SHAREit.

Scientific Hub of Applied Research in Engineering & Information Technology

Received: 11.01.2021 Revised: 09.02.2021 Accepted: 24.02.2021

check for updates

Research Article

Image Segmentation by Combining Artificial Intelligence and Intense Investigation Techniques

N. Shanmugapriya[*1], S. Pannirselvam[1]

[1] *Department of Computer Science, Erode Arts and Science College (Autonomous), Erode – 638009, India.*

Various electronic brain insight also artificial intelligence researchers archaic grabbed towards Image semantic segmentation. Various implementations including independent steering, inside exploration, and even implicit or aggravated fact networks greatly necessitate precise also proficient segmentation mechanisms. Also, deep learning methodologies are highly demanded in nearly all domain or execution aim correlated to electronic brain insight, comprising linguistic dissection or site recognizing. The image segmentation is attained by series process and universal segmentation method does not prevail for low resolution image in present method, but there arises assessing difficulty for performance comparison of these segmentation methods. The research work intends at combining machine learning and deep learning approaches and thereby yielding semantically precise predictions and exhaustive segmentation along with attaining computationally effectual way. Image color segmentation is achieved by machine learning and semantic labeling through deep learning. Two algorithms are utilized in this process. First algorithm deals with detecting super pixels using Modified Support Vector Machine (MSVM) based machine learning and these super pixels segmentation is done depending on textures and colors. The second algorithm utilizes convolutional Neural Network (CNN) for training color categories and thus achieving object classification into semantic labels. It is thereby substantiated that suggested semantic image segmentation methods offers great predictive accuracy in contradiction with other segmentation approaches.

Keywords: Image semantic segmentation, feature extraction, Modified Support Vector Machine (MSVM), classification, convolutional Neural Network (CNN).

1. Introduction

The object recognition precision besides efficiency can be greatly enhanced by ideal object automatic segmentation of semantic region. Apart from this, image level classification and detection are considered as other key image tasks [1]. Every image is treated as an identical category which corresponds to classification whereas detection corresponds to object localization and recognition.

Correspondence should be addressed to
N. Shanmugapriya; shanmugapriya.n@gmail.com

The pixel-level prediction usually confers to image segmentation due to its pixel classification into its category. Besides, task named instance segmentation helps in merging detection and segmentation in a combined way. Linguistic figure dissection similarly called as portrait labeling or site analyzing, associates fixing linguistic labels problem (e.g., “person” or “dog”) to each image component. The semantic image segmentation process comprises of image pixel classification into a specimen, where every specimen relates to class [2]. This job is considered as scene understanding concept or an improved image global context.

The scene understanding is regarded as one among key steps and challenging task in computer vision. Massive application such as figure correcting, aggravated fact

and hard-driving vehicles are greatly facilitated by successful image segmentation techniques [3]. Grouping generation done by present image segmentation algorithms are quite challenging which is equivalent to semantic regions. Intraclass diversity as well as inter-class ambiguity are the key solution for addressing this challenge. Particularly, object itself comprises quite high distinction fields or contrast amid point and surrounding is trivial [4]. Texture is regarded as beneficial characterization for an extensive range of image in this segmentation process. Human visual structures greatly utilize appearance for identification and explanation. Generally, shade is typically a component assets whereas surface assessments can be done from components groups [5]. Texture features extractions have been suggested by enormous number of methods.

Texture feature extraction classification can be normally divided into geographical texture quality production techniques and phantom texture quality production techniques [6]. Texture features extraction can be accomplished through pixel statistics computation or local pixel structures determination in original image domain while second helps in image transformation into frequency domain and feature computation from transformed image [7]. Moreover, the challenging issues such as image partitioning, image features organization need to focused. Generally, there are primarily three techniques for an image transformation into regions set: structured framework technique, unverified figure dissection and divert stage warner.

Recently, several motivating and innovative image segmentation algorithms have been suggested besides these algorithms classification is divided into five main classifications such as thresholding, template matching, region growing, edge detection and clustering [8]. These are validated to be effective in several applications, nevertheless nothing are largely suitable to entire images, in addition dissimilar algorithms are typically not similarly appropriate for specific application. Region growing algorithms is mainly meant for image feature information spatial repartition [9]. On the whole, improved performance is obtained when compared with thresholding methods for numerous images sets. Nevertheless, classic region growing processes are fundamentally sequential [10]. Regions is dependent on order of pixels scanning and pixels value, which in turn depends on problems revealed above, several approaches and their respective improvements have been anticipated for ensuring accuracy and image segmentation rapidity [11]. In order to mitigate the problems, exploiting information on other domains, particularly artificial intelligence is greatly required. In recent times, intelligent methodologies, such as social web work and brace line appliance (BLA) are suggested for image segmentation. Also, machine learning and machine learning based algorithms has been focused by many researches recently.

