In [38]:
```
import numpy as np
import pandas as pd
import os
import seaborn as sns
import warnings
warnings.filterwarnings("ignore")
import matplotlib.pyplot as plt
import scipy.stats as stats
from termcolor import colored
from statannot import add_stat_annotation
```

In [39]:
```
#import dataframe
path = 'normalized.pickle'
df = pd.read_pickle(os.path.join(r'C:\Users\Teresa\Desktop\TESE\Textural Analysis', path))
```

In [40]:
```
#separate df according to phase
G1 = df.loc[df['Automatic Label'] == 0]
S_G2 = df.loc[df['Automatic Label'] == 1]

#Check number of nuclei in each phase
print(G1.shape)
print(S_G2.shape)
```

```
(2291, 50)
(1262, 50)
```

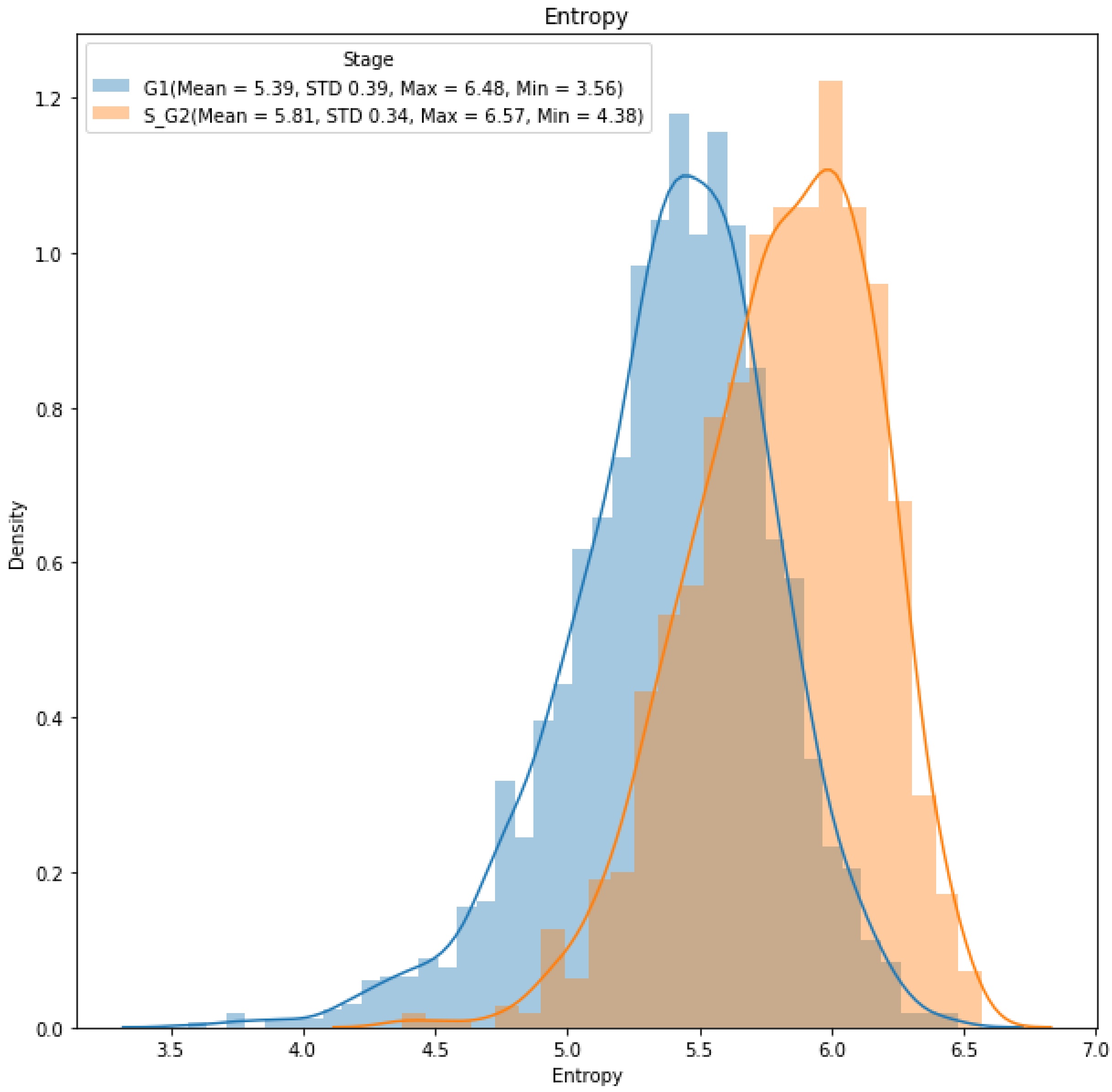
In [41]:
```
#data included in dataframe
list(df.columns)
```

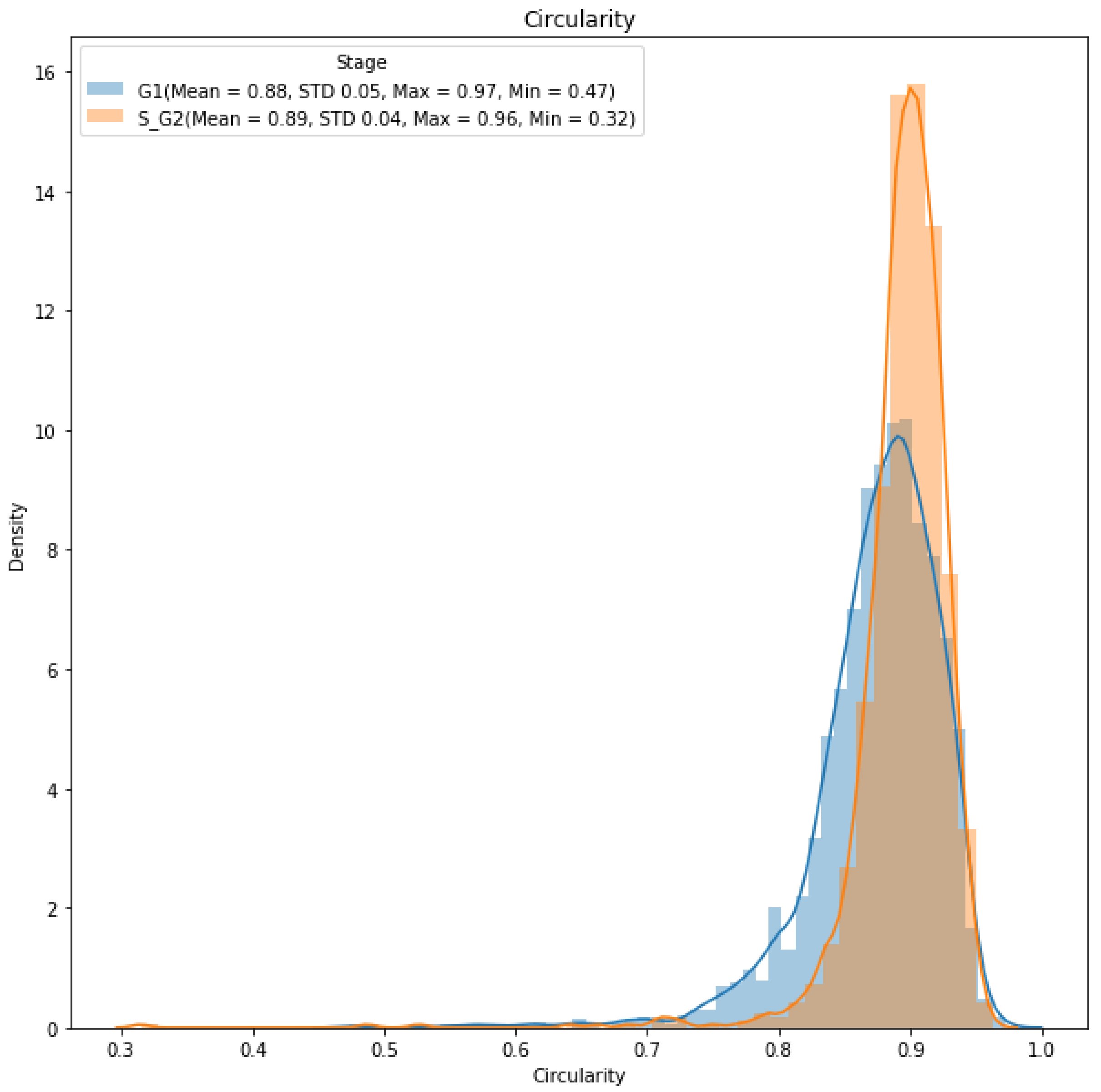
Out[41]:
```
['Area',
 'Image',
 'Mean Green',
 'Mean Intensity',
 'Mean Red',
 'Normalized Mean Green',
 'Normalized Mean Red',
 'Normalized Total Green',
 'Normalized Total Red',
 'Nucleus Patch',
 'Total Green',
 'Total Intensity',
 'Total Red',
 'norm_area',
 'norm_intensity',
 'Automatic Label',
 'Mean',
 'Std',
 'Variance',
 'Skewness',
 'Kurtosis',
 'Uniformity',
 'Invariant Uniformity',
 'GLCM Entropy',
 'GLCM Invariant Entropy',
 'Correlation',
 'Invariant Correlation',
 'Dissimilarity',
 'Invariant Dissimilarity',
 'Contrast',
 'Invariant Contrast',
 'Homogeneity',
 'Invariant Homogeneity',
 'Energy',
 'Invariant Energy',
 'BB Area',
 'Centroid',
 'Weighted Centroid',
 'Centroid Divergence',
 'Eccentricity',
 'Equivalent Diameter',
 'Major Axis Length',
 'Minor Axis Length',
 'Max Intensity',
 'Min Intensity',
 'Orientation',
 'Perimeter',
 'Solidity',
 'Entropy',
 'Circularity']
```

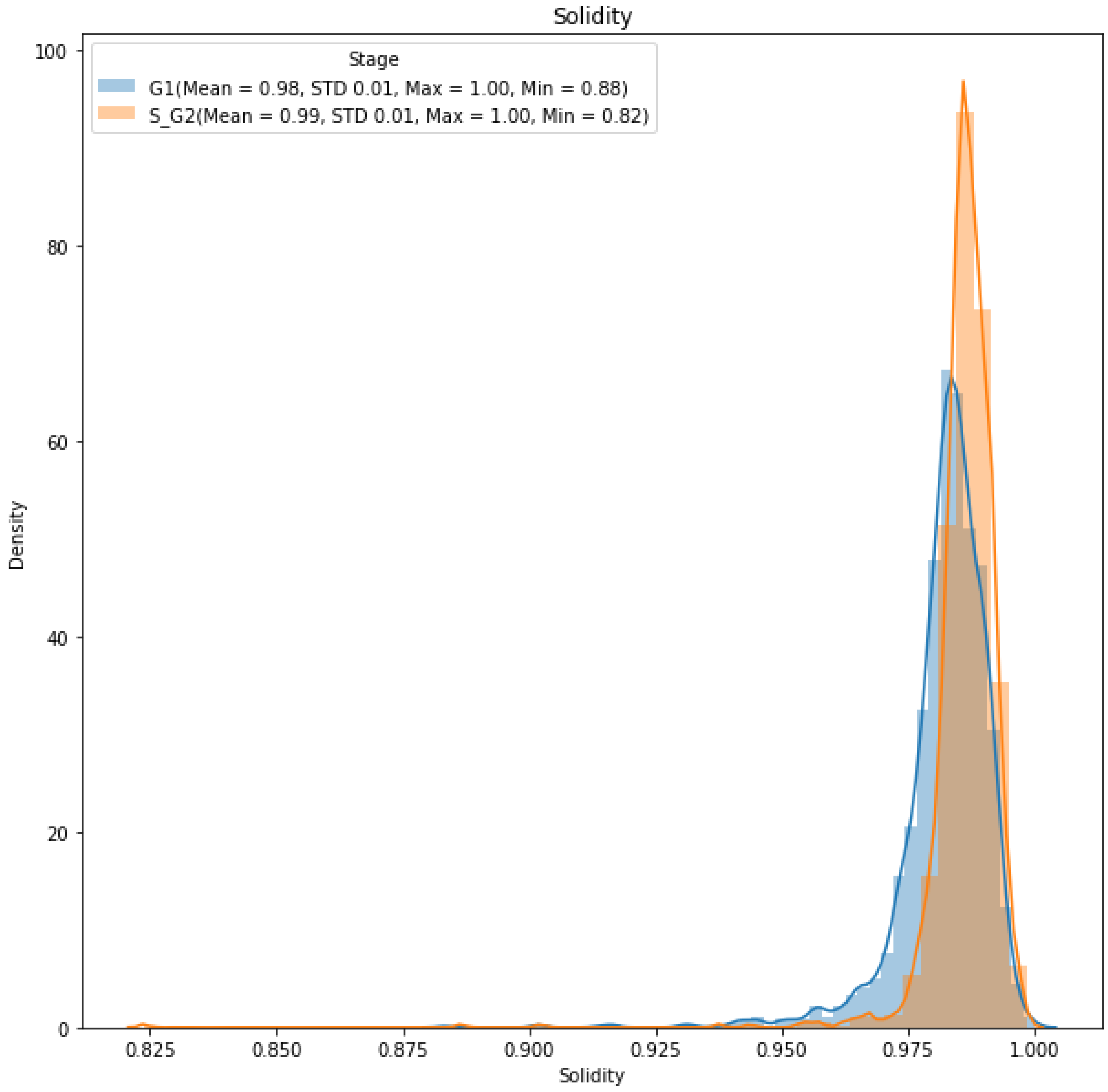
In [42]:
```
#utest

def plt_hist_and_stats_utest(columns, types, histogram = True):
    for column in columns:
        #create a new figure
        plt.figure()
        plt.rcParams['figure.figsize'] = (10,10)
        for subtype in types:

            tp = eval(subtype)
            #subset to the type

            #compute some statistics
            aux = tp[column].describe()

            #Draw the density plot
            plt.rcParams['figure.figsize'] = (10,10)
            sns.distplot(tp[column], hist = histogram, kde = True,
                         label = subtype+r'(Mean = %0.2f, STD %0.2f, Max = %0.2f, Min = %0.2f)' % (aux['mean'], aux['st
d'], aux['max'], aux['min']))
            plt.legend(prop = {'size': 10}, title = 'Stage')
            plt.title(column)
            plt.xlabel(column)
            plt.ylabel('Density')


        subset1  = eval(types[0])
        subset2 = eval(types[1])
        u_statistic, pVal = stats.mannwhitneyu(subset1[column], subset2[column], alternative = 'two-sided')
        print('THE P-VALUE IS:')
        print(pVal)
        if pVal < 0.05:
            aux = '\033[1m' + ' is ' + '\033[0m'
            aux = colored(aux, 'blue')
            hipothesis =  'H1: The difference' + aux + 'statistically significant (at significance level: 0.05).'
        else:
            aux = '\033[1m' + ' is not ' + '\033[0m'
            aux = colored(aux, 'blue')
            hipothesis ='H0: The difference' + aux + 'statistically significant (at significance level: 0.05).'
        print('For feature ' + column + ' - ' + hipothesis)
        #print(u_statistic)
```

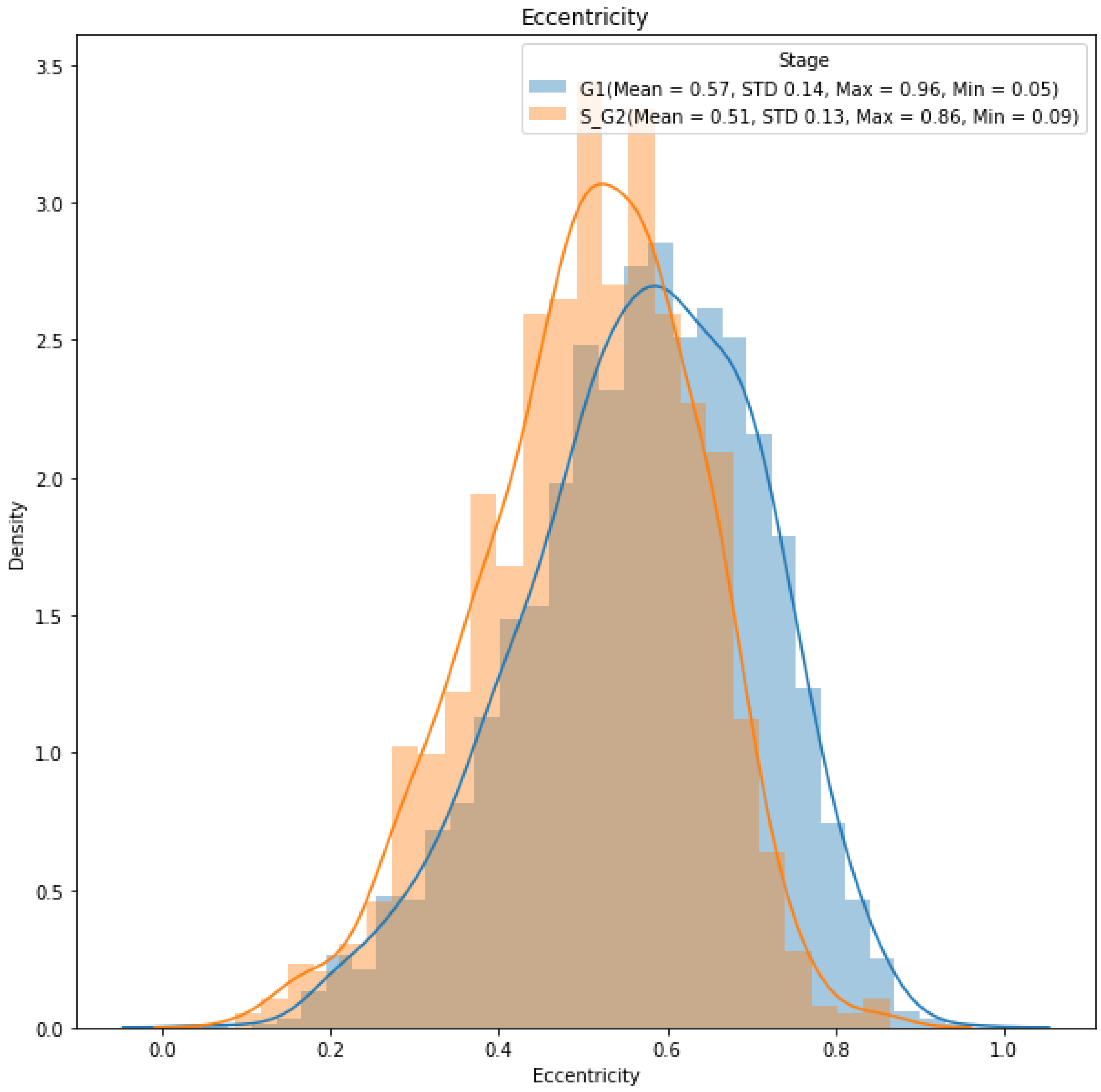
In [43]:
```
#utest

types = ['G1','S_G2']
columns = ['Entropy',
            'Circularity',
            'Solidity',
            'Eccentricity',
            'Perimeter',
            'Area',
            'Mean Intensity',
            'Variance',
            'Skewness',
            'Kurtosis',
            'Invariant Uniformity',
            'GLCM Invariant Entropy',
            'Invariant Correlation',
            'Invariant Dissimilarity',
            'Invariant Contrast',
            'Invariant Homogeneity',
            'Invariant Energy',
            'Equivalent Diameter',
            'Centroid Divergence']

plt_hist_and_stats_utest(columns, types, True)
```

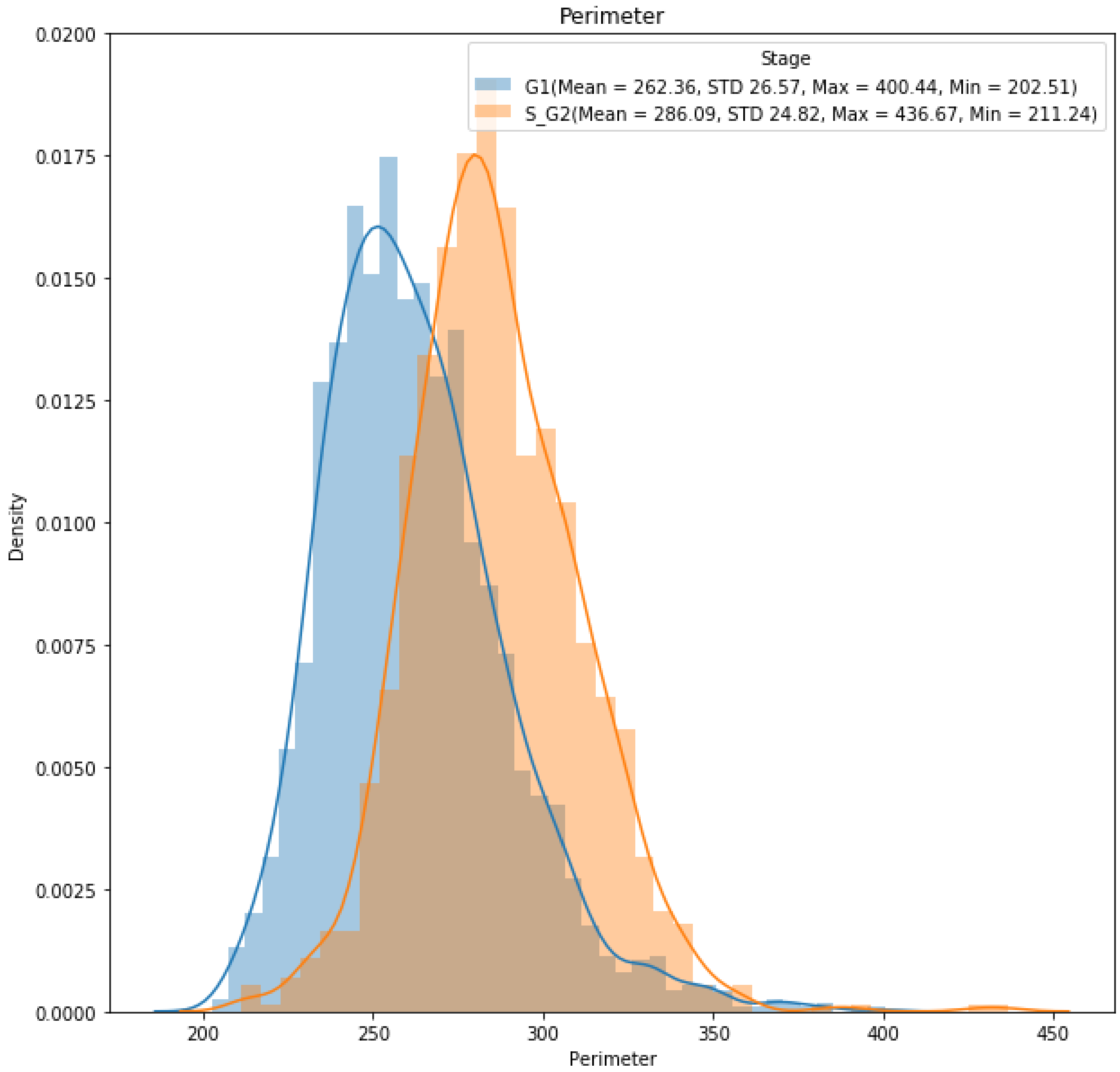
```
THE P-VALUE IS:
3.616021368134499e-189
For feature Entropy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
2.6768406762283908e-42
For feature Circularity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
3.205597269953735e-61
For feature Solidity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.6783604860866635e-39
For feature Eccentricity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
8.065038876911337e-147
For feature Perimeter - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
6.087580336900195e-181
For feature Area - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
4.8301901786625234e-107
For feature Mean Intensity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
7.206003202766363e-103
For feature Variance - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
7.624210301004186e-127
For feature Skewness - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
6.816372232033188e-141
For feature Kurtosis - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
2.4427628801449847e-54
For feature Invariant Uniformity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
7.830668595122856e-244
For feature GLCM Invariant Entropy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
3.6611469127177655e-231
For feature Invariant Correlation - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
3.438335276147601e-41
For feature Invariant Dissimilarity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
9.488027440859892e-44
For feature Invariant Contrast - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.8548852666700498e-06
For feature Invariant Homogeneity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
3.1721833059713423e-12
For feature Invariant Energy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
6.087580336900195e-181
For feature Equivalent Diameter - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
0.005067838448174236
For feature Centroid Divergence - H1: The difference is statistically significant (at significance level: 0.05).
```


Entropy
Stage
G1(Mean = 5.39, STD 0.39, Max = 6.48, Min = 3.56)
S_G2(Mean = 5.81, STD 0.34, Max = 6.57, Min = 4.38)
Density
0.0
0.2
0.4
0.6
0.8
1.0
1.2
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
Entropy

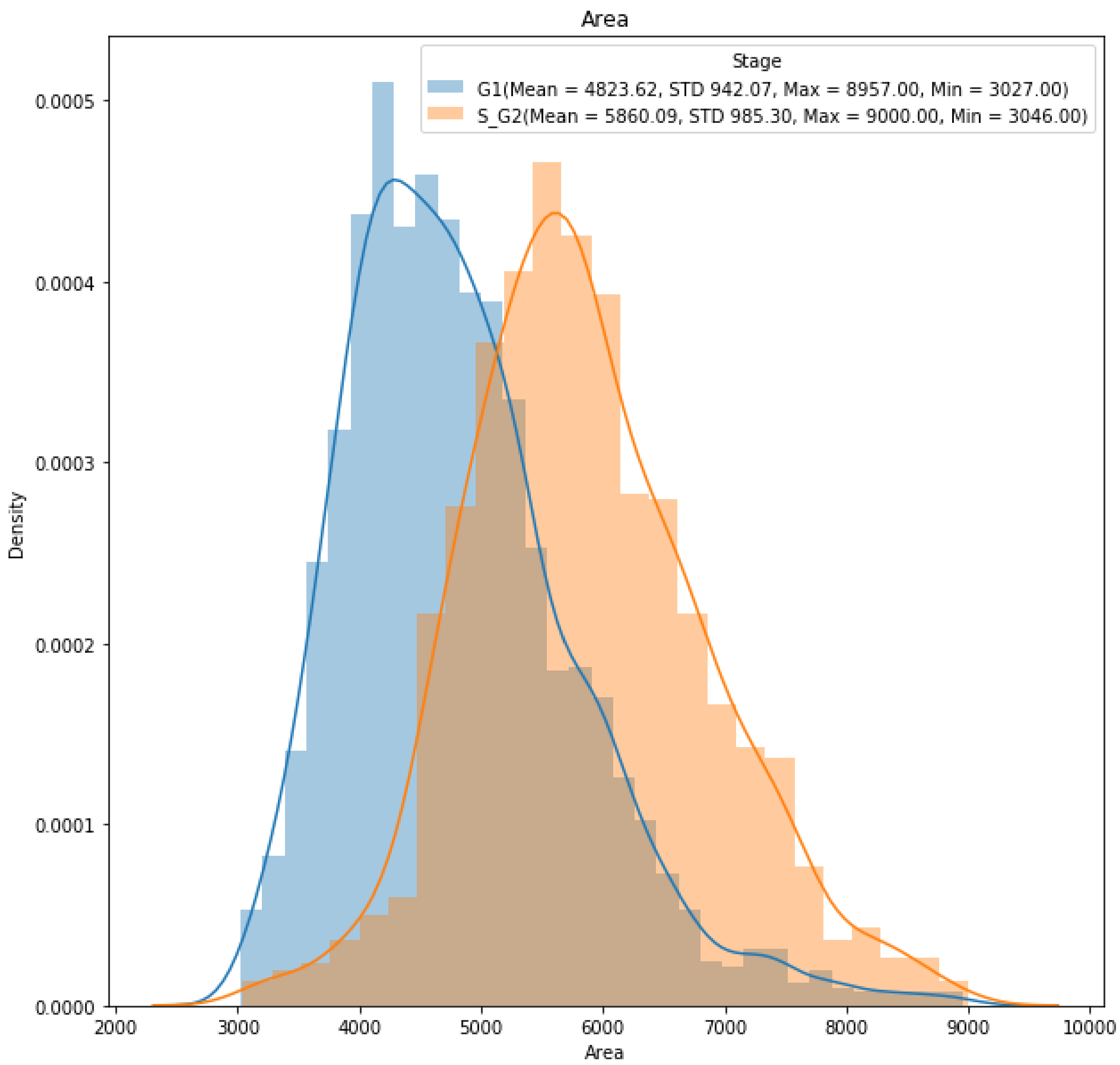

Circularity
Stage
G1(Mean = 0.88, STD 0.05, Max = 0.97, Min = 0.47)
S_G2(Mean = 0.89, STD 0.04, Max = 0.96, Min = 0.32)
Density
0
2
4
6
8
10
12
14
16
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Circularity


Solidity
Stage
G1(Mean = 0.98, STD 0.01, Max = 1.00, Min = 0.88)
S_G2(Mean = 0.99, STD 0.01, Max = 1.00, Min = 0.82)
Density
0
20
40
60
80
100
0.825
0.850
0.875
0.900
0.925
0.950
0.975
1.000
Solidity

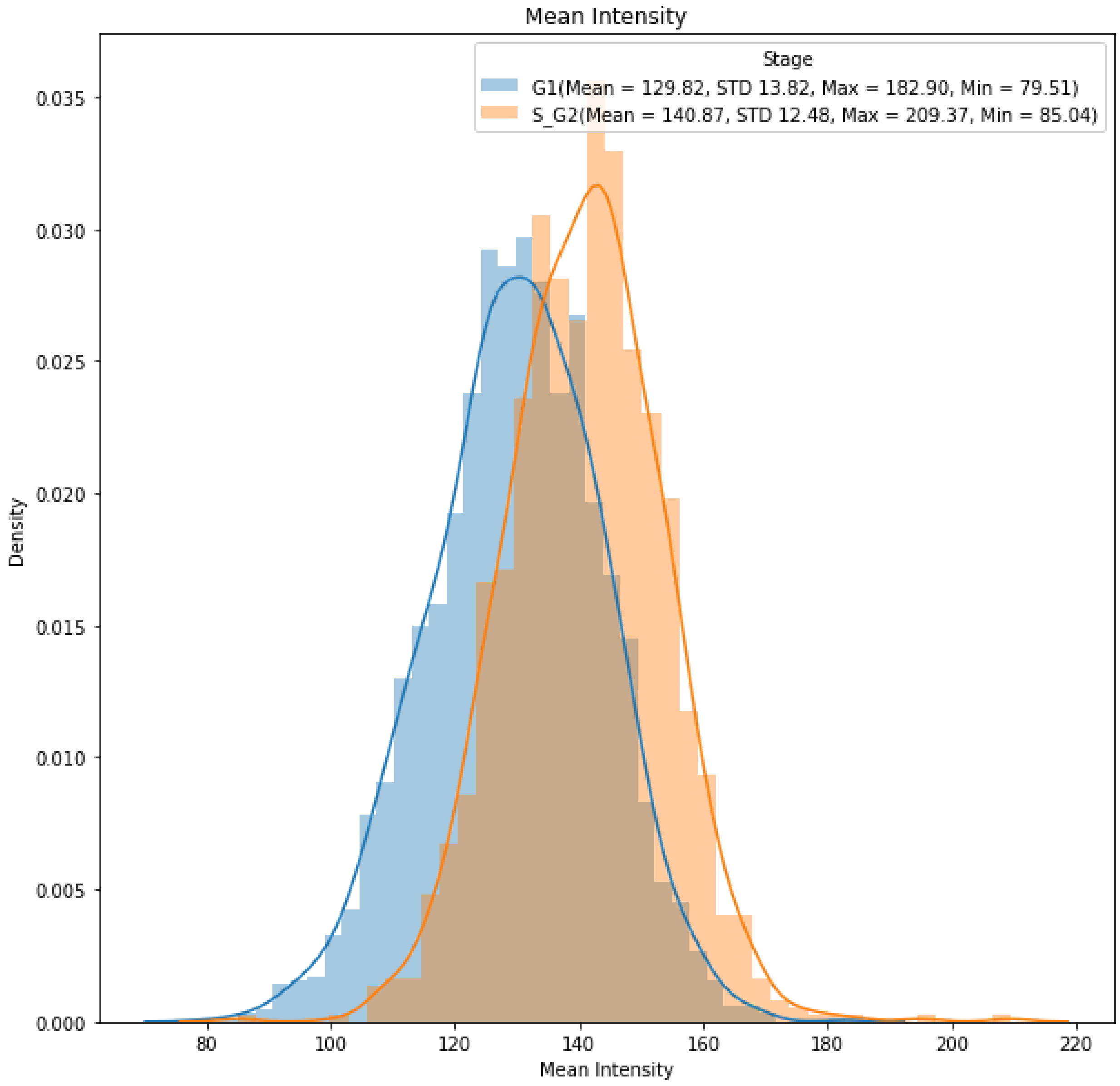

Eccentricity
Stage
G1(Mean = 0.57, STD 0.14, Max = 0.96, Min = 0.05)
S_G2(Mean = 0.51, STD 0.13, Max = 0.86, Min = 0.09)
Density
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
0.0
0.2
0.4
0.6
0.8
1.0
Eccentricity

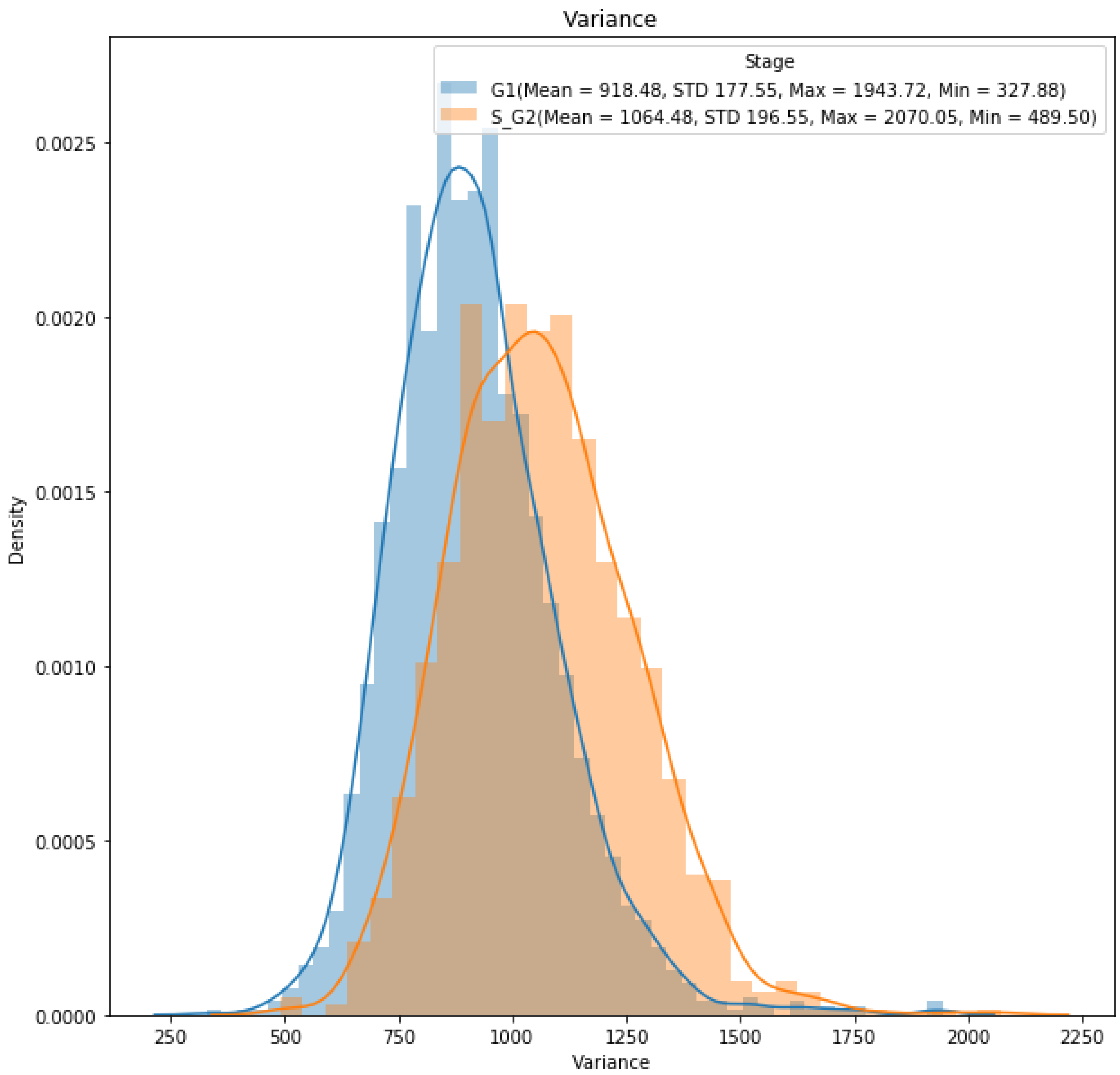


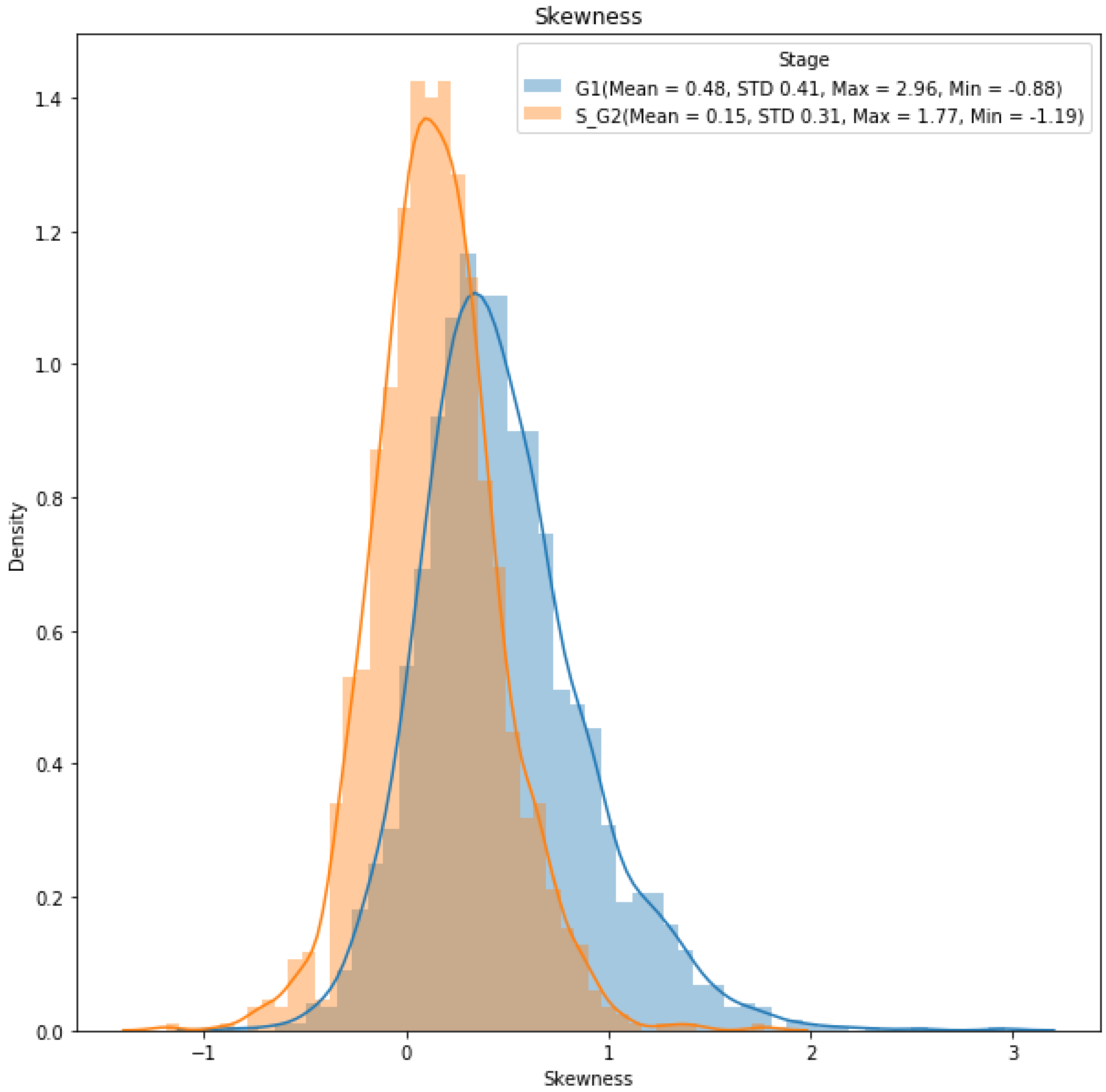
Perimeter
Stage
G1(Mean = 262.36, STD 26.57, Max = 400.44, Min = 202.51)
S_G2(Mean = 286.09, STD 24.82, Max = 436.67, Min = 211.24)
0.0200
0.0175
0.0150
0.0125
0.0100
0.0075
0.0050
0.0025
0.0000
Density
200
250
300
350
400
450
Perimeter

