BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 69(4), 2021, Article number: e137728
DOI: 10.24425/bpasts.2021.137728

An enhanced performance evaluation of workflow computing and scheduling using hybrid classification approach in the cloud environment

P. THARANI* **and A.M. KALPANA**

Department of Computer Science and Engineering, Government College of Engineering, Salem-636011, Tamil Nadu, India

Abstract. Workflow scheduling is the major problem in cloud computing consisting of a set of interdependent tasks which is used to solve the various scientific and healthcare issues. In this research work, the cloud-based workflow scheduling between different tasks in medical imaging datasets using Machine Learning (ML) and Deep Learning (DL) methods (hybrid classification approach) is proposed for healthcare applications. The main objective of this research work is to develop a system which is used for both workflow computing and scheduling to minimize the makespan, execution cost and to segment the cancer region in the classified abnormal images. The workflow computing is performed using different Machine Learning classifiers and the workflow scheduling is carried out using Deep Learning algorithm. The conventional AlexNet Convolutional Neural Networks (CNN) architecture is modified and used for workflow scheduling between different tasks to improve the accuracy level. The AlexNet architecture is analyzed and tested on different cloud services Amazon Elastic Compute Cloud- EC2 and Amazon Lightsail with respect to Makespan (MS) and Execution Cost (EC).

Key words: cloud; workflow scheduling; machine learning; CNN; AlexNet.

1. INTRODUCTION

The developing information technologies require cloud computing for efficient resource handling process as the back-end system. The Amazon EC2 is the best example for the infrastructure-based cloud model which can be functioned with several numbers of Virtual Machines (VM). Each VM provides the executing platform for the remote users to execute the developed programs to eliminate the problems of resource constraint environment at the user end [1–2]. By providing number of VM in cloud environment, the Quality of Service (QoS) of the system will be improved. These VMs can be maintained with the advanced computer processing elements by the Cloud Service Providers (CSP). Based on the services of the VMs, CSP allocates the paying fee for the customers to use the maintained VMs in cloud environment. Most of the CSP provide storage space for the customers in order to save the executed data in cloud environment, which helps the customers to operate the executed data from anywhere in the world through the internet [3–4], as depicted in Fig. 1.

In order to reduce the power consumption of the VMs in the cloud environment, task scheduling activity is important. Several researchers provide different task scheduling methods for VMs in a cloud in order to reduce the power consumption and to reduce the execution time of each task in the cloud. Most of these methods used soft computing techniques such as Machine Learning methods and evolutionary approaches such as Particle Swarm Optimization (PSO), Genetic Algorithm (GA) for resolving the problems during the workflow scheduling process. The makespan time of the workflow scheduling process of these methods is high and hence the accuracy level of the scheduling functions will not be optimum. Hence, this research work develops the hybrid classification approach which combines both Machine Learning (ML) and Deep Learning (DL) methods for the effective workflow scheduling process in different cloud environments. This proposed method can function in two modes as user end and remote end. In user end, ML algorithms are used, and the DL algorithms are used in remote end at cloud system.

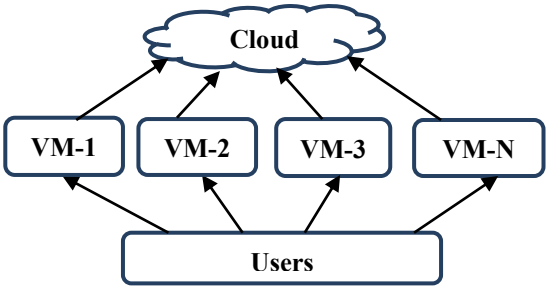

Fig. 1. Cloud System

The main research goals of this work are:

- To develop a cloud-based system used for both workflow computing and scheduling to minimize the Makespan (MS) and Execution Cost (EC) in cloud environments for healthcare applications using hybrid ML and DL approach.
- To identify the cancer region in the classified abnormal images using threshold method.

*e-mail: tharanisiva05@gmail.com

Manuscript submitted 2020-12-16, revised 2021-05-21, initially accepted for publication 2021-05-24, published in August 2021

2. LITERATURE REVIEW

Muhammad Sardaraz *et al*. [5] used the GA process for workflow scheduling on heterogeneous model. The multi-objective scheduling function was applied on the number of multi-behavior tasks in order to allocate the scheduling time period between the set of multi-objects in a cloud system. The scheduling objects of these tasks were used, and resource allocation was done on these scheduling tasks in cloud environment. Hosseinzadeh *et al*. [6] analyzed various scheduling methods between the multi-objective tasks in cloud system environment. The authors discussed the several methodologies for improving the scheduling accuracy with simulation environments. The extensive results were compared between different workflow scheduling methods. Chen *et al*. [7] resolved the unbalancing problems in cloud based multi-object tasks functionalities with resource constrain domains. The authors developed Reference Queue based Cloud Service Architecture (RQCSA) for proving an effective solution for resource allocations in the dynamic cloud resource environment. The multi-objective results between numbers of task functions were compared with state-of-the-art method on various cloud platforms. Hu *et al*. [8] developed an effective flow for the multi-task workflows between numbers of multi-object-oriented tasks in cloud system. The authors used GA method for allocating the static resources between the tasks. The simulation was done in the real time cloud environment with MS and EC parameters. Grochowski *et al*. [9] improves Deep Neural Network (DNN) training and focuses on the most popular Convolutional Neural Networks (CNN) structures, namely Visual Geometry Group (VGG) based Neural Networks (NN) and the experiment was conducted on skin cancer detection.

