Semantic Image with Convolutional Neural Networks and Deep Learning

Shivam Singh[1], Vishesh Yadav[2], Suraj Khokarale[3], Devarsh Patil[4], Prof. Sagar Dhanake[5]
Students, Department of Computer Engineering [1,2,3,4]
Assistant Professor, Department of Computer Engineering [5]
Dr. D. Y. Patil Institute of Technology, Pune, Maharashtra, India

Abstract: *Semantic image segmentation is a vast area of interest for computer vision and machine learning researchers. Many vision applications need accurate and efficient image segmentation and segment classification mechanisms for assessing the visual contents and perform the real-time decision making. The application area includes remote sensing, autonomous driving, indoor navigation, video surveillance and virtual or augmented reality systems etc. The segmentation and classification of objects generate the specific performance parameters for various applications which require detailed domain analysis. There are broad range of applications where remote sensing image scene classification play an important role and has been receiving remarkable attention. This demand coincides with the rise of deep learning approaches in almost every field or application target related to computer vision, including semantic segmentation or scene understanding. This survey paper provides a review of different traditional methods of image segmentation and classification. By comparing these methods with semantic image segmentation using deep learning it is assumed to show the far better result.*

Keywords: Image segmentation, Region-based, CNN, U-Net

I. INTRODUCTION

An image is a way of transferring information, and the image contains lots of useful information. Understanding the image and extracting information from the image to accomplish some works is an important area of application in digital image technology, and the first step in understanding the image is the image segmentation. In practice, it is often not interested in all parts of the image, but only for some certain areas which have the same characteristics. Image segmentation is one of the hotspots in image processing and computer vision. It is also an important basis for image recognition. It is based on certain criteria to divide an input image into a number of the same nature of the category in order to extract the area which people are interested in. And it is the basis for image analysis and understanding of image feature extraction and recognition. There are many commonly used image segmentation algorithms. The first is the threshold segmentation method. Threshold segmentation is one of the most commonly used segmentation techniques in region-based segmentation algorithms. Its essence is to automatically determine the optimal threshold according to a certain criterion, and use these pixels according to the gray level to achieve clustering. Followed by the regional growth segmentation. The basic idea of the regional growth algorithm is to combine the pixels with similar properties to form the region, that is, for each region to be divided first to find a seed pixel as a growth point, and then merge the surrounding neighbourhood with similar properties of the pixel in its area. Then is the edge detection segmentation method. Edge detection segmentation algorithm refers to the use of different regions of the pixel Gray or colour discontinuity detection area of the edge in order to achieve image segmentation.

- The Purpose of Image Segmentation is to partition the image into meaningful regions with respect to a particular application.
- The segmentation is based on measurements taken from the image and might be grey level colour, texture, depth or motion.
- Usually, Image Segmentation is an initial and vital step in a series of processes aimed at overall image understanding.
- Image segmentation is the first step in image analysis.

Motivation:

Image segmentation is an important image processing, and it seems everywhere if we want to analyse what inside the image. For example, if we seek to find if there is a chair or person inside an indoor image, we may need image segmentation to separate objects and analyse each object individually to check what it is. Image segmentation usually serves as the pre-processing before image pattern recognition, image feature extraction and image compression. Researches of it started around 1970, while there is still no robust solution, so we want to find the reason and see what we can do to improve it.

II. DESCRIPTION OF THE PROBLEM

Problem Statement:

What could satisfy the growing need for efficiently processing and analyzing the information contained in digital images which is a continuous challenge? Since images may contain hundreds of thousands of pixels, any sequential processing upon them becomes easily tedious, even for the fastest processors, so how could we minimize the time and resource consumption?

Objectives:

- The objective of image segmentation is to extract information that is represented in the form of data from image via feature measurement and object representation.
- The aim of image segmentation is depending on the accuracy of feature extraction.

