# Power Line Segmentation in Aerial Images Using Convolutional Neural Networks

Sumeet Saurav, Prashant Gidde(✉), Sanjay Singh, and Ravi Saini

CSIR - Central Electronics Engineering Research Institute (CSIR-CEERI), Pilani, India
sumeet@ceeri.res.in, prashantpsgl@gmail.com

**Abstract.** Visual inspection of transmission and distribution networks is often carried out by various electricity companies on a regular basis to maintain the reliability, availability, and sustainability of electricity supply. Till date the widely used technique for carrying out an inspection is done manually either using foot patrol and/or helicopter operated manually. However, recently due to the widespread use of quadcopters/UAVs powered by deep learning algorithms, there have been requirements to automate the visual inspection of the power lines. With this objective in mind, this paper presents an approach towards automatic autonomous vision-based power line segmentation in optical images captured by Unmanned Aerial Vehicle (UAV) using deep learning backbone for data analysis. Power line segmentation is often considered as a first step required for power line inspection. Different state-of-the-art semantic segmentation techniques available in the literature have been used and a comparative analysis has been done in terms of the Jaccard index on two different power line databases. This paper also presents a new power line database captured using UAV along with the baseline results. Experimental results show that out of the four deep learning-based segmentation architectures used in our experiments the Nested U-Net architecture out-performed others in terms of line segmentation accuracy in various background scenarios.

**Keywords:** U-Net · Nested U-Net · Transfer learning · Unmanned Aerial Vehicle (UAV) · Semantic segmentation · Power line inspection

## 1 Introduction

Visual inspection of the transmission and distribution networks is often carried out by the electricity companies regularly to maintain the reliability, availability, and sustainability of electricity supply. Traditional methods used for inspection of power networks which are being followed from decades makes use of field surveys and airborne surveys [1]. Moreover, during emergency situations or on regular basis the inspection is usually carried out by a team of inspectors traveling either on foot or by helicopters to visually inspect the power lines with the help of binoculars and sometimes with Infrared (IR) and corona detection cameras [2]. The major limitation involved in using the above-mentioned methodology is that the method is quite slow, it is expensive and involves danger and is also limited by the visual observation skill of


B. Deka et al. (Eds.): PReMI 2019, LNCS 11941, pp. 623–632, 2019.
https://doi.org/10.1007/978-3-030-34869-4_68

the inspectors [3]. Therefore, in order to overcome these limitations of the traditional methods of power line inspection, recently a number of studies have been conducted to automate the visual inspection by using automated helicopters, flying robots, and/or climbing robots [3]. In this paper, we also propose a methodology for automatic inspection of the power lines in images captured by UAVs using Deep Learning (DL) as a backbone for carrying out the analysis.

The rest of the paper is structured as follows: Sect. 2 details different existing relevant literature reviews followed by a description of the proposed methodology in Sect. 3. Section 3 also highlights different blocks used in our proposed methodology which includes discussion about different pre-processing and post-processing techniques and deep learning architectures used in our experiments. Next, in Sect. 4, we highlight the experimental setup which includes a description of the databases, details about different experiments carried out on the databases and discussion related to the results. Finally, we conclude the paper with a brief conclusion mentioned in Sect. 5.

## 2 Related Works

Although a vision-based approach powered by deep learning algorithms seems to be the most impeccable approach for power line inspection, but there are only a few works available in the literature dealing with power line inspection. This is mainly attributed to the lack of power line database available in the literature for performing experiments. From the best of our knowledge, the first work related to the use of computer vision in power line inspection has been reported in [4]. In this work, the authors have done a survey study related to the use of computer vision in the detection of power lines, an inspection of power lines, detection and inspection of insulators, power line corridor maintenance, and pylon detection. In another work reported in [5], the authors have devised a method called CBS (Circle Based Search) for power line detection. This method was validated using several tests on real and synthetic images, obtaining satisfactory results in both cases. In [6] author proposed method, named PLineD for power line detection and inspection using UAV captured visual camera images. The database consists of 82 images from various background scenarios of power lines captured using hexacopter UAV. In [7], proposed a technique for a multi-class classification for power infrastructure detection and classification using deep learning approaches. The database consists of 150 pictures taken from UAV. The proposed method achieved 75% F-score for multi-class classification and an 88% F-score achieved for pylon detection. For power line recognition, an F-score of 70% obtained on 11 unseen images. In [8], the authors have proposed secure autonomous navigation approaches for transmission line inspection. The tower detection performed using a faster region-based convolution neural network and power line segmentation achieved using a fully convolutional neural network. Finally constructed UAV platform was evaluated in a practical environment.