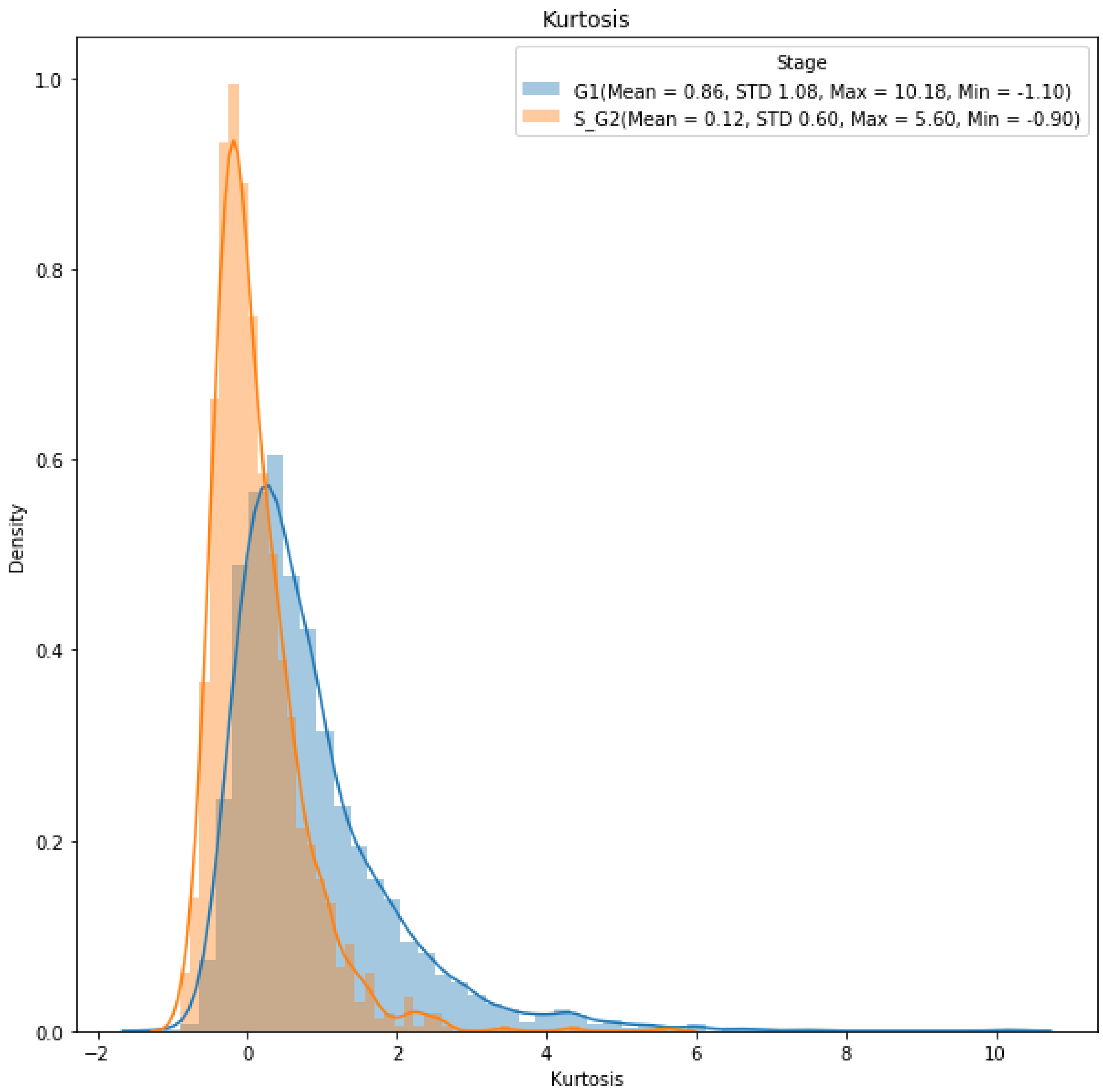


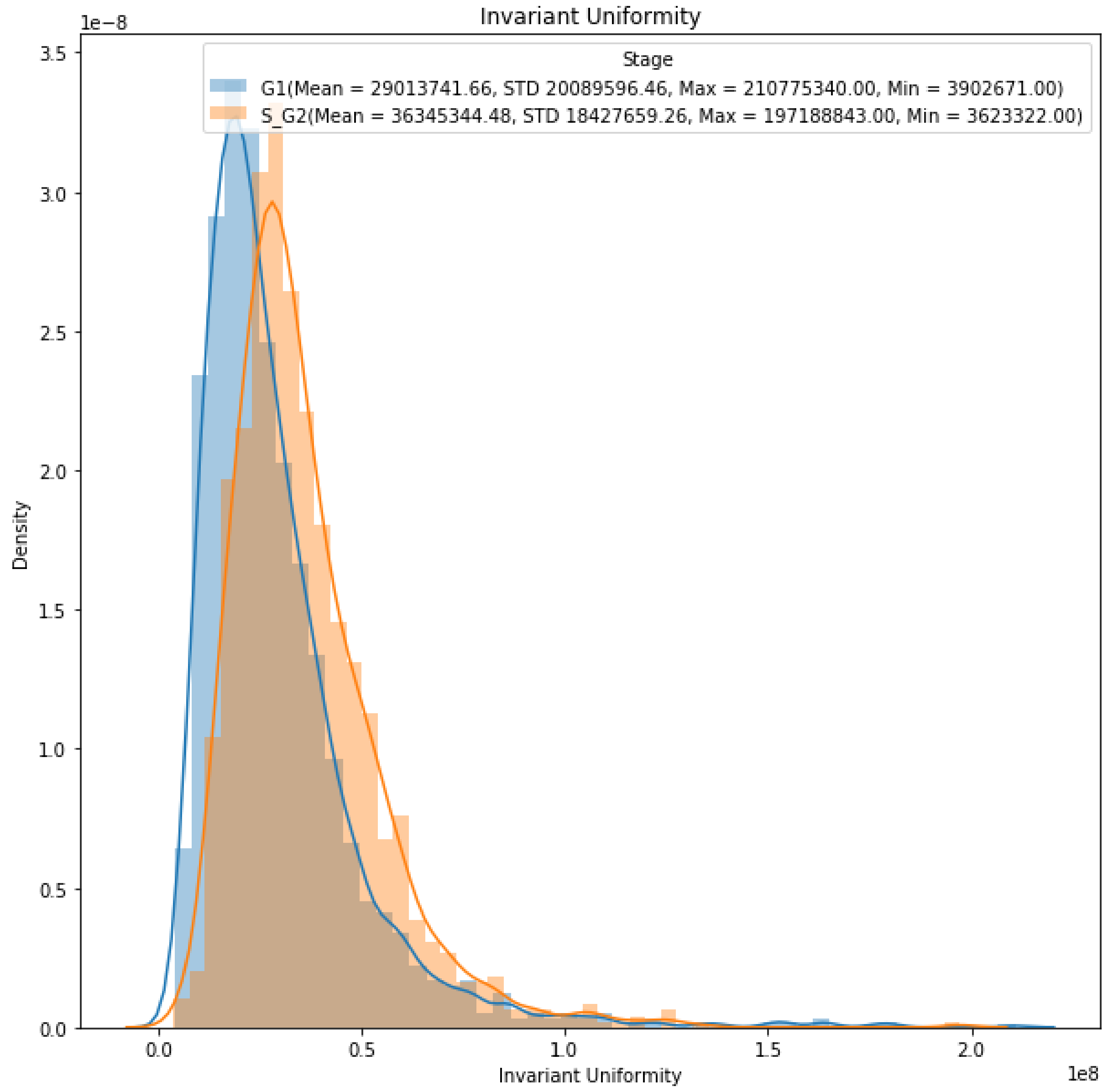
Area
Stage
G1(Mean = 4823.62, STD 942.07, Max = 8957.00, Min = 3027.00)
S_G2(Mean = 5860.09, STD 985.30, Max = 9000.00, Min = 3046.00)
0.0005
0.0004
0.0003
0.0002
0.0001
0.0000
Density
2000
3000
4000
5000
6000
7000
8000
9000
10000
Area

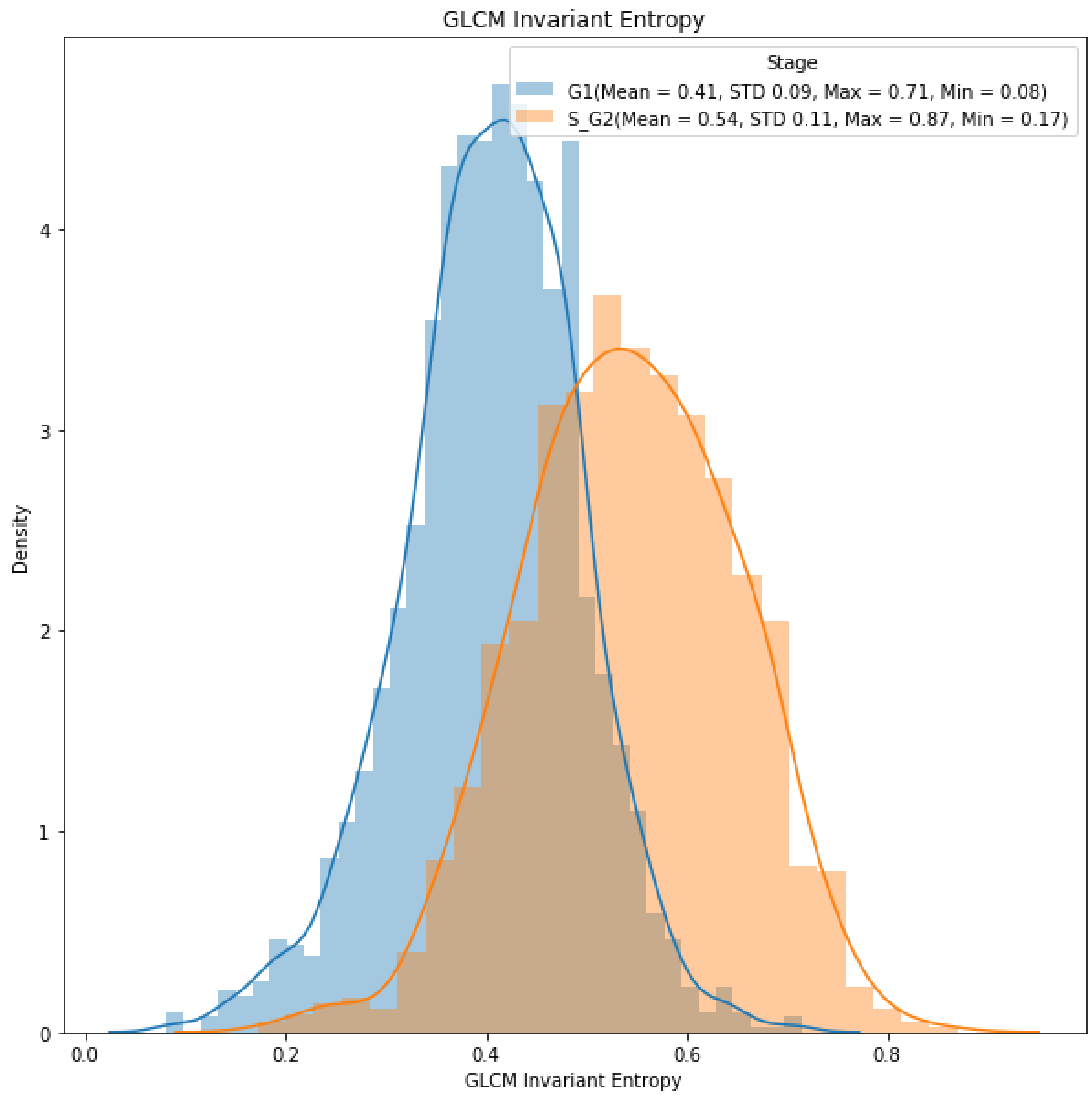

Mean Intensity
Stage
G1(Mean = 129.82, STD 13.82, Max = 182.90, Min = 79.51)
S_G2(Mean = 140.87, STD 12.48, Max = 209.37, Min = 85.04)
Density
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
80
100
120
140
160
180
200
220
Mean Intensity

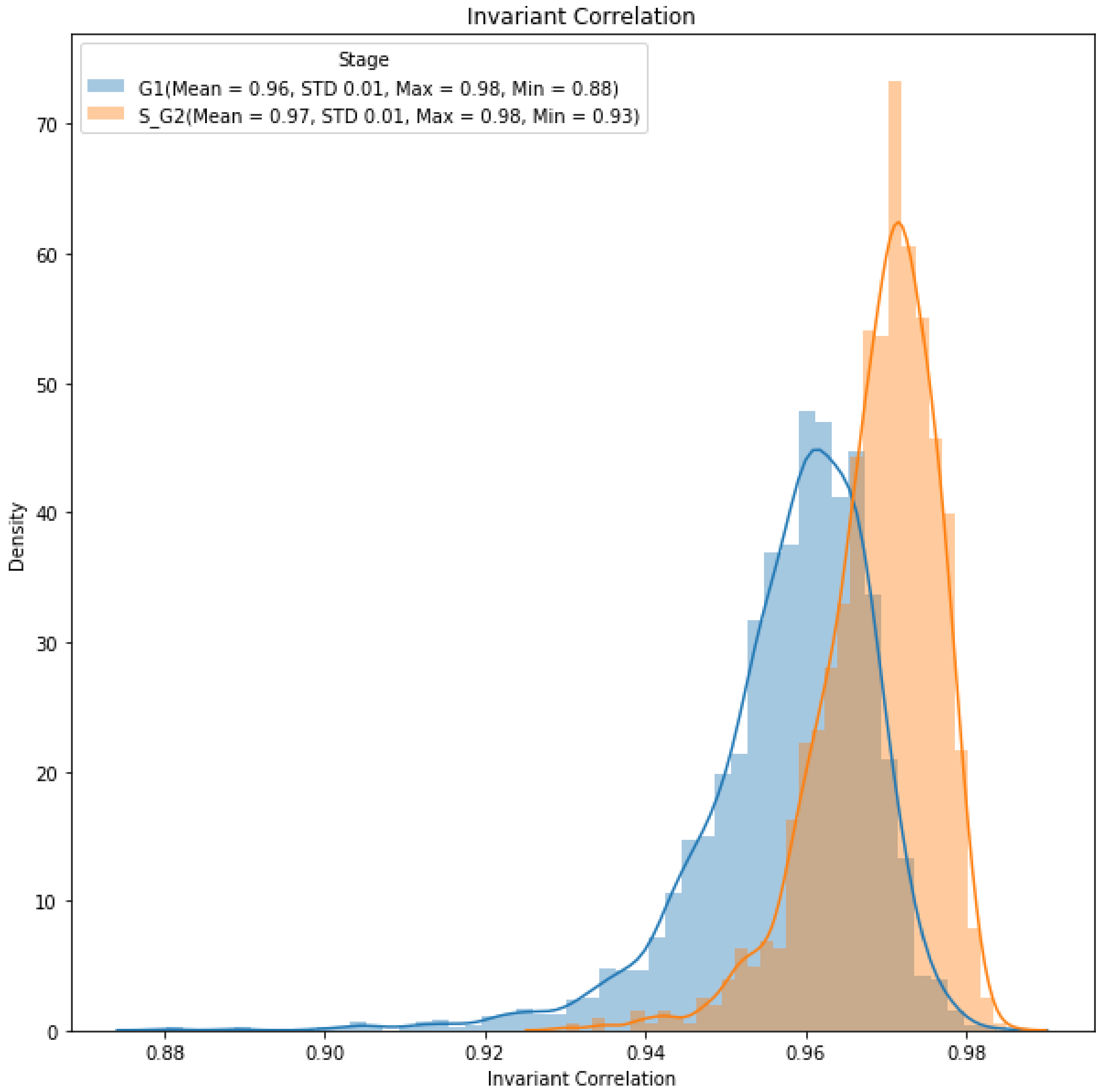

Variance
Stage
G1(Mean = 918.48, STD 177.55, Max = 1943.72, Min = 327.88)
S_G2(Mean = 1064.48, STD 196.55, Max = 2070.05, Min = 489.50)
Density
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
250
500
750
1000
1250
1500
1750
2000
2250
Variance


Skewness
Stage
G1(Mean = 0.48, STD 0.41, Max = 2.96, Min = -0.88)
S_G2(Mean = 0.15, STD 0.31, Max = 1.77, Min = -1.19)
Density
Skewness

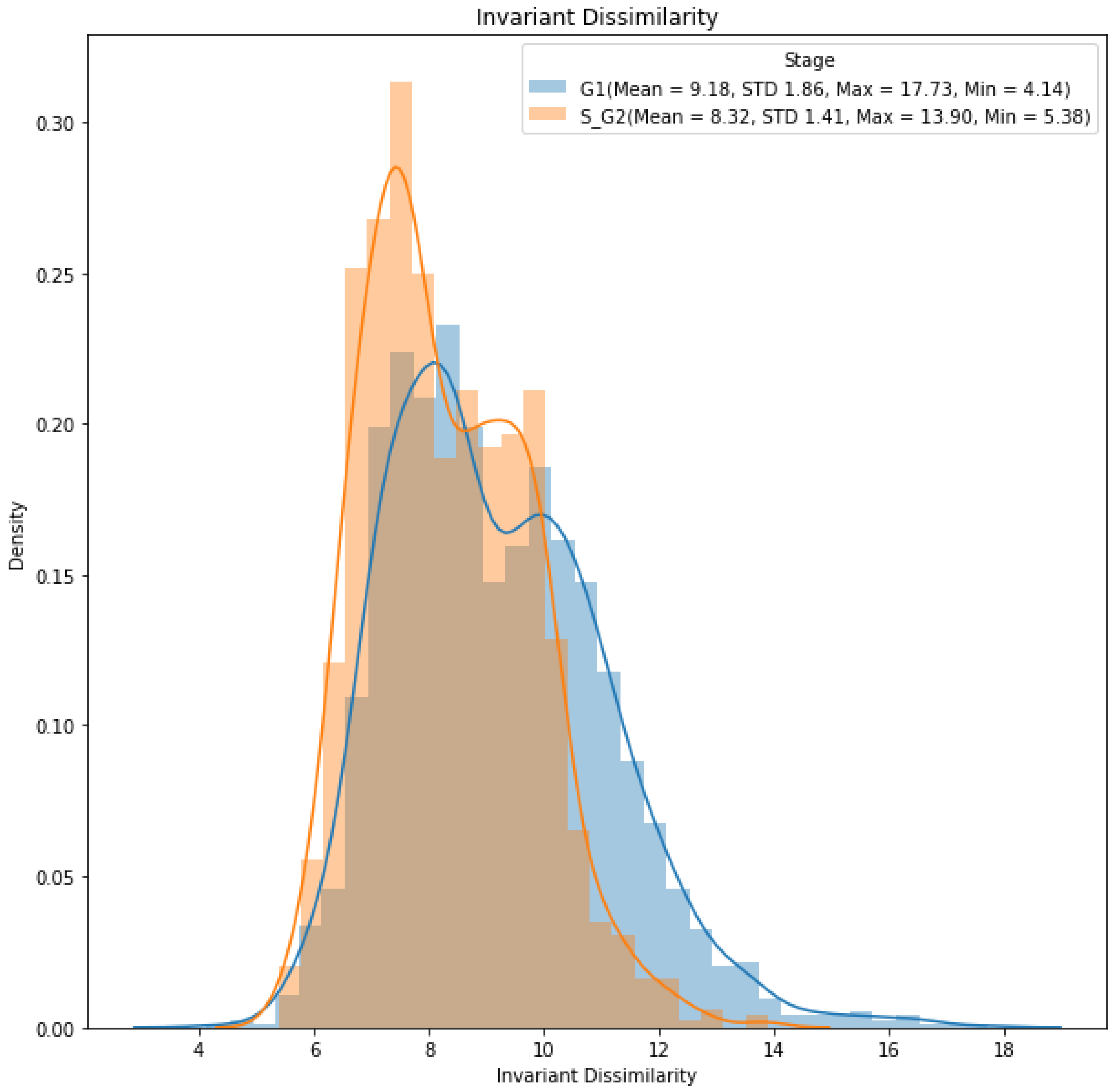

Kurtosis
Stage
G1(Mean = 0.86, STD 1.08, Max = 10.18, Min = -1.10)
S_G2(Mean = 0.12, STD 0.60, Max = 5.60, Min = -0.90)
Density
Kurtosis


Invariant Uniformity
1e−8
Stage
G1(Mean = 29013741.66, STD 20089596.46, Max = 210775340.00, Min = 3902671.00)
S_G2(Mean = 36345344.48, STD 18427659.26, Max = 197188843.00, Min = 3623322.00)
Density
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
1e8
Invariant Uniformity

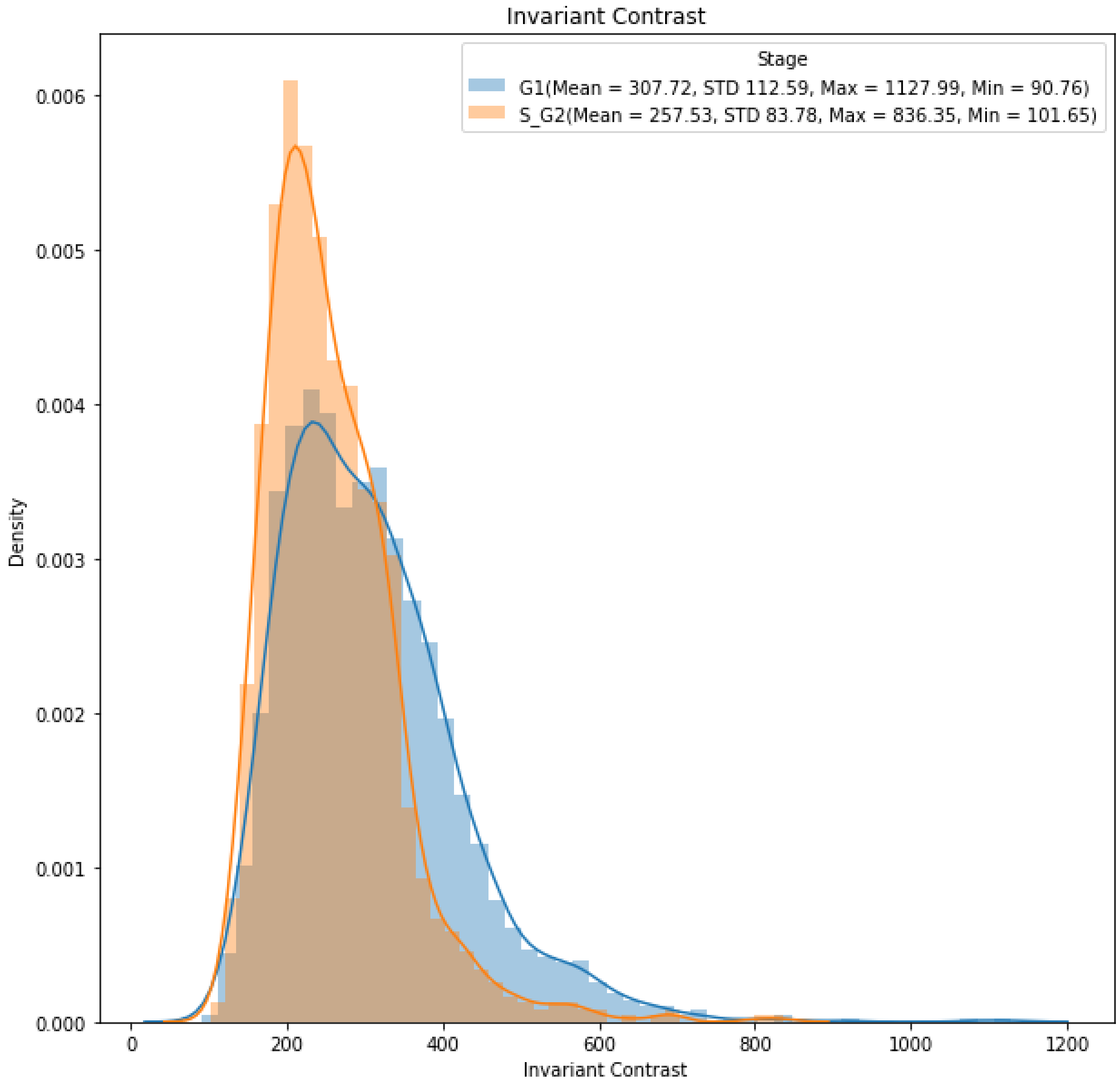

GLCM Invariant Entropy
Stage
G1(Mean = 0.41, STD 0.09, Max = 0.71, Min = 0.08)
S_G2(Mean = 0.54, STD 0.11, Max = 0.87, Min = 0.17)
Density
4
3
2
1
0
0.0
0.2
0.4
0.6
0.8
GLCM Invariant Entropy

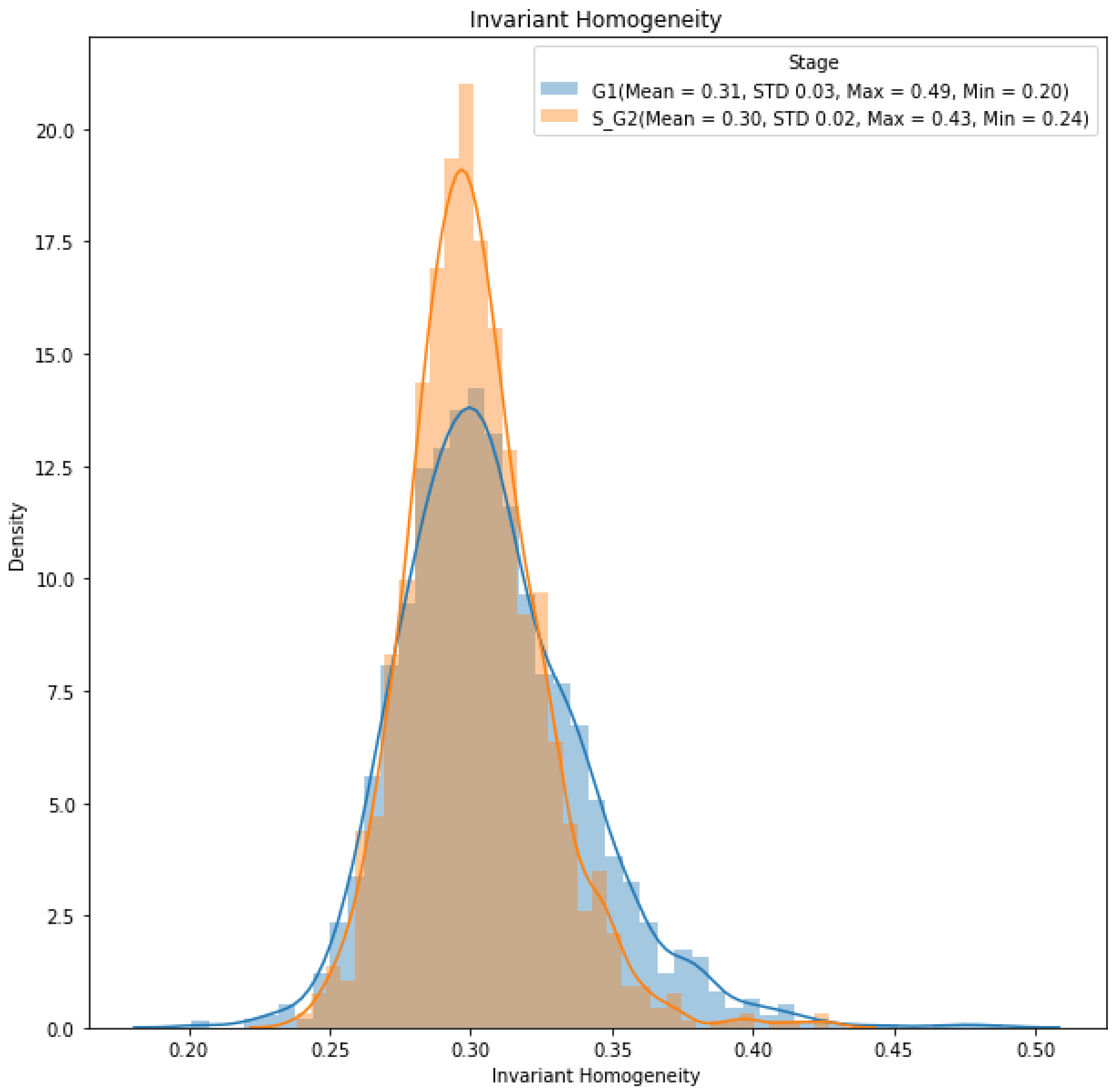


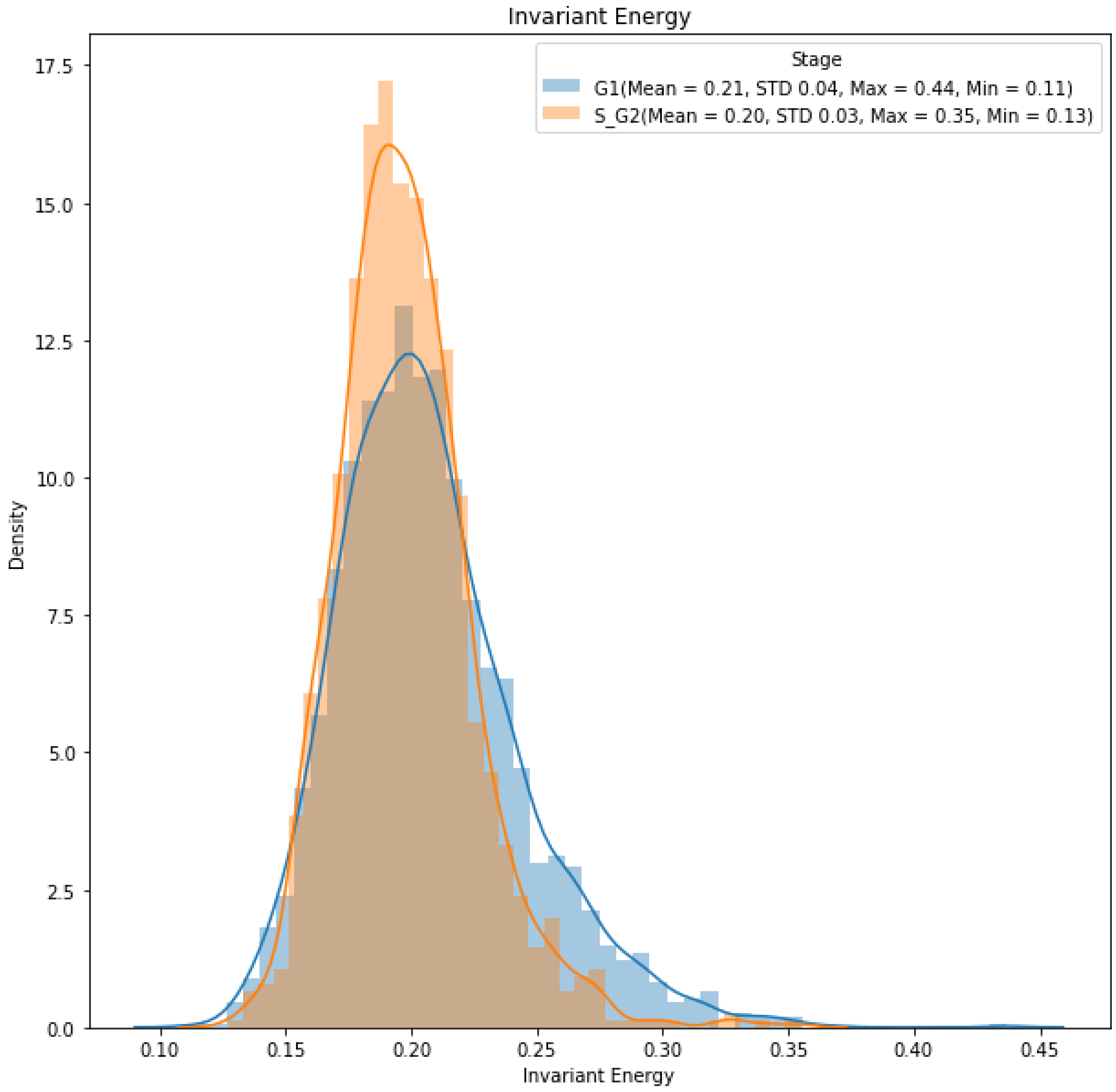
Invariant Correlation
Stage
G1(Mean = 0.96, STD 0.01, Max = 0.98, Min = 0.88)
S_G2(Mean = 0.97, STD 0.01, Max = 0.98, Min = 0.93)
Density
Invariant Correlation

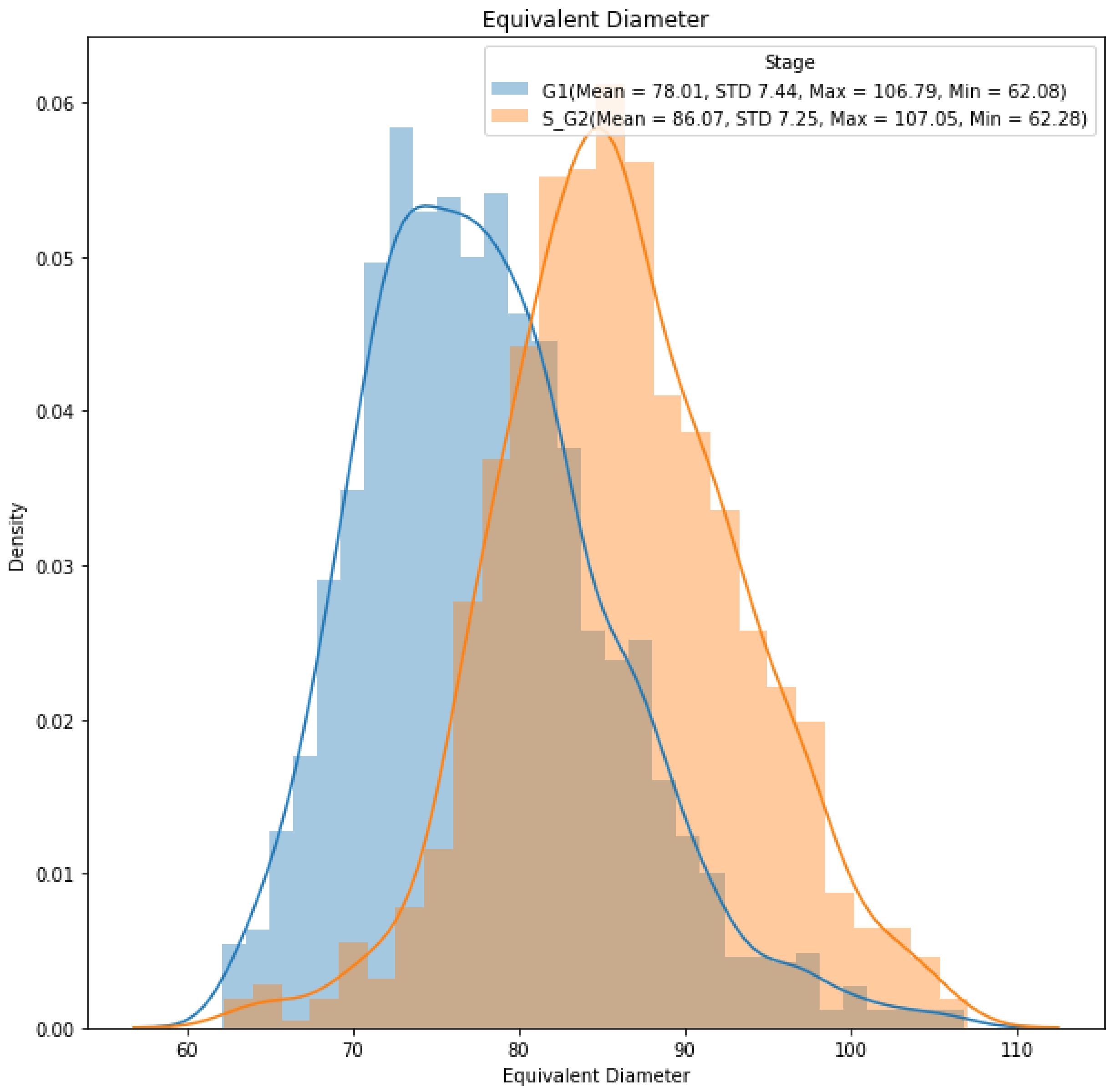


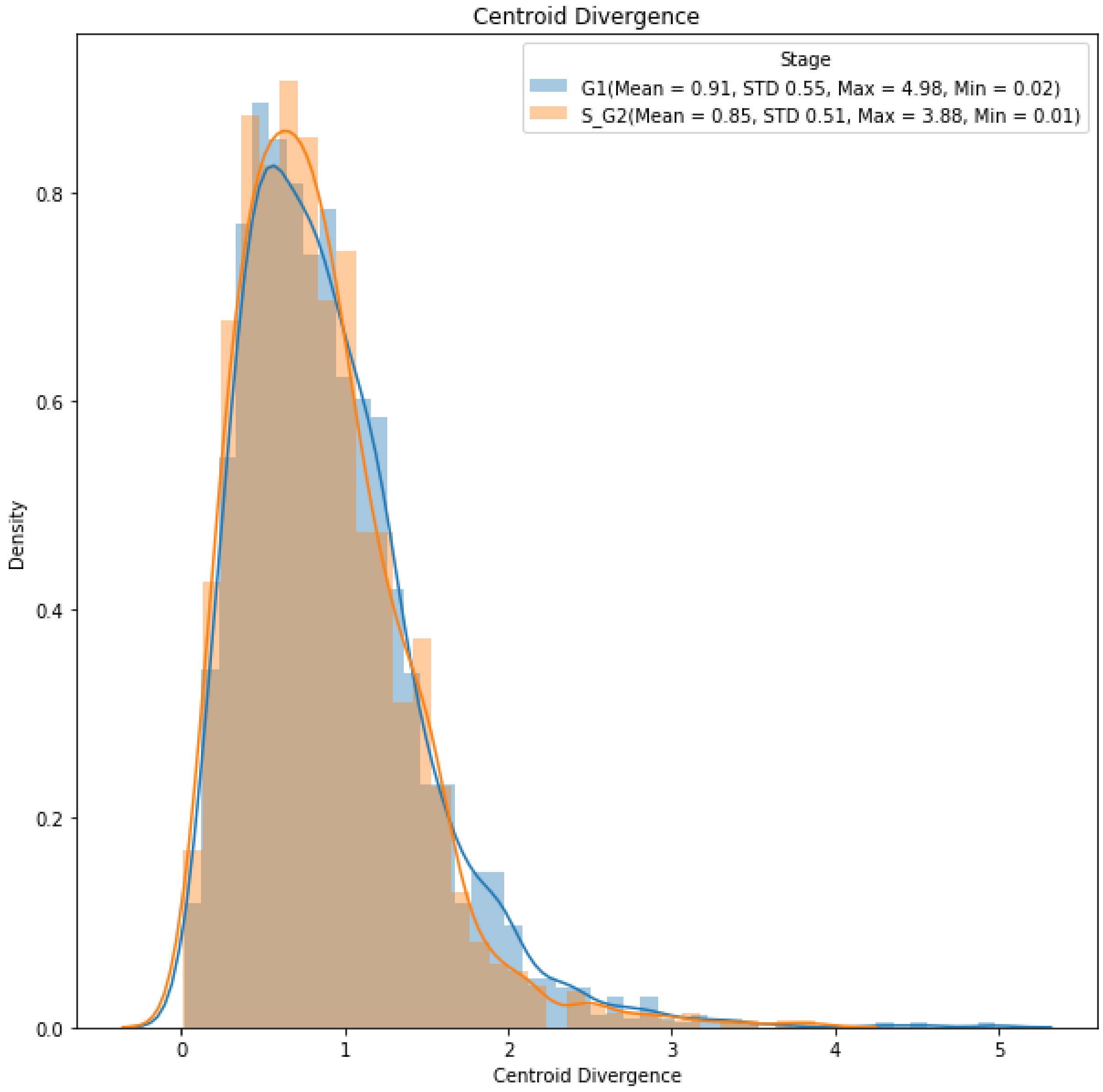
Invariant Dissimilarity
Stage
G1(Mean = 9.18, STD 1.86, Max = 17.73, Min = 4.14)
S_G2(Mean = 8.32, STD 1.41, Max = 13.90, Min = 5.38)
Density
Invariant Dissimilarity


