DOI: 10.15849/IJASCA.240730.06

Int. J. Advance Soft Compu. Appl, Vol. 16, No. 2, July 2024
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Custom Convolutional Neural Network Model for Identification of Nutritional Deficiencies in Children

Shilpa Ankalaki[1], Vidyadevi G Biradar[2], Kushal G[3] and Kavya N[4]

[1]Department of Computer Science and Engineering, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal, Karnataka 576104, India
e-mail: shilpa.ankalaki@manipal.edu
[2,3,4]Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India
e-mail: vidyadevi.g.biradar@nmit.ac.in
e-mail: kushalkowsh@gmail.com
e-mail: kavyan1402@gmil.com

Abstract

Undernutrition occurs when there are deficiencies in essential vitamins, while overnutrition refers to consuming an excessive amount of nutrients, leading to issues such as obesity, diabetes, and related health problems. Malnutrition often stems from the economic and social status of parents. In children, malnutrition can significantly result in physical and mental growth issues. Therefore, there is a need for early prediction of malnutrition in children to mitigate its adverse effects. Children image comprises great amount of information which can be analyzed for distinguishing between a nourished or malnourished children. The success stories of convolutional neural networks in image classification are the motivation to develop a deep learning model for the classification. In this work a custom designed convolutional neural network model is proposed for classification of children into category nourished or malnourished. The dataset consists of 630 and 1530 nourished and malnutrition children's images respectively. The model is trained on children's images database that includes both web scraped images and synthetic images. The convolutional neural network model is optimized by selecting optimal activation function through experimentation. The model is trained for 100 epochs with Adam optimizer and various activation functions. CNN with Relu with weight decay activation function obtained considerably good results with accuracy of 93.44%, precision 0.89, 0.85 recall, 0.87 F1 score for nourished class; and precision 0.95, 0.96 recall, and 0.96 F1 score for malnourished category. The model obtained a good accuracy of 92.83% for Leaky ReLu with precision 0.8, 0.9 recall, 0.85 F1 score for nourished and precision 0.97, 0.94 recall and 0.95 F1 score malnourished. Further, to develop trust in the model visualization of activation maps, convolutional layers and filters are implemented.

Keywords: *Deep Learning, Nutritional Value, Health, activation maps, activation functions, convolutional neural networks, hyper parameters*

Received 25 February 2024; Accepted 13 April 2024

1 Introduction

Nourishing children is vital for global health as it promotes physical and cognitive development, ensuring a strong foundation for future productivity. Adequate nutrition during childhood strengthens immune systems, reducing susceptibility to diseases and contributing to long-term well-being. Investing in children's nutrition is a cost-effective strategy to break the cycle of poverty and enhance the overall health of communities worldwide. Early life malnutrition is a pervasive global difficulty that poses severe threats to the physical and cognitive development of younger individuals. Early detection and intervention are paramount to mitigate its consequences efficiently. Historically, dietary assessments have depended on manual measurements and medical reviews performed by healthcare experts. However, those techniques are often aid-intensive, prone to human mistakes, and might not be quite easily available in deprived areas. Malnutrition is a condition that occurs due to fewer intakes or over intake of nutrients [1]. Infant malnutrition is a standard worldwide health trouble, jeopardizing the bodily and cognitive development of children. Conventional strategies for assessing nutritional reputation are regularly useful, however, their availability in deprived regions is doubtful. There is a need for a scalable and green approach to classify children as nourished or malnourished based on pics to facilitate early intervention [2].

Machine learning (ML) algorithms are exploited in the existing research work on prediction of malnutrition among children and women. The prediction of malnutrition in children below 5 years is addressed using various ML algorithms [3]. Malnutrition in women is a public health issue that has adverse effect on child and as results the child may have stunted growth, low birth weight, diabetes etc. ML algorithms are used for prediction of major risks associated with malnutrition in women [4]. Good nutrition is vital to human health, and malnutrition in children is a global problem that leads to mortality and morbidity in children, Afghanistan is such one country which suffers from malnutrition at severe to acute level, to predict malnutrition ML models were used [5].

Children images contain significant information which can be analyzed for identification of nutritional state of the child. Deep Learning models such as convolutional neural network (CNN) are widely used for the task of image classification. This works proposes a tool to analyze children’s images for prediction of malnutrition status using CNN model. Identification of malnutrition in children using images allows for quick and accurate assessment, enabling timely intervention to prevent long-term health issues. Image-based diagnostics provide objective data, reducing subjectivity in nutritional evaluations and ensuring more targeted and effective interventions.

Children images contain significant information which can be analyzed for identifi-cation of nutritional state of the child. Deep Learning models such as convolutional neural network (CNN) are widely used for the task of image classification. This works proposes a tool to analyze children’s images for prediction of malnutrition status using CNN model. Identification of malnutrition in children using images allows for quick and accurate assessment, enabling timely intervention to prevent long-term health issues. Image-based diagnostics provide objective data, reducing subjectivity in nutritional evaluations and ensuring more targeted and effective interventions. The objectives and contribution of this paper are as follows:

- Design of a custom designed convolutional neural network (CNN) model with optimization on model hyper parameters for identification of malnutrition in children

- There is a paucity of children image databases, therefore this paper proposes a method of generating children's images which are synthetic and suitable for model training.
- Quantitative analysis of model performance using metrics accuracy, F1-score, Recall and Precision.
- Qualitative analysis of model behavior in taking decisions in the process of classification of children images into nourished and malnourished by visualization of convolutional layers, filter, and activation maps.

The research carried out is articulated into different sections, Section 2 gives insight into literature surveys. Section 3 discusses methodology of the proposed method, and Section 4 illustrates the performance analysis of proposed CNN model. Section 5 provides qualitative performance analysis, Section 6 reports results and comparative analysis of model performance. Finally, conclusions and directions for future work are given.

2 Related Work

The study on malnutrition prediction among children is extensively carried out using ML algorithms. The ML methods such as logistics regression, Support Vector Machine (SVM), Random Forest algorithms are widely used in prediction of nutritional status in children and women. These algorithms are explored on the dataset of children pertaining to specific regions. The most popularly used benchmark datasets employed by researchers to evaluate ML algorithms are United States Agency for International Development (USAID), CAESAR and Bangladesh Demographic and Health Survey 2014 (BDHS) and Afghanistan dataset. ML algorithms like Generalized linear models (GLM) [2], Ridge [2], Lasso[2], Elastic Net,[2] Neural Networks (NN)[2][4][7-8], Random Forest(RF)[2-5][8], k- nearest neighbor (kNN) [3][7][8], Linear Discriminant Analysis[3] Support Vector Machine (SVM)[3-4][7], Linear regression (LR) [3][7], Naïve Bayesian (NB) [4], Decision Tree [4][8], J48[5], ResNet[6], Inception Net [6], Enhanced Gradient Boosting [8] and hybrid ensemble model [8] are employed for malnutrition prediction in given literature.

