**NEWS**
OF THENATIONAL ACADEMY OF SCIENCES OF THE REPUBLIC OF KAZAKHSTAN
**PHYSICO-MATHEMATICAL SERIES**
ISSN 1991-346X
Volume 2, Number 336 (2021), 102 – 108 **https://doi.org/10.32014/2021.2518-1726.27**

UDC 0004.5
IRSTI 20.53.19

**Zh. H. Spabekova, A. G. Karelova, A. E. Qami, Z. S. Abilkaiyr**

Kazakh National University named after Al-Farabi, Almaty, Kazakhstan.
E-mail: qamiaiko@gmail.com

# RECOGNITION OF IDENTIFICATION DATA OF BANK CARDS

**Abstract.** This article describes the recognition of bank card information. Recognizing an object with a camera is one of the most important tasks at the moment. Recognizing credit card data at the same time is a rather complex algorithmic task, but at the moment the implementation of this task is very relevant and in-demand due to the increase in the number of payment transactions via mobile devices. The implementation of this task can save a person from having to enter most of the data when making online payments. The fundamental difficulties of this problem are discussed and methods for solving it are proposed. The problem under consideration is solved for the case of application on mobile devices, which imposes strict requirements for computational complexity. The article presents the results of a formal analysis of the performance and accuracy of the proposed algorithm. The error spectrum of the recognition system as a whole shows that the proposed algorithm solves the problem with the required accuracy.

The main question that was investigated at this work: is it possible to use the Tesseract OCR library for text recognition from video images, for example, timecode? That is, digital time data embedded in the footage images. This is important for the automation of individual procedures for video technical expert studies.

Object recognition by the camera is one of the most important tasks at the moment. The fundamental difficulties of this problem are discussed and methods for its solution are proposed. The article presents the results of a formal analysis of the performance and accuracy of the proposed algorithm. The spectrum of errors of the recognition system as a whole shows that the proposed algorithm solves the problem with the required accuracy.

**Keywords:** Python, recognition, bank card, embossed numbers, OpenCV.

**Introduction.** The development of banking technologies has led to the widespread introduction of Internet banking, i.e. remote banking customer service via the Internet, as well as mobile banking, which provides the ability to manage bank accounts using a tablet computer or smartphone. Currently, one client of the bank may have two or more bank cards, which he can use for payment. Filling in all data fields when paying for services or making purchases using mobile banking requires time and attention. Therefore, the improvement of mobile banking involves the development of algorithmic and software to automate the entry of bank card details into the system based on the analysis of its image obtained using a mobile device.

A bank card is a typical example of a document with a flexible form, therefore, when processing it, algorithms for detecting and recognizing data that are used on other types of documents with a flexible form can be used [1]. The work [2] presents an approach for the segmentation and recognition of text characters on business cards with a plain background. In this case, such basic steps are used as the selection of edges by the Sobel operator, line thinning, projective transformation, adaptive binarization, segmentation of words and text symbols with their subsequent recognition. The application of this approach is limited by the ability to work only with a simple background and rather significant computational costs, which imposes a limitation on its implementation. The paper proposes an algorithm for processing business cards at the first stage of which a rough background removal is performed based on block analysis of the input image, at the second stage the obtained connected components are classified to determine text regions and exclude logo images. Then adaptive binarization is applied to the result, which allows you to separate the text from the background for further recognition. This algorithm can also be used only for images of cards with a monotone background.

Development uses python for preprocessing and segmentation of text characters, and the Tesseract library for their recognition. In this article, preprocessing uses several operations with high computational complexity to improve the input image and segmentation of characters, therefore, a simplified version of the algorithm is implemented on a mobile device with the absence of certain preprocessing stages, which led to a significant decrease in the quality of its work. At the first step, the size of the input image is reduced, preprocessing operations are performed using the OpenCV library, and character recognition is based on Tesseract. However, the results of experiments in the works are presented for images of business cards with a single-color background, on which text symbols are visible.

The capabilities of the algorithms for recognizing bank card numbers under the control of python software also do not include the recognition of the user's first and last name. The use of preprocessing of card images ensures the accuracy of digit recognition up to 90%. It should also be noted that these algorithms are designed to work with embossed cards with numbers and letters embossed on the front side. In [3], a method for constructing the reliability function of recognition of images of embossed symbols is presented and its effectiveness is shown to increase the probability of their correct recognition. However, non-embossed cards have become widespread, which can be issued to a client on the day of his application to the bank. But due to the lack of a procedure for extruding alphanumeric information, its localization and recognition are much more difficult and require effective preprocessing to localize characters and ensure their correct recognition.