Invariant Contrast
Stage
G1(Mean = 307.72, STD 112.59, Max = 1127.99, Min = 90.76)
S_G2(Mean = 257.53, STD 83.78, Max = 836.35, Min = 101.65)
Density
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0
200
400
600
800
1000
1200
Invariant Contrast


Invariant Homogeneity
Stage
G1(Mean = 0.31, STD 0.03, Max = 0.49, Min = 0.20)
S_G2(Mean = 0.30, STD 0.02, Max = 0.43, Min = 0.24)
Density
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
20.0
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Invariant Homogeneity


Invariant Energy
Stage
G1(Mean = 0.21, STD 0.04, Max = 0.44, Min = 0.11)
S_G2(Mean = 0.20, STD 0.03, Max = 0.35, Min = 0.13)
Density
Invariant Energy
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45


Equivalent Diameter
Stage
G1(Mean = 78.01, STD 7.44, Max = 106.79, Min = 62.08)
S_G2(Mean = 86.07, STD 7.25, Max = 107.05, Min = 62.28)
Density
Equivalent Diameter
0.00
0.01
0.02
0.03
0.04
0.05
0.06
60
70
80
90
100
110



In [44]:
```
#auxiliary function

def stats_pval_table(column, types):
    import scipy.stats as stats
    subset1  = eval(types[0])
    subset2 = eval(types[1])
    u_statistic, pVal = stats.mannwhitneyu(subset1[column], subset2[column], alternative = 'two-sided')
    return pVal
```

In [45]:
```
#get a table with p-values

types = ['G1','S_G2']

columns = ['Entropy',
            'Circularity',
            'Solidity',
            'Eccentricity',
            'Perimeter',
            'Area',
            'Mean Intensity',
            'Variance',
            'Skewness',
            'Kurtosis',
            'Invariant Uniformity',
            'GLCM Invariant Entropy',
            'Invariant Correlation',
            'Invariant Dissimilarity',
            'Invariant Contrast',
            'Invariant Homogeneity',
            'Invariant Energy',
            'Equivalent Diameter',
            'Centroid Divergence']

pval_table = pd.DataFrame(columns =[
            'Entropy',
            'Circularity',
            'Solidity',
            'Eccentricity',
            'Perimeter',
            'Area',
            'Mean Intensity',
            'Variance',
            'Skewness',
            'Kurtosis',
            'Invariant Uniformity',
            'GLCM Invariant Entropy',
            'Invariant Correlation',
            'Invariant Dissimilarity',
            'Invariant Contrast',
            'Invariant Homogeneity',
            'Invariant Energy',
            'Equivalent Diameter',
            'Centroid Divergence'])

pval_list = []
for column in columns:
    pval = stats_pval_table(column, types)
    pval_list.append(pval)
comparison = types[0] + ' vs ' + types[1]

i=0
res = {}
for label in columns:
    res[label] = pval_list[i]
    i = i+1


row = len(pval_table)
pval_table.loc[row] = res
```

In [46]:
```
#visualize
pval_table
```

Out[46]:

	Entropy	Circularity	Solidity	Eccentricity	Perimeter	Area	Mean Intensity	Variance	Skewness	Kurtosis	Invariant Uniformity	GLCM Invariant Entropy
0	3.616021e-189	2.676841e-42	3.205597e-61	1.678360e-39	8.065039e-147	6.087580e-181	4.830190e-107	7.206003e-103	7.624210e-127	6.816372e-141	2.442763e-54	7.830669e-244

In [47]:
```
#test for different p-values
pval_table < 0.05
```

Out[47]:

	Entropy	Circularity	Solidity	Eccentricity	Perimeter	Area	Mean Intensity	Variance	Skewness	Kurtosis	Invariant Uniformity	GLCM Invariant Entropy	Invariant Correlation	Inva Dissimi
0	True	True	True	True	True	True	True	True	True	True	True	True	True	

In [48]:
```
pval_table < 0.01
```

Out[48]:

	Entropy	Circularity	Solidity	Eccentricity	Perimeter	Area	Mean Intensity	Variance	Skewness	Kurtosis	Invariant Uniformity	GLCM Invariant Entropy	Invariant Correlation	Inva Dissimi
0	True	True	True	True	True	True	True	True	True	True	True	True	True	

In [49]:
```
pval_table < 0.001
```

Out[49]:

	Entropy	Circularity	Solidity	Eccentricity	Perimeter	Area	Mean Intensity	Variance	Skewness	Kurtosis	Invariant Uniformity	GLCM Invariant Entropy	Invariant Correlation	Inva Dissimi
0	True	True	True	True	True	True	True	True	True	True	True	True	True	

In [50]:
```
pval_table < 0.0001
```

Out[50]:

	Entropy	Circularity	Solidity	Eccentricity	Perimeter	Area	Mean Intensity	Variance	Skewness	Kurtosis	Invariant Uniformity	GLCM Invariant Entropy	Invariant Correlation	Inva Dissimi
0	True	True	True	True	True	True	True	True	True	True	True	True	True	

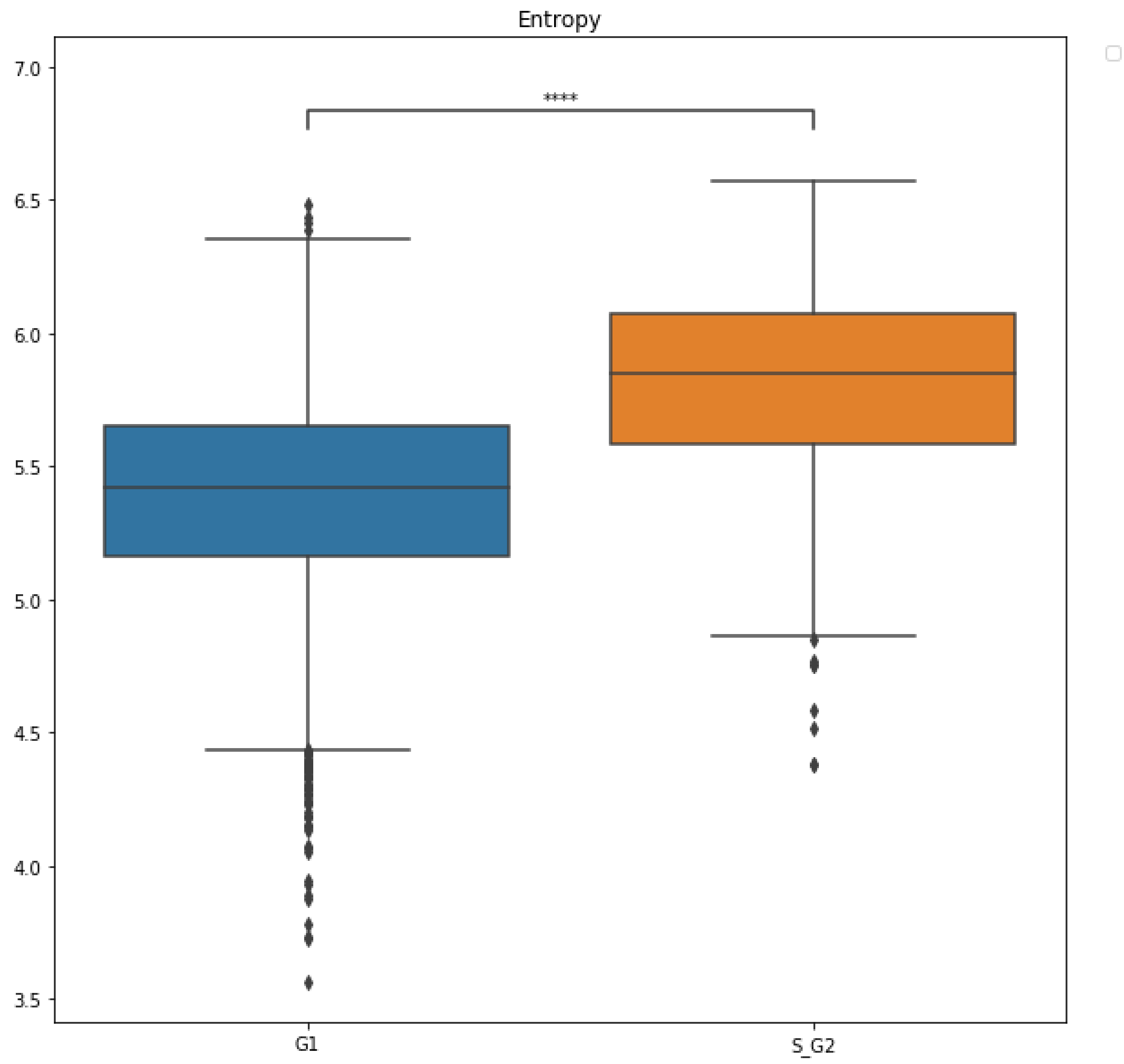
In [51]:
```
#another auxiliary function

def obtain_columns(columns, typee):
        import scipy.stats as stats
        subset  = eval(typee)
        subset1 = subset[columns]
        return subset1
```

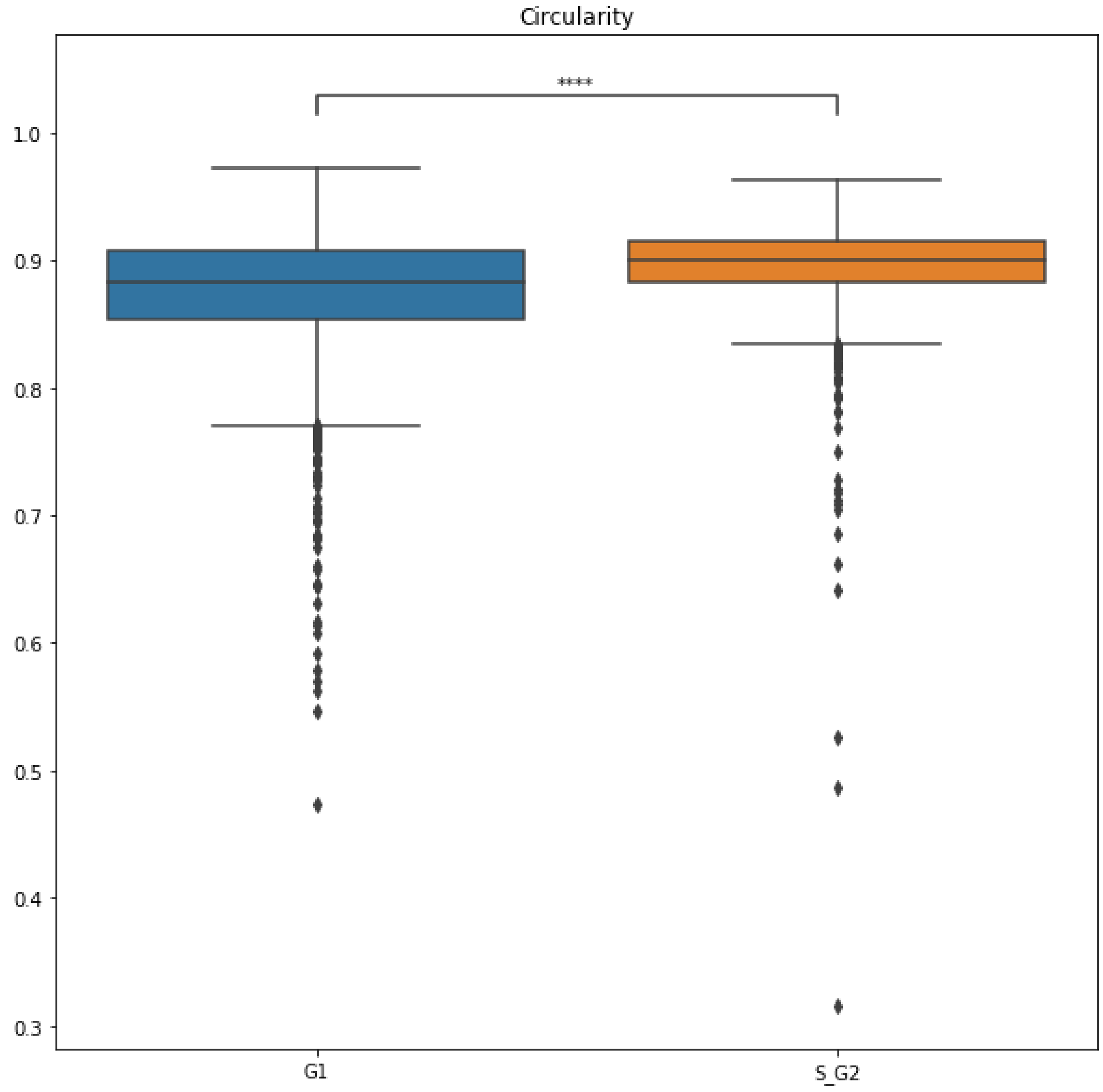
In [52]:
```
#obtain whisker plots

types = ['G1','S_G2']
columns = ['Entropy',
            'Circularity',
            'Solidity',
            'Eccentricity',
            'Perimeter',
            'Area',
            'Mean Intensity',
            'Variance',
            'Skewness',
            'Kurtosis',
            'Invariant Uniformity',
            'GLCM Invariant Entropy',
            'Invariant Correlation',
            'Invariant Dissimilarity',
            'Invariant Contrast',
            'Invariant Homogeneity',
            'Invariant Energy',
            'Equivalent Diameter',
            'Centroid Divergence']


for column in columns:
    plt.figure()
    filename = []
    table = pd.DataFrame(columns = [])
    for tp in types:
        columnn = obtain_columns(column, tp)


        a = []

        for i in np.arange(len(columnn.to_numpy())):
            a.append(columnn.to_numpy()[i])



        table_aux = pd.DataFrame({tp: a})
        table = pd.concat([table,table_aux], axis=1)


    ax = sns.boxplot(data=table)
    add_stat_annotation(ax, data=table,
    box_pairs=[(types[0], types[1])],
                test='Mann-Whitney', text_format='star', loc='inside', verbose=2)
    plt.rcParams['figure.figsize'] = (10,10)
    plt.legend(loc='upper left', bbox_to_anchor=(1.03, 1))
    plt.title(column)
```

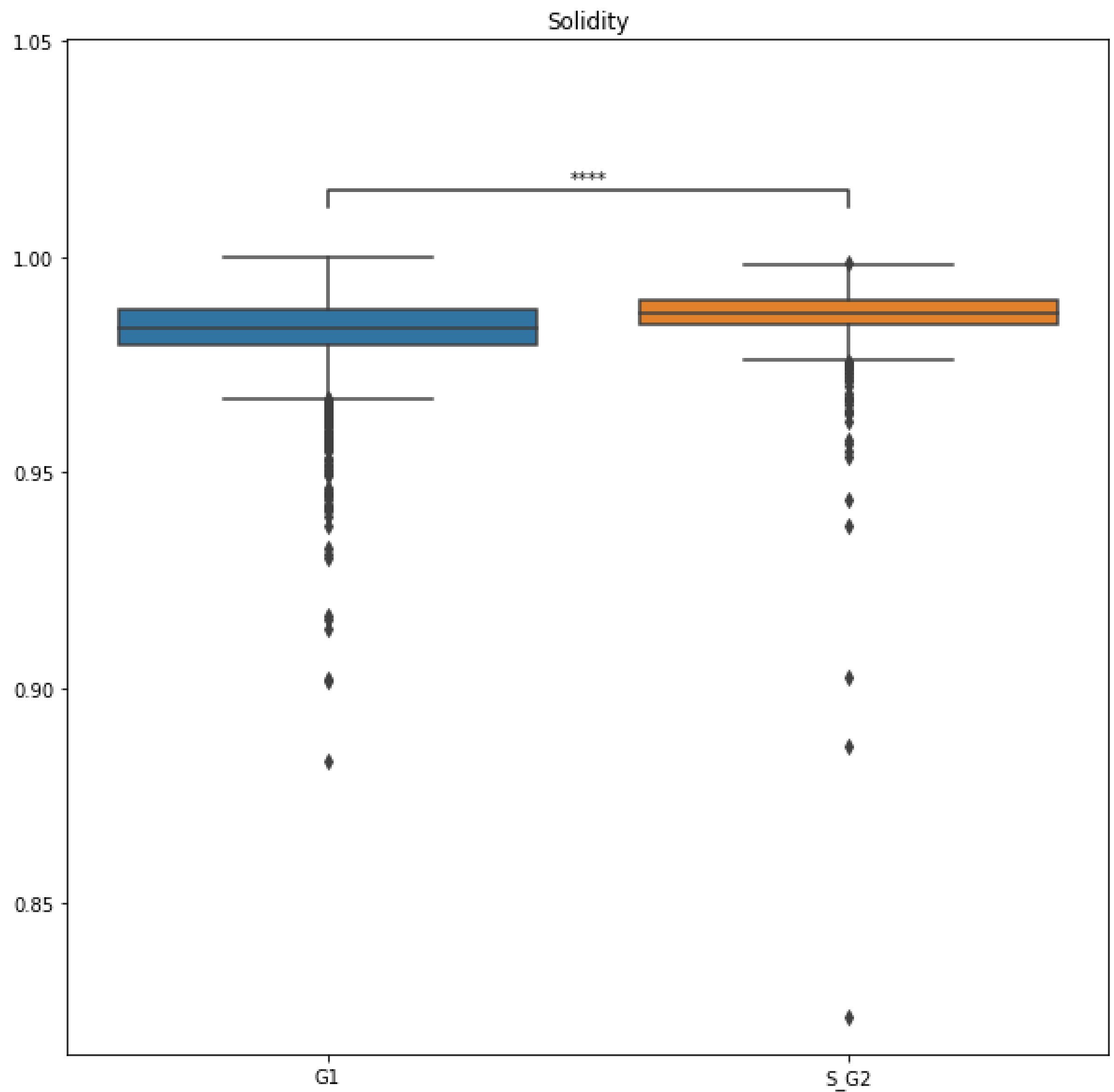
```
No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=3.616e-189 U_stat=5.872e+05

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=2.677e-42 U_stat=1.047e+06

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=3.206e-61 U_stat=9.626e+05

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=1.678e-39 U_stat=1.830e+06
p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=8.065e-147 U_stat=6.905e+05

No handles with labels found to put in legend.
No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=6.088e-181 U_stat=6.063e+05

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=4.830e-107 U_stat=8.025e+05

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=7.206e-103 U_stat=8.154e+05

No handles with labels found to put in legend.
```

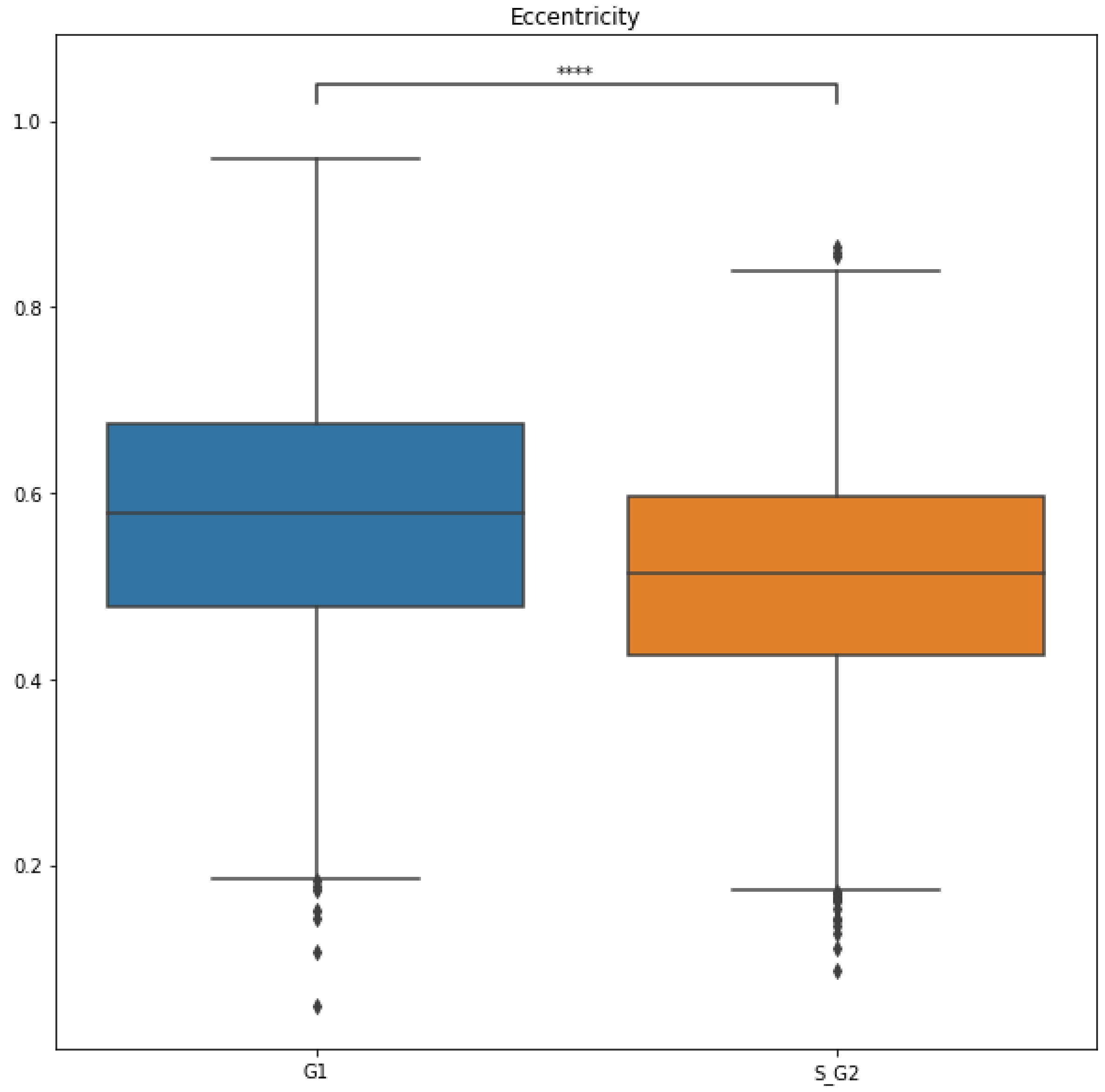
```
p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=7.624e-127 U_stat=2.147e+06

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=6.816e-141 U_stat=2.185e+06
p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=2.443e-54 U_stat=9.914e+05

No handles with labels found to put in legend.
No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=7.831e-244 U_stat=4.698e+05

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=3.661e-231 U_stat=4.957e+05

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=3.438e-41 U_stat=1.839e+06

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=9.488e-44 U_stat=1.852e+06

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=1.855e-06 U_stat=1.585e+06

No handles with labels found to put in legend.
```

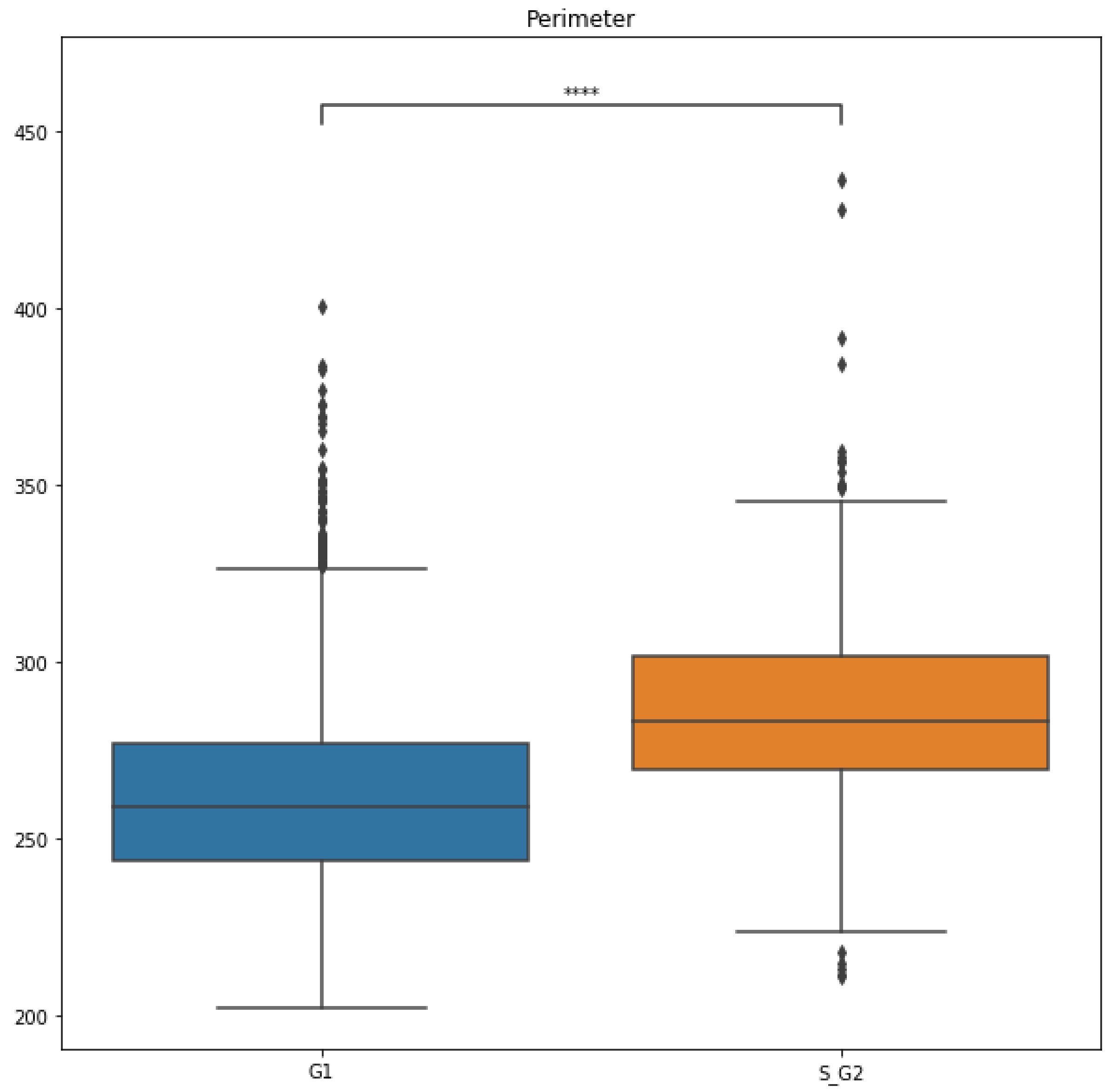
```
p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=3.172e-12 U_stat=1.650e+06

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=6.088e-181 U_stat=6.063e+05

No handles with labels found to put in legend.

p-value annotation legend:
ns: 5.00e-02 < p <= 1.00e+00
*: 1.00e-02 < p <= 5.00e-02
**: 1.00e-03 < p <= 1.00e-02
***: 1.00e-04 < p <= 1.00e-03
****: p <= 1.00e-04

G1 v.s. S_G2: Mann-Whitney-Wilcoxon test two-sided with Bonferroni correction, P_val=5.068e-03 U_stat=1.528e+06
```

