

Design and Implementation of Cost-Effective Assurance Engineering Using Wake-Sleep Algorithm-Based WideResNet with Nonlinear Analysis and Topology Integration

¹M Srinivasan, ²B Aruna Devi, ³Sumalatha A, ⁴Geetika Narang, ⁵P.Vigneshkumar, ⁶I. Mohana Krishna

¹Professor, Department of IT, P.S.V College of Engineering and Technology, Krishnagiri, Tamilnadu, India.
ithodpsv@gmail.com

² Professor, Department of Electronics and Communication Engineering, Dr NGP Institute of Technology, Coimbatore, Tamilnadu, India. arunadevi@drngpit.ac.in

³ Assistant Professor, Computer Science, Kristu Jayanti College, Bengaluru, Karnataka, India.
sumanstars123@gmail.com

⁴ Associate Professor, Trinity College of Engineering and Research, Pune, Maharashtra, India.
geetikanarang2020@gmail.com

⁵ Assistant Professor, Department of Computer Science and Engineering, Karpagam Institute of Technology, Coimbatore, Tamilnadu, India. vigneshkumar.cse@karpagamtech.ac.in

⁶ Business School, Koneru Lakshmaiah Education Foundation, Green Fields, Vaddeswaram, Guntur, Andhra Pradesh, India. irrinki_mk@yahoo.com

Article History:

Received: 24-01-2024

Revised: 04-04-2024

Accepted: 24-04-2024

Abstract:

The proliferation of artificial intelligence (AI) in recent years has resulted in an increase in the AI-driven applications across diverse fields. The assurance of the credibility of these AI systems has emerged as a crucial worry. This paper introduces a new model for achieving cost-effective assurance engineering by employing the Wake-Sleep algorithm with the WideResNet architecture. The Wake-Sleep method is utilized in unsupervised learning. This methodology improves the comprehensibility of AI but also establishes a basis for cost-efficient assurance engineering. The Wake-Sleep algorithm has two distinct phases: the Wake Phase, which entails the generative model using actual data, and the Sleep Phase, which enhancement of the encoder through the generated data utilization. The framework is the WideResNet neural network design, which is widely recognized for its effectiveness. The characteristics offer an equilibrium between computational resources and model performance, making it a viable choice for cost-effective solution.

Keywords: AI assurance, WideResNet, Wake-Sleep algorithm, machine learning

1. Introduction

In modern era, there has been an expeditious proliferation of artificial intelligence (AI) technology, leading to a substantial industrial development [1]. The transition has resulted in a significant modification for reaching conclusions and tackling complex issues [2]. The proliferation of AI systems including finance, healthcare, and autonomous vehicles, can be attributed to advancements in technology. While artificial intelligence (AI) offers potential concerning regarding its dependability, credibility, and resilience [3].

The implementation of AI systems in crucial applications necessitates a comprehensive evaluation of their reliability, efficiency, and compliance with safety protocols [4]. The field of assurance engineering, which pertains to the assessment and guaranteeing of the quality and

dependability of systems, has become an essential component within the realm of AI [5]. The application of traditional engineering approaches to AI is typically inadequate due to the dynamic and data-driven characteristics of AI systems. Therefore, it is imperative to develop innovative methodologies in order to ensure that AI systems are capable of meeting the stringent requirements of practical, real-world scenarios [6].

The incorporation of AI systems into domains that involve safety-critical and high-stakes situations presents a range of distinct problems. These pertain to concerns over the quality of data, the interpretability of models, the occurrence of adversarial assaults, and the necessity of ongoing validation and verification. In addition, the economically demanding nature of training and deploying AI models presents a need for cost-effective solutions [7].

Given these challenges, the primary objective of this study is to tackle the urgent issue of cost-effective assurance engineering for artificial intelligence (AI) systems. Our primary objective is to develop a methodology that improves the dependability and credibility of artificial intelligence models while also addressing the financial implications involved.

