Jurnal Ilmiah Teknik Elektro Komputer dan Informatika (JITEKI)
Vol. 10, No. 2, June 2024, pp. 463-475
ISSN: 2338-3070, DOI: 10.26555/jiteki.v10i2.29177

Addressing Overfitting in Dermatological Image Analysis with Bayesian Convolutional Neural Network

Mulki Indana Zulfa[1], Andreas Sahir Aryanto[2], Bintang Abelian[1], Waleed Ali[3]
[1]Department of Electrical Engineering, Universitas Jenderal Soedirman, Purbalingga, 53371, Indonesia
[2]Angusta Systems, Eveleigh, Sydney, New South Wales, 2015, Australia
[3]Information Technology Department, King Abdulaziz University, Rabigh, 25732, Saudi Arabia

ARTICLE INFO

Article history:

Received June 19, 2024
Revised July 04, 2024
Published July 22, 2024

Keywords:

Skin cancer;
Overfitting;
Transfer learning;
Validation accuracy;
Model uncertainty

ABSTRACT

VGG, ResNet, and DenseNet are popular convolutional neural network (CNN) designs for transfer learning (TL), aiding dermatological image processing, particularly in skin cancer categorization. These TL-CNN models build extensive neural network layers for effective image classification. However, their numerous layers can cause overfitting and demand substantial computational resources. The Bayesian CNN (BCNN) technique addresses TL-CNN overfitting by introducing uncertainty in model weights and predictions. Research contributions are (i) comparing BCNN with three TL-CNN architectures in dermatological image processing and (ii) examining BCNN ability to mitigate overfitting through weight perturbation and uncertainty during training. BCNN uses flipout layers to perturb weights during training, guided by the KL divergence and Binary Cross Entropy (BCE) loss function. The dataset used is the ISIC Challenge 2017, categorized as malignant and benign skin tumors. The simulation results show that three TL-CNN architectures, namely VGG-19, ResNet-101, and DenseNet-201, obtained training accuracies of 96.65%, 100%, and 97.70%, respectively. However, all three were only able to achieve a maximum validation accuracy of around 78%. In contrast, BCNN can produce training and validation accuracy of 81.30% and 80%, respectively. The difference in training and validation accuracy values produced by BCNN is only 1.3%. Meanwhile, the three TL-CNN architectures are trapped in an overfitting condition with a difference in training and validation values of around 20%. Therefore, BCNN is more reliable for dermatological image processing, especially for skin cancer images.

Corresponding Author:

Mulki Indana Zulfa, Department of Electrical Engineering, Jenderal Soedirman University, Purbalingga, 53371, Indonesia
Email: mulki_indanazulfa@unsoed.ac.id

1. INTRODUCTION

Machine learning research in dermatological image processing is crucial in early detection and precise skin cancer diagnosis. Dermatological analysis plays a crucial role in identifying and skin cancer classification and other potentially harmful skin abnormalities. Skin cancer exhibits a significant incidence rate and is frequently detected at an advanced stage. Skin cancer is a common and serious disease, with melanoma being the most fatal type. While melanoma only accounts for about 1% of skin cancer cases, it is responsible for a significant number of fatalities [1], [2]. By 2024, it is estimated that around 100.000 people will have been affected by melanoma [1]. The disease can affect any part of the skin, particularly areas that have been exposed to UV light for extended periods of time [3], [4]. This can cause the skin to become darker and increase in size. Early detection is key in preventing death from melanoma, and it is crucial to distinguish between benign and malignant types [5].

The advancement of machine learning research aimed at creating an automated system capable of detecting initial indications of skin cancer holds significant clinical significance [6], [7]. Machine learning techniques, namely transfer learning (TL) using convolutional neural networks (CNN) like VGG, ResNet, and DenseNet, are employed to expedite diagnosis, enhance accuracy, and minimize human error in interpreting dermatological images [8], [9]. Hence, it is crucial to continue conducting thorough and focused research on the application of machine learning in analyzing dermatological images, particularly in the classification of skin cancer. This research has significant implications for enhancing patient care and contributing to the battle against skin cancer [10]. However, these models often suffer from overfitting, necessitating the exploration of alternative approaches such as Bayesian Convolutional Neural Network (BCNN) to enhance performance and reliability.

Regarding the problem of overfitting that is typically encountered by conventional CNN models in the field of dermatological image analysis [11], the BCNN technique has emerged as a potentially useful solution. The uncertainty that is introduced into the model weights and predictions by BCNN contributes to the regularization of the model and the improvement of its generalization [12]. During the training process, BCNN introduces perturbations by utilizing flipout layers, which results in an increase in resilience [13]. This method makes use of Bayesian inference to give a probabilistic framework that is better able to manage the variety and complexity of skin cancer images. As a result, it is particularly well-suited for clinical applications in where accuracy and reliability are of the utmost importance. Despite the fact that BCNN provides a number of obstacles, it does offer a number of major advantages, including the reduction of overfitting and the improvement of model generalization. The computational complexity and training time of BCNN are higher than those of regular CNNs, which may limit their applicability in environmental settings with limited resources. In addition, because of the probabilistic nature of BCNN, it is necessary to do thorough validation and careful tweaking of hyperparameters in order to guarantee consistently dependable results [14]. The overfitting of TL-CNN models in the context of dermatological image analysis is the key issue that is addressed in this paper. This overfitting hinders the ability of these models to generalize successfully to novel data that they have not before encountered. By conducting an exhaustive analysis and comparison of the performance of BCNN with that of typical TL-CNN architectures, specifically in the classification of skin cancer, the purpose of this research is to fill the gap that has been identified. The objective of this study is to provide evidence that demonstrates how BCNN may successfully reduce the effects of overfitting, hence improving the precision and dependability of automated skin cancer detection systems.

Lei *et al*. put forward a method for enhancing underwater images by focusing on pixel-based techniques. Their goal was to tackle issues like low contrast and color distortion by improving contrast and correcting color [15]. Transform-based methods for image enhancement are also being utilized. Hesseini *et al*. introduced a transform-based method for enhancing strain images while preserving spatial resolution [16]. Deng *et al*. halo-free method is another example that produced halo-free images for all datasets and even achieved the highest entropy for certain datasets [8]. Other image enhancement methods include Yang pixel-based approach, which enhances images and addresses discontinuity in the lower gray region [17]. Global contrast, local contrast, and bright contrast stretching are additional pixel-based techniques that enhance leukemic images' interpretability [18]. Mouzai *et al*. proposed a self-supervised pixel stretching method to improve X-ray images for accurate diagnoses [9]. Mahmood *et al*. proposed a transform-based method to overcome issues like overstretching and unnatural image appearance [10]. There are still more image enhancement methods proposed such as fuzzified contrast enhancement and object-based multilevel [19], [20].