III. LITERATURE REVIEW

Sr No	Title	year	methodology	Author
1	MIScnn : a framework for medical image segmentation with CNN & deep learning.	2021	Python Library for Image Segmentation using Keras for deep learning model	Dominik Muller,Frank Kramer
2	A review of deep learning based medical image segmentation methods.	2021	Image segmentation using CNN architecture(LeNet and AlexNet)	Xiangbin liu , liping Song,shuai liu , yudong zhang
3	The KiTs19 Challenge Data	2020	Collection of CT scans of patients who had renal tumor	Nicholas Heller, Matthew Patterson
4	Neuro Image:Clinical	2019	Image segmentation using MICCAI dataset and CNN architecture	Sergi Valverde, Arnau oliver
5	TernausNet: U-net with VGG11 encoder pre-trained on imageNet for image segmentation	2019	Image segmentation using U-net (CNN) architecture.	Vladimir Iglovikov, Alexey shvets

- MIScnn is a framework for medical image segmentation with CNN & deep learning, which uses keras for deep learning models innovated by Dominik Muller,Frank Kramer.
- Image segmentation using CNN architecture(LeNet and AlexNet) a study by Xiangbin liu , liping Song,shuai liu , yudong zhang gives us the deep knowledge of different CNN architectures and their weaknesses .
- The KiTs19 Challenge Data by Nicholas Heller, Matthew Patterson is Collection of CT scans of patients who had renal tumor this data set is then used to train Deep Learning model for medical image segmentation .
- Neuro Image:Clinical by Sergi Valverde, Arnau oliver innovated the deep learning CNN model for faster image segmentation , model being trained using MICCAI dataset.
- TernausNet: U-net with VGG11 encoder pre-trained on imageNet for image segmentation by Vladimir Iglovikov, Alexey Shvets used the U-net architecture for image Segmentation.

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

IJARSCT
Impact Factor: 6.252

Volume 2, Issue 5, May 2022

IV. SYSTEM DESIGN AND FLOW

A description of the neural network architecture used is shown below in the figure. The overall system communication is described in the figure below. The Descriptive version of system architecture is shown in figure

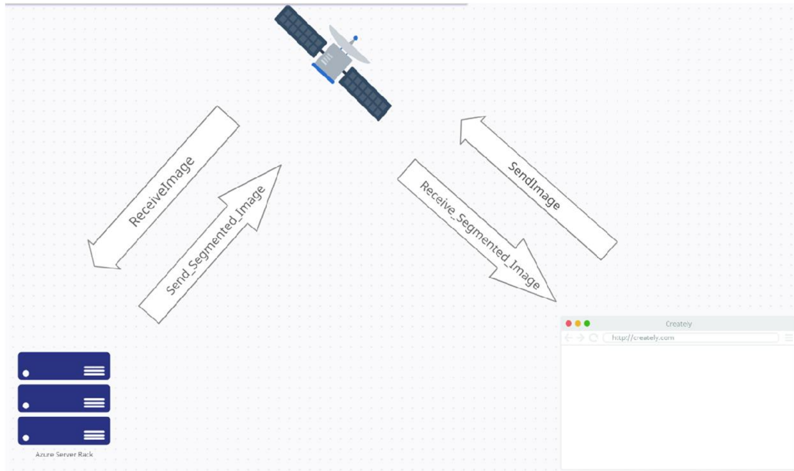

Fig. Communication Diagram

An image is a way of transferring information, and the image contains lots of useful information. Understanding the image and extracting information from the image to accomplish some works is an important area of application in digital image technology, and the first step in understanding the image is the image segmentation. In practice, it is often not interested in all parts of the image, but only for some certain areas which have the same characteristics. Image segmentation is one of the hotspots in image processing and computer vision. It is also an important basis for image recognition. It is based on certain criteria to divide an input image into a number of the same nature of the category in order to extract the area which people are interested in. And it is the basis for image analysis and understanding of image feature extraction and recognition.

Aerial image processing is similar to scene understanding, but it involves semantic segmentation of the aerial view of the landscape. This type of technology is very useful in times of crisis like a flood, where drones can spread to survey different areas to locate people and animals who need rescues. Self-driving cars require image capturing sensors that could enable them to visualize the environment, make decisions and navigate accordingly. These are Some of the real-world implementations of the Semantic Segmentation.