2. Literature Review

This section gives an outline of current techniques reviews for semantic images classification. The semantic image segmentation methods are discussed at the same time by considering both different features distinguishability and relationship amid neighboring pixels.

Chen et al [12] suggested a methodology for learning multi-scale features at every pixel location in a soft manner. A contemporary semantic image segmentation model is utilized for combined training of multi-gauge store figures besides observation copy. The suggested copy not merely outclasses median and better merging, however, permits for diagnostically visualizing features significance at diverse positions and scales. As well, with extra supervision to output at every scale is crucial for attaining outstanding performance after multi-scale features integration. Three interesting datasets, containing PASCAL-Person-Part, PASCAL VOC 2012 and a subset of MS-COCO 2014 are utilized for demonstrating model efficacy by carrying out extensive experiments.

He et al [13] suggested a residual learning basis for network training in an easy way which is significantly deeper than those used earlier. The layers reformulation is done plainly as learning residual functions with reference to layer inputs, apart from learning unreferenced functions. Comprehensive empirical substantiation infers that residual networks are at ease for optimization, and accuracy can be achieved from extensively expanded intensity. On the figure lace meta data, remaining lace evaluation is done with to 152 coating - 8× depth greater than VGG nets however yet

possesses lesser complexity. These remaining lace ensemble attains 3.57% mistake on figure lace test set in addition attained 1st place on ILSVRC 2015 grading mission.

Lin et al [14] utilized Conditional Random Fields (CRFs) with CNN-based pair wise potential functions for capturing semantic correlations amid neighboring patches. The suggested deep structured model effectual piecewise training is exploited for avoiding recurrent costly CRF deduction for back generation. The patch-environment scene capturing is done revealing that web plan with classic multi-gauge figure store and gliding monolith dividing is effectual for execution enhancement. The experimentation is carried out for various prevalent linguistic dissection metadata, comprising NYUDv2, PASCAL VOC 2012, PASCAL-Scene, and SIFT-flow. Above all, a convergence-re uniting outcome of 78.0 is attained on inspiring PASCAL VOC 2012 metadata.

Liu et al [15] suggested a methodology for enabling deterministic end-to-end computation in a single forward pass by technique named Convolutional Neural Network (CNN), specifically Deep Parsing Network (DPN)1. DPN is extended to contemporary CNN architecture particularly for modelling unary terms along with additional layers devising for mean field algorithm (MF) approximation of pair wise terms which holds various appealing properties. Initially, it is quite dissimilar that existing mechanisms that joined CNN and MRF, where MF several iterations were necessitated for every training image in the course of back-propagation, DPN might attain great performance through approximating MF one iteration. Second, DPN signifies different sorts of pairwise terms, creating several prevailing works as its distinctive circumstances. Third, DPN creates MF r to be parallelized easily and speeded up in Graphical Processing Unit (GPU). DPN thorough assessment is done on PASCAL VOC 2012 dataset, where a single DPN model produces better high-tech segmentation accuracy of 77.5%.

Ghiasi et al [16] offered two contributions: First one is that low apparent convolutional feature maps spatial resolution , high-dimensional feature depiction comprising substantial sub-pixel localization information. (2) A multi-resolution reconstruction architecture is developed based on Laplacian pyramid deploying skip connections from greater resolution feature maps and multiplicative gating to consecutively enhance segment boundaries rebuilt from lower-resolution maps. This methodology produces contemporary semantic segmentation outcomes on PASCAL VOC and Cityscapes segmentation benchmarks deprived of resorting to further complex random-field inference or instance detection driven architectures.

Long et al [17] suggested a system comprising arbitrary size input besides correspondingly-sized output with proficient inference along with learning achieved by fully convolutional" networks. The fully convolutional networks space is defined with their utilization to geographically thick forecast mission and bring out relation to earlier figures. The present grading system (Alex Net, VGG net, and Google Net) is adapted into totally complication system for transferring their accomplished depictions via modification to segmentation task. And then cut planning that merges linguistic data from an extensive, rough coating with display data from surface is suggested along with fine layer for producing precise and exhaustive segmentations.

