

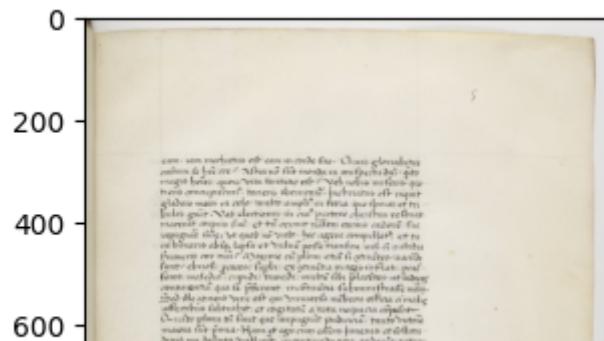
```
%load_ext autoreload
%autoreload 2

The autoreload extension is already loaded. To reload it, use:
%reload_ext autoreload

import cv2
import matplotlib.pyplot as plt
import numpy as np
from sklearn.metrics import mean_squared_error, mean_absolute_error

from functions.utils import *
from functions.cart import *
from functions.crop import *
from functions.analyse import *

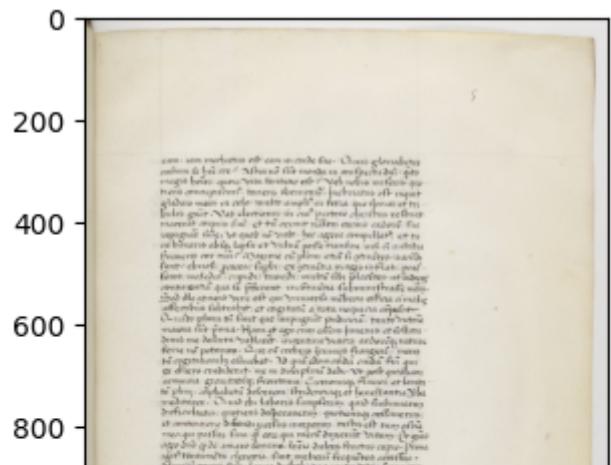
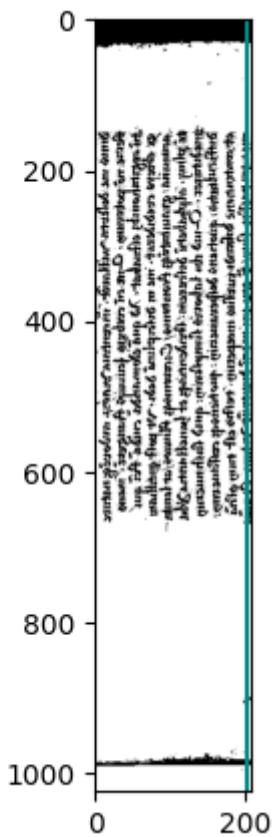
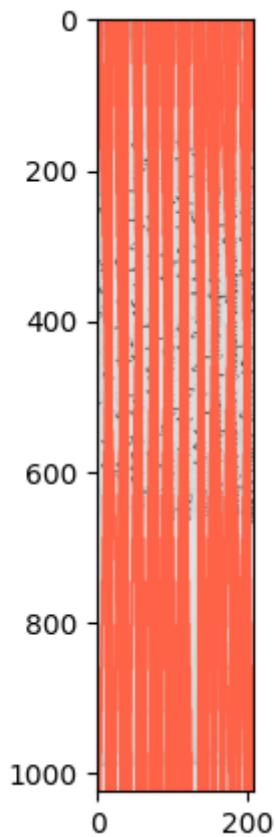
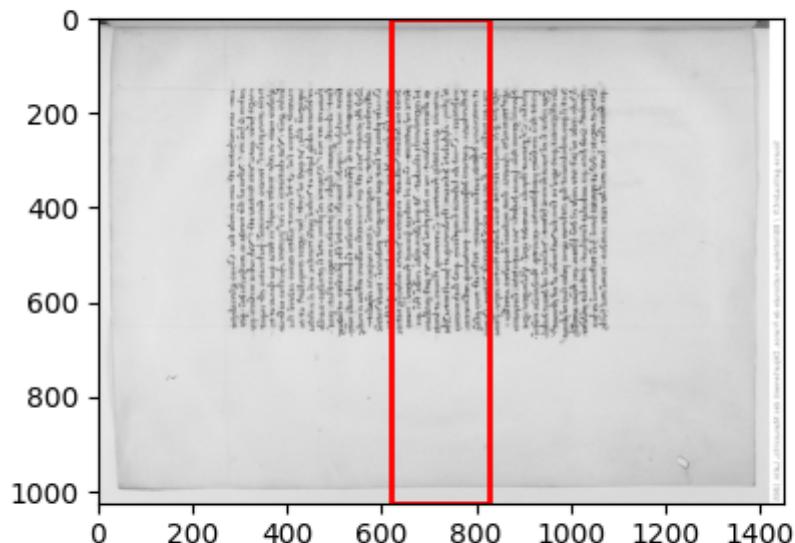
image = cv2.imread("../ImagesCodicologie/Corpus de jeu/Emilie Cottreau-Gabillet/BnF Lat 1865, f. 5r.jpeg")
image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
plt.imshow(image)
plt.show()
```

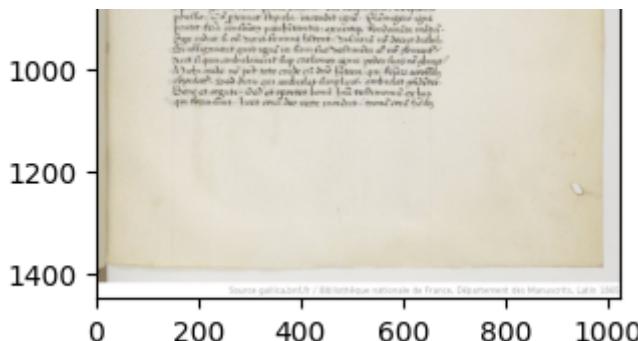


```
imgs = reframe(image)
```

69 candidate lines amongst 69 lines found

rmax = 0.3779296875



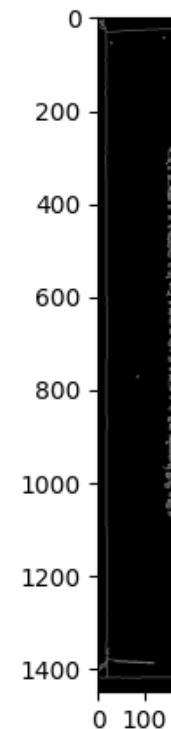
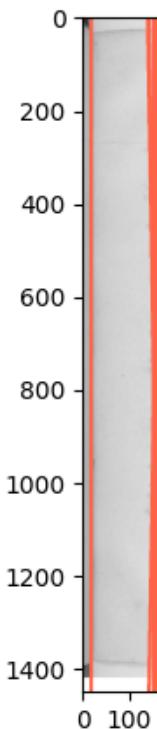
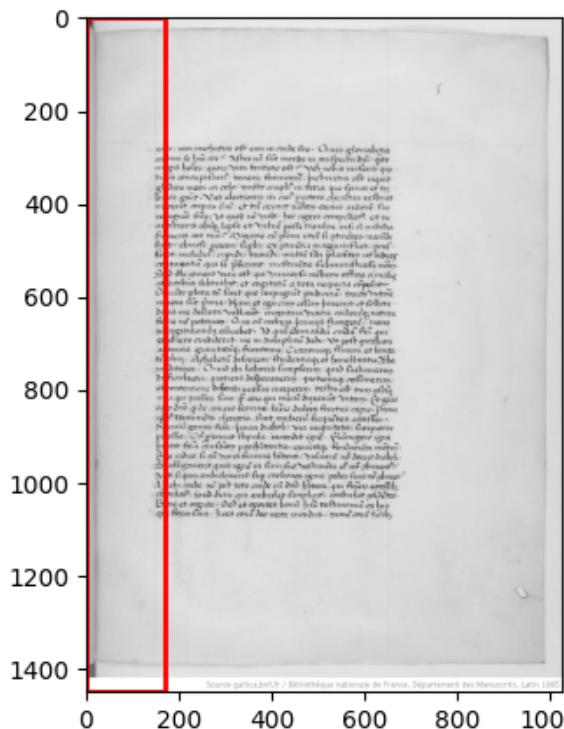


-----1-----

19 candidate lines amongst 19 lines found

rmax = 0.9793103448275862

r = 0.9793103448275862 with split at 18 with risk = 80261121.96904385



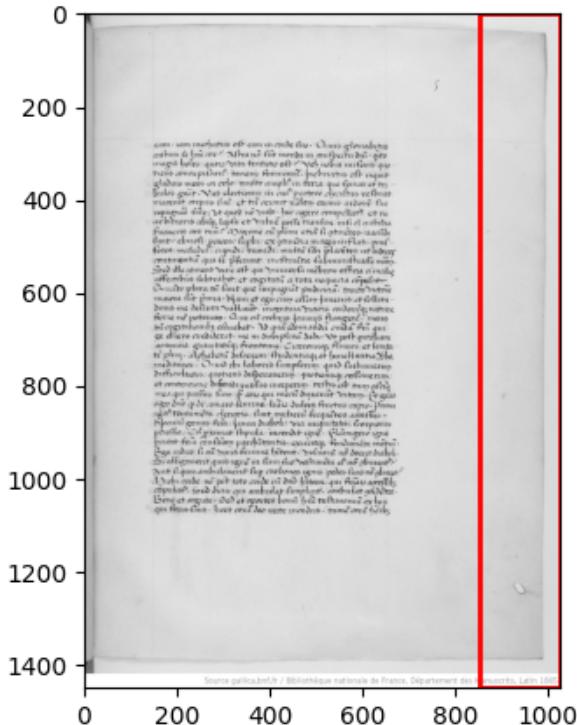
18

-----r-----

5 candidate lines amongst 5 lines found

rmax = 0.92

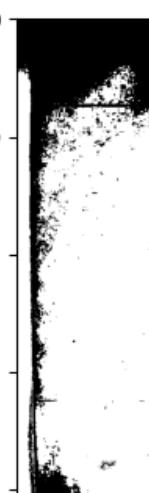
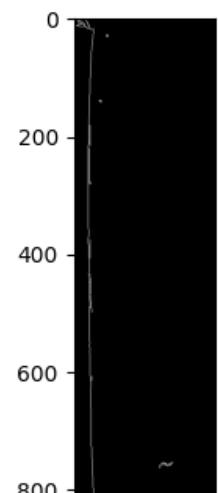
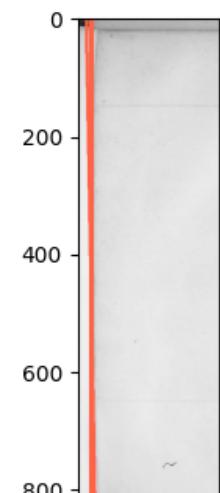
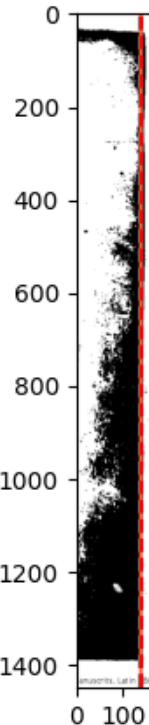
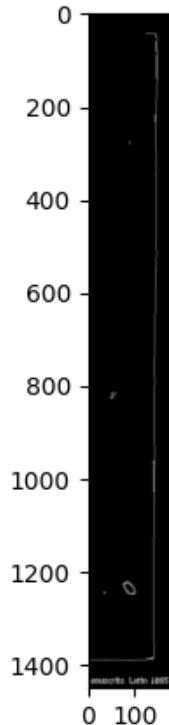
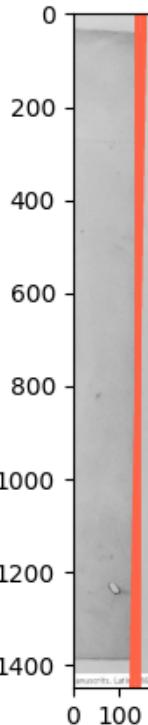
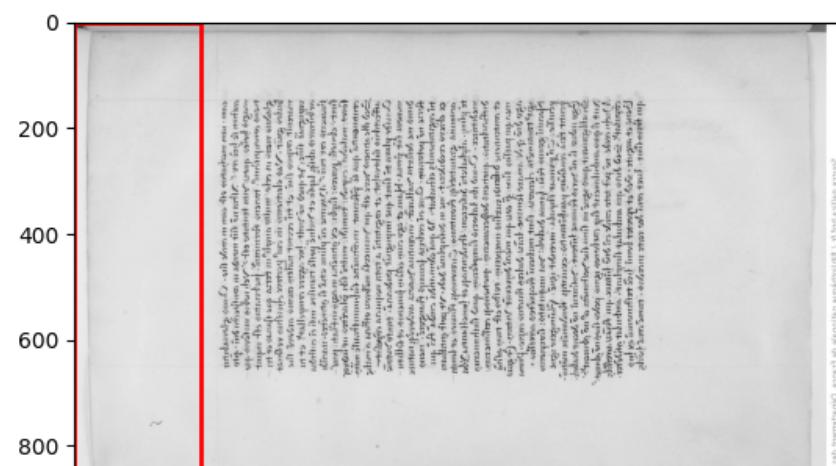
r = 0.7206896551724138 with split at 989 with risk = 20636324.276266098

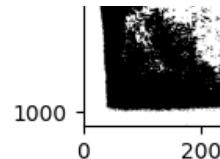
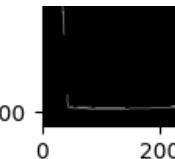
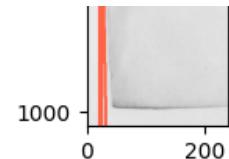
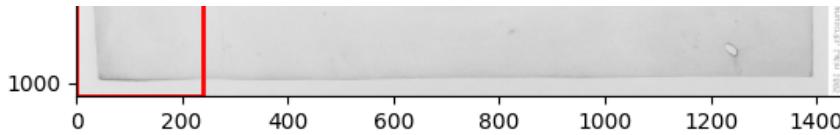


989

-----U-----

3 candidate lines amongst 3 lines found

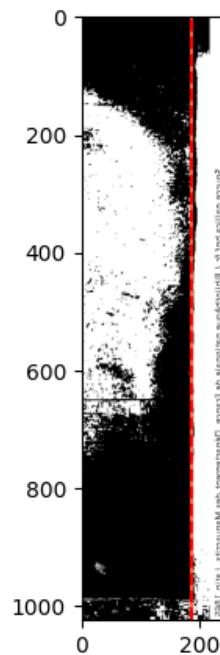
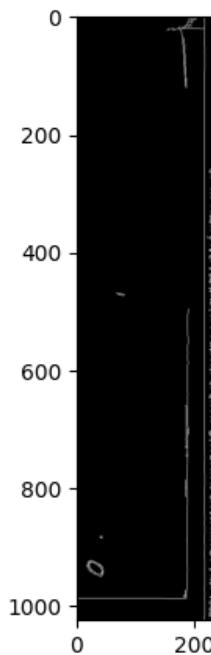
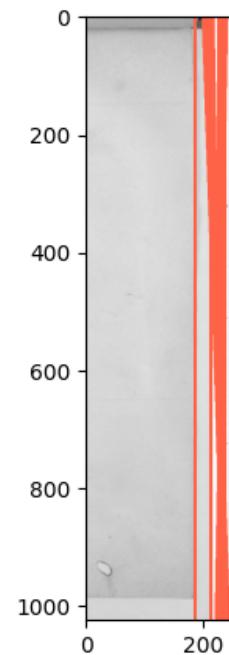
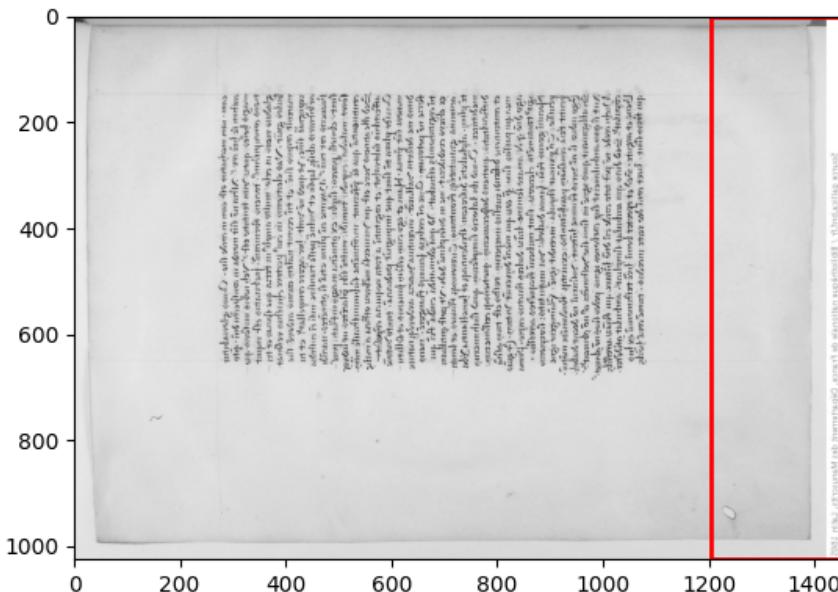
 $r_{max} = 0.6455078125$ $r = 0.0$ with split at -1 with risk = -1



0

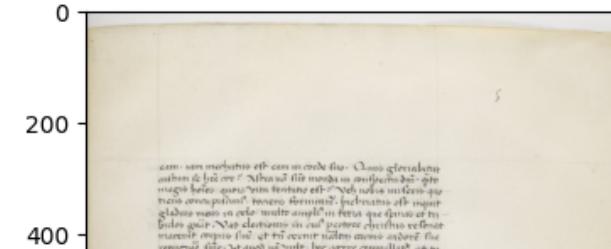
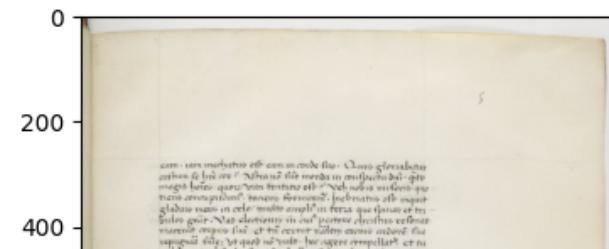
-----d-----

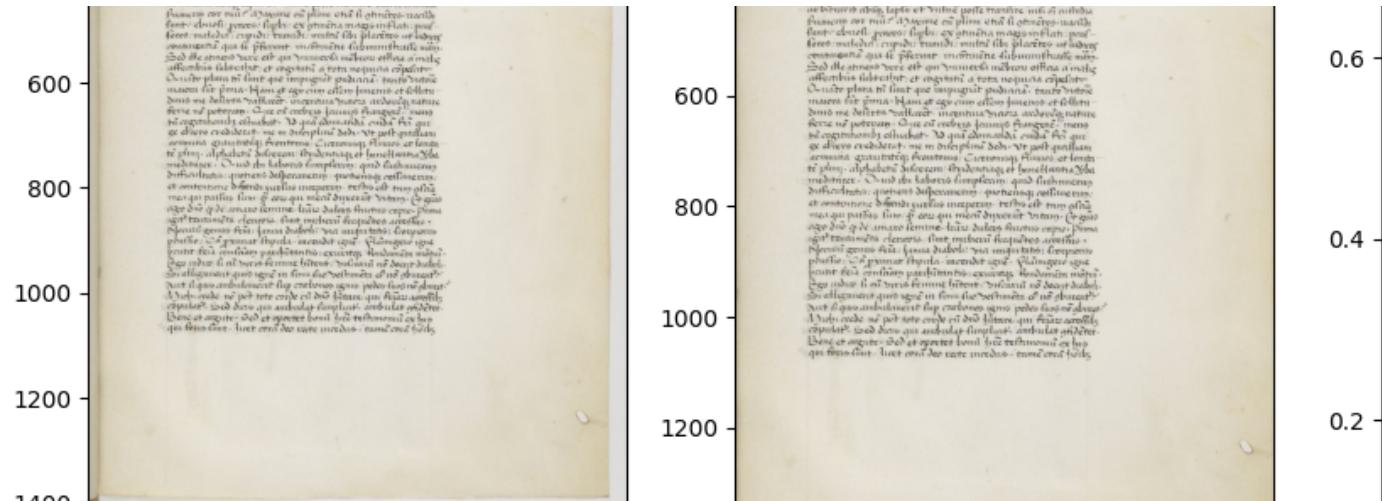
17 candidate lines amongst 18 lines found

 $r_{\max} = 0.9287109375$ $r = 0.9287109375$ with split at 1391 with risk = 52685248.75181506

1391

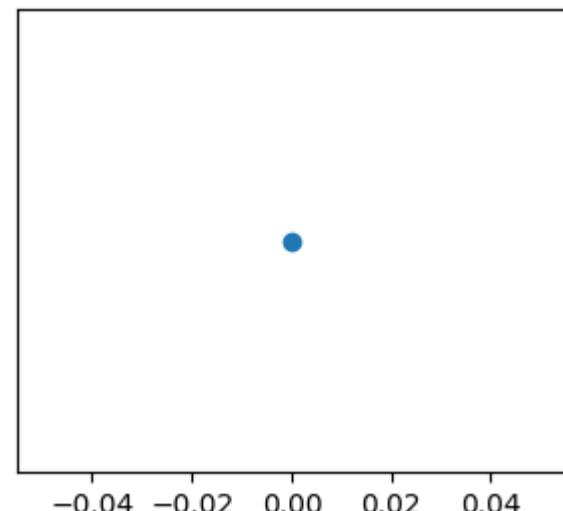
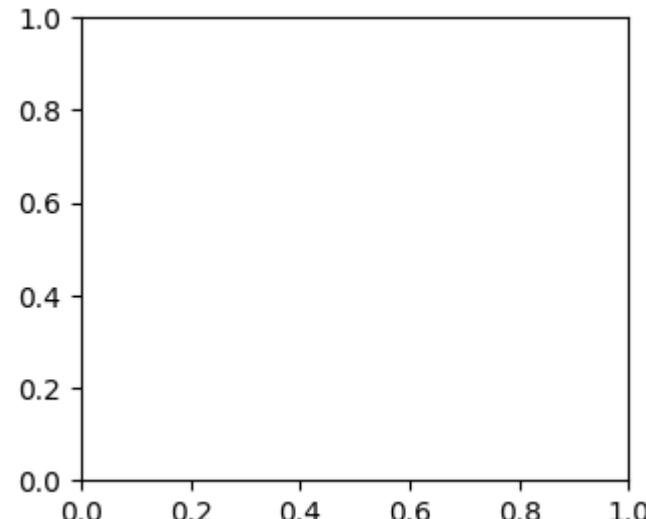
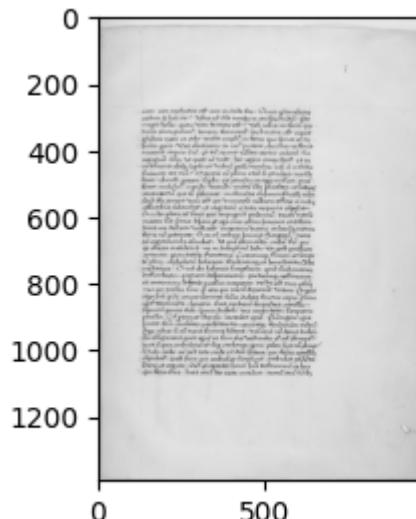
[18, 989, 0, 1391]





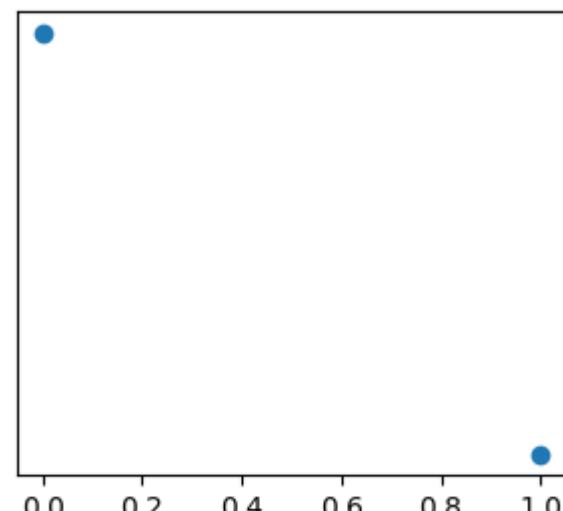
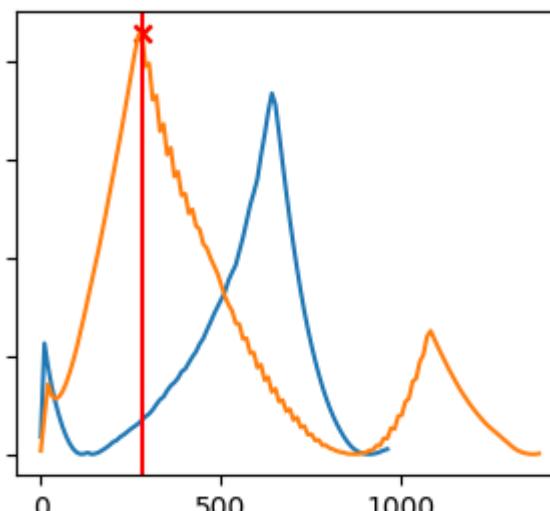
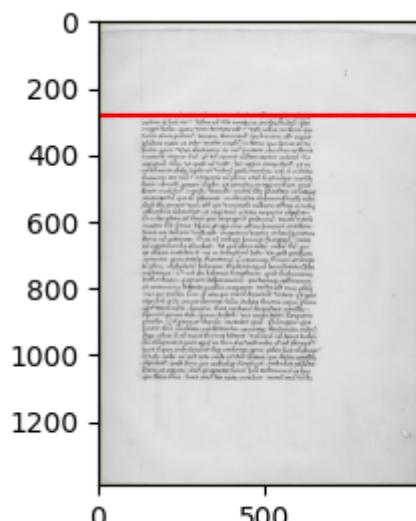
```
allres = []
for im in imgs:
    res = cart_regression_bi_simul(im,mean_squared_error,max_depth=10,rate=10)
    print(res)
    show_areas(im,res)
    allres.append(res)
```

DEPTH = 10



---> new split at 281

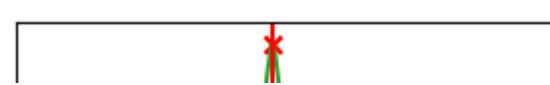
DEPTH = 9

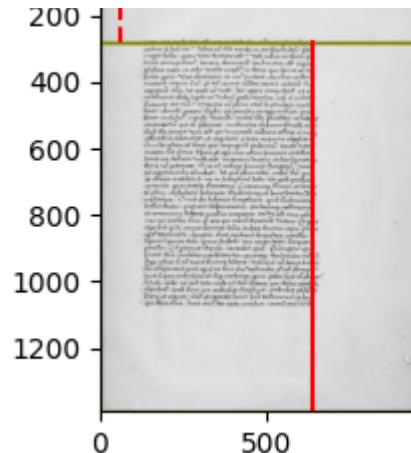


diff = 42.87992795044488

---> new split at 641

DEPTH = 8

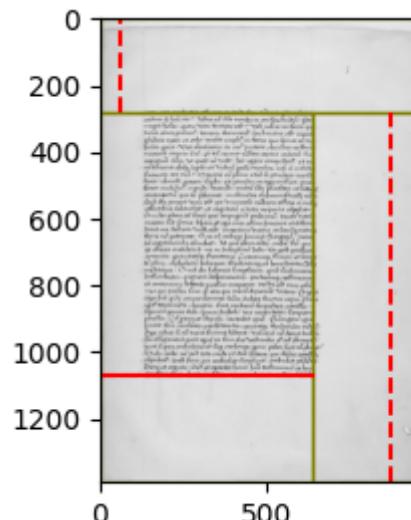
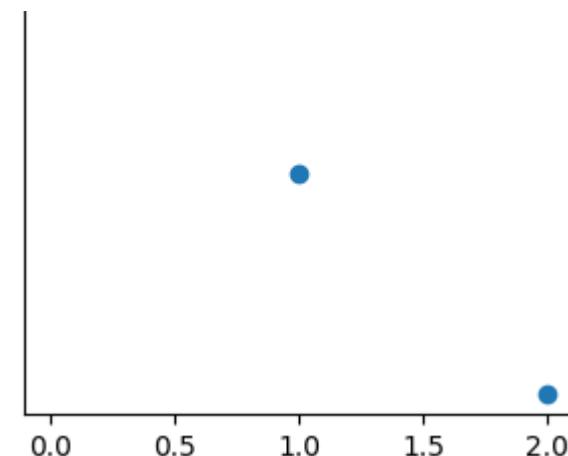
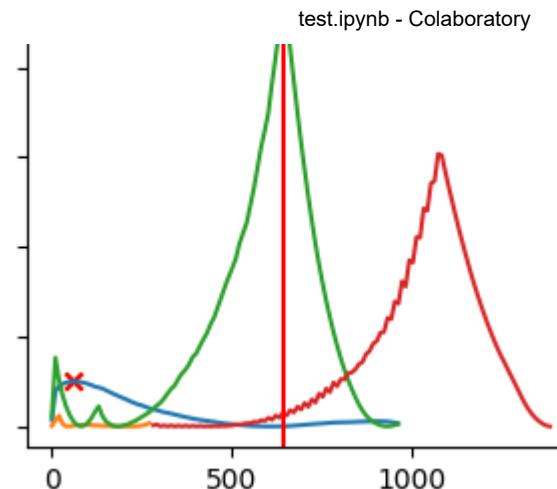




diff = 46.86488214114013

---> new split at 1072

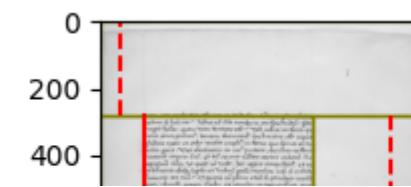
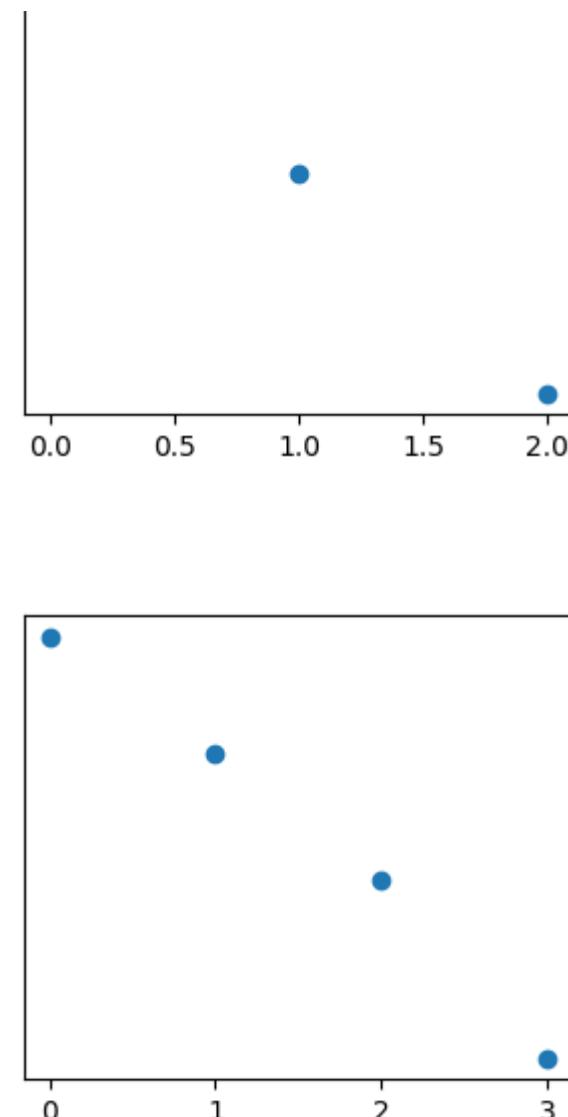
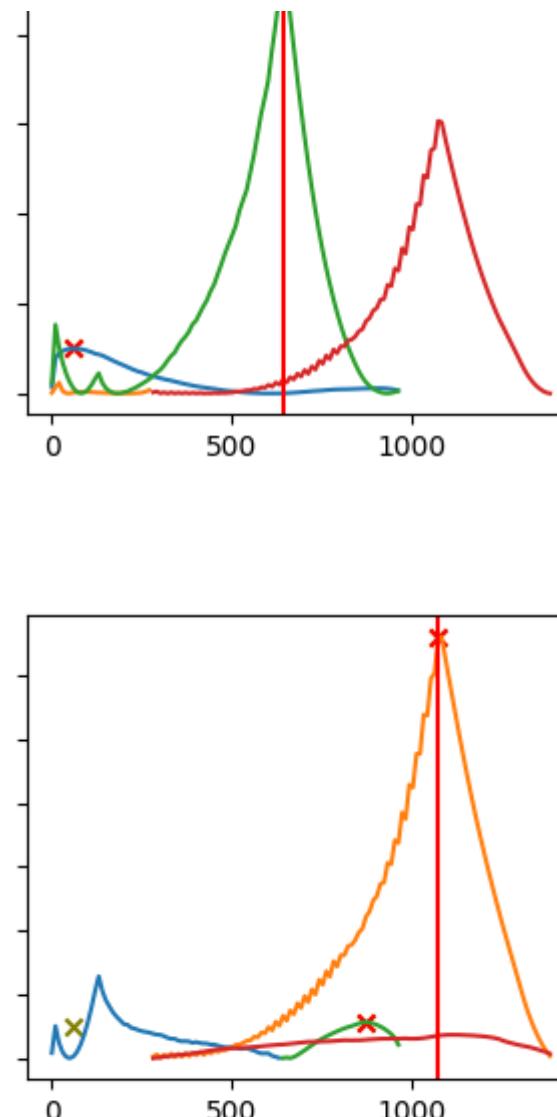
DEPTH = 7

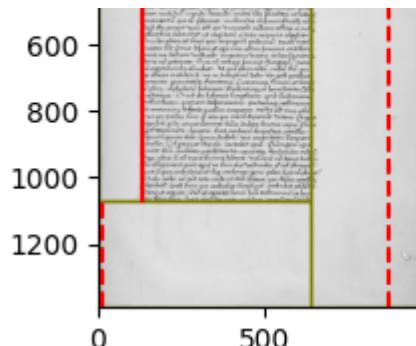


diff = 66.09388599351303

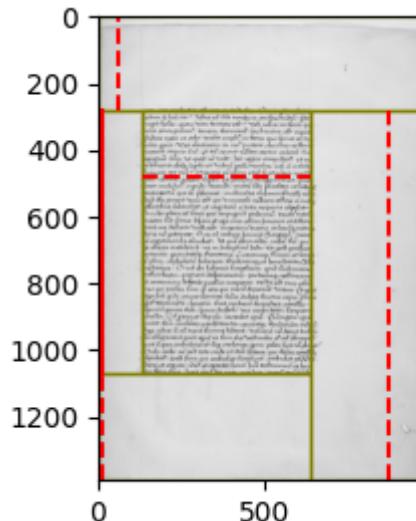
---> new split at 131

DEPTH = 6

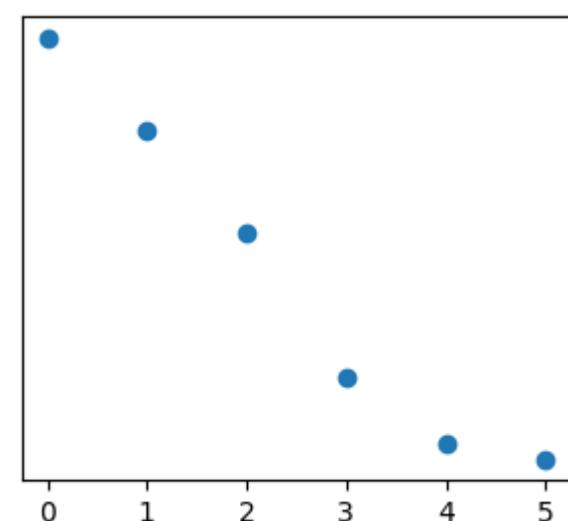
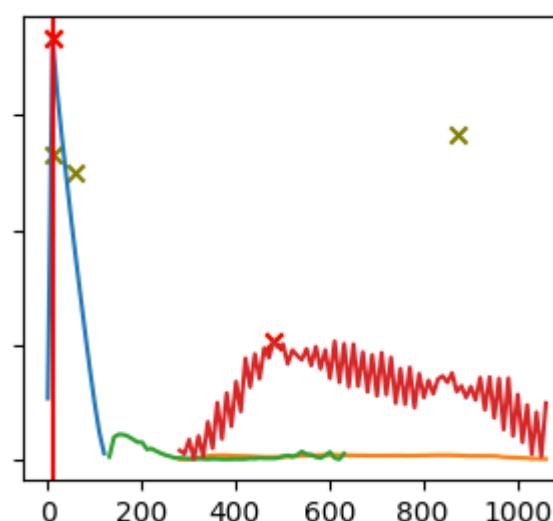
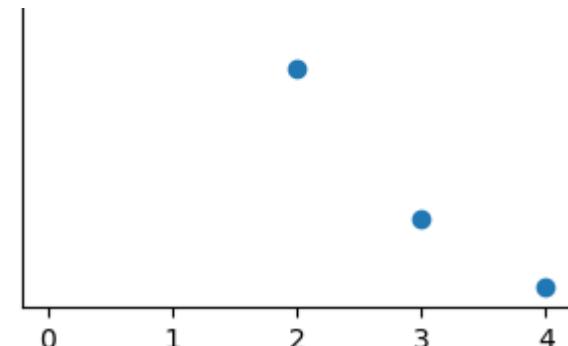
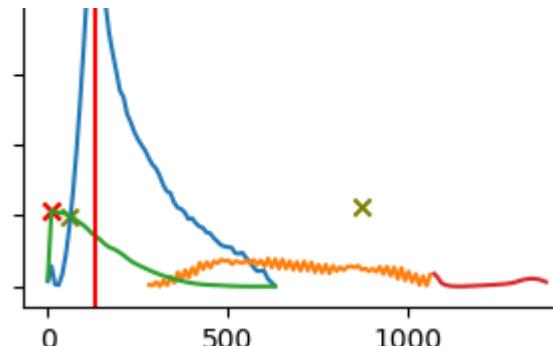


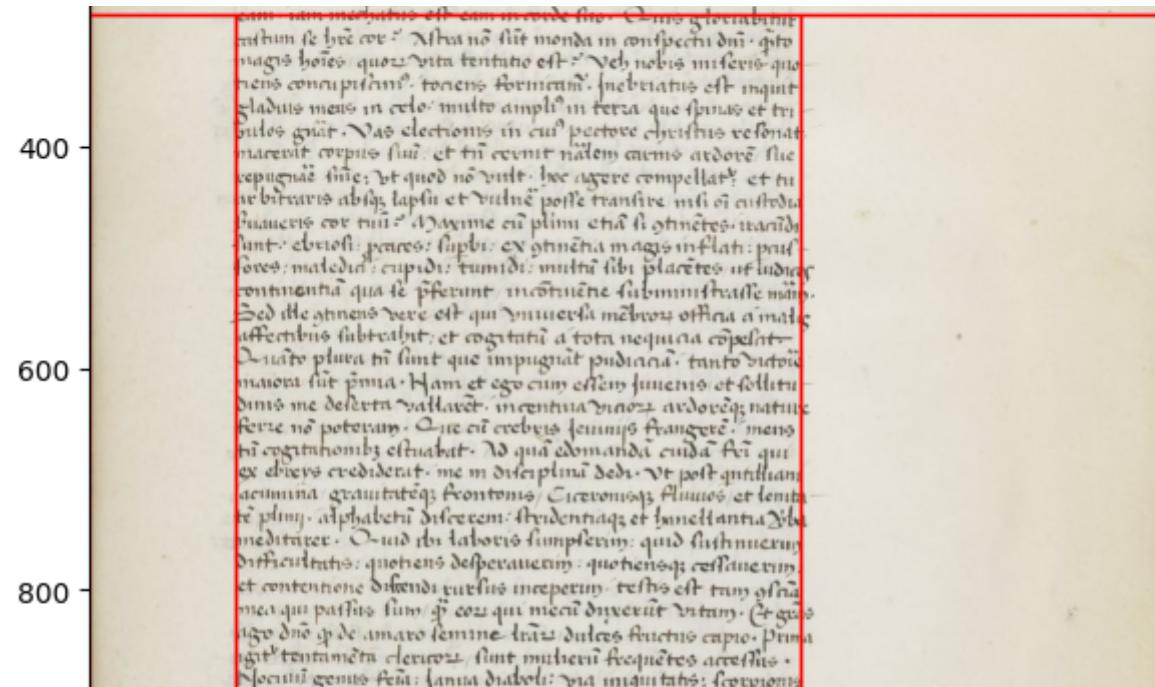


```
diff = 29.69565309885445
---> new split at 11
DEPTH = 5
```

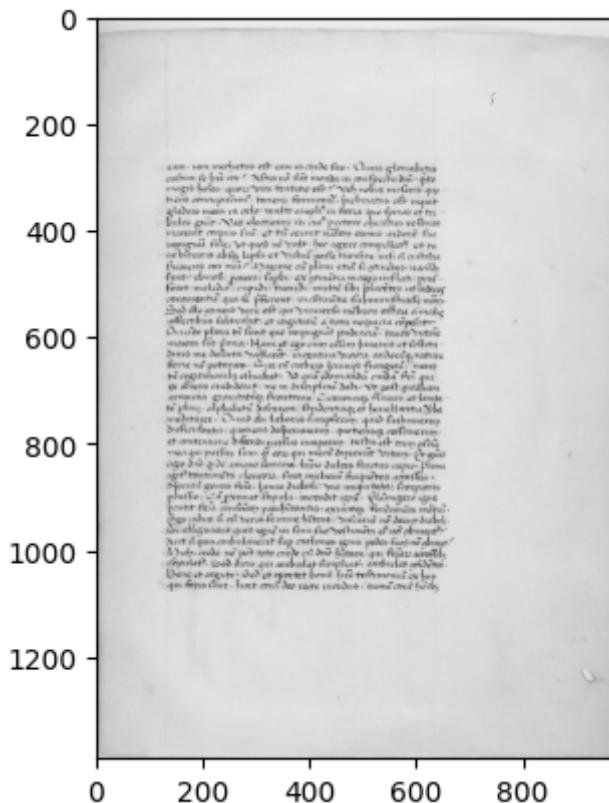


```
diff = 7.354660980473
[[0, 0, 971, 281], [0, 281, 131, 1072], [131, 281, 641, 1072], [0, 1072, 641, 1391], [641, 281, 971, 1391]]
```

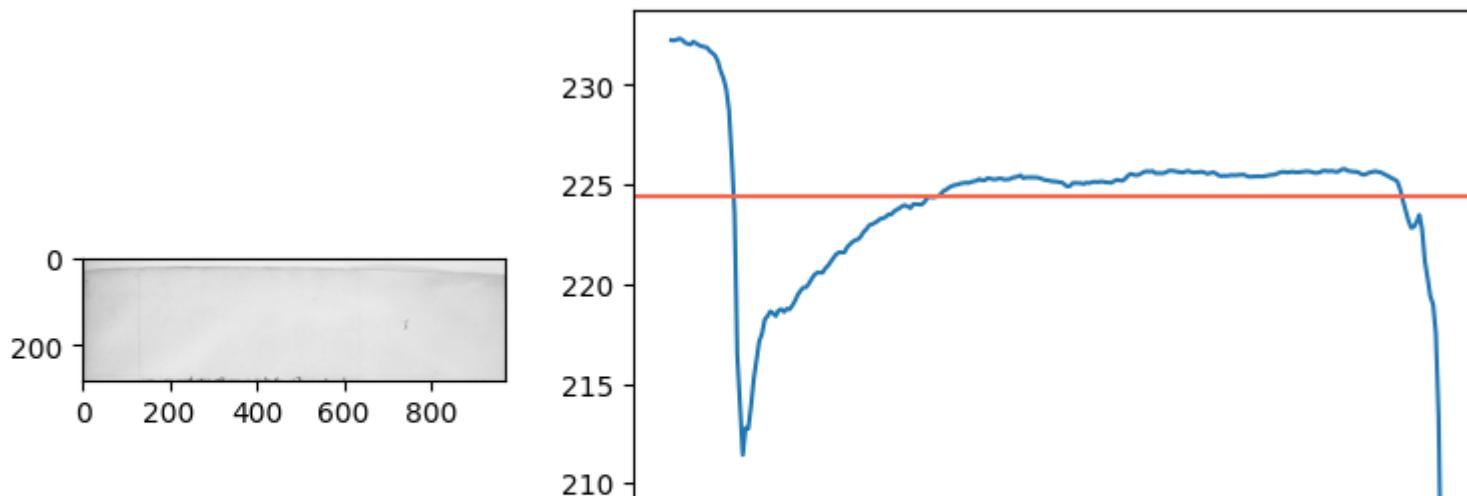


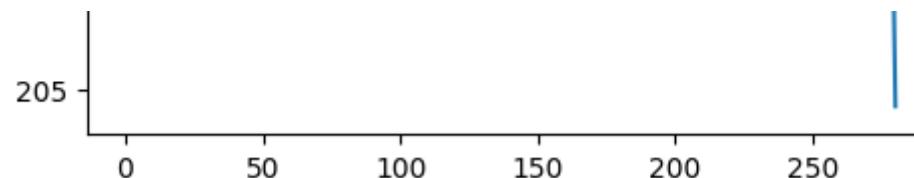


check_text(allres, imgs)



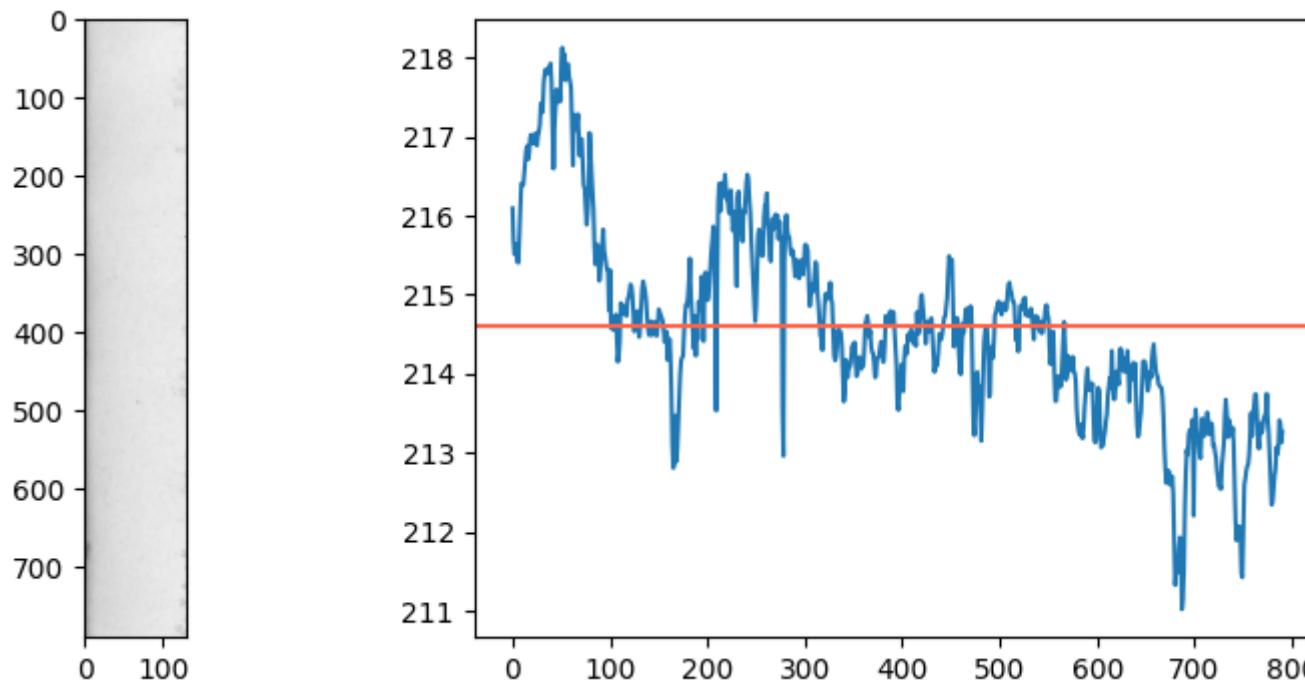
Analyzing area 0 0 971 281





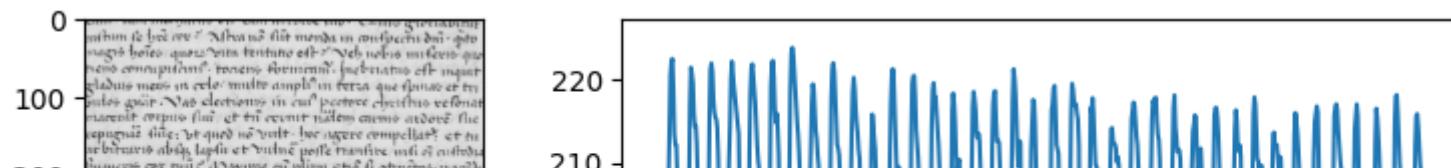
[23, 1, 169] [71, 2, 15]
29.93697082575694 74.55348564770276
---> not text

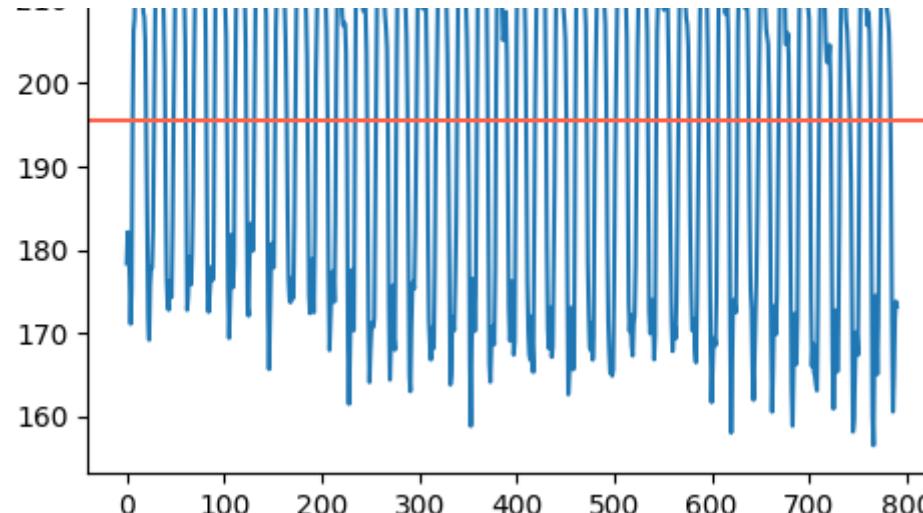
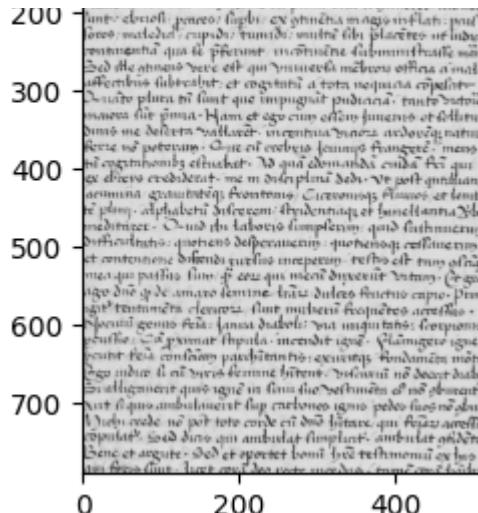
Analyzing area 0 281 131 1072



