JOURNAL OF
EMERGING INVESTIGATORS

Article

The Utilization of Artificial Intelligence in Enabling the Early Detection of Brain Tumors

Shanzeh Haider[1], Odysseas Drosis[2]

[1] The American International School of Johannesburg, Midrand, South Africa

[2] Computer Science, École polytechnique fédérale de Lausanne, Lausanne, Switzerland

SUMMARY

Diagnosing brain tumors is challenging due to their location and varied presentations, which may mimic common disorders. A cancer diagnosis can be missed even when advanced imaging is conducted due to interpretive error or an incompatible clinical history as presented. Machine learning, when applied to radiological imaging, can aid in alerting physicians of the presence of tumors and improve diagnostic evaluation. Enhanced evaluation of malignant tumors can lead to earlier detection and positively improve prognosis, quality of life, and treatment. We aimed to enhance brain tumor diagnosis using machine learning. In this study, we developed two machine learning models: a logistic regression model and a neural network model. We hypothesized that, while both of our techniques would demonstrate a high diagnostic accuracy, the neural network model would produce more successful results due to its greater complexity. Applying a dataset sourced from Kaggle, an online data science resource, into the algorithms demonstrated with test accuracies of 68% in the logistic regression model and 84% in the neural network model. Overall, the models suggested a promising future for machine learning applications to brain tumor diagnoses.

INTRODUCTION

An intracranial (brain) tumor is an abnormal mass of tissue caused by uncontrolled proliferation that is not regulated by innate control mechanisms (1). According to Johns Hopkins Medicine, 30 out of 100,000 American adults develop a brain tumor (2). The National Cancer Institute estimates that there will be 25,400 new brain tumor cases in 2024, and that 18,760 of these cases will result in death (NEW). Magnetic resonance imaging (MRI) is a vital tool in diagnosing tumors. MRI scanning is a non-invasive medical imaging procedure that produces precise images of the human body's organs and tissues, such as the brain, using a magnetic field and radio waves (4). The images generated can be used together with the patient's clinical history and presentation to diagnose a tumor.

Misdiagnosis in the primary stages of brain tumors can occur due to misleading symptoms and lack of diagnostic MRI imaging. The symptoms of brain tumors may mimic other disease presentations, including Alzheimer's disease, migraines, Lyme disease, meningitis, and Multiple Sclerosis (5). Therefore, a physician may not detect the tumor within the imaging scan or may not even order imaging given the patient's clinical presentation. This can lead to delayed diagnosis and treatment, worsening survival rates.

There is an urgent need to develop and evaluate modalities that can facilitate earlier diagnosis of intracranial tumors. Artificial intelligence (AI) classifying brain tumors presents an opportunity to address this problem. Machine learning uses computers to perform tasks that usually call for human intelligence. Machine learning is a subfield of AI that can analyze enormous volumes of data quickly by identifying patterns. Machine learning algorithms include neural networks and logistic regression models (6). Neural networks feature interconnected nodes that layer to approximate the function of brain neurons (**Figure 1**). Neural networks can cluster and categorize raw data algorithmically, identify hidden patterns and correlations, and continually learn and improve with time (7). Logistic regression is a classification algorithm that uses mathematical concepts, such as logistic functions, to make predictions (**Figure 2**). Based on a given dataset of independent factors, logistic regression calculates the likelihood that an event will occur (6). Logistic regression allows computers to demonstrate intelligence in terms of analytical ability and deduction. Combining machine learning capabilities with physiological imaging can enhance the detection of characteristics that define tumors and enable more accurate and precise diagnosis (8).

Previous studies have demonstrated the potential application of AI in diagnosing intracranial tumors and other medical conditions. A study by Amin Ul Haq et al. proposed a machine learning model using convolutional neural networks to improve the accuracy of brain tumor detection by AI, with an accuracy of 99.90% (9). In another recent study by the NIH, machine learning aided in diagnosing breast cancer using mammograms and neural networks and suggested