Mostajabi et al [18] suggested semantic segmentation by presenting only feed-forward architecture. The small image elements (superpixels) mapping is done into rich feature representations extracted from nested region sequence of increasing extent. These regions are attained via "zooming out" from superpixel to scene-level resolution. The image statistical structure and label space are greatly utilized short of setting up explicit structured prediction mechanisms, hence complex as well as expensive inference are avoided. As an alternative superpixels classification are done by feedforward multilayer network which attains 69.6% average accuracy on PASCAL VOC 2012 test set.

Dong et al [19] utilized unsupervised segmentation and supervised segmentation for establishing a segmentation system. There exist two-grade methodology, i.e., intensity deduction besides intensity gathering which is exploited for unsupervised segmentation. Image color projection is done into a small prototypes set by automatic-assembling chart (AAC) teaching in intensity reduction. The resembled heating (RH) searches classical groups from AAC frameworks in color clustering. This two-grade

methodology is benefitted by SOM and SA, thereby achieving near-classical distribution with cheap mathematical price. The learning as well as pixel classification mainly comprised in supervised segmentation. Color prototype is well-defined for curved part representation in intensity area. A Graded Framework Training (GFT) is utilized for generating diverse color prototype sizes from device intensity sample. These intensity frameworks offer better evaluate for device intensity.

Wang et al [20] utilized component wise base angle unit (BAU) allocation for color image segmentation. Initially, image component-grade color feature and texture feature is greatly utilized as BAU model (grader) input which are abstracted through confined analogy type and Gabor strainer. Here, SVM model (grader) training is achieved through FCM with uprooted component-grade quality. Lastly, intensity figure segmentation is attained with directed BAU type (grader). The complete benefit is not considered for image segmentation apart from local information of color image benefit, along with BAU classifier proficiency. The suggested approach is considered to offer very effective segmentation outcomes besides computational behavior, decreased time and intensity figure segmentation eminence is increased when compared with classic segmentation methods.

Jafari et al [21] suggested an approach for cross-modality influence (CMI) quantification. Also, relationship amid final accuracy and cross-modality effect is validated, though not a simple linear one. The larger cross-modality impact need not essentially helps in translation into an enhanced accuracy. In addition, beneficial balance amid cross-modality effects might be attained through network architecture and conjecture besides the relationship can be exploited for understanding diverse network design ranges. A Convolutional Neural Network (CNN) architecture is suggested which integrates contemporary outcomes for depth estimation and semantic labeling. Also, enhanced outcomes can also be attained via balancing cross-modality influences amid depth and semantic prediction for both tasks through NYU-Depth v2 benchmark.

Mousavian et al [22] suggested a novel model for concurrent depth estimation and semantic segmentation from distinct RGB image. The model training parts feasibility can be demonstrated for every task and fine tuning complete, combined model on both tasks at the same time by single loss function. Likewise, deep CNN is coupled with fully connected CRF for capturing contextual interfaces and interactions amid semantic and depth cues which helps in final outcomes accuracy enhancement. The suggested model training and evaluation is done on NYU Depth V2 dataset which outclasses the conventional approach in terms of outcomes on depth estimation task.

There are seven major classes of image segmentation methods: region-based, boundary-based, edge detection-based, template matching, cluster-based, threshold-based, graph-based, besides stated classes combination. Image segmentation possesses considerable influence on overall algorithmic performance. Nevertheless, these methods need not ensure regions separation with similar intensities, but be appropriate to diverse regions. An efficient image segmentation method is chiefly focused in color-texture natural images and performance measurements in a proficient way.

3. Proposed Methodology

Image segmentation is regarded as a series process for fragmenting visual input into segments for image analysis simplification. A segment generally signifies objects or parts of objects, and comprises pixels sets, or “super-pixels”. Image segmentation categorizes pixels into higher components, abolishing necessity for individual pixels consideration as units of observation. The research work intends at combining machine learning and deep learning approaches and thereby yielding semantically precise predictions and exhaustive segmentation along with attaining computationally effectual way. Image color segmentation is achieved by machine learning besides semantic labeling through deep learning. Two algorithms are utilized in this process.

