

<http://dx.doi.org/10.17703/JCCT.2024.10.2.453>

JCCT 2024-3-52

딥러닝 기반의 의료 OCR 기술 동향

Trends in Deep Learning-based Medical Optical Character Recognition

윤성연*, 최아린*, 김채원**, 오수민**, 손서영**, 김지연***, 이현희****, 한명은****, 박민서*****

Sungyeon Yoon*, Arin Choi*, Chaewon Kim**, Sumin Oh**, Seoyoung Sohn**,
Jiyeon Kim***, Hyunhee Lee****, Myeongeun Han*****, Minseo Park*****

요약 광학 문자 인식(Optical Character Recognition, OCR)은 이미지 내의 문자를 인식하여 디지털 포맷(Digital Format)의 텍스트로 변환하는 기술이다. 딥러닝(Deep Learning) 기반의 OCR이 높은 인식률을 보여줌에 따라 대량의 기록 자료를 보유한 많은 산업 분야에서 OCR을 활용하고 있다. 특히, 의료 산업 분야는 의료 서비스 향상을 위해 딥러닝 기반의 OCR을 적극 도입하였다. 본 논문에서는 딥러닝 기반 OCR 엔진(Engine) 및 의료 데이터에 특화된 OCR의 동향을 살펴보고, 의료 OCR의 발전 방향에 대해 제시한다. 현재의 의료 OCR은 검출한 문자 데이터를 자연어 처리(Natural Language Processing, NLP)하여 인식률을 개선하였다. 그러나, 정형화되지 않은 손글씨(Handwriting)나 변형된 문자에서는 여전히 인식 정확도에 한계를 보였다. 의료 데이터의 데이터베이스(Database)화, 이미지 전처리(Pre-processing), 특화된 자연어 처리를 통해 더욱 고도화된 의료 OCR을 발전시키는 것이 필요하다.

주요어 : 광학 문자 인식, 딥러닝, 의료 인공지능

Abstract Optical Character Recognition is the technology that recognizes text in images and converts them into digital format. Deep learning-based OCR is being used in many industries with large quantities of recorded data due to its high recognition performance. To improve medical services, deep learning-based OCR was actively introduced by the medical industry. In this paper, we discussed trends in OCR engines and medical OCR and provided a roadmap for development of medical OCR. By using natural language processing on detected text data, current medical OCR has improved its recognition performance. However, there are limits to the recognition performance, especially for non-standard handwriting and modified text. To develop advanced medical OCR, databaseization of medical data, image pre-processing, and natural language processing are necessary.

Key words : Optical Character Recognition, Deep Learning, Medical AI

1. 서론

광학 문자 인식(Optical Character Recognition, OCR)은 인쇄, 촬영, 스캔 이미지 내의 문자를 인식하여 기계가 읽고 편집할 수 있는 디지털 포맷(Digital

Format)의 텍스트로 변환하는 기술이다[1, 2]. OCR은 반복적인 문서 관련 업무 프로세스를 자동화할 수 있다는 장점을 제공한다. 최근 의료, 금융, 법률 등 대량의 기록 자료를 보유한 산업 분야에서 자료의 디지털화를 위해 OCR을 도입하고 있다[3]. 또한, 딥러닝(Deep

*준회원, 서울여자대학교 데이터사이언스학과 대학원생

**준회원, 서울여자대학교 데이터사이언스학과 학부생

***준회원, 서울여자대학교 디지털미디어학과 학부생

****준회원, 서울여자대학교 데이터과학전공 학부생

*****정회원, 서울여자대학교 데이터사이언스학과 조교수 (교신저자)

접수일: 2024년 1월 2일, 수정완료일: 2024년 1월 21일

게재확정일: 2024년 2월 1일

Received: January 2, 2024 / Revised: January 21, 2024

Accepted: February 1, 2024

*****Corresponding Author: mpark@swu.ac.kr

Dept. of Data Science, Seoul Women's Univ, Korea

Learning)을 활용한 OCR의 문자 인식률과 인식 속도의 향상으로 그 적용 분야가 점차 확장되고 있다[4].