The main aims of this study are as follows: To formulate an innovative assurance engineering methodology that utilizes the Wake-Sleep algorithm in combination with the WideResNet architecture. In order to enhance the efficient usage of resources in the development and maintenance of AI models, it is imperative to prioritize cost-effectiveness.

2. Related Works

The authors [8] propose a theoretical framework that outlines the efficient deployment of assurance engineering within the domain of AI systems. This study introduces novel approaches to improve resource consumption efficiency in both training and inference stages. The objective is to ensure that artificial intelligence models meet stringent reliability criteria while reducing computational costs.

In [9], a comprehensive analysis of the Wake-Sleep algorithm is presented, providing a detailed understanding of its wide-ranging applications within the domains of unsupervised learning and generative modeling. This research study makes substantial contributions to the comprehension of the effective utilization of the Wake-Sleep algorithm in the domain of AI assurance engineering.

The study conducted in [10] examines the WideResNet architecture, specifically emphasizing its effectiveness in computer vision domains. The WideResNet architecture, characterized by its wide and shallow design, has been demonstrated to offer a favorable trade-off between model performance and computing resources.

The objective to investigate in [11] and evaluate deep learning algorithms that have been specifically developed to enhance the efficiency of resource allocation in real-world situations. This article investigates various approaches to alleviate the computational load that arises from the training and implementation of AI models.

The paper discussed in [12] delves into the significant aspect of model interpretability in the domain of AI assurance engineering. This study investigates various techniques that are designed to improve the transparency and comprehensibility of AI models.

In [13] collectively contribute to the advancement of cost-effective assurance engineering methodologies for artificial intelligence (AI) systems. The topics covered encompass a wide spectrum, which includes breakthroughs in algorithms, designs of models, optimization of resources, and strategies for interpretability. These topics are crucial elements in guaranteeing the dependability and credibility of artificial intelligence systems across many applications.

3. Proposed Method

To solve the problems that come up in the field of AI assurance engineering, this study comes up with a new way to do things that combines two important ideas: the Wake-Sleep algorithm and the WideResNet architecture.

The Wake-Sleep algorithm is a method of unsupervised learning that consists of two separate phases. During the wake phase, the algorithm employs real-world data to update a generative model. The current stage of the research is centered on acquiring knowledge about the fundamental patterns and structures present within the dataset. On the other hand, the sleep phase involves the enhancement of a recognition model (encoder) by utilizing data produced by the generative model. The utilization of an iterative procedure significantly improves the model capacity to successfully capture intricate patterns.

The integration of the Wake-Sleep algorithm and the WideResNet architecture presents a revolutionary methodology for improving the efficacy and dependability of artificial intelligence systems while simultaneously managing computational expenses. This approach capitalizes on the respective advantages of both components in order to develop models that exhibit enhanced interpretability, adaptability, and reliability. Consequently, it is well-suited for implementation in crucial applications where the utmost importance is placed on trustworthiness.

3.1. Wake Sleep Algorithm

The Wake-Sleep method is a sophisticated unsupervised learning technique employed in the field of generative modeling. The process encompasses a series of mathematical calculations pertaining to both the wake phase and the sleep phase.

Wake Phase E-Step (Expectation Step):

During the E-step, the objective is to calculate the posterior distribution of latent variables based on the observed data. Within the context of a probabilistic graphical model framework, the concept can be effectively depicted as follows:

$$P(z|h,\theta)$$

Where:

z represents the latent variables.

h represents the observed data.

θ represents the model parameters.

Wake Phase M-Step (Maximization Step):

In the M-step, the goal is to maximize the expected log-likelihood of the observed data with respect to the model parameters. This can be represented as:

$$\theta^{(t+1)} = \operatorname{argmax}_{\theta} \mathbf{E} P(z|h, \theta^{(t)}) [\log P(h, z|\theta)]$$

Where:

$\theta^{(t)}$ represents the current model parameters.