- First algorithm deals with detecting super pixels using Modified Support Vector Machine (MSVM) based machine learning and these super pixels segmentation is done depending on textures and colors. Image color feature and texture feature serves as SVM design (variety) input, which are obtained by internal geographical correlation plan design. Lastly, color image segmentation is attained with trained SVM model (classifier).

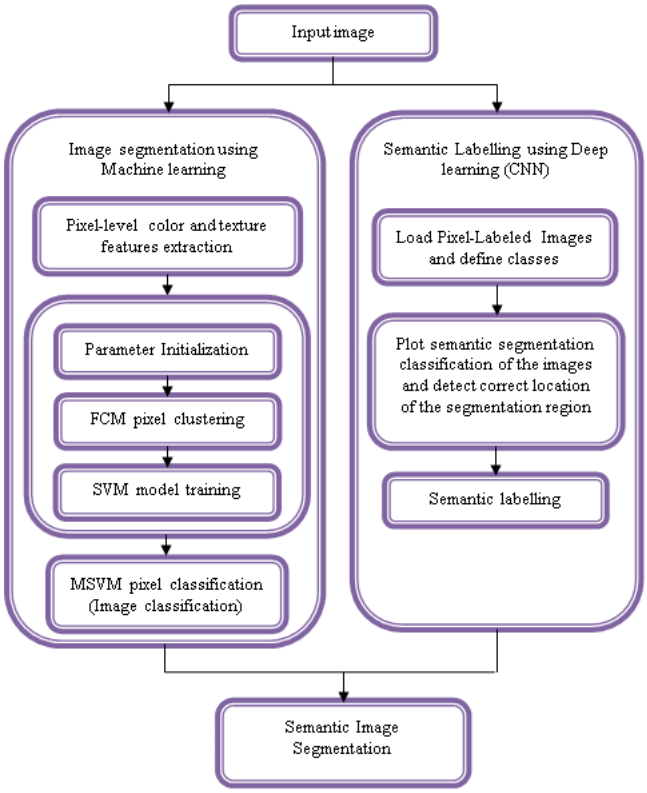

Fig. 1. The overall process of the proposed semantic image segmentation technique.

- The second algorithm utilizes convolutional Neural Network (CNN) for training color categories and thus achieving object classification into semantic labels. Deep Convolution Neural Network (DCNN) is greatly utilized for this classification model. Multiple layers are also deployed for building an enhanced feature space. The 1st order features (e.g. edge) is learnt by first layer and Second layer helps in learning higher order features (edges combinations etc.). Deep Convolution Neural Network (DCNN) offer outcomes in color segmentation form.

3.1. component-level color quality representation

Throughout this research, every image pixel detection is done from a related sector that belongs to a body/segment of an object. Classification process defines sculpture analysis issue, and the segmentation tends to designate a label for a pixel or a region individually. Hence, the extraction of efficient pixel-level image feature is considered being crucial. At this point, the pixel level color feature extraction is performed through local spatial similarity measure system. During describing the image, color plays a vital role, since it is highly ascendant and discernible base-level visual feature. In accordance with local spatial similarity measure model, a novel component-level color quality is introduced in this segment.

Consider X_i= (X_i^R ,X_i^G ,X_i^B) as RGB components of the ith pixel. Then, the computation of pixel-level color feature 〖CF〗_i^k (k = R,G,B) of component X_i^k (k = R,G,B) might be defined as follows, in which every color component might be considered as a gray scale image.

3.2. Component-level color quality representation

During figure segmentation process, Consistency is taken as a common feature. It is frequently used in combination with color briefing that helps to attain optimal distribution outcomes than feasible with fair intensity lone. Here, Steerable filter is applied to the image for procuring the pixel-level texture feature and extracting filter responses local energy , termed as pixel texture feature. During most of these processes, the application of capricious inclination filters below maladaptive power, besides process yield scrutiny as a purpose of both intention and period predominantly rely on it. At many orientations, the response of the filter can be obtained by applying various forms of the same filter, which differ from each other by getting minor rotation in angle. As a class of filters, the Steerable filter synthesizes a arbitrary orientation filter as a linear combination of a ‘‘basis filters’’ set. The image is divided into direction span range found through fundamental strainers that possesses the different orientations in order to identify the edges situated at these orientations in an image. Consequently, a filter might be adaptively “steered” towards any orientation, besides the output of filter can be analytically determined as an orientation function.

