BAREKENG: Journal of Mathematics and Its Applications
December 2024 Volume 18 Issue 4 Page 2423–2434
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol18iss4pp2423-2434

ADVANCEMENTS IN ALZHEIMER'S DIAGNOSIS THROUGH MRI USING BAYESIAN CONVOLUTIONAL NEURAL NETWORKS AND VARIATIONAL INFERENCE

Alifia Ardha Nareswari[1*], Dina Tri Utari[2]

[1,2]*Department of Statistics, Faculty of Mathematics and Natural Sciences, Universitas Islam Indonesia*
Jln. Kaliurang Km 14.5, Sleman, Yogyakarta, 55584, Indonesia

Corresponding author's e-mail: [*] ***20611177@students.uii.ac.id***

Article History:

Received: 28th, March 2024
Revised: 31st, May 2024
Accepted: 27th, July 2024
Published:14th, October 2024

Keywords:

Alzheimer;
Bayesian Convolutional Neural Network;
Variational Inference.

ABSTRACT

Alzheimer's disease is one of the brain disorders that can be deadly in older. The disease is less treated and less recognized, but Alzheimer's disease is now a significant public health problem. Early detection of the disease can significantly reduce symptoms. However, the lack of medical personnel makes handling this disease complex. Therefore, an automatic diagnosis of Alzheimer's disease is needed with a Magnetic Resonance Imaging (MRI) examination to get an accurate diagnosis of the disease. This study classified the type of Alzheimer's disease with deep learning methods using the Bayesian Convolutional Neural Network (BCNN) and the Variational Inference (VI) technique. It aims to determine image classification and accuracy level at the level of Alzheimer's disease by using 2,400 brain MRI images, divided into three classes (non-demented, very mild demented, and mild demented) based on severity. The data was acquired from the kaggle.com website. We use a dataset scenario of 80% for training and 20% for testing, 100×100 pixels, kernel size 3×3, and optimizer Adam with epoch 200. The accuracy of the image classification process is 80%. The non-demented label predicts that the uncertainty is 0.371, and the other uncertainty prediction is 0.002. The ability to anticipate uncertainty enables clinicians to make informed decisions regarding the reliability of the model's output and the need for additional validation or confirmation.

How to cite this article:

A. A. Nareswari and D. T. Utari., "ADVANCEMENTS IN ALZHEIMER'S DIAGNOSIS THROUGH MRI USING BAYESIAN CONVOLUTIONAL NEURAL NETWORKS AND VARIATIONAL INFERENCE," *BAREKENG: J. Math. & App.*, vol. 18, iss. 4, pp. 2423-2434, December, 2024.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

As per the World Health Organization (WHO), Alzheimer's disease is the predominant type of Dementia, accounting for 60-70% of cases [1]. The majority of these cases manifest in individuals aged 65 years or above. By 2023, the global prevalence of Dementia is projected to exceed 55 million individuals, with over 60% of this population residing in low- and middle-income nations. Presently, the management of Alzheimer's disease solely entails alleviating symptoms, with no pharmaceutical interventions capable of providing a cure. The condition of Dementia will have physical, psychological, social, and economic effects not only on the individual with Dementia but also on their caregivers, families, and communities. Frequently, individuals lack knowledge or comprehension of Dementia, resulting in the stigmatization and challenges associated with its diagnosis and treatment. Hence, the identification of Alzheimer's disease can be accomplished through a rapid and uncomplicated system that utilizes an individual's demographic and clinical information.

Integrating advanced technologies, such as magnetic resonance imaging (MRI) and deep learning algorithms, has led to notable advancements in diagnosing Alzheimer's disease. The utilization of digital technology and the internet has emerged as viable solutions for addressing these challenges. The utilization of Artificial Intelligence (AI) represents a burgeoning domain within scientific inquiry. AI can identify objects, offering valuable and streamlined insights, such as identifying disease types through image classification. Deep learning techniques, capable of automatically extracting features, can be employed in the object recognition process. The Bayesian Neural Network (BNN) is a deep learning technique that can extract image features or images. It can classify images, generate uncertainty estimates, and predict by incorporating probability elements into the neural network.

According to [2], applying BCNN in combination with variational inference has demonstrated the potential to improve the precision and effectiveness of Alzheimer's disease detection. The techniques above utilize deep learning to examine complex patterns within medical images, explicitly focusing on neuroimaging applications [3]. CNNs have been utilized in MRI analysis to create frameworks that can identify distinct features of Alzheimer's disease from imaging data [4]. Furthermore, the utilization of multimodal MRI imaging has yielded significant findings regarding the structural alterations linked to Alzheimer's disease, underscoring the significance of employing sophisticated imaging methodologies to comprehend the pathology of this ailment [5]. Moreover, research has shown that ensemble learning and transfer learning methods are particularly effective in enhancing the categorization of Alzheimer's disease based on MRI data [6]. According to [7], these methodologies improve the precision of diagnosis and play a significant role in the timely identification of Alzheimer's disease, a critical aspect for the successful management and treatment of the condition. Moreover, incorporating Bayesian techniques in the MRI analysis has demonstrated promise in enhancing datasets and the efficacy of classifiers for Alzheimer's disease [8].