Nasr *et al*. [10] extended cloud based effective system for performing workflow scheduling between numbers of tasks. This method used Chemical Reaction Optimization (CRO) approach for workflow computations and Ant Colony Optimization (ACO) for workflow scheduling. The integration of these two methods for workflow scheduling improved the performance and the authors tested the methodology on CloudSim toolkit. The performance of this method was compared with PSO technique. Swiderska-Chadaj *et al*. [11] reveals the DL methods for the detection and segmentation of the damaged regions in images. This technique is based on CNNs and uses the U-net model to achieve the pixel-wise segmentation of these unwanted regions.

Cui *et al*. [12] solved the task scheduling problems in the cloud environment using an evolutionary technique. The authors used the top-down modeling method to allocate the priority to each task in the system. The population diversity of each task was improved by the GA functionality. The scheduling time period was determined by the fitness function in GA. The authors analyzed the effectiveness of the method with the help of the scheduling cost function on cloud environment. Wang *et al*. [13] developed two evolutionary approaches for workflow scheduling in cloud environment using Adaptive Genetic Algorithm (AGA) and Load Balancing Genetic Algorithm (LBGA). The fitness model was improved using AGA method and the overall fitting property was improved using LBGA. The authors used a greedy algorithm to improve the population in GA during workflow scheduling. The multi-functionality fitness model was utilized in order to improve the scheduling accuracy of this proposed system.

3. MATERIALS AND METHODS

In this section, the materials and methodologies are discussed with the cloud services.

3.1. Materials

Three different datasets are utilized for analyzing the effectiveness of the proposed method. They are given as brain dataset [14], lung dataset [15] and breast dataset [16]. In this research work, 1500 noncancer affected images (healthy brain images) and 1200 cancer affected brain Magnetic Resonance Imaging (MRI) images are obtained from the brain dataset. 890 of noncancer affected images and 850 of cancer affected Computer Tomography (CT) lung images are obtained from the Lung dataset. The breast dataset consists of 1600 of noncancer affected images and 1200 of cancer affected MRI images.

In this work, 60% of the images in individual dataset are used for training. The remaining 40% of the images are used for testing/validation. Hence, 900 noncancer affected brain images and 720 cancer affected brain images are used for training the proposed model. 600 noncancer affected brain images and 480 cancer affected brain images are used for testing. 534 noncancer affected lung images and 510 cancer affected lung images are used for training. 356 noncancer affected lung images and 340 cancer affected lung images are used for testing. 960 noncancer affected breast images and 720 cancer affected breast images are used for training and 640 noncancer affected breast images and 480 cancer affected breast images are used for testing. All the images in these individual datasets have the image resolution of 256×256 pixels with grey scale format. These datasets are also having manual cancer segmented images and the images in these datasets are license free for the researchers.

The two types of clouds from Amazon (Amazon Elastic Compute Cloud-EC2 and Amazon Lightsail cloud) [17] were used for the classification and segmentation process of the cancer affected images in cloud environments. These two clouds can be utilized as public cloud for executing many application-oriented programs and support both reliable, secure services for the user-end applications. The performance accuracy may be varied with respect to the cloud service environments.

3.2. Methods

The cloud environment-based workflow scheduling between different tasks in medical imaging datasets using ML and DL methods (Hybrid classification approach) is proposed for healthcare applications. Figure 2 shows the DL based workflow computations and scheduling between medical imaging datasets in healthcare application.

The DL based workflow computations and their scheduling between tasks have the following sub-blocks.

- Workflow computations using ML algorithms
- Workflow scheduling using DL algorithm
- Segmentation

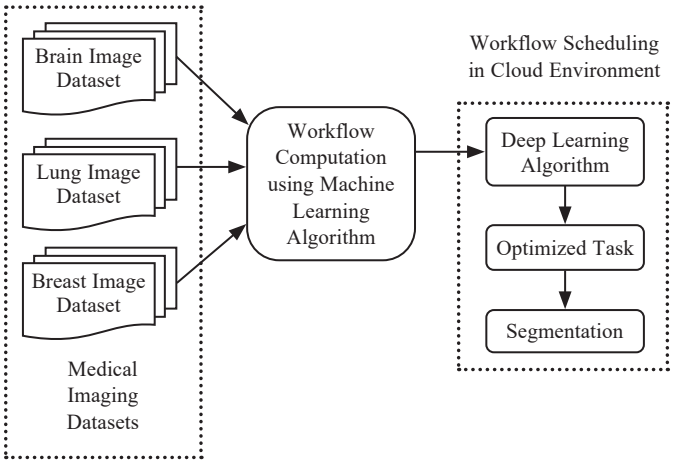

Fig. 2. DL based workflow computations between medical imaging datasets

3.3. Workflow computations using ML algorithm

Three medical application tasks are considered. Each task has a separate dataset which acquires the scanned images from the patient in a health center and stores in the respective datasets. The datasets are brain, lung, and breast. The images in these datasets are classified into either cancer affected or normal images using ML algorithms.