Effective classification tasks require image enhancement as a crucial step [19]. This process allows models, such as the conventional CNN, to accurately identify the class of an image [21]. However, it is a common issue for machine learning models to overfit and require a large amount of data [22], [23]. Transfer Learning CNN (TL-CNN) is a state-of-the-art technique for CNNs that reuses obtained knowledge from source data to not depend on extensive labeled data, thus overcoming this limitation. TL-CNN still has its weaknesses, such as the catastrophic forgetting dilemma that leads to drastic changes in weights that can overfit the model [22], [23]. To tackle this issue, Bayesian Convolutional Neural Network (BCNN) is utilized. The BCNN can prevent overfitting and has been proven to produce highly accurate results in numerous studies [24], [25]. Bargagna *et al*. devised a method that utilizes BCNN to analyze medical images in scenarios where resources are scarce. Their model achieved an impressive accuracy rate of 78.28% [26]. Zain *et al*. applied BCNN to the identification of tuberculosis and achieved remarkable accuracy rates of 96.42% and 86.46% on two distinct datasets [27]. Thiagarajan *et al*. also employed BCNN in their method for classifying breast histopathology images and achieved an accuracy rate of 88% in the data test [26]. Feng *et al*. utilized BCNN for seismic facies classification and achieved an accuracy rate close to 100% [28]. Joshaghani *et al*. utilized BCNN to classify

hyperspectral remote sensing images and their model achieved the highest accuracy rate while also being more resilient to overfitting [29]. Chandra *et al*. proposed the use of BCNN for analyzing unstructured data such as graphs and achieved comparable accuracy rates to those of the canonical Graph Convolutional Neural Network [30]. Mo *et al*. proposed the application of BCNN to predict or fill the one-year gap between Gravity Recovery and Climate Experiment and its Follow-On and their model demonstrated superior performance [31].

Dermatological image analysis has made strides recently, highlighting the ongoing difficulty of overfitting in machine learning algorithms [32]. Though successful in many situations, transfer learning methods frequently require assistance in generalizing to fresh data, which can result in overfitting. Because BCNN can include uncertainty in model weights and predictions, it is becoming a potential method among the several that researchers have been investigating to address this problem. Improving classification accuracy and reducing overfitting in dermatological image processing increases diagnostic accuracy and has significant clinical practice consequences. This study aims to demonstrate, using BCNN, how improved model generalization and dependability may lead to better patient outcomes and more informed clinical choices [33]. This study highlights the therapeutic efficacy and prospective impact of accurate and early identification of skin cancer in reducing the death rate linked with melanoma and other skin malignancies. This work aims to address a significant gap in current research by conducting a comparative analysis of TL-CNN and BCNN for the purpose of skin cancer categorization.

However, it is crucial to avoid overfitting the model. A few approaches have been suggested for CNN models to utilize BCNN. Therefore, the main paper contribution are (i) to compares BCNN and three TL-CNN architectures in dermatological image processing, and (ii) to examines BCNN ability to mitigate overfitting through weight perturbation and uncertainty during training. The second section of this work offers a comprehensive account of the simulation technique used, which encompasses a thorough discussion of the dataset, the TL-CNN architecture, and a complete elucidation of BCNN. The following section covers the simulation results and offers a comprehensive analysis, followed by a brief overview of the conclusions.

2. METHODS

2.1. Dataset

The experimental data originated from International Skin Imaging Collaboration (ISIC) Challenge 2017 Datasets, containing over 2000 ldermoscopy images of skin lessions [34]. These images were categorized as malignant and benign skin tumors. The dataset was seperated into training and validation sets. The training set contained 374 malignant and 1.626 benign images, while the validation set contained 120 maligant and 30 benign images. The split of this dataset is based on the split given by the source of the dataset. This imbalanced dataset was strategically chosen to research about BCNN capability to mitigate the overfitting problem due to imbalanced dataset is one of potential catalyst for overfitting problem [35].

2.2. Preprocessing

Raw ISIC images typically contain a diverse artifacts, including hairs and pixel noise of various intensities. Preprocessing steps are essential to reduce these artifacts and make the images better for the next analysis. To remove non-uniform pixel noise, the median filter offers a well-established solution [36], [37], [38]. Apart from noise reduction, the median filter has the additional benefit of smoothing the underlying pixel values (signal). Therefore, the median filter is first applied to each color channel separately to effectively removing noise artifacts. Segmentation of foreground objects in many cases requires image preprocessing techniques to enhance the difference between the object and the background. Contrast stretching, a well-known image processing method, can be utilized to improve the pixel intensity distribution, thus allowing for a smoother segmentation process [39], [40]. Arslan, *et al*. [41] proposed the contrast stretching method based on the mean and standard deviation of pixels intensities defined in (1) and (2). The resulting value were then mapped to the 0-255 range for each RGB channel separately before being combined back into a single image.

$$Lowintensity = \overline{x} - (\sigma * N) \tag{1}$$

$$Highintensity = \overline{x} + (\sigma * N) \tag{2}$$

To remove hair artifacts, a two-step procedure is done. First, RGB photos are transformed to greyscale. Grayscale images with a single intensity channel are computationally more efficient than RGB images when performing morphological processes [42]. Second, hair pixels are reduced using a Bottom-hat filter, a morphological technique that combines dilation and erosion. This filter makes use of cross-shaped structural pieces measuring 25x25. The use of odd-sized structural pieces makes it easier to find the center, which is

critical for morphological operations [43]. The size 25×25 was experimentally determined through studies to strike a compromise between efficiently eradicating hair artifacts and reducing potential harm to existing skin lesions. Furthermore, the Otsu thresholding method was applied to isolate the region of interest (ROI) containing the lesion. The resulting segmentation mask is then used for pixel multiplication (AND) with the original image to preserve the lesion area.

2.3. Transfer Learning

Transfer learning has emerged as a dominat approach for medical image classification tasks either for convolutional neural networks or deep neural networks, that consistently achieving superior performance [44], [45], [46]. These technique utilizes the knowledge (pre-trained) extracted from large-scale source dataset and applies it to new target task in this case medical image classification problem [47], [48], [49]. In this study, we employed pre-trained weights from well-established architectures, including VGG, ResNet, and DenseNet, all trained on the extensive ImageNet dataset. Furthermore, the whole model architecture in this work uses 100 epochs, with a learning rate of 0.001 and a batch size of 16 [26].

VGG, Convolutional neural networks introduced by the Visual Geometry Group (VGG) for image classification task. There are two variants of VGG models: VGG-16 and VGG-19, named for their number of convolutional layers (16 and 19, respectively). The convolutional layers of VGG utilized the 3 x 3 filter with a stride 1 and for max-pooling layer is with a 2 x 2 filter of stride 2. Both VGG-16 and VGG-19 is followed by three Fully-Connected (FC) layers: the first FC layers has 4096 channels each, the third performs 1000 classification by using soft-max activation function. All hidden layers are utilized the ReLU non-linear activation layer [50]. In this study, instead using the 1000 classification and soft-max activation, the sigmoid activation layer used due to the classification task only for two classification (maligant and benign).

ResNet, Residual Networks (ResNets) are convolutional neural networks that utilize residual learning, also known as skip connections, to address the degradation problem that blocks the network ability to learn complex features when stacking many layers in deep learning [51]. Unlike VGG-19 models, which rely on max-pooling layers, ResNets employ residual learning through skip connections. These connections, mathematically represented by the formula H(x) := F(x) + x, allow the network to learn the identity mapping (where the output is the same as the input) in addition to the desired transformation (represented by F(x)) [52]. Inspired by VGG-19 architectures, these models utilize global average pooling instead of max-pooling. In this study, we use ResNet-50 and ResNet-101 pre-trained weights with sigmoid activation for binary classification for the final fully connected layer.

