

challenge1

April 8, 2018

Author: Tyler Chase

Date: 04/08/2018

1 Coding Challenge 1

1.1 Part 1: Parse the DICOM Images and Contour Files

In [1]: # Parse all DICOM Images and Contour Files Pipe to Numpy Arrays

```
# Suppress warning due to old version of pydicom
import warnings
warnings.filterwarnings('ignore')

from parsing import Pipe

dataAddress = '/Users/tylerchase/Documents/professional/job_applications_2018/' + \
              'Arterys/challenges/challengePhase1/data/final_data'
output = Pipe(dataAddress)
```

In [2]: # Some unit tests of the contour parsing

```
import numpy as np

print('Count', 'Height', 'Width')
print('DCM Array Shape:', np.shape(output.dcmArray))
print('Contour Array Shape:', np.shape(output.contourArray))
print('Label Array Shape:', np.shape(output.labels))

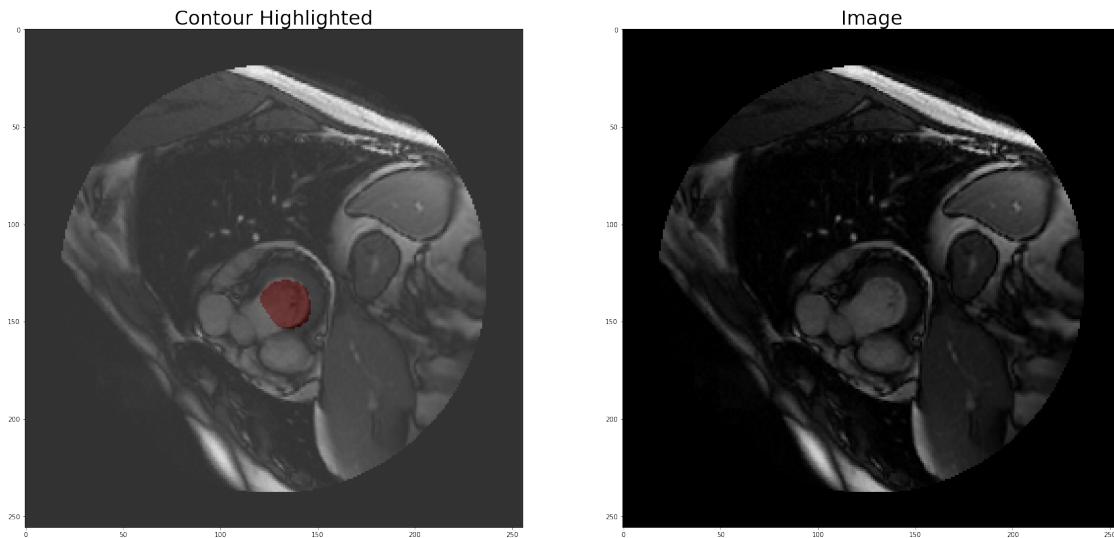
print('\nContour Max:', int(np.amax(output.contourArray)))
print('Contour Min:', int(np.amin(output.contourArray)))

('Count', 'Height', 'Width')
DCM Array Shape: (96, 256, 256)
Contour Array Shape: (96, 256, 256)
Label Array Shape: (96,)

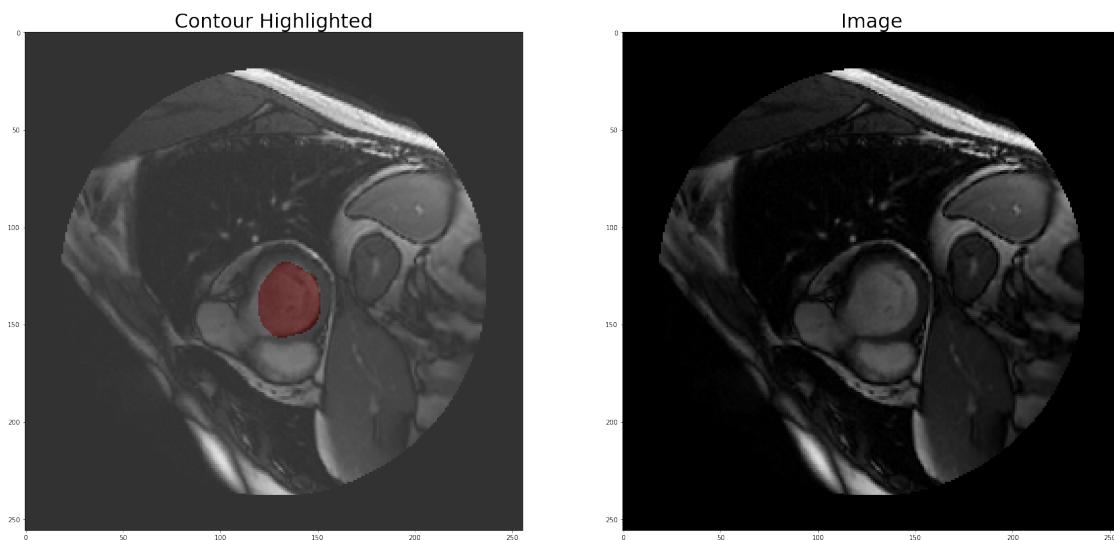
Contour Max: 1
Contour Min: 0
```

```
In [3]: # Plot contours over DICOM images as a sanity check (unit test)
output.plotPipedData()
```

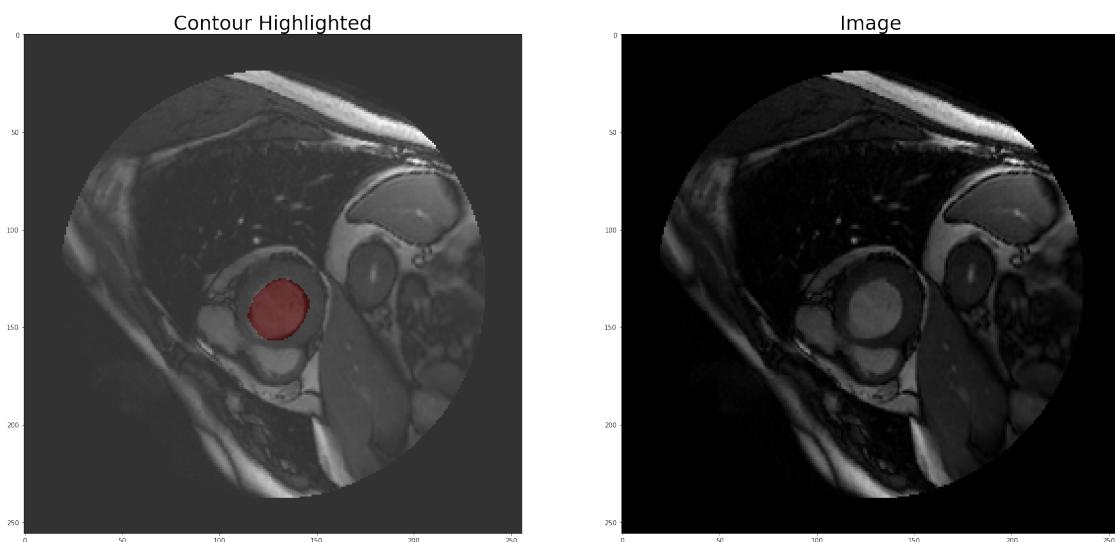
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 48



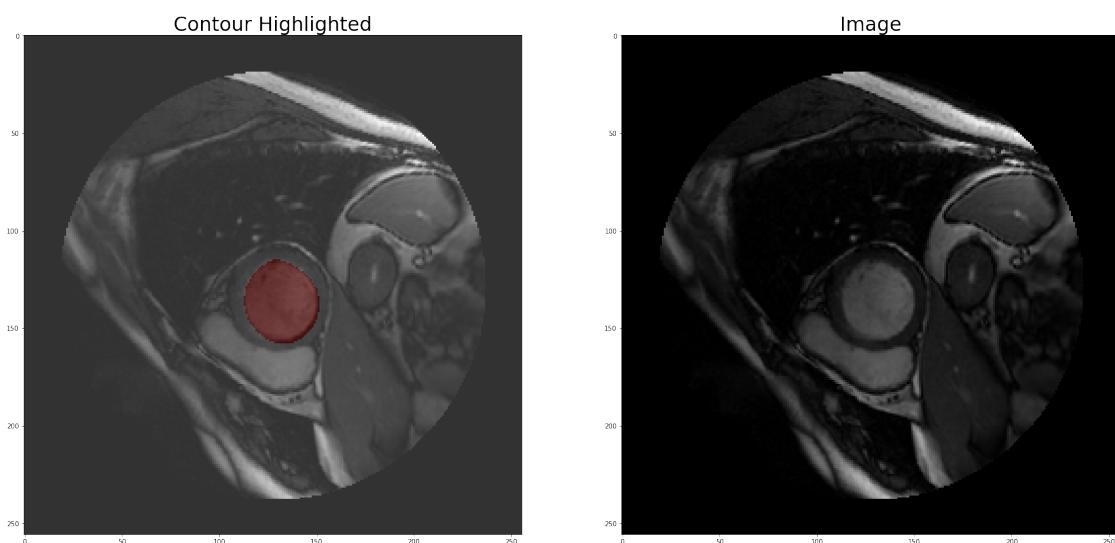
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 59



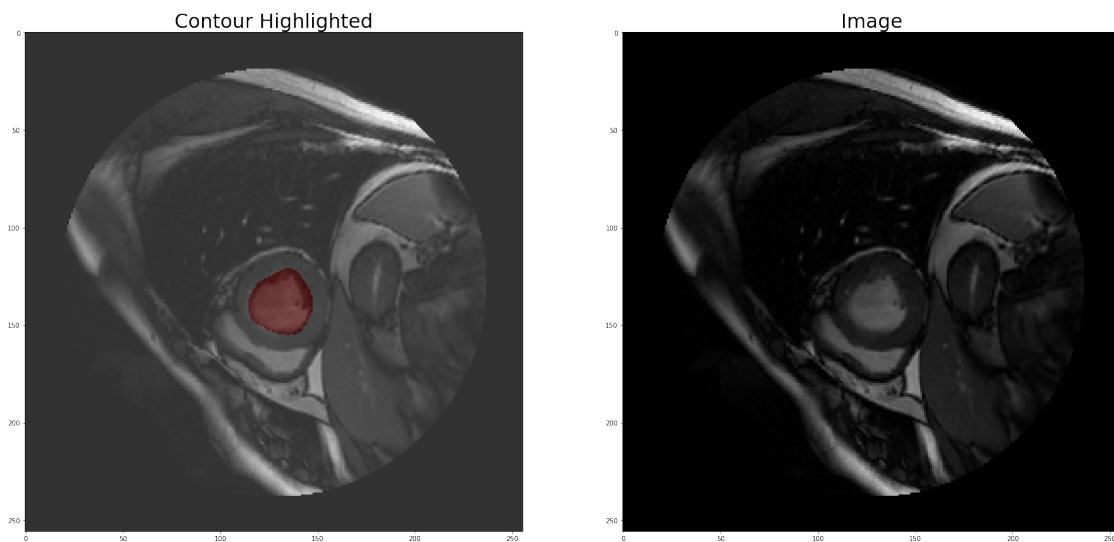
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 68



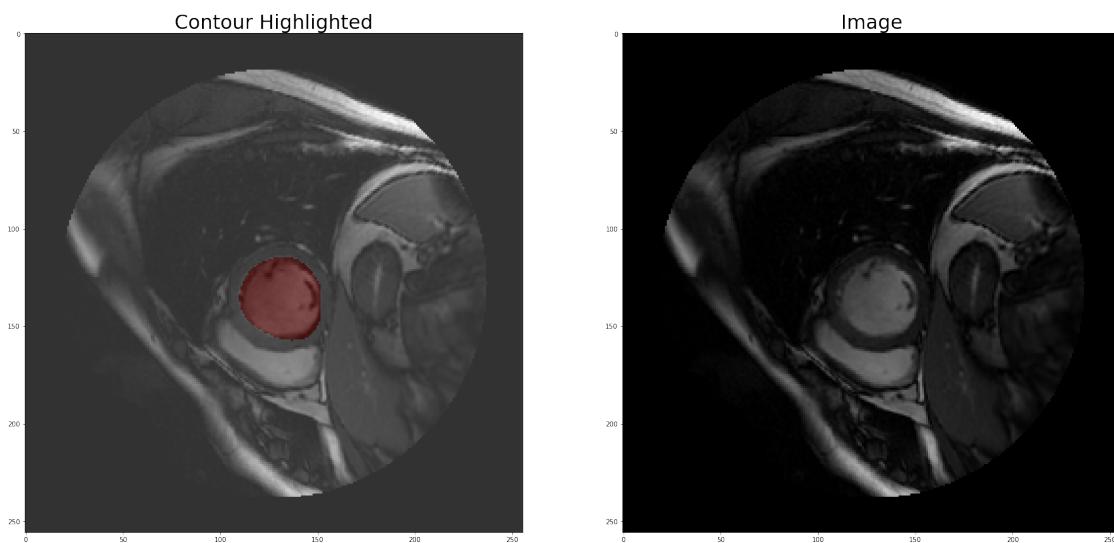
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 79



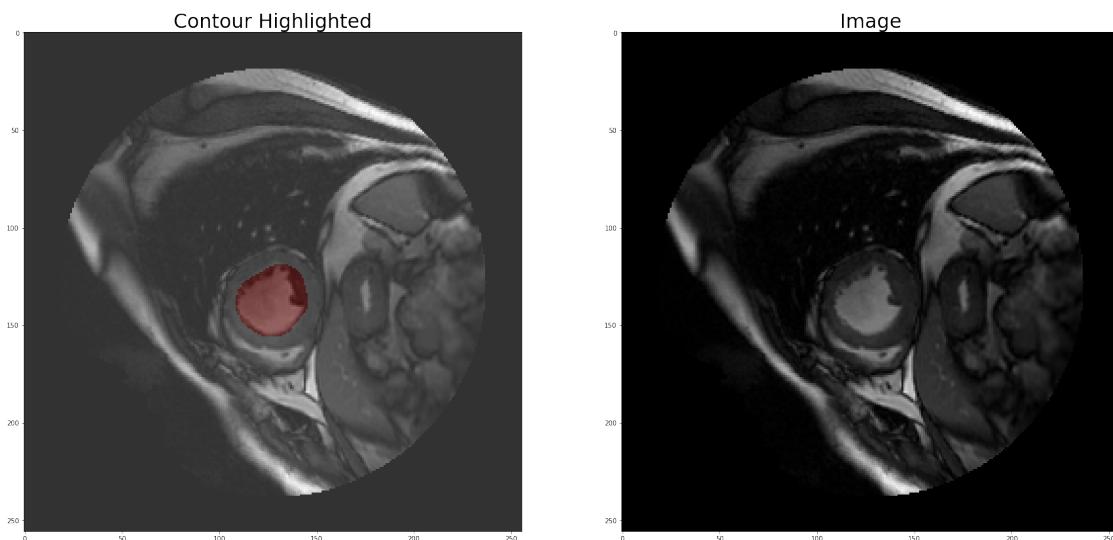
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 88



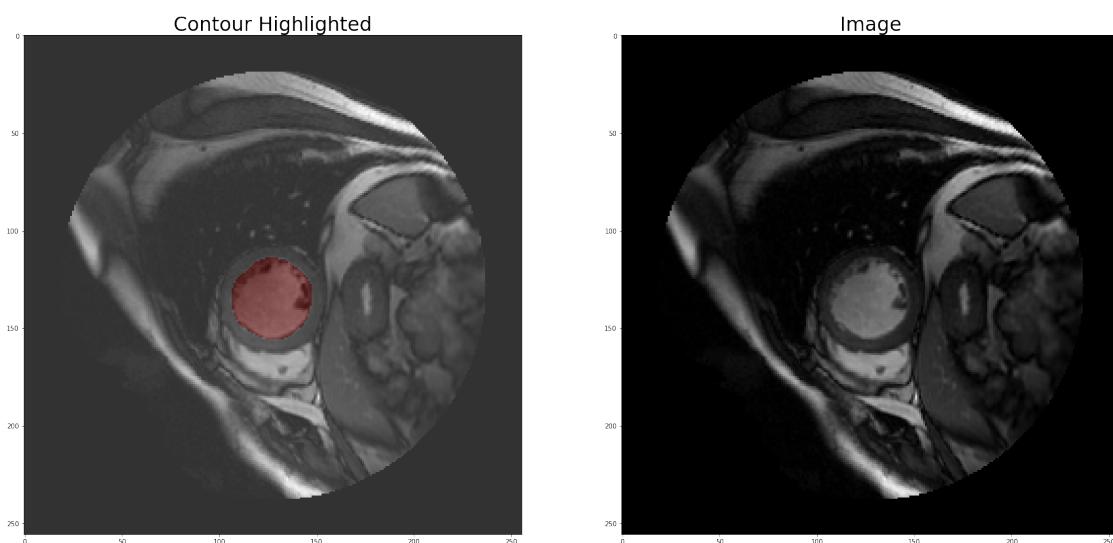
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 99



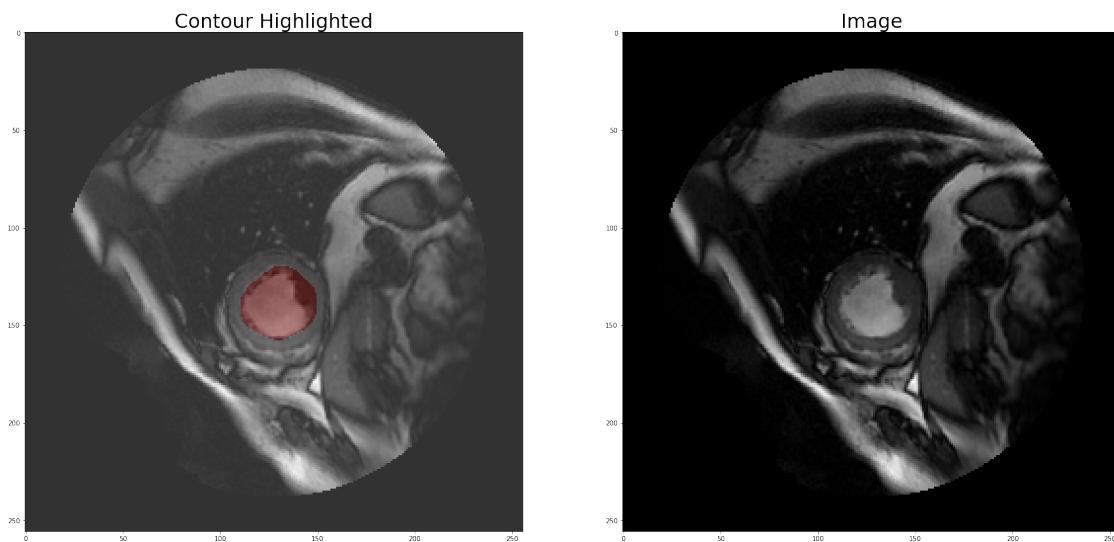
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 108



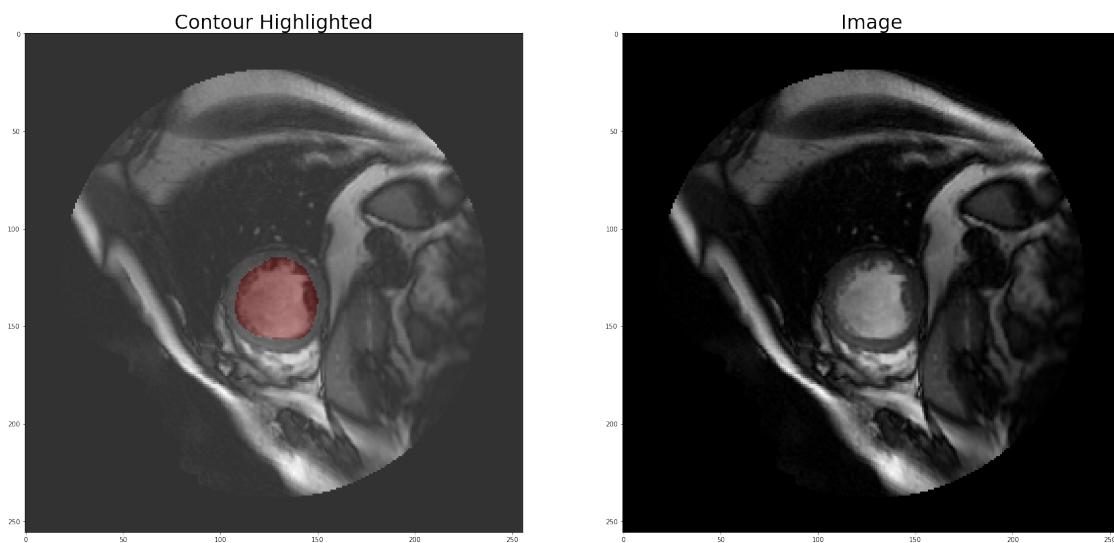
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 119



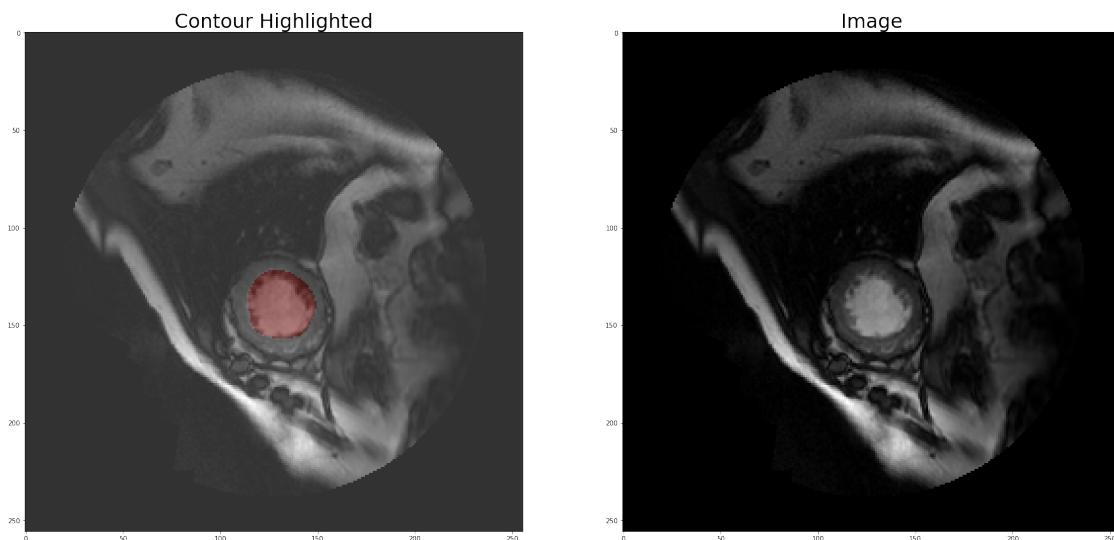
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 128



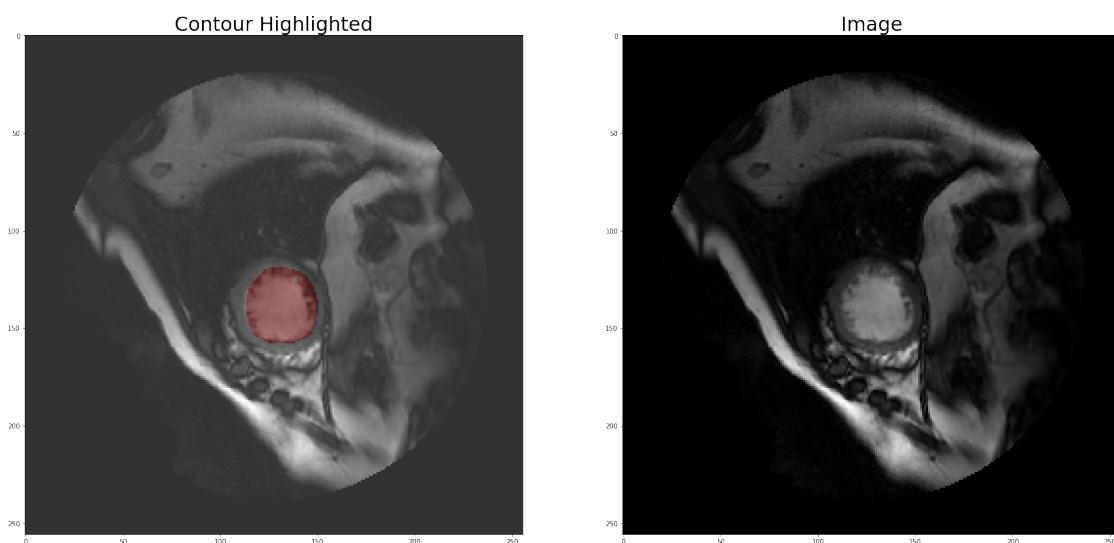
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 139



