

온라인 한글 필기 인식 알고리즘 개발

정다빈*, 이강은*, 정민진*, 문창진*, 김성석**, 김재현***, 양순옥****

*서경대학교 소프트웨어학과

**서경대학교 소프트웨어학과, 교신저자

***서경대학교 컴퓨터공학과

****가천대학교 교양학부

e-mail:{zxzx1227, gang3039, alswws4390, jin1004boy}@naver.com, {sskim03, statsr}@skuniv.ac.kr, soyang@gachon.ac.kr

Development of Algorithm for Online Handwriting Hangul Recognition

Dabin Jeong*, Kang Eun Lee*, Min Jin Jeong*, Changjin Moon*,
Sungsuk Kim**, Jaehyun Kim***, Sun Ok Yang****

*Dept. of Software Engineering, Seokyeong University

**Dept. of Software Engineering, Seokyeong University, corresponding author

***Dept. of Computer Engineering

****College of Liberal Arts, Gachon University

요 약

본 논문은 기계학습 기반 온라인 한글 필기 인식 시스템의 첫 구현 결과를 담고 있다. 한글의 글자는 최소한 하나의 모음을 포함하고 있으며, 이 모음은 대개 직선으로 필기한다는 사전 지식을 활용하여 인식에 적용하고자 한다. 이를 위해 사용자가 온라인으로 필기하면 획 데이터를 획득하여 중성에 해당하는 모음을 찾는 알고리즘을 개발하였다. 제안한 알고리즘에서는, 우선 필기한 글자를 포함하는 사각형 R 과 각 획을 둘러싸는 사각형 SR 을 생성한 후, 직선을 판별하고, 이 직선들이 모음을 구성하는 후보군을 찾는 과정으로 구성되어 있다. 아직 초기 연구이므로, 다양한 경우에 대한 분석이나 실험 결과는 없지만, 이를 활용하여 온라인 필기 인식 모델에 적용하여 인식 성능을 높이기 위한 추후 연구의 기반으로 활용하고자 한다.

키워드 : 온라인 한글 필기 인식, 기계학습, 전처리, 자음/모음

1. 서론

인공지능 관련 연구가 활발하게 진행되면서 새로운 응용뿐만 아니라 전통적인 분야에서도 이를 활용하여 문제 해결 방법을 찾으려고 하고 있다. 필기 인식 역시 필기자마다 독특하게 필기한 것을 컴퓨터가 식별하도록 하는 주제로서, 인공지능 분야의 오래된 응용 중 하나이다[1]. 필기 인식은 온라인 필기인식과 오프라인 필기 인식으로 구분할 수 있다. 오프라인 필기 인식은 사용자의 필기를 이미지 형식으로 저장하여 이를 식별하는 것이다. 이에 반해 온라인 필기 인식은 사용자가 필기할 때 펜의 좌표 정보나 방향, 접촉 여부 등에 대한 정보를 활용하여 인식하는 것이다. 보다 많은 정보를 활용할 수 있다는 측면에서 온라인 필기 인식 기법이 상대적으로 단순하면서도 더 좋은 인식률을 보일 수 있다.

과거에는 한글 글자의 특성값(feature)을 추출하여 분류하고자 하였지만, 최근에는 딥러닝(Deep Learning)이라는 프레임워크를 활용한 데이터 기반의 해결 방법이 널리 연

구되고 있다. 한글 인식 분야의 경우, 10여년 전까지도 주로 은닉마코프 모델(Hidden Markov Model, HMM)이나 동적 베이저안 네트워크(Bayesian Network Model) 등에 기반한 온라인 필기 인식을 시도하였다[2, 3]. 하지만 영어, 중국어, 아랍어 등에서 컨볼루션 뉴럴 네트워크(Convolutional Neural Network, CNN), 순환 신경망(Recurrent Neural Network, RNN) 등을 이용하여 높은 인식률을 보인 연구결과들이 발표되었다. 이에 따라 최근 한글 필기 연구에서도 CNN을 활용한 오프라인 필기 인식 연구[4, 5]나 RNN을 이용한 연구[1] 결과들이 발표되었다.

