

Intelligent Resource Allocation and Capacity Computation through RaI Representation in the Cloud using Deep Learning Techniques

^{1*}Santhoshini Banda, ²A.Sampath Dakshina Murthy, ³Bommideni Revathi, ⁴Narasimham Challa

ABSTRACT--There are some situations where cloud computing is used to enhance the ability to the business goals, when and where to offload the resources like hardware, software, networks to cloud. So that one can offload the resources for processing as image based computation includes segmentation, deep learning for object recognition. Intelligent Agent algorithm also uses to collect performance metric in continuous period of time. Dynamic cloud allocation mechanism is implemented in processing of images parallelly. By adopting suitable mechanism one can automatically add images to cloud in real-time to know the number of available cloud instances. Queue length can be known with this. The proposed intelligent cloud resource procedure through RaI (Resources as Images) in the cloud improves overall response time, optimal utilization of cloud in order to access, allot and to determine the capacity of the resources.

Key words--RaI, Intelligent Agent algorithm, Machine learning, Cloud computing.

I. INTRODUCTION

Hardware and software both can be referred by cloud in overall internet. As day by day cloud service providers(CSP) is going to increase then the storage of data in cloud will also increase. So, consumers or customers face many difficulties like price to get cloud instance, migration between CSP's and dynamically changing resource offerings. To assist customers, Intelligent cloud resource allocation(ICRA) through Resources as Images (RaI) and AI agent is introduced for effective sharing of the resources. It discloses total information regarding present status of CSP resources, its offerings and evaluates it to get the best and suitable configuration based on customer need.. It is based on raw computer resources like storage, processing power, databases and other applications. CSP resources can be used by customers at a particular amount of time based on agreement than purchasing those resources completely. Customers can either increase or remove the information in the cloud according to their requirements [1-3]. Many researches are going on, to make use of efficient cloud when there is increasing of CSP. Techniques like Load Balancing aimed to reduce energy cost, dynamic pricing model to maximize revenue. User has to balance less cost with more appropriate quality services. As day by day number of CSP's are increasing, customers choosing CSP based on less cost, Proving the quality in services and position of the model in the market. Furthermore setting prices to the customer at a time, CSP's extend their service by

^{1*} Asst Professor of CSE, Stanley college of Engineering and Technology for Women, Hyderabad, santhoshini.banda@gmail.com

² Asst Professor, Department of ECE, Vignan's Institute of Information Technology, Visakhapatnam, sampathdakshinamurthy@gmail.com

³Asst Professor of CSE, Stanley college of Engineering and Technology for Women, Hyderabad, revathibommideni@stanley.edu.in

⁴ Professor, Department of CSE, Vignan's Institute of Information Technology, Visakhapatnam, drchallan@gmail.com

providing resources dynamically in AWS with according to cost. Spot pricing will be done based on cloud utilization and demand of customer. In addition to this, customer face several challenges of choosing cloud resource with minimum cost by comparing with all available CSP's. Customer must reevaluate configurations from all CSP's because it may change. If customer want to move from one cloud to other cloud with same features but more suitable configuration is called Migration.

II. EXISTING SYSTEM

It includes various techniques like segmentation, deep learning, indexing, decision making algorithm, scheduling management, and dynamic cloud allocation. As this mobile application uses large memory space to store images and videos, it should need large physical storages like ram, chips. To achieve this, certain heavy computation load should be offloaded to cloud for easy processing. With this cost of cloud allocating resources may increase as the number of instances increasing. So, we have to maintain dynamic memory allocation and storage [4]. To process in average time, if number of active user connection increasing then usage time of CPU, memory, and status of queue also increases. The data processing mechanism in the Intelligent cloud is used to give more performance and feasibility. Then it is to perform classification and estimate the number of capabilities of each resource objects. The overall time taken to compute capacities, the second level ingredient testing algorithm technique undergoes two stages in the processing. Overall time taken in computation is T_{mx} .