DenseNet, Densely Connected Convolutional Networks is the one approach to increasing the depth of deep convolutional networks by addressing the problems by addressing the issues of gradient vanishing, redundant features, and reducing the number of parameters. Instead combine features through summation like ResNet, this model combine features by concatenating them. This introduced L(L+1)/2 connection in an L-layer network [53]. In this study, we use DenseNet-121 and DenseNet-169 pre-trained weights with sigmoid activation for binary classification for the final fully connected layer.

2.4. Bayesian Neural Network

Neural Networks (NNs) work by learning and optimizing weights to minimize a loss value based on the training data. This fundamental process is crucial to the training of neural networks, where the weights are initialized and then adjusted iteratively to reduce the loss until it converges to a minimum value, indicating that an ideal weight matrix has been obtained [54], [55]. However, this can lead to issues such as slow convergence, time consumption, and getting stuck in local minima. Bayesian inference offers a powerful probabilistic framework for NNs. It allows us to estimate a posterior distribution over the possible weights that is typically expressed in the form of a posterior distribution [56]. The use of Bayesian inference is one method for improving medical image classification [57], [58], [59]. To estimate the posterior distribution we can use Bayes' theorem defined in (3):

$$P(w|D) = \frac{P(D|w) * P(w)}{P(D)} \tag{3}$$

Varitional Inference is a practical approach to approximate the posterior distribution [25]. The idea iis to introduce a tractable distribution $q(w|\varphi)$, called the variational distribution, parametrized by set of parameters φ. The values of the parameters φ are then learned such that the variational distribution is as close as possible to the exact posterior. To measure this closeness, we utilize the Kullback-Leibler Divergence (KL Divergence).

This function calculates the divergence between two probability distributions. It leverages concepts from Shannon information theory to quantify the difference between the distributions [60] as defined in (4):

$$KL[q(w|\varphi) \parallel P(w|D)] = \int q(w|\varphi) log\left(\frac{q(w|\varphi)}{P(w)}\right) dw + \int q(w|\varphi) log\big(P(D)\big) dw - \int q(w|\varphi) log\big(P(D|w)\big) dw \quad (4)$$

To apply the varitional inference in NNs, we can use one of the simplest method that is Flipout layer. Introduced by Wen *et al*., flipout layers offer an efficient method for decorrelating the gradients between different training examples in neural networks [61]. Flipout apply a specific type of perturbation to the weights during the training process. Even though Flipout layers don't fully implement a Bayesian CNN with variational inference, by introducing randomness through the weight perturbation, Flipout injects uncertainty into the weights during training. It samples from these weight distributions (kernel and bias posteriors) for each training step, introducing randomness into the forward pass. This is analogous to integrating over the weight distributions in variational inference. The mathematical expression of weight perturbation and forward defined in (5) and (6), respectively. In this study, the Flipout layer will be used immediately on each convolutional layer of the BCNN.

$$w_i' = w_i + \varepsilon_i \quad (5)$$

$$y = \sigma\big(\sum(w_i' * z_i)\big) \quad (6)$$

2.5. Regularization Effect

Standard neural networks with a single weight estimate can become overly reliant on specific weight values during training. This overreliance can lead to overfitting, where the network performs well on the training data but struggles with unseen data [62]. Flipout layers offer a solution to this problem. These layers introduce a specific type of perturbation to the weights during training. This perturbation effectively changes the weights used in each forward pass, even for the same training example. Consequently, the gradients calculated during backpropagation also vary across these multiple passes. The variations in gradients introduced by Flipout layers, known as decorrelated gradients, play a crucial role in addressing overfitting [61]. By preventing the network from relying on a single set of weight values, Flipout discourages the network from memorizing specific training examples. This helps the network generalize better to unseen data. Additionally, the decorrelated gradients encourage weight sharing across different neurons. This promotes a more efficient representation of the data and reduces overfitting. Finally, the multiple forward passes with perturbed weights can be seen as an implicit form of ensemble learning. The network explores different weight configurations, leading to a more robust model.

2.6. Loss Function

In the training domain of neural networks, loss functions serve as a basis for guiding the optimization process [63]. These functions measure the difference between the network predicted output and the true output (ground truth) [64], [65]. This section discusses two loss functions used in this study: binary cross entropy and KL divergence. Binary Cross Entropy (BCE) is a basic loss function specifically designed for classification task that involve binary outputs (0 or 1) [66]. This function mathematically quantifies the average difference between the network predicted probability and ground truth label associated with each data point. Thus in principle, BCE measures the information-theoretic difference between the ideal probability distribution representing the true label and the probability distribution generated by the network prediction. The mathematical formulation of BCE is defined in (7).

$$BCE = -\Big(y_i * log\big(p(x_i)\big) + (1 - y_i) * log\big(1 - p(x_i)\big)\Big) \quad (7)$$

For a class label of 1, the loss is minimized as $p(x_i)$ approaches 1. This signifies high confidence in the positive class, indicating that the network is correctly learning to identify examples belonging to that class. Conversely, for a class label of 0, the loss is minimized as $p(x_i)$ approaches 0. This signifies high confidence in the negative class, indicating that the network is effectively distinguishing examples that do not belong to the positive class. The negative sign ensures that the loss function always yields non-negative values, with lower values indicating better network performance. KL divergence (KLD), also known as Kullback-Leibler

divergence, works as a more general measure of the information-theoretic difference between two probability distributions. The KL divergence presents a unique perspective. In contrast to functions such as binary cross entropy, it doesn't directly measure the difference between the predicted label and the actual label. Instead, KL divergence measures the information gain (in bits) required to represent a sample from the true distribution (p(x)) using a code designed for the predicted distribution (q(x)) [67]. In simpler terms, it tells that how much additional information would be required to encode data from the true distribution if a code meant for a different distribution was used. The mathematical formulation of KLD defined in (8). However, a larger KL divergence doesn't necessarily translate to a better loss value in the context of training. This is because KL divergence is not a symmetric measure. In scenarios like variational inference within Bayesian neural networks, a smaller KL divergence is generally preferred as it indicates a closer match between the approximate and true posterior distributions.

$$KLD(p \parallel q) = \sum \left(p(x) * log \left(\frac{p(x)}{q(x)} \right) \right) \quad (8)$$

Both could be chosen depending on the conditions involved in the studies. For example, BCE is commonly used in simple binary classification applications because to its simplicity and computational efficiency. On the other hand, KL divergence could be helpful in model regularization by punishing departures from the prior distribution, which leads to generalization [68]. When deciding between these loss functions, there are tradeoffs to consider as well as computational factors. BCE is computationally less intensive and easy to deploy, making it ideal for large-scale binary classification tasks. KL divergence, in contrast, can be computationally difficult since it requires calculating the divergence across distributions; however, it provides a deeper foundation for models where comprehending uncertainty is crucial, such as probabilistic modeling and variational inference [68]. The selection between BCE and KL divergence is thus determined by the task's unique needs, with BCE being better for efficiency in binary classification and KL divergence providing advantages in scenarios requiring robust uncertainty estimates.

