The Forward-Forward Algorithm as a feature extractor for skin lesion classification: A preliminary study

Abel Reyes-Angulo [1]   **Sidike Paheding** [1]

Abstract

Skin cancer, a deadly form of cancer, exhibits a 23% survival rate in the USA with late diagnosis. Early detection can significantly increase the survival rate, and facilitate timely treatment. Accurate biomedical image classification is vital in medical analysis, aiding clinicians in disease diagnosis and treatment. Deep learning (DL) techniques, such as convolutional neural networks and transformers, have revolutionized clinical decision-making automation. However, computational cost and hardware constraints limit the implementation of state-of-the-art DL architectures. In this work, we explore a new type of neural network that does not need backpropagation (BP), namely the Forward-Forward Algorithm (FFA), for skin lesion classification. While FFA is claimed to use very low-power analog hardware, BP still tends to be superior in terms of classification accuracy. In addition, our experimental results suggest that the combination of FFA and BP can be a better alternative to achieve a more accurate prediction.

1. Introduction

Skin cancer is one of the most common types of cancer (Linares et al., 2015; Siegel et al., 2023), and early detection and diagnosis can greatly improve the prognosis of patients (Miller et al., 2019). Dermatologists use visual inspection to diagnose skin cancer, which can be challenging because there are numerous skin lesions and the visual similarity between different types of them. Therefore, Computer-aided diagnosis (CAD) systems are required to assist dermatolo-

[1]Department of Applied Computing, Michigan Technological University, Houghton, USA. Correspondence to: Abel Reyes-Angulo <areyesan@mtu.edu>, Sidike Paheding <spahedin@mtu.edu>.

gists in the early detection and diagnosis of skin cancer, aiming to ensure timely treatment (Bhinder et al., 2021). In the past few years, deep learning (DL) has revolutionized the automation of clinical decisions through computational system assistance. Techniques such as backpropagation (Rumelhart et al., 1985), convolutional neural networks (CNNs) (LeCun et al., 1998), and recently proposed transformers (Vaswani et al., 2017) have boosted confidence in the use of AI in a real clinical setup. However, the computational cost and hardware constraints related to the implementation of state-of-the-art DL architectures limit the feasibility of actual deployment implementation (Imteaj et al., 2022). The Forward-Forward Algorithm (FFA) (Hinton, 2022) has been presented as an alternative to optimizing the training process of neural networks, considering the "mortal computational" cost. Although results reported by the use of FFA have not outperformed traditional mechanisms such as backpropagation, the trade-off could be less requirement in hardware implementation. In this work, we explore the performance of FFA for skin lesion classification and also compare it to a backpropagation-based trained model. Finally, we propose to combine the two techniques during the training process, where the FFA is used as a feature extractor and backpropagation to improve model prediction.

2. Related works

Skin lesion classification has been a significant research area in recent years, and numerous deep-learning techniques have been designed for this task (Zhang et al., 2019; Mahbod et al., 2019). Most existing methods use traditional CNNs, which learn the feature representation through backpropagation (Harangi, 2018; Albahar, 2019). However, these methods can be limited by the quality of the feature representation. In order to overcome this limitation, some researchers have suggested methods that use adversarial training to improve the feature representation. For example, GAN-based methods have been implemented for skin lesion classification, which use a generator network to generate realistic skin lesion images and a discriminator network to classify the generated images as real or fake (Rashid et al., 2019).

In contrast to these methods, we herein propose to combine

the FFA with traditional backpropagation to enhance feature representation and improve classification accuracy in skin lesion classification. Our approach does not require adversarial training or pre-training on a large dataset, making it computationally efficient and suitable for large-scale datasets. Our approach means to provide a baseline on the performance of DL architecture using the FFA strategy in contrast with backpropagation, in addition to exploring the combination of the two techniques.

3. Methodology

Our proposed framework consists of two stages: feature representation with FFA and final prediction with backpropagation.