Existing approaches use static images as input data, on which in some cases noise, glare, significant changes in brightness, insufficient sharpness, etc. can occur, which leads to the need to obtain a new image, which may also be unsuitable for correct detection cards, information fields on it, and (or) their further recognition. Improving the efficiency of data processing can be achieved through the use of video, which is easily obtained using a mobile phone. However, the presence of not one image, but their sequence leads to the need to develop other approaches to processing. In [4], it was proposed to use video sequences to detect document edges by mobile devices. At the same time, image segmentation is performed in the Lab color space using morphological processing, linking the boundaries of individual segments based on the Hough transformation and detecting lines that limit the document as a whole. However, to reduce computational costs, a reduction of the input image frame size from 1280 × 720 to 180 × 100 pixels is applied. Such a reduction in the frame is not suitable for solving the problem of localizing and recognizing information on a bank card, since the probability of character recognition will be low due to its small size.

In this paper, an algorithm is proposed for detecting and recognizing bank card details by video sequences. The algorithm allows almost real-time recognition of information fields on the front of a bank card, for embossed and non-embossed types of cards.

**Setting of problems**

**2. Algorithm for recognizing bank card details by video sequences**

In the developed algorithm, a sequence of frames obtained with a video camera enters the rectangular areas detection unit, which detects all rectangles in the frame and returns only one that satisfies the bank card parameters.

This area is converted to a grayscale image, which is transferred to the segmentation unit to highlight and index areas on the card image that correspond to information fields such as the card number, its expiration date, and the owner's first and last name. For segmented blocks, filtering contrast is improved. This is followed by adaptive binarization and morphological processing. The next step is to refine the boundaries of the symbol areas and uses the sliding window method. Thereafter, the segmented areas are passed to the OCR unit to recognize digital and text characters. Finally, the data evaluation unit processes the received information to display or reject the result. If the received card number and expiration date are valid, the result is considered correct for them. Recognition of the owner name field data stops after the first result is obtained. If information is not read from all three information fields in the current frame, then only the missing ones are processed in the next frame.

**Methods**

**2.1. Card detection.** When developing a card detection algorithm, it is taken into account that the frame size of the input video sequence varies and can reach a resolution of 3840 × 2160 pixels; the devices also provide automatic focus on the object, brightness correction, and white balance. Thus, the resulting image will, as a rule, have the following acceptable characteristics: sharpness, brightness, and sufficient frame resolution for automatic extraction of bank card data.

At the first stage, all rectangular areas in the image are detected and localized. Further, to determine the contour describing the bank card, the analysis of the aspect ratio of the sides of the obtained objects is used. Since the height $m_o$ and width $n_o$ of the card are standardized, the ratio of the dimensions of its sides is a constant value, it is advisable to use it as a criterion when finding the card outline from all detected rectangular areas. Then a rectangular region $r_i$ with dimensions $m_i \times n_i$ can be assigned to the card image in the frame if the condition is met:

$$\frac{no}{mo} - e < \frac{n_i}{m_i} < \frac{no}{mo} + e \quad (1)$$

where $e$ is the coefficient of permissible deviation of the aspect ratio of the card sides.

An unambiguous match is considered to be the rectangle with the longest side $m_{max}$ of all detected ones. The dimensions of the rectangular area found at this stage and its location relative to the entire image $p_i(x_i, y_i)$ are used to extract the card area from the input frame.

**2.2. ROI segmentation.** On the resulting image of card $I$ with dimensions $m_I \times n_I$, t is necessary to find areas containing information about the bank card number ($I_C$). The size and location of these areas are defined by ISO / IEC 7811-5.4: 2018, which can be used to segment these fields, formally represented as: $C(x_C, y_C, m_C, n_C)$ – card number;

Since the values of $m_I$ and $n_I$ can vary during filming, it is necessary to provide normalization to correctly determine $C$. For this, scaling factors are used, which are multiplied by the parameters of these areas, in width ($C_{sw}$) and in height ($C_{sh}$):

$$C_{sw} = \frac{n_I}{n_0}, C_{sh} = \frac{m_I}{m_0}. \quad (2)$$

The bank card number contains 16 digits, divided into 4 equal groups of 4 digits each. Based on this, the $I_C$ image selected at the previous stage with the area of the $C_I$ card number is divided into 4 equal areas $C_{I1}$, $C_{I2}$, $C_{I3}$, $C_{I4}$.