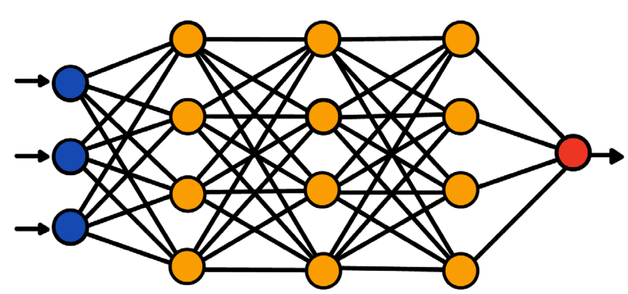

Figure 1. Neural network schematic. (The blue circles represent input nodes, the orange circles represent hidden layer nodes and the red circle represents the output node. The arrows represent the direction. The lines represent edges).

https://doi.org/10.59720/23-098

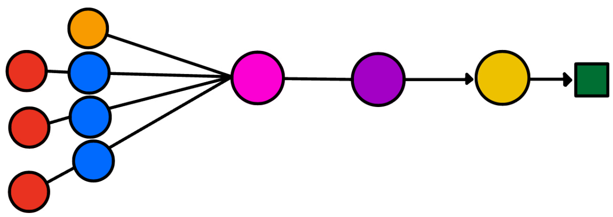

Figure 2. Linear regression schematic. (The red circles represent input nodes, the orange circle represents a desirable bias node, and the blue circles represent weights. The pink, purple, and yellow circles all represent various functions which are performed by the model. The green square represents the predicted class of 0 or 1. The lines and arrow represent the direction of the model.)

further research may accurately diagnose cancer using the scanning technology available (10). Another used two AI models to diagnose 3 specific types of brain tumors from 3264 MRI scans and achieved 96.47% and 95.63% train accuracy using a 2D convolutional neural network and an auto-encoder network, respectively (11). These findings emphasized the potential role of AI in medical imaging detection, as the model was able to detect tumor presence and differentiate between tumor types.

We utilized a logistic regression and a neural network model to improve early brain tumor detection and expand on other research that has been conducted in this area. The dataset used in our model consisted of brain tumor MRI images and was found on Kaggle, an online data science resource (12). We used two common machine learning models, neural networks and logistic regression, to explore how the models differ in their capacity to detect brain tumors. Our primary goal was to evaluate the two models to determine what approach would be best used in the field's ongoing research phase. We hypothesized that while both models would produce high detection accuracy rates, the neural network model would achieve better accuracy than the logistic regression model. We hypothesized this because neural network models have more layers, nodes, and weights than linear regression models, allowing them to interact more precisely with image prediction data (13). In this study, we used an MRI Dataset, and created neural networks and logistic regression models (12). We experimented with the correlation between the number of hidden layers with model performance for the neural networks model, to obtain the highest possible accuracy. We found the performance of the models supported our hypothesis, with the neural network obtaining roughly 15% higher accuracy rates than the linear regression model.

RESULTS

To investigate the use of machine learning for brain tumor detection, we created neural networks and logistic regression models using an MRI dataset from Kaggle (12). The MRI dataset consists of 255 brain MRI images, of which 99 were consistent with no tumor, whereas 156 were consistent with a brain tumor (**Figure 3**). The images ranged in dimension from 232 x 309 pixels to 512 x 512 pixels (12). The stages of the tumors shown in the images were unknown.

Since there were many images in the dataset that had different dimensions, we developed a low-memory algorithm that was trained fast to make predictions. This low memory algorithm was desirable as it allowed the models to process large amounts of images in a short amount of time. We resized the images, split them into training and testing datasets, then used these to respectively train and test the models. Even with a simple model, we obtained a test accuracy of 68% for the logistic regression model and 84% on the neural network model (**Figure 4**). For the neural networks model, we altered the number of hidden layers to experiment if we could obtain a higher test accuracy. The test accuracy for 8 hidden layers is 84.20%, whereas the test accuracy for 10 hidden layers is 82.60% (**Figure 5**). Thus, we decided to keep our model with 8 hidden layers due to the correlation between hidden layers and test accuracy that is presented in Figure 5. The train accuracy was 70% for the logistic regression model and 95% for the neural network model (**Figure 4**).

