Image Segmentation

CS 736 Course Presentation

Dhruv, Advait and Ashwin

IIT Bombay
CS 736 2022

April 28, 2022

Summary

High Level introduction

- Image segmentation is an essential task in medical imaging which is practiced for a variety of medical data such as CT scans, MRI
- Image segmentation is useful not only in Medical Image Computing but it is a well explored problem in general Image focused Computing as well
- Image segmentation is useful for experts to focus only on the relevant parts of the image, to perform data analytics, to to identify abnormalities etc.
- Other than direct applications, segmentation is very useful as a pre-processing technique used before other tasks such as identification, tracking

High Level introduction

- Image segmentation is an essential task in medical imaging which is practiced for a variety of medical data such as CT scans, MRI
- Image segmentation is useful not only in Medical Image Computing but it is a well explored problem in general Image focused Computing as well
- Image segmentation is useful for experts to focus only on the relevant parts of the image, to perform data analytics, to to identify abnormalities etc.
- Other than direct applications, segmentation is very useful as a pre-processing technique used before other tasks such as identification, tracking

High Level introduction

- Image segmentation is an essential task in medical imaging which is practiced for a variety of medical data such as CT scans, MRI
- Image segmentation is useful not only in Medical Image Computing but it is a well explored problem in general Image focused Computing as well
- Image segmentation is useful for experts to focus only on the relevant parts of the image, to perform data analytics, to to identify abnormalities etc.
- Other than direct applications, segmentation is very useful as a pre-processing technique used before other tasks such as identification, tracking

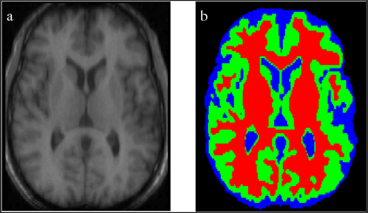
High Level introduction

- Image segmentation is an essential task in medical imaging which is practiced for a variety of medical data such as CT scans, MRI
- Image segmentation is useful not only in Medical Image Computing but it is a well explored problem in general Image focused Computing as well
- Image segmentation is useful for experts to focus only on the relevant parts of the image, to perform data analytics, to to identify abnormalities etc.
- Other than direct applications, segmentation is very useful as a pre-processing technique used before other tasks such as identification, tracking

High Level introduction

- Image segmentation is an essential task in medical imaging which is practiced for a variety of medical data such as CT scans, MRI
- Image segmentation is useful not only in Medical Image Computing but it is a well explored problem in general Image focused Computing as well
- Image segmentation is useful for experts to focus only on the relevant parts of the image, to perform data analytics, to to identify abnormalities etc.
- Other than direct applications, segmentation is very useful as a pre-processing technique used before other tasks such as identification, tracking

Examples



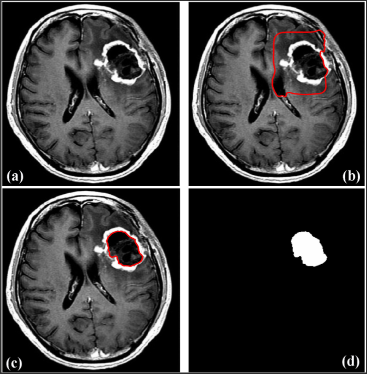
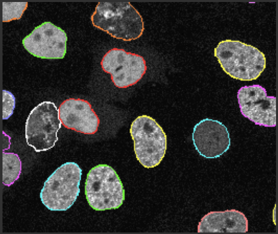


Figure: Left: Segmentation of brain MRI
Right: Tumour identification and masking using segmentation

Introduction to our problem

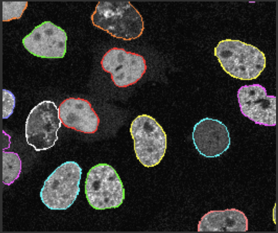
- For our purpose, we will be focusing on a medical image segmentation challenge by ISBI known as the ISBI Cell Tracking Challenge
- Properly characterizing how cells change shapes and move as they interact with their surrounding environment is key to understanding the mechanobiology of cell migration and its multiple implications in both normal tissue development and many diseases