특히, 코로나 19 이후 의료 산업 분야는 의료 서비스 향상을 위해 딥러닝 기반의 OCR을 적극 도입하는 모습을 보여주고 있다[5]. 2000년대 초반부터 전자의무기록(Electronic Medical Record, EMR)으로 병원정보시스템이 개선되었으나, 여전히 건강 상태, 의료 기록, 검사결과 등의 다양한 종이 기록이 의료 산업 분야에서 발생하고 있다. 의료 산업 분야에서의 OCR은 다양한 형태로 존재하는 환자 관련 기록의 빠른 검색과 효율적인 관리를 위해 사용되고 있으며, 의료진이 더욱 정확하고 신속한 의료 서비스를 제공할 수 있게 도와주는 형태로 고도화되고 있다.

환자의 상태를 진단하고 치료 계획을 세우는 데에 있어 핵심적인 자료로 사용되는 의료 데이터를 효과적으로 관리하고 활용하기 위해서는 높은 정확도를 가지는 의료 특화 OCR이 필수적이다. 그러나 의료 데이터는 의학 전문 용어를 포함하기 때문에, 의료계 종사자가 아닐 경우 이를 해석하는 데에 한계가 존재한다. 이를 해결하기 위해 딥러닝 기반의 오픈소스(Open Source) OCR 엔진(Engine)에 자연어 처리(Natural Language Processing, NLP)를 접목한 의료 OCR이 제안되고 있다[6-13].

본 논문에서는 딥러닝 기반 OCR 엔진 및 서비스(Service)와 의료 데이터에 특화된 OCR 기술 동향에 대해 살펴보고, 향후 의료 OCR 기술의 발전 방향을 제시하고자 한다.

II. OCR 엔진 및 서비스

딥러닝 기반 오픈소스 OCR 엔진으로는 Tesseract OCR[14], Easy OCR[15], docTR OCR[16], Keras OCR[17]이, 기업에서 제공하는 딥러닝 기반 OCR 서비스로는 Naver Clova OCR[18], Google Cloud Vision API[19]가 대표적으로 활용된다. 본 장에서는 딥러닝 기반 OCR 엔진 및 서비스의 특성에 대해 살펴보고자 한다.

1. Tesseract OCR

Tesseract OCR은 구글(Google)에서 개발된 C++ 기반 오픈소스(Open Source) 엔진이다. 100개 이상의 언

어를 지원하기 때문에 다양한 언어로 된 데이터에 활용이 가능하고, 인쇄체(Print)와 손글씨(Handwriting) 모두 정확하게 인식한다는 장점이 있다. Adaptive Thresholding 기술을 활용하여 밝기값이 다른 이미지에서도 효과적으로 문자 영역과 배경 영역을 구분한다. Tesseract 4 버전부터는 LSTM(Long Short-Term Memory)[20]을 활용하여 단어 별 인식 기술을 제공하고 있다. 또한, 모듈(Module)화 된 구조를 가지고 있어 필요에 따라 확장하거나 사용자 정의가 가능하다. 그러나, 데이터의 품질에 민감하게 반응하고, 다중 언어 인식 기술을 제공하지 않기 때문에 데이터에 사용된 언어 별로 각각 인식해야 한다는 한계가 존재한다. GPU(그래픽처리장치, Graphic Processing Unit)를 사용하지 않기 때문에 다른 OCR 엔진에 비해 처리 속도 또한 느리다.