The equation for the steering constraint is given by,

$$F_\theta(m,n) = \sum_{k=1}^{N} b_k(\theta) A_k(m,n) \quad (1)$$

Here, in accordance with the arbitrary orientation θ, the interpolation function is denoted by b_k (θ) that drives the orientations of filter. The basis filters A_k (m,n) represent the rotated versions of the impulse response at θ.

$$I(m,n) * F_\theta(m,n) = \sum_{k=1}^{N} b_k(\theta)(I(m,n) * A_k(m,n) \quad (2)$$

In which, the basis filters A_k (m,n) represent the impulse response rotated versions at θ. The following steps

define the strategy developed for pixel-level texture feature extraction.

Step 1: Color space transformation.

From RGB color space, the color image I is converted to YCbCr color space, where Y denotes light element, whereas Cb and Cr indicate saturation elements.

Step 2: Application of dirigible strainer to light element Y.

In this phase, Dirigible strainer decay is applied with four inclination sub-bands. Here, determined that a one-level decomposition is enough since the images are solidly small in size. Among those, solely the four orientation bands are utilized. Identification of regions that encompasses the dominant orientation (0°, 45°, 90°, 135°) is the core objective, besides it is four coefficients supreme , through which orientation is determined at a provided pixel location.

Step 3: component texture quality extraction.

Consider S_1 (x,y),S_2 (x,y),S_3 (x,y) and S_4 (x,y) as the Dirigible sub band factor at position (x, y) corresponding to 0°, 45°, 90°, and 135° directions, correspondingly. In that, the four coefficients maximum (in actual value) at location (x, y) is signified by TF(x, y), and it is component quality feature at position (x, y)

3.3. Component-based intensity figure analysis using the SVM and FCM

Image segmentation is known as the most common classification issue, and it is possible to resolve it through all the existing classification methods. Recently introduced an innovative machine learning approach, namely Modified Support Vector Machine (MSVM). Since the efficiency of SVM based classification not directly relies on classified entities dimension , it is considered being an impressive method. Hence, the SVM can be employed to deal with the challenges of color image segmentation. Throughout this study, image pixel-level color feature and texture feature are fed into SVM framework (classifier) as an input, and the local spatial similarity measure system is used to extract them.

Conversely, in the context of target classes overlapping, SVM turns out to be inefficient [23]. Moreover, the SVM lacks its performance, when the feature count for each data point goes beyond the quantity of training data samples. For avoiding the overlapping, the SVM framework (classifier) is trained through FCM accompanied by extracted pixel-level features.

A. *Pixel-level color and texture feature extraction.*

The local spatial similarity measure model besides Steerable filter are used for extracting the features of pixel-level color and texture in above section.

B. *FCM-based SVM training sample selection.*

Training sample selection plays a vital role in determining the generalization level of SVM classification rules for an unknown sample. As represented by the earlier works, this factor can be considered as significant as regards accurate classifications, when compared to the selection of classification algorithms. There are several methods to select the training pixels in terms of SVM-based image segmentation. Whereas, for identifying and labelling the small patches of similar pixels within an image, the widespread sampling method is utilized. Nevertheless, adjacent pixels need to be correlated spatially/possess identical values. Thus, the spectral variability of each class may get underestimated and classification may get degraded if the process of training samples collection follows this method. So that the random sampling can be taken as the simple approach for minimizing the impact of spatial correlation. Two types of random sampling methods are utilized, namely equal sample rate (ESR), and equal sample size (ESS). In ESR, certain percentage of pixels are assigned as training data, which are arbitrarily sampled from each class. Whereas, in ESS, a fixed number of samples are assigned as training data, which are arbitrarily sampled from each class. In this research, presented a FCM based feature extraction technique for evading the overlapping issue in SVM.

As an unsupervised method, Fuzzy c-means (FCM) clustering has been implemented to feature analysis successfully. A image representation can be done in different feature spaces, besides by combining identical data points in feature space into clusters, the FCM algorithm can accomplish the classification of image [24]. To attain this clustering, a cost function that relies on pixels distance to cluster centres in feature domain has been minimized iteratively. During which the FCM clustering algorithm further involved in selecting the training samples to perform with SVM classifiers.

C. *SVM model training.*

The training set generated in the preceding phase is exploited for training the SVM model (classifier).

D. *SVM pixel classification.*

The trained SVM model (classifier) is used for predicting remaining image pixels class labels (test samples). The training set (class labels provided via FCM clustering) and test set (class labels trough SVM) are merged together for acquiring the entire label vector, besides for returning it as a clustering solution, i.e. the image segmentation outcomes.