Negotiation

It is a process by which two or more process come to one agreement which consists of proposal, concessions, etc is accepted by all. Main functionality in this situation to users is to provide cloud computing with less cost and all desired features and QoS. Based on user's preference one ICRA agent will communicate with other agent for mutual agreement. By this agents can improve their communication socially and their response can be improved [5]. **Agreement at service level** is a process of understanding between parties and accepting conditions and terms is referend as SLA (service level agreements). It consist of company name under agreement and their quality of service. WSAG(web service agreement negotiation) describes steps for SLA negotiation.

Negotiation contains template, offers and agreements. series for description of the services they are offering configurations, guarantee they provide in QoS soon. Based on requirements of user, service providers sends template to customer. An instantaneous template is called offer. It is chosen when a specific configuration is selected by customer with agreement. An Agreement is created based on the acceptance from cloud providers and customers. An Counter-offer is generated new configuration when the offer is rejected.

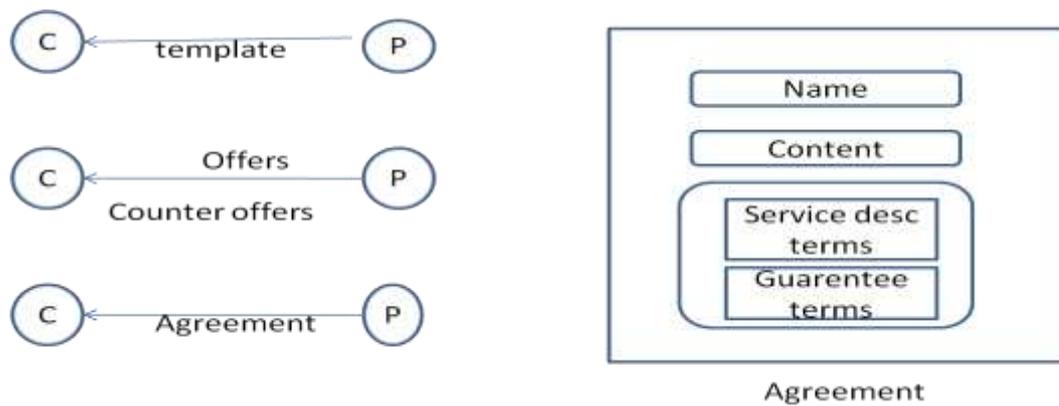


Figure 1: SLA template

Micro agreements consists of terms and conditions for particular period of time. Customers may sign agreement for long period of time like one year or more. In the case, customer should not migrate to other CSP. They should stay until their completion of time period. This CSP have reliable income for a period and accuracy may be improved. But the main disadvantage is, customers should not react to any changes in the market even new product releases and cost may decrease. In order to overcome this, micro-agreements are introduced to maintain between one hour or some minutes. By this customer may accept agreement in hourly basis and they can react to release of new product in the market with high efficiency.

III. PROPOSED SYSTEM

To assist customers, ICRA through RaI provides all available cloud resources, choosing appropriate configuration based on requirement, negotiating service agreement with CSP, monitoring service agreement for violation and assisting migrations between CSP's. ICRA describes various information and status of all available CSP's like price tag of CSP in the market, accessibility, guarantee, locality and reputation. User gives his requirements by showing their configuration then, ICRA checks and matches the appropriate and best cloud instance from CSP's by removing agreements at service level. It provides services lifetime to ensure no agreement violations. If violation occurs, severe action will be taken. ICRA frequently monitors availability of better resource or new release of better configuration and gives notifications to users [6-9].

Initially consumer/customer who want CSP will directly contact ICRA with all requirements written in SLA offer to specify software or hardware requirements, priorities, range of options, dependencies etc. Let us take example that consumer has requirement of SaaS. With SLA, one can specify CPU, Version and storage should be of strict specific range. So, cost may be reduced for getting SaaS. If any changes in requirements occur, then consumer has to update the specifications through ICRA agent. Changes may be of availability of new resource servers in the market or any requirements change in their plan or computing needs may increase. To enable those changes, consumers monitor in the cloud and inform ICRA agent [10].