## 3 Proposed Methodology

The block diagram representation of the proposed methodology used in this work has been shown in Fig. 1. During training, we provide the CNN architecture the training and corresponding ground-truth mask images. Once the training gets completed the trained network should be able to predict the binary mask corresponding to any unseen test image containing power lines. The predicted mask obtained from the trained model is overlying on the original input image in order to visualize the segmented power line image as shown in the figure below.

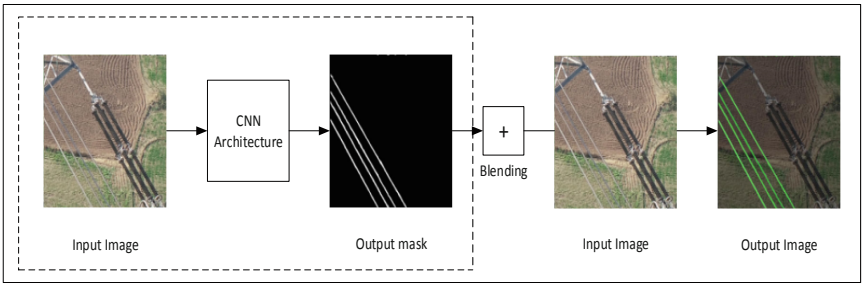


**Fig. 1.** The methodology used for power line segmentation

### 3.1 Datasets and Ground-Truth Preparation

In the proposed work, different experiments have been carried out on two power line databases. The first database (referred to in this work as SR-RGB database) is generated with the cooperation of Turkish Electricity Transmission Company (TEIAS) and has been obtained from [9]. In this database, videos were captured from actual aircraft flown over Turkey at 21 different locations during different days of the season. The captured images available in the database have different background scenarios, and weather conditions along with different lighting conditions. Due to varying conditions, the database contains several difficult scenes where low contrast causes invisibility for the power line. At present, the database contains 4000 Infrared (IR) and 4000 Visible light (VL) images having a resolution size of 128 × 128. Out of 4000 images from each category, 2000 images contain power lines and the rest of the images do not contain power lines. We have only used the RGB images available in the database in our experiments. Moreover, since the size of the images was too small, therefore the images were first super-resolved to a size of 512 × 512 using the technique presented in [10]. Sample images from the database and there corresponding super-resolved version has been shown in Fig. 2. As can be seen from the figure, the quality of the super-resolved image is comparable to that of the original image and hence we used the super-resolution technique presented in [10].

The second database used in experiments is an in-house database that consists of 530 power line images captured by a UAV and we refer to this database as the NAL-RGB database. The dataset consists of various background scenarios such as agrarian and rural areas. The resolution of the images in the database is 5472 × 3078 pixels. Therefore, due to the limitation of the GPU memory, we first cropped the images to a size of 512 × 512 pixels, which are non-overlapping image patches from the original input aerial image as shown in Fig. 3. Using this operation we obtained a total of 40227 images. From the total images, the images which do not contain any power lines were removed and the final database consists of 3568 images containing power lines.

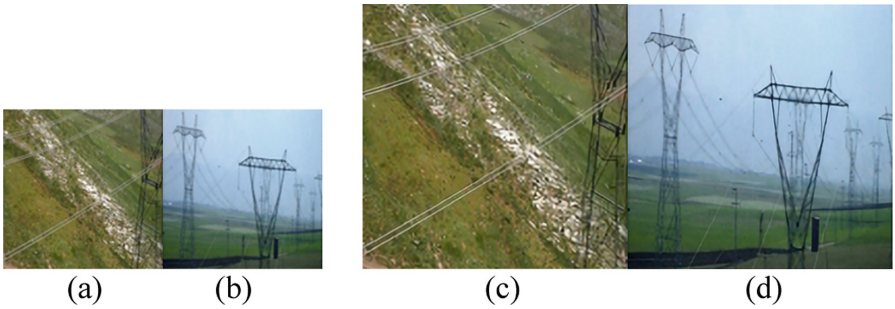


**Fig. 2.** (a) and (b) are the original images; (c) and (d) are 4x super-resolved images

Once the database gets prepared, the next step involved in the generation of ground-truth required for training the deep learning architectures. To generate the ground–truth mask from the images available in both the datasets we used the VGG image annotator (VIA) [11] which is publically available free of cost. The sample image along with its annotation has been shown in Fig. 4 and the sample annotated image with its binary mask is shown in Fig. 5.