There is an urgent requirement for sophisticated computational methods that yield precise diagnoses and provide valuable information regarding the certainty and dependability of predictions. Incorporating uncertainty estimation into the model framework, BCNN coupled with VI presents a promising approach to tackle this challenge. Nevertheless, despite their potential, the utilization of BCNN with VI in diagnosing Alzheimer's disease from MRI data has not been thoroughly investigated. This study aims to close this divide by examining the effectiveness of BCNN with VI in improving the precision and dependability of Alzheimer's disease diagnosis using MRI analysis. This study aims to enhance the diagnostic tools for early detection and monitoring of Alzheimer's disease by utilizing the inherent uncertainty estimation capabilities of BCNN with VI.

2. RESEARCH METHODS

2.1 Convolutional Neural Network (CNN)

The Convolutional Neural Network (CNN) is a type of neural network architecture that operates feed-forwardly. It consists of multiple convolution layers, which may follow an optional batch unification layer, activation function, and normalization. Furthermore, the architecture comprises fully connected layers. The image diminishes in size as it traverses the network, primarily due to reaching its maximum capacity. The final layer will present the probability prediction of the class [9]. In mathematical terms, the image that is the focus of a CNN is known as a tensor. The terms vector and matrix are frequently employed in the context of tensors [10]. The provided image depicts CNN's architectural design.

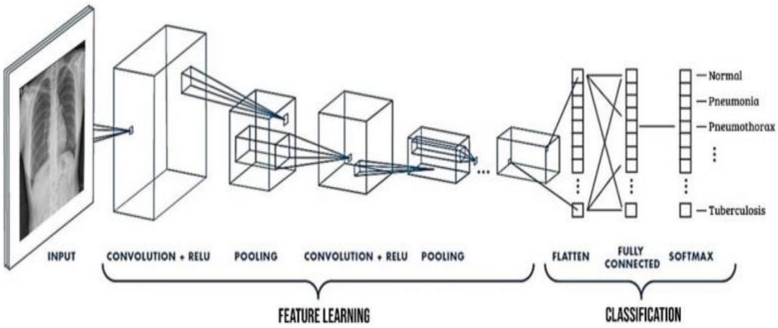

Figure 1. CNN Architecture [11]

Convolutional Neural Network (CNN) models can facilitate the development of image classifications to predict and classify images. The typical model architecture consists of multiple layers with initial values, weights, and biases. Subsequently, it utilizes training data to input weights and biases [9]. The CNN architecture comprises feature learning and classification [11]. The initial component of the architecture is feature learning, which encompasses multiple layers. The initial layer is called the convolution layer, while the subsequent layer is the pooling layer. An activation function is applied to each layer between the first and second types. The classification layer comprises multiple layers, with the neurons in each layer fully interconnected with previous layers. This layer receives the vector input from the image feature extraction layer. According to reference [12], the transformation yields accuracy values for each classification class as its output.

2.1.1 Convolutional Layer

Once the data is inputted into the CNN layer, it undergoes a convolution process using multiple filters to identify image data. The convolution process involves multiplying the dots between the input feature map in the image and a filter [13]. Convolution on image data is performed to extract features from the input image. Convolutional operations linearly alter the input data, considering the spatial information in the data. The selection of the convolution kernel is contingent upon the weights assigned to the layer. The convolution kernel can be trained using inputs from the CNN [14].

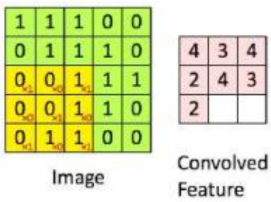

Figure 2. Convolution Operation [14]

2.1.2 Pooling Layer

The pooling layer enhances the CNN detection process by reducing the dimensions of the convolution map feature through matrix calculations, thereby accelerating the process. Max pooling refers to selecting the highest value from the feature map, whereas average pooling involves calculating the average value of the feature map. The procedure above occurs within the pooling layer [13]. The predominant method employed for pooling layers involves the utilization of a 2x2 filter, which is implemented in a two-step manner and applied to every input slice [15]. For instance, when employing pooling (2x2) with a stride of two, the algorithm will select the maximum and average values for the 2x2 pixel region during each filter shift. This process results in the generation of the subsequent image.

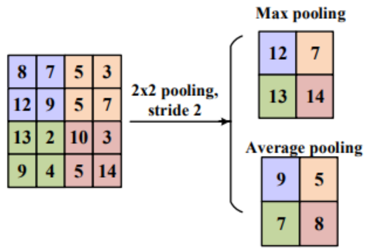

Figure 3. Pooling Layer Operation [16]

2.1.3 Flatten

Flattening is the process of transforming a matrix into a one-dimensional vector. The feature map acquired from the preceding layer is transformed into a one-dimensional vector to facilitate classification using the fully connected SoftMax layers [17].