2. Easy OCR

Easy OCR은 2020년에 공개된 파이토치(PyTorch) 기반의 오픈소스 엔진이다. 80개 이상의 언어의 다중 언어 인식 기술을 지원한다. 라이브러리(Library) 형태로 제공하기 때문에 쉽게 사용이 가능하다. 문자 영역 인식(Detection) 기술과 문자 인식(Recognition) 기술을 각각 혹은 동시에 활용할 수 있으며, 필요에 따라 수정이 가능하다. 문자 영역 인식에는 네이버(Naver)에서 개발한 CRAFT(Character Region Awareness for Text Detection)[21]를 활용하여 다각형(Polygon) 형태의 문자 영역을 검출한다. 문자 인식에는 ResNet(Residual Neural Network), LSTM, CTC(Connectionist Temporal Classification)[22] 기술을 결합하여 문자의 의미를 고려해 문자를 인식한다. 또한 GPU를 사용하도록 설정할 수 있어 빠르게 인식이 가능하다. 그러나, 사전 학습(Pre-trained)된 데이터 셋의 개수의 차이로 인해 언어 별 성능의 격차가 크며, 필기체나 손글씨에 대한 인식률이 낮다.

3. docTR OCR

docTR OCR은 민디(Mindee)에서 공개한 파이썬(Python) 기반의 오픈소스 엔진이다. 텐서플로우(Tensorflow)와 파이토치 프레임워크 버전을 제공하고 있다. 문자 영역 인식에는 LinkNet[23]과 DBNet[24]을 활용하고, 문자 인식에는 CRNN(Convolutional to

Recurrent Neural Network)[25], SAR(Show, Attend and Read)[26], MASTER(Multi-Aspect Non-local Network for Scene Text Recognition)[27] 기술을 결합하여 다양한 위치 및 형태의 문자를 인식한다. 그러나 영어와 숫자만 인식할 수 있다는 한계가 존재한다.

4. Keras OCR

Keras OCR은 2019년 공개된 케라스(Keras) 기반의 오픈소스 엔진이다. 문자 영역 인식에는 네이버의 CRAFT를 활용하고, 문자 인식에는 CRNN을 활용하여 다양한 형태의 문자를 검출해낸다. 그러나 영어와 숫자만 인식할 수 있다는 한계가 존재한다.

5. Naver Clova OCR

Naver Clova OCR은 네이버에서 개발한 OCR 서비스이다. 한국어, 영어, 일본어, 중국어를 지원하며 특히 한국어와 영어에 높은 인식 정확도를 보인다. 특정 양식을 갖춘 문서 혹은 이미지에서 문자를 인식할 때 해당 양식에 맞는 템플릿을 사전에 만들어두는 방식을 통해 데이터의 구조를 파악하여 필요한 영역의 정보만을 추출하고 가공하는 데에 특화되어 있다. 클라우드(Cloud) 기반으로 서비스가 제공되기 때문에 접근성이 뛰어나며, API(Application Programming Interface) 형태로 다양한 애플리케이션(Application)이나 플랫폼(Platform)에 연동이 가능하다는 장점이 있다. 그러나, 제공되는 언어가 한정적이고, 언어 별 인식률의 차이가 크다.

6. Google Cloud Vision API

구글의 Google Cloud Vision API는 이미지 및 비디오 데이터에서 텍스트 정보를 추출하고 해석하는 기능을 제공하는 클라우드 기반 서비스이다. Google Cloud Vision API는 35개 언어의 다중 언어 인식 기술을 제공한다. AutoML(Auto Machine Learning)을 활용하여 이미지를 분류하는 머신러닝(Machine Learning) 모델을 활용하기 때문에 인쇄체와 손글씨 모두 높은 인식 정확도를 보일 뿐만 아니라, 표, 서명, 로고 등과 같은 특정한 양식 또한 명확하게 인식하고 추출할 수 있는 장점이 있다. 선행 학습된 API를 통해 객체 인식, 필기체 인식 등이 가능하며, Google Cloud Platform과 통합되어 클라우드에서 제공되는 다른 서비스와 연동하여 사

용할 수 있다. 그러나 사전에 훈련된 모델을 기반으로 OCR 서비스가 제공되기 때문에 특정한 도메인에 대한 최적화가 어렵다는 한계가 있다.