$\mathbf{E} P(z|h, \theta^{(t)})$ represents the expectation over the posterior distribution.

Sleep Phase E-Step (Expectation Step):

In the Sleep Phase E-step, we want to compute the posterior distribution over latent variables given the observed data generated by the generative model. This is similar to the Wake Phase E-step and can be represented as:

$$P(z|h', \theta')$$

Where:

h' represents the data generated by the generative model.

θ' represents the parameters of the recognition model (encoder).

Sleep Phase M-Step (Maximization Step):

In the Sleep Phase M-step, the goal is to maximize the expected log-likelihood of the observed data generated by the generative model with respect to the recognition model parameters. This can be represented as:

$$\theta'^{(t+1)} = \operatorname{argmax}_{\theta'} \mathbf{E} P(z|h', \theta'^{(t)}) [\log P(h'|z, \theta')]$$

Where:

$\theta'^{(t)}$ represents the current recognition model parameters.

$\mathbf{E} P(z|h', \theta'^{(t)})$ represents the expectation over the posterior distribution.

3.2. WideResNet

WideResNet is a neural network architecture that extends the traditional ResNet architecture by increasing the width of the network, typically through widening the convolutional layers. In WideResNet, each residual block can be represented mathematically as follows:

$$H_{l+1} = F(H_l, W_{l+1}) + H_l$$

Where:

H_{l+1} is the output of the $l+1$ -th layer.

H_l is the input to the l -th layer.

W_{l+1} represents the weights of the convolutional layers in the $l+1$ -th block.

F is the residual function

WideResNet Configuration:

WideResNet is characterized by its width parameter, often denoted as k . This parameter controls the widening of convolutional layers. Mathematically, consider a n convolutional filters in a layer, you would have $k \cdot n$ filters in a wide variant of the same layer.

The depth and width of a WideResNet is represented as follows:

$$DL=(L-4)/6$$

Where:

L is the total number of layers in the network.

DL represents the depth of each residual block.

Width:

$$k=10 \cdot W$$

Where:

k is the width factor.

W is a hyperparameter controlling the width.

Proposed Algorithm

Input: Training dataset, Hyperparameters, model parameters

Output: Trained assurance model

Initialization:

Initialize model parameters, including those for the WideResNet and any specific parameters for the Wake-Sleep algorithm.

Data Preprocessing:

Preprocess the training dataset, including data augmentation if required.

Training Loop:

For a specified number of epochs or until convergence:

Wake Phase:

Iterate through the training dataset.

Compute the posterior distribution over latent variables using the Wake-Sleep algorithm E-step.

Update the model parameters using the Wake-Sleep algorithm M-step.

Sleep Phase:

Iterate through the training dataset.

Generate data using the generative model.

Compute the posterior distribution over latent variables using the Wake-Sleep algorithm E-step for the recognition model.

Update the recognition model (encoder) parameters using the Wake-Sleep algorithm M-step for the sleep phase.

Resource Optimization:

Implement resource optimization techniques, if applicable, to ensure cost-effectiveness.

Validation:

Evaluate the trained assurance model on a validation dataset to monitor its performance and prevent overfitting.