[101, 1, 1, 14, 1, 2, 7, 2, 1, 6, 8, 1, 1, 5, 12, 67, 37, 9, 4, 9, 1, 6, 1, 2, 11, 4, 1, 8, 21, 1,
36.75883827541377 19.62369756628041
---> not text

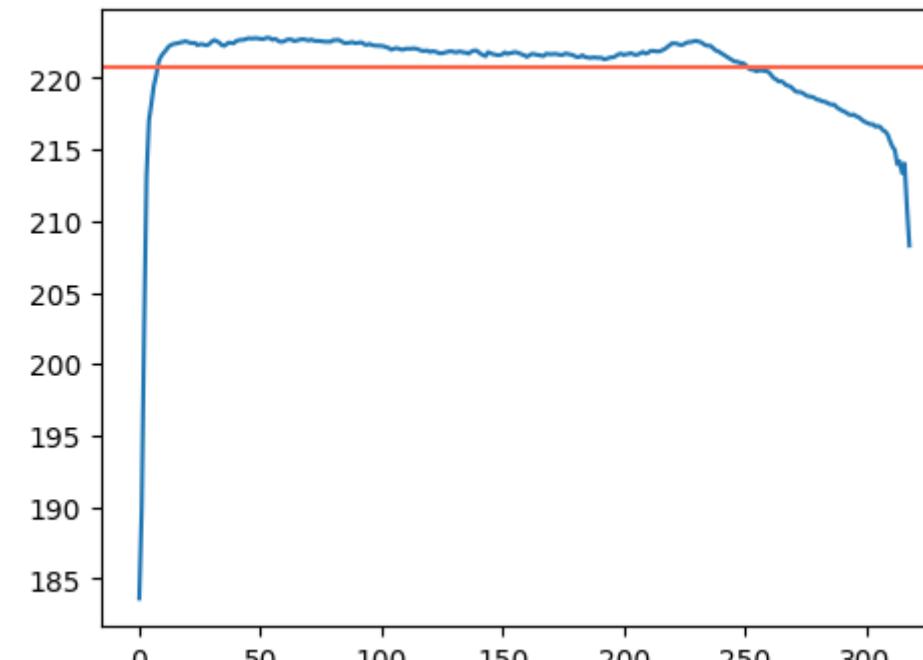
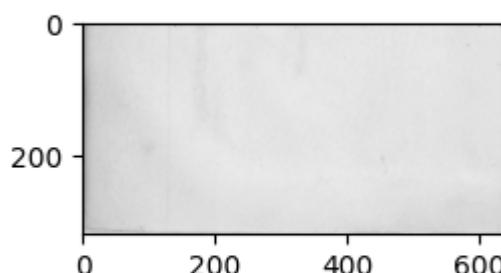
Analyzing area 131 281 641 1072





----->>> Contains 39 lines of text

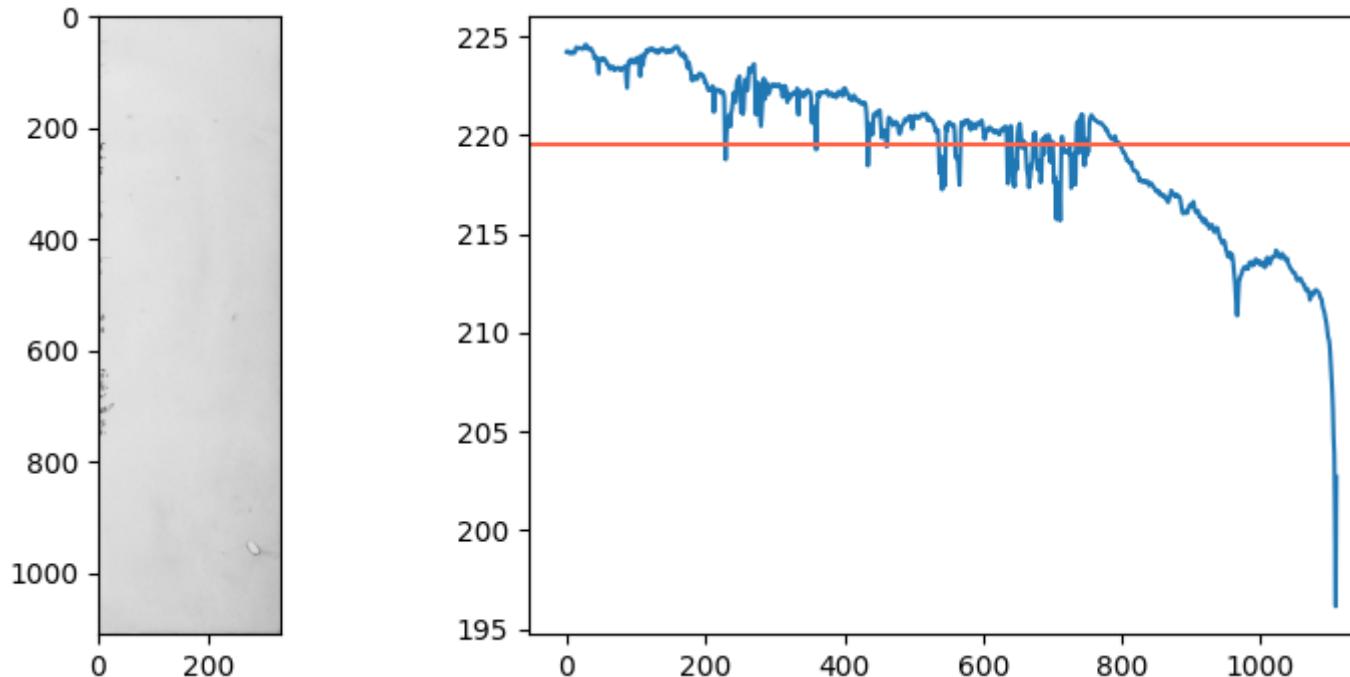
Analyzing area 0 1072 641 1391



〔8. 681 〔243〕

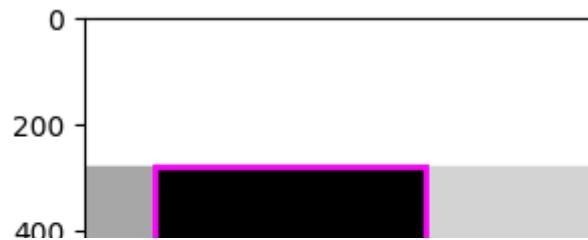
L-- -- L--
30.0 0.0
---> not text

Analyzing area 641 281 971 1391

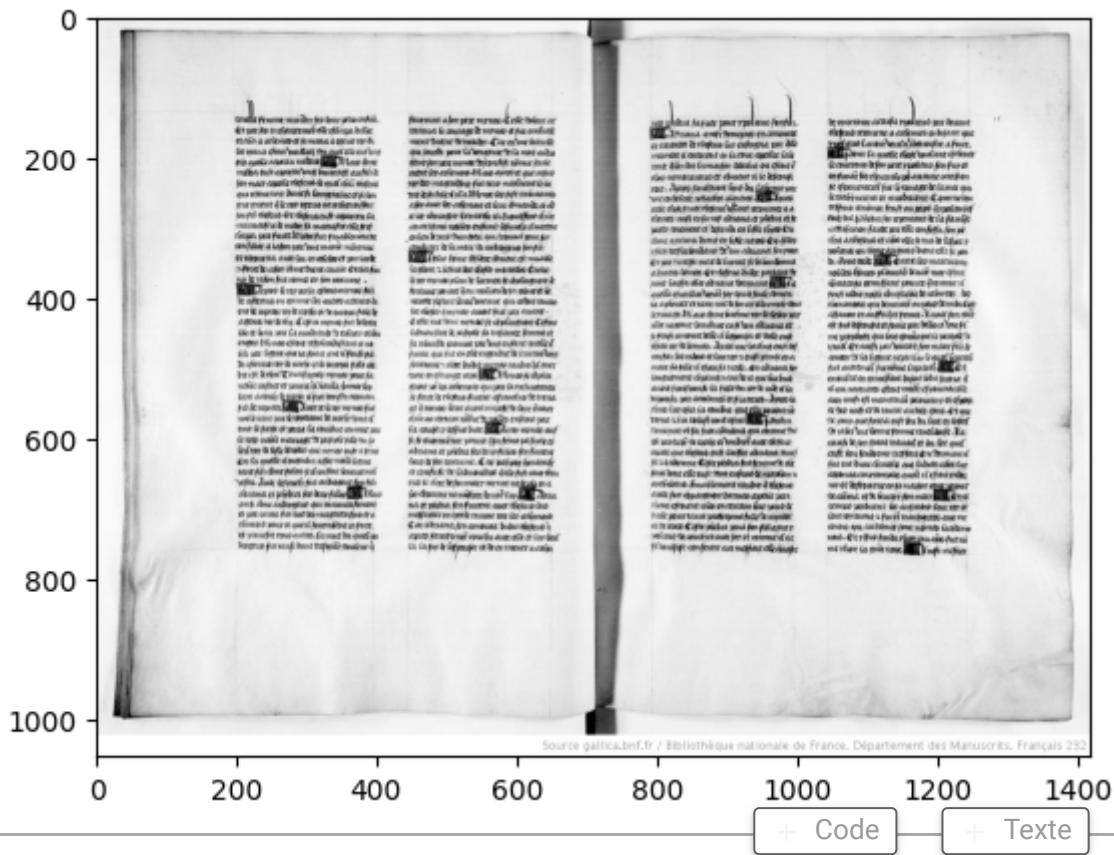


[229, 129, 73, 25, 74, 13, 67, 4, 6, 3, 5, 1, 6, 2, 3, 3, 6, 1, 44] [1, 2, 2, 1, 11, 8, 3, 8, 2, 1:
68.78027160293026 57.0958275008388

---> not text

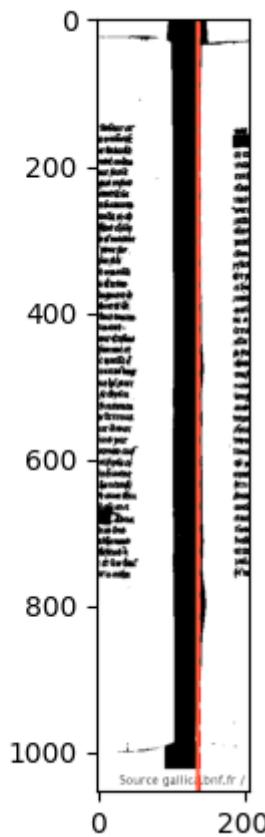
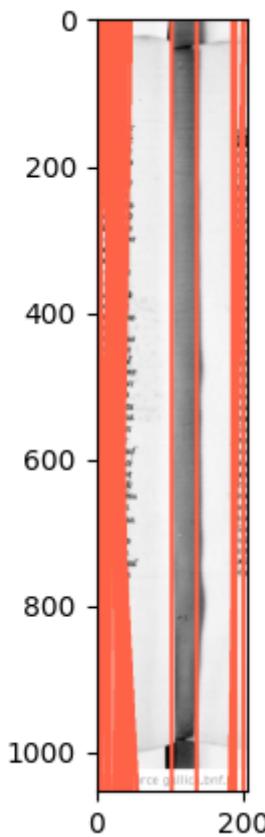
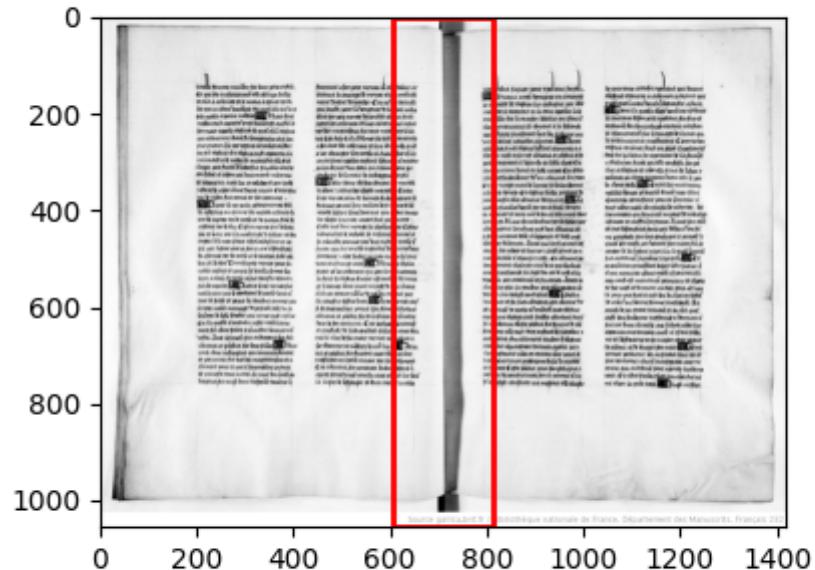


```
image = cv2.imread("../ImagesCodicologie/Corpus de jeu/Emilie Cottreau-Gabillet/BnF Fr 232, ff. 10v-11r.jpeg")
image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
plt.imshow(image)
plt.show()
```

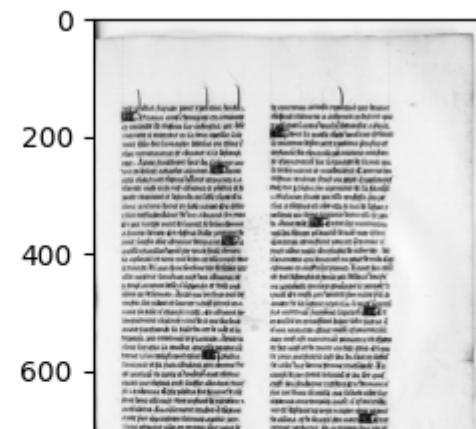
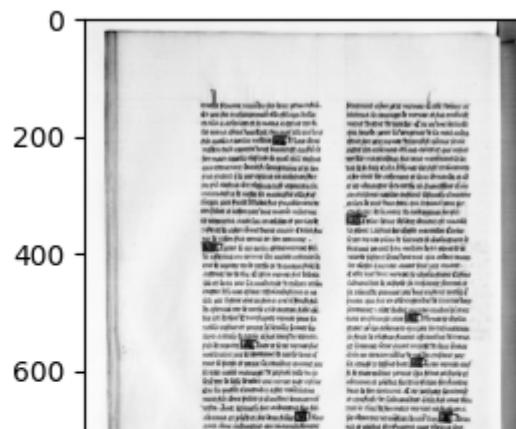


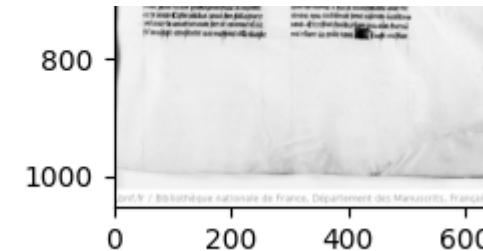
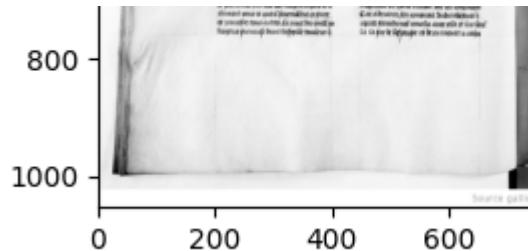
```
imgs = reframe(image)
```

126 candidate lines amongst 126 lines found
 $r = 0.9146110056925996$ with split at 743



$r_{max} = 0.9146110056925996$



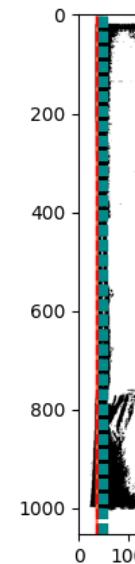
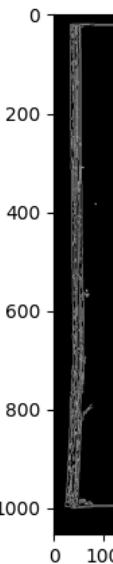
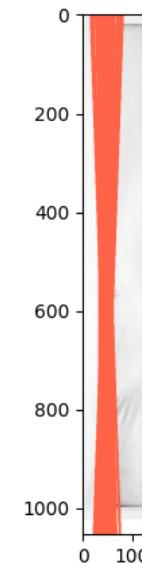
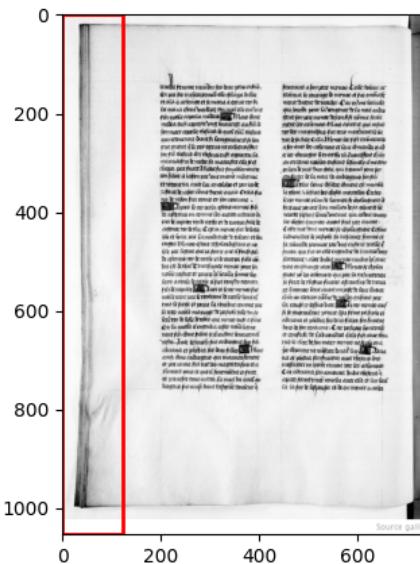


-----1-----

33 candidate lines amongst 52 lines found

rmax = 0.9364326375711575

r = 0.715370018975332 with split at 37 with risk = 155241227.32348263



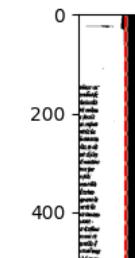
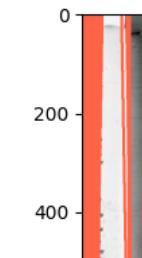
37

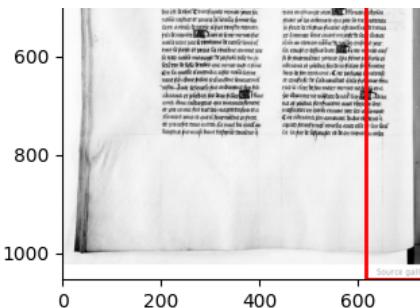
-----r-----

50 candidate lines amongst 67 lines found

rmax = 0.9791271347248577

r = 0.9791271347248577 with split at 711 with risk = 252644748.2644291

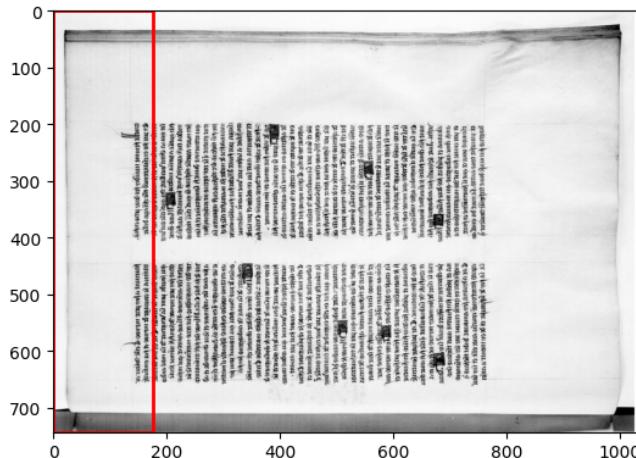




711

-----U-----

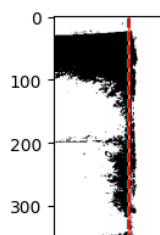
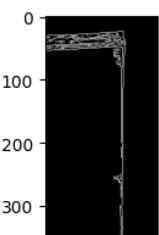
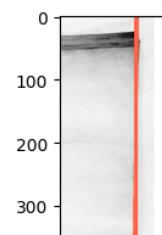
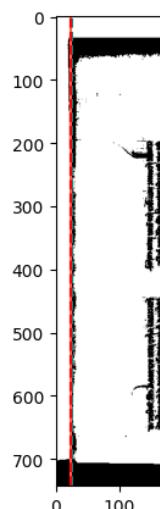
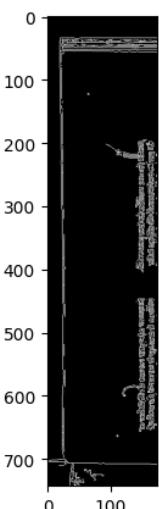
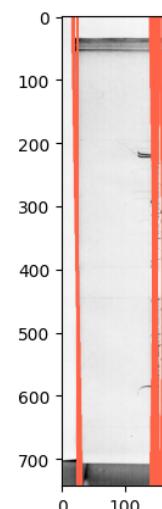
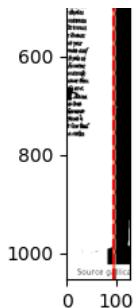
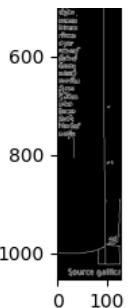
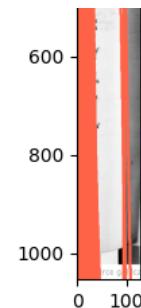
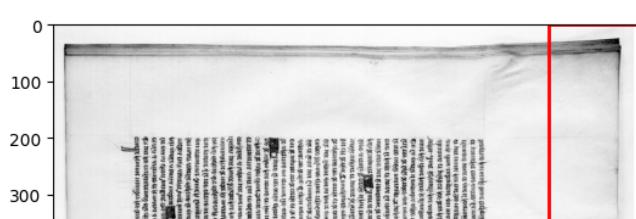
16 candidate lines amongst 20 lines found

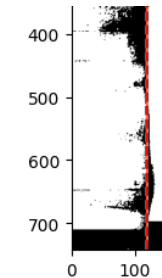
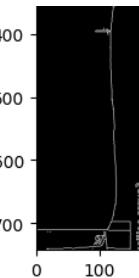
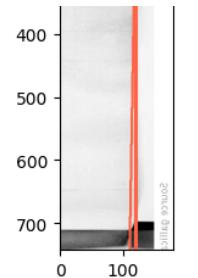
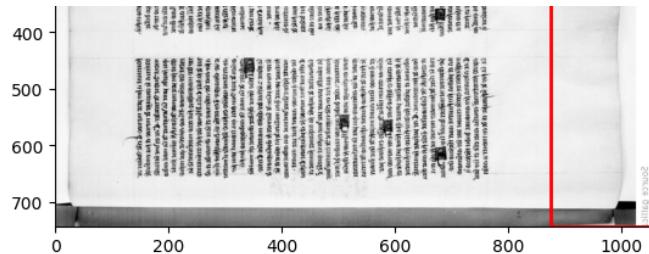
 $r_{max} = 0.7429340511440108$ $r = 0.7321668909825034$ with split at 24 with risk = 244803181.67098528

24

-----d-----

3 candidate lines amongst 3 lines found

 $r_{max} = 0.8088829071332436$ $r = 0.7711978465679676$ with split at 995 with risk = 192749336.61994454



995

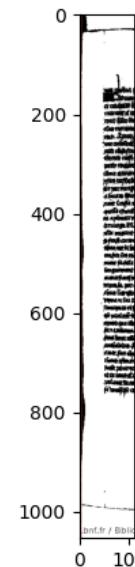
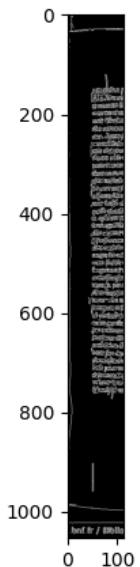
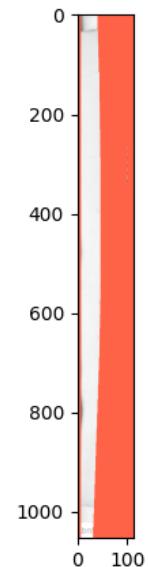
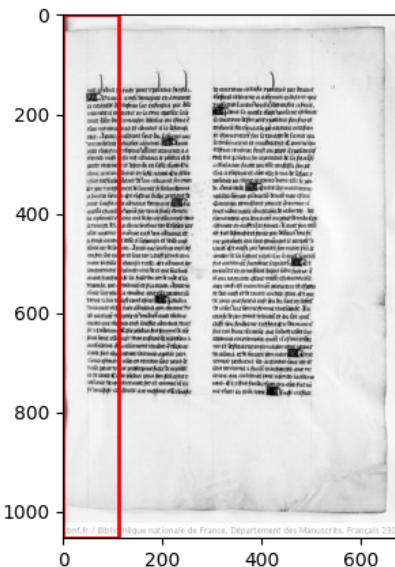
[37, 711, 24, 995]

-----1-----

78 candidate lines amongst 157 lines found

rmax = 0.9013282732447818

r = 0.9013282732447818 with split at 1 with risk = 301377964.99273473



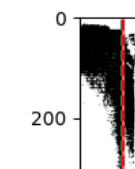
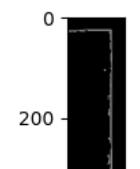
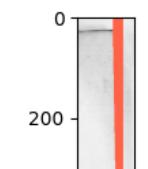
1

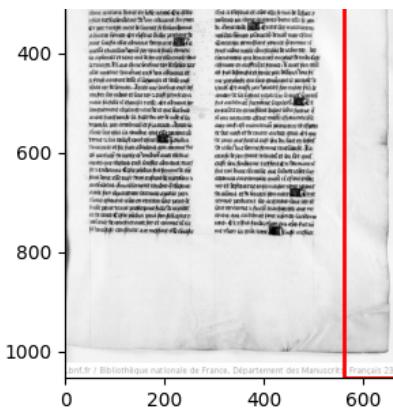
-----r-----

6 candidate lines amongst 6 lines found

rmax = 0.9335863377609108

r = 0.9127134724857685 with split at 647 with risk = 29844017.557092696

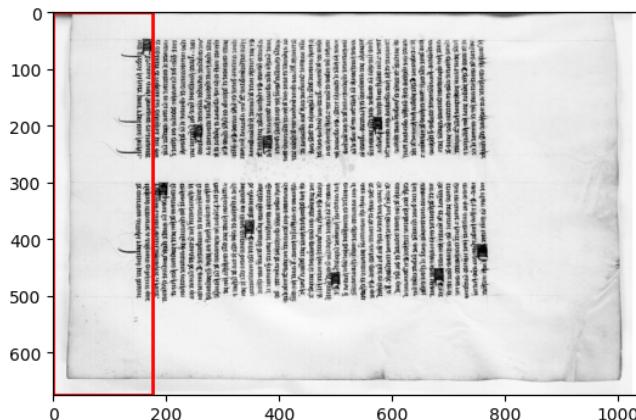




647

-----U-----

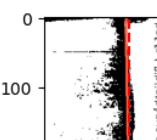
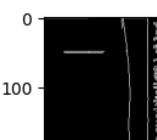
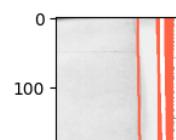
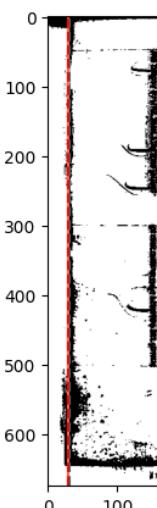
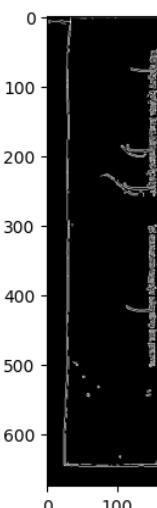
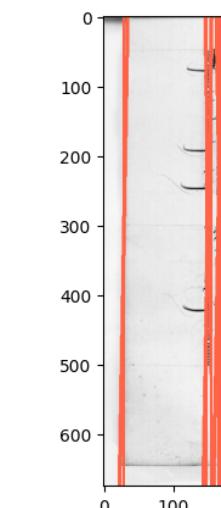
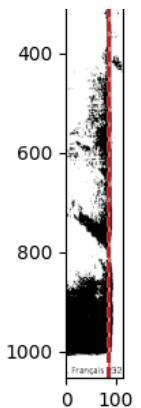
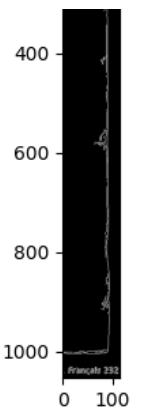
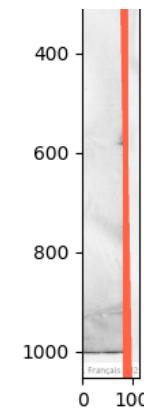
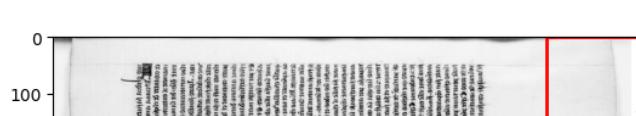
13 candidate lines amongst 15 lines found

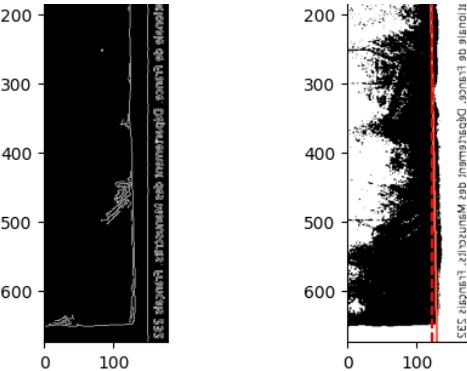
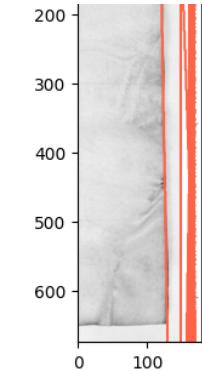
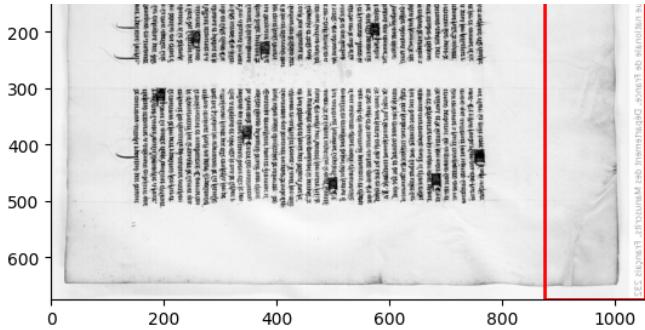
 $r_{max} = 0.8088888888888889$ $r = 0.8088888888888889$ with split at 29 with risk = 139841098.1885421

29

-----d-----

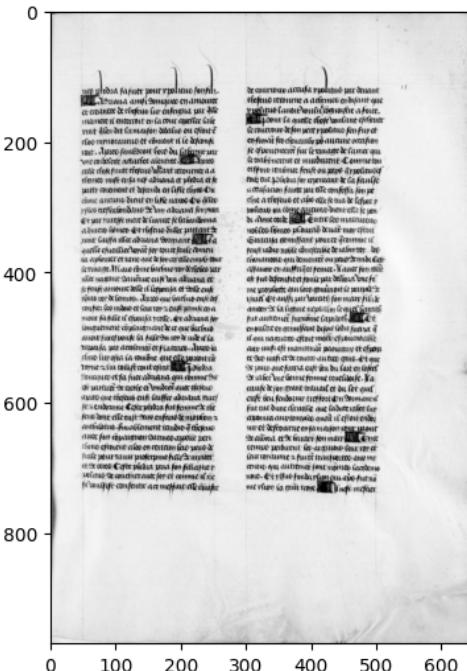
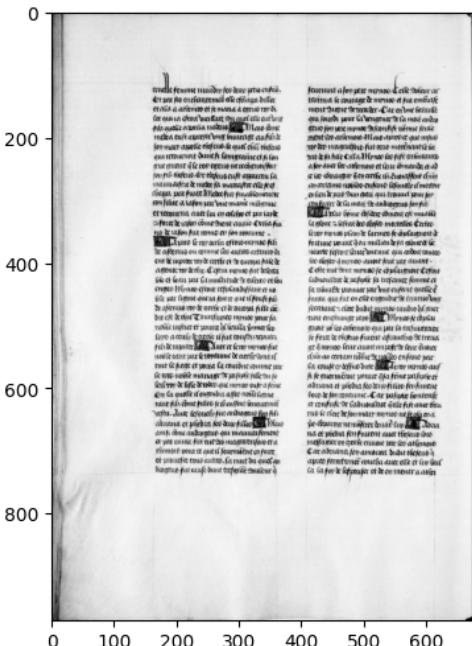
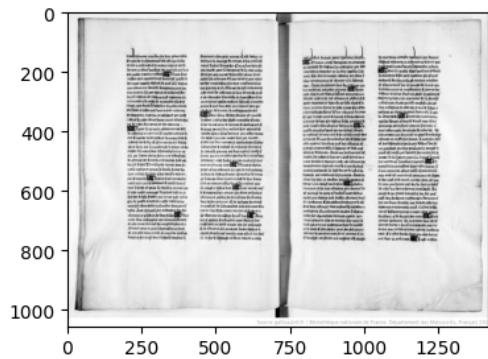
10 candidate lines amongst 10 lines found

 $r_{max} = 0.9081481481481481$ $r = 0.9081481481481481$ with split at 998 with risk = 40582010.98997111



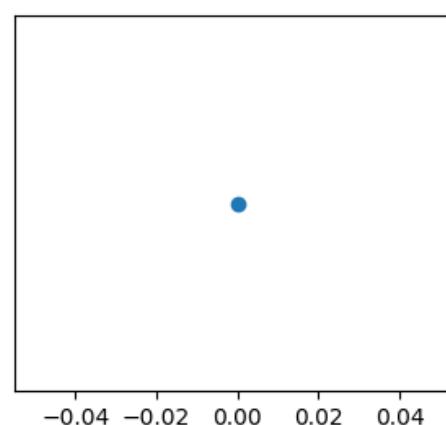
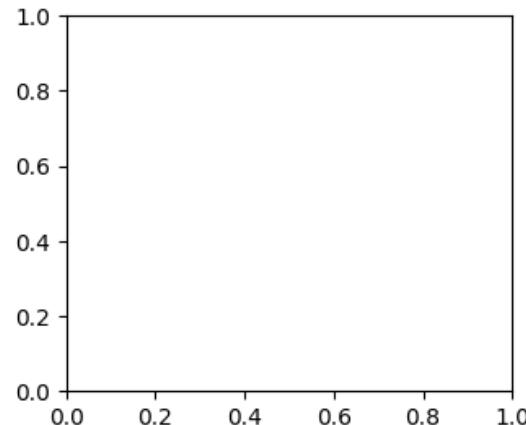
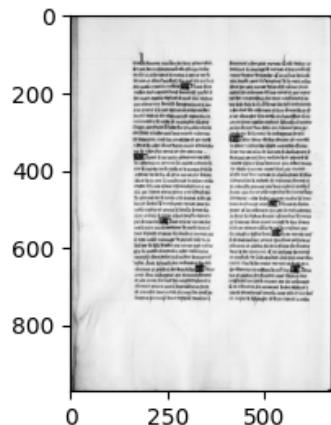
998

[1, 647, 29, 998]



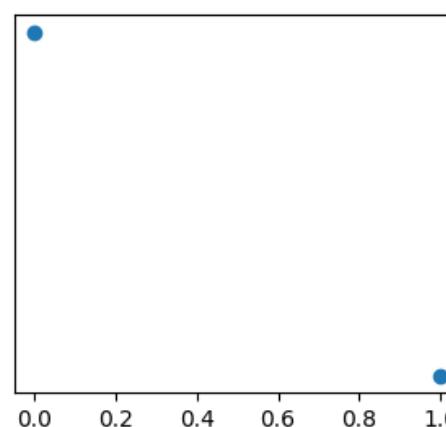
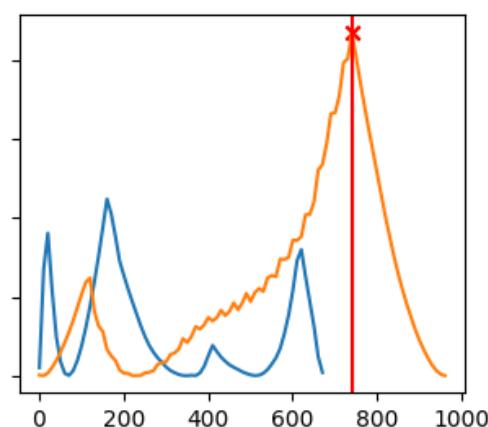
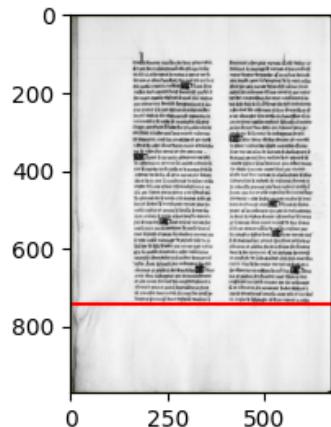

```
allres = []
for im in imgs:
    res = cart_regression_bi_simul(im,mean_squared_error,max_depth=10,rate=10)
    print(res)
    show_areas(im,res)
    allres.append(res)
```

DEPTH = 10



---> new split at 741

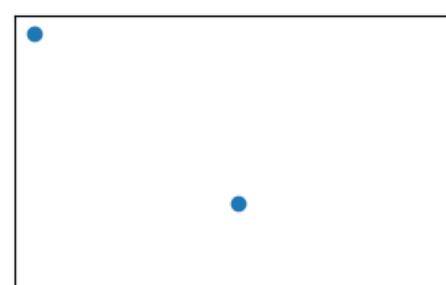
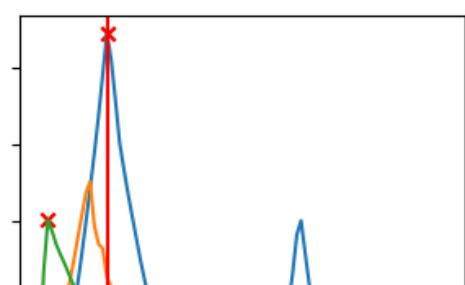
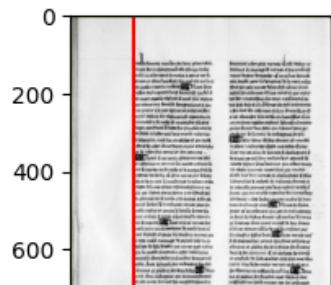
DEPTH = 9

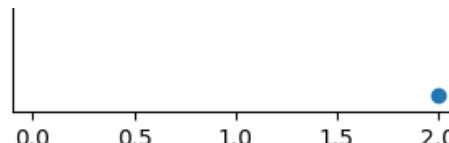
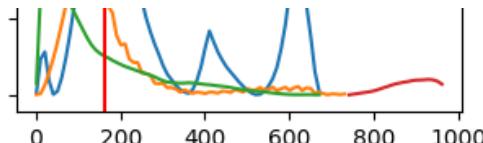
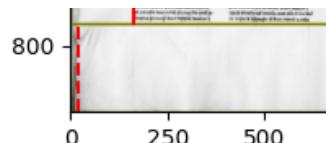


diff = 87.02801820011427

---> new split at 161

DEPTH = 8

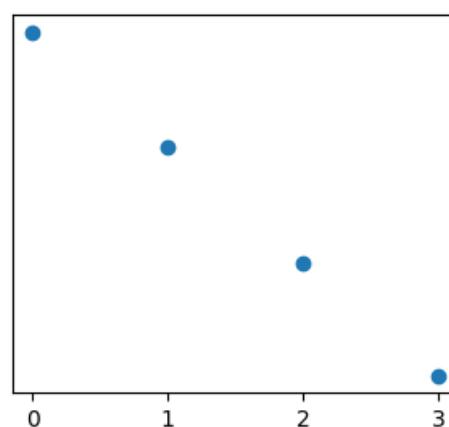
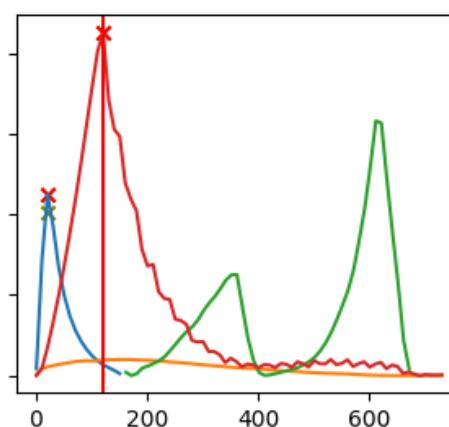
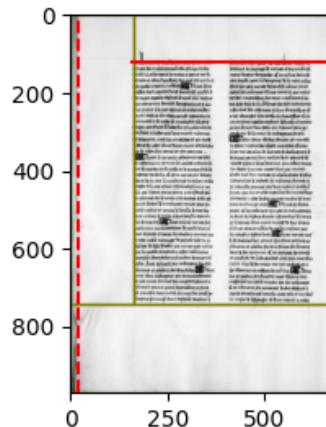




diff = 88.86382472235755

---> new split at 121

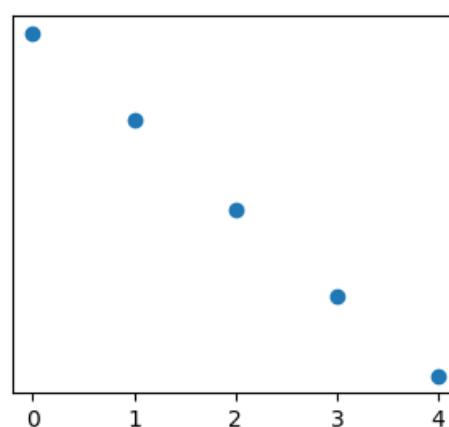
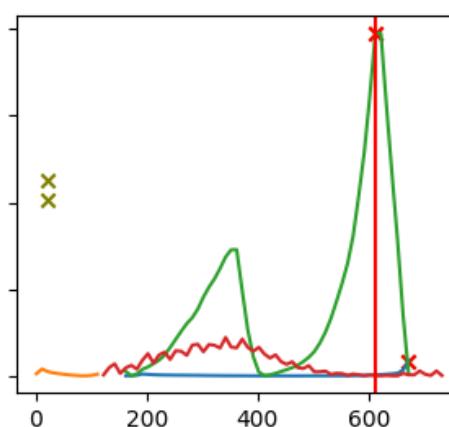
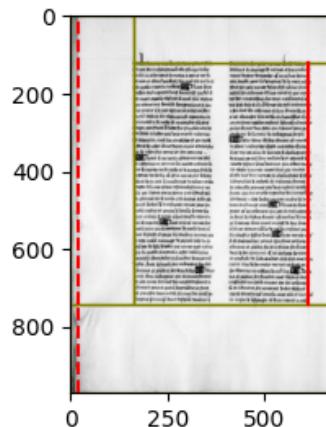
DEPTH = 7



diff = 85.22154777960168

---> new split at 612

DEPTH = 6

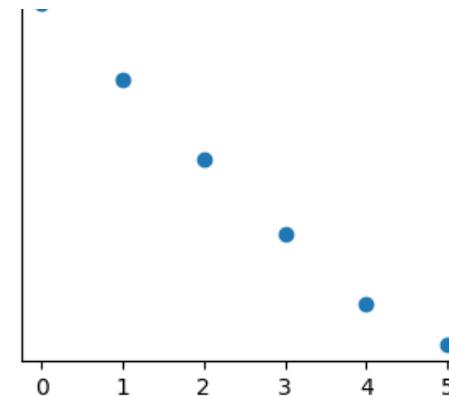
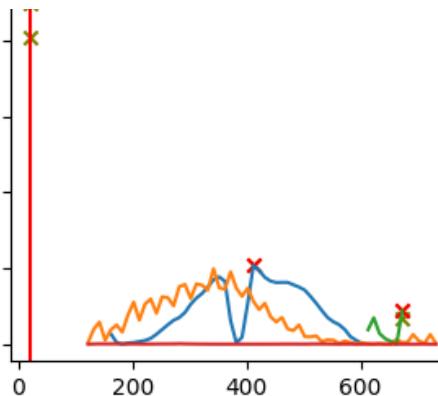
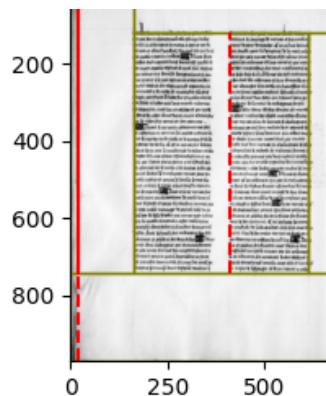


diff = 79.09582573012426

---> new split at 21

DEPTH = 5

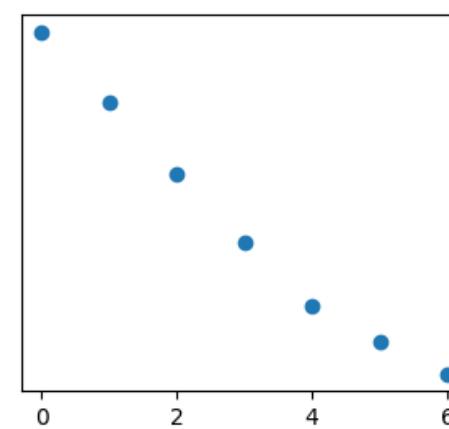
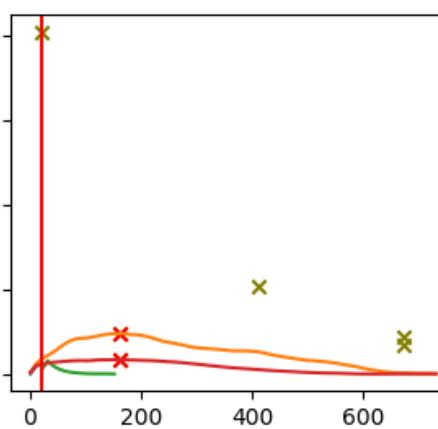
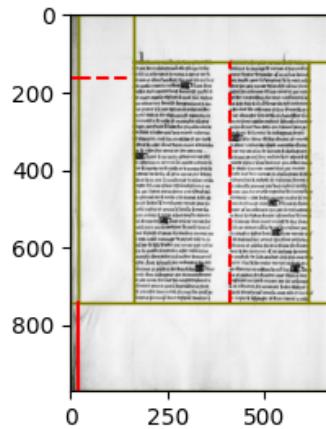




diff = 44.98233804809888

---> new split at 21

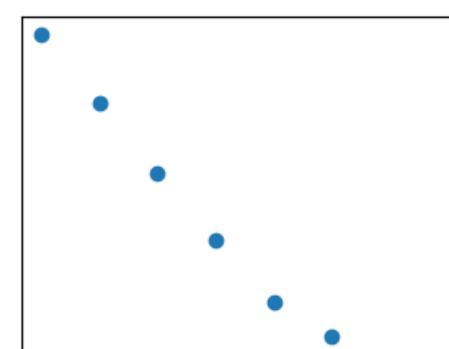
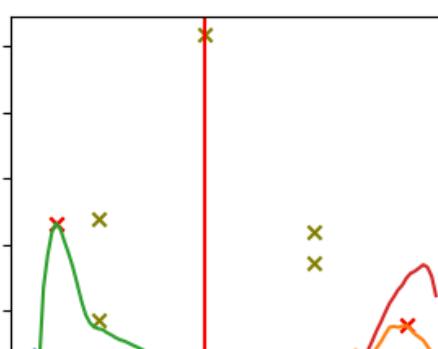
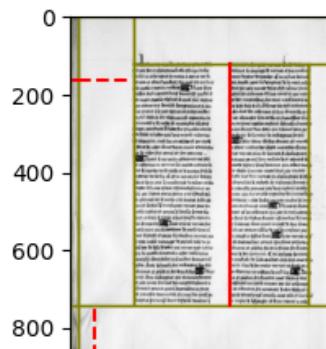
DEPTH = 4



diff = 40.41011461939706

---> new split at 412

DEPTH = 3

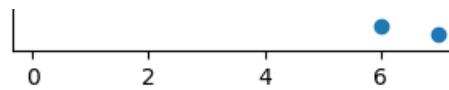
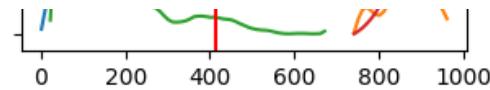
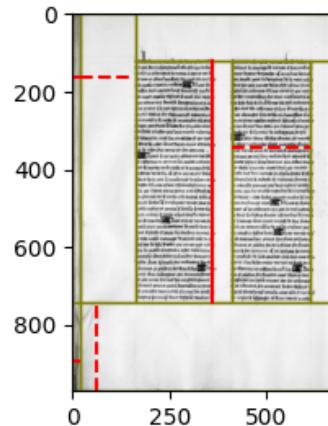




diff = 10.373191873936094

---> new split at 362

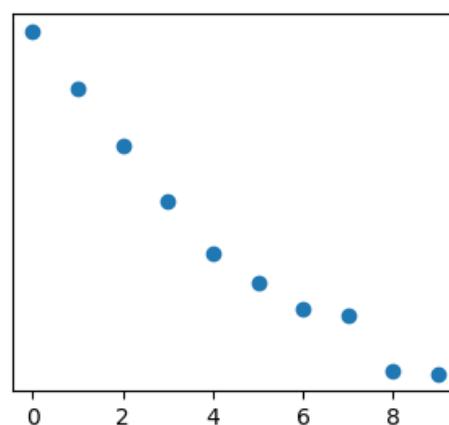
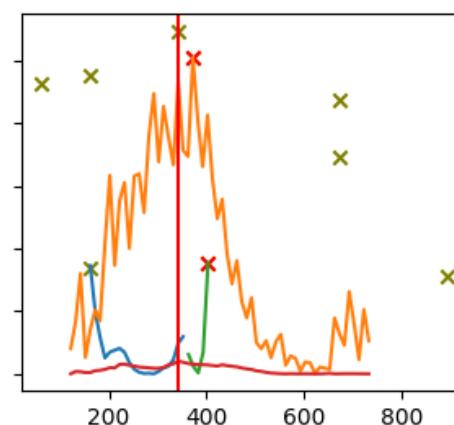
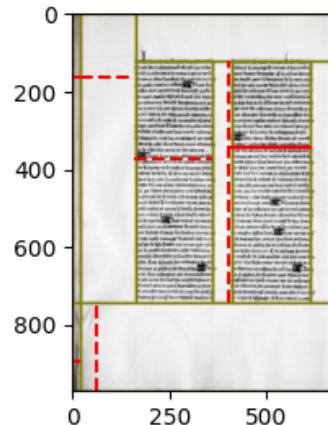
DEPTH = 2