III. 의료 OCR 기술

박정민 외[6]는 Google Cloud Vision API를 이용해 문서에 기재된 의약품명을 추출하고, 추출한 의약품명을 기준으로 사용자에게 의약품 제조 업체명, 고유코드, 효능, 복용 방법, 사용 전 주의사항, 사용 중 주의사항, 이상 반응, 보관법, 사진 등 9개의 정보를 안내하는 시스템을 제안하였다. 실험에 사용한 100개의 데이터 셋에서, 제안 시스템은 99.50%의 높은 정확도를 보였다.

Hsu 외[7]는 Tesseract OCR을 이용해 수면 패턴을 기록한 문서에서 무호흡증(Apnea hypopnea index, AHI)과 관련된 내용을 인식하고, 의료 용어로 학습한 BERT(Bidirectional Encoder Representations from Transformers)[28] 기반 모델인 ClinicalBERT[29]를 활용하여 무호흡증 환자의 데이터를 분류하는 방법을 제안하였다. 실험에 사용한 2015년 6월 1일부터 2018년 5월 31일까지의 무호흡증 환자의 995개 수면 패턴 보고서에서, 제안 방법은 94.76%의 정확도를 보였다.

Abdulaziz 등[8]은 Tesseract OCR을 이용해 University of California, Irvine(UCI)에서 수집한 심장(Heart) 또는 신장(Kidney) 질환을 가진 환자의 건강검진 데이터를 관리하는 연구를 진행하였다. 또한, Bag of Words(BoW)[30]와 AdaBoost(Adaptive Boost)[31]를 통해 환자 데이터를 관리하는 방법을 제안하였다. 실험에 사용한 879명의 환자 데이터 셋에서, 평균 93.30%의 정확도를 보여주었다.

Silva 등[9]은 Tesseract OCR로 당뇨병(Diabetes Mellitus) 환자 데이터를 디지털화하고, 스페이스(spacy) 파이썬(Python) 라이브러리를 활용해 주요 단말 토큰화[32, 33]하는 방법을 제안하였다. 또한 Random Forest[34]를 활용하여 당뇨병 진행 정도에 따라 환자 데이터를 분류하였다. 실험에 사용한 768개 데이터 셋에서, 사이킷런(Sci-kit learn) 오픈소스를 활용해 구현한 제안 방법은 약 90%의 정확도를 보였다.

이남화 등[10]은 Easy OCR을 활용하여 혈액학적 모니터링 장치(Hemodynamic Monitoring Device)의 혈압(BP, Blood Pressure), SVV(Stroke Volume Variation),

LCWI(Left Cardiac Work Index), SI(Stroke Index), SVRI(Systemic Vascular Resistance Index) 등 5개의 수치 정보를 추출하고, 각 수치의 변동에 따라 주입할 약물의 종류와 양을 결정하는 자동화 시스템을 제안하였다. 제안한 시스템이 구현된 환경에서, EasyOCR이 혈액학적 모니터링 장치의 정보를 인식하고, 이를 기준으로 각 환자 별 약물 주입량을 결정하는 데에 약 1.15 초가 소요되었다.

Batra 등[11]은 Tesseract OCR, Easy OCR, docTR OCR을 활용하여 PNG 형태의 의료 문서를 디지털화하고, 이미지 이진화(Image Binarization), 밝기 변환(Brightness Transformation), 감마 보정(Gamma Correction), 시그모이드 변환(Sigmoid Stretching), 이미지 평활화(Image Smoothing), 이미지 선명화(Image Sharpening) 등 6가지 전처리(Pre-processing) 방법을 결합하여 인식 성능을 개선하는 방법을 제안하였다. 실험 결과, 이미지 평활화 전처리를 한 의료 문서를 Tesseract OCR로 디지털화할 경우 92.46%의 가장 높은 정확도를 보임을 증명하였다.

Tangkawanit 등[12]은 Tesseract OCR, Easy OCR, Keras OCR을 활용하여 혈액 검사 결과지의 14가지 수치를 디지털화하고, 해당 수치를 기반으로 암을 진단하는 시스템을 제안하였다. 실험 결과, Tesseract OCR을 활용할 경우 81.50%의 가장 높은 정확도를 보임을 증명하였다.