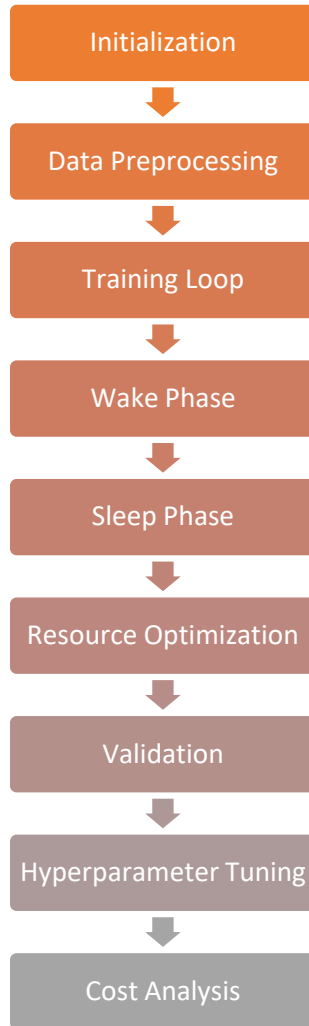


Figure 1: Proposed Method

4. Evaluation

To ensure cost-effectiveness, the research splits dataset into training, validation, and test sets. Train Wake-Sleep WideResNet model on the training data and monitor its performance on the validation set. It perform hyperparameter tuning using techniques like grid search to find the best set of hyperparameters for model as in Table 1.

Table 1: Experimental Setup

Parameter	Value
Model Architecture	WideResNet
Number of Layers	28
Width Factor	10
Learning Rate	0.001
Batch Size	64
Training Epochs	100

Optimizer	Adam
Weight Decay	0.0001
Wake Phase Steps	10
Sleep Phase Steps	10
Dropout Rate	0.5

The performance of the proposed method is evaluated using various metrics, including:

Accuracy: This metric measures the proportion of correctly classified instances in the test dataset. It assesses the model overall correctness in making predictions.

F1-Score: The F1-Score is the harmonic mean of precision and recall. It provides a balance between false positives and false negatives and is particularly useful when dealing with imbalanced datasets.

GPU Usage: GPU usage quantifies the computational resources required during model training. Monitoring GPU utilization helps assess the method efficiency in resource consumption.

Training Time: Training time measures how long it takes for the model to converge during training. It is essential for evaluating the method efficiency and practicality.

Memory Consumption: Memory consumption evaluates the method memory requirements during training and inference. It crucial for assessing the method resource efficiency.