3.4. Semantic Labeling using Convolutional Neural Networks (CNN)

Every image pixel is categorized into meaningful classes of objective with the help of Semantic segmentation. In accordance with the learning levels of representations, deep learning is regarded as the admirable branch of machine learning. Whereas, CNN is a type of deep neural network. The suggested gives a detained analysis of CNN algorithm process comprising of both forward process and back propagation. The highest speed up and parallel effectiveness are theoretically assessed through estimating forward and backward computing actual time.

4. Result and Discussion

The proposed semantic image segmentation performance is evaluated in accordance with the detailed experiments performed in the images of natural scenery. Several performance parameters have been measured, for which the rates of true positive (TP), false positive (FP), true negative (TN), and false negative (FN) have calculated initially. Accordingly, Precision was taken as the first metric, which is rate of rescued cases that were applicable. Further, Revocation is another metric that indicates the ratio of relevant instances that were recovered. Though the accuracy and revocation measures are contradicted with each other, yet both of them prove its significance in the evaluation of prediction system performance. So, in order to derive as a single metric, these two measures can be unified with equal weights, namely F-measure. Finally, accuracy is considered as an ultimate metric that signifies the ratio of appropriately predicted instances from the overall instances predicted.

Precision: This measure refers to as the proportion between appropriate positive observations obtained and anticipated beneficial survey.

$$\text{Precision (P)} = TP/TP+FP \qquad (3)$$

Sensitivity or Revocation: It defines the proportion between appropriately recognized positive observation and the total observation.

$$\text{Recall (R)} = TP/TP+FN \qquad (4)$$

F-measure: It refers to as Precision and Recall weighted average. Therefore, it considers both false positives and false negatives.

$$\text{F1 Score} = 2*(R * P) / (R + P) \qquad (5)$$

Accuracy: This metric is estimated on the basis of positives and negatives that is specified by,

$$\text{Accuracy} = (TP+FP)/(TP+TN+FP+FN) \qquad (6)$$

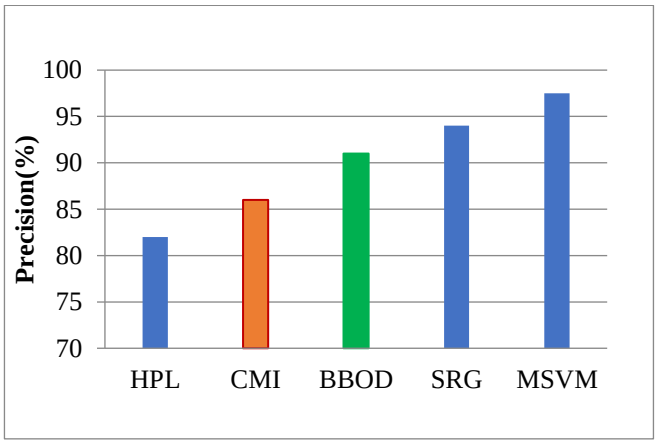

Fig.2. Accuracy observation outcomes between the advanced and actual methods for Linguistic Figure Dissection

In Fig. 2, the accuracy outcomes for the Semantic Image Segmentation process obtained by proposed and prevailing approaches using proposed MSVM and CNN based method are compared. In the figure, the graphs represent the efficiency of the proposed MSVM and CNN method to deliver comparatively greater precision values than the prevailing segmentation methods.

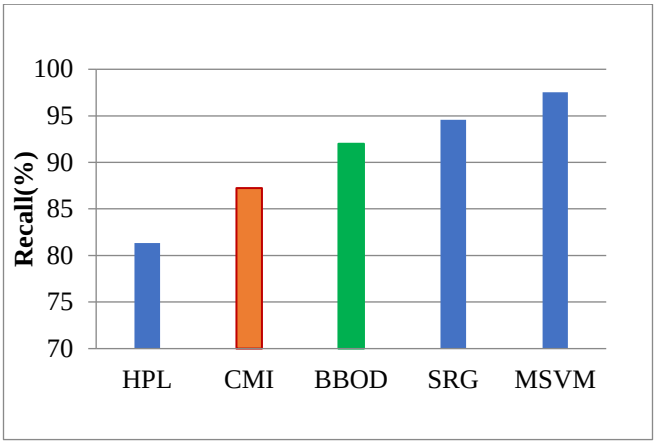

Fig.3. Recall observation outcomes between the advanced and actual methods for Linguistic figure Dissection

Fig. 3 compares the Recall rates for the Semantic Image Segmentation process attained by proposed and existing approaches using proposed MSVM and CNN based method. The graphs represent the proficiency of the proposed MSVM and CNN method to provide comparatively higher Recall rates than the prevailing segmentation methods.