Workflow Computation Algorithm

Input: Brain, Breast and Lung images.
Output: Tumor or not of individual images.
Start;
Step 1:
The medical grey scale images in each individual task are enhanced using Adaptive Histogram Equalization (AHE) method [18].
Step 2:
Convert these spatial domain pixels of the enhanced images into the multi-domain pixels in order to improve the classification rate using Gabor transform [19].
Step 3:
Compute the Local Binary Pattern (LBP) and Local Ternary Pattern (LTP) features [20] from the multi-domain converted images and these features are stored in the Feature Index Matrix (FIM).
Step 4:
(a) Cancer affected brain images are detected by applying the Co-Active Adaptive Neuro Fuzzy Inference System (CANFIS) classifier [21] on the computed FIM.
(b) Cancer affected lung images are detected by applying the Neural Network (NN) classifier [22] on the computed FIM.
(c) Cancer affected breast images are detected by applying the Adaptive Neuro Fuzzy Inference System (ANFIS) classifier [23] on the computed FIM.
Step 5:
The task weight is computed using the following procedures.

(i) The FIM is computed for all the images in each individual task. Find the alpha index using the following equation.

$$\alpha = (\mathrm{MaFV} - \mathrm{MiFV})/\mathrm{N} \tag{1}$$

Where, MaFV is the Maximum Feature Value and MiFV is the Minimum Feature Value in the computed FIM. N is the number of computed features in FIM.

(ii) Determine the heuristic index (h) of the FIM using the following equation.

$$\mathrm{h} = \sum \mathrm{FIM}/\mathrm{N} \tag{2}$$

(iii) Compute the weight of the individual task using alpha index and heuristic index as stated below.

$$\mathrm{W} = (\alpha - \mathrm{h})/\mathrm{h} \tag{3}$$

End.

The computed weights of the individual task along with the task counts, size of task, MS and EC of each task are fed into the DL algorithm for selecting the optimized task in cloud environment. The cancer regions are segmented in the images which are available in optimized task only.

3.4. Workflow scheduling using DL algorithm

Scheduling is the process of selecting the optimized task between the set of tasks. This scheduling process is executed in Remote Cloud Server (RCS). After scheduling, the optimized task is applied to segmentation module which is used to locate the cancer regions in the images of optimized task. The DL algorithm – CNN is utilized to perform the scheduling between the tasks. In the existing research works, the authors mainly used PSO [1] and GA [2] for selecting the optimized task among the set of tasks. These conventional methods required large number of input parameters for scheduling process. Moreover, these methods consumed high span time of each individual task which makes the scheduling process more complex. In order to eliminate such limitations in existing works, CNN algorithm is utilized for optimized scheduling process.

The conventional AlexNet CNN architecture is modified and used for the scheduling process. The input parameters for this proposed AlexNet CNN architecture are the task counts, image counts in each individual task, task size, weight, individual task makespan and execution cost. These parameters are computed and given to the AlexNet CNN architecture in matrix format along with the data augmented dataset.

3.5. Data augmentation

This method is used to increase the images count in each individual task for improving the classification accuracy of the CNN architecture. In this work, only the tumor affected images in each individual task are data augmented and they are fed with the input task parameters to the CNN architecture for selecting the optimized task in cloud environment. This work consists

of data augmentation methods as right shifting, left shifting, color space transformation (selecting luminance component) and flipping process [24]. These methods are applied to the individual task (cancer affected images only) and the data augmented images are stored in a separate dataset. The brain dataset consists of 1200 cancer affected images, lung dataset consists of 850 cancer affected images and breast dataset consists of 1200 cancer affected images. The new dataset after the data augmentation process consists of 4800 cancer affected brain images (4 data augmentation methods * 1200 brain images), 3400 cancer affected lung images and 4800 cancer affected breast images. The total number of images after data augmentation method is about 13,000.

At the end of the data augmentation process, task counts as 3, image counts in each individual task as 658 (2700 + 1740 + 2800 = 7240), task size as 474, 480, 640 (256 * 256 * 7240), task weight, individual task makespan and execution cost along with these data augmented 13,000 images are fed into the Modified AlexNet CNN architecture for selecting the optimized task.

3.6. CNN architecture

The modified AlexNet CNN architecture consists of the following blocks.

- Convolutional Layers (CL);
- Pooling Layers (PL);
- Fully Connected Neural Networks (FCNN).

In this proposed work, the CNN architecture is used to produce the optimized task. There are many conventional CNN architectures available as LeNET, AlexNet, GoogLeNet etc., these CNN architectures differ based on the number of internal blocks in each architecture. The AlexNet CNN architecture is able to train huge number of parameters and attain high classification accuracy. This architecture also achieved the best performance with minimum computations and training time which implies better efficiency compared to other CNN architectures.

Figure 3a shows the conventional AlexNet architecture and Fig. 3b shows the proposed AlexNet architecture. The conventional AlexNet CNN architecture [3] used five numbers of CL, three numbers of PL and three numbers of FCNN layers as illustrated in Fig. 3a.

In the conventional CNN architecture, first CL is designed with 16 numbers of Convolutional Filters (CF) with the kernel size of 3×3, the second CL is considered with 32 numbers of CF with the kernel size of 3×3 and the third CL is designed with 64 numbers of CF with the kernel size of 3×3 and the fourth CL is designed with 128 numbers of CF with the kernel size of 3×3 and the fifth CL is designed with 256 numbers of CF with the kernel size of 3×3. The size of all CF in all CL in conventional CNN architecture is fixed.