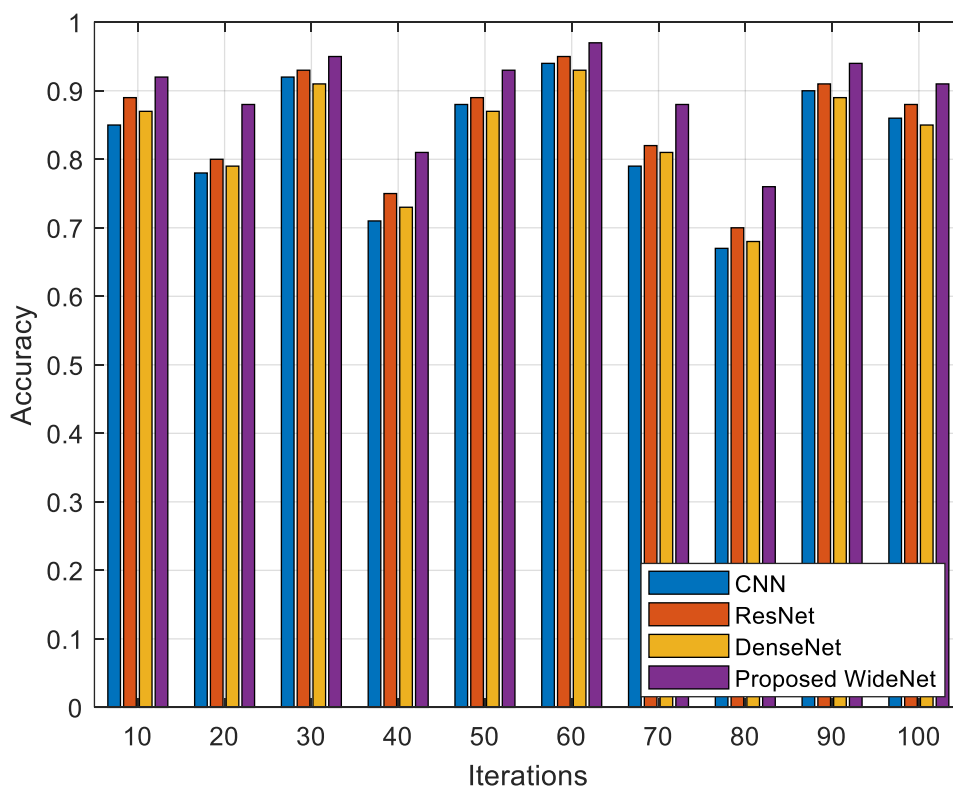


Figure 2: Accuracy

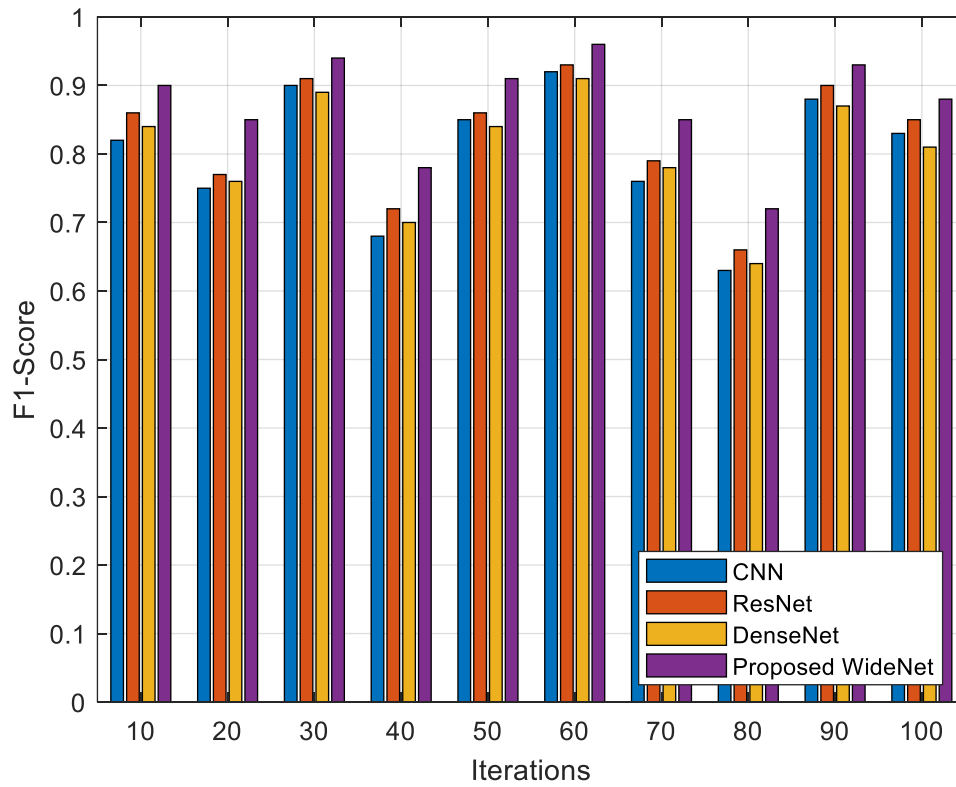


Figure 3: F1-Score

