© 2019 IJSRSET | Volume 6 | Issue 2 | Print ISSN: 2395-1990 | Online ISSN : 2394-4099
Themed Section : Engineering and Technology
DOI : https://doi.org/10.32628/IJSRSET

Review on Cloud-Based Image Classification

K. Sunil Manohar Reddy*[1], Prof. BK Tripathi[2], Dr. S. K. Tyagi[3]
*[1]Assistant Professor, Department of CSE, MECS, Hyderabad, Telangana, India
[2]Professor, Department of CSE, HBTU, Kanpur, Uttar Pradesh, India
[3]Professor, Department of CSE, CCSU, Meerut, Uttar Pradesh, India

ABSTRACT

Automated quantitative image analysis is important for all fields of bioscience research. Though many software programs and algorithms have been developed for bio image processing, a sophisticated knowledge of image processing methods and high-performance computing resources are needed to use them. So, a cloud-based image analysis platform was developed, known as IMACEL that includes morphological analysis and machine learning-based image classification. The distinctive click-based UI of IMACEL's morphological analysis framework allows researchers with lesser resources to evaluate particles quickly and quantitatively without previous knowledge of image processing. As all image processing and machine learning algorithms are executed on high-performance virtual machines, users can access a similar analytical environment from anywhere. A validation review of the morphological analysis and image classification of IMACEL was done. The output indicates that this platform is an accessible and potentially powerful tool for the quantitative evaluation of bioimages,which will reduce the barriers to life science research.

Keywords: CNN, Cloud Based Image Analysis, Morphological Analysis, IMACEL

I. INTRODUCTION

Latest improvements in microscopic and image processing methods have led to new findings in the life sciences. With the emergence of imaging devices like microscopes, MRI, and CT, image data in the life sciences are increasingly detailed. Specifically, the introduction of visualization techniques like the usage of fluorescence microscopy and fluorescent probes, aid the analysis of biological structures and transform molecular imaging. Hence, it's becoming critical to analyse these bio image data efficiently and rapidly in quantitative studies [1,2]. Usually, the evaluation of large and detailed images is very tedious and laborious, and is a burden for researchers. In addition to new trends in imaging devices, many of open source and commercial image analysis software (e.g., ImageJ [3], ImagePro, and Photoshop) and libraries for programming languages (e.g., OpenCV and Bioconductor) have come into play; however, their usage needs specialist knowledge.

Machine learning is also used to analyse huge quantities of bioimage data. Using this method, it is feasible to automate or semi-automate analysis for the target extraction and classification of diverse and humongous numbers of biological images [4,5]. Deep learning-based CNNs are expected to be helpful for single-cell experiments with high-throughput and high-content screening [6,7]. A report on using nonlinear dimensionality reduction together with deep learning to reconstruct cell cycle and disease progression has illustrated the efficiency of applying machine learning techniques to objective biological prediction [8]. For example, we proposed a system earlier that joins machine learning and active learning [9] for sub cellular localization, mitotic phase classification, and the discrimination of apoptosis in images of plant and human cells. This system

achieved an accuracy level greater than or equal to that of the annotators [10]..

II. CLOUD BASED IMAGE ANALYSIS

Even though advanced image processing and machine learning methods are needed in life science studies, several research labs are ill-equipped to do the bioimage analysis that uses advanced imaging techniques and many computing resources. For generic morphological analysis, like counting a number, measuring an area, and extracting various features of a shape, researchers require information about signal/background setting, noise reduction filtering, binaiysation setting, and particle analyzer function in de facto-standard image processing software ImageJ, and should manually select specific algorithms for each research purpose and tune the parameters manually. For classification analysis, all software and analytical environments need skills for programming languages to input commands. So, even image processing plays vital role in quantitative data research for life sciences, present image processing solutions are very sophisticated for many researchers to use. Hence, user-friendly software for image analysis is required to expand the utilization of imaging techniques throughout the life sciences.

IMACEL is a cloud-based image analysis framework developed for automatic classification and morphological analysis. As all image processing and machine learning techniques are performed by virtual machines in the cloud, it's not needed to set up powerful computers in laboratory. IMACEL's target data includes different types of microscopic bioimages. The most vital feature in IMACEL is that the new UI for researchers with less knowledge of image processing. IMACEL suggests multiple candidates for morphological analysis, permits users to choose the most efficiently processed images (S1 Movie). This allows users to know appropriate procedures quickly and easily. In addition to morphological analysis, IMACEL can do automatic image classification from uploaded annotated images using random forests and a deep learning algorithm.