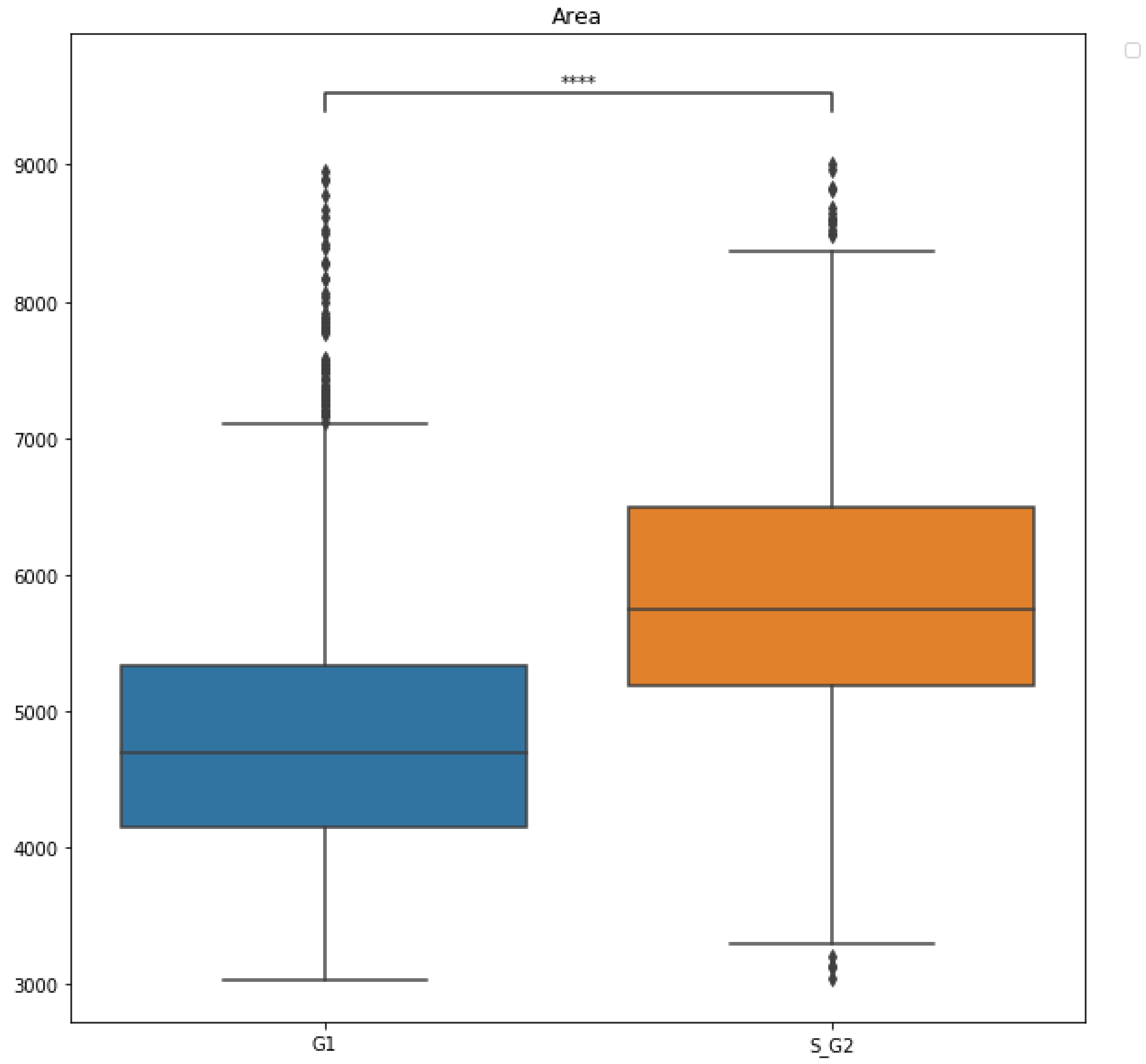


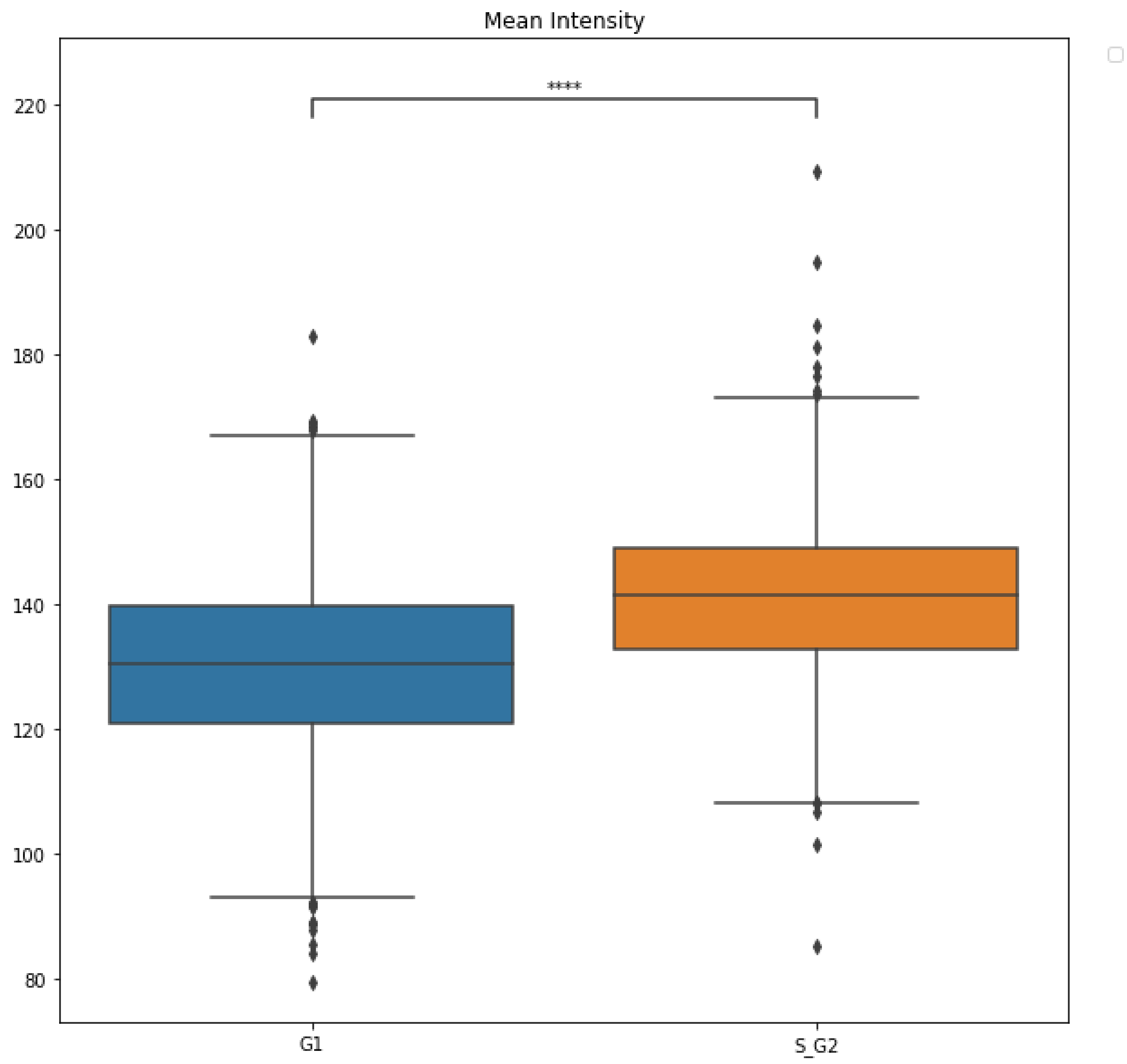

Circularity
****
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
G1
S_G2

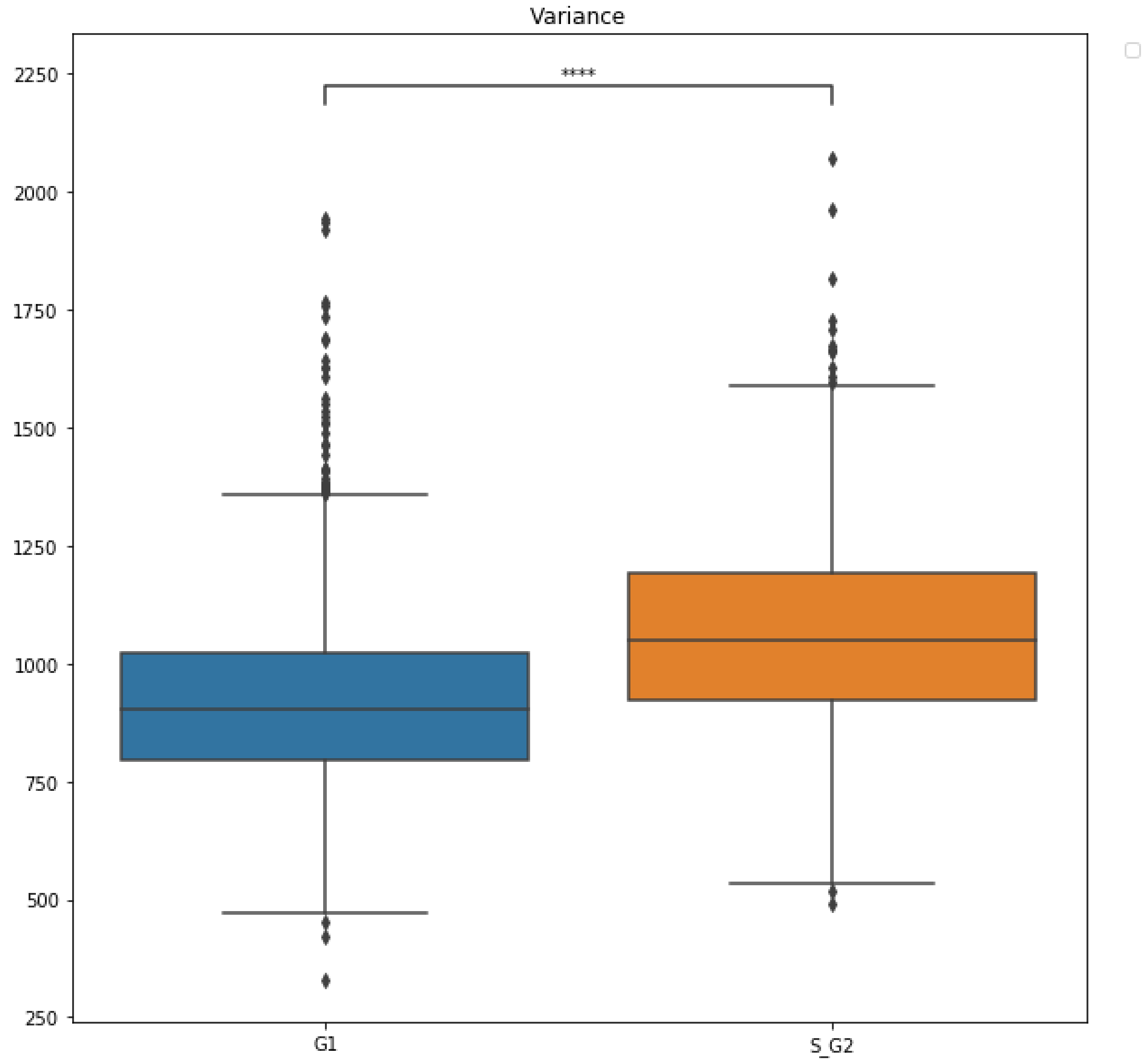

Solidity
****
1.05
1.00
0.95
0.90
0.85
G1
S_G2

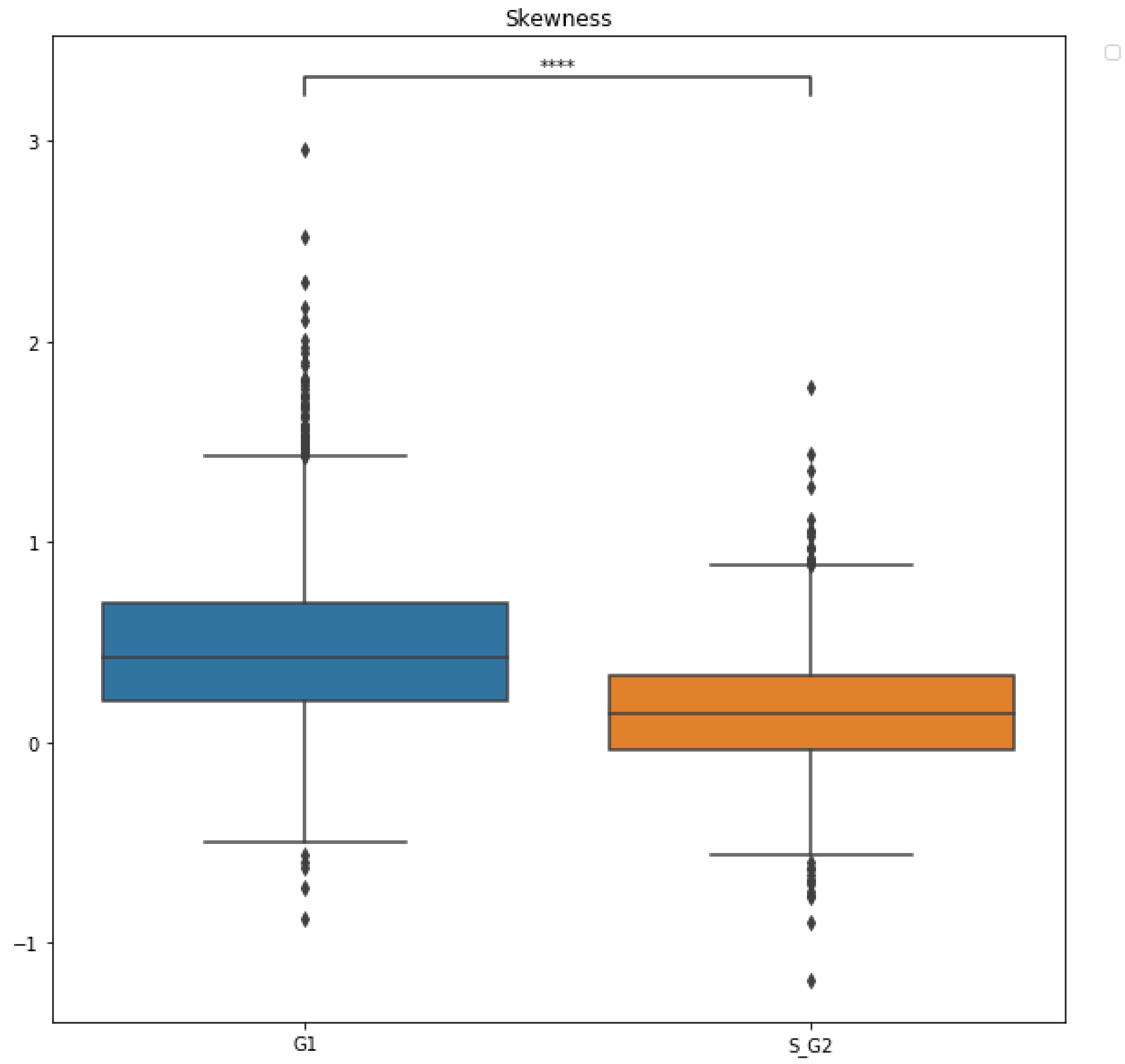

Eccentricity
****
1.0
0.8
0.6
0.4
0.2
G1
S_G2

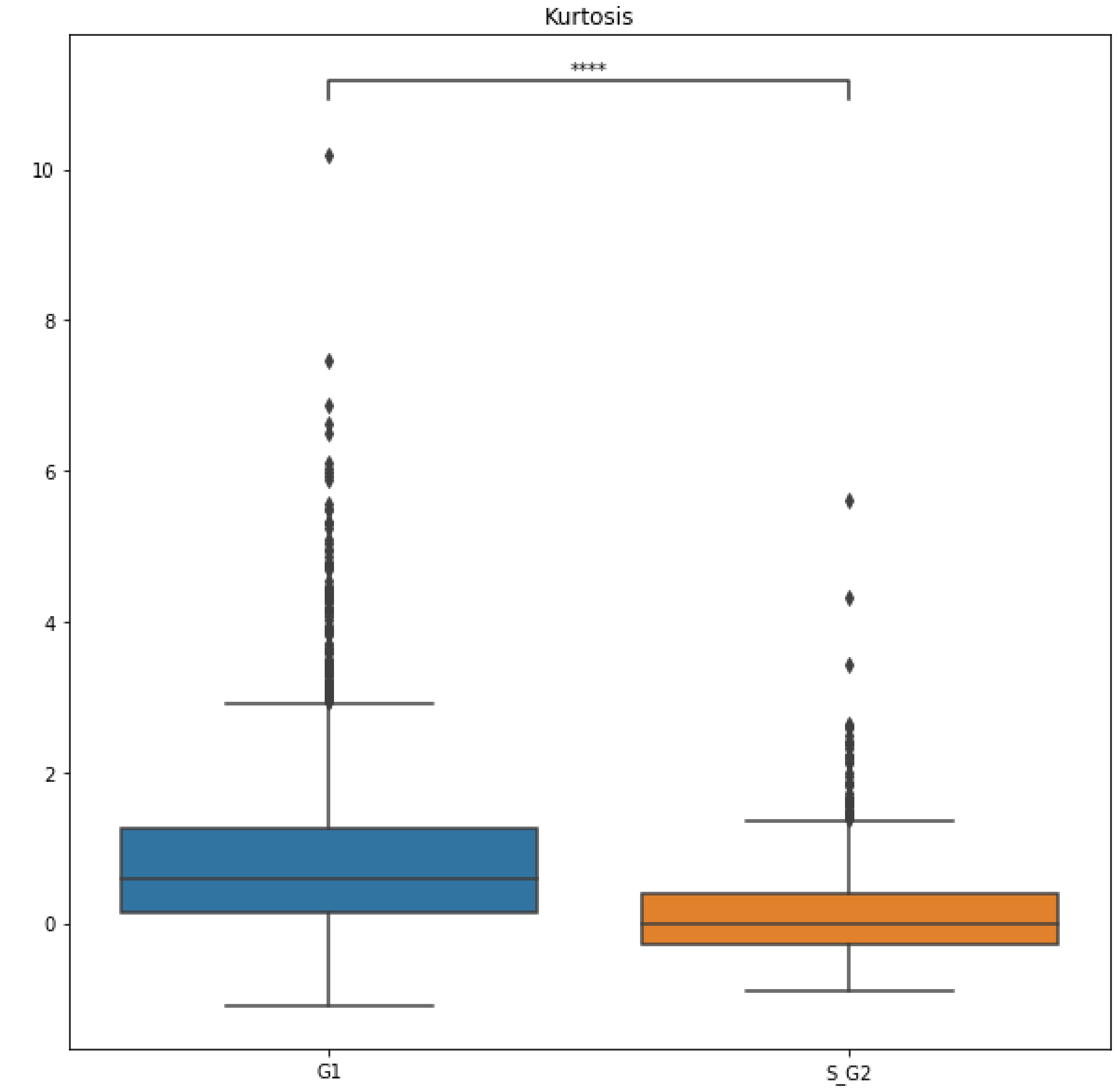

Perimeter
****
450
400
350
300
250
200
G1
S_G2

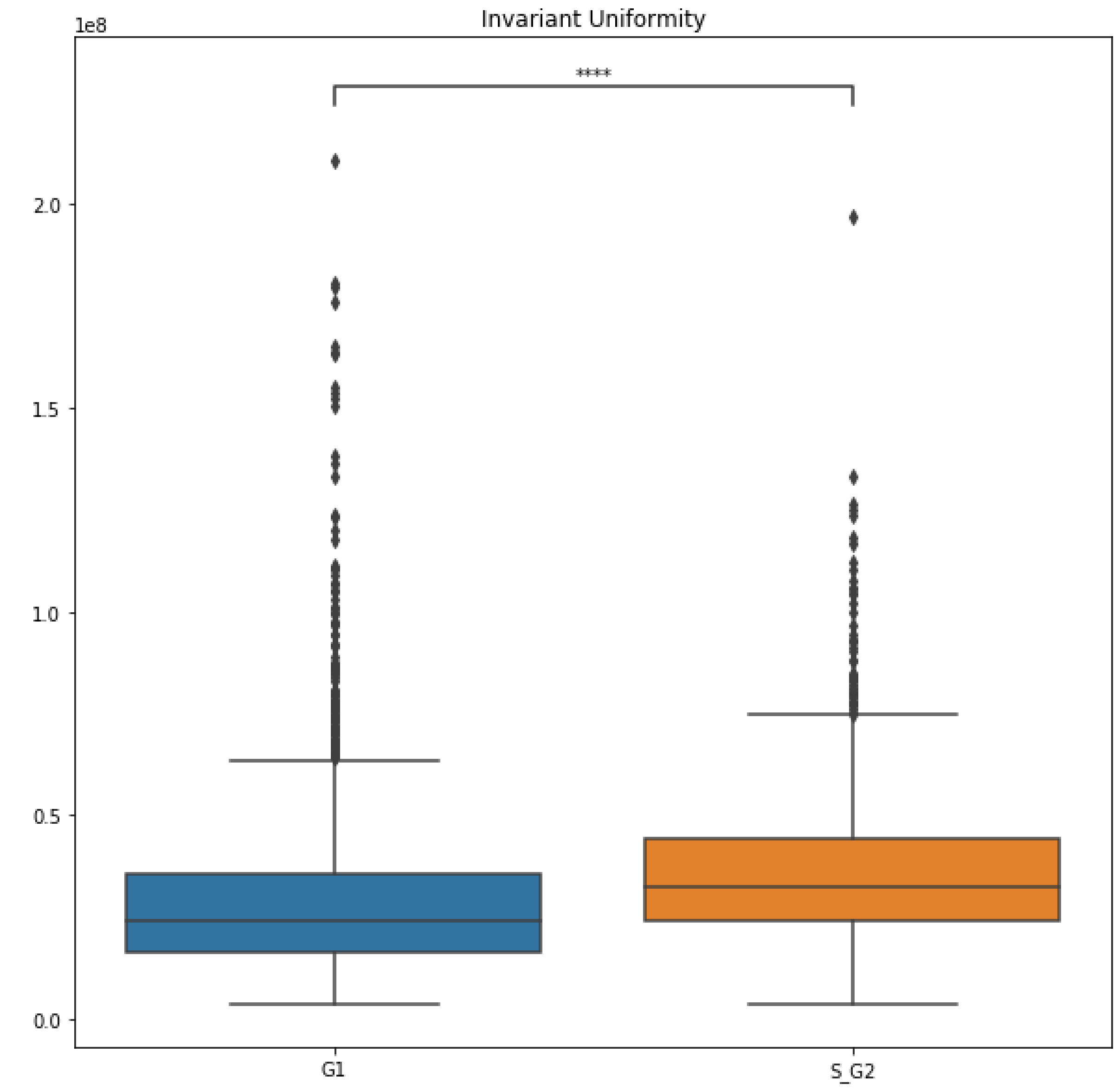

Area
****
9000
8000
7000
6000
5000
4000
3000
G1
S_G2

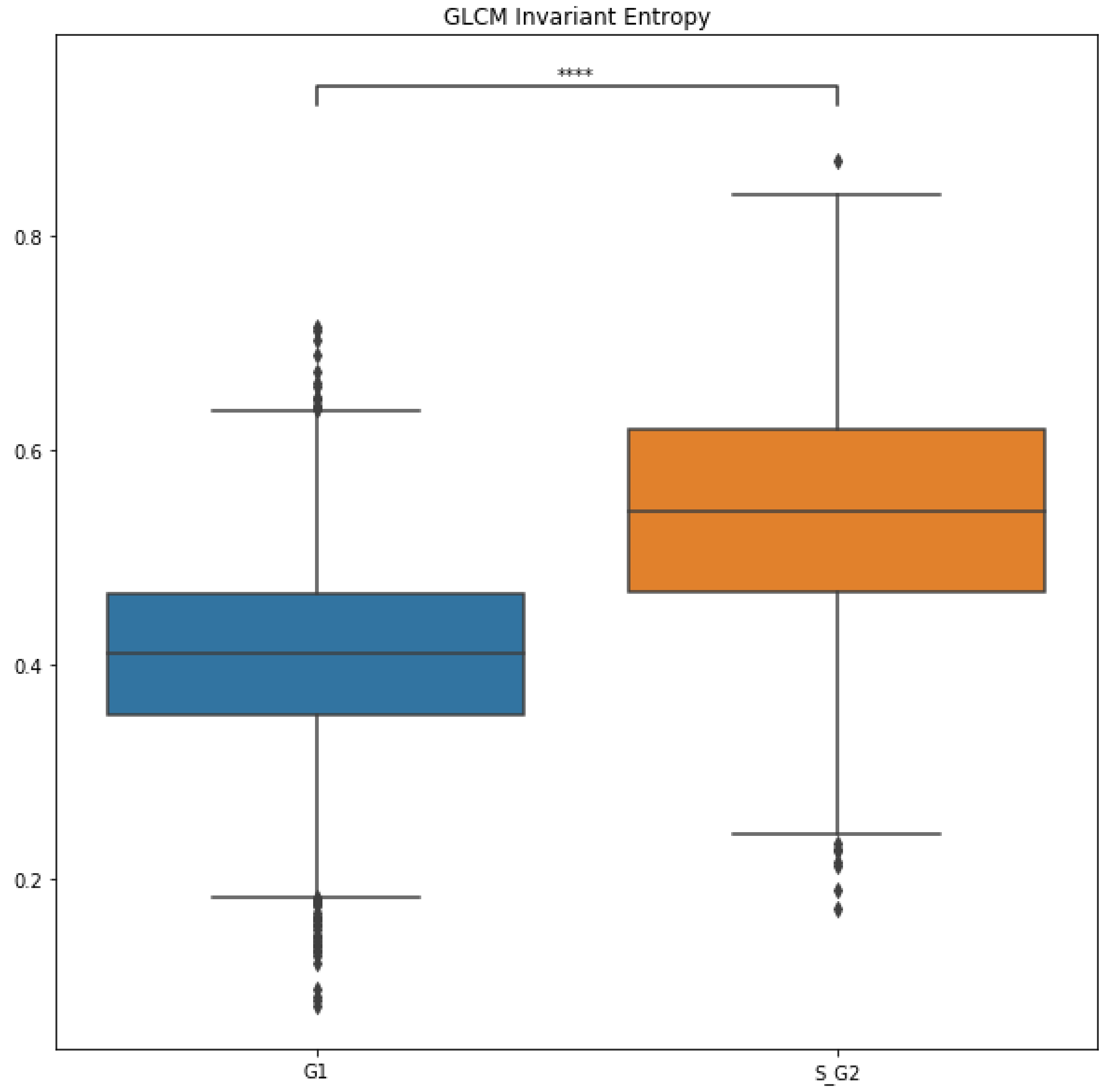

Mean Intensity
****
220
200
180
160
140
120
100
80
G1
S_G2

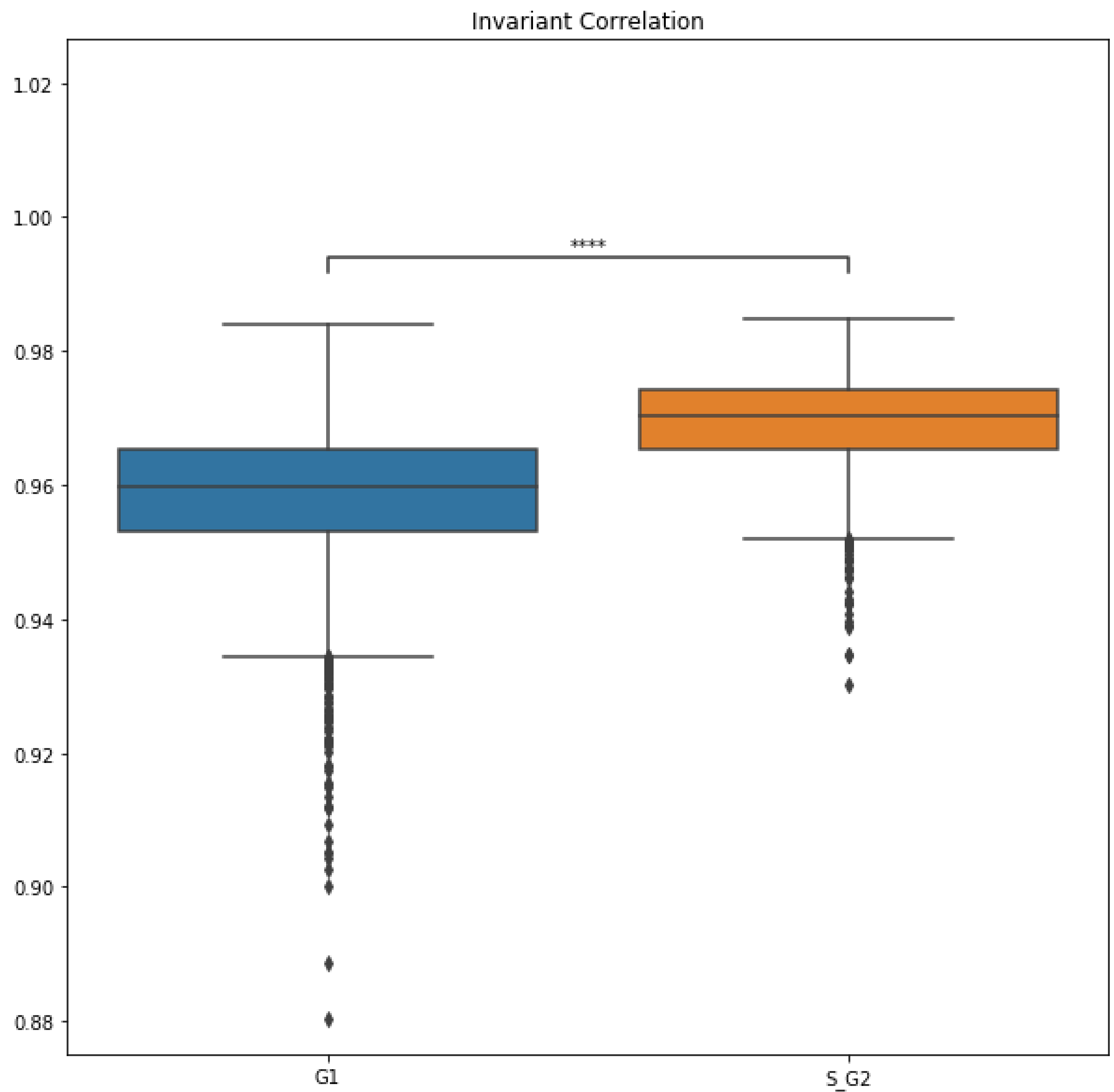

Variance
****
2250
2000
1750
1500
1250
1000
750
500
250
G1
S_G2

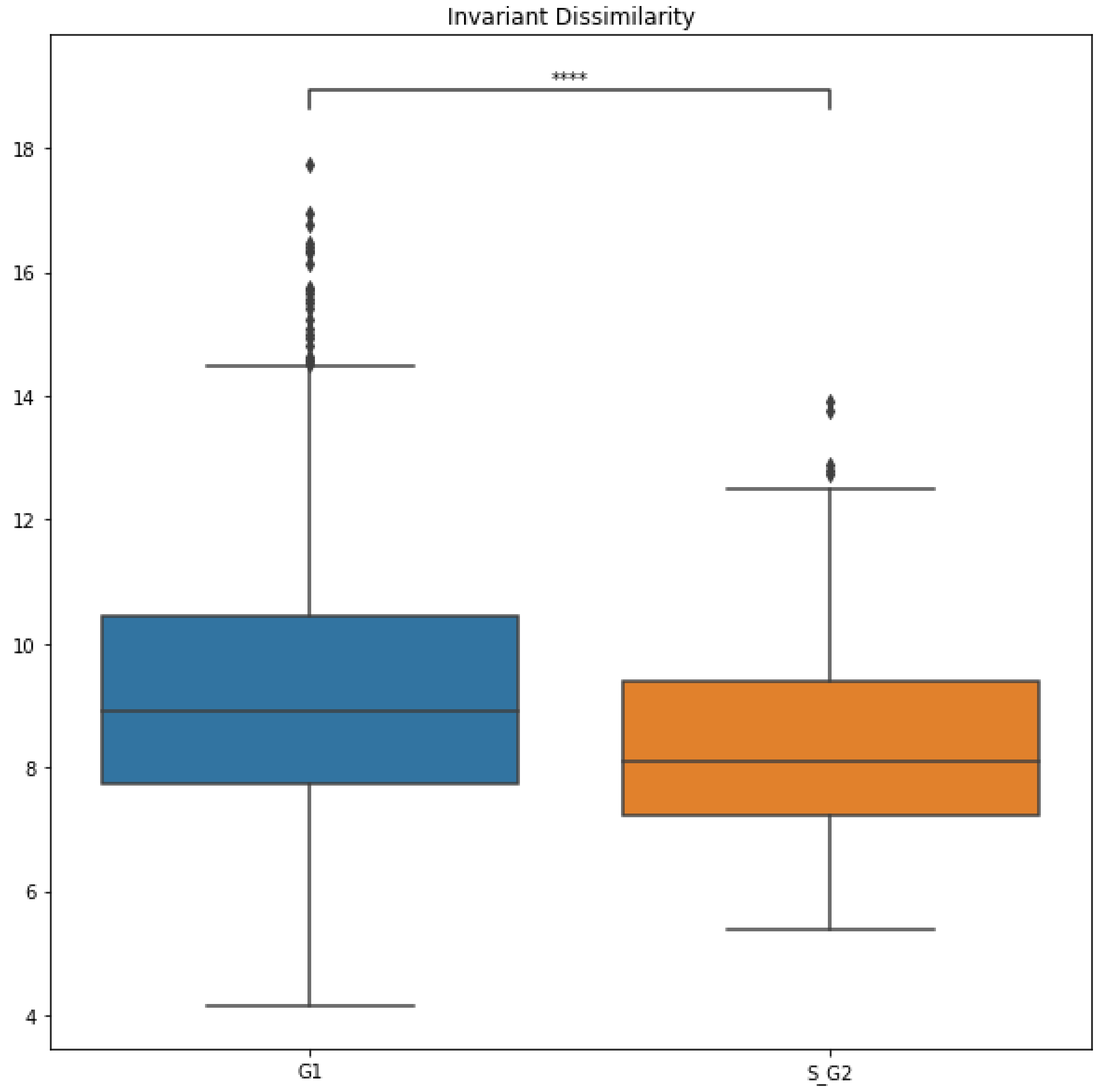

Skewness
****
3
2
1
0
−1
G1
S_G2

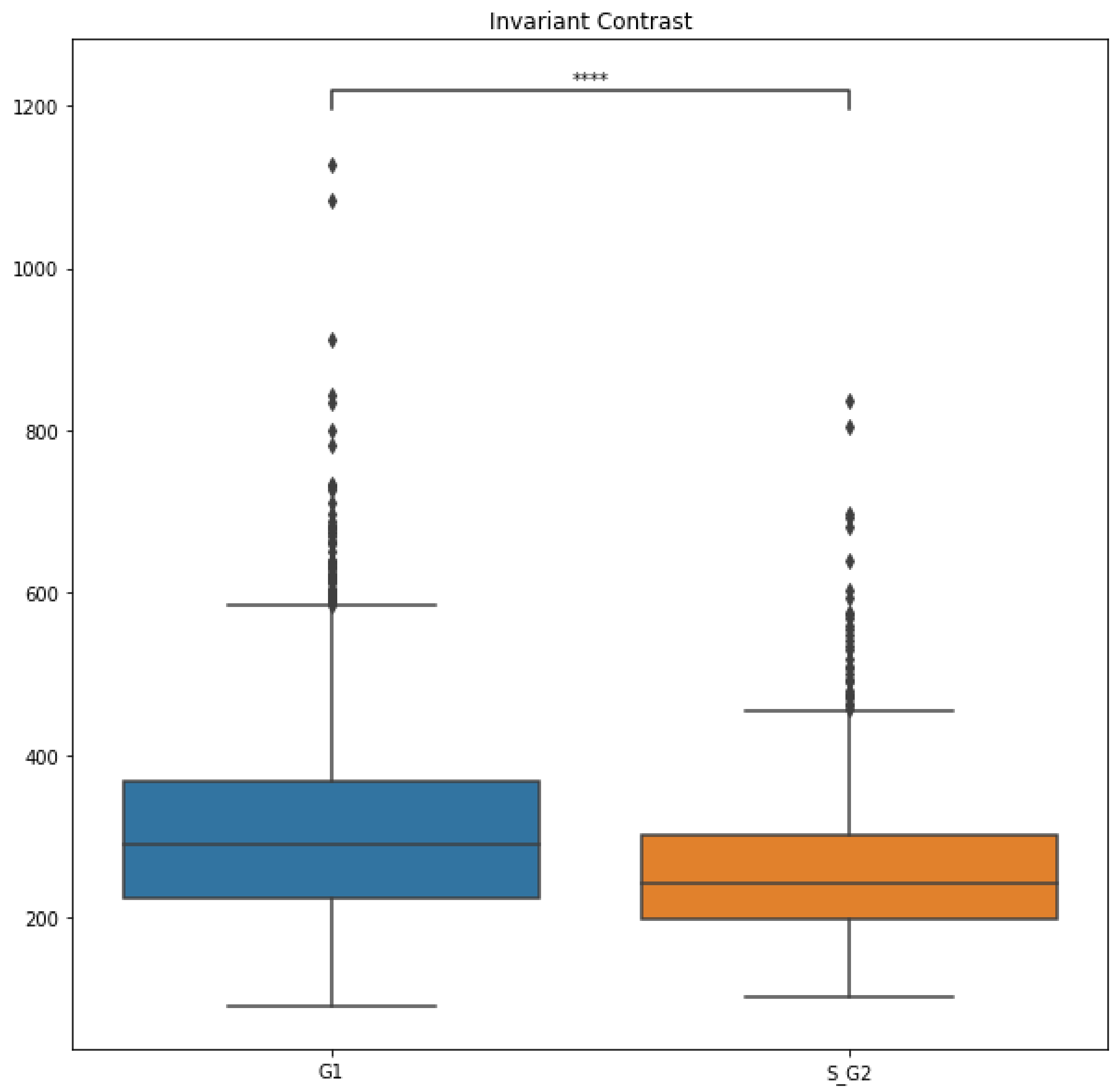

Kurtosis
****
10
8
6
4
2
0
G1
S_G2

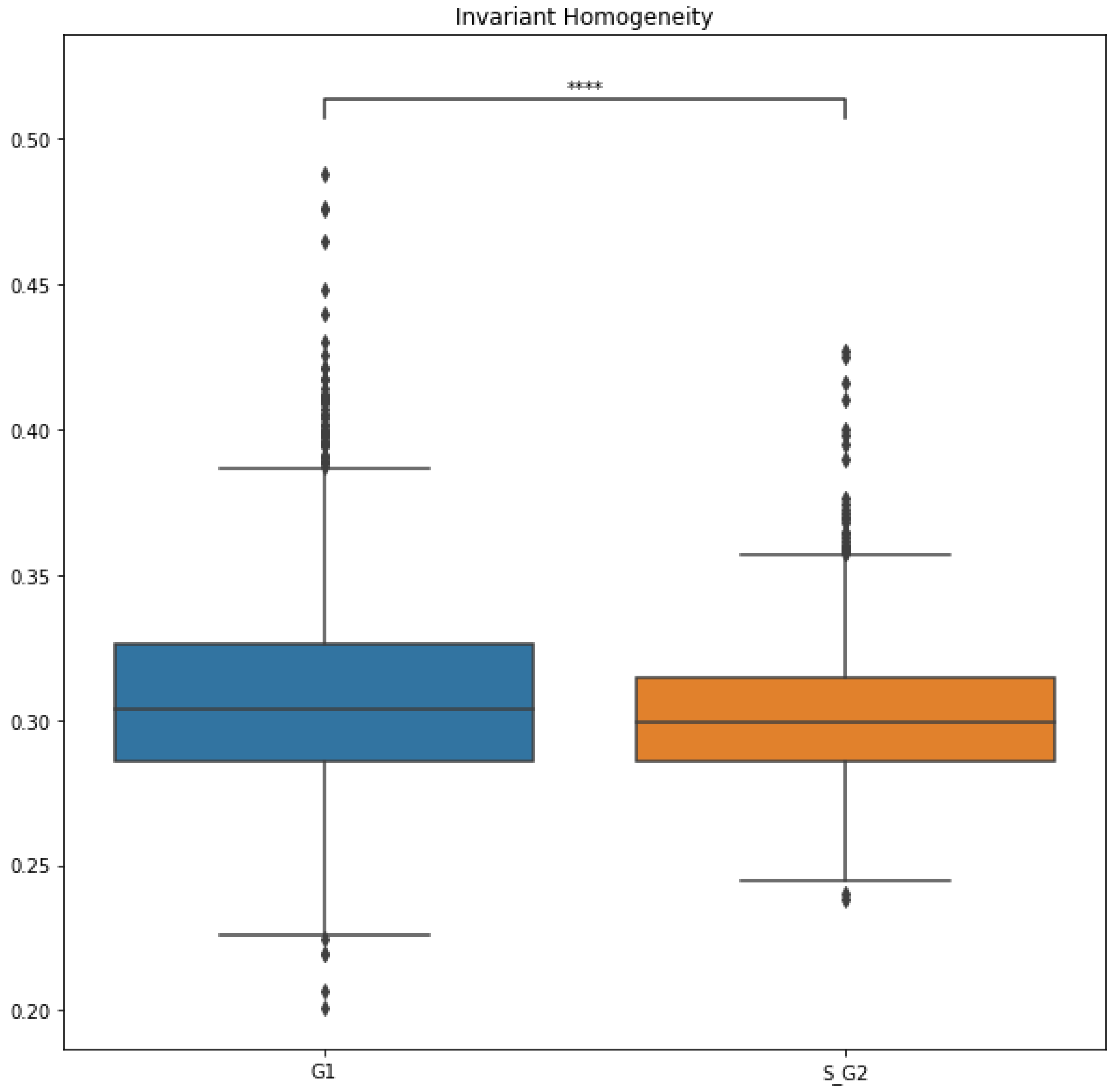

Invariant Uniformity
1e8
****
2.0
1.5
1.0
0.5
0.0
G1
S_G2

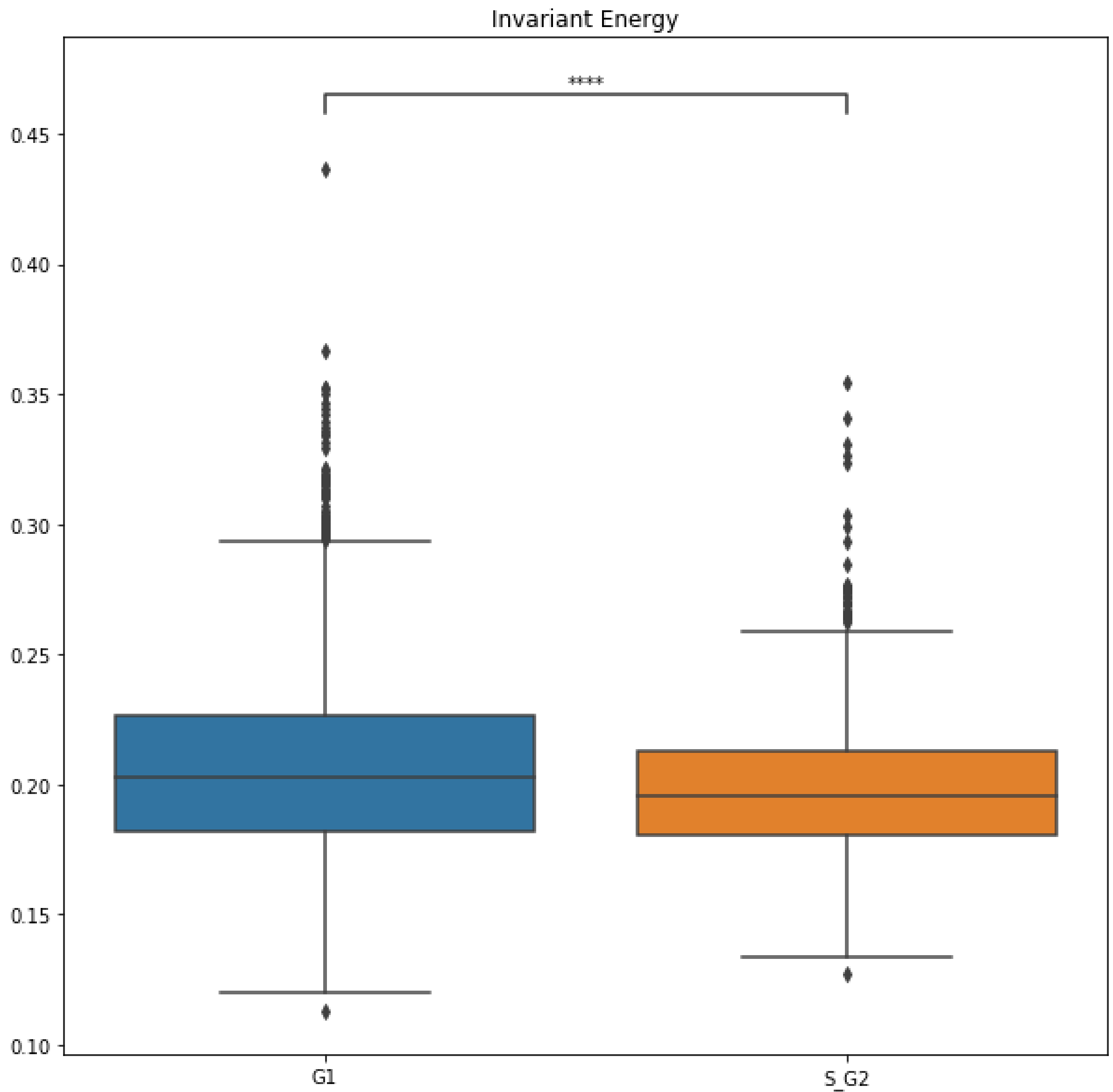

GLCM Invariant Entropy
****
0.8
0.6
0.4
0.2
G1
S_G2

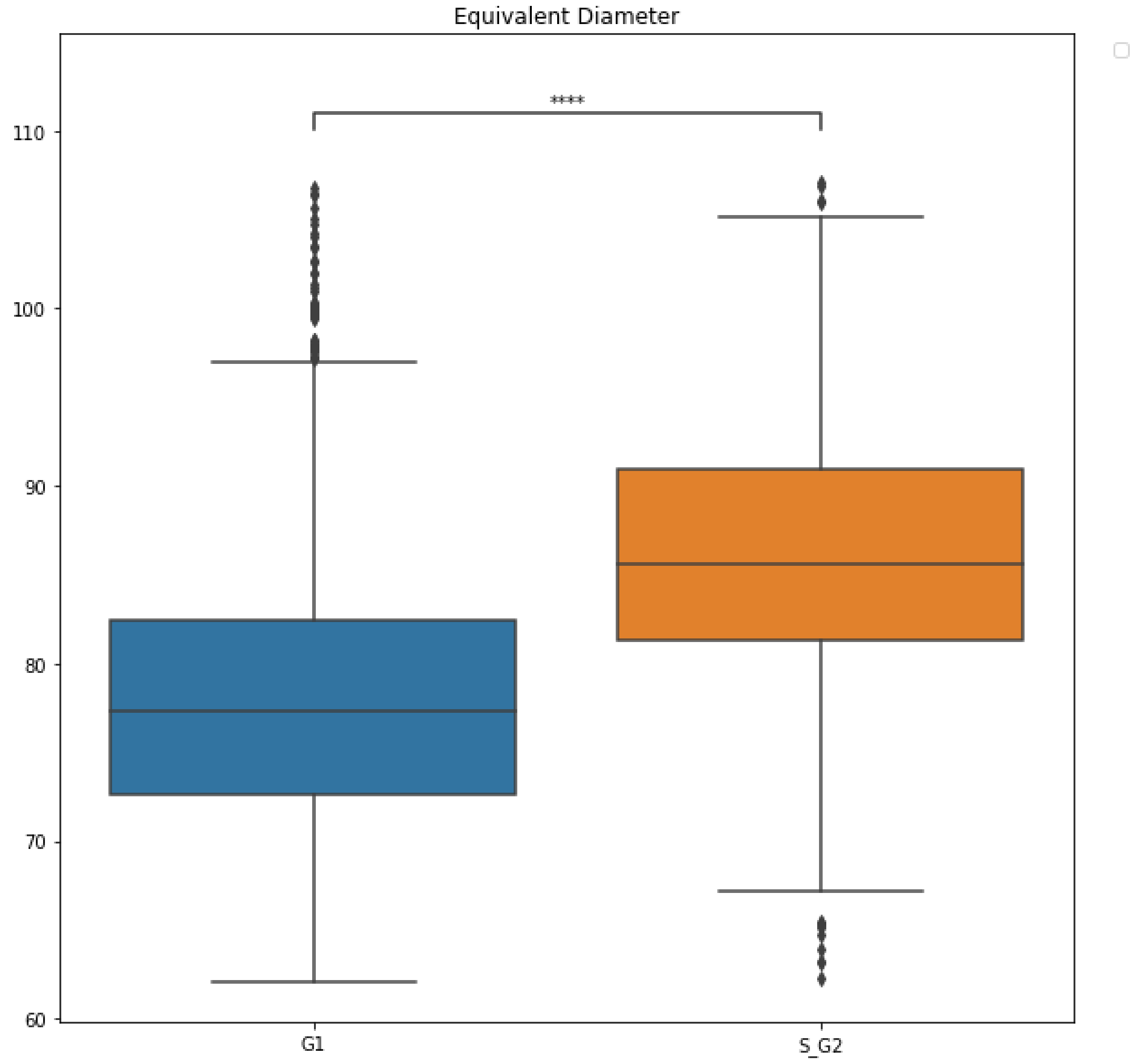

Invariant Correlation
****
1.02
1.00
0.98
0.96
0.94
0.92
0.90
0.88
G1
S_G2

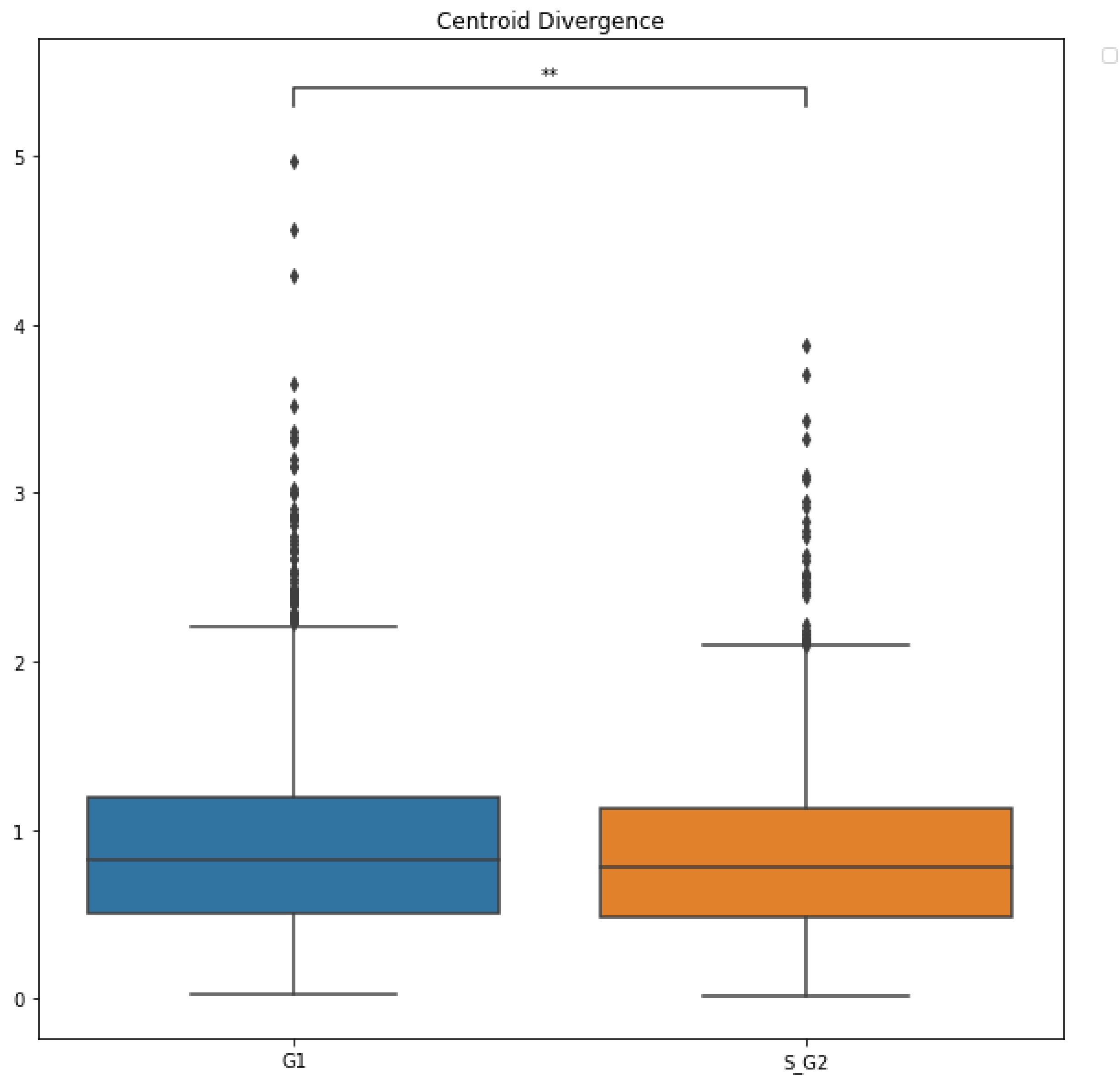

Invariant Dissimilarity
****
18
16
14
12
10
8
6
4
G1
S_G2


Invariant Contrast
****
1200
1000
800
600
400
200
G1
S_G2


Invariant Homogeneity
****
0.50
0.45
0.40
0.35
0.30
0.25
0.20
G1
S_G2


Invariant Energy
****
0.45
0.40
0.35
0.30
0.25
0.20
0.15
0.10
G1
S_G2


Equivalent Diameter
****
110
100
90
80
70
60
G1
S_G2

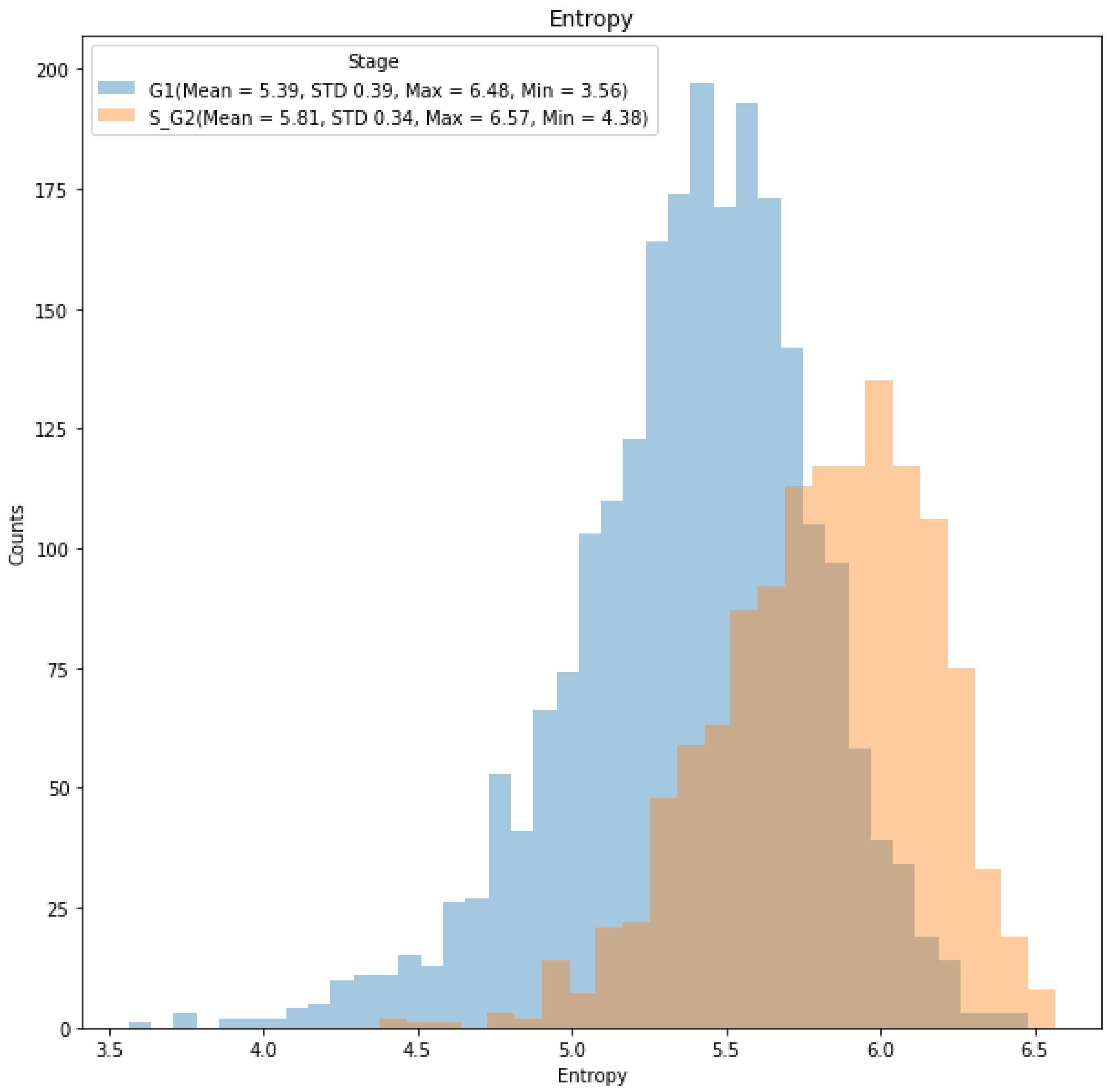

Centroid Divergence
**
5
4
3
2
1
0
G1
S_G2

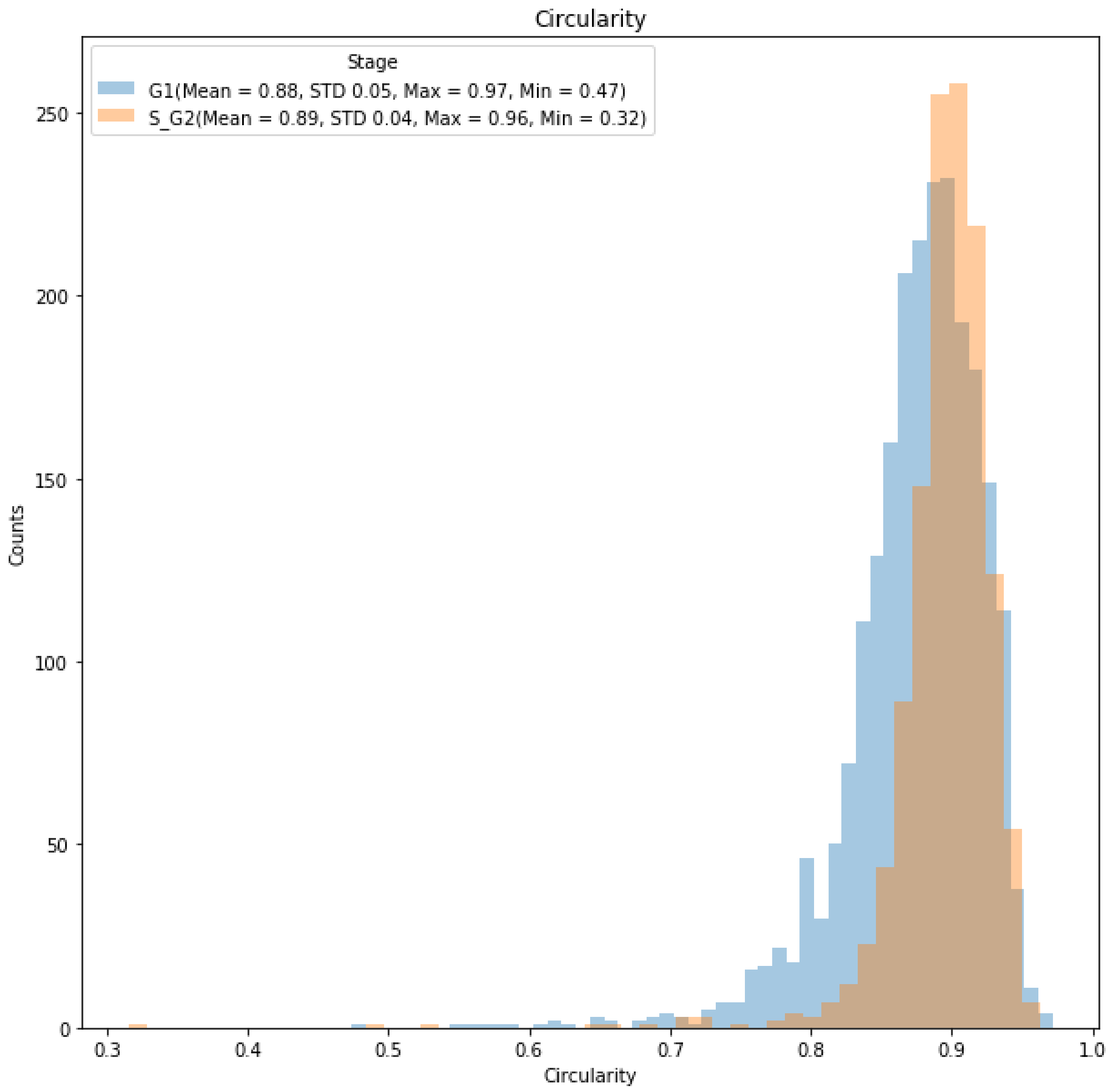
In [54]:
```
#t-test for same data

def plt_hist_and_stats_ttest(columns, types, histogram = True, equal_var = True):
    for column in columns:
        #create a new figure
        plt.figure()
        for subtype in types:

            tp = eval(subtype)
            #subset to the type

            #compute some statistics
            aux = tp[column].describe()

            #Draw the density plot
            sns.distplot(tp[column], hist = histogram, kde = False,
                         label = subtype+r'(Mean = %0.2f, STD %0.2f, Max = %0.2f, Min = %0.2f)' % (aux['mean'], aux['st
d'], aux['max'], aux['min']))
            plt.legend(prop = {'size': 10}, title = 'Stage')
            plt.title(column)
            plt.xlabel(column)
            plt.ylabel('Counts')

        subset1  = eval(types[0])
        subset2 = eval(types[1])
        u_statistic, pVal = stats.ttest_ind(subset1[column], subset2[column], axis = 0, equal_var = equal_var)
        print('THE P-VALUE IS:')
        print(pVal)
        if pVal < 0.05:
            aux = '\033[1m' + ' is ' + '\033[0m'
            aux = colored(aux, 'blue')
            hipothesis =  'H1: The difference' + aux + 'statistically significant (at significance level: 0.05).'
        else:
            aux = '\033[1m' + ' is not ' + '\033[0m'
            aux = colored(aux, 'blue')
            hipothesis ='H0: The difference' + aux + 'statistically significant (at significance level: 0.05).'
        print('For feature ' + column + ' - ' + hipothesis)
        #print(u_statistic)
```

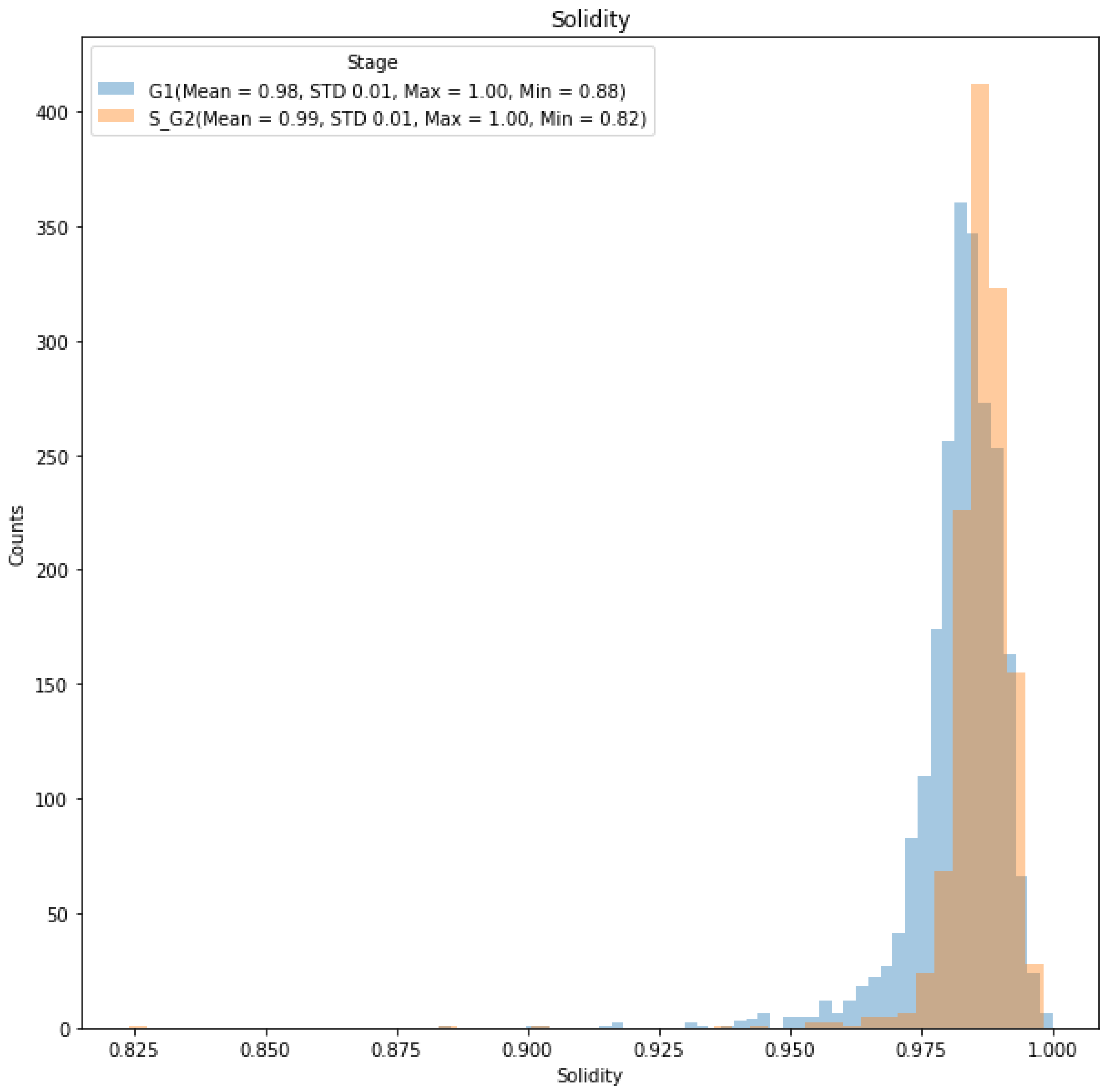
In [55]:
```
#analysing data for t-test, with and without assumption of equal variance
#also, plots of histogram counts rather than density

types = ['G1','S_G2']
columns = ['Entropy',
            'Circularity',
            'Solidity',
            'Eccentricity',
            'Perimeter',
            'Area',
            'Mean Intensity',
            'Variance',
            'Skewness',
            'Kurtosis',
            'Invariant Uniformity',
            'GLCM Invariant Entropy',
            'Invariant Correlation',
            'Invariant Dissimilarity',
            'Invariant Contrast',
            'Invariant Homogeneity',
            'Invariant Energy',
            'Equivalent Diameter',
            'Centroid Divergence']


plt.rcParams['figure.figsize'] = (10,10)
print ('---------------- EQUAL VAR TRUE ----------------')
plt_hist_and_stats_ttest(columns, types, equal_var = True)
print ('---------------- EQUAL VAR FALSE ----------------')
plt_hist_and_stats_ttest(columns, types, equal_var = False)
```

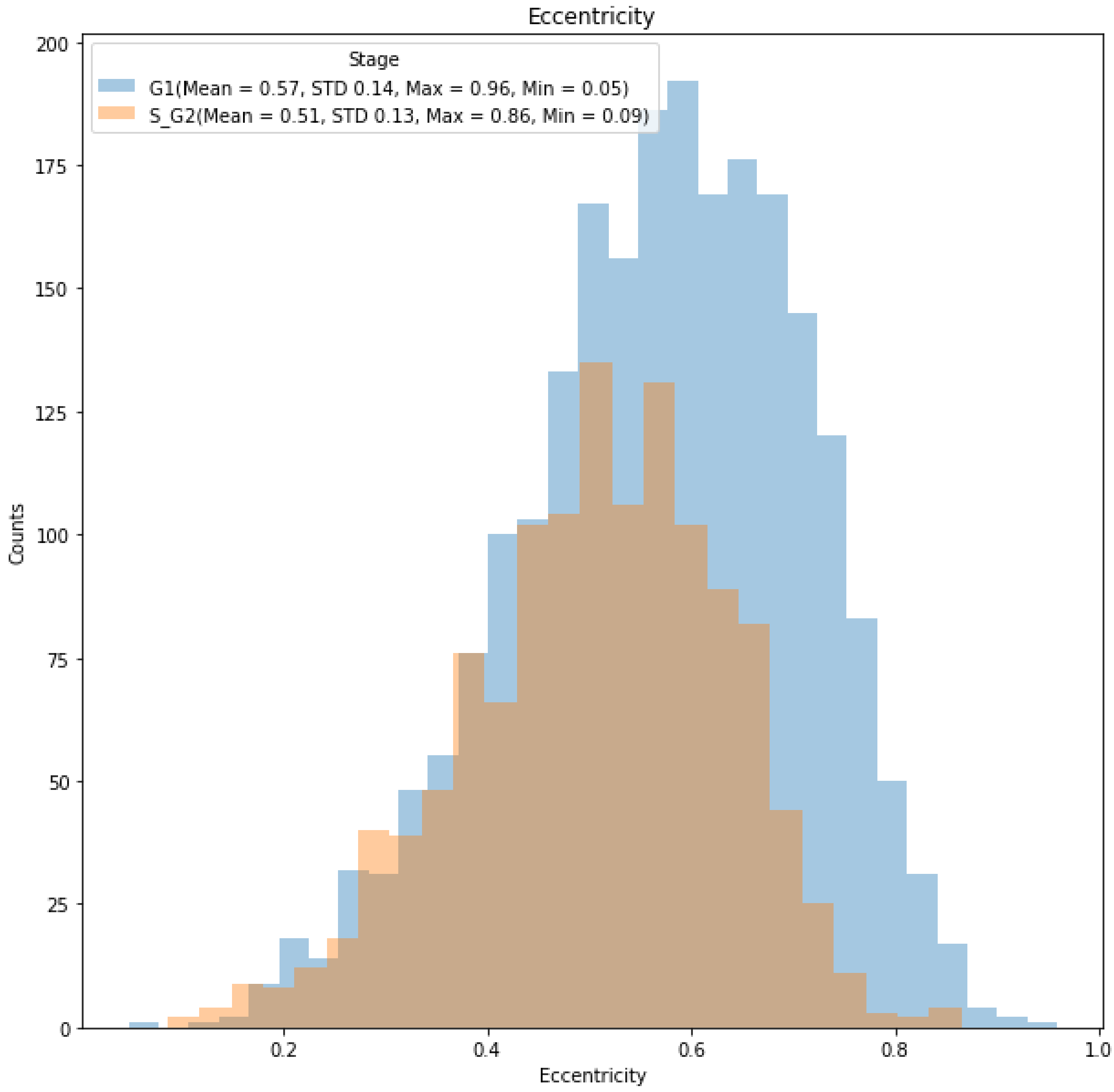
```
---------------- EQUAL VAR TRUE ----------------
THE P-VALUE IS:
1.09848854863564e-199
For feature Entropy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.159104141051883e-33
For feature Circularity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
