one was discovered. Thus, reducing the computational costs.

In the best model developed through G.A., Bagging, as an ensemble algorithm, was applied to reduce the variance of prediction error, resulting in approximately 99% accuracy when classifying the three Glioma grades. Evaluations of the results indicated that the classification of brain tumours was expertly done.

Most of the datasets used in this research were obtained from four databases that are available online for research purposes. These are IXI, REMBRANDT, TCGA-GBM, and TCGA-LGG, and they include images of subjects with no lesions, and subjects with low and high gliomas. In addition, each also provides an MRI brain tumour dataset containing T1-weighted images from 233 patients with Meningioma, Glioma, and Pituitary brain tumour types.

VGG-19 CNN

Another study on brain tumour classification by Sajjad et al. [18] firstly suggested a deep learning technique to utilize segmenting of the tumorous regions from MRI images. The proposed system was efficiently trained using extensive data augmentation in the second step. Lastly, the brain tumour grade classification was done using the augmented data, and the pre-trained VGG-19 CNN model. Evaluations were performed on the original as well as augmented data, where the performance metrics of sensitivity, specificity, and accuracy showed significant improvement.

According to the authors [18], there were challenges in accessing MRI datasets on the internet, and this led to extensive use of data augmentation. They expanded the data using various parameters and different techniques to fill the data gaps and make the system transformational and noise invariant. The desired accuracy was achieved using eight different augmentation techniques, namely, rotation, flipping, skewness, shears, Gaussian blur, sharpening, edge detection, and emboss.

The proposed method employed uses VGG-19 CNN [18] architecture to fine-tune the brain tumour grade classification. VGG-19 architecture consists of 19 weighted layers, in which there are 16 convolutional and three fully connected layers. In VGG-19 architecture, the first two convolutional layers are followed by max-pooling, and a similar combination is repeated for the subsequent two layers. The next eight layers are designed as a combination of four convolutional layers trailed by max pooling. Further, the last three layers are fully connected, resulting in 4096, 4096, and 1000 features, respectively. The motivation behind using VGG-19 is that there are 3×3 kernels in all the convolutional layers with one stride; other CNN models ignore the critical patterns in the MRI. Plus, larger sized kernels increase the number of parameters required.

Novel CCN Architecture

In the same vein, Alqudah et al. [19] researched how to grade brain tumours using cropped tumour lesions and uncropped brain images. The cropped and uncropped images consist of three types of brain tumours: class glioma, meningioma, and pituitary tumours. The T1-Weighted MRI images are fed into a novel CNN architecture and trained to calculate the weight of networks. Following this, a confusion matrix for cropped, uncropped, and segmented cases is generated. The comparison between the CNN architecture outputs with the original image label was carried out based on the generated confusion matrices. The overall results showed that both cropped and uncropped have high accuracy, high sensitivity, and high specificity.

The dataset proposed in this method [19] is available for free online. It contains 3064 T1 weighted and contrast-enhanced brain MRI images and includes three classes – glioma, meningioma, and pituitary tumour. In this research, the cropped and uncropped brain tumour images were stored as a database, and three folders created; each one consists of the images for a specific class. The database was partitioned into 70% data for training and the rest for testing. The proposed CNN classifier is a powerful tool, and its overall performance has an accuracy of 98.93% and a sensitivity of 98.18% for the cropped lesions. In comparison, the results for the uncropped lesions are 99% for accuracy and 98.52% for sensitivity, and the results for segmented lesion images are 97.62% for accuracy and 97.40% for sensitivity.

U-Net CNN

Mlynarski et al. [25] proposed a CNN-based segmentation model that was trained using weakly annotated images in addition to fully annotated images. The authors were able to achieve their goal by exploiting the representation learning ability of CNNs to learn from weakly annotated images while supervising the training using fully annotated images to learn features relevant for the segmentation task.

The dataset [25] was obtained from the BRATS 2018 database, and it contains 285 multi-sequence MRI images of patients diagnosed with low-grade gliomas or high-grade gliomas. The images were obtained from 19 different imaging centres. To normalize the intensities in each image, they were divided by a median of nonzero voxels and multiplied by a fixed constant. The model aims to take advantage of all available voxel-wise and image-level annotations.