diff = 82.84362319593697

---> new split at 342

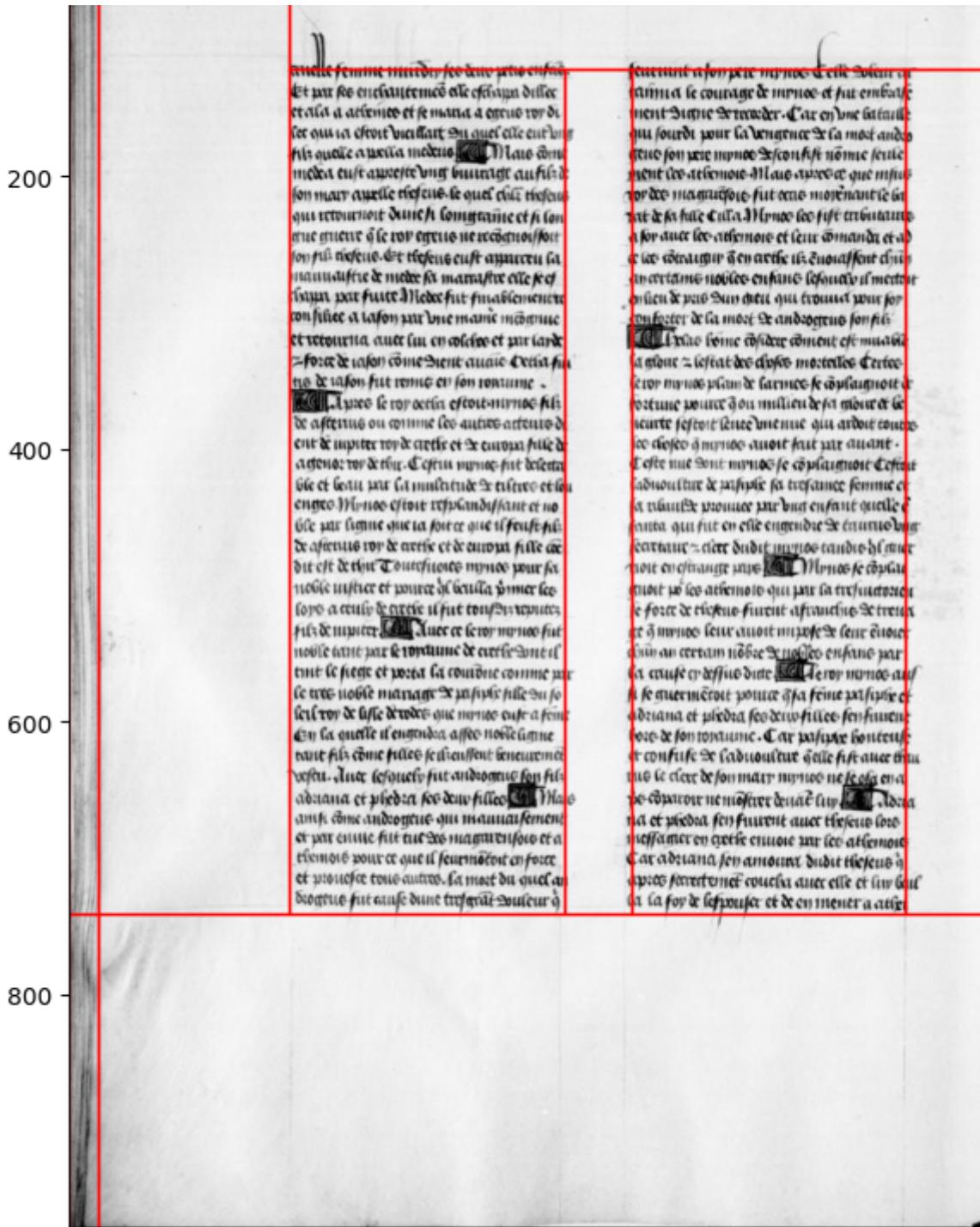
DEPTH = 1

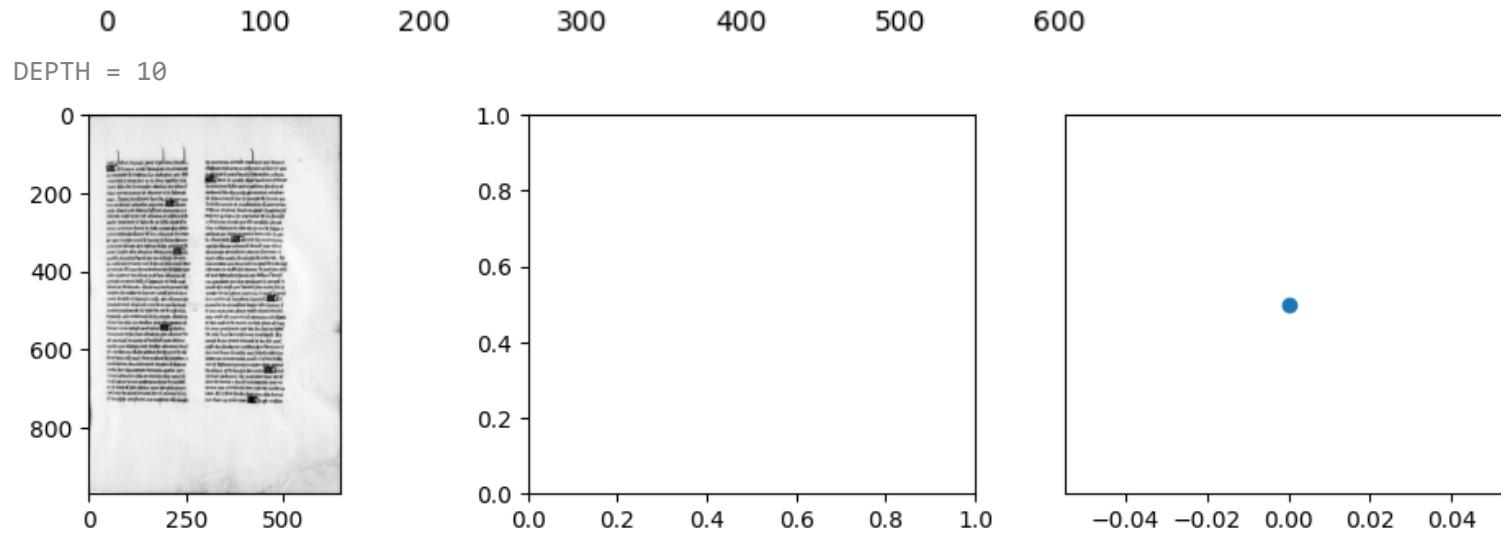


diff = 5.479288972991526

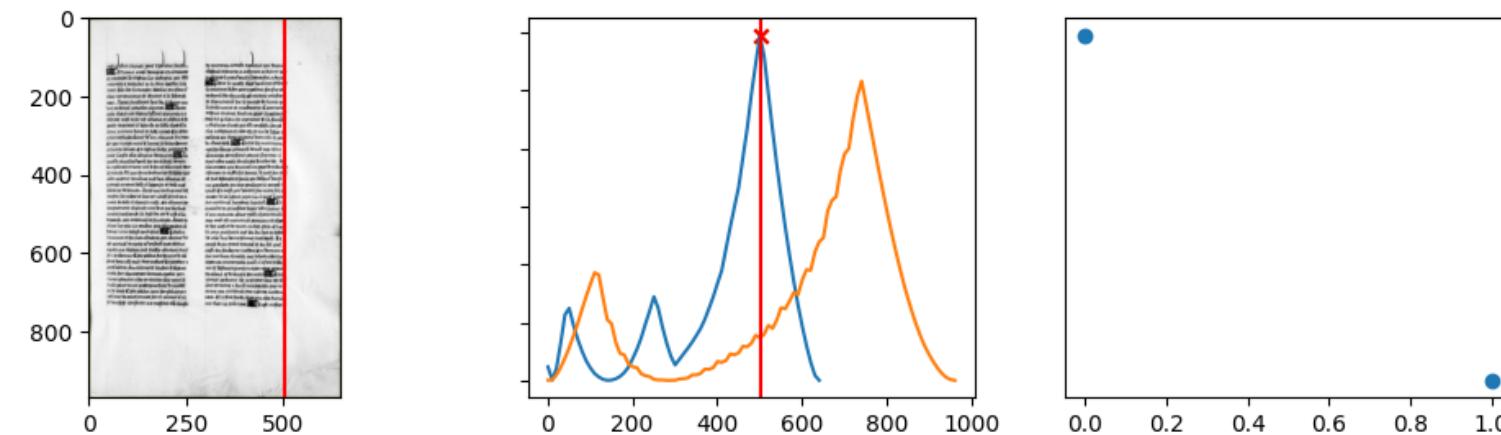
[[0, 0, 21, 741], [21, 0, 161, 741], [161, 0, 674, 121], [161, 121, 362, 741], [362, 121, 412, 741]



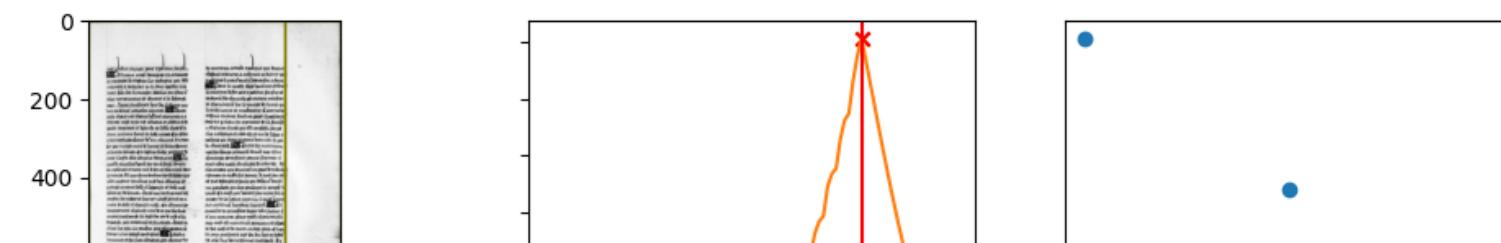


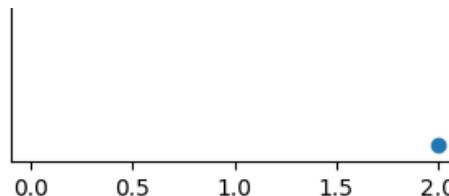
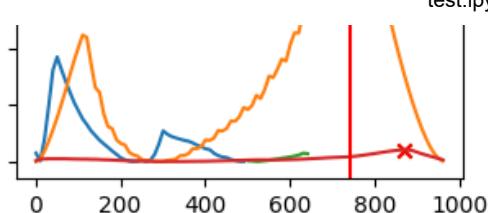
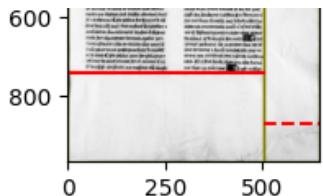


---> new split at 501
DEPTH = 9

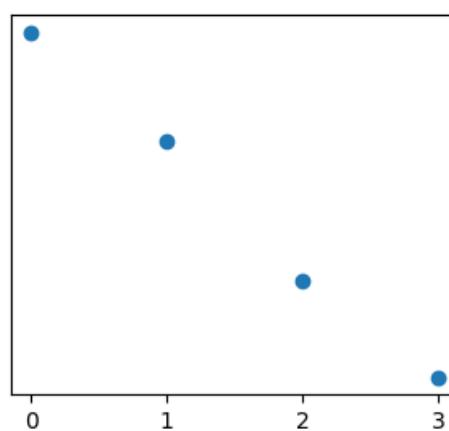
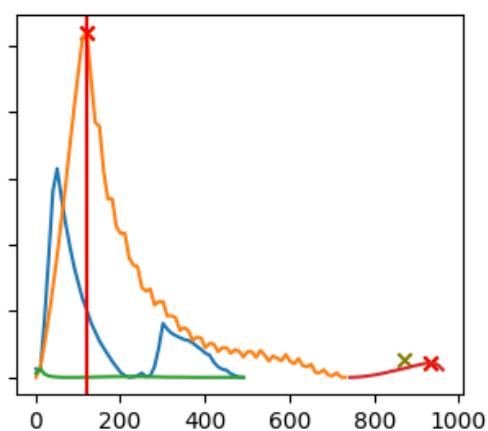
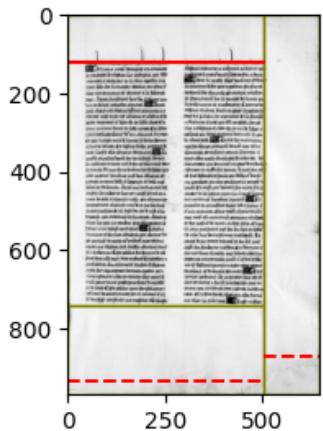


diff = 119.0557166443341
---> new split at 741
DEPTH = 8

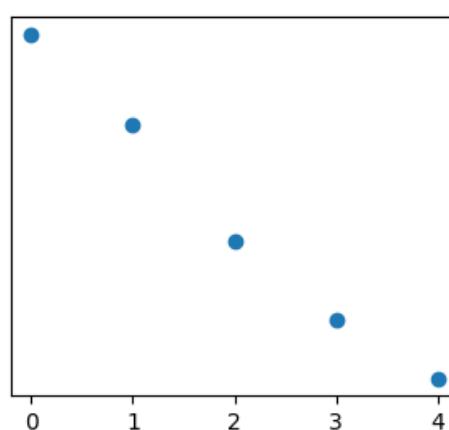
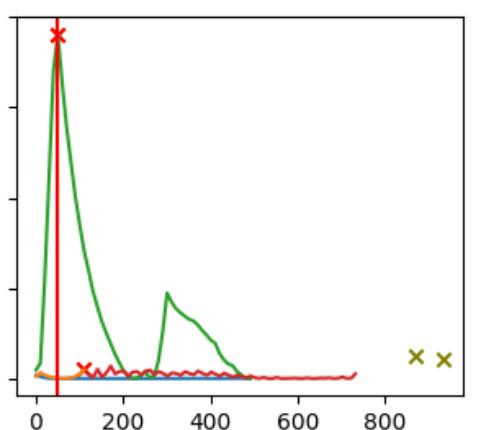
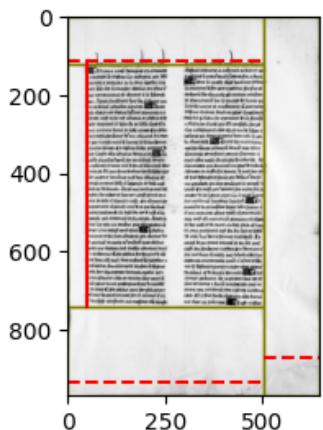




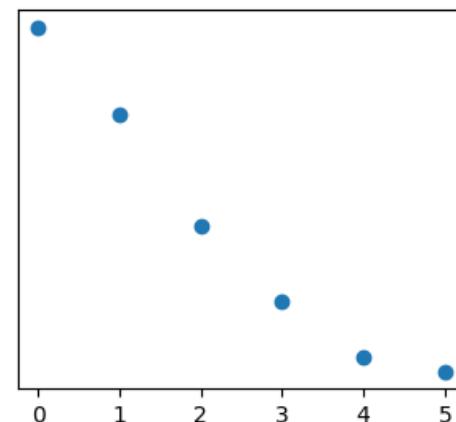
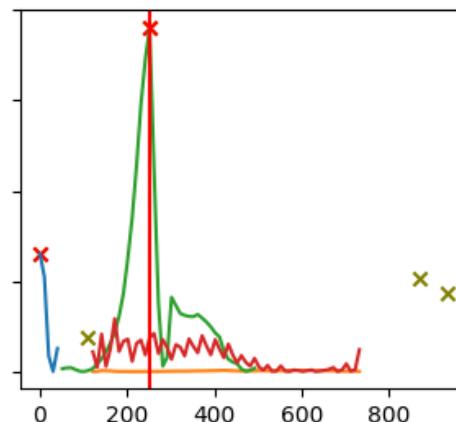
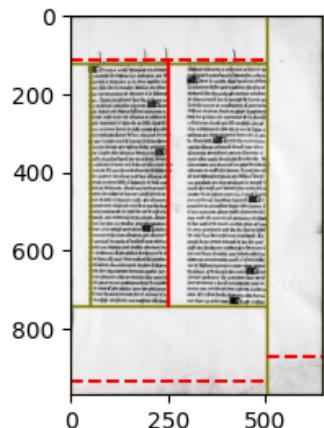
diff = 151.75651353701505
---> new split at 121
DEPTH = 7



diff = 104.11189827771909
---> new split at 51
DEPTH = 6



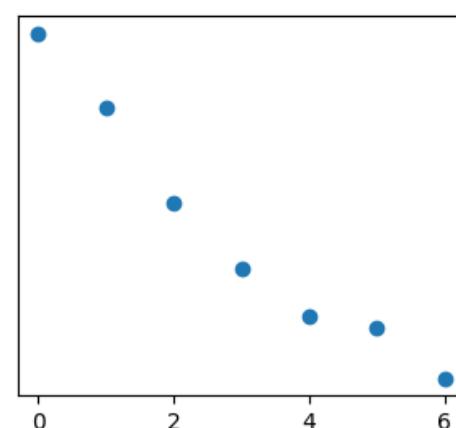
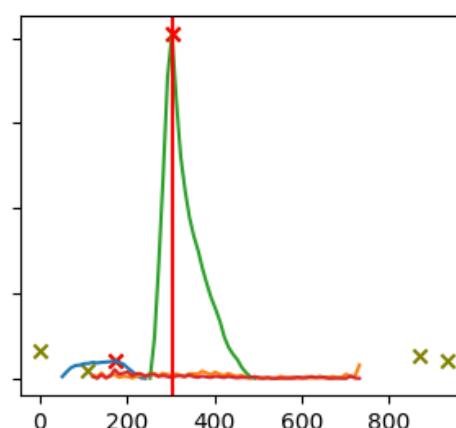
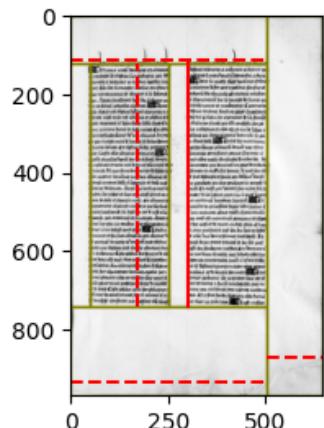
diff = 76.19319261768942
---> new split at 252
DEPTH = 5



diff = 19.056875516814216

---> new split at 303

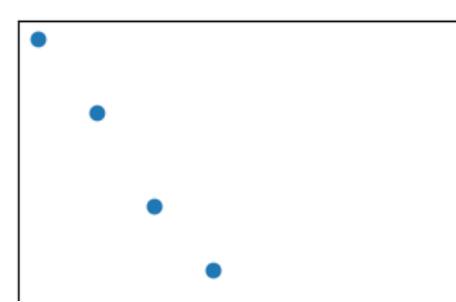
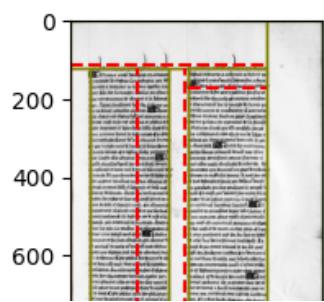
DEPTH = 4

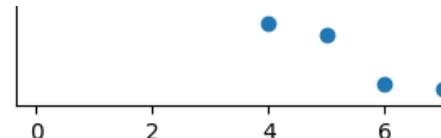
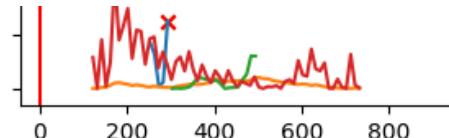
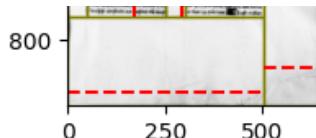


diff = 81.14222466994556

---> new split at 1

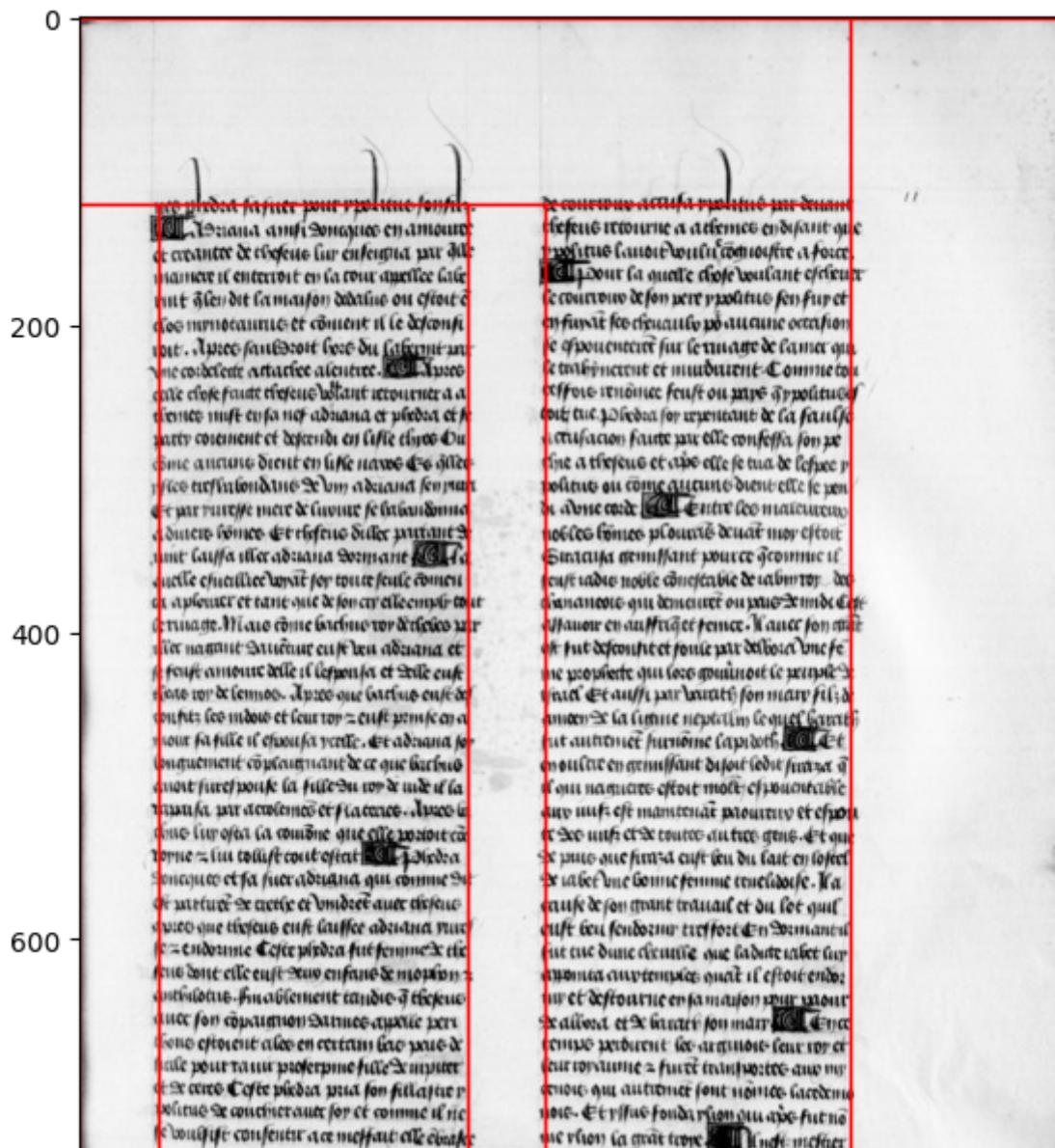
DEPTH = 3

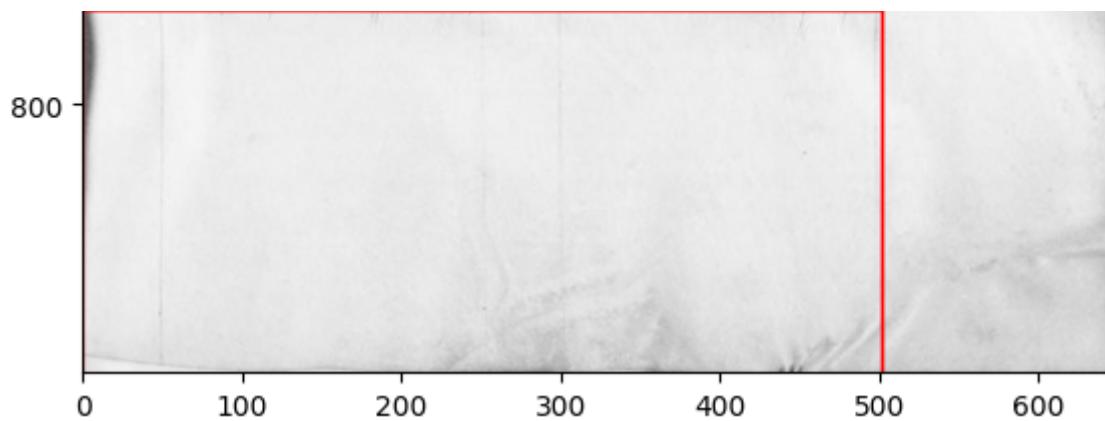




diff = 6.499509110011786

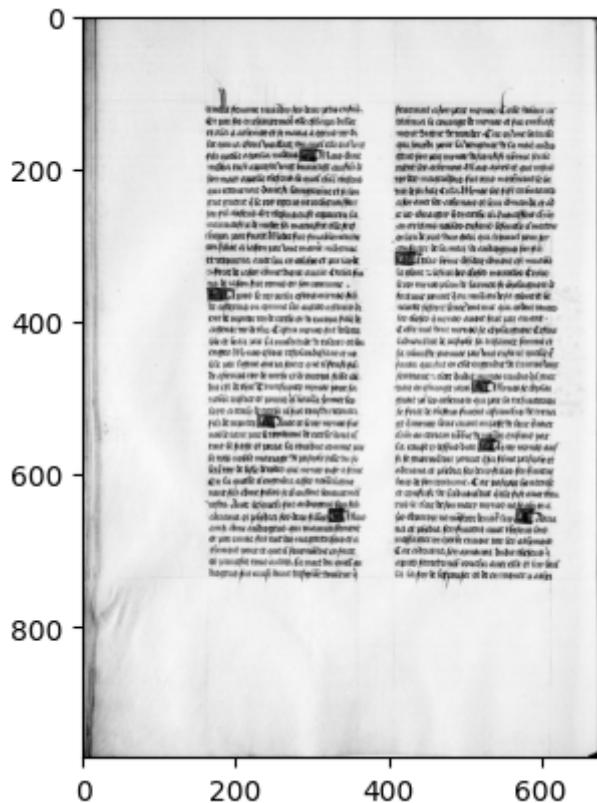
[[0, 0, 501, 121], [0, 121, 51, 741], [51, 121, 252, 741], [252, 121, 303, 741], [303, 121, 501, 741]]



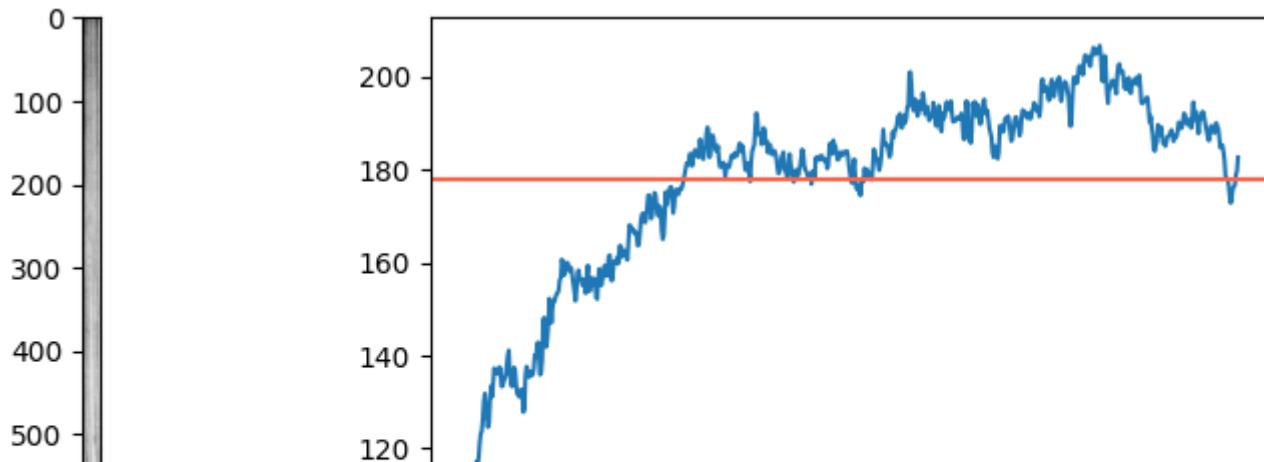


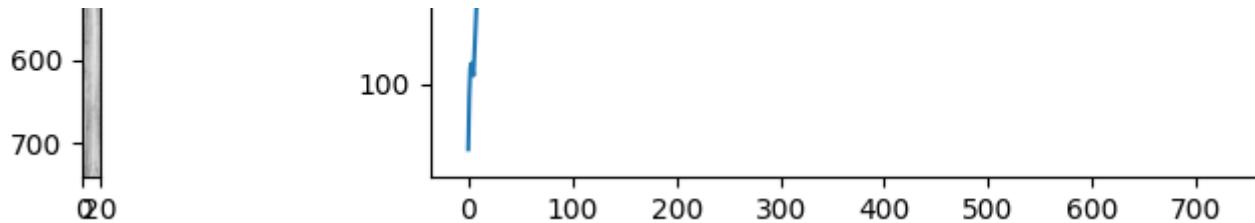
```
check_text(allres,imgs)
```

IMAGE 0



Analyzing area 0 0 21 741



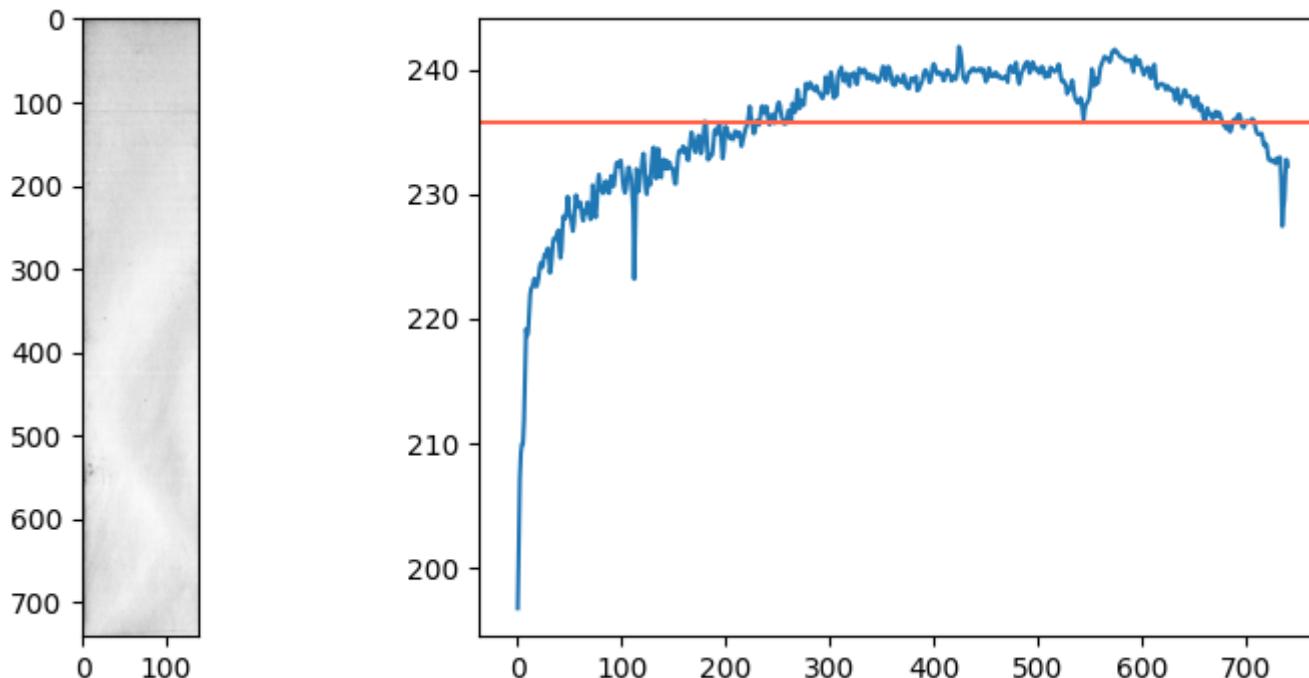


[208, 1, 1, 1, 2, 8, 2, 8] [63, 41, 16, 37, 3, 6, 341, 3]

67.76141509000531 106.71076562371765

---> not text

Analyzing area 21 0 161 741



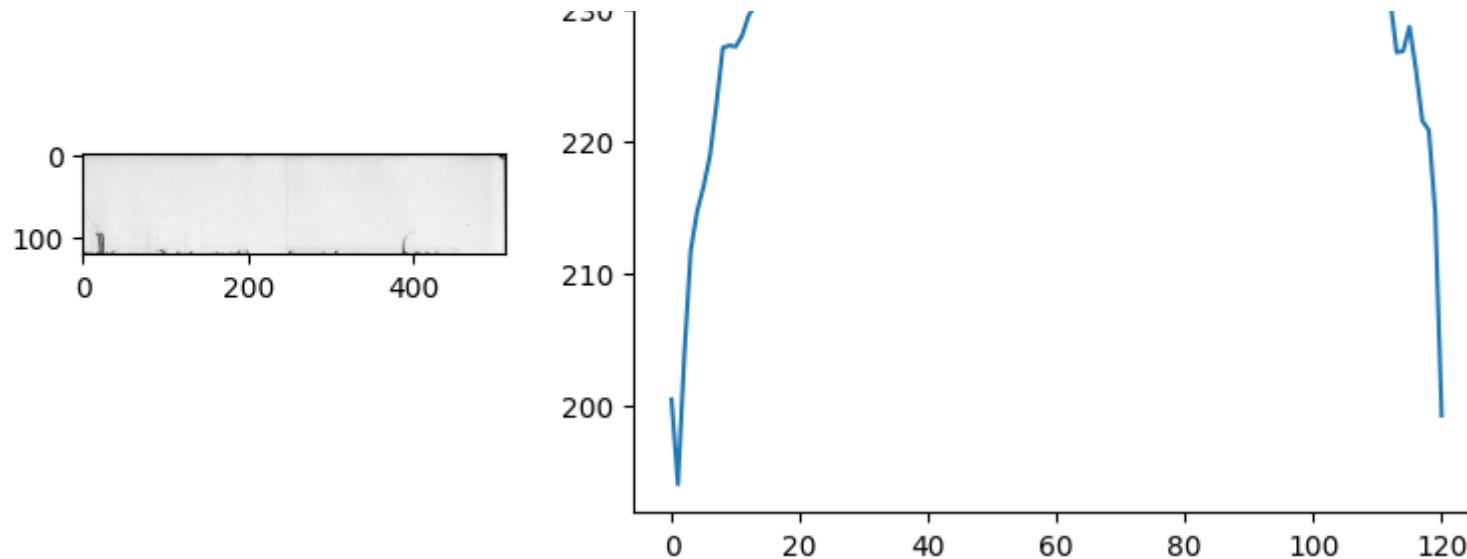
[180, 41, 4, 2, 1, 1, 3, 10, 4, 32] [1, 4, 1, 9, 4, 8, 419, 8, 9]

52.46293930004303 129.98556425072843

---> not text

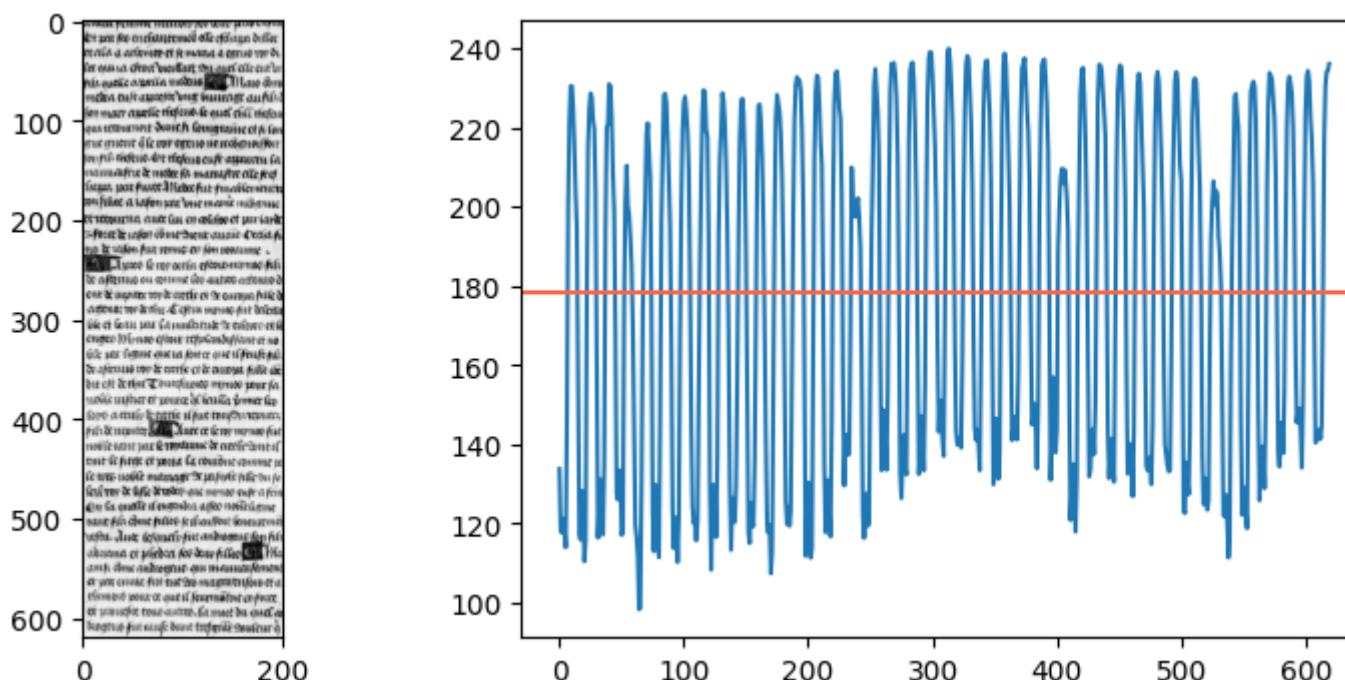
Analyzing area 161 0 674 121



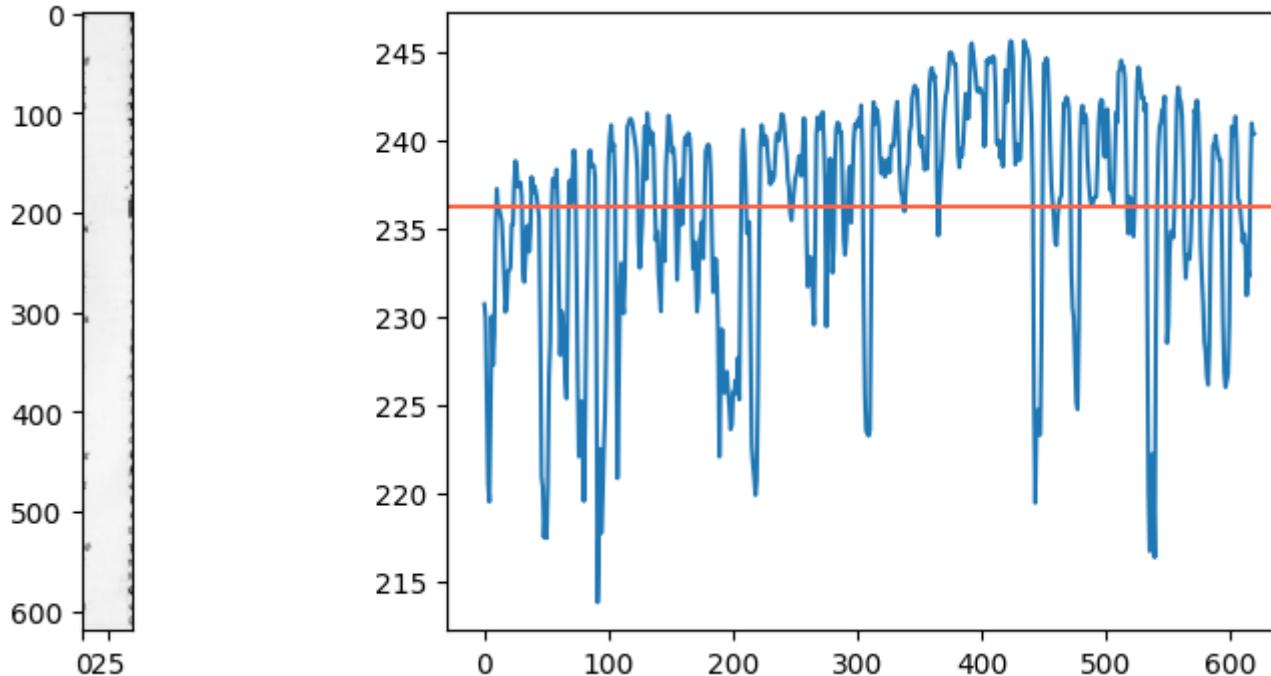


[21, 1, 9] [11, 79]
8.219218670625303 34.0
---> not text

Analyzing area 161 121 362 741



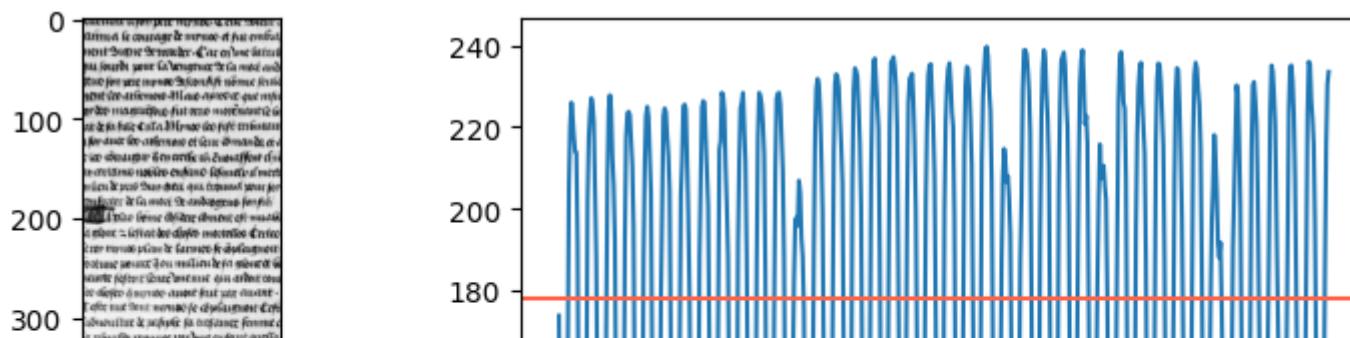
Analyzing area 362 121 412 741

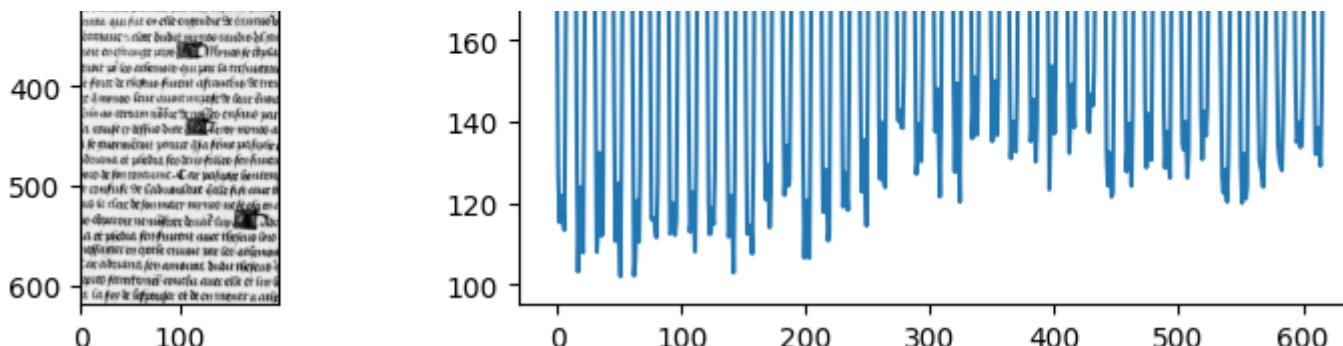


[10, 13, 7, 11, 9, 10, 10, 8, 4, 8, 3, 1, 10, 24, 11, 2, 8, 2, 2, 3, 1, 7, 2, 1, 7, 4, 7, 2, 2, 8,
4.50994366225729 13.454564508836732

---> not text

Analyzing area 412 121 612 741

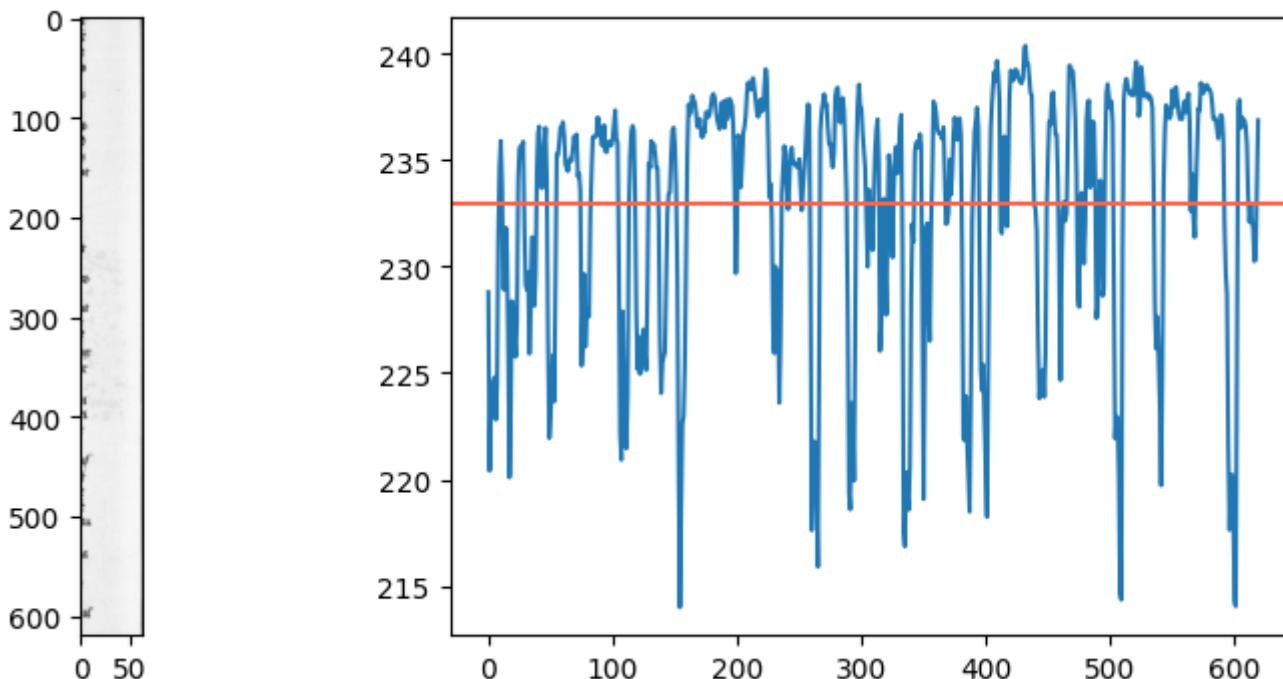




[8, 7, 8, 7, 7, 7, 8, 7, 7, 7, 7, 8, 8, 7, 7, 6, 6, 7, 7, 7, 7, 7, 6, 7, 7, 7, 7, 7, 7, 7, 7, 0.6053499339507735 0.9319499109533074

----->>> Contains 41 lines of text

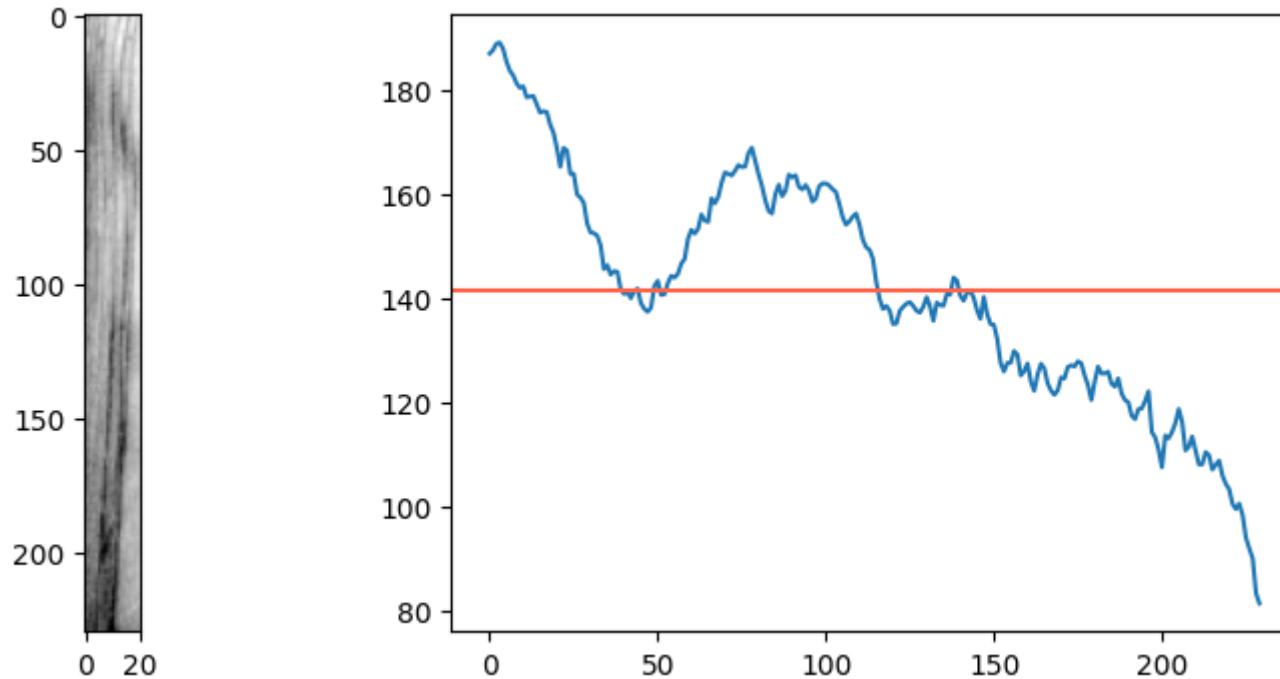
Analyzing area 612 121 674 741



[9, 12, 10, 7, 8, 9, 10, 8, 8, 2, 8, 1, 1, 8, 7, 2, 3, 2, 3, 2, 10, 8, 3, 9, 9, 2, 1, 11, 4, 2, 2, 3.52723743547757 8.727414123744156

