

test_images

August 22, 2019

```
[1]: import cv2
import numpy as np
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
import os

from sklearn.cluster import KMeans
```

1 Script below binarizes images and demonstrates with a batch of test images. The before and after are displayed sequentially.

First test performed on 08-21-2019 Last test performed on 08-21-2019

Once the script is done processing, save as .pdf so that it can be compared to previous scripts and attempts at the problem.

```
[13]: test_dir = 'Test_Images/'

for filename in os.listdir(test_dir):

    # read in image
    img = cv2.imread(test_dir + filename)
    #img = cv2.resize(img, (600,150))

    plt.figure(figsize=(20,5))
    plt.imshow(img);

    # convert to grayscale
    gray_img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

    # dilate the image in order to get ride of the math
    dilated_img = cv2.dilate(gray_img, np.ones((10,10), np.uint8))

    # show dilated image
    # plt.figure(figsize=(20,10))
    # plt.imshow(dilated_img, cmap='gray');
```

```

### Step 2: Blur ###
# suppress anything else with a blur function

blur_img = cv2.medianBlur(dilated_img, 21)

### Step 3: Calculate the difference between the original and background
→ just made ###
# identical bits will be black (close to zero difference), text will be
→ white (large difference)

diff_img = 255 - cv2.absdiff(gray_img, blur_img)

### Step 4: Apply Simple Threshold ###
thresh = 180
maxValue = 255
ret, thresh2 = cv2.threshold(diff_img, thresh, maxValue, cv2.THRESH_BINARY)

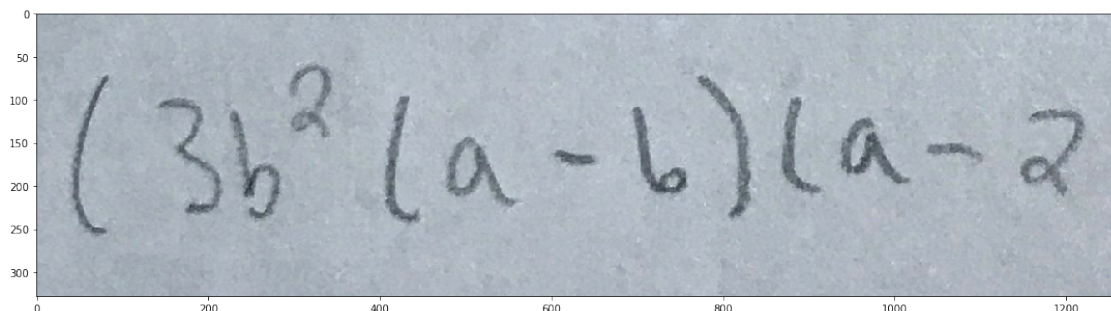
### Other Option - Step 4: Apply Otsu Binarization ###
thresh = 0
maxValue = 255
ret, thresh3 = cv2.threshold(thresh2, thresh, maxValue, cv2.
→ THRESH_BINARY+cv2.THRESH_OTSU)

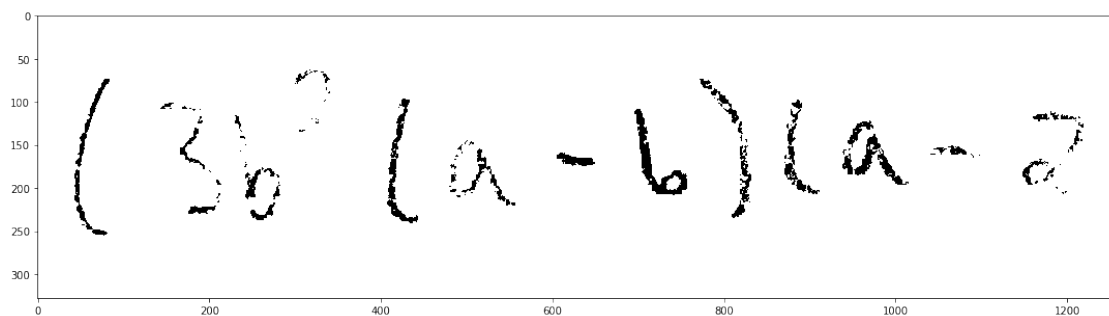
plt.figure(figsize=(20,5))
plt.imshow(thresh3,cmap='Greys_r');

