

APPLICATION OF OPTICAL MARK RECOGNITION TECHNIQUES TO SURVEY ANSWER SHEETS AT DALAT UNIVERSITY

Thai Duy Quy^{a*}, Phan Thi Thanh Nga^a, Nguyen Van Huy Dung^a

^aThe Faculty of Information and Technology, Dalat University, Lam Dong, Vietnam

**Corresponding author: Email: quytd@dlu.edu.vn*

Article history

Received: November 18th, 2020

Received in revised form: December 22nd, 2020 | Accepted: December 29th, 2020

Available online: February 5th, 2021

Abstract

In this paper, we examine some image processing techniques used in optical mark recognition, and then we introduce an application that collects data automatically from survey answer sheets at Dalat University. This application is constructed with the Aforge framework. Two types of survey answer sheets are used as input forms for our application: the teaching quality and the administrative quality survey answer sheets. Results show that our application has good performance in recognizing handwritten marks, with an accuracy of 98.9% per 667 answer sheets. Moreover, this application is clearly a time-saving solution for administrative staff because the inputting process is now nine times faster than before.

Keywords: Computer vision; Image processing; Optical mark recognition; Survey answer sheet.

DOI: [http://dx.doi.org/10.37569/DalatUniversity.11.1.791\(2021\)](http://dx.doi.org/10.37569/DalatUniversity.11.1.791(2021))

Article type: (peer-reviewed) Full-length research article

Copyright © 2021 The author(s).

Licensing: This article is licensed under a CC BY-NC 4.0

1. INTRODUCTION

Nowadays, automation techniques help to enhance the speed and efficiency of information processing and communication. Since their inception, automation techniques have undergone many development stages and have made great advances in technical and scientific calculations as well as in administrative management (Ngô & Đỗ, 2000). One of the focus areas for automation is image recognition, in which information is automatically retrieved from handwritten data. This technique is used in optical character recognition, optical mark recognition (OMR), invoice identification, postal code recognition, automatic map recognition, music recognition, face recognition, and fingerprint identification, etc. Each type of application has its own processing techniques based on the characteristics of the input data and serves different purposes in many areas of life. This article mainly explores and examines some techniques in optical mark recognition.

Optical mark recognition is a technique that uses a computer to retrieve data from handwriting or hand-filled answer sheets (Bergeron, 1998; Cip & Horak, 2011; Kumar, 2015; Popli et al., 2014; Surbhi et al., 2012; Yunxia et al., 2019). The technique is used for collecting information from surveys and answers to multiple choice questions. The technique can also be integrated with image scanners, which are specialized in scanning and identifying different types of answer sheets.

The OMR technique was invented in the 1960s by American scientists. IBM's computer systems were used to process questionnaires after images were scanned into the computer (Yunxia et al., 2019). Today, this technique has been researched and applied in many different fields, such as exam marking, timekeeping, survey evaluations, vote identification, etc. (Surbhi et al., 2012). The main concepts concerning the objects used in mark recognition, such as data areas, personal areas, and calibration points are discussed by Cip and Horak (2011). For effective optical mark identification, de Elias et al. (2019), Kumar (2015), and Surbhi et al. (2012) have proposed several general techniques, such as binary transformation, image rotation, and shifting. Yunxia et al. (2019) used a convolution neural network and the Tensorflow library to study identification methods for answer sheets with various characteristics.

Domestically, the OMR technique has been studied by Ngô and Đỗ (2000) by applying preprocessing techniques on images of the MarkRead system. Mai (2014) developed a recognition application used for survey answer sheets at the Vietnam National University of Forestry. In addition, some commercial identity systems have been built, such as TickREC and IONE. However, these versions are commercial and cannot be applied to the current survey questionnaires at Dalat University.

Currently, there are many identity libraries that run on different platforms (e.g., Python, C#, Java, etc.). However, in our research, we wanted to find libraries that support the .Net platform, as we aim to develop web applications later. Thus, we examined some .Net-based libraries, such as ImageProcessor, Csharp Image Library, EmguCV, and Aforge.Net. The ImageProcessor and Csharp Image Library simply process images

without convolution operations. The EmguCV library, developed from OpenCV, also supports image processing, but does not have strong built-in support for the convolution operations matrix. We examined the Aforge library and found that it is not only a free library that supports many techniques for image preprocessing, but that it also supports image convolution, which makes it suitable for our application.

2. METHODOLOGY

2.1. The survey answer sheets

We selected two types of survey answer sheets that are used at Dalat University, namely, the student survey on teaching quality and the student survey on the administration and departments (Figure 1). These answer sheets are much used each semester to help the university's teaching and administration become more effective. After receiving the students' answers, the staff must manually process the results in a Microsoft Excel file and then make a statistical summary based on the numbers. Due to the large number of survey answer sheets, this task is time consuming and boring.

