International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

IMAGE SEGMENTATION USING DEEP LEARNING TECHNIQUES

G.Mythri Sri, G. Naga Thanishq Reddy, G. Nikhitha, G. Nithin Kumar

AI&ML, Malla Reddy University

---***---

Abstract: Color image segmentation is an important task in computer vision, which involves partitioning an image into different regions based on their color similarity. In recent years, deep learning has emerged as a powerful tool for image segmentation. The use of convolutional neural networks (CNNs) has shown promising results in various applications, including color image segmentation. The process of color image segmentation using deep learning typically involves training a CNN on a large dataset of labeled images, where each pixel in the image is assigned, a label corresponding to its segment. During training, the network learns to extract features from the input image and use them to classify each pixel into its corresponding segment. The trained network can then be used to segment new images.

Key Words: Image segmentation, Convolutional Neural Networks (CNNs), Numerical digits, Recognition accuracy, robustness, Computational efficiency, Pre-processing techniques, Communication accessibility, Assistive technology.

1. INTRODUCTION

Image processing refers to a set of methods used to carry out specific operations on the image to extract some meaningful information from it. Digital image processing manipulates the images using a computer system. Computer vision is the field that focuses on gaining high-level understanding from digital images. Image segmentation is the subfield of digital image processing, and computer vision provides set of methods for splitting a digital image into various fragments. This process helps to analyze the image in a more meaningful form. The objective of image segmentation is to localize boundaries and objects presented in an image. For that, a label is assigned to every pixel in an image where pixels with the same label share a definite set of characteristics. Numerous methods are available in the literature that is successfully employed for image segmentation. Some of them are edge-detection, graph-cut, k-means clustering, histogram-based methods, thresholding, and watersheds. However, in recent years, Deep Learning (DL) methods show remarkable performance for image segmentation tasks with higher accuracy. Deep learning is the field of machine learning algorithms that contains models that are inspired by human brains. It is one of the revolutionary advancements in machine learning and artificial intelligence [3]. Deep learning is successfully applied to many fields, including agriculture, healthcare, video surveillance, etc. [4][5]. Image segmentation is useful in many applications such as real time object detection, medical image analysis, video surveillance, augmented reality, etc. This paper provides a comprehensive review of different deep learning methods available for image segmentation. Section II describes the background details of image segmentation tasks. Section III describes the various deep learning models available for image segmentation tasks. Section IV covers the details regarding open datasets publically available for image segmentation. Section V explains the various applications, and section VI provides the details about different performance evaluation metrics used to evaluate the deep learning model for image segmentation.

2. RELATED WORK

Computers need deep learning approaches to figure out how tasks can be carried out without specific programming. To perform those functions, computers must learn from the knowledge received. When a necessary computer task is carried out, algorithms can be programmed to tell the system to take specific measures to solve the problem. In practical practice, it helps the machine not to specify any required steps, but to construct its algorithm. In the machine training field, computers are taught how todo tasks where an algorithm is not completely satisfactory. If there are several potential responses, one way is to identify some of the right choices as valid. This method is used to improve the algorithms used as computer training data to determine the right reactions.For example, the MNIST handwritten numbers dataset was often used to train a digital character recognition system.

2.1. Methods of the deep learning:

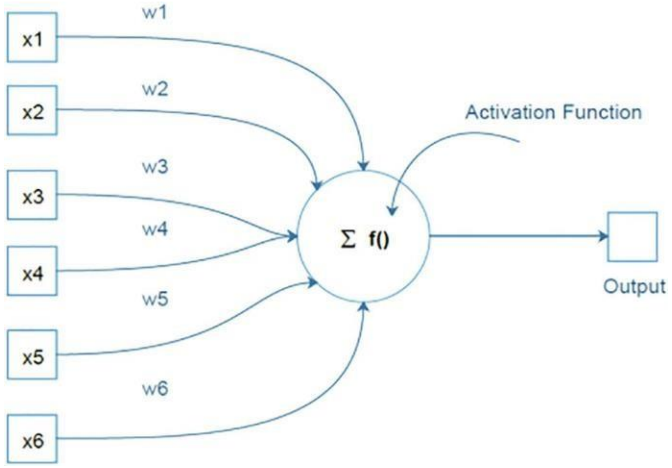
The output of a deep learning analysis algorithm running on data is referred to as ‘‘pattern”. A model is what was learned from an algorithm.It is ‘‘the thing” that is saved after the execution of a training data algorithm and represents the rules, numbers, and any data structure unique to the algorithm needed to make predictions.Deep learning consists of a building model that will practice on some training data and will then process the additional information for predictions. Many approaches for deep learning systemshave been used and researched. Artificial Neural networks Decision tree Support vector Machine Regression Analysis Bayesian Networksless robust to variations in lighting conditions, hand orientation, and deformations.