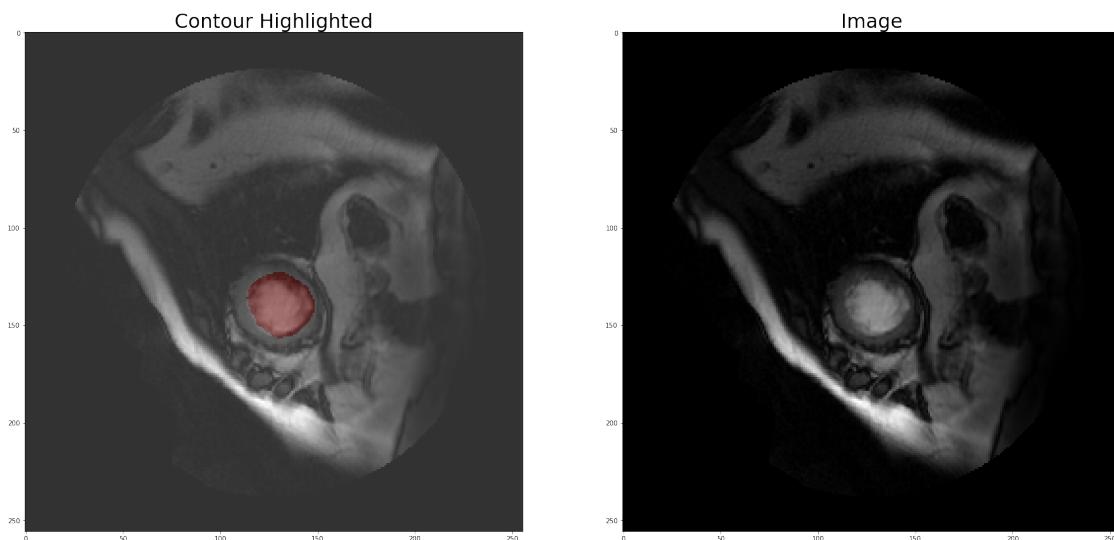
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 148



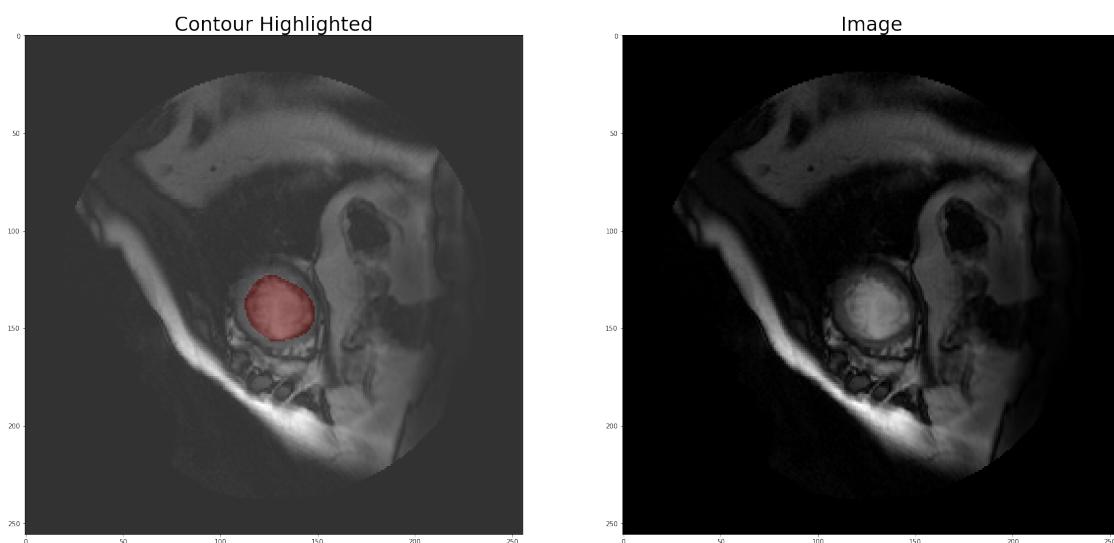
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 159



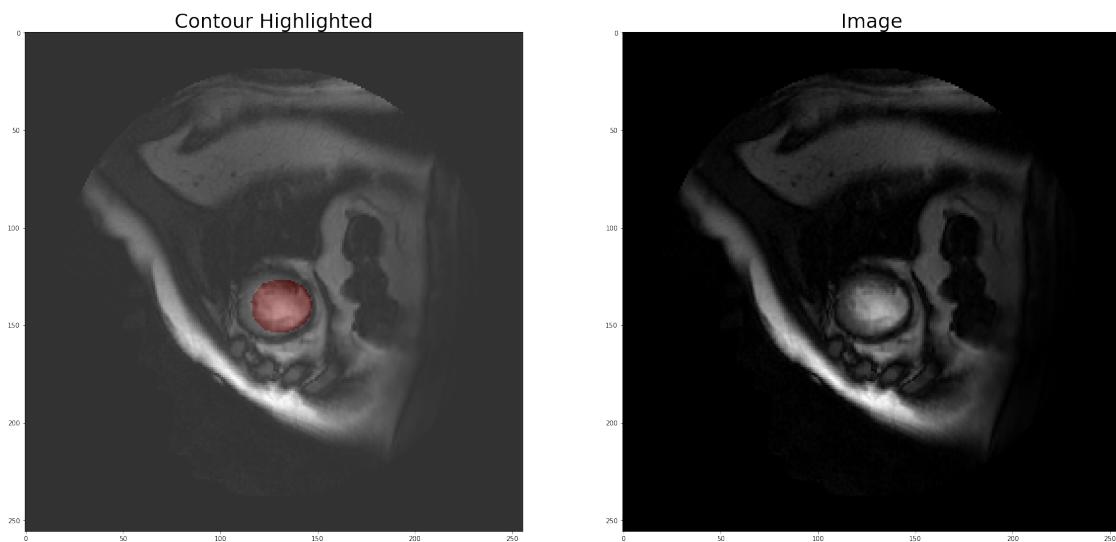
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 168



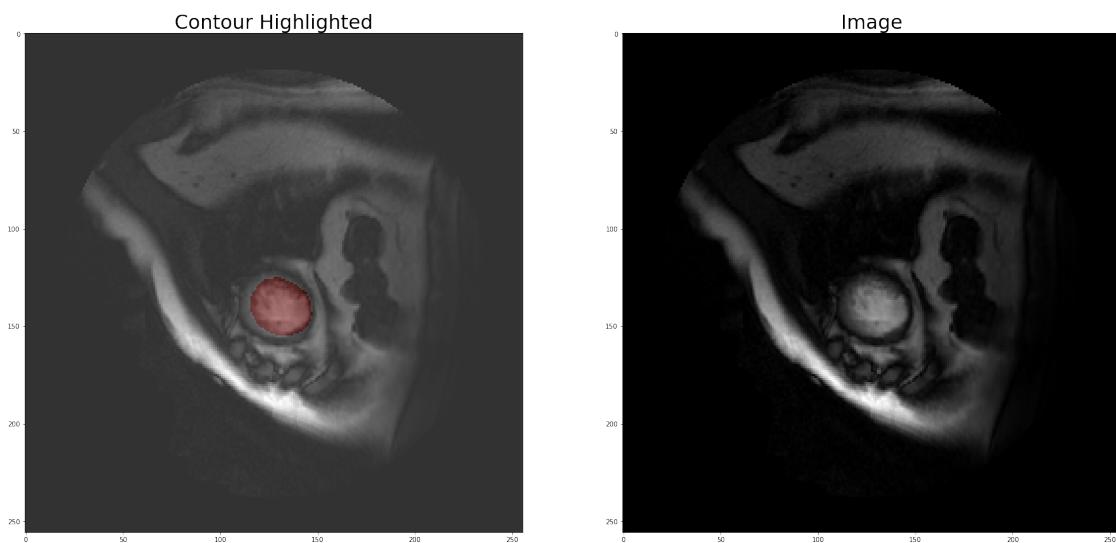
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 179



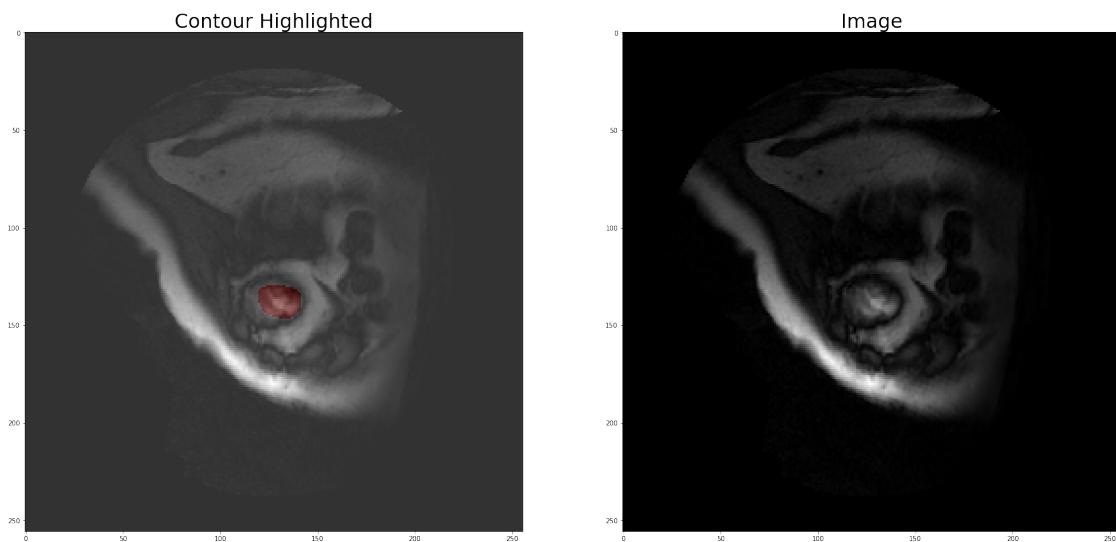
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 188



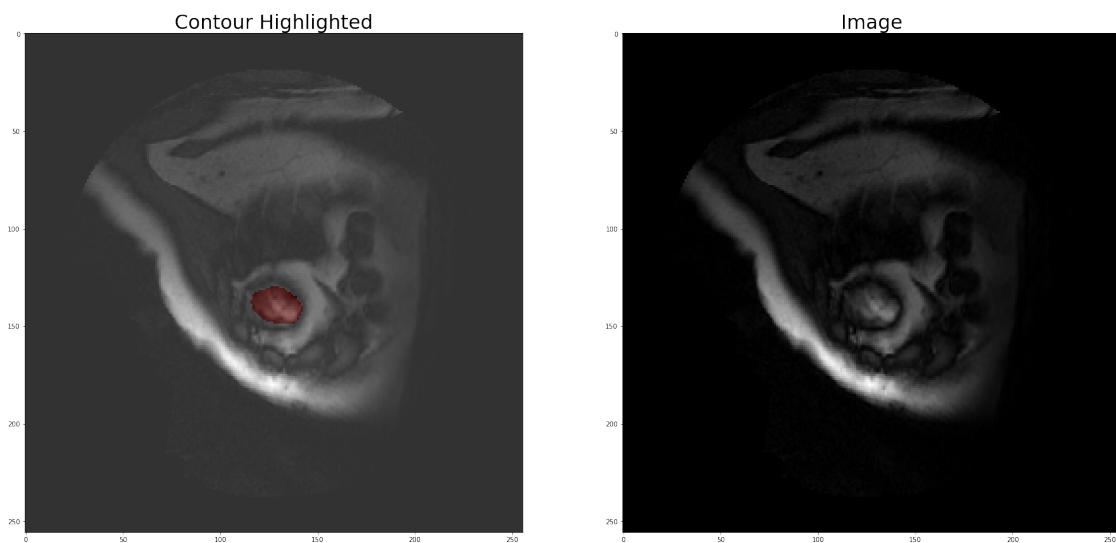
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 199



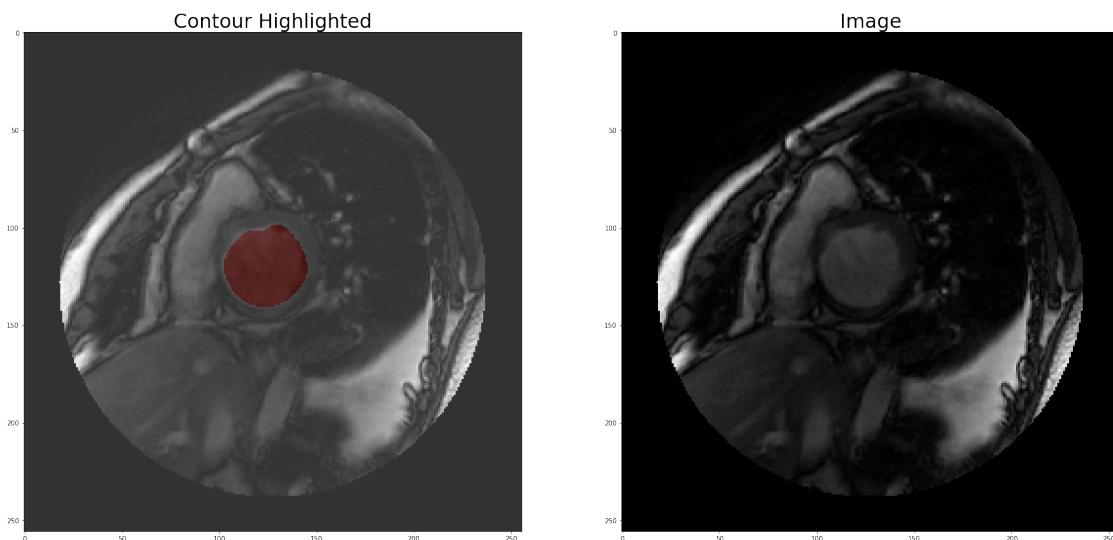
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 208



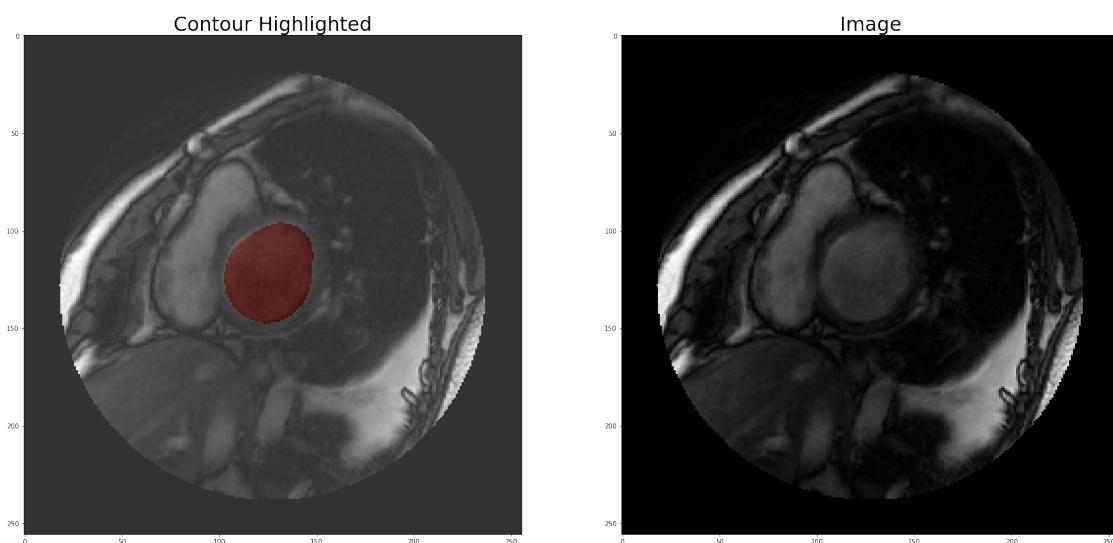
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 219



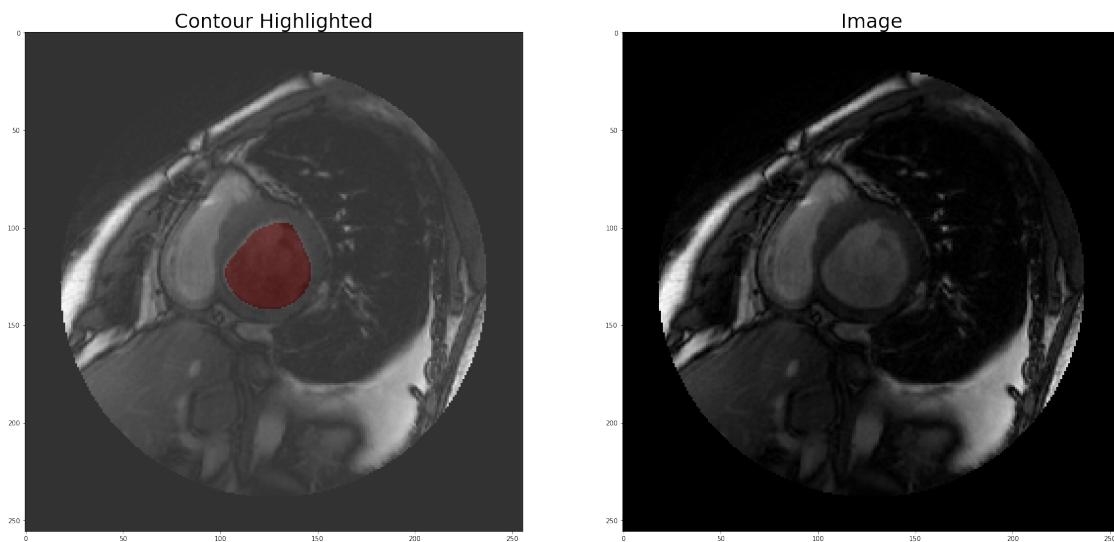
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 47



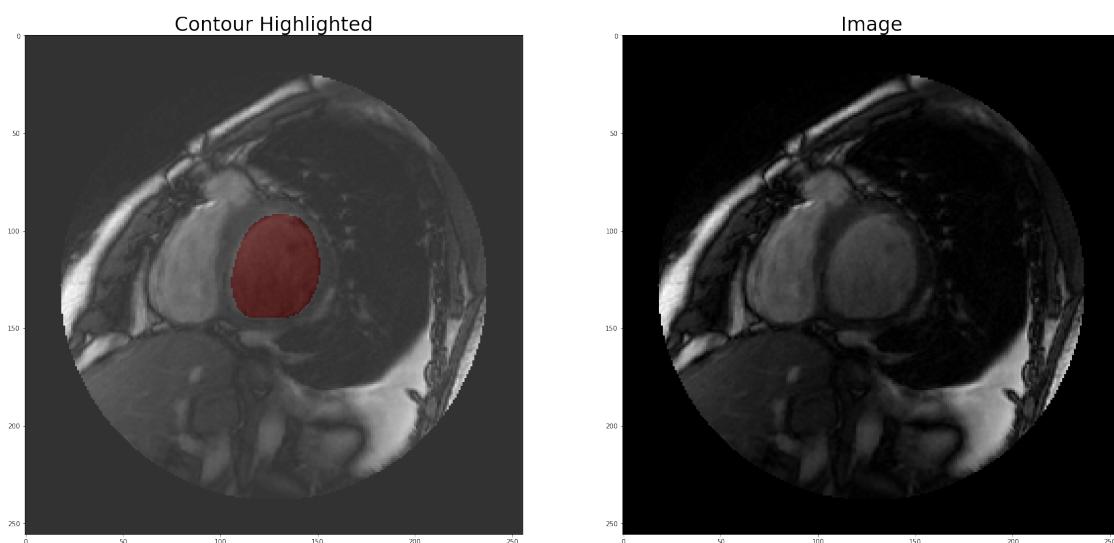
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 60



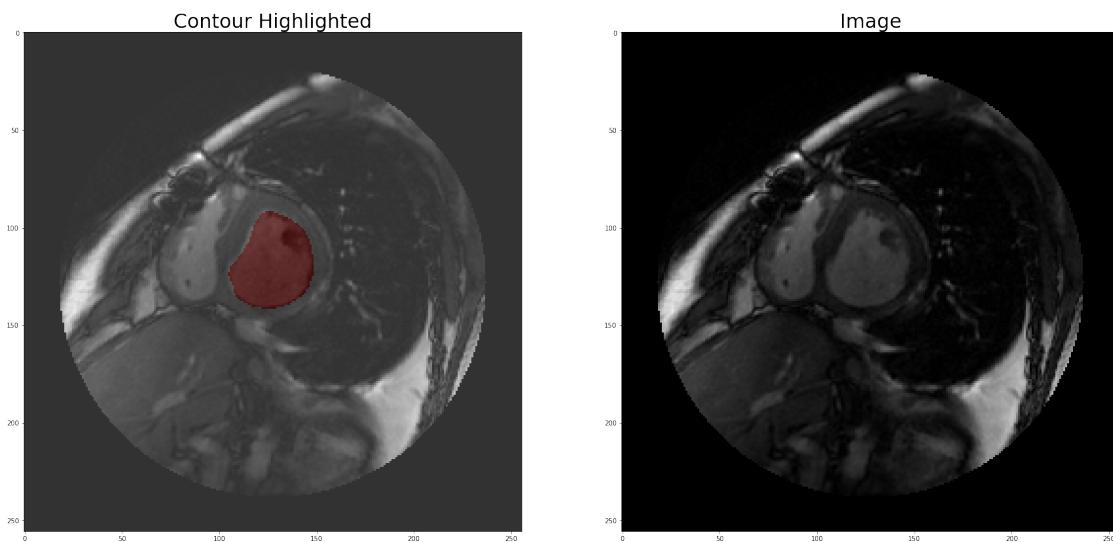
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 67



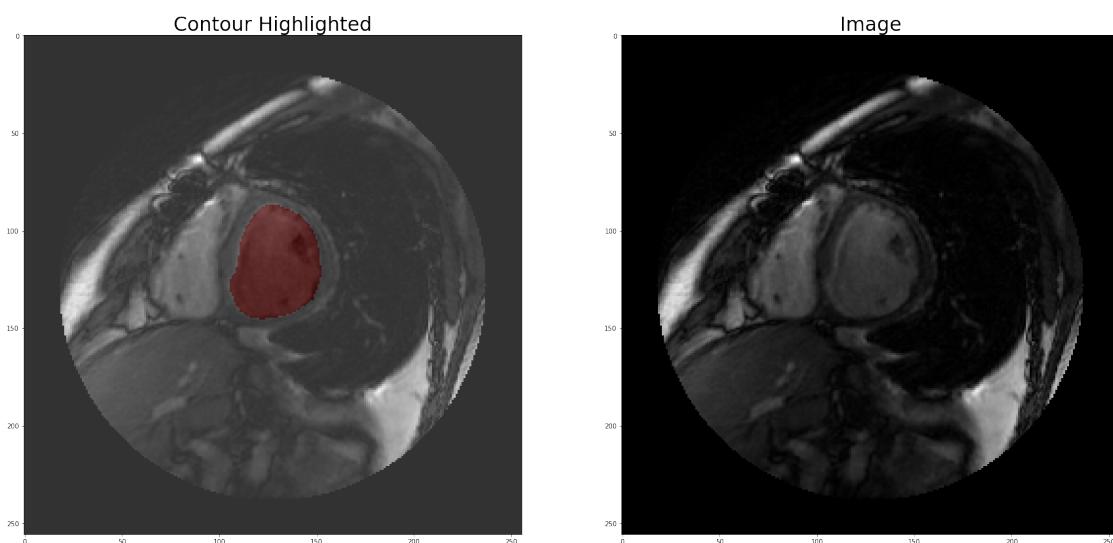
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 80



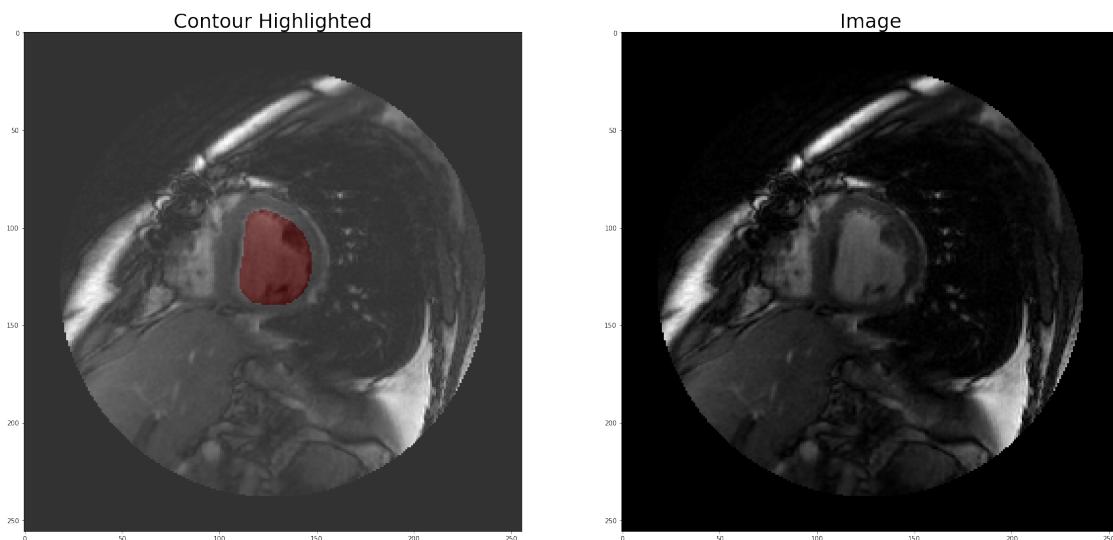
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 87



