*Journal of Discrete Mathematical Sciences & Cryptography*
ISSN 0972-0529 (Print), ISSN 2169-0065 (Online)
Vol. 26 (2023), No. 3, pp. 899–911
DOI : 10.47974/JDMSC-1777

# Impact of computer vision based secure image enrichment techniques on image classification model

A. Shubha Rao *
K. Mahantesh
*Department of ECE*
*SJB Institute of Technology*
*Bangalore 560060*
*Karnataka*
*India*

**Abstract**

Growing advancement in Artificial Intelligence has led to developing systems which can think, analyze, recognize patterns and are capable enough to convert data into information has turned out to be a major prerequisite. An effort to provide enriched data with Computer Vision based Image Enrichment techniques and its impact on performance of Deep learning algorithms (VIRNet) is analyzed in the research. At first, various image enrichment technique like sharpening, denoising, color enhancement which makes the dominant features more visible which can be easily captured by the classification model. Later, popular computer vision techniques for image segmentation which differentiates the background and foreground area like Deeplabv3, Mask R-CNN and Region based segmentation methods are applied and its impact over Classification model is analyzed. The experimental study was conducted on popular image dataset Caltech 101 and Caltech-256, from the results it is clearly evident the image enrichment based on CLAHE outweighs the popular segmentation methods. Besides image segmentation and enrichment techniques, implementing an AES encryption in the process helps in protecting the sensitive data and ensures the data is kept private and secure. An image classification algorithm using deep learning equipped with AES encryption is implemented, which adds a layer of security to an image processing system.

***Subject Classification:*** *Primary 68T07, Secondary 68T45.*

***Keywords:*** *Deep learning, Computer vision, Image segmentation.*

* *E-mail:* **mail2shugar@gmail.com**

© Taru Publications

## 1. Introduction

With the growing development of technology in the field of Artificial Intelligence and Robotics, each day comes as a challenge to break the barriers built in the respective field for the future advancement [1]. Image classification is one such area which plays a major role in most of the state-of-the-art based technology leading from Self Driving Automobile to Remote Sensing to Robotic Assistance in various fields [2]. Effort to accurately, efficiently classify images has paved its long way from the traditional hand crafted feature extraction based Machine learning approach to self-learning convolution based Deep Learning Algorithms. The renowned Convolution Neural Network (CNN) learns the ideal filters to extract best features from the image which will result in accurate Classification [3].

On the journey instead of passing the images directly to the model, computer vision based image enrichment techniques are applied to the images which makes the dominant features more visible leading to improved image classification. Various feature enrichment techniques like reducing the noise, sharpening, color contrast and channel adjustment and extracting the foreground using segmentation are applied. The impact of computer vision based feature enrichment methods on Deep Learning model against the Original Image is analyzed and discussed. The flow diagram of the research conducted is outlaid in Figure 1.

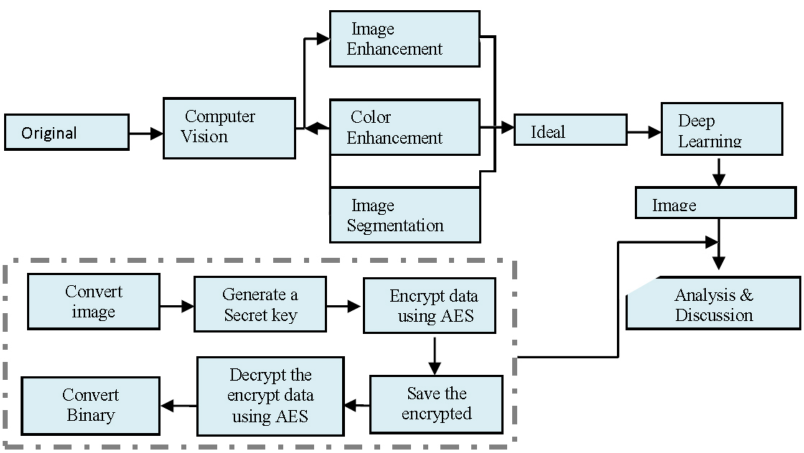



**Figure 1**

**Flow diagram of analysis impact of computer vision based secure image enrichment on Image Classification**

## 2. Literature Review

Scene text recognition which combines features extracted by atrous convolution at multiple scales based on regression is proposed [4]. A robot to fetch the ball and carry was implemented with the help of computer vision based deep learning techniques that could recognize the ball efficiently from the image feeds. Many such innovative real- life applications of image processing can be used for securing text information [5]. Another popular application of computer vision is for change detection in videos. A much generalized lighter ChangeDet model was designed where maximum multi-spatial receptive (MMSR) field blocks are used for background estimation based on contrast change from temporal history [6].