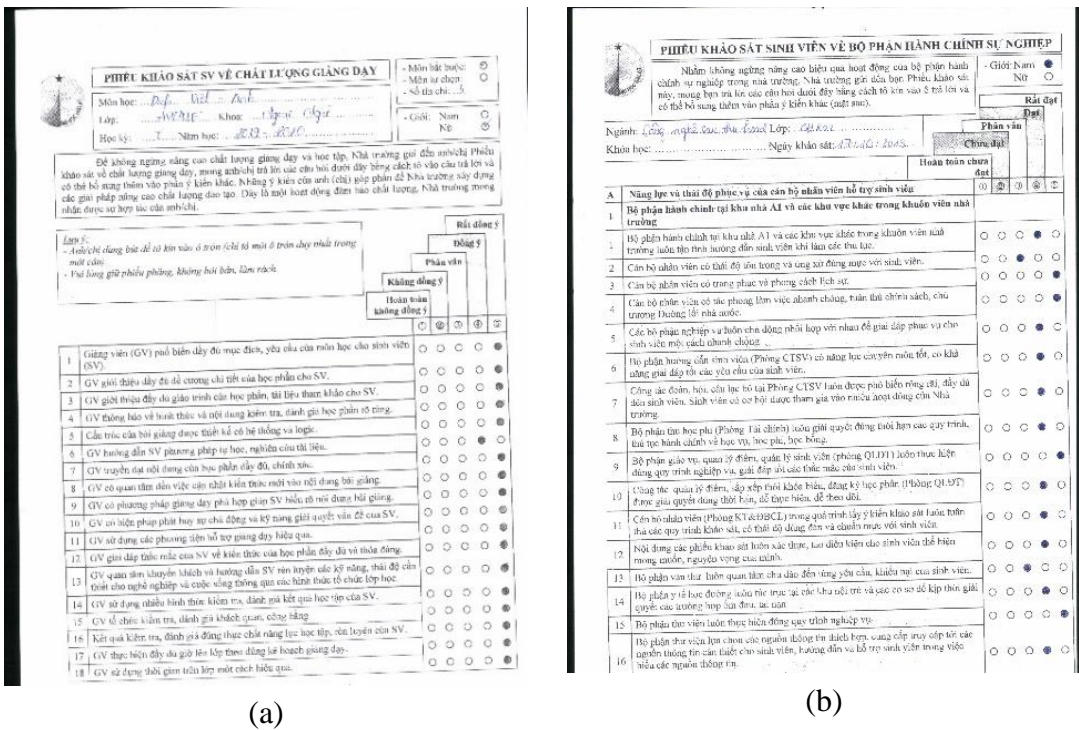


Figure 1. Two types of survey answer sheets used at Dalat University

Note: a) The student survey on teaching quality; b) The student survey on the administration and departments.

The survey answer sheets have two sides. The front side consists of a logo and personal information, followed by an introduction and then lines of questions. There are 23 questions on the front and back sides. A comment section is at the bottom of the back side. These surveys are not designed for machine reading, so there are no calibration points for optical mark recognition. After examining the answer sheets, we decided to

apply a number of convolution techniques for image preprocessing based on the characteristics of the scanned images. After the preprocessing, we continue by applying the OMR method to detect handwriting and to build an application.

2.2. Convolution techniques

Convolution is a technique of image processing used to transform the image matrix to a result matrix related to the original image. This technique is used in transformations on images, such as smoothing, boundary extraction, and filtering. The convolution formula is represented as follows:

$$k(x, y) * f(x, y) = \sum_{u=-m/2}^{m/2} \sum_{v=-n/2}^{n/2} k(u, v) f(x-u, y-v) \tag{1}$$

where $f(x,y)$ is an image matrix and $k(x,y)$ is a filter matrix with dimensions $(m \times n)$.

An important component in the convolution Equation (1) is the filter, which is called the kernel matrix. The filter's anchor point is located at the center of the matrix, and it determines the corresponding matrix area on the image for convolution (Kim, 2016). The convolution method moves the kernel matrix over the pixels around the anchor point, then calculates the result matrix with the convolution Equation (1) (Figure 2).

