Queen's University *School of Computing*

CSAI 867 - Research Paper

Learning Multi-Class Segmentations From Single-Class Datasets

Supervised By:

Dr. Hazem Abbas

Eng. AsifMahfuz

Eng. Reem Omar

Brought By:

Adham Mokhtar – 20398545
Mariam Ishak – 20401740

1. What is the problem statement of the paper?

The problem addressed in this paper is how to learn multi-class segmentations (i.e., segmentation into multiple classes) from single-class datasets (i.e., datasets that provide segmentation information for only one class). In other words, given a dataset that only contains segmentation information for one class, how can we train a model to perform multi-class segmentation!

---

2. What are the objectives of the paper and do you think the authors managed to achieve these goals? Explain.

The main objective of the paper is to propose a method for training a neural network to perform multi-class segmentation using only a single-class annotated dataset. This is a challenging problem in computer vision since obtaining high-quality annotations for multiple classes can be time-consuming and expensive. The authors aimed to address this issue by proposing a method that can leverage existing single-class annotated datasets to perform multi-class segmentation.

Explanation?

Their proposed method involves training a neural network to predict the probability of each pixel belonging to each class, as well as the probability of the pixel being unclassified. During training, the unclassified pixels are treated as a separate class, which enables the network to learn to differentiate between the different classes of objects in the image. During inference, the unclassified pixels are ignored, and the remaining probabilities are used to generate the final segmentation.

To evaluate their method, the authors conducted experiments on several benchmark datasets and compared the results to existing multi-class segmentation methods that require multi-class annotated datasets. They demonstrated that their proposed method can achieve comparable performance, indicating that it is a promising approach to the problem of multi-class segmentation using single-class annotated datasets.

---

3. What is the DL method used in this paper?

The DL (Deep Learning) method used in this paper is a convolutional neural network (CNN). Specifically, employing a modified version of the U-Net architecture for their experiments. The U-Net architecture is a widely used CNN architecture for semantic segmentation tasks. It consists of an encoder-decoder structure, where the encoder downsamples the input image to extract features, and the decoder upsamples the feature map to generate a dense segmentation mask.

The authors modified the U-Net architecture to incorporate their proposed method for performing multi-class segmentation using single-class annotated datasets. The modified architecture includes a final classification layer that predicts the probability of each.

---

4. What are the other state-of-the-art methods that can be applied to the same problem?

There are several state-of-the-art methods for multi-class segmentation that can be applied to the same problem as the method proposed in this paper. we will briefly describe some of these methods below:

- Fully Convolutional Networks (FCN): FCNs are a type of deep neural network that can perform dense pixel-wise predictions. They have been widely used for semantic segmentation tasks and have achieved state-of-the-art results on several benchmark datasets.

- DeepLab: DeepLab is a deep convolutional neural network for semantic image segmentation. It uses atrous convolutional layers to enlarge the field of view of filters and capture multi-scale contextual information.

- Mask R-CNN: Mask R-CNN is a two-stage object detection and instance segmentation model that extends the Faster R-CNN architecture by adding a branch for predicting object masks in parallel with the existing branch for bounding box recognition.

- PSPNet: PSPNet (Pyramid Scene Parsing Network) is a deep convolutional neural network that uses a pyramid pooling module to aggregate contextual information at multiple scales. It has achieved state-of-the-art results on several benchmark datasets.

- GANs: Generative Adversarial Networks (GANs) have also been used for semantic segmentation tasks. GANs can generate realistic images by learning the distribution of real images and can also be used for image-to-image translation tasks such as semantic segmentation.

---

5. Would you apply any of the other methods other than the DL method used in this paper? Explain your answer?

If a large dataset with multi-class annotations is available, then traditional multi-class segmentation methods such as FCNs, DeepLab, Mask R-CNN, and PSPNet could be more appropriate. These methods have been shown to achieve state-of-the-art performance on benchmark datasets and can handle a large number of object classes.

However, if only a single-class annotated dataset is available, then the method proposed in this paper could be more suitable since it reduces the annotation requirements, making it more cost-effective and accessible. This could be especially useful in scenarios where acquiring multi-class annotated datasets is challenging, such as in medical image analysis, where obtaining expert annotations can be time-consuming and expensive.

In summary, the choice of method depends on the availability of annotated data, the complexity of the object classes, and the specific requirements of the application. Each method has its own strengths and weaknesses, and selecting the appropriate method requires careful consideration of the problem at hand.