CNN 혹은 RNN 기반 연구들은 대개 대용량 데이터를 수집한 후 학습이 진행되므로, 초기에 글자별로 다수의 글자 데이터를 획득하는 것이 중요하다. 인식 연구를 위해 CIFAR[6]나 AI Hub[7]와 같은 사이트에서는 대용량의 다양한 이미지를 제공해주고 있다.

기계학습에서 한글 필기 인식용 학습을 위해서는 각 글자별로 많은 필기 데이터를 획득한 후 모델을 완성해나가야 한다. 이때, 획득한 필기 데이터를 이용하여 학습하면서 모델이 완성되는데, 그 출력은 가능한 한글 글자 중 하나를 선택(classification)하는 문제가 된다. 하지만 이론

* 이 논문은 정부(과학기술정보통신부, 교육부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임 (No. 2019M3E7A1113102).

적으로 모델의 출력에 해당하는 한글의 가능한 글자의 개수는 초성(19)×중성(21)×종성(20)=8,379개이다. 출력의 개수가 많다는 것은 모델의 복잡도 개선이나 성능에서 좋지 않은 결과가 될 수 있다.

이에 본 연구에서는 한글의 특성을 고려한 학습 모델을 제안한다. 즉, 한글은 초성(자음), 중성(모음), (선택적으로) 종성(자음)으로 구성되므로, 인식 단위를 자음/모음으로 한정하도록 한다. 따라서 사용자가 필기하면 이를 초성, 중성, 종성으로 구분한 후 이를 인식하도록 한다. 본 논문에서는 초기 연구 결과로서 온라인 필기 인식을 위한 아이디어를 제안하고 있다.

2. 온라인 한글 필기 인식 알고리즘

온라인 필기는 사용자가 획을 쓸 때마다 관련된 데이터를 수집하여 인식에 활용할 수 있다. 오프라인 인식에 비해 활용할 수 있는 정보가 많으므로 상대적으로 모델이 단순화될 수 있고 인식률도 높다. 본 연구도 한글 인식 모델 구축을 위한 기본 아이디어를 온라인 필기부터 구현하도록 한다.

2.1 학습 데이터 수집

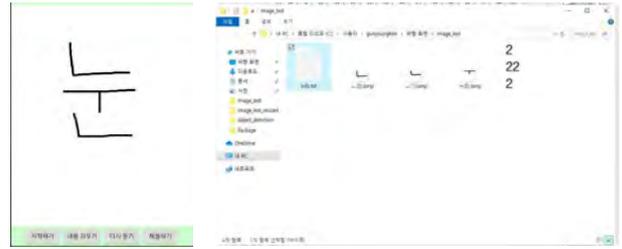
한글은 초성 19개, 중성 모음 21개, 종성 20개를 이용하여 하나의 글자를 구성한다(겹자음/겹모음 포함). 이에 따라 학습 단위도 자음/모음 단위로 수행하려고 한다.

초성 (19)	ㄱ ㄴ ㄷ ㄹ ㄲ ㄴㅇ ㄷㅇ ㄹㅇ ㅈ ㅊ ㅋ ㆁ 교 흥 기 기 ㅂ ㅅ ㅈ
중성 (21)	ㅏ ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ ㅞ ㅟ 에 에 ㅑ ㅓ ㅕ ㅗ ㅛ ㅜ ㅠ ㅡ ㅣ ㅞ ㅟ
종성 (20)	ㄱ ㄴ ㄷ ㄹ ㄲ ㄴㅇ ㄷㅇ ㄹㅇ ㅈ ㅊ ㅋ ㆁ 교 흥 기 기 ㅂ ㅅ ㅈ ㅊ ㅋ ㆁ

기계학습을 위해서는 아주 많은 사용자의 필기 데이터를 수집하는 과정이 필요한데, 이 작업 자체가 매우 부하가 클 수 있다. 이에 따라 사용자의 필기 데이터를 수집할 때부터 초성/중성/종성으로 구분하여 데이터를 수집하고자 한다. 사용자가 필기하기 전에 서버는 사용자용 앱에 필기해야 할 데이터를 미리 넘겨주며, 이때 필기할 글자에 대한 사전 분석 정보도 함께 보내준다. 이 정보에는 각 자음/모음에 대한 (최소, 최대) 획 정보를 포함하고 있으므로, 사용자가 필기할 때마다 구분이 용이하게 한다. 이러한 접근 방식을 활용하면, 글자 하나마다 자음, 모음을 1개 이상씩 수집할 수 있어서 데이터 수집이 용이하다(그림 1 참조). (그림 2)는 이러한 과정을 거쳐 수집한 자음 ‘ㄱ’ 데이터의 일부를 보여주고 있다.