Cell tracking in action

Introduction to our problem

- For our purpose, we will be focusing on a medical image segmentation challenge by ISBI known as the ISBI Cell Tracking Challenge
- Properly characterizing how cells change shapes and move as they interact with their surrounding environment is key to understanding the mechanobiology of cell migration and its multiple implications in both normal tissue development and many diseases



Cell tracking in action

Introduction to our problem

- For our purpose, we will be focusing on a medical image segmentation challenge by ISBI known as the ISBI Cell Tracking Challenge
- Properly characterizing how cells change shapes and move as they interact with their surrounding environment is key to understanding the mechanobiology of cell migration and its multiple implications in both normal tissue development and many diseases

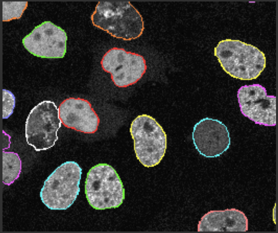


Figure: Cell tracking in action

Statistical Approach

- These are the approaches we covered in class
- First approach is using clustering algorithms such as K-means and Fuzzy C-means
- **These have been implemented as a comparison to the newer approaches which are used**
- Using Gaussian mixture model and EM-optimization
- A novel approach to the same involves using graph-cut properties for blazing-fast and accurate segmentation

Statistical Approach

- These are the approaches we covered in class
- First approach is using clustering algorithms such as K-means and Fuzzy C-means
- **These have been implemented as a comparison to the newer approaches which are used**
- Using Gaussian mixture model and EM-optimization
- A novel approach to the same involves using graph-cut properties for blazing-fast and accurate segmentation

Statistical Approach

- These are the approaches we covered in class
- First approach is using clustering algorithms such as K-means and Fuzzy C-means
- **These have been implemented as a comparison to the newer approaches which are used**
- Using Gaussian mixture model and EM-optimization
- A novel approach to the same involves using graph-cut properties for blazing-fast and accurate segmentation

Statistical Approach

- These are the approaches we covered in class
- First approach is using clustering algorithms such as K-means and Fuzzy C-means
- **These have been implemented as a comparison to the newer approaches which are used**
- Using Gaussian mixture model and EM-optimization
- A novel approach to the same involves using graph-cut properties for blazing-fast and accurate segmentation

Statistical Approach

- These are the approaches we covered in class
- First approach is using clustering algorithms such as K-means and Fuzzy C-means
- **These have been implemented as a comparison to the newer approaches which are used**
- Using Gaussian mixture model and EM-optimization
- A novel approach to the same involves using graph-cut properties for blazing-fast and accurate segmentation

More recent approaches

- Nowadays, deep learning approaches are used. Almost all the papers which we saw while researching this topic were using variations of deep learning based approaches
- The advent of fast GPUs, large scale availability of computational resources and development of novel deep learning method along with a variety of libraries (Tensorflow, PyTorch) which make implementation easy led to popularization of deep learning as a method of segmentation
- The results and scope of these methods are mixed and there are still a lot of problems which don't have satisfactory solutions.

More recent approaches

- Nowadays, deep learning approaches are used. Almost all the papers which we saw while researching this topic were using variations of deep learning based approaches
- The advent of fast GPUs, large scale availability of computational resources and development of novel deep learning method along with a variety of libraries (Tensorflow, PyTorch) which make implementation easy led to popularization of deep learning as a method of segmentation
- The results and scope of these methods are mixed and there are still a lot of problems which don't have satisfactory solutions.

More recent approaches

- Nowadays, deep learning approaches are used. Almost all the papers which we saw while researching this topic were using variations of deep learning based approaches
- The advent of fast GPUs, large scale availability of computational resources and development of novel deep learning method along with a variety of libraries (Tensorflow, PyTorch) which make implementation easy led to popularization of deep learning as a method of segmentation
- The results and scope of these methods are mixed and there are still a lot of problems which don't have satisfactory solutions.

