ISSN No:-2456-2165

System for Identifying Texts Written in Kazakh Language

Absenova B.M., Zhanibay S.B., Zheksebay D.M., Temesheva S.A., Turmaganbet U.K.
Faculty of physics and technologies, Al-farabi named Kazakh National UniversityAlmaty, Kazakhstan

Abstract:- Recently, image-based text extraction has become a prominent and hard study subject in computer vision. In this article, the texts written in Kazakh are classified based on factors such as writing style and diversity of writing, and a text recognition system based on correctly defined terms is developed. Text matching is accomplished by using the EasyOCR library to the input picture to extract the areas containing the words. The words in the text are then determined using the locations acquired. To initialize the text and recognize the defined text, two distinct functions are utilized. As a consequence, a method for recognizing Kazakh words as graphics is developed. The accuracy of proper text identification revealed a result of 91%.

Keywords:- *Text, detection, optical character recognition, scene,detector.*

I. INTRODUCTION

In any social network, e-book, or website, the problem of proper text recognition is quite high. The writing may be found on documents, road signs, billboards, and other items like vehicles and phones. In current gadgets, automatically discovering and recognizing text from pictures is a system that plays a vital part in many tough challenges, such as machine translations or image and videotapes. [1]. Last year, the objective of document identification and identification of documents in immediate settings piqued the interest of the computer vision and paper analysis communities. Furthermore, recent advances [2-5] in other areas of computer vision have enabled the development of even more powerful text detection and identification systems than previously [6-8].

Text detection and identification in pictures is gaining popularity in computer vision and image comprehension due to its multiple potential applications in image search, scene understanding, visual assistance, and so on. [9] Text wall reading has various real-world applications [10], such as assisting visually impaired persons [11], robotic vision [12], unmanned vehicle navigation [13], human-computer interface [14], geolocation.

Technical text recognition is divided into two stages: (a) Text detection, which identifies and locates text in natural situations and/or video. In a nutshell, it is the process of distinguishing between text and non-text regions. (b) Text recognition is the representation of a text's semantic meaning.

(1) A text screen, which is often a click camera and depicts a typical scenario. Due to unpredictable conditions, such as advertising holdings, signs, storefronts, text on buses, front panels, and many others, this generates an unorganized and confusing environment. (2) Captions or graphic text are manually added to photographs and/or films to supplement visual and aural material, making text extraction easier than in a natural text situation. Difficulties with the diversity and complexity of natural pictures for text extraction are addressed from three perspectives: (1) natural text change owing to variable font size, style, color, scale, and orientation, (2) backdrop ambiguity, and the existence of diverse problems, such as characters. All of these elements impede the separation of the real text from the natural visual, which might lead to mistakes. (3) Interference elements including noise, low quality, distortion, and uneven illumination also cause issues when detecting and extracting text from a natural backdrop. [17]

Difficulties with the diversity and complexity of natural pictures for text extraction are addressed from three perspectives: (1) natural text change owing to variable font size, style, color, scale, and orientation, (2) backdrop ambiguity, and the existence of diverse problems, such as characters. All of these elements impede the separation of the real text from the natural visual, which might lead to mistakes.

(3) Interference elements including noise, low quality, distortion, and uneven illumination also cause issues when detecting and extracting text from a natural backdrop. [17]

Using the benefits of deep image representation, a variety of deep models for filtering/classifying textual components [23], [24], [25] or word/character identification [26], [25], [27], [24] have recently been constructed. For text detection, the author of [23] utilized a two-layer CNN with an MSER detector. The CNN model was utilized for detection and identification tasks in studies [25, 24]. These text classifiers are often trained using a binary text/non-text label, which is significantly less useful for learning.

In this paper, we focus on the challenge of identifying text in the Kazakh language, with the goal of correctly localizing the exact places of text lines or phrases in an image. Text information in photographs can include a wide range of text patterns on a very complicated background. For instance, the text may be very tiny, in a different font, and with a complicated underlining. This presents basic challenges for this activity, as determining the right components of symbols is difficult. To tackle this challenge, a variety of text detection technologies must be used.