Authors of [1], employed AlexNet CNN model for prediction of malnutrition on dataset consisting of children's images. Despite achieving a 96% accuracy with AlexNet, the researchers noted the limitation of the small dataset. To enhance this work can be enhanced further by using larger datasets.

The performance of comparison of various state-of-the techniques for malnutrition detection is discussed in Table 1.

Table 1 . Current Advancements in malnutrition prediction: An Overview of State-of-the-Art Methods and Approaches

Reference	Objective of work	Dataset	Model use and Performance
[1]	Malnutrition detection in children	Malnutrition 250 and Normal children 250 images	AlexNet Accuracy=0.96 F1=0.92 Precision= 0.95

[2]	Classification of undernutrition among children under five in Ethiopian administrative regions.	EDHS	Accuracy (95% CI) GLM- 0.356 (0.344, 0.369) Ridge - 0.649 (0.636, 0.661) Lasso - 0.683 (0.671, 0.695 Elastic-Net- 0.682 (0.670, 0.694) NN- 0.656 (0.644, 0.668) RF- 0.688 (0.676, 0.700)
[3]	Malnutrition status prediction in children	BDHS	**Accuracy (95% CI)** k-NN- 65.59, LDA- 68.57, SVM- 66.74, RF- 68.23, LR-68,13
[4]	Prediction of malnutrition among women	BDHS	NB- 83.0% SVM- 80.7% DT – 85.1% ANN- 81.2% RF- 85.5%
[5]	Edematous malnutrition prediction in afghan children	Data from the secondary sources was gathered from two hospitals located in Afghanistan.	RF- 97.14% J48-94.51%
[6]	Human body height and waist prediction for identification of malnutrition	CAESAR	**ResNet** Average MAE For height prediction- 10 For waist circumference –60 **Inception** Average MAE For height prediction- 9 For waist circumference –59
[7]	malnutrition detection in newborn babies	UNICEF	**For age-related weight** KNN – 95.9%, SVM-97.4% Logistic Regression- 95.9%, Naïve Bayes – 95.4% Neural Network with ReLu – 95.4% Neural Network with Sigmoid- 96.4% **Height-for-age** KNN – 94.3%, SVM-96.2% Logistic Regression- 94.8%, Naïve Bayes – 93.8% Neural Network with ReLu – 97.4% Neural Network with Sigmoid- 97.9%

[8]	Develop an ensemble hybrid model to improve malnutrition prediction accuracy	Demographic and Health Survey data	RF- 0.81% DT- 0.60% EGB- 0.79% KNN- 0.74% Majority voting-based hybrid ensemble – 0.96%
[9]	Malnutrition prediction in 0-59 months of children	BDHS	Class wise accuracy of ANN Wasting – 85.97 Underweight- 69.97 Stunting- 67.27

A thorough analysis of malnutrition detection highlights the predominant use of survey data in existing research. However, the literature review underscores a significant gap in studies pertaining to malnutrition detection through visual analysis. The proposed study intends to bridge this gap by focusing on the classification of malnourished children using a convolutional neural network that assesses their visual attributes.

ML algorithms need handcrafted data which are manually prepared by careful selection of feature sets and this process is time-consuming. In this paper a method of predicting malnutrition status in children by analyzing child image is proposed. The proposed method merely needs children's image and does not depend on any other pathological reports. ML model faces problems in classification of high dimensional data such as image. CNN models are advanced tools for addressing image classification problems, as they possess the intelligence to automatically select features from the provided images for classification purposes. In this study, a custom designed CNN proposed for predicting malnutrition in children using images of children. The detailed methodology of the proposed work is presented in the following section.

3 Problem Formulations or Methodology

In this section proposed methodology for malnutrition prediction in children is illustrated. The proposed method does not need body parameters or pathological reports etc. and it is based on images captured by the camera. DL architectures such as CNN model is explored for prediction for classification of children as nourished or malnourished. Fig. 1 shows various steps of the proposed methodology.

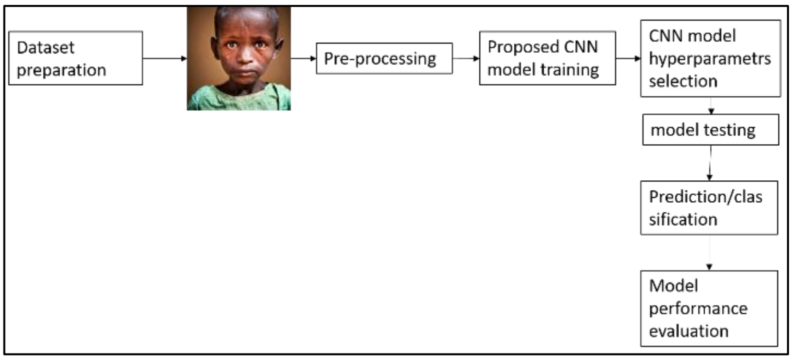


Fig.1. Methodology of proposed work

The various steps are as follows: firstly, image dataset preparation is conducted, followed by image preprocessing and model training. The model's hyperparameters are then investigated through exhaustive experimentation, identifying suitable ones. The model's performance is evaluated by computing accuracy, F1-score, recall, and precision, followed by comparative performance analysis. Qualitative analysis of the model is performed by visualizing CNN layers and activation maps to instill confidence in the model's functionality. Further elaboration of the methodology is provided in the subsequent sections.

3.1 Dataset

Most of the existing methods present in literature are based on ML algorithms which use body parameters stored on csv file. In this work the method of classification children as nourished and malnourished is based on child image. Due to the unavailability of child images in public repository this work uses synthetic images for model training. In this work the children's dataset is generated using stable diffusion (SD) model and the resulting images are segregated, and manual annotated.

SD models enhance the generation of synthetic images by iteratively refining noise levels, contributing to realistic and diverse image synthesis through their ability to capture complex data distributions [10]. Generative models are widely used for tasks such as image synthesis and data augmentation, Diffusion models are successfully applied in synthesis of images which look like realistic ones as appearing images [11]. Synthetic data offers a promising solution, especially when generated from diffusion-based methods like SD [12].

Synthetic images are generated using SD model with the WordNet taxonomy and these synthetics images are utilized for training the model. To investigate the capabilities of the SD model [13]. In this work children's images are generated by SD model by inputting precise text prompt to the model for nourished children category and malnourished children category and the resulting synthetic images are segregated and manually annotated. This image dataset contains both images of nourished children and malnourished children cleaning images.

Additionally, images scraped from specific websites are added to this dataset. The dataset comprises a total of 630 nourished children's images in which 504 are used for model training and 126 are used for testing. And there are a total of 1530 number of images pertaining to malnourished category, in which 1153 are used for model training and 377 images are used for testing. To validate the image dataset generated by SD model medical practitioners are consulted and expert opinion about visual looks of synthetic images appearing as realistic images is taken.