We are building a model which can be used for semantic segmentation. We use the Tensorflow Library to create DNN model. We make it easier to use by simply deploying the model over the web so anyone with internet connection and browser access can use it. The overall user experience is smooth. The user just uploads the image and gets the segmented image.

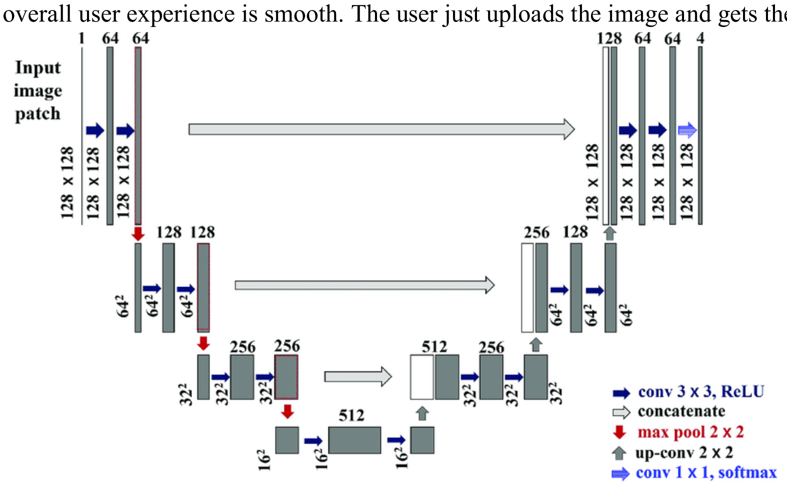

Fig. U-Net Architecture (128x128)

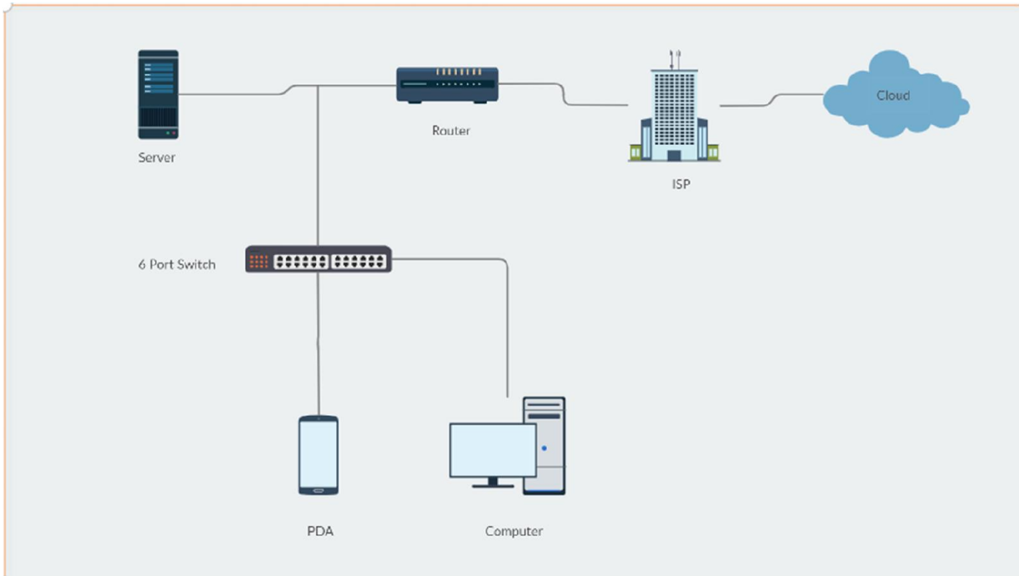

Fig. System Architecture

VI. PROJECT IMPLEMENTATION

CNN (Convolutional Neural Network / ConvNet) Algorithm

A Convolutional Neural Network is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlaps to cover the entire visual area.

CNNs are regularized versions of multilayer perceptrons. Multilayer perceptrons usually mean fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. The "full connectivity" of these networks makes them prone to over fitting data. Typical ways of regularization, or preventing over fitting, include: penalizing parameters during training (such as weight decay) or trimming connectivity (skipped connections, dropout, etc.) CNNs take a different approach towards regularization: they take advantage of the hierarchical pattern in data and assemble patterns of increasing complexity using smaller and simpler patterns embossed in their filters. Therefore, on a scale of connectivity and complexity, CNNs are on the lower extreme.