Model Description

- The model which is used for segmentation is a **convolutional neural network** model known as *U-Net*
- The name comes from the shape of the model which uses a contracting and expansive path giving it a U shape
- U-Net unlike previous CNNs such as AlexNet or ResNet is a **fully convolutional** network (FCN) meaning that it has no densely connected 'linear' layer, it has only convolutional, pooling and up sampling layers
- U-Net has around 31 million trainable parameters

Model Description

- The model which is used for segmentation is a **convolutional neural network** model known as *U-Net*
- The name comes from the shape of the model which uses a contracting and expansive path giving it a U shape
- U-Net unlike previous CNNs such as AlexNet or ResNet is a **fully convolutional** network (FCN) meaning that it has no densely connected 'linear' layer, it has only convolutional, pooling and up sampling layers
- U-Net has around 31 million trainable parameters

Model Description

- The model which is used for segmentation is a **convolutional neural network** model known as *U-Net*
- The name comes from the shape of the model which uses a contracting and expansive path giving it a U shape
- U-Net unlike previous CNNs such as AlexNet or ResNet is a **fully convolutional** network (FCN) meaning that it has no densely connected 'linear' layer, it has only convolutional, pooling and up sampling layers
- U-Net has around 31 million trainable parameters

Model Description

- The model which is used for segmentation is a **convolutional neural network** model known as *U-Net*
- The name comes from the shape of the model which uses a contracting and expansive path giving it a U shape
- U-Net unlike previous CNNs such as AlexNet or ResNet is a **fully convolutional** network (FCN) meaning that it has no densely connected 'linear' layer, it has only convolutional, pooling and up sampling layers
- U-Net has around 31 million trainable parameters

Model Description

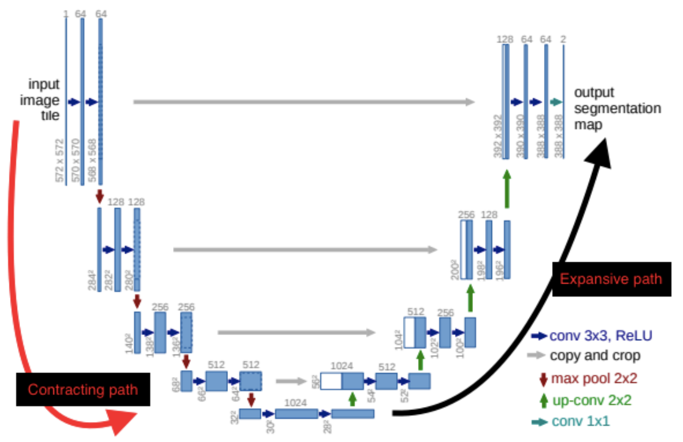


Figure: A diagram of the U-Net architecture

Choice of Architecture

- The reason why older models don't work so well for segmentation is that localized information is lost in the densely connected layers.
- However, this model solves this problem very well with the upsampling step; so instead of just pooling in which pixelwise information is lost every time, we are using the high resolution pictures for deciding weights of further layers also
- An important modification of this network is the presence of feature channels which allow the network to propagate context information to higher resolution layers.
- At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes.
- In total the network has 23 convolutional layers.

Choice of Architecture

- The reason why older models don't work so well for segmentation is that localized information is lost in the densely connected layers.
- However, this model solves this problem very well with the upsampling step; so instead of just pooling in which pixelwise information is lost every time, we are using the high resolution pictures for deciding weights of further layers also
- An important modification of this network is the presence of feature channels which allow the network to propagate context information to higher resolution layers.
- At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes.
- In total the network has 23 convolutional layers.

Choice of Architecture

- The reason why older models don't work so well for segmentation is that localized information is lost in the densely connected layers.
- However, this model solves this problem very well with the upsampling step; so instead of just pooling in which pixelwise information is lost every time, we are using the high resolution pictures for deciding weights of further layers also
- An important modification of this network is the presence of feature channels which allow the network to propagate context information to higher resolution layers.
- At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes.
- In total the network has 23 convolutional layers.

Choice of Architecture

- The reason why older models don't work so well for segmentation is that localized information is lost in the densely connected layers.
- However, this model solves this problem very well with the upsampling step; so instead of just pooling in which pixelwise information is lost every time, we are using the high resolution pictures for deciding weights of further layers also
- An important modification of this network is the presence of feature channels which allow the network to propagate context information to higher resolution layers.
- At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes.
- In total the network has 23 convolutional layers.