6.407380375951467e-36
For feature Solidity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
7.592393362449388e-39
For feature Eccentricity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
2.7093929415119073e-137
For feature Perimeter - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.2893942970113184e-185
For feature Area - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.8255181581966345e-114
For feature Mean Intensity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.5726742723081895e-105
For feature Variance - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
6.9057134494061e-125
For feature Skewness - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
8.764991364009976e-106
For feature Kurtosis - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
2.1430633359082267e-26
For feature Invariant Uniformity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
3.4343672040729804e-296
For feature GLCM Invariant Entropy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
8.723324608205045e-205
For feature Invariant Correlation - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.4891860949336786e-45
For feature Invariant Dissimilarity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.3461631989835842e-42
For feature Invariant Contrast - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
7.505999031870302e-09
For feature Invariant Homogeneity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.873189523806329e-15
For feature Invariant Energy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
6.452891134286194e-189
For feature Equivalent Diameter - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
0.0021203159398922467
For feature Centroid Divergence - H1: The difference is statistically significant (at significance level: 0.05).
---------------- EQUAL VAR FALSE ----------------
THE P-VALUE IS:
5.562577365054583e-208
For feature Entropy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
5.71398356899168e-38
For feature Circularity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
2.362134627910443e-38
For feature Solidity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.2258521694598226e-40
For feature Eccentricity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
6.979709316783e-139
For feature Perimeter - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
9.410336559733953e-174
For feature Area - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.1821416690924502e-118
For feature Mean Intensity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
3.643996555947263e-97
For feature Variance - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
3.486280837906461e-143
For feature Skewness - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.809336197064544e-140
```

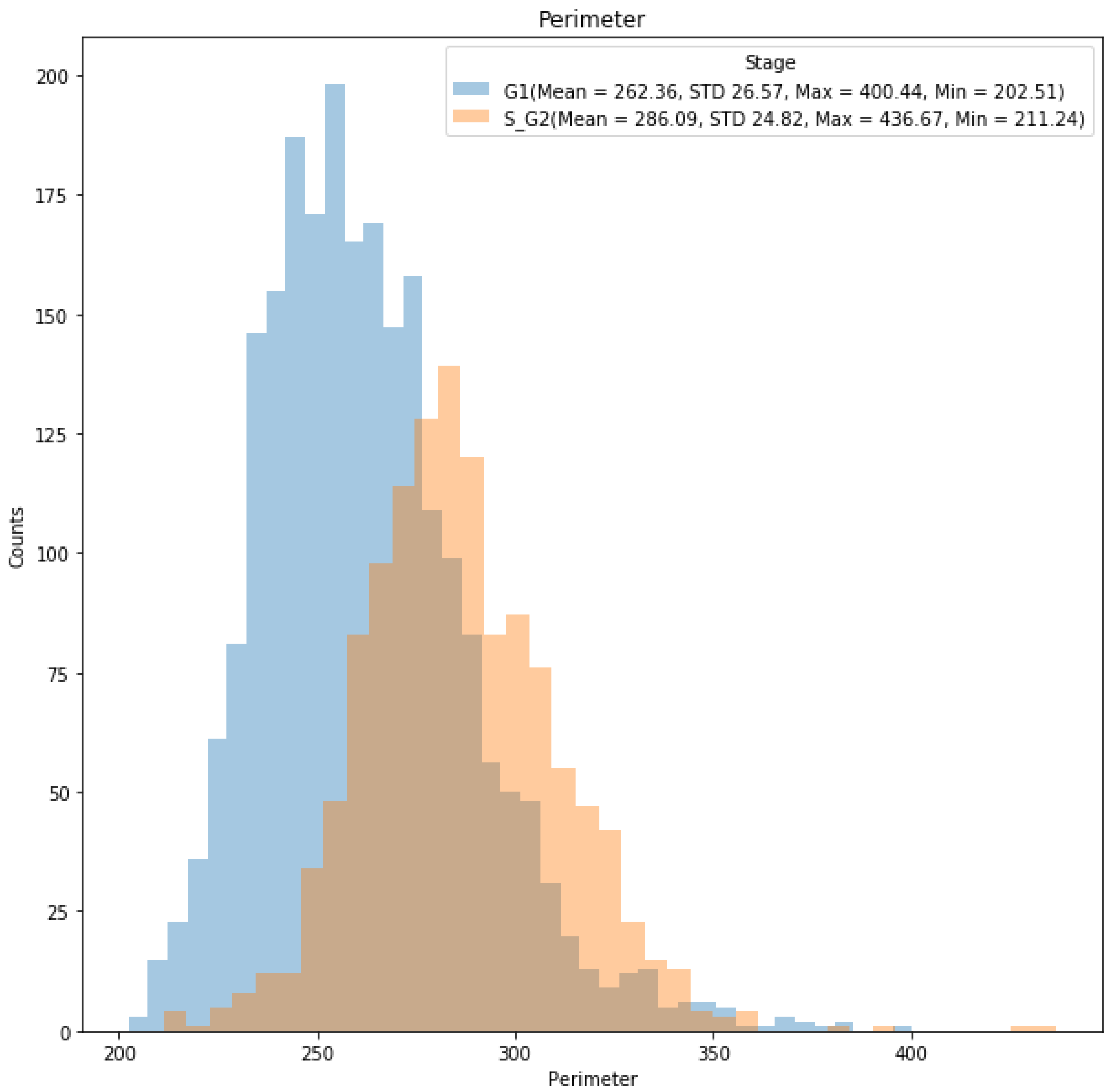
```
For feature Kurtosis - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.59253716860003e-27
For feature Invariant Uniformity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
8.573486689100333e-249
For feature GLCM Invariant Entropy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.4391833119813733e-245
For feature Invariant Correlation - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.0554710099227322e-52
For feature Invariant Dissimilarity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.1962730302653597e-49
For feature Invariant Contrast - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
2.7229064935388033e-10
For feature Invariant Homogeneity - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
4.8518144998974725e-18
For feature Invariant Energy - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
1.3935335938505521e-184
For feature Equivalent Diameter - H1: The difference is statistically significant (at significance level: 0.05).
THE P-VALUE IS:
0.0017044406727256703
For feature Centroid Divergence - H1: The difference is statistically significant (at significance level: 0.05).
```

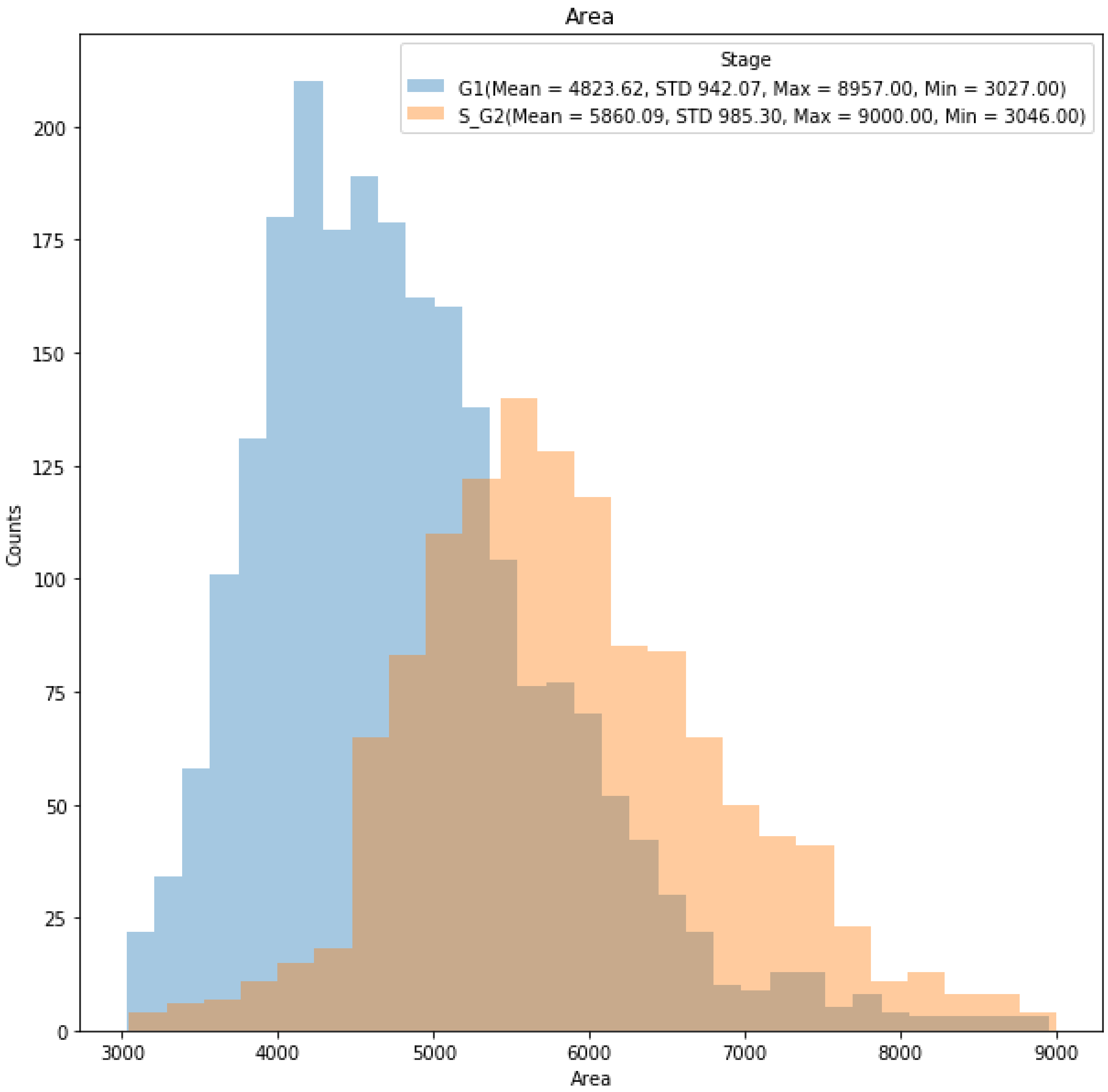


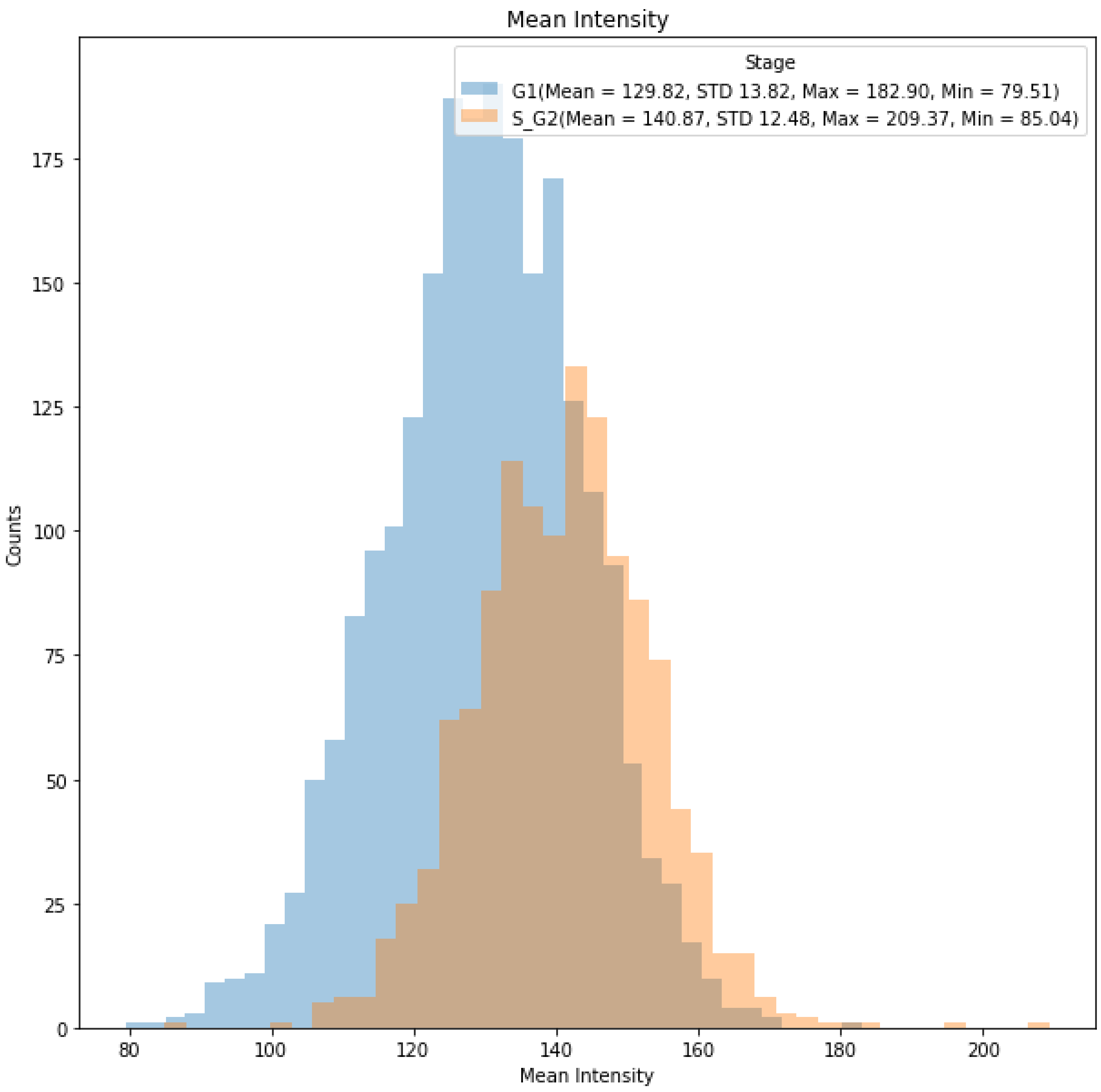

Circularity
Stage
G1(Mean = 0.88, STD 0.05, Max = 0.97, Min = 0.47)
S_G2(Mean = 0.89, STD 0.04, Max = 0.96, Min = 0.32)
Counts
Circularity

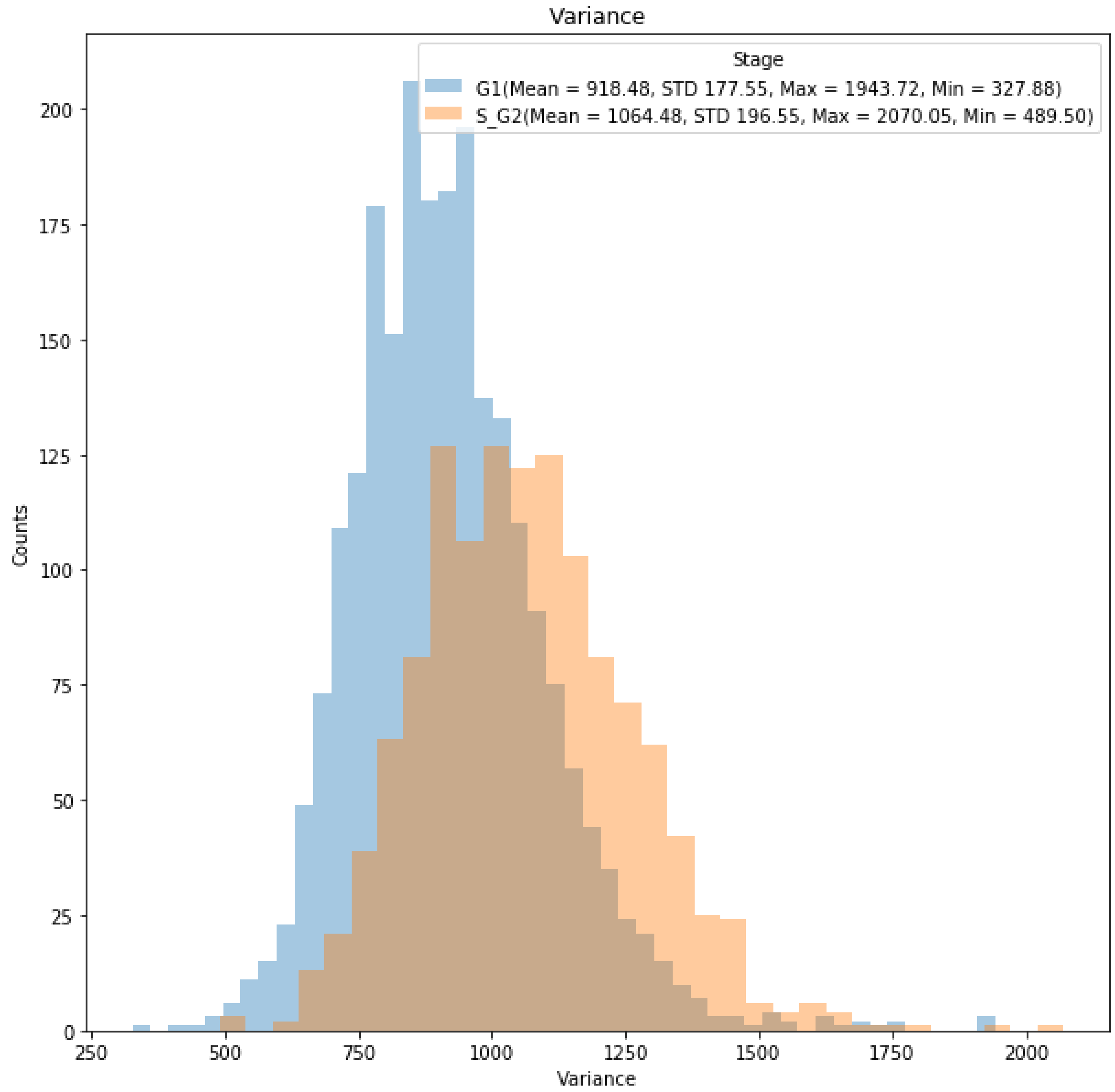

Solidity
Stage
G1(Mean = 0.98, STD 0.01, Max = 1.00, Min = 0.88)
S_G2(Mean = 0.99, STD 0.01, Max = 1.00, Min = 0.82)
Counts
Solidity

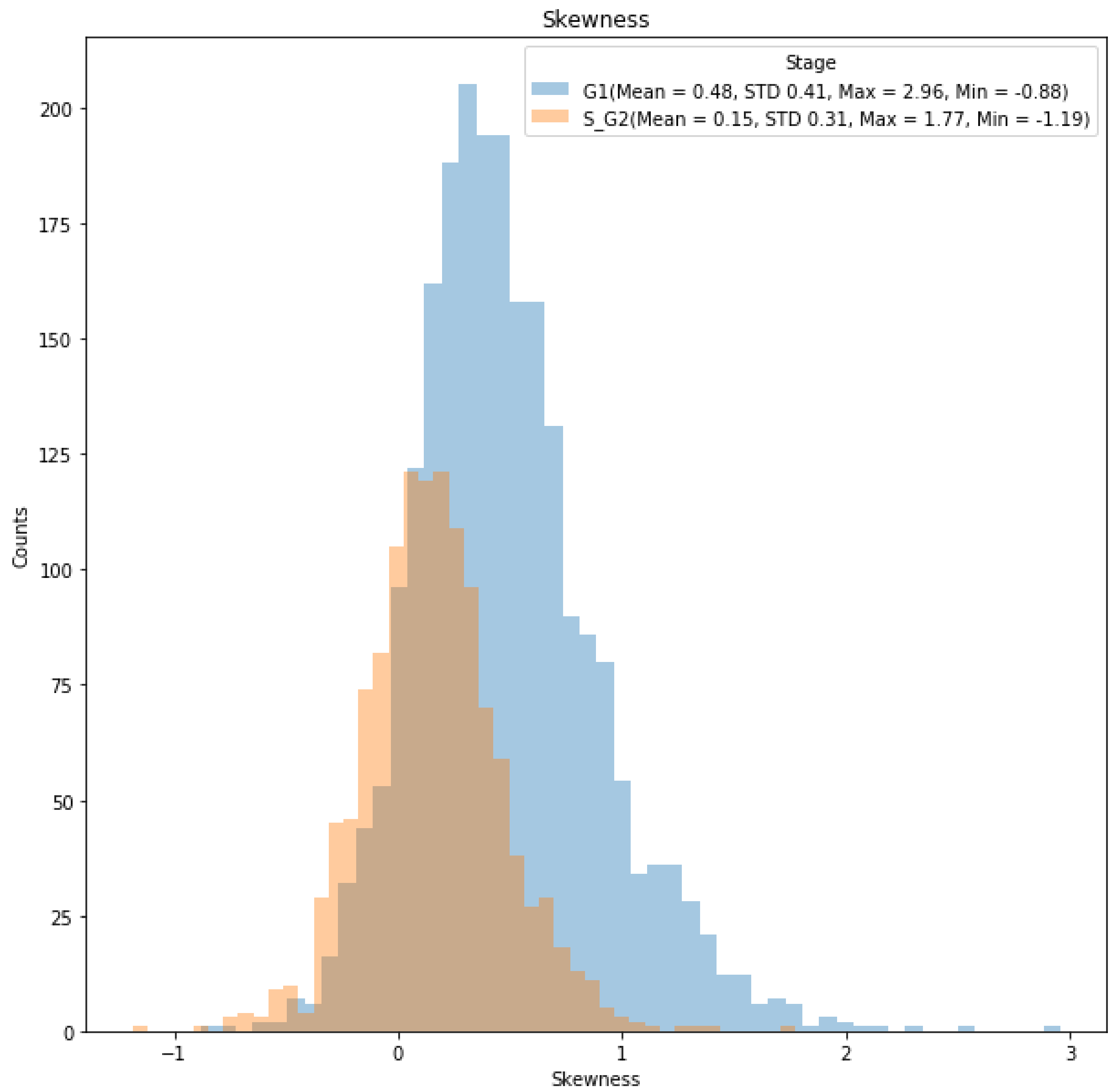

Eccentricity
Stage
G1(Mean = 0.57, STD 0.14, Max = 0.96, Min = 0.05)
S_G2(Mean = 0.51, STD 0.13, Max = 0.86, Min = 0.09)
Counts
0
25
50
75
100
125
150
175
200
0.2
0.4
0.6
0.8
1.0
Eccentricity