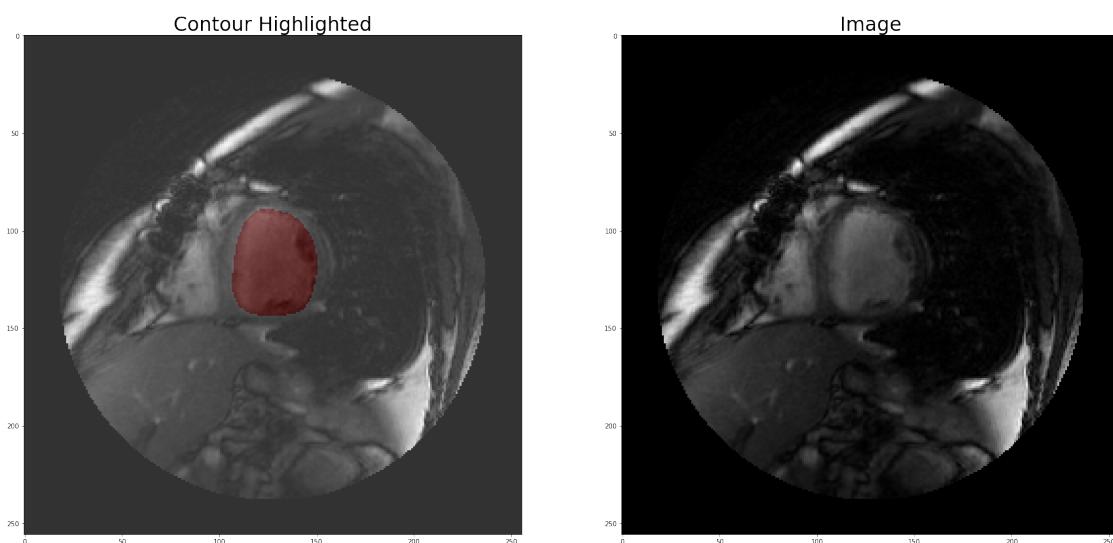
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 100



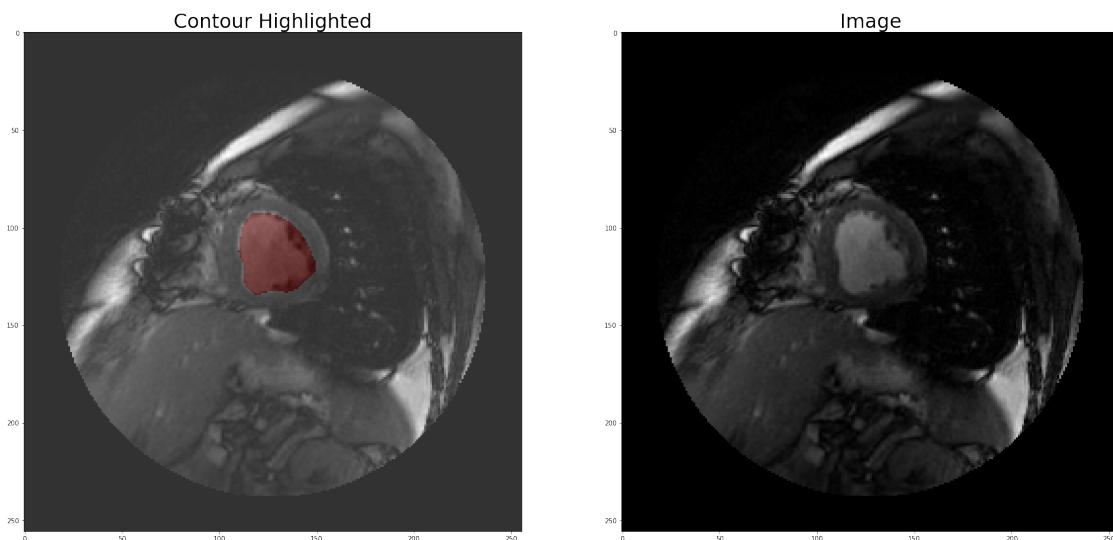
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 107



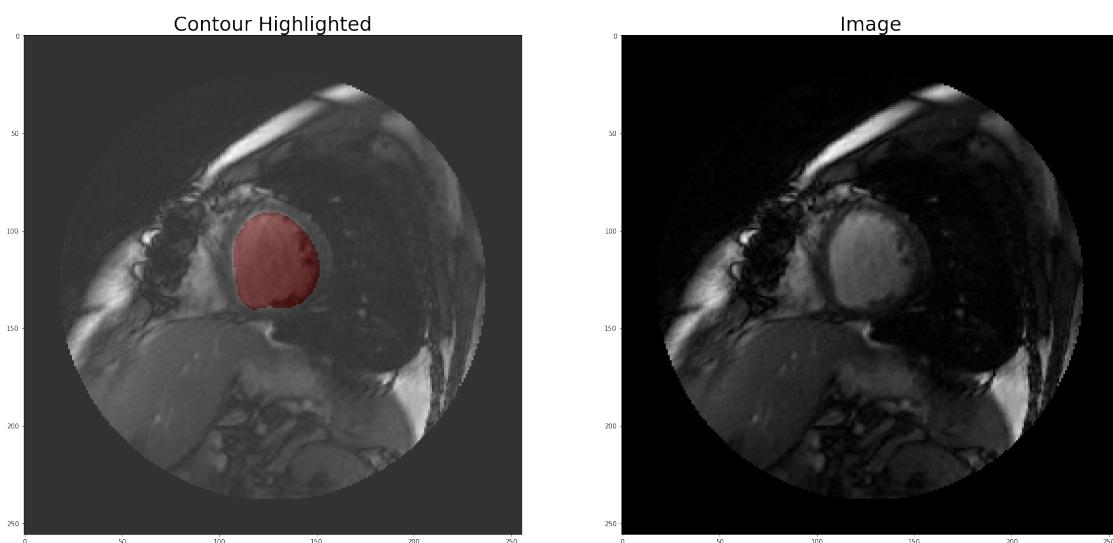
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 120



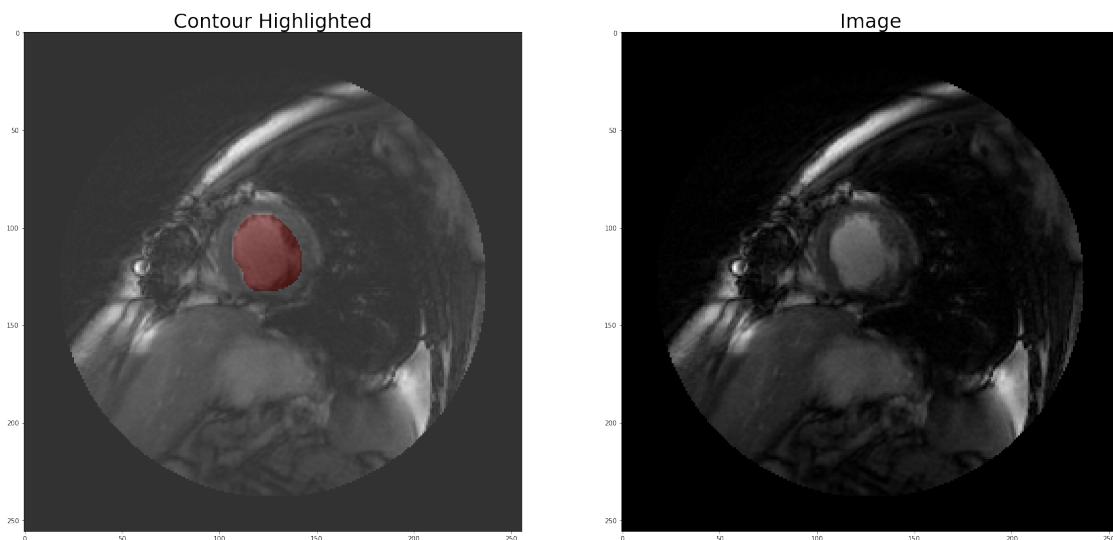
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 127



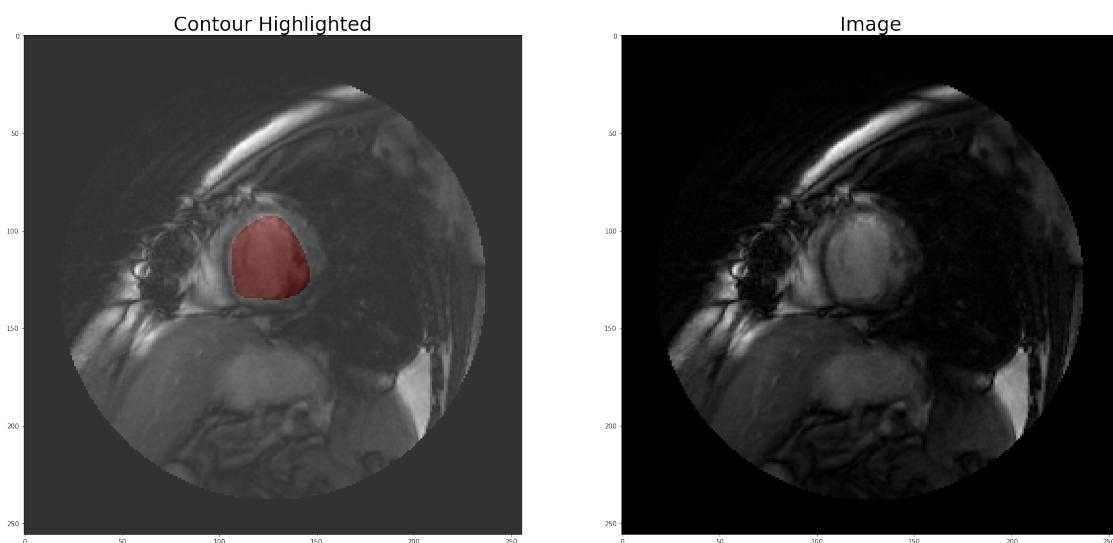
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 140



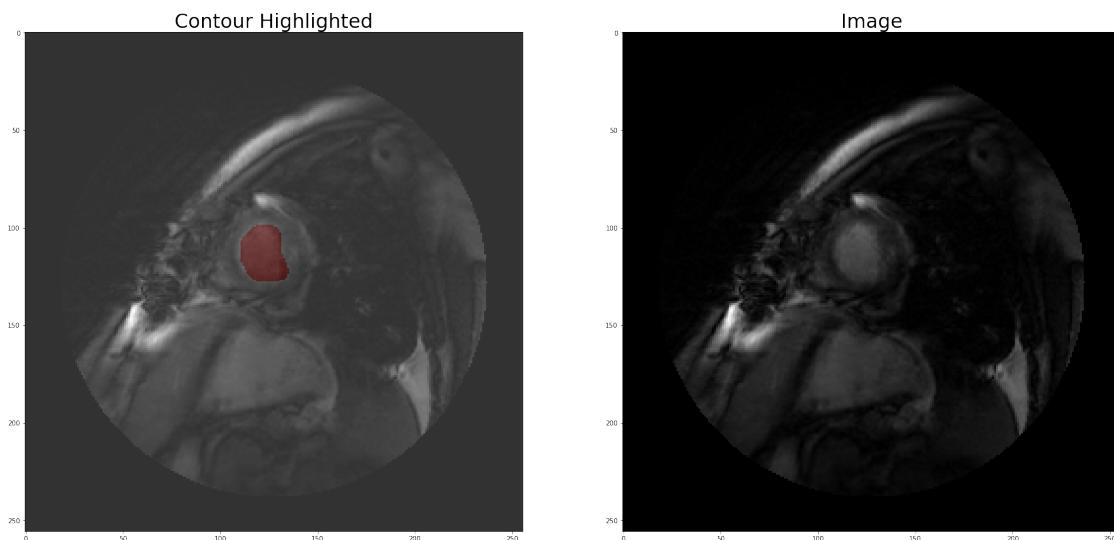
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 147



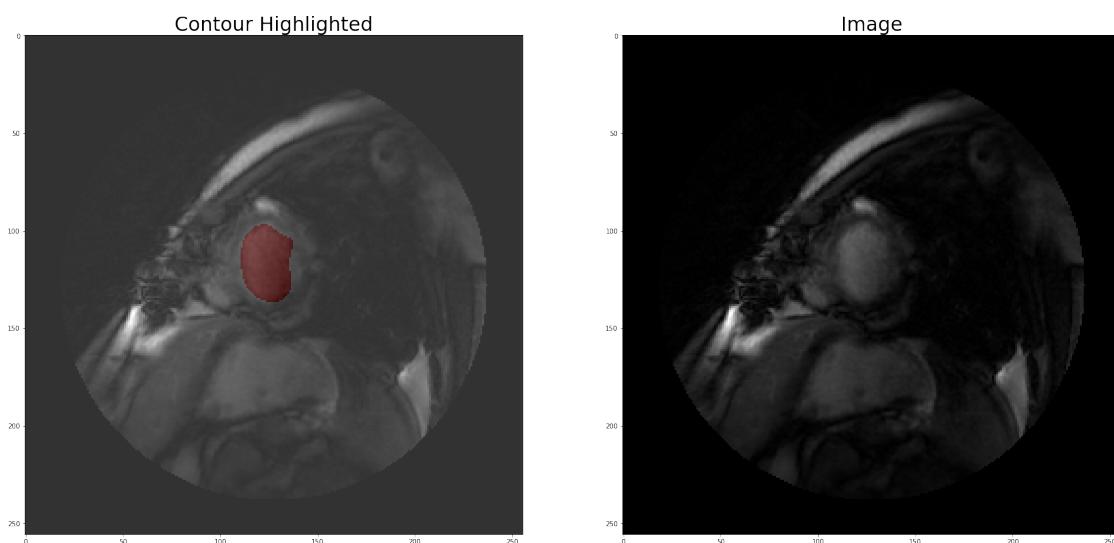
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 160



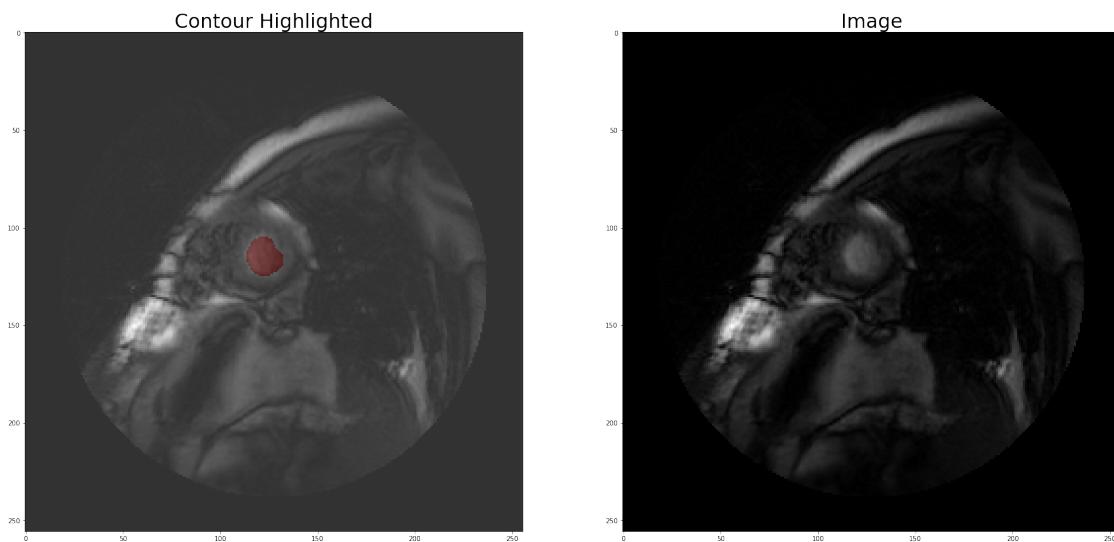
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 167



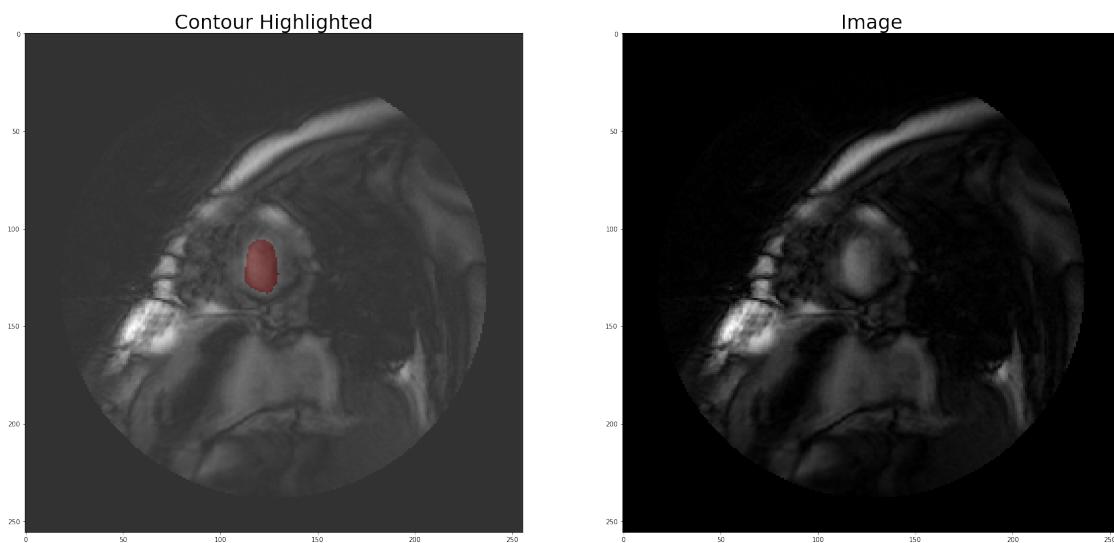
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 180



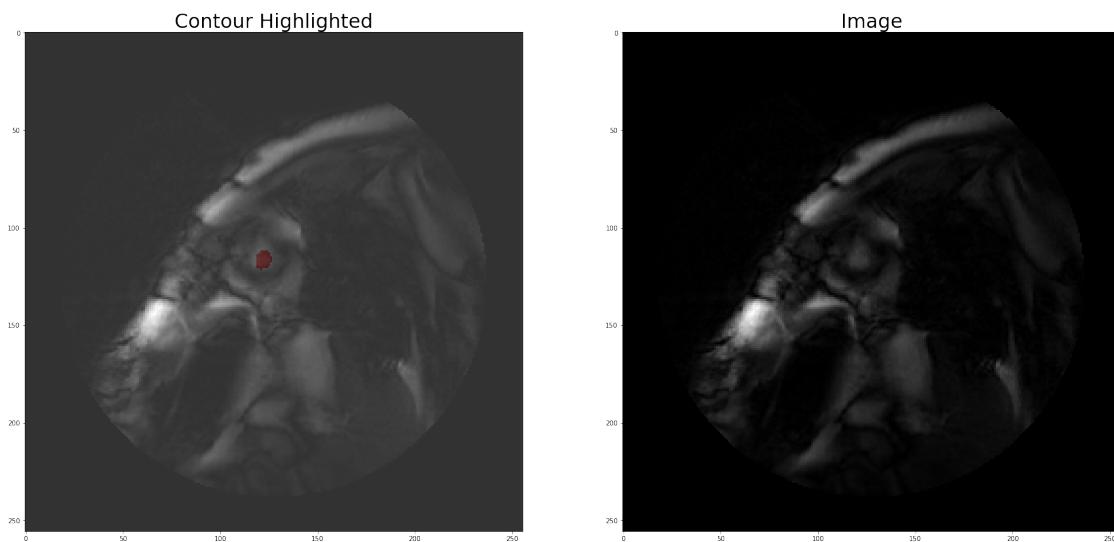
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 187



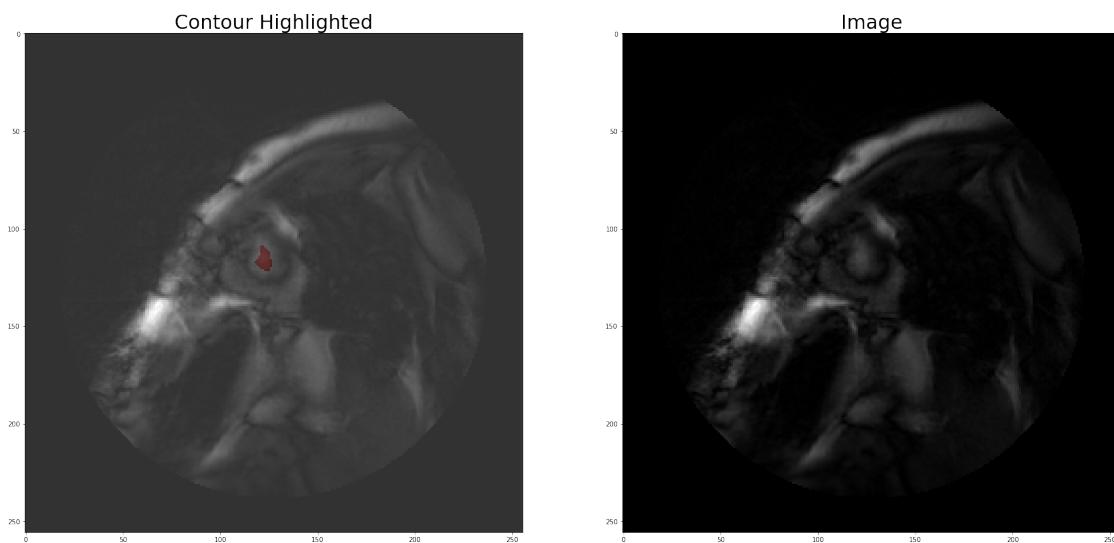
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 200



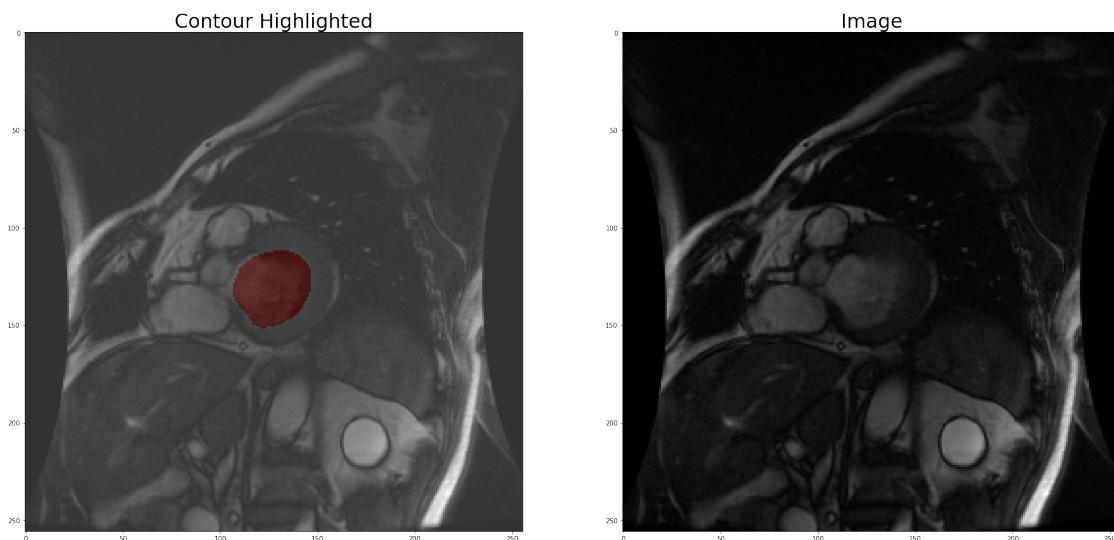
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 207