The confusion matrices, alongside the positive predictive values (PPV) and false discovery rates (FDR), provided more details about the performance of both models. Of 255 MRI scans (156 cancerous and 99 non-cancerous), our logistic regression model correctly predicted 106 as cancerous, incorrectly predicted 15 as cancerous, correctly predicted 84 as non-cancerous, and incorrectly predicted 50 as non-cancerous (**Table 1**). The PPV was 87.6%, meaning that out of all the cancer positive predictions made by the model, 87.6% were actually cancerous. The FDR was 12.4%, showing that among all scans predicted as having cancer, approximately 12.4% were non-cancerous. Furthermore, of 255 MRI scans (156 cancerous and 99 non-cancerous), our neural networks model correctly predicted 131 as cancerous, incorrectly predicted 8 as cancerous, correctly predicted 91 as non-cancerous, and incorrectly predicted 25 as non-cancerous (**Table 2**). The PPV for the neural networks model was 94%. Furthermore, the FDR was 5.76%.

DISCUSSION

Our study aimed to understand the application of artificial intelligence in brain tumor detection, through the creation of logistic regression and neural networks models based on an MRI image dataset. For the logistic regression model, we obtained a train accuracy of 70% and a test accuracy of 68%. For the neural network model, we obtained a train accuracy of 95%, and a test accuracy of 84%. These results supported

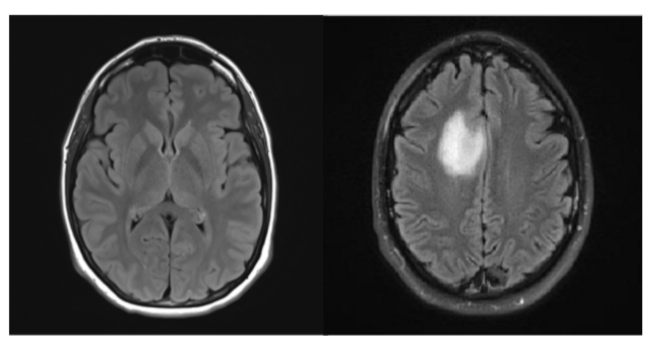

Figure 3. Sample MRI images without (left) and with (right) brain tumors. Scans were sourced from Kaggle (10).

https://doi.org/10.59720/23-098

Total MRI Scans 156 + 99 = 255		Predicted Condition	
		Cancer 139	Non-Cancer 116
Actual Condition	Cancer 156	131	25
	Non-Cancer 99	8	91

Table 1: Confusion Matrix for the Logistic Regression Model. Dataset was sourced from Kaggle (10). Predicted Condition refers to the outcomes predicted by the model; Actual Condition refers to the outcomes supported by the dataset. Green boxes show correct predictions by the model; red boxes show incorrect predictions by the model. Positive Predictive Value was 87.6%; False Discovery Rate was 12.4%.

our hypothesis as the neural network achieved greater accuracies than the logistic regression model We attributed this finding to the overall higher number of parameters in this model, which led it to more successful predictions (14). Additionally, the neural networks model obtained a higher PPV than the logistic regression model, showing that the neural networks model is better at identifying cancerous brain tumor scans. However, the overall performance of the logistic regression model was impressive compared to the typical performance of these models, which is often low due to their simple structure (15). Nonetheless, it may be possible to further improve both models in accuracy measurements by implementing further code and parameters.

Early in our workflow, we resized the scans to be the same dimensions to allow the models to train faster and more accurately. However, resizing some scans also could have reduced the quality of the images, which may have decreased the accuracy of both the logistic regression and the neural network models due to increased blur in the scans (13). In addition, the size of the images we sourced from the dataset varied, and the resizing resulted in disproportionate dimensions for some images, with lower quality for the already smaller images. The resizing method may have caused part of the difference between train and test set percentages for both models. Removing images that were distorted by resizing may be a way to overcome this issue.Furthermore, we were unable to verify whether the images from the dataset were early stage or late stage tumor images. Thus, some conclusions may not necessarily be applicable specifically to earlier diagnosis and treatment of brain tumors. This limitation could be addressed in a further study with training and test data sets specific to early stage and late stage tumors.

The accuracies and confusion matrices for each model help us understand the real-world applications of this study. Since we did not perform a statistical analysis or cross validation on the values we attained, we based our comparison and analysis of the data on other studies within the artificial intelligence field. The results are promising when compared to the study led by Amin Ul Haq, which maintained a 99.90% accuracy (9). The train accuracy of our neural network model (95%) was only 4.9% less from the accuracy found by Amin Ul Haq, which demonstrates the potential of our model. The lack of cross validation and statistical analysis is a limitation, however, it can be addressed in a further investigation through the implementation of such techniques.