The contributions of this study are as below:

- A tool that enables life science researchers is presented with limited image processing experience and computing resources to automatically and quantitatively analyse microscopic image data.
- The morphological analysis of the system is verified by checking the number and size of stress granules in images using the batch process function. Moreover, we examine the classification analysis of cell cycle progression using machine learning methods on the IMACEL platform.

The adoption of IMACEL in life science analysis has the ability to enhance the quality and quantity of research, especially for researchers who would not otherwise have the experience and resources to do such investigations

III. METHODS

Implementation and architecture of the IMACEL platform

IMACEL is a cloud-based image processing platform which works on Windows, Mac OS X, and Linux. The image processing core modules of IMACEL were developed using Python 3 and OpenCV, and computation is done on a virtual machine using the Microsoft Azure service.

A virtual machine with the standard D2 v2 instance type (2 vCPU, 7 GB RAM) was utilized in this study. Azure Storage was used as image storage server. To connect to the storage server from a web application server, the Azure Storage SDK for Python was used. The database and web server used URLs for their connections to the storage server.

Security of the IMACEL platform

IMACEL used SSL/TLS to create a secure connection between the web browser, web server, application server, and storage server. To permit restricted access to resources in the storage server, a shared access signatures (SAS) provided by Azure Storage was used.

Cloud-based image processing

In order to use IMACEL, researchers upload images to the web server with a browser, and the images are processed by high-performance virtual machines running on the Microsoft Azure platform that are able to communicate with the system's database. Processed image data are reverted back to the researchers through the browser. The maximum data size for uploading images depends on the sort of browser. For example, Internet Explorer 11 has a limitation of 4 GB for file uploading.

IMACEL Interface with a click-based UI

The IMACEL platform includes a novel click-based interface designed for researchers who does not have advanced image processing knowledge. Researchers can upload images to IMACEL, mentioning the imaging method (e.g., fluorescence, or electron microscopy) and imaging target (e.g., bacteria, yeast, or brain tissue) to enable the IMACEL particle analyser to provide practical suggestions. In image processing, users click on the most suitable processed image shown in the browser. This clickable UI permits researchers of all skill levels to extract particles quantitatively and objectively from raw input images.

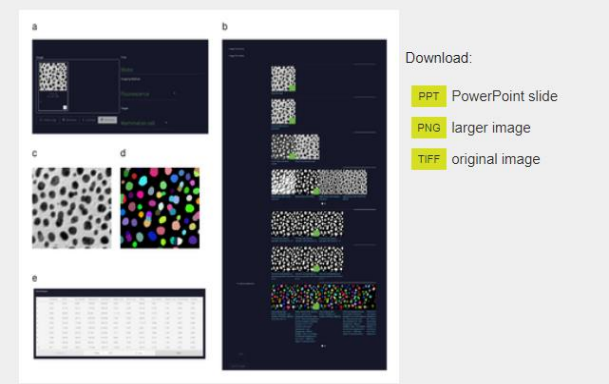

Fig.1. Interface of the IMACEL particle analyser

(a) The image title, imaging method, and specimen type should be provided to start each image processing process. (b) Click-based UI of the IMACEL particle analyser. Users click on the image to choose the most suitable processed image for each procedure, such as noise reduction, binarisation, and post processing. (c) Input image, (d) segmentation output image, and (e) quantitative segmentation output.

Many watershed algorithms are available at the last stage of the procedure. In addition, various morphological features of particles, like number, area, roundness, fitted ellipse long and short axes, centred coordinates, and solidity, are obtained automatically.
The IMACEL platform is devised for scientific image processing with attention on bio imaging. So, all suggested procedures are suitable for maintaining bio image integrity. To enable the archiving of image processing procedures in each researcher's experimental notes, an image processing report is also provided.

Classification algorithms in the IMACEL classifier

Two classification algorithms are performed in the current version of IMACEL: a random forest and a deep learning algorithm. The CNN architecture of AlexNet that was first place in the Image Net Large Scale Visual Recognition Challenge 2012 (ILSVRC2012), is used. AlexNet comprises of eight layers: five convolution layers and three fully connected layers. Moreover, the version used in IMACEL was pre-trained on the data used in ILSVRC2012.