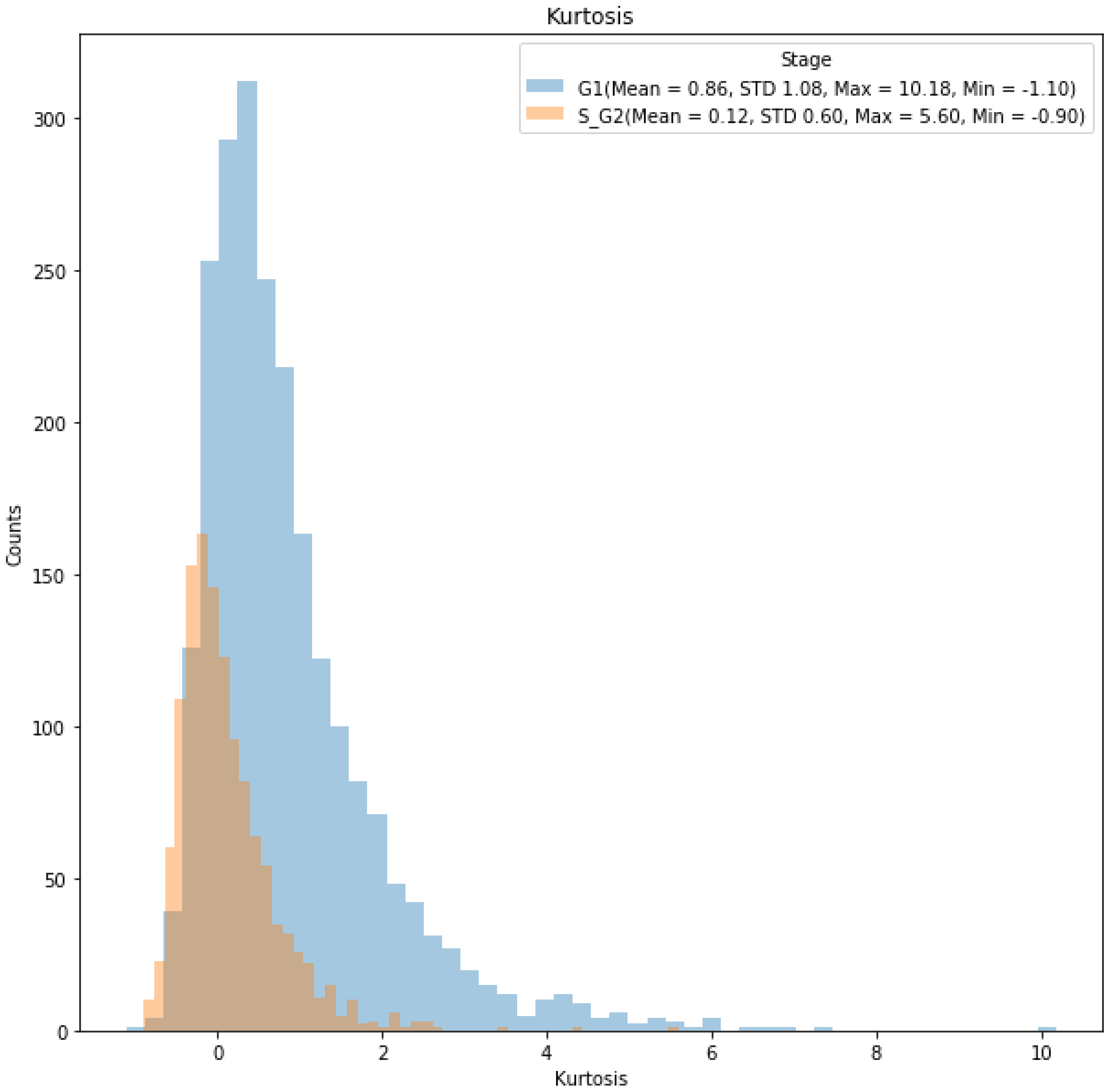

Perimeter
Stage
G1(Mean = 262.36, STD 26.57, Max = 400.44, Min = 202.51)
S_G2(Mean = 286.09, STD 24.82, Max = 436.67, Min = 211.24)
Counts
0
25
50
75
100
125
150
175
200
200
250
300
350
400
Perimeter

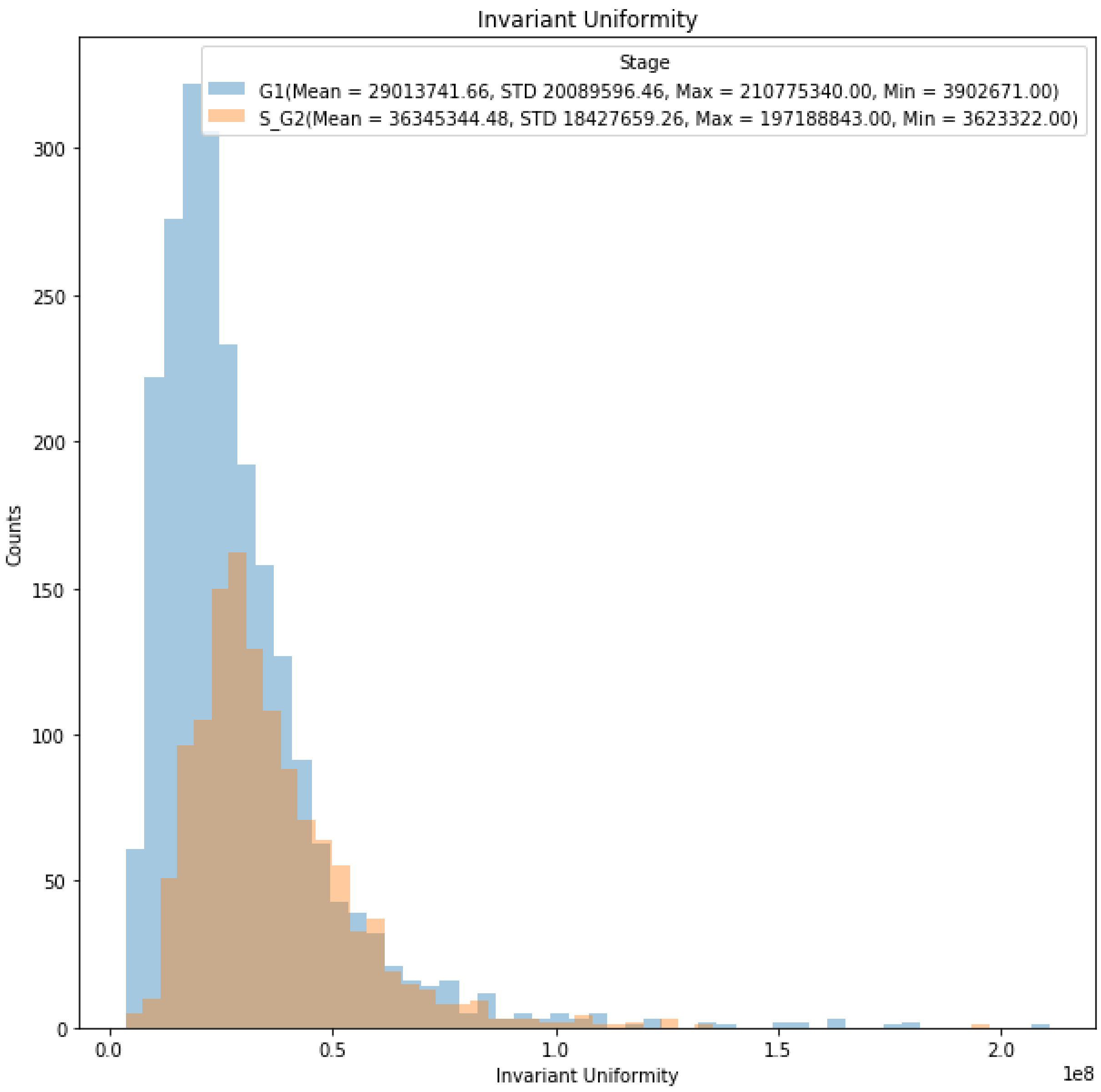

Area
Stage
G1(Mean = 4823.62, STD 942.07, Max = 8957.00, Min = 3027.00)
S_G2(Mean = 5860.09, STD 985.30, Max = 9000.00, Min = 3046.00)
Counts
0
25
50
75
100
125
150
175
200
3000
4000
5000
6000
7000
8000
9000
Area

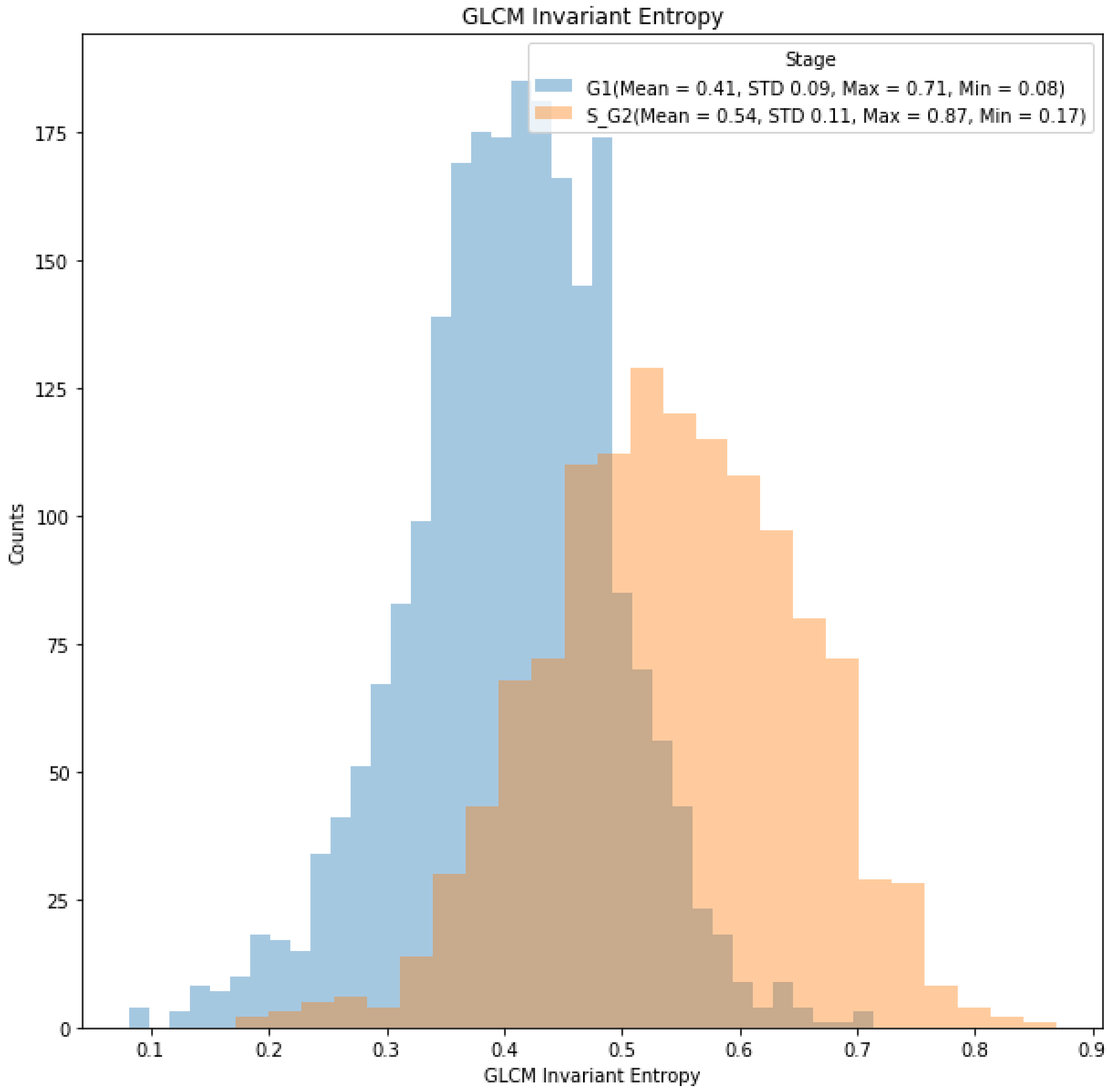

Mean Intensity
Stage
G1(Mean = 129.82, STD 13.82, Max = 182.90, Min = 79.51)
S_G2(Mean = 140.87, STD 12.48, Max = 209.37, Min = 85.04)
Counts
0
25
50
75
100
125
150
175
80
100
120
140
160
180
200
Mean Intensity

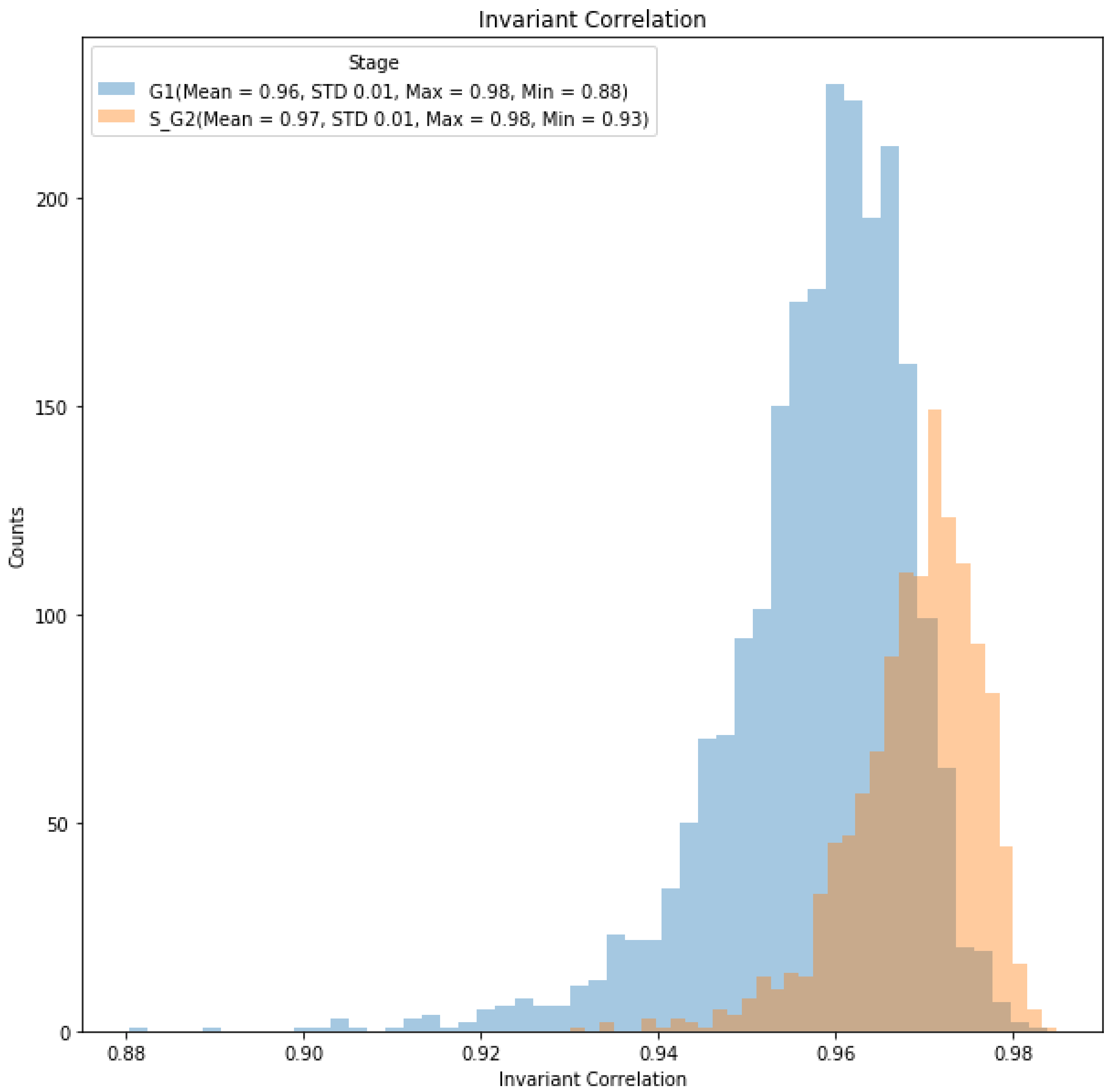

Variance
Stage
G1(Mean = 918.48, STD 177.55, Max = 1943.72, Min = 327.88)
S_G2(Mean = 1064.48, STD 196.55, Max = 2070.05, Min = 489.50)
Counts
Variance

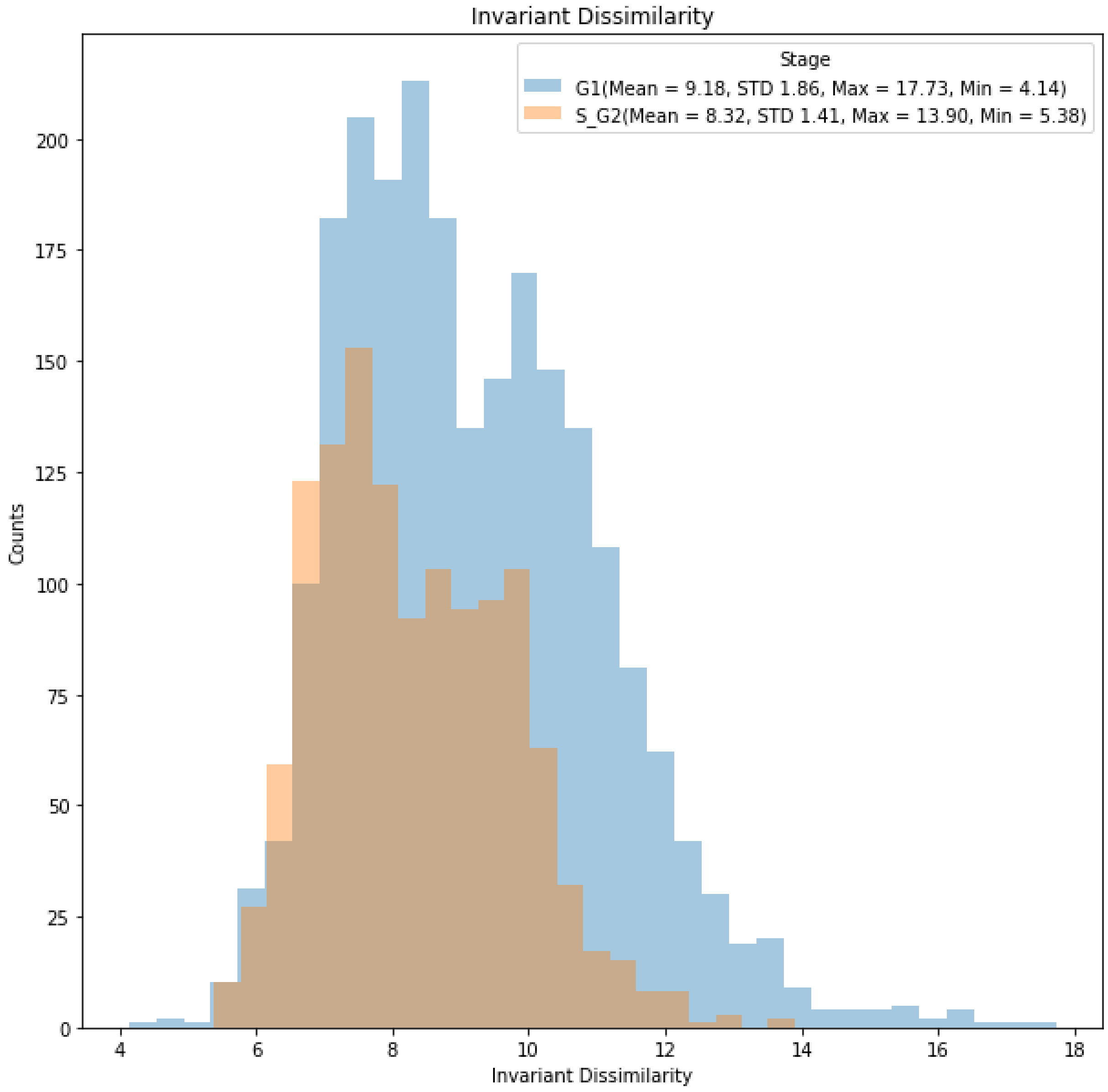

Skewness
Stage
G1(Mean = 0.48, STD 0.41, Max = 2.96, Min = -0.88)
S_G2(Mean = 0.15, STD 0.31, Max = 1.77, Min = -1.19)
Counts
Skewness

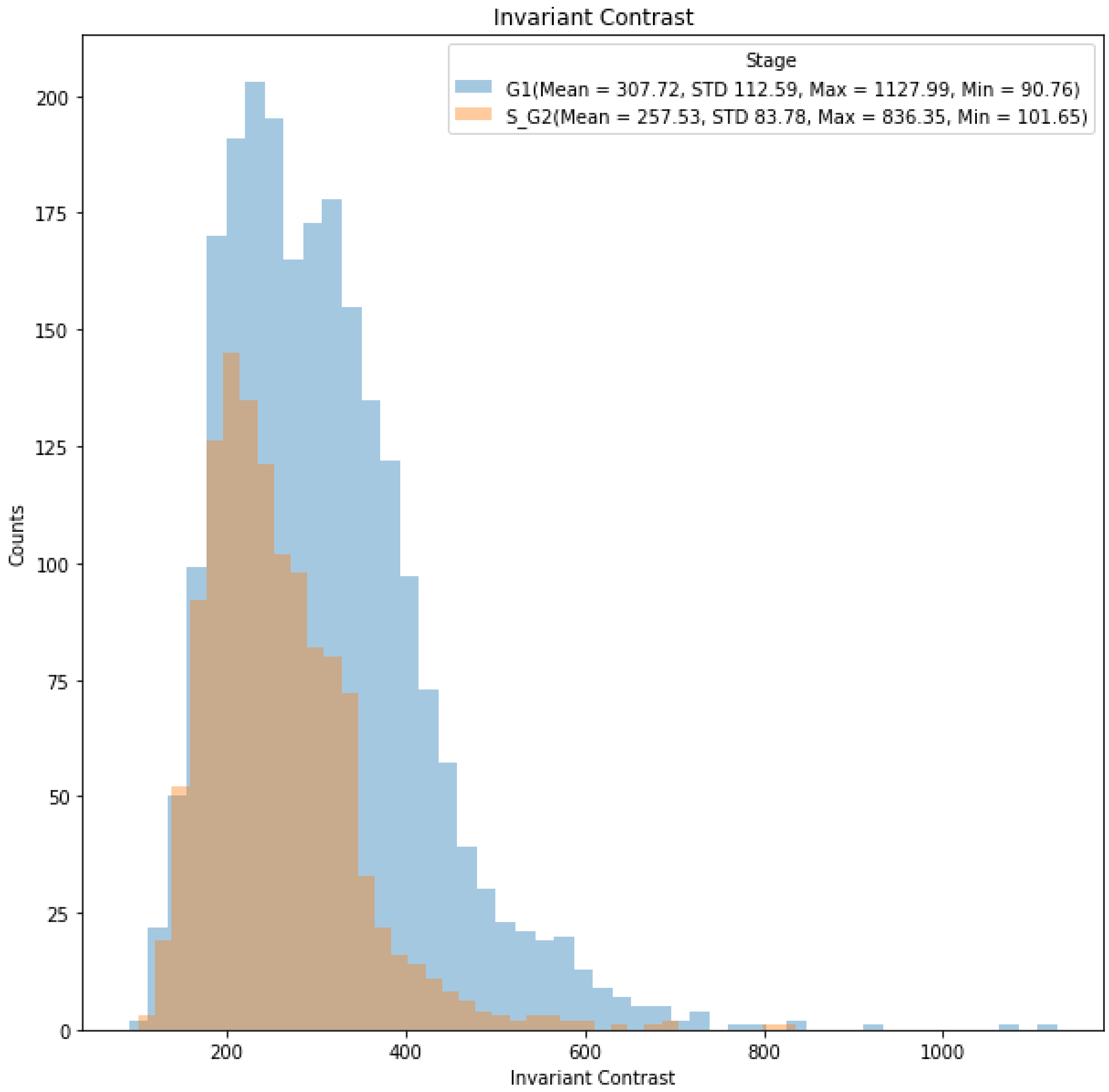

Kurtosis
Stage
G1(Mean = 0.86, STD 1.08, Max = 10.18, Min = -1.10)
S_G2(Mean = 0.12, STD 0.60, Max = 5.60, Min = -0.90)
Counts
Kurtosis

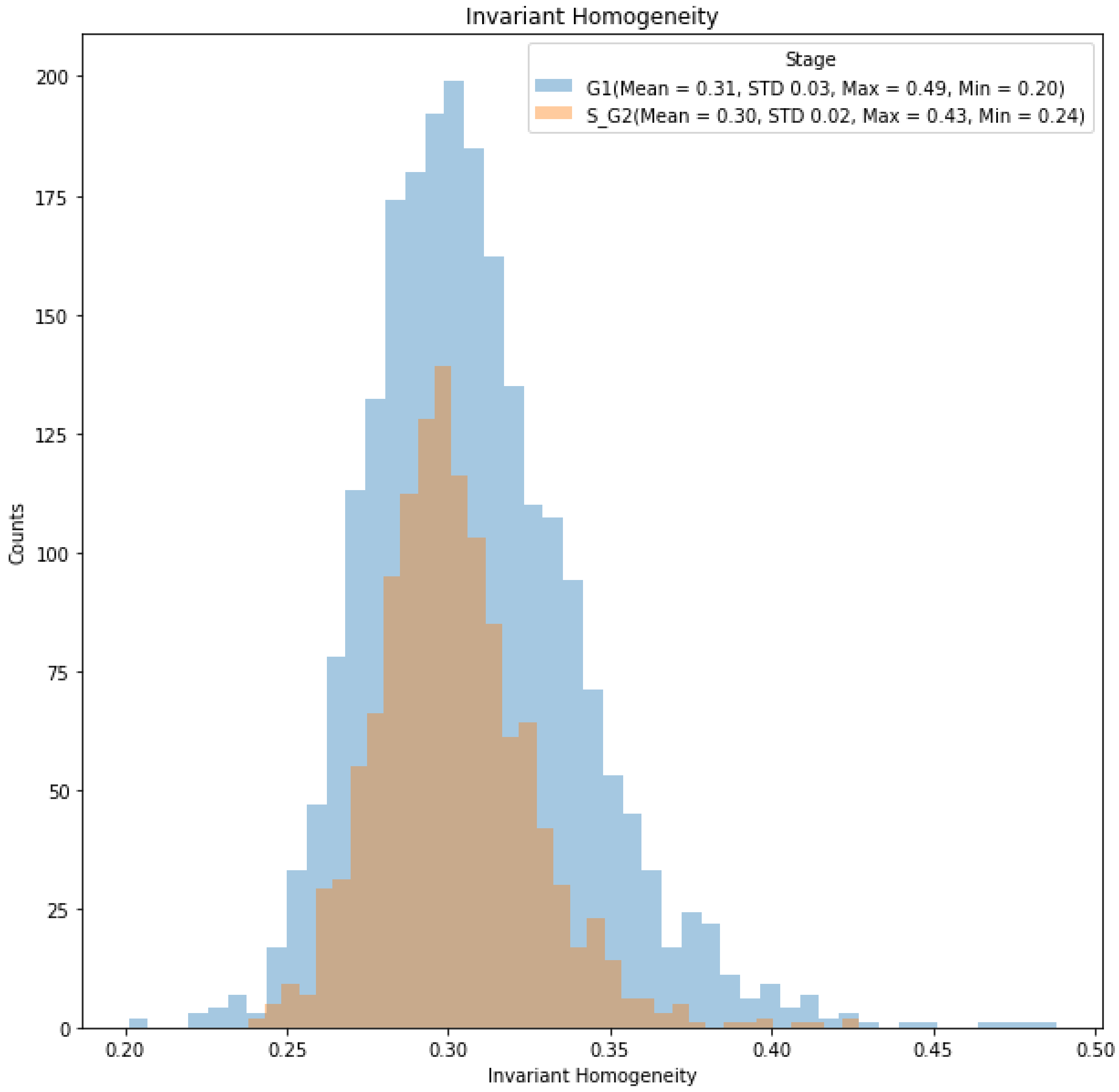

Invariant Uniformity
Stage
G1(Mean = 29013741.66, STD 20089596.46, Max = 210775340.00, Min = 3902671.00)
S_G2(Mean = 36345344.48, STD 18427659.26, Max = 197188843.00, Min = 3623322.00)
Counts
Invariant Uniformity
1e8

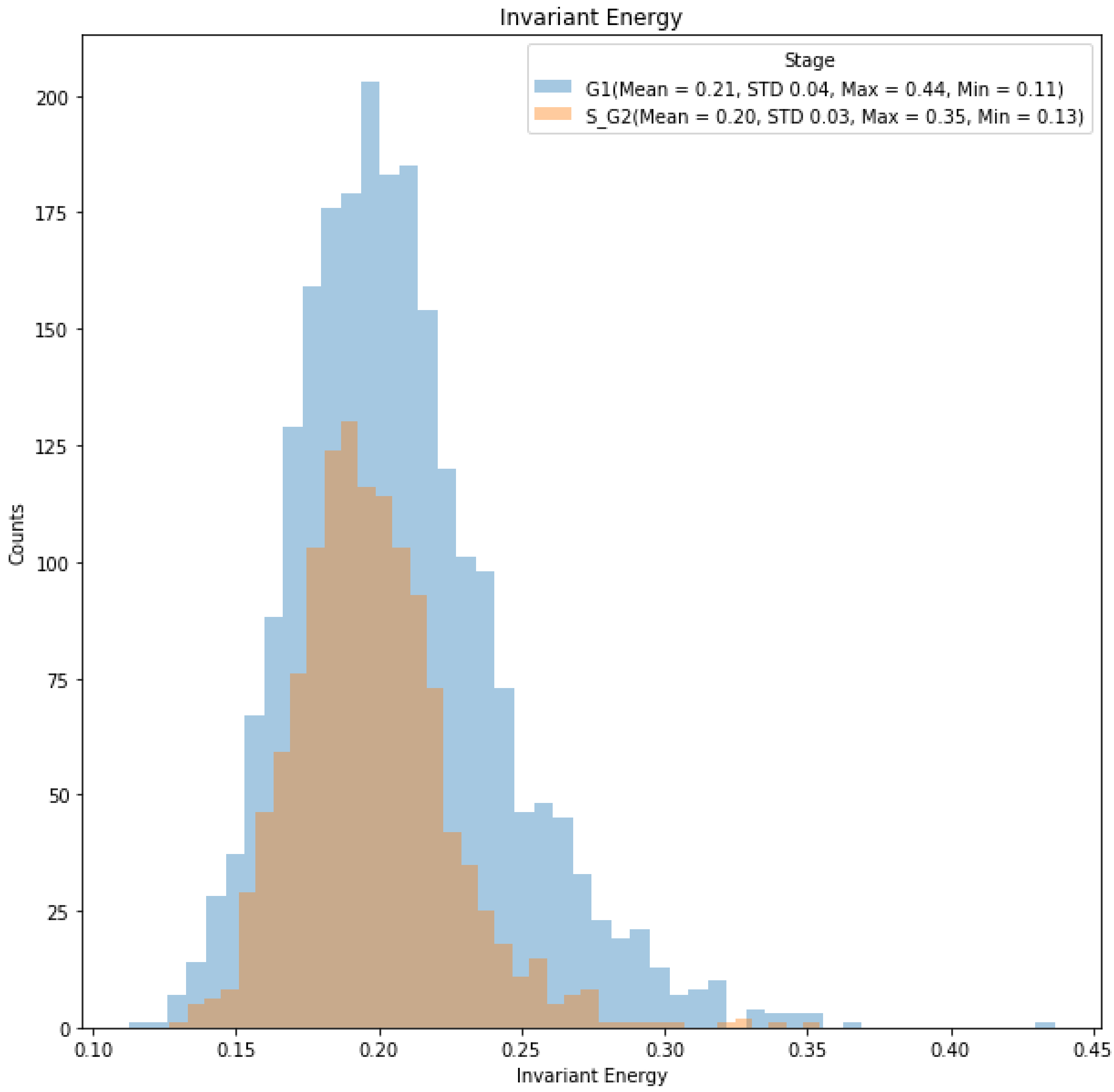

GLCM Invariant Entropy
Stage
G1(Mean = 0.41, STD 0.09, Max = 0.71, Min = 0.08)
S_G2(Mean = 0.54, STD 0.11, Max = 0.87, Min = 0.17)
Counts
GLCM Invariant Entropy

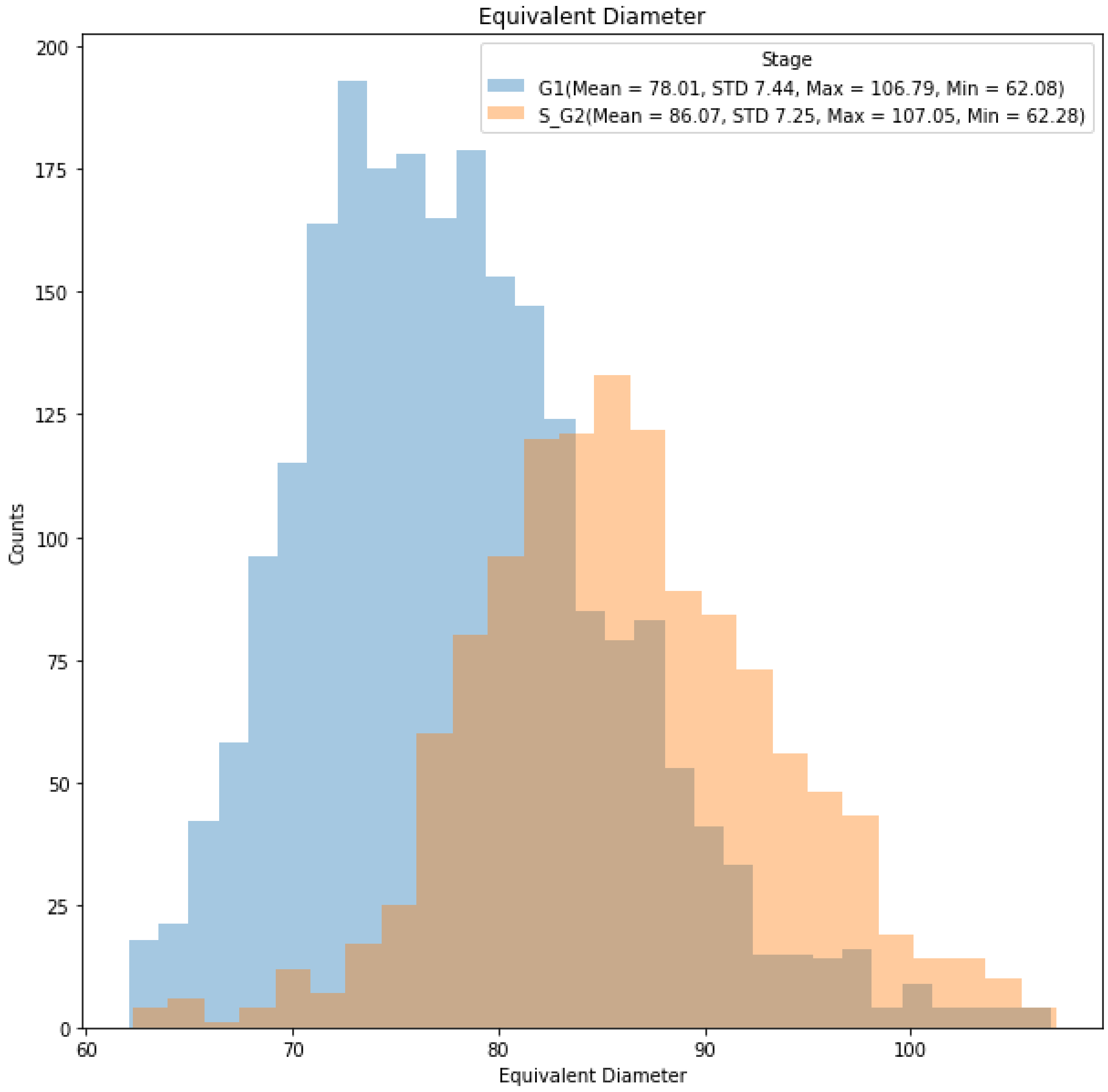

Invariant Correlation
Stage
G1(Mean = 0.96, STD 0.01, Max = 0.98, Min = 0.88)
S_G2(Mean = 0.97, STD 0.01, Max = 0.98, Min = 0.93)
Counts
Invariant Correlation

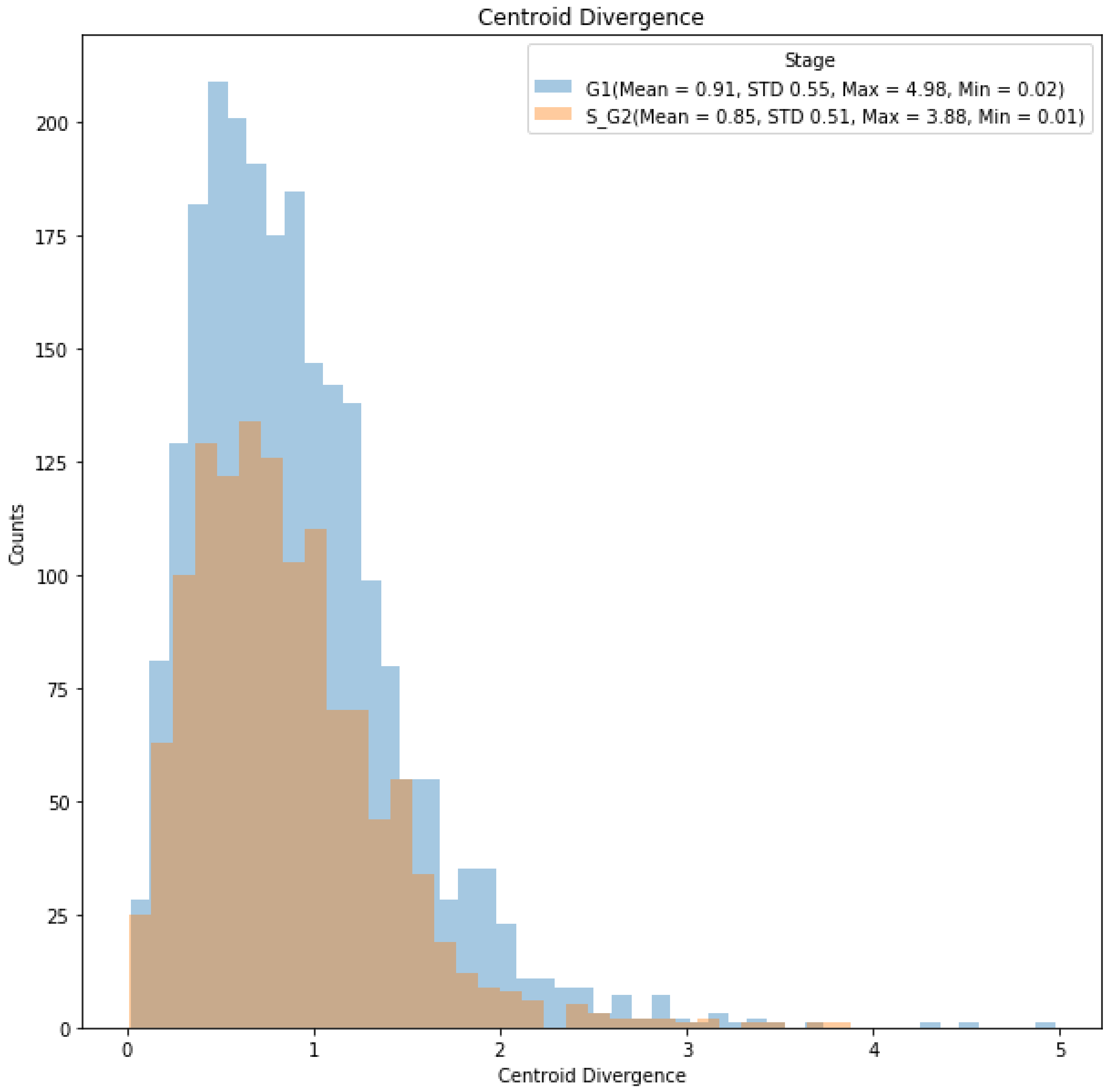

Invariant Dissimilarity
Stage
G1(Mean = 9.18, STD 1.86, Max = 17.73, Min = 4.14)
S_G2(Mean = 8.32, STD 1.41, Max = 13.90, Min = 5.38)
Counts
Invariant Dissimilarity