2.1.4 Fully Connected Layer

In the CNN, the fully connected layer is the final stage, where all processes are interconnected to facilitate computation. Following multiple data mapping layers, the two-dimensional feature map obtained from the preceding process is transformed into a one-dimensional format, facilitating classification [13].

2.1.5 Activation Function

The activation function is a non-linear function that facilitates modifications in neural network operations, enabling data representation at higher dimensions and facilitating classification through straightforward pathways.

1) SoftMax
The SoftMax activation function is employed to classify multiple classes. The equation in SoftMax is expressed in the following form [12].

$$f_j(Z) = \frac{e^{zj}}{\sum_k e^{zk}} \tag{1}$$

The SoftMax activation algorithm yields more comprehensible outcomes and offers superior probabilistic interpretation compared to alternative classification algorithms. SoftMax can be used to compute probabilities for all labels. A vector with an accurate value will be extracted from the label and transformed into a vector with a range of zero to one. This vector will have a value of one when all the vectors are added together [12].

2) Swish
The Swish activation function is employed to compute functions within automated search-based learning frameworks. The function is smooth, indicating that it does not exhibit abrupt changes in direction, as observed in the case of Rectified Linear Unit (ReLU) near $x = 0$. Conversely, the function exhibits a gradual curve from 0 to 0 and subsequently increases. It is possible to describe the activation function of Swish as follows [18].

$$f_{(x)} = x \cdot sigmoid\ (x) = \frac{x}{1 + e^x} \tag{2}$$

3) Rectified Linear Unit (ReLU)
ReLU activation is an activation layer on the CNN model that uses the function:

$$f(x) = \max(0, x) \tag{3}$$

which means that this function thresholds with a zero value to the pixel value in the input image. With this activation, all pixel values in the image that are less than zero will be made 0 [12].

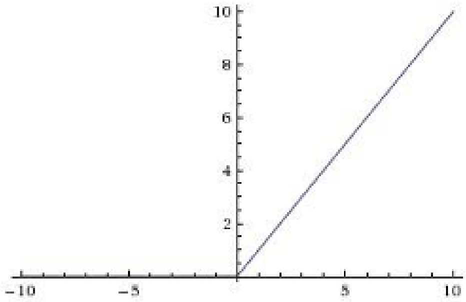

Figure 4. ReLU Activation Function [12]

2.1.6 ADAM Optimizer

The Adaptive Moment Optimization (ADAM) algorithm optimizes stochastic objective functions on first-order gradients by employing adaptive estimation of low-order moments. The technique above is employed to compute the velocity of individual adaptive estimation for parameters that deviate from the estimated moments of the initial and subsequent gradients. ADAM has the advantage of having a limited step size due to its step size hyperparameters. This allows it to work on sparse gradients and naturally perform step-size annealing forms without requiring the objective to be stationary [19]. The mathematical formula for ADAM is as follows [20]:

$$x_t = \delta_1 * x_{t-1} - (1-\delta_1) * g_t \tag{4}$$

$$y_t = \delta_2 * y_{t-1} - (1-\delta_2) * g_t^2 \tag{5}$$

$$\triangle \omega_t = -\eta \frac{x_t}{\sqrt{y_{t+\epsilon}}} * g_t \tag{6}$$

$$\omega_{t+1} = \omega_t + \triangle \omega_t \tag{7}$$

2.2 Bayesian Neural Network

Bayes' theorem is a significant theory in statistics that allows for the inference of the probable outcome of an event. Conventional machine-learning models incorporate uncertainty by integrating machine-learning theory and Bayesian theory [21]. In evaluating models, epistemic uncertainty is a valuable tool [22]. Epistemic uncertainty is a direct indicator of the dependability of predictions and can be employed as a metric for assessing the robustness of a model [21]. The uncertainty resulting from not knowing the actual function that generates the data or more specifically the uncertainty surrounding the model parameters is known as epistemic uncertainty [23]. Increasing training data can decrease epistemic uncertainty, but not aleatoric uncertainty. This is crucial for applications related to safety crises that use tiny data sets [24]. Furthermore, BNN can enhance CNN's performance by reducing uncertainty in both estimate accuracy and prediction accuracy. The BNN is considered a feasible approach for probabilistic learning due to its utilization of a precise and efficient deep variance learning algorithm [25]. The diagram presented below illustrates the ability of BNN to evaluate epistemic uncertainty by manipulating weight parameters and biases in Convolutional Neural Networks, such as Artificial Neural Networks (ANN), from fixed values to probability distributions [21].

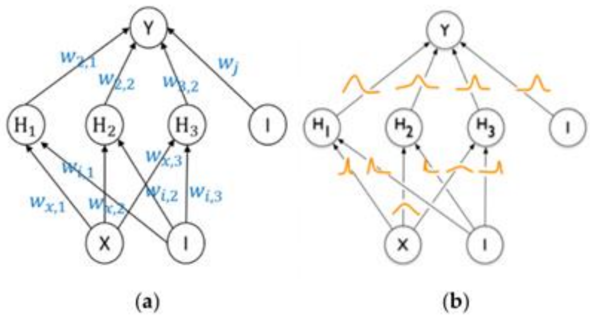

Figure 5. Illustration of ANN and BNN [26]

Based on Figure 5 each weight of ANN has a fixed value while each weight of BNN is given by a distribution $w_{i,j}$.