Cell culture of mammalian and plant cells

The tobacco (*Nicotianatabacum*) BY-2 cell line was diluted 95-fold with a changed Linsmaier and Skoog medium assisted with 2,4-D at weekly intervals. The cells were agitated on a rotary shaker at 130 rpm at 27 °C in the dark. The cell cycle progression was synchronised with 5 mg–1 aphidicolin (Sigma). A transgenic BY-2 cell line, stably expressing an RFP-Histone H2B fusion protein, can be maintained and

synchronized by procedures same as those used for the original BY-2 cell line.

African green monkey kidney fibroblast-derived COS7 cells were acquired from the RIKEN BioResource Center and refined in high glucose Dulbecco's modified Eagle's medium (Gibco) supplemented with 10% qualified heat inactivated fatal bovine serum from USDA-approved regions (Gibco), 50 U/mL penicillin-50 μg/mL streptomycin (Gibco), 2 mM L-glutamine (Gibco), 1 mM sodium pyruvate (Gibco), MEM nonessential amino acids (Gibco), and 55 μM 2-mercaptoethanol (Gibco) at 37 °C in 5% CO2.

Stress treatment and immunofluorescence labelling

COS7 cells cultured onto a 35-mm glass-based dish (IWAKI) were treated with 0.5 μM sodium arsenite (Fluka) for 15 min or 60 min and fixed with 3% paraformaldehyde (Sigma Aldrich) and 0.1% glutaraldehyde (Sigma Aldrich) at 37°C in 5% CO2 for 10 min. COS7 cells were permeabilised with 0.2% Triton-X 100 (SIGMA) and blocked at 37°C in 5% CO2 for 30 min with 10% goat serum (Life Technologies) and then incubated for 30 min with primary antibodies, rabbit polyclonal anti-PABP antibody (Abcam), diluted in TOYOBO. After washing with 0.2% Triton-X 100 and PBS, the cells were incubated with Alexa 488 labeled goat anti-rabbit secondary antibodies diluted in Can Get Signal Solution A.

Microscope

For the observation of the cell cycle in BY-2 cells, the cells were imaged using fluorescent microscopy (FSX100, Olympus, Tokyo, Japan). To obtain the nuclear regions, the images were processed manually using ImageJ.

To observe the COS7 cells, they were imaged using fluorescence microscopy (N-STORM, Nikon, Tokyo, Japan). Noise in the fluorescent microscopy image was minimized with a difference of Gaussian filter using ImageJ

Manual evaluation of the number and size

To evaluate the number and size of stress granules, boundaries were derived manually using ImageJ software. Manual evaluations were done by two researchers who were not involved in this study to avoid biases that could overestimate the differences in treatment effect and underestimate the differences between results of manual evaluation and IMACEL particle analyzer.

Authors of this paper are experts in plant cell division. So, they annotated the training data for the classification of cell cycle progression in tobacco BY-2 suspension-cultured cells.

Statistical information

To analyze the differences in the number of stress granules, the Mann–Whitney U test was calculated using free statistical software R and RStudio versions 3.3.1 and 1.1.383, respectively.

IV.RESULTS

As illustrated in the below figure, IMACEL is a cloud-based image processing platform. Researchers upload images to the web server using a web browser. Image processing and image classification are done by high-performance virtual machines, and the processed image data are sent back using the browser. IMACEL has the following two independent functions: a particle analyzer for morphological analysis and a classifier for bioimage classification.

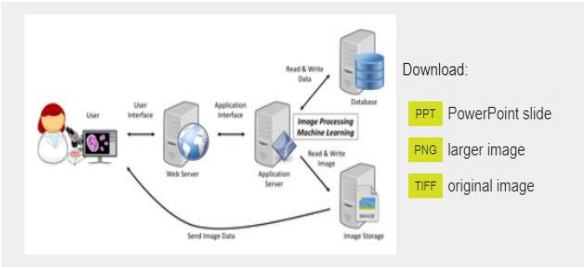

Fig.2. Architecture of IMACEL, a cloud-based image processing and machine learning platform for life science researchers.

The entire procedure of image processing is done in the cloud using high-performance virtual machines. The public-domain images used in this figure were obtained from Openclipart.