2.7. Experimental Design

The BCNN utilized a specific hyperparameter configuration detailed in Table 1, Table 2 is the hyperparameters for all model training, and Table 3 is the computational resource used in this study. The input data for all networks consisted of RGB images sized 50x50x3, with the corresponding binary output classifying images as cancerous or non-cancerous [26]. The BCE loss served as the primary loss function for all networks. Additionally, experiments with BCNN explored the use of KL Divergence for comparative analysis. The following sections describe the training and validation results each network architecture achieves. Then, the training and validation accuracy is compared for the evaluation metrics. Furthermore, the evaluation metric compares each model's training and validation accuracy.

Table 1. Hyperparameters for Network Architecture

Network	conv1	conv2	conv3	conv4	conv5	conv6	FC1
BCNN	16	32	32	64	128	256	512

Table 2. Hyperparameters for Training

Network	Learning rate	Batch size	Layers learnt	Epochs
TL-CNN	0.001	16	2	100
BCNN	0.001	16	All	100

Table 3. Computational Resources

Language	Python Version 3.9
GPU	RTX 3050Ti 4GB
Memory	16 GB RAM

3. RESULTS AND DISCUSSION

3.1. Result

The accuracy for traing and validation for each epoch is shown in Fig. 1 – Fig. 3 for TL-CNN. The results revealed that training accuracy increased with each epoch. However, validation accuracy stopped improving

after a few epochs, suggesting overfitting. In other words, the TL-CNN model failed to generalize its predictions to unseen data and became overly reliant on the training data. This is evidenced by the continued rise in training accuracy beyond 15-20 epochs, while validation accuracy remained stagnant. In contrast, the BCNN architecture incorporates Flipout layers. These layers introduce weight perturbation during training, which acts as a regularizer and helps prevent overfitting. Consequently, the BCNN inherently addresses the overfitting issue observed with the TL-CNN models. Fig. 4 illustrates the accuracy for the BCNN training and validation data, using different loss functions. The main difference compared to TL-CNN is that the accuracy for both datasets increases steadily with each epoch. This shows that BCNN, most likely due to its Flipout layer structure, effectively overcomes the overfitting problem. Table 3 details the training and validation accuracy achieved by each network architecture. While the TL-CNN exhibits consistently higher training accuracy, often surpassing the BCNN by 17-19%, a significant drop in its validation accuracy is evident. This highlights the issue of overfitting in TL-CNN, where the model memorizes training data specifics and fails to generalize well to unseen data. The validation accuracy of TL-CNN remains below 80%, falling short of the threshold considered acceptable for medical diagnosis.

However, the BCNN demonstrates a more balanced approach. Its validation accuracy consistently hovers around 80%, exhibiting a minimal difference of only 1.3% compared to its training accuracy. This suggests that the regularization effect from BCNN effectively discourages the network from memorizing specific training examples, making it a more suitable candidate for medical applications where reliable performance on unseen data is crucial. Beyond its superior performance in avoiding overfitting, the BCNN demonstrates surprising efficiency in terms of training time. Table 4 reveals that the BCNN achieves excellent results with remarkably faster training compared to TL-CNN architectures. Notably, the BCNN total training time for both BCE and KL Divergence loss functions falls within a range of 1.67 to 1.7 minutes, significantly less than the TL-CNN, which can take over 5 minutes, particularly for ResNet and DenseNet models. This advantage extends to individual training steps as well. The BCNN requires only 12.50 milliseconds per step, whereas the TL-CNN can take anywhere from two to ten times longer.

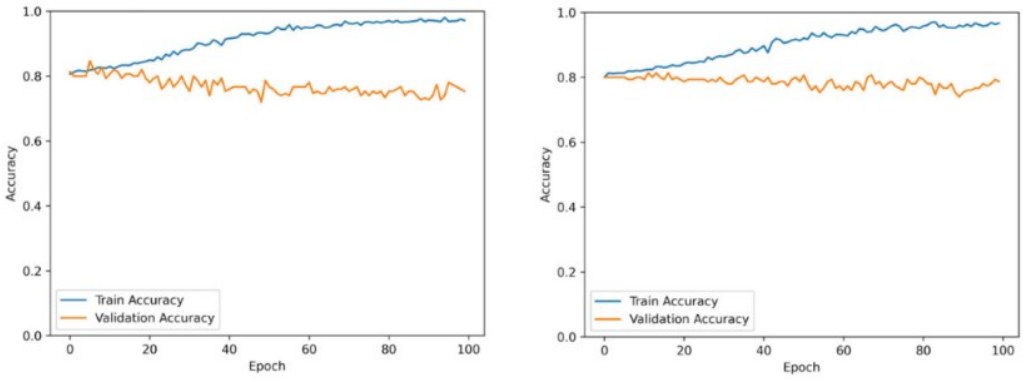


Fig. 1. Skin cancer classification accuracy used TL-CNN: (a) VGG-16 and (b) VGG-19

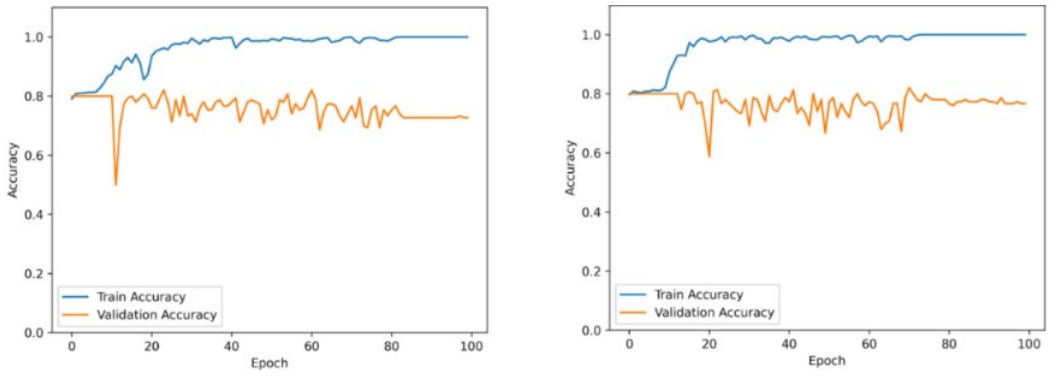
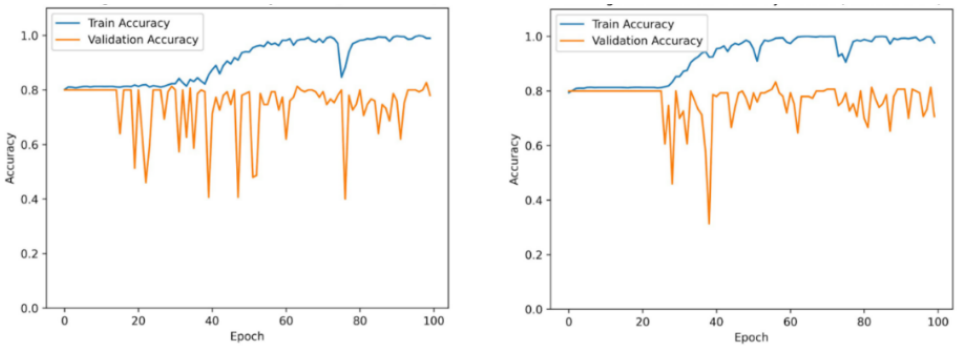
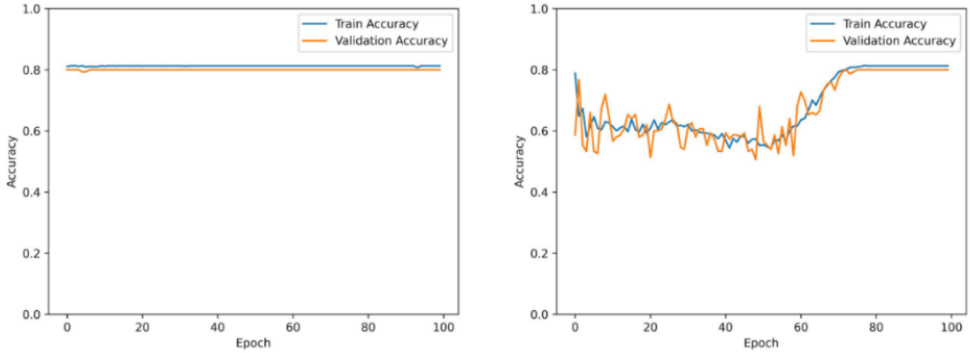