Application of deep learning in securing and monitoring maritime activity is explored, where the text is extracted from the images and converted into a format feasible for machines with the EAST model. In the next step, segmented regions are processed using OCR model [7]. A review of performance of various deep learning models is conducted for the analysis of MRI brain tumor image segmentation [8]. Most of the real life event based images contains text that gets occluded by other objects. A deep learning based algorithm which first eliminates the false positives and later predicts the missing text based on NLP is proposed [9]. A method to improve the retrieval performance of skin blisters images by indexing the images based on LBP and DFT that reduces both memory and time complexity [10].

## 3. Computer Vision Methods

In order to improve the performance of image classification, various computer vision based image enrichment techniques are applied on original image making the dominant features easily noticeable for the deep learning model.

### 3.1 *Image Enrichment Techniques*

**Image Sharpening:** Image Sharpening using Unsharp Mask (*USM*) is the most traditional method used to make the edge more dominant and clearly visible. In the process unsharp mask (*M*) is created, by smoothing the image using Gaussian filter (*G*) which preserves the constant version of the image by removing the spikes. Later, appropriate weighted version of the mask is subtracted from the original image (*O*) to get contrast enhanced edges while preserving the acutance of the image. Application of the Unsharp Mask to original image is shown in Figure 2.

$$G = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

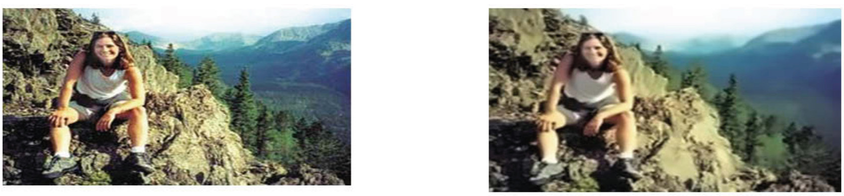
$$USM = w_1 O - w_2 M \tag{2}$$

**Image Denoisning:** Most of commonly used noise removal techniques compute the noise only by considering smaller local neighborhood around the pixel. A much improved FastNIMeans Denoising method for colored images, which finds average of the noise from all the similar areas in the image that needs to be removed [11]. The method is able to efficiently retain the better texture from the original image. Example of FastNIMeans denoising technique is shown in Figure 3.

(a) (b)

**Figure 2**

**(a) Original Image (b) USM sharpened Image**



(a) (b)

**Figure 3**

**(a) Original Image (b) FastNIMeans Denoised Image**

3.2 *Color Enhancement Techniques*

Many color based image enhancement techniques have been proposed including the Multi-Scale Retinex with Color Restoration (MSRCR), Multi-Scale Retinex with Chromaticity Preservation (MSRCP), but the most efficient among all the others is the Contrast Limited Adaptive Histogram Equalization (CLAHE) method. Adaptive Histogram equalization (AHE) is a technique which computes and performs histogram equalization for distant, similar smaller section of the image. AHE increases the local contrast and enhances the edges, but it ends up amplifying even the undesirable noise inherently present. To overcome the disadvantage, the idea of Contrast Limited Adaptive Histogram Equalization (CLAHE) method was proposed. CLAHE performs the enhancement proportional to the Cumulative distribution function (*F*) for the given pixel neighbourhood by chipping off the histogram [12]. The chopped histogram is later redistributed equally among all the bins.

$$F(x) = P(X \le x) = \Sigma_{x_i} p(x_i) = \int_{-\infty}^{x} f_X(t)\,dt \tag{3}$$

3.3 *Image Segmentation Techniques*

Yet, above all the pre-mentioned computer vision techniques – Image Segmentation alone stands as a major advancement in the field. With the invention of various kinds of convolution as an important technique to better understand an image for semantic feature extraction various state-of-the-art architectures have been proposed each performing better than the other [13]. Some of the popular image segmentation techniques which are studied are:

**Deeplabv3plus:** An architecture which is mainly based on dilated convolution, it is able to extract larger feature maps with its spread out filters at various scales achieves greater visibility while maintaining the same number of parameters. Together merged with methods like Atrous Spatial Pyramid Pooling, depthwise convolution, improves object boundaries with encoder decoder architecture, deeplabv3plus based transfer learning offers various kind of segmentation approaches with significant performance.