**2.3. Segment display enhancement.** Using the operations of mathematical morphology, the boundaries of symbols are underlined on the grayscale image obtained from the previous stage. To determine the color of the background and symbols, the average brightness value is calculated. If the result is greater than 127, then the background color is light and the character color is dark, and vice versa.

If the color of the characters is defined as light, a morphological transformation WhiteTopHat is performed on the image, which subtracts the open image from the original one, thus emphasizing the details of the outlines of light characters. The BlackTopHat transform is used to emphasize the detail of the borders of dark characters by subtracting the original image from the closed image. The structuring element $b$ for the cores of both filters has a rectangular shape and its size $n_b \times m_b$ , taking into account the experimentally determined scaling factors, is calculated as:

$$n_b = 0.06n_I, m_b = 3n_b \quad (3)$$

where $n_I$ is the width of the image received for transformation.

**2.4.** Localization of borders. Since in the images of bank cards, despite their small size, the brightness-contrast characteristics of the fragments can differ significantly, the brightness levels of the pixels of the background areas on which the symbols are located also change. Taking into account this feature and the need to preserve the boundaries of symbols, the image binarization method with an adaptive threshold is used based on the analysis of the local region according to the expression:

$$bin_{x,y}\begin{cases}255, if\ Y_{x,y} > T(x,y)\\ 0, otherwise\end{cases}$$

where $T(x, y)$ is an adaptive binarization threshold, calculated for each pixel as $T(x, y) = = corr(I, G) - c$; $corr(I, G)$ – cross-correlation of a local $k \times k$ image fragment with a Gaussian window; $c$ is a constant. The $k \times k$ Gaussian window coefficient matrix is defined as[6]:

$$G_i = \alpha \cdot e^{\frac{-\left(i-\frac{(k-1)}{2}\right)^2}{2\cdot\sigma^2}}, \tag{4}$$

$$\sigma = 0.3\cdot ((k-1)\cdot 0.5 - 1) + 0.8,$$

where $i = 0 \ldots k - 1$, α is the scaling factor chosen in such a way that $\sum_i G_i = 1$

To reduce the amount of noise and remove uninformative details from the image, as well as to roughly separate the characters from the background of the image, morphological operations of closure and erosion are applied sequentially.

The shape of the structuring element $b$ of size $m_b \times n_b$ ($n_b = \frac{w_{symb}}{3}$) of the kernels of both filters is set taking, into account the peculiarities of the font characters according to ISO / IEC 7811–7.1: 2018 in the form ellipse. The thickness of the symbol $w_{symb}$ does not depend on the type of the processed fragment of the card image and is defined for OCR-B. The projection of the font width (mm) by the size in pixels, relative to the size of the bank card area, is a constant and is calculated based on its full width:

$$w_{symb} = 0.004884 \cdot n_0 \tag{5}$$

Next, the edges of the character area on the processed image are refined based on the sliding vertical window method. The height of the window varies according to the font size of the characters. $H_{dc}$ – card number font height (4.0 mm). Then their projections to the height in pixels relative to the size of the full card area $H_{dc}$. The general view of the calculation is represented by the formula:

$$h = \frac{H \cdot m_0}{54.0} \tag{6}$$

where 54.0 mm is the height of the card (ISO / IEC 7811-6.3: 2018), and $m_O$ is the height of the detected area of the card (pix).

The received character area from the previous step, black and white lists of characters, as well as the language identifier, are sequentially transferred to the Tesseract recognition system.

The white list, consisting of a set of numbers from 0 to 9 and the language identifier "eng", are transmitted as parameters for displaying the bank card number. For an image with an expiration date: the numbers 0 through 9 and the "/" character form the white list, and "eng" is the language identifier. For the name of the cardholder: "rus" and "eng" are language identifiers, and a set of characters, including punctuation marks, special characters, and numbers, is blacklisted. This approach helps to increase the speed of character recognition by the Tesseract library and speed up the entire processing process as a whole.

**Results.** Testing of the quality characteristics of the developed algorithm was performed using a database of 180 bank cards, which were placed on a complex background, the information fields of the details were not blocked by foreign objects and were visually recognized.