Fig. 2. Skin cancer classification accuracy used TL-CNN: (a) ResNet50 and (b) ResNet101



(a) (b)

Fig. 3. Skin cancer classification accuracy used TL-CNN: (a) DenseNet169 and (b) DenseNet201



(a) (b)

Fig. 4. Skin cancer BCNN used loss function: (a) Binary Cross Entrophy and (b) KL Divergence

Table 4. A Comparison of TL-CNN and BCNN

Network	Training Accuracy	Validation Accuracy
VGG-16	97.15%	75.33%
VGG-19	96.65%	78.67%
ResNet50	100.00%	72.67%
ResNet101	100.00%	76.67%
DenseNet169	98.95%	78.00%
DenseNet201	97.70%	70.67%
BCNN (BCE)	81.30%	80.00%
BCNN (KL Divergence)	81.30%	80.00%

Table 5. A Time Consumption Comparison of TL-CNN and BCNN

Network	Total Training Time (min)	Average Training Time/step (ms)
VGG-16	1.70	21.26
VGG-19	3.35	24.28
ResNet50	5.18	52.68
ResNet101	10.20	92.80
DenseNet169	9.46	91.23
DenseNet201	12.03	113.56
BCNN (BCE)	1.70	12.50
BCNN (KL Divergence)	1.67	12.50

3.2. Discussion

The findings of our study demonstrate that both TL-CNN and BCNN methodologies exhibit significant promise in skin cancer classification. TL-CNN leverages the collective knowledge acquired from models trained on extensive datasets, enabling the identification of intricate patterns in skin cancer dataset. TL-CNN exhibits excellent training accuracy but is prone to overfitting during validation. However, BCNN provides benefits in measuring model uncertainty, enabling clinicians to assess the degree of confidence in

categorization conclusions. The medical setting places great emphasis on the issue of interpretability, and BCNN offers a more transparent framework for delivering uncertainty estimates.

By employing the BCE loss function, BCNN can generate accurate probability predictions for binary classification. Nevertheless, the failure to consider model uncertainty in a medical setting can diminish the comprehensibility of model predictions. Conversely, KL Divergence enables BCNN to consider the uncertainty of the model when producing predictions, enhancing the dependability and assurance in clinical decision-making. During the conducted simulation, it was demonstrated that utilizing the KL Divergence loss function resulted in a training loss value that was tenfold reduced compared to employing BCE. Hence, when selecting a loss function for BCNN, it is crucial to take into account both the model performance and the requirement for interpreting and evaluating uncertainty within the medical domain. Additional research is essential to understand the consequences of selecting this loss function in skin cancer categorization and other medical fields.

It is important to note that TL-CNN and BCNN have their limits. TL-CNN is susceptible to overfitting when used with tiny datasets, whereas BCNN often demands more excellent processing resources. BCNN, or Bayesian Convolutional Neural Network, utilizes a Bayesian methodology to compute the posterior distribution. Theoretically, this could necessitate multiple samplings of the posterior distribution, which would demand more processing resources than traditional methods. Additionally, integrating these models with medical technology and formulating strategies to enhance model explainability are areas of research that show promise. We are confident that TL-CNN and BCNN possess significant promise in aiding clinical skin cancer diagnosis. These two models require further refinement to enhance their capacity to respond to the variability and intricacy of additional medical data.

Both BCNN and TL-CNN have their limitations [69]. Small datasets may lead TL-CNN to overfit, emphasizing the model's reliance on data and the need for big, diverse datasets for efficient generalization [70]. Although BCNN provides better uncertainty estimation, it frequently requires more processing power. Due to the necessity of performing uncertainty inference in neural networks, which involves continuously sampling from the posterior distribution during training and inference, there is an additional computational requirement [71]. It also means that selecting suitable priors for consistent performance may require careful consideration and possibly more advanced techniques.

Nevertheless these limitations, TL-CNN and BCNN demonstrate a great deal of promise for supporting clinical skin cancer diagnosis. In order to better handle the complexity and variety of new medical data, these models need to be substantially improved. Promising research areas include integrating these models with medical technology and developing techniques to improve model explain ability. Enhancing these models' precision and reliability can ultimately result in better patient outcomes and more knowledgeable clinical judgment, highlighting the significance of this research for the medical sector.

4. CONCLUSION

The paper performs a comparative analysis of three TL-CNN architectures along with BCNN for the purpose of classifying skin cancer. TL-CNN architectures, such as VGG, ResNet, and DenseNet, effectively address this issue, achieving training accuracy ranging from 96.65% to 100%. However, these three TL-CNN designs exhibit overfitting issues, with the most significant disparity between training and validation accuracy values being 21.82% for VGG, 27.33% for ResNet, and 27.03% for DenseNet. The BCNN model demonstrated distinct outcomes by effectively mitigating the overfitting issue, achieving training and validation accuracy rates of 81.30% compared to 80% when employing both the BCE loss function and KL Divergence. Furthermore, the BCNN exhibits remarkable efficacy in terms of the duration required for training. The simulation results demonstrate that BCNN exhibits a training process that is 34% more expedient than the training period achieved by TL-CNN.

REFERENCES

[1] A. C. Society, "Cancer Statistics Center," 2023. https://cancerstatisticscenter.cancer.org/ (accessed Apr. 01, 2024).
[2] M. N. C. I. Bethesda, "PDQ® Adult Treatment Editorial Board. PDQ Melanoma Treatment," 2023. https://www.cancer.gov/types/skin/hp/melanoma-treatment-pdq (accessed Apr. 01, 2024).
[3] Z. Apalla, D. Nashan, R. B. Weller, and X. Castellsagué, "Skin Cancer: Epidemiology, Disease Burden, Pathophysiology, Diagnosis, and Therapeutic Approaches," *Dermatol. Ther. (Heidelb).*, vol. 7, no. S1, pp. 5–19, Jan. 2017, https://doi.org/10.1007/s13555-016-0165-y.
[4] M. Shorfuzzaman, "An explainable stacked ensemble of deep learning models for improved melanoma skin cancer detection," *Multimed. Syst.*, vol. 28, no. 4, pp. 1309–1323, Aug. 2022, https://doi.org/10.1007/s00530-021-00787-5.
[5] A. C. Geller, S. M. Swetter, and M. A. Weinstock, "Focus on Early Detection to Reduce Melanoma Deaths," *J. Invest. Dermatol.*, vol. 135, no. 4, pp. 947–949, Apr. 2015, https://doi.org/10.1038/jid.2014.534.