3.1. Backpropagation

Introduced by Rumelhart et al. (Rumelhart et al., 1986; 1985), the backpropagation algorithm, utilized for training artificial neural networks, entails calculating the gradient of the error function concerning the network weights. This gradient information is then used to update the weights by using gradient descent, aiming to minimize the global loss function of the task. The authors demonstrated that this algorithm could facilitate the learning of distributed representation of input data, enabling effective generalization to new data.

The backpropagation algorithm consists of two main stages: the forward pass and the backward pass. During the forward pass, input data is propagated through the network, generating corresponding predictions. Subsequently, during the backward pass, the gradients of the loss function are calculated with respect to the network's parameters by propagating the error backward through the neural network. These gradients are then utilized to update the parameters using an optimization algorithm, such as stochastic gradient descent (SGD) or Adam (Amari, 1993; Bock et al., 2018).

3.2. The Forward-Forward Algorithm

Backpropagation   **Forward-Forward Algorithm**

Local update
Forward Pass
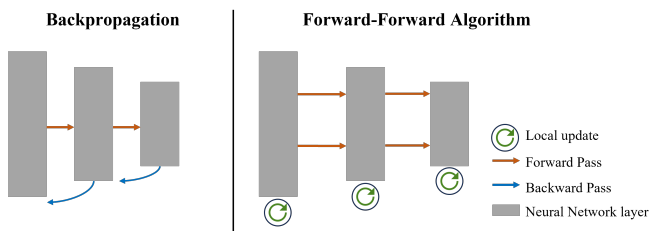Backward Pass
Neural Network layer

Figure 1. A brief illustration of the backpropagation and the forward-forward algorithm.

Introduced by Geoffrey Hinton (Hinton, 2022), the forward-

forward algorithm (FFA) is an innovative learning procedure for neural networks. It draws inspiration from the Boltzmann machines (Hinton et al., 1986) and Noise Contrastive Estimation (Gutmann & Hyvärinen, 2010), Its primary objective is to replace the traditional forward and backward passes of backpropagation with two parallel forward passes. The first forward pass uses real data, and the second forward pass uses negative data, which the network generates itself. Each layer possesses its individual objective function, which essentially aims to maximize the goodness for positive data while minimizing the goodness for negative data. Figure 1 illustrates the main difference between the traditional backpropagation algorithm and the newer FFA. Mathematically, the goodness is computed as follows

$$prob(\text{positive-data}) = \sigma(\sum_j y_j^2 - \theta)) \qquad (1)$$

where y_i represents activity of the jth hidden unit, θ a given threshold, and σ a logistic distribution function. The FFA has been shown to work well on a few small problems (Ororbia, 2023), but it has not yet been tested on large-scale problems. In addition, the FFA has been claimed to be superior in hardware efficiency, through low power consumption when compared with the backpropagation and the gradient computation (Hinton, 2022; Kendall et al., 2020).

3.3. Feature representation in FFA

In the FFA, pairs of data points are created by corrupting real data points with noise. The network is then trained to distinguish between the real data points and the corrupted data points. This training process forces the network to learn representations that are robust to noise. The proposed method utilizes the FFA algorithm as the initial stage of learning, and then the final classification is trained by using regular backpropagation.

4. Implementation details

The proposed methodology was tested with two skin lesion classification benchmark datasets: ISIC 2016 and HAM 10000 datasets, details about the dataset are provided in section 5. The ISIC 2016 provides training and testing datasets with the correspondent classification ground truth (i.e. benign or malign skin lesion). However, the HAM 10000 dataset provides only the training dataset with the corresponding type of skin lesion ground truth. For the last case, we randomly split the dataset in training and testing with a ratio of 8:2, respectively. In order to alleviate the computational constraints all the images, from the aforementioned datasets, were downsampled to 64×64 pixels size, maintaining the original RGB channels. The inputs were flattened and normalized prior to the training process. The architecture for all the experiments was composed of 3 fully