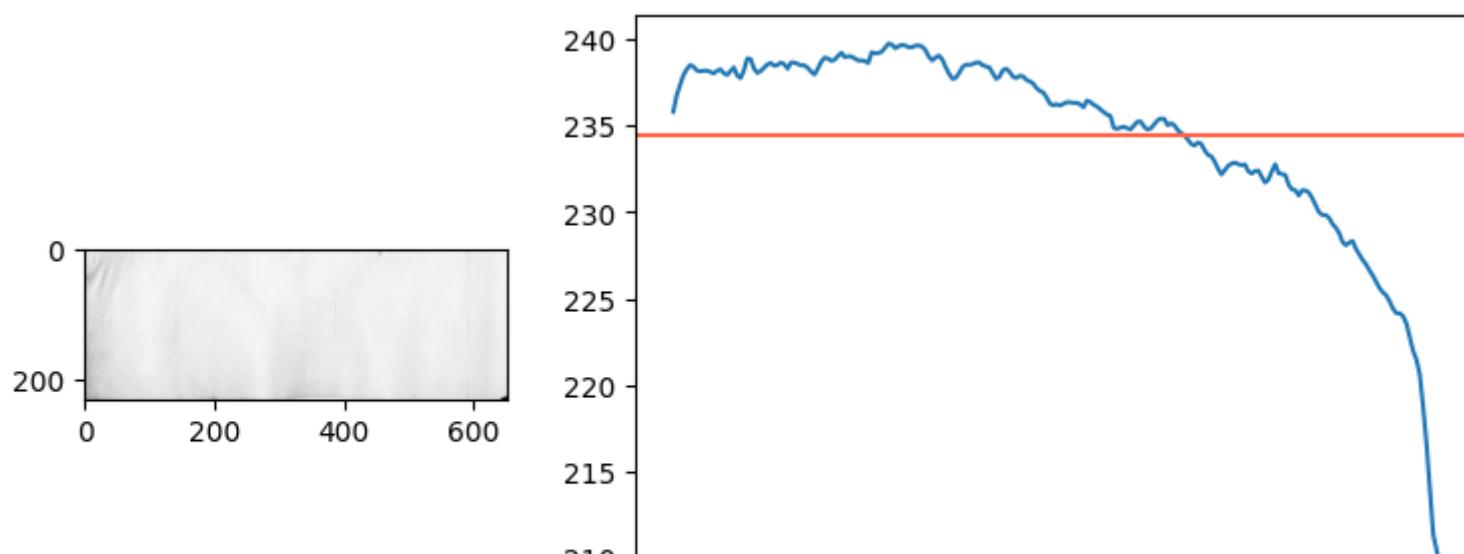
--> not text

Analyzing area 0 741 21 971



[40, 1, 2, 63, 2] [4, 4, 2, 22, 90]
33.5952377577537 25.476263462289754
---> not text

Analyzing area 21 741 674 971



[152] [78]

0.0 0.0

---> not text

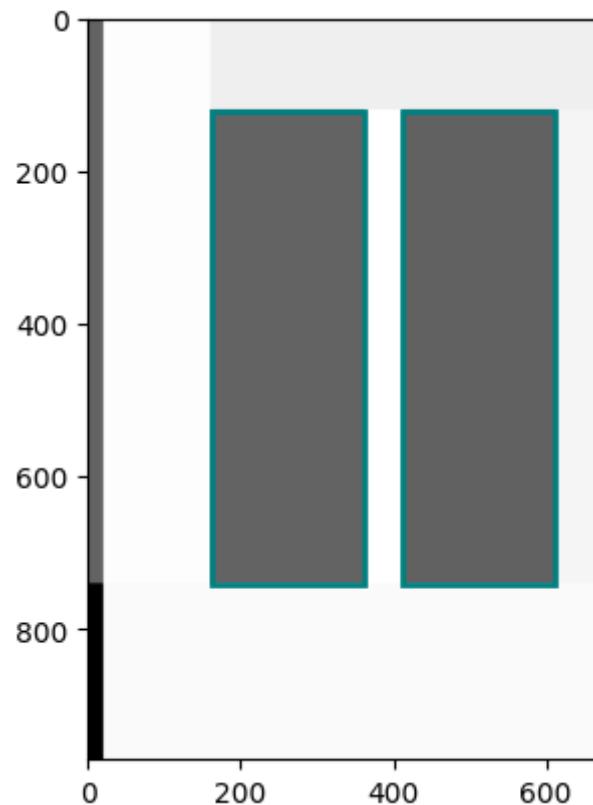
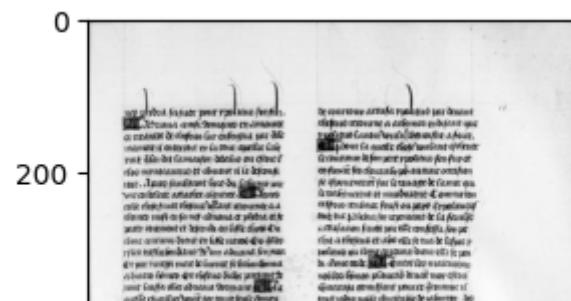
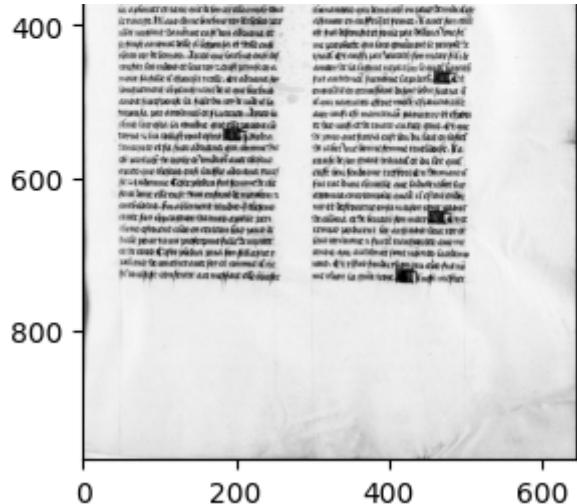
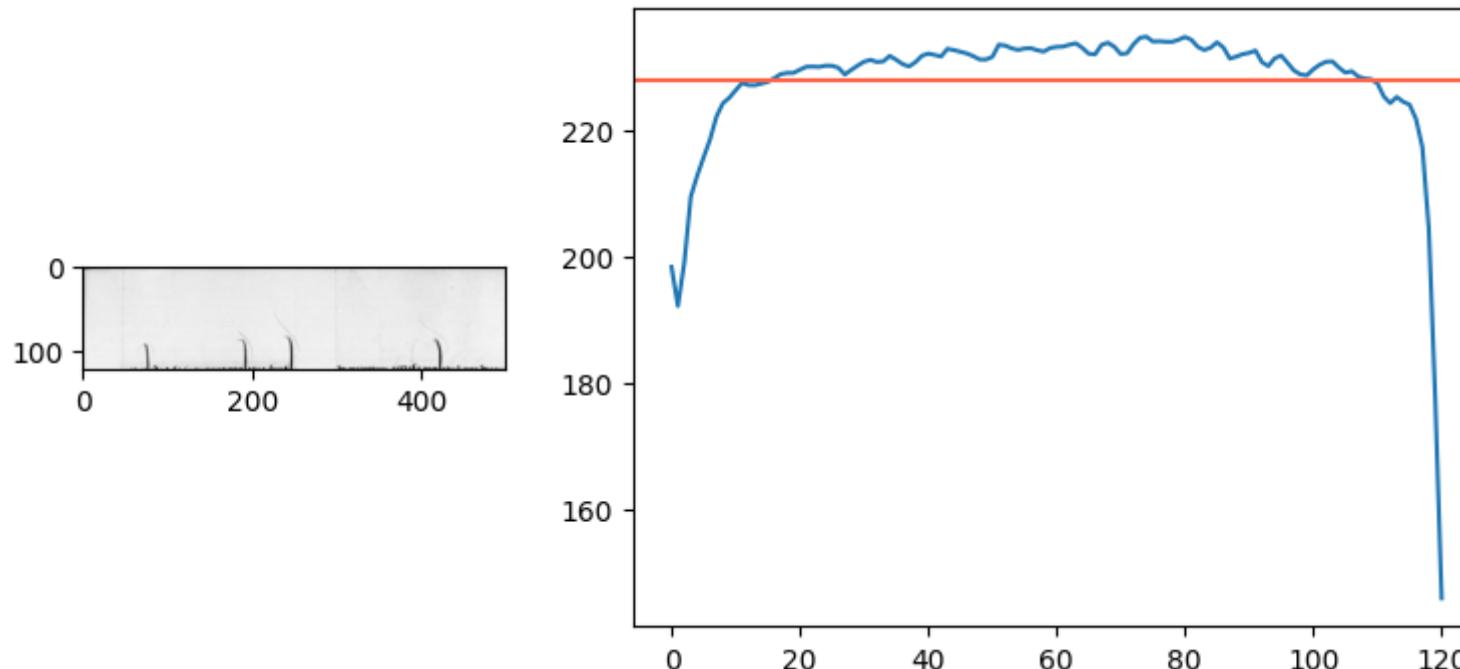


IMAGE 1





Analyzing area 0 0 501 121

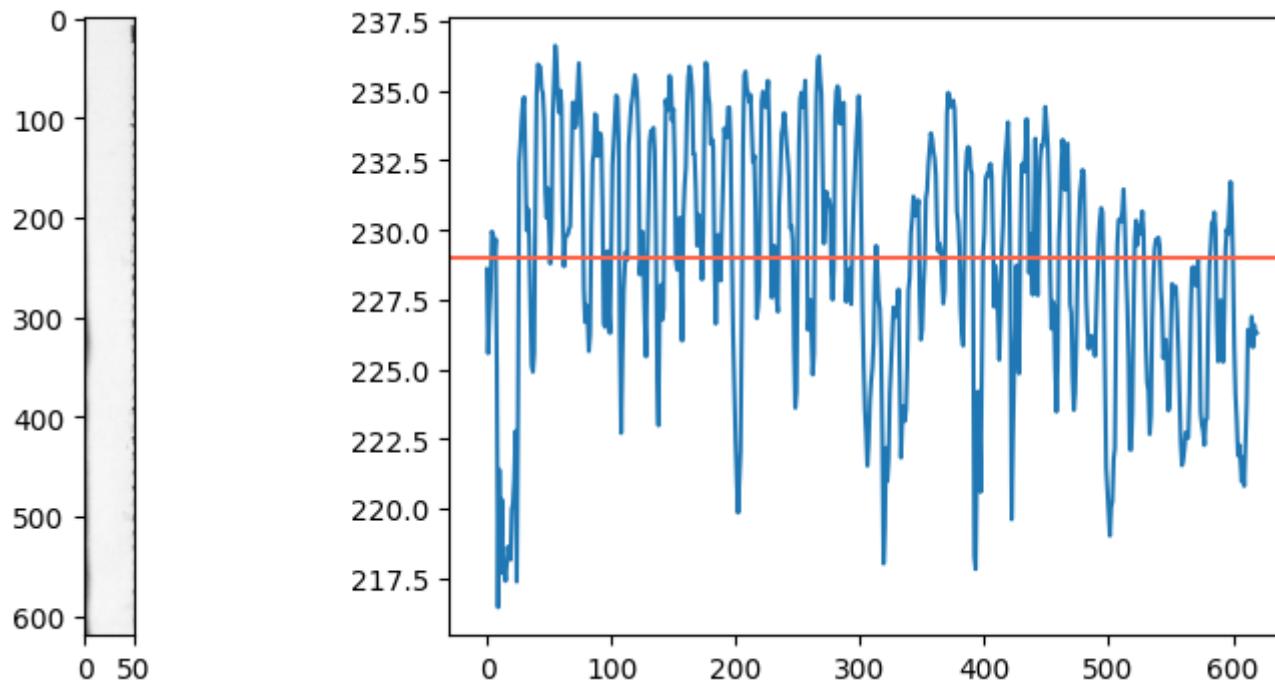


[16, 11] [94]

2.5 0.0

---> not text

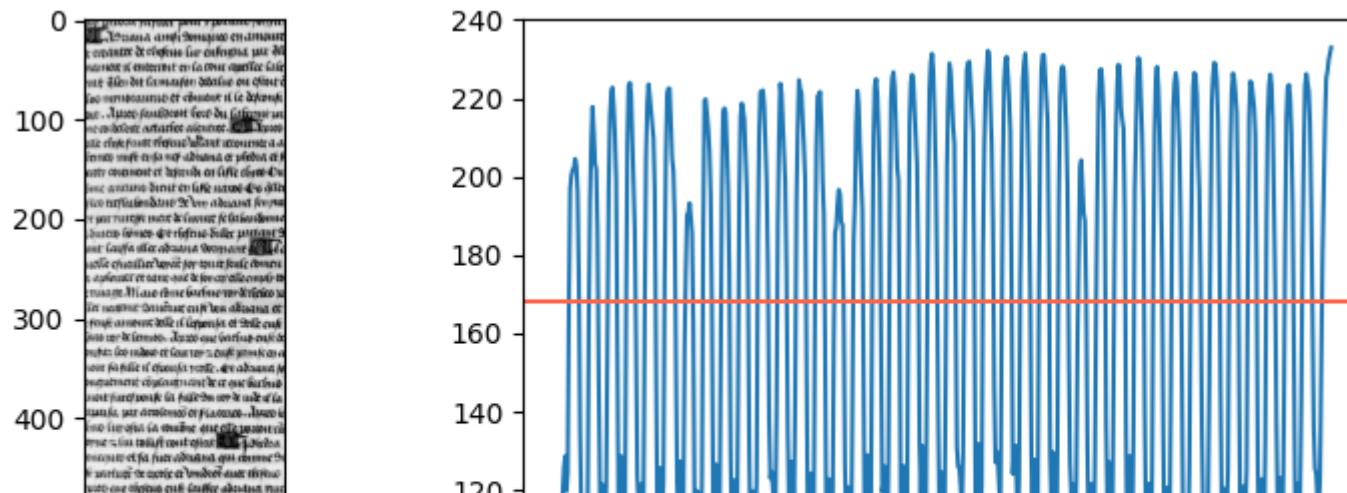
Analyzing area 0 121 51 741

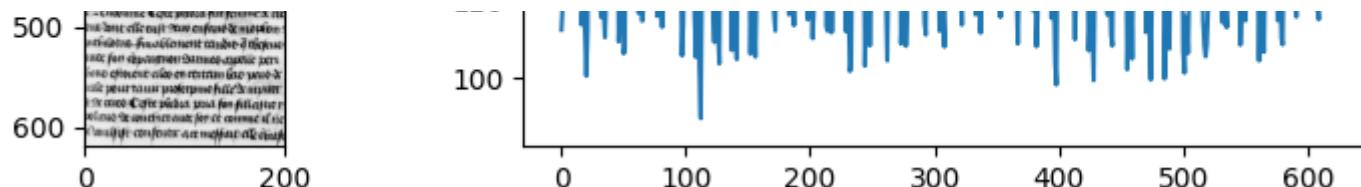


[4, 18, 4, 1, 1, 6, 3, 2, 5, 1, 2, 4, 6, 2, 2, 1, 2, 2, 8, 3, 2, 2, 3, 6, 3, 5, 11, 27, 5, 3, 3, 8,
7.2100866649747095 3.8014941831721063

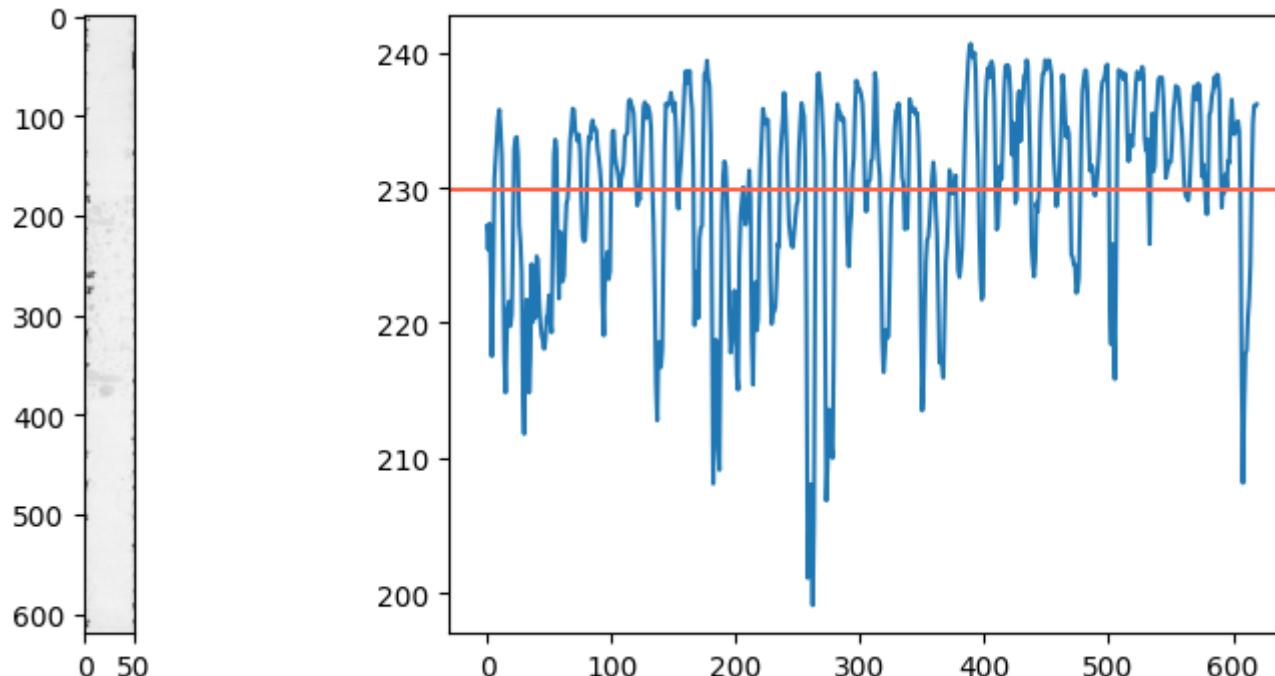
---> not text

Analyzing area 51 121 252 741





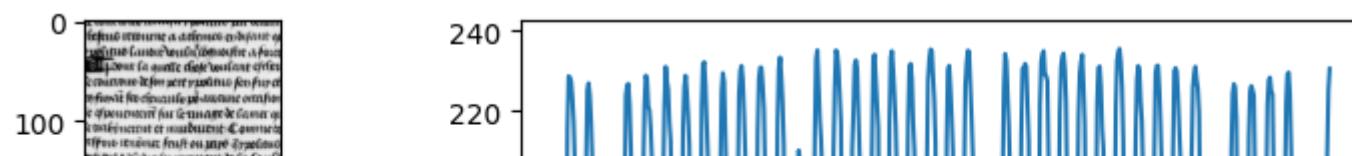
Analyzing area 252 121 303 741

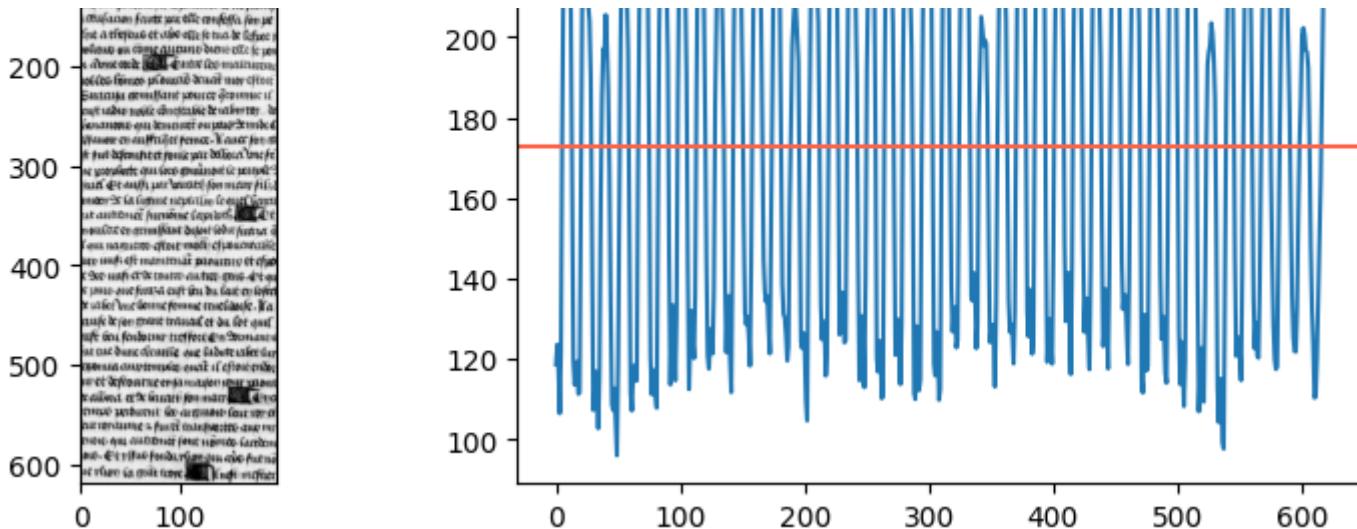


```
[6, 9, 28, 9, 5, 7, 3, 9, 2, 7, 9, 13, 4, 8, 8, 8, 8, 8, 5, 2, 9, 3, 9, 11, 3, 6, 4, 3, 2, 7, 3, 8  
4.647512775668293 5.794339910636931
```

--> not text

Analyzing area 303 121 501 741

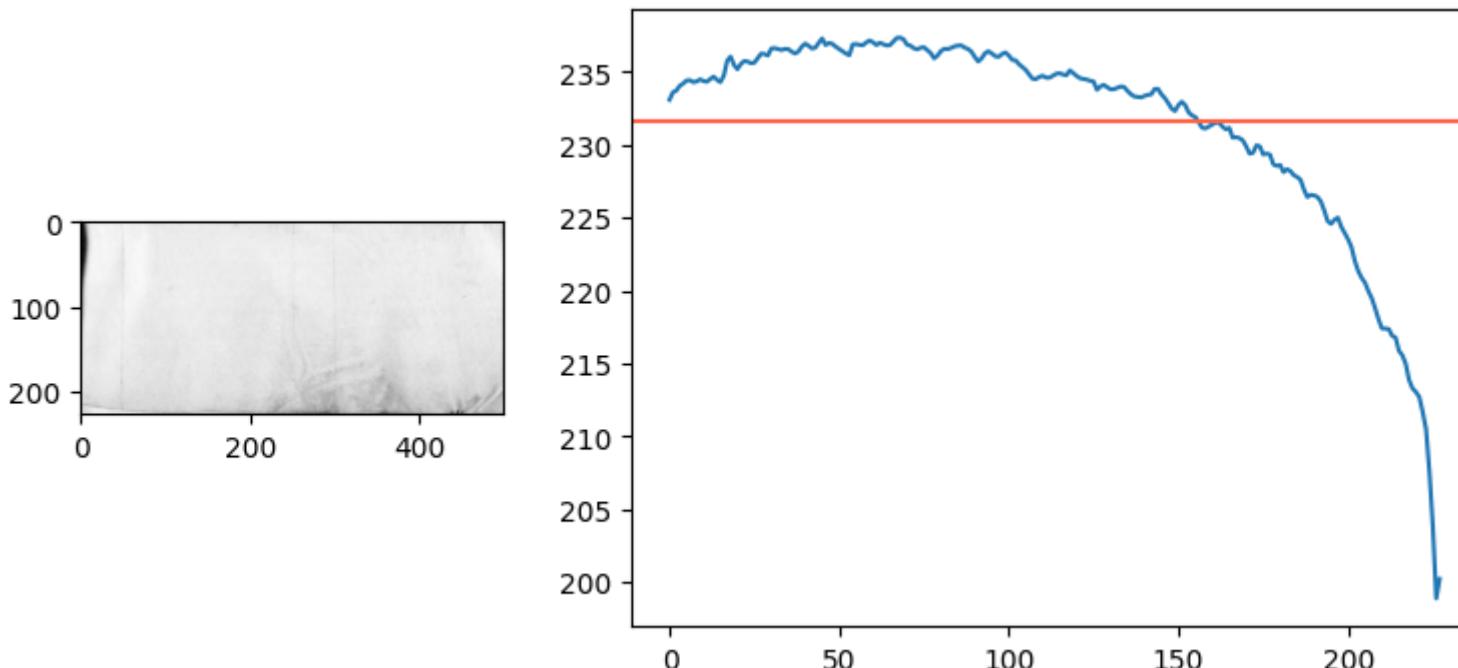




[5, 8, 8, 9, 8, 8, 7, 7, 7, 7, 7, 7, 7, 8, 8, 8, 6, 7, 7, 7, 8, 8, 7, 6, 8, 7, 6, 7, 6, 7, 7, 7, 7, 7, 7, 8, 7, 0.7789619230571372 0.9165021574661432

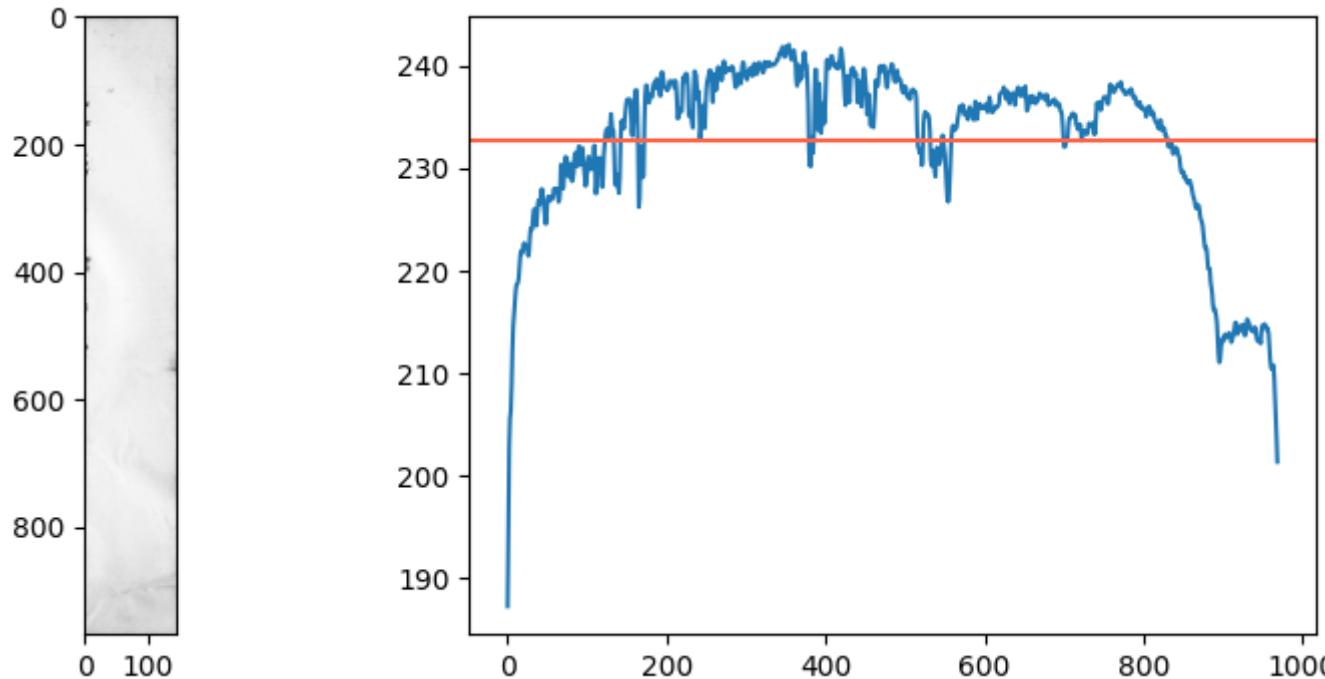
----->>> Contains 41 lines of text

Analyzing area 0 741 501 969



↳ ↳
0.0 0.0
---> not text

Analyzing area 501 0 646 969

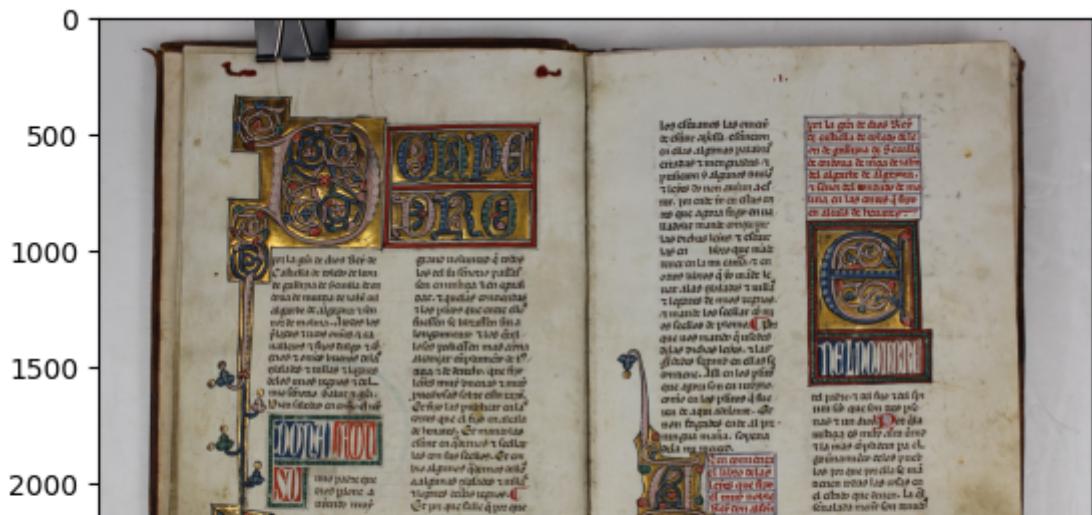


[123, 9, 8, 4, 1, 7, 15, 10, 5, 1, 139] [11, 21, 207, 1, 131, 9, 1, 141, 18, 107]

48.2287935669722 71.02682591809942

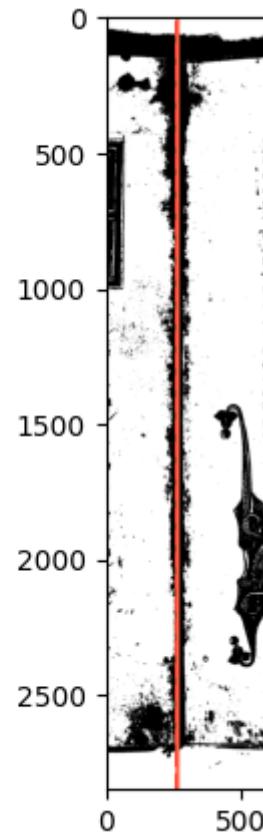
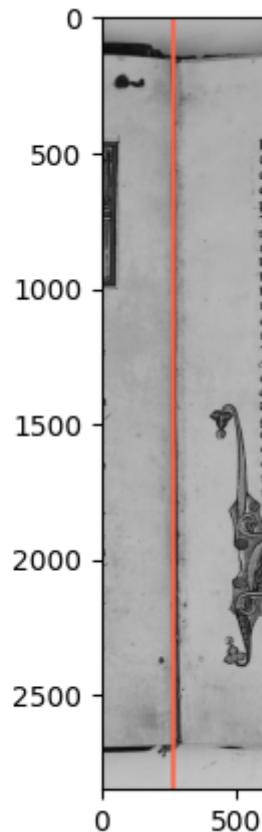
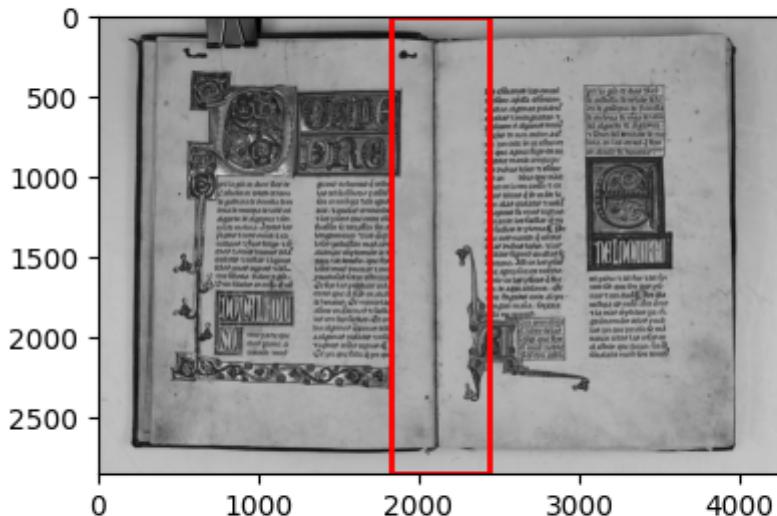
---> not text

```
image = cv2.imread("../ImagesCodicologie/Corpus de jeu/Foronda/FF-RBMEZIII9, fol.1.jpg")
image = cv2.cvtColor(image, cv2.COLOR_BGR2RGB)
plt.imshow(image)
plt.show()
```

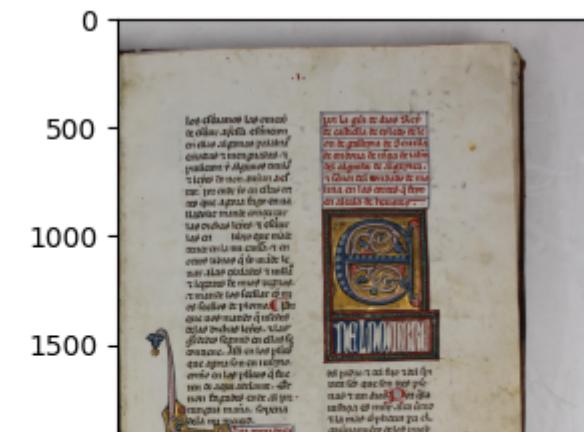
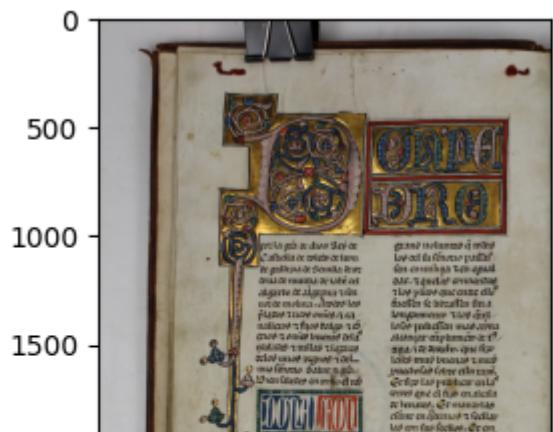


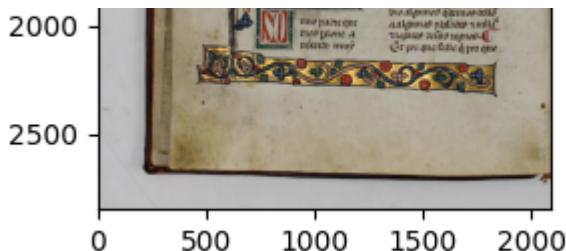
```
imgs = reframe(image)
```

643 candidate lines amongst 645 lines found
 $r = 0.9266151685393258$ with split at 2094



$r_{max} = 0.9266151685393258$



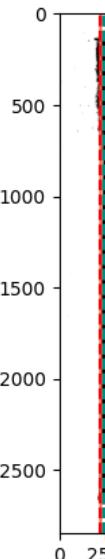
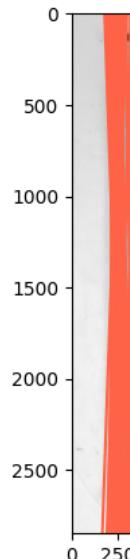
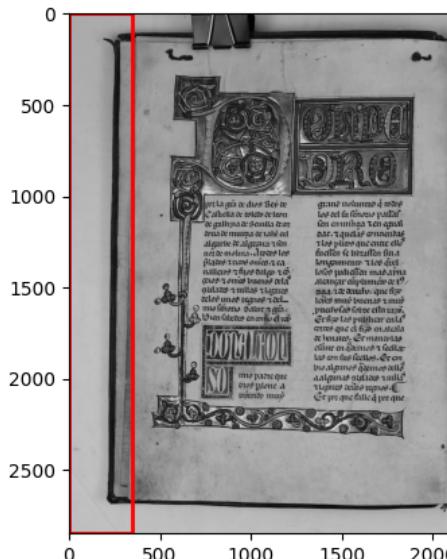


-----1-----

60 candidate lines amongst 76 lines found

rmax = 0.913623595505618

$r = 0.7447331460674158$ with split at 227 with risk = 1174753956.5567145



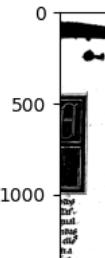
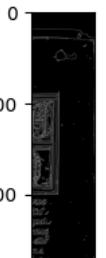
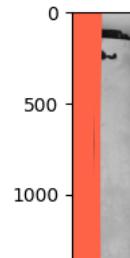
227

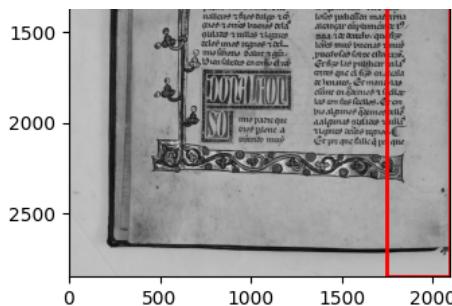
-----r-----

117 candidate lines amongst 271 lines found

rmax = 0.9224016853932584

$r = 0.8862359550561798$ with split at 2086 with risk = 1672939558.4326963

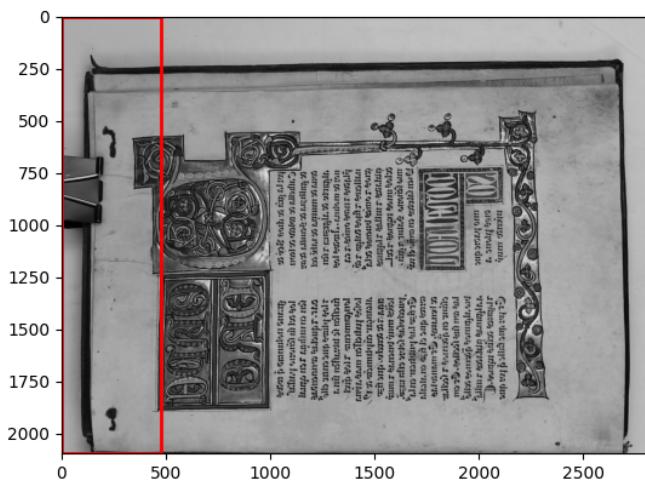




2086

-----U-----

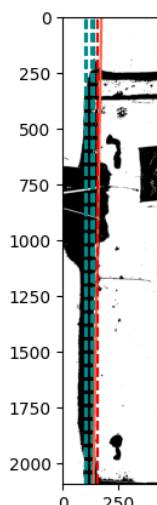
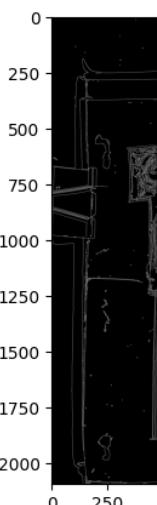
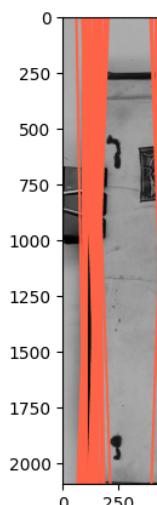
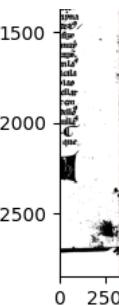
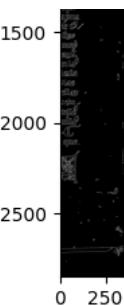
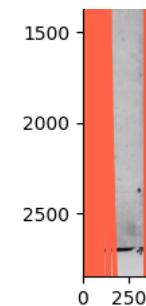
79 candidate lines amongst 123 lines found

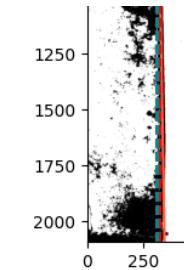
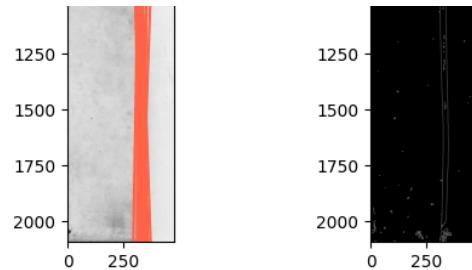
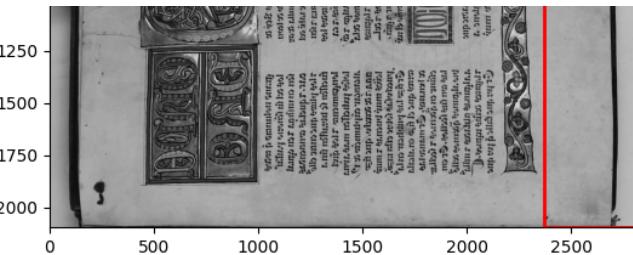
 $r_{max} = 0.8744030563514804$ $r = 0.7363896848137537$ with split at 155 with risk = 2108269159.3852994

155

-----d-----

21 candidate lines amongst 26 lines found

 $r_{max} = 0.8629417382999045$ $r = 0.8042024832855779$ with split at 2698 with risk = 1263617105.0120268



2698

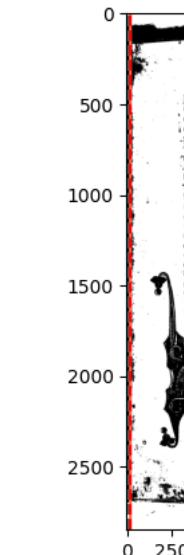
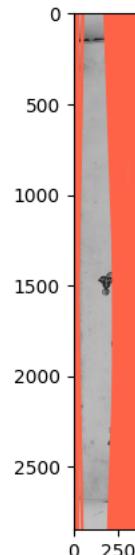
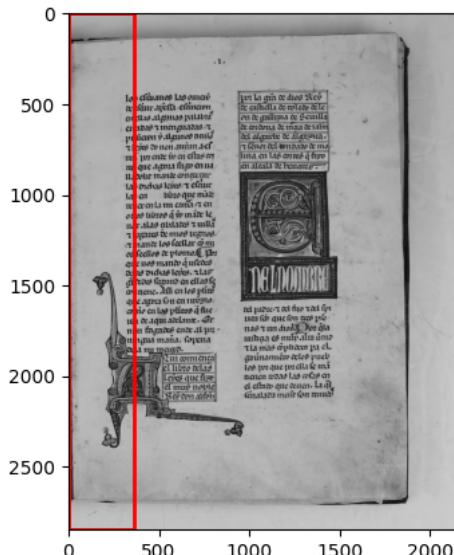
[227, 2086, 155, 2698]

-----1-----

139 candidate lines amongst 554 lines found

rmax = 0.9283707865168539

r = 0.7612359550561798 with split at 20 with risk = 1353812636.0060759



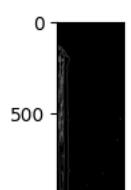
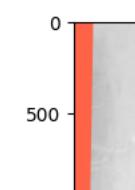
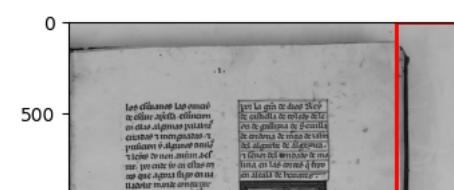
20

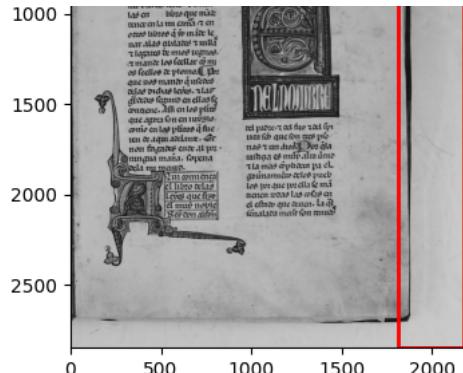
-----r-----

42 candidate lines amongst 46 lines found

rmax = 0.9287219101123596

r = 0.9002808988764045 with split at 1881 with risk = 764135737.7339078

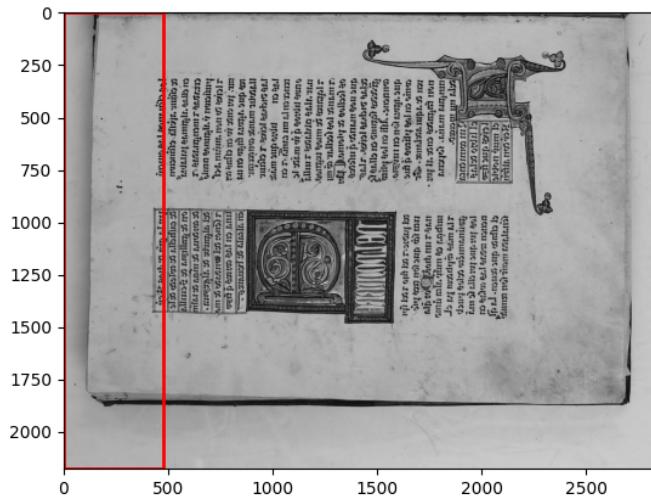




1881

-----U-----

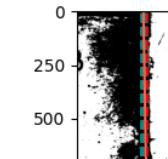
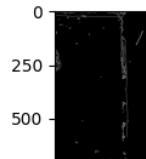
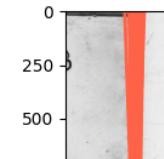
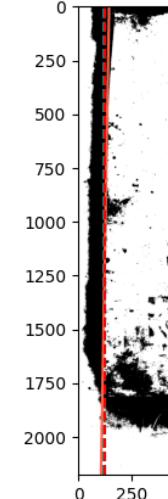
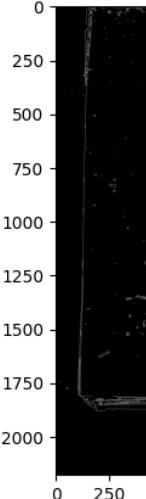
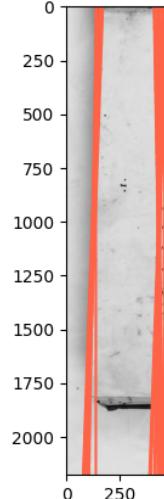
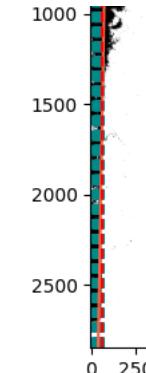
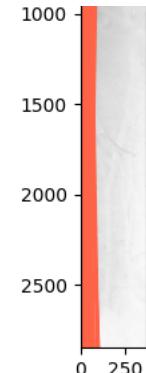
33 candidate lines amongst 48 lines found

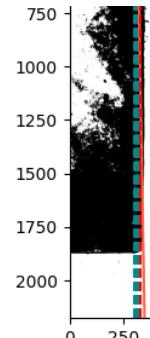
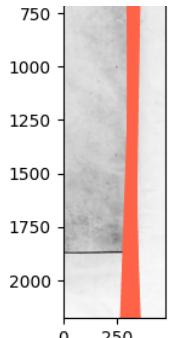
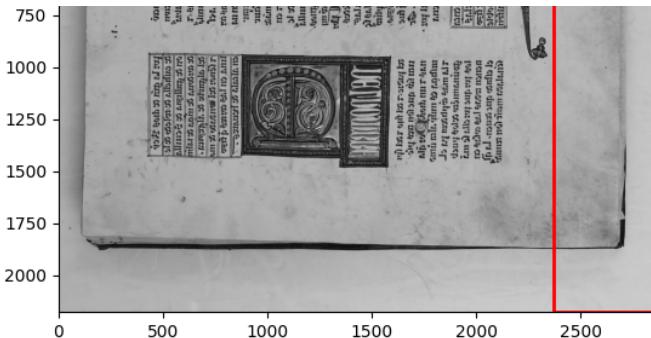
 $r_{max} = 0.8530762167125804$ $r = 0.8530762167125804$ with split at 126 with risk = 564598641.5757909

126

-----d-----

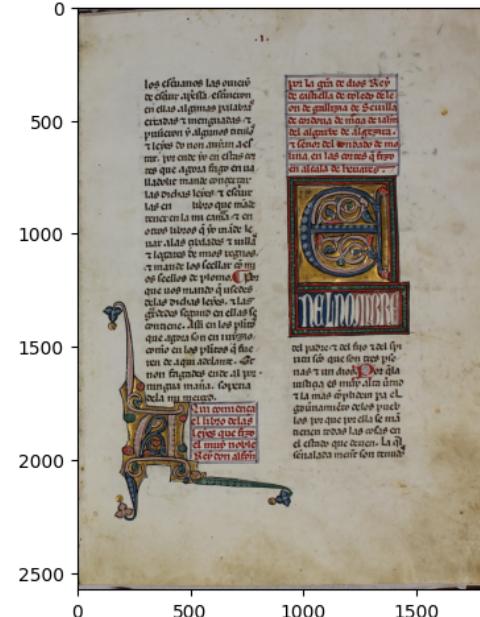
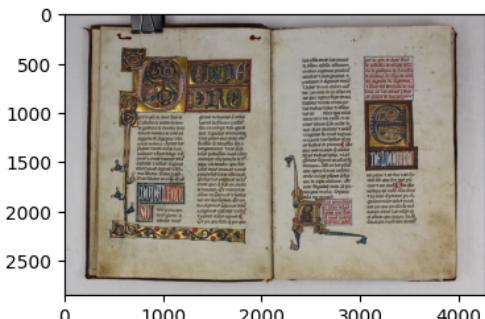
29 candidate lines amongst 33 lines found

 $r_{max} = 0.859504132231405$ $r = 0.7295684113865932$ with split at 2701 with risk = 444419433.5240047