The BCNN algorithm involves a series of steps to calculate the uncertainty in classification predictions. These steps are as follows [21]:

1) There is a dataset $D = \{x_i, y_i\}$ and $P(D|w)$ follows on a categorical distribution.

$$P(D|w) = \prod_{i=1}^{n} P(y_i|x_i, w) \tag{8}$$

2) Maximize likelihood to get Maximum Likelihood Estimation (MLE) on parameters w.

$$w_{MLE} = \underset{w}{argmax} \sum_i^n log\, P(y_i|x_i, w) \tag{9}$$

3) Estimate the Maximum a Posteriori (MAP) of w to maximize the likelihood of the dataset multiplied by the distribution $P(w)$.

$$w_{MAP} = \underset{w}{argmax} \sum_i^n log\, P(y_i|x_i, w) + log\, P(w) \tag{10}$$

The *w-point* estimates provided by MLE and MAP are not always reliable. The purpose of using Bayesian inference is to find the posterior weight distribution based on training data $P(w|D)$, so that uncertainty parameters can be quantified.

4) $P(w|D)$ in analytical solutions are not found, hence the approximate algorithm is considered by being used to estimate the $q(w|\theta)$ posterior distribution with Kullback-Leibler (KL) divergence.

$$KL[q(w|\theta)P(w|D)] = E_{q(w|\theta)} log \frac{q(w|\theta)}{P(w|D)} \tag{11}$$

However, the manual calculation of KL divergence may be complicated depending on the model and prior distribution or posterior distribution used, then use Tensor Flow Probability (TFP) to facilitate the calculation of KL divergence in VI.

5) If VI is considered a normal distribution, then it θ can represent as. This is in line with the concept of $\theta = \mu, \sigma^2$ a normal mean field where the researcher treats each parameter independently in a variational distribution. The goal of VI is to find distributions from variational distributions to approximate posterior distributions that are difficult to solve and therefore cannot be optimized in complex models such as BNN. In this case, it is often optimized using a destination function based on the estimation of stochastic gradients in Monte Carlo samples (MC) of a variational distribution called Monte Carlo Variational Inference (MCVI) [27].

6) BNN prediction system using Bayes' theorem based on training model.

$$P(y|x, D) = \int P(y|x, w)P(w|D)dw \tag{12}$$

3. RESULTS AND DISCUSSION

3.1 Alzheimer's Dataset

This study used a dataset comprising 2400 brain MRI images obtained from individuals diagnosed with Alzheimer's disease. The images were categorized into three distinct classes: non-demented, very mildly demented, and mildly demented. Alzheimer's disease is defined as individuals with type A who do not exhibit initial indications of Alzheimer's disease. Thus, they are not impacted by cognitive deterioration that surpasses the typical range for their age and condition. Type B individuals exhibit initial indications of Alzheimer's disease, resulting in mild cognitive impairments such as challenges in retaining new information or recalling existing information. However, they are still capable of performing routine tasks without substantial hindrances. Type C individuals exhibit distinct cognitive symptoms that significantly impact their capacity to engage in daily activities, including challenges in information retention, cognitive clarity, and other related functions.

Nevertheless, they can operate autonomously with aid or modifications. The data was acquired from the kaggle.com website, specifically from the Alzheimer's Dataset repository, which consists of four classes of images. The repository owner, Sarvesh Dubey, is identified as the owner [28]. Researchers opt to utilize three classes out of a total dataset of 6400, consisting of 2400 samples, due to one class having a smaller sample size than the others. This approach ensures class balance and mitigates potential bias in the classification model. Presented below is an MRI scan depicting the cerebral region of an individual diagnosed with Alzheimer's disease. Presented below is an MRI scan depicting the brain of individuals diagnosed with Alzheimer's disease.

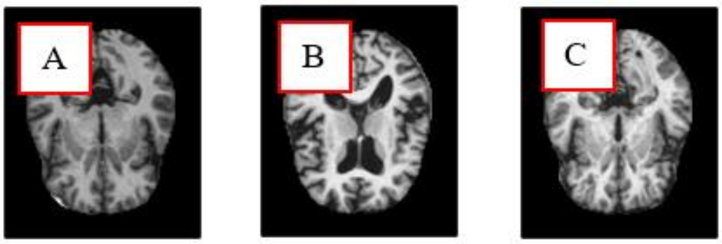

Figure 6. **Brain MRI (A: *Non-Demented,* B: *Very Mild Demented,* C: *Mild Demented*)**

3.2 Pre-processing

The purpose of this stage is to standardize all the features present in the brain MRI image to facilitate the analysis. In this stage, the brain MRI images are subjected to feature equalization to facilitate the subsequent analysis. The element equalization process is employed to enhance the clarity of brain MRI images, as the elements within each image exhibit variations. Subsequently, the image sizes are adjusted to 100x100 pixels. Subsequently, we performed dataset preprocessing by employing thresholding, which assigns a numerical value ranging from 0 to 1 for black images and 0 to 255 for white images. Using the brain MRI image, we employ the grayscale color mode to generate a gray hue.