Drukker 등[13]은 Easy OCR을 활용하여 Slow-flow HD모드 초음파 영상에서 태아의 골 열지수(TIb, Bone Thermal Index) 정보를 수집하는 시스템을 제안하였다. 4주~10주 임신부의 185개 초음파 영상을 활용한 실험 결과, 세계 초음파 의학 및 생물학 연맹(WFUMB, World Federation for Ultrasound in Medicine and Biology)의 제한기준 0.7보다 크게 감소된 수치인 약 0.32의 골 열지수를 보였다.

IV. 결 론

본 논문에서는 딥러닝(Deep Learning) 기반 광학 문자 인식(Optical Character Recognition, OCR) 엔진(Engine)과 의료 데이터에 특화된 OCR 기술의 동향을 살펴보았다. 의료 데이터는 환자의 상태를 진단하고 치료 계획을 세우는 데에 있어 핵심적인 자료로 사용된

다. 의료 데이터를 기반으로 학습된 의료 특화 OCR 기술을 활용하는 경우, 의료 데이터의 보안 및 관리 수준을 크게 향상시킬 수 있다. 특히, 다양한 종이 기록의 수작업 처리 한계를 극복하고 페이퍼리스(Paperless) 의료 시스템으로의 전환을 가속화할 수 있다는 장점이 있다.

최근 높은 정확도가 필수적인 의료 데이터의 특성을 고려하기 위해 OCR로 검출한 문자 데이터를 자연어 처리(Natural Language Processing, NLP)하는 방법이 제안되고 있음을 확인하였다. 그러나, 현재 OCR 기술은 정형화되지 않은 손글씨(Handwriting)나 초점, 화질, 그림자 등 외적 요인에 의해 변형된 문자에서는 여전히 한계를 보인다.

이러한 한계를 극복하기 위해 다양한 의료 데이터의 데이터베이스(Database, DB)화, 정교한 이미지 전처리(Pre-processing), 의료 데이터에 특화된 자연어 처리가 제공되는 의료 OCR 기술이 필요하다. 의료 용어 데이터베이스의 구축을 통해 의료 분야의 특수 용어를 체계적으로 관리하는 데에 도움을 줘 의료 데이터의 효율적인 활용에 기여할 수 있을 것으로 기대한다. 또한, 높은 성능의 의료 특화 OCR가 실제 의료 산업 분야에 상용화되는 경우 의료 업무 자동화와 효율성의 증대로 인한 업무 프로세스의 혁신과 의료 비용 절감, 환자들의 의료 서비스에 대한 접근성 향상에 도움이 될 것으로 기대한다.

References

- [1] J. Memon, M. Sami, R. A. Khan, and M. Uddin, "Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Literature Review (SLR)," *IEEE Access*, Vol. 8, pp. 142642-142668, 2020. DOI:10.1109/ACCESS.2020.3012542
- [2] Z. Raisi, M. A. Naiel, P. Fieguth, S. Wardell, and J. Zelek, "Text detection and recognition in the wild: A review," *arXiv preprint arXiv:2006.04305*, 30 Jun 2020. DOI: 10.48550/arXiv.2006.04305
- [3] S. Ahn, H. Hwang, and J. Hee, "A Case Study on the Application of AI-OCR for Data Transformation of Paper Records," *Journal of Information Management Society*, Vol. 39, No. 3, pp. 165-193, 2022. DOI:10.3743/KOSIM.2022.3.165