3.2 Preprocessing of images

The input images are preprocessed for feature extraction by CNN model, they are,

- Resize - It is commonly needed for CNN models that all input photos have precisely the same dimensions; therefore, this resizing step makes sure of that. Additionally, it standardizes input size for model architectural compatibility and lowers computational complexity. Input images are loaded and resized to (224,224).
- Image to tensor conversion- Images are converted from their native format, such as JPEG, into PyTorch tensors.
- Normalization - The image tensors' pixel values are standardized to have a mean of [0.485, 0.456, and 0.406] and a standard deviation of [0.229, 0.224, and 0.225].

3.3 Proposed CNN model architecture

CNNs has DL architecture, emerges as a robust choice for challenges related to images. The undeniable impact of CNN models has significantly fueled the growing enthusiasm for DL in recent years. A typical CNN architecture comprises convolutional layers, responsible for feature extraction, and pooling layers, which reduce spatial dimensions. Fully connected layers follow, facilitating high-level reasoning and output generation. Convolutional layers, which house filters or kernels that scan the image to extract important features like edges, patterns, and textures, make up the core of the CNN.

The proposed CNN architecture uses convolution layers followed by batch normalization and max pooling layers. The detailed architecture description is given in Table 2.

Table 2. Model Architecture of each Layer

Layer Type	Output Size	Kernel Size	Activation Function	Additional Info
Input Image	3 x 224 x 224	-	-	RGB image
Convolution (Covn2d)	32 x 224 x 224	3 x 3	ReLU	32 filter, Padding = 1
Batch Normalization	32 x 224 x 224	-	-	Normalize Activations
Convolution (Covn2d)	32 x 224 x 224	3 x 3	ReLU	32 filter, Padding = 1
Batch Normalization	32 x 224 x 224	-	-	Normalize Activations
Max Pooling (MaxPool2d)	32 x 112 x 112	2 x 2	-	Reduce spatial dimensions by 2 x 2
Convolution (Covn2d)	64 x 112 x 112	3 x 3	ReLU	64 filter, Padding = 1
Batch Normalization	64 x 112 x 112	-	-	Normalize Activations
Convolution (Covn2d)	64 x 112 x 112	3 x 3	ReLU	64 filter, Padding = 1
Batch Normalization	64 x 112 x 112	-	-	Normalize Activations
Max Pooling (MaxPool2d)	64 x 56 x 56	2 x 2	-	Reduce spatial dimensions by 2 x 2
Convolution (Covn2d)	128 x 56 x 56	3 x 3	ReLU	128 filter, Padding = 1
Batch Normalization	128 x 56 x 56	-	-	Normalize Activations
Convolution (Covn2d)	128 x 56 x 56	3 x 3	ReLU	128 filter, Padding = 1
Batch Normalization	128 x 56 x 56	-	-	Normalize Activations
Max Pooling (MaxPool2d)	128 x 28 x 28	2 x 2	-	Reduce spatial dimensions by 2 x 2
Convolution (Covn2d)	256 x 28 x 28	3 x 3	ReLU	256filter, Padding = 1
Batch Normalization	256 x 28 x 28	-	-	Normalize Activations
Convolution (Covn2d)	256 x 28 x 28	3 x 3	ReLU	256 filter, Padding = 1
Batch Normalization	256 x 28x 28	-	-	Normalize Activations
Adaptive Average Pooling (MaxPool2d)	256 x 4 x 4		-	Convert spatial dimensions by 4 x 4
Fully Connected (Linear)	512	-	ReLU	Hidden Layer with 5112 units
Fully Connected (Linear)	256	-	ReLU	Hidden Layer with 256 units
Fully Connected (Linear)	128	-	ReLU	Hidden Layer with 128 units
Fully Connected (Softmax)	2	-	ReLU	Output layer with Sigmoid activation function with 2 Units

The specialty of the model architecture is its deep stack of convolutional layers followed by means of absolutely linked layers. This allows it to research hierarchical functions from

the input pictures and capture complicated patterns. Max Pooling layer reduces computing complexity while maintaining essential information. SoftMax activation function is used in output layer for predicting the classification as "nourished" or "malnourished." In the model's training phase, batch normalization layers are used to speed up the training of the model to improve its robustness and efficiency. The adaptive pooling layer at the end of the convolutional layers guarantees that the version can manage to enter images of numerous sizes and continually produces a hard and fast-length feature representation. Batch normalization and dropout layers are blanketed to improve training stability and prevent overfitting.
Activation function introduces nonlinearity to better fit the model and enhance its accuracy. The performance of the model depends on the type of activation function as different activation functions influence model performance in different ways [15]. In this paper, CNN model performance is evaluated as widely used activation functions. An exhaustive survey on activation function is discussed [16]. The effect of various activations on CNN model performance in facial expression recognition is explored [17].

3.2.1 Activation Functions used with CNN.

The performance analysis of proposed custom CNN model is optimized by selecting the right activation function for the given classification problem, that classification of children images. In this work activation functions—ReLU, Leaky ReLU, Sigmoid, Tanh, Hard Tanh, and Swish are used. These functions add non-linear characteristics and allow the model to capture complex correlations in the data.

ReLU (Rectified Linear Unit) - ReLu, as given in equation (1) improves the efficiency of DL model, especially in the hidden layers of CNNs in addition to avoiding vanishing gradient problems.

$$ReLU(x) = \max(0, x) \qquad (1)$$

Leaky ReLU - it addresses the dying ReLU problem by allowing a small gradient for negative values. It is represented by equation (2).

$$LeakyReLU = \begin{cases} \alpha x \; if \; x < 0 \\ x \; if \; x \geq 0 \end{cases} \qquad (2)$$

Where α is a small constant, when x is negative, it returns a small negative value scaled byα, and when x is positive, it returns x itself.

Sigmoid - used in the output layer of binary classification networks, it maps any input into a value between 0 and 1. It is given in equation (3)

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad (3)$$

Tanh - Hyperbolic Tangent – it outputs values between -1 and 1, offering a zero-cantered output which can sometimes make learning easier for the subsequent layers. It is given in equation (4).

$$tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \qquad (4)$$

(e) Hard Tanh - Values below a certain negative threshold are clipped to -1, and values above a positive threshold are clipped to 1. Its characteristic equation is given in the equation (5).

$$HardTanh(x) = \begin{cases} -1\, if\ x < -1 \\ x\, if -1 \le x \le 1 \\ 1\, if\ x > 1 \end{cases} \quad (5)$$

(f) Swish - it is a smooth, non-monotonic function represented by the equation (6) Swish has the advantage of being bounded below (like Leaky ReLU but unbounded above, and its smoothness aids in optimization environments [11].

$$Swish(x) = x.\sigma(x) \quad (6)$$

Where, $\sigma(x)$ is sigmoid function.