```

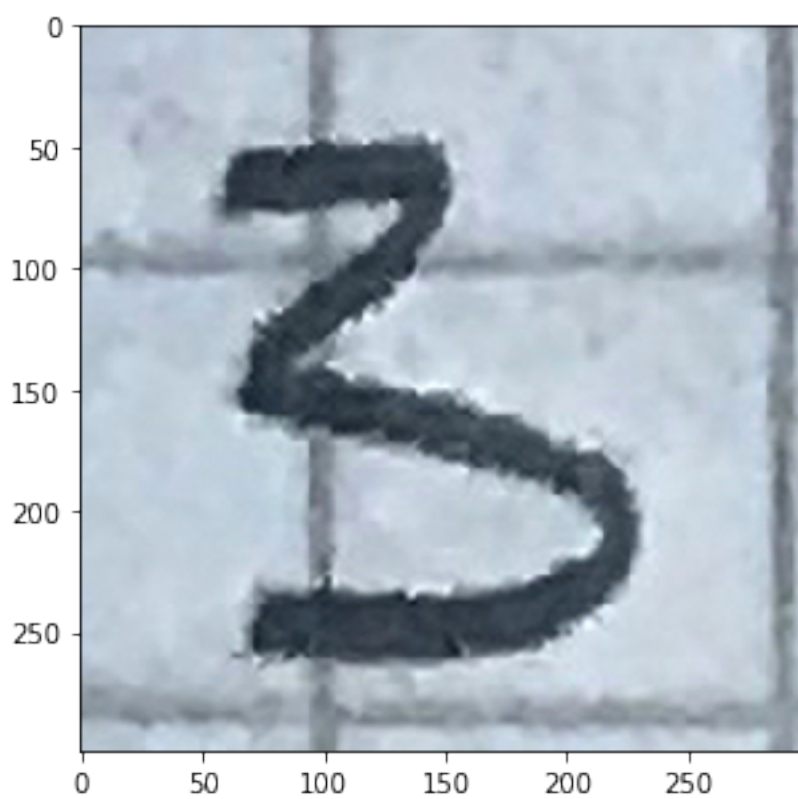
/Users/utoarca/anaconda3/lib/python3.7/site-packages/ipykernel_launcher.py:9:
RuntimeWarning: More than 20 figures have been opened. Figures created through
the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
closed and may consume too much memory. (To control this warning, see the
rcParam `figure.max_open_warning`).

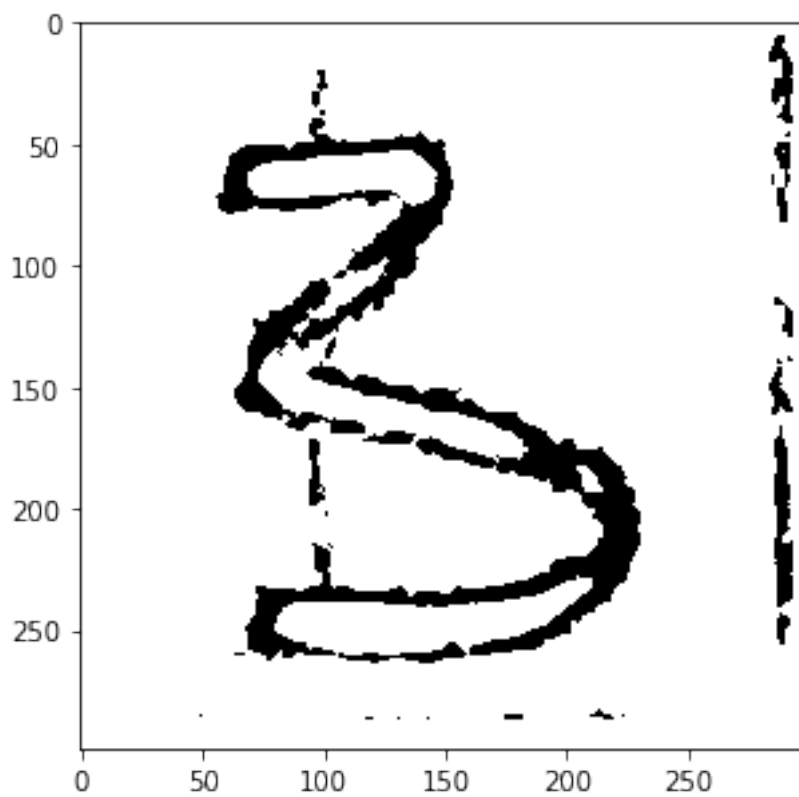
if __name__ == '__main__':
/Users/utoarca/anaconda3/lib/python3.7/site-packages/ipykernel_launcher.py:43:
RuntimeWarning: More than 20 figures have been opened. Figures created through
the pyplot interface (`matplotlib.pyplot.figure`) are retained until explicitly
closed and may consume too much memory. (To control this warning, see the
rcParam `figure.max_open_warning`).