connected hidden layers (784, 500, 500 units) with ReLU activation and batch normalization. A final softmax activation function provides the probability distribution among the classes corresponding to the classification task on each dataset. In the case of the FFA, the positive and negative data are generated by the network, and the one-hot-encoder representation of the label is overlying in the first n pixels of the input image, where n represent the number of classes available. The positive data embed the original ground truth in the corresponding input image; meanwhile, the negative data overlay a wrong label in the same input image. Figure 2 illustrates the overlay technique utilized in this work to generate the positive and the negative data. The architecture is implemented within a Tensorflow/Keras 2.9 version environment. Each experiment is run using a single NVIDIA GeForce RTX 3070 graphic card with 8GB of dedicated GPU. All experiments are run for a total of 250 epochs with checkpoint callback, and batch size of 64 samples per iteration. The mean squared error (MSE) loss was used along with the Adam optimizer with an initial learning rate of 1×10^{-3}.
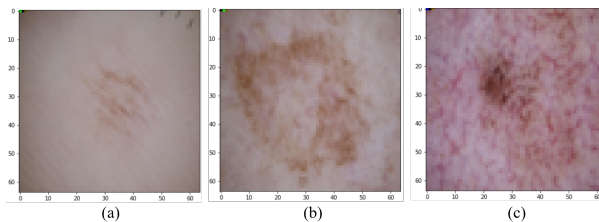

Figure 2. Illustration of the generated positive and negative data when FFA is used over the HAM 10000 dataset. The images contain embedded label information in the first 7 pixels: (a) sample with label 1, (b) sample with label 4, and (c) sample with the label 2.

5. Experiments

In this section, we provide more details about the experiments performed to evaluate the proposed method, including the description of the datasets, performance metrics used as evaluation criteria, and a discussion of the experimental results.

5.1. Datasets

In order to evaluate the robustness of the proposed methodology, we utilize two widely used and publicly available benchmark datasets related to skin lesion classification: ISIC 2016 and HAM 10000 datasets.

5.1.1. ISIC 2016

Presented in the International Skin Imaging Collaboration (ISIC) 2016 challenge, the ISIC 2016 dataset (Gutman et al.,
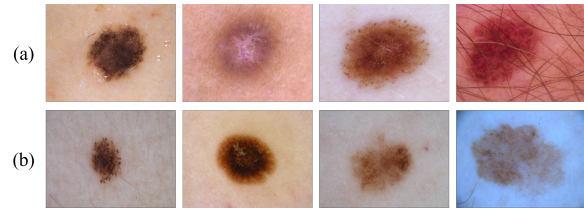

Figure 3. Samples from the two benchmark datasets utilized in the work: (a) The ISIC 2016 and (b) the HAM 10000 dataset.

2016) is a comprehensive collection of dermoscopic images that is widely used in the field of computer-aided diagnosis for skin lesions. It consists of 900 training images and 379 test images of dermoscopic skin lesions. The dataset includes binary masks for both lesion segmentation and classification purposes, enabling researchers to evaluate algorithms for these tasks. The dataset encompasses three main types of skin lesions: melanoma (Malign), seborrheic keratosis, and nevus (Benign). These lesion types cover a range of conditions commonly encountered in dermatology practice. The images in the dataset were obtained from patients across multiple countries, resulting in a diverse and challenging dataset.

5.1.2. HAM 10000

The Human Against Machine with 10000 training images (HAM10000) dataset (Tschandl et al., 2018), widely used in dermatology and computer vision, comprises 10,015 dermoscopic images of skin lesions obtained from diverse clinical sources. It was specifically created for the ISIC 2018 challenge initiated by the International Skin Imaging Collaboration, with the aim of advancing automated analysis of skin lesions. Within the HAM10000 dataset, images represent seven distinct types of skin lesions: melanoma(MEL), melanocytic nevus(NV), basal cell carcinoma(BCC), actinic keratosis(AKIEC), benign keratosis(BLK), dermatofibroma(DF), and vascular lesions(VASC). Figure 4 shows the numerical by-class distribution of the samples in the HAM 100000 dataset. Figure 3 shows samples of the two mentioned datasets.