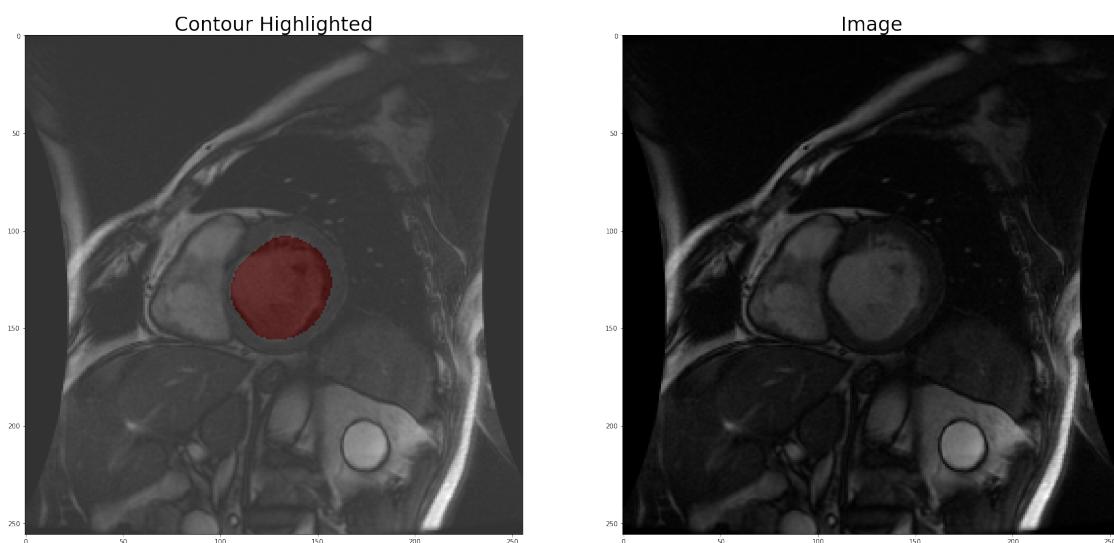
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 220



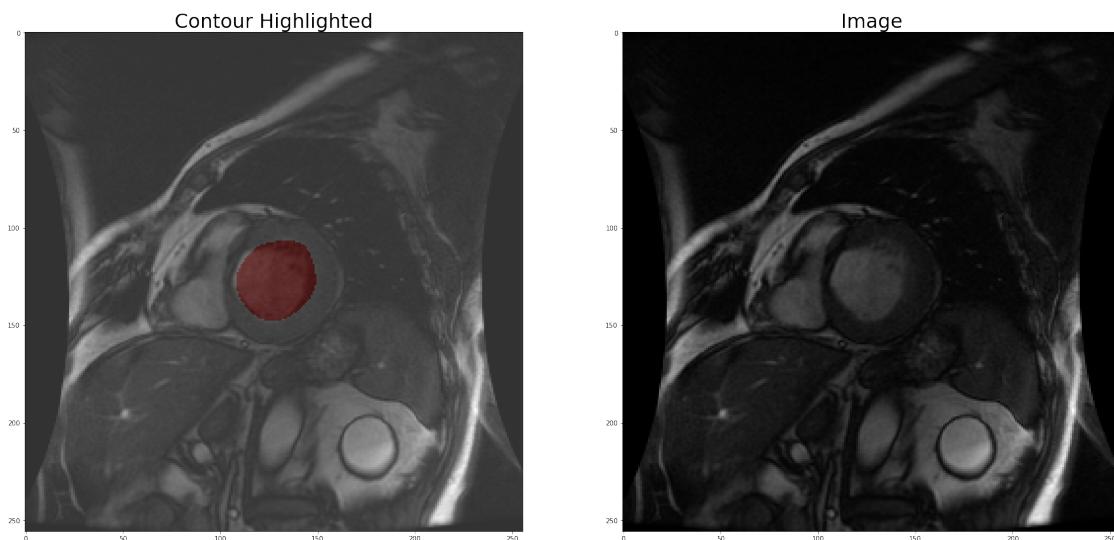
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 7



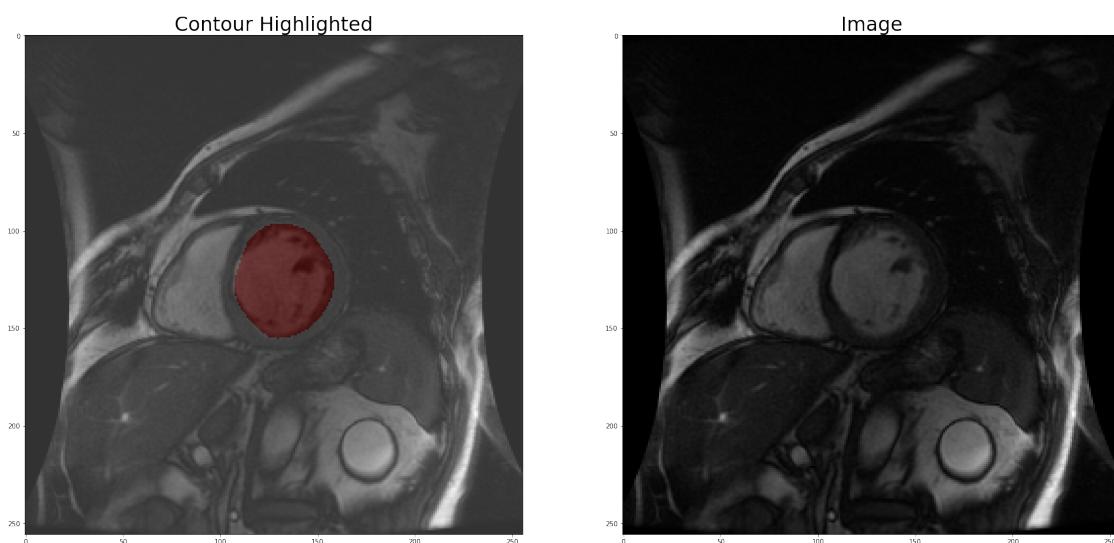
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 20



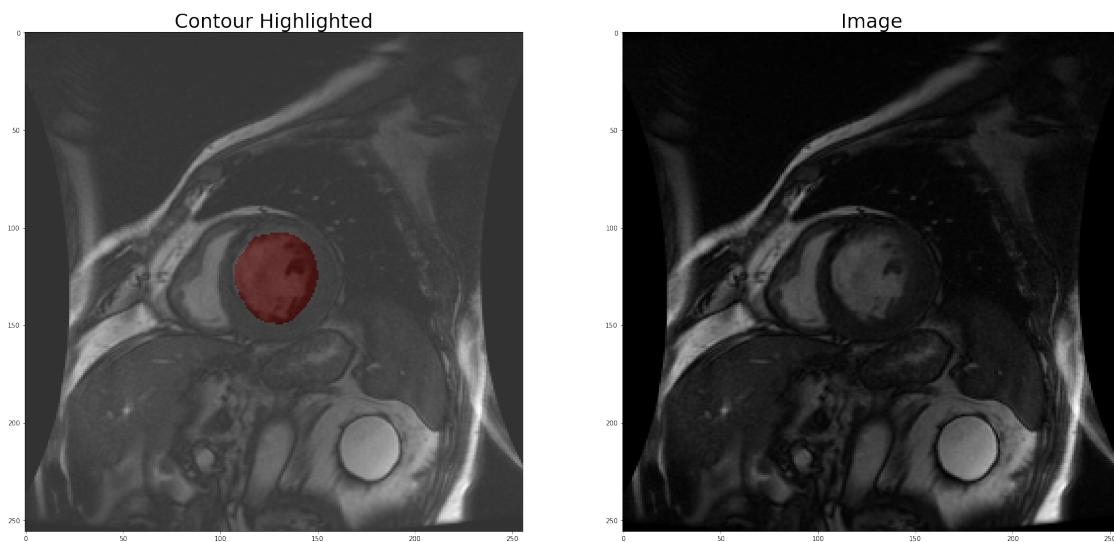
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 27



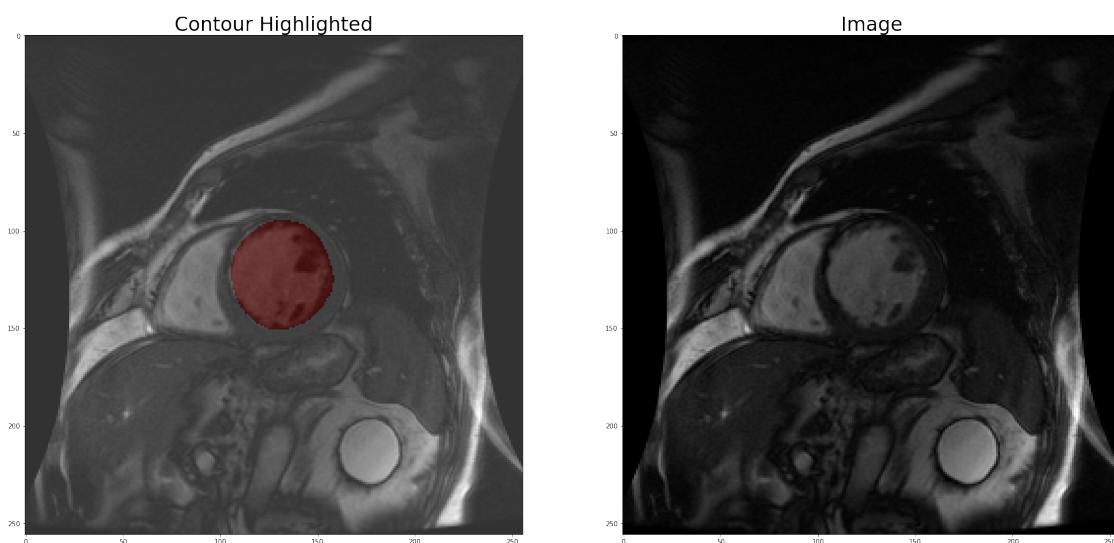
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 40



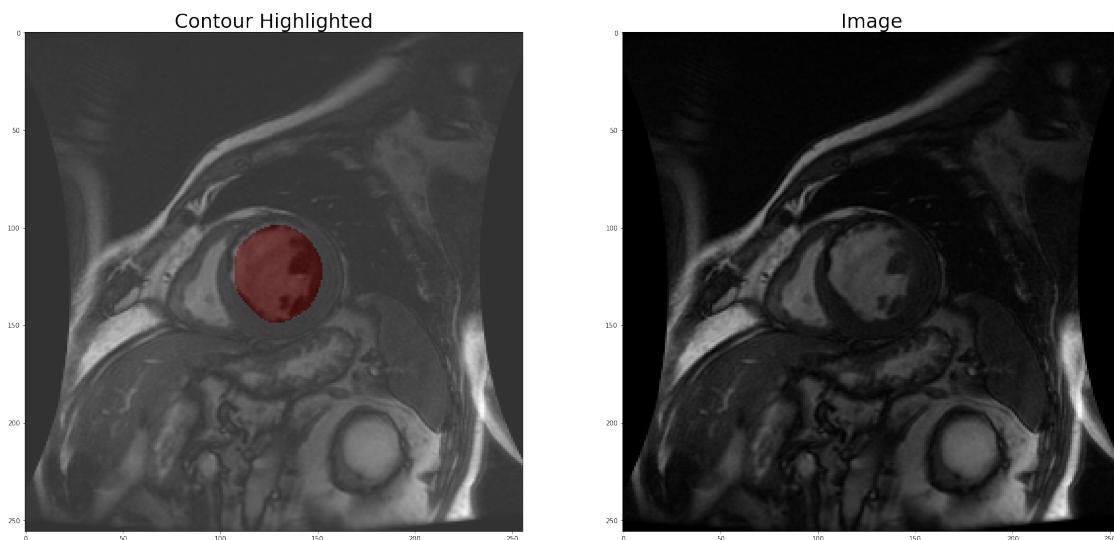
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 47



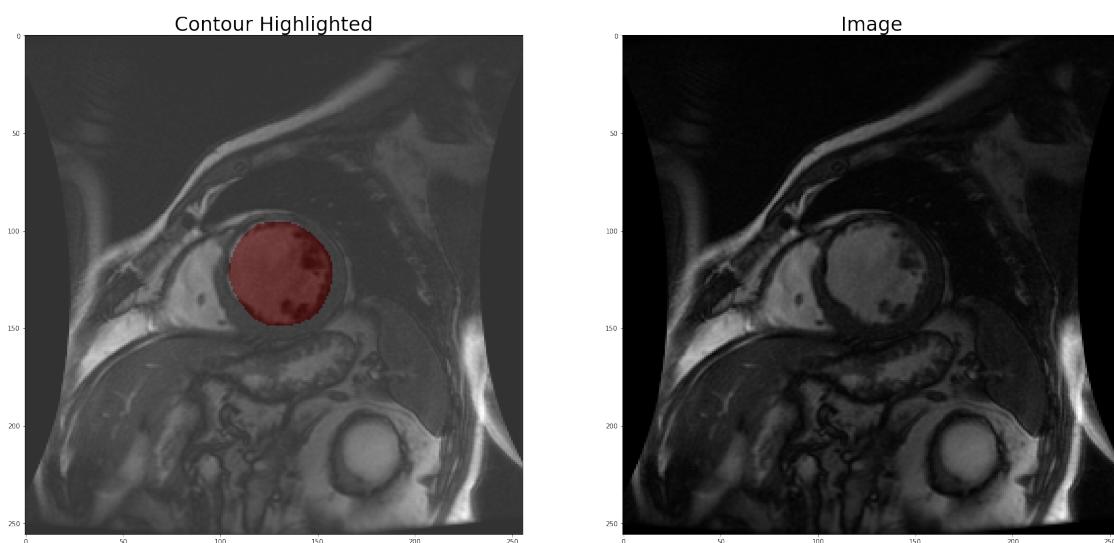
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 60



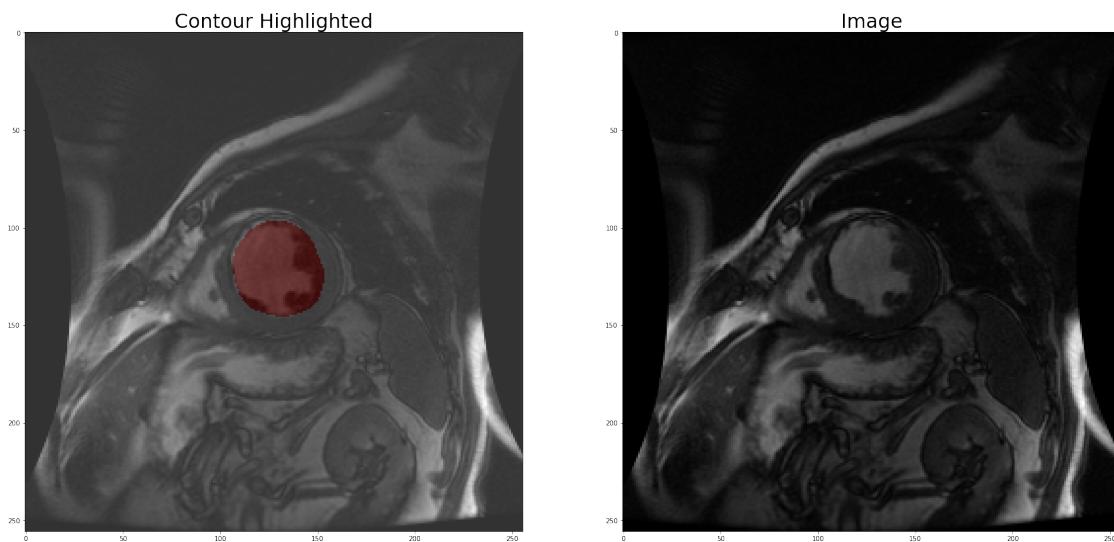
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 67



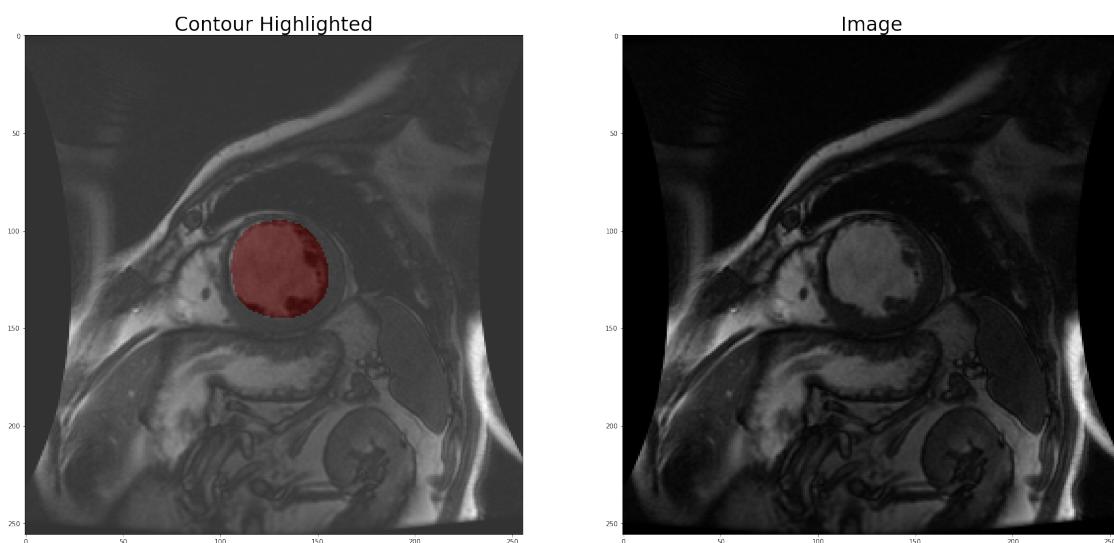
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 80



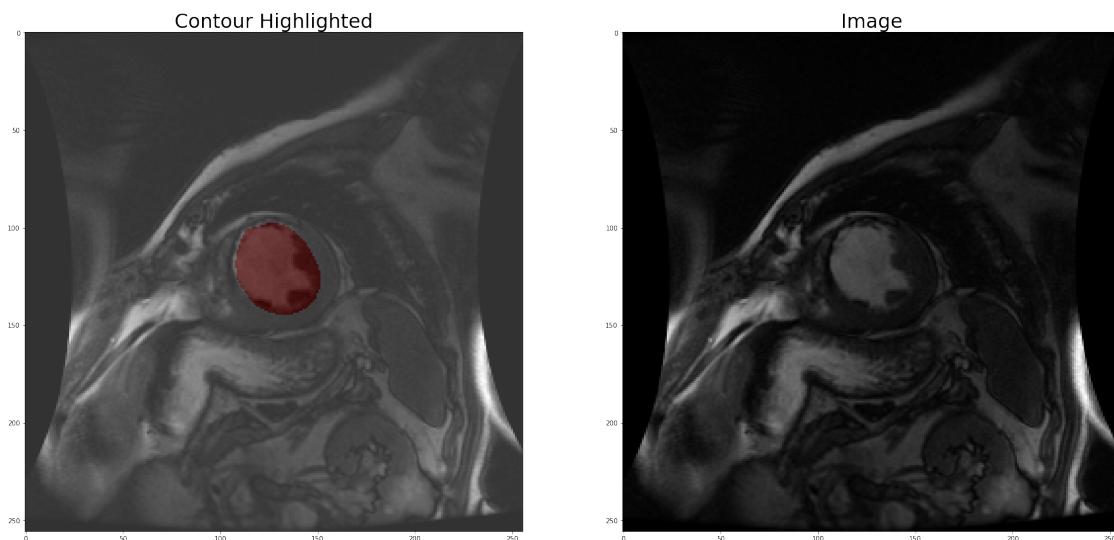
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 87



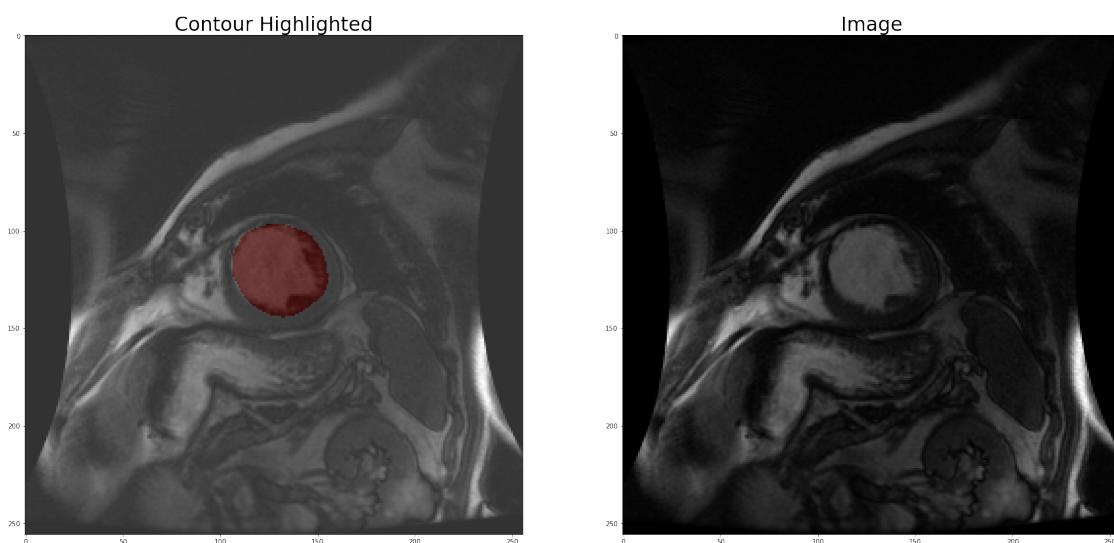
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 100



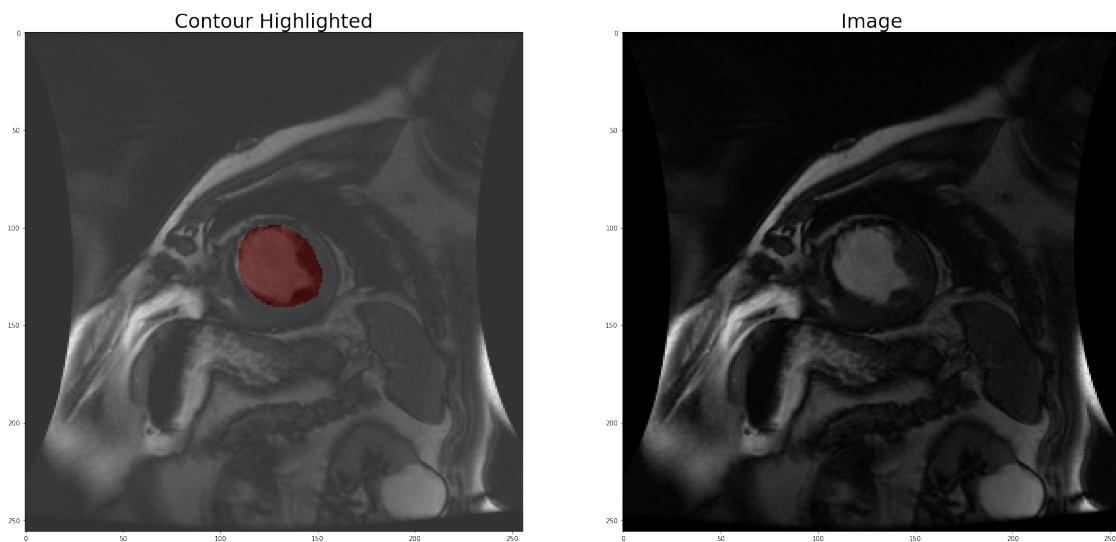
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 107



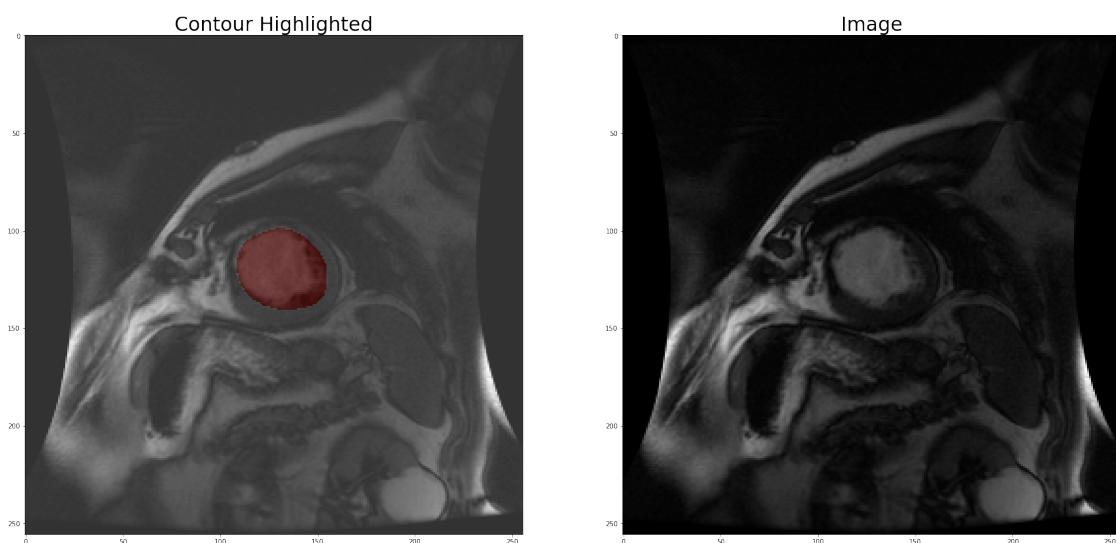
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 120