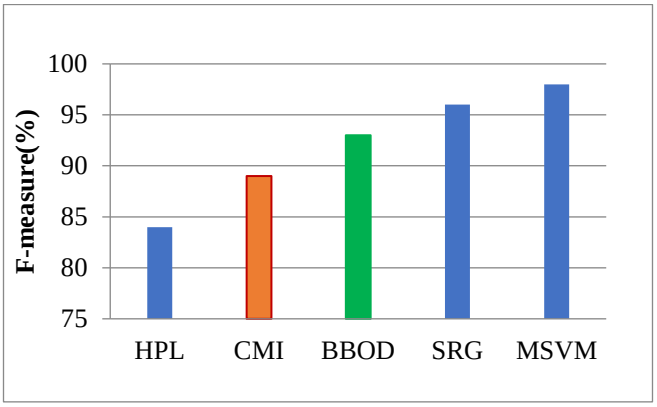

Fig.4. F-measure observation outcomes between the advanced and actual methods for Linguistic Figure Dissection

Fig. 4 compares the F-measure rates obtained by proposed and existing approaches using proposed MSVM and CNN based method for the Semantic Image Segmentation process. The graphs epitomize that the proposed MSVM and CNN method capable of providing comparatively higher F-measure rates than the prevailing segmentation methods.

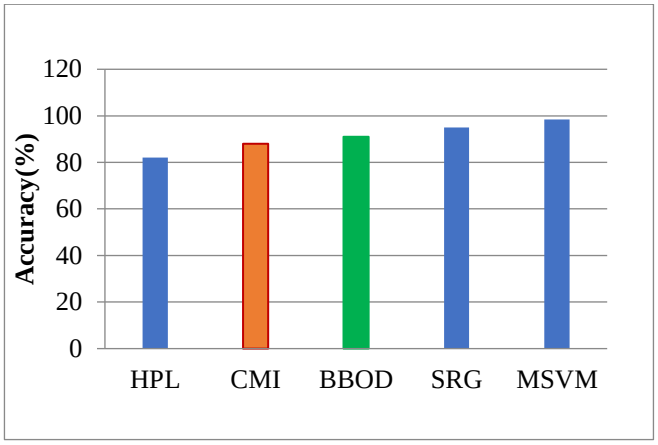

Fig.5. Exactness observation outcomes between the advanced and actual methods for Linguistic Figure Dissection

In Fig. 5, the Exactness rates obtained by proposed and existing approaches using proposed MSVM and CNN based method for the Semantic Image Segmentation process have been compared. The graphs indicate that suggested MSVM and CNN method possesses capability for achieving maximum Accuracy which is superior to the prevailing segmentation methods.

5. Conclusion

During this study, machine learning and deep learning have been deployed for processing the semantic image segmentation. Besides, based on MSVM, a novel type is presented for color texture image segmentation. Initially, local geographical analogy standard model is used image component-grade intensity quality and consistency quality production that have been fed into the SVM framework (classifier) as an input. Subsequently, the FCM accompanied by the extracted pixel-feature is involved to train the SVM system (classifier). In this classification framework, proposed a Deep Convolution Neural Network (DCNN) that includes multiple layers that builds an enhanced feature space. By means of this deep convolution neural network, a solution is provided in pattern of semantic labelling. Grounded the outcomes attained on the segmentation database, empirical findings depict suggested algorithm efficacy for obtaining quantitative outcomes that is superior to other conventional segmentation techniques. Hence, it can be concluded that suggested semantic image segmentation proposal is proficient to give maximum exactness more over optimal segmentation outcomes.

REFERENCES

[1] Guo, Y., Liu, Y., Georgiou, T., & Lew, M. S. (2018). A review of semantic segmentation using deep neural networks. International journal of multimedia information retrieval, 7(2), 87-93.

[2] Tran, T., Kwon, O. H., Kwon, K. R., Lee, S. H., & Kang, K. W. (2018, December). Blood cell images segmentation using deep learning semantic segmentation. In 2018 IEEE International Conference on Electronics and Communication Engineering (ICECE) (pp. 13-16). IEEE.