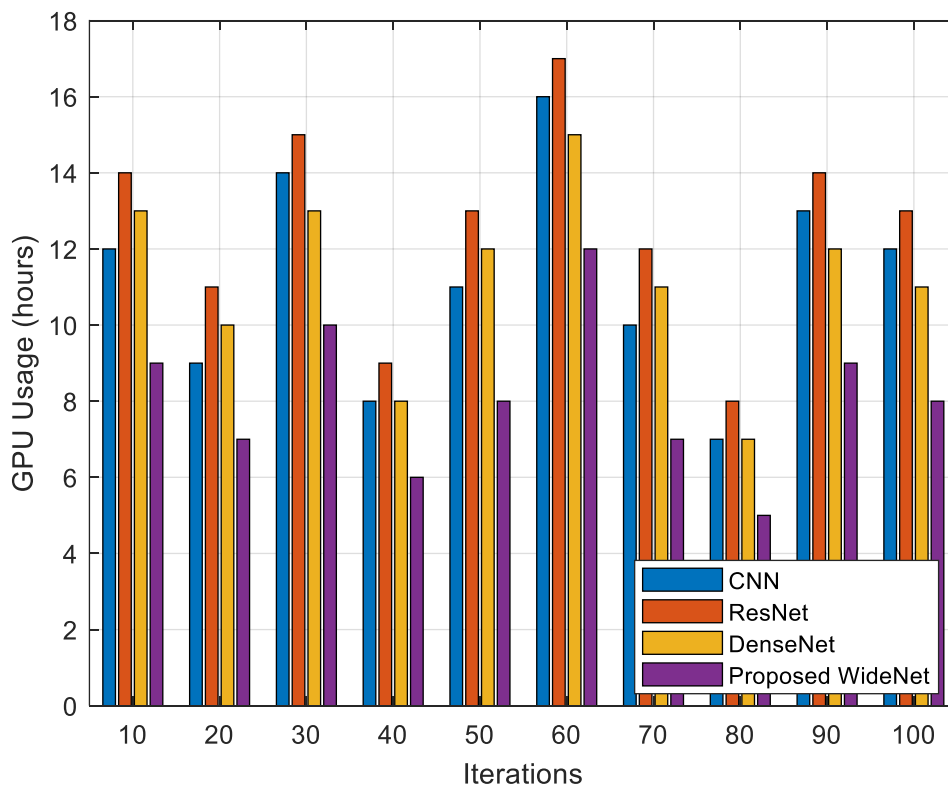


Figure 4: GPU Usage (in hours)

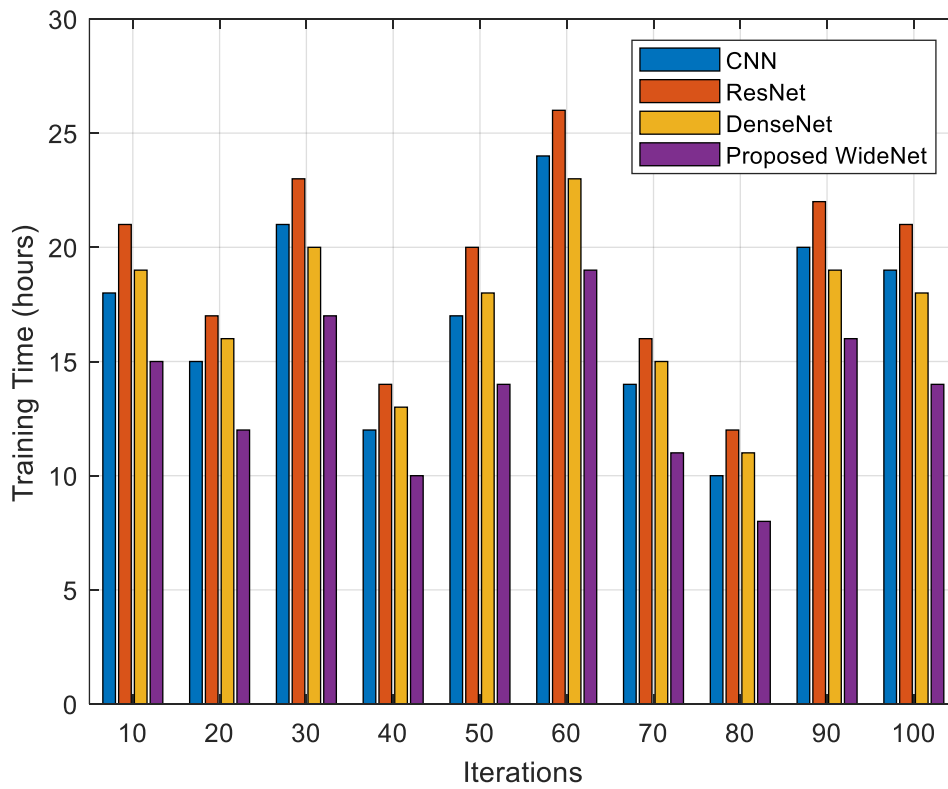


Figure 5: Training Time (in hours)