5.2. Performance metrics

5.2.1. THE ERROR RATE %:

The error rate is a commonly used performance metric that measures the accuracy of the model's predictions. It represents the proportion of incorrectly classified instances out of the total number of instances. The error rate, denoted as ϵ, is a straightforward and intuitive metric that provides a simple measure of the model's performance in image classification
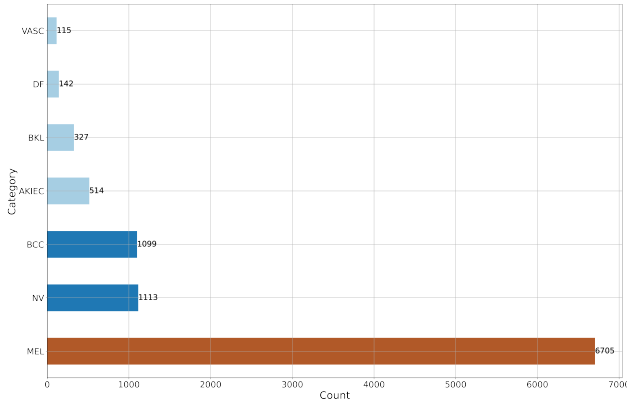
Figure 4. Distribution of the samples per class in the HAM 10000 dataset.

Table 1. Error rate (%) on skin lesion datasets.

	ISIC 2016		HAM 10000	
	Testing error rate %	Training error rate %	Testing error rate %	Training error rate %
FFA	29.52	22.48	37.21	20.42
BP	24.81	17.33	31.66	13.97
FFA+BP	**23.31**	**16.85**	**30.44**	**13.43**

Table 2. ROC-AUC on skin lesion classification benchmarks.

	ISIC 2016		HAM 10000	
	Testing ROC-AUC	Training ROC-AUC	Testing ROC-AUC	Training ROC-AUC
FFA	0.5632	0.6684	0.6912	0.7824
BP	0.6216	0.8240	0.8360	0.9324
FFA+BP	**0.6495**	**0.8472**	**0.8483**	**0.9414**

tasks and can be mathematically expressed as

$$\epsilon = \frac{FP + FN}{TP + TN + FP + FN} \times 100\% \qquad (2)$$

5.2.2. THE ROC-AUC SCORE:

The ROC-AUC measures the classifier's ability to distinguish between positive and negative instances across various classification thresholds. The ROC curve (receiver operating characteristic curve) demonstrates the true positive rate (sensitivity) against the false positive rate (1-specificity), and the ROC-AUC represents the area under this ROC curve. A higher ROC-AUC indicates better classification performance. The ROC-AUC can be mathematically expressed as

$$\text{ROC AUC score} = \int_{-\infty}^{\infty} \text{TPR}(f) \cdot \text{FPR}(f)\text{d}f \qquad (3)$$

where $\text{TPR}(f) = \frac{\text{TP}(f)}{\text{P}}$ and $\text{FPR}(f) = \frac{\text{FP}(f)}{\text{N}}$ are knowing as the true positive rate and the false positive rate, respectively

5.2.3. EXPERIMENTAL RESULTS AND DISCUSSION

Experimental results are reported in Table 1 in terms of the error rate %, and in Table 2 in terms of ROC-AUC score. In both cases, it shows superior performance of the model when it is trained with the backpropagation method in contrast when it is trained with FFA. However, reported results suggest a better performance can be achieved when both techniques are combined, and the results are consistent over the two datasets. It is important to remark on the following points: (1) The architecture configuration was not intended to achieve state-of-the-art performance but provides a comparative baseline when a DL architecture is trained using either FFA, backpropagation, or both. (2) The use of only

fully connected layers limits the size of the architecture due to computational constraints. Therefore, this is expected to be addressed through the implementation of convolutional layers.

6. Conclusion

In this work, we investigated the performance of FFA for the skin lesion classification task. We compared the prediction accuracy of the FFA, backpropagation, and the combination of both. Experimental results showed that the backpropagation algorithm yields better classification accuracy than FFA. However, our results also suggest promising insights into the use of the FFA in the neural network as a feature extractor. A contrastive learning stage with FFA complemented with classification learning using traditional backpropagation could enhance the performance for skin lesion classification when comparing the similar architecture setup using either FFA or the backpropagation method only. One limitation of our study is that we did not include the results of the energy efficiency of FFA in terms of hardware implementation, which will be our future work.