2.2. Convolutional neural networks (CNN)

ANN is an information processing paradigm mainly inspired by biological nervous systems. It is composed by a high number of processing units, called neurons, working in unison to solve a specific task. Learning process in ANN involves, like in a biological system, the adjustments of the connections between the processing units. Convolutional neural networks is a part of machine learning techniques which deals with knowledge representation and deep learning methodology is a sub field of convolutional neural networks which is more adaptive for analyzing the visual images. This deep learning algorithms help in better understanding of the parameters in the form of dividing the images into layers such that each individual layer is scrutinized and can be more perfectly analyzed when compared to the conventional analysis procedure. The number of layers that are formulated upon the application of deep leaning approach includes of horizontal layer,vertical layer, input layer and output layer. The number of hidden layers can be identified by using the product of horizontal vertical and the number of channels being used.normalization techniques, such as mean subtraction and standard deviation scaling, were applied to enhance the image contrast and reduce the impact of lighting variations.

2.2.1. CNN neural network model

Typically, fully linked neural networks aren't working well on photos. That's because every single pixel Is an input, as we add more layers, the parameter quantity increases exponentially.What's This The special structure renders one image distinguishable from another. Areas within close proximity And areas in the vicinity are highly important in photos. CNN can be used to reach a low level and low Representation of contents of images and end to end structure. a) Convolution In CNN, one tile (3x3 or 5x5 pixels) of the input feature map is extracted and added by a convolution Filters. Filters slide over the input Grid of the function from left to right and from top to bottom,one pixel at a time, extracting each respective map. The CNN ''learns" the optimal values for the filter matrices during training, which allow it to extract Relevant features. As the number of filters used increases, so does the amount of features that the CNN does remove. b) Activation Function An Activation function is a neural network component. The activation function determines whether a Fires on nerves, or not. To ensure that our network is not nonlinear, we need to make sure Because these are nonlinear activation functions. We can use a stage feature as our enabling Function which gives zero or one output.If the output reaches a certain threshold then Neurons are shot, and we got one. If the output value is below the threshold then it is not Shot, and we got a blank. c) Rectified Linear Units The ReLU is one of the Activation functions that is best known The ReLU or ''Set Linear" Unit ''output, a zero for every x value less than zero. For any value of x greater than or equal to The function returns x, then zero. The network applies after every CNN convolution process a Transformation of the rectified linear unit (ReLU) to the transformed element, to implement Nonlinearity inside pattern. d) Pooling Pooling is the next step after ReLU, in which the CNN reduces convolved sampling Feature, and number of feature map dimensions,while retaining the most critical feature Information [10].It's information and this method is called max pooling, and is one of the Usual algorithms. Also, there are other pools, such as the regular pool and min pool. Max pooling works exactly the same Convolutionary style. It slides over the feature map, extracting tiles of a given size and there is forming each extracted tile, output the maximum value to a new function map from that tile and It discards all other beliefs.



There are two parameters to Limit pooling operations.

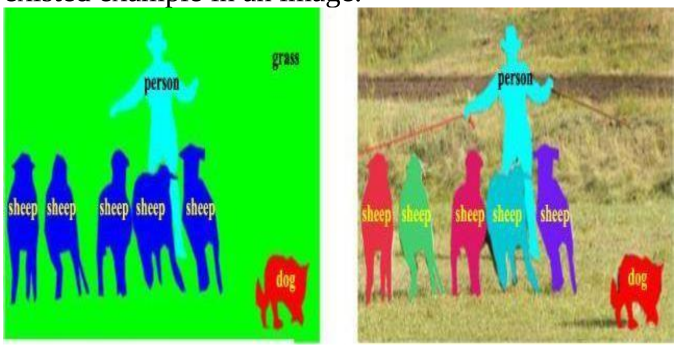
International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