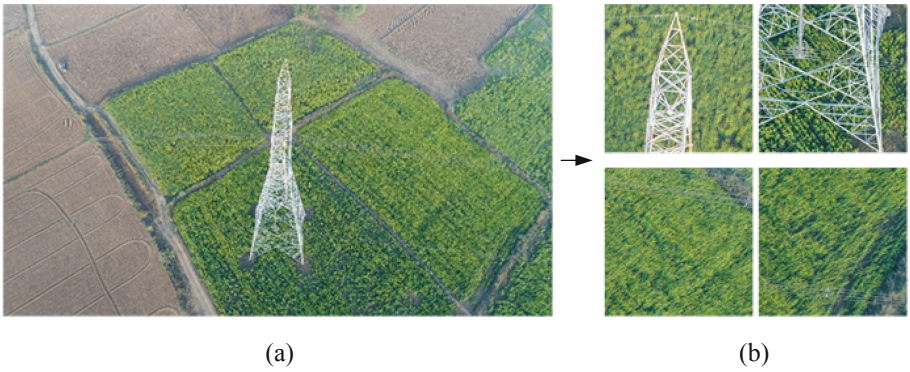


**Fig. 3.** (a) Original image with size 5472 × 3078 (b) Cropped image with size 512 × 512

### 3.2 Deep Learning Architectures

Four different state-of-the-art deep learning inspired architectures used for semantic segmentation have been used in this work. The first architecture called U-Net is proposed for biomedical image segmentation [12]. The architecture consists of an Encoder and a Decoder as shown in Fig. 6. As shown in the figure, the encoder consists of four blocks, wherein each block contains two 3 × 3 convolutional layers with ReLU activation function followed by a 2 × 2 max pool layer with stride 2 used for the down-sampling operation. The number of feature channels gets doubled after each successive down-sampling operation, starting with 64 feature maps for the first block, 128 for the second, and so on. The purpose of this contracting path is to capture the context of the input image in order to be able to do segmentation. The decoder also consists of four blocks, wherein each block contains the up-sampling (de-convolution) layer and the concatenation layer followed by two 3 × 3 convolution layer. After each up-sampling operation, it halves the number of feature channels and concatenates the higher resolution features from the encoder and up-sampled feature from decoder for better localization. The final layer consists of a 1 × 1 convolution layer to map the feature vector to the desired class. The output of this model is a pixel-by-pixel mask that shows the class of each pixel.