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

4 Results and Discussion

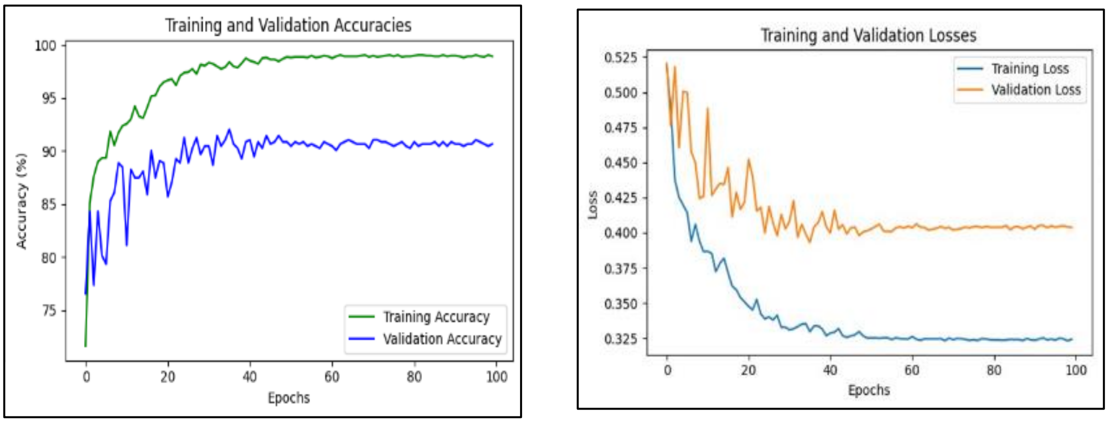
In this proposed methodology for classification of children as healthy and malnourished is experimented using a custom designed CNN which takes as input the children's images and does image classification as healthy child or malnourished child. The model is trained for 100 epochs with Adam optimizer, learning rate 0.0001 and cross entropy loss function. The number of training samples are 1657 images, and the number of validation samples are 503 images. The performance of the model is analyzed for various activation functions that are Swish, LeakyReLu, Tanh and HardTanh. The performance of the CNN model is evaluated by computing metric which includes model accuracy, F1-Score, Recall and Precision. Further, model visualization is implemented by visualizing weight filters, convolutional layers, and activation maps to under the focus on model on image features. Table 3. Shows parameters used for training model.

Table 3. Training parameters of the model

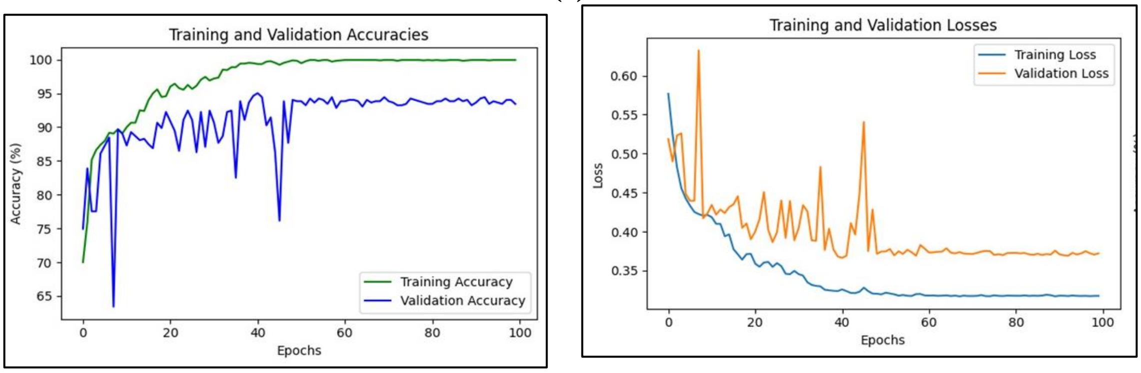
Parameters	**Value**
Epochs	100
optimizer	Adam
Loss function	Cross Entropy Loss
Learning rate	0.0001
No. of training samples	1657
No. of validation samples	503
Augmentation	Transform (Resize, Normalize - Mean & Standard Deviation)

Initially the model was trained for 50 epochs with Adam optimizer, cross entropy loss function, ReLu activation function convolutional layers and softmax for final classification layer with learning rate of 0.001, the model performance achieved was 89.66% accuracy, and for healthy category F1 value 0.79, Precision 0.78 and Recall 0.80, and for malnourished category F1 value 0.93, Precision 0.94 and Recall 0.93. The model prediction accuracy and performance were further observed by increasing number of epochs to 100 and decreasing learning rate to 0. 0001.there is an increase in accuracy to 90.66% and F1 value 0.79, precision 0.83 and recall 0.81 for healthy images, and F1 0.95, Precision 0.93 and Recall 0.94 for malnutrition category. Further, weight decay was applied for model regularization, with weight decay model accuracy improving to 93.44%. Activation functions add non-linear characteristics and allow the model to capture complex correlations in the data, different activations impact model performance in different ways. In this work activation functions—ReLU, Leaky ReLU, Sigmoid, Tanh, HardTanh, and

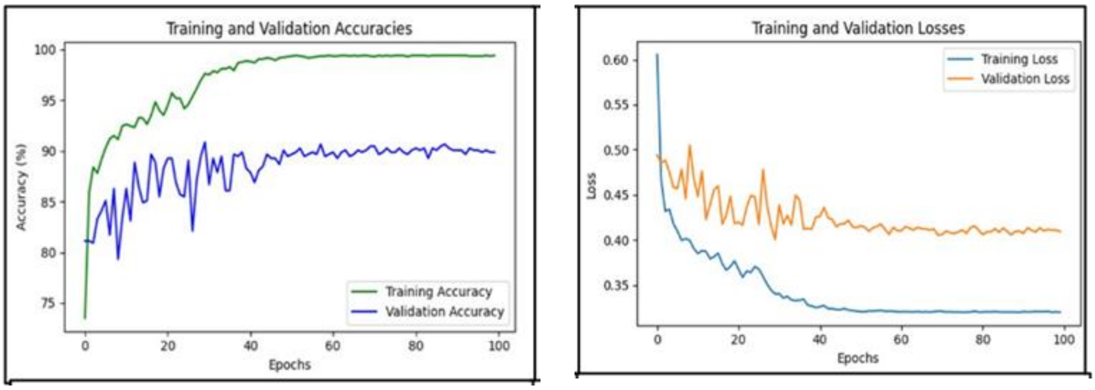
Swish are explored for improving the model performance further. Fig.2 shows loss and accuracy of model for various activation functions.



