ANALISIS GAMBAR KE TEKS KARAKTER RECOGNITION

(Tugas Proyek Khusus)

Oleh

PUTRA PRIBOWO 1617051005



FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM UNIVERSITAS LAMPUNG BANDAR LAMPUNG

2019

DATA GAMBAR YANG DIGUNAKAN GAMBAR 1

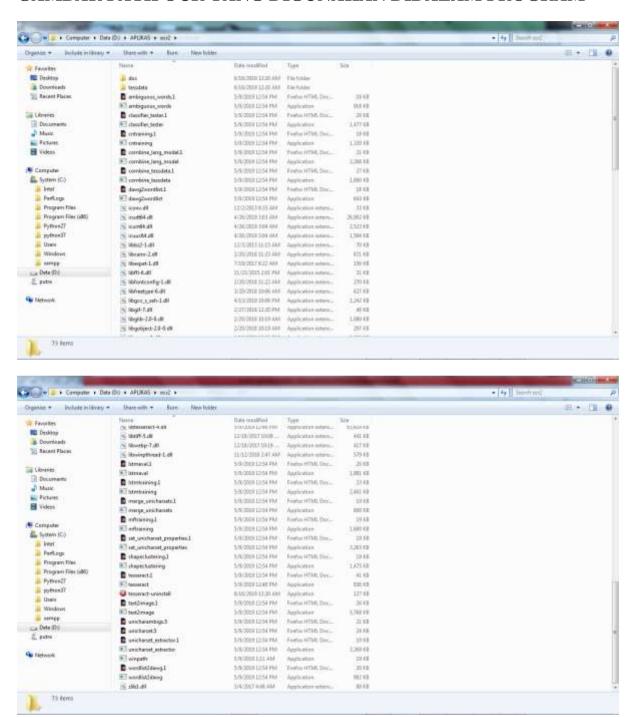


GAMBAR 2



Volwas-2002 senen € 15 Met de bus naar Keukenhof Geldig op connexxion trajecten: Haarlem - Lisse en v.v. Leiden - Lisse en v.v. Alleen geldig op de dag van afstempeling Nº 010848 connexxion brengt mensen bij elkaar

GAMBAR PATH OCR YANG DIGUNAKAN DIDALAM PROGRAM



SOURCODE

```
import cv2
import pytesseract
import numpy as np
import argparse
import os
from PIL import Image
import sys
#untuk ekstrak path tesseract ocr
pytesseract.pytesseract_cmd = r'D:\APLIKAS\ocr2\tesseract'
#untuk memasukkan path gambarnya
def get_string(img_path):
  # membaca gambar dan menjadi RGB
  img = cv2.imread(img_path, cv2.IMREAD_COLOR)
  # merubah ke grayscale
  img_c = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
  # untuk mengurangi noise
  kernel = np.ones((1, 1), np.uint8)
  img_nd = cv2.dilate(img_c, kernel, iterations=1)
  img_ne = cv2.erode(img_nd, kernel, iterations=1)
  # menulis image sebelum merubahnya ke teks
  cv2.imwrite("thres.png", img_ne)
```

```
# menggunakan tesseract
  result = pytesseract.image_to_string(Image.open("thres.png"))
  return result
if __name__ == '__main__':
  from sys import argv
  if len(argv)<2:
    print("python image-to-text.py nama file")
  else:
    print('****************')
    print('Nama: Putra Pribowo')
    print('Npm: 1617051005')
    print('*****************************
    print('--- Memulai ---')
    for i in range(1,len(argv)):
      print(argv[i])
      # memberikan informasi gambar ke teks
      print("\n")
      print('******************************
      print("hasil string saja")
      print('*************************)
      print(get_string(argv[i]))
      print("\n")
      # memberikan informasi kotak estimasi
```

```
print('***************')
 print("gambar ke boxes")
 print('****************')
 print(pytesseract.image_to_boxes(Image.open(argv[i])))
 print("\n")
 # memberikan informasi including boxes, confidences, baris and nomor baris
  print('*************************)
 print("gambar ke data")
 print('****************')
  print(pytesseract.image_to_data(Image.open(argv[i])))
  print("\n")
 # memberikan informasi tentang orientasi dan skrip deteksi
  print('*****************************
 print("gambar ke osd")
 print('****************')
 print(pytesseract.image_to_osd(Image.open(argv[i])))
 print()
 print()
print('-----')
```

HASIL GAMBAR 1

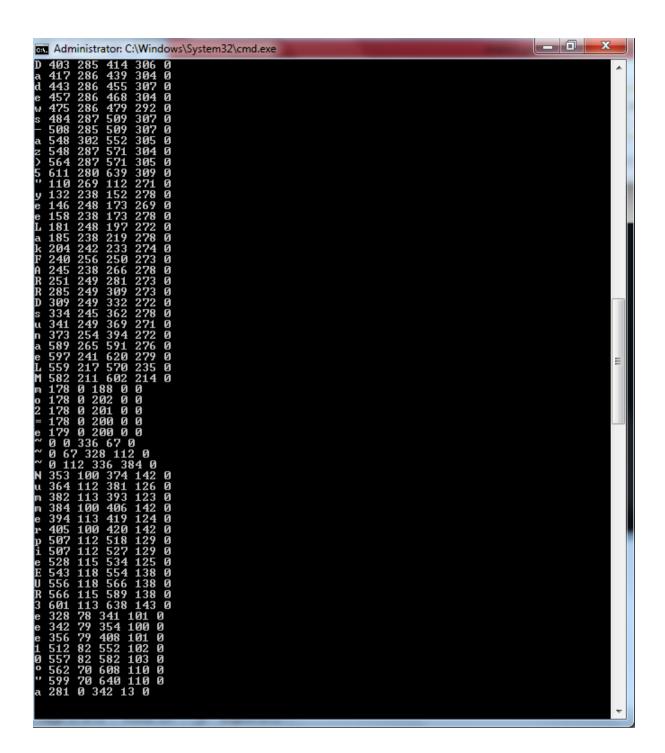
```
_ 0 X
Administrator: C:\Windows\System32\cmd.exe
Microsoft Windows [Version 6.1.7601]
Copyright (c) 2009 Microsoft Corporation. All rights reserved.
C:\python37>python image-to-text.py gambar1/111.png
Nama: Putra Pribowo
Npm : 1617051005
********
--- Memulai ---
gambar1/111.png
*******************
hasil string saja
********
First Founder
Rabobank
 *******************
gambar ke boxes
   ambar ke boxes
*************************
59 518 856 1288 0
512 470 528 502 0
533 470 540 503 0
547 470 559 494 0
562 470 579 494 0
563 470 596 499 0
614 470 630 501 0
633 469 652 493 0
658 470 699 493 0
682 470 699 493 0
684 469 724 503 0
705 470 723 503 0
730 470 765 493 0
140 131 152 140 0
175 131 242 170 0
175 131 242 170 0
265 134 278 159 0
62 0 382 128 0
502 76 771 454 0
514 11 547 55 0
546 11 574 43 0
553 11 608 61 0
574 11 636 57 0
607 11 662 61 0
634 12 692 59 0
691 15 721 47 0
721 15 751 61 0
380 2 910 4 0
o
u
n
d
o
b
 <del>***********</del>
gambar ke data
 _
<del>XXXXXXXXXXXXXXXXXXXX</del>
level
left
                                     width
0
                   page_num
top y
                                                                                              par_num line_num
                                                         block_num
                                                                                                                                                       word_num
                                                        height conf
0 0
                                                                                              text
0
                                                                                                                                                       910
                                                                                                                                                                          1290
```

r ke	data	(****							
pa	age_nur	n	block_n	um		line_nur	n	word_nu	n
1	оp	width 0	height 0	conf Ø	text 0	0	0	910	1290
1		1	0	0	0	59	2	797	770
1		1	1	0	0	59	2	797	770
1		1	1	1	0	59	2	797	770
1		1	1	1	1	59	2	797	770
1		2	0	0	0	512	787	253	34
1		2	1	0	0	512	787	253	34
1		2	1	1	0	512	787	253	34
1		2	1	1	1	512	787	84	33
1	irst	2	1	1	2	614	787	151	34
1	ounder	3	0	0	0	140	1120	102	39
1		3	1	0	0	140	1120	188	45
1		3	1	1	0	140	1120	102	39
1		3	1	1	1	140	1150	12	9
i		3	1	1	2	175	1120	103	39
1	3	4	0	0	0	62	836	709	454
1		4	1	0	0	62	836	709	454
1		4	1	1	0	62	836	709	454
1		4	1	1	1	62	836	709	454
1		5	0	0	0	514	1229	237	50
1		5	1	0	0	514	1229	237	50
1		5	1	1	0	514	1229	237	50
1	abobanl	5	1	1	1	514	1229	237	50
1	anonam	` 6	0	0	0	380	1286	530	2
1		6	1	0	0	380	1286	530	2
1		6	1	1	0	380	1286	530	2
1		6	1	1	1	380	1286	530	2

Administrator:	C:\Windows	\System32\	cmd.exe				-		3
1 1 1 Rabo 1 1 1	5	0	0	0	514	1229	237	50	*
1	5	1	Ø	9	514	1229	237	50	
1	5	1	1	Ø	514	1229	237	50	
1	5	1	1	1	514	1229	237	50	
Rabo 1	рап <i>к</i> 6	0	Ø	Ø	380	1286	530	2	
1	6	1	0	Ø	380	1286	530	2	
1	6	1	1	Ø	380	1286	530	2	
1	6	1	1	1	380	1286	530	2	
ientation tate: Ø ientation ript: Lati ript confi Sukse	confidence n dence: 1.								

HASIL GAMBAR 2

```
_ 0 X
Administrator: C:\Windows\System32\cmd.exe
Nama: Putra Pribowo
Npm : 1617051005
<del>*****************</del>
--- Memulai ---
gambar2/222.png
<del>********</del>
m2 S
*******
gambar ke boxes
```



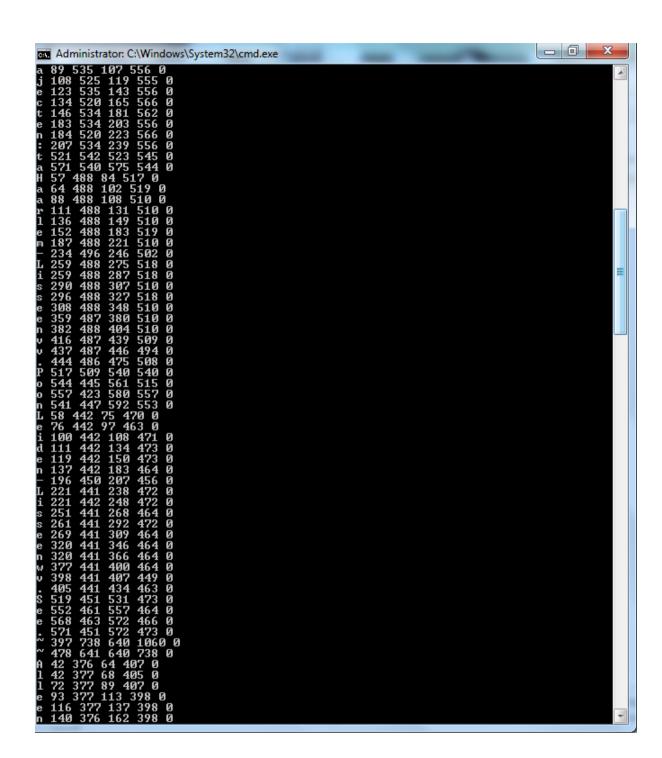
	istrator: C:\V		stem32\cmd	.exe					_ 0 X	
gambar	xxxxxxx ke data									^
level	page_nu	m	block_n		par_num	line_nu	m	word_nu	m	
left 1	top 1	width 0	height 0	conf Ø	text Ø	0	0	640	384	
-1 2 -1	1	1	0	0	0	437	10	140	25	
.5	1	1	1	0	0	437	10	140	25	
-1 4	1	1	1	1	0	437	10	140	25	
-1 5	1	1	1	1	1	437	13	28	22	
5 96 5 95 2	BR 1 232549	1	1	1	2	486	10	91	25	
2	232549 1	2	Ø	0	0	391	45	233	31	
-1 3 -1	1	2	1	0	Ø	391	45	233	31	
4	1	2	1	1	0	391	45	233	31	
-1 5 74 5 48 2 -1	1 Of wifter	2	1	1	1	391	45	94	31	
74 5 40	Afgifte 1 Kiso	2	1	1	2	535	50	40	14	
2	1	3	0	0	0	110	75	529	98	
3 -1	1	3	1	0	0	110	75	529	98	
	1	3	1	1	0	138	75	501	32	
4 1 5 0 5 0 5 4 5 3 4 1	1	3	1	1	1	138	103	5	4	
5	; 1 Dadews-	3	1	1	2	403	77	106	22	
5 42	1 az)	3	1	1	3	548	79	23	18	
5 5	1 5	3	1	1	4	611	75	28	29	ш
4 -1	1	3	1	2	0	110	105	510	41	
5 0	1,	3	1	2	1	110	113	2	2	
5 58	1 yee	3	1	2	2	132	106	41	40	Ш
5 31	yee 1 Lak	3	1	2	3	181	106	52	40	
5 9 5 9 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5	1 FAR	3	1	2	4	240	106	41	40	
5 33	1 RD	3	1	2	5	285	111	47	24	
5 18	1 sun	3	1	2	6	334	106	60	33	
5 18 5 52	1 ae	3	1	2	7	589	105	31	38	
32 4 −1	1	3	1	3	0	559	149	43	24	
5 20	1 LM	3	1	3	1	559	149	43	24	÷
										_

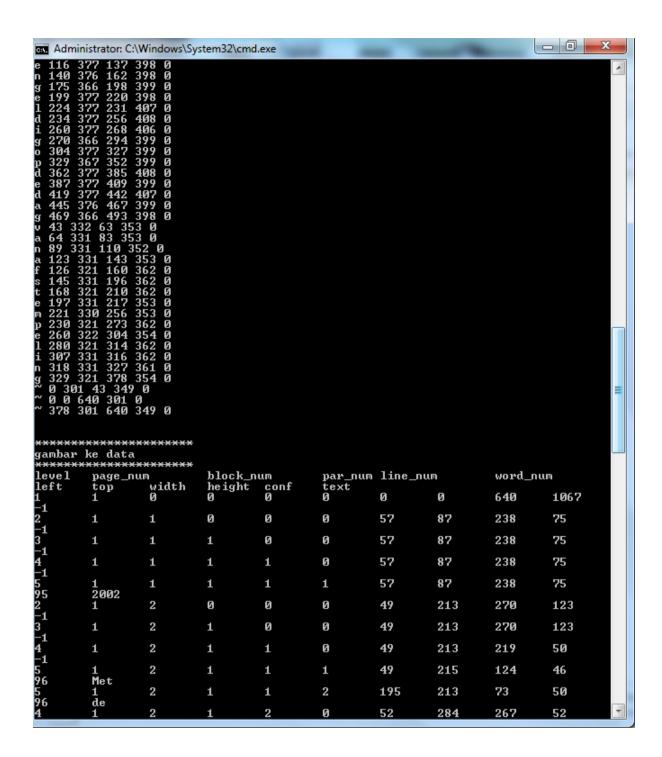
Administrator: C	\Window:	s\System32\	cmd.exe				-		3
1,	3	1	2	1	110	113	2	2	B
1 8 yee	3	1	2	2	132	106	41	40	
1 1 1 Lak	3	1	2	3	181	106	52	40	
1 Lak 1 8 FAR	3	1	2	4	240	106	41	40	
1	3	1	2	5	285	111	47	24	
3 RD	3	1	2	6	334	106	60	33	
8 sun 1	3	1	2	7	589	105	31	38	
2 ae 1	3	1	3	Ø	559	149	43	24	
1 0 LM	3	1	3	1	559	149	43	24	
U LM	4	0	0	0	230	182	134	24	
1 1	4	1	0	0	230	182	134	24	
1 1	4	1	1	0	230	182	25	24	ı
1	4	1	1	1	230	182	25	24	
mo 1	4	1	2	Ø	259	183	24	23	ı
1 1	4	1	2	1	259	183	24	23	ı
$egin{matrix} 1 & 1 \\ 2 \\ 1 \end{bmatrix}$	4	1	3	Ø	286	184	24	22	
1 1	4	1	3	1	286	184	24	22	
5 = 1	4	1	4	Ø	341	184	23	21	ı
1 1	4	1	4	1	341	184	23	21	ı
e 1	5	0	ø	Ø	Ø	9	336	384	ı
1 1	5	1	Ø	Ø	ø	9	336	384	ı
1 1	் 5	1	1	ø	Ø	9	336	384	ı
1 1	ਂ 5			1	9	9	336	384	
5 Î	6	9	9	9	328	241	312	73	
1 1	6	1	ø	0	328	241	312	73	:
1 1	6	1	1	ø	353	241	285	43	
1 1	6	1	1	1	353	242	67	42	
	. 6	1	1	2	507	255	27	17	
5 Nummer 1 0 pie 1 EUR	6	1	1	3	543	246	46	23	
ĒUR	6								
5 Nummer 1 9 pie EUR 1 3	ь	1	1	4	601	241	37	30	
1	6	1	2	0	328	274	312	40	Ļ

Administrator: C	:\Windows	\System32\	cmd.exe				-	
2 1	4	ì	3	Ø	286	184	24	22
1	4	1	3	1	286	184	24	22
- 1	4	1	4	Ø	341	184	23	21
1	4	1	4	1	341	184	23	21
e 1	5	0	0	0	0	0	336	384
1	5	1	0	0	Ø	Ø	336	384
1	5	1	1	Ø	Ø	Ø	336	384
1	5	i	1	1	Ø	Ø	336	384
1	6	Ø	0	0	328	241	312	73
ī	6	1	0	0	328	241	312	73
1	6	ī	1	0	353	241	285	43
î	6	i	1	1	353	242	67	42
Numme: 1		1	1	2	507	255	27	17
pie	6	i	1	3	543	246	46	23
ĒUR 1	6	i	1	4	601	241	37	30
3	6	1	2	9	328	274	312	40
1	6	i	2	1	328	283	80	23
eee								
100"	6	1	2	2	512	274	128	40
1	7	0	0	9	281	371	61	13
1	7	3 1 S	0	9	281	371	61	13
11 S	7	1	1	9	281	371	61	13
1 a	7	1	1	1	281	371	61	13
**************************************	******** d degree	** es: 0						
ientation coript: Cyril ript: Cyril ript confide Sukses	onfidend lic ence: 0.	ce: 0.27 .00 -						

HASIL GAMBAR 3

```
_ 0 X
 Administrator: C:\Windows\System32\cmd.exe
Alleen geldig op de dag
van afstempeling eos
Nº 910848 &
  ******
  gambar ke boxes
              57 908 107 980 0
117 907 171 979 0
180 906 234 978 0
244 905 295 977 0
49 807 108 852 0
113 806 138 852 0
113 806 173 851 0
195 805 229 854 0
235 804 268 839 0
52 734 87 783 0
93 733 158 767 0
181 733 214 766 0
257 731 252 766 0
220 732 252 766 0
220 732 252 766 0
220 732 252 766 0
220 732 252 766 0
220 732 252 766 0
220 732 252 766 0
255 731 319 766 0
104 647 144 690 0
153 647 194 690 0
163 645 243 707 0
295 646 336 688 0
295 646 336 688 0
295 646 388 707 0
295 646 388 707 0
295 646 386 600 0
317 580 153 601 0
251 579 269 601 0
218 570 238 601 0
251 579 366 601 0
218 570 238 601 0
251 579 366 600 0
321 579 342 600 0
323 577 366 600 0
321 579 365 600 0
321 579 365 600 0
323 577 366 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
321 579 365 600 0
323 577 473 599 0
450 577 473 599 0
450 577 473 599 0
450 577 473 599 0
450 577 473 599 0
450 577 473 599 0
450 577 473 599 0
54 535 69 562 0
002MetdebusnaarKeukenhofGeldigopconnexxion-
```



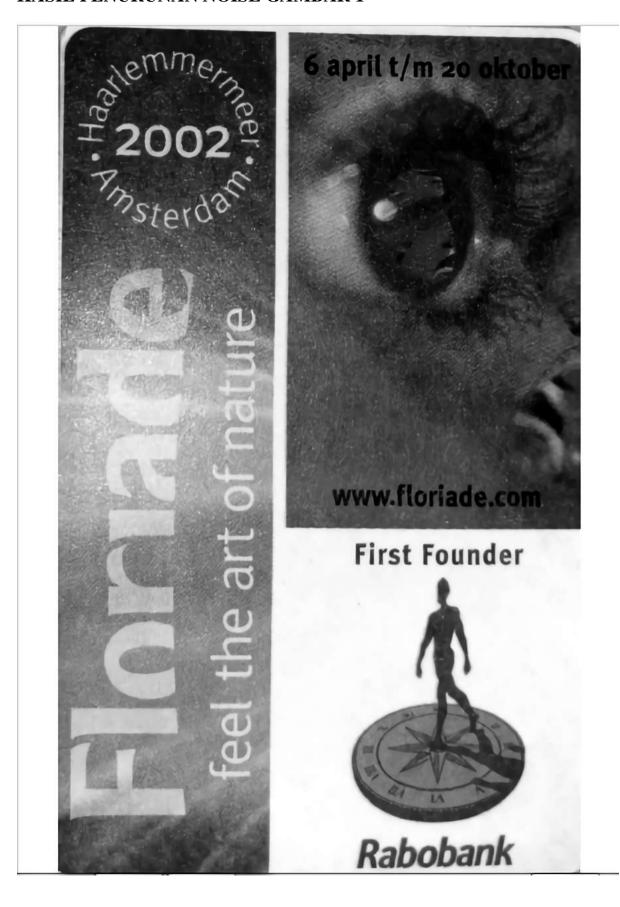


Adr	ministrator: C:					-	-	-		2
	1	2	1	1	0	49	213	219	50	
	1 Met	2	1	1	1	49	215	124	46	
	1	2	1	1	2	195	213	73	50	
	de 1	2	1	2	Ø	52	284	267	52	
	1 bus	2	1	2	1	52	284	106	50	
	1 naar	2	1	2	2	181	301	138	35	
	1	3	0	0	0	56	360	422	62	
	1	3	1	0	0	56	360	422	62	
	1	3	1	1	0	56	360	422	62	
	9		1	1	1	56	360	422	62	
	1 Keuken	hof								
	1	4	0	0	0	55	432	583	86	
	1	4	1	0	Ø	55	432	585	86	
	1	4	1	1	0	55	432	583	86	
	1	4	1	1	1	55	456	125	42	
	Geldig 1	4	1	1	2	192	466	46	31	
	op				3		461		29	
	1 connex		1	1		251		222		
	1	4	1	1	4	516	432	122	86	
	1	5	0	0	0	54	501	538	143	
	1	5	1	0	0	54	501	586	125	
	1	5	1	1	0	54	501	521	46	
	1	5	1	1	1	54	501	185	46	
	trajec	ten:				521	522	54	5	
	1 ta	5	1	1	2					
	1	5	1	2	0	57	510	535	134	
	1 Haarle	5	1	2	1	57	548	164	31	
	1	5	1	2	2	234	565	12	6	
	1	5	1	2	3	259	549	89	30	
	1 Lisse 1	5	1	2	4	359	557	45	23	
	en									
	1	5	1	2	5	416	558	59	23	
	1 Poon	5	1	2	6	517	510	75	134	
	1	5	1	3	0	58	594	514	32	
	1	5	1	3	1	58	594	125	31	

1	5	1	2	4	359	557	45	23
en 1	5	1	2	5	416	558	59	23
1	5	1	2	6	517	510	75	134
Poon 1	5	1	3	Ø	58	594	514	32
1 Le iden	5	1	3	1	58	594	125	31
Leiden 1	5	1	3	2	196	611	11	6
- 1 Lisse	5	1	3	3	221	595	88	31
Lisse 1	5	1	3	4	320	603	46	23
en 1	5	1	3	5	377	603	57	23
ωο. 1	5	1	3	6	519	594	53	22
See.	6	Ø	Ø	Ø	397	7	243	419
1	6	1	Ø	Ø	397	7	243	419
1	6	1	1	Ø	397	7	243	419
1	6	1	1	1	397	7	243	419
1	7	Ø	Ø	Ø	42	659	451	87
1	7	1	Ø	Ø	42	659	451	87
1	7	1	1	Ø	42	659	451	42
1	7	1	1	1	42	660	120	31
Alleen 1	7	1	1	2	175	659	119	42
geldig 1	7	1	1	3	304	668	48	32
op 1	7	1	1	4	362	659	47	31
de 1	7	1	1	5	419	660	74	41
dag 1	7	1	2	Ø	43	705	335	41
1	7	1	2	1	43	714	67	22
van 1	?.	1	2	2	123	705	255	41
afstem 1	peling 8	Ø	Ø	Ø	0	718	640	349
1	8	1	Ø	Ø	Ø	718	640	349
1	8	1	1	Ø	Ø	718	640	349
1	8	1	1	1	Ø	718	640	349

dministrator: C:\\	Windows\	System32\ci	md.exe				-	
Lisse 1	5	1	3	4	320	603	46	23
en 1	5	1	3	5	377	603	57	23
₩0. 1	5	1	3	6	519	594	53	22
See. 1	6	0	0	0	397	7	243	419
1	6	1	0	0	397	7	243	419
1	6	1	1	0	397	7	243	419
1	6	1	1	1	397	7	243	419
1	7	Ø	0	0	42	659	451	87
1	7	1	0	0	42	659	451	87
1	7	1	1	0	42	659	451	42
1	7	1	1	1	42	660	120	31
Alleen 1	7	1	1	2	175	659	119	42
geldig 1	7	1	1	3	304	668	48	32
op 1	7	1	1	4	362	659	47	31
de 1	7	1	1	5	419	660	74	41
dag 1	7	1	2	0	43	705	335	41
1	7	1	2	1	43	714	67	22
van 1	?.	1	2	2	123	705	255	41
afstemp 1	eling 8	Ø	0	0	0	718	640	349
1	8	1	0	0	Ø	718	640	349
1	8	1	1	0	Ø	718	640	349
1	8	1	1	1	Ø	718	640	349
	8 ******	1						
ntation in te: 0 ntation con pt: Latin pt confiden	fidenc	e: 5.49						
Sukses -								

HASIL PENURUNAN NOISE GAMBAR 1



HASIL PENURUNAN NOISE GAMBAR 2

➢ / Nederla	andse Spoorweg	en	BR 23	2549
Geldig op/van	Geldig tot en me	nt Afgift	edatum	Klasse
21.02.03		e Spoorwer 21	02.03	2
Enke	下年.北部·历	Ti ation		
SCHI				Т
	ERDAM	c.		te Sont
Via				
Afgiftepunt 2	Reductie	Nummer		rijs EUR
881.78.2	259	338522	3.	.10
Bijz.				
and Spoore				

2002

Met de bus naar Keukenhof

Geldig op connexxion trajecten:

Haarlem - Lisse en v.v. Leiden - Lisse en v.v.

Alleen geldig op de dag van afstempeling

Nº 010848

connexxion

brengt mensen bij elkaar

Volwassenen



DALAM MENGGUNAKAN SORCODENYA PADA GAMBAR 1

Python image-to-text.py gambar1/111.png

ANALISIS GAMBAR 1

Dalam analisis gambar tersebut sebelumnya membuat ukuran gambar 900x1500 pixel dan mengurangi noise beberapa kali, dan hasilnya tidak semua terbaca tulisan hanya sebagian, dikarenakan kurangnya pixel yang lebih tinggi didalam yang dikonvert dan tingkat akurasi pembacaannya mencapai 25 persen.

Tulisan dalam gambar tersebut mendapatkan hasil first founder dan Rabobank tulisannya dapat terbaca. Dan terdapat suatu nilai yang ditampung yang ditampilkan pada sistem dan menghasilkan nilai akurasi yang menggunakan tulisan apa saja yang digunakan didalam gambar.

DALAM MENGGUNAKAN SORCODENYA PADA GAMBAR 2

Python image-to-text.py gambar2/222.png

ANALISIS GAMBAR 2

Saya menggunakan ukuran gambar 1500x900 pixel dan mengurangi noise beberapa kali dengan converter online dari sedikit kebaca tulisan menjadi agak lebih dari sebelumnya, dan hasilnya tidak semua terbaca tulisan hanya sebagian, dikarenakan kurangnya pixel yang lebih tinggi didalam gambar atau kurang besar.

Tulisan dalam gambar tersebut mendapatkan hasil BR232549 aigittedatum klasse tulisannya dapat terbaca mencapai 35 persen. Dan terdapat suatu nilai yang ditampung berbentuk boxes yang ditampilkan pada sistem dan menghasilkan nilai akurasi yang menggunakan tulisan apa saja yang digunakan didalam gambar yang dapat terlihat pada gambar 2 diatas hasil running programnya.

DALAM MENGGUNAKAN SORCODENYA PADA GAMBAR 3

Python image-to-text.py gambar3/333.png

ANALISIS GAMBAR 3

Dalam analisis gambar 3 ini saya menggunakan ukuran pixel gambar 640x1067 pixel. Didalam gambar saya mengurangi noise pada gambar 1 x di convert di online. Dan hasil dari tulisan pada gambar tersebut dapat terbaca 80 persen tulisan pada gambar tersebut. Tulisan yang terbaca adalah 2002 Met de bus naar Keukenhof gelding op connexxion trajecten: haarlem – lisse en vv leiden – lisse en vv Allen geldig op de dag van afstempeling no 010848.

Yang dapat dilihat pada hasil gambar 3 dan dapat dilihat nilai dari boxes dan akurasi tulisan yang digunakan pada gambar tersebut.