The values of the coefficients that are used in the calculations were experimentally obtained using the ROC analysis technique [7]. The calculation of the permissible deviation coefficient of the aspect ratio of the sides of the bank card ($e$) was carried out taking into account its exact dimensions by ISO / IEC 7811-5.4: 2018 on a sample of 214 images, 178 of which had card objects, 36 had no cards, or their borders have been blurred or distorted. As a result of the experiment, for each coefficient $e$, which varied from 0 to 0.02 with a step of 0.0002, the proportions of true (TPR) and false (FPR) detection were calculated. For $e = 0.011$, the maximum TPR value is 0.974194 and the minimum FPR = 0.111111

The size of the local region for binarization was determined taking into account the scaling factor since this parameter depends on the input dimensions of the detected area of the bank card (its width ($W_I$, pixels)): $k = W_I \cdot k'$.

The $k'$ values varied from 0.002 to 0.03 with a step of 0.001. At this stage, the maximum value of the fraction of true character recognition TPR = 0.83447 and the fraction of false recognition of characters FPR = 0.140864 was obtained for $k'$ = 0.016. The constant c = 13 is defined similarly.

According to the results obtained using the ROC analysis, after setting the optimal coefficients and threshold values in the implementation of the algorithm, the proportion of truly positive classification of all data fields on the card is TPR = 0.88, and the proportion of false-positive classification is FPR = 0.14. Figure 1– 3 shows examples of bank card detection from a set of video sequences, indicating keyframes and detection diagrams. Card detection on the diagram corresponds to – 1, absence – 0. Video capture was carried out at a constant frequency of 30 frames / s.

Figure 1 - An example of frames of a video sequence

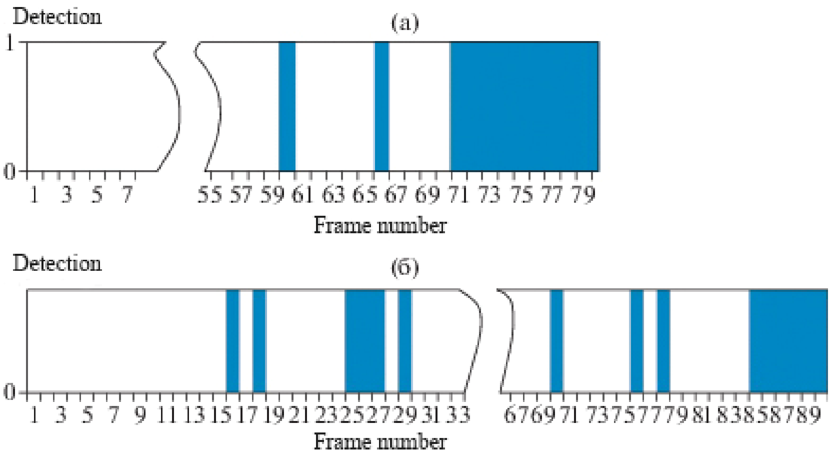


Figure 2 – Diagram of detecting a card on a video sequence

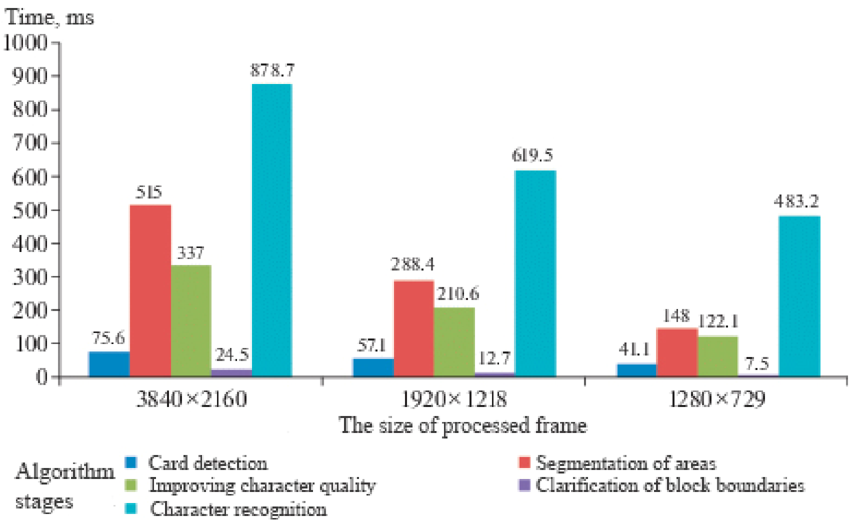


Figure 3 – Algorithm time-consuming distribution

The time spent for each stage of processing the algorithm of one frame of a video sequence of different resolutions is shown in figure 3.