U-Net Algorithm:

The U-Net architecture stems from the so-called “fully convolutional network” first proposed by Long, Shelhamer, and Darrell.

The main idea is to supplement a usual contracting network by successive layers, where pooling operations are replaced by up sampling operators. Hence these layers increase the resolution of the output. A successive convolutional layer can then learn to assemble a precise output based on this information.

One important modification in U-Net is that there are a large number of feature channels in the up sampling part, which allow the network to propagate context information to higher resolution layers. As a consequence, the expansive path is more or less symmetric to the contracting part, and yields a u-shaped architecture. The network only uses the valid part of each convolution without any fully connected layers. To predict the pixels in the border region of the image, the missing context is extrapolated by mirroring the input image. This tiling strategy is important to apply the network to large images, since otherwise the resolution would be limited by the GPU memory.

We are going to use flask for web deployment of our deep learning model. As we have used python for our deep learning model flask allows us to create a web service using python. Flask is also written in python and therefore it is easier to use our Deep Learning model which is also written in python.

The static folder contains all the static files such as saved images, CSS. The template Folder contains the html files for the web pages. And then we finally have aap.py which is our main flask app. We use app.py to launch the web app.

```
v static
  > model
  > preds
  > styles
  > uploads
> templates
app.py
custom_unet_v1_13-0...
Dockerfile
Procfile            M
requirements.txt
```

```python
app = Flask(__name__)

upload_folder = 'static/uploads'
pred_folder = 'static/preds'
app.config['UPLOAD_FOLDER'] = upload_folder
app.config['PRED_FOLDER'] = pred_folder
```

We first initialize our flask app, Before we can use files from our local folder we have to first configure those folders in our flask app.

```python
model = tf.keras.models.load_model('custom_unet_v1_13-04-22.h5')
##################helper functions ###############################
def allowed_file(filename):
    return '.' in filename and filename.rsplit('.', 1)[1].lower() in ALLOWED_EXTENSIONS

def preprocess_image(path):
    '''
    Preprocess the image so it is compatible with the model.
    Our model input takes shape (1, 128, 128, 3)
    '''
    img = tf.keras.utils.load_img(path)
    img = tf.keras.preprocessing.image.smart_resize(img,(128,128))
    img = tf.keras.utils.img_to_array(img)
    img = img/255.0
    img = np.array([img])
    return img

def create_mask(pred_mask):
    '''
    Create a mask from the prediction.
    '''
    pred_mask = tf.argmax(pred_mask, axis=-1)
    pred_mask = pred_mask[..., tf.newaxis]
    return pred_mask[0]
#######################################################
```

We don't have to create and train model every time. We will use our saved model to load a saved model we already have a utility function in the tensorflow library. As the user will be uploading image, the uploaded image might not be compatible with our model and therefore we might need to perform some preprocessing steps so we can feed this processed image to our model.

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 5, May 2022

```python
@app.route('/',methods=['GET','POST'])
def index():
    return render_template('index.html')

@app.route('/predict',methods=['GET','POST'])
def predict():
    if request.method == 'POST':
        if 'file' not in request.files:
            flash('No file part')
            return redirect(request.url)
        file = request.files['file']
        if file.filename == '':
            flash('No selected file')
            return redirect(request.url)
        if file and allowed_file(file.filename):
            filename = secure_filename(file.filename)
            file.save(os.path.join(app.config['UPLOAD_FOLDER'], filename))
            img = preprocess_image(os.path.join(app.config['UPLOAD_FOLDER'], filename))
            pred_mask = model.predict(img)
            pred_mask = create_mask(pred_mask)
            tf.keras.preprocessing.image.save_img(os.path.join(app.config['PRED_FOLDER'], filename),pred_mask)
            return render_template('pred.html',filename=filename)
        else:
            flash('File extension not allowed')
            return redirect(request.url)
```

By routes we mean the different urls user can visit. We have two route the first one is home page where user can upload the image and also download the pre-Trained model. Next, we have predict route where we output the segmented image.