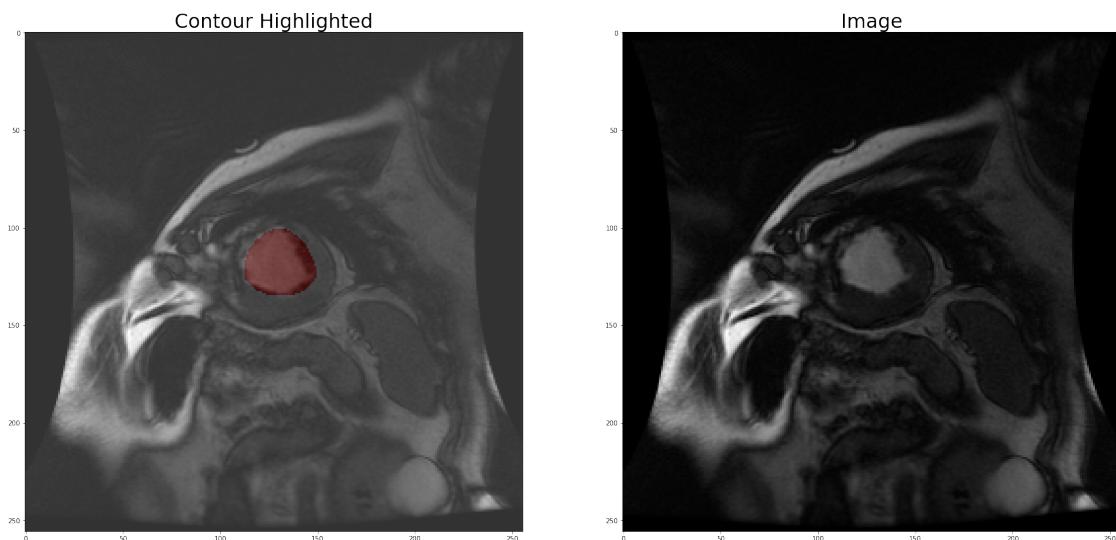
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 127



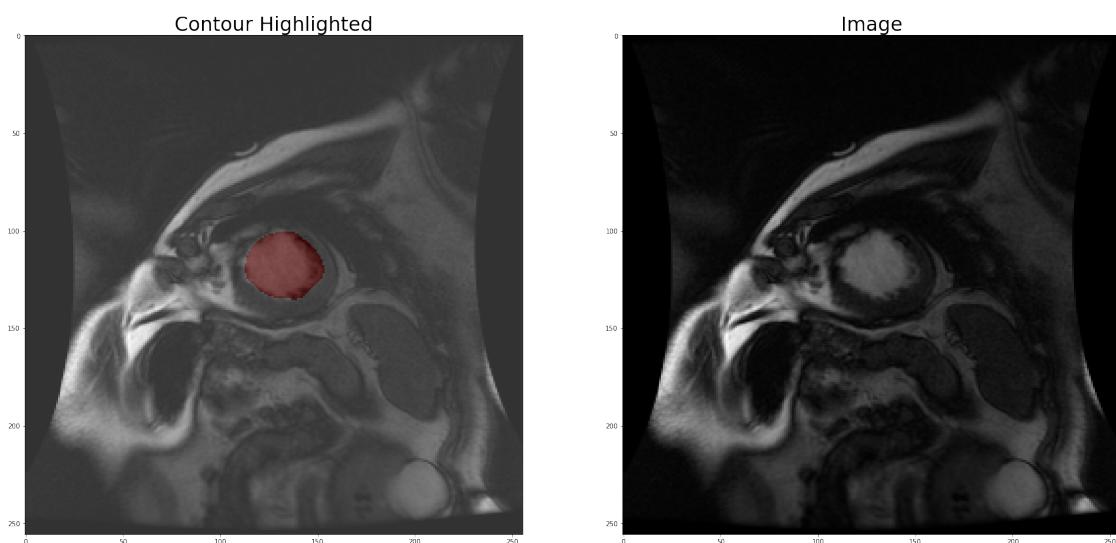
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 140



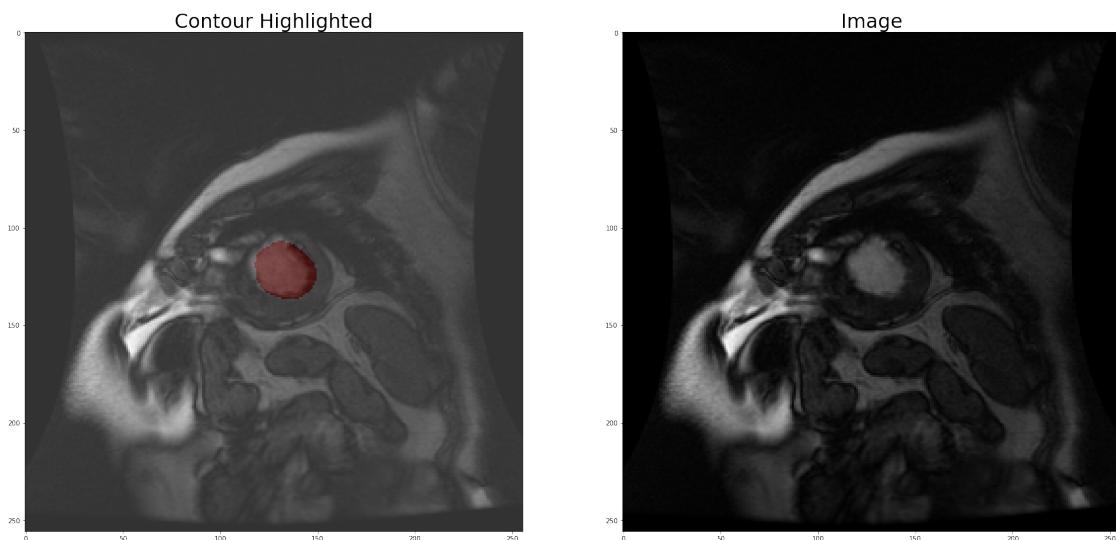
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 147



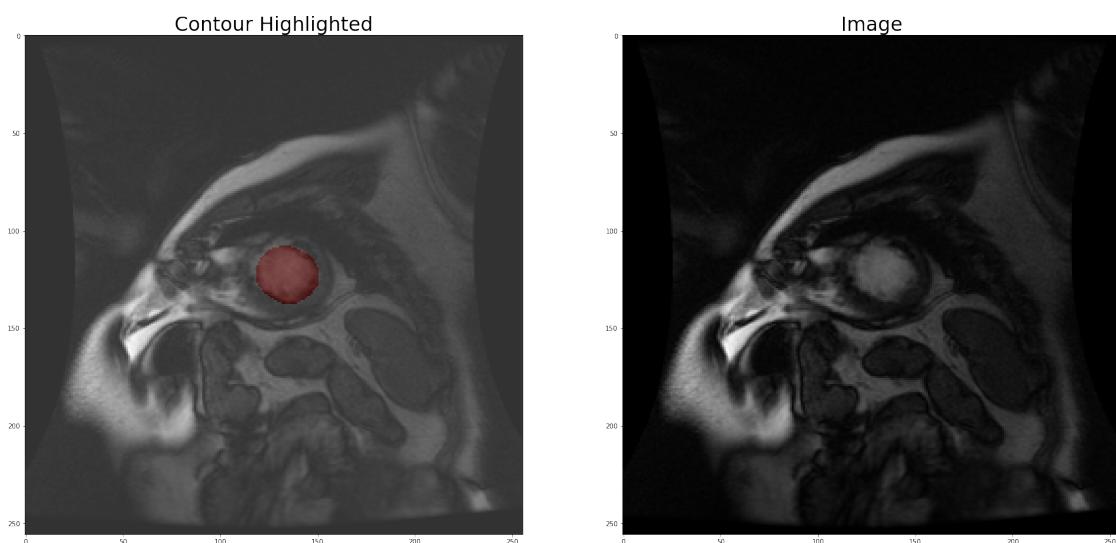
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 160



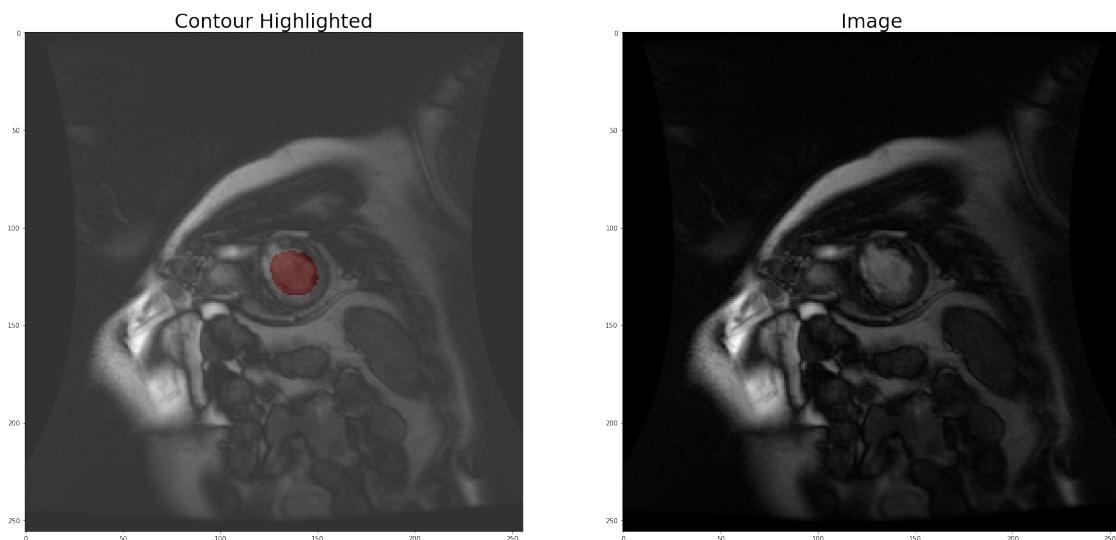
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 167



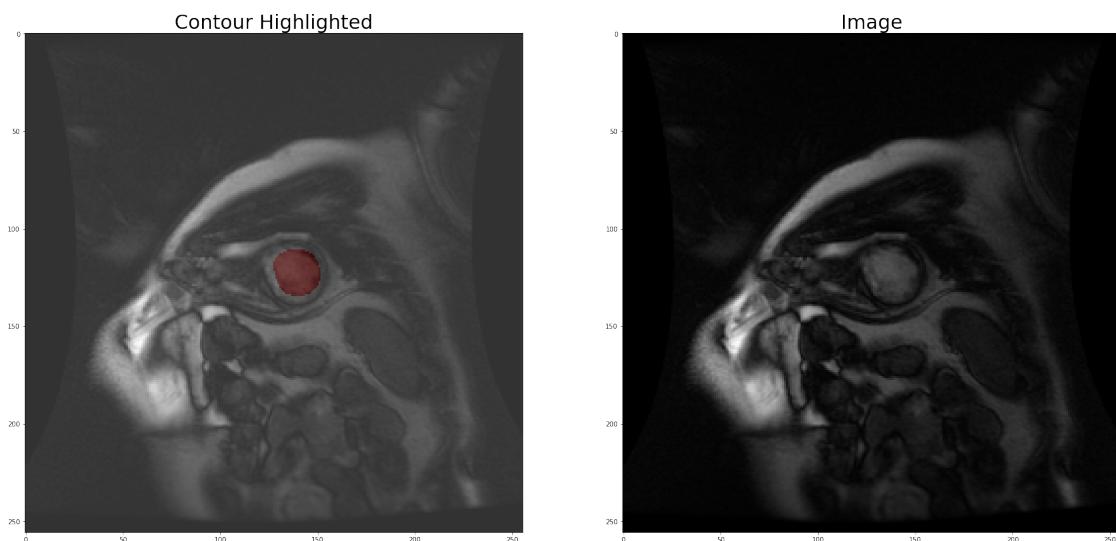
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 180



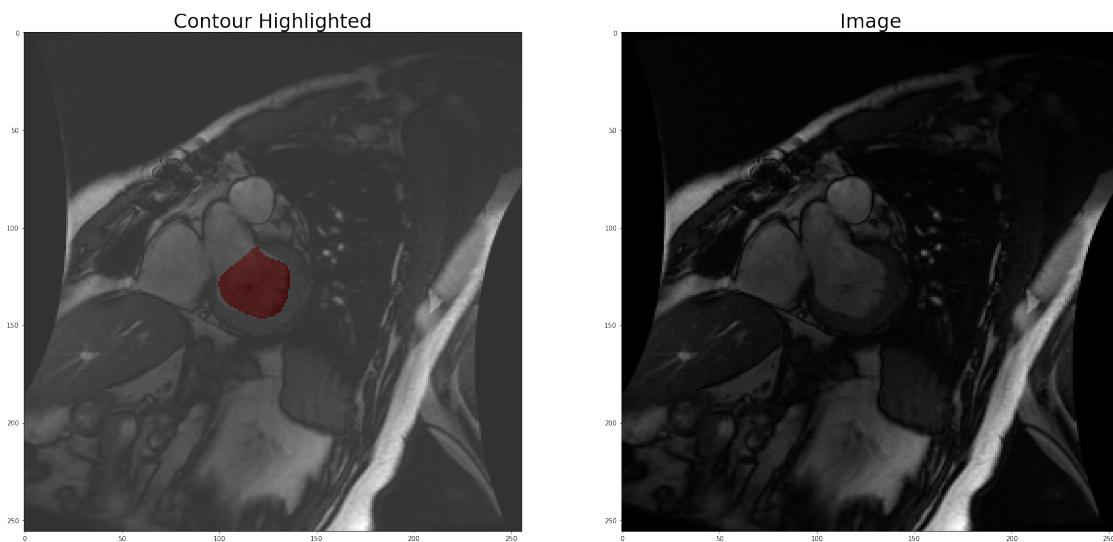
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 187



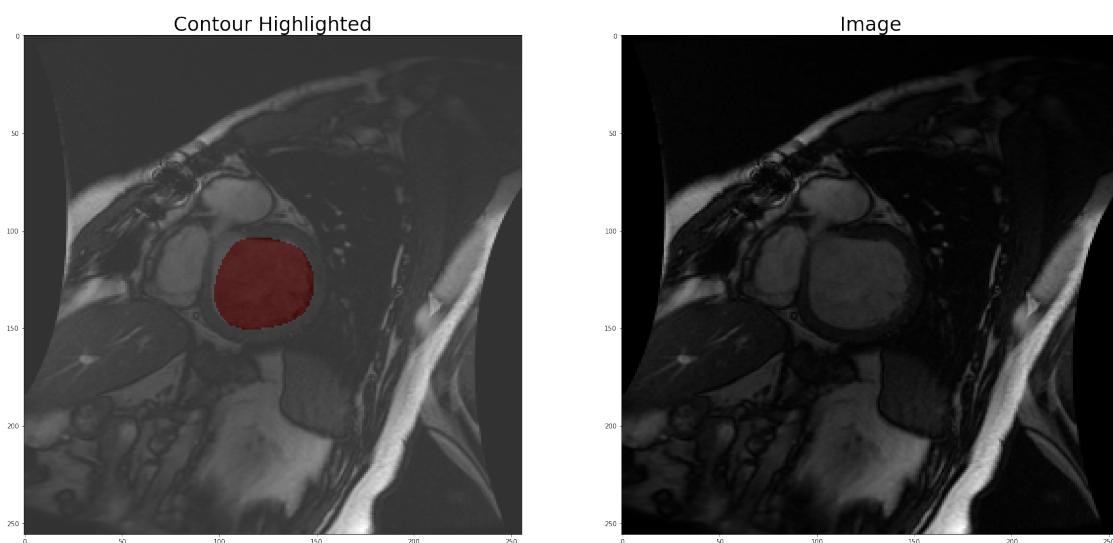
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 200



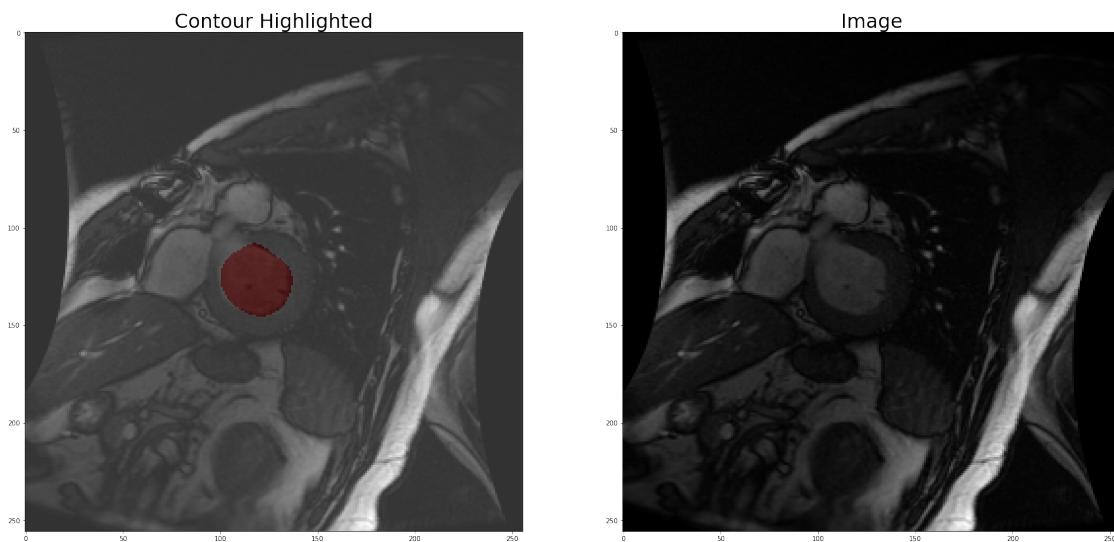
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 27



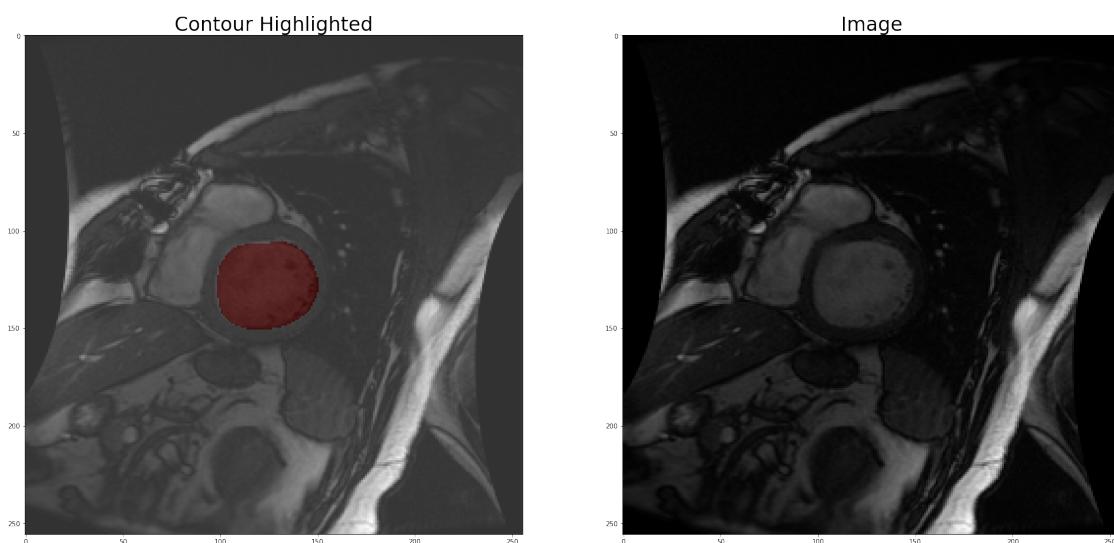
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 40



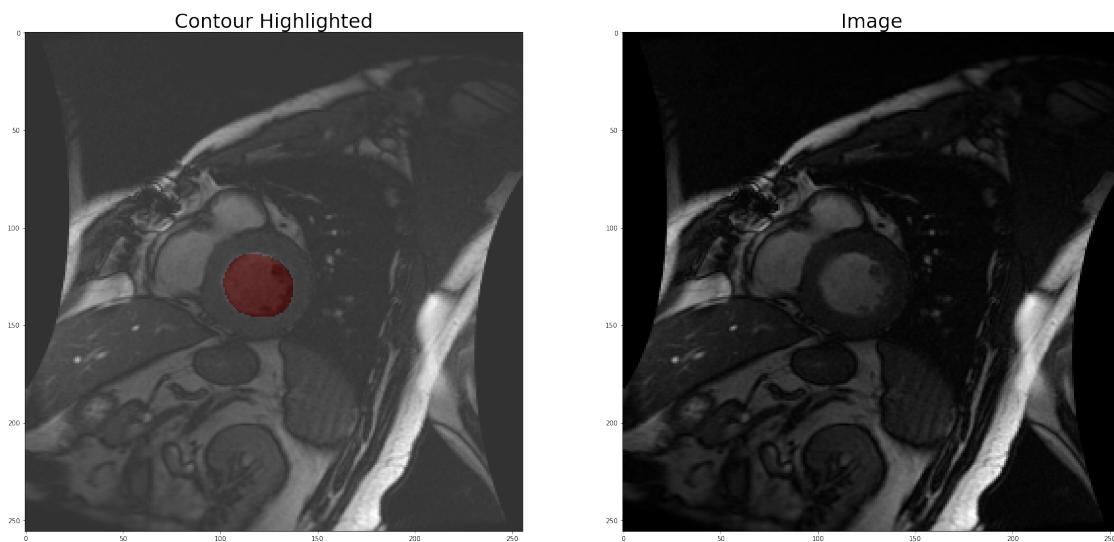
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 47



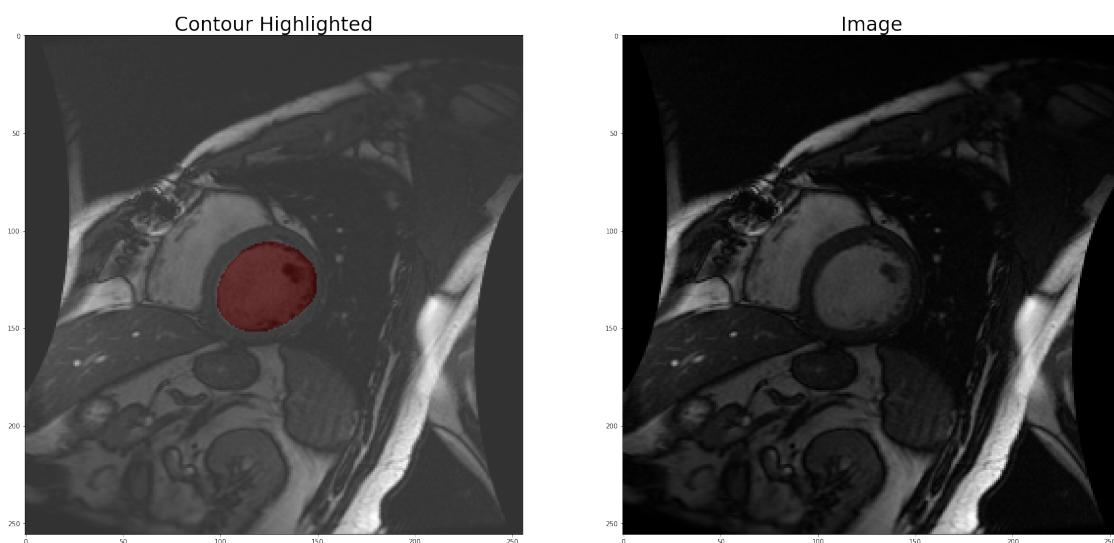
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 60



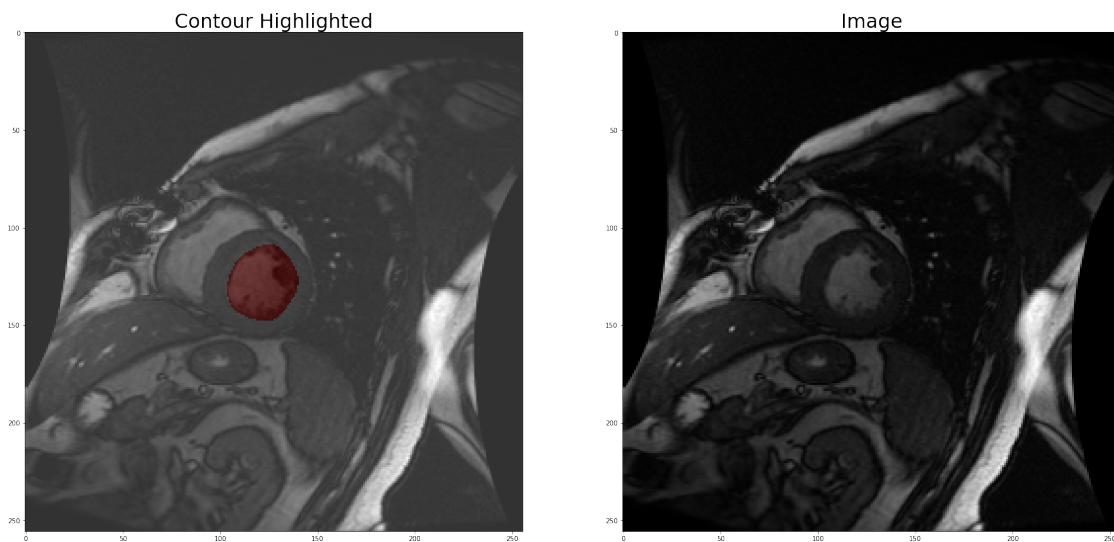
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 67