Validation of the IMACEL particle analyser

The morphological analysis of the IMACEL particle analyzer is validated by finding how similar its obtained features were to those of a manual evaluation. We focused on immunologically labeled stress granules as the shapes of the organelles are oval and traced easily by manual analysis. It's been reported that a treatment of sodium arseniteactivates the evolution of stress granules in a time-dependent manner [11–13]. So, COS7 cells treated with 0.5 μM sodium arsenite for 15 min and 60 min were examined with respect to the size and number of stress granules developed during treatment. We verified that the stress granules were segmented well by the IMACEL particle analyzer. As expected, there were significant variations in the number and size of stress granules between the 15 min and 60 min treatments, and the morphological analysis of IMACEL produced results that were very much similar to those of the manual evaluation. The batch process of the IMACEL particle analyzer (65 images each for specimen treated for 15 min and 60 min) was finished in nearly 5 min. By contrast, manual evaluation by tracing each stress granule took nearly 16 h. These results show that the IMACEL particle analyzer can evaluate the morphology of particles quantitatively and quickly, with high accuracy.

(a) Input image, binarised image, and output image of the IMACEL particle analyzer. Comparison of the distribution of the number (b) and size (c) of stress granules against stress treating time evaluated using manualanalysis and IMACEL. Asterisks express significant differences (Mann–Whitney U test) between cells treated with 0.5 μM sodium arsenite for 15 min and 60 min (in number: $p = 2.568 \times 10–13$ and $p < 2.2 \times 10–16$, in size: $p < 2.2 \times 10–16$ and $p < 2.2 \times 10–16$). (d) Total time spent on manual evaluation versus the computational time of the IMACEL particle analyzer. We measured 65 images each for specimen treated for 15 min and 60 min.

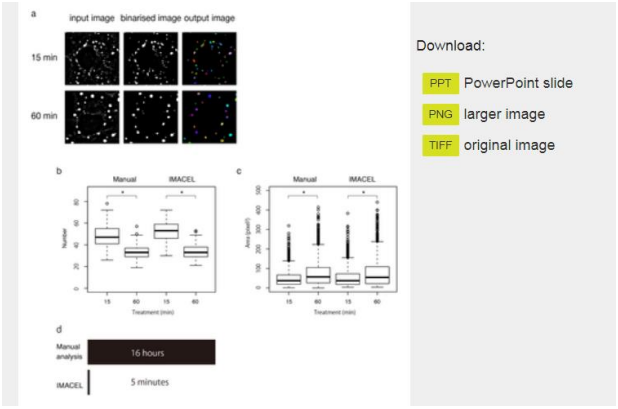

Fig. 3. Results of the IMACEL particle analyser for extracting and evaluating stress granules in COS7 cells.

Validation of the IMACEL classifier

To validate the IMACEL classifier, a classification of cell cycle progression in tobacco BY-2 suspension-cultured cells was done using two machine learning methods: random forests and deep learning. With its highly synchronized cell cycle progression [14], this cell cycle is very appropriate for bioimage classification. Moreover, synchronized BY-2 cells are one of the extremely suitable suspension-cultured cells for observing every cell cycle. Nucleuses and chromosomes were visualized using histone H2B-RFP [14,15], and the image features were obtained using the LPX296 feature extractor (formerly the KBI feature extractor [10]) and a higher-order local autocorrelation feature extractor.

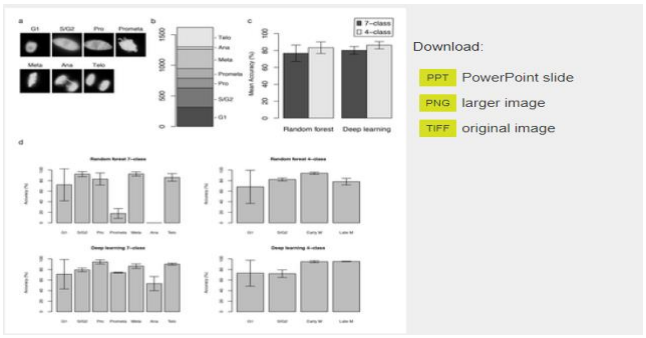

Fig.4. Results of the IMACEL classifier for cell cycle classification with nucleuses visualized using fluorescent images.