The proposed system [25] takes as input a multi-modal image of dimensions 300 × 300. It extends U-Net, which is, according to the authors, "one of the most used architectures for segmentation tasks in medical imaging." The different image modalities correspond to channels of the data layer and are the input of the first convolutional layer of the network. The U-Net is composed of a contraction part and a symmetric expanding part connected by concatenation, which are connected by concatenations between layers at the same scale, to combine low-level and local features with high-level and global features. The authors believe "the design is well suited for the tumour segmentation tasks since the classification of a voxel as tumour requires the comparison of its value with its close neighbourhood but also considering a large spatial context". The output layer generates classification scores for each pixel which are fed into Softmax for normalization. The final layer of the U-Net produces pixel-wise classification scores that are then normalized by the SoftMax function during the training phase. "Batch normalization was applied in all convolutional layers except the final layer."

Improvements in current CNN methods in comparison to other CNN methods

Several attempts at brain tumour classification have been made using mixed CNN solutions [20, 23, 24, 25], from ensemble models to single CNN. Generally, the input to the CNN architecture was MRI. Once completed, the CNN output fed to another ML algorithm to improve classification. The output classification was binary, indicating the tumour was malignant or benign in most cases. Although the performance of the mixed techniques is better in terms of accuracy, their binary output is more of underperformance when compared to the method in the news article. However, one of the improvements in the technique employed in the news article is that it not only diagnoses the three types of brain tumours (glioma, meningioma, and pituitary) like the other method, but it does much more by detecting 13 different brain tumours. The technique is quicker and faster, and the results are within 150s [5].

The CNN procedures [17, 18, 19, 25] shared above, when compared to the technique from the journal, can be considered juvenile, and their classification is limited to very few classes. Also, they rely heavily on old stored data that was tested many times. No tests can ever be 100% accurate, and if compared to other tests on the same data, the mechanism did not offer much overall improvement.

The procedure from the journal is phenomenal and relied on new technologies – the input data was improved to RGB photos, unlike the other processes that used shades of black and white through MRI or Xray. Also, it was able to predict rare tumours and not just the common ones. Overall, the journal technique is world-leading compared to the other procedures reviewed.

Conclusion

Considering the innovative transformation in deep learning, CNN has achieved great success in the field of computer vision [20], which was inspired by Fukushima's biological structure of the visual cortex [6]. CNNs are neural networks with at least one hidden convolutional layer between the input and output layers. They have non-linear properties and can extract higher-level representative features [21]. A combination of deep learning methods with CNN has shown excellent results on a wide range of medical imaging tasks, such as brain tumour classifications [17, 18, 19, 22].

The authors [5] were able to demonstrate how A.I. can be resourceful in the surgical room through near real-time diagnoses of brain tumours, all while preserving the tissue's image integrity for downstream analysis. The advent of this technology can save lots of lives because it will aid surgeons in removing just the right amount of tumorous tissue from the diagnosed patient. The authors hope the A.I. algorithm can further be developed to predict key molecular alterations in brain tumours. It was suggested that deep learning techniques could similarly be applied to other medical areas, namely dermatology, head and neck surgery, as well as breast and gynaecology surgeries.

Author details

[1]Department of Computer Science, St John's, Elizabeth Avenue, A1C 5S7 Newfoundland and Labrador, Canada. [2]Department of Computer Science, St John's, Elizabeth Avenue, A1C 5S7 Newfoundland and Labrador, Canada.

