

Enhancing Brain Stroke Detection: A Novel Deep Neural Network with Weighted Binary Cross Entropy Training



Ahmed Nidham Qasim^{1*}, Sameer Alani², Sarmad Nozad Mahmood¹, Swash Sami Mohammed¹, Dlnya Abdulahad Aziz¹, Karimeh Ibrahim Mohammad Ata³

¹ Electronic and Control Engineering Department, Technical Engineering College – Kirkuk, Northern Technical University, Mosul 41001, Iraq

² Computer Center, University of Anbar, Anbar 31002, Iraq

³ Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang 43400, Malaysia

Corresponding Author Email: sameer.h@uoanbar.edu.iq

Copyright: ©2024 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/ria.380304>

ABSTRACT

Received: 7 December 2023

Revised: 1 March 2024

Accepted: 15 April 2024

Available online: 21 June 2024

Keywords:

deep neural network, binary cross entropy (BCE), brain stroke identification, unbalanced dataset

This article introduces and evaluates a novel Deep Neural Network (DNN) designed specifically for brain stroke detection. Highlighting the innovation of using a weighted Binary Cross Entropy (BCE) loss function to address dataset imbalances, this work also aims to propose a less computationally complex model that makes predictions using very basic demographic and health-related information. In the realm of stroke detection using brain imaging, various methods, including Convolutional Neural Networks (CNNs) and Hybrid Algorithms (OzNet-mRMR-NB and Linearized Scattering Operator with ML Algorithms), have shown potential but suffer from computational complexity and time consumption. Analyzing factors such as Gender, Age, Hypertension, Heart Disease, Work Type, Residence Type, Average Glucose Level, Body Mass Index (BMI), and Smoking Status allows the proposed model to avoid the computational burdens of image-based feature extraction, focusing instead on accessible data to improve prediction accuracy. The selected model architecture features six hidden layers with dimensions (31, 64, 128, 64, 32, 16) to effectively process these inputs and accurately identify stroke risk. By using weighted BCE as the loss function, the model counteracts dataset bias towards more common non-stroke cases, emphasizing the identification of critical stroke instances during training. Furthermore, the proposed model demonstrated significant improvements in detecting stroke instances, particularly in the minority class, by employing weighted BCE. While accuracy slightly decreased from 95.36% with standard BCE to 75.36% with weighted BCE, this reduction is outweighed by substantial gains in Recall (from 0 to 34.91), Precision (from 0 to 100), and F1-Score (from 0 to 51.75). In addition to making a significant methodological contribution, this study has the potential to improve stroke prediction greatly, allowing for more efficient and rapid treatment.

1. INTRODUCTION

The integration of artificial intelligence in stroke detection is a revolutionary advancement in medical diagnostics, facilitating swift and precise identification of strokes [1]. This technology not only expedites the diagnostic process but also improves accuracy, allowing healthcare practitioners to make prompt and well-informed judgments [2, 3]. Integrating AI into stroke diagnosis is crucial for promptly commencing therapy, hence greatly improving patient outcomes and survival rates. Cerebrovascular accidents, caused by disruptions in blood circulation to the brain, need immediate medical attention to minimize long-term harm [4]. Deep learning expedites the analysis of medical imaging such as MRIs or CT scans, aiding in stroke identification, and contributing to swift diagnosis [5, 6]. Additionally, deep

learning algorithms use comprehensive patient data to predict stroke likelihood, thus enhancing stroke prevention.

Since stroke detection can lead to much better patient outcomes, it is imperative to enhance it [7]. Strokes represent a significant global health concern, and prompt and accurate diagnosis is necessary to initiate effective treatment [8]. A noteworthy development in stroke diagnosis using neuroimaging is the construction of a customized DNN using a weighted BCE loss function. This methodology addresses the challenges associated with acquiring stroke characteristics as well as the issue of imbalanced datasets. This leads to increased diagnostic accuracy and makes stroke diagnosis more rapid and accurate. Innovations in deep learning immediately improve clinical decision-making and offer a chance for early intervention that might potentially prevent long-term consequences and even save lives [9, 10].

The recommended architecture's layers are arranged to carefully analyze complex patterns in imaging data, which is essential for identifying strokes. It is specifically made for the diagnosis of brain strokes. Comprehensive feature extraction becomes easier with this architecture, which is essential for differentiating between stroke episodes and non-stroke situations. The proposed strategy uses weighted BCE in a new way to boost the sensitivity of the model to stroke patients, which are the underrepresented class, hence addressing the common issue of dataset imbalance [11]. This approach greatly improves the model's accuracy in identifying strokes, hence addressing the challenges presented by intricate feature landscapes and discrepancies in data that impede existing models. The suggested model shows a notable improvement over the conventional strategy, combining weighted BCE with a specific framework. This development significantly advances the domains of neuroimaging and stroke recognition while also improving diagnostic processes in clinical settings [12].

Moreover, within the rapidly developing field of medical imaging and diagnostics, deep learning techniques have emerged as an indispensable tool. It has completely changed how diseases like brain strokes are diagnosed and examined. Gaidhani et al. [13] demonstrated the capability of DNNs to reliably diagnose brain strokes using medical imaging, which aligns with the novel approaches suggested in the primary article. Simultaneously, Dourado et al. [14] produced an Internet of Things (IoT) system that uses deep learning to quickly identify strokes in CT images, demonstrating the usefulness of deep learning in clinical environments. In addition to these results, Sirsat et al. [15] undertook a comprehensive examination of deep learning in stroke detection, creating a fundamental comprehension of the function of artificial intelligence in this field. In addition, Vamsi et al. [16] extended the use of artificial intelligence (AI) to include the prediction of stroke severity. This study showcases the wider implications of AI in stroke care, going beyond only the detection of strokes. Last but not least, Abood et al. [17] shown the flexibility of deep learning in predictive maintenance tasks, emphasizing its possibility to be used in many areas, such as healthcare. These studies highlighted the substantial progress and wide-ranging uses of AI in improving stroke diagnostic and treatment methods.