(a) Representative images of every cell cycle in suspension-cultured plant cells. Nuclear regions were visualized with RFP-Histone H2B. (b) Distribution of number of dataset images in every class. (c) Mean accuracy of cell cycle classification in seven-class and four-class classification using random forests and deep learning. For four-class classification, the prophase, prometaphase, and metaphase were merged into the early mitotic phase. Anaphase and telophase were joined into the late mitotic phase. (d) Accuracy of each cell cycle classification with bars representing the standard deviation based on three independent experiments.

The classification dataset was prepared with 1,619 images of seven classes. To avoid over fitting, the mean accuracy was calculated using three-fold cross-validation. The UI of the actual IMACEL classifier was shown in S3 Movie.

The random forest IMACEL classifier [10] identified seven cell cycle classes with a mean accuracy of approx.ly 76.69% and four classes with a mean accuracy of approx.ly 83.31%. S/G2 and metaphase were classified with high accuracy, but prometaphase and anaphase was classified with relatively low accuracy.

By contrast, the deep learning technique in IMACEL managed to identify seven cell cycle classes with a mean accuracy of approx.ly 80.17% and four classes with a mean accuracy of approx.ly 86.21%. The mean accuracies of prometaphase and anaphase classification increased when deep learning classification was used.

These results showthat IMACEL can automatically classify images without needing the researchers to have advanced knowledge of several image processing and machine learning techniques.

V. DISCUSSION

The development of the IMACEL platform was based on two design features. The first one is that of a novel clickable-based UI. Current image processing software, like ImageJ or Photoshop CC, needs researchers to actively choose the desired function from a list of image processing process. As there is lot of flexibility in the function selection, errors can be made if inappropriate image processing procedures are used. For instance, a nonlocal mean filter [18] that is an effective noise reduction techniques, does smoothing using similar intensity distributions from distant regions irrespective of whether the regions are biologically identical or not. So, when such filtering is executed in image processing software, researchers should avoid using it. By contrast, the IMACEL particle analyzer effectively restricts the functions that can be chosen by those unfamiliar with image processing. Additionally, batch processing is easily done without the need to write macro functions in a programming language.

The second concept is the cloud-based image processing platform. Generally, machine learning requires extensive computing resources. The construction of an analytical environment is verycomplicated for many biological researchers. Moreover, high-performance machines are costly to establish in everylab. In IMACEL, as image processing and machine learning are done on high-performance virtual machines, users can freely use their own analytical environment with a browser from anywhere. In addition, as IMACEL saves previous analytical images, the platform can play the role of an image management tool.

This platform is developed for researchers in the broad field of life sciences. Microscopic images are quite often observed than MRI images in few life science journals. So, the focus of this validation study is on (fluorescence) microscopic images. However, in a similar study, a prototype version IMACEL was used to classify transmitted electron microscopic images of tumorigenic cancer stem cells into two categories (ABCGS2+ and ABCGS2-) [19]. Though, on the IMACEL platform, the image acquisition tools or type of image that may be used for analysis are not restricted. In fact, an extension to the image

processing platform that is focused on MRI, CT, and X-ray images for specific fields of life science is being developed.

Compared with manual labeling, the classification technique in IMACEL does not seem to be highly accurate. There are various reasons for this performance in this study. First, the number of images in our dataset was small, especially for the anaphase cell images. Deep learning is renowned to perform better with a large number of images, and if one class has few examples, the resulting dataset can be imbalanced and affect the accuracy. Second, every cell cycle image was obtained using cheap fluorescent microscopy rather than a more advanced method, like confocal laser scanning microscopy, and high levels of image noise could affect the output. Third, transfer learning might have affected the result. AlexNet was trained using both microscopic images and general images. Note that the above poor study conditions were chosen to assess the IMACEL platform as it's aimed at researchers who don't have advanced computing skills or equipment.

A current version of the IMACEL platform, all microscopic images employed in this study and detailed documentations will be sharedwith interested researchers on request. Presently, three-dimensional reconstruction and the extraction of the surface area and volume for three-dimensional images are developed. In addition, tracking or kinetic analysis for time-sequential observations is under development.

VI.CONCLUSION

In conclusion, we developed a new cloud-based image processing platform known as IMACEL that contains morphological analysis and image classification functions. The validation experiments show that particles can be extracted easily and quickly with high accuracy. In addition, IMACEL enables researchers to perform image classification based on machine learning without previous knowledge of image processing.

VII.REFERENCES

[1] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: Pereira F, Burges CJC, Bottou L, Weinberger KQ, editors. Advances in Neural Information Processing Systems 25. Curran Associates; 2012. p. 1097–105.