As can be seen from the graphs (figure 3), the time spent on the character recognition operation has a non-linear dependence on the size of the processed frame, in contrast to other operations.

**Discussion.** The experiments were carried out for 180 cards under various shooting conditions, of which for 104 cases the cards were embossed with symbols.

An analysis of the recognition results for all the cards in the database used shows that the developed approach significantly improves recognition on cards with non-embossed characters, which does not provide for the recognition of the cardholder's name. In general, the proportion of truly positive classification by the proposed algorithm of all three data fields on the card is TPR = 0.88, for the card number and validity period – TPR = 0.925.

Analysis of the results obtained during research shows that detection and recognition errors occur due to the influence of a combination of factors: the location of the card object on a complex background, similar to the image of the card itself; the presence of glare of artificial lighting or a significant difference in brightness; too far away from the camera; blurring the outlines of the card if the camera or the card moves during video recording. Further improvement of the algorithm is planned by improving the character recognition algorithm.

**Conclusion.** This article presents an algorithm for recognizing the details of the front side of bank cards.

Improving the efficiency of card detection and data recognition in comparison with existing approaches is provided by analyzing the sequence of images obtained from the video sequence. The proposed approach includes the detection of card boundaries in the frame based on the Viola-Jones algorithm and the OverFeat method, segmentation of information fields taking into account their location on the card, improvement of segment images using histogram normalization and morphological processing, determination of the boundaries of symbol blocks based on adaptive binarization and morphological processing, their refinement using a sliding vertical window, character recognition using the Tesseract library. The software implementation of the developed algorithm is performed using python, the OpenCV, and Tesseract libraries. At the same time, the share of truly positive classification by the proposed algorithm of all three data fields on the card is TPR = 0.88, for the card number and validity period – TPR = 0.925.

**Ж. Х. Спабекова, А. Карелова, А. Е. Қами, Ж. С. Әбілқайыр**

Әл-Фараби атындағы Қазақ ұлттық университеті, Алматы, Қазақстан.

**БАНК КАРТАЛАРЫНЫҢ ИДЕНФИКАЦИЯЛЫҚ МӘЛІМЕТТЕРІН ТАНУ**

**Аннотация.** Бұл мақалада банк картасының ақпаратын тануы туралы сипатталған. Нысанды камера арқылы тану қазіргі кездегі ең маңызды міндеттердің бірі болып табылады. Несие карталарының деректерін тану бір уақытта өте күрделі алгоритмдік міндет болып табылады, бірақ қазіргі уақытта мобильді құрылғылар арқылы төлем операциялары санының көбеюіне байланысты өте өзекті және сұранысқа ие мәселе болып табылады. Бұл міндетті жүзеге асыру адамды төлемдерді онлайн режимінде жүзеге асырған кезде көптеген мәліметтер енгізуден құтқара алады. Бұл мәселенің негізгі қиындықтары талқыланып, оны шешудің әдістері ұсынылады. Қарастырылып отырған мәселе мобильді құрылғыларға қолдану жағдайында шешіледі, бұл есептеу қиындығына қатаң талаптар қояды. Мақалада ұсынылған алгоритмнің өнімділігі мен дәлдігін ресми талдау нәтижелері келтірілген.

Осы жұмыста зерттелген негізгі сұрақ: Tesseract OCR кітапханасын видео кескіндерден мәтінді, мысалы, уақыт кодын тану үшін пайдалануға бола ма? Яғни, түсірілген кескін шеңберлеріне енгізілген уақыттың сандық деректері. Бұл бейне техникалық сараптамалық зерттеулер жүргізудің жеке рәсімдерін автоматтандыру үшін маңызды.

Заттарды камерамен тану - қазіргі кездегі маңызды міндеттердің бірі. Бұл мәселенің негізгі қиындықтары талқыланып, оны шешудің әдістері ұсынылады. Мақалада ұсынылған алгоритмнің өнімділігі мен дәлдігін ресми талдау нәтижелері келтірілген. Жалпы тану жүйесінің қателіктерінің спектрі ұсынылған алгоритм есепті қажетті дәлдікпен шешетіндігін көрсетеді.

