Image Segmentation and Object Identification using Deep Learning

by Icicta- 221

Image Segmentation and Object Identification using Deep Learning

1st Prince Kumar

Master of Computer Application

Department of Computer Science and Applications

Reva University

Bangalore, India

princesinha2018@gmail.com

2nd Nagaraju S

Assistant Professor, School of Computer Science and

Applications

Reva University

Bangalore, India

nagaraju.sms@gmail.com

Abstract – **Often based on the properties of the picture's pixels, image segmentation is a widely used method in digital image processing and analysis to divide an image into various portions or areas.** In order to simplify the picture and make each segment more processable or sustainable to analysis, image segmentation is a technique for dividing a digital image into smaller groups known as image segments.

The first stage in image analysis and data extraction is picture segmentation. Image identification or image recognition refers to the task of identifying what is present in an image. This can include recognizing specific objects, scenes, or patterns within an image. Algorithms for computer vision can utilize them to swiftly identify them. Computer vision algorithms may be used to extract elements like color, texture, and form in order to recognize abstract art images. These are characteristics may then be used to group photographs into subcategories depending on how closely they resemble other pictures with the same characteristics.

Keywords – *FCN, U-Net, Mask R-CNN, RNN, DeepLab, SegNet, CNN, SVM, KNN, DT, DBN, RF, Edge-based, Threshold-based, Region-based, Cluster-based, Graph-based, Watershed-based and sematic segmentation, instance segmentation, panoptic segmentation, deep learning, pattern recognition.*

I. INTRODUCTION

An image is split into parts or regions, each of which represents a different item or element of the picture, using a process called image segmentation in computer vision. Image segmentation aims to make an image's representation more straightforward and/or transform it into something more relevant and understandable. Image segmentation may be used for a variety of tasks, including object recognition, tracking, picture reduction and image editing. The technique of categorizing or recognizing an image based on it's content is known as picture identification. Utilizing computer vision techniques, this requires studying the image and determining it's features, such as shapes, colors, and textures. In uses such as pictures identification, facial recognition, and search.

TABLE I. ALGORITHMS OF IMAGE SEGMENTATION

SI. No.	NAME OF ALGORITHM
1.	Thresholding-Based Segmentation.
2.	Region Growing-Based Segmentation.
3.	Watershed Transform-Based Segmentation.
4.	Graph-Based Segmentation.
5.	Cluster-Based Segmentation.
6.	Mean-Shift-Based Segmentation.
7.	Active Counters-Based Segmentation.
8.	Edge-Based Segmentation.

TABLE II. ALGORITHMS OF IMAGE IDENTIFICATION

SI. No.	NAME OF ALGORITHM
1.	Convolutional Neural Networks. [CNNs]
2.	Support Vector Machines. [SVMs]
3.	K-Nearest Neighbors. [KNNs]
4.	Decision Trees. [DTs]
5.	Deep Belief Networks. [DBNs]
6.	Random Forests. [RFs]

II.   METHODOLOGY

Methodologies used in Image Segmentation –

- Fully convolutional neural networks. [FCN]
- Convolutional networks for biomedical image segmentation. [U-Net]
- Regions with convolutional neural networks. [Mask R-CNN]
- Deep labeling for semantic image segmentation. [DeepLab]
- Semantic segmentation model. [SegNet]

Methodologies used in Image Identification –

- Convolutional neural networks. [CNN]
- Transfer Learning. [TL]
- Recurrent neural networks. [RNN]
- Siamese networks. [SN]
- Generative adversarial networks. [GAN]

1) **Methodologies used in Image Segmentation –**

FCN –
- The main use of this architectural type is semantic segmentation.
- A neural network that exclusively uses convolution, pooling, and subsampling or up-sampling operations is referred to as a completely convolution network [FCN].
- An FCN is just a CNN with fewer completely linked layers.

U-Net –
- The convolutional neural network known as U-Net was created for the segmentation of biological images.
- An encoder-decoder design with skip connections makes up the U-Net FCN type.
- The encoder collects features from the input picture, and the decoder generates a segmentation map by up sampling the output.
- The skip connections aid in combining data and increase segmentation accuracy.

Mask R-CNN –
- For pixel-level instance segmentation, the Faster R-CNN object detection model has been updated into the Mask R-CNN variation.
- Each object proposal is given a segmentation mask after being generated by an area proposal network.
- Convolutional neural network characteristics and rectangular region suggestions are combined in Mask R-CNN.
- A two-stage detection technique is Mask R-CNN.
- A subset of areas in an image that potentially contain an item are found in the first step.

DeepLab –
- The cutting-edge semantic segmentation model DeepLab was created and released by Google.
- Modern deep learning or machine learning models may be found at DeepLab.
- The dense prediction is obtained by simply up-sampling the output of the last convolution layer and calculating pixel-wise loss.
- In order to extract multi-scale contextual information from photos, DeepLab is a deep learning architecture that employs several convolutional layers.
- It may be applied to instance segmentation and semantic segmentation.