[6] A. S. Panayides *et al.*, "AI in Medical Imaging Informatics: Current Challenges and Future Directions," *IEEE J. Biomed. Heal. Informatics*, vol. 24, no. 7, pp. 1837–1857, Jul. 2020, https://doi.org/10.1109/JBHI.2020.2991043.
[7] C. Rat *et al.*, "Use of Smartphones for Early Detection of Melanoma: Systematic Review," *J. Med. Internet Res.*, vol. 20, no. 4, p. e135, Apr. 2018, https://doi.org/10.2196/jmir.9392.
[8] X. Deng, Y. Zhang, X. Zhao, and H. Li, "Halo-free image enhancement through multi-scale detail sharpening and single-scale contrast stretching," *Signal Process. Image Commun.*, vol. 113, p. 116923, Apr. 2023, https://doi.org/10.1016/j.image.2023.116923.
[9] M. Mouzai, A. Mustapha, Z. Bousmina, I. Keskas, and F. Farhi, "Xray-Net: Self-supervised pixel stretching approach to improve low-contrast medical imaging," *Comput. Electr. Eng.*, vol. 110, p. 108859, Sep. 2023, https://doi.org/10.1016/j.compeleceng.2023.108859.
[10] A. Mahmood, S. A. Khan, S. Hussain, and E. M. Almaghayreh, "An Adaptive Image Contrast Enhancement Technique for Low-Contrast Images," *IEEE Access*, vol. 7, pp. 161584–161593, 2019, https://doi.org/10.1109/ACCESS.2019.2951468.
[11] N. F. Razali, I. S. Isa, S. N. Sulaiman, N. K. A. Karim, and M. K. Osman, "CNN-Wavelet scattering textural feature fusion for classifying breast tissue in mammograms," *Biomed. Signal Process. Control*, vol. 83, p. 104683, May 2023, https://doi.org/10.1016/j.bspc.2023.104683.
[12] D. Huang, R. Zuo, and J. Wang, "Geochemical anomaly identification and uncertainty quantification using a Bayesian convolutional neural network model," *Appl. Geochemistry*, vol. 146, p. 105450, Nov. 2022, https://doi.org/10.1016/j.apgeochem.2022.105450.
[13] H. Lu, S. Cantero-Chinchilla, X. Yang, K. Gryllias, and D. Chronopoulos, "Deep learning uncertainty quantification for ultrasonic damage identification in composite structures," *Compos. Struct.*, vol. 338, p. 118087, Jun. 2024, https://doi.org/10.1016/j.compstruct.2024.118087.
[14] Y. Zhu, S. Tang, and S. Yuan, "Multiple-signal defect identification of hydraulic pump using an adaptive normalized model and S transform," *Eng. Appl. Artif. Intell.*, vol. 124, p. 106548, Sep. 2023, https://doi.org/10.1016/j.engappai.2023.106548.
[15] X. Lei, H. Wang, J. Shen, Z. Chen, W. Zhang, "A novel intelligent underwater image enhancement method via color correction and contrast stretching," *Microprocess. Microsyst.*, p. 104040, Jan. 2021, https://doi.org/10.1016/j.micpro.2021.104040.
[16] Z. Hosseini, A. Khadem, and M. Hassannejad Bibalan, "A novel stretching factor estimator based on an adaptive bisection method for ultrasound strain imaging," *Biomed. Signal Process. Control*, vol. 92, p. 106083, Jun. 2024, https://doi.org/10.1016/j.bspc.2024.106083.
[17] C.-C. Yang, "Image enhancement by modified contrast-stretching manipulation," *Opt. Laser Technol.*, vol. 38, no. 3, pp. 196–201, Apr. 2006, https://doi.org/10.1016/j.optlastec.2004.11.009.
[18] R.Ravindraiah and M. V. Srinu, "Quality Improvement for Analysis of Leukemia Images through Contrast Stretch Methods," *Procedia Eng.*, vol. 30, no. 2011, pp. 475–481, 2012, https://doi.org/10.1016/j.proeng.2012.01.887.
[19] R. Kumar and A. K. Bhandari, "Fuzzified Contrast Enhancement for Nearly Invisible Images," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 5, pp. 2802–2813, May 2022, https://doi.org/10.1109/TCSVT.2021.3098763.
[20] B. Xu, Y. Zhuang, H. Tang, and L. Zhang, "Object-based multilevel contrast stretching method for image enhancement," *IEEE Trans. Consum. Electron.*, vol. 56, no. 3, pp. 1746–1754, Aug. 2010, https://doi.org/10.1109/TCE.2010.5606321.
[21] Y. Qi *et al.*, "A Comprehensive Overview of Image Enhancement Techniques," *Arch. Comput. Methods Eng.*, vol. 29, no. 1, pp. 583–607, Jan. 2022, https://doi.org/10.1007/s11831-021-09587-6.
[22] M. Iman, H. R. Arabnia, and K. Rasheed, "A Review of Deep Transfer Learning and Recent Advancements," *Technologies*, vol. 11, no. 2, p. 40, Mar. 2023, https://doi.org/10.3390/technologies11020040.
[23] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, "Transfer learning: a friendly introduction," *J. Big Data*, vol. 9, no. 1, p. 102, Oct. 2022, https://doi.org/10.1186/s40537-022-00652-w.
[24] P. Thiagarajan, P. Khairnar, and S. Ghosh, "Explanation and Use of Uncertainty Quantified by Bayesian Neural Network Classifiers for Breast Histopathology Images," *IEEE Trans. Med. Imaging*, vol. 41, no. 4, pp. 815–825, Apr. 2022, https://doi.org/10.1109/TMI.2021.3123300.
[25] L. V. Jospin, H. Laga, F. Boussaid, W. Buntine, and M. Bennamoun, "Hands-On Bayesian Neural Networks - A Tutorial for Deep Learning Users," *IEEE Comput. Intell. Mag.*, vol. 17, no. 2, pp. 29–48, 2022, https://doi.org/10.1109/MCI.2022.3155327.
[26] F. Bargagna *et al.*, "Bayesian Convolutional Neural Networks in Medical Imaging Classification: A Promising Solution for Deep Learning Limits in Data Scarcity Scenarios," *J. Digit. Imaging*, vol. 36, no. 6, pp. 2567–2577, Dec. 2023, https://doi.org/10.1007/s10278-023-00897-8.
[27] Z. Ul Abideen *et al.*, "Uncertainty Assisted Robust Tuberculosis Identification With Bayesian Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 22812–22825, 2020, https://doi.org/10.1109/ACCESS.2020.2970023.
[28] R. Feng, N. Balling, D. Grana, J. S. Dramsch, and T. M. Hansen, "Bayesian Convolutional Neural Networks for Seismic Facies Classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 10, pp. 8933–8940, Oct. 2021, https://doi.org/10.1109/TGRS.2020.3049012.
[29] M. Joshaghani, A. Davari, F. N. Hatamian, A. Maier, and C. Riess, "Bayesian Convolutional Neural Networks for Limited Data Hyperspectral Remote Sensing Image Classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 20, pp. 1–