**Fig. 4.** Images with their annotations

**Fig. 5.** Images with their binary masks

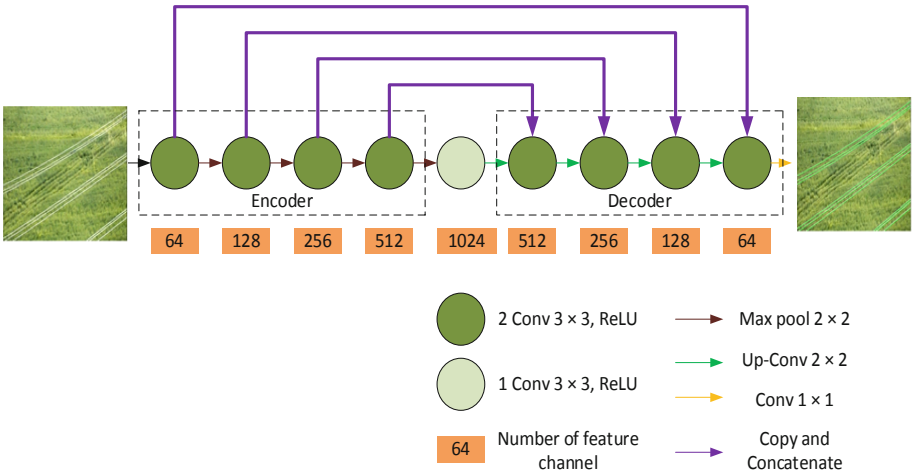


**Fig. 6.** U-Net architecture.

The second architecture called UNet-11 is an improved version of existing U-Net architecture. It consists of the VGG11 network as an encoder and further details of the architecture can be found in [13]. VGG11 network contains 7 convolutional layers along with 5 max pool layers. Each successive convolution layer is followed by the ReLU activation function. All convolutional layer uses 3 × 3 kernels wherein max pool layer is used to reduce the size of the feature map by 2. The third architecture called UNet-16 is also an improved version of U-Net and is reported in [13]. It consists of the VGG16 network as an encoder in U-Net architecture. VGG16 network contains 13 convolutional layers, with each successive layer followed by the ReLU activation function. All convolutional layer uses 3 × 3 kernels wherein max pool layer is used to reduce the size of the feature map by 2. The final architecture called Nested U-Net is an improved and modified version of U-Net architecture [14]. Nested U-Net as the name implies makes use of nested and dense skip connection between encoder and decoder apart from the typical skip connection used in U-Net. Dense skip connection is used to improve the flow of gradient. Nested U-Net consists of dense convolution block helpful in collecting the semantic level of feature map from the encoder part. The block-level representation of the Nested U-Net architecture has been shown in Fig. 7. In the figure, the green box indicates the dense convolution block which follows the consecutive dense convolution layer. The red line and orange line indicate the skip and nested connection between each dense convolution layer. Deep supervision is performed after each convolution block visualized using the blue line. In deep supervision, dense convolution layer 0_1, 0_2, 0_3 and 0_4 are added which is finally used for pixel-wise segmentation. Further details about the Nested U-Net could be found in [14].

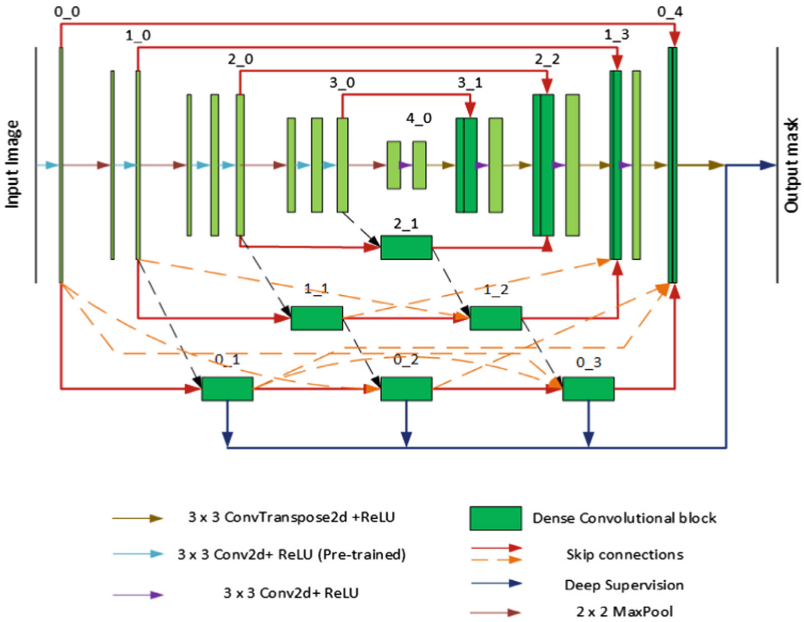


**Fig. 7.** Nested U-Net architecture (Color figure online)

## 4 Experimental Results and Discussion

In this section, we discuss different experiments performed on two power line database. All this experiment is performed in the PyTorch environment on GTX GeForce 1080Ti GPU. For U-Net and Nested U-Net, we used the open-source implementation available at github.com/Nested-UNet, whereas for UNet-11 and UNet-16 we implemented the architecture on our own. The evaluation metric used is called the Jaccard index (Intersection over Union) which is defined as a similarity measure between a finite number of sets.

### 4.1 Results on SR-RGB Database

As discussed in Sect. 3.1 the SR-RGB dataset consists of 2000 images out of which 499 images were blurred and were found difficult to annotate so we removed those images. After removing blurred images, the SR-RGB dataset consists of 1501 power line images. We selected 1200 images for training and 301 images were used for validation purposes.

The training and validation plot of different deep learning architectures such as U-Net, UNet-11, UNet-16, and Nested U-Net has been shown in Fig. 8. These plots have been obtained by training these networks using a batch size of 4, learning rate value of 1e-4 with Adam optimizer. The values of these hyperparameters have been obtained by performing a number of experiments using different values of these hyperparameters.