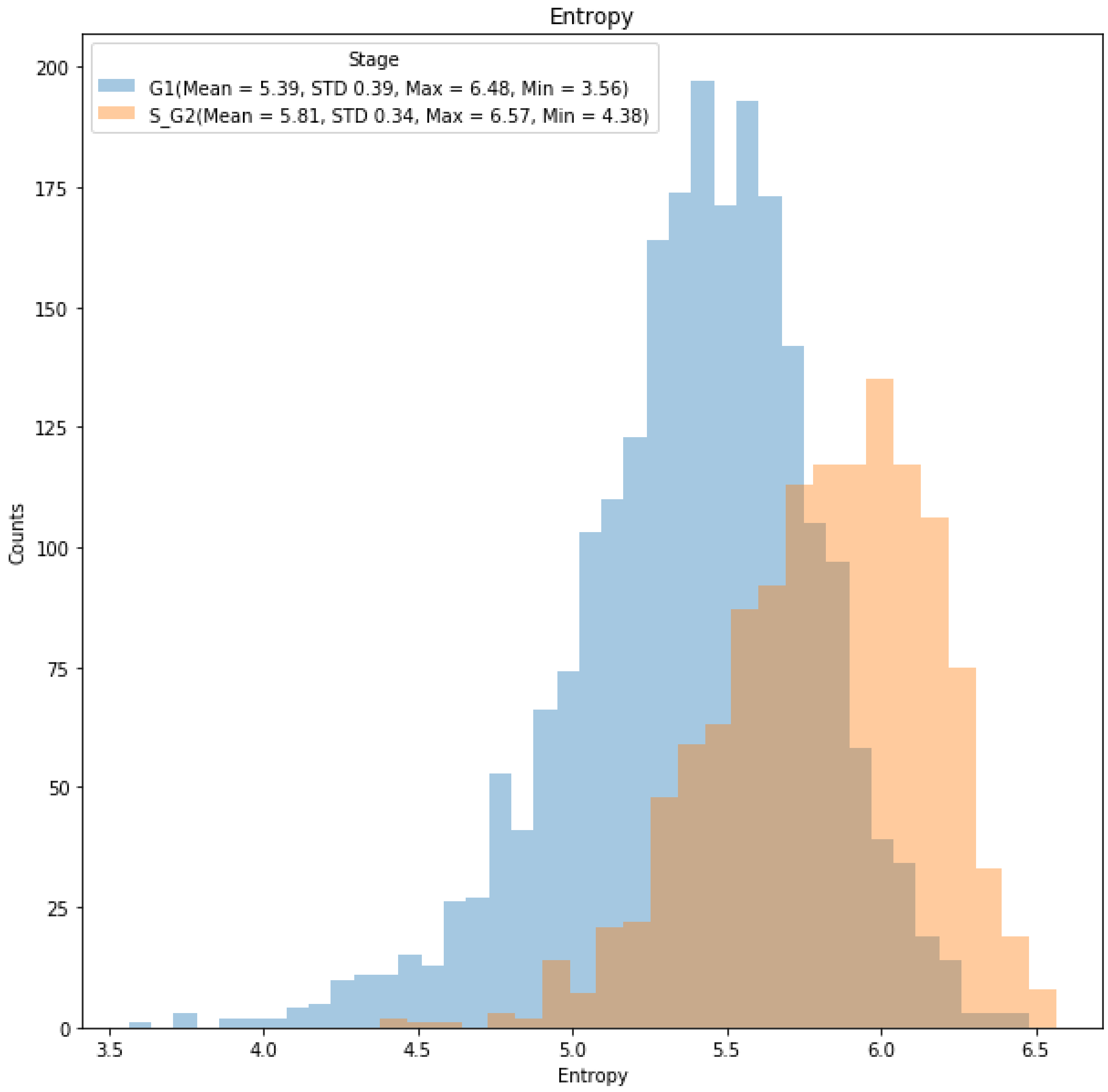

Invariant Contrast
Stage
G1(Mean = 307.72, STD 112.59, Max = 1127.99, Min = 90.76)
S_G2(Mean = 257.53, STD 83.78, Max = 836.35, Min = 101.65)
Counts
Invariant Contrast

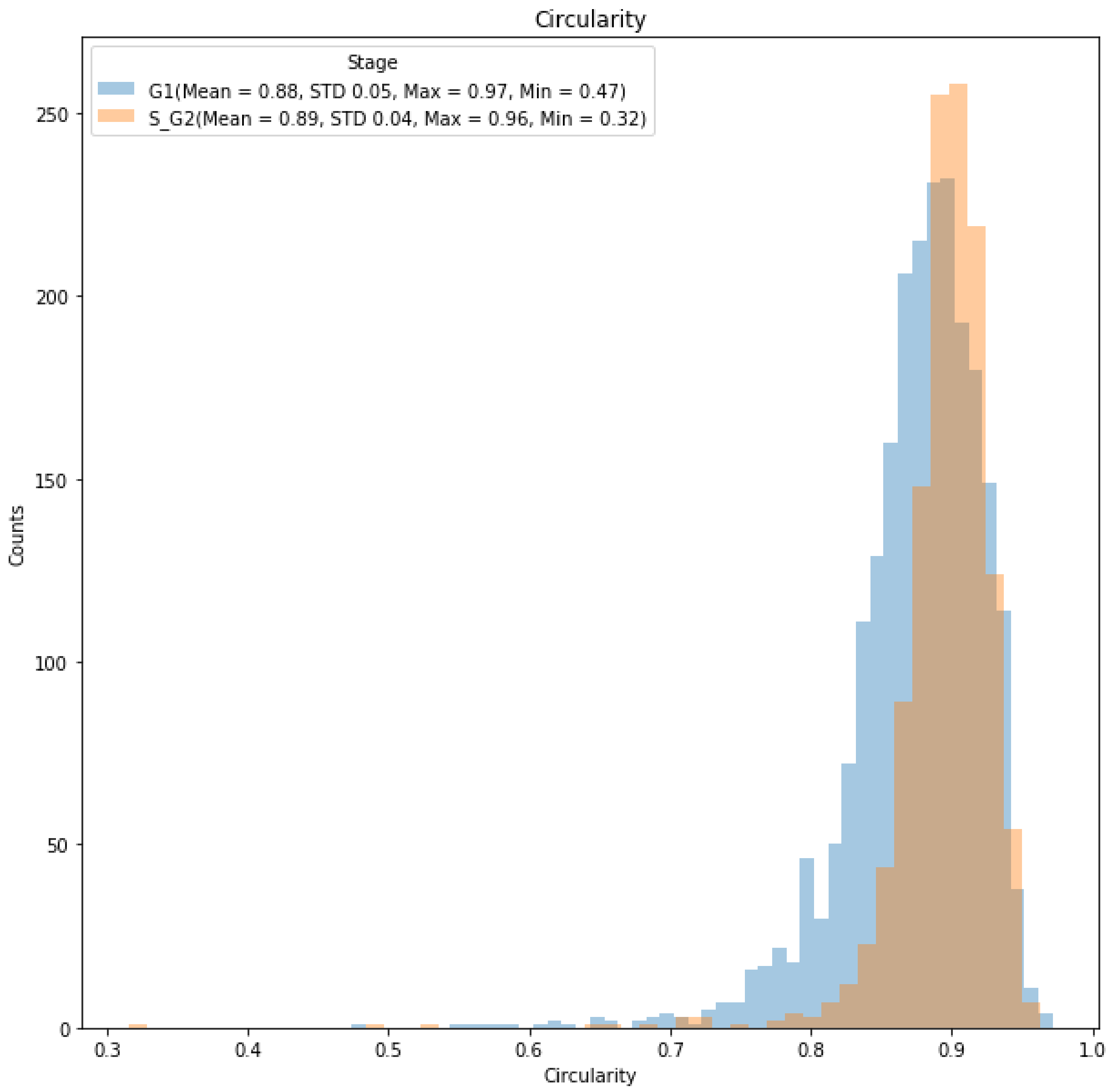

Invariant Homogeneity
Stage
G1(Mean = 0.31, STD 0.03, Max = 0.49, Min = 0.20)
S_G2(Mean = 0.30, STD 0.02, Max = 0.43, Min = 0.24)
Counts
0
25
50
75
100
125
150
175
200
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Invariant Homogeneity

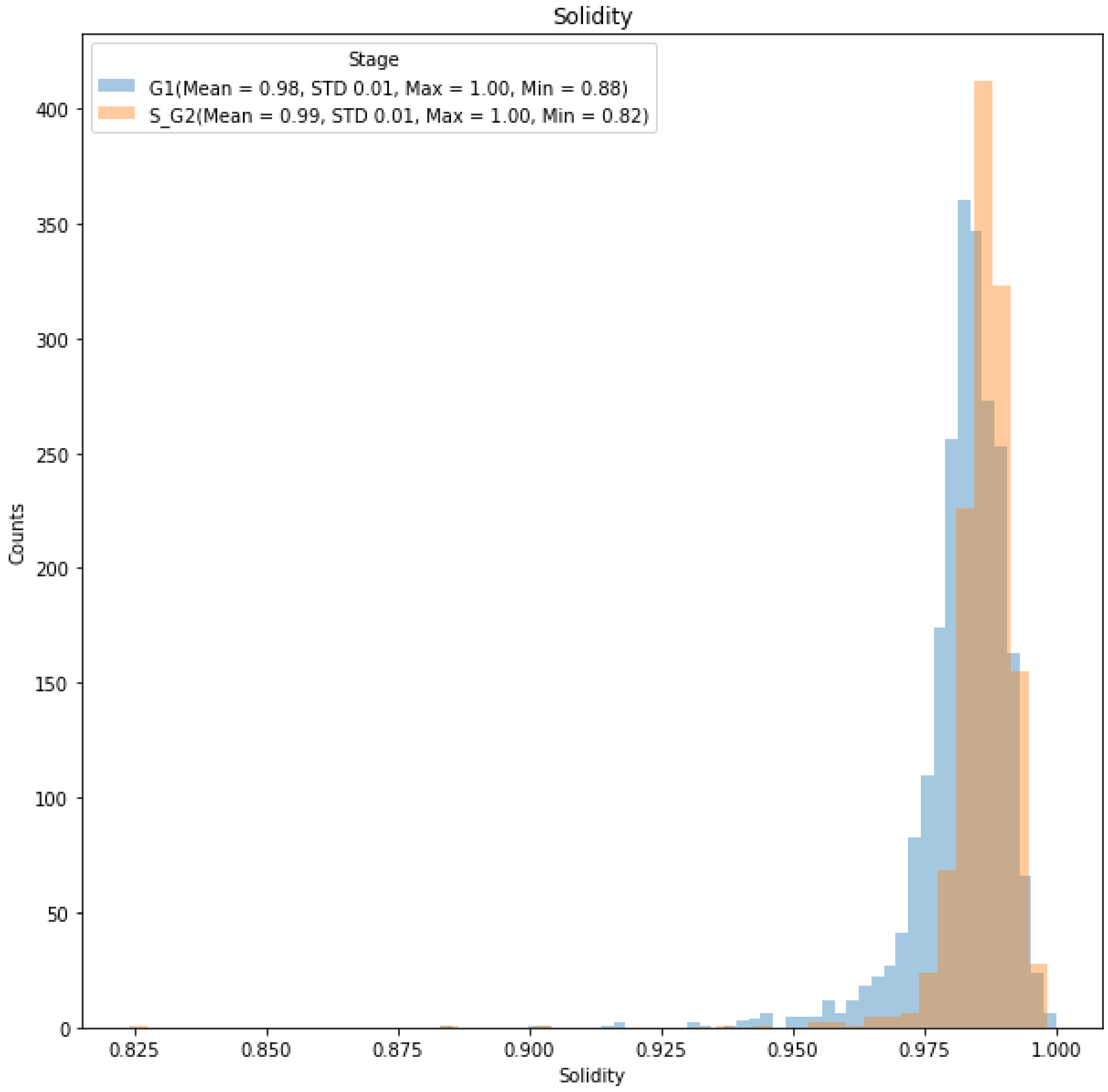

Invariant Energy
Stage
G1(Mean = 0.21, STD 0.04, Max = 0.44, Min = 0.11)
S_G2(Mean = 0.20, STD 0.03, Max = 0.35, Min = 0.13)
Counts
0
25
50
75
100
125
150
175
200
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
Invariant Energy

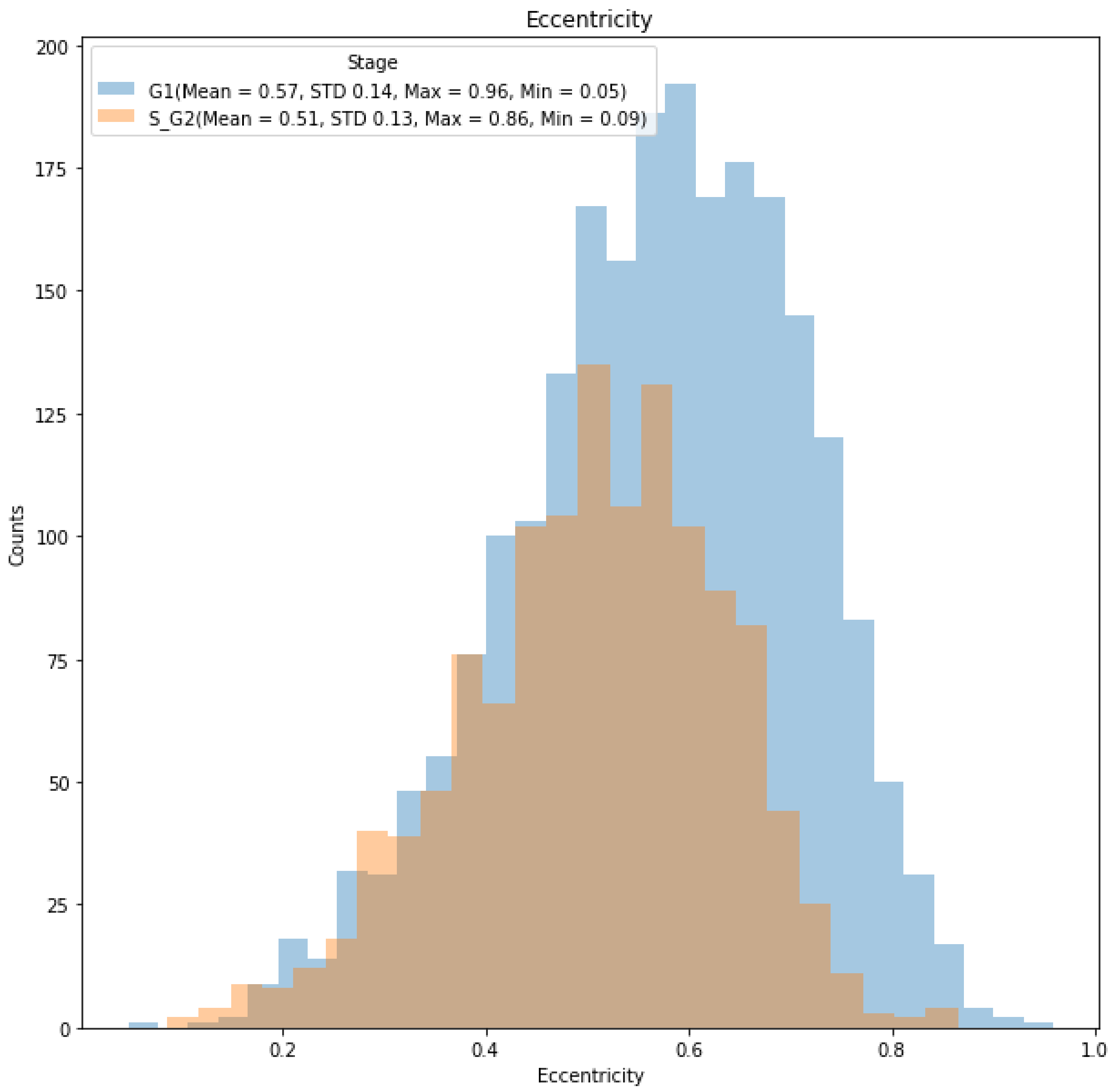

Equivalent Diameter
Stage
G1(Mean = 78.01, STD 7.44, Max = 106.79, Min = 62.08)
S_G2(Mean = 86.07, STD 7.25, Max = 107.05, Min = 62.28)
Counts
Equivalent Diameter

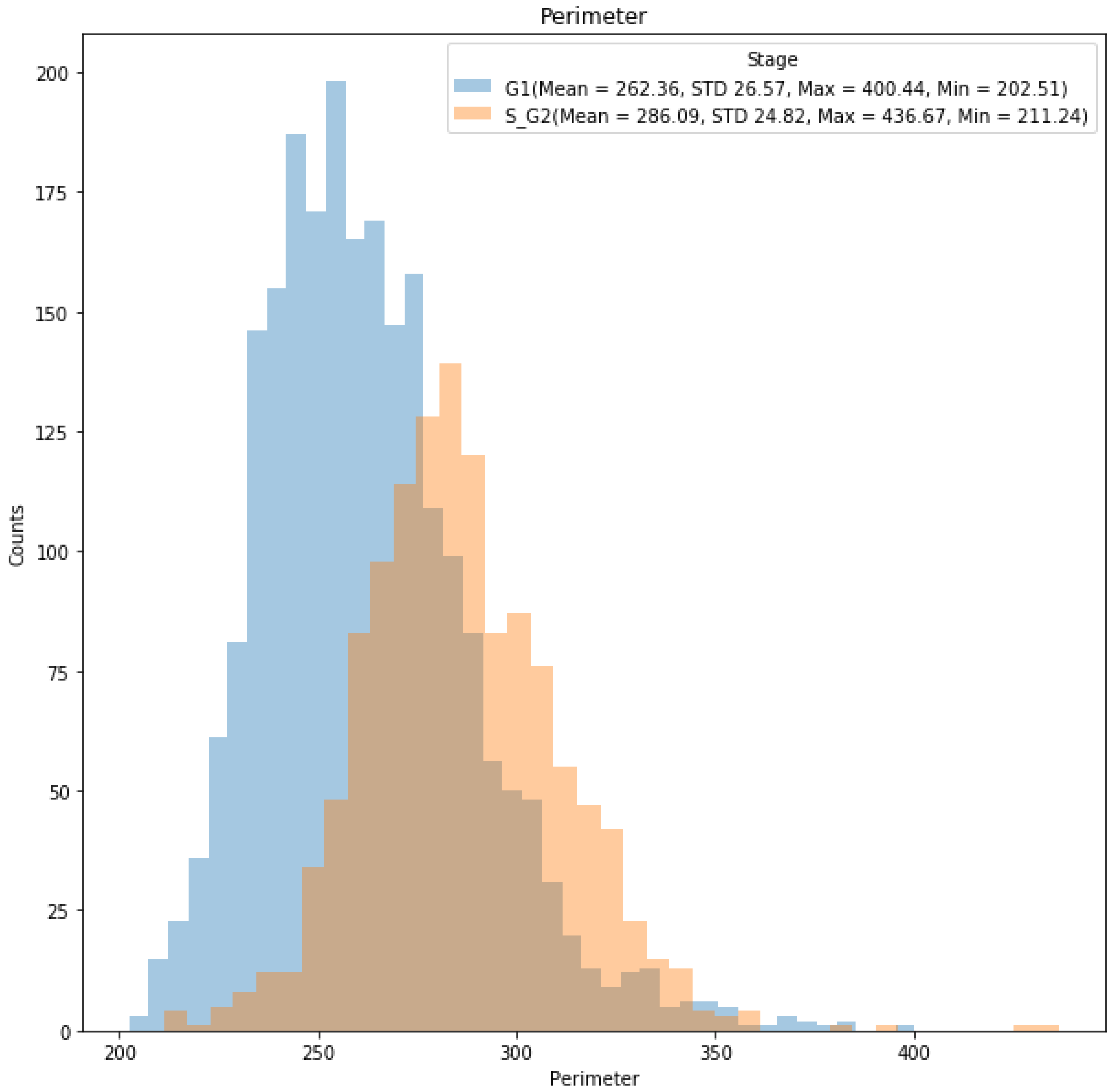

Centroid Divergence
Stage
G1(Mean = 0.91, STD 0.55, Max = 4.98, Min = 0.02)
S_G2(Mean = 0.85, STD 0.51, Max = 3.88, Min = 0.01)
Counts
Centroid Divergence

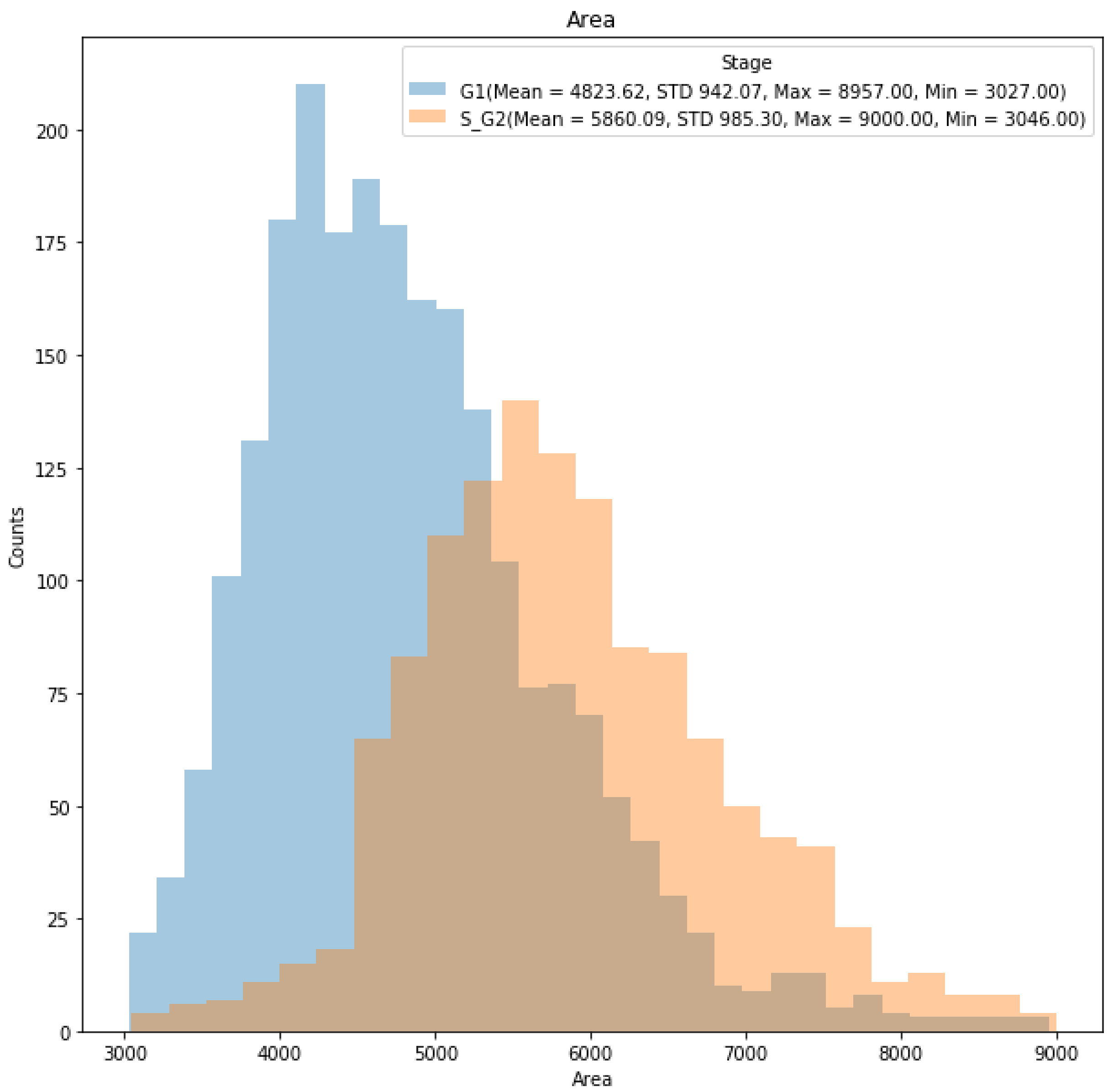

Entropy
Stage
G1(Mean = 5.39, STD 0.39, Max = 6.48, Min = 3.56)
S_G2(Mean = 5.81, STD 0.34, Max = 6.57, Min = 4.38)
Counts
Entropy

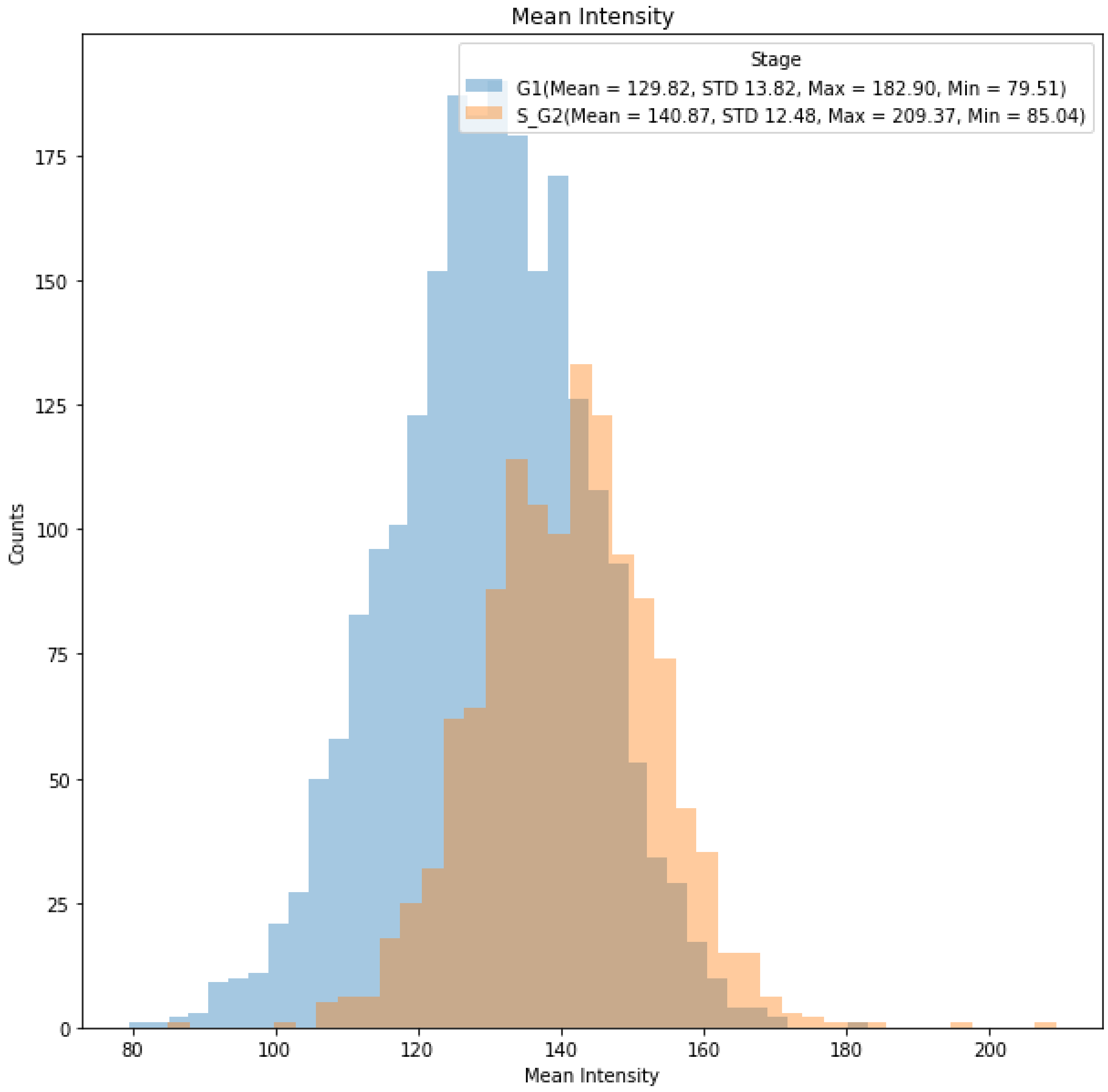

Circularity
Stage
G1(Mean = 0.88, STD 0.05, Max = 0.97, Min = 0.47)
S_G2(Mean = 0.89, STD 0.04, Max = 0.96, Min = 0.32)
Counts
Circularity

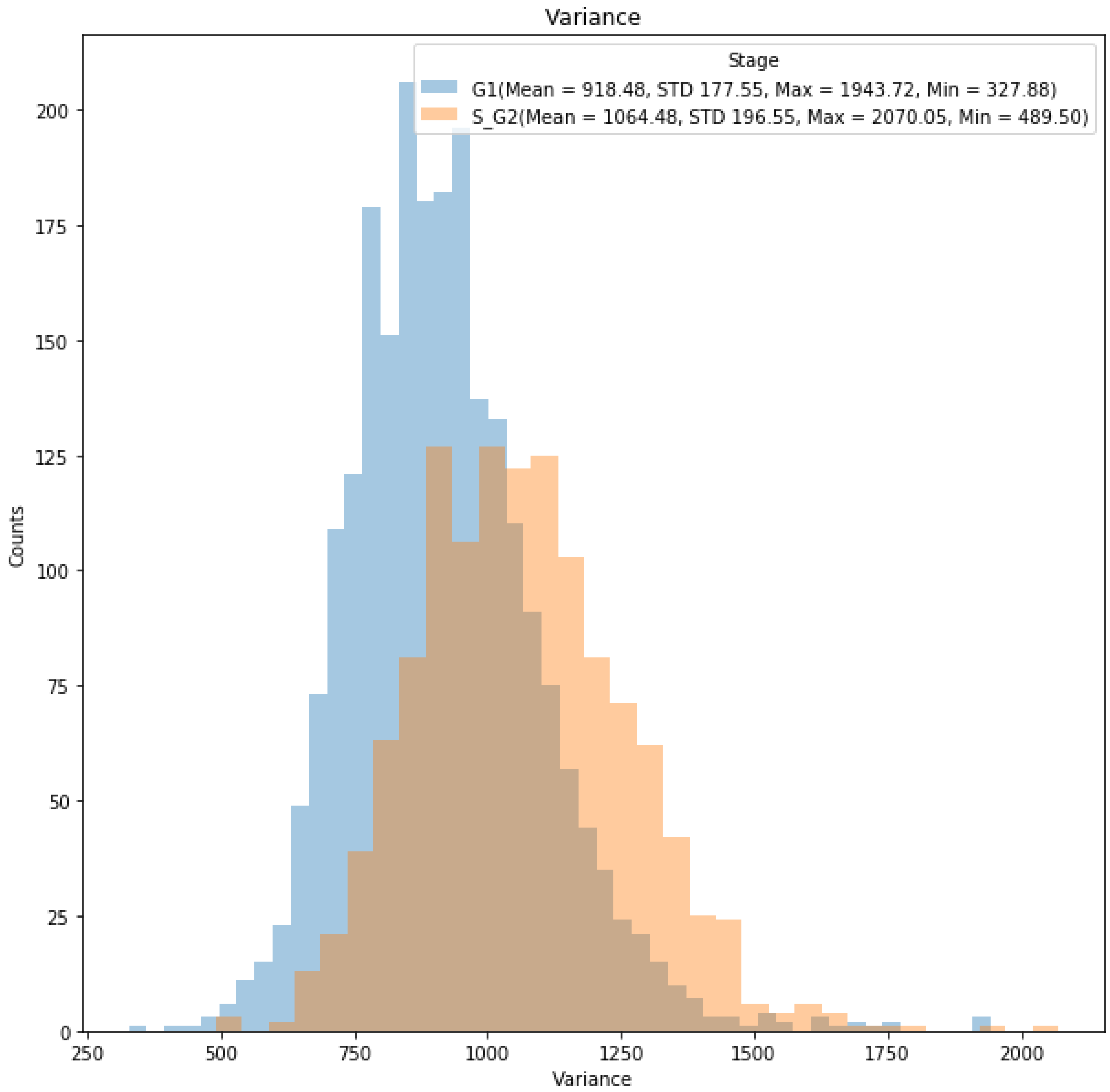

Solidity
Stage
G1(Mean = 0.98, STD 0.01, Max = 1.00, Min = 0.88)
S_G2(Mean = 0.99, STD 0.01, Max = 1.00, Min = 0.82)
Counts
Solidity
0
50
100
150
200
250
300
350
400
0.825
0.850
0.875
0.900
0.925
0.950
0.975
1.000

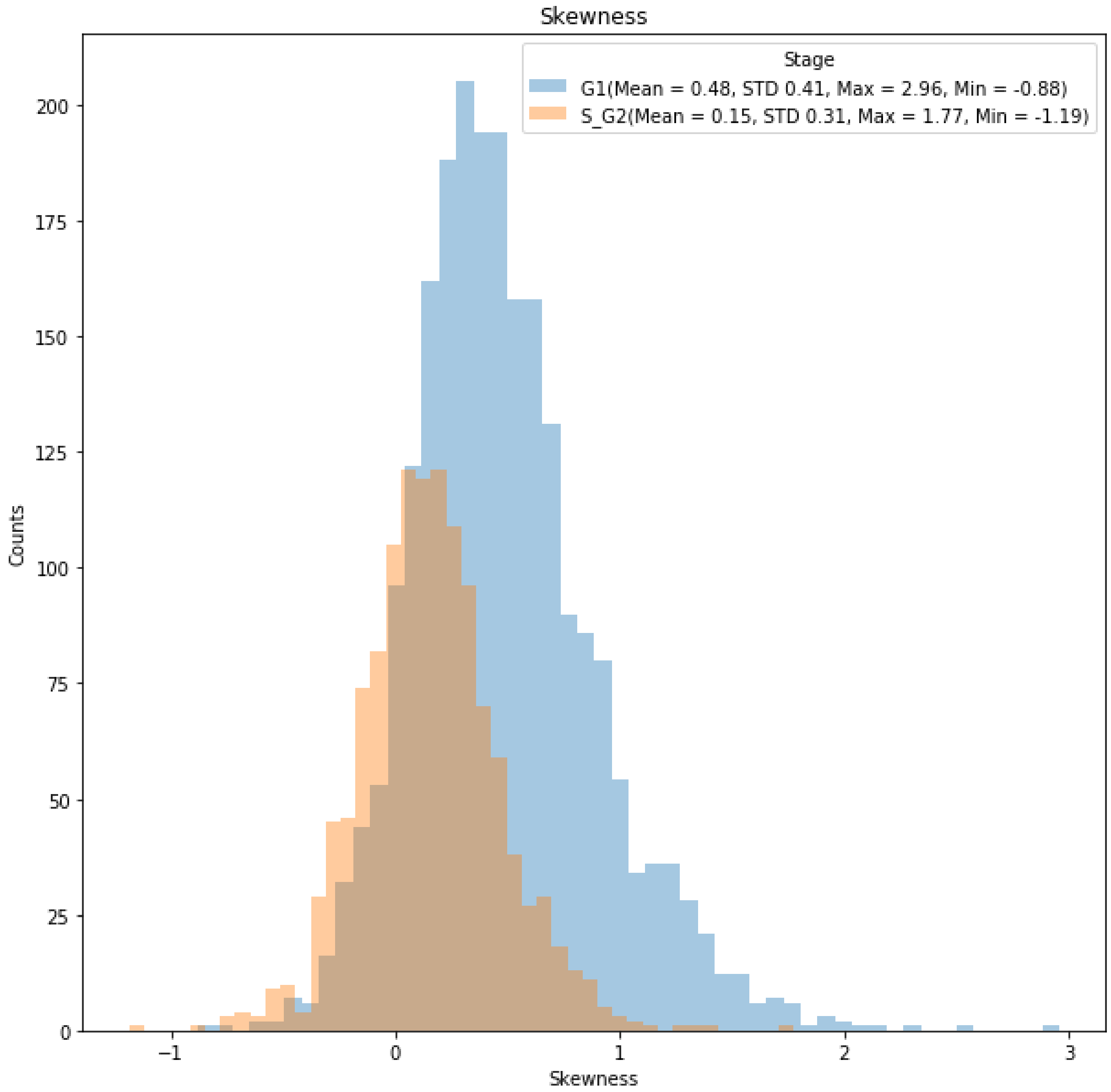

Eccentricity
Stage
G1(Mean = 0.57, STD 0.14, Max = 0.96, Min = 0.05)
S_G2(Mean = 0.51, STD 0.13, Max = 0.86, Min = 0.09)
Counts
Eccentricity
0
25
50
75
100
125
150
175
200
0.2
0.4
0.6
0.8
1.0

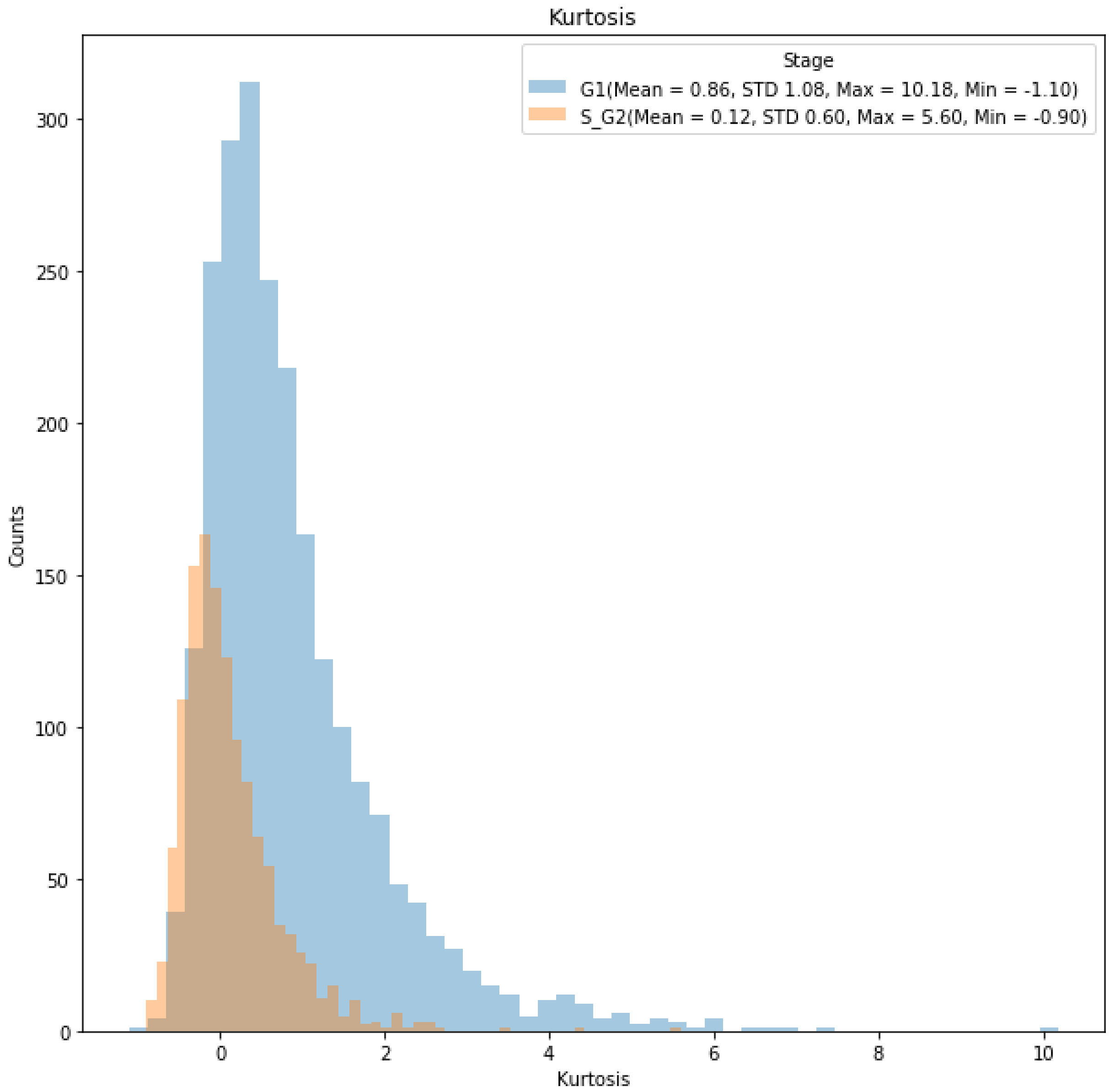

Perimeter
Stage
G1(Mean = 262.36, STD 26.57, Max = 400.44, Min = 202.51)
S_G2(Mean = 286.09, STD 24.82, Max = 436.67, Min = 211.24)
Counts
Perimeter
0
25
50
75
100
125
150
175
200
200
250
300
350
400