The use of deep learning in stroke diagnosis and treatment exemplifies a crucial transition towards more precise, effective, and nuanced healthcare diagnoses and therapies. The revolutionary model developed in this study, using an innovative strategy for addressing unbalanced data, signifies a substantial advancement in the capacity to accurately identify strokes. This progress, along with the confirming and supplementary discoveries, highlights the crucial and growing significance of AI in medical diagnosis. It highlights the transformative potential of deep learning technologies in advancing stroke detection and enhancing patient care outcomes.

The remaining sections of the paper were structured as follows: An introduction that focused on AI-based brain stroke detection alongside the works in the same field was provided in Section 1. The dataset was clarified, and the comprehensive methodology employed in the research was outlined in Section 2. The obtained results were presented in Section 3, and finally,

the concluding remarks of the paper were provided in Section 4.

2. METHODOLOGY

The dataset, provided by Kaggle [18], discusses the impact of different factors on people who may or may not experience a stroke. Only 5% of the dataset represents the samples of situations which had brain stroke, the remaining 95% of samples in this dataset do not have a brain stroke, making it an unbalanced dataset. An imbalanced dataset is characterized by an uneven distribution of samples across distinct classifications. This discrepancy could pose challenges for machine learning model training, particularly for classification tasks. There are serious problems with classifying an uneven dataset that could compromise the reliability and efficiency of machine learning algorithms [19]. As a result of having an abundance of learning examples, AI models trained on imbalanced datasets display a bias towards the dominant class. As such, it could be challenging for the model to anticipate the minority class which is frequently the class that is most interesting accurately. On the other hand, the model might not accurately capture the underlying patterns and characteristics of the minority class because there is insufficient data from that class. As a result, the minority class may have a low recall rate, which would suggest that the model is not able to accurately identify a sizable proportion of positive examples. The accuracy metric is commonly employed to evaluate the efficacy of a machine learning model, especially in datasets with uniform distribution. It measures the proportion of correctly identified cases to all instances in the dataset. Numerous metrics are frequently used in AI binary classification to evaluate a model's efficacy. As an example, the formulas for recall, accuracy, precision, and F1-score are given in the equations [20].

$$Accuracy = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \quad (1)$$

$$Precision = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \quad (2)$$

$$Recall = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \quad (3)$$

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

where,

- True Positives (TP): Cases where the model accurately predicts a positive outcome while the observations are positive.
- True Negatives (TN): Cases in which the model accurately predicts a negative outcome even when the actual data is negative.
- False Positives (FP): Irregular cases where the model predicts a positive outcome when it is negative.
- False Negatives (FN): Cases in which the model predicts negative values when they are positive.

Accuracy is a metric that evaluates the overall correctness of a model. Precision, on the other hand, assesses the accuracy of the positive predictions made by the model. Recall indicates the model's capability to capture all the positive occurrences. The F1-Score, which is the harmonic means of precision and recall, offers a balanced evaluation of both metrics. Nonetheless, the traditional accuracy metric can be misleading on unbalanced datasets. In many real-world applications, the minority class is the class of interest, therefore poor performance there might be hidden by a high accuracy. While accuracy may not be the best choice due to its sensitivity to class imbalances, several metrics are more informative and suitable for evaluating models on unbalanced datasets (Recall, Precision, and F1-Score). Noteworthy that the occurrences of the presence or absence of brain stroke conditions are not equal but are distributed unevenly. In this dataset, 95% of the total records are labeled as 'NO' for stroke, indicating the absence of stroke conditions, while the remaining 5% are labeled as 'YES' for stroke, signifying the presence of stroke conditions (Figure 1). This distribution provides insights into the prevalence of the condition within the dataset, indicating that instances without stroke conditions are much more common than those with stroke conditions.

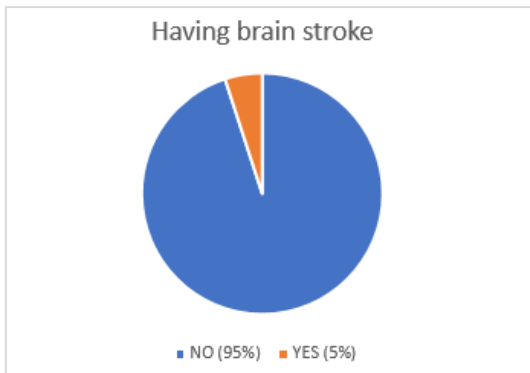


Figure 1. The percentage of brain strokes in the dataset

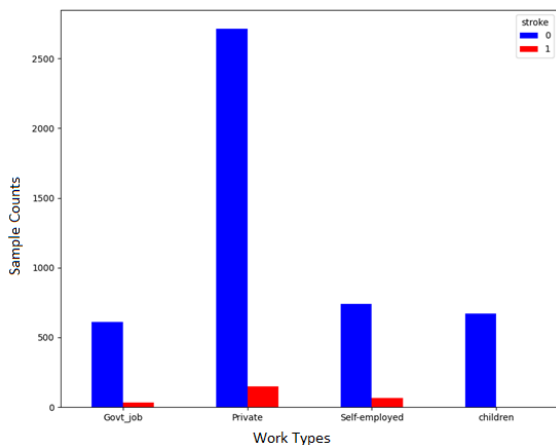


Figure 2. Brain stroke relation with work type

Additionally, the relationship between work type and stroke risk is multifaceted. Sedentary jobs involving prolonged sitting, high-stress occupations, and irregular work hours, like shift work, may increase cardiovascular risks, including stroke. Conversely, physically active jobs may have a protective effect. Environmental factors in certain occupations could also contribute. However, strokes result from a combination of genetic, lifestyle, and environmental factors, with overall

health and lifestyle choices playing a significant role. Moreover, the proposed dataset analyzed the effects of several work types, such as private, government, self-employed, and children-dealing-based jobs, on the likelihood of having a stroke (see Figure 2).

Noteworthy that the chart in Figure 3 illustrates that the percentage of women experiencing strokes is consistently higher than that of men across all age groups. This aligns with broader data indicating that women face a greater lifetime risk of strokes than men. This could be because there are more men in the general population experiencing strokes.