5, 2023, https://doi.org/10.1109/LGRS.2023.3287504.
[30] R. Chandra, A. Bhagat, M. Maharana, and P. N. Krivitsky, "Bayesian Graph Convolutional Neural Networks via Tempered MCMC," *IEEE Access*, vol. 9, pp. 130353–130365, 2021, https://doi.org/10.1109/ACCESS.2021.3111898.
[31] S. Mo *et al.*, "Bayesian convolutional neural networks for predicting the terrestrial water storage anomalies during GRACE and GRACE-FO gap," *J. Hydrol.*, vol. 604, p. 127244, Jan. 2022, https://doi.org/10.1016/j.jhydrol.2021.127244.
[32] T. I. Zohdi and M. Zohdi-Mofid, "Rapid machine-learning enabled design and control of precise next-generation cryogenic surgery in dermatology," *Comput. Methods Appl. Mech. Eng.*, vol. 417, p. 116220, Dec. 2023, https://doi.org/10.1016/j.cma.2023.116220.
[33] S. M. Abd-Alhalem, H. S. Marie, W. El-Shafai, T. Altameem, R. S. Rathore, and T. M. Hassan, "Cervical cancer classification based on a bilinear convolutional neural network approach and random projection," *Eng. Appl. Artif. Intell.*, vol. 127, p. 107261, Jan. 2024, https://doi.org/10.1016/j.engappai.2023.107261.
[34] N. C. F. Codella *et al.*, "Skin lesion analysis toward melanoma detection: A challenge at the 2017 International symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC)," in *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, Apr. 2018, vol. 3, no. 1, pp. 168–172, https://doi.org/10.1109/ISBI.2018.8363547.
[35] V. Hertel, C. Chow, O. Wani, M. Wieland, and S. Martinis, "Probabilistic SAR-based water segmentation with adapted Bayesian convolutional neural network," *Remote Sens. Environ.*, vol. 285, p. 113388, Feb. 2023, https://doi.org/10.1016/j.rse.2022.113388.
[36] M. Monajati and E. Kabir, "A Modified Inexact Arithmetic Median Filter for Removing Salt-and-Pepper Noise From Gray-Level Images," *IEEE Trans. Circuits Syst. II Express Briefs*, vol. 67, no. 4, pp. 750–754, Apr. 2020, https://doi.org/10.1109/TCSII.2019.2919446.
[37] A. H. Fredj and J. Malek, "Design and Implementation of a Pipelined Median Filter Architecture," in *2019 IEEE International Conference on Design & Test of Integrated Micro & Nano-Systems (DTS)*, pp. 1–6, Apr. 2019, https://doi.org/10.1109/DTSS.2019.8915329.
[38] P. A. Lyakhov, A. R. Orazaev, N. I. Chervyakov, and D. I. Kaplun, "A New Method for Adaptive Median Filtering of Images," in *2019 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*, pp. 1197–1201 Jan. 2019, https://doi.org/10.1109/EIConRus.2019.8657050.
[39] R. Naseem, Z. A. Khan, N. Satpute, A. Beghdadi, F. A. Cheikh, and J. Olivares, "Cross-Modality Guided Contrast Enhancement for Improved Liver Tumor Image Segmentation," *IEEE Access*, vol. 9, pp. 118154–118167, 2021, https://doi.org/10.1109/ACCESS.2021.3107473.
[40] S. You and M. Reyes, "Influence of contrast and texture based image modifications on the performance and attention shift of U-Net models for brain tissue segmentation," *Front. Neuroimaging*, vol. 1, Oct. 2022, https://doi.org/10.3389/fnimg.2022.1012639.
[41] A. Javaid, M. Sadiq, and F. Akram, "Skin Cancer Classification Using Image Processing and Machine Learning," *Proc. 18th Int. Bhurban Conf. Appl. Sci. Technol. IBCAST 2021*, pp. 439–444, 2021, https://doi.org/10.1109/IBCAST51254.2021.9393198.
[42] S. Bianco and C. Cusano, "Quasi-Unsupervised Color Constancy," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 12204–12213, Jun. 2019, https://doi.org/10.1109/CVPR.2019.01249.
[43] M. Aiguier, I. Bloch, and R. Pino-Pérez, "Abstract Mathematical morphology based on structuring element: Application to morpho-logic," *arXiv preprint arXiv:2005.01715*, pp. 1–40, May 2020, [Online]. Available: http://arxiv.org/abs/2005.01715.
[44] A. W. Salehi *et al.*, "A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope," *Sustainability*, vol. 15, no. 7, p. 5930, Mar. 2023, https://doi.org/10.3390/su15075930.
[45] L. Alzubaidi *et al.*, "Novel Transfer Learning Approach for Medical Imaging with Limited Labeled Data," *Cancers (Basel).*, vol. 13, no. 7, p. 1590, Mar. 2021, https://doi.org/10.3390/cancers13071590.
[46] I. Kandel, M. Castelli, and A. Popovič, "Musculoskeletal Images Classification for Detection of Fractures Using Transfer Learning," *J. Imaging*, vol. 6, no. 11, p. 127, Nov. 2020, https://doi.org/10.3390/jimaging6110127.
[47] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "DeTrac: Transfer Learning of Class Decomposed Medical Images in Convolutional Neural Networks," *IEEE Access*, vol. 8, pp. 74901–74913, 2020, https://doi.org/10.1109/ACCESS.2020.2989273.
[48] R. Godasu, D. Zeng, K. Sutrave, "Transfer learning in medical image classification: Challenges and opportunities," *Transfer,* vol. 5, 2020, https://core.ac.uk/download/pdf/326833449.pdf.
[49] R. Jha, V. Bhattacharjee, and A. Mustafi, "Transfer Learning with Feature Extraction Modules for Improved Classifier Performance on Medical Image Data," *Sci. Program.*, vol. 2022, pp. 1–10, Aug. 2022, https://doi.org/10.1155/2022/4983174.
[50] A. Z. Karen Simonyan, "Very Deep Convolutional Networks For Large-Scale Image Recognition," *Am. J. Heal. Pharm.*, vol. 75, no. 6, pp. 398–406, 2018, [Online]. Available: https://click.endnote.com/viewer?doi=arxiv%3A1409.1556&token=WzQxNTM2OSwiYXJ4aXY6MTQwOS4xNTU2Il0.1ERLUjt9JINMPEYW74pwJFdB7PA.
[51] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, Jun. 2016, https://doi.org/10.1109/CVPR.2016.90.