SegNet –
- SegNet is a deep learning architecture that creates excellent segmentation maps using an encoder-decoder network with skip links.
- It is renowned for having a little computational overhead and performing effectively with little data.
- Image segmentation using a deep convolutional encoder-decoder architecture.
- A semantic segmentation model is SegNet.
- The decoder up-samples the output using pooling indices from the decoder while the encoder collects features from the input picture.

2) **Methodologies used in Image Identification –**

Convolutional neural networks [CNN] – A well-liked deep learning architecture for image recognition

is CNN. They gain the ability to identify features in an images and categorize them according to their visual qualities.

Transfer learning [TLs] – Using pre-trained models that have been trained on huge datasets, like ImageNet, is a technique called transfer learning. On smaller datasets of photos, these pre-trained models may then be adjusted to recognize certain items or patterns.

Recurrent neural networks [RNNs] – Another deep learning architecture that may be applied for image recognition is RNNs. They are very helpful for seeing patterns in succession, such in video data.

Siamese networks [SNs] – A specific kind of neural network known as a siamese network may be trained to compare two photos and decide whether they are similar or dissimilar. They may be used for things like facial recognition and picture similarity searches.

Generative adversarial networks [GANs] – A GAN, a specific type of neural network, can generate new images from a collection of input photos. They can be used for activities like image creation and style elaboration.
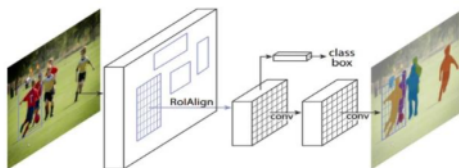
III. ADVANTAGE
ADVANTAGES OF IMAGE SEGMENTATION –
- Accuracy.
- Speed.
- Adaptability.
- Generalization.
- Automation.

ADVANTAGES OF IMAGE IDENTIFICATION –
- High accuracy.
- Generalization.
- Adaptability.
- Automation.
- Scalability.

Fig 1 : Mask R-CNN Algorithms.

Mask R-CNN Algorithms – Regions with convolutional neural network [R-CNN], a machine learning or deep learning strategy, combines rectangular area suggestions with convolutional neural network attributes. A machine learning or deep learning approaches for picture segmentation issues is called Mask R-CNN. By providing each picture's bounding box, classes, and associated binary image mask, it can distinguish between the many images inside an image or video. Faster R-CNN was used to build Mask R-CNN. For each candidate, F-RCNN provides two outputs : a class label and a bounding box. The model also gains a third branch that produces the object mask. The third branch functions in tandem with the current branch to recognize boundary boxes.

- Create a large number of potential regions for the initial sub-segmentation.
- Combines smaller sections using a greedy technique.
- Create the final candidate region proposals using the produced region.

IV. ARCHITECTURE
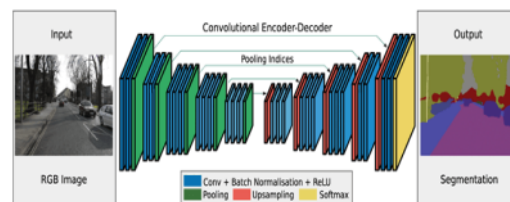ARCHITECTURE OF IMAGE SEGMENTATION –

Fig 2 : Architecture of Image Segmentation.

A common deep learning application in computer vision is image segmentation, which involves breaking an image up into several areas or segments that share comparable properties. The architecture for deep learning-based picture segmentation generally entails the following steps –

- Data Preprocessing.
- Convolutional Neural Network.
- Encoder-Decoder.
- Contracting Path.
- Expanding Path.
- Skip Connections.
- Activation Functions.
- Loss Function and Optimization.
- Post-Processing.
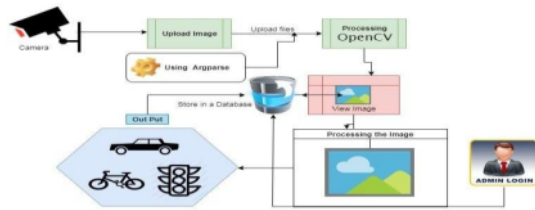
ARCHITECTURE OF IMAGE IDENTIFICATION –

Fig 3 : Architecture of Image Identification.

One common use of deep learning particularly in computer vision, is object detection or identification. Following are typically phases in the architecture of object recognition using deep learning –

- Data Preprocessing.
- Convolutional Neural Network.
- Convolutional Layers.
- Pooling Layers.
- Fully Connected Layers.
- Activation Function.
- Loss Function and Optimization.
- Transfer Learning.