A handwritten mathematical expression $(3b^2(a-b))(a-a^2)$ is shown on a coordinate grid. The x-axis ranges from 0 to 1200 with major ticks every 200 units. The y-axis ranges from 0 to 300 with major ticks every 50 units. The expression is written in black ink, with the first part $(3b^2(a-b))$ positioned roughly between x=100 and x=800, and the second part $(a-a^2)$ positioned roughly between x=800 and x=1100.





$$(1/2x - 5)(x^2 - 3x + 1)$$

$$(1/2x - 5)(x^2 - 3x + 1)$$

$$\left(\frac{(x^3)^{1/2}}{x^{7/2}} \right)$$

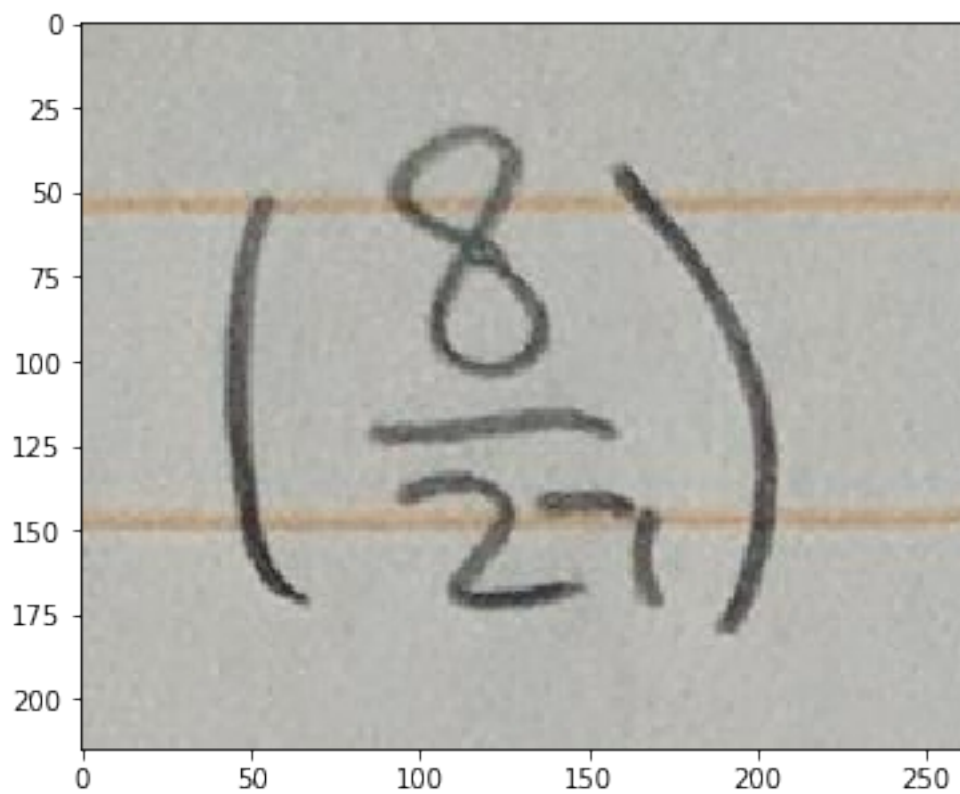
$$\left(\frac{(x^3)^{1/2}}{x^{7/2}} \right)$$

$$\lim_{x \rightarrow 8} - \frac{7(x-6) \sqrt{x+9} (3 + \sqrt{x+5})}{x-4}$$

$$\lim_{x \rightarrow 8} - \frac{7(x-6)\sqrt{x+9}(\sqrt{x+9} - \sqrt{x+5})}{x-4}$$

$$9x^3 + 58x^2 + (-41)x + (-12)$$

$$9x^3 + 58x^2 + (-41)x + (-12)$$



A handwritten expression $(\frac{8}{21})$ is shown on a coordinate grid. The x-axis ranges from 0 to 250 with major ticks every 50 units. The y-axis ranges from 0 to 200 with major ticks every 25 units. The expression is centered around x=125 and y=100.

A handwritten mathematical expression is shown on lined paper. The expression is $8 \int_0^1 x dx - 2 \int_0^1 x^3 dx$. The paper has horizontal red lines. The expression is written in dark ink.