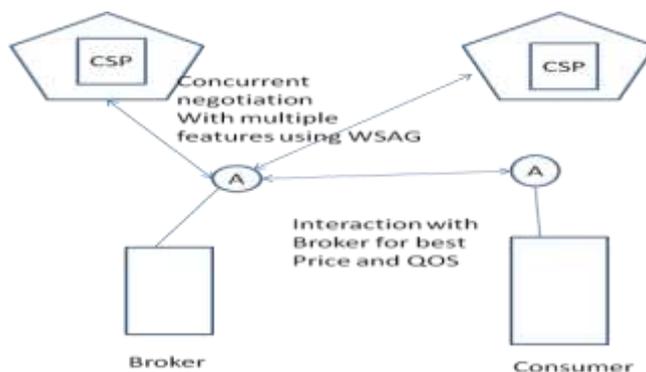


Figure 2: Overview working and communication between ICRA agent(A) and CSP

CSP acts as an interface to interact with ICRA for migration and negotiation. For negotiating, CSP must provide templates to give information. It should be updated regularly for availability of pricing. Based on request, SLA template is sent to ICRA agent. It performs a negotiation session for approving policy. This policy contains threshold value and rules for acceptance. In transfer of virtual data into cloud which is called as migration should be assisted by cloud service provider(CSP) after creating a SLA with consumer. **ICRA agent** is independent and trusted third party between consumer and SLA. It provides information about CSPs offerings in the market, test these offers, on behalf of customers agents can negotiate between SLA and CSP, checks whether any new cloud resources available and also supports migrations to new clouds. There are two approaches to identify resources when user captures the image. Those are SVM, which classifies parallelly and Deep learning techniques to achieve higher accuracy [11-15].

For computing the accessed value and capacity of resource, we propose the customer's touch on the resource as reference to the Calibration. When User touches the resource, a picture is taken. Resource recognition and capacity computation are essential parts. The Map Reduce algorithm which classifies parallelly in cloud is used to execute SVM in AWS. Decision making mechanism is used for scalability aspect. When more number of customers using the application at same time, segmentation is done initially to identify the resource items.

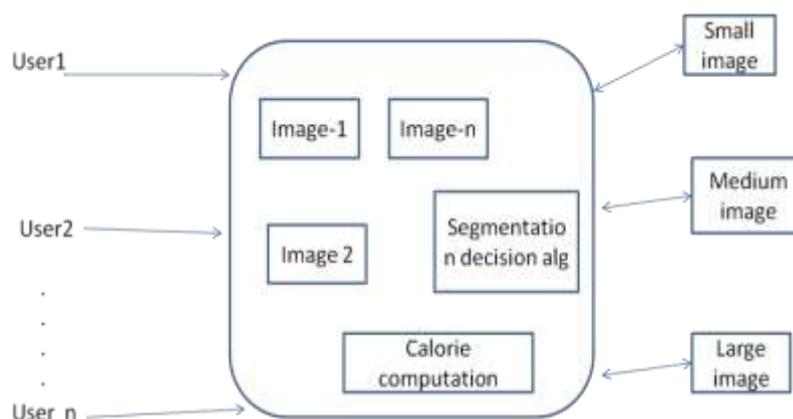


Figure 3: Intelligent Decision algorithm

But we can see two weaknesses in the Decision Making Algorithms. It checks the performance, further optimizes efficiently and allocates the resource images to remaining cloud. It should be extended at every division

of resources (single,multiple,hybrid). Single resource items performs two steps i.e Deep learning(drawing out feature, categorization) and computation. Multiple resource items performs initially segmentation to smaller objects and then followed by deep learning on each smaller part. Finally after integration, it gives the calories value. Hybrid items has extra step like recognizing the resource with deep learning techniques [16-19].

IV. DESIGN METHODOLOGY

There are two techniques for offloading schemas. First is self reliant multi cloud offloading provides balancing in high communication cost and second one is multi cloud offloading which reduces communication cost with smaller stability.