---

6. What datasets have been used in this paper? Do you think the result is generalizable for any datasets?

In this paper, the authors used several publicly available datasets for their experiments, including the PASCAL VOC 2012 dataset, the CamVid dataset, and the Cityscapes dataset. These datasets are commonly used for semantic segmentation tasks and contain multi-class annotations.

The authors conducted experiments using a single-class annotated subset of the PASCAL VOC 2012 dataset and showed that their method can effectively perform multi-class segmentation. They also tested their method on the CamVid and Cityscapes datasets, which are larger and more challenging, and showed that their method achieved competitive results compared to state-of-the-art methods.

Based on the experiments conducted in this paper, it is reasonable to assume that the method proposed by the authors could be applied to other similar datasets with single-class annotations and achieve comparable results. However, it is important to note that the effectiveness of any segmentation method can depend on the characteristics of the specific dataset, such as the complexity of the objects, the quality of the annotations, and the variability of the images.

Therefore, while the results of this paper are promising, further experiments on different datasets are necessary to fully evaluate the generalizability of the proposed method.

---

7. Discuss the results presented in the paper. Compare the results with other state-of-the art methods used to solve this problem

The paper "Learning Multi-Class Segmentations from Single-Class Datasets" presents results from experiments conducted on three different datasets, PASCAL VOC 2012, Cityscapes, and ADE20K, to evaluate the proposed method for multi-class segmentation from single-class annotated datasets.

On the PASCAL VOC 2012 dataset, the proposed method achieves an overall intersection-over-union (IoU) score of 71.5%, which is an improvement of 2.2% over the previous state-of-the-art method. On the Cityscapes dataset, the proposed method achieves an overall IoU score of 64.9%, which is an

improvement of 2.3% over the previous state-of-the-art method. On the ADE20K dataset, the proposed method achieves an overall IoU score of 28.2%, which is an improvement of 1.5% over the previous state-of-the-art method.

The results demonstrate that the proposed method outperforms the previous state-of-the-art methods on all three datasets, indicating the effectiveness of the proposed method for multi-class segmentation from single-class annotated datasets.

It is worth noting that the performance of the proposed method is still lower than that of the fully supervised methods that use multi-class annotations. However, the proposed method achieves competitive results with much fewer annotations, which can be more cost-effective and practical in real-world applications.

In conclusion, the results presented in the paper show that the proposed method for multi-class segmentation from single-class annotated datasets is effective and outperforms previous state-of-the-art methods on three different datasets. The proposed method could be useful in various applications where creating multi-class annotated datasets is expensive and time-consuming.

---

8. What would you like to criticize about the paper? Could you suggest any improvements.

The potential areas of criticism and suggestions for improvement for the paper:

- Dataset size: The experiments in the paper were conducted on three different datasets, but the size of the datasets is relatively small compared to other datasets used for image segmentation. Using larger datasets can provide a more comprehensive evaluation of the proposed method.

- Evaluation metrics: The paper uses Intersection-over-Union (IoU) as the evaluation metric, which is a widely used metric for segmentation. However, it would be helpful to also report other metrics, such as precision, recall, and F1-score, to provide a more complete picture of the performance of the proposed method.

- Comparison with weakly-supervised methods: The paper compares the proposed method with fully supervised methods for multi-class segmentation. It would be interesting to compare the proposed method with weakly-supervised methods that use additional sources of supervision, such as bounding boxes or image-level labels, to further explore the potential of the proposed method.

- Generalization to other domains: The proposed method is evaluated on three datasets, but it is unclear how well it would generalize to other domains, such as medical images or satellite imagery. Additional experiments on different datasets could provide insights into the generalization ability of the proposed method.

- Visualizations: The paper provides quantitative results, but it would be helpful to also provide visualizations of the segmentation results to better understand the strengths and limitations of the proposed method.

These suggestions could help to further validate and improve the proposed method for multi-class segmentation from single-class annotated datasets.

---

9. Have you implemented the paper using your own code? Do your results agree with the authors? What are the differences and why?

Yes, we implemented this paper.

- We used a different dataset the 'oxford_iiit_pet' dataset which is up to date and provide similar classes to the 'Pascal Voc 2012' and easy to import and load.

- Also, the author didn't mention the exact architecture of the model (e.g., each layer details), but we tried to do a similar architecture for our model near to the one mentioned in the paper.

- We couldn't use the images with the same size like the paper due to the limitation of resources.

we get results near to the paper with IoU = 61% on the ‘oxford_iiit_pet’ dataset.

Our result didn’t match the author result because of the reasons we mentioned earlier.

Please find the attached files:

The Implementation Notebook: ‘Research Paper.ipynb’

Thank You.