The dataset comprises 800 data points for each of the three classes: non-dementia, very mild dementia, and mild dementia. We partitioned datasets in a ratio of 80% for training data and the remaining for testing. In addition, 10% of the training data will be allocated for validation. BCNN design utilizes data training and validation for division purposes, while model testing is conducted using data testing. There are 1728 data for training, 480 data for testing, and 192 data in the validation process.

3.3 Bayesian Convolutional Neural Network Model

We conducted multiple experiments during this phase to compare the models with the BCNN to obtain optimal classification outcomes. Determining the optimal BCNN architecture involves researchers assessing the significance of data usage scenarios, selecting appropriate optimizers, and comparing epoch values. We employ Adam's optimizer type due to its demonstrated efficacy in training neural network models. The accuracy and loss values will be determined by researchers using both the training and validation data. We utilize this comparison to determine the optimal BCNN model, ensuring it achieves a sufficiently high level of accuracy. In this comparative analysis of epoch values, we employed epoch values ranging from 2 to 200. The subsequent findings pertain to the comparison of epoch values.

Table 1. **Accuracy and Loss of Training Process**

Epoch	Training		Validation	
	Accuracy	Loss	Accuracy	Loss
2	0.3478	6.4322	0.3385	6.4013
10	0.5220	5.6538	0.3073	6.3355
50	0.6696	3.5785	0.2552	5.7546
90	0.8513	1.6856	0.3542	4.8659
150	0.9404	0.7627	0.3021	6.4123
200	0.9583	0.5532	0.5469	3.4294

According to the data presented in **Table 1.**, the epoch of 200 yields the highest accuracy for both testing and validation. Two key performance indicators accuracy and loss are used to evaluate a classification model's effectiveness. A score of 1 denotes extreme accuracy, whereas a score of 0 denotes extreme inaccuracy in the model. The second is the loss value, which displays the deviation between the goal value and the model prediction. Better predictions are shown by lower loss values. A low loss number is ideal since it shows how effectively the model works with training data and how well it can generalize to new data. Conversely, a large loss number suggests that the model's predictions deviate significantly from the intended outcome, which may be the result of underfitting, inaccurate hyperparameter settings, or unrepresentative data. This suggests that the model may perform poorly in terms of prediction and struggle to generalize. The provided diagram illustrates the architecture of the BCNN.

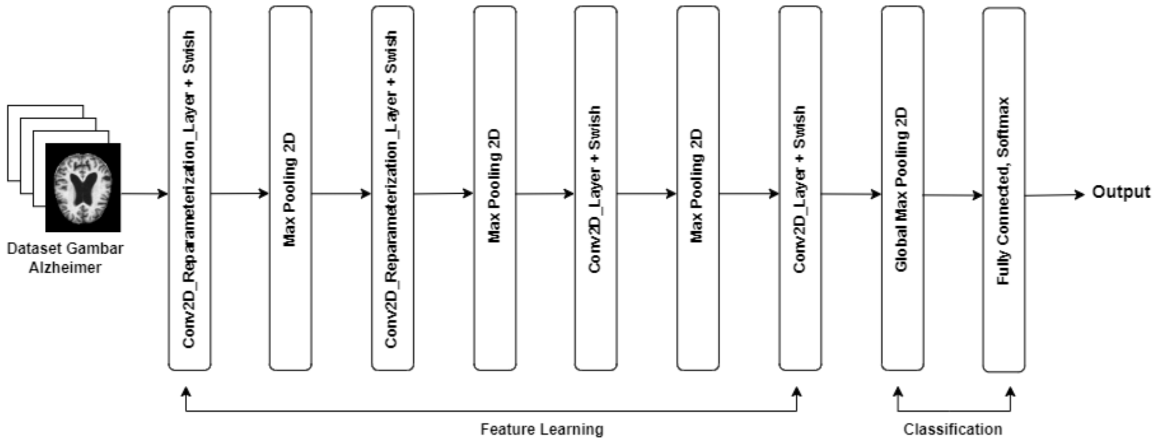

Figure 7. BCNN Architecture Scheme

This study employs the CNN model consisting of four convolutional layers and three pooling layers to extract features from brain MRI images. The optimal model is achieved with a total parameter count of 102,726. The calculation for determining the best model is as follows:

Table 2. Best Model Parameter

No	Name	Size	Parameter
1.	Conv2d_Reparameterization_1	(None, 100, 100, 16)	320
2.	MaxPool2d_3	(None, 50, 50, 16)	0
3.	Conv2d_Reparameterization_2	(None, 50, 50, 32)	9280
4.	MaxPool2d_4	(None, 25, 25, 32)	0
5.	Conv2d_4	(None, 25, 25, 64)	18496
6.	MaxPool2d_5	(None, 12, 12, 64)	0
7.	Conv2d_5	(None, 12, 12, 128)	73856
8.	Global_MaxPool2d_1	(None, 128)	0
9.	Dense_Reparameterization	(None, 3)	774
10.	One_Hot_Categorical	(None, 3)	0
Total Params: 102726			
Trainable Params: 102726			
Non - Trainable Params: 0			

The convolution process will decrease the image's dimensions, resulting in the final image size of 12x12 pixels with 128 filter parameters before entering the entire connection. Subsequently, the matrix data will undergo a conversion into a vector format to facilitate the complete connection process, thereby yielding a total of 128 neurons that are eligible for forwarding. In addition, the Monte Carlo estimation of KL Divergence will be employed in the Variational Inference technique to categorize images into three distinct classes in patients diagnosed with Alzheimer's disease based on pre-trained models.

3.4 Model Evaluation

A confusion matrix is a matrix with dimensions $N \times N$ that is utilized to assess the accuracy of a prediction model. N represents the total number of classes in the target image. The confusion matrix is utilized in the machine learning model to compare the actual target values. The table will display the quantity of training data for both the accurate and inaccurate target classification. The confusion matrix results obtained are as follows.

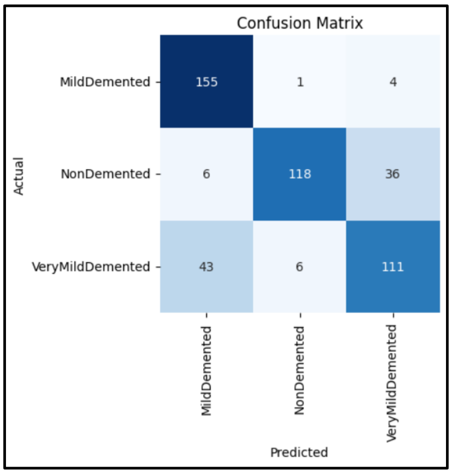

Figure 8. Confusion Matrix Results

It is possible to compute the precision, recall, and f1-score values for every class using the confusion matrix shown in **Figure 8.** In particular, the values of precision, recall, and f1-score for very mild dementia are 0.74, 0.69, and 0.71, whereas the values for no dementia are 0.94, 0.74, and 0.83. For mild dementia, the f1-score, recall, and precision were 0.76, 0.97, and 0.80. The classification accuracy achieved using the optimal BCNN model with 480 testing data is 0.80. According to the accuracy findings, the correct prediction for the image of very mild dementia is 111, the prediction rate for non-dementia is 118, and the prediction rate for mild dementia is 155.

3.5 Uncertainty Prediction

The outcome of the optimal BCNN model is additionally employed to annotate data from external sources. This procedure aims to determine the categorization of novel unidentified data observed through image patterns. Uncertainty is anticipated through the utilization of stochastic elements present within the dataset. The objective of this procedure is to introduce a degree of uncertainty in the projected label, which can be elucidated as an approximation of the model's confidence level in its predictions. The graph has been modified to display a 95% confidence interval for the estimated probability of the model. We conducted an uncertainty prediction by utilizing the sample images of non-dementia within the dataset.

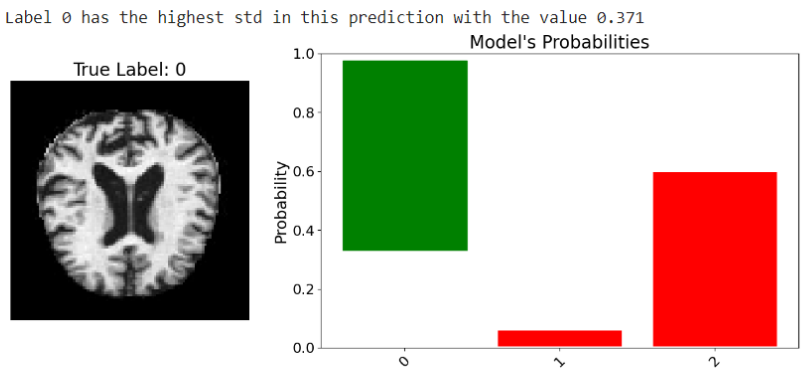

Figure 9. Uncertainty Prediction (1)

Figure 9. displays a longer green bar, indicating a high level of uncertainty in the prediction, with a value of 0.371. Therefore, as the stem increases, so does the level of uncertainty. If the plot exhibits red bars, it signifies that the class in question is not predictable, whereas green bars indicate a predicted class. In the

prediction, the model assigns a probability of 0 (non-dementia), indicating that the model is confident that the image is non-dementia.