VII. Discussion and Results

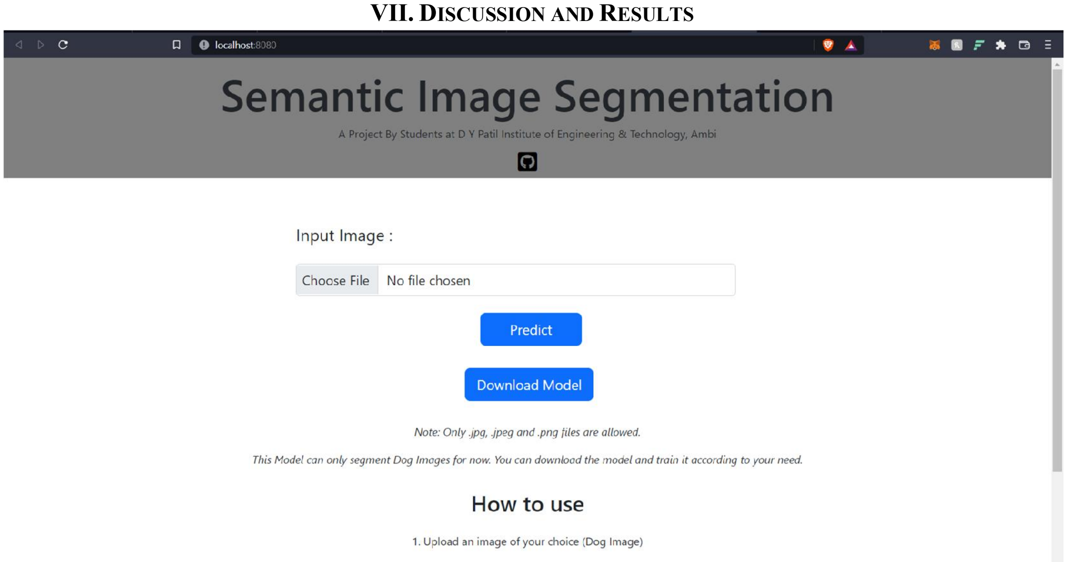

Fig. Webpage1

DOI: 10.48175/568

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 2, Issue 5, May 2022
Impact Factor: 6.252

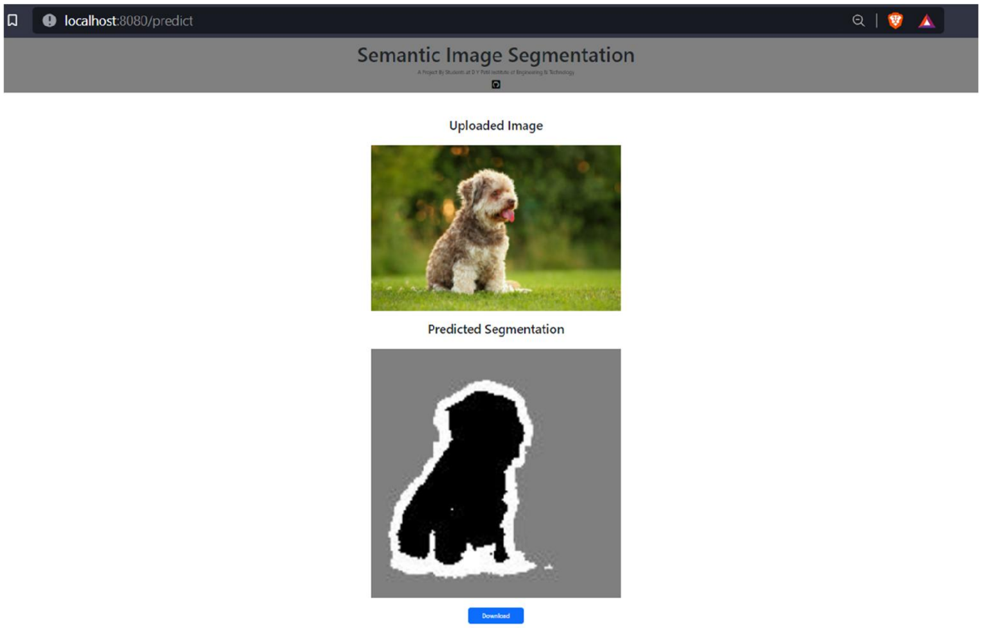

Fig. Webpage 2

VIII. ADVANTAGES & DISADVANTAGES

Advantages:-

- Better Understanding of the image.
- Faster processing of the image.
- Responsiveness to slight changes in what target output is required.
- Makes image processing faster by extracting the required ROI(Region of Interest).
- It is easy to understand and less time-consuming.