[3] Kemker, R., Salvaggio, C., & Kanan, C. (2018). Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. ISPRS journal of photogrammetry and remote sensing, 145, 60-77.

[4] Al-Kafri, A. S., Sudirman, S., Hussain, A., Al-Jumeily, D., Natalia, F., Meidia, H., ... & Al-Jumaily, M. (2019). Boundary delineation of MRI images for lumbar spinal stenosis detection through semantic segmentation using deep neural networks. IEEE Access, 7, 43487-43501.

[5] Zhu, Y., Sapra, K., Reda, F. A., Shih, K. J., Newsam, S., Tao, A., & Catanzaro, B. (2019). Improving semantic segmentation via video propagation and label relaxation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 8856-8865).

[6] Zlateski, A., Jaroensri, R., Sharma, P., & Durand, F. (2018). On the importance of label quality for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 1479-1487).

[7] Zhang, H., Zhang, H., Wang, C., & Xie, J. (2019). Co-occurrent features in semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 548-557).

[8] Noh, H., Hong, S., & Han, B. (2015). Learning deconvolution network for semantic segmentation. In Proceedings of the IEEE international conference on computer vision (pp. 1520-1528).

[9] Lu, Z., Fu, Z., Xiang, T., Han, P., Wang, L., & Gao, X. (2016). Learning from weak and noisy labels for semantic segmentation. IEEE transactions on pattern analysis and machine intelligence, 39(3), 486-500.

[10] Li, A., Lu, Z., Wang, L., Han, P., & Wen, J. R. (2016). Large-scale sparse learning from noisy tags for semantic segmentation. IEEE transactions on cybernetics, 48(1), 253-263.

[11] Liu, G., Han, P., Niu, Y., Yuan, W., Lu, Z., & Wen, J. R. (2017, May). Graph-boosted convolutional neural networks for semantic segmentation. In 2017 International Joint Conference on Neural Networks (IJCNN) (pp. 612-618). IEEE.

[12] Chen, L. C., Yang, Y., Wang, J., Xu, W., & Yuille, A. L. (2016). Attention to scale: Scale-aware semantic image segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3640-3649).

[13] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[14] Lin, G., Shen, C., Van Den Hengel, A., & Reid, I. (2016). Efficient piecewise training of deep structured models for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3194-3203).

[15] Liu, Z., Li, X., Luo, P., Loy, C. C., & Tang, X. (2015). Semantic image segmentation via deep parsing network. In Proceedings of the IEEE international conference on computer vision (pp. 1377-1385).

[16] Ghiasi, G., & Fowlkes, C. C. (2016, October). Laplacian pyramid reconstruction and refinement for semantic segmentation. In European conference on computer vision (pp. 519-534). Springer, Cham.

[17] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3431-3440).

[18] Mostajabi, M., Yadollahpour, P., & Shakhnarovich G. (2015). Feedforward semantic segmentation with zoom-out features. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3376-3385).

[19] Dong, G., & Xie, M. (2005). Color clustering and learning for image segmentation based on neural

networks. IEEE transactions on neural networks, 16(4), 925-936.

[20] Wang, X. Y., Wang, T., & Bu, J. (2011). Color image segmentation using pixel wise support vector machine classification. Pattern Recognition, 44(4), 777-787.

[21] Jafari, O. H., Groth, O., Kirillov, A., Yang, M. Y., & Rother, C. (2017, May). Analyzing modular CNN architectures for joint depth prediction and semantic segmentation. In 2017 IEEE International Conference on Robotics and Automation (ICRA) (pp. 4620-4627). IEEE.

[22] Mousavian, A., Pirsiavash, H., & Košecká, J. (2016, October). Joint semantic segmentation and depth estimation with deep convolutional networks. In 2016 Fourth International Conference on 3D Vision (3DV) (pp. 611-619). IEEE.

[23] Vishwanathan, S. V. M., & Murty, M. N. (2002, May). SSVM: a simple SVM algorithm. In Proceedings of the 2002 International Joint Conference on Neural Networks. IJCNN'02 (Cat. No. 02CH37290) (Vol. 3, pp. 2393-2398). IEEE.

[24] Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. Computers & Geosciences, 10(2-3), 191-203.

[25] Wu, J. (2017). Introduction to convolutional neural networks. National Key Lab for Novel Software Technology. Nanjing University. China, 5, 23