2.2 기계 학습 모델

기계학습 인식을 위한 모델은 일반적으로 CNN을 기반으로 한다. 서버의 한글 초성, 중성, 종성 분리 알고리즘을



(a) 사용자가 ‘눈’을 필기한 상태
(b) 사용자의 필기로부터 분리하여 저장한 자음(초성/중성 ‘ㄴ’ 2개)과 모음(‘ㅛ’) 데이터
(그림 1) 필기 데이터로부터 자/모음 분리 저장



(그림 2) 수집한 필기 데이터의 예 - ‘ㄱ’

기반으로 최종 분리된 28×28 크기(784개 픽셀)의 비트맵 이미지를 입력으로 받아 학습을 진행한다. 사용된 모델은 CNN 신경망을 기반으로 두고 있지만, 해당 모델을 조금 더 변형한 Deep CNN 모델을 바탕으로 구현하였다. 모델의 세부적 구조는 4개의 ‘Conv-Relu-MaxPooling 레이어와 3개의 Fully-Connected 레이어를 바탕으로 최종 40개 (19개 자음+21개 모음)의 클래스를 출력으로 설정한다. 이렇게 분리 및 전처리된 이미지 데이터를 위의 모델에 입력하여 학습시켰다.

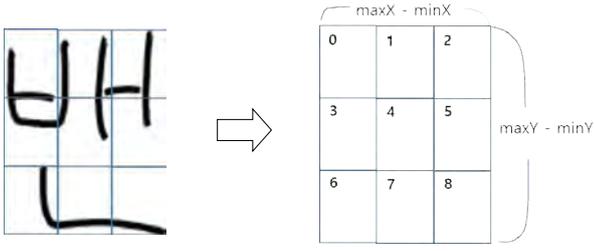
2.3 온라인 필기로부터 초성, 중성, 종성 추출하기

한글의 글자에는 하나의 모음(겹모음 포함)이 포함되어 있으며, 모음은 대개 직선으로 획을 그린다. 따라서 제안한 알고리즘에서는 온라인 한글 필기 인식을 위해서 먼저 모음을 추출하고 나머지 부분에 해당하는 자음을 분리하여 글자를 인식하는 방식을 사용한다. 앞서 언급하였듯이, 온라인 필기는 하나의 글자에 대한 획, 그리고 각 획에 대한 좌표 데이터를 활용하여 인식을 진행하도록 한다1).

이를 위해 우선 두 가지 종류의 사각형 데이터를 생성해야 한다. 먼저 필기한 글자를 포함하는 가장 작은 사각형 R을 생성한 후(그림 3-(a)) 이를 가로, 세로 3등분 한다(그림 3-(b)).

다음으로 각 획별로 획을 둘러싸는 사각형 SR(Stroke Rectangle)을 생성한다. 이는 하나의 획별로 하나의 SR이

1) 본 논문에서는 하나의 글자 단위로 필기하고 인식한다고 가정한다.

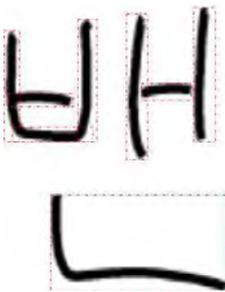


(a) 사용자 필기 ‘밴’을 둘러싼 사각형 R

(b) 사각형 R을 가로 세로 3등분한 후 각 세부 영역 번호 부여

(그림 3) 사용자 필기에 대한 R 생성

생성됨을 의미한다. 예를 들면, ‘밴’의 경우 총 획수 (sNum)가 8이며, 따라서 (그림 4)와 같이 8개의 사각형 SR이 생성됨을 알 수 있다.