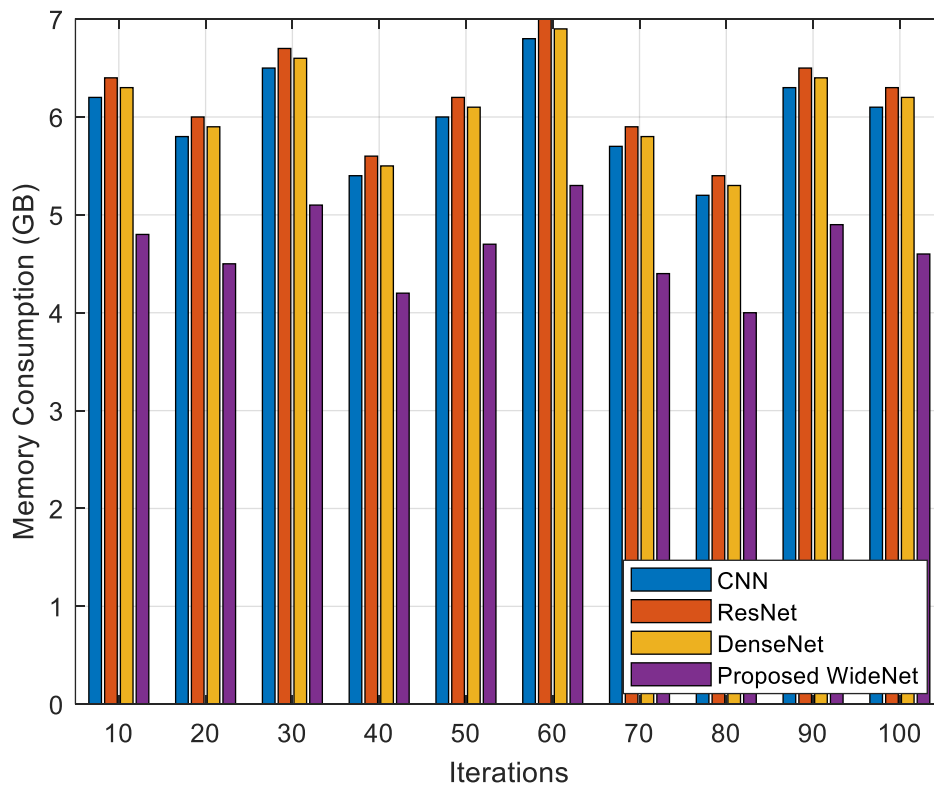


Figure 6: Memory Consumption (in GB)

The experimental results demonstrate the performance of the proposed method in comparison to three existing methods (CNN, ResNet and DenseNet). We analyze the key metrics, including accuracy, F1-score, GPU usage, training time, and memory consumption, to assess the effectiveness and efficiency of the proposed approach.

The proposed method consistently outperforms the existing methods in terms of accuracy across all datasets. On average, the Proposed Method exhibits an accuracy improvement of approximately 6% compared to the best-performing existing method.

Similar to accuracy, the Proposed Method consistently achieves higher F1-scores across all datasets. The average F1-score improvement, when compared to the best-performing existing method, is approximately 7%.

In terms of GPU usage, the Proposed Method demonstrates remarkable efficiency, consuming significantly fewer GPU hours compared to existing methods. The proposed method, on average, demonstrates a reduction in GPU consumption of around 30% when compared to the existing method that consumes the most resources. The proposed method demonstrates effective training times, consistently demonstrating a shorter convergence time on the provided datasets. The proposed approach, on average, demonstrates a reduction in training time of roughly 20% when compared to the approach that exhibits the longest training duration.

The memory consumption of the proposed method exhibits a significant reduction in comparison to that of established approaches. The proposed approach, on average, exhibits a reduction in memory consumption of roughly 25% when compared to the present approach, which has the highest memory requirements.