**Түйін сөздер**: Python, тану, банктік карта, рельефтік нөмірлер, OpenCV.

**Ж. Х. Спабекова, А. Карелова, А. Е. Қами, Ж. С. Әбілқайыр**

Казахский национальный университет имени аль-Фараби, Алматы, Казахстан

**РАСПОЗНАВАНИЕ ИДЕНТИФИКАЦИОННЫХ ДАННЫХ БАНКОВСКИХ КАРТ**

**Аннотация.** В данной статье описывается распознавание информации банковской карты. Распознавание объекта с помощью камеры – одна из самых важных задач на данный момент. Распознавание данных кредитных карт одновременно является достаточно сложной алгоритмической задачей, однако реализация этой задачи очень актуальна и востребована в связи с увеличением количества платежных операций с помощью мобильных устройств. Реализация этой задачи сможет избавить человека от необходимости вводить большую часть данных при совершении онлайн-платежей. Обсуждаются фундаментальные трудности этой проблемы и предлагаются методы ее решения. Рассматриваемая задача решается для применения на мобильных устройствах, что предъявляет жесткие требования к вычислительной сложности. В статье представлены результаты формального анализа производительности и точности предложенного алгоритма.

Основной вопрос, который был исследован в данной работе: можно ли использовать библиотеку Tesseract OCR для распознавания текста из видеоизображений, например, timecode? То есть цифровые временные данные, встроенные в отснятые кадры изображений. Это важно для автоматизации отдельных процедур проведения видеотехнических экспертных исследований.

Распознавание объектов камерой – одна из важнейших задач на данный момент. Обсуждаются фундаментальные трудности этой проблемы и предлагаются методы ее решения. В статье представлены результаты формального анализа производительности и точности предложенного алгоритма. Спектр ошибок системы распознавания в целом показывает, что предложенный алгоритм решает поставленную задачу с требуемой точностью.

**Ключевые слова:** Python, распознавание, банковская карта, эмбоссированные цифры, OpenCV.

**Information about authors:**

Spabekova Zhanara Khalilaevna, Senior Lecturer of the Department "Informatics", Al-Farabi Kazakh National University, zhanaraspabekova@gmail.com, https://orcid.org/0000-0001-5970-3286 ;

Abilkaiyr Zhaina Serzhankyzy, Master's student of Al-Farabi Kazakh National University, abilkhair.zhaina.98@gmail.com, https://orcid.org/0000-0001-6748-428X;

Qami Ayaulym.Yerbolkyzy, Master's student of Al-Farabi Kazakh National University, qamiaiko@gmail.com, https://orcid.org/0000-0001-9840-9178;

Karelova Ayagul Gabidenkyzy, Master's student of Al-Farabi Kazakh National University, karelovaayagul@gmail.com, https://orcid.org/0000-0003-3960-4226

**REFERENCES**

[1] Sheshkus A., Nikolaev D., Ingacheva A., Skoryukina N. Approach to the recognition of flexible forms on the example of the credit card date recognition. Proceedings SPIE 9875, Eighth International Conference on Machine Vision (ICMV 2015), Barcelona, Spain, 8 December 2015, session 3. P. 83–88.

[2] Christian T., Gustavsson D. Content recognition of business cards. Summer Project, IT University of Copenhagen, Copenhagen, Denmark, 5 September 2007.

[3] Арлазаров В.В., Булатов К.Б., Карпенко С.М. Метод определения надежности распознавания тисненых символов. Труды Института системного анализа РАН. 2013. Т. 63. № 3. С. 117–122.

[4] Puybareau E., Geraud T. Real-Time Document Detection in Smartphone Videos. Proceedings of the 24th IEEE International Conference on Image Processing (ICIP), France, October 2018. P. 1498–1502.

[5] Sermanet P., Eigen D., Zhang X., Mathieu M., Fergus R., LeCun Y. OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks. International Conference on Learning Representations (ICLR), Banff, Canada 14–16 April 2014. P. 1055–1061.

[6] Copyright 2011–2014, OpenCV dev team, 2019. OpenCV documentation / Image filtration. https://docs.opencv.org/2.4/modules/imgproc/doc/filtering.html#getgaussiankernel

[7] Mason S.J., Graham N.E. Areas beneath the relative operating characteristics (ROC) and relative operating levels (ROL) curves: Statistical significance and interpretation. / S.J. Mason, N.E. Graham // Q. J. R. Meteorol. Soc. 2002. V. 128. P. 2145–2166.