Disadvantages:-

- Depending on the segmentation method used, if the image is noisy the image segmentation process might consume greater time.
- If the wrong Image segmentation method is used the segmented image might miss the ROI(region of interest).

IX. APPLICATIONS

Medical Imaging

- CT scan
- Magnetic Resonance Imaging(MRI)

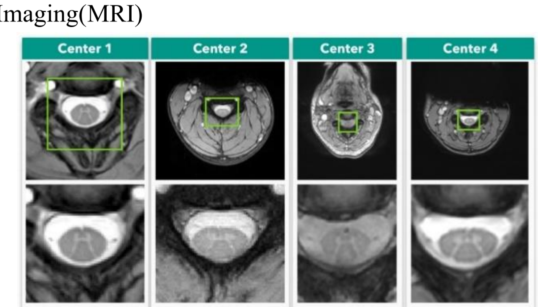

Fig. Medical Image Segmentation

Computer guided Surgery (Da Vinci robot heart surgery)

- Image segmentation is used to identify the region of operation.

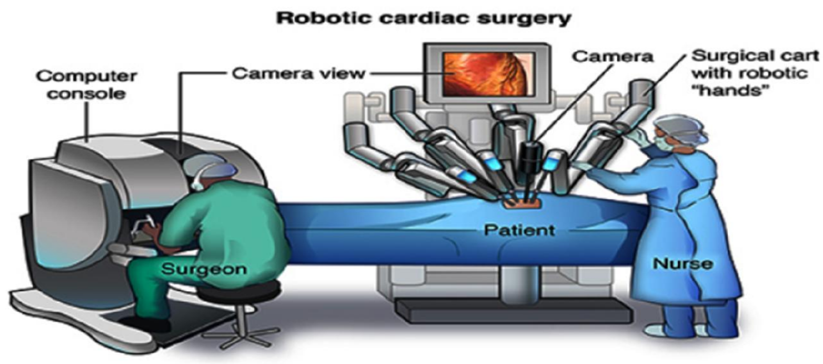

Fig. Robotic Surgery

Self Driving Car

- Self-Driving Cars use semantic image segmentation to differentiate between different objects / person on the road.