ISSN No:-2456-2165

II. METHODOLOGY

A. Selecting a Template

One of the most pressing challenges at the present is textual definition. To determine the text written in the Kazakh language with true precision, neural network models are utilized. The Python programming environment is used to design a system for identifying text written in Kazakh in this article.

The easyocr approach is quite useful when programming. As the name implies, EasyOCR is a Python library that enables computer vision engineers to do optical character recognition with minimal effort. When it comes to optical character recognition, EasyOCR is presently the most user-friendly option:

The EasyOCR package has few dependencies, which simplifies the setup of the OCR development environment.

To do OCR, just two lines of code are required: one to start the Reader class and the other to recognize the picture using thetext reader method.

Following that, a simple Python script will be written to do optical character recognition with the EasyOCR module.

Python and the PyTorch libraries are used to implement EasyOCR. EasyOCR currently only enables the recognition ofprinted text.

Optical character recognition with EasyOCR

Because the putText function in OpenCV cannot show non-ASCII characters, a quick and handy method is built to parse these potentially unpleasant characters.

The cleanup text helper method merely checks that the sequence number of characters in the text string given is less than 128, and removes any other characters. The command line options are now decided when the input data is ready to work.

Three command line parameters are accepted by the script:

- image: path to an input picture with text for OCR.
- langs: a list of language codes separated by commas (no spaces). The script assumes English by default (en).
- gpu: a graphics processing unit.
- langs is a comma-separated list of Python languages for the EasyOCR engine. The —image input is then loaded.

Note. EasyOCR, unlike Tesseract, can operate with the BGR OpenCV default color channel order. As a result, changing the color channels after loading the

To use EasyOCR for optical character recognition, an instance of the Reader object is first constructed, with lang andthe logical value —gpu passed to the constructor. When the input picture is passed, the readtext method is invoked.

The EasyOCR output consists of three tuples:

- box: The coordinates of the localized text's bounding box.
- text: OCR string
- prob: The likelihood of OCR findings
- When you select an EasyOCR result, the bounding boxcoordinates are first unpacked.

III. RESULTS AND DISCUSSION

For testing, the EasyOCR library was utilized, and the implementation was done in Python. The Reader string from the EasyOCR library was used to initialize the text, and the readtext method was used to identify the image. Figures 1 and 2 illustrate the texts for recognizing the terms in the text, as well as the tests that were performed.

The first major issue with this categorization is that, other from the 42 classes chosen (from A to ; from A to ;), all other items are not recognized as text [14]. There are several subclasses of non-textual objects. This makes handwritten text recognition more challenging.