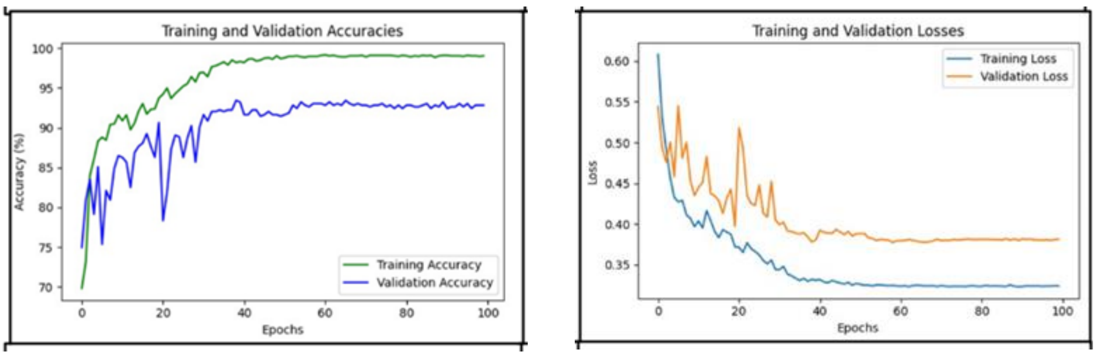
(a)



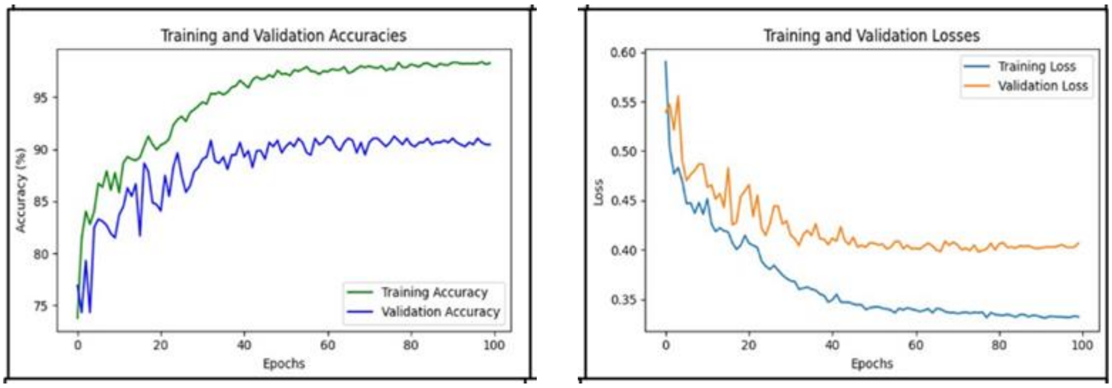
(b)



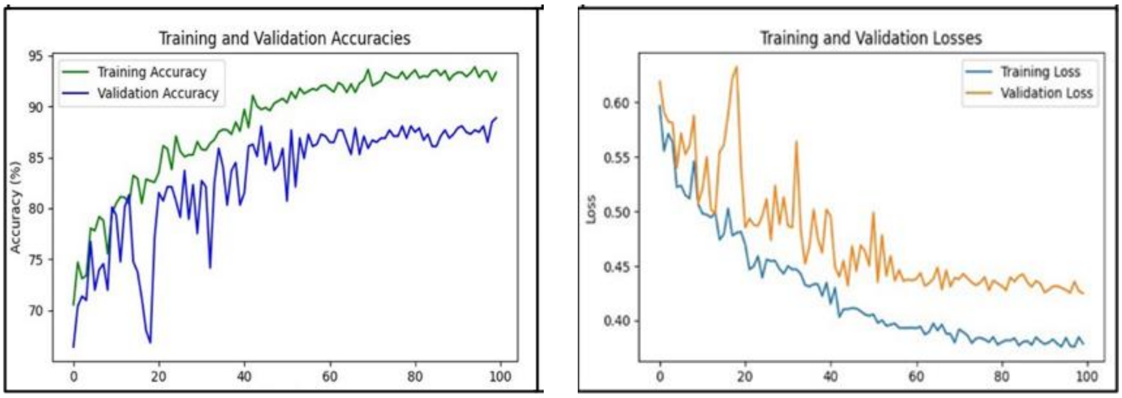
(c)



(d)



(e)



(f)

Fig.2. Accuracy and Loss graphs of model for activation functions
(a) ReLu (b) ReLu with weight decay (c)Swish (d) LeakyReLu (e) Tanh and (f) HardTanh.

The observations of model performance are summarized here, the graph of ReLu and ReLu with weight decay indicates that the model generalizes well to unseen data. Training Accuracy Increases steadily, indicating the model's improving performance on the training set. Around 93.44 % accuracy was obtained as per setting the parameter show in the Experimentation Table 6. Validation Accuracy also increases but seems to plateau earlier than the training accuracy. It maintains a relatively high value, suggesting good generalization.swish activation function’s validation loss decreases similarly to the training loss and achieves accuracy of 89.86%. Leaky ReLU Activation Function- during the first few epochs, loss decreases quickly before levelling out and achieves accuracy of 92.83%. Tanh demonstrates how well the model fits the issue it is meant to address with 90.46% accuracy, and HardTanh- model learns fast and is capable of accurately predicting labels and achieves accuracy of 88.87%. Table. 6 shows the comparative analysis of model performance for various activation functions.

Table 4. Performance Evaluation for Different Activation function

Activation Function	**Nourished**			**Malnourished**			**Accuracy**
	Precision	**Recall**	**F1 - Score**	**Precision**	**Recall**	**F1 - Score**	
ReLU	0.79	0.83	0.81	0.95	0.93	0.95	0.9066
ReLU with	0.89	0.85	0.87	0.95	0.96	0.96	0.9344

weight decay							
Swish	0.87	0.76	0.81	0.91	0.96	0.93	0.8986
Leaky ReLU	0.8	0.9	0.85	0.97	0.94	0.95	0.9283
Tanh	0.87	0.77	0.82	0.92	0.96	0.93	0.9046
HardTanh	0.81	0.76	0.78	0.92	0.93	0.92	0.8887

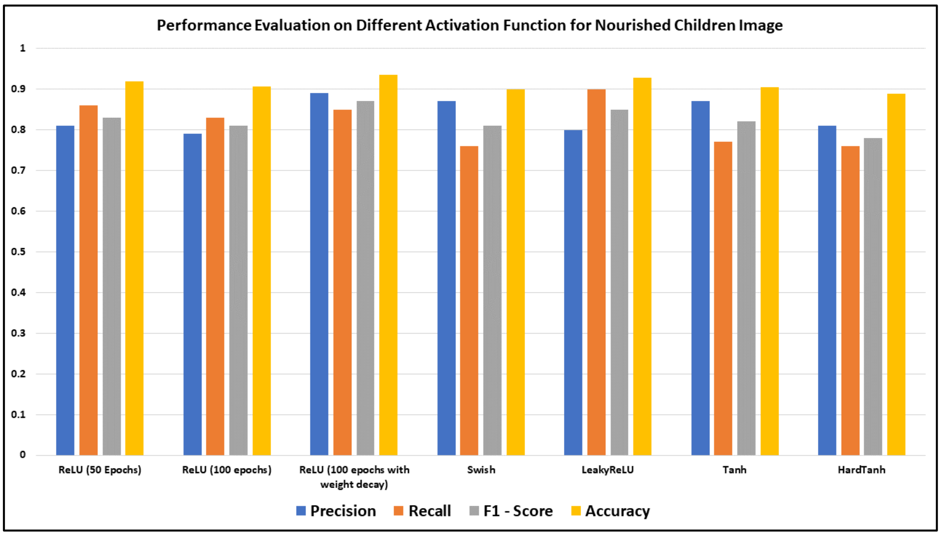


Figure 3. Model Performance for Nourished Children class.