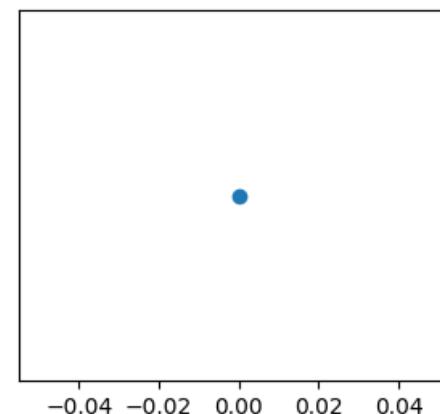
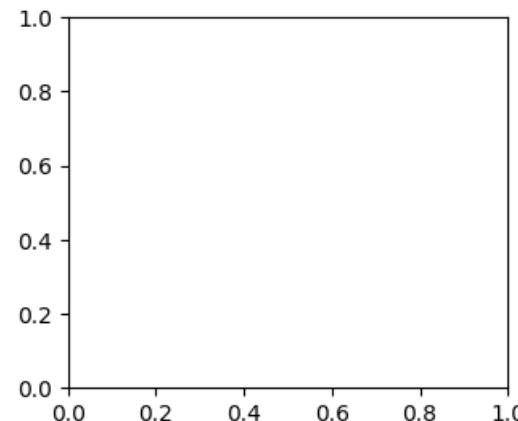
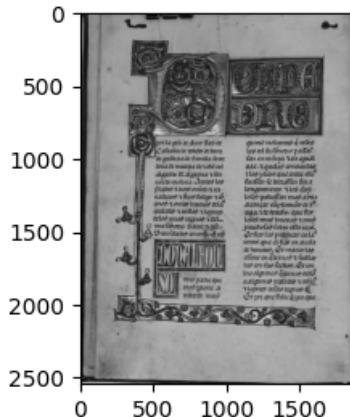
2701

[20, 1881, 126, 2701]



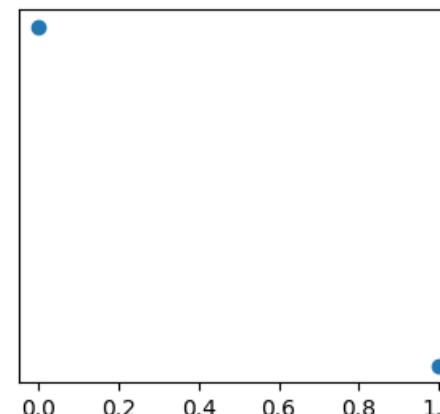
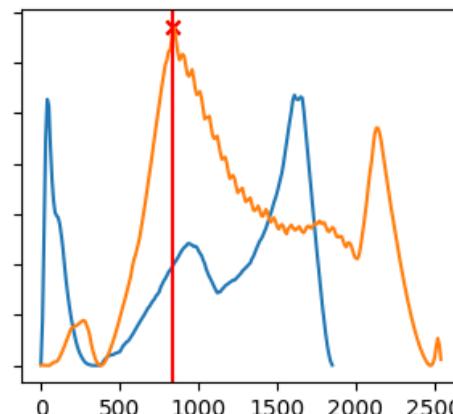
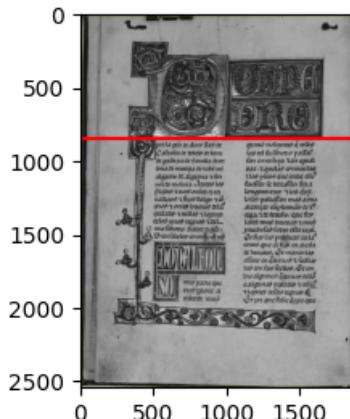

```
allres = []
for im in imgs:
    res = cart_regression_bi_simul(im,mean_squared_error,max_depth=20,rate=10)
    print(res)
    show_areas(im,res)
    allres.append(res)
```

DEPTH = 20



---> new split at 841

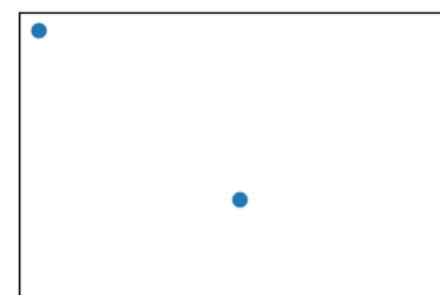
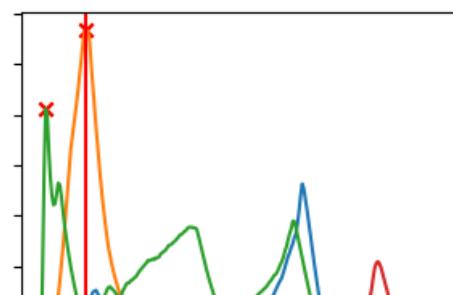
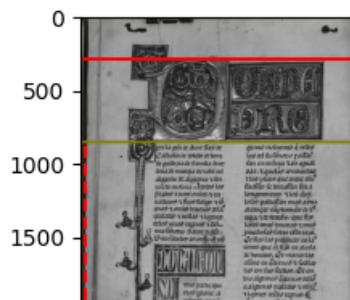
DEPTH = 19

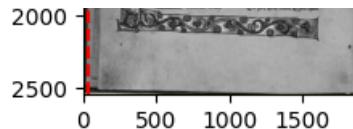


diff = 670.5115926021554

---> new split at 281

DEPTH = 18

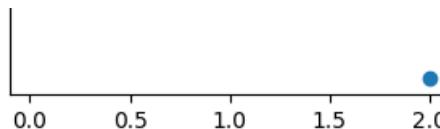
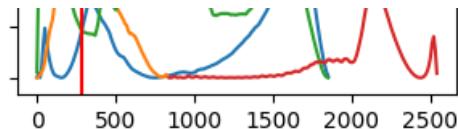
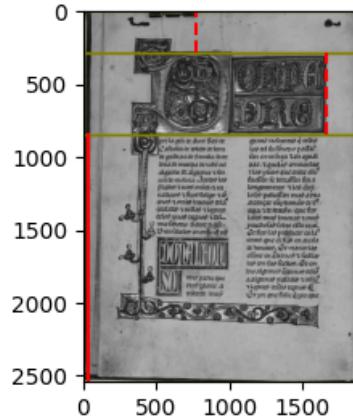




diff = 667.9500899707618

---> new split at 31

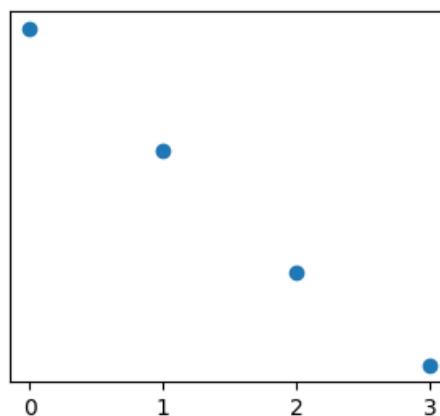
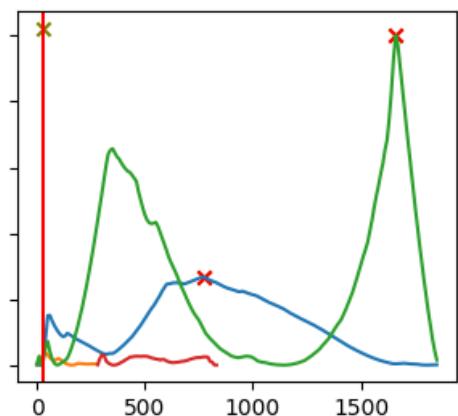
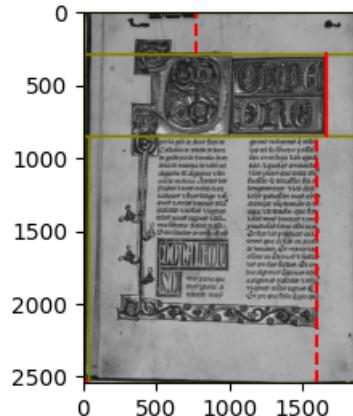
DEPTH = 17



diff = 511.0572349476315

---> new split at 1661

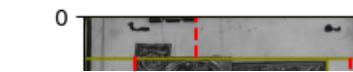
DEPTH = 16

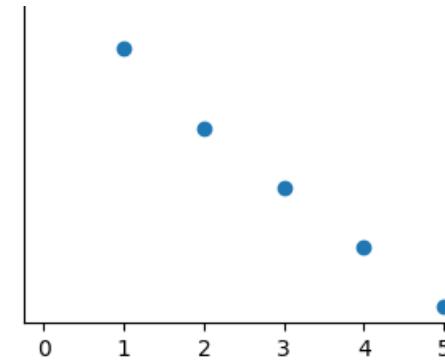
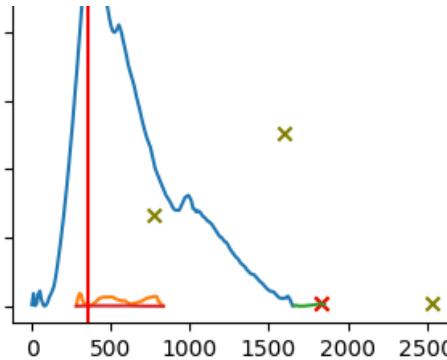
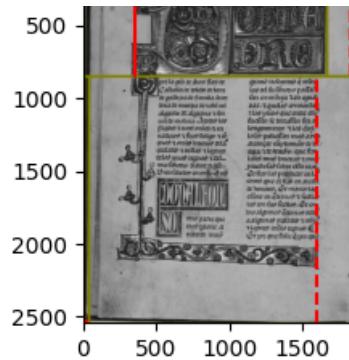


diff = 500.42792421812374

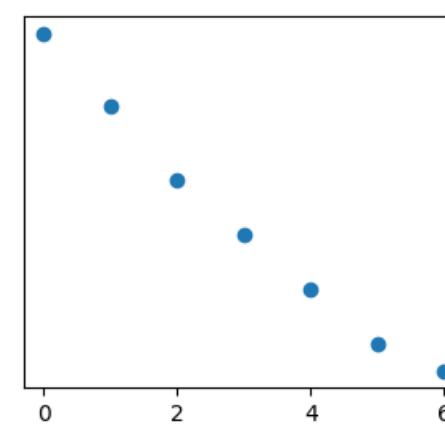
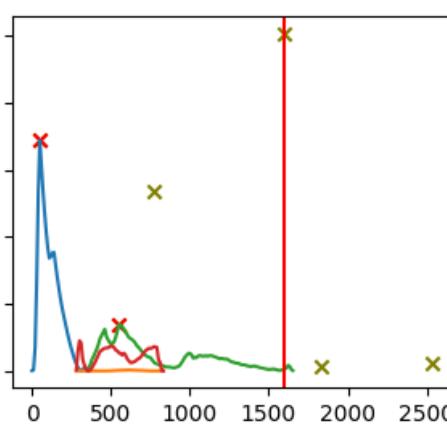
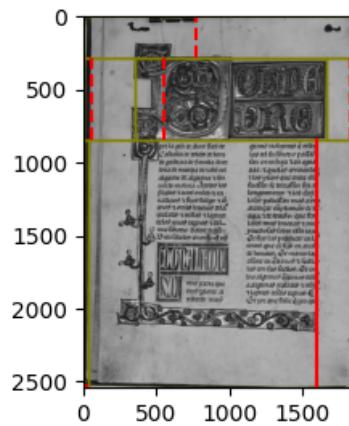
---> new split at 351

DEPTH = 15

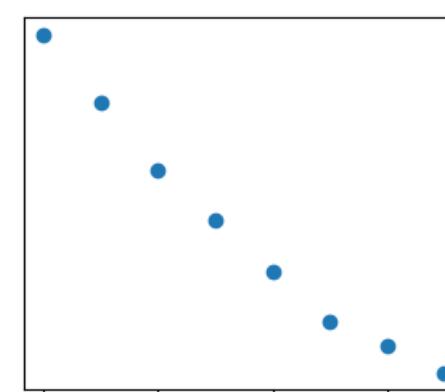
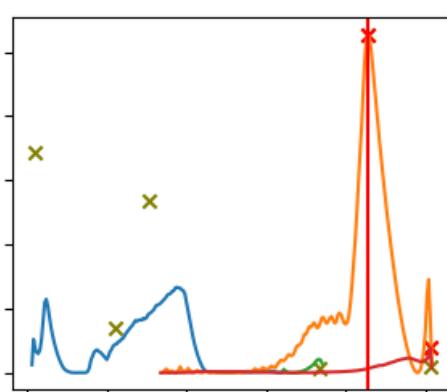
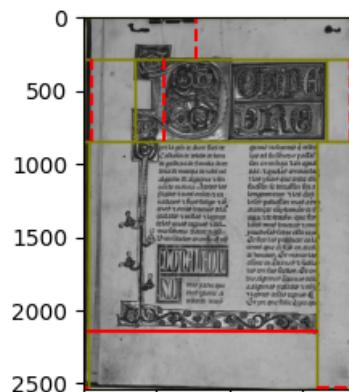




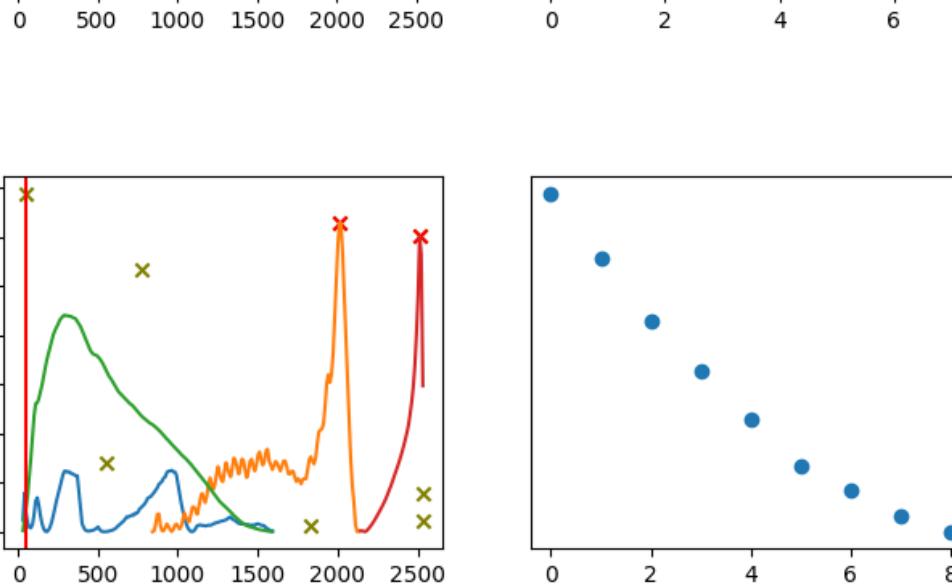
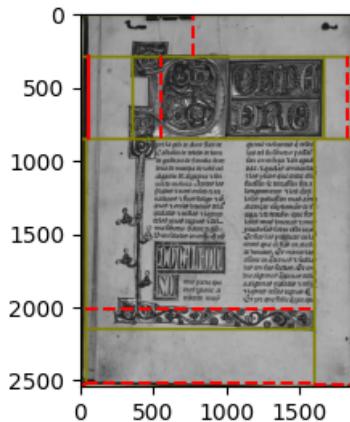
```
diff = 491.16927585299345
---> new split at 1602
DEPTH = 14
```



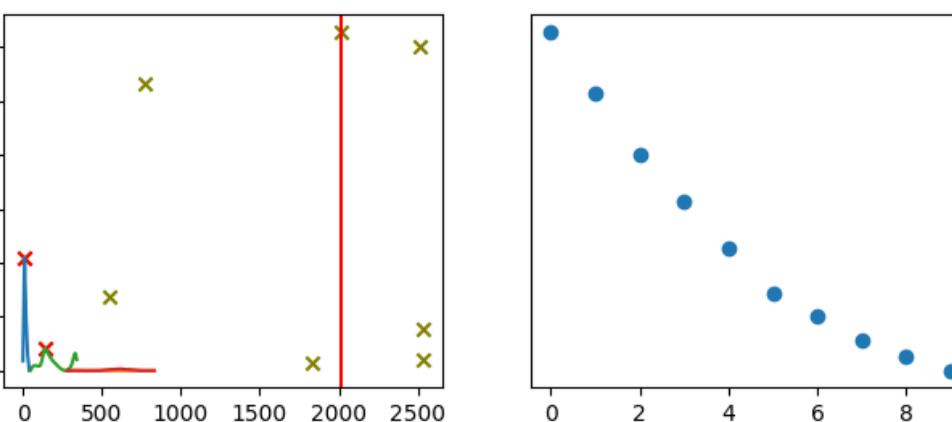
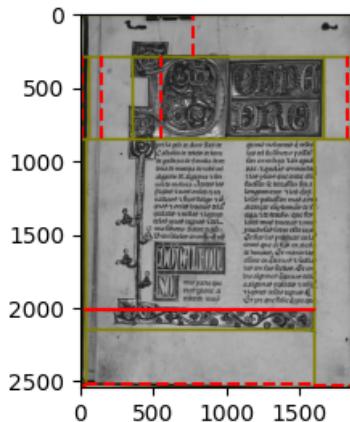
```
diff = 251.79222384681998
---> new split at 2142
DEPTH = 13
```



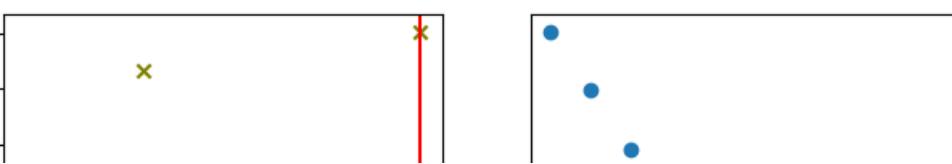
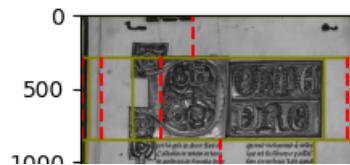
```
0 500 1000 1500
diff = 263.59640190947357
---> new split at 51
DEPTH = 12
```

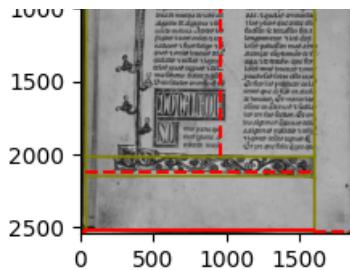


```
0 500 1000 1500
diff = 172.04123349004294
---> new split at 2012
DEPTH = 11
```



```
0 500 1000 1500
diff = 157.25287890072832
---> new split at 2513
DEPTH = 10
```

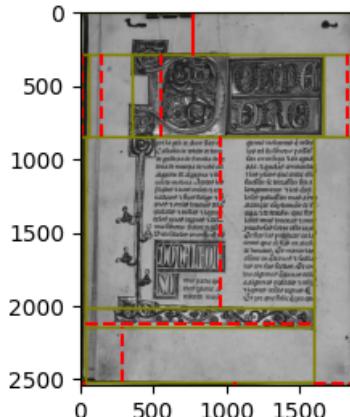




diff = 150.57386627243432

---> new split at 771

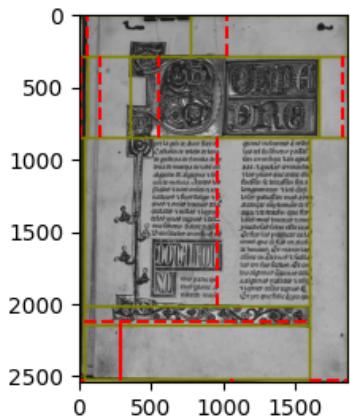
DEPTH = 9



diff = 133.46645872679073

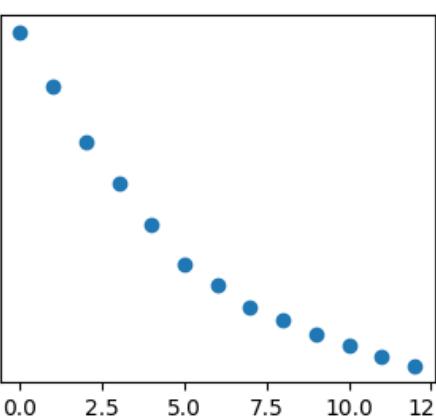
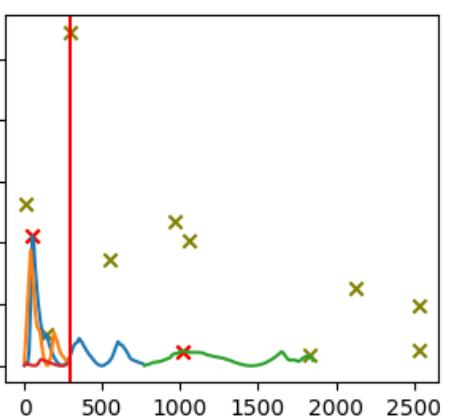
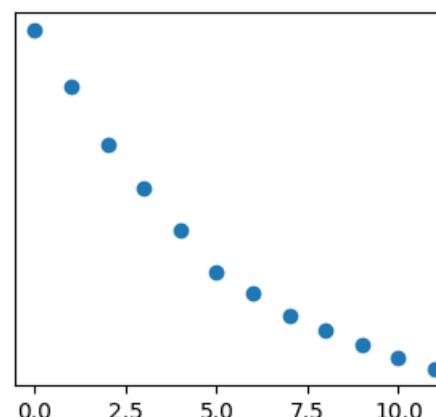
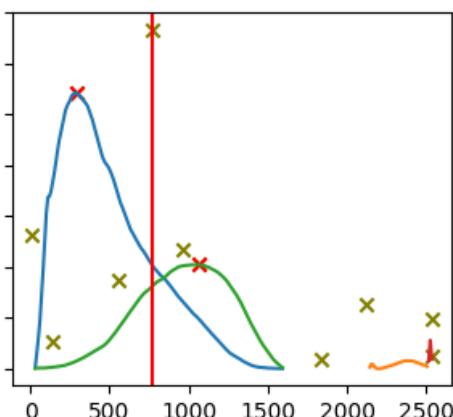
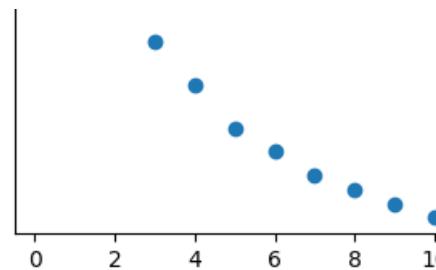
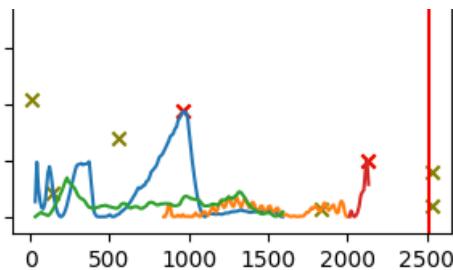
---> new split at 292

DEPTH = 8

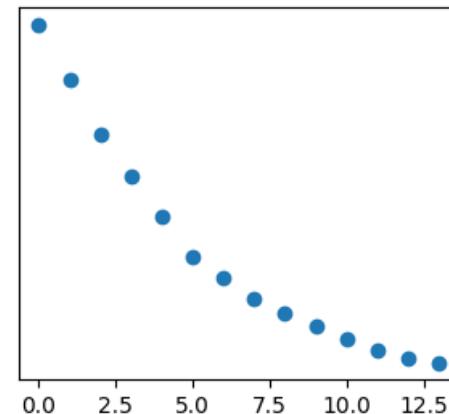
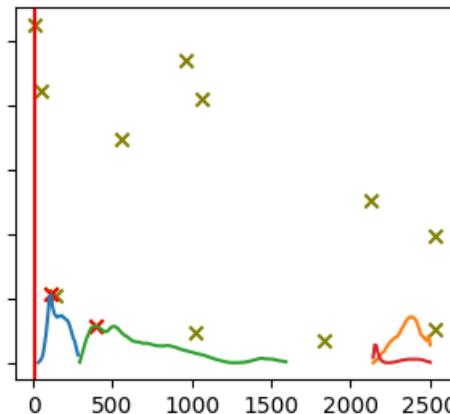
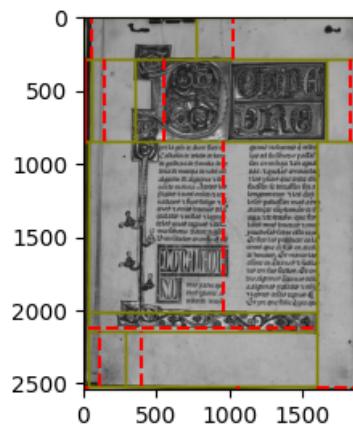


diff = 108.66003561545858

---> new split at 11



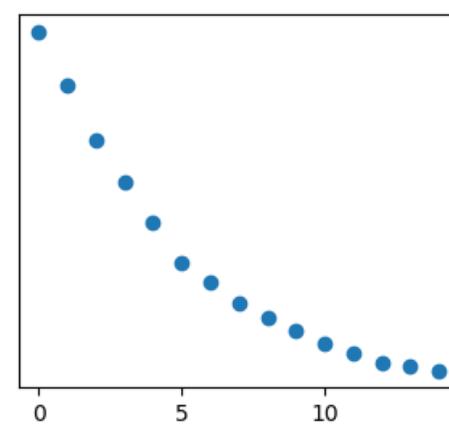
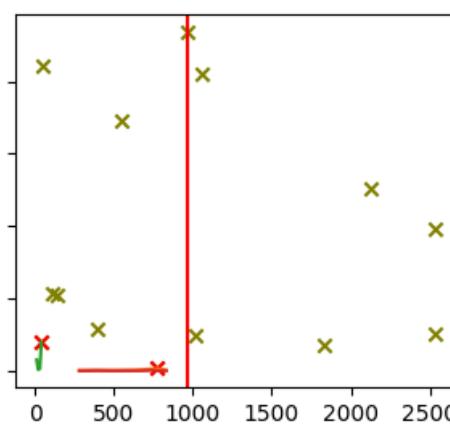
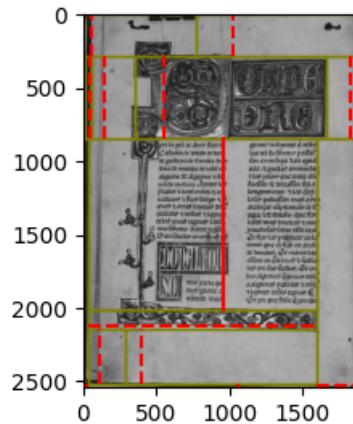
--> new split at 1123
DEPTH = 7



diff = 52.462722383055734

--> new split at 962

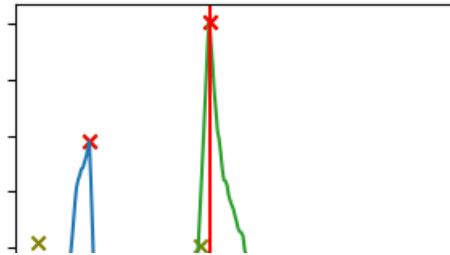
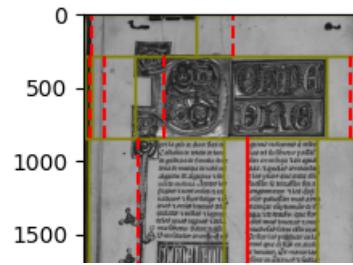
DEPTH = 6

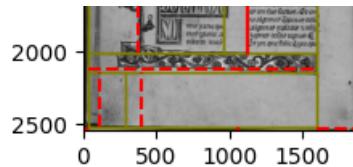


diff = 46.868798692717064

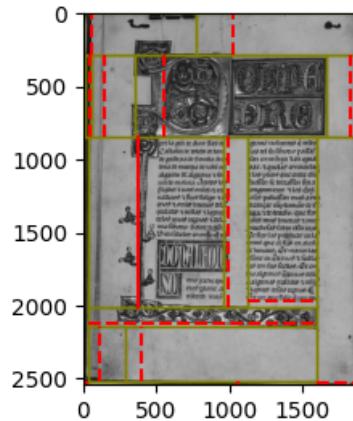
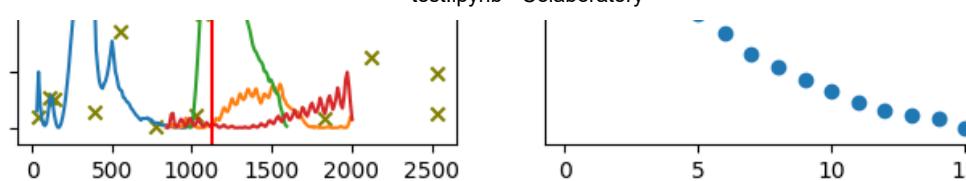
--> new split at 1123

DEPTH = 5

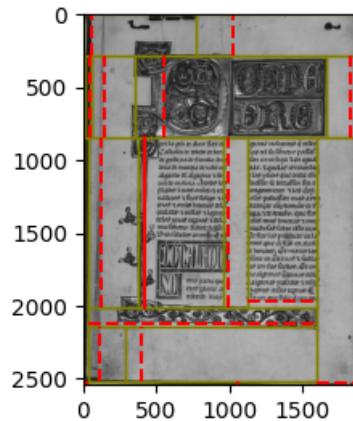
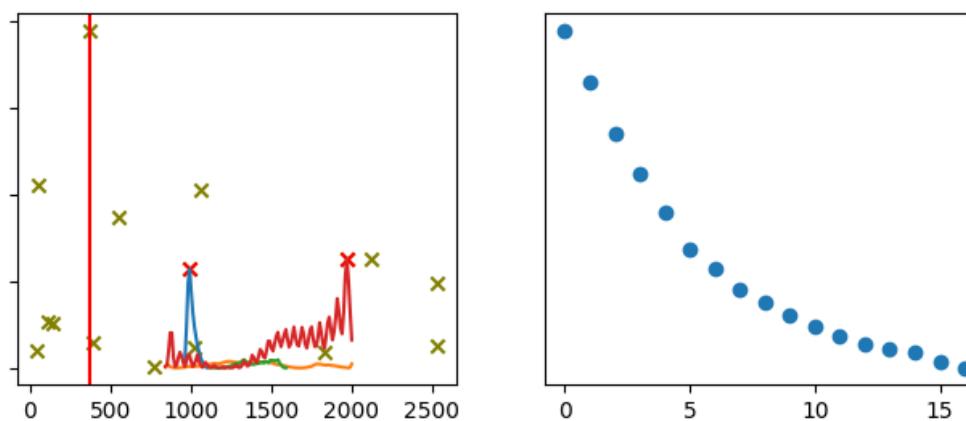




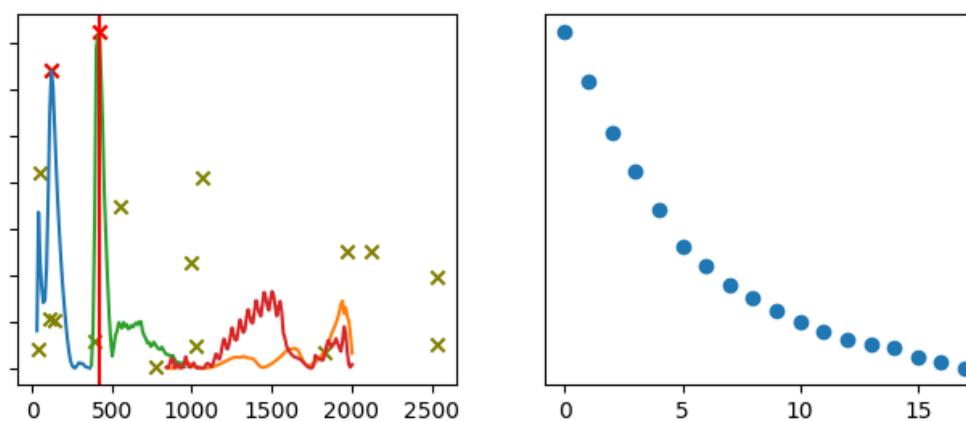
diff = 120.79684262422052
---> new split at 372
DEPTH = 4

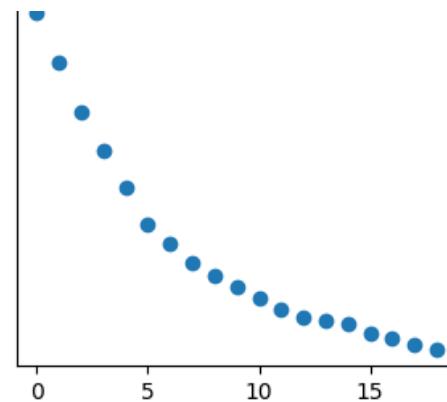
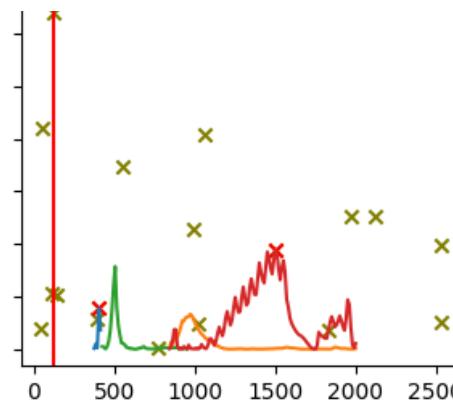
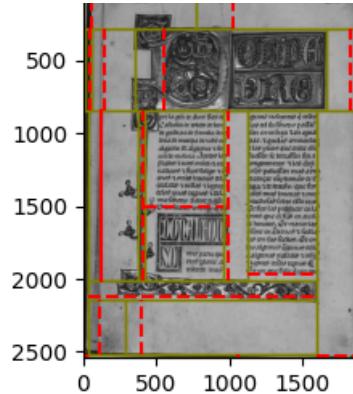


diff = 77.87258099501105
---> new split at 423
DEPTH = 3

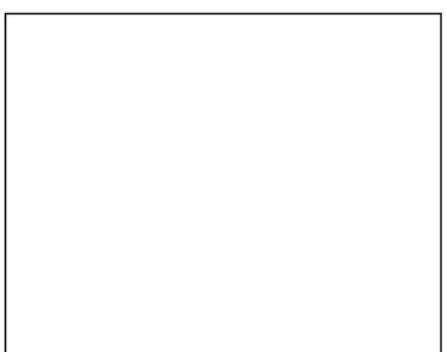
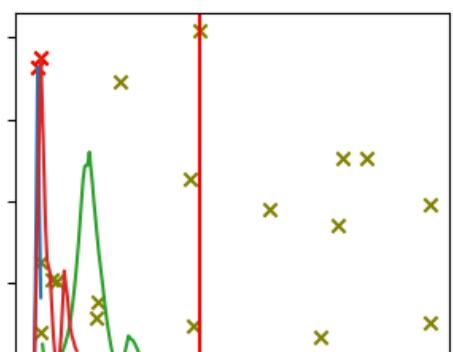
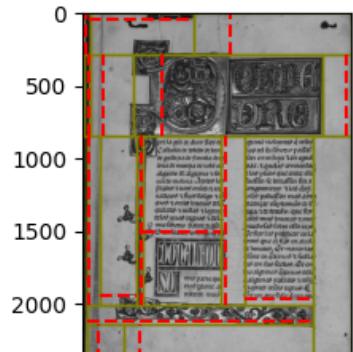
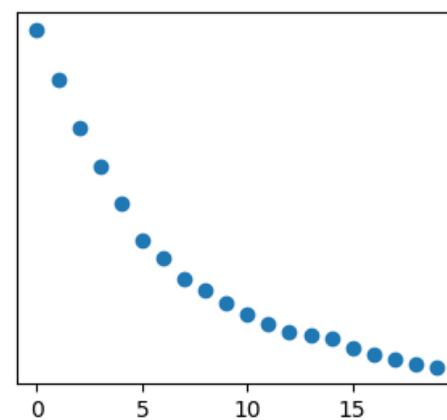
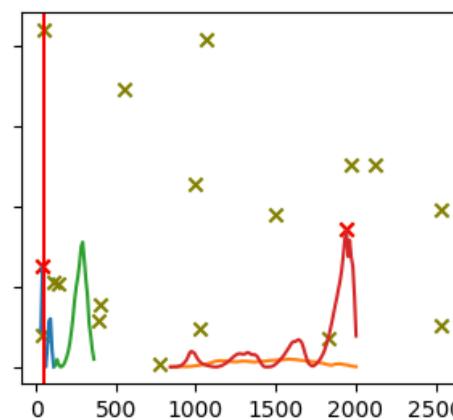
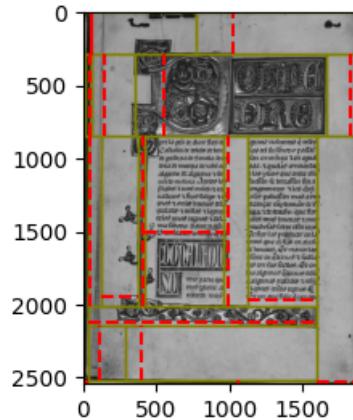


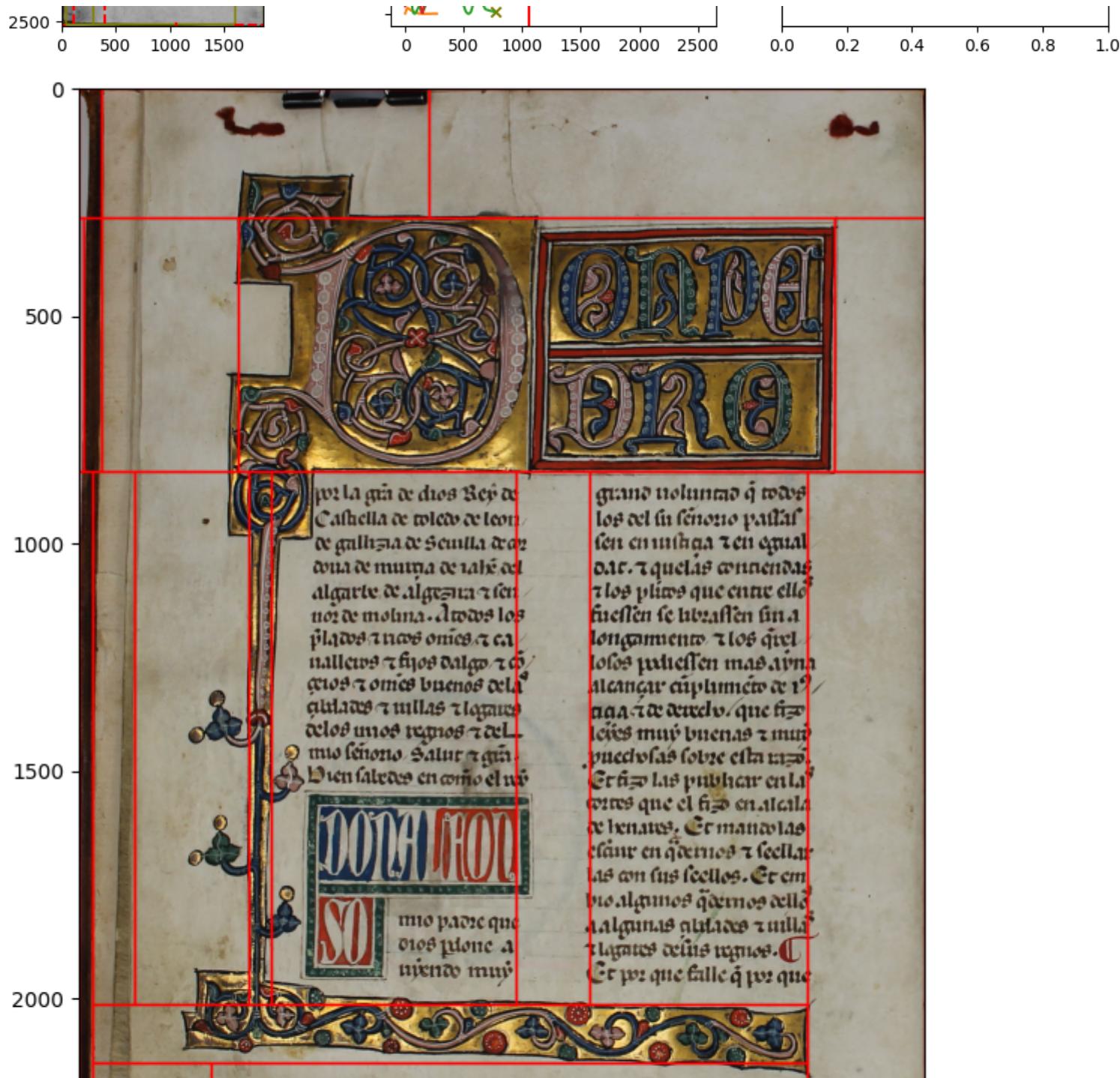
diff = 72.52165587502168
---> new split at 122
DEPTH = 2

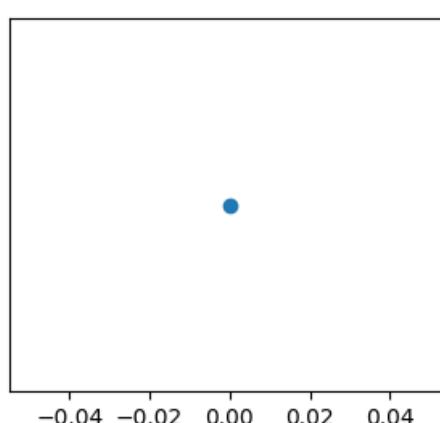
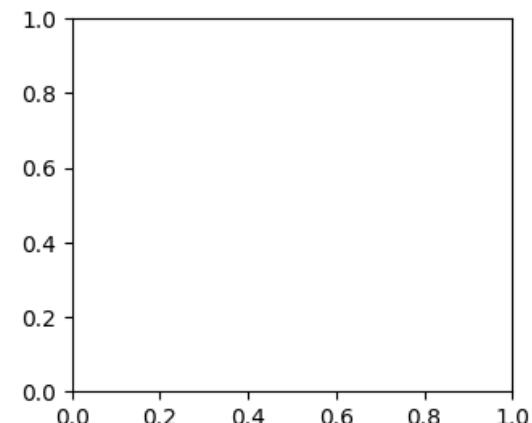
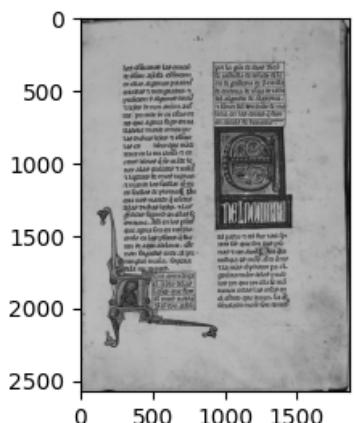
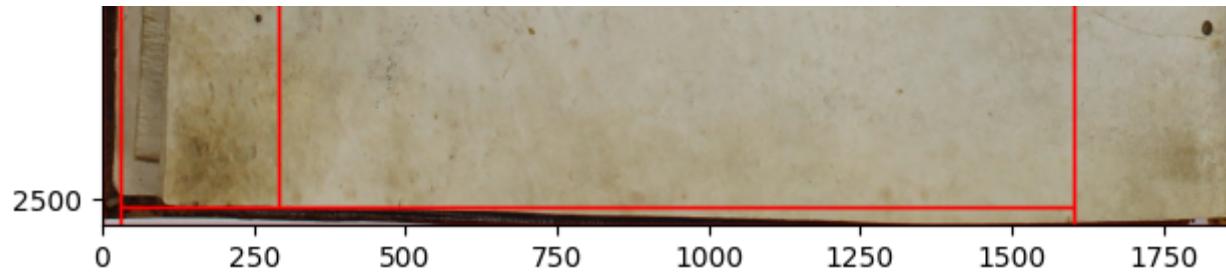




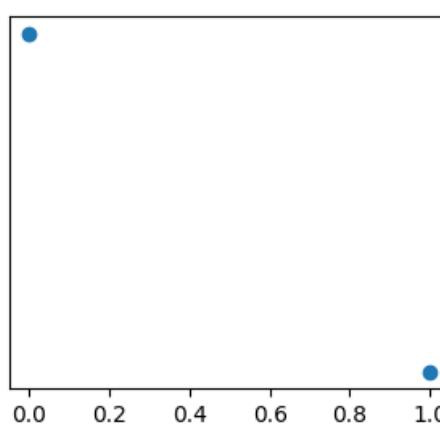
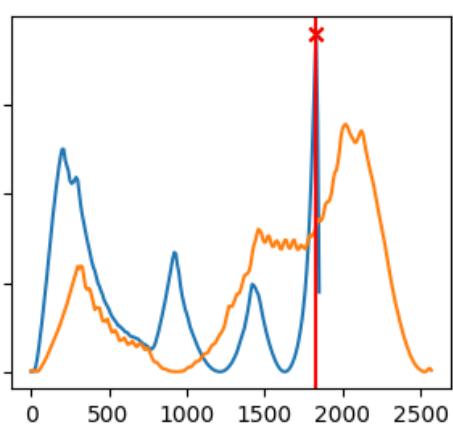
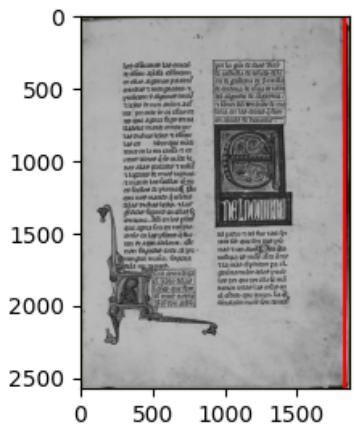
```
diff = 63.99648623164103
---> new split at 51
DEPTH = 1
```





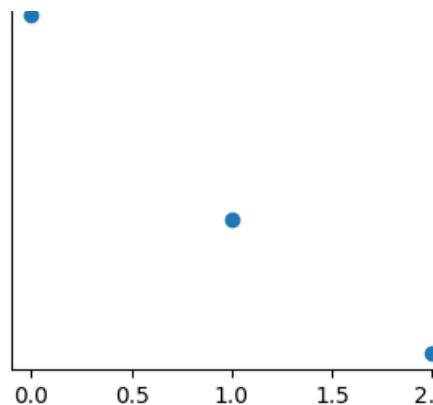
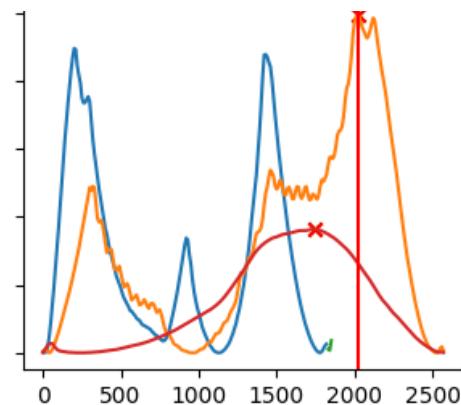
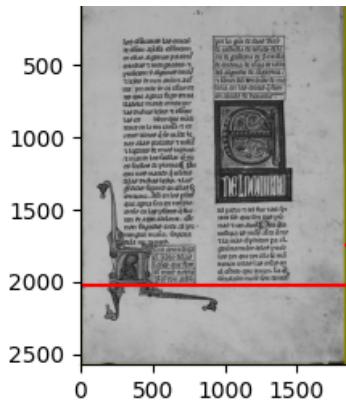


---> new split at 1831
DEPTH = 19



diff = 379.9946107613614
---> new split at 2021
DEPTH = 18

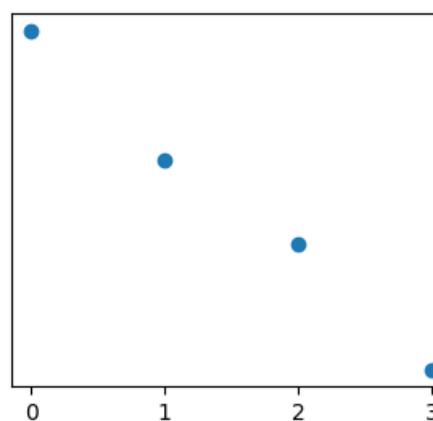
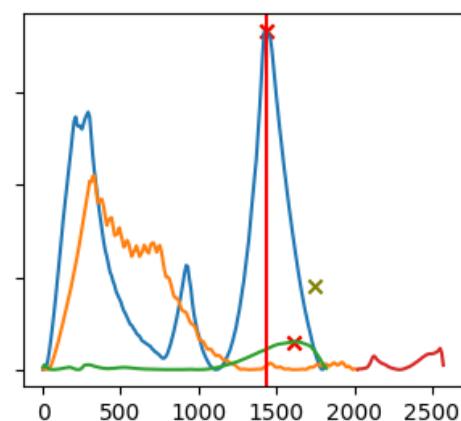
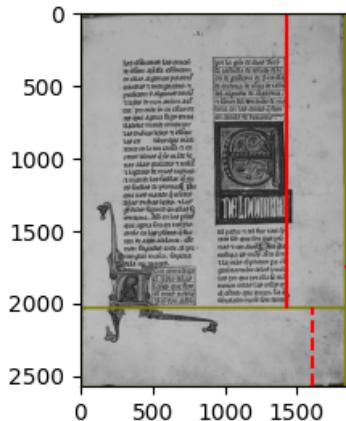




diff = 248.74578951224794

---> new split at 1431

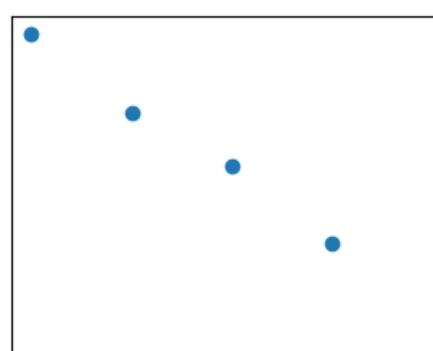
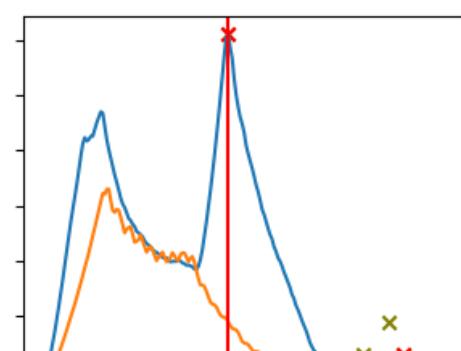
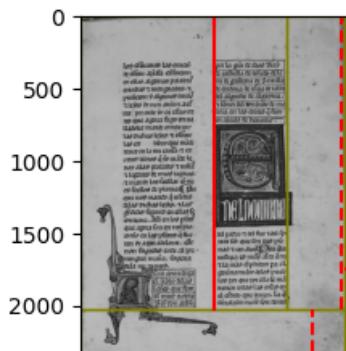
DEPTH = 17