**Mask RCNN:** It's an upgrade to Faster RCNN with very little overhead, uses Region Proposal Network to extract the probable regions containing object. Mask RCNN replaces the RoI pool with RoI Align to better preserve the spatial information more accurately, the result of which is fed into series of convolution layer to produce the pixel wise mask for the object.

**ML based segmentation:** Traditional machine learning based segmentation is quite effective when it comes to segmenting images of different classes, as opposed to deep learning model which perform well mostly on pre-trained set of classes. Hence for the sake of comparing the result, edges are first detected using canny, later contours are found, followed by defining the background and foreground area, generating mask and finely cutting the image to obtain the segmented version.

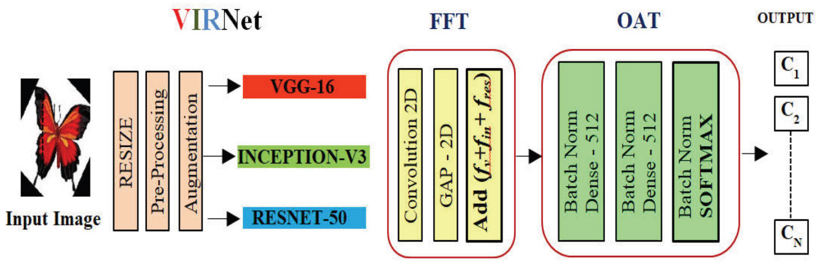



**Figure 4**

**VIRNet architecture used for Image classification**

3.4 *VIRNet: Deep Learning Model*

Deep learning model which is being used for the study and analysis of impact of computer vision techniques on image classification task is the VIRNet [14]. Since feature extraction is a major part based on which classifier classifies, VIRNet employs best models at hand as feature extractor - VGG16, Inception-v3 and Resnet-50. The extracted semantic features are carefully fused together using a Feature Fusion Technique (FFT) to create a final fused feature vector (*Ffused* ). The features are passed onto the One for All Top (OAT) Module, series of batch normalization with Dense layer pushes the learning further, fits itself as the ideal classifier top for all the 3 models. VIRNet is an architecture designed to extract rich set of semantic features, which can be trained faster even from a smaller training sample with efficient outcome. Figure 4 shows the detailed architecture of VIRNet.

3.5 *Dataset Caltech-101 & Caltech-256 used for image classification.*

The digital, public dataset was collected by California Institute of Technology to promote computer vision applications in 2003. Caltech-101 contains 9,146 annotated images of resolution around 300*200 pixels belonging to 101 object categories and a single background class, with a minimum of 50 images in each class. Super collection Caltech-256 dataset contains 30,607 images belonging to 256 different object categories and a single clutter class [15]. The minimum number of images per category was increased to 80 images.

3.6 *Implementing AES for enhanced security*

After the image classification by a deep learning algorithm, implementing AES encryption in image processing can provide an additional layer of security for sensitive image data. The classified image needs to be converted to binary data before AES encryption. This can be achieved by reading the image file and converting it to binary data using a suitable programming language. A secret key needs to be generated using a secure key generation method, and the AES algorithm can be used to encrypt the binary data. The encrypted binary data can then be saved to a file or database for future use. To decrypt the encrypted data, the same secret key is used to decrypt the binary data using the AES algorithm. Finally, the decrypted binary data can be converted back to an image format, such as JPEG or PNG, to view the decrypted image. Implementing AES encryption in image processing after image classification can help to secure the image data and prevent unauthorized access to it. This added layer of security can be especially important for sensitive image data, such as medical images or confidential documents.