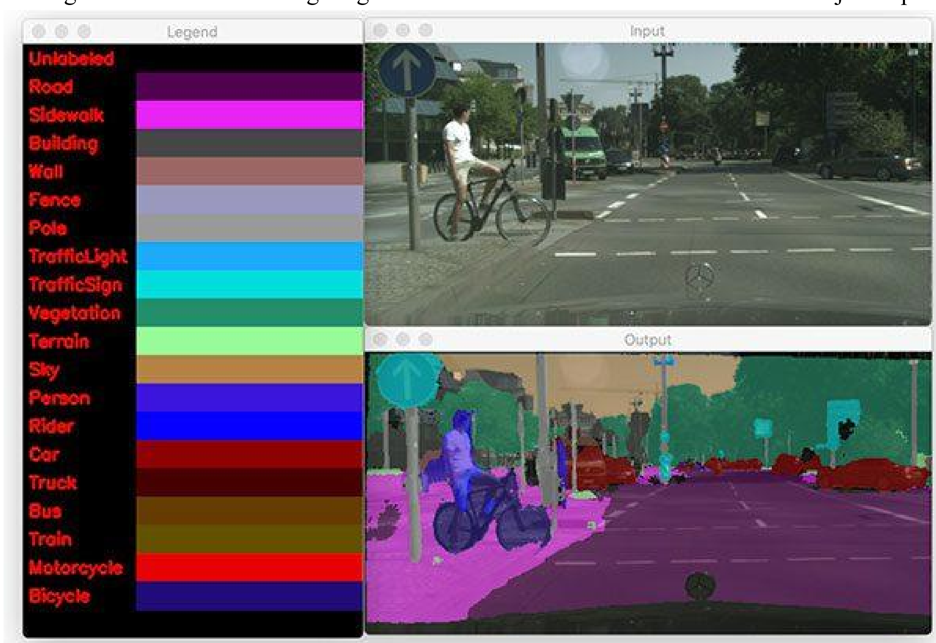

Fig. Self Driving Car Segmentation Image

X. CONCLUSION& FUTURE WORK

Semantic image segmentation is a key application in image processing and computer vision domain. Besides briefly reviewing on traditional semantic image segmentation, we comprehensively list recent progress in semantic image segmentation, especially based on DCNN, in the following aspects: 1. fully convolutional network, 2. up-sample ways. Till now, more and more methods are emerging to make semantic image segmentation more accurate or faster or both on accuracy and speed. We hope this review on recent progress of semantic image segmentation can make some help to researchers related to this area. Semantic segmentation is the process of classifying each pixel belonging to a particular label. It doesn't different across different instances of the same object. For example, if there are 2 cats in an image, semantic segmentation gives same label to all the pixels of both cats. In this article, a novel deep learning architecture was developed to address several shortcomings of the original U-Net architecture. We Overcame some of the short comings by using the pre-trained up-sampler and down-sampler. We made it easy to use, so any one with internet and browser access can use our image segmentation model from anywhere, without requiring heavy hardware requirement. We believe more efficient models are to be expected in the future.

The Future Works include creating a switch-load ML framework which is highly customizable and is easy to deploy over the Internet. By switch-load we mean, there are many other ML model which could also perform Semantic image Segmentation so we it should be easy to load these models in our framework. We also propose an Image pre-processing pipeline which is highly customizable so we can pre-process the same data to work with different models with ease. We also Propose to add a model section on our website where user can switch the model with ease and compare the segmentation image accuracy.

ACKNOWLEDGEMENTS

The completion of our project brings with it a sense of satisfaction, but it is never complete without those people who made it possible and whose constant support has crowned our efforts with success. One cannot even imagine our completion of the project without guidance and neither can we succeed without acknowledging it. It is a great pleasure that we acknowledge the enormous assistance and excellent co-operation to us by the respected personalities.

REFERENCES

[1]. Aggarwal P, Vig R, Bhadoria S, Dethe CG. Role of segmentation in medical imaging: a comparative study. Int J Comput Appl. 2011;29:54–61.

[2]. Gibelli D, Cellina M, Gibelli S, Oliva AG, Termine G, Pucciarelli V, et al. Assessing symmetry of zygomatic bone through three-dimensional segmentation on computed tomography scan and "mirroring" procedure: a contribution for reconstructive maxillofacial surgery. J Cranio-Maxillofac Surg. 2018;46:600–4. https://doi.org/10.1016/j.jcms.2018.02.012.

[3]. Cellina M, Gibelli D, Cappella A, Toluian T, Pittino CV, Carlo M, et al. Segmentation procedures for the assessment of paranasal sinuses volumes. Neuroradiol J. 2020. https://doi.org/10.1177/1971400920946635.

[4]. Hu X, Luo W, Hu J, Guo S, Huang W, Scott MR, et al. Brain SegNet: 3D local refinement network for brain lesion segmentation. BMC Med Imaging. 2020;20:17. https://doi.org/10.1186/s12880-020-0409-2.

[5]. Sun R, Wang K, Guo L, Yang C, Chen J, Ti Y, et al. A potential fi eld segmentation based method for tumor segmentation on multi-parametric MRI of glioma cancer patients. BMC Med Imaging. 2019;19:48. https://doi. org/10.1186/s12880-019-0348-y.

[6]. Claudia C, Farida C, Guy G, Marie-Claude M, Carl-Eric A. Quantitative evaluation of an automatic segmentation method for 3D reconstruction of intervertebral scoliotic disks from MR images. BMC Med Imaging. 2012;12:26. https://doi.org/10.1186/1471-2342-12-26.

