

# Hi Big Data, AI

AI 기반 문서 인식 자동화 솔루션 도입 제안



Socioloy 201527014 류명기  
e-Business 201646006 이성범

# Table of Content

1

## 문제제기 및 도입 필요성

왜 AI기반 문서 인식 솔루션을 도입해야 할까요?

2

## 활용사례 (Use Case)

실제 우리 주변에서 활용되고 있는 문서 인식 솔루션

3

## 기대효과

AI기반 문서 인식 솔루션을 도입하게 된다면 달라지는  
한신의 모습!

4

## 프로젝트 설명

현재 진행하고 있는 손 글씨 인식 솔루션 설명

01.

# 문제제기 및 도입의 필요성

앞으로 다가 올 디지털 트랜스포메이션 시대!  
한신은 왜 AI 기반 문서 인식 솔루션을 도입해야 할까요?



# 문제제기



현재 대한민국은 기관의 업력이 길수록  
“아날로그 데이터 축적량”이 많음



코로나19 영향으로 비대면 거래 비중이  
늘어나는 등 언택트 문화가 자리잡으면서  
확대되는 ‘페이퍼리스(paperless)’



가치 있는 데이터를 제대로 선별하지 않고  
스캔, PDF 파일화를 통한 “단순적재형태”는  
효율적인 데이터베이스 구축에 어려움만 증가



앞으로 제 4차 산업혁명시대의 교육과 행정은  
온라인을 주축으로 이루어질 것



빠르게 변화하는 디지털 시대에서 뒤처지지  
않으려면, 문서에서 필요한 정보를 쉽게 찾고  
빠르게 접근할 수 있어야 함



따라서 종이 형태로 존재하는 오프라인 문서  
및 행정자료를 파일 형태의 온라인 데이터로  
저장하는 과정이 필요함



문서를 얼마나 빠르고 효율적으로 처리 하는 것은  
전체 생산성에 상당한 영향을 미칠 것

**“AI 기반 문서 인식 솔루션은  
미래 교육 경쟁력의  
기본적인 초석”**

02.

## 활용사례 (Use Case)

우리 주변에서는 AI 기반 문서 인식 솔루션이  
어떻게 사용되고 있을까요?!



## 활용사례 1

### AI 문자인식(OCR) 솔루션 텍스트스코프(Textscope™)



스마트폰 및 맥스로 촬영한  
저화질 문서 OCR지원



회전된 입력이미지  
자동 회전 인식



문서 종류 자동분류  
및 처리 최적화



비정형 문서를 통한  
중요 정보 추출

 Lomin

통장 사본, 팩스 사본, 사업자 등록증 등  
오프라인 문서를 AI 문자인식 솔루션 활용 하여  
일일이 입력 할 필요 없이, 간단하게 내 폴더에 !

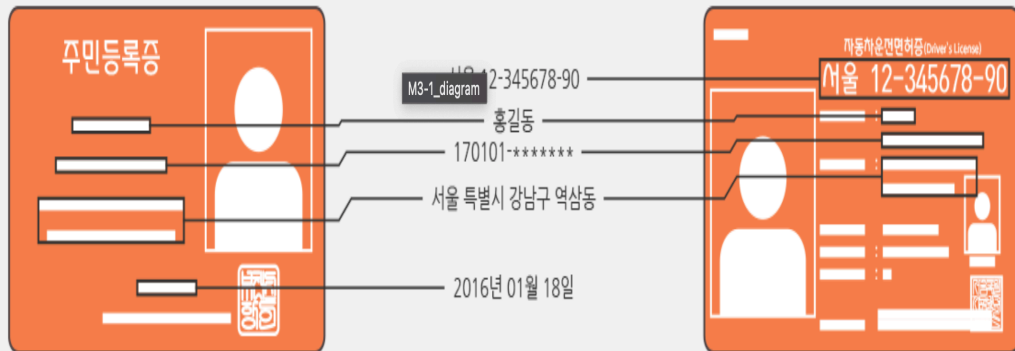
## 활용 사례 및 기능 설명 (영상)



<https://youtu.be/i6ajuLybWLQ>



## 활용사례 2



신분증에서도 필요한 정보만  
추출하여 보관 !

03.

# 기대효과

AI 기반 문서 자동화 솔루션을 적용하면 무엇이 달라지며  
한신대학교에는 어떻게 적용 할 수 있을까요?





## 효율성

문서를 분류하고  
데이터를 추출하는  
과정의 효율성

## 데이터의 자산화

아날로그 데이터 수집 역량을 강화  
메타데이터 수집 자동화로 가치 있는 데이터의 자산화

## 미래 교육 준비

데이터 가치를 끌어올리는 동시에  
그와 관련한 AI 서비스를 개발 하고  
운영하는데 있어서 중요한 초석으로 작용

# 기대효과

# 한신대학교, AI 문서 자동화 솔루션을 만나다

## 도서관 책 보관

매년 버려지는 한신대학교  
도서관의 오래된 책들,  
이제 문서화하여 보관해요 !

## 교내 다양한 영수증 처리

교내에서 처리되는 다양한 영수증,  
이제 문서 자동화로 간단하게 파일 변환해서 정리해요 !

## 손 시험지 관리 및 보관

매 시험마다 알 수 없는 글씨들,  
처리가 곤란한 시험지들은 이제 안녕~  
인식 한 번으로 평가와 보관이 끝 !

## 오프라인 문서들의 처리 및 보관

팩스, 통장사본, 처방전 등 오프라인 문서, 손으로 다시 한  
번 입력하는 번거로움 없이 바로 문서로 저장 !

# 04.

## 프로젝트

일반 문서를 넘어, 손 글씨까지 인식하여 문서화 하는 프로젝트





현재의 문서 자동화 솔루션과 함께

ICR(Intelligence Character Recognition, 지능형 필기체 문자 인식 기술)  
기술을 접목한다면 아날로그 데이터의 문자를 추출하고  
디지털 데이터로 변환하는 업무의 효율성 증가



# 프로젝트 진행상황

한글의 경우 총 11,172개의 문자로 구성

실제 손 글씨 데이터를 확보가 쉽지 않음

현재 프로젝트의 가능성을 확인하고자 6만개의  
0~9까지의 숫자 손 글씨가 적힌 **MNIST** 데이터 셋을  
활용하여 가능성 확인

# 개발환경

Python을 개발 언어로 선택

Google Colab에서 진행

딥 러닝의 모델의 빠른 학습을 위해서 GPU를 사용

모델은 Python의 딥 러닝 라이브러리 keras 활용

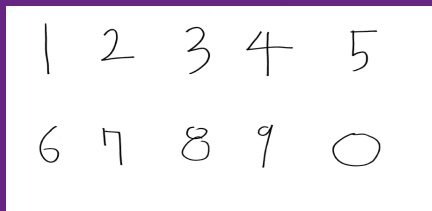




# 전체 프로세스

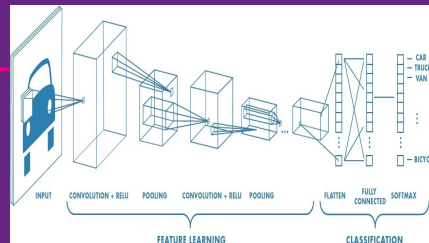


MNIST  
데이터 셋



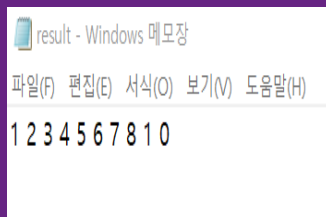
나의  
손글씨  
데이터

CNN  
모델



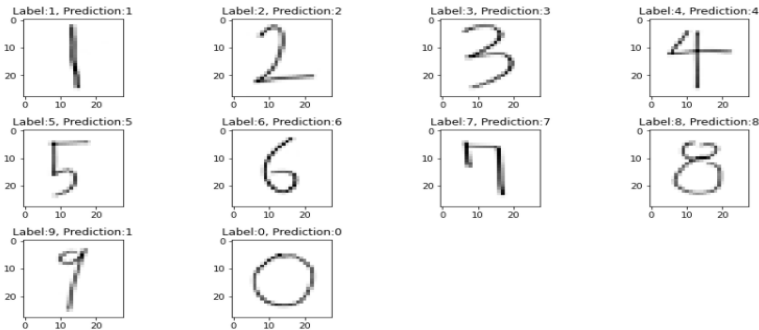
학습된  
CNN  
모델

텍스트  
파일로  
출력



## 결과

직접 손 글씨를 작성 한 후 모델에  
넣어본 결과 9를 제외한 모든 숫자를  
정확하게 예측했으며, 예측한 데이터  
또한 txt 파일로 자동 저장이 되었음



result - Windows 메모장

파일(F) 편집(E) 서식(O) 보기(V) 도움말(H)

1 2 3 4 5 6 7 8 1 0

# 이제는 문서 인식 자동화 솔루션을 도입할 때!

## 교육 및 행정 처리의 효율성

문서자동화로 기존에  
낭비 되던 시간과  
자원은 이제 그만

## 교육의 질적 발전

언제 어디서나 활용 가능 한  
디지털 문서로 교육의 질적 발전



## 새로운 가치에 투자

남는 자원과 시간을 새로운  
미래 교육에 투자



**Thanks !**