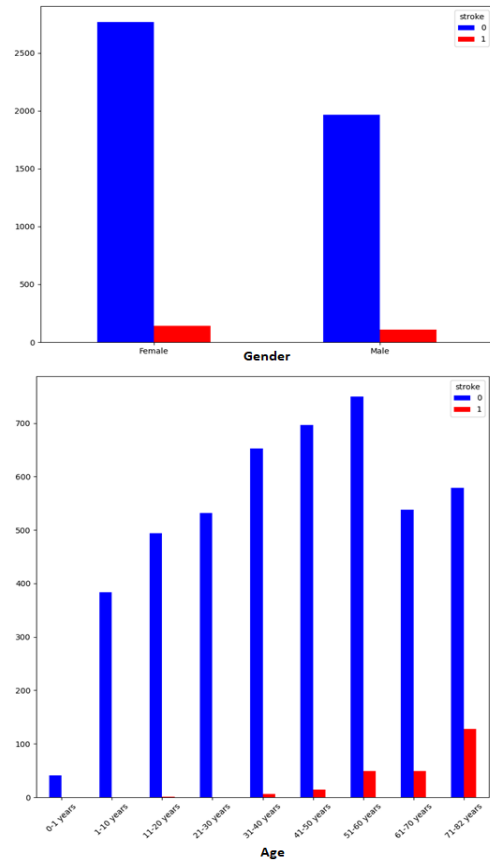


Figure 3. Brain stroke relation with gender and age

Figure 4 depicts a bar graph illustrating the number of people categorized by their smoking status. Alternatively, the interpretation of the graph proposes that individuals are more inclined to quit smoking once they have started. This inference arises from the observation that the number of current smokers is lower than the number of those who have formerly smoked.

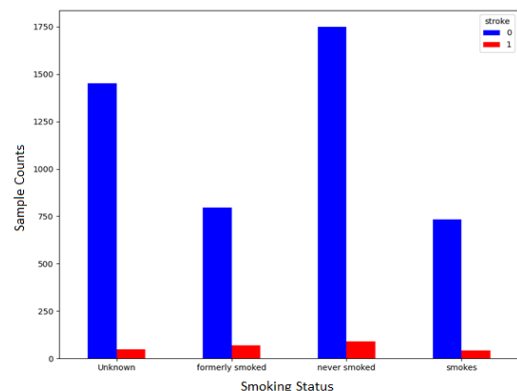


Figure 4. Brain stroke relation with smoking

Moreover, Figure 5 indicates that individuals residing in rural areas have a lower stroke rate compared to those in urban areas. One possibility is that rural areas exhibit a lower prevalence of stroke risk factors, including high blood pressure, high cholesterol, diabetes, and obesity. They might also have reduced exposure to air pollution and other environmental toxins.

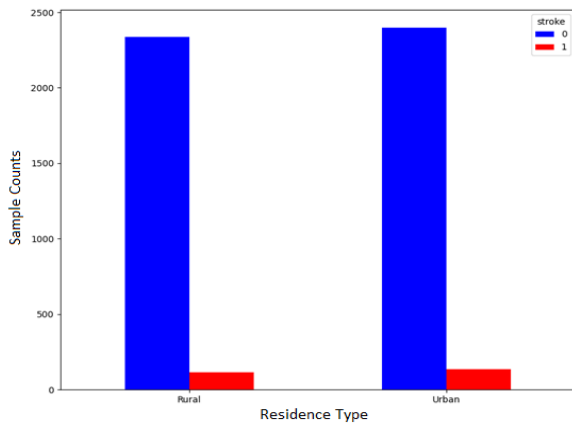


Figure 5. Brain stroke relation with residence type

Furthermore, Figure 6 shows the number of people with and without hypertension who have strokes. People with hypertension are more likely to have strokes than people without hypertension. This is consistent with medical knowledge, as hypertension is a major risk factor for stroke.

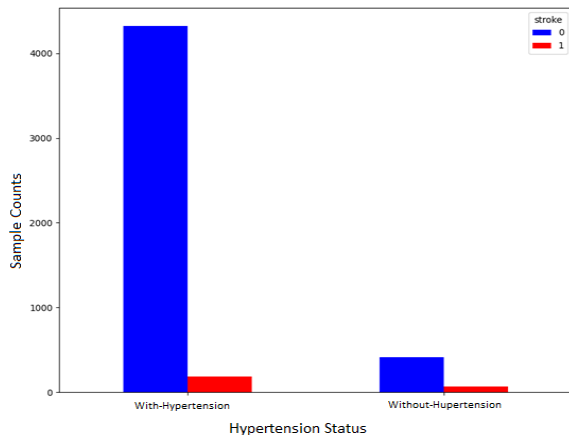


Figure 6. Brain stroke relation with hypertension

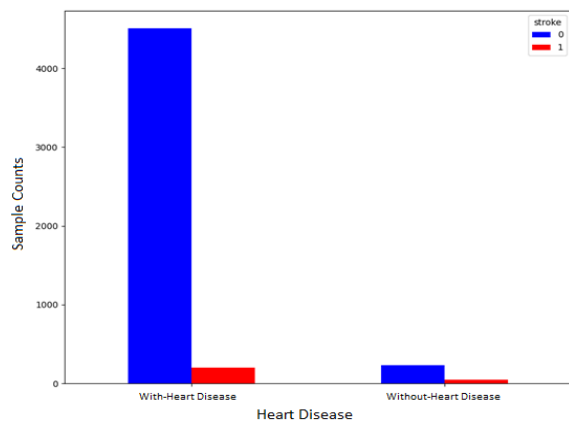


Figure 7. Brain stroke relation with heart disease

Notably, the incidence of stroke is greater in those with heart disease compared to those without heart disease (Figure 7). This suggests that there is a relationship between heart disease and stroke.

A higher glucose level is associated with an increased risk of stroke, as shown in Figure 8. Individuals who have had a stroke tend to have significantly higher average glucose levels compared to those who have not had a stroke.