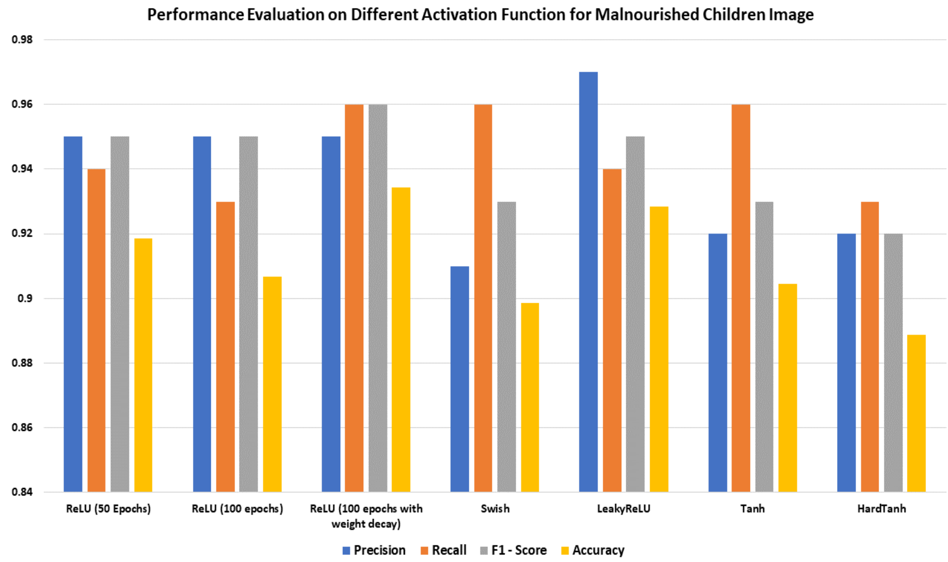


Figure 4. Model Performance for Malnourished Children class.

Figure 3 and 4 shows performance of various activation function for nourished and malnourished children classification. The ReLu with dactivation function gave considerably good results however, the model gave the best accuracy of 93.44% with precision 0.89, 0.85recall, 0.87 F1 score for nourished and precision 0.95, 0.96 recall, and 0.96 F1 score for malnourished category. The model gave good accuracy of 92.83% for

Leaky ReLu with precision 0.8, 0.9 recall, 0.85 F1 score for nourished and precision 0.97, 0.94 recall and 0.95 F1 score malnourished.

5 Qualitative analyses of model.

DL networks are like black boxes, although these networks make the right predictions, the analysis of how predictions are done by the model is not clear. CNN layer explanations give pictures of the inner working of CNN and help in quality analysis of performance of the model [18]. CNN model visualization techniques are important tools that help understand the features considered by network model for taking decisions [19]. CNNs process images and they preserve spatial associations for which model learns. 2-D filter weights of the model can be examined to know the features that the model uses, further, activation maps of convolutional layers are inspected to understand the features extracted from the image. In this work visualization of CNN filters, visualization of feature maps and convolutional layers is implemented. Figure 5. and Figure 6 depicts convolutional layers visualization images of the CNN model presented in this work.

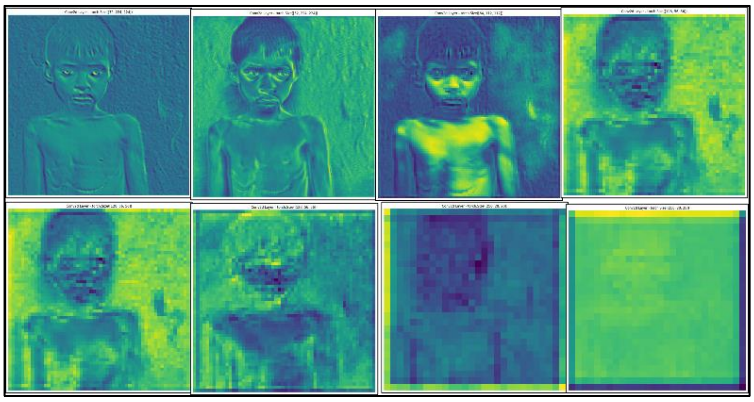


Figure 5. Visualization of convolutional layers- layer 1 to layer 8 of CNN model for sample malnourished child

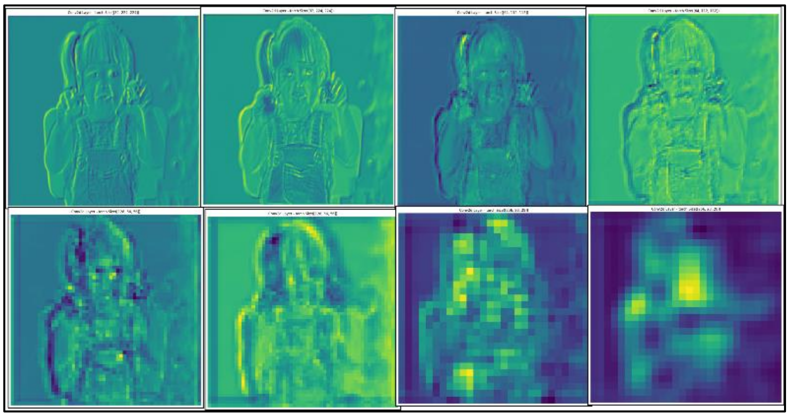


Figure 6. Visualization of convolutional layers- layer 1 to layer 8 of CNN model for sample malnourished child

Filters in CNN are the weights applied to small portions of the input data to produce feature maps. The visualized filter’s function iterates through each filter in a specific convolutional layer. For each filter, it displays the weights as grayscale images. Filters show the learned patterns and features that CNN extracts. Filters help in recognizing patterns, providing insights into what the network focuses on, Figure 7, shows filters weights visualization for convolutional layer 1 and convolutional layer 6, due to the limitation of the space only two layers filter weights are shown.