The function of CL is to convolve the input parameters with the Convolutional Kernel (CK). Hence, the convolved sequences are having higher size responses. This will increase the computational time of the other consequent sub-blocks in CNN architecture. The size of this CL response should be reduced and for this purpose, the CL responses are passed through the PL which reduces the size of the CL responses. The conventional AlexNet architecture used Average PL which selects the average response of the 3×3 window which was placed over the CL responses. The size of PL is equal to the size of the CF. Finally, three numbers of FCNN are used in conventional AlexNet architecture which produces the classified responses. The first FCNN with 2048 numbers of neurons, the second FCNN with 1024 numbers of neurons and the third FCNN with 524 numbers of neurons were considered. The average PL in conventional AlexNet architecture yields to errors during the size reduction process of the CL responses. This reduces the classification accuracy of the system. Hence, the

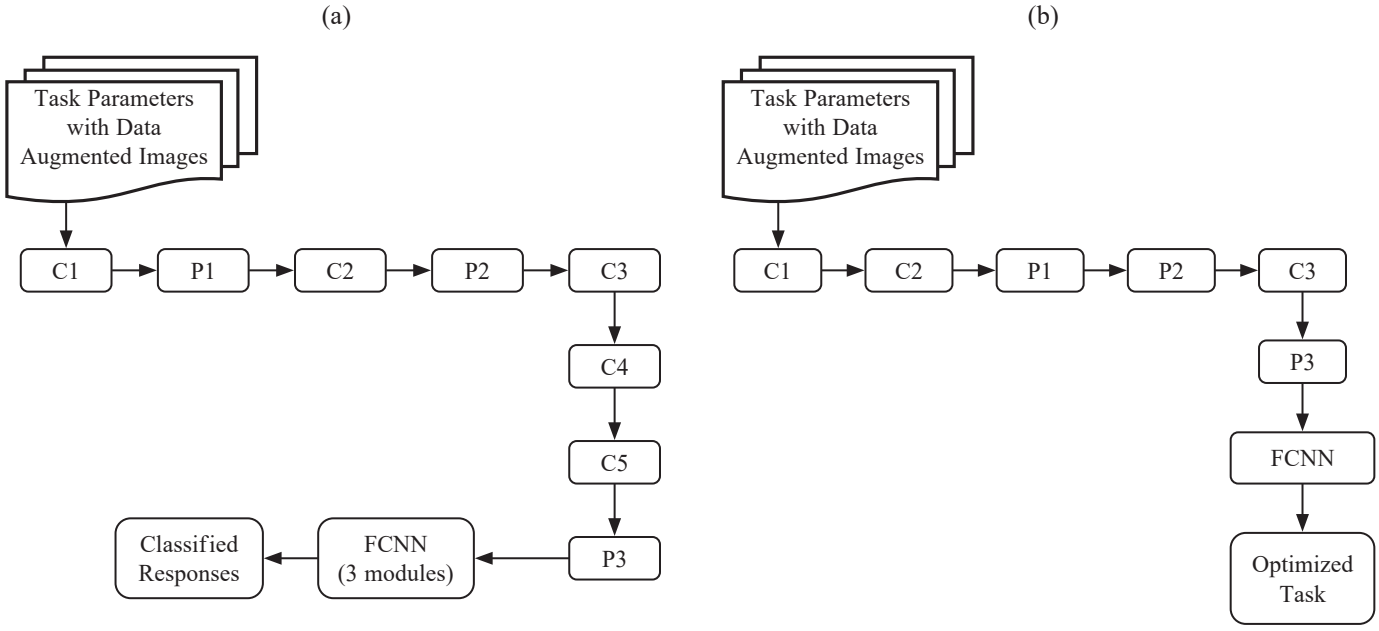

Fig. 3. CNN architectures: a) Conventional AlexNet architecture [3], b) Proposed AlexNet architecture

proposed AlexNet architecture is designed with three CL and three PL and one FCNN as depicted in Fig. 3b.

In the proposed design, Rectifier Linear Unit (ReLU) is integrated with each CL for eliminating the negative responses in the convolved sequences to reduce the errors propagation through the next layers. In this AlexNet architecture, Max-PL is used instead of Average-PL in the conventional AlexNet architecture. The Max-PL selects the maximum value of 3 * 3 window which is placed over the CL responses. The value is not changed during the PL process which improves the classification accuracy of the system.

The task parameters with data augmented images are convolved with first CL. The first CL is designed with 64 numbers of CF with the kernel size of 5×5. The convolved responses may have the negative values which can be converted into positive values using ReLU module. The output parameters from this ReLU module are again convolved with the second CL which is designed with 128 numbers of CF with the kernel size of 7×7. Also, ReLU module is placed after second CL to convert the negative convolved output into positive convolved output. The size of kernel in each CF is chosen after several trails in order to produce the optimized output results. The size of the output parameters from the second CL is reduced by passing these into two consecutive PL. The size reduced parameters from the second PL are again convolved with the third CL which is designed with 256 CFs with the kernel size of 9×9. The size of the output convolved parameters from the third CL is reduced using third PL and subsequentially passed to the FCNN. Two FCNN layers are used, and each layer is designed with three layers one input, one output and 16 hidden layers. Each hidden layer in FCNN is designed with 36 numbers of neurons. The input and output layer are designed with 15 and 2 numbers of neurons respectively. The number of hidden layers and the number of neurons in each hidden layer is chosen after several trails in order to obtain the optimum classification results.