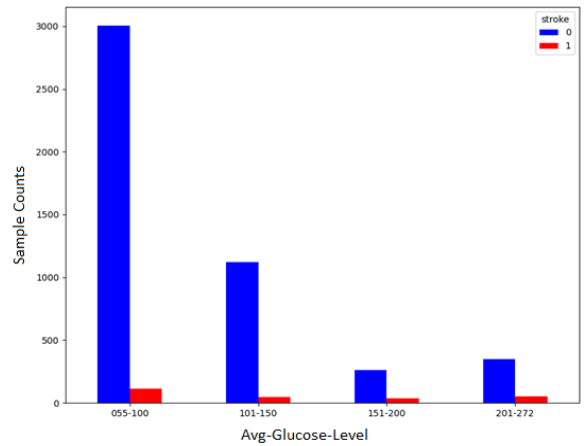


Figure 8. Brain stroke relation with glucose level

The correlation between marital status and the occurrence of strokes is presented in Figure 9. The data reveals a higher percentage of stroke cases among married individuals compared to their single counterparts. Married individuals may lead more stress-inducing lifestyles, further contributing to an increased risk of strokes.

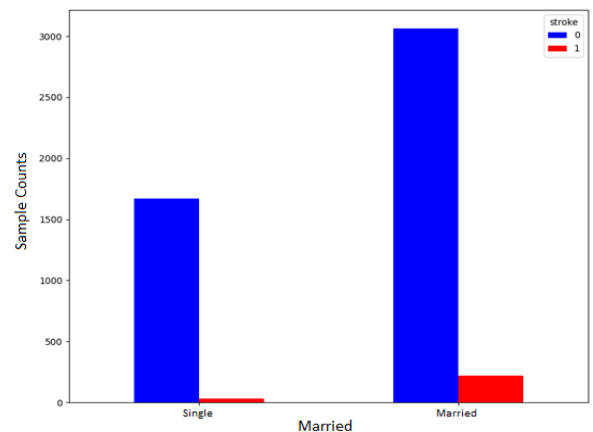


Figure 9. Brain stroke relation with married people

The proposed methodology notably did not use a selective feature extraction process; instead, all features available in the dataset were employed to train the model. This strategy is based on the dataset's comprehensive coverage of a broad range of factors relevant to stroke prediction. The aspects include demographic, health-related, and lifestyle factors, which together contribute to a comprehensive comprehension of stroke risk.

Statistical characteristics, including mean age, average blood sugar, and body mass index (BMI), were scaled from 0 to 1. To avoid any one feature's size from disproportionately impacting the model during training, it is necessary to normalize the data such that the ranges of features are consistent with one another.

On the other hand, factors such as gender, hypertension, heart disease, marital status, employment type, residence type, and smoking status were encoded using categorical encoding. This procedure converts categorical information into a format that DNNs can effectively handle, enabling the model to reliably comprehend these characteristics and use them for stroke prediction.

The decision to include all characteristics and use these specific pre-processing approaches was driven by the goal of enhancing the model's ability to effectively learn from diverse input points. By normalizing numerical features and encoding categorical features categorically, it is ensured that the model is trained on data that faithfully reflects the complex nature of stroke risk, thus improving its predictive accuracy.

In this study, a specialized model designed for the detection of brain strokes is introduced. The architecture of the proposed artificial intelligence model is illustrated in Figure 10.

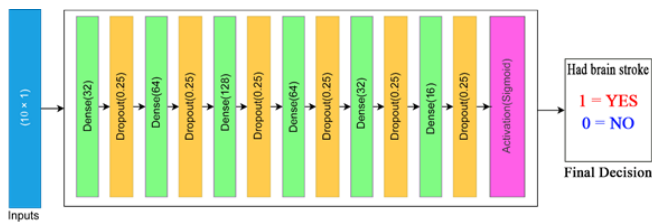


Figure 10. The proposed model architecture

The model is provided with information from thirty inputs (List 1). The proposed model comprises six hidden dense layers with sizes of (32, 64, 128, 64, 32, 16) respectively. Following each dense layer is a dropout layer with a probability of 0.25 to mitigate the overfitting problem during model training. The final layer utilizes the Sigmoid activation function due to the binary nature of the classification problem, specifically in detecting brain strokes. It is worth mentioning that the weighted BCE is employed to address the issue of an unbalanced dataset. This loss function, an adaptation of the normal BCE, incorporates class weights to account for the imbalance in the dataset. Distinct weights are assigned to the positive and negative classes, mitigating the effects of imbalance during training. This strategic weighting enables the model to prioritize the optimization of the minority class. Moreover, the BCE loss for a single sample is typically defined in Eq. (5) [19].

$$BCE = -[y \cdot \text{Log}(p) + (1 - y) \cdot \text{Log}(1 - p)] \quad (5)$$

where, y represents the true label (ground truth) of the sample, taking values of either 0 or 1, and p signifies the predicted probability of the positive class. Now, in handling class weights, the BCE loss with class weights is modified in Eq. (6) [19].

$$\text{Weighted BCE} = -[w_{pos} \cdot y \cdot \text{log}(p) + w_{neg} \cdot (1 - y) \cdot \text{Log}(1 - p)] \quad (6)$$

where, w_{pos} is the weight assigned to the positive class (the minority class), and w_{neg} is the weight assigned to the negative class (the majority class).

In determining class weights for an imbalanced binary dataset based on class frequencies, utilize the following formula. In the context of a binary classification problem with "Class 0" as the majority class and "Class 1" as the minority

class, initiate by calculating the class frequencies: $N0$ for the number of samples in Class 0 (majority class) and $N1$ for the number of samples in Class 1 (minority class). Subsequently, derive the class weights as follows: for Class 0 (majority class), $w_{neg} = N1/N0$, and for Class 1 (minority class), $w_{pos} = N0/N1$. The class weight for Class 0 (w_{neg}) represents the ratio of the number of samples in Class 1 ($N1$) to the number of samples in Class 0 ($N0$), while the class weight for Class 1 (w_{pos}) is the inverse of this ratio.