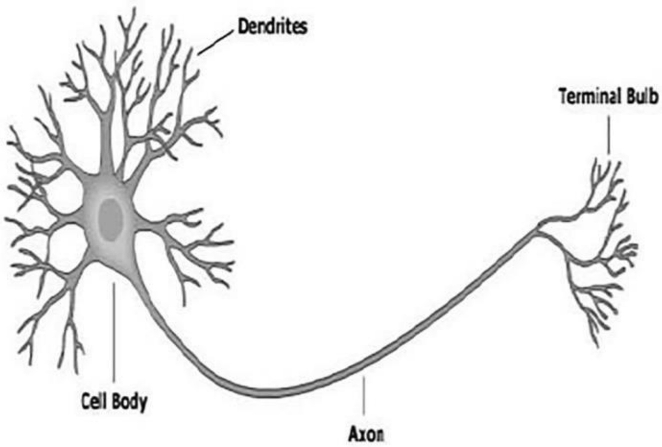
Max-pooling size Typical filter is 2x2 pixels. The distance between each extracted tile is called stride, in pixel. It is distinct from those of convolution, Where convolution filters slide pixel by pixel over the feature map but in max pooling, step Determines the locations from which each tile is drawn. For a filter size of 2x2, specifies a phase of 2 That the max pooling operation extracts from the function map all non-overlapping 2x2 tiles. e) Fully Connected Layers At the end of a convolutional neural network one or more completely connected layers are. Four layers Is completely connected when each node in the first layer connects to each node in the second layer Cap. Their task is to conduct classification based on the characteristic that the convolutions collect.

Mostly, neurons are fully connected to the end. This final layer is completely linked and contains a Activation function SoftMax which gives a probability value of 0 to 1 for each of the model's ranking labels attempt to predict.

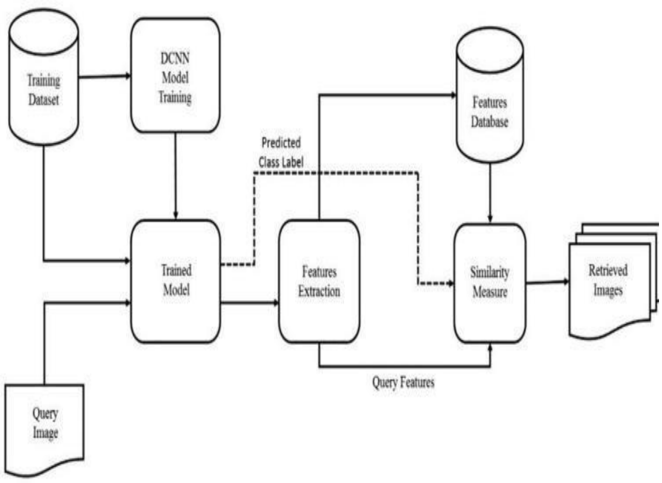
3. IMPLEMENTATION AND ALGORITHM

In computer vision, image segmentation plays a pivotal role. It divides the image into multiple segments that represent objects. For that, image pixels are sorted to make larger segments. These distinct segments are containing each pixel with similar attributes. The segmentation process can be of two types; semantic segmentation and instance segmentation. In semantic segmentation, all the pixels of an image are classified into meaningful classes of objects. It does not differentiate distinct instances of the same object. Semantic segmentation defines the process of coupling each pixel of an image with a class label. Instance, segmentation identifies the boundaries between different objects instead of only classifying them. For that, each instance of an object contained in an image is masked independently. Instance segmentation derives from the object detection process where a bounding box is drawn to detect each object instance of an image with a label for the classification task. Instance segmentation performs object detection by adding a segmentation mask for each existed example in an image.





3.1CNN Architecture



International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Algorithm selection:

The first step is to choose the supervised algorithm to use. Every method has different strength and weak point. The choice depends on the particular problem and on the kind and amount of available data. Some of these algorithms are: Support Vector Machine (SVM), Decision Tree, Artificial Neural Network and Deep Learning. In this work the focus will be on Deep Learning, extension of ANN, for reasons that will be explained in the next sections. Training: The training phase is probably the most important one, as the final performances depend on the predictive model built.A known dataset is selected; must be as more representative of the problem as possible. Using dataset not general enough can lead to over fitting and to bad performances. This set, the training set, must provide an output (label) for each listed input. The algorithm is trained with the selected dataset. The aim of this phase is trying to build a model able to t the data provided, that is predict the correct output for each input provided as best as possible. Validation: The validation phase is important to test the performances achieved by the prediction model built in the previous phase. Another known dataset, called test set, is prepared. The dataset must provide, as the training set, reliable input and output for each example. An important property of this set is that it should be as independent as possible from the training one. The previously trained algorithm is here used to predict the input data of Training andImplementation

The CNN algorithm is trained using a supervised learning approach. The dataset is divided into training and validation sets, and the network parameters are optimized through backpropagation and gradient descent algorithms. The model is implemented using deep learning frameworks, such as TensorFlow which provide convenient tools for building and training CNN models. The training process involves iteratively updating the model weights to minimize the classification error. Various optimization techniques, such as dropout regularization and batch normalization, are employed to improve model performance and prevent overfitting.