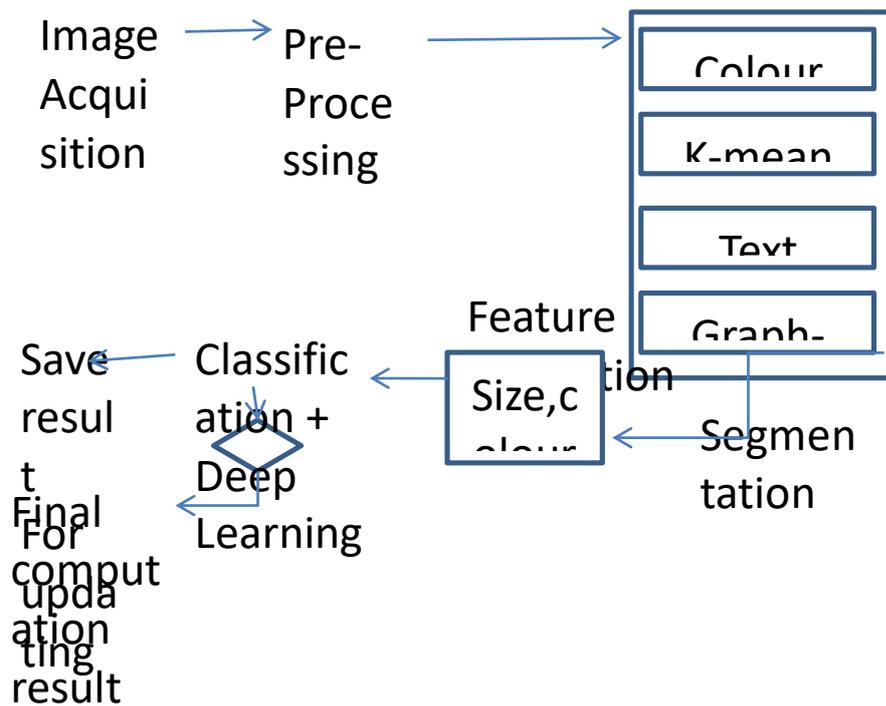


Figure 4: Overall processing of resources marked as images

Single and multiple resource identity can be done following way. Initially user clicks images and system captures it. Preprocessing will be done. Then segmentation is performed based on features for extracting various resource partitions and save them as single resource image. Image classification is applied to analyze what type of resource it is. Then, it contains with the user by showing the ingredients for confirmation of all detected resources. Users have to observe and perform modifications if only detection is wrong. Once, user finalizes it, then it goes to next step i.e detection of calories among each resource and shows the result on screen.

Hybrid identity performs two stages of recognition testing. In the initial step whole picture of hybrid resource is treated as one, then it is tested. In the second level ingredient of that particular resource will be identified.