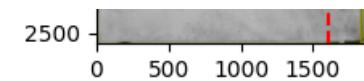


diff = 367.0210682956804

---> new split at 931

DEPTH = 16

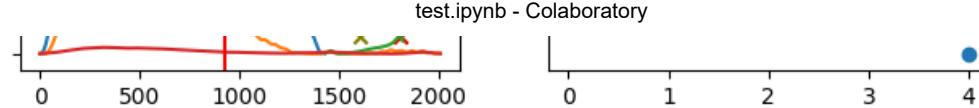
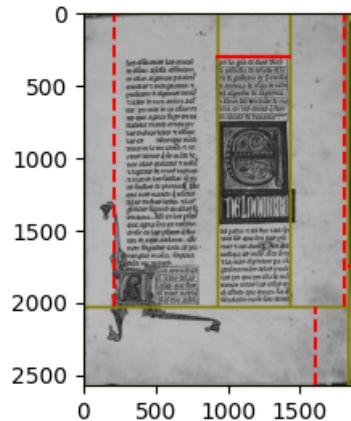




diff = 611.3688794841

---> new split at 301

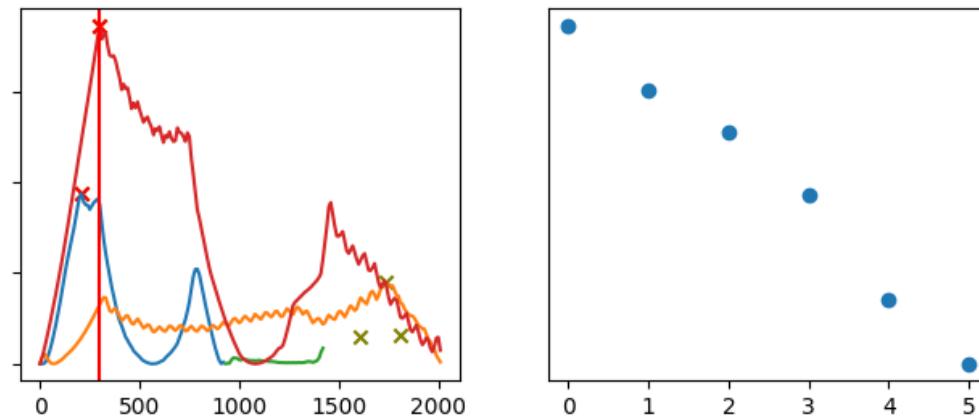
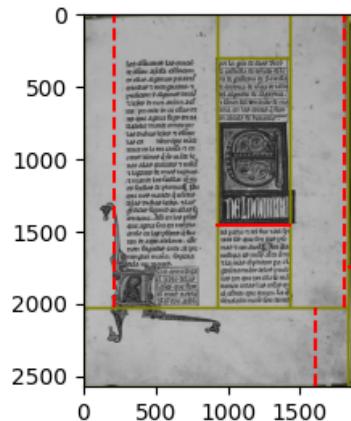
DEPTH = 15



diff = 372.99707603494244

---> new split at 1462

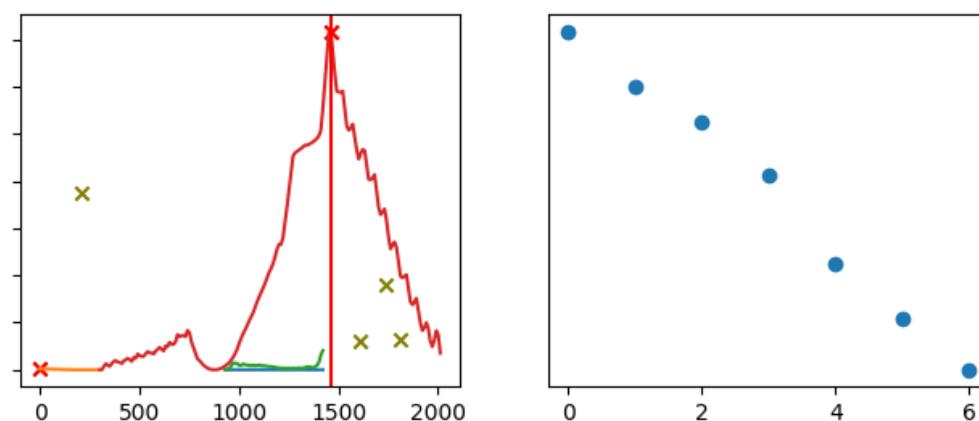
DEPTH = 14

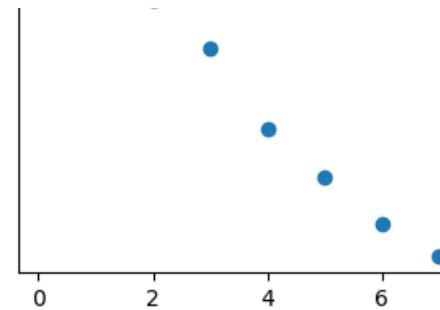
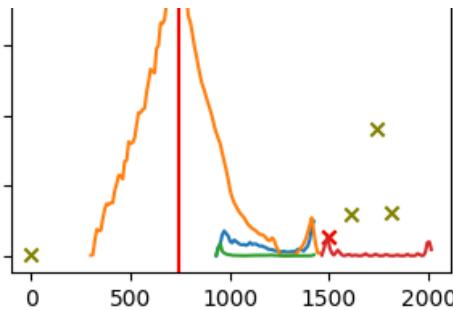
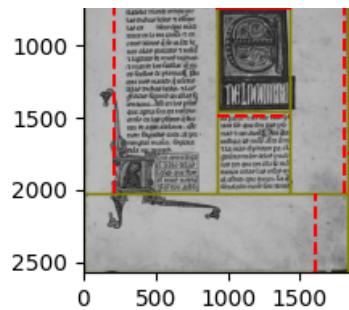


diff = 358.51659977362124

---> new split at 742

DEPTH = 13

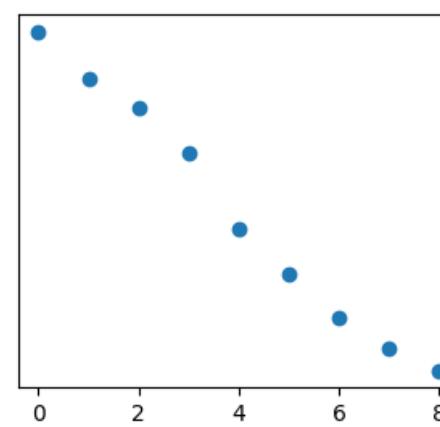
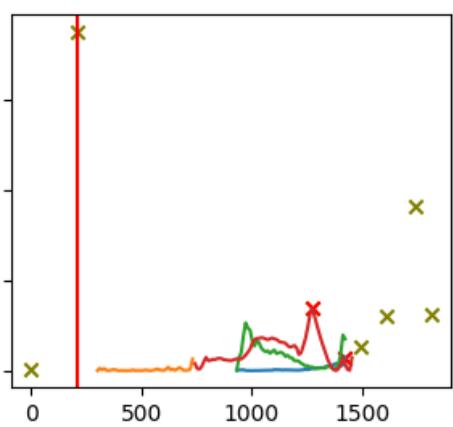
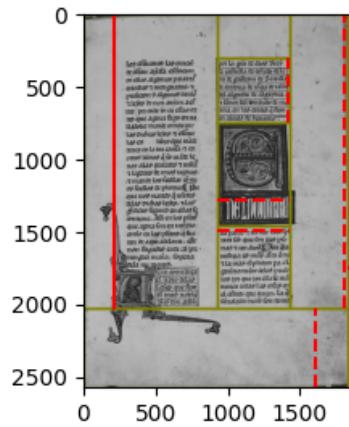




diff = 240.87164593621856

---> new split at 211

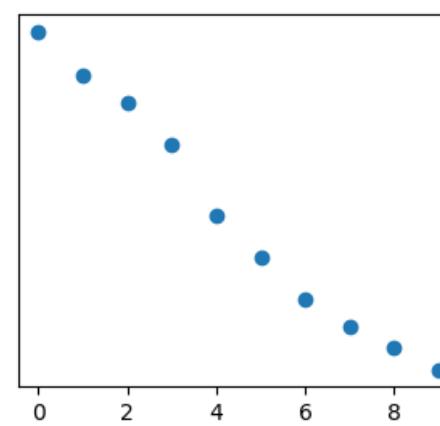
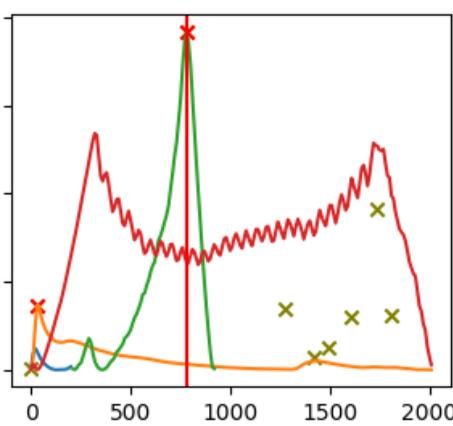
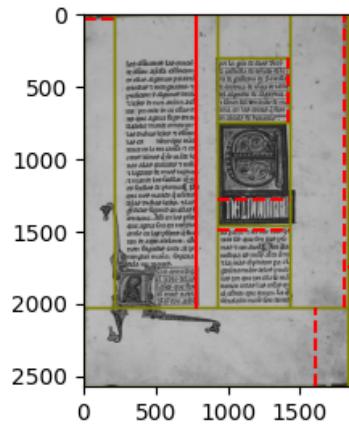
DEPTH = 12



diff = 187.3428396010786

---> new split at 782

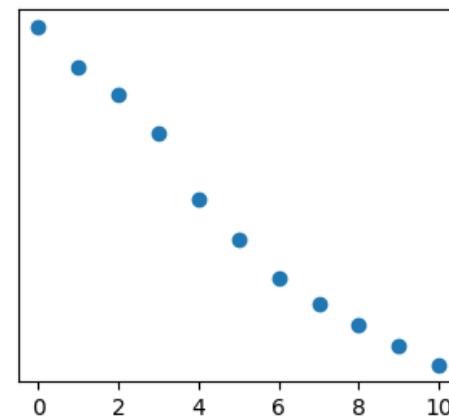
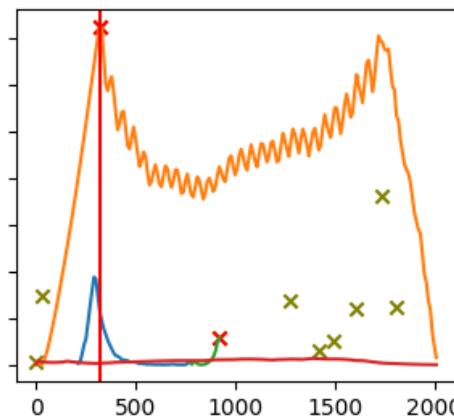
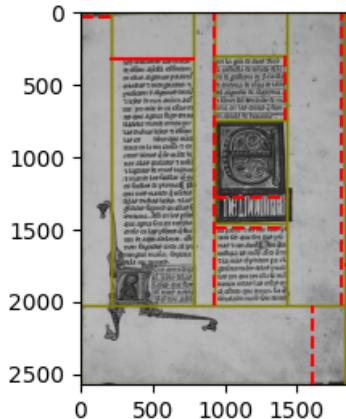
DEPTH = 11



```
diff = 191.94570592394393
```

---> new split at 321

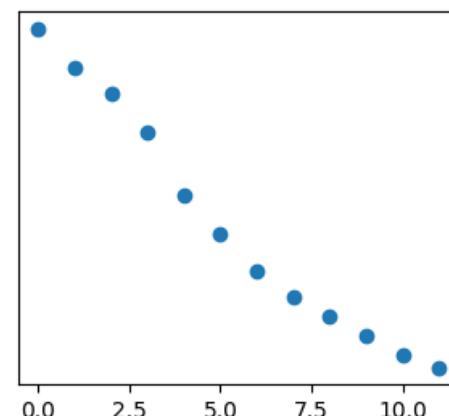
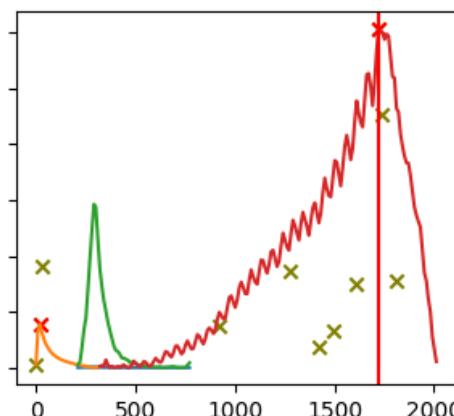
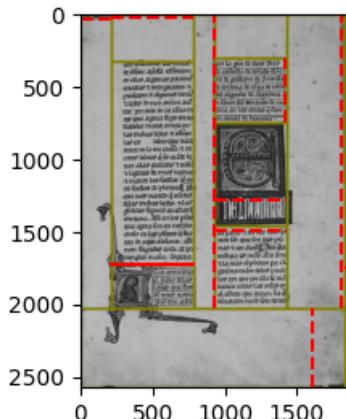
DEPTH = 10



```
diff = 181.4050631641112
```

---> new split at 1722

DEPTH = 9

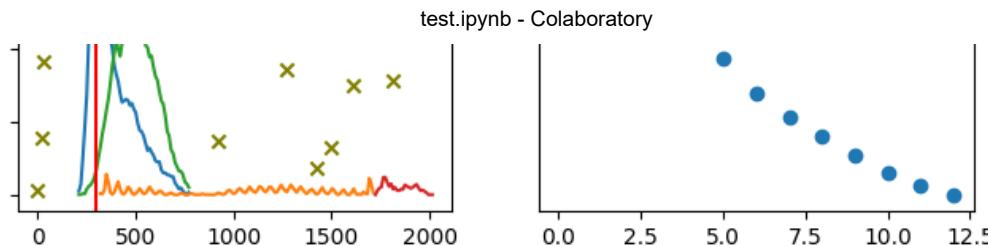
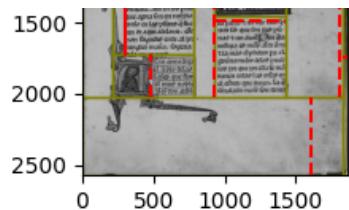


```
diff = 121.40531243320038
```

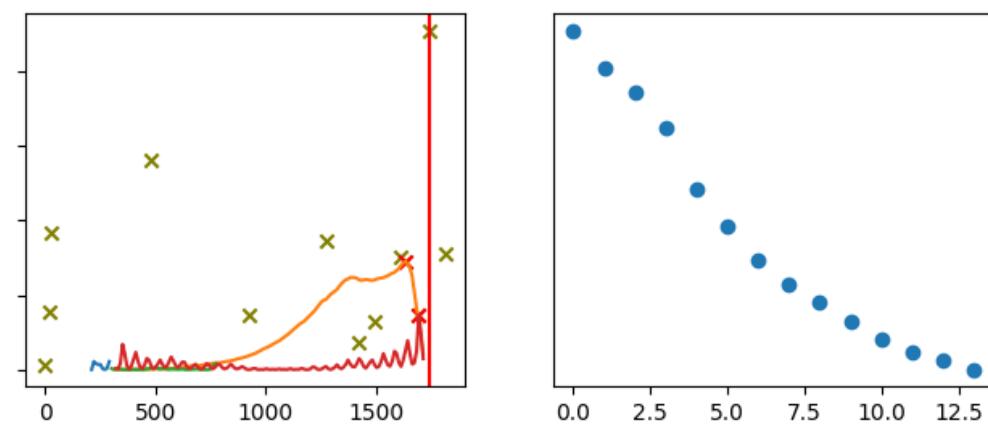
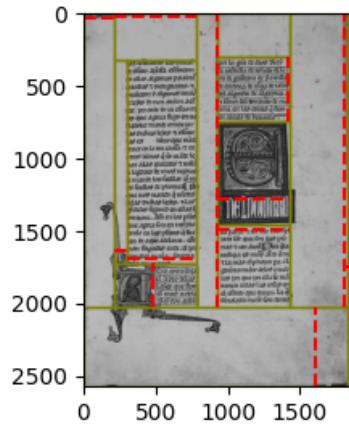
---> new split at 302

DEPTH = 8

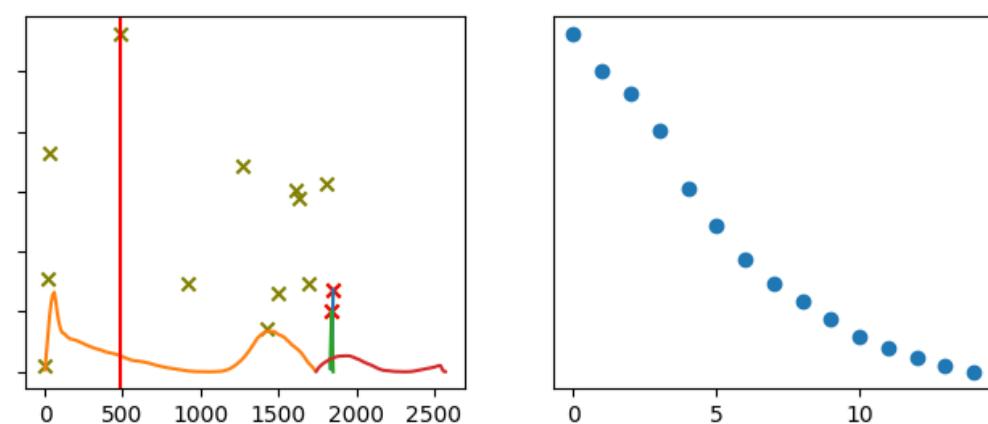
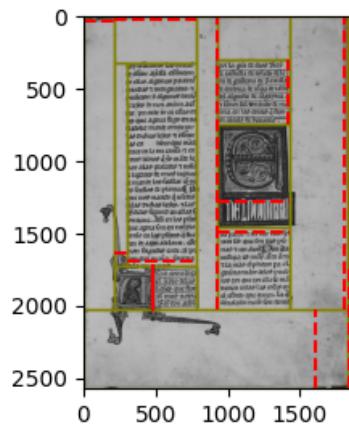




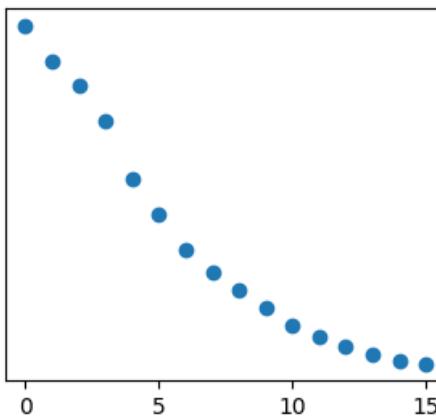
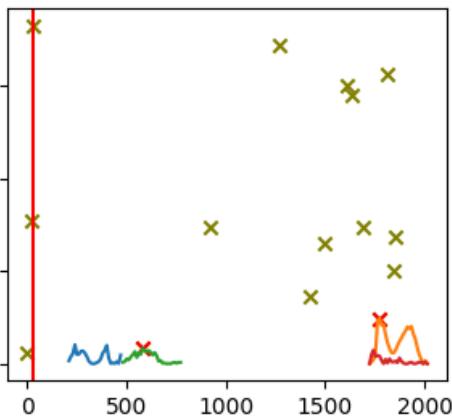
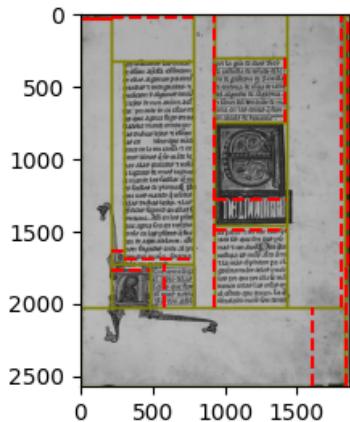
diff = 91.51289483761411
---> new split at 1741
DEPTH = 7



diff = 90.87923178068559
---> new split at 482
DEPTH = 6



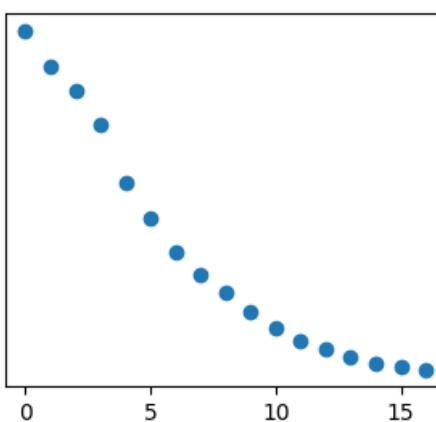
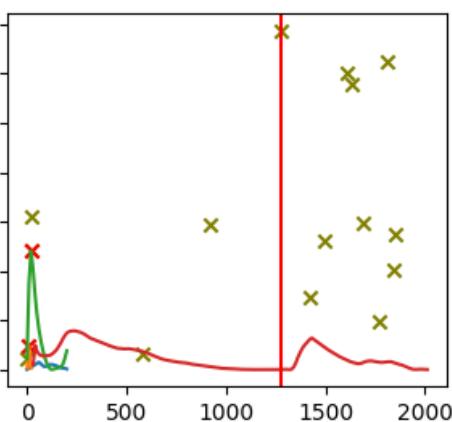
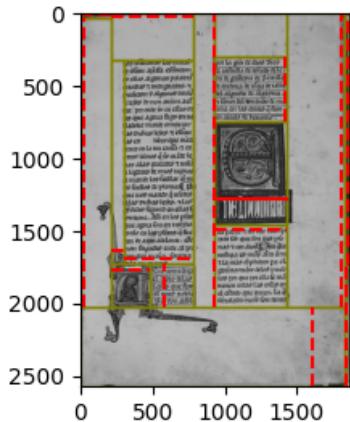
diff = 56.30874143492292
---> new split at 31
DEPTH = 5



diff = 36.52265293290475

---> new split at 1273

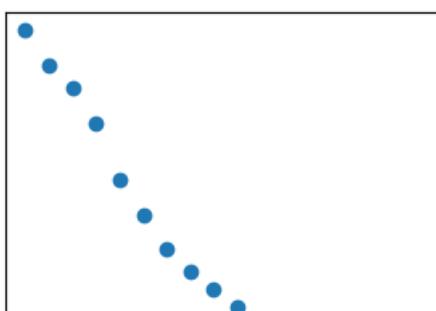
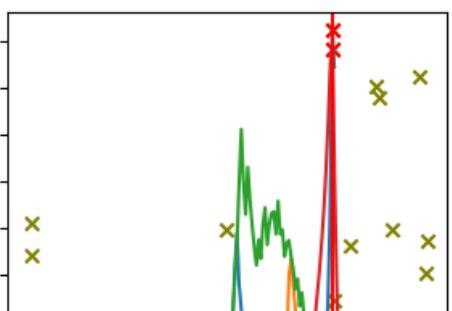
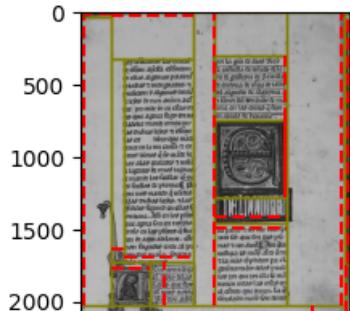
DEPTH = 4

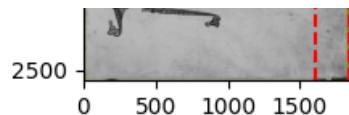


diff = 34.36638615534321

---> new split at 1412

DEPTH = 3

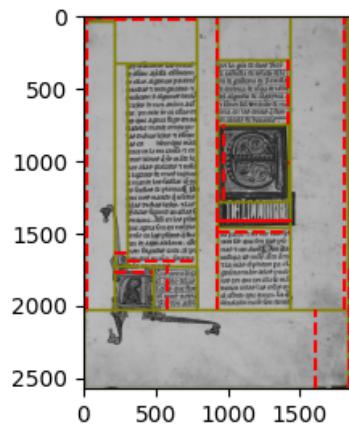




diff = 36.281329975740846

---> new split at 1414

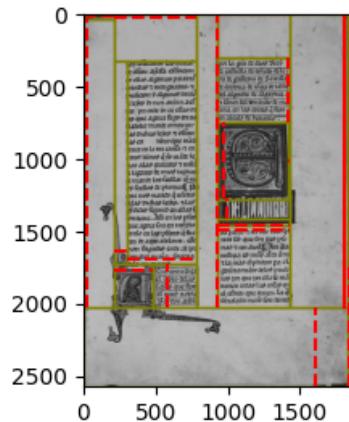
DEPTH = 2



diff = 34.22643661016173

---> new split at 1812

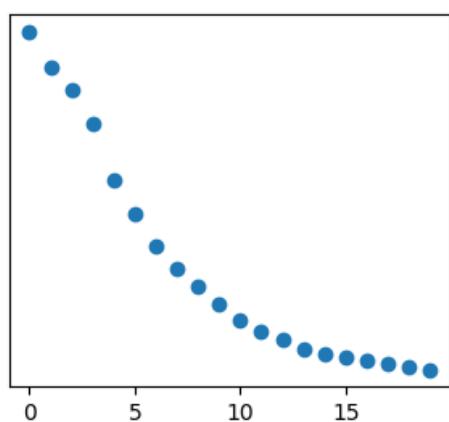
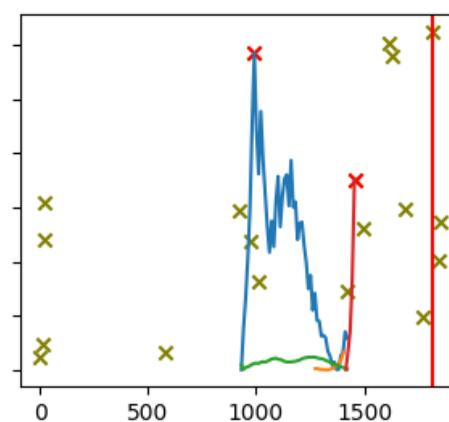
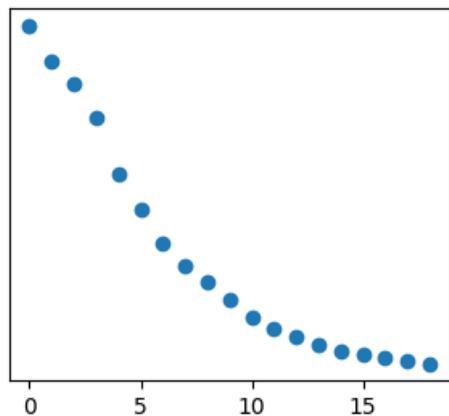
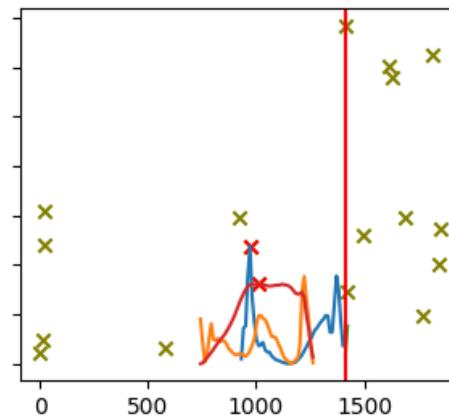
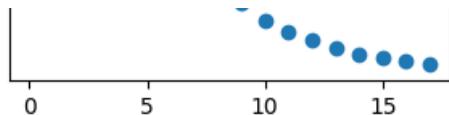
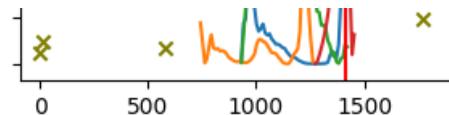
DEPTH = 1

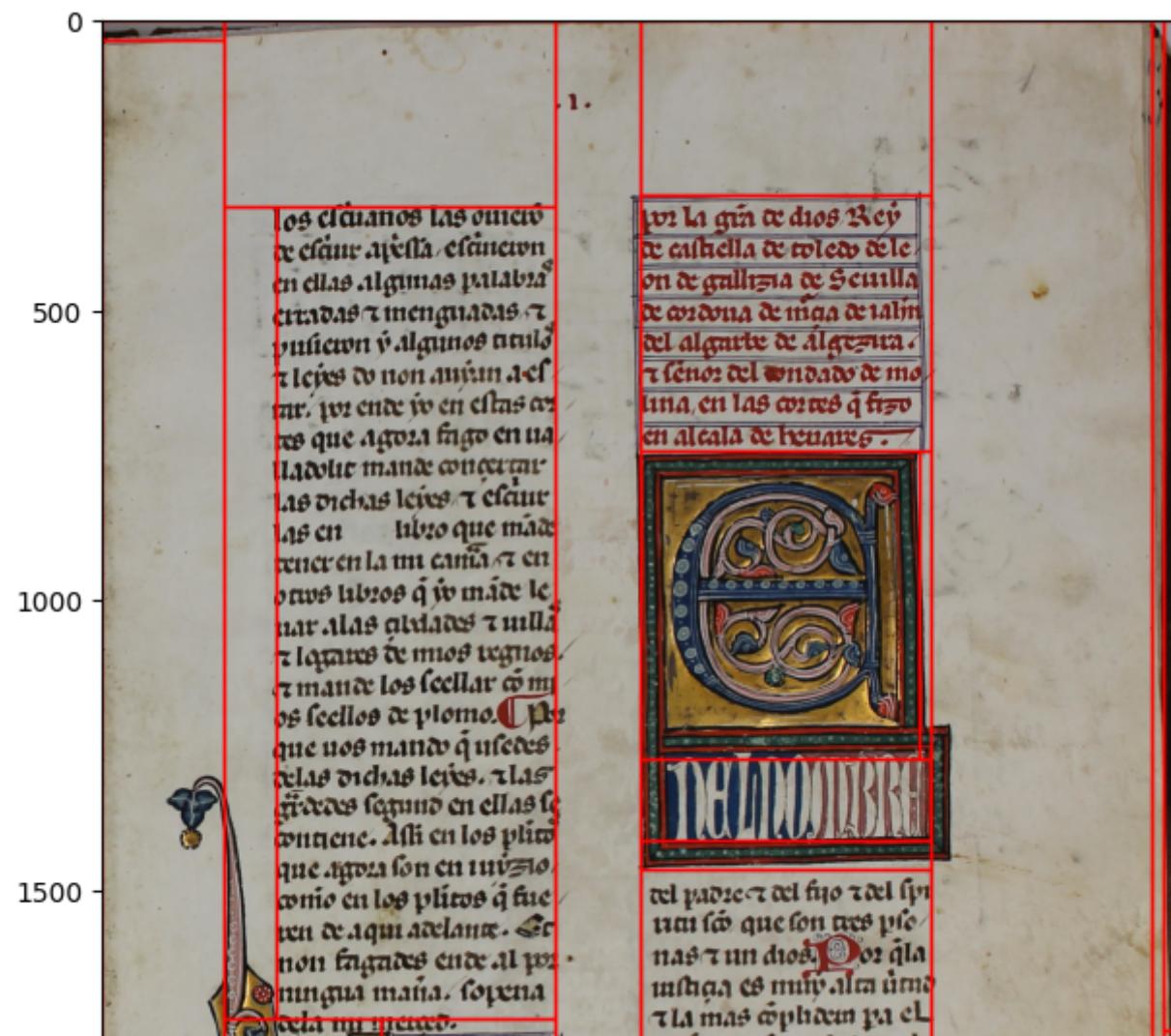
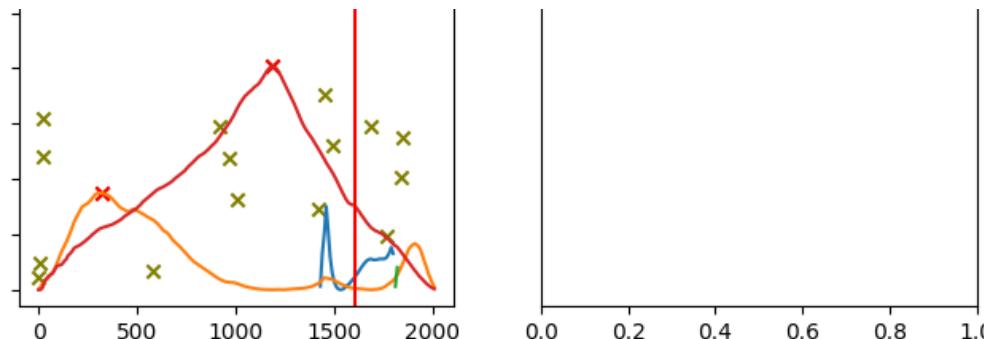
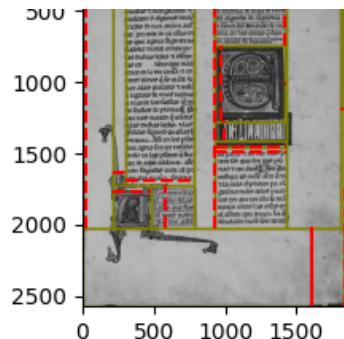


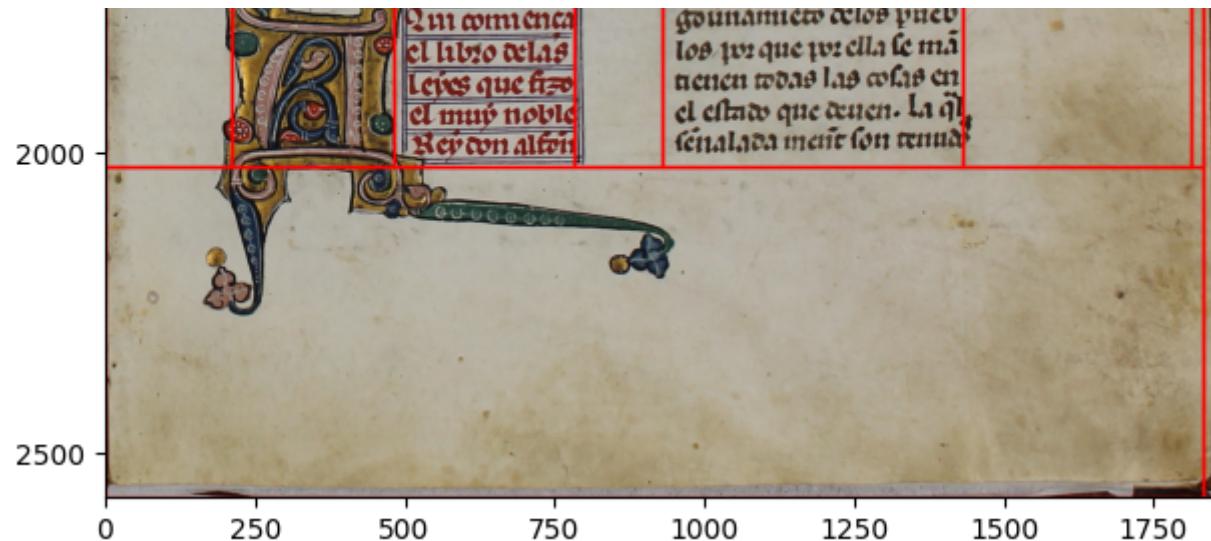
diff = 31.22829796527276

---> new split at 1611

[[0, 0, 211, 31], [0, 31, 211, 2021], [211, 0, 782, 321], [211, 321, 302, 1722], [302, 321, 782, 1722]]



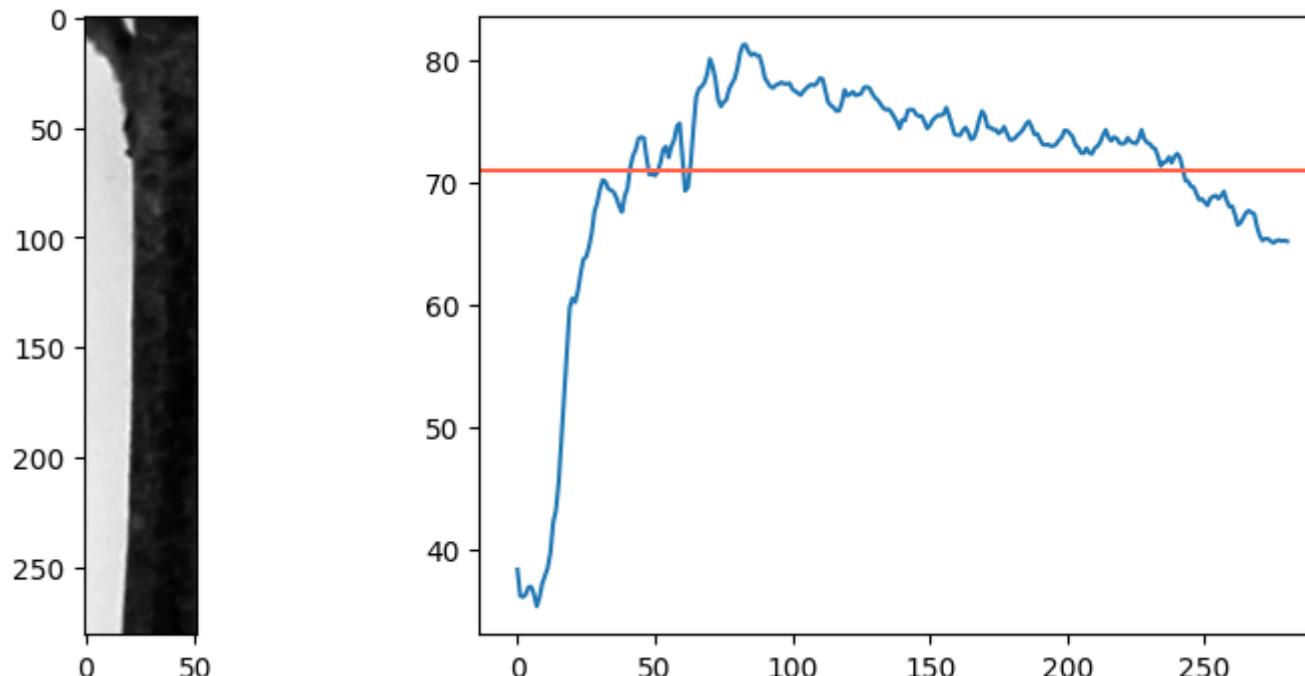





```
check_text(allres,imgs)
```

IMAGE 0

Analyzing area 0 0 51 281

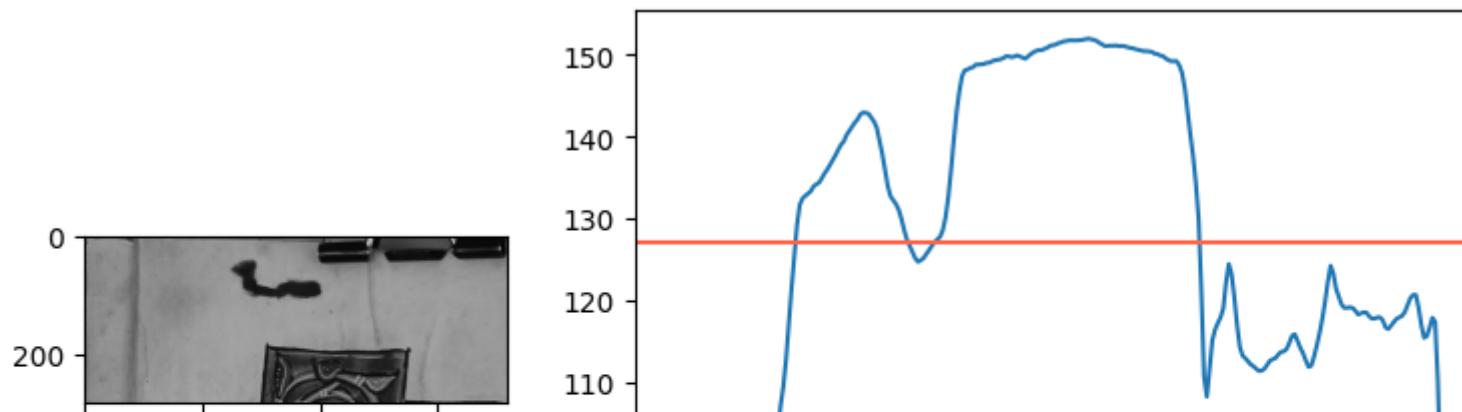


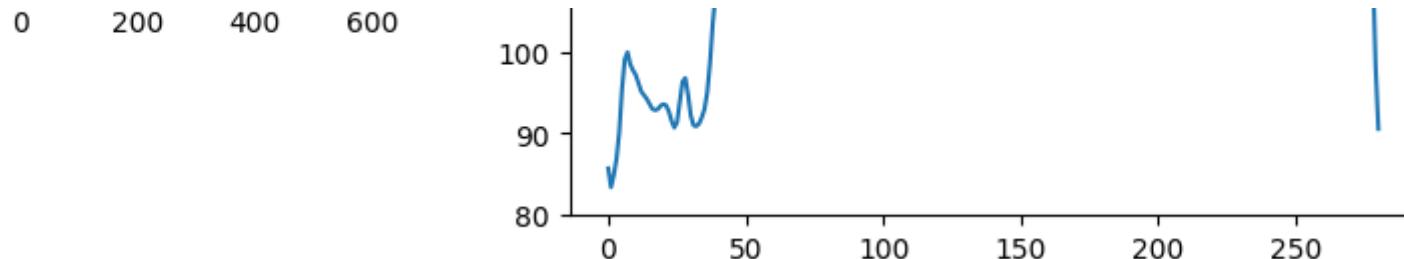
[41, 4, 2, 38] [7, 9, 180]

18.2944663764757 81.08568855777749

---> not text

Analyzing area 51 0 771 281



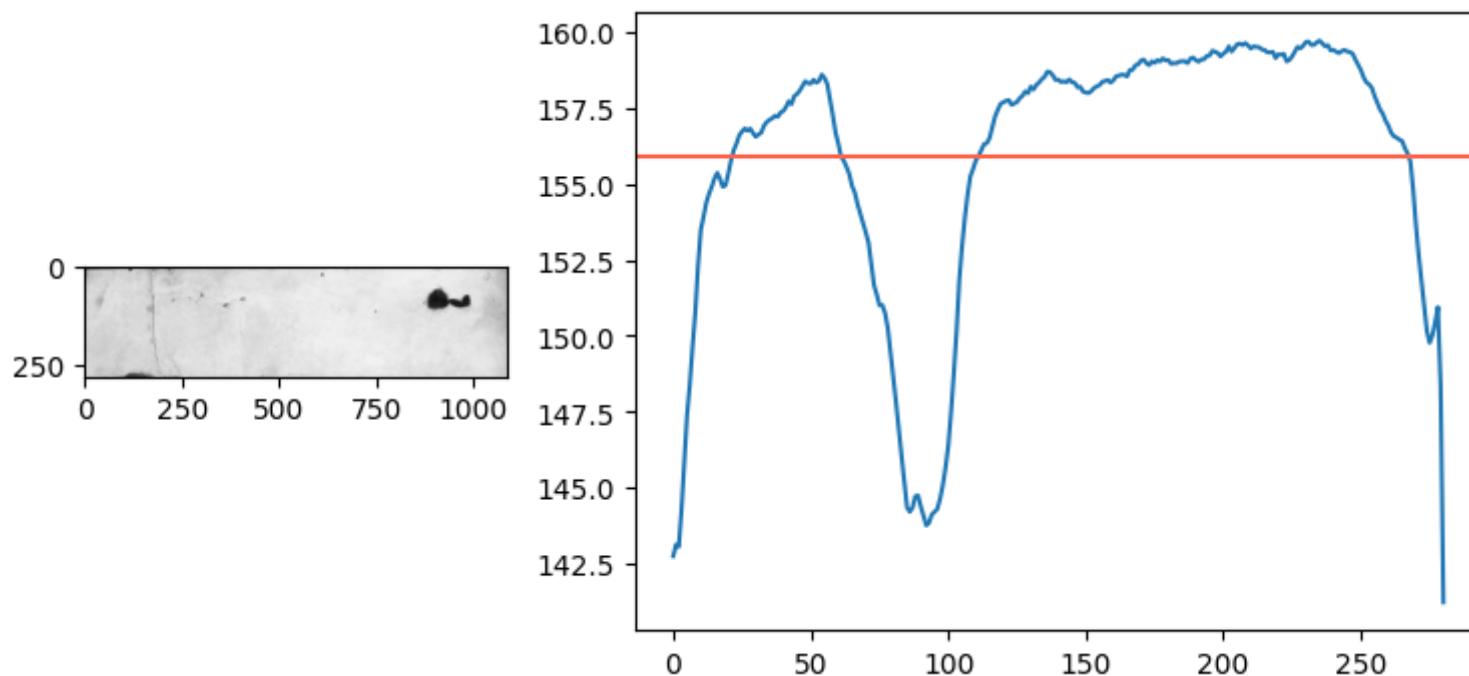


[45, 10, 89] [41, 96]

32.321303604073066 27.5

---> not text

Analyzing area 771 0 1859 281



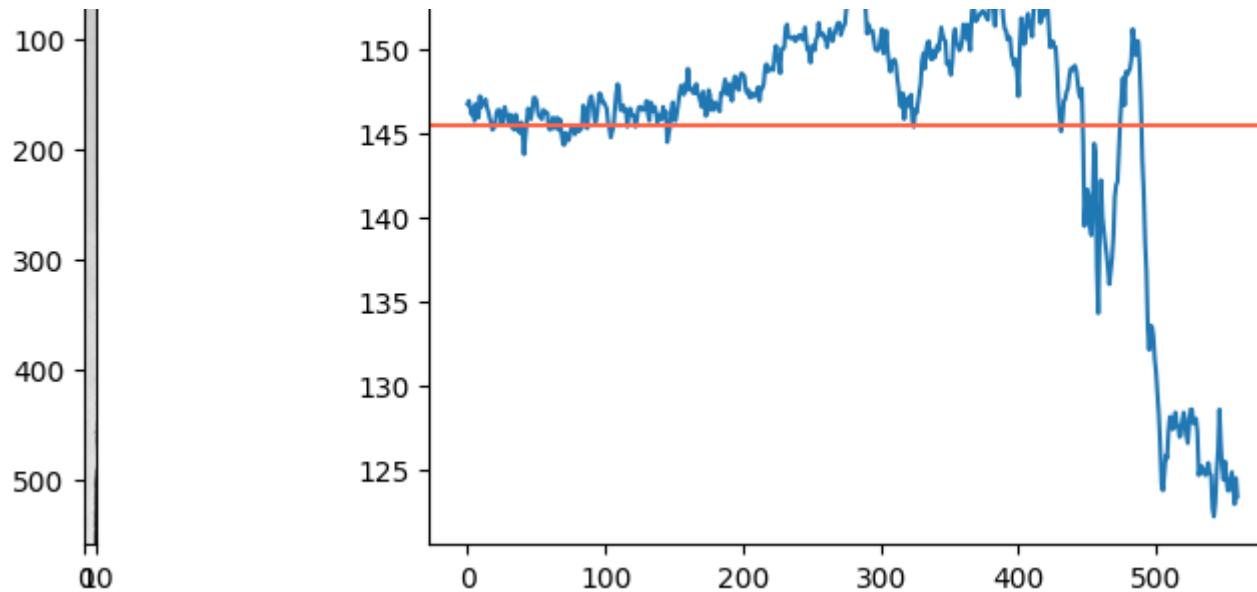
[22, 49, 13] [40, 157]

15.297058540778355 58.5

---> not text

Analyzing area 0 281 11 841



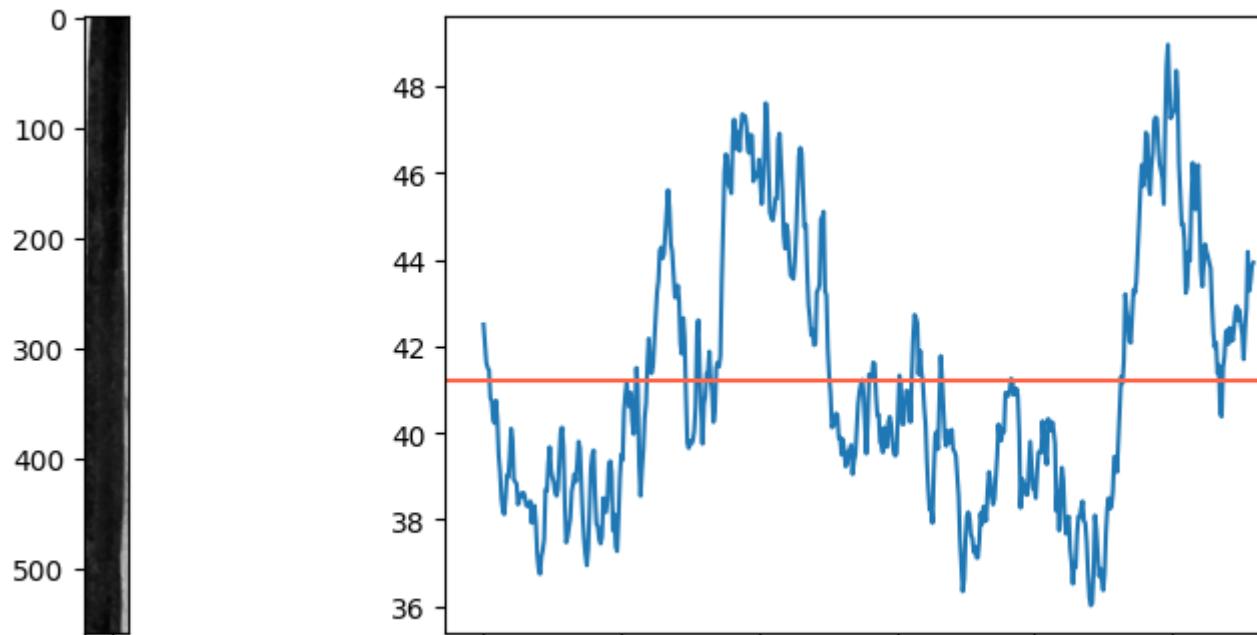


```
[18, 12, 1, 1, 16, 2, 1, 1, 1, 3, 15, 9, 5, 22, 176, 106, 15, 16] [2, 3, 3, 3, 2, 2, 2, 1, 12, 1, 1]
```

```
15.940853280928838 42.87222528433794
```