(그림 4) 8개의 획 각각을 둘러싼 8개의 사각형 SR

이와 같이 사용자가 필기한 글자에 대하여 하나의 R과 여러 SR이 생성되면 모음을 찾기 위한 알고리즘을 수행한다. 알고리즘에서는 먼저 직선에 해당하는 획을 찾고자 하는데, 획이 직선이라는 것은 각 획을 둘러싼 사각형 SR의 가로, 세로 비율이 일정 비율 이하임을 의미한다. 연구팀에서 수집한 글자 데이터를 이용하여 직선에 해당하는 획들을 추출하여 기울기를 계산하였더니, 평균 0.164 이하이면 직선이라고 판단할 수 있었다. 필기로부터 R과 SR 생성이 완료되면 다음과 같은 모음에 구성하는 획들을 찾기 위한 알고리즘을 수행한다.

```

i = 0 ;
while (i < sNum) { // sNum : 총 획수
    seti = { }
    if sR[i]이 직선인 경우
        seti = seti U { sR[i] } ;
        for (j = i + 1 ; j < sNum ; j++ ) {
            if seti 내의 모든 원소 sR[]과 sR[j]에
            대하여 isOneVowel 수행 {
                seti = seti U { sR[j] }
            } // end of if 절
        } // enf of for-loop
    }

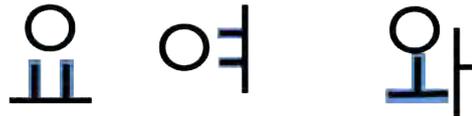
```

```

i++ ;
}

```

알고리즘에서 isOneVowel 함수는 SR이 직선으로 판정되면 다른 직선인 SR과 비교를 수행하여 하나의 모음을 구성하는 직선인지를 판정한다. 보다 구체적으로 두 SR s1과 s2가 주어졌을 때, 두 직선이 모두 가로(혹은 세로 직선이라면 (그림 5)의 첫 두 경우를 판별하고자 하며, 두 직선이 가로와 세로일 경우에는 (그림 5)의 세 번째 경우를 판별하고자 한다.

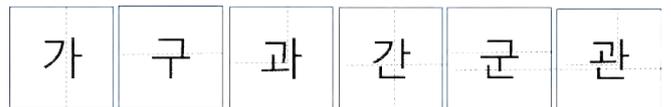


(그림 5) 두 직선이 같은 모음에 해당되는 경우

이때 (그림 3-(b))에서 보여준 영역은 직선이라고 판단된 획이 모음이 아닌 경우를 배제하기 위해 사용된다. 즉, 어떤 직선은 자음을 구성하는 직선일 수도 있으므로, 다음과 같은 4가지 경우를 고려하여 모음에 해당하는 후보를 추출하고자 한다.

- ① R의 0과 만나는 사각형은 초성이다(0번과 만나지 않았다면 1, 3번과 비교)
- ② R의 6, 7, 8과 만나는 직선 아닌 사각형이 존재한다면 중성이 있다고 판정한다
- ③ R의 6, 7, 8과 만나는 직선이 있다면(모음이 될 확률이 있다면), 이 직선은 1개 혹은 2개 혹은 예외적으로 3개(‘ㄱ’)의 획으로만 구성된다
- ④ R의 2와 만나는 직선이 있다면 모음일 확률이 높다 (예외 ‘ㅃ’으로 이루어진 ‘ㅍ’. 단 ‘ㅈ’, ‘ㄷ’으로 만들어진 ‘보’, ‘포’ 같은 글자는 (경우 ①)에 의해 걸러질 수 있음)

한글에서 모음이 위치할 수 있는 곳은 (그림 6)과 같이 6가지이며, R을 가로, 세로 3등분한 것은 이를 고려하여 모음에 해당하는 직선을 판별하기 위함이다.

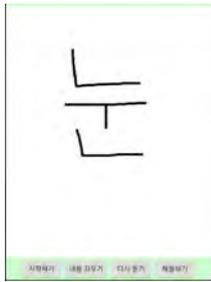


(그림 6) 모음의 위치에 따른 한글의 글자 패턴 6가지

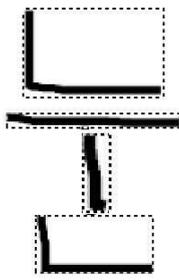
이와 같은 알고리즘을 수행하여 모음을 구성하는 획들로 구성된 후보 이미지를 하나 이상 찾아낸 다음, 이를 28x28 이미지로 작성하여 학습된 모델에 적용하면 일정 확률을 가진 자음/모음을 출력한다. 그 출력값이 모음에 해당된다면, 그 획들을 배제한 나머지 부분이 초성이나 중

성에 해당하는 자음으로 결정하여 모델에 적용하고자 한다.