Choice of Architecture

- The reason why older models don't work so well for segmentation is that localized information is lost in the densely connected layers.
- However, this model solves this problem very well with the upsampling step; so instead of just pooling in which pixelwise information is lost every time, we are using the high resolution pictures for deciding weights of further layers also
- An important modification of this network is the presence of feature channels which allow the network to propagate context information to higher resolution layers.
- At the final layer a 1x1 convolution is used to map each 64-component feature vector to the desired number of classes.
- In total the network has 23 convolutional layers.

Results of Training

- We trained the model for 5 epochs and used the standard method of dropout to prevent overfitting.
- As a comparative study, we implemented the U-Net method but we also used K-means clustering to compare the results. We planned to use graph cuts as well however we were falling short of time
- The results with only the given data were not at all satisfactory, however with data augmentation we got really good results.
- K-means could identify the cell boundaries correctly, however the segmentation was not ideal in the sense that organelles were segmented in a different class than the cell cytoplasm which is not an ideal scenario.

Results of Training

- We trained the model for 5 epochs and used the standard method of dropout to prevent overfitting.
- As a comparative study, we implemented the U-Net method but we also used K-means clustering to compare the results. We planned to use graph cuts as well however we were falling short of time
- The results with only the given data were not at all satisfactory, however with data augmentation we got really good results.
- K-means could identify the cell boundaries correctly, however the segmentation was not ideal in the sense that organelles were segmented in a different class than the cell cytoplasm which is not an ideal scenario.

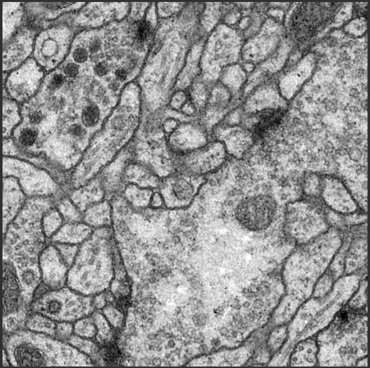
Results of Training

- We trained the model for 5 epochs and used the standard method of dropout to prevent overfitting.
- As a comparative study, we implemented the U-Net method but we also used K-means clustering to compare the results. We planned to use graph cuts as well however we were falling short of time
- The results with only the given data were not at all satisfactory, however with data augmentation we got really good results.
- K-means could identify the cell boundaries correctly, however the segmentation was not ideal in the sense that organelles were segmented in a different class than the cell cytoplasm which is not an ideal scenario.

Results of Training

- We trained the model for 5 epochs and used the standard method of dropout to prevent overfitting.
- As a comparative study, we implemented the U-Net method but we also used K-means clustering to compare the results. We planned to use graph cuts as well however we were falling short of time
- The results with only the given data were not at all satisfactory, however with data augmentation we got really good results.
- K-means could identify the cell boundaries correctly, however the segmentation was not ideal in the sense that organelles were segmented in a different class than the cell cytoplasm which is not an ideal scenario.