3. RESULTS AND DISCUSSION

The study purposefully incorporated a 0.25 dropout rate following every layer in the model to address the possibility of overfitting. The rationale for the choice was that this specific dropout rate strikes the best balance between the model's complexity and its ability to work with new data. The choice not to experiment with other dropout rates was made after doing preliminary studies and reviewing existing literature. These sources suggest that a dropout rate of 0.25 is often effective for models with comparable complexity and data characteristics. Consequently, the focus was directed toward refining other model parameters and designs, determining that a dropout rate of 0.25 was sufficient for the objectives. This approach facilitated the attainment of strong model performance, guaranteeing that it did not excessively match the training data, hence preserving its capacity to be applied to new data and its effectiveness in forecasting stroke.

This research presents an AI model that was built and trained utilizing the KERAS and COLAB frameworks. Eighty percent will be used for training, and twenty percent will be used for validation. The training process spans 100 epochs, following the configurations outlined in Table 1.

Table 1. Training parameters

Parameter	Value
Learning rate	1×10^{-4}
Epochs	100
Optimizer	Adam Optimizer
Beta 1	0.9
Beta 2	0.999
Loss	Weighted BCE
Metric	Accuracy, Precision, Recall, F1-Score
Train dataset size	80%
Test dataset size	20%
Trainable parameters	21,665

The normal (BCE) model computes the cross-entropy between predicted and actual labels without incorporating any weighting. Conversely, the weighted BCE model assigns distinct weights to various examples based on their significance. Accordingly, the weighted BCE model proves superior for this task, evident in its higher validation accuracy (Figure 11). This implies that the model effectively captures underlying patterns in the data, demonstrating a reduced likelihood of overfitting to the training data.

The Recall is a quantitative measure that indicates the proportion of correctly detected positive instances. It is shown in Figure 12, which compares the performance of a standard BCE model with that of a weighted BCE model on both the training and validation sets. The superior performance of the weighted BCE model can be attributed to its strategic weighting of positive examples during training, enhancing its ability to accurately identify them.

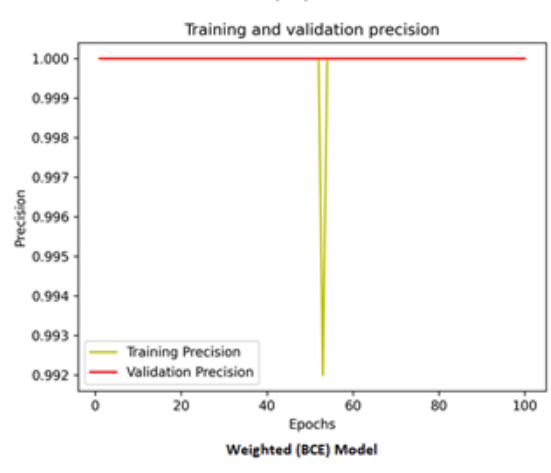
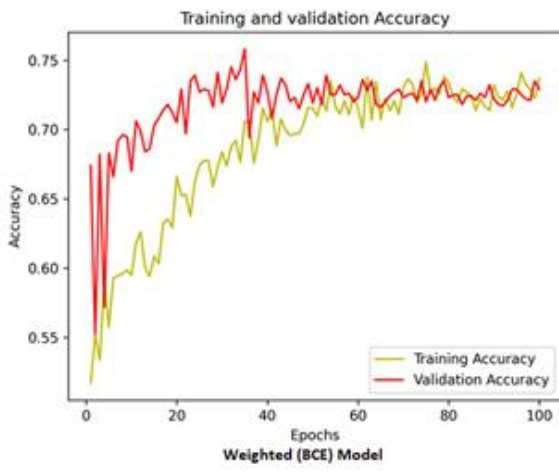
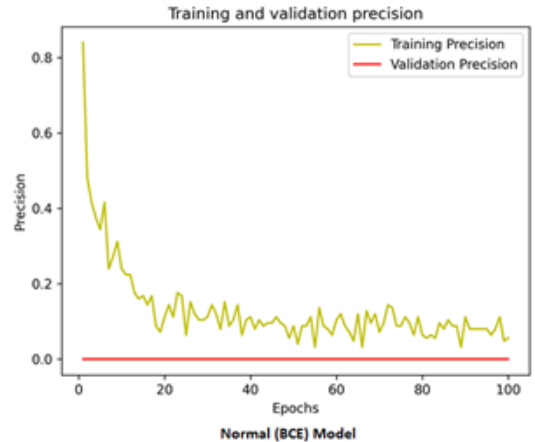
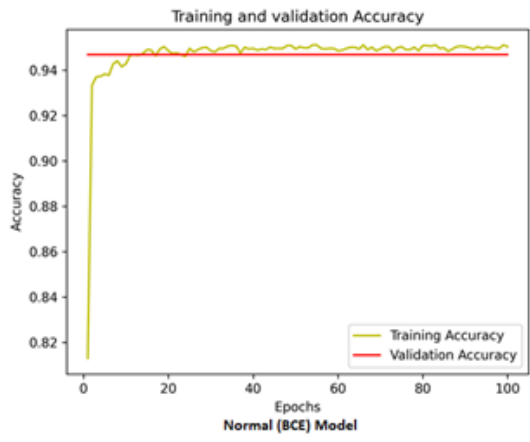


