



손글씨 인식 모델

김현주



1

주제 선정 배경



2

진행 상황



3

추후 계획

주제 선정 배경

1

주제 선정 배경

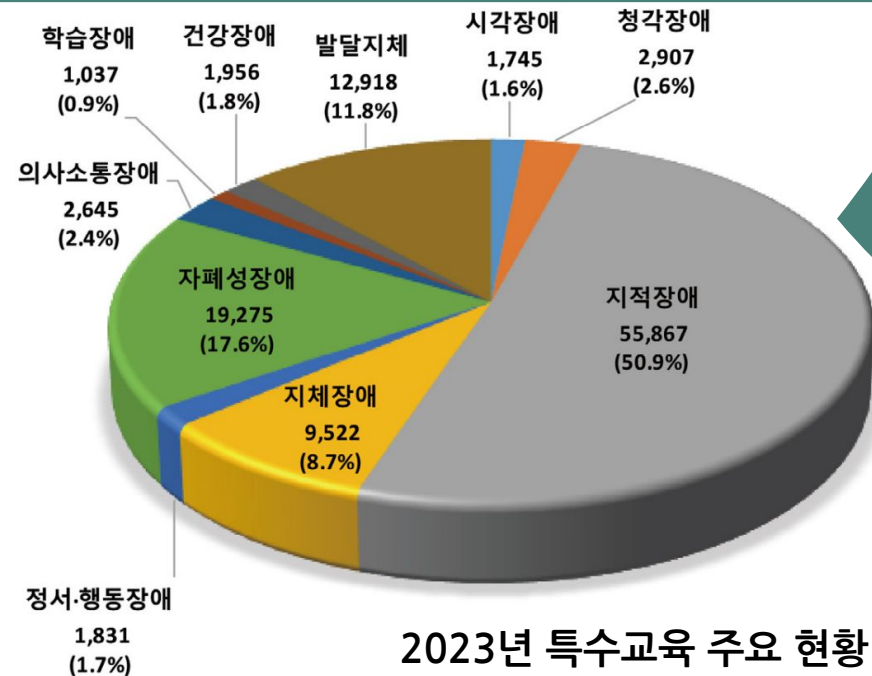
장애 유형은 지적장애(41.9%)가 가장 많으며 다음으로는 자폐성장애(22.8%), 지체장애(11.8%), 발달지체장애(11.3%) 순으로 나타났다. 특히 자폐성장애 학생은 2018년(2천73명)에 비해 2022년(3천44명)에 46.8%나 증가했다. Feb 9, 2023



복지뉴스

<http://m.bokjinews.com/news/articleView> ;

서울 특수교육 학생 4년간 4.9% 증가..."특수학급 259개 확대 ...



지체장애 아동은 어깨, 팔, 손의 **운동능력이 부족**하여
글씨를 잘 쓰지 못하는 경우가 많고 전혀 쓰지 못하거나
연필을 잡고 쓰기는 해도 매우 느리고 힘들어서 **쉽게 피로**해짐

글씨 쓰기가 학교 활동에서 차지하는 비중이 높기 때문에 가능하면
직접 쓸 수 있도록 지도하는 것이 좋음

2

목표

손글씨 인식 모델을 활용해

지체 장애 학생의 **수업 참여율이 증가**하도록 도움을 줄 수 있는
프로그램 개발

하지만 데이터 수집에 어려움이 있어

먼저 **비장애인의 손글씨 데이터**를 학습하고 잘 **인식**할 수 있는
모델을 개발할 예정

3

데이터 출처

1



#공공 #인쇄체 #타자체 #수기 #행정 #외교 #문화 #과학기술 #문자인식 #OCR #AI

NEW OCR 데이터(교육)

분야

 교육

유형

 텍스트, 이미지

구축년도 : 2022 갱신년월 : 2023-11 조회수 : 31,465 다운로드 : 312 용량 : 127.70 GB

다운로드

↓ 샘플 데이터

관심데이터 등록

42

<https://aihub.or.kr/aihubdata/data/view.do?currMenu=&topMenu=&aihubDataSe=data&dataSetSn=71300>

2



#자연어처리 #광학글자인식 #한글 문자 이미지 데이터 #문서 이미지 분석/분류 #로봇 프로세스 자동화 #인쇄체

다양한 형태의 한글 문자 OCR

분야

 한국어

유형

 이미지

구축년도 : 2020 갱신년월 : 2021-06 조회수 : 33,104 다운로드 : 2,534 용량 : 39.61 GB

다운로드

↓ 샘플 데이터

관심데이터 등록

83

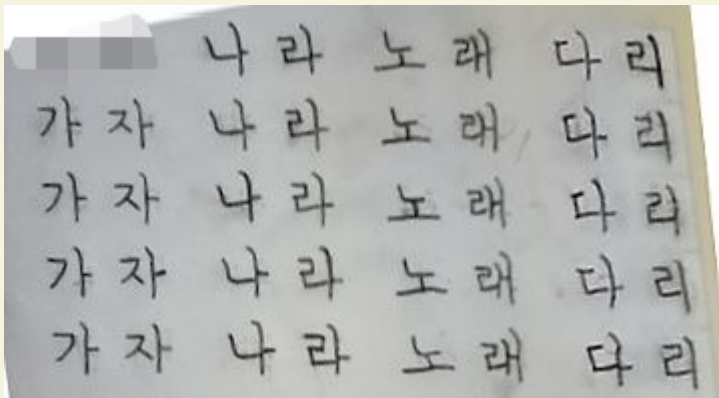
<https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=data&dataSetSn=91>



진행 상황

1

교육용 OCR - 데이터 확인



```
{
  "Annotation": {
    "object_recognition": 1,
    "text_language": 0
  },
  "Dataset": {
    "category": 0,
    "identifier": "OCR(edu)",
    "label_path": "OCR(edu)/E/G1",
    "name": "대규모 OCR 데이터(교육)",
    "src_path": "OCR(edu)/E/G1",
    "type": 1
  },
  "Images": {
    "acquisition_location": "타임교육C\u0026P",
    "application_field": "자유형",
    "area": 2,
    "background": 0,
    "class": 1,
    "data_captured": "2022.10.25 02:48:54",
    "dpi": 300,
```

```
    "grade": 1,
    "group": 1,
    "height": 196,
    "identifier": "EDU_E1_032060",
    "make_id": "R154",
    "section": 1,
    "type": "png",
    "width": 358
  },
  "Bbox": [
    {
      "data": "가자",
      "id": 1,
      "type": 1,
      "x": [15, 15, 78, 78],
      "y": [44, 75, 44, 75]
    },
  ],
```


1

교육용 OCR

```
# 파일 정보를 읽어서 데이터프레임 생성
df_list = []
path = './data/labels/'
grades = os.listdir(path)
for grade in grades:
    files = os.listdir(path+grade)
    for file in files:
        with open(path+grade+'/'+file, 'r') as f:
            data = json.load(f)
            # 파일명, 파일의 x,y좌표값, 손글씨 라벨 리스트
            file_name = data["Images"]["identifier"]
            file_type = data["Images"]["type"]
            data_num = len(data["Bbox"])
            info = data["Bbox"]
            for i in range(data_num):
                data_list = []
                label = info[i]['data']
                x_start = info[i]['x'][0]
                y_start = info[i]['y'][0]
                x_end = info[i]['x'][-1]
                y_end = info[i]['y'][-1]

                data_list.append(grade+'/'+file_name+'.'+file_type)
                data_list.append(label)
                data_list.append(x_start)
                data_list.append(y_start)
                data_list.append(x_end)
                data_list.append(y_end)

            df_list.append(data_list)
```

```
{
  "Annotation": {
    "object_recognition": 1,
    "text_language": 0
  },
  "Dataset": {
    "category": 0,
    "identifier": "OCR(edu)",
    "label_path": "OCR(edu)/E/G1",
    "name": "대규모 OCR 데이터(교육)",
    "src_path": "OCR(edu)/E/G1",
    "type": 1
  },
  "Images": {
    "acquisition_location": "타임교육C\u0026P",
    "application_field": "자유형",
    "area": 2,
    "background": 0,
    "class": 1,
    "data_captured": "2022.10.25 02:48:54",
    "dpi": 300,
```

```
    "grade": 1,
    "group": 1,
    "height": 196,
    "identifier": "EDU_E1_032060",
    "make_id": "R154",
    "section": 1,
    "type": "png",
    "width": 358
  },
  "Bbox": [
    {
      "data": "가자",
      "id": 1,
      "type": 1,
      "x": [15, 15, 78, 78],
      "y": [44, 75, 44, 75]
```

1

교육용 OCR

	file_name	label	x_start	y_start	x_end	y_end
0	E1/EDU_E1_032060.png	가자	15	44	78	75
1	E1/EDU_E1_032060.png	가자	17	83	79	113
2	E1/EDU_E1_032060.png	가자	21	118	81	151
3	E1/EDU_E1_032060.png	가자	21	156	86	187
4	E1/EDU_E1_032060.png	나라	97	7	163	38
...
366233	M3/SCH_M3_010692.png	\angle CDA	436	498	614	592
366234	M3/SCH_M3_010692.png	(\angle BAD	204	601	442	693
366235	M3/SCH_M3_010692.png	=	441	621	482	677
366236	M3/SCH_M3_010692.png	\angle DCB)	479	595	669	678
366237	M3/SCH_M3_010692.png	\therefore	146	501	202	558

전체 이미지 크롭 후 저장하기

```
folder_path = './data/images/'
```

```
for i in range(len(dataDF)):
```

```
    img_path = folder_path+'/' + dataDF['file_name'][i]
```

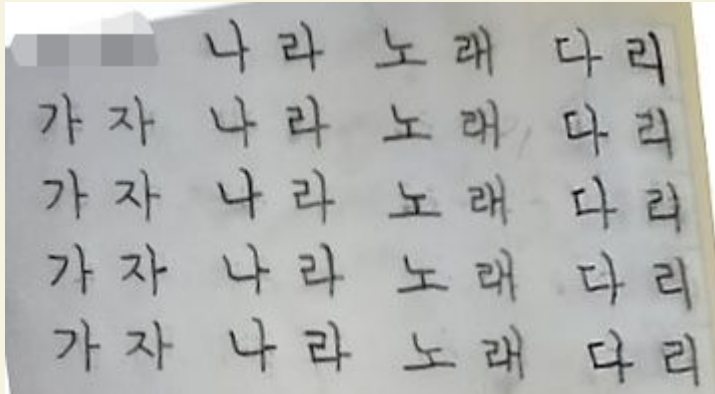
```
    img = Image.open(img_path)
```

```
    crop_img = img.crop(tuple(dataDF.iloc[i])[-4:])
```

```
    crop_img.save(f"./crop_image/{i}_{dataDF['label'][i]}"+'.png')
```

1

교육용 OCR



행 번호를 붙여서
이미지 저장



1

교육용 OCR

	file_name	label	x_start	y_start	x_end	y_end
0	E1/EDU_E1_032060.png	가자	15	44	78	75
1	E1/EDU_E1_032060.png	가자	17	83	79	113
2	E1/EDU_E1_032060.png	가자	21	118	81	151
3	E1/EDU_E1_032060.png	가자	21	156	86	187
4	E1/EDU_E1_032060.png	나라	97	7	163	38
...
366231	M3/SCH_M3_010692.png	\angle CDA	436	498	614	592
366232	M3/SCH_M3_010692.png	(\angle BAD	204	601	442	693
366233	M3/SCH_M3_010692.png	=	441	621	482	677
366234	M3/SCH_M3_010692.png	\angle DCB)	479	595	669	678
366235	M3/SCH_M3_010692.png	\therefore	146	501	202	558

라벨의 고유값과 각 값의 빈도수 출력

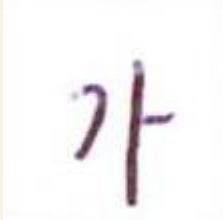
```
len(dataDF['label'].unique()), dataDF['label'].value_counts()
```

110,295

단어로 학습을 진행하면 경우의 수가 너무 많음
→ 한 글자씩 학습 진행

2

한 글자씩 학습



```
{  
  "info" : {  
    "name" : "Korean OCR Data Set",  
    "description" : "Korean OCR Data Set (letter handwriting)",  
    "date_created" : "2020-12-22 13:38:11",  
    "text" : "가"  
  },  
  "image" : {  
    "file_name" : "00130001001.jpg",  
    "width" : 111,  
    "height" : 110,  
    "dpi" : 300,  
    "bit" : 24  
  },  
  "text" : {  
    "type" : "letter",  
    "output" : "handwrite",  
    "letter" : {  
      "value" : "가"  
    }  
  }  
}
```


2

한 글자씩 학습

```
# 각 글자 이름의 폴더 생성
for label in labels:
    os.mkdir(f"./letter/{label}")
```

앞과 같이 데이터프레임을 만든 후
각 라벨의 이름과 같은 폴더를 생성

```
# 이미지 읽어서 저장하기
file_path = './data/handwrite_image/'
for i in range(len(dataDF)):
    img_path = file_path+dataDF['image_path'][i]
    img = Image.open(img_path)
    img.save(f"./letter/{dataDF['label'][i]}/{dataDF['image_path'][i][:3]}"+'.jpg')
```

새로 생성한 폴더에 이미지 다시 저장

2

한 글자씩 학습

```
1 # 이미지 데이터셋 생성
2 imgDS = ImageFolder(root=IMG_PATH, transform=transConvert)
✓ 1.9s
```

```
1 print(f'imgDS.classes : {imgDS.classes}')
2 print(f'imgDS.class_to_idx : {imgDS.class_to_idx}')
3 print(f'imgDS.targets : {imgDS.targets}')
✓ 0.0s
```

```
imgDS.classes : ['가', '각', '간', '갸', '갈', '깁', '깃', '감', '
imgDS.class_to_idx : {'가': 0, '각': 1, '간': 2, '갸': 3, '갈': 4,
imgDS.targets : [0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0,
```



```
1 len(imgDS)
✓ 0.0s
```

276886

ResNet50 활용해
전이학습 수행 중



[EPOCH 26/50]

```
-----
- [TRAIN] LOSS : 2.2796462045297203 F1SCORE : 0.3401109931126041 ACCURACY : 0.34063307157266043
- [VALID] LOSS : 6.697066157488104 F1SCORE : 0.0875964167666127 ACCURACY : 0.08814444371129497
scheduler.num_bad_epochs : 0
```

[EPOCH 27/50]

```
-----
- [TRAIN] LOSS : 2.213554882424666 F1SCORE : 0.35142941541445666 ACCURACY : 0.35202064042286463
- [VALID] LOSS : 6.874000700548988 F1SCORE : 0.08595958971894593 ACCURACY : 0.08651730239473107
scheduler.num_bad_epochs : 1
```



추후 계획

1

성능 개선

- 더 많은 손글씨 데이터 + 인쇄체도 함께 학습
- 데이터에 맞는 다른 모델도 사용하여 학습

(가능하다면 Swin Transformer...)

2

CRAFT 모델 사용

CRAFT 모델은 텍스트 검출을 위해 단어를 바로 예측하는 것이 아니라,

문자의 위치를 나타내는 **region score**,

문자간 거리를 나타내는 **affinity score**를 예측



2

CRAFT 모델 사용

입력 이미지는 한 글자가 아닌 **단어** 또는 **문장 단위**

→ 입력된 이미지의 위치를 찾아 한글자씩 자르기 위해
CRAFT 모델 사용

→ 최종 시연을 할 때는 글자의 **위치를 찾는 모델**과
그 위치의 **글자를 인식**하는 모델 2가지 사용할 계획

2

CRAFT 모델 사용

이미지 업로드 → 이미지 자르기 →
한 글자씩 읽기 → 결과 통합 후 출력

3

웹페이지 구현



모델부터 잘
만들겠습니다...

The background is a light cream color with several abstract, organic shapes in teal, orange, and red. A large teal shape is in the top right, a smaller orange circle is in the top left, and a large orange shape with a white hole is in the bottom left. A red shape is in the middle right, and another red shape is in the bottom center.

감사합니다

김현주