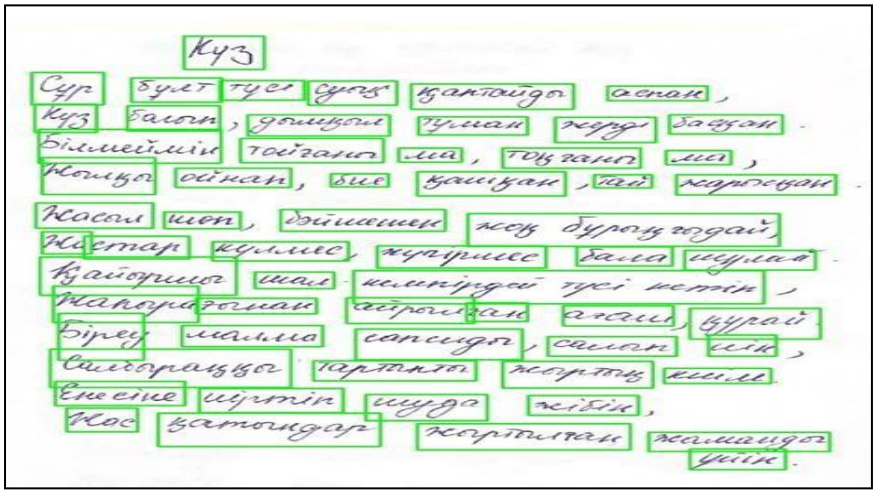

Fig. 1: First text recognition using the EasyOCR package

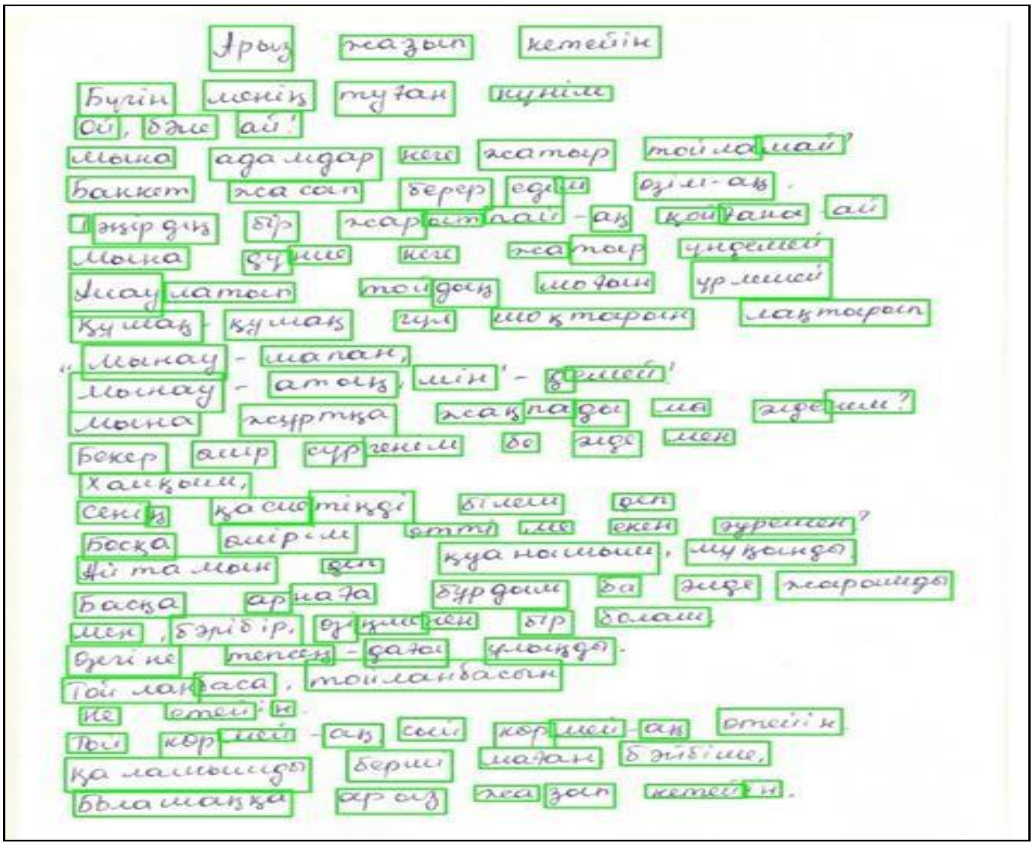

Fig. 2: Second text recognition using the EasyOCR package

To identify the inscription-text on the photograph, a collection of texts written in Kazakh is required. As a result, writings written in various styles were used as input. The "bbox" function of the EasyOCR package was used to calculatethe coordinates of the words in the provided text in order to determine the text in the photos. The words in the supplied text were detected independently as green rectangles, encircled, and chopped using the same coordinates as shown in Figure 1. Words are detected within the rectangle provided by the "Text" function. The "Prob" function returns the identified words as separate pictures.