Figure 11. Brain stroke training and validation accuracy

Figure 13. Brain stroke training and validation precision

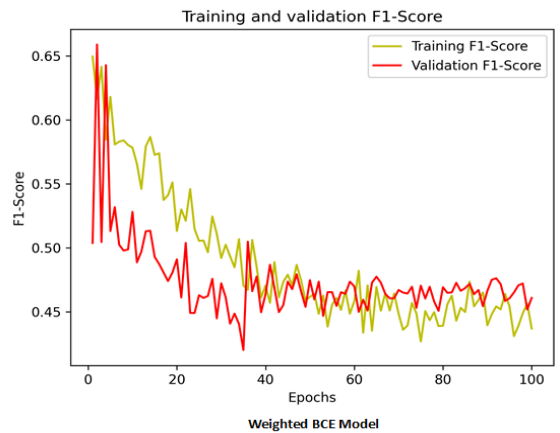
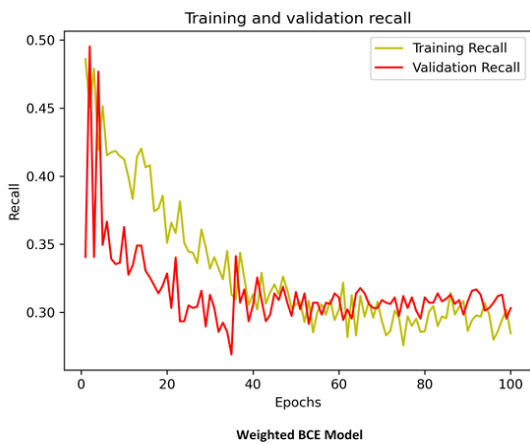
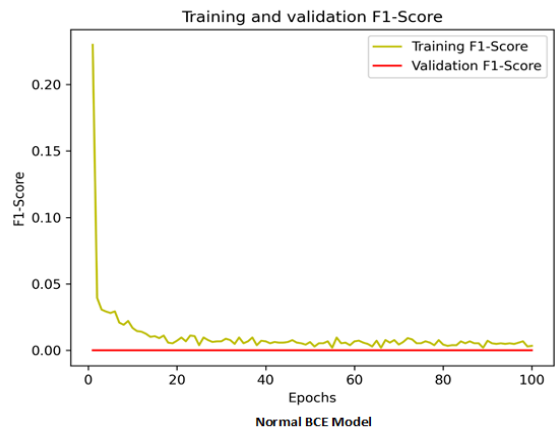
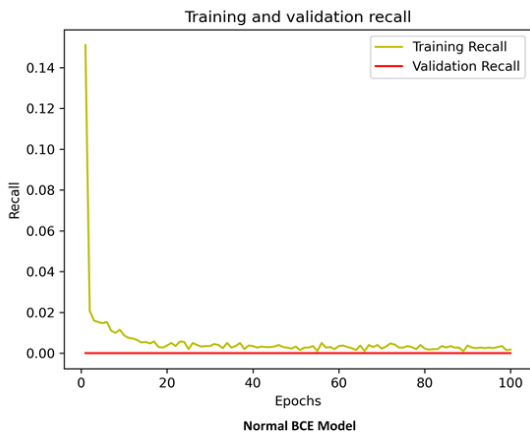


Figure 12. Brain stroke training and validation accuracy

Figure 14. Brain stroke training and validation F1-score

As depicted in Figure 13, the superiority of the weighted BCE model is evident over the normal BCE model in terms of both training and validation precision. Consequently, the weighted BCE model stands as a preferred choice for classification tasks, particularly when dealing with imbalanced datasets, as it demonstrates the capacity to effectively learn from the data and generalize well to new instances.

The observed differences in F1 scores highlight the trade-off between training performance and the model's ability to generalize. The normal BCE model, with a higher training F1-score, struggles to extend its performance to new instances. On the other hand, the weighted BCE model, with a more balanced performance between training and validation F1-scores, indicates a better capacity to generalize beyond the training data (Figure 14). This emphasizes the importance of considering both training and validation metrics to assess a model's overall performance and its ability to generalize to real-world scenarios.

Weights are commonly employed to tackle class imbalance, where one class contains substantially more samples than the other. By assigning a higher weight to the minority class, the weighted BCE loss function enables the model to prioritize those samples. As depicted in Figure 15, the weighted BCE model exhibits lower training and validation losses compared to the normal BCE model. This discovery suggests that the weighted BCE model performs exceptionally well in acquiring knowledge from the training data and has improved skills in extrapolating to novel, unfamiliar data.

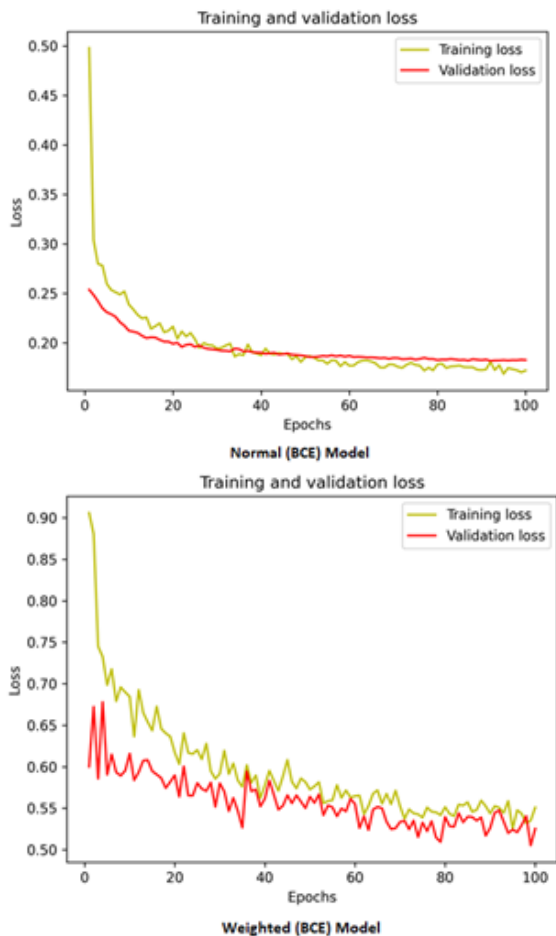


Figure 15. Brain stroke training and validation loss

Considering the data presented in Table 2, the Weighted

BCE variant of the Proposed Model demonstrates superior overall performance. This deduction is made based on its maximum F1 score, which signifies a weighted mean of accuracy, recall, and precision. Furthermore, the Proposed Model (Weighted BCE) has the greatest recall out of all the models, suggesting its improved capacity to properly detect positive cases. In contrast, the remaining models in the table, including Weighted Random Forest and Weighted Logistic Regression, exhibit lower overall performance. This is attributed to their lower F1 scores, diminished recall, and reduced precision in comparison to the Proposed Model.