(그림 7)은 개발한 시스템에 의한 결과 화면을 보여주고 있다. (a)와 같이 앱에서 사용자가 '눈'을 필기하면, R과 각 획별로 SR을 생성한 후, 하나의 모음 후보 이미지를 추출하여 모델에 적용하여 'ㅏ'를 식별한다. 나머지 획들로부터 두 개의 이미지를 추출하여 'ㄴ'을 식별하여 최종적으로 (d)와 같은 식별 결과를 보여준다.



(a) 사용자의 필기



(b) 필기로부터 생성된 SR



(c) 획 데이터 기반으로 추출된 자음/모음



(d) 인식 결과 화면

(그림 7) 제안 시스템으로 식별을 수행한 결과 화면

추출하여 활용하려고 한다. 비스듬한 필기는 직선 구별 알고리즘을 변경하면 처리가 가능할 것으로 보인다. 마지막으로 흘려쓰기는 초성/중성/종성 내에서의 흘려쓰기와, 초성-중성 혹은 중성-종성 흘려쓰기로 구분하여 처리하고자 한다. 이러한 개념을 올바르게 구현하여 추후 인식률을 높은 온라인 필기 인식 시스템 개발을 완료하고자 한다.

참고문헌

[1] 김병희, 장병탁, "순환신경망을 이용한 한글 필기체 인식," 정보과학회 컴퓨팅의 실제 논문지, 제25권, 5호, pp. 316-321, 2017

[2] B.-K. Sin and J. Kim, "Ligature Modeling for Online Cursive Script Recognition," IEEE Trans. Pattern Anal. Mach. Intell, Vol. 19, No. 6, pp. 623-633, 1997.

[3] S. Cho and J. Kim, Bayesian Network Modeling of Hangeul Characters for On-line Handwriting Recognition, Proc. ICDAR, pp. 207-211, 2003.

[4] I. J. Kim and X. H. Xie, "Handwritten Hangeul Recognition Using Deep Convolutional Neural Networks," Int. J. Doc. Anal. Recognit., Vol. 18, No. 1, pp. 1-13, 2015.

[5] I. J. Kim, C. B. Choi, and S. H. Lee, "Improving Discrimination Ability of Convolutional Neural Networks by Hybrid Learning," Int. J. Doc. Anal. Recognit., Vol. 19, No. 1, pp. 1-9, 2016.

[6] <https://www.cs.toronto.edu/~kriz/cifar.html>

[7] <https://aihub.or.kr/>

3. 결론 및 향후 연구과제

본 논문에서는 현재 개발이 시작된 필기 인식의 초기 연구 결과인 온라인 필기 인식 시스템의 동작 원리를 설명하고 있다. 온라인 필기 시스템으로부터 획득할 수 있는 획과 관련된 좌표 정보를 이용하여 사각형 R과 각 획을 둘러싸는 사각형 SR을 생성한 후 직선에 해당하는 획을 판별하고, 직선 중에서 하나의 모음이 될 수 있는 경우를 찾아서 모음 후보로 만드는 것이 목적이다.

하지만 개발 초기이므로 인식률이 높지는 않으며, 자음과 모음을 구성하는 직선을 구분하는 것이 완벽하지 않다. 또한 사용자가 필기를 비스듬하게 하거나 흘려쓰기를 하는 경우도 아직 인식 알고리즘에서 올바르게 반영하지 못하였다.

이를 위해 우선 학습 데이터를 충분히 확보하고자 한다. 최근의 상황으로 인하여 여러 사용자의 다양한 필기 데이터를 획득하는 것이 곤란하므로 다른 접근방법을 취하고자 한다. 즉, 여러 사이트에서 이미지 형식으로 배포해주는 한글 필기 데이터로부터 자음/모음만 수동적으로