A handwritten mathematical expression is shown on lined paper. The expression is $8 \int_0^1 x dx - 2 \int_0^1 x^3 dx$. The paper has horizontal red lines. The expression is written in dark ink.

$$f'(x) = 4x = \frac{(8x+4) \cdot 7 - 8x \cdot 3}{(6x+1)^2}$$

$$f'(x) = 4x = \frac{(8x+4) \cdot 7 - 8x \cdot 3}{(6x+1)^2}$$

$$f'(x) = 2x = \frac{(7x+2) \cdot \frac{d}{dx} 1x - 2x \cdot \frac{d}{dx} (3x+1)}{(5x+1)^2}$$

$$f'(x) = 2x = \frac{(7x+2) \cdot \frac{d}{dx} 1x - 2x \cdot \frac{d}{dx} (3x+1)}{(5x+1)^2}$$

Handwritten formula on a grid background:

$$\frac{x^3 + 3x^2h + 3xh^2 + h^3 + 2 - (x^3 + 2)}{h}$$

Digitized version of the binomial expansion formula:

$$\frac{x^3 + 3x^2h + 3xh^2 + h^3 + 2 - (x^3 + 2)}{h}$$

Handwritten expression on a grid background:

$$(22 - 16\sqrt{3})$$

Digitized version of the expression:

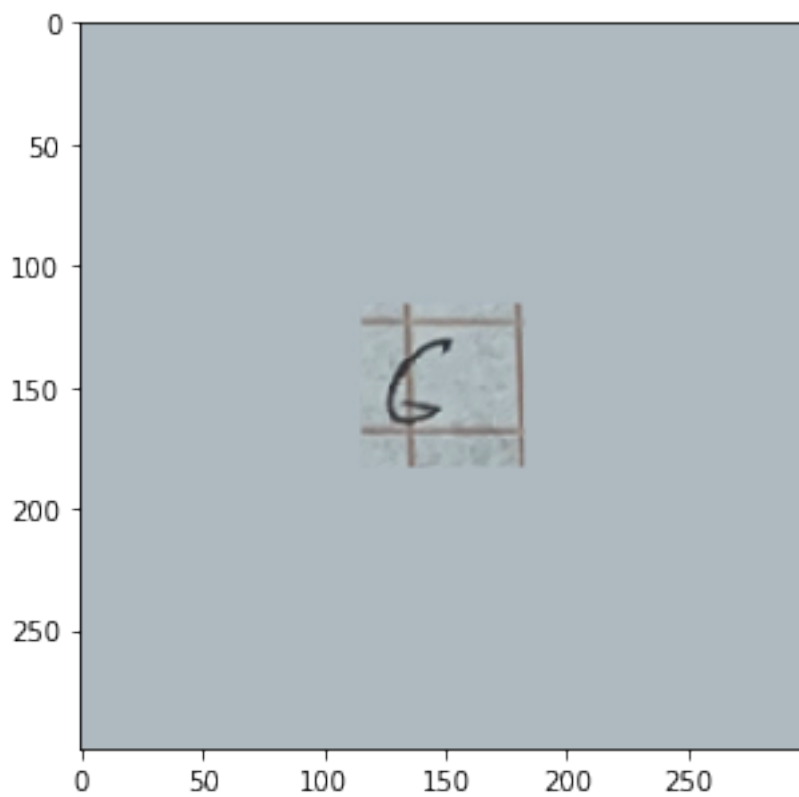
$$(22 - 16\sqrt{3})$$

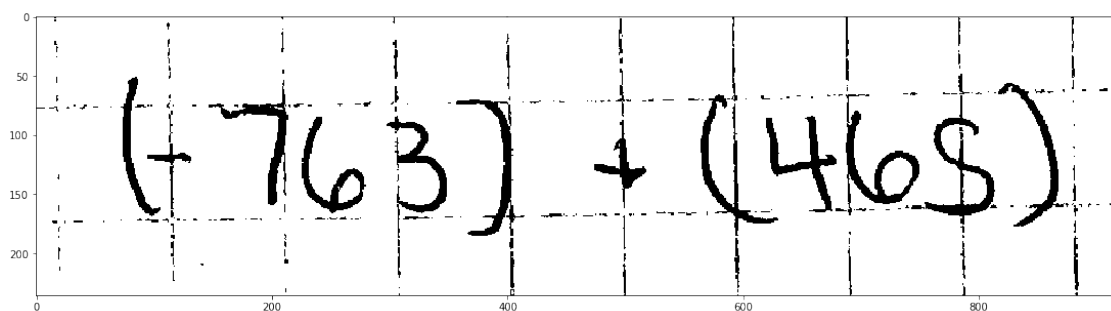
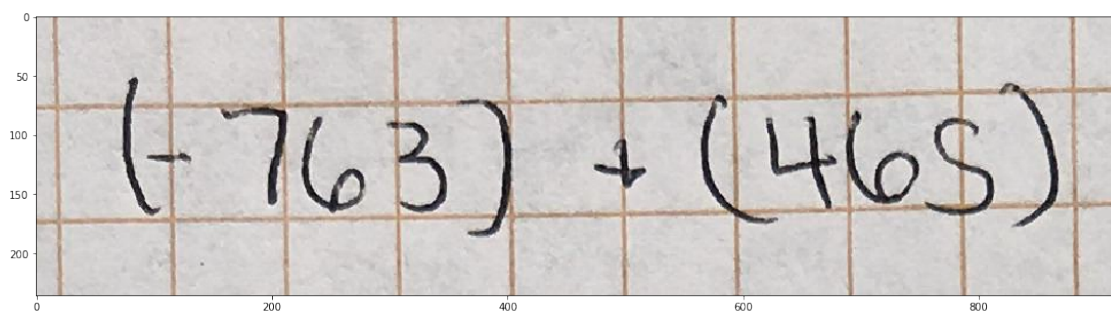
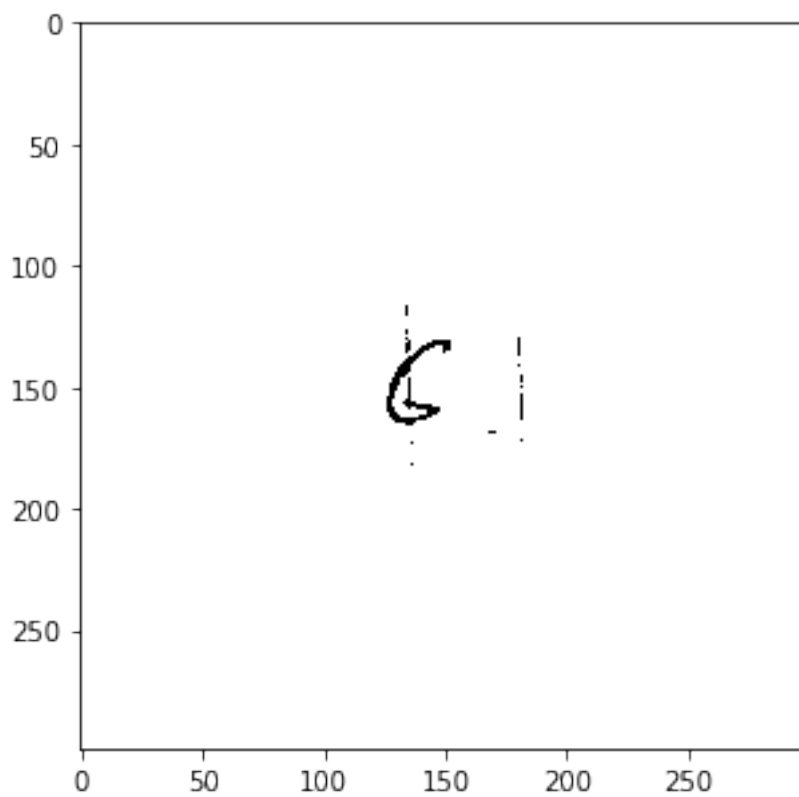
A photograph of a handwritten mathematical equation on a blue grid background. The equation is $\sum_{n=1}^{20} \frac{18}{n^2 - 4} = \infty$. The numbers 20, 18, and the infinity symbol are written in a larger, bolder script than the rest of the equation. The summation symbol is also large and stylized. The grid lines are light blue and spaced at intervals of 50 units on both the x and y axes.

$$\sum_{n=1}^{20} \frac{18}{n^2 - 4} = \infty$$

A photograph of the same handwritten mathematical equation on a white background. The equation is $\sum_{n=1}^{20} \frac{18}{n^2 - 4} = \infty$. The handwriting is consistent with the first image, but the background is plain white. The grid lines are absent, but the equation is centered horizontally and vertically within the frame.