References

1. Yadav, S., Jadhav, S.: Deep convolutional neural network-based medical image classification for disease diagnosis. Journal of Big Data **6**, 1–30 (2019)
2. Muhammad, I.R., Saeeda, N., Ahmad, Z.: Deep Learning for Medical Image Processing: Overview, Challenges, and Future. Lecture Notes in Computational Vision and Biomechanics **26**, 1–10 (2018)
3. Yamashita, R., Nishio, M., Togashi, K.: Convolutional neural networks: an overview and application in radiology. Insights into Imaging **9**, 611–629 (2018)
4. A.I. comes to the operating room https://www.nytimes.com/2020/01/06/health/artificial-intelligence-brain-cancer.html
5. Todd, C.H., Balaji, P., Arjun, R.A., Esteban, V.S., Akshay, M., Siri, S.K., Daniel, G.E., Randy, S.D., Zia, U.F., Spencer, L., Petros, D.P., Tamara, M., Shah, A.H., Garton, H.J., Maher, C.O., Heth, J.A., McKean, E.L., Sullivan, S.E., Hervey-Jumper, S.L., Patil, P.G., Thompson, B.G., Sagher, O., McKhann, G.M., Komotar, R.J., Ivan, M.E., Snuderl, M., Otten, M.L., Johnson, T.D., Sisti, M.B., Bruce, J.N., Muraszko, K.M., Trautman, J., Freudiger, C.W., Canoll, P., Lee, H., Camelo-Piragua, S., Orringer, D.A.: Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. Nature Medicine **26**, 52–58 (2020)
6. Wang, W., Xin, W., Weizheng, W., Yang, Y.: Development of convolutional neural network and its application in image classification: a survey. Optical Engineering **58**, 1–19 (2019)
7. He, K., Zhang, X., Ren, S., Sun, J.: Deep Residual Learning for Image Recognition. 1–12 (2015)
8. Ge, D., Yao, X., Xiang, W., Wen, X., Lui, E.: Design of High Accuracy Detector for MNIST Handwritten Digit Recognition Based on Convolutional Neural Network. 12th International Conference on Intelligent Computation Technology and Automation (ICICTA) **58**, 658–662 (2019)
9. Singh, M.K., Baluja, G.S., Sahu, D.P.: Understanding the Convolutional Neural Network & it's Research Aspects in Deep Learning. International Journal for Research in Applied Science & Engineering Technology **55**, 867–887 (2019)
10. Suganuma, M., Shirakawa, S., Nagao, T.: Genetic Programming Approach to Designing Convolutional Neural Network Architectures. Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence 5369–5373 (2018)
11. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. Communications of the ACM **60**, 84–90 (2017)

12. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A., Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision **115**, 211–252 (2015)
13. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. Nature **521**, 436–444 (2015)
14. Ranjeet, K., Amit, D.: Localization, and Classification of Brain Tumor using Machine Learning & Deep Learning Techniques. International Journal of Innovative Technology and Exploring Engineering **8**, 60–66 (2019)
15. Ngo, L.: Using a Deep Learning Network to Diagnose Congestive Heart Failure. Radiology **290**, 523–524 (2019)
16. Sobhaninia, Z., Rezaei, S., Noroozi, A., Ahmadi, M., Zarrabi, H., Karimi, N., Emami, A., Samavi, S.: Brain Tumor Segmentation Using Deep Learning by Type-Specific Sorting of Images. Department of Electrical and Computer Engineering, Isfahan University of Technology 1–4 (2018)
17. Kabir, A.A., Ayati, M., Kazemi, F.: Magnetic resonance imaging-based brain tumor grades classification and grading via convolutional neural networks and genetic algorithms. Biocybernetics and Biomedical Engineering **39**, 63–74 (2019)
18. Sajjad, M., Khan, S., Muhammad, K., Wu, W., Ullah, A., Baik, S.: Multi-grade brain tumor classification using deep CNN with extensive data augmentation. Journal of Computational Science **30**, 174–182 (2019)
19. Alqudah, A., Alquraan, H., Qasmieh, I.A., Alqudah, A., Al-Shar, W.: Brain Tumor Classification Using Deep Learning Technique - A Comparison between Cropped, Uncropped, and Segmented Lesion Images with Different Sizes. International Journal of Advanced Trends in Computer Science and Engineering **8**, 3684–3691 (2019)
20. Sun, L., Zhang, S., Chen, H., Luo, L.: Brain Tumor Segmentation and Survival Prediction Using Multimodal MRI Scans with Deep Learning. Frontiers in Neuroscience **13**, 1–9 (2019)
21. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Lui, T., Wang, X., Wang, G., Cai, J., Chen, T.: Recent advances in convolutional neural networks. Pattern Recognition **77**, 354–377 (2018)
22. Agravat, R.R., Raval, M.S.: Brain Tumor Segmentation and Survival Prediction. 1–9 (2019)
23. Ari, A., Hanbay, D.: Deep learning based brain tumor classification and detection system. TURKISH JOURNAL OF ELECTRICAL ENGINEERING & COMPUTER SCIENCES **26**, 2275–2286 (2018)
24. Sultan, H., Salem, N., Al-Atabany, W.: Multi-Classification of Brain Tumor Images Using Deep Neural Network. IEEE Access **7**, 69215–69225 (2019)
25. Mlynarski, P., Delingette, H., Criminisi, A., Ayache, N.: Deep learning with mixed supervision for brain tumor segmentation. Journal of Medical Imaging **6**, 1–13 (2019)