Overview ofresults	The number of words in the given text	The number of incorrectly Identified words	The averagevalue of the error
According tothe first text	233	23	9,87%
According tothe second text	319	31	9,7%

Table 1: Overview of results

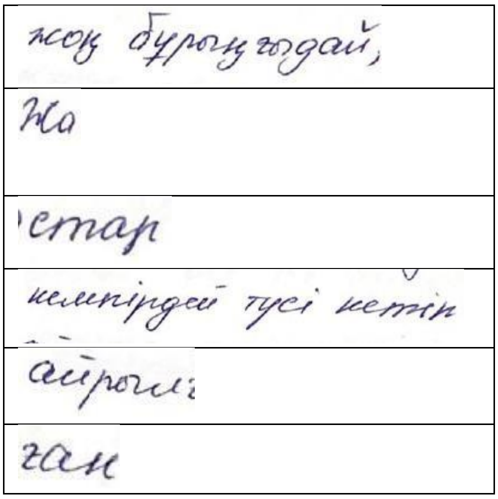

Table 2: Example of errors on the first text

ISSN No:-2456-2165

The results of utilizing OCR with the EasyOCR library are shown above. The —langs argument instructs text (and, eventually, EasyOCR) to identify Kazakh letters and words.

As input, several texts written in various scripts were provided. The algorithm was able to recognize each individual word in the provided text during text identification. However, due to a variety of factors, each individual word in the given text was incorrectly identified. Tables 2 and 3 show the list of words that were incorrectly identified for the given texts.

Some inaccuracies were created as a result of the first text's results, as displayed in the second table. That is, the Kazakh text recognition system encountered difficulty during text recognition in the following cases:

Even if the words are separated in the text;

If letters unique to the Kazakh language are written in aconfusing manner;

If the letters, {з, р, д, й, ү, ұ, у} are jumbled up in the textidentification area, and so on.

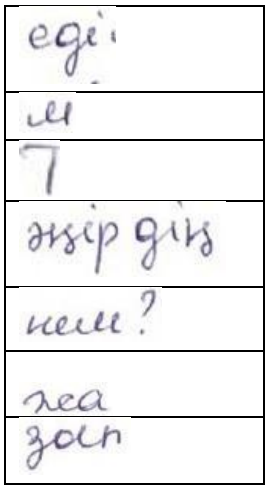

Table 3: Example of errors on the second text

The second text resulted in errors, as shown in table 3. In the following situations, the system had difficulty identifying the text for this text:

- Even if the words in the text are printed in different sizes;
- It is tough to write letters together;
- If words are typed using special symbols when writing the text (!?)
- Even if the letters in the text's words are written at a distancefrom one another;

Although the identification of individual words in Kazakh text works systematically, it was discovered that external human factors are an impediment to identification.

IV. CONCLUSION

A technique for detecting individual words in Kazakh-language texts was created in this study. During the system's development, efforts were made to detect words written in the Kazakh language with varying numbers of symbols, writing styles, underlining, and sizes. The findings were obtained using the EasyOCR package and the Python programming environment. The system produced the following results: it identified texts in the Kazakh language and outputted the information as individual recognized words as independent pictures. Many environmental and personal elements of the individual were discovered to impact the high accuracy of identification for the identification of words in the provided text. A system based on recognizing individual letters is being developed to improve the system for identifying texts in the Kazakh language and eliminate errors. In addition, neural network models will be used to enhance the system.

REFERENCES

[1.] Unless there are six authors or more give all authors' names; do not use "et al.". Papers that have not been published,

[2.] Bartz C., Yang H., Meinel C. STN-OCR: A single neural network for text detection and text recognition //arXiv preprint arXiv:1707.08831. – 2017.

[3.] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 770–778, 2016

[4.] M. Jaderberg, K. Simonyan, A. Zisserman, and K. kavukcuoglu. Spatial transformer networks. In Advances in Neural Information Processing Systems 28, pages 2017–2025. Curran Associates, Inc., 2015.

[5.] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi. You only look once: Unified, real-time object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 779– 788, 2016.

[6.] S. Ren, K. He, R. Girshick, and J. Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In C. Cortes, N. D. Lawrence, D. D. Lee, M. Sugiyama, and R. Garnett, editors, Advances in Neural Information Processing Systems 28, pages 91–99. Curran Associates, Inc., 2015.