The proposed AlexNet architecture is modified from the conventional AlexNet architecture with respect to the following aspects.

- The AlexNet architecture is designed with a smaller number of CL, PL and FCNN layers which minimizes the MS and EC.
- The errors in convolved sequences are reduced in AlexNet architecture by implementing ReLU with each CL.
- The number of internal layers and the number of neurons in each FCNN is reduced in the CNN architecture.

3.7. Segmentation

The threshold method is used as the segmentation technique for detecting the cancer regions in the classified abnormal images. The histogram count for each pixel in classified abnormal image is computed. The average between the minimum and maximum histogram value is computed and noted as threshold 1. The difference between the minimum and maximum histogram value is determined and noted as threshold 2.

The pixels that lie between the computed thresholds (threshold 1 and threshold 2) are segmented as cancer pixels. Figure 4a, 4c and 4e shows the source images of a brain, lung, and breast. Figures 4b, 4d and 4f shows the cancer region segmented images, respectively.

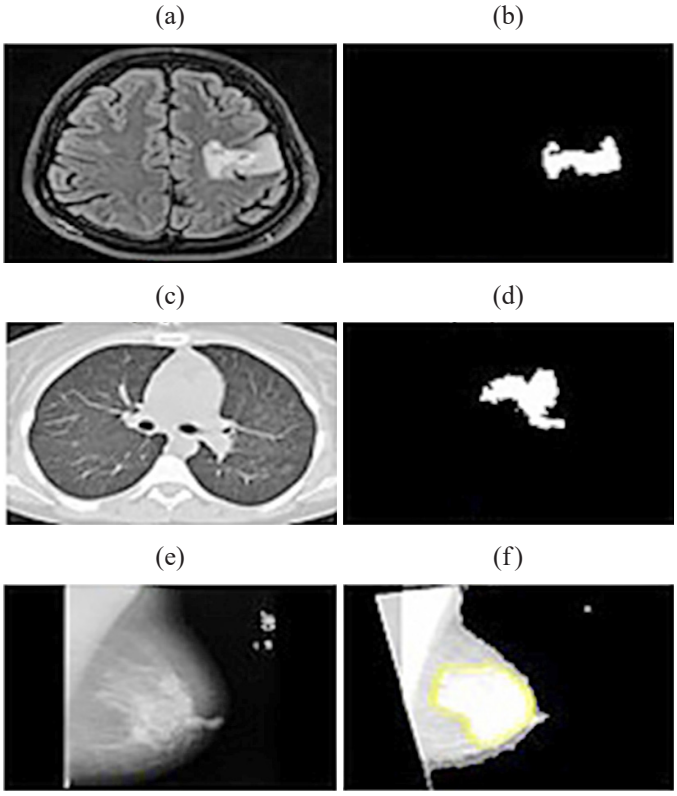

Fig. 4. a) Brain image, b) Cancer region segmented brain image, c) Lung image, d) Cancer region segmented lung image, e) Breast image, f) Cancer region segmented breast image

4. RESULTS AND DISCUSSIONS

The workflow scheduling methodology using DL classification approach is tested in MATLAB R2018 version with inbuilt DL simulation toolbox. The cloud integration setup is established in MATLAB simulating tool through reliable hardware architectures with Intel Core i7-eigth generation processor and 8 GB RAM with Quad core 1.8 GHz and Windows 10 operating system.

The workflow scheduling is analyzed with respect to four different system environments. The system A, B, C, D consists of 100, 500, 1000, 10000 sets of task functions. Each system environment is tested by ML approaches (Support Vector Machine- SVM and NN) and DL approach (CNN) with respect to MS and EC. Both MS and EC are measured in seconds. The time for executing the individual task function is measured by MS and the average time for executing all task functions in cloud environment is measured by EC.

Table 1 is the analysis of CNN based workflow scheduling approach with respect to various systems on different cloud environments using conventional AlexNet architecture in terms of MS and EC.

Table 1
Analysis of CNN based workflow scheduling approach with respect to various systems on different cloud environments (Conventional AlexNet architecture)

Task Environmental system	Approaches	Number of tasks	MS (s)	EC (s)
Amazon EC2 Cloud Environment				
System A	SVM	100	45.7	42.1
	NN		39.1	32.7
	CNN		29.1	22.8
System B	SVM	500	56.9	48.9
	NN		51.6	37.1
	CNN		32.7	27.7
System C	SVM	1000	61.7	56.9
	NN		56.8	49.1
	CNN		45.1	43.9
System D	SVM	10000	67.9	65.8
	NN		61.6	60.7
	CNN		48.1	45.8
Amazon Lightsail Cloud				
System A	SVM	100	56.9	41.9
	NN		41.8	32.9
	CNN		33.9	29.1
System B	SVM	500	61.7	51.9
	NN		55.6	44.5
	CNN		41.9	29.8
System C	SVM	1000	67.9	61.9
	NN		58.8	55.6
	CNN		45.6	42.1
System D	SVM	10000	72.9	69.1
	NN		61.9	56.9
	CNN		55.6	45.9