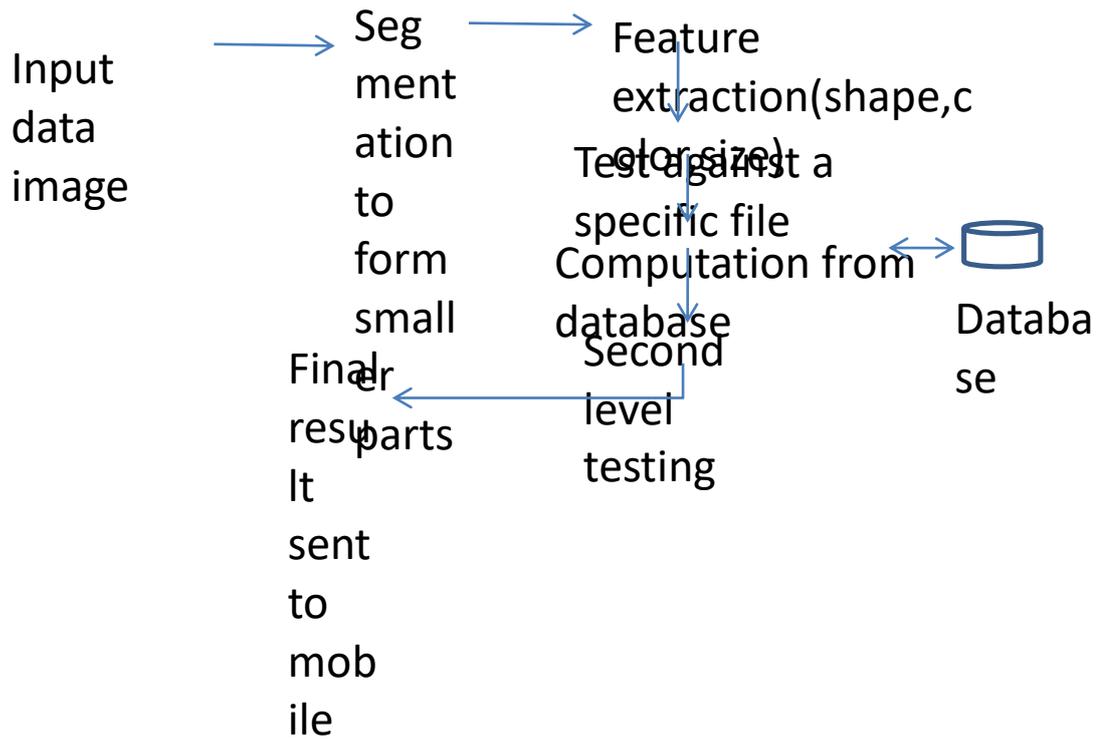


Figure 5: Process of computation

DNN (Deep Neural Network) is one the algorithm for feature extraction and classification to detect the objects more accurately even with small datasets. Further provides flexibility to combine features. Initially, DNN performs unsupervised learning without labelling of images. In order to get correct and error-free outcome, DNN analyzes the vertex/nodes which are unseen and edge parameters which is under Supervised Learning. To calculate the grade(i.e inclination or decination) of cost function rapidly , we use the Back propagation algorithm which is a one of the supervised learning. Modifications can be made by customer accordingly either to wait and bias, to get the proper outcome by training the deep neural network. Labelling these training inputs as $x_1, x_2, x_3, \dots, x_n$. Stochastic gradient descent algorithm chooses minimum batch of inputs and train those weights and bias are modifies further for accurate result. DNN is arranged with many layers that are hidid in step by step manner. RBM(restricted boltzmann machine) can be detected by making together of two layers that hidid which forms a stack of RBM further. A combination of neuron can be formed from hidid layer. I^k is input vector, h^k is hidden neuron and O^k is ouput neuron of k layer. Upper features can get from below features. Forecasting of layer is based on probablistic model.

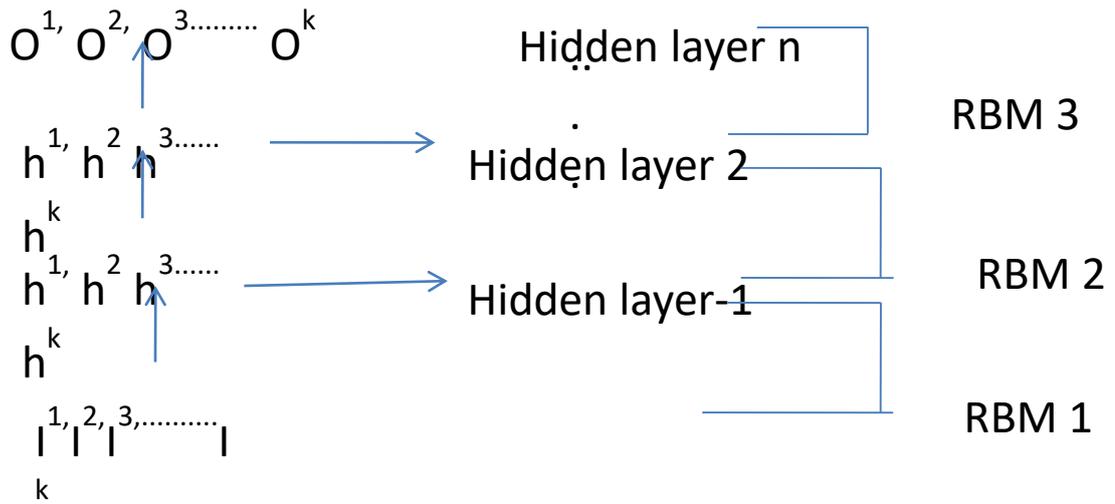
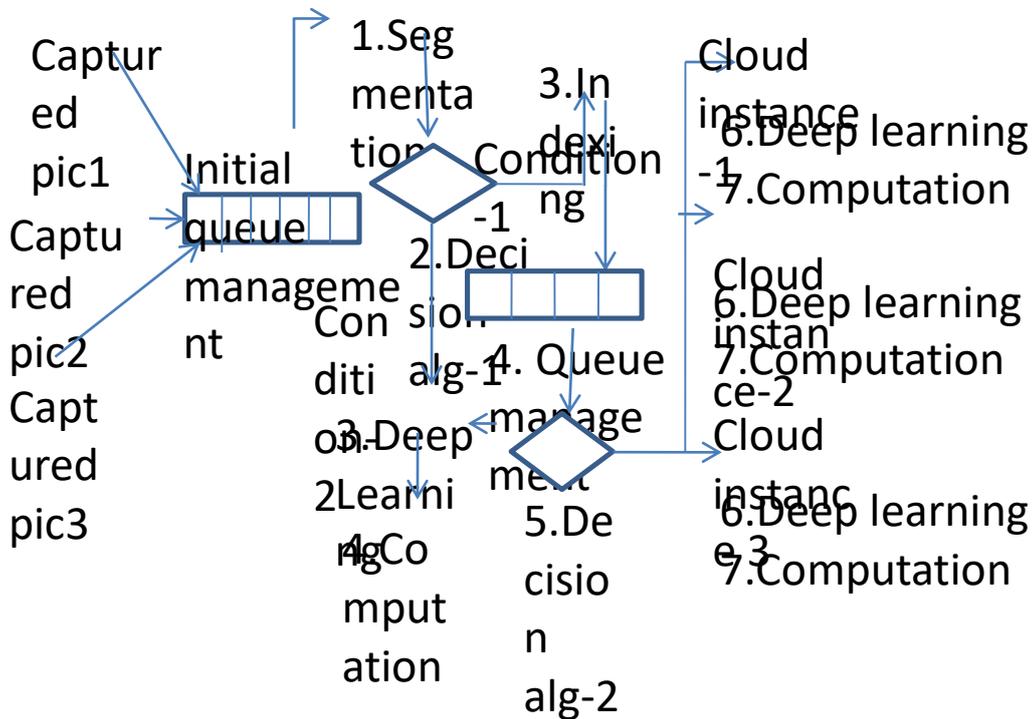