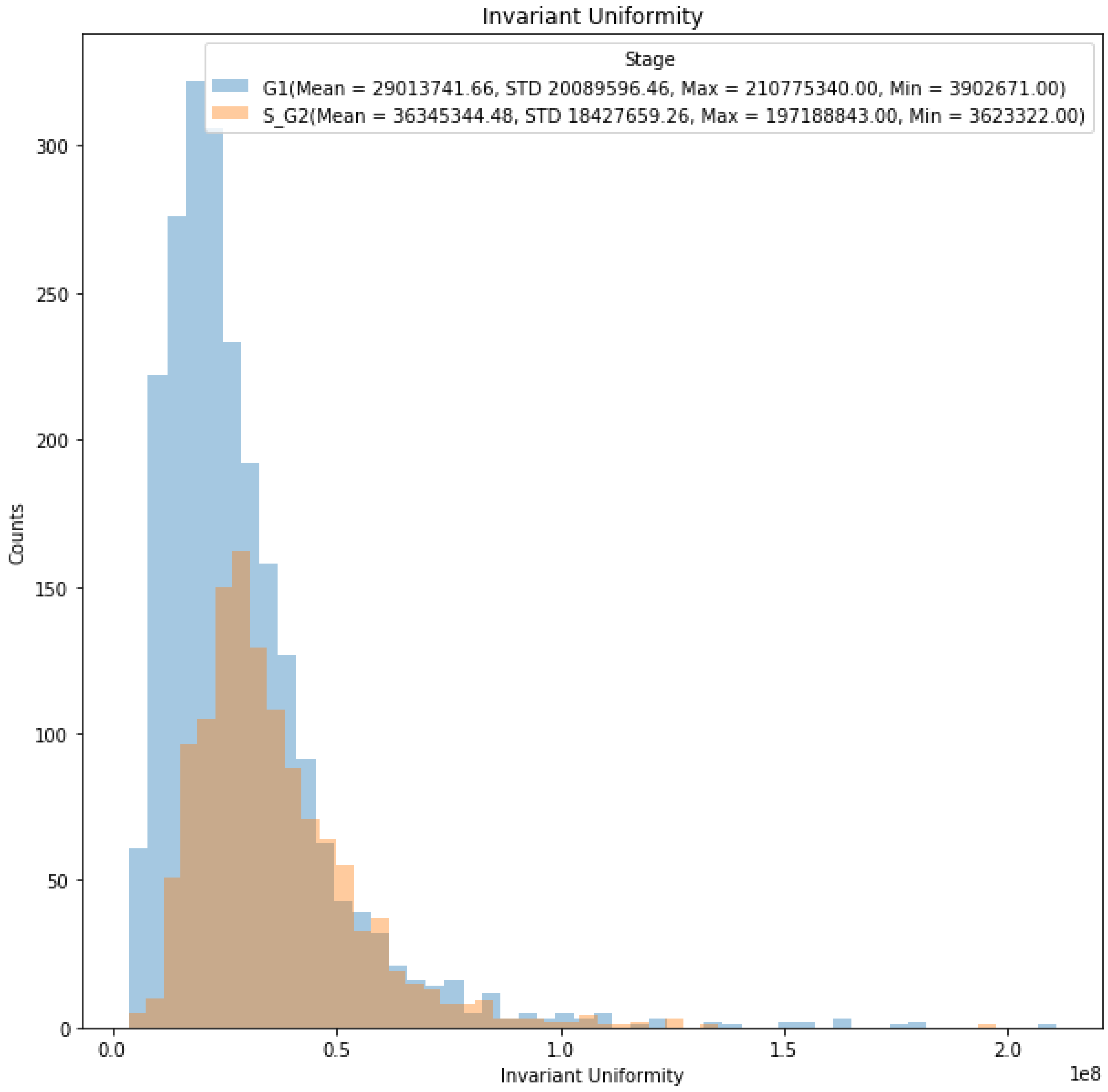

Area
Stage
G1(Mean = 4823.62, STD 942.07, Max = 8957.00, Min = 3027.00)
S_G2(Mean = 5860.09, STD 985.30, Max = 9000.00, Min = 3046.00)
Counts
Area
0
25
50
75
100
125
150
175
200
3000
4000
5000
6000
7000
8000
9000

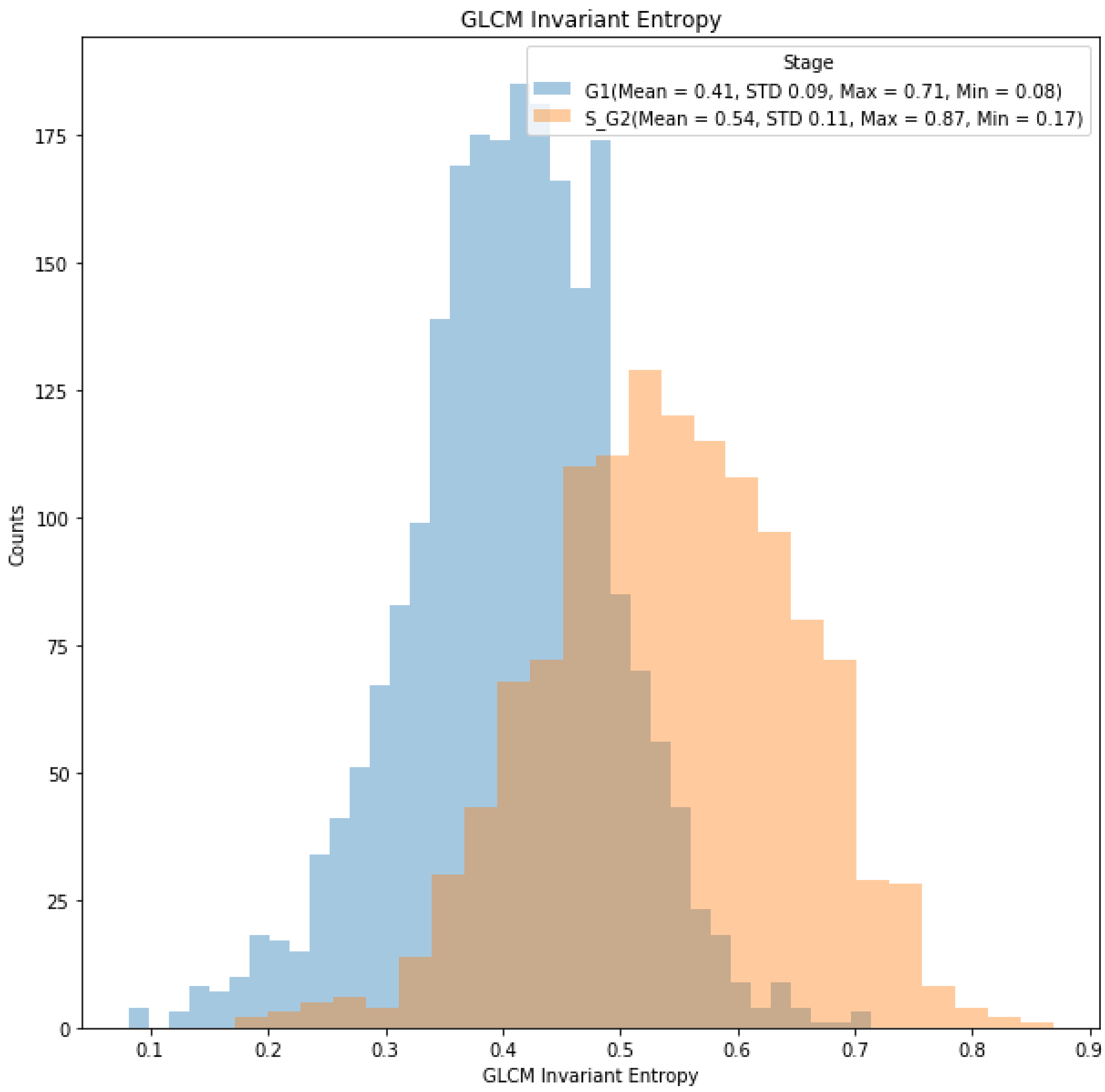

Mean Intensity
Stage
G1(Mean = 129.82, STD 13.82, Max = 182.90, Min = 79.51)
S_G2(Mean = 140.87, STD 12.48, Max = 209.37, Min = 85.04)
Counts
Mean Intensity

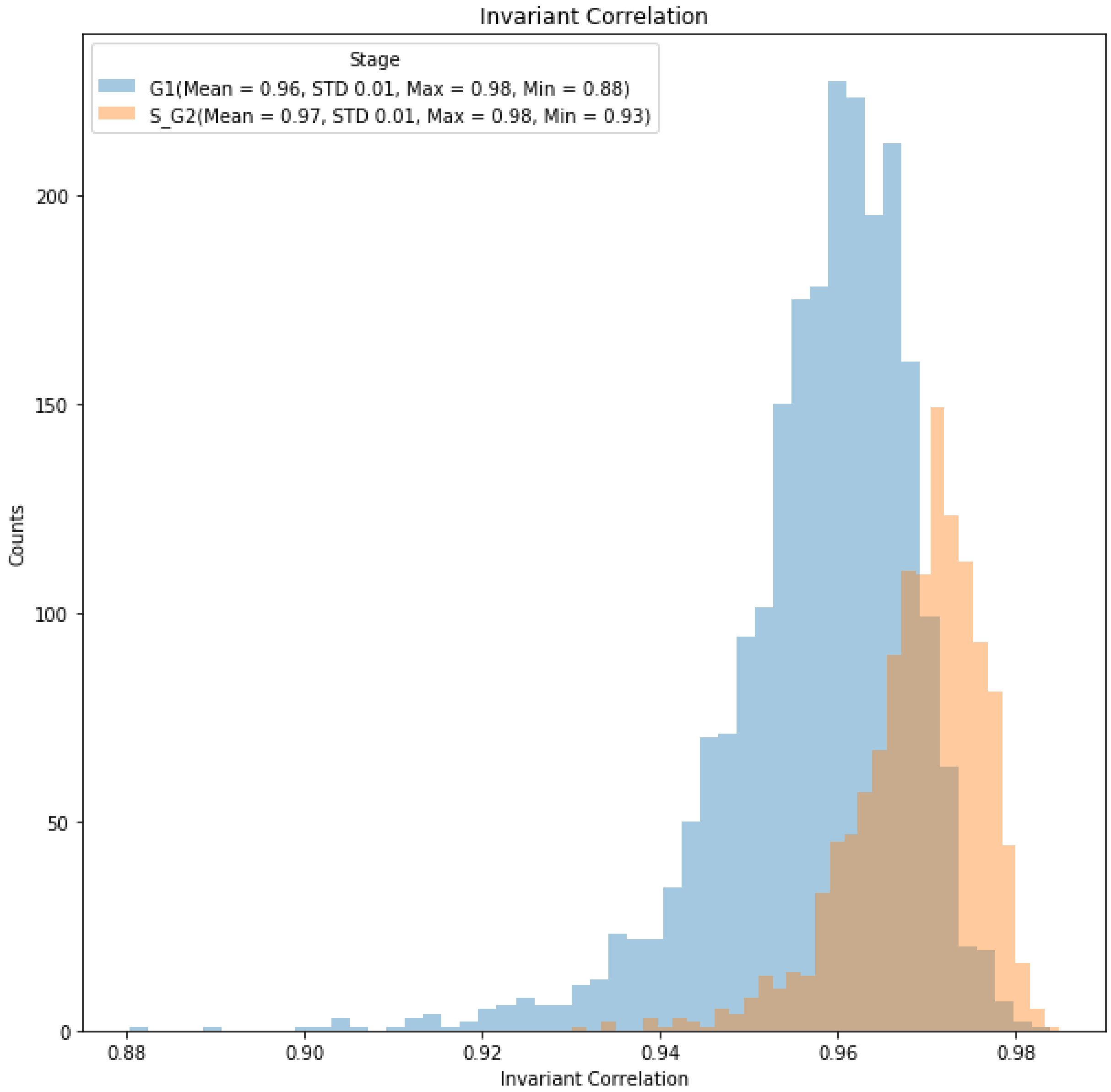

Variance
Stage
G1(Mean = 918.48, STD 177.55, Max = 1943.72, Min = 327.88)
S_G2(Mean = 1064.48, STD 196.55, Max = 2070.05, Min = 489.50)
Counts
Variance

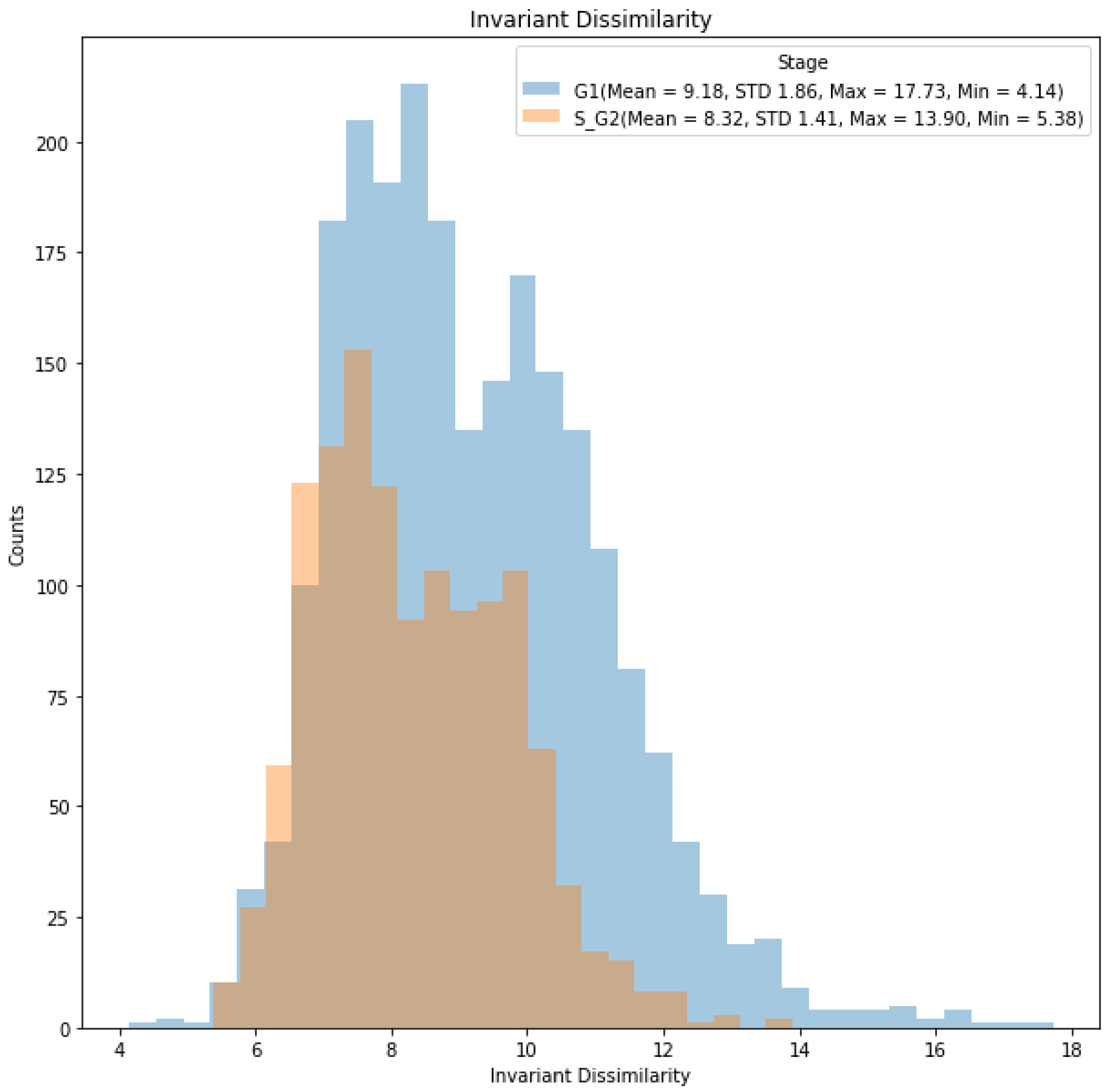

Skewness
Stage
G1(Mean = 0.48, STD 0.41, Max = 2.96, Min = -0.88)
S_G2(Mean = 0.15, STD 0.31, Max = 1.77, Min = -1.19)
Counts
Skewness

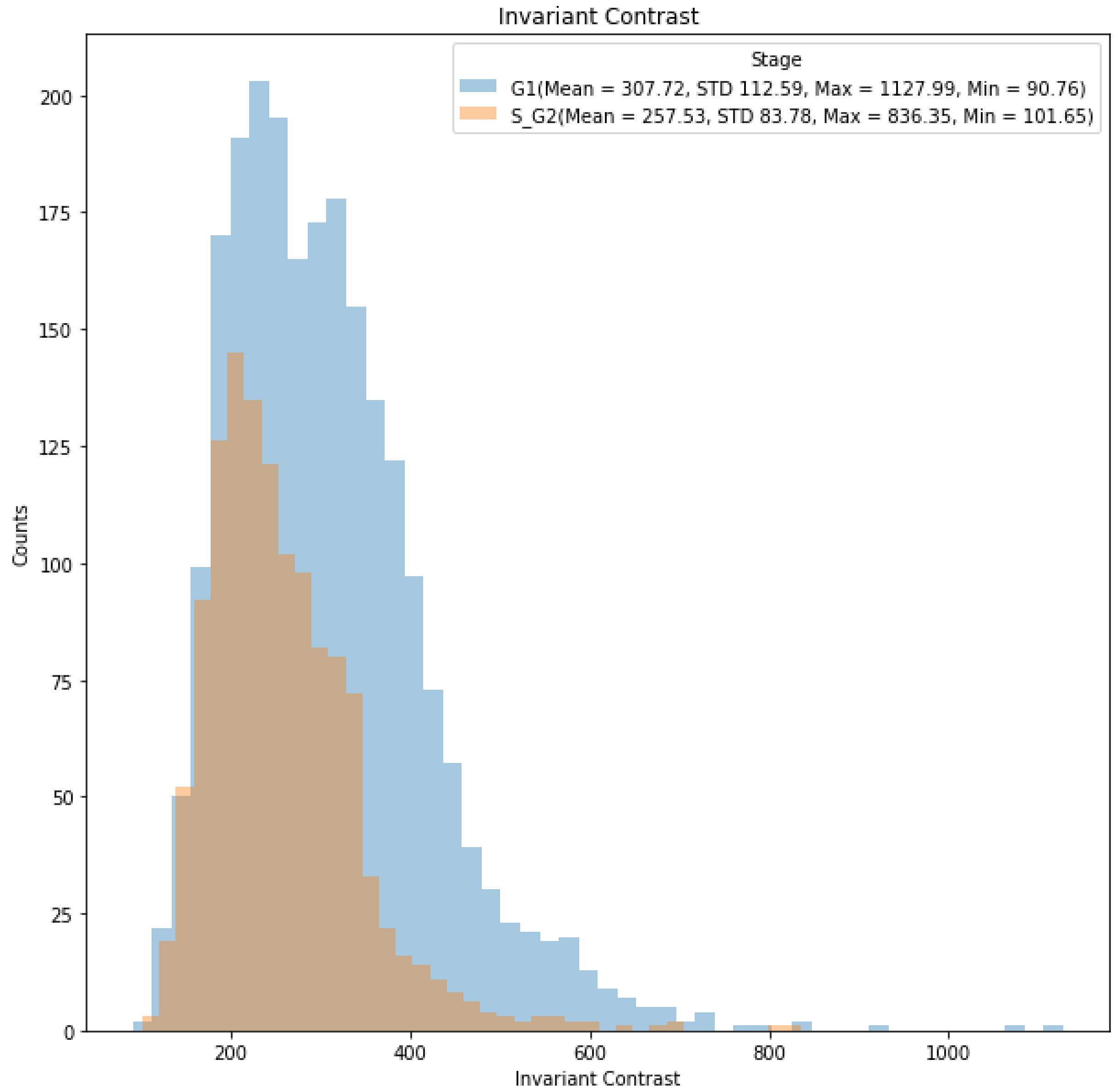

Kurtosis
Stage
G1(Mean = 0.86, STD 1.08, Max = 10.18, Min = -1.10)
S_G2(Mean = 0.12, STD 0.60, Max = 5.60, Min = -0.90)
Counts
Kurtosis

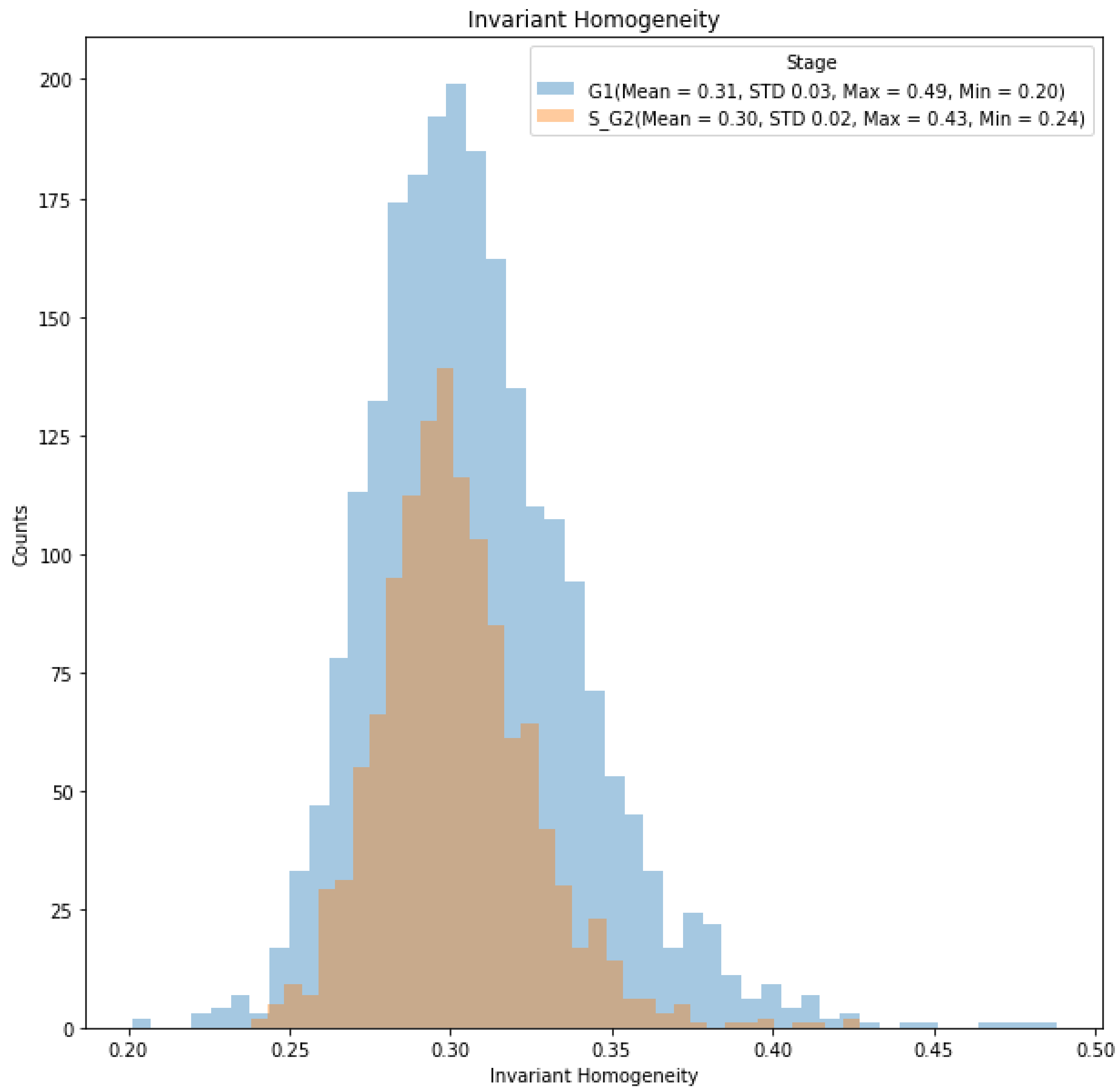

Invariant Uniformity
Stage
G1(Mean = 29013741.66, STD 20089596.46, Max = 210775340.00, Min = 3902671.00)
S_G2(Mean = 36345344.48, STD 18427659.26, Max = 197188843.00, Min = 3623322.00)
Counts
Invariant Uniformity
1e8

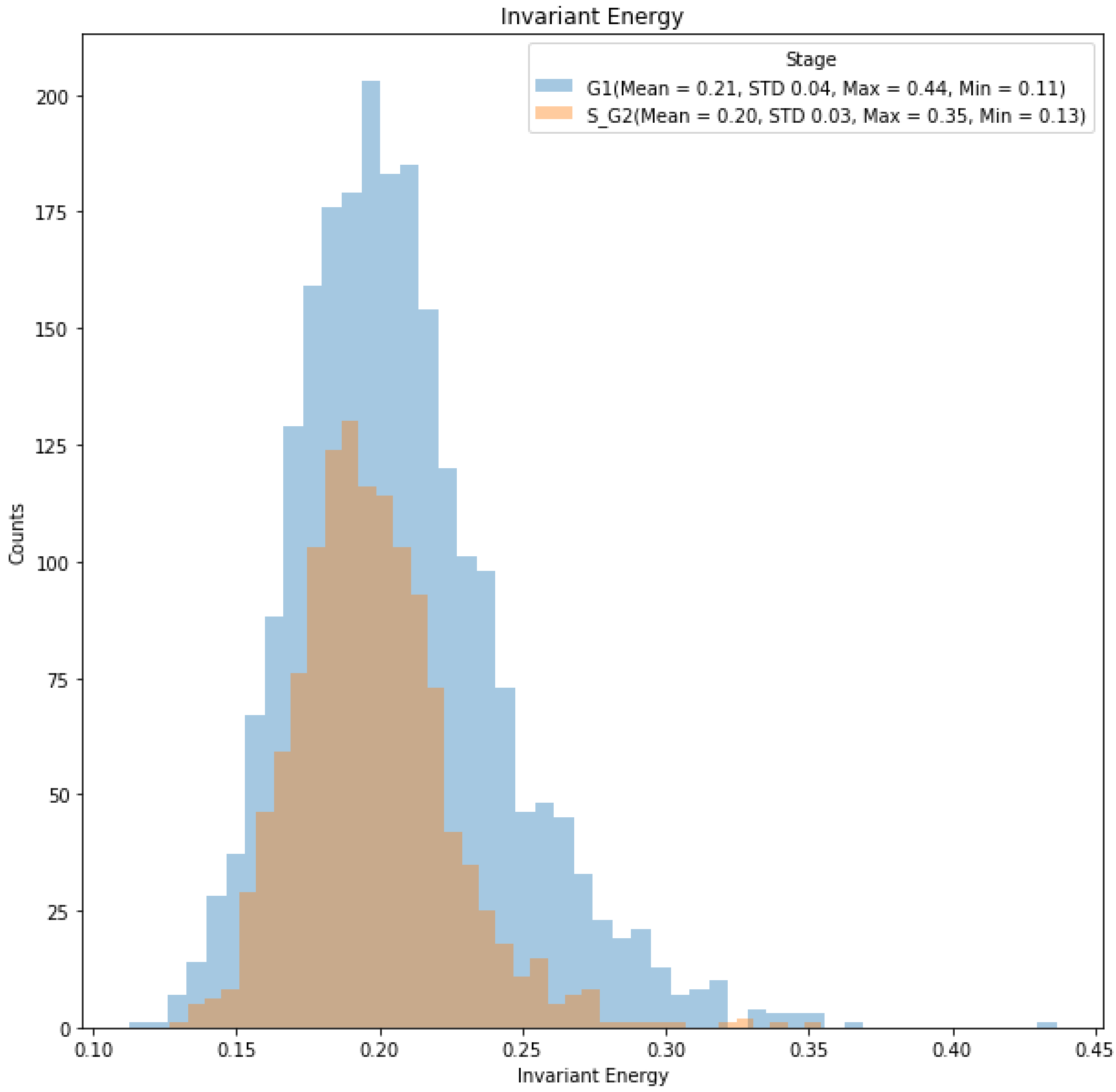

GLCM Invariant Entropy
Stage
G1(Mean = 0.41, STD 0.09, Max = 0.71, Min = 0.08)
S_G2(Mean = 0.54, STD 0.11, Max = 0.87, Min = 0.17)
Counts
GLCM Invariant Entropy

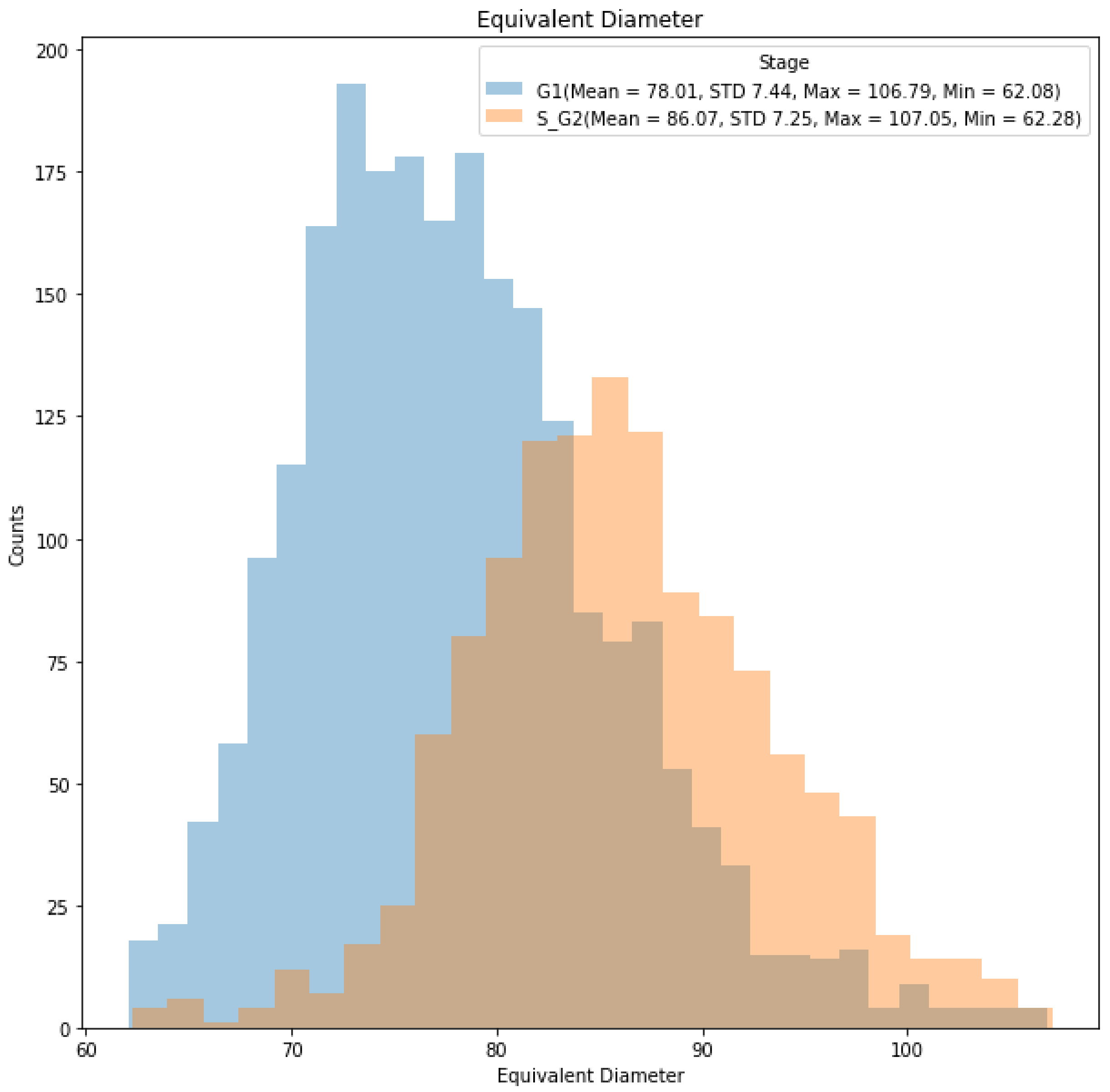

Invariant Correlation
Stage
G1(Mean = 0.96, STD 0.01, Max = 0.98, Min = 0.88)
S_G2(Mean = 0.97, STD 0.01, Max = 0.98, Min = 0.93)
Counts
0
50
100
150
200
0.88
0.90
0.92
0.94
0.96
0.98
Invariant Correlation

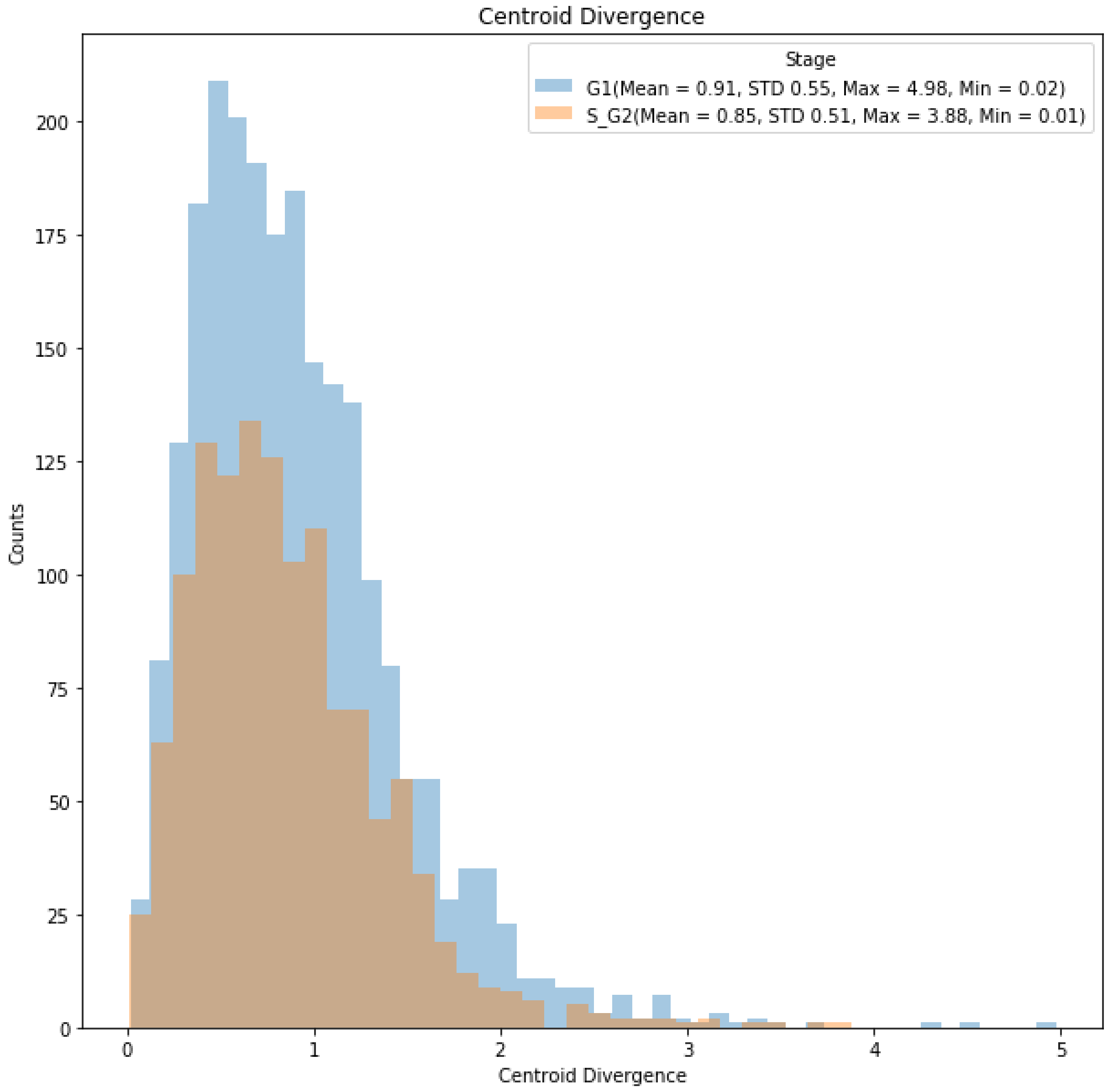

Invariant Dissimilarity
Stage
G1(Mean = 9.18, STD 1.86, Max = 17.73, Min = 4.14)
S_G2(Mean = 8.32, STD 1.41, Max = 13.90, Min = 5.38)
Counts
0
25
50
75
100
125
150
175
200
4
6
8
10
12
14
16
18
Invariant Dissimilarity


Invariant Contrast
Stage
G1(Mean = 307.72, STD 112.59, Max = 1127.99, Min = 90.76)
S_G2(Mean = 257.53, STD 83.78, Max = 836.35, Min = 101.65)
Counts
0
25
50
75
100
125
150
175
200
200
400
600
800
1000
Invariant Contrast


Invariant Homogeneity
Stage
G1(Mean = 0.31, STD 0.03, Max = 0.49, Min = 0.20)
S_G2(Mean = 0.30, STD 0.02, Max = 0.43, Min = 0.24)
Counts
0
25
50
75
100
125
150
175
200
0.20
0.25
0.30
0.35
0.40
0.45
0.50
Invariant Homogeneity


Invariant Energy
Stage
G1(Mean = 0.21, STD 0.04, Max = 0.44, Min = 0.11)
S_G2(Mean = 0.20, STD 0.03, Max = 0.35, Min = 0.13)
Counts
0
25
50
75
100
125
150
175
200
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0.45
Invariant Energy


Equivalent Diameter
Stage
G1(Mean = 78.01, STD 7.44, Max = 106.79, Min = 62.08)
S_G2(Mean = 86.07, STD 7.25, Max = 107.05, Min = 62.28)
Counts
0
25
50
75
100
125
150
175
200
60
70
80
90
100
Equivalent Diameter



In [46]:

```
#not used, dataset for each phase not the same size

#paired t-test, to check how a change of phase causes differences on nuclei features


def plt_hist_and_stats_ttest_paired(columns, types, histogram = True):
    for column in columns:
        #create a new figure
        plt.figure()
        for subtype in types:

            tp = eval(subtype)
            #subset to the type

            #compute some statistics
            aux = tp[column].describe()

            #Draw the density plot
            sns.distplot(tp[column], hist = histogram, kde = False,
                         label = subtype+r'(Mean = %0.2f, STD %0.2f, Max = %0.2f, Min = %0.2f)' % (aux['mean'], aux['st
d'], aux['max'], aux['min']))
            plt.legend(prop = {'size': 10}, title = 'Stage')
            plt.title(column)
            plt.xlabel(column)
            plt.ylabel('Counts')

        subset1  = eval(types[0])
        subset2 = eval(types[1])
        u_statistic, pVal = stats.ttest_rel(subset1[column], subset2[column])
        print('THE P-VALUE IS:')
        print(pVal)
        if pVal < 0.05:
            aux = '\033[1m' + ' is ' + '\033[0m'
            aux = colored(aux, 'blue')
            hipothesis =  'H1: The difference' + aux + 'statistically significant (at significance level: 0.05).'
        else:
            aux = '\033[1m' + ' is not ' + '\033[0m'
            aux = colored(aux, 'blue')
            hipothesis ='H0: The difference' + aux + 'statistically significant (at significance level: 0.05).'
        print('For feature ' + column + ' - ' + hipothesis)
        #print(u_statistic)
```

In [ ]:
```
types = ['G1','S_G2']
columns = ['Entropy',
            'Circularity',
            'Solidity',
            'Eccentricity',
            'Perimeter',
            'Area',
            'Mean Intensity',
            'Variance',
            'Skewness',
            'Kurtosis',
            'Invariant Uniformity',
            'GLCM Invariant Entropy',
            'Invariant Correlation',
            'Invariant Dissimilarity',
            'Invariant Contrast',
            'Invariant Homogeneity',
            'Invariant Energy',
            'Equivalent Diameter',
            'Centroid Divergence']


print ('---------------- PAIRED T TEST ----------------')
plt_hist_and_stats_ttest_paired(columns, types)
```