

Multi Model Recognition and Mining Alphabet Identification using NLP

Mayur Sakhare¹, Rushikesh Bhalerao², Vishal Sonawane³, Nikita Karpe⁴, Prof. Niranjana Bhale⁵

Students, Department of Information Technology^{1,2,3,4}

HOD, Department of Information Technology⁵

Matoshri College of Engineering and Research Center, Nashik, Maharashtra, India

Abstract: Optical character recognition (OCR), usually abbreviated to OCR, is the mechanical or electronic conversion of scanned or photographed images of typewritten or printed text into machine-encoded or computer-readable text. It is widely used as a data entry method for various original paper data sources, including passport documents, invoices, bank statements, receipts, business cards, mail, and other printed records. OCR serves as a common technique for digitizing printed texts, enabling electronic editing, efficient storage, online display, and utilization in machine processes such as machine translation, text-to-speech conversion, key data extraction, and text mining. OCR is a field of research encompassing pattern recognition, artificial intelligence, and computer vision. Optical Character Recognition or OCR is also used for the electronic translation of handwritten, typewritten, or printed text into machine-translated images. It finds widespread application in recognizing and searching text from electronic documents or publishing text on websites. In our proposed methodology, we developed our system on a Windows 11 PC, utilizing PYTHON as the frontend software

Keywords: OCR, python, AI, Image processing, NLP, Photo, image, character

REFERENCES

- [1]. Chaudhuri, K. Mandaviya, P. Badelia, and S. K. Ghosh, Optical character recognition systems, vol. 352, 2017. doi: 10.1007/978-3-319-50252-6_2.
- [2]. P. Divya et al., "Web based optical character recognition application using flask and tesseract," Mater. Today Proc., 2021, doi: 10.1016/j.matpr.2020.10.850.
- [3]. A. T. Sahlol, C. Y. Suen, H. M. Zawbaa, A. E. Hassanien, and M. A. Elfattah, "Bio-inspired BAT optimization algorithm for handwritten Arabic characters recognition," 2016 IEEE Congr. Evol. Comput. CEC 2016, pp. 1749–1756, 2016, doi: 10.1109/CEC.2016.7744000.
- [4]. A. Dobroczeni, R. Takacs, B. M. Cermak, and Shchokin, "Design Of Machines And Structures," vol. 18, no. 1, pp. 2–4, 2014.
- [5]. Bartz, C., Yang, H. and Meinel, C., "STN-OCR: A single neural network for text detection and text recognition". 2017. arXiv preprint arXiv:1707.08831.
- [6]. H. Singh and A. Sachan, "A Proposed Approach for Character Recognition Using Document Analysis with OCR," Proc. 2nd Int. Conf. Intell. Comput. Control Syst. ICICCS 2018, pp. 190–195, 2019, doi: 10.1109/ICCONS.2018.8663011
- [7]. Sushruth Shastry, Gunasheela G, Thejus Dutt, Vinay D S and Sudhir Rao Rupanagudi, A novel algorithm for Optical Alphabet Recognition (OCR). 978-1-4673-5090-7/13, 2013, IEEE
- [8]. Lulu Zhang, Xingmin Shi, Yingjie Xia, Kuang Mao, "A Multi-filter Based License Plate Localization and Recognition Framework". 978-1-4673-4714-3/13, 2013 IEEE
- [9]. Jieun Kim, and Ho-sub Yoon "Graph Matching Method for Alphabet Recognition in Natural Scene Images. 978-1-4244-8956-5/11, 2011, IEEE
- [10]. Ibrahim El Khatib, Yousef Samir-Mohamad Omar, and Ali Al Ghouwayel, "AN EFFICIENT ALGORITHM FOR AUTOMATIC RECOGNITION OF THE LEBANESE CAR LICENSE PLATE. ISBN: 978-1-4799-5680-7/15, 2015, IEEE.

- [11]. Feng Yanga, and Fan Yangb, “Alphabet Recognition Using Parallel BP Neural Network”. 978-14244-1724-7/08, 2008, IEEE
- [12]. Rókus Arnold, and Póth Miklós “Alphabet Recognition Using Neural Networks”. 11th IEEE International Symposium on Computational Intelligence and Informatics, 18–20 November, 2010, Budapest, Hungary
- [13]. Amarjot Singh, Ketan Bacchuwar, Akash Choubey, and Devinder Kumar, “An OMR Based Automatic Music Player”. 978-1-61284-840-2/11, 2011, IEEE.
- [14]. Shan Du, Member, IEEE, Mahmoud Ibrahim, Mohamed Shehata, Senior Member, IEEE, and Wael Badawy, Senior Member, IEEE “Automatic License Plate Recognition (ALPR):A State-of-the-Art Review, IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, VOL. 23, NO. 2, FEBRUARY 2013.
- [15]. Imran Shafiq Ahmad, Boubakeur Boufama, Pejman Habashi, William Anderson and Tarik Elamsy, “Automatic License Plate Recognition: A Comparative Study”. IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2015
- [16]. Rejean Plamondon, Fellow, IEEE, and Sargur N. Srihari, Fellow, IEEE, “On-Line and Off-Line Handwriting Recognition:A Comprehensive Survey” IEEE Transactions on Pattern Analysis And Machine Learning Intelligence, Vol. 22, No. 1, Jan 2000
- [17]. Hadar I. Avi-Itzhak, Thanh A. Diep, and Harry Garland, “High Accuracy Optical Alphabet Recognition Using Neural Networks with Centroid Dithering IEEE Transactions on Pattern Analysis And Machine Learning Intelligence, Vol. 17, No. 2, Feb 1995