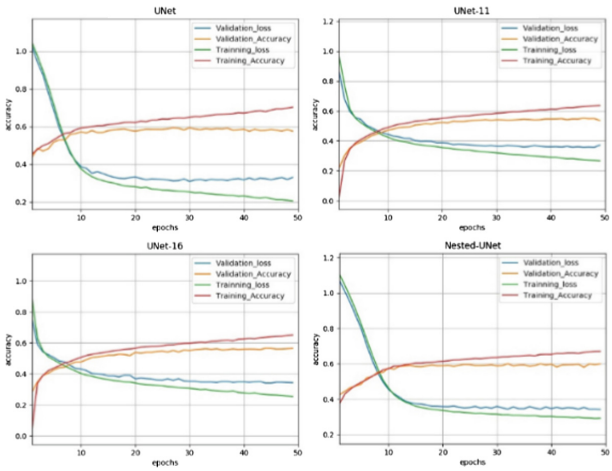


**Fig. 8.** Training and validation plot of different deep learning architectures on SR-RGB database.

The Jaccard index value obtained on validation images from the database corresponding to U-Net, UNet-11, UNet-16, and Nested U-Net is 0.59, 0.59, 0.60, and 0.59 respectively. From here, we can find out that all the architectures performed equally

well on this dataset. The segmented output image obtained after performing the blending operation on the predicted output mask corresponding to the input test image using the Nested U-Net architecture has been shown in Fig. 9. From the out segmented result, it is clear that the trained model is perfectly capable of segmenting the power lines contained in the input test image.

### 4.2 Results on NAL-RGB Database

The NAL-RGB database consists of 3568 RGB images having a resolution of 512 × 512 pixels. From 3568 images, we have used 2850 images for training and the remaining 718 images have been used for validation/testing purposes.

The training and validation plot of different deep learning architectures such as U-Net, UNet-11, UNet-16, and Nested U-Net has been shown in Fig. 10. These plots have been obtained by training these networks using a batch size of 4, learning rate value of 1e-4 with Adam optimizer. The values of these hyperparameters have been obtained by performing a number of experiments using different values of these hyperparameters.

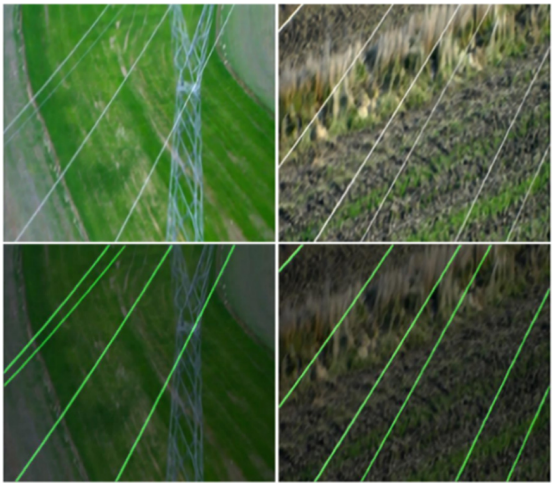

**Fig. 9.** Visual results obtained using the Nested U-Net trained model on SR-RGB database.

The Jaccard index value obtained on validation images from the database corresponding to U-Net, UNet-11, UNet-16, and Nested U-Net is 0.64, 0.66, 0.67, and 0.70 respectively. From here, we can find out that the Nested U-Net performed well compared to the other three architectures. This is mainly attributed to the deep supervision used in Nested U-Net architecture. The segmented output image obtained after performing the blending operation on the predicted output mask corresponding to the input test image using the Nested U-Net architecture has been shown in Fig. 11. From the segmented output result, it is clear that the trained model is perfectly capable of segmenting the power lines contained in the input test image.