- [4] G. Min, A. Lee, K. S. Kim, J. E. Kim, H. S. Kang, and G. H. Lee, "Recent Trends in Deep Learning-Based Optical Character Recognition," *Electronics and Telecommunications Trends*, Vol. 37, No. 5, pp. 22–32, Oct 2022. DOI: 10.22648/ETRI.2022.J.370503
- [5] D. Gifu, "AI-backed OCR in Healthcare," *Procedia Computer Science*, Vol. 207, No. 2, pp. 1134–1143, Oct 2022. DOI: 10.1016/j.procs.2022.09.169
- [6] J. M. Park, S. K. Choi, J. Y. Kim, S. H. Jung, and C. B. Sim, "Implementation of a Drug Information Retrieval System Through OCR API pErformance Comparison," *The Journal of The Korea Institute of Electronic Communication Sciences (KIECS)*, Vol. 18, No. 5, pp. 989–998, 31 Oct 2023. DOI: 10.13067/JKIECS.2023.18.5.989
- [7] E. Hsu, I. Malagaris, Y. F. Kuo, R. Sultana, and K. Roberts, "Deep learning-based NLP data pipeline for EHR-scanned document information extraction," *JAMIA open*, Vol. 5, No. 2, pp. 1–12, 2022. DOI: 10.1093/jamiaopen/ooac045
- [8] W. A. Qader and M. M. Ameen, "Diagnosis of Diseases from Medical Check-up Test Reports Using OCR Technology with BoW and AdaBoost algorithms," *In 2019 International Engineering Conference (IEC)*, pp. 205–210, 23–25 June 2019. DOI: 10.1109/IEC47844.2019.8950605
- [9] W. A. J. R. Silva, H. M. K. Shirantha, L. J. M. V. N. Balalla, R. A. D. V. K., N. Kuruwitaarachchi, and D. Kasthurirathna, "Predicting Diabetes Mellitus Using Machine Learning and Optical Character Recognition," *In 2021 6th International Conference for Convergence in Technology (I2CT)*, pp. 1–6, 02–04 Apr 2021. DOI: 10.1109/I2CT51068.2021.9417941
- [10] N. Lee, M. Jeong, Y. Kim, J. Shin, I. Joe, S. Jeon, and B. Ko, "IoT-based Architecture and Implementation for Automatic Shock Treatment," *KSII Transactions on Internet and Information Systems*, Vol. 16, No. 7, pp. 2209–2224, June 2022. DOI: 10.3837/tiis.2022.07.005
- [11] P. Batra, N. Pulkit, D. Kurmi, J. Temburne, P. Sahare, and T. Diwan, "OCR-MRD: performance analysis of different optical character recognition engines for medical report digitization," *International Journal of Information Technology*, Vol. 16, No.1, pp. 447–455, 24 November 2023. DOI:10.21203/rs.3.rs-2513255/v1
- [12] S. Tangkawanit, J. Pooksook, J. Ieamsaard, and P. Sornkhom, "OCR Application for Cancer Care," *2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, 7–10 November 2022. DOI:10.23919/APSIPAASC55919.2022.9980078
- [13] L. Drukker, R. Droste, C. Ioannou, L. Impey, J. A. Noble, and A. T. Papageorgiou, "Function and safety of SlowFlowHD ultrasound Doppler in obstetrics," *Ultrasound in Medicine & Biology*, Vol. 48, No. 6, pp. 1157–1162. DOI: 10.1016/j.ultrasmedbio.2022.02.012
- [14] R. Smith, "An overview of the Tesseract OCR engine," *In Proceedings of the Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*, Vol. 2, pp. 629–633, 23 - 26 Sep 2007.
- [15] EasyOCR, JaidedAI, 2023. Available online: <https://github.com/JaidedAI/EasyOCR> (accessed on 20 February 2024)
- [16] H. Feng, Y. Wang, W. Zhou, J. Deng, and H. Li, "Doctr: Document image transformer for geometric unwarping and illumination correction," *Proceedings of the 29th ACM International Conference on Multimedia*, pp. 273–281, October 2021. DOI:10.48550/arXiv.2110.12942
- [17] Keras-ocr, 2019. Available online: <https://keras-ocr.readthedocs.io/en/latest/> (accessed on 20 February 2024)
- [18] Naver CLOVA OCR, 2023. Available online: <https://clova.ai/ocr/?lang=ko> (accessed on 20 February 2024)
- [19] Cloud Vision API, Detect text in images, 2023. Available online: <https://cloud.google.com/vision/docs/ocr?hl=ko> (accessed on 20 February 2024)
- [20] Y. Wang, M. Huang, L. Zhao, and X. Zhu, "Attention-based LSTM for aspect-level sentiment classification," *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pp. 606–315, January 2016. DOI:10.18653/v1/D16-1058
- [21] Y. Baek, B. Lee, D. Han, S. Yun, and H Lee, "Character region awareness for text detection," *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 9365–9374, DOI: 10.48550/arXiv.1904.01941
- [22] A. Graves, S. Fernandez, F. Gomez, and J. Schmidhuber, "Connectionist temporal classification: labelling unsegmented sequence data with recurrent neural networks,"