When considering the reliability of data collected from machine learning models, it is imperative to evaluate the differences between the percentage accuracy for train and test sets in both models, along with percentage accuracies (13). These differences show the variation in how the model reacts to new data compared to data it has seen before in training, and it is optimal to obtain a higher test accuracy than train accuracy (14). For example, the test accuracy being lower than the train accuracy for the neural network model may be another source of error in how the model was trained: overfitting (16). The neural network model performed well on the training data but not as well on the data from the test set, indicating that it may have overfit the training set. This means

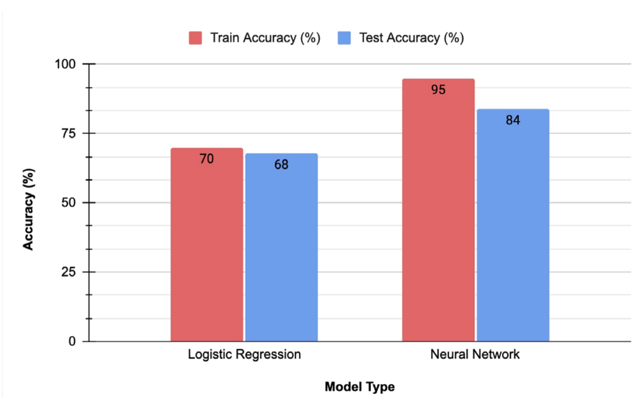

Figure 4. Percent Accuracies for logistic regression and neural network models. 255 MRI images were used in collaboration with both a logistic regression and neural networks model to understand the application of machine learning to brain tumor detection. The outputs were brain tumor positive or negative. The train set consisted of 75% of the images and the test set consisted of 25%. This data was collected through running the codes for both models and reading output values.

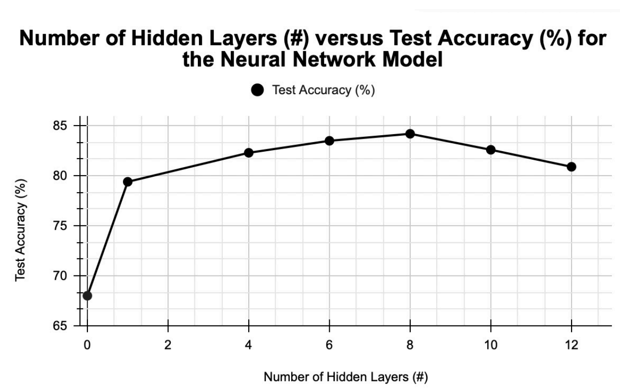

Figure 5. Line Graph Showing the Relationship Between the Number of Hidden Layers and the Percent Test Accuracy for the Neural Network Model. Dataset used for the model was sourced from Kaggle. Varying numbers of hidden layers were added to the neural network model, resulting in gradual changes to model test accuracy. Overfitting was noted after 8 hidden layers.

https://doi.org/10.59720/23-098

Total MRI Scans 156 + 99 = 255		Predicted Condition	
		Cancer 139	Non-Cancer 116
Actual Condition	Cancer 156	131	25
	Non-Cancer 99	8	91

Table 2: Confusion Matrix for Neural Networks Model. Dataset was sourced from Kaggle (10). Predicted Condition refers to the outcomes predicted by the model; Actual Condition refers to the outcomes supported by the dataset. Green boxes show correct predictions by the model; red boxes show incorrect predictions by the model. Positive Predictive Value was 94%; False Discovery Rate was 5.76%.

that the model could have memorized data it previously saw and was unable to generalize to unknown examples identified in the test set. But, this is not conclusive in terms of the model, as the train and test accuracies are relatively close in value, and further steps including retraining the model under different parameters could be used to assess this further. Additionally, the decrease in test accuracy between 8 and 10 hidden layers in the neural networks model may be indicative of overfitting. Potential future experiments could include developing neural network models which explicitly combat overfitting, for instance using a larger dataset.