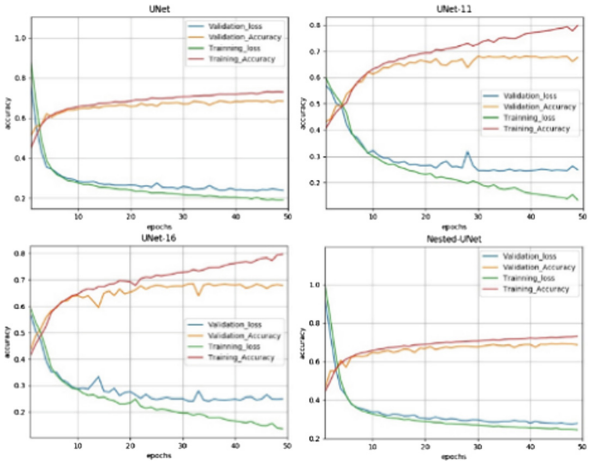


**Fig. 10.** Training and validation plot of different deep learning architectures on the NAL-RGB database.

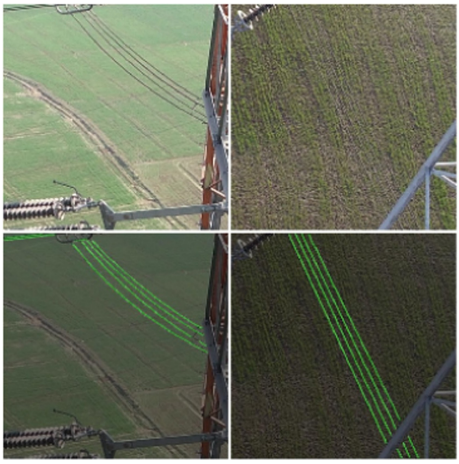

**Fig. 11.** Visual results obtained using the Nested U-Net trained model on NAL-RGB database.

## 5 Conclusion

In this paper, we have presented a methodology for the automatic segmentation of the power line in UAV image using deep learning backbone for data analysis. Power line segmentation is often considered as the first step required for power line inspection. We have also introduced a new database captured using UAV and presented baseline results using different deep learning architectures available in the literature for semantic segmentation. Different deep learning architectures were trained and validated on two power line database and a comparative analysis was done using the Jaccard index as the

evaluation metric. From the experiments, we found that the Nested U-Net performed relatively well compared to the other deep learning inspired image segmentation architectures. This is mainly due to the deep supervision used in Nested U-Net architecture. Thus, the proposed methodology could potentially be used for automatic inspection of power lines in UAVs captured images.

## References


1. Matikainen, L., et al.: Remote sensing methods for power line corridor surveys. ISPRS J. Photogr. Remote Sens. **119**, 10–31 (2016)
2. Katrasnik, J., Pernus, F., Likar, B.: A survey of mobile robots for distribution power line inspection. IEEE Trans. Power Delivery **25**(1), 485–493 (2009)
3. Nguyen, V.N., Jenssen, R., Roverso, D.: Automatic autonomous vision-based power line inspection: a review of current status and the potential role of deep learning. Int. J. Electr. Power Energy Syst. **99**, 107–120 (2018)
4. Mirallès, F., Pouliot, N., Montambault, S.: State-of-the-art review of computer vision for the management of power transmission lines. In: Proceedings of the 2014 3rd International Conference on Applied Robotics for the Power Industry, pp. 1–6. IEEE, October 2014
5. Cerón, A., Prieto, F.: Power line detection using a circle based search with UAV images. In: 2014 International Conference on Unmanned Aircraft Systems (ICUAS), pp. 632–639. IEEE, May 2014
6. Santos, T., et al.: PLineD: vision-based power lines detection for unmanned aerial vehicles. In: 2017 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC), pp. 253–259. IEEE, April 2017
7. Varghese, A., Gubbi, J., Sharma, H., Balamuralidhar, P.: Power infrastructure monitoring and damage detection using drone captured images. In: 2017 International Joint Conference on Neural Networks (IJCNN), pp. 1681–1687. IEEE, May 2017
8. Hui, X., Bian, J., Zhao, X., Tan, M.: Vision-based autonomous navigation approach for unmanned aerial vehicle transmission-line inspection. Int. J. Adv. Rob. Syst. **15**(1), 1729881417752821 (2018)
9. Yetgin, Ö.E., Gerek, Ö.N.: Powerline image dataset (infrared-IR and visible light-VL). Mendeley Data, v7 (2017). http://dx.doi.org/10.17632/n6wrv4ry6v.7
10. Yamanaka, J., Kuwashima, S., Kurita, T.: Fast and accurate image super resolution by deep CNN with skip connection and network in network. In: Liu, D., Xie, S., Li, Y., Zhao, D., El-Alfy, E.S. (eds.) Neural Information Processing. LNCS, vol. 10635, pp. 217–225. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-70096-0_23
11. Dutta, A., Zisserman, A.: The VGG image annotator (VIA). arXiv preprint arXiv:1904.10699 (2019)
12. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
13. Iglovikov, V., Shvets, A.: Ternausnet: U-net with VGG11 encoder pre-trained on imagenet for image segmentation. arXiv preprint arXiv:1801.05746 (2018)
14. Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J.: UNet++: a nested U-net architecture for medical image segmentation. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS 2018. LNCS, vol. 11045, pp. 3–11. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_1