[2] Chaib S, Yao H, Gu Y, et al. Deep feature extraction and combination for remote sensing image classification based on pre-trained CNN models. International Conference on Digital Image Processing. 2017:104203D.

[3] A.A.M. Al-Saffar, H. Tao, M.A. Talab, *Review of deep convolution neural network in image classification* (International conference on radar, antenna, microwave, electronics, and telecommunications. IEEE, Jakarta, 2018), pp. 26–31.

[4] J.Y. Lee, J.W. Lim, E.J. Koh, A study of image classification using HMC method applying CNN ensemble in the infrared image. Journal of Electrical Engineering & Technology **13**(3), 1377–1382 (2018).

[5] Kutsuna N, Higaki T, Matsunaga S, Otsuki T, Yamaguchi M, Fujii H, et al. Active learning framework with iterative clustering for bioimage classification. Nat Commun. 2012 Aug;3:1032. pmid:22929789.

[6] L. Yang, A.M. Maceachren, P. Mitra, et al., Visually-enabled active deep learning for (geo) text and image classification: a review. ISPRS Int. J. Geo-Inf. **7**(2), 65 (2018).

[7] A.B. Said, I. Jemel, R. Ejbali, et al., *A hybrid approach for image classification based on sparse coding and wavelet decomposition* (Ieee/acs, international conference on computer systems and applications. IEEE, Hammamet, 2018), pp. 63–68.

[8] M.Z. Afzal, A. Kölsch, S. Ahmed, et al., *Cutting the error by half: investigation of very deep CNN and advanced training strategies for document image classification* (Iapr international conference on document analysis and recognition. IEEE computer Society, Kyoto, 2017), pp. 883–888.

[9] B. Kieffer, M. Babaie, S. Kalra, et al., *Convolutional neural networks for histopathology image classification: training vs. using pre-trained networks* (International conference on image processing theory. IEEE, Montreal, 2018), pp. 1–6.

[10] Mou L, Ghamisi P, Zhu X X. Unsupervised spectral-spatial feature learning via deep residual conv-Deconv network for hyperspectral image classification IEEE transactions on geoscience & Remote Sensing. 2018,(99):1–16.

[11] Reddick WE, Glass JO, Cook EN, et al. Automated segmentation and classification of multispectral magnetic resonance images of brain using artificial neural networks. IEEE Trans Med Imaging. 1997;16(6):911–918. [PubMed] [Google Scholar].

[12] S. Roychowdhury, J. Ren, *Non-deep CNN for multi-modal image classification and feature learning: an azure-based model* (IEEE international conference on big data. IEEE, Washington, D.C., 2017), pp. 2893–2812.

[13] S.A. Quadri, O. Sidek, Quantification of biofilm on flooring surface using image classification technique. Neural Comput. &Applic. **24**(7–8), 1815–1821 (2014).

[14] Sachin R, Sowmya V, Govind D, et al. Dependency of various color and intensity planes on CNN based image classification. International Symposium on Signal Processing and Intelligent Recognition Systems. Springer, Cham, Manipal, 2017:167–177.

[15] Sachin R, Sowmya V, Govind D, et al. Dependency of various color and intensity planes on CNN based image classification. 2017.

[16] X. Fu, L. Li, K. Mao, et al., in *Chinese High Technology Letters*. Remote sensing image classification based on CNN model (2017).

[17] Shima Y. Image augmentation for object image classification based on combination of pre-trained CNN and SVM. International Conference on Informatics, Electronics and Vision & 2017, International sSymposium in Computational Medical and Health Technology. 2018:1–6.

[18] X. Wang, C. Chen, Y. Cheng, et al, Zero-shot image classification based on deep feature extraction. United Kingdom: IEEE Transactions on Cognitive & Developmental Systems, **10**(2), 1–1 (2018).

[19] Z. Yan, V. Jagadeesh, D. Decoste, et al., HD-CNN: hierarchical deep convolutional neural network for image classification. EprintArxiv 4321-4329 (2014).

Cite this article as :

K. Sunil Manohar Reddy, Prof. BK Tripathi, Dr. S. K. Tyagi, "Review on Cloud-Based Image Classification", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 6 Issue 2, pp. 291-298, March-April 2019.
Journal URL : http://ijsrset.com/IJSRSET207481