V. IMPLEMENTATION
IMPLEMENTATION OF IMAGE SEGMENTATION

- **Data Gathering and Data Preparation** – Gather a sizable dataset of segmented photos and get it ready for model training. The photos may need to be resized, cropped, and normalized in addition to being labelled with the appropriate segmented parts.
- **Model Selection** – Select a model architecture that is suitable for jobs involving picture segmentation. Fully convolutional networks [FCNs], U-Net, and Mask R-CNN are some prominent deep learning models for image segmentation, whereas conditional random fields [CRFs] and graph-cut based techniques are well-liked machine learning models.
- **Model Training** – Utilize the provided dataset to train the chosen model. In order to do this, the data must be divided into training and validation sets. The training set is then used to modify the model's parameters in order to minimize the desired loss function.

This could need a substantial investment of time and computing power.

- **Hyperparameter Tuning** – Improve the model's performance by fine-tuning the hyperparameters. This might entail altering the regularization parameters, batch size, or learning rate.
- **Model Evaluation** – Assess the performance of the trained model with a different testing dataset. Metrics like intersection over union [IoU], dice coefficient, or pixel precision may need to be calculated in order to do this.
- **Post-Processing** – The resultant segmentation map should be posted-processed to enhance its quality or get rid of any artefacts. To eliminate noise, this may include using morphological processes like erosion and dilation or thresholding.
- **Deployment** – Utilize the trained model for problems involving picture segmentation in practical contexts. The might entail optimizing the model's performance for inference and integrating it with other software systems, such as online applications or mobile apps.

IMPLEMENTATION OF IMAGEI DENTIFICATION

- **Data Collection and Data Preparing** – Gather a sizable picture collection and get it ready for model training. The photos may need to be resized, cropped, and normalized in addition to being assigned to the appropriate classes.
- **Model Selection** – Pick a model architecture that can handle tasks involving picture identification. Convolutional neural network [CNNs], like VGG, ResNet, and Inception are some prominent deep learning models for image recognition, and SVMs Random Forests, K-Nearest Neighbors [KNNs] are common machine learning models.
- **Training the Model** – Utilize the provided dataset to train the chosen model. In order to do this, the data must be divided into training and validation sets. The training set is then used to modify the model's parameters in order to minimize the desired loss function.

This could need a substantial investment of time and computing power.

- **Fine-Tuning the Model** – Improve the model's performance by fine-tuning the hyperparameters. This might entail altering the regularization parameters, batch size, or learning rate.
- **Evaluation** – Assess the performance of the trained model with a different testing dataset. Calculating measures of the like accuracy, precision, recall, and F1 score may be necessary for this.
- **Deployment** – Utilize the trained model for tasks involving picture identification in practical contexts. This might entail optimizing the model's performance for inference and integrating it with other software systems, such as online applications or mobile apps.

VI.   RESULT

An image that has been segmented using deep learning and machine learning is one in which the original image has been separated into several areas or segments depending on the objects or features seen in the image. CNNs and encoder-decoder architectures like U-Net, FCNs, Mask R-CNNs, model has attained cutting-edge performance in instance segmentation tasks, and the VGG and ResNet models have excelled in picture classification tasks on sizable benchmark datasets like ImageNet are often used in machine learning and deep learning models for image segmentation. These models are developed using sizable datasets of annotated pictures, where each pixel is assigned a class or segmentation mask that corresponds to it.

Image Segmentation – The DeepLab V3+ model outer performed humans with an overall mean in intersection over union [mIoU] score of 82.1% in the pascal VOC 2012 segmentation task. The U-Net model has long history of success in biomedical picture segmentation tasks and has demonstrated excellent accuracy on a number of datasets, including the ISBI cell tracking challenge and the Kaggle Data science Science Bowl. With a mAP of 39.8% for object recognition and 36.7% for instance segmentation tasks on the COCO dataset.

Image Identification – In 2014, the VGG16 model outer performed the current state-of-the-art accuracy of 88.7% by achieving top-5 test accuracy of 92.0% on the ImageNet Large Scale Visual Recognition Challenge dataset. [ILSVRC] on the ILSVRC 2015 dataset, the ResNet-152 model attained a top-5 error rate of 3.5% which was much lower than the winning model from the previous year. The inception V4 model outer performed all earlier models in the competition, achieving top-1 accuracy of 80.2% on the ILSVRC 2017 dataset.
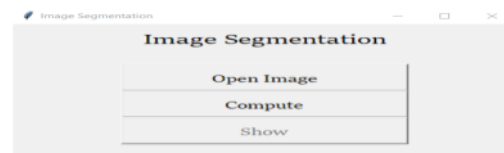

Fig 4 : User Interface.