- Proceedings of the 23rd international conference on Machine learning*, pp. 369–376, January 2006. DOI:10.1145/1143844.1143891
- [23] A. Chaurasia and E. Culurciello, “Linknet: Exploiting encoder representations for efficient semantic segmentation,” *In 2017 IEEE visual communications and image processing (VCIP)*, pp. 1–4, 10–13 December 2017. DOI: 10.1109/VCIP.2017.8305148
- [24] M. Liao, Z. Wan, C. Yao, K. Chen, and X. Bai, “Real-time scene text detection with differentiable binarization,” *In Proceedings of the AAAI conference on artificial intelligence*, Vol. 34, No. 7, pp. 11474–11481, April 2020. DOI: 10.1609/aaai.v34i07.6812
- [25] B. Shi, X. Bai, and C. Yao, “An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition,” *IEEE transactions on pattern analysis and machine intelligence*, Vol. 39, No. 11, pp. 2298–2304, November 2017. DOI: 10.1109/TPAMI.2016.2646371
- [26] H. Li, P. Wang, C. Shen, and G. Zhang, “Show, attend and read: A simple and strong baseline for irregular text recognition,” *In Proceedings of the AAAI conference on artificial intelligence*, Vol. 33, No. 01, pp. 8610–8617, July 2019. DOI: 10.1609/aaai.v33i01.33018610
- [27] L. Ning, Y. Wenwen, Q. Xianbiao, C. Yihao, G. Ping, X. Rong, and B. Xiang, “Master: Multi-aspect non-local network for scene text recognition,” *Pattern Recognition*, 15 April 2021. DOI:10.1016/j.patcog.2021.107980
- [28] J. Devlin, M. W. Chang, K. Lee, K. Toutanova, “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” *arXiv preprint arXiv:1810.04805*, 2018. DOI: 10.48550/arXiv.1810.04805
- [29] E. Alsentzer, J. R. Murphy, W. Boag, W. H. Weng, D. Jin, T. Naumann, and M. B. A. McDermott, “Publicly available clinical BERT embeddings,” *In Proceedings of the 2nd Clinical Natural Language Processing (ClinicalNLP) Workshop in North American Chapter of the Association for Computational Linguistics (NAACL)*, 2019. DOI: 10.48550/arXiv.1904.03323
- [30] J. Polpinij and A. K. Ghose, “An ontology-based sentiment classification methodology for online consumer reviews,” *In 2008 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology*, Vol. 1, pp. 518–524, 09–12 Dec 2008. DOI: 10.1109/WIIAT.2008.68
- [31] P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” *In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 1, 8–14 December 2001. DOI: 10.1109/CVPR.2001.990517
- [32] S. Yoon and M. Park, “Media-based Analysis of Gasoline Inventory with Korean Text Summarization,” *The Journal of the Convergence on Culture Technology (JCCT)*, Vol. 9, No. 5, pp. 509–515, Oct 2023. DOI: 10.17703/JCCT.2023.9.5.509
- [33] H. S. Lee, “Research of Late Adolescent Activity based on Using Big Data Analysis,” *The International Journal of Advanced Culture Technology (IJACT)*, Vol. 10, No. 4, pp. 361–368, Dec 2022. DOI: 10.17703/IJACT.2022.10.4.361
- [34] M. R. Segal, “Machine learning benchmarks and random forest regression,” *Center for Bioinformatics and Molecular Biostatistics*, 2004.