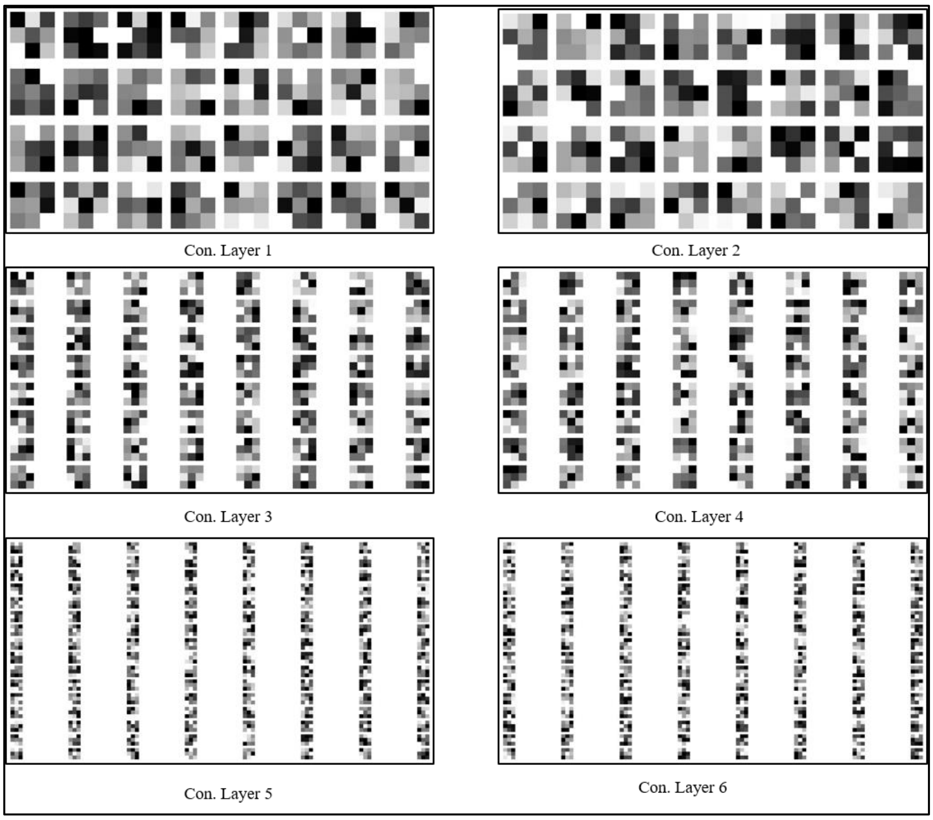


Figure 7. Filter weights visualization for convolutional layer 1 and 6

Activation maps represent the output of applying filters to the input data. The visualize activation map functions take a specific layer and an input image or previous activation map. It visualizes the activation maps produced by each filter in the layer, showcasing the regions where filters respond strongly in the input data. While activation maps highlight areas in the input that trigger certain filters, providing insights into what the network focuses on. These visualizations can be vital for debugging, interpreting, and refining a CNN's performance.

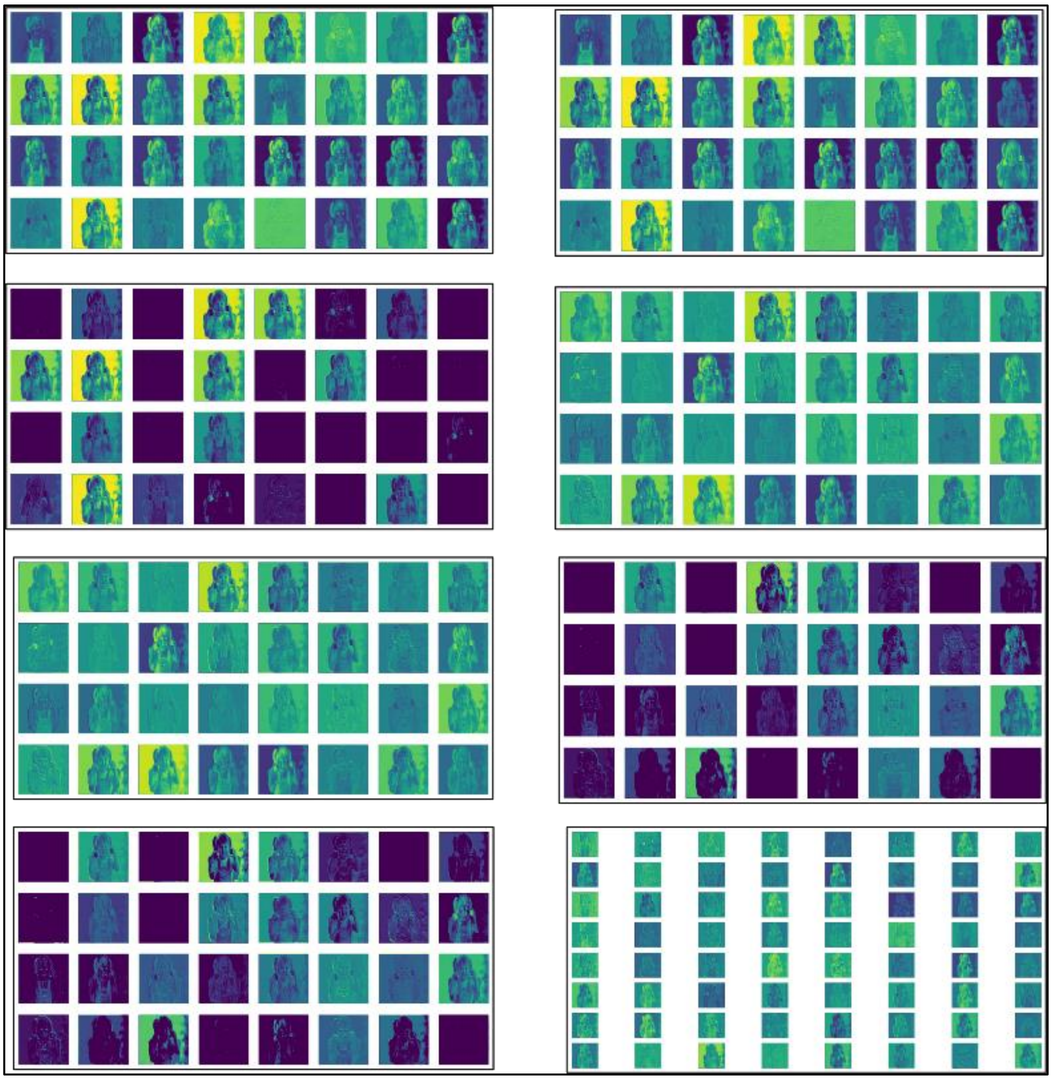


Fig.8. Activation Mapping at Convolution Layer 1-8 from top left to bottom right images.

Activation maps shown in Fig.8 clearly indicate that the model is extracting the appropriate features to understand the malnutrition status in the children. Therefore, it indicates that the model's behavior is satisfactory.