Figure 6: Deep Neural Network working

Capacity computation is performed with Auto- Calibration approach. It is used to find the distance between user and resource at how much distance user has captured the pic of object. This approach will measure or calculate the dimension of the resource taken from database. Smaller particles features can be calculated from dimensions of the block.



Condition-1: sends to other cloud instance, Condition-2: sends to same cloud instance

Figure 7: Proposed Cloud Agent Algorithm

Procedure for single and mutiple computation:

Proposed Cloud Agent Captured images are sent to first queue. Then it is segmented to smaller parts based on single or multiple images based on that it will process further. If it is single image, then it will be under deep learning or otherwise it is sent to second queue within a certain limit amount of time because new images should be loaded in first queue. If it is single, decision algorithm-1 is executed else it goes to algorithm-2. Based on the current availability of customer participation and characteristics of cloud agent (average time consumption, usage time of CPU and memory). When the decision finalizes, remaining procedure like Deep learning and capacity computation on cloud agent can be done. It it decides to perform functionality on remote clouds, then it will be indexed to maintain consistency. After indexing, it sends to queue and prioritize them for offloading image sequentially to other cloud servers, then after capacity computation is done.

Decision Algorithm-1:

This algorithm says about the result and presentation of local cloud, to check whether usage of cloud is in committed threshold level. Initially image in the cloud process locally to queue, then later goes to the instance of cloud. Space in the disk can be verified based on cpu usage of cloud is in lesser threshold value. If usage of disk is not in range of threshold value, then rejection can be done to accept further images and removes existing images for disk space. The reason for setting threshold is for evaluating accurately by not effecting the length of queue. If threshold value is greater, then it resends the image. If it is lesser than threshold value, it may proceeds for further computation steps.