---> not text

Analyzing area 11 281 51 841

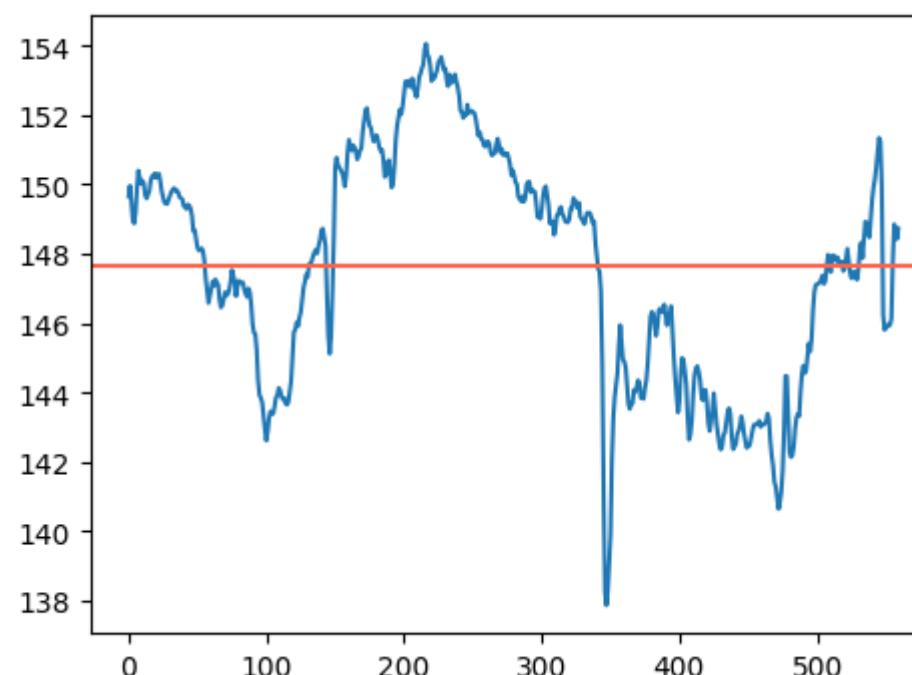
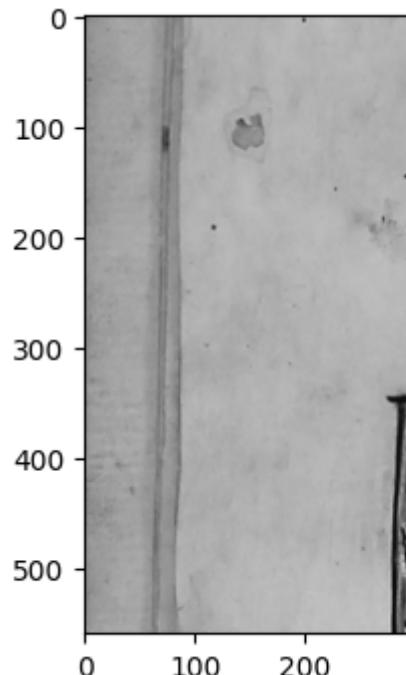


025

0 100 200 300 400 500

[5, 1, 28, 3, 4, 83, 1, 5, 1, 8, 2, 2, 1, 70, 23] [106, 7, 8, 4, 3, 23, 4, 17, 8, 13, 49, 78, 1, 2
31.02870626121718 25.177238397674465
---> not text

Analyzing area 51 281 351 841

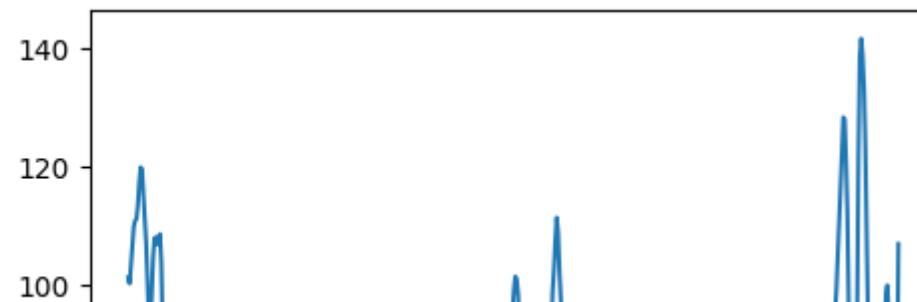


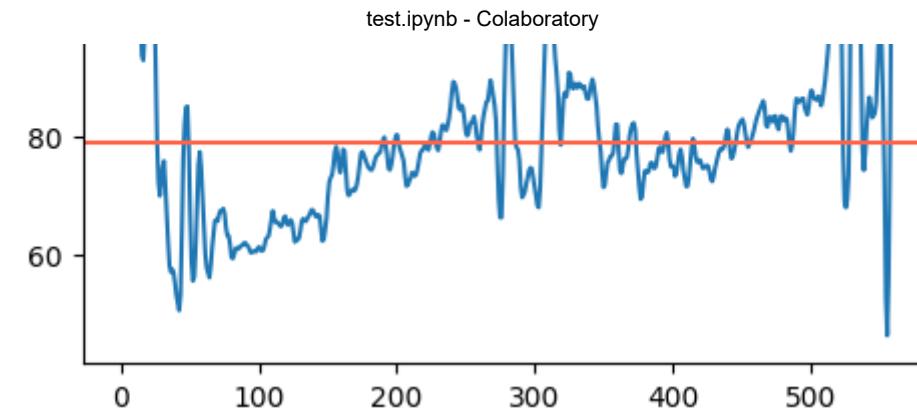
[56, 13, 192, 2, 8, 3, 17, 5] [75, 5, 166, 2, 2, 7, 7]

57.76235621854745 60.83584469702052

---> not text

Analyzing area 351 281 1661 841

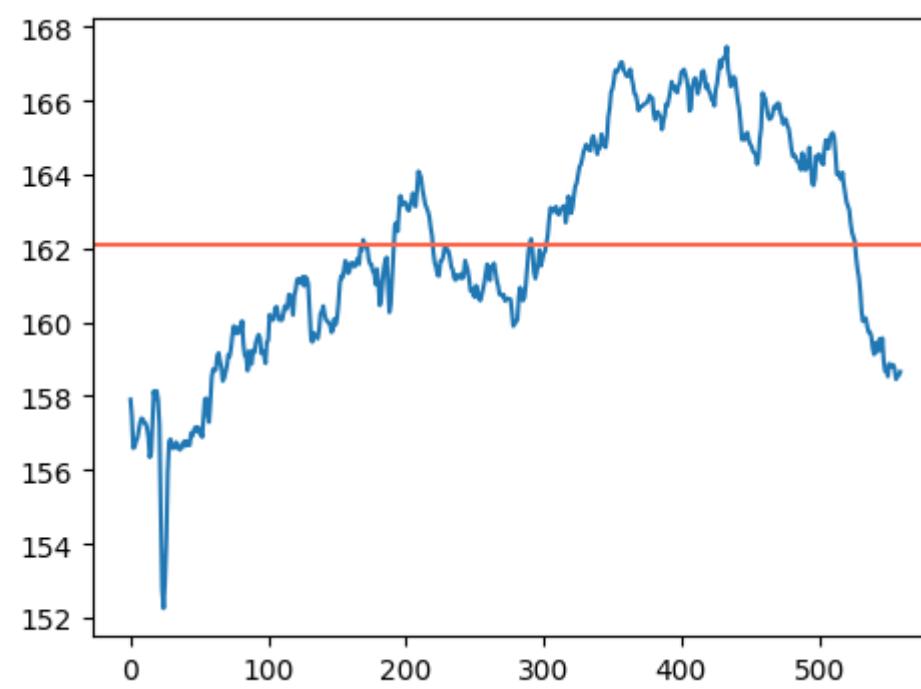
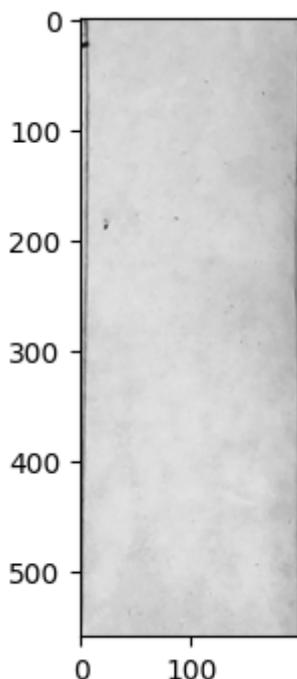




[27, 3, 2, 3, 4, 27, 12, 8, 13, 28, 3, 5, 3, 1, 3, 8, 28, 36, 9, 13, 1] [19, 141, 7, 23, 3, 2, 5, 1]
29.63321616024828 10.75768986151128

---> not text

Analyzing area 1661 281 1859 841

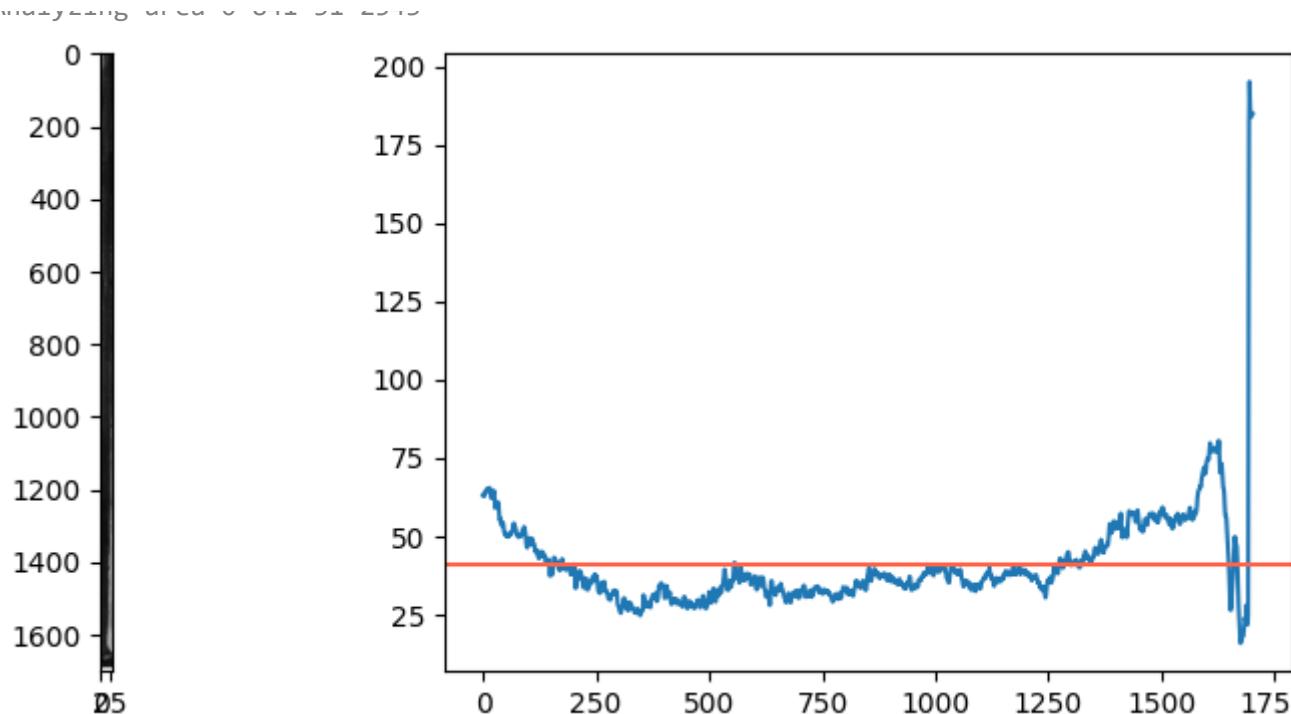


[169, 22, 70, 10, 33] [1, 28, 2, 225]

57.7075385023482 93.58151526877516

---> not text

Analyzing area 0 841 31 2543

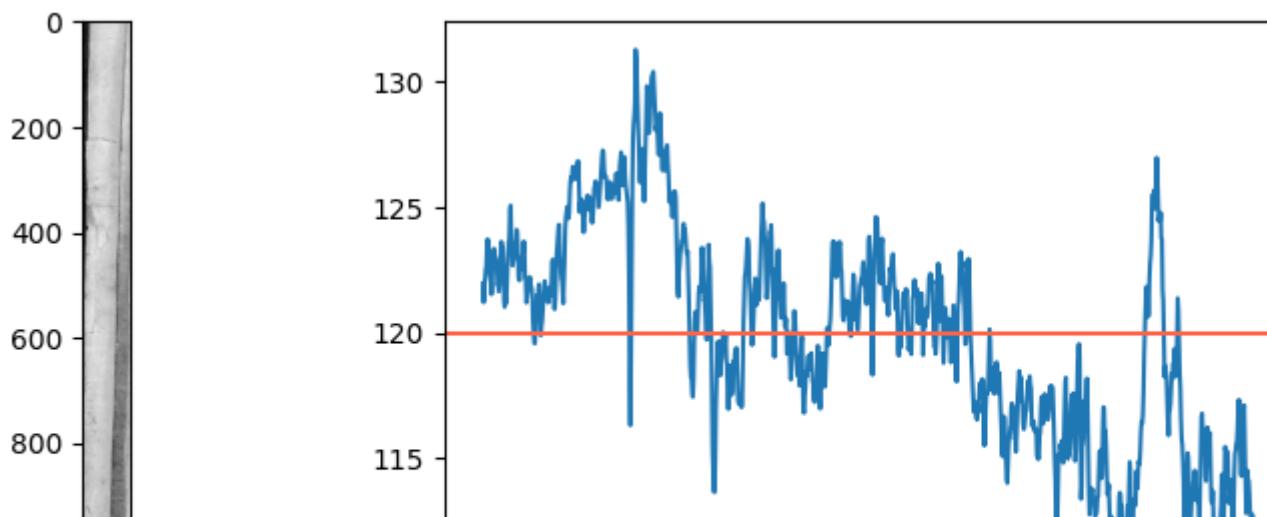


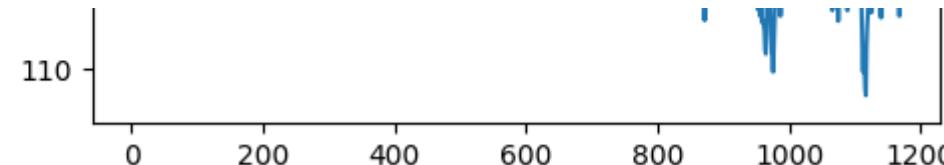
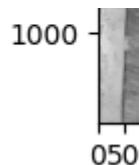
[147, 11, 6, 1, 3, 7, 17, 4, 8, 327, 9, 10] [8, 7, 23, 352, 715, 3, 1, 2, 7, 10, 24]

216.55656028311145 93.13774148479706

---> not text

Analyzing area 31 841 122 2012

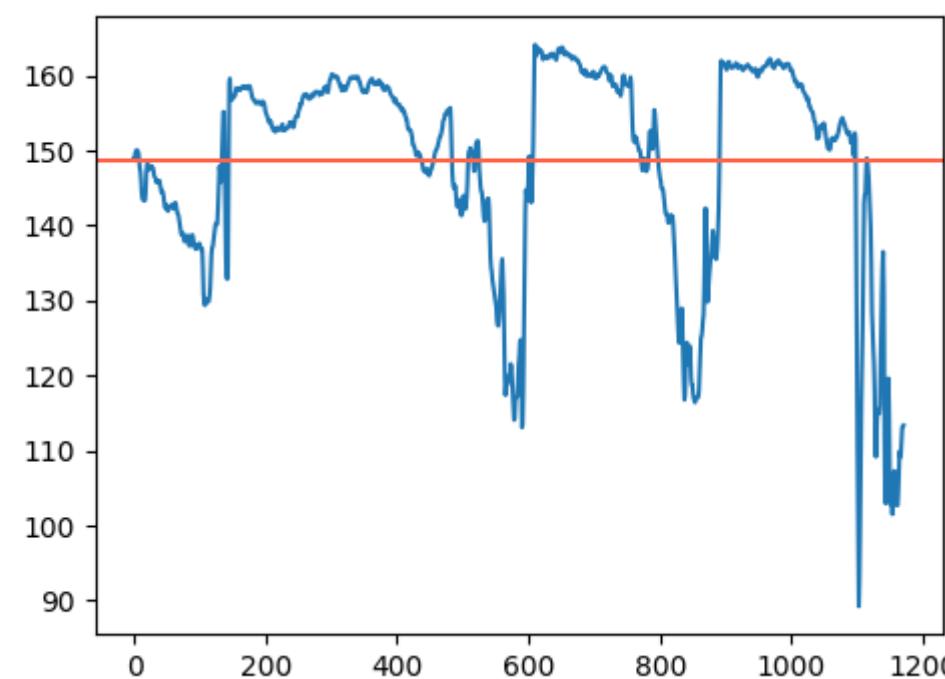
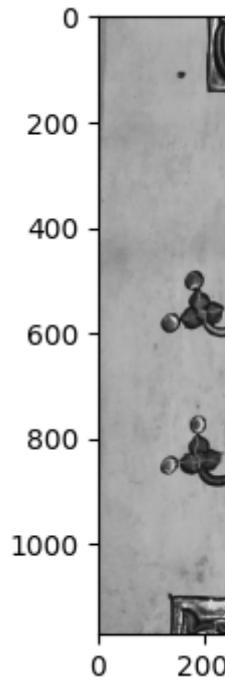




[77, 8, 134, 89, 18, 6, 1, 12, 31, 17, 1, 1, 4, 1, 32, 31, 37, 9, 17, 16, 9, 3, 4, 4, 11, 6, 1, 29,
46.299700936157734 29.870907348562366

---> not text

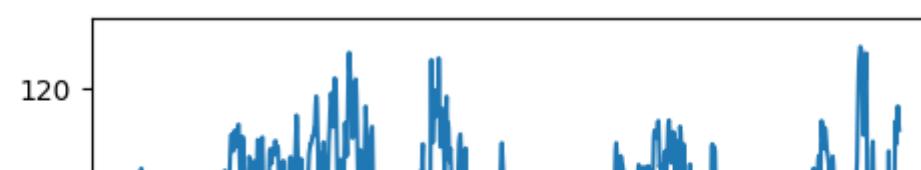
Analyzing area 122 841 372 2012

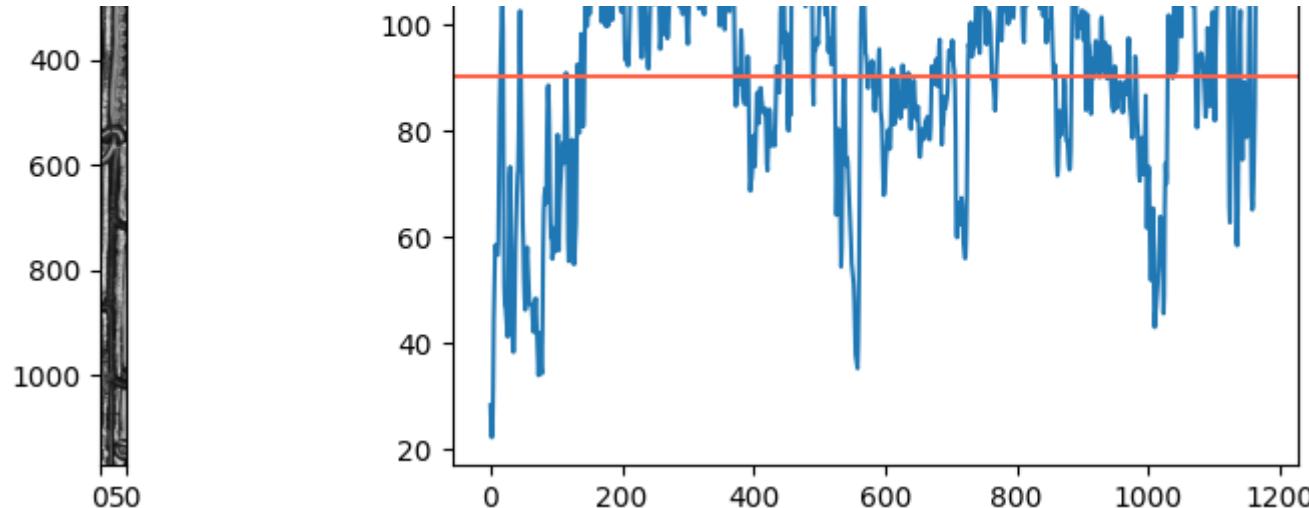


[9, 5, 293, 28, 7, 6, 1, 165, 1, 15, 207, 2] [125, 6, 18, 25, 3, 76, 5, 4, 6, 93, 16, 55]
39.63373983531372 96.44468045926062

---> not text

Analyzing area 372 841 423 2012

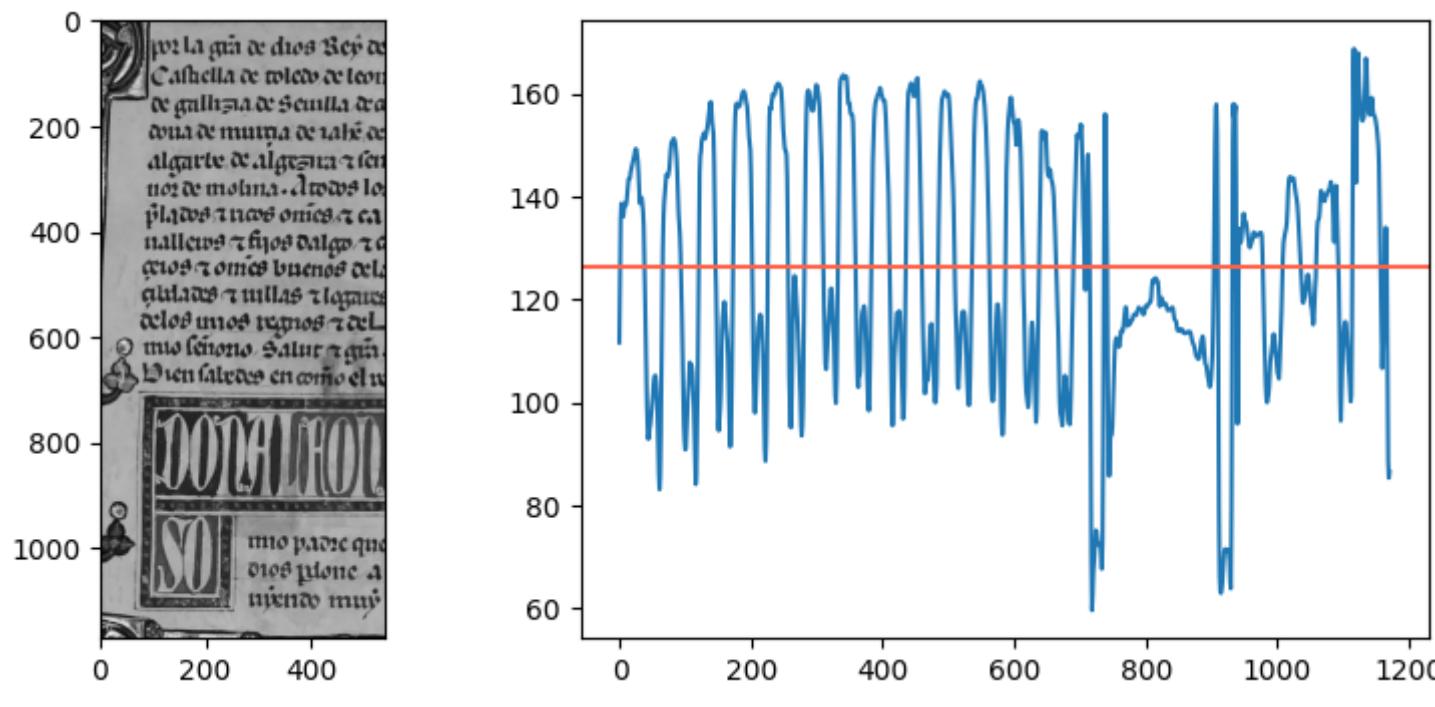




[15, 24, 67, 16, 3, 3, 5, 2, 6, 42, 2, 5, 2, 13, 24, 3, 7, 18, 4, 8, 7, 35, 1, 1, 1, 12, 20, 6, 1,
13.490978853169501 34.603081710866284

---> not text

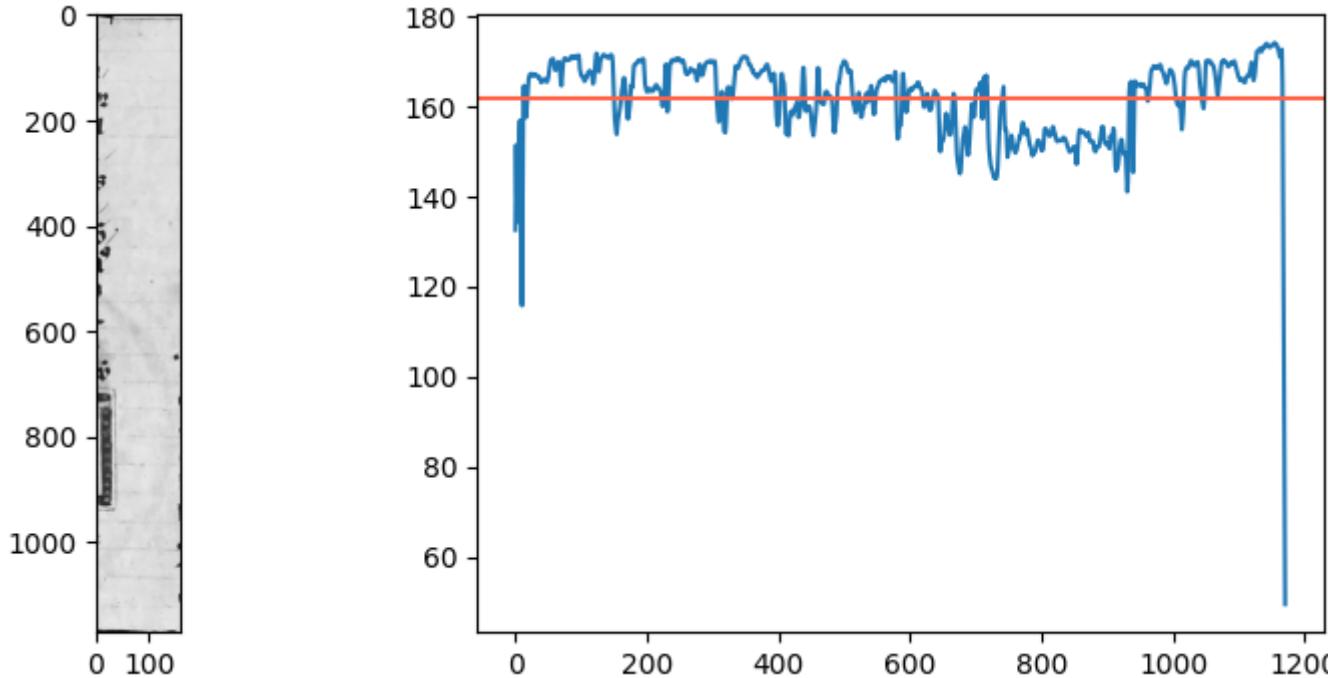
Analyzing area 423 841 962 2012



29.82995790960639 11.16062481735676

---> not text

Analyzing area 962 841 1123 2012

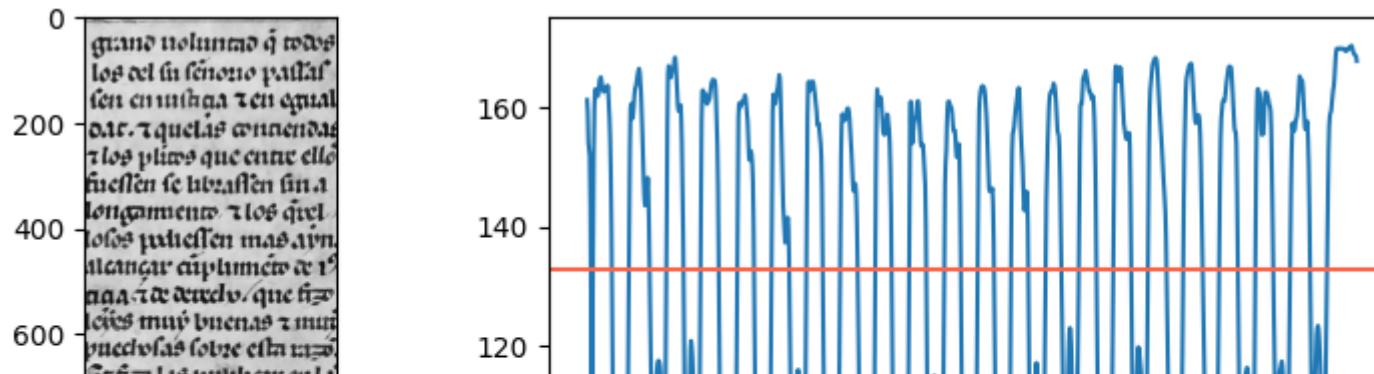


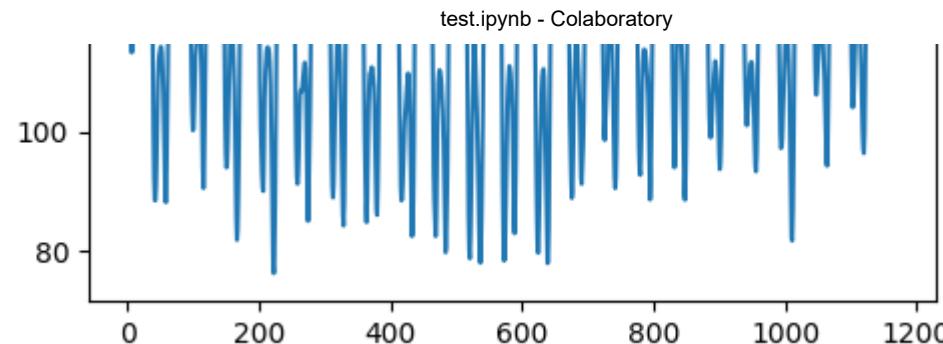
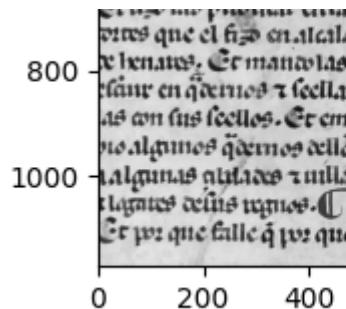
[13, 5, 11, 6, 4, 2, 7, 7, 1, 7, 26, 18, 8, 9, 8, 9, 3, 5, 9, 4, 5, 8, 23, 32, 3, 23, 190, 4, 3, 11]

32.319148252235856 32.592172984542145

---> not text

Analyzing area 1123 841 1602 2012



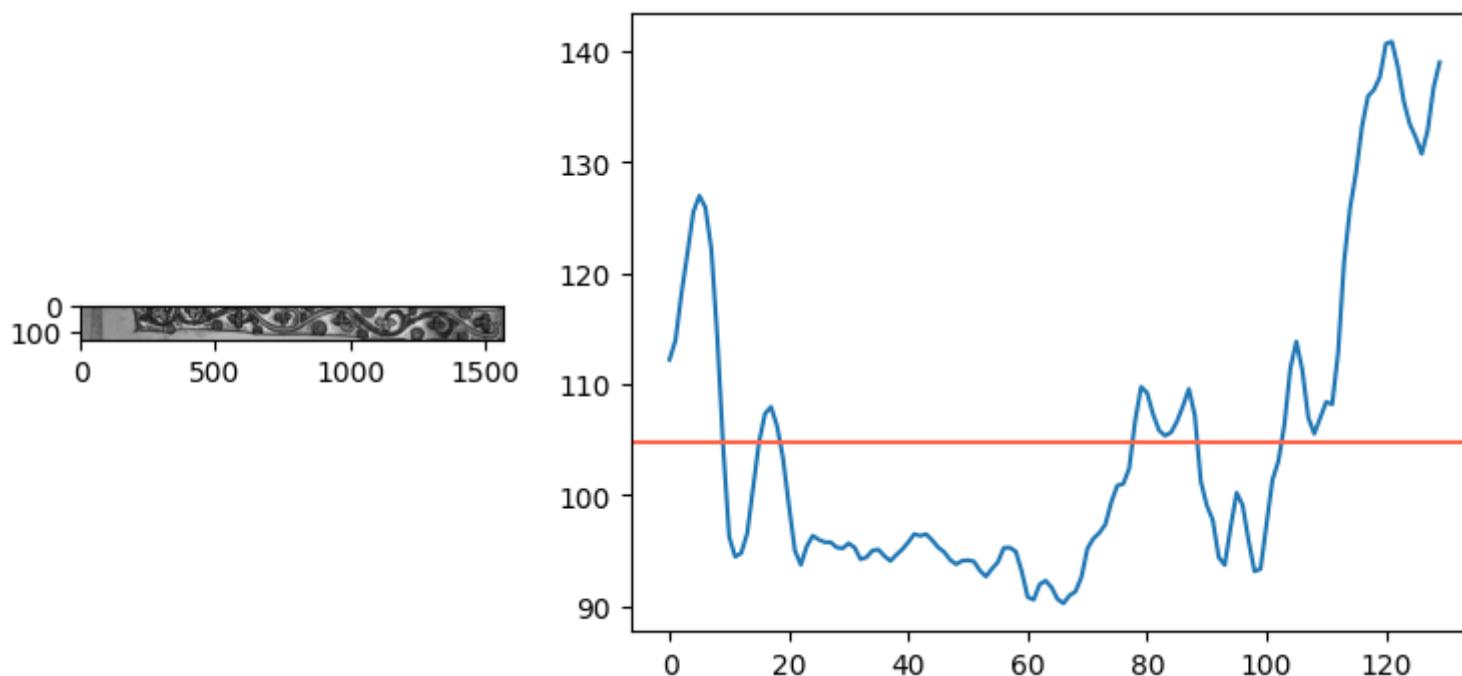


[6, 29, 31, 26, 28, 25, 27, 25, 27, 24, 26, 26, 25, 26, 24, 27, 28, 28, 29, 29, 27, 31, 46] [3, 26,

4.875516501564198 6.1959191399386295

----->>> Contains 22 lines of text

Analyzing area 31 2012 1602 2142



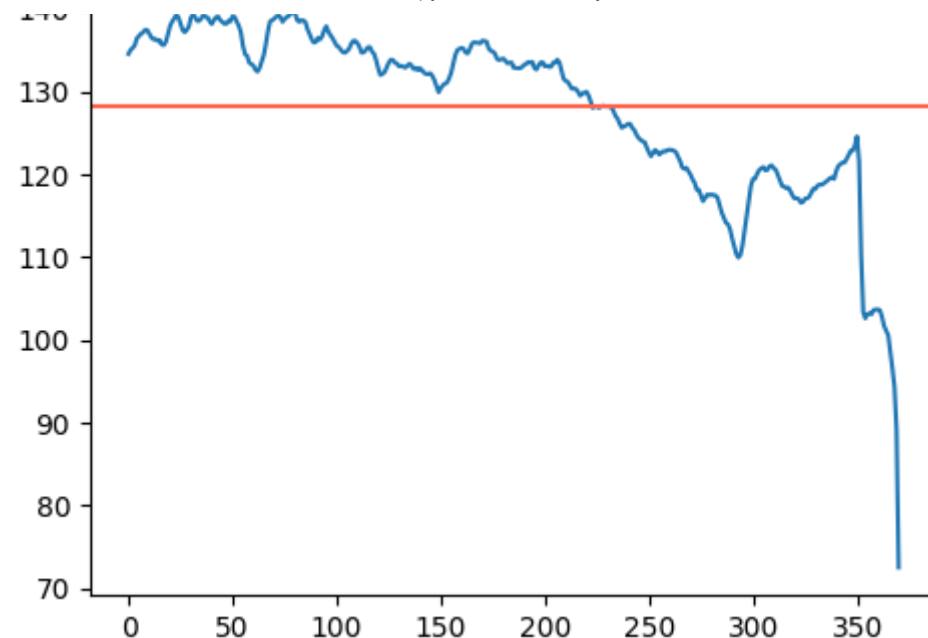
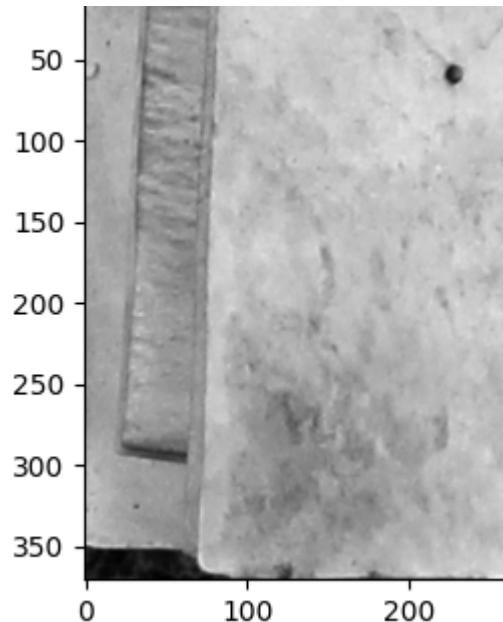
[9, 4, 11, 27] [6, 59, 14]

23.32857094256359 8.613216588476108

--> not text

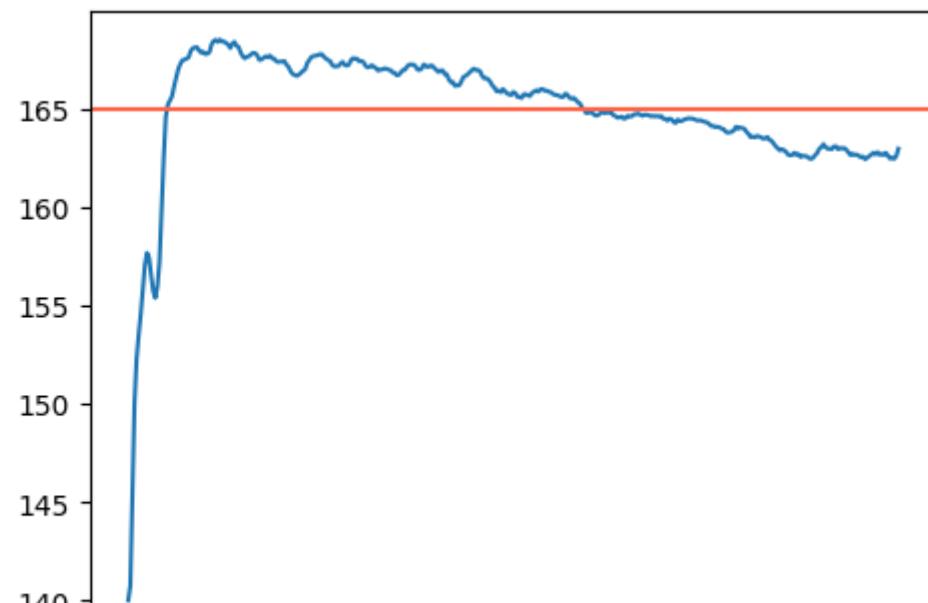
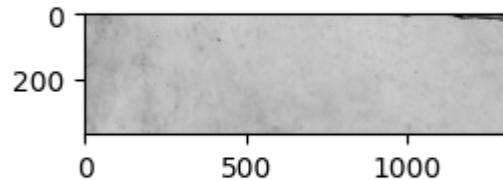
Analyzing area 31 2142 292 2513





[223, 1, 1] [1, 3, 142]
66.001683480213 104.65180361560904
---> not text

Analyzing area 292 2142 1602 2513

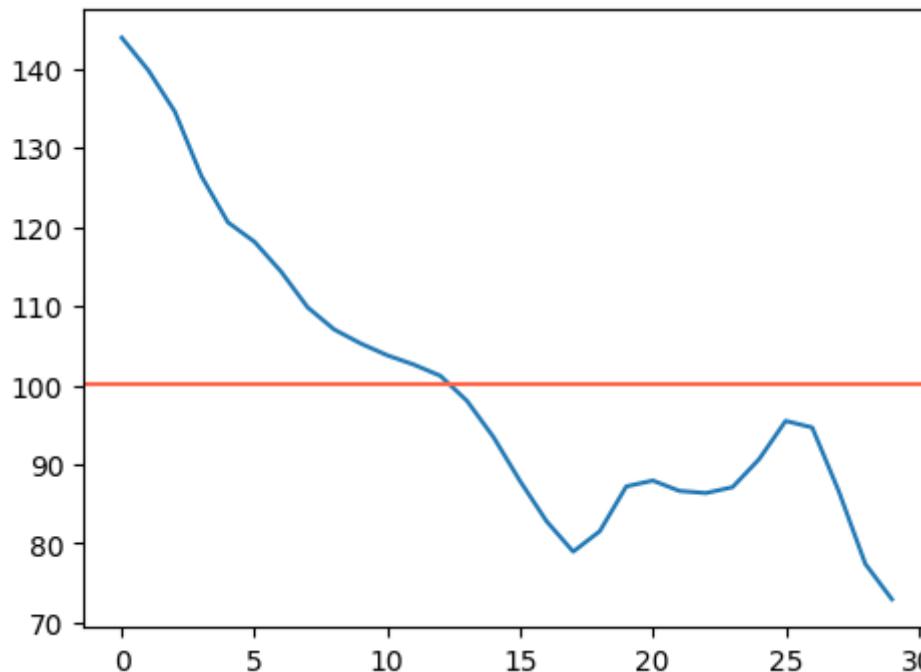


[19, 151] [201]

66.0 0.0

---> not text

Analyzing area 31 2513 1602 2543

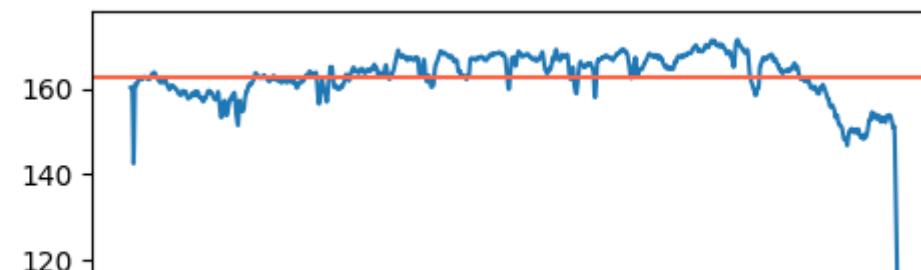
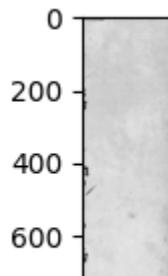


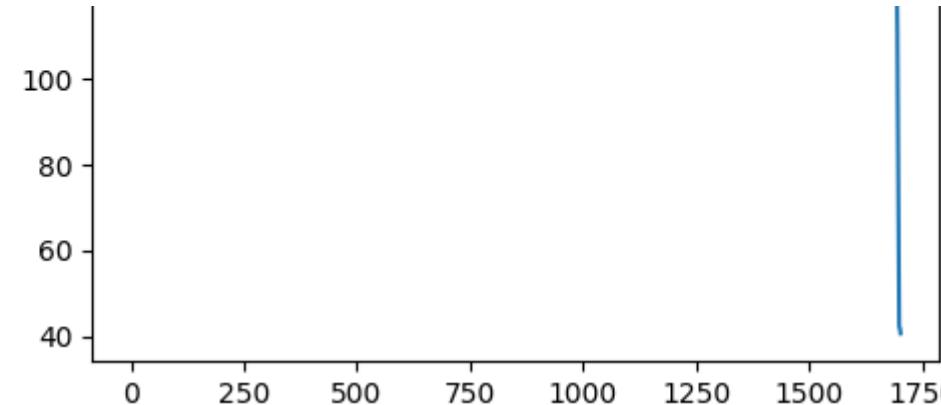
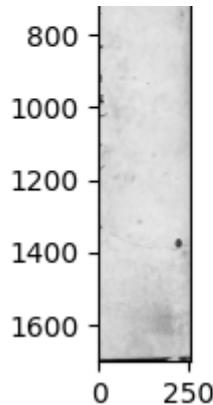
[13] [17]

0.0 0.0

---> not text

Analyzing area 1602 841 1859 2543





[22, 4, 5, 2, 212, 11, 2, 5, 22, 35, 28, 27, 4, 3, 4, 10, 2, 6, 2, 10, 5, 1, 19, 2, 1, 215] [1, 9, 55.13754989373416 59.132745581445825

---> not text

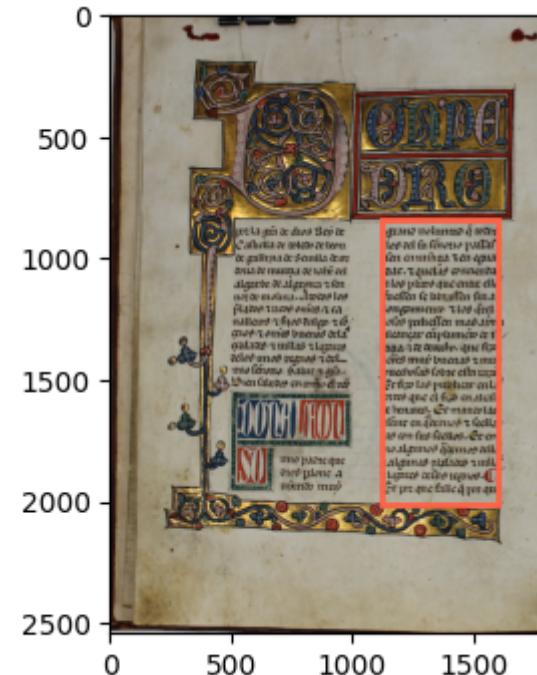
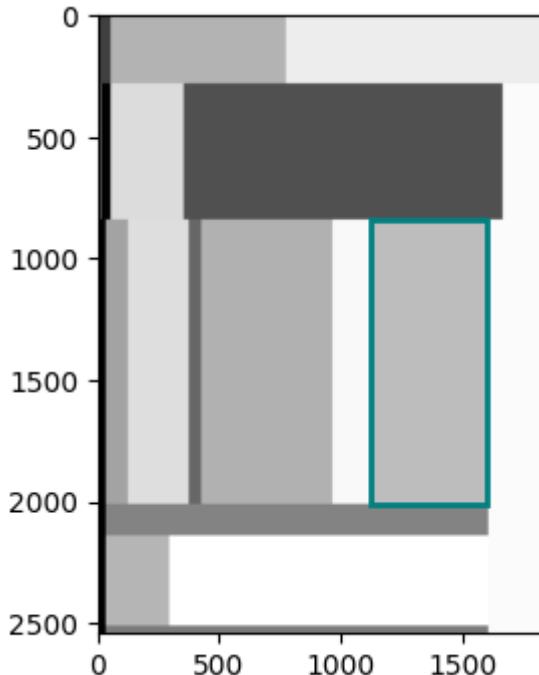
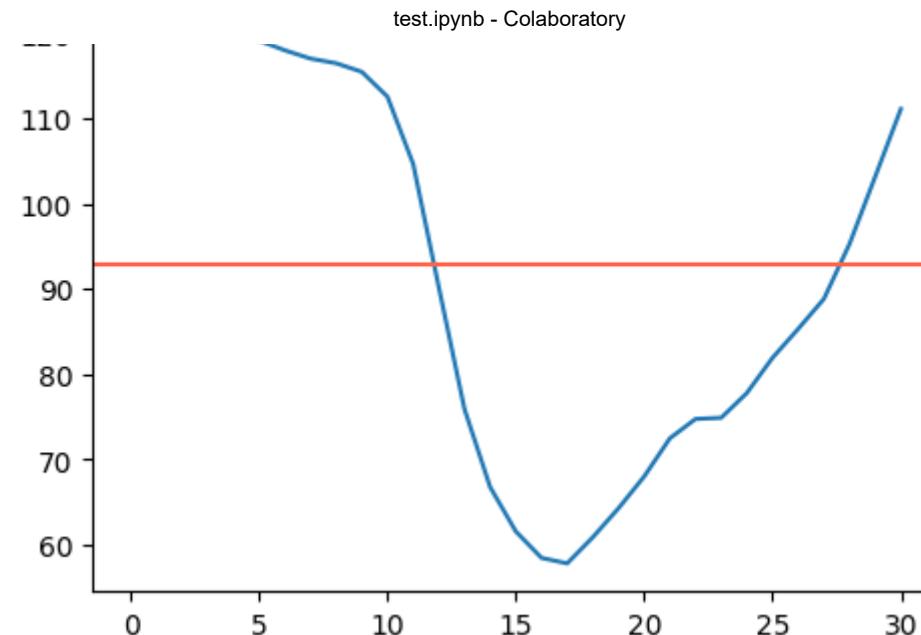
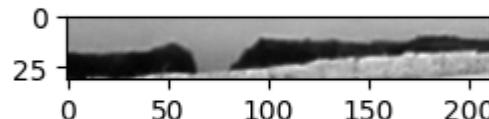


IMAGE 1

Analyzing area 0 0 211 31



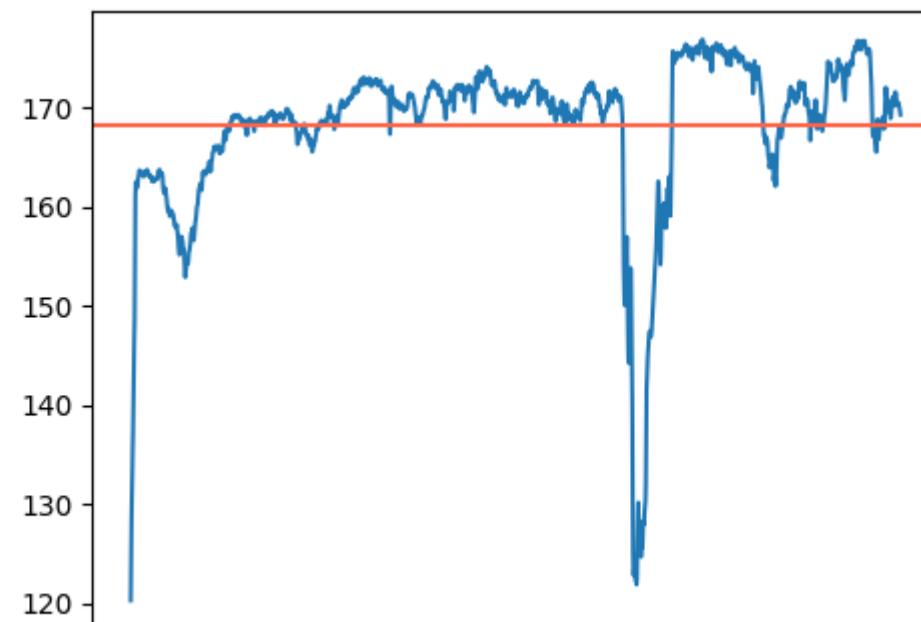
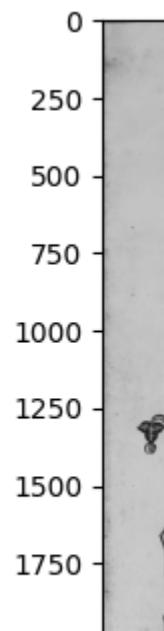


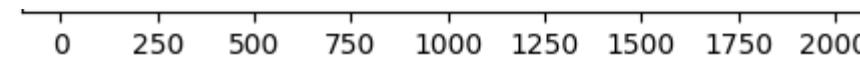
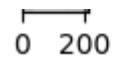
[12, 3] [16]