Moreover, after completing 100 epochs of training, the proposed model's performance was evaluated under two scenarios: one employing normal BCE and the other using weighted BCE. The evaluation was conducted on the test dataset, with the results for each scenario detailed in Table 2. When normal BCE was applied, the model achieved notably high accuracy, but the Recall, Precision, and F1-score metrics all registered zero values. This phenomenon can be attributed to the significant imbalance between the majority and minority classes, where the majority class dominates the dataset with a 95% prevalence, compared to the 5% representation of the minority class. In highly imbalanced datasets, a model might achieve elevated accuracy by predominantly predicting the majority class. This leads to an abundance of true negatives and true positives for the majority class, consequently resulting in a misleadingly high accuracy value.

Nonetheless, the model's predictive accuracy for the minority class may be compromised due to its limited exposure during training. Consequently, an increase in false negatives occurs, resulting in diminished recall and F1-score. Moreover, since the model struggles to accurately identify positive instances (manifested by low true positives), precision also experiences a notable decline. In contrast, the adoption of weighted BCE reveals a reduction in Accuracy from 95.36% to 75.36%. Nevertheless, this compromise is followed by significant improvements in the measures of Recall, Precision, and F1-score, all of which are no longer limited to zero. This enhancement suggests that the suggested AI model is trained more efficiently using weighted BCE, showcasing an enhanced capacity to forecast occurrences from the underrepresented class.

The dataset was obtained and examined from Kaggle. The influence of several parameters on the probability of having a stroke, uncovering a substantial disparity in class distribution. just 5% of the samples represented cases of stroke, while the other 95% represented instances without strokes. The analysis highlighted the asymmetry, which poses a hurdle in machine learning, particularly for classification tasks. The essay accurately identified possible biases and difficulties associated with imbalanced datasets, emphasizing the insufficiency of depending just on accuracy measurements and pushing for metrics like precision, recall, and F1-Score in such situations.

The dataset explored the correlation between stroke incidence and variables such as occupation, gender, smoking habits, residential location, hypertension, cardiovascular disease, glucose levels, and marital status. In addition, the research presented a specialized DNN designed to detect strokes. This model incorporates 10 specific input characteristics. The design consisted of six concealed dense layers with dropout layers to mitigate overfitting, culminating in the use of the Sigmoid activation function for binary classification. The weighted BCE loss function was used to address the imbalance by allocating different weights to the

positive and negative classes.

The AI model was trained using KERAS and Google Colab. It used an 80-20 split of training and validation data and was trained for 100 epochs. The model's training with an inadequate amount of data from the minority class led to subpar predictions for that specific class. More specifically, there was an increase in the number of false negatives,

resulting in decreased recall and F1 scores. Moreover, the model exhibited challenges in correctly detecting positive occurrences, resulting in a decrease in the number of true positives and a reduced level of accuracy. Nevertheless, the use of weighted BCE enhanced the model's capacity to accurately forecast instances from the underrepresented class.

Table 2. Performance scores on test dataset

Algorithm	Accuracy (%)	Recall (%)	Precision (%)	F1-Score (%)
Normal SVM	95.36	0	0	0
Weighted SVM	73.42	62.26	12.79	21.22
Normal Random Forest	95.36	0	0	0
Weighted Random Forest	94.68	1.88	50	3.65
K=2	94.38	18.86	20	3.44
Normal Logistic Regression	95.36	0	0	0
Weighted Logistic Regression	73.42	69.81	12.93	21.82
Normal BCE	95.36	0	0	0
Proposed Weighted BCE	75.36	34.91	100	51.75

Although the total accuracy dropped from 95.36% to 75.36%, there were notable improvements in recall, precision, and F1-score, all of which were originally zero. This indicates that the use of weighted BCE has improved the model's capacity to precisely detect instances belonging to the minority class. Results of the evaluation were presented for scenarios using standard BCE and weighted BCE. Although the standard BCE approach achieved great accuracy, zero results for recall, precision, and F1-score have been achieved owing to the imbalance. However, the use of weighted BCE resulted in a decrease in accuracy but increased recall, precision, and F1-score. This proves that the model performed better in predicting occurrences from the minority class.

The achieved findings gave a concise summary of the suggested deep DNN model for stroke detection, highlighting the significance of the six hidden dense layers and the use of weighted BCE to address dataset imbalance. The model's improved performance with weighted BCE was shown, exhibiting higher accuracy, F1-score, and recall in comparison to situations without weighted BCE. The review confirmed the efficacy of the suggested model in tackling the unbalanced dataset for stroke detection.

4. CONCLUSIONS

This article presents the ultimate effects of employing the weighted BCE compared to the standard BCE when dealing with the difficulties of the unbalanced datasets, particularly the challenges of neuroimaging data in stroke detection. This study proved that the standard BCE realized a creditable accuracy with 95.36%. However, it presented an obvious insufficiency to achieve the minority class as demonstrated in the zero results of recall, precision, and F1-score metrics. Despite the decrease in overall accuracy to approximately 75.36%, the implementation of the weighted BCE represents a significant improvement in the effectiveness of the model in the minority class. This is confirmed due to the increased results of recall to 34.91%, accuracy to 100%, and F1 – score to 51.75%, which demonstrate the applicability of the weighted BCE. The above improvements highlight the increased sensitivity and specificity of the model in identifying stroke cases. This underscores the critical importance of

weighted BCE in enhancing model performance for clinical use. This modification has effectively improved the detection of stroke cases while maintaining a satisfactory degree of accuracy. This represents critical progress towards developing AI solutions that are fair and accurate, and thus have significant clinical relevance.

Additionally, weighted BCE significantly improves the accuracy of stroke prediction by facilitating the identification of underrepresented groups in imbalanced datasets, including stroke cases. Traditional modeling techniques that rely on accuracy often ignore these critical and uncommon conditions. For this reason, they provide measures of accuracy that appear high but are unable to identify diseases such as strokes, which may underestimate the patient's treatment. Alternatively, weighted BCE recalibrates model performance and significantly enhances F1 results, memory and accuracy, raising the bar for stroke detection accuracy and reliability. This improvement is essential to advance equitable healthcare technology as well as rapid intervention, which may successfully reduce brain damage and improve patient outcomes. These developments are consistent with the goals of precision medicine, which seeks to provide personalized medical treatment based on the unique needs of each patient. Significant progress is emerging in the application of artificial intelligence in neuroimaging to predict strokes by incorporating weighted BCE. This breakthrough improves diagnostic accuracy and reliability and encourages broader integration of AI into personalized healthcare solutions.