## 4. Results and Discussion

The study is conducted using latest version of Python programming language, Tensorflow-gpu and CUDA. Since the model was designed to learn faster, it will be trained 20 epochs using Adam Optimizer with Categorical Cross Entropy while monitoring and recording the best model version for the validation accuracy. The performance of the model is measured on dataset in terms of accuracy (%).

4.1 *Result of Image Enrichment Techniques*

At first, the results of applying various image enrichment techniques based on USM sharpening and fastNMI denoising to the original image was evaluated on deep learnring model for classification. The various combination of methods that were evaluated are: (a) USM based sharpening an image (b) fast NMI based denoising an image (c) fastNMIDenoising the USM sharpened image and (d) using both Gaussian blur and fastNMI denoised version as a mask on original image. From the results in Table 1. it can be concluded that USM based sharpening the image makes the edges more dominant, improves the performance of the classification model on both Caltech101 and Caltech256 datasets.

**Table 1**

**Result of Image Enrichment Techniques on test dataset using VIRNet model**

| Image Enrichment Techniques Accuracy (%) | Caltech-101 30-Train | Caltech-256 30-Train |
|---|---|---|
| Original Image | 91.08 | 74.25 |
| USMsharpened image | 92.23 | 77.39 |
| fastNMIDenoise image | 91.64 | 73.86 |
| Denoise an USM sharpened image | 92.02 | 74.83 |
| Both Gaussian blur and denoise as mask | 89.25 | 71.24 |

### 4.2 *Result of Color Enhancement Techniques*

Next, the impact of CLAHE based image enrichment technique to the original image for classification task was analyzed. In addition to RGB color space, the application of CLAHE color equalization technique was also studied by converting the image into LAB and HSV color spaces. From the results it is evident that application of CLAHE for an image in LAB color space improves the deep learning models accuracy. The results are tabulated in Table 2.

### 4.3 *Result of Image Segmentation Techniques*

Later, the impact of computer vision based segmented image are passed onto the VIRNet deep learning model for classification and the results are analyzed. Some of the popular models which were applied to segment the image are: (a) DeepLabv3 plus architecture - background

**Table 2**

**Result of Color Enhancement Techniques on test dataset using VIRNet model**

| Color Enhancement Techniques Accuracy (%) | Caltech-101 30-Train | Caltech-256 30-Train |
|---|---|---|
| Original Image | 91.08 | 74.25 |
| CLAHE in RGB image | 92.48 | 75.28 |
| CLAHE in LAB image | 92.56 | 80.12 |
| CLAHE in HSV image | 92.31 | 76.03 |

changed to black (b) Mask RCNN based segmentation -cropping the image for coordinates of bounding box of the object (c) Machine Learning based segmentation. Since the whole buzz of computer vision concentrates so heavily on image segmentation it was much required to analyze the impact on popular segmentation models on the task of image classification. It is a huge disappointment that deep learning model fails to learn and extract features from segmented image. The comparison of results of segmented image versus the original image is tabulated in Table 3.

From the results of the whole experiential study conducted on the impact of the computer vision based image enrichment techniques on Image Classification, it is clearly evident that enriched features of image make a significant impact on the deep learning model. The proposed computer vision based application of CLAHE color equalization improves the accuracy of the model when compared to original image. A comparison of proposed method with the previous work on Caltech101 and Caltech256 is tabulated in Table 4 and Table 5 respectively. Sharpening the images with visible clear edges makes quite an impact on deep learning model next to CLAHE based application. It is worth mentioning, impact of merging the best methods like Sharpen with CLAHE, Denoising sharpened image with CLAHE was conducted. But merging the methods is analogous to one method trying to negate the feature made dominant by the other, hence the model failed to show any significant improvement.

Failure with the impact of segmented images in classification task, pushes to the edge of rethinking which features exactly the deep learning model rely on and what causes the negative impact. It is evident the segmented image by highlighting the object present in the image and converting the rest to black representing background suffers badly. Since most of the popular models are pre-trained with certain set of preset classes, the performance drops on any new classes which is unseen by the model. Henceforth, developing segmentation algorithms with classification task

**Table 3**

**Result of Image segmentation Techniques on test dataset using VIRNet model**

| **Image Segmentation Techniques Accuracy (%)** | **Caltech-101 30-Train** | **Caltech-256 30-Train** |
|---|---|---|
| Original Image | 91 | 74.25 |
| DeepLabv3+ segmented image | 60 | 32.56 |
| Mask RCNN segmented image | 77 | 56.31 |
| ML based segmented image | 77 | 58.12 |

in mind such that it preserves the core essence of the image while making the features more dominant like CLAHE which makes a significant impact on performance is essential.