0.0 4.5

---> not text

Analyzing area 0 31 211 2021

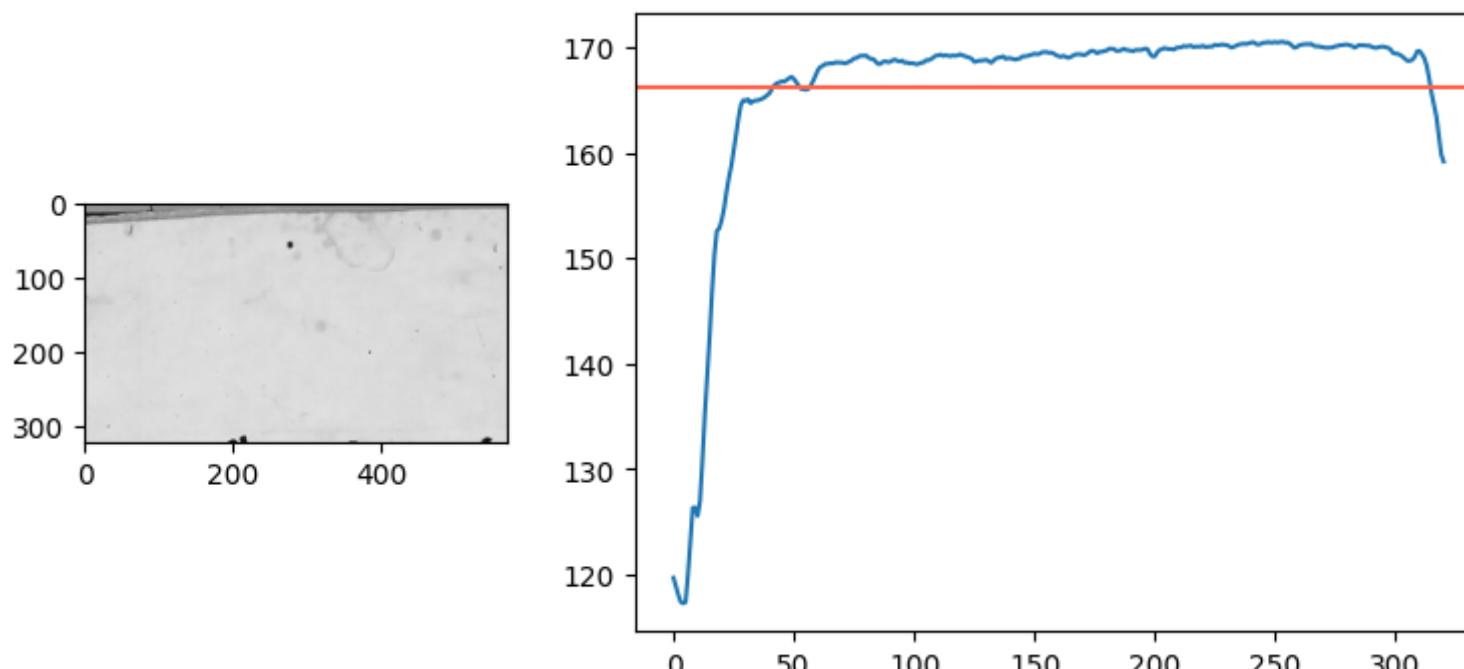




[258, 9, 1, 3, 1, 1, 20, 35, 1, 1, 2, 2, 127, 42, 8, 3, 2, 3, 13, 5, 1, 2] [36, 13, 1, 24, 77, 3, 1]
57.6886055180013 130.3895942205195

---> not text

Analyzing area 211 0 782 321

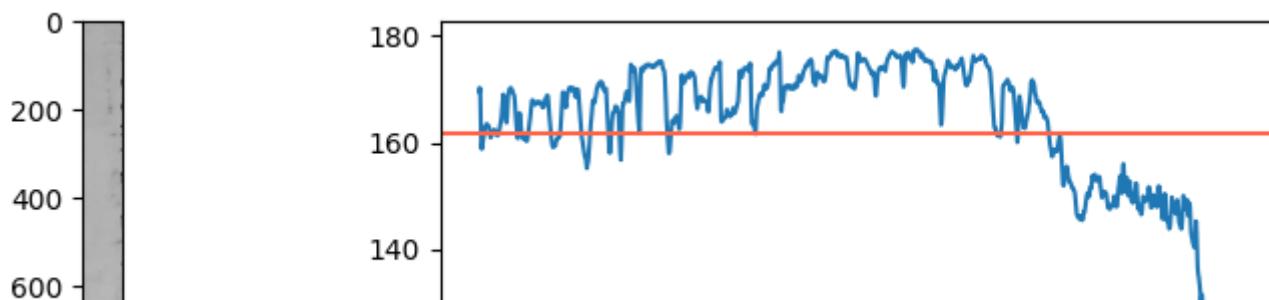


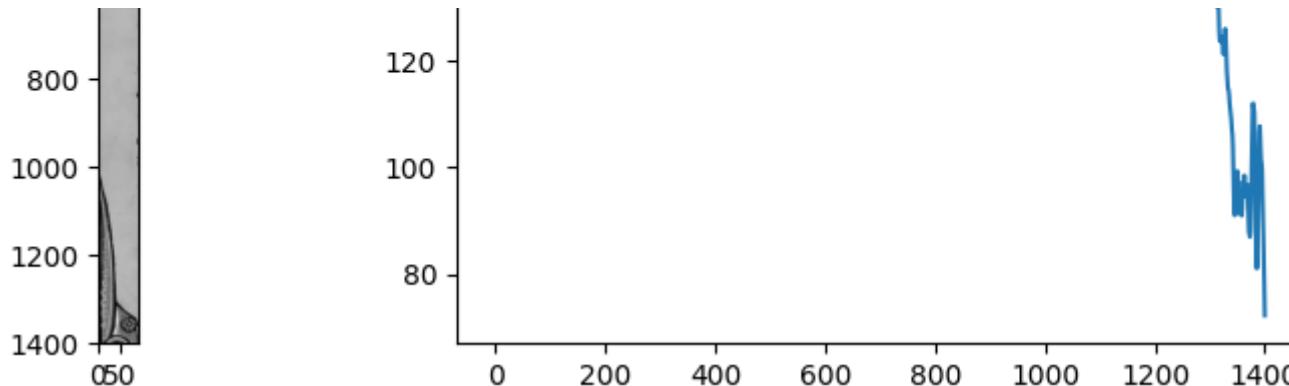
[42, 4, 6] [11, 258]

17.46106780494506 123.5

---> not text

Analyzing area 211 321 302 1722

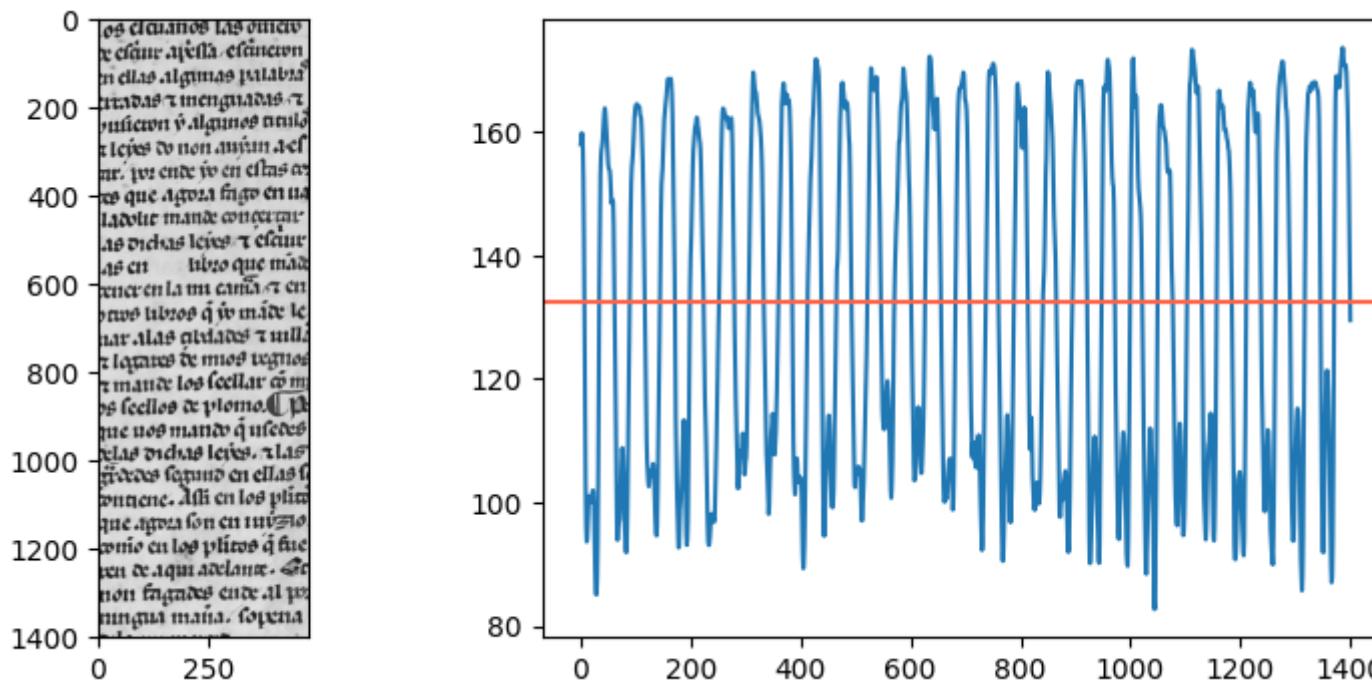




[4, 10, 7, 1, 34, 6, 41, 39, 33, 15, 83, 153, 438, 28, 55] [6, 4, 1, 3, 2, 12, 17, 16, 5, 4, 7, 1,
89.5927576438086 107.09426792420882

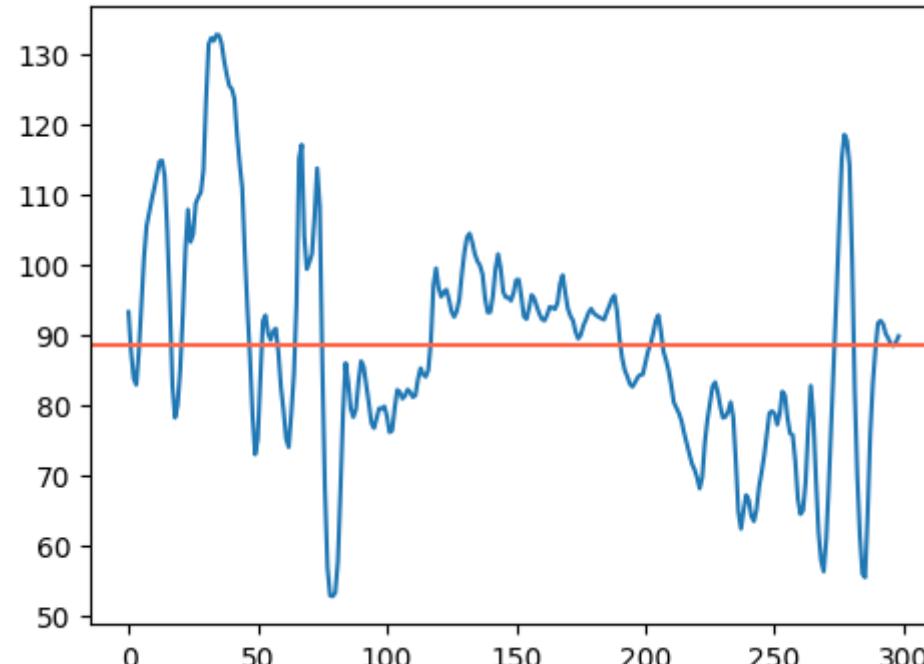
---> not text

Analyzing area 302 321 782 1722



[7, 28, 27, 27, 25, 28, 28, 25, 25, 28, 25, 30, 29, 26, 23, 24, 25, 27, 24, 22, 27, 28, 28, 27, 27,
5.080693168525563 4.178460674420875
----->>> Contains 27 lines of text

Analyzing area 211 1722 482 2021

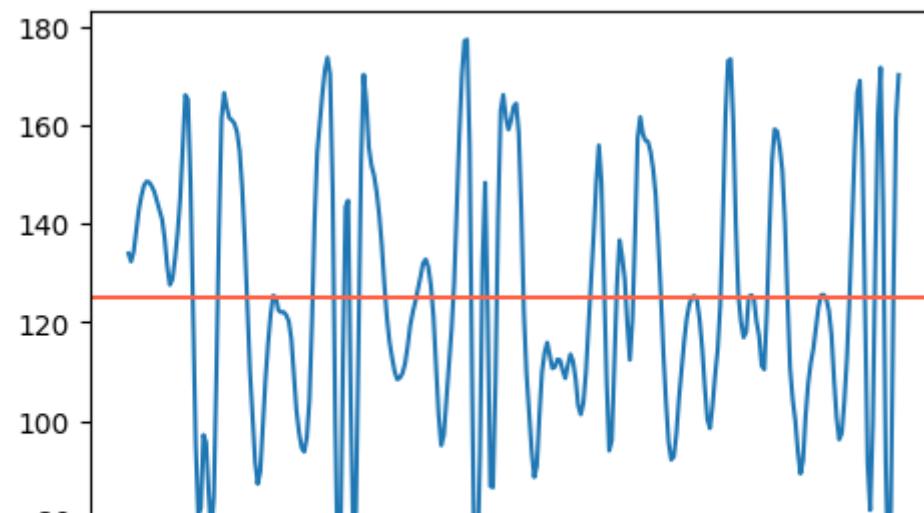
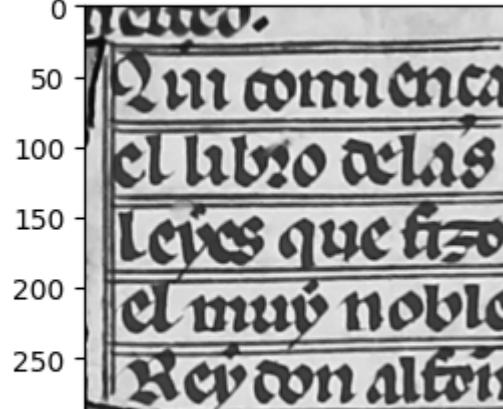


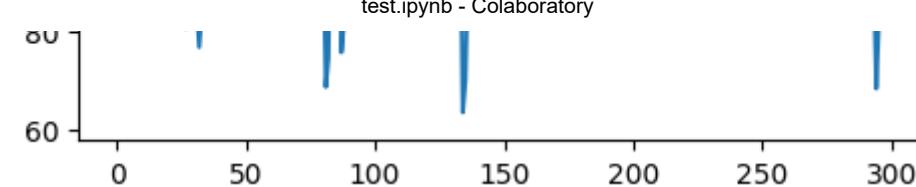
[1, 12, 26, 6, 10, 74, 4, 7, 5, 2] [4, 4, 5, 7, 42, 12, 67, 9, 2]

21.136826443408328 20.895214763193987

---> not text

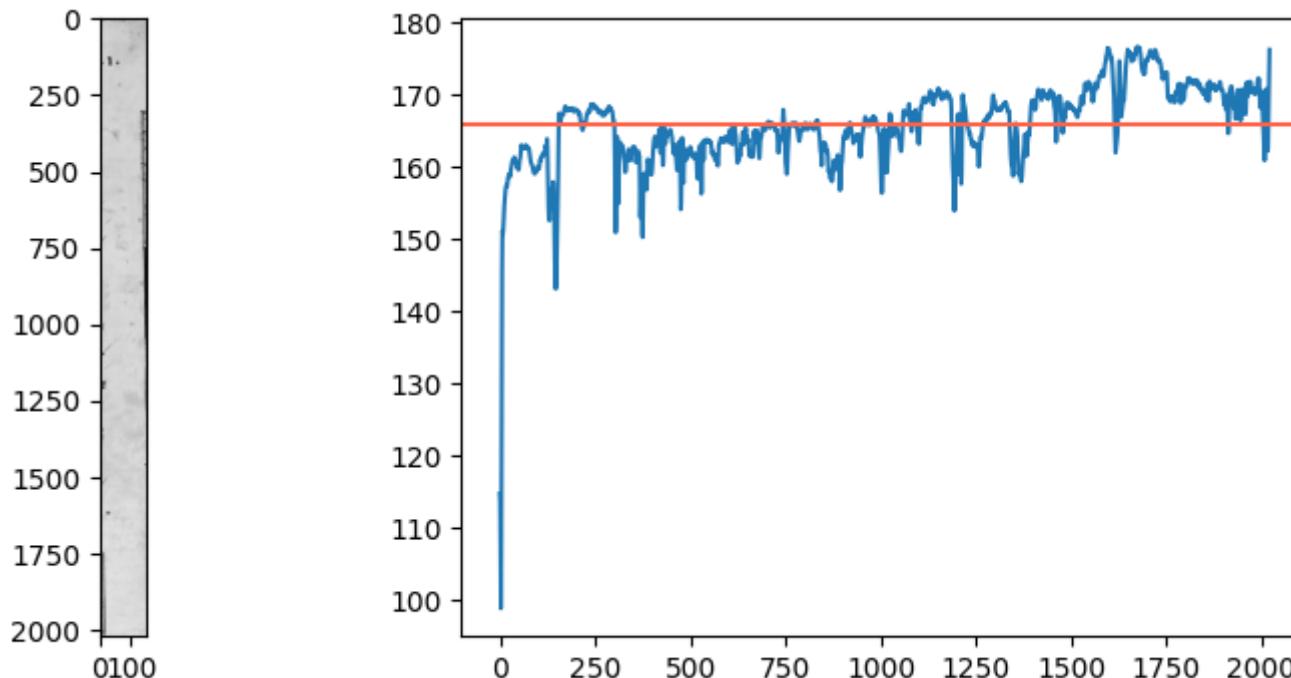
Analyzing area 482 1722 782 2021





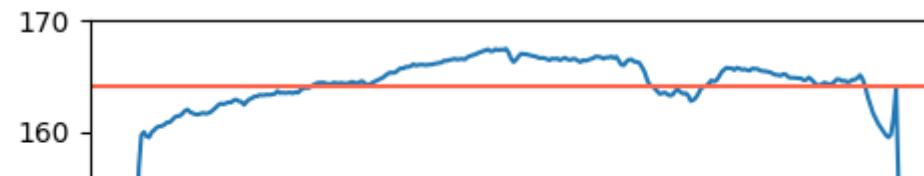
```
[25, 11, 2, 8, 2, 10, 6, 7, 2, 10, 6, 4, 11, 2, 7, 2, 7, 2, 6, 4, 3] [10, 10, 14, 4, 4, 12, 8, 4, 4
5.467174773134658 5.151227755816107
----->>> Contains 20 lines of text
```

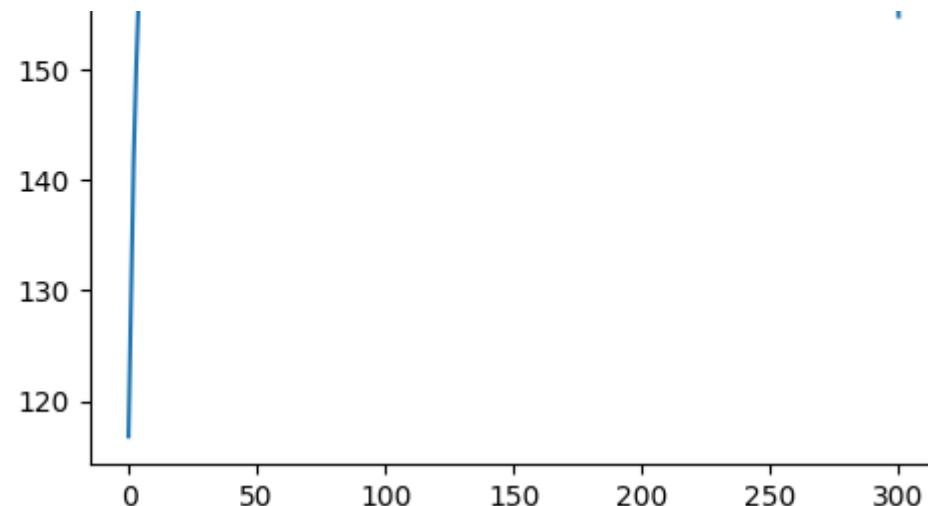
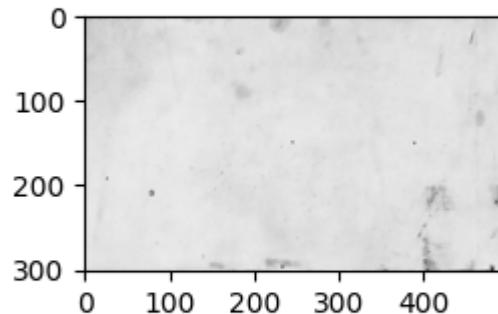
Analyzing area 782 0 931 2021



```
[155, 7, 313, 88, 4, 24, 13, 1, 2, 20, 5, 1, 13, 81, 35, 1, 31, 30, 4, 7, 16, 9, 40, 15, 37, 3, 3,
60.56874045660419 57.674496570239725
---> not text
```

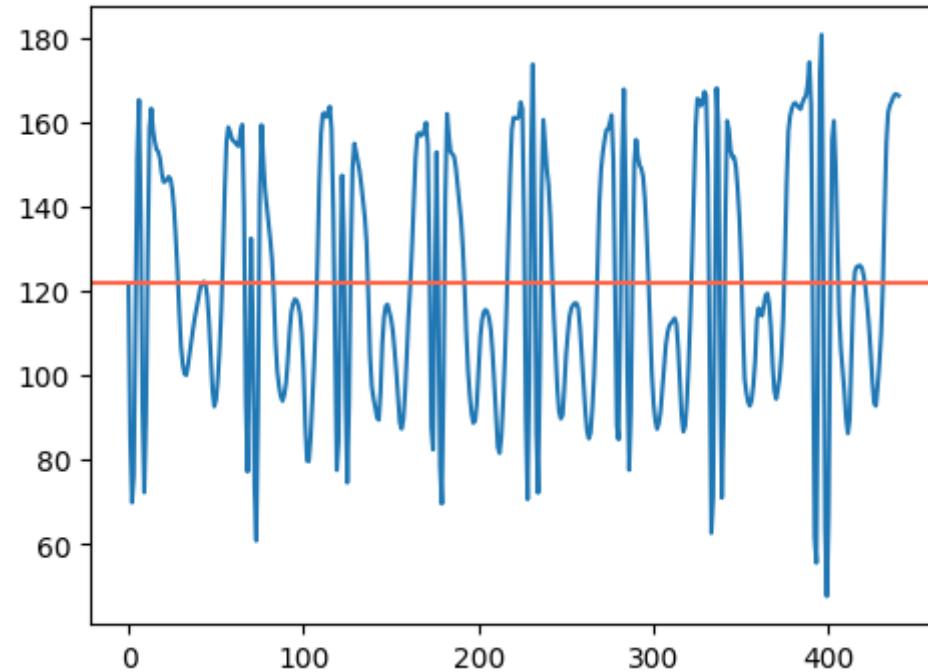
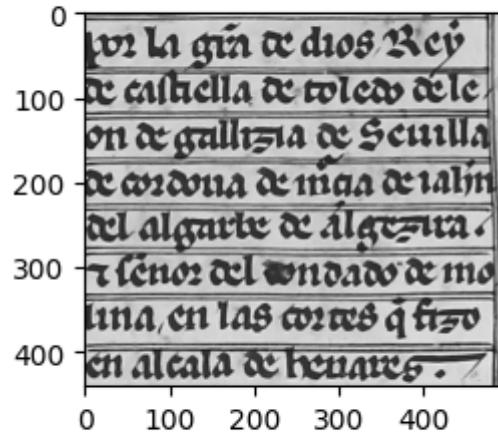
Analyzing area 931 0 1431 301





[72, 20, 14] [133, 62]
26.042699979499478 35.5
---> not text

Analyzing area 931 301 1431 742

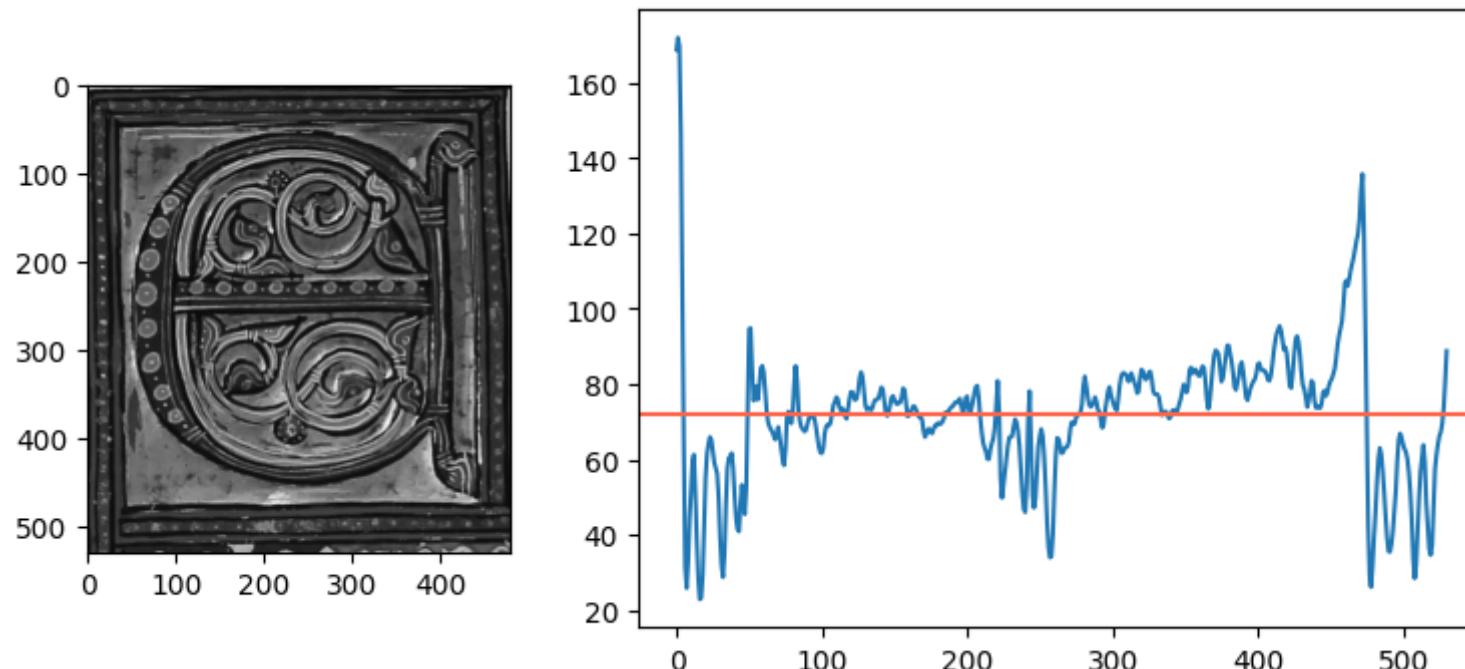


[5, 3, 14, 10, 3, 4, 25, 3, 3, 25, 3, 4, 25, 3, 3, 25, 4, 3, 24, 4, 3, 24, 4, 4, 9, 11] [3, 18, 1,

8.740844745441102 4.6257895535878895

---->>> Contains 26 lines of text

Analyzing area 931 742 1412 1273

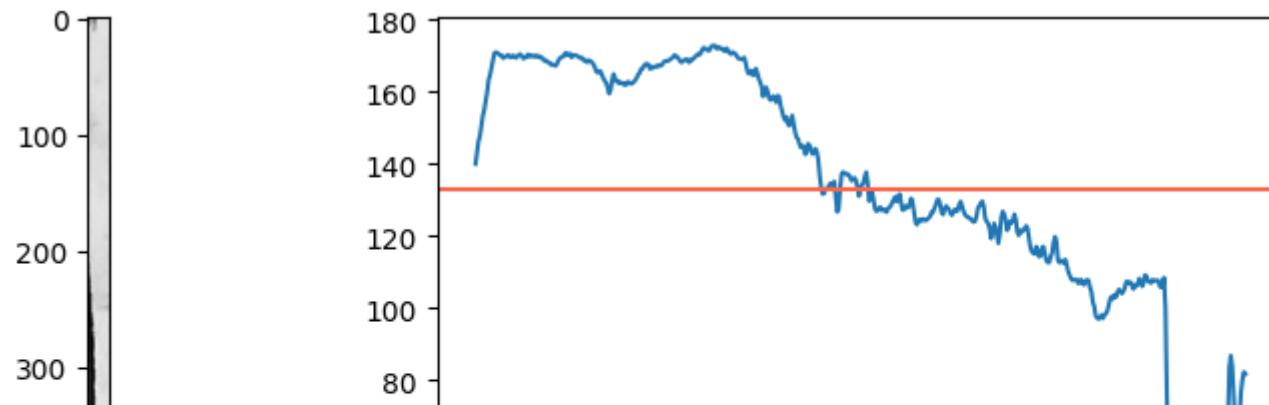


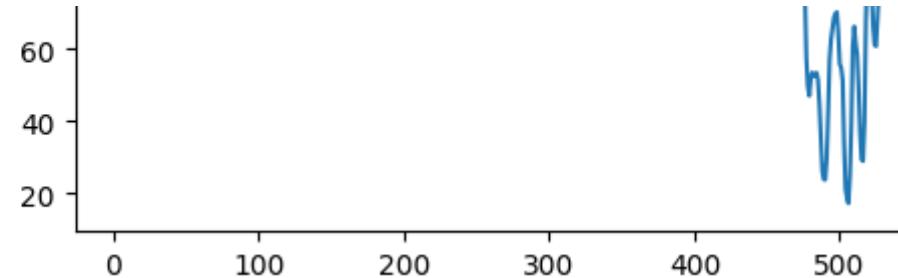
[5, 14, 1, 5, 9, 27, 13, 6, 24, 3, 1, 15, 39, 3, 134, 3] [44, 14, 2, 22, 2, 1, 2, 18, 11, 20, 33, :

16.11403804830503 31.491814412637453

---> not text

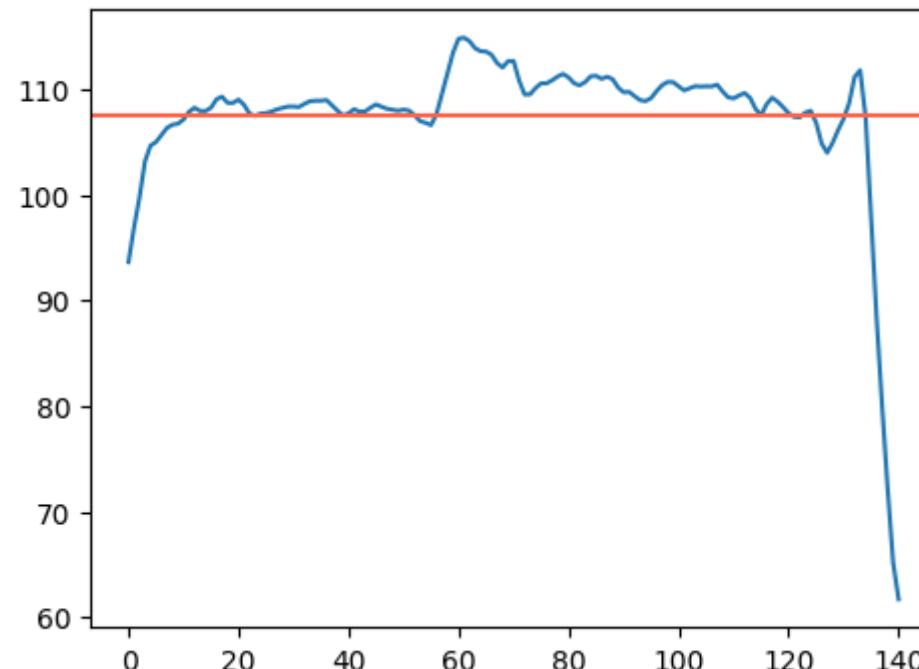
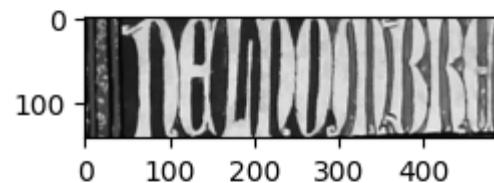
Analyzing area 1412 742 1431 1273





[239, 7, 12, 5] [2, 4, 2, 260]
111.43159336561602 100.05842043526371
---> not text

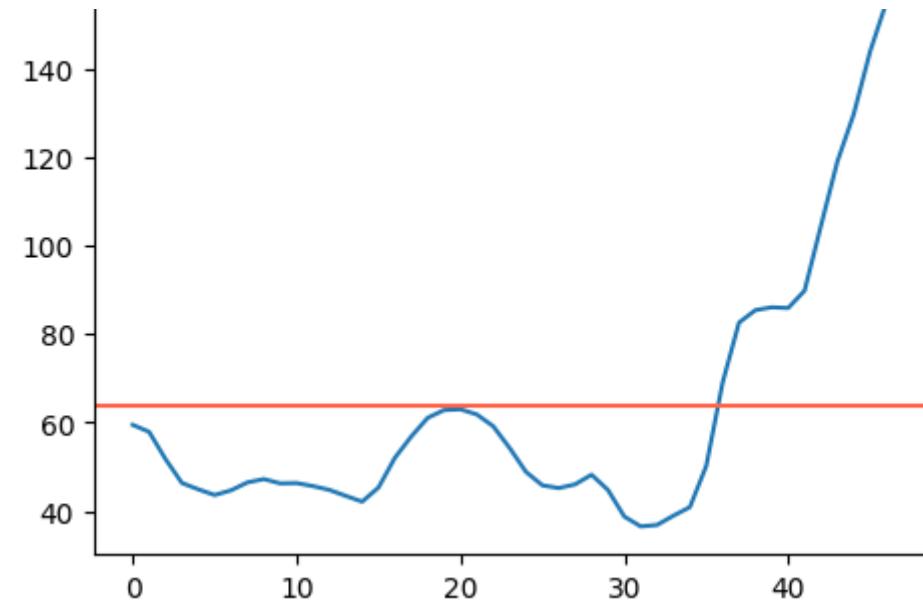
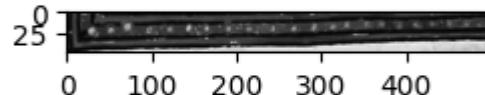
Analyzing area 931 1273 1431 1414



[11, 1, 3, 2, 6, 7] [12, 29, 65, 2, 3]
3.415650255319866 23.489572154468885
---> not text

Analyzing area 931 1414 1431 1462



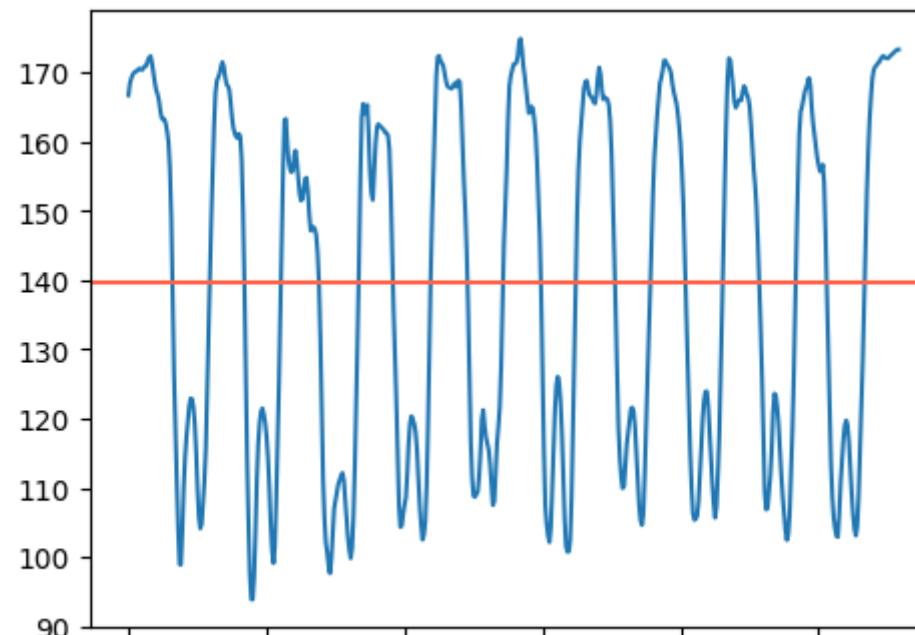
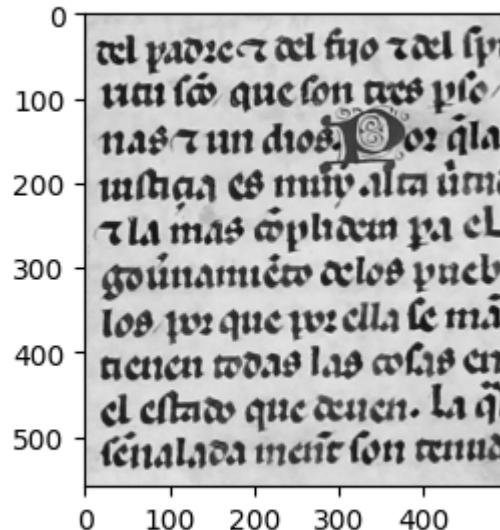


[36] [12]

0.0 0.0

---> not text

Analyzing area 931 1462 1431 2021

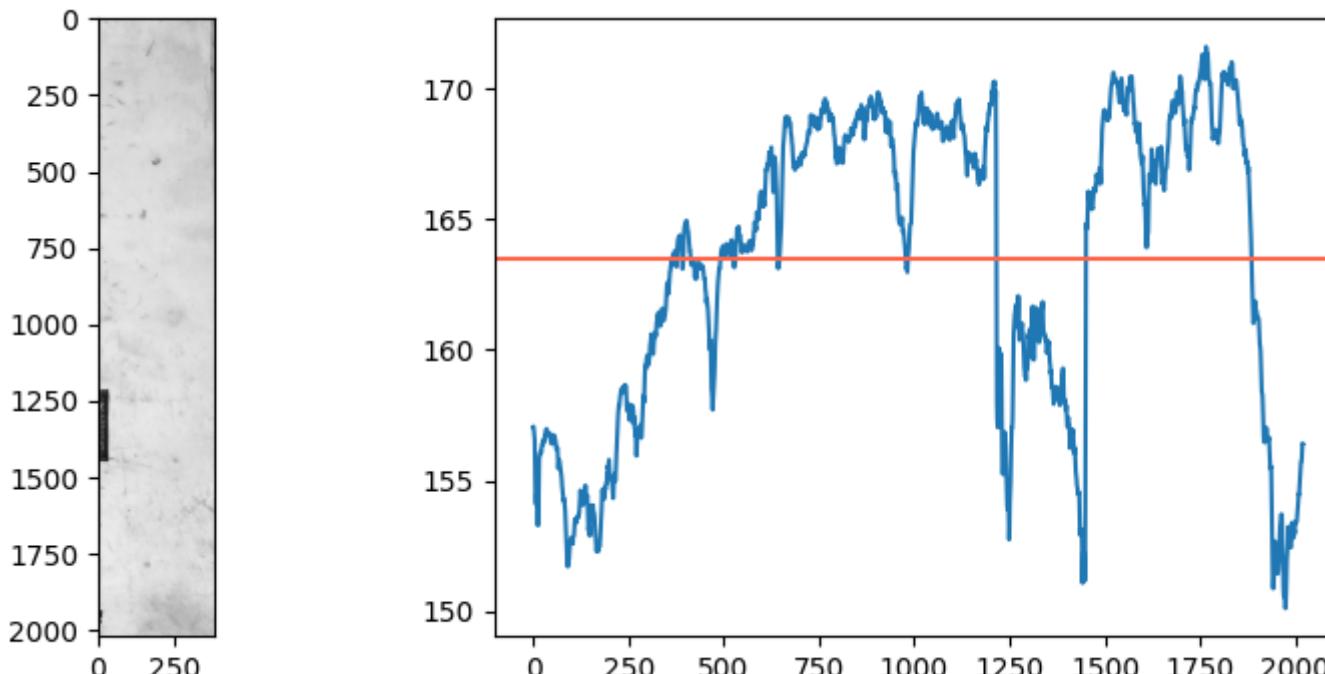


[32, 25, 28, 25, 27, 27, 29, 25, 26, 22, 25] [27, 27, 28, 27, 26, 25, 26, 27, 27, 28]

0.8717797887081347 2.4995867427031855

----->>> Contains 10 lines of text

Analyzing area 1431 0 1812 2021



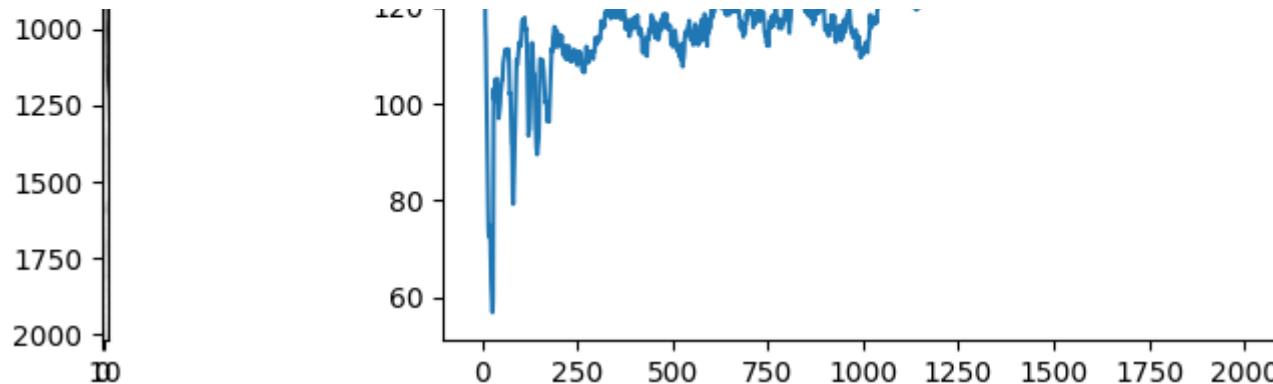
[366, 1, 3, 4, 79, 1, 6, 5, 6, 235, 135] [2, 5, 14, 19, 10, 21, 111, 331, 232, 435]

116.56942031281667 150.6047808006107

--> not text

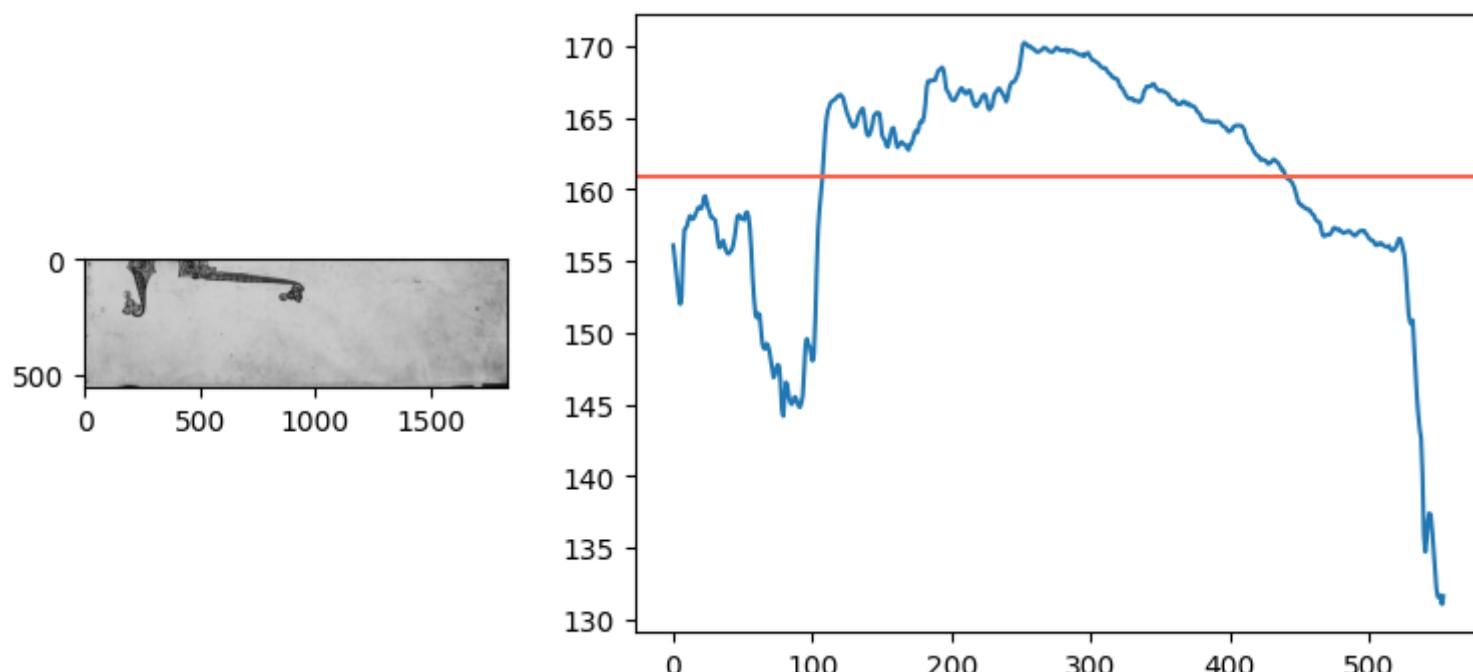
Analyzing area 1812 0 1831 2021





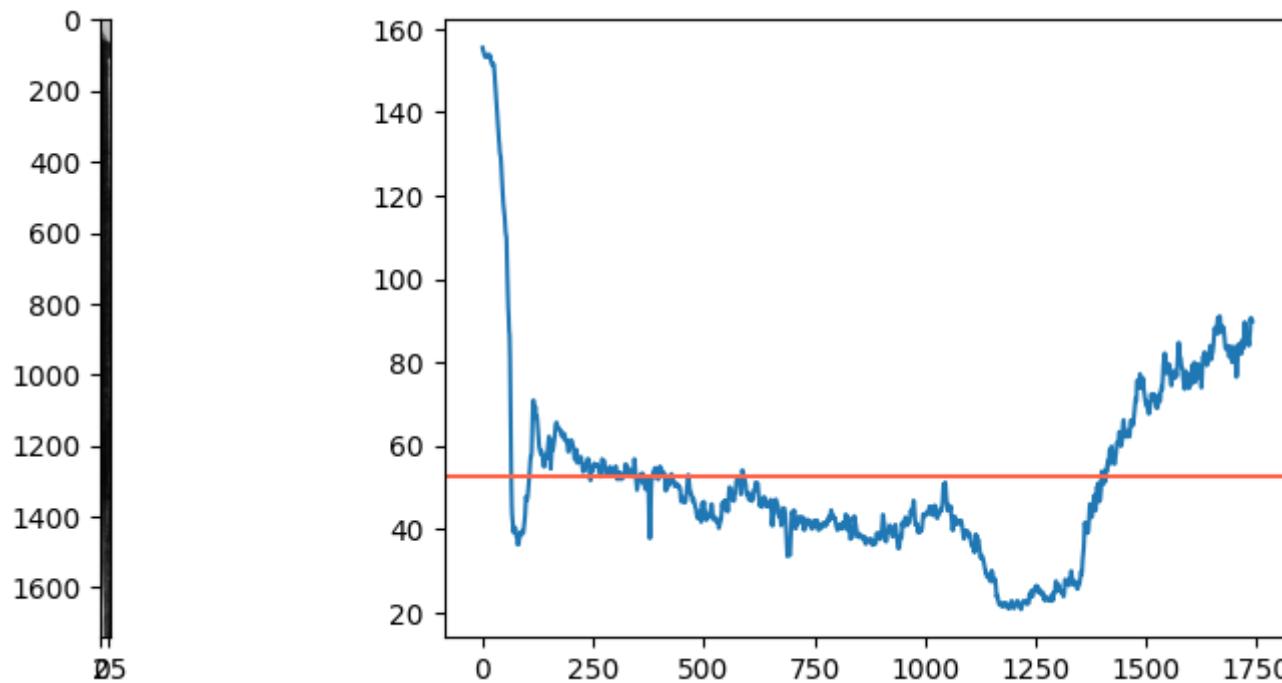
[7, 833] [1181]
0.0 413.0
---> not text

Analyzing area 0 2021 1831 2575



[108, 114] [332]
3.0 0.0
---> not text

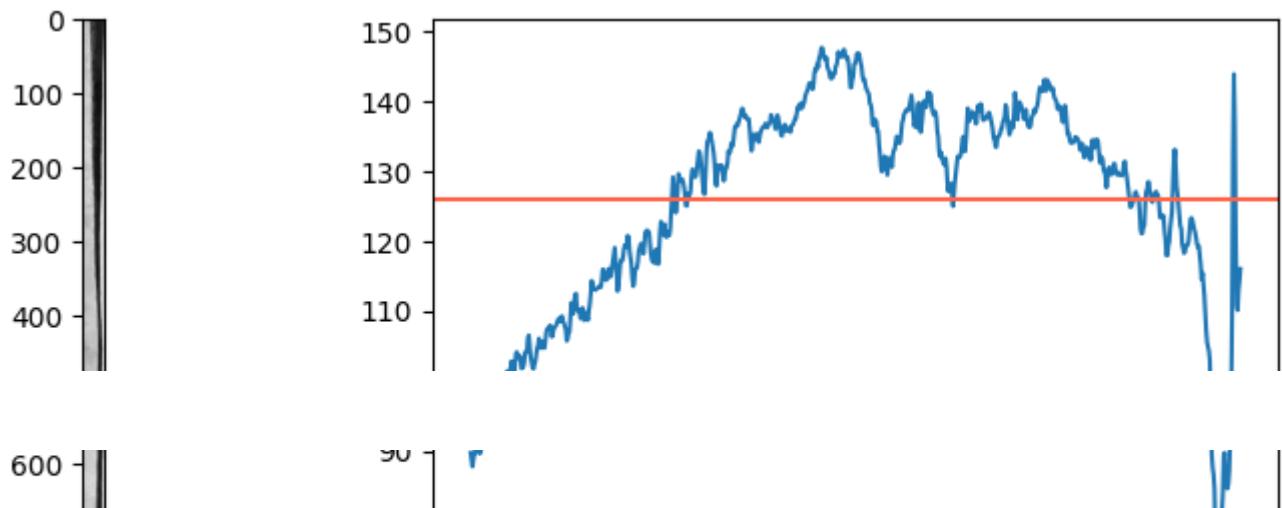
```
analyzing 0 25 100 200 400 600 800 1000 1200 1400 1600 1800 2000
```

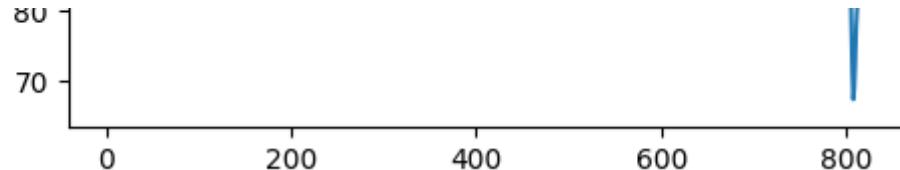
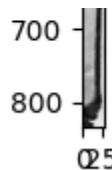


[65, 137, 22, 35, 6, 1, 18, 10, 2, 2, 1, 10, 14, 1, 2, 1, 5, 3, 333] [41, 2, 1, 4, 2, 2, 1, 10, 1,
183.23577370544243 77.08932833366944

---> not text

Analyzing area 1831 1741 1861 2575





[218, 3, 2, 1, 4, 7, 2, 15, 59, 6] [3, 9, 287, 191, 6, 6, 6, 6, 3]

64.26515385494693 99.66288856954073

---> not text