References

Albahar, M. A. Skin lesion classification using convolutional neural network with novel regularizer. *IEEE Access*, 7:38306–38313, 2019.

Amari, S.-i. Backpropagation and stochastic gradient descent method. *Neurocomputing*, 5(4-5):185–196, 1993.

Bhinder, B., Gilvary, C., Madhukar, N. S., and Elemento, O. Artificial intelligence in cancer research and precision medicine. *Cancer discovery*, 11(4):900–915, 2021.

Bock, S., Goppold, J., and Weiß, M. An improvement of the convergence proof of the adam-optimizer. *arXiv preprint arXiv:1804.10587*, 2018.

Gutman, D., Codella, N. C., Celebi, E., Helba, B., Marchetti, M., Mishra, N., and Halpern, A. Skin lesion analysis

toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1605.01397*, 2016.

Gutmann, M. and Hyvärinen, A. Noise-contrastive estimation: A new estimation principle for unnormalized statistical models. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pp. 297–304. JMLR Workshop and Conference Proceedings, 2010.

Harangi, B. Skin lesion classification with ensembles of deep convolutional neural networks. *Journal of biomedical informatics*, 86:25–32, 2018.

Hinton, G. The forward-forward algorithm: Some preliminary investigations. *arXiv preprint arXiv:2212.13345*, 2022.

Hinton, G. E., Sejnowski, T. J., et al. Learning and relearning in boltzmann machines. *Parallel distributed processing: Explorations in the microstructure of cognition*, 1 (282-317):2, 1986.

Imteaj, A., Mamun Ahmed, K., Thakker, U., Wang, S., Li, J., and Amini, M. H. Federated learning for resource-constrained iot devices: Panoramas and state of the art. *Federated and Transfer Learning*, pp. 7–27, 2022.

Kendall, J., Pantone, R., Manickavasagam, K., Bengio, Y., and Scellier, B. Training end-to-end analog neural networks with equilibrium propagation. *arXiv preprint arXiv:2006.01981*, 2020.

Langley, P. Crafting papers on machine learning. In Langley, P. (ed.), *Proceedings of the 17th International Conference on Machine Learning (ICML 2000)*, pp. 1207–1216, Stanford, CA, 2000. Morgan Kaufmann.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

Linares, M. A., Zakaria, A., and Nizran, P. Skin cancer. *Primary care: Clinics in office practice*, 42(4):645–659, 2015.

Mahbod, A., Schaefer, G., Wang, C., Ecker, R., and Ellinge, I. Skin lesion classification using hybrid deep neural networks. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1229–1233. IEEE, 2019.

Miller, K. D., Nogueira, L., Mariotto, A. B., Rowland, J. H., Yabroff, K. R., Alfano, C. M., Jemal, A., Kramer, J. L., and Siegel, R. L. Cancer treatment and survivorship statistics, 2019. *CA: a cancer journal for clinicians*, 69 (5):363–385, 2019.

Ororbia, A. Learning spiking neural systems with the event-driven forward-forward process. *arXiv preprint arXiv:2303.18187*, 2023.

Rashid, H., Tanveer, M. A., and Khan, H. A. Skin lesion classification using gan based data augmentation. In *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 916–919. IEEE, 2019.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. Learning internal representations by error propagation. Technical report, California Univ San Diego La Jolla Inst for Cognitive Science, 1985.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. Learning representations by back-propagating errors. *nature*, 323(6088):533–536, 1986.

Siegel, R. L., Miller, K. D., Wagle, N. S., and Jemal, A. Cancer statistics, 2023. *CA: a cancer journal for clinicians*, 73(1):17–48, 2023.

Tschandl, P., Rosendahl, C., and Kittler, H. The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, 5(1):1–9, 2018.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., and Polosukhin, I. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Zhang, J., Xie, Y., Xia, Y., and Shen, C. Attention residual learning for skin lesion classification. *IEEE transactions on medical imaging*, 38(9):2092–2103, 2019.