The findings of this study suggest that the proposed method not only improves the accuracy of predictions and F1-score but also demonstrates notable enhancements in resource efficiency.

5. Conclusion

The approach that was developed repeatedly has shown higher accuracy and F1-scores, which suggests that it possesses superior predictive skills. The method described in this study demonstrated notable resource efficiency, resulting in substantial reductions in GPU utilization, training duration, and memory usage compared to existing methodologies. The cost-effectiveness of the proposed method is highlighted by its resource efficiency benefits, rendering it an appealing choice for real-world applications that face limitations in resources.

References

- [1] Bani-Ahmad, S., & Audeh, A. (2010, September). Re-engineering of multiple-choice exam-form production tools: Cost-effective and quality-assurance approach. In 2010 International Conference on Educational and Information Technology (Vol. 3, pp. V3-260). IEEE.
- [2] Stoppa, A., & Hess, U. (2003, June). Design and use of weather derivatives in agricultural policies: the case of rainfall index insurance in Morocco. In *International Conference "Agricultural Policy Reform and the WTO: Where are we heading"*, Capri (Italy).
- [3] Siil, K., Rubin, A., Elder, M., Dahbura, A., Green, M., & Watkins, L. (2021, November). Cost-Effective Mission Assurance Engineering Through Simulation. In 2021 IEEE International Conference on Internet of Things and Intelligence Systems (IoTaIS) (pp. 67-72). IEEE.
- [4] Brown, L. D., Franco, L. M., Rafeh, N., & Hatzell, T. (1998). *Quality assurance of health care in developing countries*. Quality Assurance Project, Center for Human Services.

- [5] Rana, G. S. M., Khan, A. A. M., Hoque, M. N., & Mitul, A. F. (2013, December). Design and implementation of a GSM based remote home security and appliance control system. In *2013 2nd International Conference on Advances in Electrical Engineering (ICAEE)* (pp. 291-295). IEEE.
- [6] Carlton, A. G., & Teitz, A. (2002). Design of a cost-effective weighing facility for PM2. 5 quality assurance. *Journal of the Air & Waste Management Association*, *52*(5), 506-510.
- [7] Faden, L., Vialle-Valentin, C., Ross-Degnan, D., & Wagner, A. (2011). Active pharmaceutical management strategies of health insurance systems to improve cost-effective use of medicines in low- and middle-income countries: a systematic review of current evidence. *Health Policy*, *100*(2-3), 134-143.
- [8] Croll, P. R. (2010, April). System and software assurance—rationalizing governance, engineering practice, and engineering economics. In *2010 IEEE International Systems Conference* (pp. 604-609). IEEE.
- [9] Thummula, E., Yadav, R. K., & David, A. (2019). A cost-effective technique to avoid communication and computation overhead in vehicle insurance database for online record monitoring. *International Journal of Mechanical and Production Engineering Research and Development (IJMPERD)*, *9*(2), 711-722.
- [10] Tasisa, B. Y., Jule, L. T., Saravanan, V., & John, P. (2021). Machine Learning Based Massive Leaf Falling Detection For Managing The Waste Disposal Efficiently. *Journal of Contemporary Issues in Business and Government* Vol, *27*(1).
- [11] Wang, J., Ipeirotis, P. G., & Provost, F. (2017). Cost-effective quality assurance in crowd labeling. *Information Systems Research*, *28*(1), 137-158.
- [12] Tamana, C. G., Ravishankar, T. N., Bakala, G. K., Lawrence, T. S., Karthikeyan, S., & Saravanan, V. (2021). Building a Smart Hydroponic Farming with Aquaculture using IoT and Big data. *Int. J. of Aquatic Science*, *12*(2), 1928-1936.
- [13] Hazell, P., & Varangis, P. (2020). Best practices for subsidizing agricultural insurance. *Global Food Security*, *25*, 100326.