Table 2
Analysis of CNN based workflow scheduling approach with respect to various systems on different cloud environments (Modified AlexNet architecture)

Task Environmental system	Approaches	Number of tasks	MS (s)	EC (s)
Amazon EC2 Cloud Environment				
System A	SVM	100	38.1	21.7
	NN		31.9	20.1
	CNN		21.2	15.8
System B	SVM	500	46.7	31.9
	NN		43.1	29.7
	CNN		28.9	17.9
System C	SVM	1000	51.9	45.9
	NN		49.8	41.8
	CNN		35.9	22.7
System D	SVM	10000	59.9	56.1
	NN		55.8	52.6
	CNN		33.9	26.7
Amazon Lightsail Cloud				
System A	SVM	100	44.7	32.9
	NN		37.1	29.7
	CNN		31.9	19.6
System B	SVM	500	51.9	46.7
	NN		47.5	39.1
	CNN		38.4	24.5
System C	SVM	1000	57.5	55.9
	NN		51.9	44.7
	CNN		42.8	29.9
System D	SVM	10000	61.7	72.9
	NN		55.9	65.8
	CNN		45.7	32.9

The workflow scheduling approach is tested using SVM, NN and CNN architectures, on Amazon EC2 and Lightsail. Table 2 is the analysis of CNN based workflow scheduling approach with respect to various systems using ML and DL classification methods in terms of MS and EC. In Amazon EC2 cloud services, the proposed workflow scheduling using SVM method in system A, B, C, D consumes MS as 38.1 s, 46.7 s, 51.9 s, 59.9 s. and EC as 21.7 s, 31.9 s, 45.9 s and 56.1 s respectively. The NN method consumes MS as 31.9 s, 43.1 s, 49.8 s, 55.8 s and EC as 20.1 s, 29.7 s, 41.8 s, 52.6 s respectively. Followed by CNN method in system A were 21.2 s as MS and 15.8 s as EC, in system B 28.9 s as MS and 17.9 s as EC, in system C 35.9 s as MS and 22.7 s as EC and in system D 33.9 s as MS and 26.7 s as EC.

In Amazon Lightsail cloud services, the proposed workflow scheduling using SVM method in system A, B, C, D consumes MS as 44.7 s, 51.9 s, 57.5 s, 61.7 s and EC as 32.9 s, 46.7 s, 55.9 s, 72.9 s. The NN method in consumes MS as 37.1 s, 47.5 s, 51.9 s, 55.9 s and EC as 29.7 s, 39.1 s, 44.7 s, 65.8 s respectively.

The workflow scheduling using CNN method in system A achieved 31.9 s as MS and 19.6 s as EC, in system B 38.4 s as MS and 24.5 s as EC, in system C 42.8 s as MS and 29.9 s as EC and in system D 45.7 s as MS and 32.9 s as EC.

Table 3 is the analysis of workflow scheduling approaches using CNN architecture with respect to various task environmental systems on different clouds. The consumed values of MS and EC are linearly increased with respect to the number of task functions increasing in cloud environment. The implementation of the CNN architecture for workflow scheduling in cloud environment reduces the MS and EC for large number of task functions.

Table 3
Analysis of workflow scheduling approach using CNN architecture with respect to various task environmental systems on different clouds

Task Environmental systems	Amazon EC2 Cloud Environment		Amazon Lightsail Cloud	
	MS (s)	EC (s)	MS (s)	EC (s)
System A	21.2	15.8	31.9	19.6
System B	28.9	17.9	38.4	24.5
System C	35.9	22.7	42.8	29.9
System D	33.9	26.7	45.7	32.9

The Amazon EC2 and Lightsail environment consist of the average MS for SVM based method consumes 49.1 s, 53.9 s and the average EC as 38.9 s, 52.1 s. The MS for NN consumes 45.1 s, 48.1 s and the EC as 36.0 s, 44.8 s. The MS for CNN consumes 29.9 s, 39.7 s and EC as 20.7 s, 26.7 s as shown in Table 4.

Table 4
Analysis of workflow scheduling approach with respect to different classifiers on different cloud services

Methodology	Amazon EC2 Cloud Environment		Amazon Lightsail cloud	
	MS (s)	EC (s)	MS (s)	EC (s)
SVM	49.1	38.9	53.9	52.1
NN	45.1	36.0	48.1	44.8
CNN	29.9	20.7	39.7	26.7

The comparisons of CNN based workflow scheduling approach with other existing techniques on Amazon EC2 cloud service in System D task functions is shown in Table 5.

Table 5
Comparisons of CNN based workflow scheduling approach with other existing techniques on Amazon EC2 cloud service in System D task functions

Methodology	Cloud service	Algorithm	MS (s)	EC (s)
Proposed method	Amazon EC2	Modified AlexNet CNN	33.9	26.7
Aida A. Nasr *et al.* (2019)	Amazon EC2	ACO	46.9	41.5
Ahmad M. Manasrah *et al.* (2018)	Amazon EC2	Hybrid GA-PSO	52.1	49.9

Figure 5 is the graphic plot of the comparisons methods of CNN architecture with conventional techniques on Amazon EC2 cloud environment.