Result of K-Means segmentation



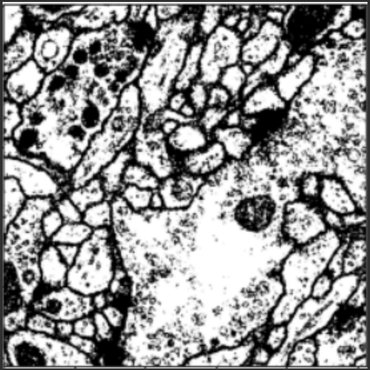


Figure: The results of K-means segmentation after 30 epochs

Result of K-Means segmentation

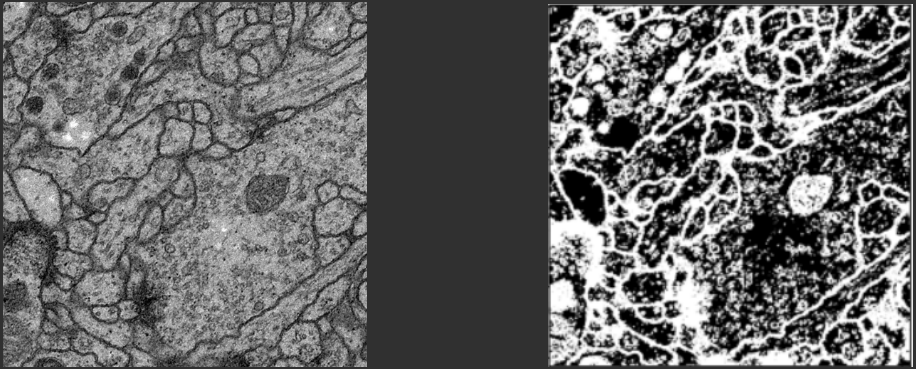
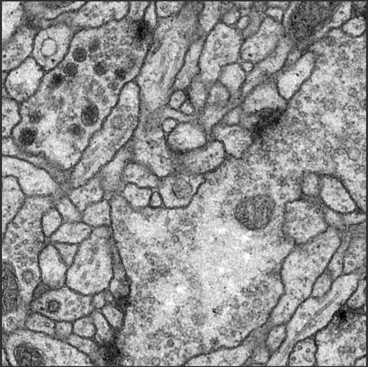


Figure: The results of K-means segmentation after 30 epochs

Result of Neural Network Prediction



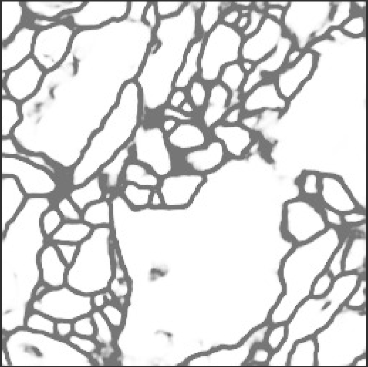


Figure: The results of U-Net segmentation after 5 epochs

Result of Neural Network Prediction

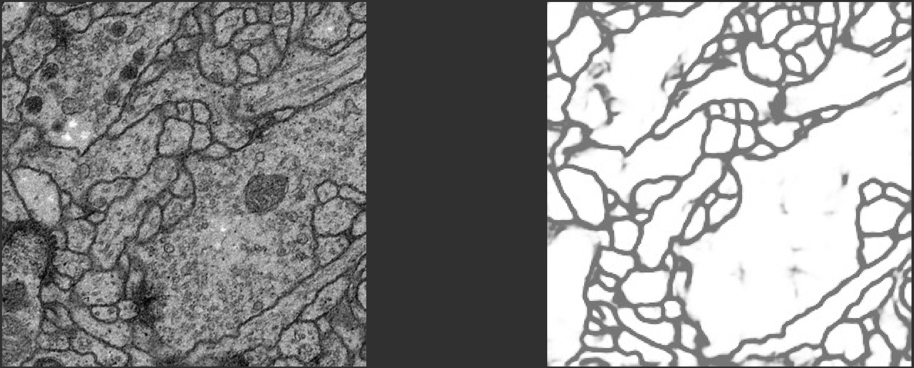


Figure: The results of K-means segmentation after 5 epochs

Results: Comparison

- The loss function used by us was **binary cross entropy loss** and the final loss after 5 epochs is 0.22747. It was obvious that increasing
- The accuracy at the end of 5 epochs was 0.9243. The accuracy on training set was not as high as we expected but we didn't increase the number of epochs to avoid overfitting.
- For K-means, the accuracy was same even after 15 or 30 epochs which means that the best fit was achieved. The results with K-Means were not very satisfactory even for hard binary segmentation which highlights the inefficiencies of the method.

Results: Comparison

- The loss function used by us was **binary cross entropy loss** and the final loss after 5 epochs is 0.22747. It was obvious that increasing
- The accuracy at the end of 5 epochs was 0.9243. The accuracy on training set was not as high as we expected but we didn't increase the number of epochs to avoid overfitting.
- For K-means, the accuracy was same even after 15 or 30 epochs which means that the best fit was achieved. The results with K-Means were not very satisfactory even for hard binary segmentation which highlights the inefficiencies of the method.