Image Segmentation –

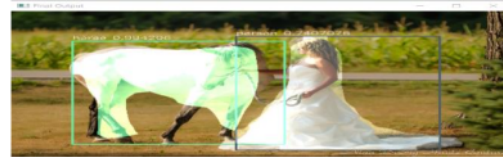Fig 5 : Final Image, Bus, Cars, Person.

Fig 6 : Final Image, Horse, Person.

Fig 7 : Final Image, Birds.

Image Identification –
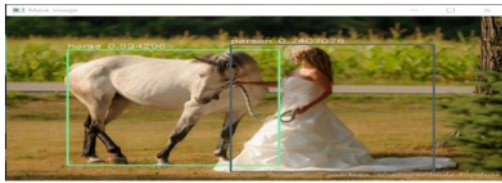

Fig 8 : Mask Image, Bus, Cars, Person.
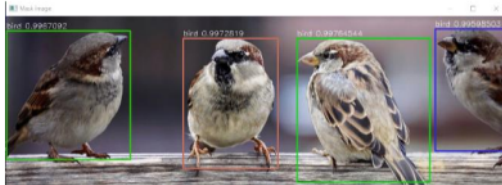
Fig 9 : Mask Image, Horse, Person.

Fig 10 : Mask Image, Birds.

CONCLUSION

In conclusion, deep learning has significantly improved performance in the critical computer vision tasks of image segmentation and image identification. CNNs, one type of deep learning and machine learning model, have displayed outstanding performance in both image segmentation and image identification tests. While VGG, ResNet, and Inception 10 are frequently used for image recognition, FCNs, U-Net, and Mask R-CNN are popular model for image segmentation. In a number of benchmark datasets, including as the Pascal VOC 2012 segmentation challenge, COCO dataset for instance segmentation, and ImageNet Large Scale Visual Recognition Challenge for image classification, deep learning models have demonstrated state-of-the-art performance. Deep learning implementation of these tasks, however, necessitates extensive data preparation, model selection, and hyperparameter tweaking. Deep learning models may also have drawbacks include overfitting to the training data, difficulty deciphering the learnt features, and sensitivity to changes in the distribution of the input data.

ACKNOWLEDGMENT

REFERENCES

Image Segmentation –

[1] Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2018). Rethinking atrous convolution for semantic image segmentation. arXiv preprint arXiv:1706.05587.

[2] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Springer, Cham.

[3] He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).

[4] 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).

[5] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018).

[6] DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4), 834-848.

Image Identification –

[1] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[2] 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).

[3] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In AAAI (Vol. 4, pp. 12).

[4] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems (pp. 1097-1105).

[5] Huang, G., Liu, Z., van der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

Image Segmentation and Object Identification using Deep Learning

5   M.S Sriya, Chandan Singh Rawat. "Oil Spill Detection using Synthetic Aperture Radar", 2022 Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), 2022
Publication

1 %

6   Submitted to University of Strathclyde
Student Paper

1 %

7   Submitted to University of Hull
Student Paper

1 %

8   Submitted to Queen Mary and Westfield College
Student Paper

1 %

9   Priyanshu Rawat, Madhvan Bajaj, Vikrant Sharma, Satvik Vats. "A Comprehensive Analysis of the Effectiveness of Machine Learning Algorithms for Predicting Water Quality", 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), 2023
Publication

1 %

10   Leah A. Groves, Blake VanBerlo, Natan Veinberg, Abdulrahman Alboog, Terry M. Peters, Elvis C. S. Chen. "Automatic segmentation of the carotid artery and internal jugular vein from 2D ultrasound images for 3D vascular reconstruction",

1 %

International Journal of Computer Assisted
Radiology and Surgery, 2020
Publication

11 Submitted to UI, Springfield
Student Paper
1 %

12 "Brainlesion: Glioma, Multiple Sclerosis,
Stroke and Traumatic Brain Injuries", Springer
Science and Business Media LLC, 2018
Publication
1 %

13 Submitted to Middlesex University
Student Paper
1 %

14 digitalcommons.unl.edu
Internet Source
1 %

15 Submitted to SRH Holding
Student Paper
<1 %

16 Zhiqiang Xu, Yujie Zhang, Jianguo Miao, Qiang
Miao. "Global attention mechanism based
deep learning for remaining useful life
prediction of aero-engine", Measurement,
2023
Publication
<1 %

17 doaj.org
Internet Source
<1 %

18 Dinh-Hung Le, Nhat-Minh Le, Khac-Hung Le,
Van-Truong Pham, Thi-Thao Tran. "DR-
Unet++: An Approach for Left Ventricle
<1 %

Segmentation from Magnetic Resonance Images", 2022 6th International Conference on Green Technology and Sustainable Development (GTSD), 2022
Publication

19	dokumen.pub Internet Source	<1 %
20	web.wpi.edu Internet Source	<1 %

Exclude quotes	On	Exclude matches	< 10 words
Exclude bibliography	On		