ACKNOWLEDGMENT

The authors would like to thank the electrical and control department of the technical engineering college of Kirkuk, Northern Technical University for their continued support to this study.

REFERENCES

- [1] Akter, B., Rajbongshi, A., Sazzad, S., Shakil, R., Biswas, J., Sara, U. (2022). A machine learning approach to detect the brain stroke disease. In 2022 4th International

- Conference on Smart Systems and Inventive Technology (ICSSIT), Tirunelveli, India, pp. 897-901. <https://doi.org/10.1109/ICSSIT53264.2022.9716345>
- [2] Quazi, S. (2022). Artificial intelligence and machine learning in precision and genomic medicine. *Medical Oncology*, 39: 120. <https://doi.org/10.1007/s12032-022-01711-1>
- [3] Balhara, S., Gupta, N., Alkhayyat, A., Bharti, I., Malik, R.Q., Mahmood, S.N. Abedi, F. (2022). A survey on deep reinforcement learning architectures, applications and emerging trends. *IET Communications*. <https://doi.org/10.1049/cmu2.12447>
- [4] Caplan, L.R., Simon, R.P., Hassani, S. (2023). Chapter 27 - Cerebrovascular disease—stroke. *Neurobiology of Brain Disorders (Second Edition)*, 457-476. <https://doi.org/10.1016/B978-0-323-85654-6.00044-7>
- [5] Khalid, H., Hussain, M., Ghamdi, M.A.A., Khalid, T., Khalid, K., Khan, M.A., Fatima, K., Masood, K., Almotiri, S.H., Farooq, M.S., Ahmed, A. (2020). A comparative systematic literature review on knee bone reports from MRI, x-rays and CT scans using deep learning and machine learning methodologies. *Diagnostics*, 10(8): 518. <https://doi.org/10.3390/diagnostics10080518>
- [6] Saad, M.A., Alhamdane, H.J., Ali, A., Alani, S. (2020). Total energy consumption analysis in wireless mobile ad hoc network with varying mobile nodes', *Indonesian Journal of Electrical Engineering and Computer Science*, 20(3): 1397-1405. <https://doi.org/10.11591/ijeecs.v20.i3.pp1397-1405>
- [7] Chen, Y., Wang, Y., Hu, F., Feng, L.F., Zhou, T.H., Zheng, C. (2021). LDNNET: Towards robust classification of lung nodule and cancer using lung dense neural network. *IEEE Access*, 9: 50301-50320. <https://doi.org/10.1109/ACCESS.2021.3068896>
- [8] Mahmood S.N., Mohammed, S.S., Ismaeel, A.G., Clarke, H.G., Mahmood, I.N., Aziz, D.A., Alani, S. (2023). Improved malaria cells detection using deep convolutional neural network. In *2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Istanbul, Turkiye, pp. 1-4. <https://doi.org/10.1109/HORA58378.2023.10156747>
- [9] Ismaeel, A.G., Janardhanan, K., Sankar, M., Natarajan, Y., Mahmood, S.N., Alani, S., Shather, A.H. (2023). Traffic pattern classification in smart cities using deep recurrent neural network. *Sustainability*, 15(19): 14522. <https://doi.org/10.3390/su151914522>
- [10] Wang, X., Zhao, Y., Pourpanah, F. (2020). Recent advances in deep learning. *International Journal of Machine Learning and Cybernetics*, 11: 747-750. <https://doi.org/10.1007/s13042-020-01096-5>
- [11] Hurtik, P., Tomasiello, S., Hula, J., Hynar, D. (2022). Binary cross-entropy with dynamical clipping. *Neural Computing and Applications*, 34: 12029-12041. <https://doi.org/10.1007/s00521-022-07091-x>
- [12] Qu, S.Y., Li, B., Zhao, S.X., Zhang, L., Wang, Y. (2023). A coordinated model pruning and mapping framework for RRAM-based DNN accelerators. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 42(7): 2364-2376. <https://doi.org/10.1109/TCAD.2022.3221906>
- [13] Gaidhani, B.R., Rajamenakshi, R.R., Sonavane, S. (2019). Brain stroke detection using convolutional neural network and deep learning models. In *2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, Jaipur, India, pp. 242-249. <https://doi.org/10.1109/ICCT46177.2019.8969052>
- [14] Dourado, C.M.J.M., Silva, S.P., Nobrega, R.V.M.D., Barros, A.C.D.S., Filho, P.P.R., Albuquerque, V.H.CD. (2019). Deep learning IoT system for online stroke detection in skull computed tomography images. *Computer Networks*, 152: 25-39. <https://doi.org/10.1016/j.comnet.2019.01.019>
- [15] Sirsat, M.S., Fermé, E., Camara, J. (2020). Machine learning for brain stroke: A review. *Journal of Stroke and Cerebrovascular Diseases*, 29(10): 105162. <https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.105162>
- [16] Vamsi, B., Bhattacharyya, D., Midhunchakkaravarthy, D. (2021). Detection of brain stroke based on the family history using machine learning techniques. *Smart Technologies in Data Science and Communication*, 17-31. https://doi.org/10.1007/978-981-16-1773-7_2
- [17] Abood, A.M., Nasser, A.R., Al-Khazraji, H. (2022). Predictive maintenance of electromechanical systems using deep learning algorithms. *Ingénierie des Systèmes d'Information*, 27: 1009-1017. <https://doi.org/10.18280/isi.270618>
- [18] Kaggle. (2022). Brain Stroke Dataset. <https://www.kaggle.com/datasets/jillanisofttech/brain-stroke-dataset>.
- [19] Bishop, C.M. (2006). *Pattern recognition and machine learning*. Information Science and Statistics, Springer New York, NY. <https://doi.org/10.1007/978-0-387-45528-0>
- [20] Chollet, F. (2021). *Deep learning with Python: Simon and Schuster*. <https://www.mangoud.com/wp-content/uploads/2020/11/Francois-Chollet-Deep-Learning-with-Python-2018-Manning.pdf>.