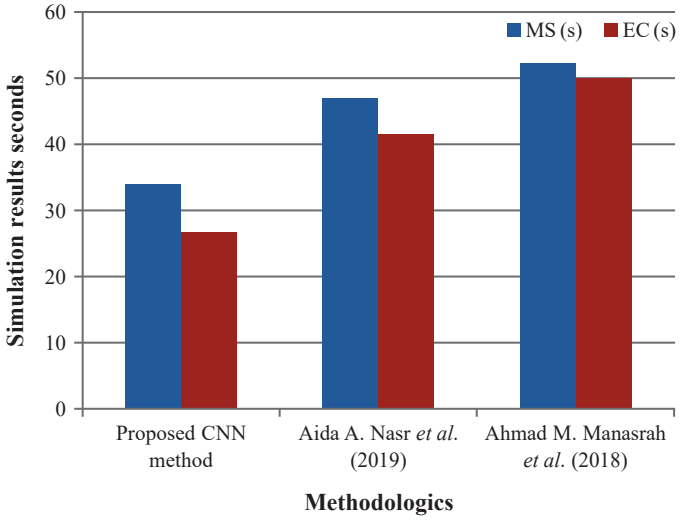

Fig. 5. Graphical plot of the comparisons methods on Amazon EC2 cloud environment

5. CONCLUSION

The hybrid classification approach which combines ML and DL algorithm has been implemented. The workflow computations are performed using different ML algorithms in user environment. The modified AlexNet CNN architecture for workflow scheduling on different cloud services Amazon EC2 and Amazon Lightsail is performed. The performance of workflow computing and scheduling is analyzed using MS and EC on different cloud services and segmentation is made in the optimized task to detect the cancer region in the classified abnormal images. The average MS and EC for the system using modified CNN architecture consumes 29.9 s and 20.7 s on Amazon EC2 cloud services. The average MS and EC for the proposed system consumes 39.7 s and 26.7s on Amazon Lightsail cloud services.

The abbreviations used in this research work are given in the following table:

Terms	Abbreviations
Adaptive Genetic Algorithm	AGA
Adaptive Histogram Equalization	AHE
Adaptive Neuro Fuzzy Inference System	ANFIS
Ant Colony Optimization	ACO
Chemical Reaction Optimization	CRO
Cloud Service Providers	CSP
Co-Active Adaptive Neuro Fuzzy Inference System	CANFIS
Computer Tomography	CT
Convoultional Filter	CF

Terms	Abbreviations
Convolutional Kernel	CK
Convolutional Layers	CL
Convolutional Neural Networks	CNN
Deep Learning	DL
Deep Neural Network	DNN
Execution Cost	EC
Elastic Compute Cloud	EC2
Feature Index Matrix	FIM
Fully Connected Neural Networks	FCNN
Genetic Algorithm	GA
Load Balancing Genetic Algorithm	LBGA
Local Binary Pattern	LBP
Local Ternary Pattern	LTP
Magnetic Resonance Imaging	MRI
Makespan	MS
Machine Learning	ML
Neural Network	NN
Particle Swarm Optimization	PSO
Pooling Layers	PL
Quality of Service	QoS
Rectifier Linear Unit	ReLU
Reference Queue based Cloud Service Architecture	RQCSA
Remote Cloud Server	RCS
Support Vector Machine	SVM
Virtual Machines	VM
Visual Geometry Group	VGG

REFERENCES

[1] A.M. Manasrah and H. Ba Ali, "Workflow scheduling using hybrid GA-PSO algorithm in cloud computing," *Wireless Commun. Mob. Comput.*, vol. 15, no. 3, pp. 1–16, 2018, doi: 10.1155/2018/1934784.

[2] S.G. Ahmad, C.S. Liew, E.U. Munir, A.T. Fong, and S.U. Khan, "A hybrid genetic algorithm for optimization of scheduling workflow applications in heterogeneous computing systems," *J. Parallel Distrib. Comput.*, vol. 87, no. 2, pp. 80–90, 2016, doi: 10.1016/j.jpdc.2015.10.001.

[3] H. Alaskar, A. Hussain, N. Al-Aseem, P. Liatsis, and D. Al-Jumeily, "Application of convolutional neural networks for automated ulcer detection in wireless capsule endoscopy images," *Sensors.*, vol. 19, no. 6, pp. 1265–1281, 2019, doi: 10.3390/s19061265.

[4] L. Teylo, L. Arantes, P. Sens, and L.M.A. Drummond, "A dynamic task scheduler tolerant to multiple hibernations in cloud environments," *J. Cluster Comput.*, vol. 21, no. 5, pp. 1–23, 2020, doi: 10.1007/s10586-020-03175-2.

[5] M. Sardaraz and M. Tahir, "A parallel multi-objective genetic algorithm for scheduling scientific workflows in cloud computing," *Int. J. Distri. Sensor Net.*, vol. 16, no. 8, pp. 1–10, 2020, doi: 10.1177/1550147720949142.

[6] M. Hosseinzadeh, M.Y. Ghafour, and H.K. Hama, "Multi-objective task and workflow scheduling approaches in cloud computing: a comprehensive review," *J. Grid Comput.*, vol. 18, no. 3, pp. 327–356, 2020, doi: 10.1007/s10723-020-09533-z.

[7] C.L. Chen, M.L. Chiang, and C.B. Lin, "The high performance of a task scheduling algorithm using reference queues for cloud-computing data centers," *Electronics*, . vol. 9, no. 1, pp. 371–379, 2020, doi: 10.3390/electronics9020371.