[52] M. Yin, X. Li, Y. Zhang, and S. Wang, "On the Mathematical Understanding of ResNet with Feynman Path Integral," *arXiv preprint arXiv:1904.07568*, 2019, [Online]. Available: http://arxiv.org/abs/1904.07568.
[53] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2261–2269, Jul. 2017, https://doi.org/10.1109/CVPR.2017.243.
[54] M. V. Narkhede, P. P. Bartakke, and M. S. Sutaone, "A review on weight initialization strategies for neural networks," *Artif. Intell. Rev.*, vol. 55, no. 1, pp. 291–322, Jan. 2022, https://doi.org/10.1007/s10462-021-10033-z.
[55] W. Cao, X. Wang, Z. Ming, and J. Gao, "A review on neural networks with random weights," *Neurocomputing*, vol. 275, pp. 278–287, Jan. 2018, https://doi.org/10.1016/j.neucom.2017.08.040.
[56] K.-T. Fang and R. Li, "Bayesian Statistical Inference on Elliptical Matrix Distributions," *J. Multivar. Anal.*, vol. 70, no. 1, pp. 66–85, Jul. 1999, https://doi.org/10.1006/jmva.1998.1816.
[57] G. Atteia, A. Alhussan, and N. Samee, "BO-ALLCNN: Bayesian-Based Optimized CNN for Acute Lymphoblastic Leukemia Detection in Microscopic Blood Smear Images," *Sensors*, vol. 22, no. 15, p. 5520, Jul. 2022, https://doi.org/10.3390/s22155520.
[58] M. Ait Amou, K. Xia, S. Kamhi, and M. Mouhafid, "A Novel MRI Diagnosis Method for Brain Tumor Classification Based on CNN and Bayesian Optimization," *Healthcare*, vol. 10, no. 3, p. 494, Mar. 2022, https://doi.org/10.3390/healthcare10030494.
[59] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "BayesCap: A Bayesian Approach to Brain Tumor Classification Using Capsule Networks," *IEEE Signal Process. Lett.*, vol. 27, pp. 2024–2028, 2020, https://doi.org/10.1109/LSP.2020.3034858.
[60] C. E. Shannon, "A Mathematical Theory of Communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, 1948, https://doi.org/10.1002/j.1538-7305.1948.tb00917.x.
[61] Y. Wen, P. Vicol, J. Ba, D. Tran, and R. Grosse, "Flipout: Efficient Pseudo-Independent Weight Perturbations On Mini-Batches," *Published as a conference paper at ICLR 2018*, pp. 1–16, 2018, [Online]. Available: https://arxiv.org/pdf/1803.04386.pdf.
[62] G. Plusch, S. Arsenyev-Obraztsov, and O. Kochueva, "The Weights Reset Technique for Deep Neural Networks Implicit Regularization," *Computation*, vol. 11, no. 8, 2023, https://doi.org/10.3390/computation11080148.
[63] S. Dräger and J. Dunkelau, "Evaluating the Impact of Loss Function Variation in Deep Learning for Classification," *arXiv preprint arXiv:2210.16003*, 2022, [Online]. Available: http://arxiv.org/abs/2210.16003.
[64] L. Li, M. Doroslovacki, and M. H. Loew, "Approximating the Gradient of Cross-Entropy Loss Function," *IEEE Access*, vol. 8, pp. 111626–111635, 2020, https://doi.org/10.1109/ACCESS.2020.3001531.
[65] K. Cho, J. H. Roh, Y. Kim, and S. Cho, "A Performance Comparison of Loss Functions," *ICTC 2019 - 10th Int. Conf. ICT Converg. ICT Converg. Lead. Auton. Futur.*, pp. 1146–1151, 2019, https://doi.org/10.1109/ICTC46691.2019.8939902.
[66] R. Wali, "Xtreme Margin: A Tunable Loss Function for Binary Classification Problems," arXiv preprint arXiv:2211.*00176*, pp. 1–10, 2022, [Online]. Available: http://arxiv.org/abs/2211.00176.
[67] I. Matychyn, "On Computation of Matrix Mittag-Leffler Function," *arXiv preprint arXiv:1706.01538 2017,* [Online]. Available: http://arxiv.org/abs/1706.01538.
[68] A. Asperti and M. Trentin, "Balancing Reconstruction Error and Kullback-Leibler Divergence in Variational Autoencoders," *IEEE Access*, vol. 8, pp. 199440–199448, 2020, https://doi.org/10.1109/ACCESS.2020.3034828.
[69] L. Zhang, J. Dong, J. Zhang, and J. Yang, "A Modified Stein Variational Inference Algorithm with Bayesian and Gradient Descent Techniques," *Symmetry (Basel).*, vol. 14, no. 6, p. 1188, Jun. 2022, https://doi.org/10.3390/sym14061188.
[70] V. Sevetlidis, G. Pavlidis, S. Mouroutsos, and A. Gasteratos, "Tackling Dataset Bias With an Automated Collection of Real-World Samples," *IEEE Access*, vol. 10, pp. 126832–126844, 2022, https://doi.org/10.1109/ACCESS.2022.3226517.
[71] A. F. Psaros, X. Meng, Z. Zou, L. Guo, and G. E. Karniadakis, "Uncertainty quantification in scientific machine learning: Methods, metrics, and comparisons," *J. Comput. Phys.*, vol. 477, p. 111902, Mar. 2023, https://doi.org/10.1016/j.jcp.2022.111902.

BIOGRAPHY OF AUTHORS

Mulki Indana Zulfa is a lecturer and researcher at the electrical engineering department at Jenderal Soedirman University, Purwokerto, Indonesia. He obtained a bachelor degree in informatics engineering from Telkom University, then continued his master and doctoral studies at ITB and UGM respectively. Since graduating from his doctoral degree, he has started to build a lot of quality research which has been published in several reputable international journals. Also, currently he actively serves as department secretary. He is also listed as head of the IT division at the engineering faculty and is active as a member of the e-learning management at the university quality assurance agency, LP3M. Various research grants have often been obtained, such as the BLU research scheme, RTA UGM research while pursuing doctoral studies, and DRTPM research. He is currently engaged in teaching and conducting research in the areas of optimization, data mining, and machine learning. He can be contacted at email:

mulki_indanazulfa@unsoed.ac.id. Orchid:0000-0003-4994-3360.

Andreas Sahir Aryanto is a graduate student with a Bachelor degree in Electrical Engineering from Jenderal Soedirman University, Purwokerto, Indonesia. He completed his degree in 3 years and 4 months, maintaining a strong GPA of 3.87. Andreas has a strong interest in research, particularly in areas like machine learning, deep learning, image processing, pattern recognition, and expert systems, especially when applied to medical image diagnosis and analysis. He is highly motivated to pursue his academic journey further and is actively seeking opportunities for graduate studies to obtain his Master and Doctoral degrees. Currently, Andreas works as an ML engineer at Angusta Systems, Eveleigh, Australia. Email: andreas.sahir.a@gmail.com.

Bintang Abelian MW is a graduate student with a Bachelor degree in Electrical Engineering from Jenderal Soedirman University with a solid GPA of 3.30. Bintang has demonstrated consistent academic performance. His project work is particularly notable, having contributed to several projects that have been referenced by faculty members for academic papers. Looking ahead, Bintang is planning to further his studies in either Computer Science or Data Science, reflecting his strong interest in data analysis research. Email: bintang.wijonarko@mhs.unsoed.ac.id.

Waleed Ali received the Ph.D. degrees in computer science from the Faculty of Computing, Universiti Teknologi Malaysia (UTM), Malaysia 2012. He has been an Associate Professor with the IT Department, Faculty of Computing and Infor mation Technology, King Abdulaziz University, Rabigh, since September 2013. He has published many papers in several high-impact factor journals, good conferences, and book chapters. His research interests include intelligent web caching, web usage mining, and machine learning techniques and their applications. Email: waabdullah@kau.edu.sa. Orchid: 0000-0003-3746-4274.