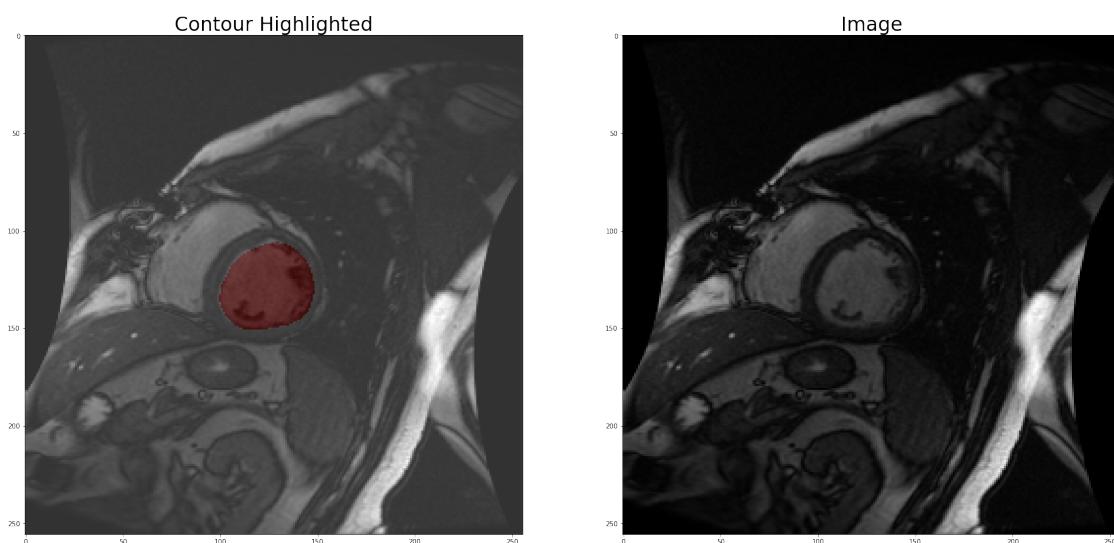
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 80



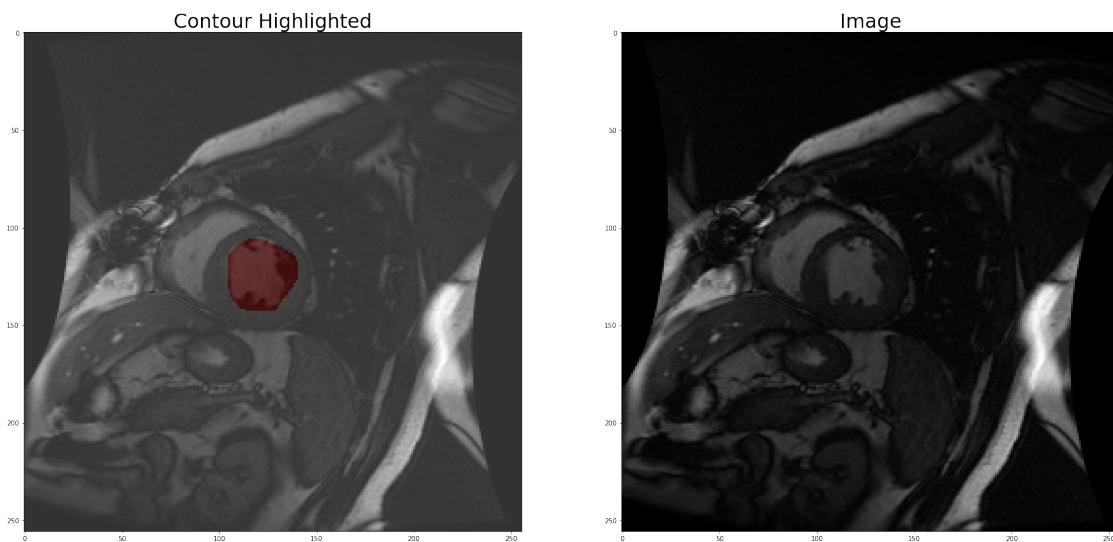
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 87



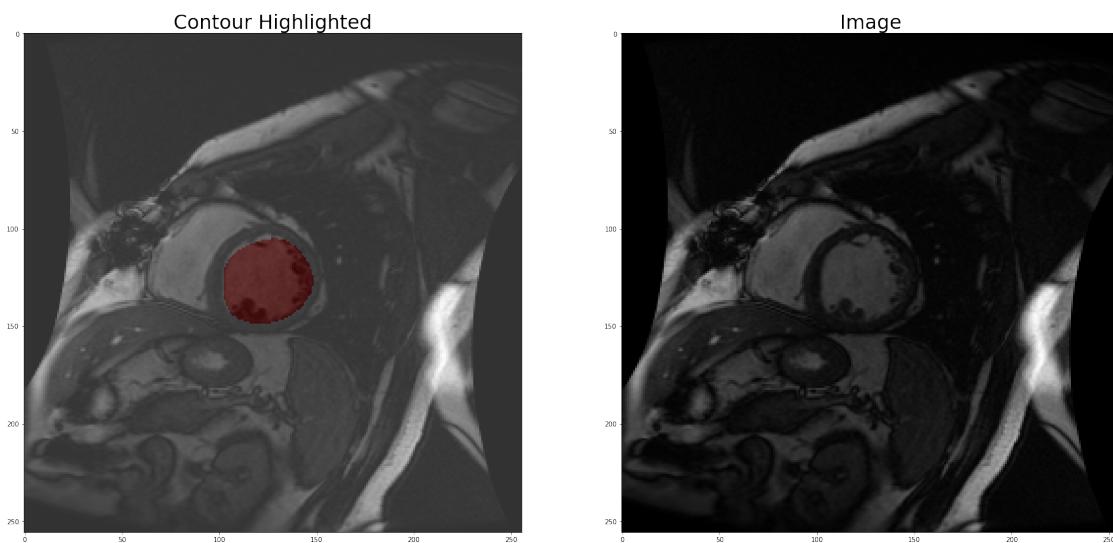
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 100



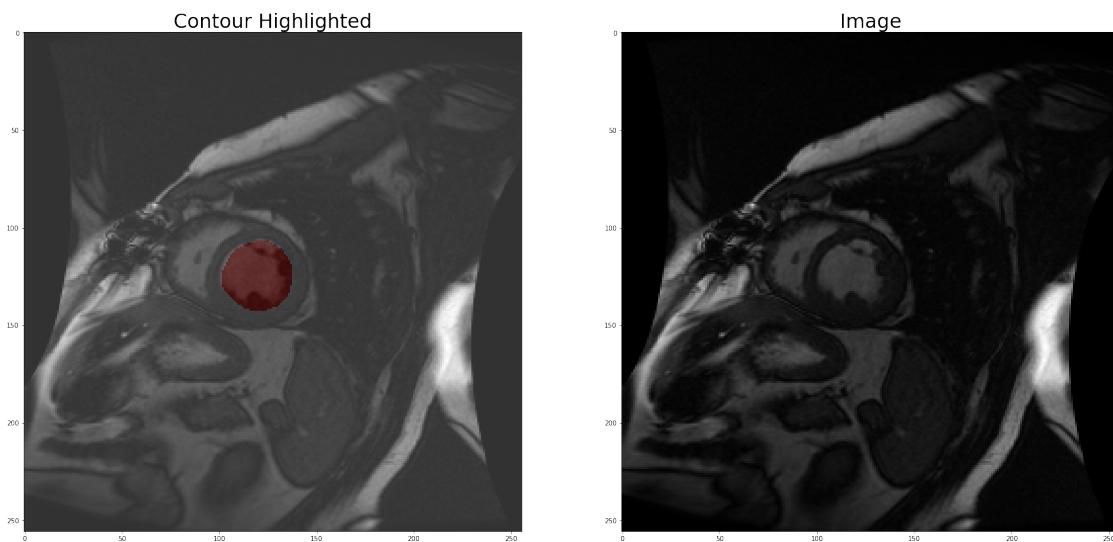
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 107



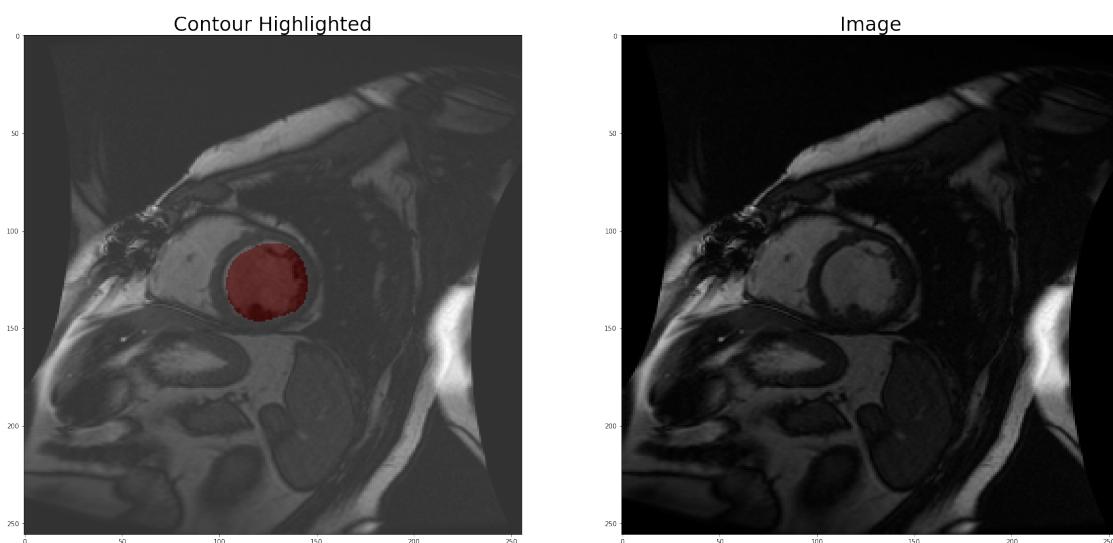
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 120



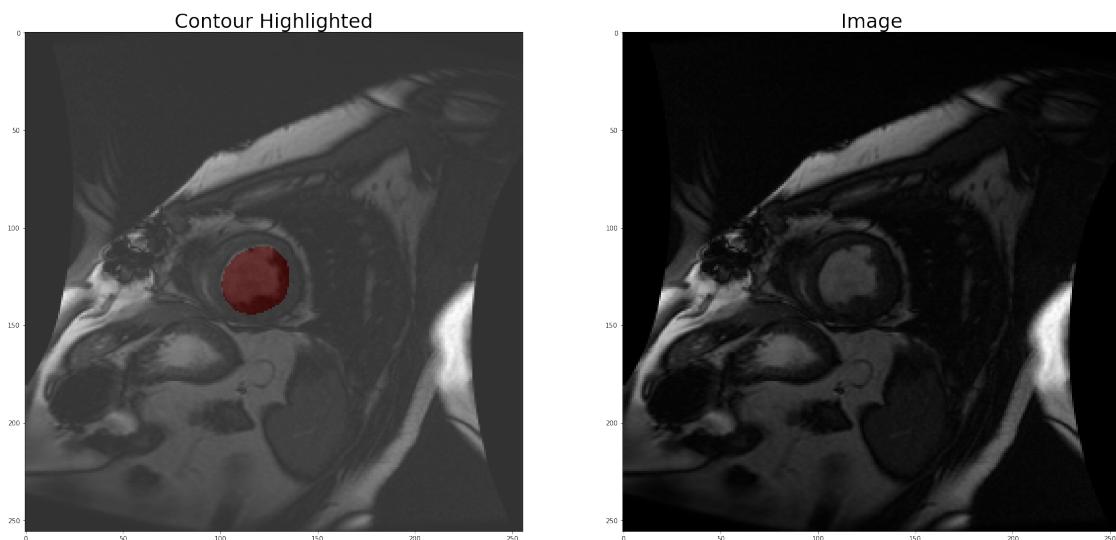
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 127



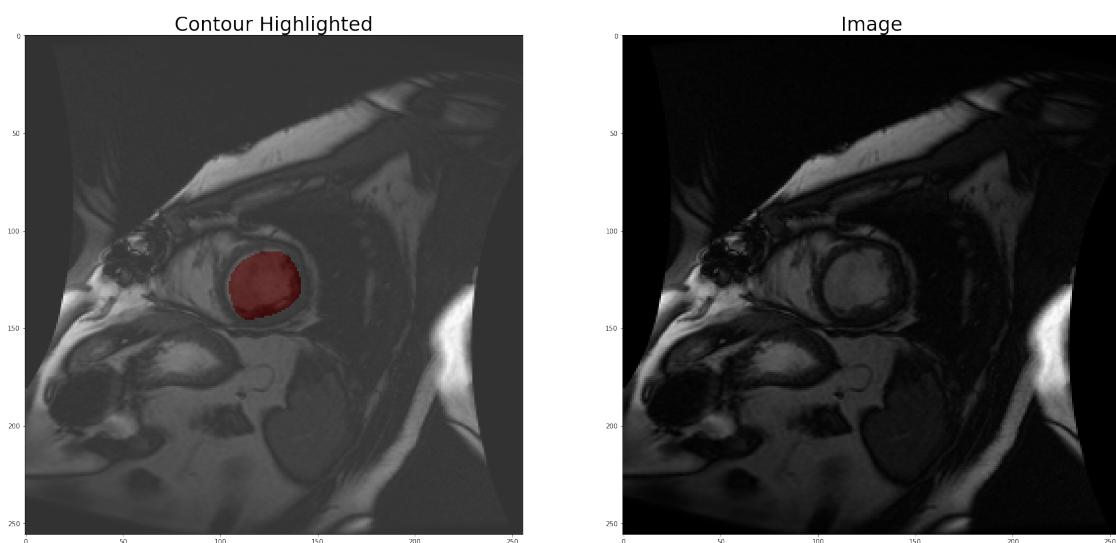
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 140



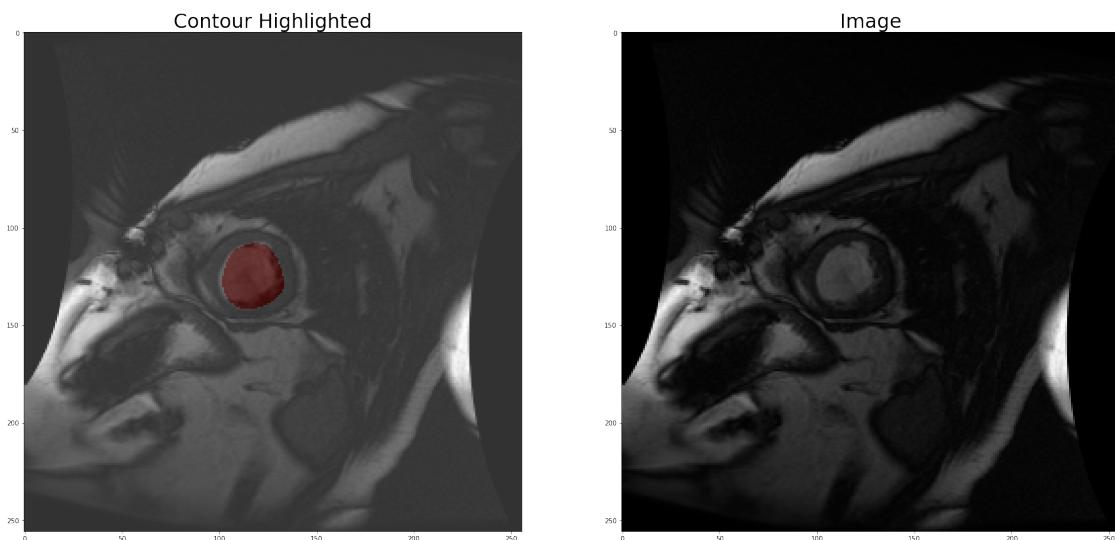
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 147



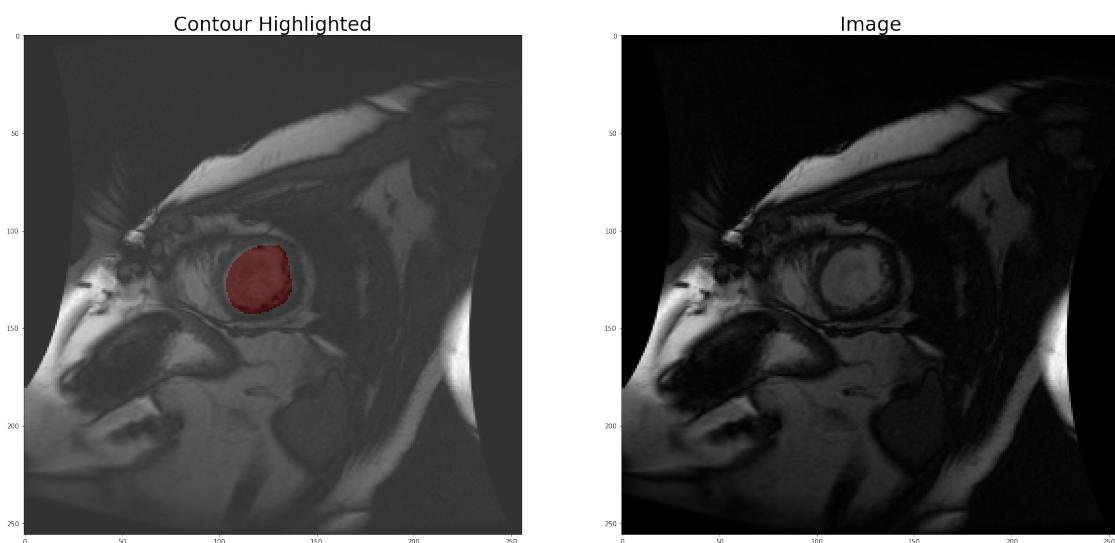
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 160



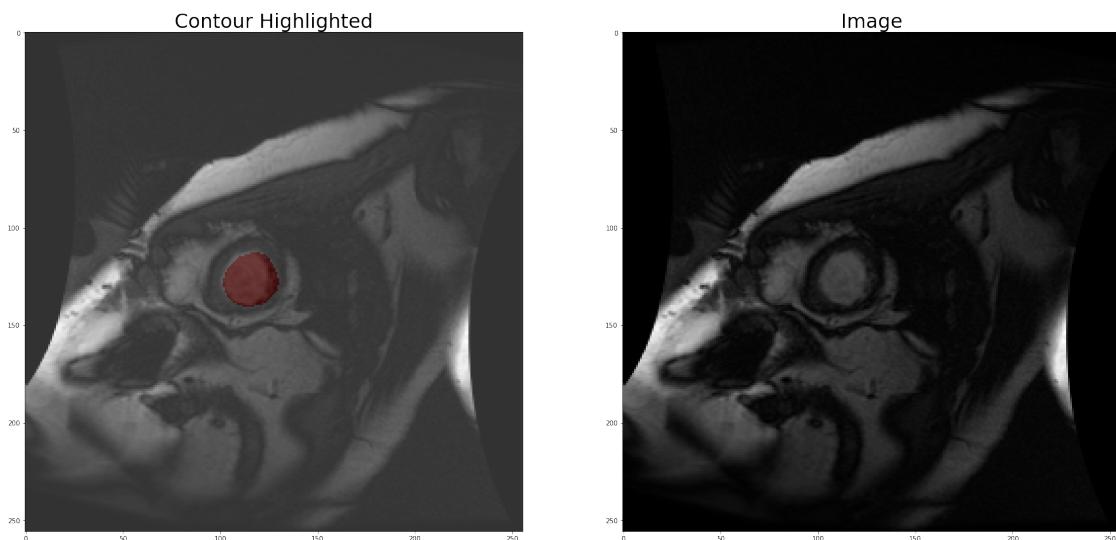
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 167



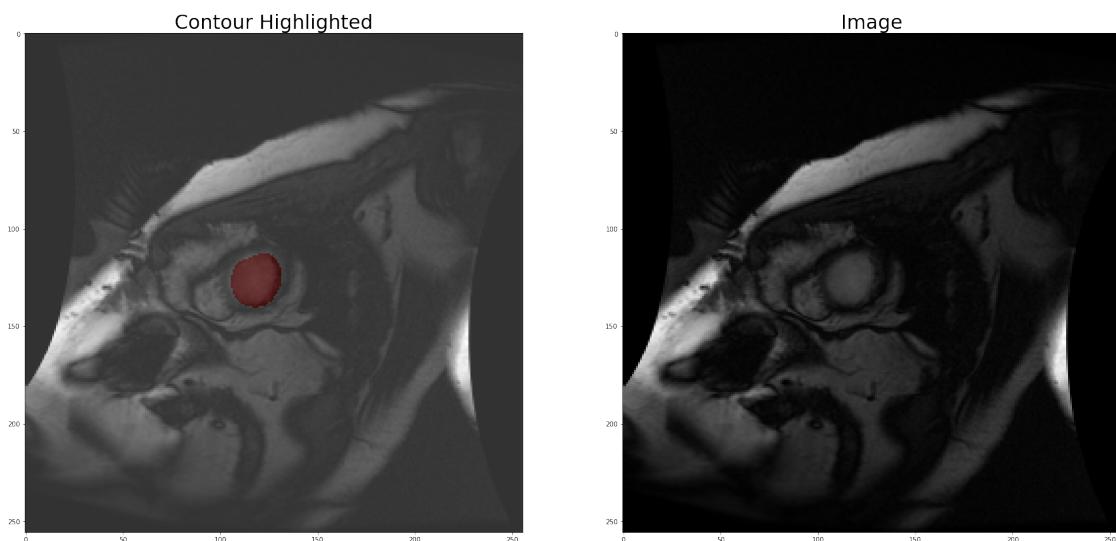
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 180



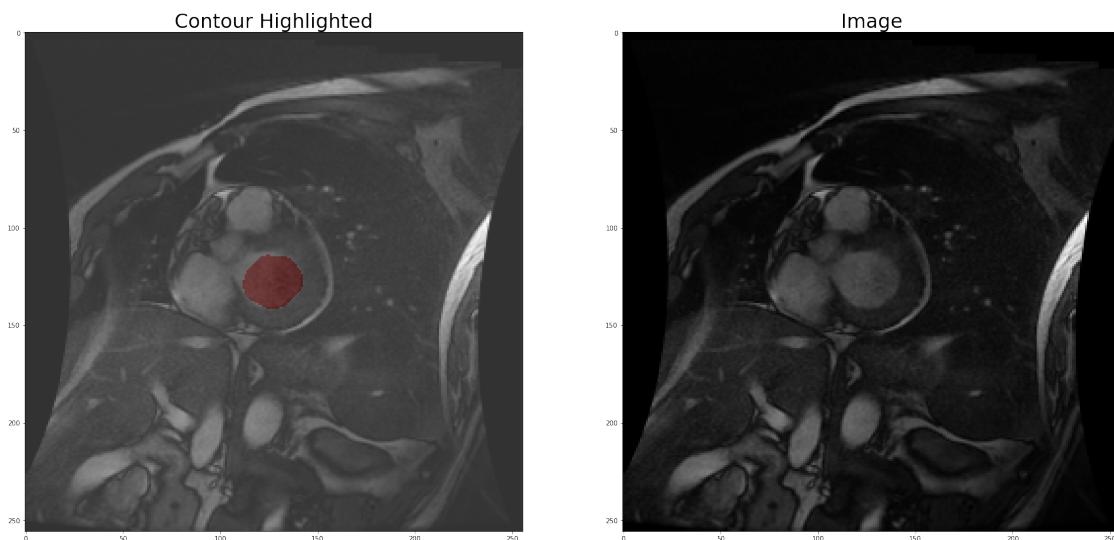
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 187



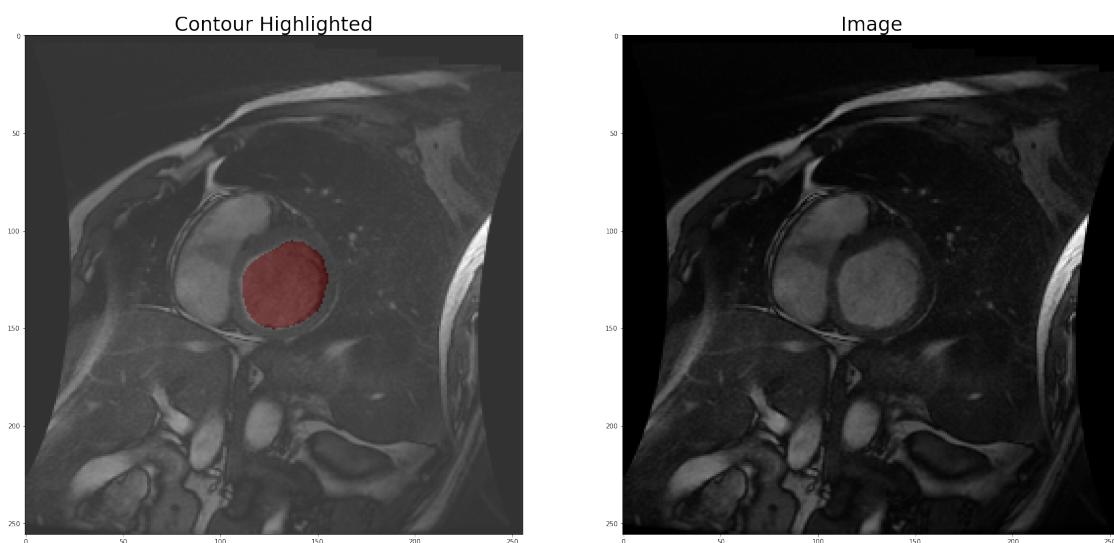
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 200