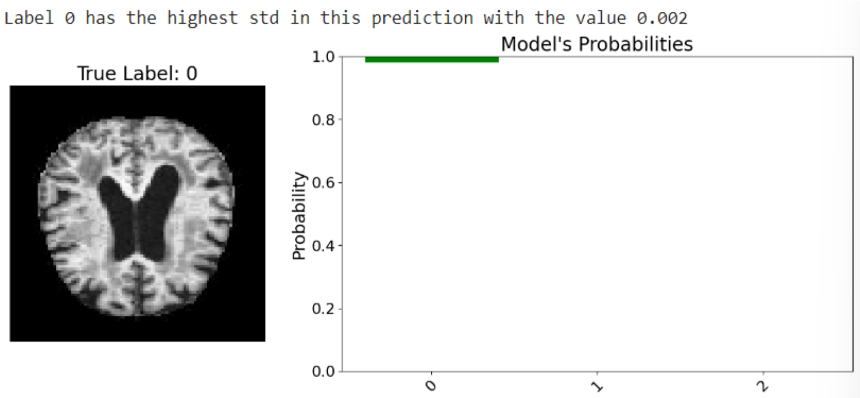

Figure 10. Uncertainty Prediction (2)

By employing various non-dementia images, it has been observed that the presence of shorter green bars signifies a low level of uncertainty in prediction, specifically at 0.002. Therefore, as the bar increases, so does the level of uncertainty. Red bars in the plot indicate an unpredictability of the class, while green bars indicate a predictability of the class. The prediction reveals that the model assigns a probability of 0 (non-dementia), indicating the model's belief that the image is not affected by any degeneration.

4. CONCLUSIONS

To effectively classify different types of Alzheimer's disease (non-dementia, very mild dementia, and mild dementia), it is crucial to identify the optimal parameter architecture. This entails obtaining a dataset scenario with a ratio of 80% for training, a pixel size of 100x100, a kernel size of 3x3, and an Adam optimizer. According to Adam, the most optimal epoch is Epoch 200, achieving an accuracy rate of 80% for testing. The best model's uncertainty prediction yields a value of 0.371 for the non-dementia label. In contrast, the second prediction of uncertainty, also based on the best model, yields a value of 0.002 for the non-dementia label.

REFERENCES

[1] WHO, "Dementia," World Health Organization. Accessed: Sep. 26, 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia

[2] G. Litjens *et al.*, "A survey on Deep Learning in Medical Image Analysis," *ScienceDirect*, vol. 42, pp. 60–88, Dec. 2017, Accessed: Mar. 25, 2024. [Online]. Available: https://doi.org/10.1016/j.media.2017.07.005

[3] L. Zhang, M. Wang, M. Liu, and D. Zhang, "A Survey on Deep Learning for Neuroimaging-Based Brain Disorder Analysis," Oct. 08, 2020, *Frontiers Media S.A.* doi: 10.3389/fnins.2020.00779.

[4] S. Murugan *et al.*, "DEMNET: A Deep Learning Model for Early Diagnosis of Alzheimer Diseases and Dementia from MR Images," *IEEE Access*, vol. 9, pp. 90319–90329, 2021, doi: 10.1109/ACCESS.2021.3090474.

[5] D. Wang *et al.*, "Application of Multimodal MR Imaging on Studying Alzheimer's Disease: A Survey," *Bentham Science*, vol. 10, no. 8, pp. 877–892, 2013.

[6] S. Ahmed *et al.*, "Ensembles of Patch-Based Classifiers for Diagnosis of Alzheimer Diseases," *IEEE Access*, vol. 7, pp. 73373–73383, 2019, doi: 10.1109/ACCESS.2019.2920011.

[7] J. Wu, H. Zhang, X. Zhu, Y. Zhang, X. Ding, and H. Yang, "Ensemble Learning-Base Multimodal Data Analysis Improving the Diagnostic Accuracy of Alzheimer's Disease," in *Proc. SPIE 12770, Optics in Health Care and Biomedical Optics XIII*, 2023. Accessed: Mar. 25, 2024. [Online]. Available: https://doi.org/10.1117/12.2687618

[8] M. Reynolds, T. Chaudhary, M. E. Torbati, D. L. Tudorascu, and K. Batmanghelich, "ComBat Harmonization: Empirical Bayes versus Fully Bayes Approaches," *bioRxiv*, 2023, doi: 10.1101/2022.07.13.499561.

[9] B. Raharjo, *Deep Learning dengan Python*. Semarang, 2022.

[10] H. A. Haay, S. Trihandaru, and B. Susanto, "INTRODUCTION OF PAPUAN AND PAPUA NEW GUINEAN FACE PAINTING USING A CONVOLUTIONAL NEURAL NETWORK," *BAREKENG: Jurnal Ilmu Matematika dan Terapan*, vol. 17, no. 1, p. 213, Apr. 2023, doi: 10.30598/barekengvol17iss1pp0211-0224.