This overfitting speculation does not apply to the logistic regression model due to the fact that the 2% difference within the model's train and test accuracy did not serve as a noticeable difference, which is the most apparent symptom of overfitting. For the neural network model, an 11% difference was calculated, which leaves open the possibility of overfitting. However, it is crucial to understand that overfitting is not the only explanation, but only the most common one for larger differences in percent accuracy. The larger difference in the neural network accuracies is possibly because of the higher complexity level of the model, or the fact that deep neural networks are highly sensitive to small changes in their input (14). In fact, in neural networks, an 11% difference in train and test accuracy is fairly small (14). The fact that the model was able to classify 84% of the test data correctly suggests that it is learning the general patterns produced by tumors, which is not suggestive of overfitting. Furthermore, by modifying the training epochs, we confirmed that the current number was most appropriate and avoided overfitting in the best possible manner. By manipulating the parameters, we were able to obtain a high accuracy rate. The model overall proved to work better than the logistic regression model, as shown in the results section.

To further this investigation a follow-up study could examine how the model performs on images from other scanners (which may have different strengths and capabilities), and having scans from different imaging techniques in the dataset. It would be interesting to find MRI scans of proportional dimensions, which would eliminate the resizing error that we encountered. We could also attempt to solve the potential overfitting in the models by utilizing validation techniques for a statistical analysis of the data, or fitting the model with different proportions of images in the train and test datasets. Pragmatically, we could implement the model onto a phone or create a device specifically for this model. This would be useful on a global scale in both urban and rural areas, as telephones are a common technology. Thus, it may help reduce geographical healthcare disparities, as individuals and communities worldwide could access this model.

MATERIALS AND METHODS

MRI Image Dataset

The dataset used in this research project was sourced from Kaggle, an online platform with datasets that are usable for machine learning. The title of the dataset is "Brain MRI Images for Brain Tumor Detection", last updated in 2019. The contributor who created the dataset, Navoneel Chakrabarty, holds an Engineering Doctorate at Eindhoven University of Technology and is based in the Netherlands. Chakrabarty stated in writing on Kaggle that the dataset's images have been compiled from various sites but did not give further information (12). The MRI dataset consisted of 255 images of the human brain, with some containing tumors and others without tumors (12). The dimensions of the images in the dataset varied, with a range from 232 x 309 pixels to 512 x 512 pixels. Of the 255 images, 156 images were categorized as consistent with a definitive diagnosis of an intracranial tumor, while 99 images were consistent with no tumor.

Data processing

To run and create code, we used PyCharm software (17). Using the Python computer language within the software, we developed two models to predict the train and test accuracies and understand how each model worked. The package used for both models was Scikit-learn. One model utilized logistic regression, while the other utilized neural networks. These models are very different in many ways, including their complexity, non-linearity, and interpretability (18). For both models, the images in the dataset were resized to be of uniform dimensions of 400 x 400 pixels, which is essential for training. Next, we split the data set into train and test sets. The train set had a higher percentage of images from the dataset (75%) than the test set (25%) so the model had more data to learn from. We used the train set to develop the model and the test set to test the model. This split created an unbiased estimator and helped us understand how well the learned model will generalize to unseen data.

Logistic Regression Model

For the model using this algorithm, the pixels that compose each image have weights and numbers associated with their position and order. The purpose of the model was to find the weight of these pixels, and then multiply them with that pixel's respective number. We then had the model perform a summation of these products, which was effectively as many terms as pixels. It was encoded into the algorithm that if the summation is positive it is 1 and if not it is 0. These two binary values helped determine whether the image presented a tumor or not.

Neural Networks Model

Our second model used a neural network. We implemented the VGG16 convolutional neural network, which has 16

https://doi.org/10.59720/23-098

layers, totaling around 138 million parameters (19). On top of the VGG16, we added another customized layer with 100 nodes to increase the capacity of the model. This information was used to determine if the images contained brain tumors or not. The code's parameters were continuously altered to improve the accuracy and optimize the model. In specific, we tested how the accuracies may change when adding different numbers of hidden layers. For example, adding one hidden layer to the current model did not affect the train and test performance. We experimented with the correlation between the test accuracy and the number of hidden layers. We chose the number of hidden layers for our model based on which amount would result in the highest test accuracy without overfitting.

ACKNOWLEDGMENTS

We would like to recognize the Inspirit AI organization as a contributor to the success of this project. Inspirit AI connected Shanzeh and Odysseas to pursue the research presented in this paper, and provided additional support where needed.