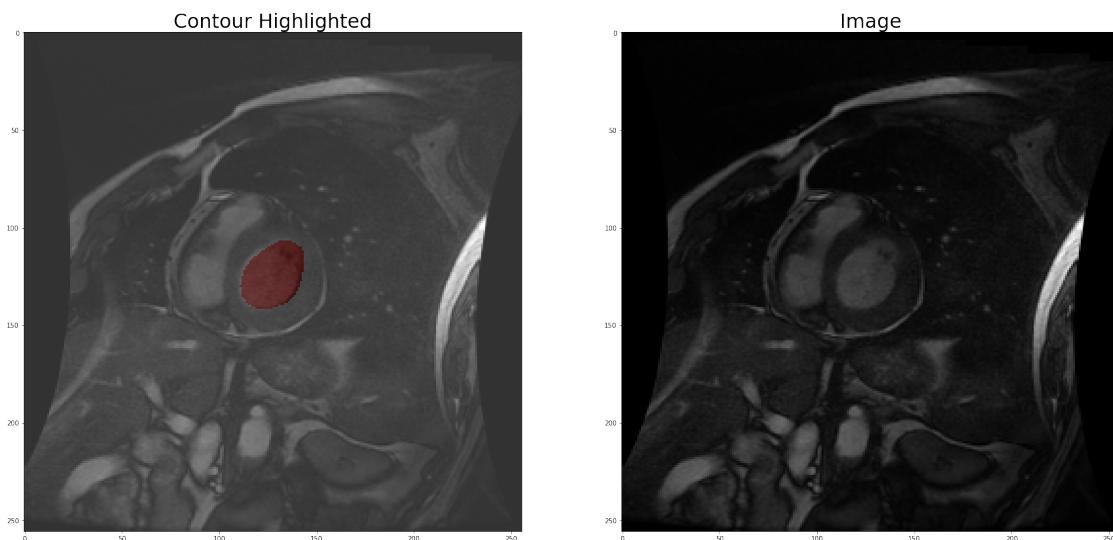
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 9



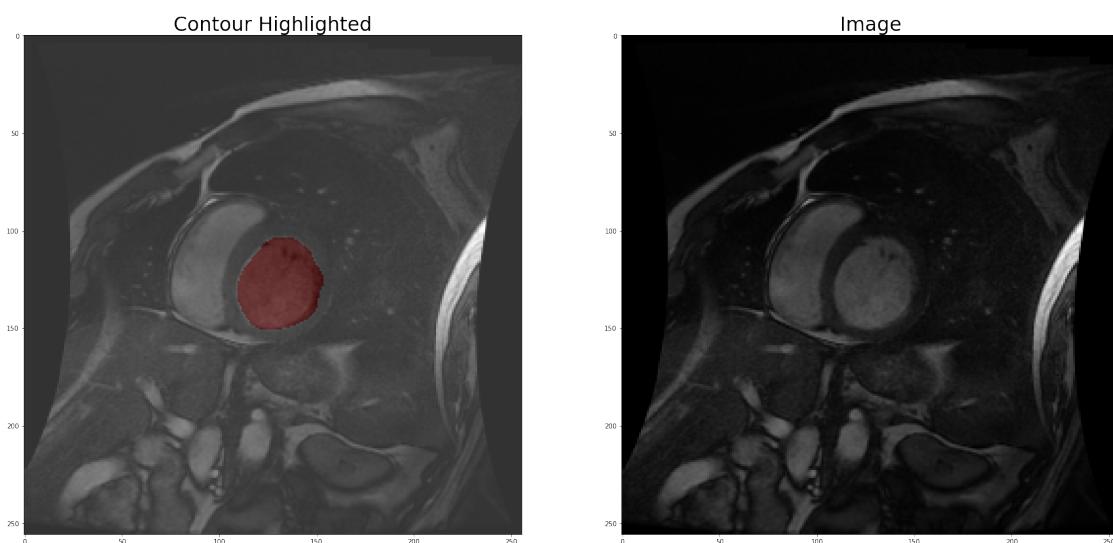
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 19



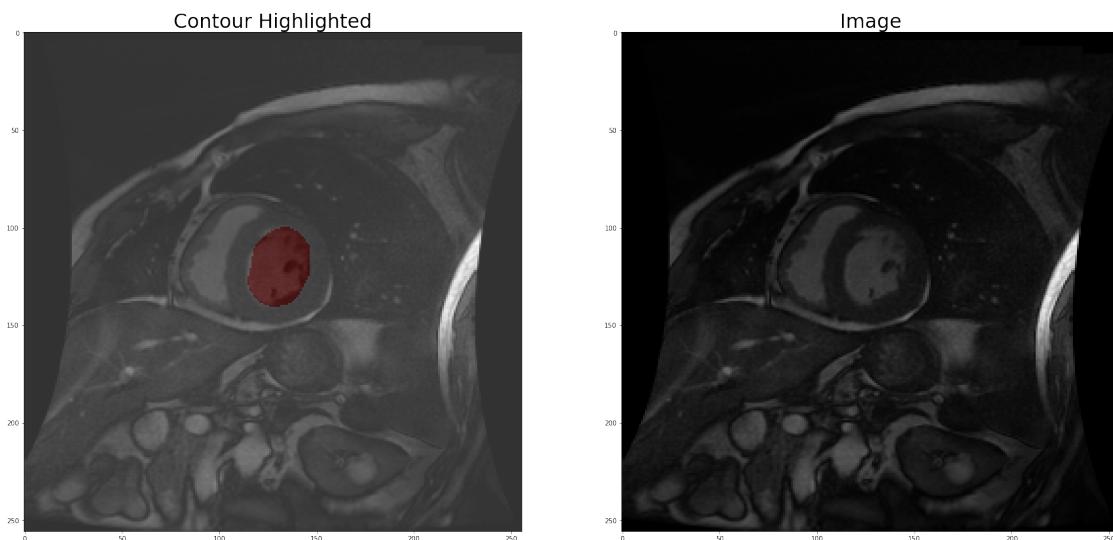
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 29



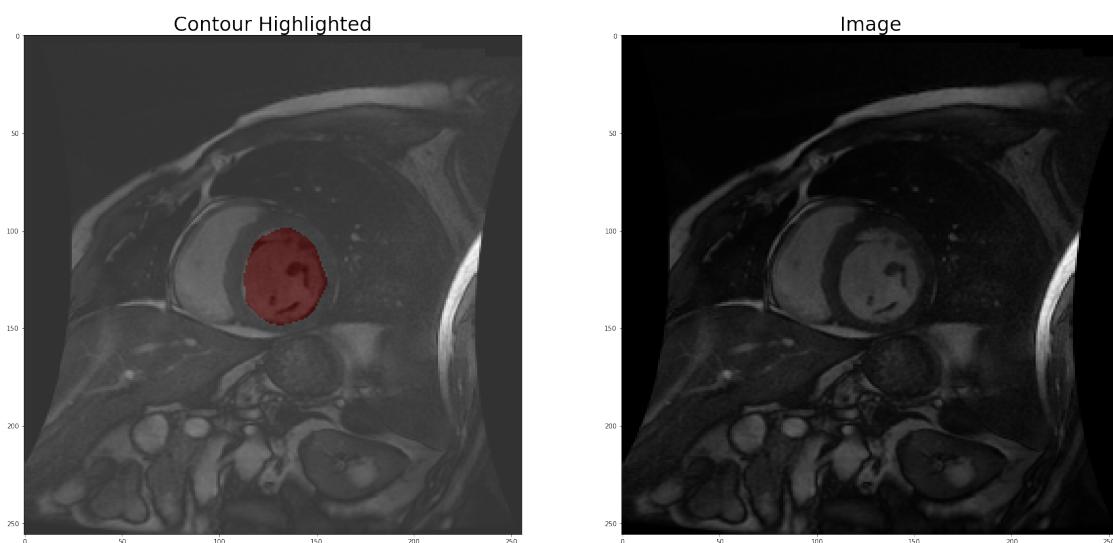
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 39



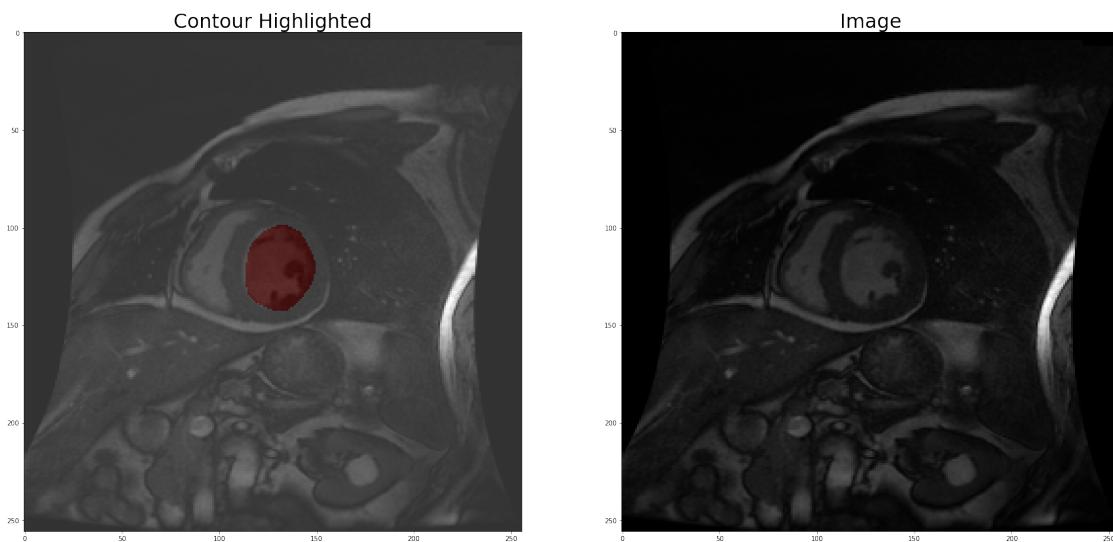
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 49



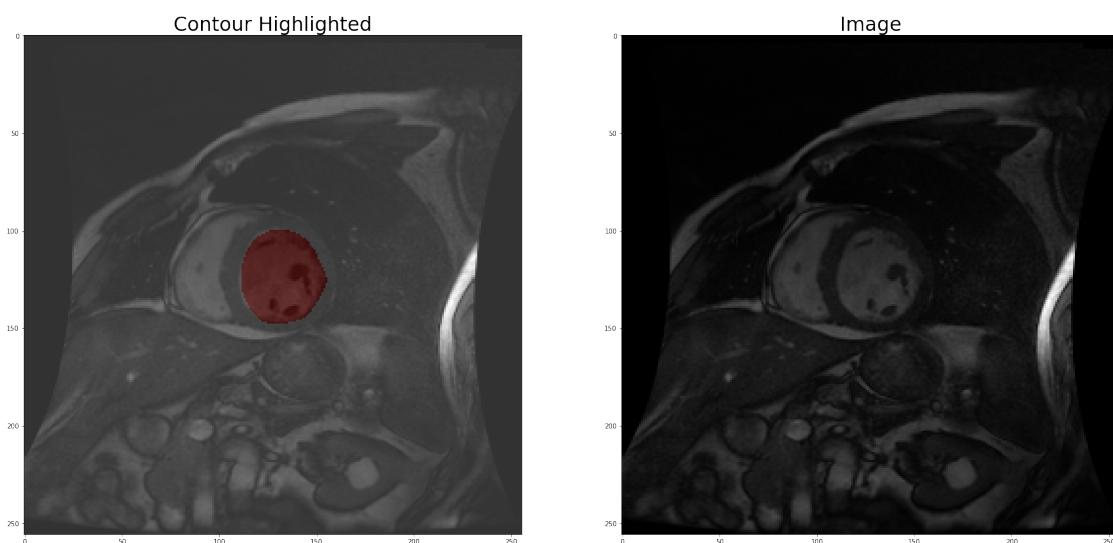
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 59



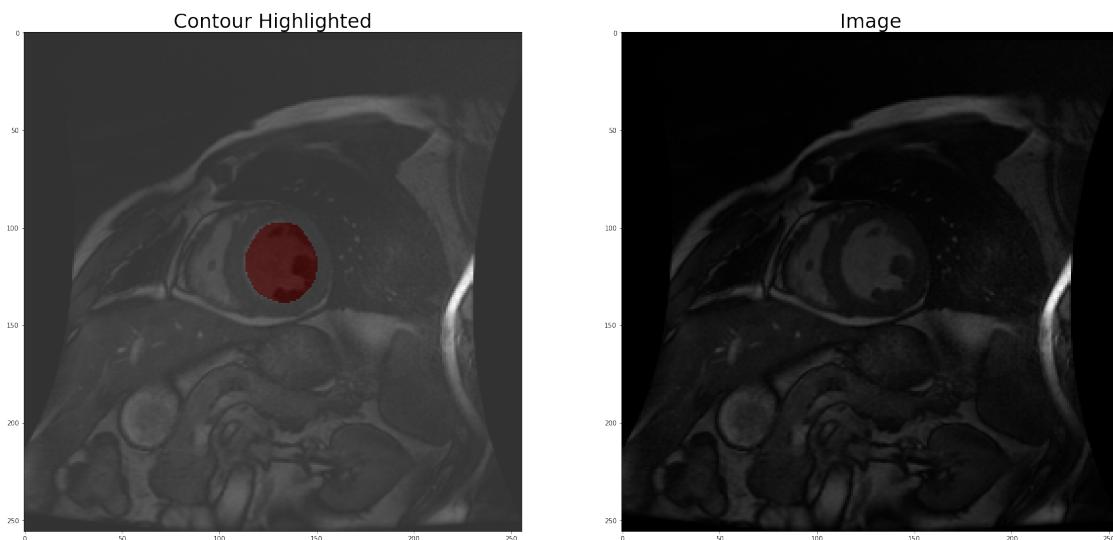
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 69



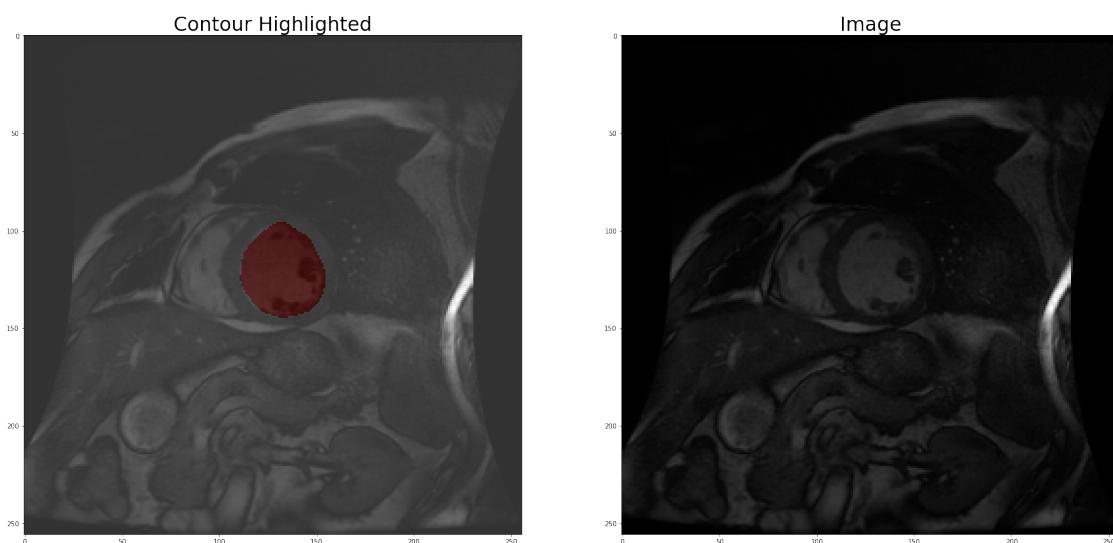
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 79



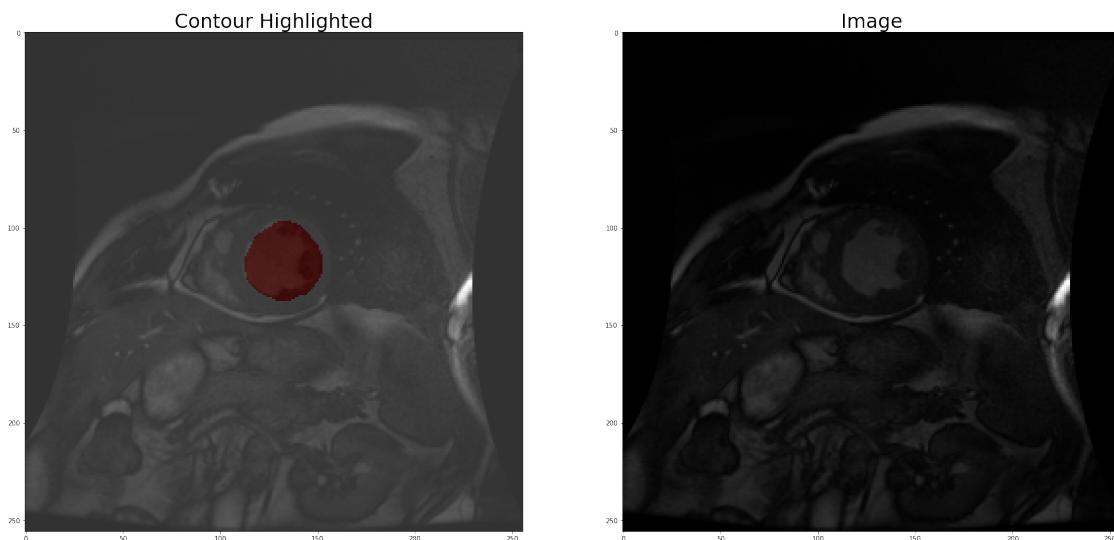
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 89



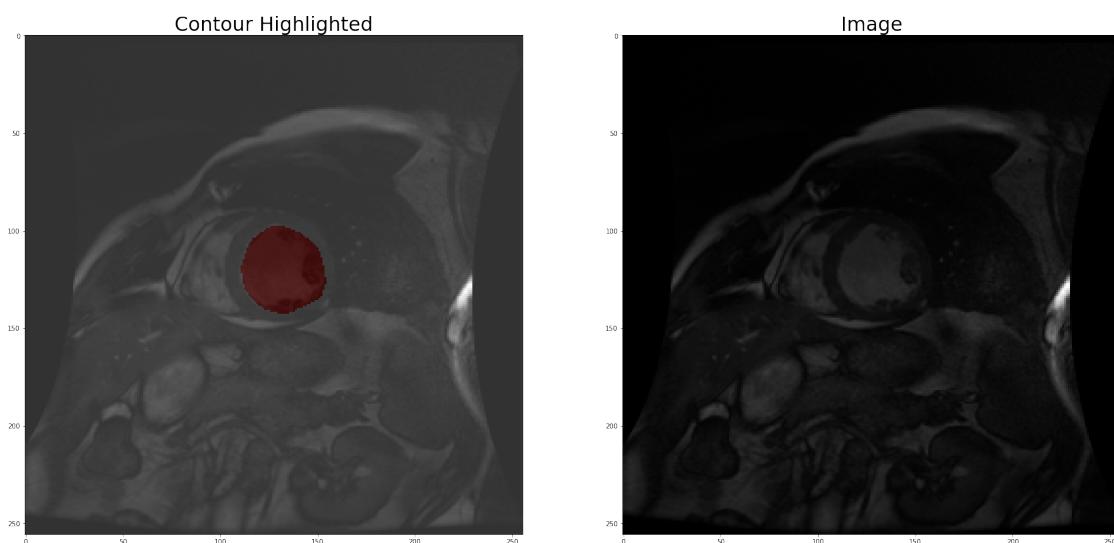
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 99



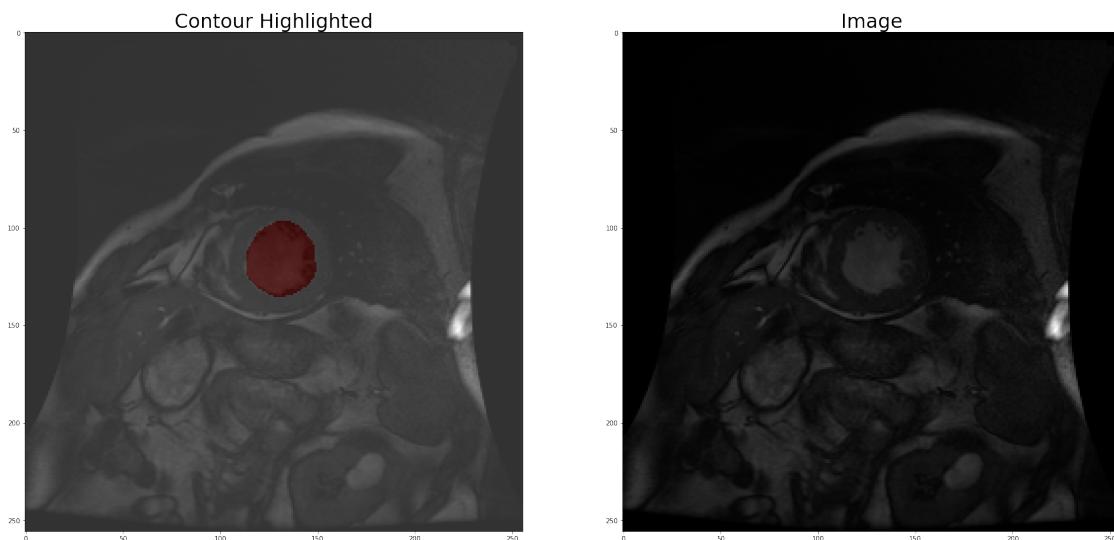
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 109



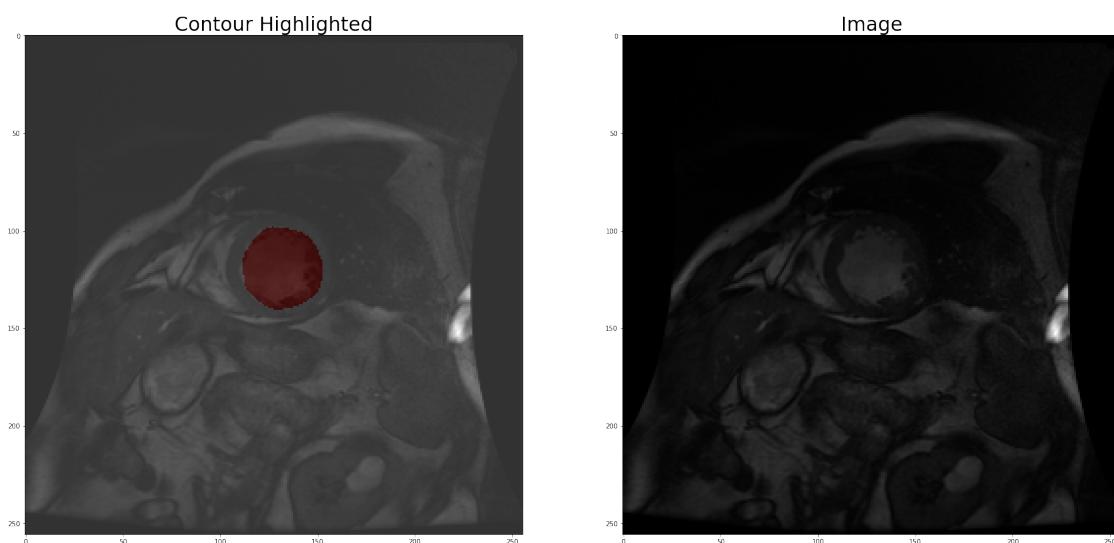
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 119