**Table 4**

**Comparative analysis of proposed method on Caltech-101 with previous work**

| Methods<br>Accuracy (%) | Caltech-101<br>15-Train | Caltech-101<br>30-Train |
|---|---|---|
| Shape matching [16] | 45 | - |
| Pyramid match kernels [17] | 49.5 | 58.2 |
| Discriminative nearest neighbour [18] | 59 | 66 |
| Local naïve bayes nearest neighbour [19] | 47.8 | 55.2 |
| Sparse localized features [20] | 33 | 41 |
| Relevance based classification [21] | - | 43.8 |
| Gaussian mixture models [22] | - | 72.3 |
| Inceptionv3 [23] | 64 | 67 |
| Ensemble Model [24] | 73.11 | 79.23 |
| VIRNet [14] | 87.74 | 91 |
| Computer Vision+VIRNet (Proposed) | 90.36 | 92.56 |

**Table 5**

**Comparative analysis of proposed method on Caltech-256 with previous work**

| Methods<br>Accuracy (%) | Caltech-256<br>15-Train | Caltech-256<br>30-Train |
|---|---|---|
| Learning dictionary [25] | 30.35 | 36.22 |
| Sparse spatial coding [26] | 30.59 | 37.08 |
| Discriminative coding for object classification [27] | 28 | 30 |
| Local naïve bayes classifier [19] | 33.5 | 40.1 |
| Caltech Institute classification [15] | 28.3 | 34.1 |
| Combined image descriptors [28] | - | 33.6 |
| Inceptionv3[22] | 58 | 59 |
| Ensemble Model [23] | 60 | 62 |
| VIRNet [14] | 70.28 | 74.25 |
| Computer Vision+VIRNet (Proposed) | 75.05 | 80.12 |

## 5. Conclusion

The purpose of the research conducted was to analyze the impact on computer vision based techniques on deep learning model (VIRNet) for image classification. The impact of various techniques which were studied are USM sharpening, fast NMI based denoising, CLAHE in different color spaces and image segmentation. The research has clearly proved that image quality makes a significant mark on the deep learning model. Enhancing significant features like edges, color contrast present in the image makes a huge impact like captioning the distinctive features and making is clearly visible for the deep learning model.

Failure of the segmented images has pushed us to rethink and rebuild the segmentation algorithms with classification in mind, the way deep learning model understands the image. Developing segmentation model which retains the core significant features like edges, the color contrast, the brightness, which are worth making a difference on classification algorithms is important for next generation of Artificial Intelligence.

It can be concluded from the research, with the application of enriched computer vision techniques it is possible to significantly improve the performance of deep learning algorithm with better generalization capability because features is all that matters. As a part of further work, various other computer vision based techniques needs to be applied and examined.

## References

[1] Sahera H Z A, Ali Akbar H & Masoud T, Modeling and prediction using an artificial neural network to study the impact of foreign direct investment on the growth rate / a case study of the State of Qatar, *Journal of Statistics and Management Systems*, 25:8, 1991-2003, (2022).

[2] Farah A. Abdulghani, Nada A.Z. Abdullah, Hybrid deep learning model for Arabic text classification based on mutual information, *Journal of Information and Optimization Sciences*, 43:8, 1901-1908, (2022).

[3] Nana Yaw Duodu, Warish Patel, Advanced EHR system in homes-in-schools' automation using internet of things and machine learning, *Journal of Information and Optimization Sciences*, 43:8, 1953-1962, (2022).

[4] Liangjun Wang, Weijie Gu, Scene text detection with improved receptive field and adaptive feature fusion, Proc. SPIE 11911, (2021).

[5] Anil Kumar & Sandeep Kumar Sharma. Information cryptography using cellular automata and digital image processing, *Journal of Discrete Mathematical Sciences and Cryptography*, 25:4, 1105-1111 (2022), DOI: 10.1080/09720529.2022.2072437.