Results: Comparison

- The loss function used by us was **binary cross entropy loss** and the final loss after 5 epochs is 0.22747. It was obvious that increasing
- The accuracy at the end of 5 epochs was 0.9243. The accuracy on training set was not as high as we expected but we didn't increase the number of epochs to avoid overfitting.
- For K-means, the accuracy was same even after 15 or 30 epochs which means that the best fit was achieved. The results with K-Means were not very satisfactory even for hard binary segmentation which highlights the inefficiencies of the method.

Results of the 2 methods

- The main problem with K-means is that it only takes into account intensities and not neighborhood information while deciding classes.
- This problem is not resolved in some of the modifications to K-means also.
- The main advantage of Neural Networks is that spatial information is contained in the feature channels and the model 'learns' the intensity information, spatial information and it also takes into account not just the difference in intensities but also the results of actual labelled segmentation.
- The graph s-t cuts method can be explored for this problem because of it's superfast convergence rate and high accuracy

Results of the 2 methods

- The main problem with K-means is that it only takes into account intensities and not neighborhood information while deciding classes.
- This problem is not resolved in some of the modifications to K-means also.
- The main advantage of Neural Networks is that spatial information is contained in the feature channels and the model 'learns' the intensity information, spatial information and it also takes into account not just the difference in intensities but also the results of actual labelled segmentation.
- The graph s-t cuts method can be explored for this problem because of it's superfast convergence rate and high accuracy

Results of the 2 methods

- The main problem with K-means is that it only takes into account intensities and not neighborhood information while deciding classes.
- This problem is not resolved in some of the modifications to K-means also.
- The main advantage of Neural Networks is that spatial information is contained in the feature channels and the model 'learns' the intensity information, spatial information and it also takes into account not just the difference in intensities but also the results of actual labelled segmentation.
- The graph s-t cuts method can be explored for this problem because of it's superfast convergence rate and high accuracy

Results of the 2 methods

- The main problem with K-means is that it only takes into account intensities and not neighborhood information while deciding classes.
- This problem is not resolved in some of the modifications to K-means also.
- The main advantage of Neural Networks is that spatial information is contained in the feature channels and the model ‘learns’ the intensity information, spatial information and it also takes into account not just the difference in intensities but also the results of actual labelled segmentation.
- The graph s-t cuts method can be explored for this problem because of it’s superfast convergence rate and high accuracy

Challenges and Further ideas

- The greatest challenge for the Neural network was finding labelled test data. Construction of the model was a little difficult.
- Since the training data was too small, we had to use Keras' data augmentation techniques to get more data.
- For further projects, one can try to improve the model and explore optimizations such as U-Net++, reduced U-Net and make the model powerful enough to handle a large number of classes.
- We wish to explore the graph cut method for the same and look for methods to implement multi-class segmentation with graph cut methods and also explore the Gaussian Mixed Model fitting method.

Challenges and Further ideas

- The greatest challenge for the Neural network was finding labelled test data. Construction of the model was a little difficult.
- Since the training data was too small, we had to use Keras' data augmentation techniques to get more data.
- For further projects, one can try to improve the model and explore optimizations such as U-Net++, reduced U-Net and make the model powerful enough to handle a large number of classes.
- We wish to explore the graph cut method for the same and look for methods to implement multi-class segmentation with graph cut methods and also explore the Gaussian Mixed Model fitting method.

Challenges and Further ideas

- The greatest challenge for the Neural network was finding labelled test data. Construction of the model was a little difficult.
- Since the training data was too small, we had to use Keras' data augmentation techniques to get more data.
- For further projects, one can try to improve the model and explore optimizations such as U-Net++, reduced U-Net and make the model powerful enough to handle a large number of classes.
- We wish to explore the graph cut method for the same and look for methods to implement multi-class segmentation with graph cut methods and also explore the Gaussian Mixed Model fitting method.

Challenges and Further ideas

- The greatest challenge for the Neural network was finding labelled test data. Construction of the model was a little difficult.
- Since the training data was too small, we had to use Keras' data augmentation techniques to get more data.
- For further projects, one can try to improve the model and explore optimizations such as U-Net++, reduced U-Net and make the model powerful enough to handle a large number of classes.
- We wish to explore the graph cut method for the same and look for methods to implement multi-class segmentation with graph cut methods and also explore the Gaussian Mixed Model fitting method.

Thank you

THANK YOU!