[7.] L. Gomez-Bigorda and D. Karatzas. Textproposals: a

textspecific selective search algorithm for word spotting in the wild. arXiv:1604.02619 [cs], 2016.
[8.] Gupta, A. Vedaldi, and A. Zisserman. Synthetic data for text localisation in natural images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2315–2324, 2016.
[9.] Shi, X. Wang, P. Lyu, C. Yao, and X. Bai. Robust scene text recognition with automatic rectification. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4168– 4176, 2016.
[10.] He, Tong, et al. "Text-attentional convolutional neural network for scene text detection." IEEE transactions on image processing 25.6 (2016): 2529-2541.
[11.] Francis, Leena Mary, and N. Sreenath. "Live detection of text in the natural environment using convolutional neural network." Future Generation Computer Systems 98 (2019): 444-455.
[12.] H. Kandemir, B. Cantürk, M. Ba¸stan, Mobile reader: Turkish scene text reader for the visually impaired, in: Signal Processing and Communication Application Conference (SIU), 2016 24th, IEEE, 2016, pp. 1857–1860
[13.] H. Wang, J. Wang, W. Chen, L. Xu, Automatic illumination planning for robot vision inspection system, Neurocomputing 275 (2018) 19–28
[14.] F. Zakharin, S. Ponomarenko, Unmanned aerial vehicle integrated navigation complex with adaptive tuning, in: Actual Problems of Unmanned Aerial Vehicles Developments (APUAVD), 2017 IEEE 4th International Conference, IEEE, 2017, pp. 23–26.
[15.] S. J. Joo, A. L. White, D. J. Strodtman, J. D. Yeatman, Optimizing text for an individual's visual system: The contribution of visual crowding to reading difficulties, Cortex 103 (2018) 291–301..
[16.] F. Sheng, C. Zhai, Z. Chen, B. Xu, End-to-end chinese image text recognition with attention model, in: International Conference on Neural Information Processing, Springer, 2017, pp. 180–189
[17.] J. Ye, Z. Xu, Y. Ding, Image search scheme over encrypted database, Future Generation Computer Systems.
[18.] Francis, Leena Mary, and N. Sreenath. "Live detection of text in the natural environment using convolutional neural network." Future Generation Computer Systems 98 (2019): 444-455.
[19.] Matas, J., et al., Robust wide-baseline stereo from maximally stable extremal regions. Image and vision computing, 2004. 22(10): p. 761- 767.
[20.] Neumann, L. and J. Matas. A method for text localization and recognition in real-world images. in Asian Conference on Computer Vision. 2010: Springer.
[21.] Epshtein, B., E. Ofek, and Y. Wexler. Detecting text in natural scenes with stroke width transform. in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. 2010: IEEE.
[22.] Kasar, T., J. Kumar, and A. Ramakrishnan. Font and background color independent text binarization. in Second international workshop on camera-based document analysis and recognition. 2007.
[23.] Francis, Leena Mary, and N. Sreenath. "TEDLESS–Text detection using least-square SVM from natural scene." Journal of King Saud University-Computer and Information Sciences 32.3 (2020): 287-299.
[24.] W. Huang, Y. Qiao, and X. Tang, "Robust scene text detection with convolutional neural networks induced mser trees," 2014, in European Conference on Computer Vision (ECCV)
[25.] T. Wang, D. Wu, A. Coates, and A. Y. Ng, "End-to-end text recognition with convolutional neural networks," 2012, in International Conference on Pattern Recognition (ICPR).
[26.] M. Jaderberg, A. Vedaldi, and A. Zisserman, "Deep features for text spotting," 2014, in European Conference on Computer Vision (ECCV).
[27.] P. He, W. Huang, Y. Qiao, C. C. Loy, and X. Tang, "Reading scene text in deep convolutional sequences," 2016, in The 30th AAAI Conference on Artificial Intelligence (AAAI-16).
[28.] Bissacco, M. Cummins, Y. Netzer, and H. Neven, "Photoocr: Reading text in uncontrolled conditions," 2013, in IEEE International Conference on Computer Vision (ICCV).