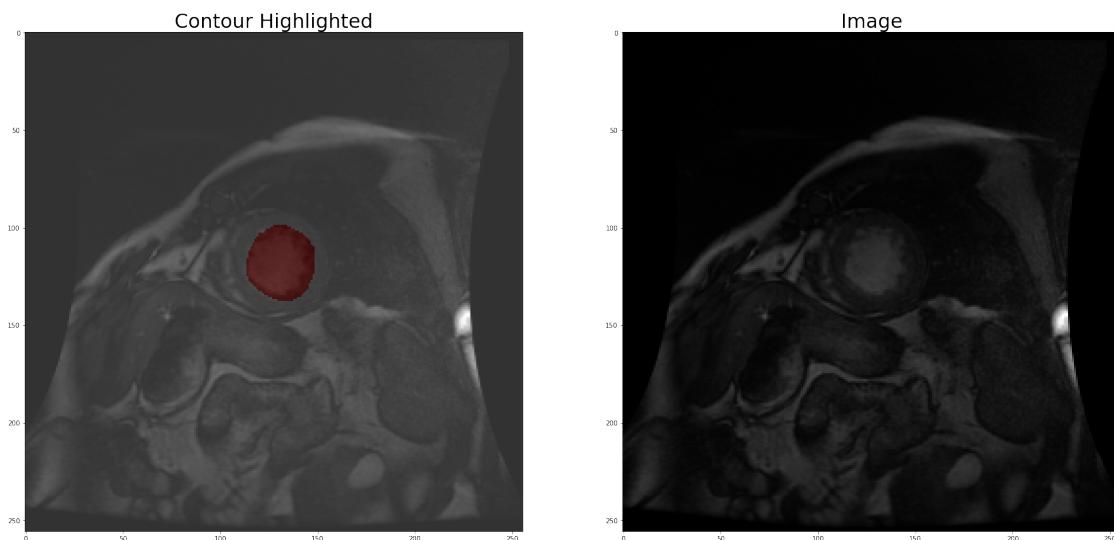
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 129



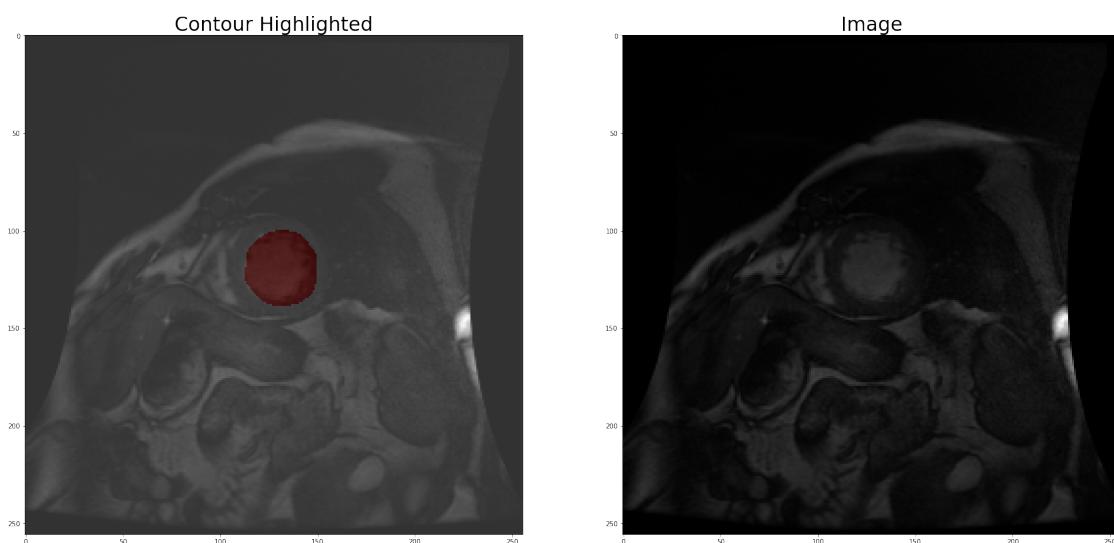
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 139



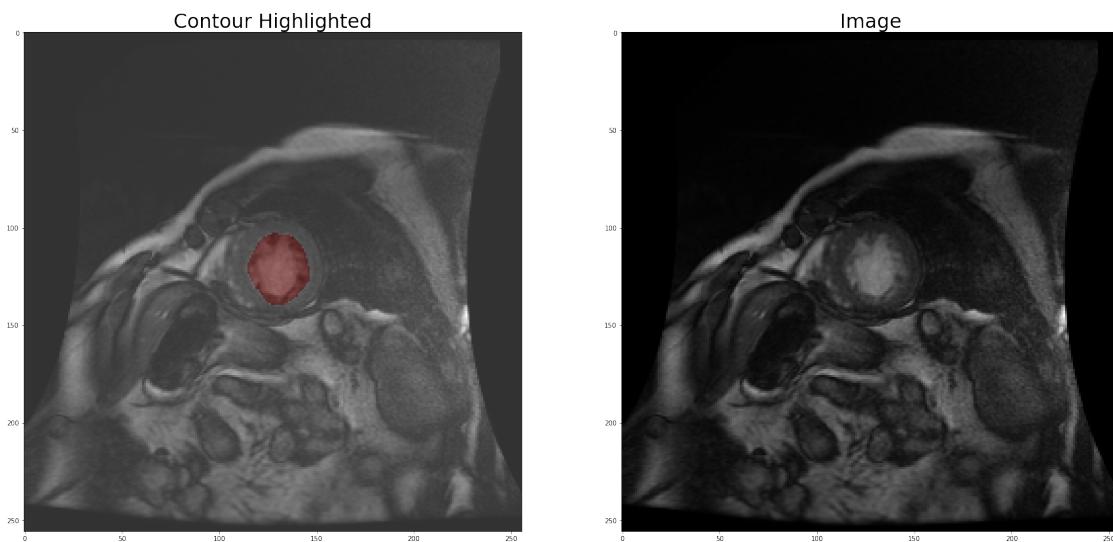
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 149



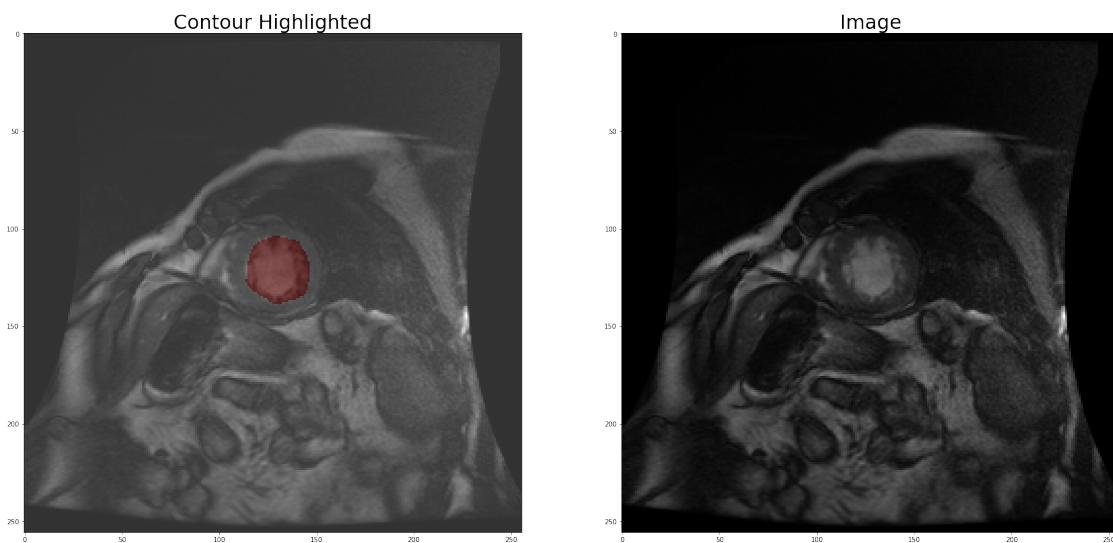
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 159



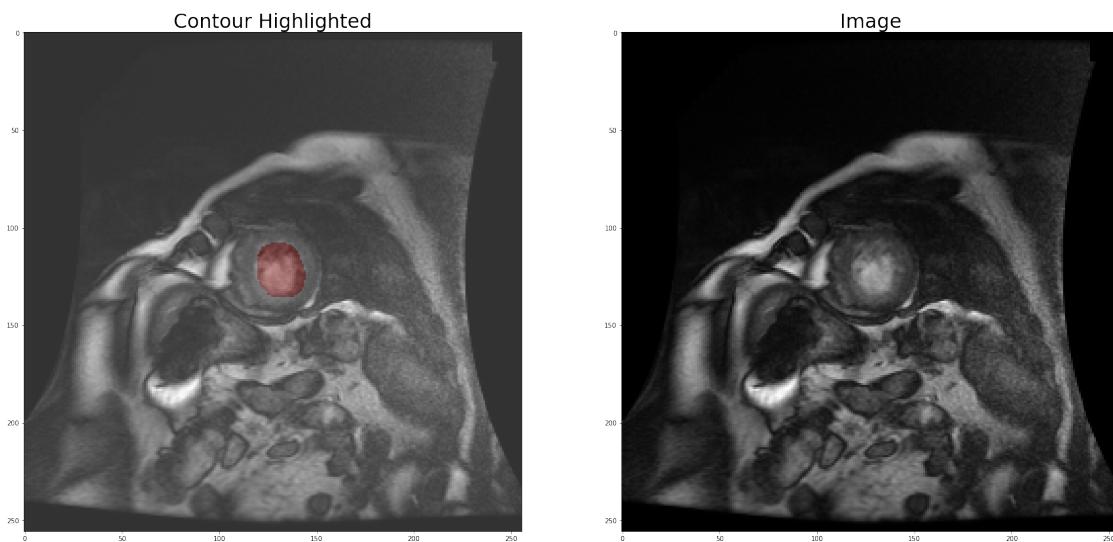
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 169



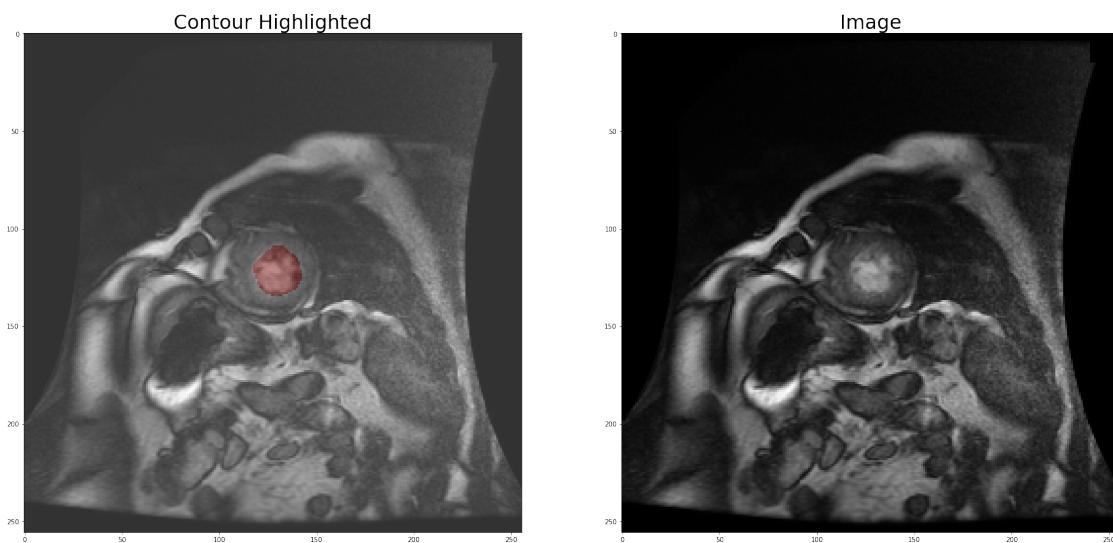
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 179



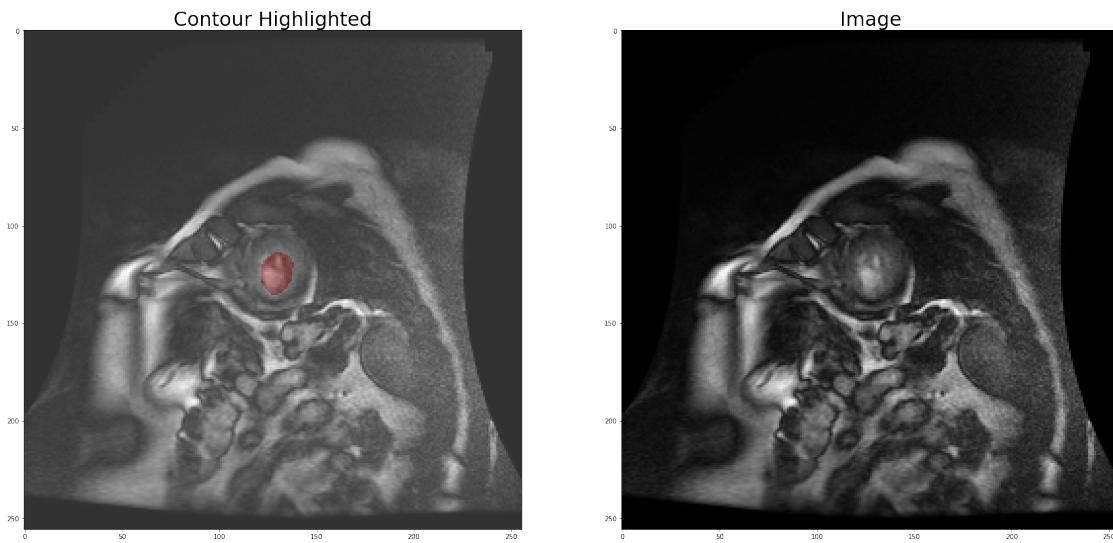
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 189



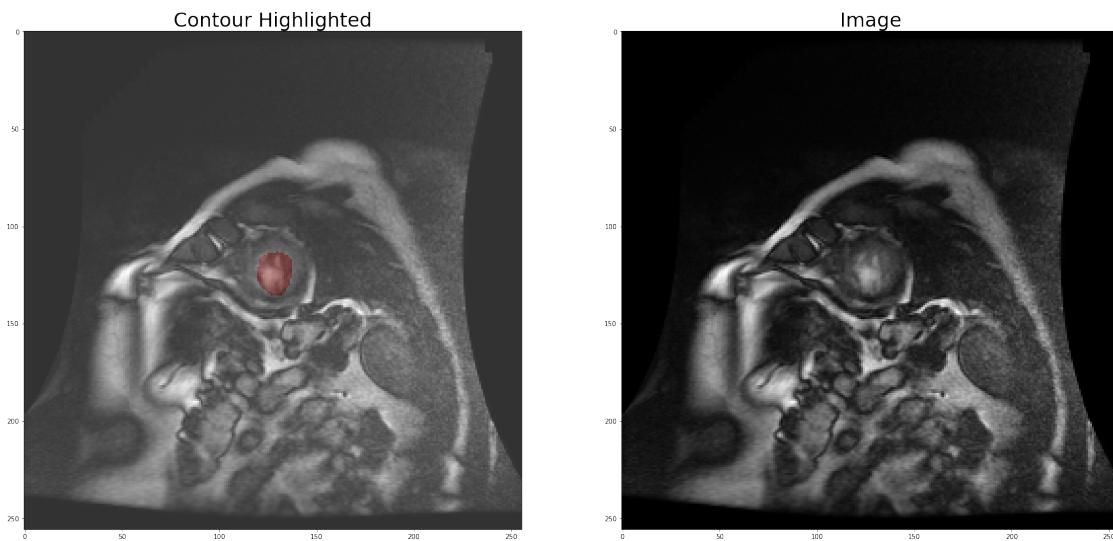
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 199



DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 209



DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 219



1.1.1 Part 1 Questions

How did you verify that you are parsing the contours correctly? I verified that I was parsing the contours correctly by a series of unit tests. We see that the dimensions of the contours match the dimensions of the DICOM images. The total number of DICOM images that have contour labels is 96 which matches the total sample count of the contour array. The maximum of the contour masks is 1 and the minimum of the contour mask is 0. Every Contour was then plotted on top of the DICOM image as a sanity check (another unit test). I am not a domain expert but the results

seam consistent with highlighting the myocardium in MRI images. Certainly the shape of the bulb of contrast seems properly highlighted for each image.

What changes did you make to the code, if any, in order to integrate it into our production code base? I wrapped the starter code in a class named Pipe and added 5 additional methods. The method named processDcmContour() processes the dcm file and contour file, returning a DICOM image and binary mask of the contour. The method named plotDcmContour() plots the DICOM image and its respective contour for visualization. The method named plotPipedData plots all scan image cross-sections with their respective contours (ordered by index id and numerical value of file). The method determineContourNumbers() returns an array of contour file numbers from a folder of contours. This method is necessary in order to know which DCM files we have contours for. Lastly, the method pipeToNumpy() processes all DCMs and contours to numpy arrays (ordered by index id and numerical value of file).

1.2 Part 2: Model training pipeline

```
In [4]: # Permute the data and sample in batches
        #(note batch statistics are output instead of feeding into neural network)
        from parsing import batchify

        batchify(output.dcmArray, output.contourArray, output.labels,
                 batch_size = 8, epochs = 2)

epoch: 1
batch: 1
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 219
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 167
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 139
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 120
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 180
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 67
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 67
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 48

epoch: 1
batch: 2
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 80
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 120
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 207
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 200
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 200
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 160
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 88
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 169
```

```

epoch: 1
batch: 3
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 27
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 29
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 140
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 59
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 180
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 188
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 159
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 127

epoch: 1
batch: 4
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 147
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 89
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 127
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 160
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 147
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 87
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 220
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 108

epoch: 1
batch: 5
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 60
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 107
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 107
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 200
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 87
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 47
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 119
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 209

epoch: 1
batch: 6
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 187
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 167
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 179
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 9

```

```
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 109
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 79
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 40
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 100

epoch: 1
batch: 7
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 80
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 147
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 47
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 87
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 149
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 27
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 120
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 208

epoch: 1
batch: 8
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 59
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 119
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 167
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 99
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 140
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 7
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 49
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 67

epoch: 1
batch: 9
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 39
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 199
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 148
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 60
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 19
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 100
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 68
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 20

epoch: 1
batch: 10
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
```

```

DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 107
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 180
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 128
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 47
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 140
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 40
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 139
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 187

epoch: 1
batch: 11
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 80
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 199
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 179
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 168
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 129
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 60
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 219
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 159

epoch: 1
batch: 12
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 100
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 160
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 99
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 69
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 187
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 127
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 79
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 189

epoch: 2
batch: 1
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 219
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 167
DCM: SCD0000501,     Contours: SC-HF-I-6,      FileNumber: 139
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 120
DCM: SCD0000201,     Contours: SC-HF-I-2,      FileNumber: 180
DCM: SCD0000401,     Contours: SC-HF-I-5,      FileNumber: 67
DCM: SCD0000301,     Contours: SC-HF-I-4,      FileNumber: 67
DCM: SCD0000101,     Contours: SC-HF-I-1,      FileNumber: 48

```

```

epoch: 2
batch: 2
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 80
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 120
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 207
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 200
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 200
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 160
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 88
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 169

epoch: 2
batch: 3
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 27
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 29
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 140
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 59
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 180
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 188
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 159
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 127

epoch: 2
batch: 4
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 147
DCM: SCD0000501, Contours: SC-HF-I-6, FileNumber: 89
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 127
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 160
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 147
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 87
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 220
DCM: SCD0000101, Contours: SC-HF-I-1, FileNumber: 108

epoch: 2
batch: 5
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 60
DCM: SCD0000201, Contours: SC-HF-I-2, FileNumber: 107
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 107
DCM: SCD0000401, Contours: SC-HF-I-5, FileNumber: 200
DCM: SCD0000301, Contours: SC-HF-I-4, FileNumber: 87

```

```
DCM: SCD0000201,      Contours: SC-HF-I-2,      FileNumber: 47
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 119
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 209

epoch: 2
batch: 6
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000301,      Contours: SC-HF-I-4,      FileNumber: 187
DCM: SCD0000401,      Contours: SC-HF-I-5,      FileNumber: 167
DCM: SCD0000101,      Contours: SC-HF-I-1,      FileNumber: 179
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 9
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 109
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 79
DCM: SCD0000301,      Contours: SC-HF-I-4,      FileNumber: 40
DCM: SCD0000401,      Contours: SC-HF-I-5,      FileNumber: 100

epoch: 2
batch: 7
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000401,      Contours: SC-HF-I-5,      FileNumber: 80
DCM: SCD0000401,      Contours: SC-HF-I-5,      FileNumber: 147
DCM: SCD0000301,      Contours: SC-HF-I-4,      FileNumber: 47
DCM: SCD0000201,      Contours: SC-HF-I-2,      FileNumber: 87
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 149
DCM: SCD0000401,      Contours: SC-HF-I-5,      FileNumber: 27
DCM: SCD0000201,      Contours: SC-HF-I-2,      FileNumber: 120
DCM: SCD0000101,      Contours: SC-HF-I-1,      FileNumber: 208

epoch: 2
batch: 8
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000101,      Contours: SC-HF-I-1,      FileNumber: 59
DCM: SCD0000101,      Contours: SC-HF-I-1,      FileNumber: 119
DCM: SCD0000201,      Contours: SC-HF-I-2,      FileNumber: 167
DCM: SCD0000101,      Contours: SC-HF-I-1,      FileNumber: 99
DCM: SCD0000301,      Contours: SC-HF-I-4,      FileNumber: 140
DCM: SCD0000301,      Contours: SC-HF-I-4,      FileNumber: 7
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 49
DCM: SCD0000201,      Contours: SC-HF-I-2,      FileNumber: 67

epoch: 2
batch: 9
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000501,      Contours: SC-HF-I-6,      FileNumber: 39
```

```
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 199
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 148
DCM: SCD0000301,     Contours: SC-HF-I-4,     FileNumber: 60
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 19
DCM: SCD0000301,     Contours: SC-HF-I-4,     FileNumber: 100
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 68
DCM: SCD0000301,     Contours: SC-HF-I-4,     FileNumber: 20

epoch: 2
batch: 10
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000301,     Contours: SC-HF-I-4,     FileNumber: 107
DCM: SCD0000401,     Contours: SC-HF-I-5,     FileNumber: 180
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 128
DCM: SCD0000401,     Contours: SC-HF-I-5,     FileNumber: 47
DCM: SCD0000401,     Contours: SC-HF-I-5,     FileNumber: 140
DCM: SCD0000401,     Contours: SC-HF-I-5,     FileNumber: 40
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 139
DCM: SCD0000201,     Contours: SC-HF-I-2,     FileNumber: 187

epoch: 2
batch: 11
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201,     Contours: SC-HF-I-2,     FileNumber: 80
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 199
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 179
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 168
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 129
DCM: SCD0000401,     Contours: SC-HF-I-5,     FileNumber: 60
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 219
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 159

epoch: 2
batch: 12
X_batch shape: (8, 256, 256)
y_batch shape: (8, 256, 256)
DCM: SCD0000201,     Contours: SC-HF-I-2,     FileNumber: 100
DCM: SCD0000201,     Contours: SC-HF-I-2,     FileNumber: 160
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 99
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 69
DCM: SCD0000401,     Contours: SC-HF-I-5,     FileNumber: 187
DCM: SCD0000301,     Contours: SC-HF-I-4,     FileNumber: 127
DCM: SCD0000101,     Contours: SC-HF-I-1,     FileNumber: 79
DCM: SCD0000501,     Contours: SC-HF-I-6,     FileNumber: 189
```

1.2.1 Part 2 Questions

Did you change anything from the pipelines built in parts 1 to better streamline the pipeline for part 2? If so what? If not, is there anything that you can imagine changing in the future? I did not change anything significantly from part 1 to better streamline part 2. In fact one of the reasons my output of part 1 is a combined numpy array is because I was looking ahead to part 2. In the future it would likely be a good idea to read the dataset in batches instead of reading the entire thing into RAM. For a very large dataset the pipeline would likely break since the entire dataset is being read into memory. Another thing that would be good to change in the future is to parallelize the pipeline from part 1. The same sort of calculation is being done many times which lends itself to parallelization.

How do you/did you verify that the pipeline was working correctly? I verified that I was parsing the contours correctly by a series of unit tests. We see that the dimensions of the contours match the dimensions of the DICOM images. The total number of DICOM images that have contour labels is 96 which matches the total sample count of the contour array. The maximum of the contour masks is 1 and the minimum of the contour mask is 0.

I verified that the parsing of the dicom files were correct by plotting them in grayscale. They appear like typical MRI cross sections.

Every Contour was then plotted on top of the DICOM image as a sanity check (another unit test). The title has the origin directory of the DICOM image, the origin directory of the contour, and the file number. The directories were spot checked with the link csv and the file numbers were checked with the respective directories. I am not a domain expert but the results seam consistent with highlighting the myocardium in MRI images. Certainly the shape of the bulb of contrast seems properly highlighted for each image.

Given the pipeline you have build, can you see any deficiencies that you would change if you had more time? If not, can you think of any improvements/enhancements to the pipeline that you could build in? There are definitely some deficiencies that I would change if I had more time like I mentioned briefly in question 1 of part 2. In the future it would likely be a good idea to read the dataset in batches instead of reading the entire thing into RAM. For a very large dataset the pipeline would likely break since the entire dataset is being read into memory. Another thing that would be good to change in the future is to parallelize the pipeline from part 1. The same sort of calculation is being done many times which lends itself to parallelization.