6 Conclusion

In this work a custom designed CNN model is presented for classification of children into categories nourished or malnourished. The model is trained on children's images database that includes both web scraped images and synthetic images. The CNN model is optimized by selecting optimal activation function through experimentation. The model is trained for 100 epochs with Adam optimizer and various activation functions. The ReLu with weight decay activation function gave considerably good results however, the model gave the best accuracy of 93.44% with precision 0.89, 0.85recall, 0.87 F1 score for nourished and precision 0.95, 0.96 recall, and 0.96 F1 score for malnourished category. The model gave good accuracy of 92.83% for Leaky ReLu with precision 0.8, 0.9 recall, 0.85 F1 score for nourished and precision 0.97, 0.94 recall and 0.95 F1 score malnourished. Further, to develop trust in the model visualization of activation maps, convolutional layers and filters are implemented. The model presented in this work provides a sturdy foundation for classifying kids' dietary repute and other recommendations totally on photos. The future work of this research will be firstly to expand the diversity of the training dataset using data augmentation approaches to improve model generalization, secondly expanding the dataset that covers a wider variety of malnutrition-related diseases and various populations may improve the model's applicability in real-world settings.

Data Availability Statement

Data used in this research work is available and will be shared upon request.

Conflict of Interest

All authors declare that they have no conflicts of interest.

References

[1] Lakshminarayanan, Arun Raj, et al. (2021). Malnutrition Detection using Convolutional Neural Network. 2021 In *Seventh International conference on Bio Signals, Images, and In-strumentation (ICBSII*) (pp. 1-5) IEEE.

[2] Fenta, H.M., Zewotir, T. & Muluneh, E.K. (2021). A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. *BMC Med Inform Decis Mak,* 21(291), 1-12. https://doi.org/10.1186/s12911-021-01652-1

[3] Talukder, Ashis, and Benojir Ahammed. (2020). Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. *Nutrition*, 78, (2020): 110861.

[4] Islam, Md Merajul, et al. (2022). Application of machine learning based algorithm for prediction of malnutrition among women in Bangladesh. *International Journal of Cognitive Computing in Engineering,* 3, 46-57.

[5] Momand, Ziaullah, Mohammad Shuaib Zarinkhail, and Mohammad Farhad Aryan. (2022). Machine learning based prediction of edematous malnutrition in afghan children. *Proceedings of International Conference on Emerging Technologies and Intelligent Systems: ICETIS 2021* (Vol. 2) Springer International Publishing, 2022.

[6] Mohammed Khan, Hezha, et al. (2021, November). Predicting Human Body Dimensions from Single Images: a first step in automatic malnutrition detection. CAIP 2021: *Proceedings of the 1st International Conference on AI for People: Towards Sustainable AI*, *CAIP 2021*, Bologna, Italy. European Alliance for Innovation, 2021.

[7] Kishore, K. Krishna, et al. (2023). Prediction of malnutrition in newborn Infants using machine learning techniques. *Research Square*. https://doi.org/10.21203/rs.3.rs-2958834/v1

[8] Khan, Md Nafiul Alam, and Rossita Mohamad Yunus. (2023). A hybrid ensemble approach to accelerate the classification accuracy for predicting malnutrition among under-five children in sub-Saharan African countries. *Nutrition,* 108: 111947.

[9] M. M. Shahriar, M. S. Iqubal, S. Mitra and A. K. Das. (2019). A Deep Learning Approach to Predict Malnutrition Status of 0-59 Month's Older Children in

Bangladesh," 2019 *IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT), Bali, Indonesia*, (pp. 145-149) IEEE. doi: 10.1109/ICIAICT.2019.8784823.

[10] Ho, Jonathan, Ajay Jain, and Pieter Abbeel. (2020). Denoising diffusion probabilistic models. *In Proceedings of the 34th International Conference on Neural Information Processing Systems (NIPS'20).* (no. 574, pp. 6840-6851) Curran Associates Inc., Red Hook, NY

[11] Lorenz, Peter, Ricard L. Durall, and Janis Keuper. (2023). Detecting images generated by deep diffusion models using their local intrinsic dimensionality. *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023. 10.1109/ICCVW60793.2023.00051

[12] Jian, Yanan, et al. (2023). Stable Diffusion for Aerial Object Detection. *arXiv preprint arXiv: 2311.12345*, PP. 1-7.

[13] Stöckl, Andreas. (2023). Evaluating a synthetic image dataset generated with stable diffusion. *Proceedings of Eighth International Congress on Information and Communication Technology. ICICT 2023. Lecture Notes in Networks and Systems*, (Vol 693. pp 805–818) Springer, Singapore.

[14] Hao, Wang, et al. (2020). The role of activation function in CNN.2020 *2nd International Conference on Information Technology and Computer Application* (ITCA). IEEE, 2020.

[15] W. Hao, W. Yizhou, L. Yaqin and S. Zhili. (2020). The Role of Activation Function in CNN," 2020 *2nd International Conference on Information Technology and Computer Application (ITCA)*, Guangzhou, China, 2020, (pp. 429-432). doi: 10.1109/ITCA52113.2020.00096.

[16] Dubey, Shiv Ram, Satish Kumar Singh, and Bidyut Baran Chaudhuri. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 503, pp. 92-108

[17] Wang, Yingying, et al. The influence of the activation function in a convolution neural network model of facial expression recognition. *Applied Sciences* 10(5): 1897, pp. 1-20.

[18] Wang, Zijie J., et al. (2020). CNN explainer: learning convolutional neural networks with interactive visualization. *IEEE Transactions on Visualization and Computer Graphics*. 27(2), pp. 1396-1406.

[19] Mohamed, Elhassan, Konstantinos Sirlantzis, and Gareth Howells. (2022). A review of visualization-as-explanation techniques for convolutional neural networks and their evaluation. *Displays* 73: 102239, pp. 1-21

Notes on contributors

Dr. Shilpa Ankalaki is currently working as Assistant Professor, Department of Computer Science and Engineering, Manipal Institute of Technology Bengaluru, Manipal Academy of Higher Education, Manipal. Her research interests include Machine Learning, Deep Learning, Data Mining and Artificial Intelligence applications. She holds a Ph.D. Degree in Computer Science and Engineering from the Visveswaraya Technological University, Belagavi, India.

***Dr. Vidyadevi G* Biradar** is a professor working in the department of Information science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, India. She has over 25 years of teaching and research experience. She has published around 20 research articles in peer reviewed journals and published 6 patents including Indian and German. She has also published books and book chapters. Her fields of research interest include Digital image processing, Computer vision, Machine learning, Deep learning, and Software systems.

Kushal G. is the student of Data Science program in Nitte Meenakshi Institute of Technology, Bengaluru. His field of interest are Data Science, Computer Vision, Machine Learning and Deep learning.

Kavya N. is the student of Data Science program in Nitte Meenakshi Institute of Technology, Bengaluru. His field of interest are Data Science, Machine Learning and Deep learning.