[11] D. Gunawan and H. Setiawan, "Convolutional Neural Network dalam Analisis Citra Medis," *KONSTELASI (Konvergensi Teknologi dan Sistem Informasi)*, vol. 2, no. 2, pp. 377–378, Dec. 2022, Accessed: Jan. 08, 2024. [Online]. Available: https://ojs.uajy.ac.id/index.php/konstelasi/article/view/5367/2568
[12] S. Ilahiyah and A. Nilogiri, "Implementasi Deep Learning Pada Identifikasi Jenis Tumbuhan Berdasarkan Citra Daun Menggunakan Convolutional Neural Network," *JUSTINDO (Jurnal Sistem dan Teknologi Informasi Indonesia)*, vol. 3, no. 2, pp. 50–52, 2018, Accessed: Jan. 08, 2024. [Online]. Available: https://core.ac.uk/download/pdf/229217808.pdf
[13] B. B. Traoré, B. Kamsu-Foguem, F. Tangara, and B. B. Traore, "Deep convolution neural network for image recognition," *Ecol Inform*, vol. 48, pp. 257–268, 2019, doi: 10.1016/j.ecoinf.2018.10.002ï.
[14] I. W. S. E. P, A. Y. Wijaya, and R. Soelaiman, "Klasifikasi Citra Menggunakan Convolutional Neural Network (CNN) pada Caltech 101," *JURNAL TEKNIK ITS*, vol. 5, no. 1, 2016, Accessed: Jan. 06, 2024. [Online]. Available: https://media.neliti.com/media/publications/191064-ID-klasifikasi-citra-menggunakan-convolutio.pdf
[15] A. Peryanto, A. Yudhana, and R. Umar, "Rancang Bangun Klasifikasi Citra Dengan Teknologi Deep Learning Berbasis Metode Convolutional Neural Network," *Jurnal Ilmiah Teknik Informatika (FORMAT)*, vol. 8, no. 2, p. 140, 2019, Accessed: Jan. 07, 2024. [Online]. Available: https://www.mathworks.com/discovery/convolutional-neural-network.html
[16] H. Yingge, A. Imran, and K. Y. Lee, "Deep Neural Networks on Chip - A Survey," *2020 IEEE International Conference on Big Data and Smart Computing, BigComp 2020*, pp. 589–592, Feb. 2020, doi: 10.1109/BigComp48618.2020.00016.
[17] A. Yusuf, R. C. Wihandika, and C. Dewi, "Klasifikasi Emosi Berdasarkan Ciri Wajah Menggunakan Convolutional Neural Network," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 3, no. 11, p. 10598, 2019, Accessed: Jan. 08, 2024. [Online]. Available: http://j-ptiik.ub.ac.id
[18] M. H. Naufal, "Pendeteksian Balon Ucapan Pada Komik Jepang (Manga) Dengan Run Length Smooth Dan EfficientNet-B6," Universitas Komputer Indonesia, 2023.
[19] D. P. Kingma and J. L. Ba, "Adam: A Method for Stochastic Optimization," Dec. 2014, Accessed: Mar. 13, 2024. [Online]. Available: http://arxiv.org/abs/1412.6980
[20] M. Yaqub *et al.*, "State-of-the-art CNN Optimizer for Brain Tumor Segmentation in Magnetic Resonance Images," *Brain Sci*, vol. 10, no. 7, pp. 1–19, Jul. 2020, doi: 10.3390/brainsci10070427.
[21] H. Yu, A. H. Seno, Z. Sharif Khodaei, and M. H. F. Aliabadi, "Structural Health Monitoring Impact Classification Method Based on Bayesian Neural Network," *MDPI*, vol. 14, no. 19, Oct. 2022, doi: 10.3390/polym14193947.
[22] J. W. Book, "Training Bayesian Neural Networks A Study of Improvements to Training Algorithms," Lund University, 2020. Accessed: Oct. 01, 2023. [Online]. Available: https://lup.lub.lu.se/student-papers/record/9008040/file/9008041.pdf
[23] H. Jordão, A. J. Sousa, and A. Soares, "Using Bayesian Neural Networks for Uncertainty Assessment of Ore Type Boundaries in Complex Geological Models," *Natural Resources Research*, vol. 32, no. 6, pp. 2495–2514, Dec. 2023, doi: 10.1007/s11053-023-10265-6.
[24] A. Kendall and Y. Gal, "What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?," 2017.
[25] D. T. Chang, "Bayesian Neural Networks: Essentials," 2021, Accessed: Oct. 03, 2023. [Online]. Available: https://arxiv.org/abs/2106.13594
[26] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra, "Weight Uncertainty in Neural Networks," *International Conference on Machine Learning*, vol. 37, 2015, Accessed: Oct. 26, 2023. [Online]. Available: https://arxiv.org/abs/1505.05424
[27] M. Fujisawa and I. Sato, "Multilevel Monte Carlo Variational Inference," *Journal of Machine Learning Research*, vol. 22, pp. 1–44, 2021, Accessed: Mar. 13, 2024. [Online]. Available: http://jmlr.org/papers/v22/20-653.html.
[28] S. Dubey, "Alzheimer's Dataset (4 class of Images)," kaggle.com. [Online]. Available: https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images