Received: April 19, 2023
Accepted: May 11, 2024
Published: February 5, 2025

REFERENCES

1. "Brain Tumors - Classifications, Symptoms, Diagnosis and Treatments." American Association of Neurological Surgeons. www.aans.org/en/Patients/Brain-Tumors. Accessed 2 Mar. 2023.
2. "Brain Tumors and Brain Cancer." *Johns Hopkins Medicine*, www.hopkinsmedicine.org/health/conditions-and-diseases/brain-tumor. Accessed 2 Sept. 2024.
3. "Cancer Stat Facts: Brain and Other Nervous System Cancer." *National Cancer Institute*, edited by Surveillance, Epidemiology, and End Results Program, seer.cancer.gov/statfacts/html/brain.html#risk. Accessed 2 Sept. 2024.
4. "MRI." Mayo Clinic. www.mayoclinic.org/tests-procedures/mri/about/pac-20384768. Accessed 2 Mar. 2023.
5. "Exploring Brain Tumours and Brain Tumour Misdiagnosis." Hatch Brenner. www.hatchbrenner.co.uk/news/exploring-brain-tumours-and-brain-tumour-misdiagnosis. Accessed 4 Mar. 2023.
6. "What Is Logistic Regression?." *IBM*. www.ibm.com/topics/logistic-regression. Accessed 10 Mar. 2023.
7. "Neural Networks: What Are They and Why Do They Matter?" SAS Analytics. www.sas.com/en_us/insights/analytics/neural-networks.html. Accessed 20 Mar. 2023.
8. Cè, Maurizio, et al. "Artificial Intelligence in Brain Tumor Imaging: A Step toward Personalized Medicine." *Current Oncology* vol. 30.3 2673-270122, Feb. 2023, https://doi.org/10.3390/curroncol30030203.
9. Haq, Amin Ul, et al. "DACBT: deep learning approach for classification of brain tumors using MRI data in IoT healthcare environment." *Sci Rep* 12, 2022, https://doi.org/10.1038/s41598-022-19465-1.
10. Dileep, Gayathri and Sanjeev G Gianchandani Gyani. "Artificial Intelligence in Breast Cancer Screening and Diagnosis." Cureus vol. 14, Oct. 2022, https://doi.org/10.7759/cureus.30318.
11. Saeedi, Soheila, et al. "MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques." BMC Medical Informatics and Decision Making, vol. 23, no. 1, 23 Jan. 2023, https://doi.org/10.1186/s12911-023-02114-6.
12. Chakrabarty, Navoneel. "Brain MRI Images for Brain Tumor Detection." *Kaggle*. www.kaggle.com/datasets/navoneel/brain-mri-images-for-brain-tumor-detection. Accessed 2 Jan. 2023.
13. Saponara, Sergio, and Alessandro De Gloria. "Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance." *Applications in Electronics Pervading Industry, Environment and Society*, 2022.
14. "How to Know If Your Machine Learning Model Has Good Performance - Obviously AI." *Obviously.ai*. www.obviously.ai/post/machine-learning-model-performance. Accessed 22 Mar. 2023.
15. "Evaluating a Linear Regression Model." *Ritchieng*. www.ritchieng.com/machine-learning-evaluate-linear-regression-model. Accessed 22 Mar. 2023.
16. Varoquaux, Gaël and Veronika Cheplygina. "Machine Learning for Medical Imaging: Methodological Failures and Recommendations for the Future." *Npj Digital Medicine*, vol. 5, no. 1, 12 Apr. 2022, https://doi.org/10.1038/s41746-022-00592-y.
17. JetBrains. *PyCharm*.
18. Dreiseitl, Stephan, and Lucila Ohno-Machado. "Logistic Regression and Artificial Neural Network Classification Models: A Methodology Review." J*ournal of Biomedical Informatics*, vol. 35, no. 5-6, Oct. 2002, pp. 352–359. https://doi.org/10.1016/s1532-0464(03)00034-0.
19. "Understanding VGG16: Concepts, Architecture, and Performance." *Datagen*, datagen.tech/guides/computer-vision/vgg16/. Accessed 21 Mar. 2023.

Copyright: © 2025 Haider and Drosis. All JEI articles are distributed under the attribution non-commercial, no derivative license (http://creativecommons.org/licenses/by-nc-nd/4.0/). This means that anyone is free to share, copy and distribute an unaltered article for non-commercial purposes provided the original author and source is credited.