[7]. Guo Y, Liu Y, Georgiou T, Lew MS. A review of semantic segmentation using deep neural networks. Int J Multimed Inf Retr. 2018;7:87–93. https ://doi.org/10.1007/s13735-017-0141-z.

[8]. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review. J Med Syst. 2018;42:226. https://doi.org/10.1007/s10916-018-1088-1.

[9]. Wang G. A perspective on deep imaging. IEEE Access. 2016;4:8914–24.

[10]. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2012;2017(42):60–88.

[11]. Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. Annu Rev Biomed Eng. 2017;19:221–48. https://doi.org/10.1146/annurev-bioen g-071516-044442.

[12]. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. Lect Notes Comput Sci (including SubserLect Notes ArtifIntellLect Notes Bioinformatics). 2015;9351:234–41.

[13]. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. UNet++: a nested U-Net architecture for medical image segmentation. 2018. http://arxiv.org/ abs/1807.10165. Accessed 19 Jul 2019.

[14]. Isensee F, Petersen J, Klein A, Zimmerer D, Jaeger PF, Kohl S, et al. nnUNet: self-adapting framework for U-Net-based medical image segmentation. 2018. http://arxiv.org/abs/1809.10486. Accessed 19 Jul 2019.

[15]. Kolařík M, Burget R, Uher V, Říha K, Dutta M. Optimized high resolution 3D dense-U-Net network for brain and spine segmentation. Appl Sci. 2019;9:404.

[16]. Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. 3D U-net: learning dense volumetric segmentation from sparse annotation. Lect Notes Comput Sci. 2016;9901:424–32.

[17]. Lee K, Zung J, Li P, Jain V, Seung HS. Superhuman accuracy on the SNEMI3D connectomics challenge. 2017; Nips:1–11. http://arxiv.org/ abs/1706.00120.

[18]. Gibson E, Li W, Sudre C, Fidon L, Shakir DI, Wang G, et al. NiftyNet: a deep-learning platform for medical imaging. Comput Methods Programs Biomed. 2018;158:113–22. https://doi.org/10.1016/j.cmpb.2018.01.025.

[19]. Neuroimaging Informatics Technology Initiative. https://nifti.nimh.nih. gov/background. Accessed 19 Jul 2019.

[20]. Roy S, Carass A, Prince JL. Patch based intensity normalization of brain MR images. In: Proceedings—international symposium on biomedical imaging. 2013.

[21]. Nyú LG, Udupa JK. On standardizing the MR image intensity scale. Magn Reson Med. 1999;42:1072–81. https://doi.org/10.1002/(SICI)1522- 2594(199912)42:63.0.CO;2-M

[22]. Isensee F, Maier-Hein KH. An attempt at beating the 3D U-Net. 2019;1–7. http://arxiv.org/abs/1908.02182.

[23]. Lin G, Shen C, Van Den HA, Reid I. Efcient piecewise training of deep structured models for semantic segmentation. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2016;2016:3194–203.

[24]. Hussain Z, Gimenez F, Yi D, Rubin D. Diferential data augmentation techniques for medical imaging classifcation tasks. Annu Symp Proc AMIA Symp. 2017;2017:979–84.

[25]. Eaton-rosen Z, Bragman F. Improving data augmentation for medical image segmentation. Midl. 2018; 1–3.

[26]. Perez L, Wang J. The efectiveness of data augmentation in image classifcation using deep learning. 2017. http://arxiv.org/abs/1712.04621. Accessed 23 Jul 2019.

[27]. Taylor L, Nitschke G. Improving deep learning using generic data augmentation. 2017. http://arxiv.org/abs/1708.06020. Accessed 23 Jul 2019.

[28]. Isensee F, Jäger P, Wasserthal J, Zimmerer D, Petersen J, Kohl S, et al. batchgenerators—a python framework for data augmentation. 2020. https://doi.org/10.5281/zenodo.3632567.

[29]. Heller N, Sathianathen N, Kalapara A, Walczak E, Moore K, Kaluzniak H, et al. The KiTS19 challenge data: 300 kidney tumor cases with clinical context, CT semantic segmentations, and surgical outcomes. 2019. http://arxiv.org/abs/1904.00445. Accessed 19 Jul 2019.