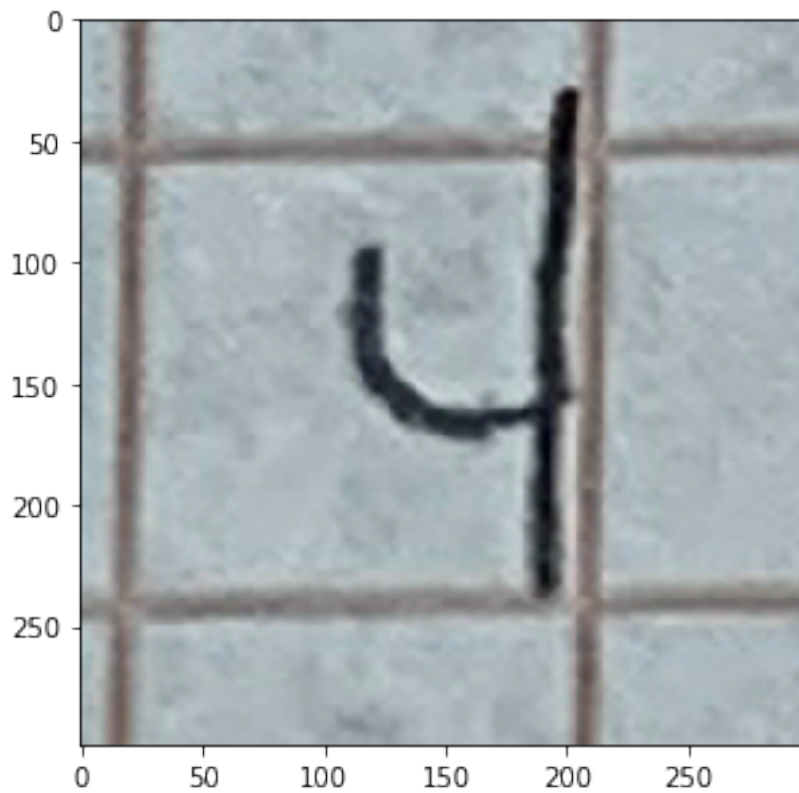
$$\sum_{n=1}^{20} \frac{18}{n^2 - 4} = \infty$$

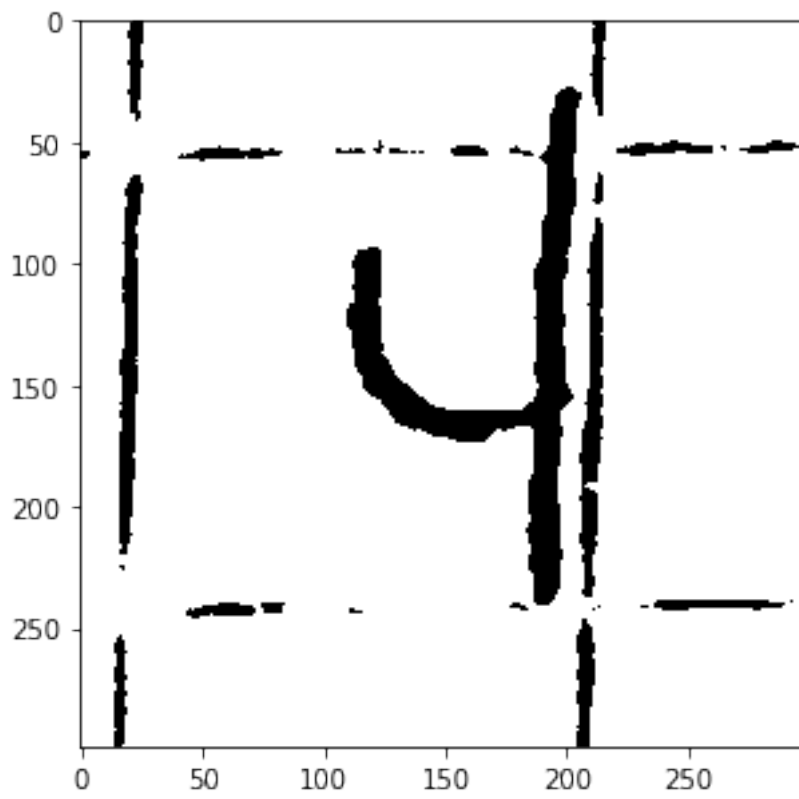




$$h'(r) = \frac{(r+1) \cdot (-1 - \frac{1}{2}r^{-1/2}) - (2-r-r^{1/2}) \cdot 1}{(r+1)^2}$$

$$h'(r) = \frac{(r+1) \cdot (-1 - \frac{1}{2}r^{-1/2}) - (2-r-r^{1/2}) \cdot 1}{(r+1)^2}$$





[]:

[]: