

# ResNet-50 모델을 이용한 손글씨 데이터 세트의 분류 성능 분석 및 비교

송제용<sup>0</sup>, 시종욱<sup>\*\*</sup>, 김성영<sup>\*</sup>

<sup>0</sup>금오공과대학교 컴퓨터공학과,

<sup>\*</sup>금오공과대학교 컴퓨터공학과,

<sup>\*\*</sup>금오공과대학교 컴퓨터·AI융합공학과

e-mail: joon6093@naver.com<sup>0</sup>, jwsi425@kumoh.ac.kr<sup>\*\*</sup>, sykim@kumoh.ac.kr<sup>\*</sup>

## Analysis and Comparison of Classification Performance on Handwritten Datasets using ResNet-50 Model

Jeyong Song<sup>0</sup>, Jongwook Si<sup>\*\*</sup>, Sungyoung Kim<sup>\*</sup>

<sup>0</sup>Dept. of Computer Engineering, Kumoh National Institute of Technology,

<sup>\*</sup>Dept. of Computer Engineering, Kumoh National Institute of Technology,

<sup>\*\*</sup>Dept. of Computer-AI Convergence Engineering, Kumoh National Institute of Technology

### ● 요약 ●

본 논문은 손글씨 인식 분야에서 가장 기본적이고 중요한 주제인 손글씨 데이터 세트에 대한 분류 성능을 분석하고 비교하는 것을 목표로 한다. 이를 위해 ResNet-50 모델을 사용하여 MNIST, EMNIST, KMNIST 라는 세 가지 대표적인 손글씨 데이터 세트에 대한 분류 작업을 수행한다. 각 데이터 세트의 특징과 도메인, 그리고 데이터 세트 간의 차이와 특징에 대해 다루며, ResNet-50 모델을 학습하고 평가한 분류 성능을 비교하고 결과에 대해 분석한 결과를 제시한다.

**키워드:** 손글씨(Handwritten), 분류(Classification), 인식(Recognition), ResNet

## I. Introduction

손글씨에 대한 인식은 컴퓨터 비전과 기계학습 분야에서 가장 기초적인 주제로 간주되고 있다. 손글씨는 개인마다 매우 다양한 형태로 나타내며, 이를 자동으로 인식하는 기술은 문자 인식, 자동 번역 등의 시스템에 활용되고 있다. 딥러닝의 등장 이후로 이러한 손글씨 인식 연구는 높은 성능을 보여주고 있다.

본 연구의 목적은 ResNet-50 모델을 사용하여 손글씨 데이터 세트에 대한 분류 성능을 평가하고 비교하는 것이다. 이를 통해 ResNet-50 모델이 손글씨 분류 문제에 대하여 어느 정도의 효과가 있는지 확인하고 다양한 손글씨 데이터 세트의 특징에 대하여 차이를 이해하고 분석하고자 한다.

## II. Handwritten Datasets

MNIST[1], EMNIST[2], KMNIST[3]은 모두 손글씨 데이터 세트로, 손으로 쓴 숫자 혹은 알파벳 등의 글자 하나에 대한 이미지만을 포함하고 있다. 또한, 제공되는 데이터의 공통점은 28x28의 그레이 스케일 이미지로 구성되어 있다는 것이다.

MNIST[1]은 가장 분류하기 쉽고 많이 알려진 손글씨 숫자 데이터 세트이다. 0부터 9까지의 10개의 클래스로 분류할 수 있으며 컴퓨터 비전 분야를 처음 접할 때 주로 사용하는 데이터 세트이다.

EMNIST[2]는 MNIST의 확장 버전으로 알파벳과 숫자를 포함하여 더 다양한 손글씨 데이터로 이루어져 있다. EMNIST는 세부적인 다양한 카테고리 구성되어 있지만 본 논문에서는 EMNIST-letters에 대해서만 다룬다. EMNIST-letters는 알파벳(A to Z)에 대한 분류를 위한 데이터 세트로, 대소문자를 결합하여 26개의 클래스로 분류한다.

KMNIST[3]은 일본의 고전적 글자인 "Kuzushiji"를 기반으로 한 손글씨 데이터 세트이다. 히라가나의 10행을 나타내는 일본어 문자에 대하여 10개의 클래스로 분류할 수 있다.

그림 1은 MNIST, EMNIST-letters, KMNIST의 데이터에 대한 예시를 나타낸다. 첫 번째는 숫자 데이터 세트인 MNIST, 두 번째는 알파벳 데이터 세트인 EMNIST-letters, 마지막은 히라가나 데이터 세트인 KMNIST를 나타낸다. 각 데이터 세트에 대해 제공되는 학습 데이터와 실험 데이터의 개수에 대한 정보는 표 1과 같다.



Fig. 1. Example of Handwritten Datasets

Table 1. Number of data for the datasets

Datasets	Trainb	Test	Sum
MNIST	60,000	10,000	70,000
EMNIST-letters	88,000	14,800	102,800
KMNIST	60,000	10,000	70,000

### III. Classification using ResNet-50

ResNet-50[4]은 깊이가 증가함에 따라 발생할 수 있는 Gradient Vanishing 문제를 해결하기 위한 구조를 갖고 있다. 50개의 층으로 구성되어 있으며, Residual Block 구조를 사용하여 입출력 사이에 연결을 추가하여 학습 과정에서 잔차에 대한 학습을 진행할 수 있다. 이전 층의 정보가 Skip-Connection을 통해 현재 층의 입력에 더해지기 때문에, Gradient가 깊은 레이어 영역까지 효과적으로 전달된다.

각 데이터 세트는 (28, 28) 크기의 그레이 스케일 이미지로 구성되어 있지만, ResNet-50 모델은 (224, 224, 3) 크기의 입력을 요구한다. 따라서, 모든 데이터 세트에 대해 이미지를 조정하고 재구성하는 과정이 필요하다. 이를 위해 (224, 224) 크기로 변환 후 입력으로 사용하여 분류 작업을 수행할 수 있도록 모델을 학습한다. 또한, Flatten 레이어 이전에서는 각 채널별로 하나의 값을 가지도록 하기 위해 Average Pooling의 Stride를 4로 설정한다.

실험에서는 Ubuntu 18.04 LTS 운영체제 환경에서 GeForce RTX 3090을 사용해 학습과 실험을 진행한다. 학습할 때는 Epoch 수를 50으로 고정하고, Validation Loss가 10번 이상 개선되지 않으면 Early Stopping 기능을 사용하여 학습을 조기 종료하도록 한다. 학습에는 Adam Optimizer를 사용하고, 초기 Learning rate는 1e-4로 설정한다. 데이터 증강 없이 제공되는 원본 데이터를 기반으로 Pretrained 모델이 아닌 처음부터 ResNet-50을 학습하여 특징 추출 및 분류 성능을 비교한다.

표 2는 본 논문에서 제시한 데이터 세트를 바탕으로 정확도를 계산하고 간단한 네트워크와의 성능 차이를 보인다. 간단한 네트워크는 컨볼루션 레이어와 Dense 레이어가 각각 2개로 구성된다. MNIST는 분류 성능이 98.72%으로 간단하게 구성된 네트워크보다 성능이 낮게 나타난다. MNIST는 특징이 간단하게 구성되어 있기 때문에 복잡한 ResNet-50의 경우 오히려 정확도가 떨어지는 결과를 볼 수 있다. EMNIST-letters의 분류 성능은 92.09%로, 간단한 네트워크보다 약 10%의 높은 성능을 보인다. 26가지의 카테고리가 존재하는 일파벳 손글씨 이미지로 구성되어 있으며, 상대적으로 분류가 어려운 문제이기에 ResNet-50을 사용한 결과가 더 높게 나타난다. KMNIST의 분류 성능은 95.20%로 간단한 네트워크보다 다소 높은 성능을 보인다. 이는 일본어 손글씨가 다른 언어의 손글씨에 비해 더 복잡한

형태를 가지기 때문이다.

Table 2. Performance Evaluation using Accuracy

Datasets	Simple Network	ResNet-50
MNIST	<b>99.15%</b>	98.72%
EMNIST-letters	82.18%	<b>92.09%</b>
KMNIST	94.69%	<b>95.20%</b>

### IV. Conclusions

본 논문에서는 ResNet-50 모델을 기반으로 여러 손글씨 데이터 세트의 특징에 따라 분류 성능이 나타남을 보였다. 이러한 결과는 손글씨 인식 기술의 발전과 응용 분야에 많은 기대를 할 수 있다.

### REFERENCES

- [1] Y. LeCun, C. Cortes, and C. Burges, "MNIST handwritten digit database", ATT Labs, Vol. 2, 2010.
- [2] G. Cohen, S. Afshar, J. Tapson, and A. Van Schaik, "EMNIST: Extending MNIST to handwritten letters", Proc. Of international joint conference on neural networks (IJCNN), pp. 2921-2926, 2017.
- [3] T. Clanuwat, M. Bober-Irizar, A. Kitamoto, A. Lamb, K. Yamamoto, and D. Ha, "Deep learning for classical japanese literature". arXiv: 1812.01718.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition", Proc. Of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.