Decision algorithm-2:

It is responsible for deciding the cloud node among the list of nodes that takes least response

CPU utilization of node $n = CU_n$ Memory utilization = MU_n Estimated time = T_n For image i_k from $i_1 i_2 i_3 \dots i_p$
do $CU_m =$ check CU for node m If $CU_m <$ threshold then $MU_m =$ check If $MU_m <$ threshold Dispatch(i_k) to node m; break; initiate Deep Learning (i_k, m); else check estimated time T_m for CU_m for each node x do $CU_x =$ check CU Check estimated time T_x for CU_x $MU_x =$ check MU Check estimated time T_y If $T_x > T_y$ $T_{cf} = T_y$; Else $T_{cf} = T_x$; If $T_m > T_{cf}$ then Dispatch (i_k) to node m; Initiate DL(i_k, m); Break; Else Dispatch (i_k) to node x; Initiate DL(i_k, y); Break; End for; Break End for

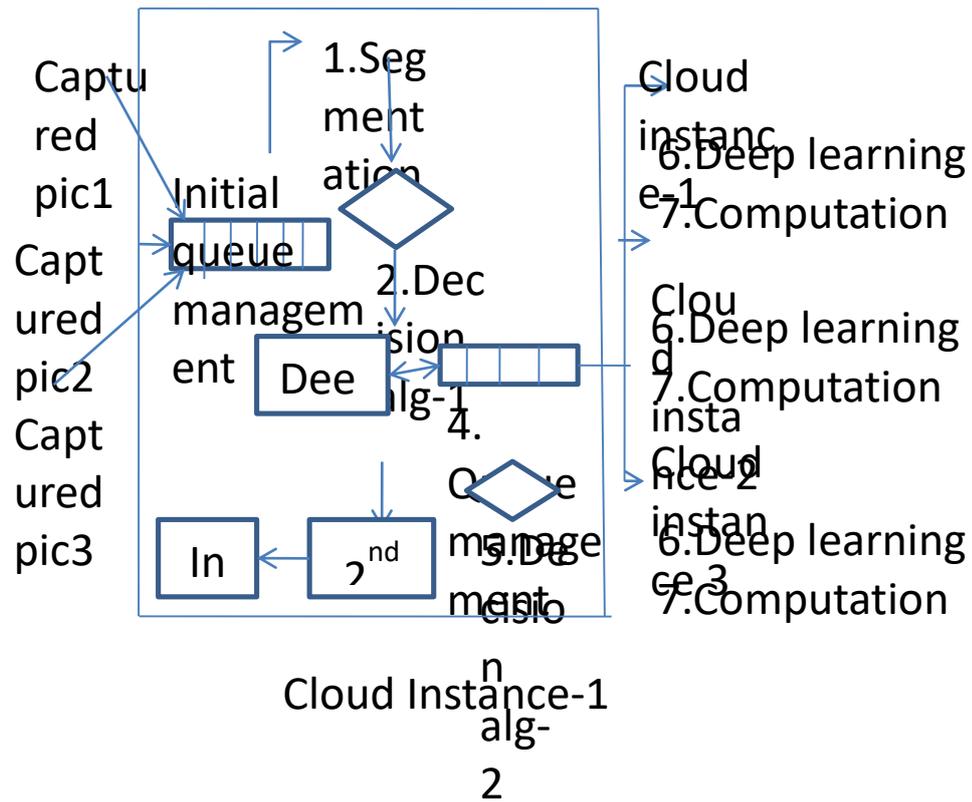


Figure 8: Architecture of Hybrid Processing

In indexing, Initially images are sent to queue and it is segmented in level-1. According to Decision algorithm-1, it finalizes to offload images to other cloud for second level testing. During this second level testing, remaining images start first level and repeats further. Then segmented images are arranged in sorted order with sequence number. For multiple images, integration of capacity estimation will be occurred after segmenting and checking the value separately. In hybrid image processing, indexing is done by level-1 and level-2.

V. CONCLUSION

The paper described about the offloading mechanism of the resources like hardware, software, networks in to cloud. The offloaded resources subjected for processing as image based computation includes segmentation, deep learning. Cloud Intelligent Agent algorithm has been introduced to measure the performance metric in continuous period of time. Dynamic cloud allocation mechanism also implemented for parallel processing of this paper depicted that the suitable mechanism to add images in to cloud in real-time to know number of available cloud instances. Further, narrated the intelligent cloud resource procedure through RaI (Resources as Images) in the cloud for single, multiple and hybrid categories of ingredients of resources to enable to compute the capacity.