[6] M. Mandal, Scene Independency Matters: An Empirical Study of Scene Dependent and Scene Independent Evaluation for CNN-Based Change Detection, IEEE Trans on Intelligent Transportation Systems, 23(3), (2022).

[7] Cornelia Nita, Marijke Vandewal, On the automatic text detection and recognition algorithms for maritime images, Proc. SPIE 11870, Artificial Intelligence and Machine Learning in Defense Applications III, (2021).

[8] Amishi Vijay, B.S.Saini. A Review of Brain Tumor Image Segmentation of MR Images Using Deep Learning Methods, AJIR proceedings, (2021).

[9] Ayush Mittal, Palaiahnakote Shivakumara, Umapada Pal, Tong Lu, Michael Blumenstein, A new method for detection and prediction of occluded text in natural scene images, Signal Processing: Image Communication, (2021).

[10] K. Sujatha et al., A Content-Based Image Retrieval System for Diagnosis and Detection of Skin Cancer Using Self-Organizing Feature Maps, Digital future of health care, CRC Press, (2021).

[11] T. Zin, S. Seta et al., Local Image Denoising Using RAISR, IEEE Access, 10, 2420-22428, (2022).

[12] Lu, P.; Huang, Q. Robotic Weld Image Enhancement Based on Improved Bilateral Filtering and CLAHE Algorithm. Electronics, 11, 3629, (2022).

[13] Rupinder Kaur, Anubha Jain, Implementation and assessment of new hybrid model using CNN for flower image classification, *Journal of Information and Optimization Sciences*, 43:8, 1963-1973, (2022).

[14] Rao, A.Shubha, Mahantesh, K, VidyaShree N. VIRNet: One for All Top based on Feature Fusion Technique. HCII, Springer, CCIS, 1580, (2022).

[15] A. Holub G. Griffin and P. Perona. Caltech 256 object category dataset. Technical Report, California Institute of Technology, (2007).

[16] T. L. Berg A. C. Berg, J. Malik. Shape matching and object recognition using low distortion correspondence. In IEEE CVPR, 1:26-33, (2005).

[17] K. Grauman, Darell. Pyramid match kernels: Discrminative classification with sets of image features. Technical report MIT-CSAIL-TR-2006-020, (2006).

[18] Maire. M-Malik Hao Zhang, Berg A.C. Svm-knn: Discriminative nearest neighbor classification for visual category recognition. IEEE-CVPR, (2006).

[19] Sancho Mc Cann and David G. Lowe. Local naive bayes nearest neighbor for image classification.IEEE-CVPR, 650–3656, (2012).

[20] Jim Mutch and David G Lowe. Muticlass object recognition with sparse, localized features. IEEE CVPR, (2006).

[21] R Rigamonti, Vincent Lepetit. On the relevance of sparsity for image classification. Computer Vision and Image Understanding, 125, (2014).

[22] Mahantesh, K., Aradhya, An Impact of Complex Hybrid Color Space in Image Segmentation. Advances in Intelligent Systems and Computing, (2014).

[23] Shubha Rao, A., Mahantesh, K. Image Classification Based on Inception-v3 and a Mixture of Handcrafted Features. Distributed Computing and Optimization Techniques. Lecture Notes in Electrical Engineering, 903,(2022).

[24] Rao,A.Shubha, K..Mahantesh. Ensemble Model for improved Image Classification. CCIS, Springer, 1697, (2022).

[25] Yu-Jin Zhang Bao-Di Liu, Yu-Xiong Wang. Learning dictionary on manifolds for image classification. Pattern Recognition, 46:1879-1890, (2012).

[26] Antonio W. Vieira Mario F. M. Sparse spatial coding: A novel approach for efficient and accurate object recognition. IEEE International Conference on Robotics and Automation (ICRA), 2592-2598, (2012).

[27] Zhang Y Zheng Y Liu B,Discriminant sparse coding for image classification. International Conference on Acoustics, Speech and Signal Processing, (2012).

[28] C Liu, S Banerji, New image descriptors based on color, texture, shape, and wavelets for object and scene image classification. Neurocomputing, (2013).