[8] Y. Hu, H. Wang, and W. Ma, "Intelligent Cloud workflow management and scheduling method for big data applications," *J. Cloud Comp*, vol. 9, no. 1, pp. 1–13, 2020, doi: 10.1186/s13677-020-00177-8.

[9] M. Grochowski, A. Kwasigroch, and A. Mikołajczyk," Selected technical issues of deep neural networks for image classification purposes," *Bull. Polish Acad. Sci. Tech. Sci.*, vol. 67, no. 2, pp. 363–376, 2019, doi: 10.24425/bpas.2019.128485.

[10] A.A. Nasr, N.A. El-Bahnasawy, and G. Attiya, "Cost-effective algorithm for workflow scheduling in cloud computing under deadline constraint," *Arab. J. Sci. Eng.*, vol. 44, no. 4, pp. 3765–3780, 2019, doi: 10.1007/s13369-018-3664-6.

[11] Z. Swiderska-Chadaj, T. Markiewicz, J. Gallego, G. Bueno, B. Grala, and M. Lorent, "Deep learning for damaged tissue detection and segmentation in Ki-67 brain tumor specimens based on the U-net model," *Bull. Polish Acad. Sci. Tech. Sci.*, vol. 66, no. 6, pp. 849–856, 2018, doi: 10.24425/bpas.2018.125932.

[12] Y. Cui and Z. Xiaoqing, "Workflow tasks scheduling optimization based on genetic algorithm in Clouds," in *IEEE 3rd Int. Conf. on Cloud Computing and Big Data Analysis (ICCCBDA)*, *Chengdu*, 2018, pp. 6–10, doi: 10.1109/ICCCBDA.2018.8386458.

[13] T. Wang, Z. Liu, Y. Chen, Y. Xu, and X. Dai, "Load balancing task scheduling based on Genetic algorithm in Cloud Computing," in *IEEE 12th Int. Conf. on Dependable*, *Autonomic and Secure Computing*, Dalian, 2014, pp. 146–152, doi: 10.1109/DASC.2014.35.

[14] X. Zhao, Y. Wu, G. Song, Z. Li, Y. Zhang, and Y. Fan, "A deep learning model integrating FCNNs and CRFs for brain tumor segmentation," *Med. Imag. Anal.*, vol. 43, no. 4, pp. 98–111, 2018, doi: 10.1016/j.media.2017.10.002.

[15] C. Liu, R. Zhao, W. Xie, and M. Pang, "Pathological lung segmentation based on random forest combined with deep model and multi-scale superpixels," *Neu. Proc*, *Let.*, vol. 52, no. 2, pp. 1631–1649, 2020, doi: 10.1007/s11063-020-10330-8.

[16] C. Zhang, J. Zhao, Ji. Niu, and D. Li, "New convolutional neural network model for screening and diagnosis of mammograms," *PLoS One.*, vol. 15, no. 8, pp. 67–80, 2020, doi: 10.1371/journal.pone.0237674.

[17] L.D. Nicanor, H.R. Orozco Aguirre, and V.M. Landassuri Moreno, "An assessment model to establish the use of services resources in a cloud computing scenario," *J. High Perf. Vis. Int.*, vol. 12, no. 5, pp. 83–100, 2020, doi: 10.1007/978-981-15-6844-2_7.

[18] V. Magudeeswaran and J. Fenshia Singh, "Contrast limited fuzzy adaptive histogram equalization for enhancement of brain images," *Int. J. Imag. Sys. and Tech.*, vol. 27, no. 1, pp. 98–103, 2017, doi: 10.1002/ima.22214.

[19] S.P. Cleary and J.S. Prell, "Liberating native mass spectrometry from dependence on volatile salt buffers by use of Gábor transform," *Int. J. Imag. Syst. Tech.*, vol. 20, no. 4, pp. 519–523, 2019, doi: 10.1002/cphc.201900022.

[20] V. Srivastava, R.K. Purwar, and A. Jain, "A dynamic threshold-based local mesh ternary pattern technique for biomedical image retrieval," *Int. J. Imag. Sysy. Tech.*, vol. 29, no. 2, pp. 168–179, 2019, doi: 10.1002/ima.22296.

[21] J.H. Johnpeter and T. Ponnuchamy, "Computer aided automated detection and classification of brain tumors using CANFIS classification method," *Int. J. Imag. Sysy. Tech.*, vol. 29, no. 4, pp. 431–438, 2019, doi: 10.1002/ima.22318.

[22] N. Kumar and D. Kumar, "Classification using artificial neural network optimized with bat algorithm", *Int. J. Innovative Tech. Exploring Eng. (IJITEE)*, vol. 9, no. 3, pp. 696–700, 2020, doi: 10.35940/ijitee.C8378.019320.

[23] A.S. Mahboob and S.H. Zahiri, "Automatic and heuristic complete design for ANFIS classifier, network: computation in neural systems," *J. Net. Comput. Neu. Syst*., vol. 30, no. 4, pp. 31–57, 2019, doi: 10.1080/0954898X.2019.1637953.

[24] C. Shorten and T.M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big. Data.*, vol. 60, no. 6, pp. 1–48, 2019, doi: 10.1186/s40537-019-0197-0.