REFERENCES

1. Satyanarayanan, Mahadev, Paramvir Bahl, Ramón Caceres, and Nigel Davies. "The case for vm-based cloudlets in mobile computing." *IEEE pervasive Computing* 8, no. 4 (2009): 14-23.

2. Hinton, Geoffrey E., Simon Osindero, and Yee-Whye Teh. "A fast learning algorithm for deep belief nets." *Neural computation* 18, no. 7 (2006): 1527-1554.
3. Hassan, Mohammad Mehedi, Biao Song, Ahmad Almogren, M. Shamim Hossain, Atif Alamri, Mohammed Alnuem, Muhammad Mostafa Monowar, and M. Anwar Hossain. "Efficient Virtual Machine Resource Management for Media Cloud Computing." *KSII Transactions on Internet & Information Systems* 8, no. 5 (2014).
4. M. Shamim Hossain and G. Muhammad, cloud – Assisted Industrial IOT – enabled Framework for health monitoring, Elsevier computer networks, 2016
5. Y. Bengio, Learning deep architectures for AI, Foundat. Trends in Mach. Learn., vol. 2, no. 1, pp. 1-127, 2009
6. Jennings, N. And Wooldridge, M., editors (1998). Applications of Intelligent Agents, chapter 1, pages 3-28. *Agent Technology: Foundations, Applications and Markets*. Springer
7. Anandasivam, A and Prem M (2009). Bid price control and dynamic pricing in clouds. In proceedings of the European Conference on Information Systems, Pages 1-14.
8. Armbrust, M Fox, A., Griffith, R Joseph, A Katz, R Konwinski, A Lee, Patterson, D Rabkin, A stoica, I., et al. (2010). A view of cloud computing. *Communications of the ACM*, 53(4):50-58.
9. Sim, K. (2010). Towards complex negotiation for cloud economy. *Advances in Grid and Pervasive computing*, Pages 395-406
10. Weiss A (2007). Computing in the clouds. *netWorker*, 11(4):16-25.
11. A.S Prasad and S>Rao, " A Mechanism Design Approach to Resource Procurement in cloud computing." *IEEE Trans. Computers*, vol. 63, no. 1, 2014, pp. 17-30
12. Y. LeCun, Y. Bengio and G. Hinton, "Deep Learning," *Nature*, vol. 521, no. 7533, 2015, pp.436-444
13. W.Wang, B. Li and B. Liang, "Dominant Resource Fairness in cloud computing systems with heterogeneous Servers," *Proc. 33rd International conference on computer communications*, 2014
14. V. Mnih et al., "Asynchronous methods for deep reinforcement learning," *Proc, 33rd International conference on Machine Learning*, 2016.
15. R.H. Hwang et al., "cost optimization of Elasticity cloud resource subscription ploicy," *IEEE Trans. Services Computing*, vol. 7, no. 4, 2014, pp.561-574
16. Dr. B. Sankara babu, A. Sampath Dakshina Murthy, Sampenga Veerraju, B. Omkar Lakshmi Jagan , K. Saikumar "Implementation of Real and Accurate Watermarking System For Security Using Logistic Regression Machine Learning Techniques", *The Journal of Research on the Lepidoptera*, Volume 51 (1): 783-792, March 2020.
17. A. Sampath Dakshina Murthy, P. Satyanarayana Murthy, V. Rajesh, Sk. Hasane Ahammad, B. Omkar Lakshmi Jagan, "Execution of Natural Random Forest Machine Learning Techniques on Multi Spectral Image Compression", *International Journal of Pharmaceutical Research* Volume 11, Issue 4, Oct - Dec, 2019.
18. K.Raju, S.Kiran Pilli, G. Siva Suresh Kumar, K. Saikumar, B. Omkar Lakshmi Jagan, "Implementation of Natural Random Forest Machine Learning Methods on Multi Spectral Image Compression", *Journal of Critical Review*, Volume 6, Issue 5, pg. 265-273, 2019.

19. Ravada Aamani, Adinarayana Vannala, A. Sampath Dakshina Murthy, K. Saikumar, B. Omkar Lakshmi Jagan, "Heart Disease Diagnosis Process using MRI Segmentation And Lasso Net Classification ML", Journal of Critical Review, Volume 7, Issue 6, pg. 717-721, 2020.