3.2 EVALUATION METHODOLOGY

A variety of evaluation metrics can measure the effectiveness of any segmentation model. After training an image segmentation model, it outputs a prediction. To identify the model performance, it is required to be evaluated. The following are some of the frequently used evaluation metrics for image segmentation tasks.

A. Intersection over Union (IoU) IoU calculates the performance using the intersection and union between the Prediction and Ground Truth values. IoU falls between 0 - 1 (0 - 100%), where 0 signifies no overlap and 1 signifies perfectly overlapping segmentation.

B. Pixel Accuracy

Pixel accuracy takes the ratio of correctly classified pixels with respect to total pixels. The following is the equation for defining pixel accuracy. Accuracy = (TP+TN) / (TP+TN+FP+FN) C. Precision Precision is **the** ratio between True Positives and all the Positives. The following is the equation for defining precision. Precision = TP/(TP + FP) D. Recall Recall defines as a measure that correctly identifies True Positives. The following is the equation for defining recall. Recall = TP / (TP + FN) E. F1 Score The weighted average of Precision and Recall is known as F1 score. It is useful when there an uneven class distribution existed in the dataset. The following is the equation for defining the F1-Score. F1 = 2* ((Precision * Recall)/(Precision + Recall))

5.Conclusion

Artificial Intelligence and deep learning models are successfully applied on many computer vision tasks. The paper presents a survey conducted for applying deep learning models for image segmentation tasks. It provides the architecture of FCN, Mask R-CNN, U-Net, SegNet and DeepLab that are successfully applied for various image segmentation tasks and provides higher accuracy and good segmentation results. The paper also provides the details of different open datasets available for experimenting with image segmentation tasks. the various performance metrics used to evaluate the model is also discussed. Moreover, it presents the research work carried out for applying deep learning models for image segmentation tasks.

6. REFERENCES

[1] R. Szeliski, Computer Vision: Algorithms and Applications. London, England: Springer, 2010.

[2] A. Voulodimos, N. Doulamis, A. Doulamis, and E. Protopapadakis, "Deep learning for computer vision: A brief review," Comput. Intell. Neurosci., vol. 2018, pp. 1–13, 2018.

[3] S. Patel, "A comprehensive analysis of Convolutional Neural Network models," Int. j. adv. sci. technol., vol. 29, no. 04, pp. 771–777, 2020.

[4] N. Ganatra and A. Patel, "A comprehensive study of deep learning architectures, applications and tools," Int. j. comput. sci. eng., vol. 6, no. 12, pp. 701–705, 2018.

[5] D. Chhachhiya, A. Sharma, and M. Gupta, "Designing optimal architecture of recurrent neural network (LSTM) with particle swarm optimization technique specifically for educational dataset," Int. J. Inf. Technol., vol. 11, no. 1, pp. 159–163, 2018.

[6] Y. Wei et al., "Learning to segment with image-level annotations," Pattern Recognit., vol. 59, pp. 234–244, 2016.

[7] Z. Wu, C. Shen, and A. van den Hengel, "Bridging category-level and instance-level semantic image segmentation," arXiv [cs.CV], 2016.

[8] T.-Y. Lin et al., "Microsoft COCO: Common objects in context," in Computer Vision – ECCV 2014, Cham: Springer International Publishing, 2014, pp. 740–755.

[9] A. Bhandare, M.V. Bhide, P. Gokhale, R. Chandavarkar, "Applications of Convolutional Neural Network", International Journal of Computer Science and Information Technologies, vol. 7, no. 5, pp. 2206-2215, 2016.

[10] T. Kalaiselvi, S. T. Padmapriya, P. Sriramakrishnan, and K. Somasundaram, "Deriving tumor detection models using convolutional neural networks from MRI of human brain scans," Int. J. Inf. Technol., vol. 12, no. 2, pp. 403–408, 2020