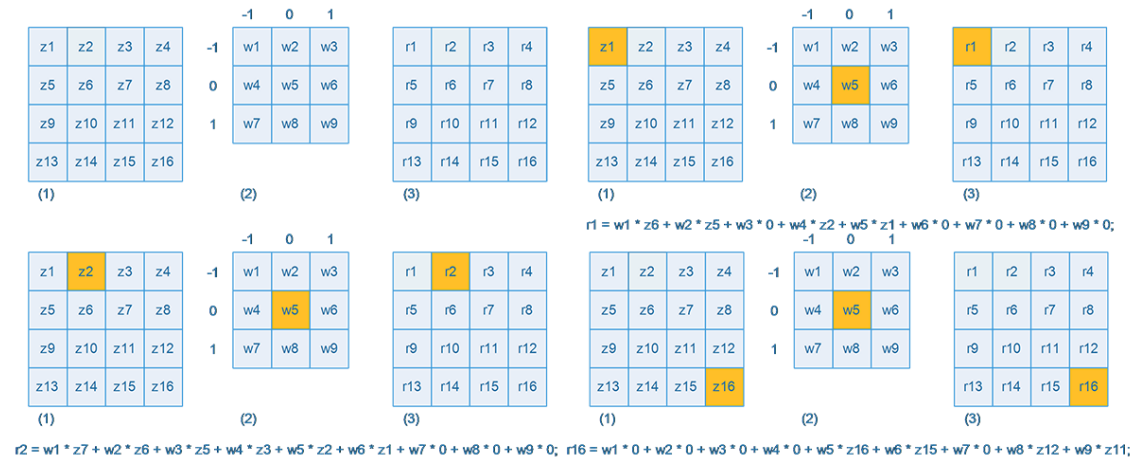


Figure 2. Convolution operation illustration

Source: Kim (2016).

2.3. Aforge platform

Aforge.NET, developed by Andrew Kirillov, is a platform support for the computer vision and artificial intelligence fields. The platform is designed for researchers and developers and is used for image processing, neuron networks, genetic algorithms, fuzzy logic, machine learning, and robotics. The platform also strongly supports programmers on .NET and complies with the GNU General Public License. Aforge.NET includes a set of online open source libraries. When users download to their computer,

they can simply add some *.dll files needed for their project. The powerful platform supports effective image processing and recognition with built-in convolution operations and basic pixel image methods.

3. RECOGNITION TECHNIQUES

3.1. Recognition diagram

Figure 3 shows a diagram of the OMR technique used in our application. The process includes the following steps: First, the answer sheets are converted to images and stored in the computer. Second, the scanned images are preprocessed to become binary images. After that, the application will determine the anchor points (also called calibration marks), which are located at certain positions on the binary image. The frame trimming step is then used to cut images by blocks based on the anchor points from the previous step. In the next step, the application uses a histogram to read the pixel image and recognize the hand-filled answers. Finally, statistical results are provided to the user.

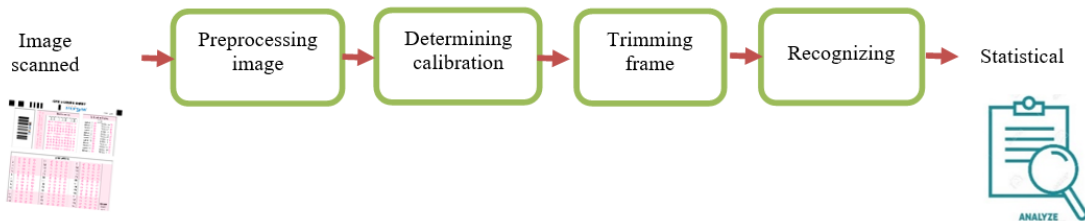


Figure 3. OMR technique diagram

3.2. Image preprocessing

Preprocessing of images is used to transform the image pixels before the recognition stage. For highly efficient and accurate recognition results, we apply several techniques, including image rotation, grayscale transformation, noise filtering, and image binarization.

- *Image rotation*: The scanning process may skew images, so the image must be rotated vertically before the recognition process. We rely on the Hough transform (Phan et al., 2017) to find the angle of inclination (θ), then rotate the image in the opposite direction ($-\theta$). This process makes the image upright and easy to identify in the next steps.
- *Grayscale image*: Grayscale is an image that has only two colors, black and white, with the colors represented by shades of gray from light to dark. We apply the transformation formula from Đỗ and Phạm (2007) to convert from color images to grayscale:

$$F(x,y) = \alpha R + \beta G + \gamma B \quad (2)$$

where the R, G, and B values represent red, green, and blue, respectively, and α , β , and γ have many possible values. According to Kumar (2015), the tuple ($\alpha = 0.2125$, $\beta = 0.7154$, and $\gamma = 0.0721$) is appropriate for mark recognition on multiple choice answer sheets. When applied to our program, we saw that Kumar's tuple gave better results than others.

- *Noise filtering*: Scanned images may have noise. To reduce this problem, we apply a convolutional filter with the median filter (Yang, 2006). This operation is supported by the Aforge library. This process helps our application reduce noise in the image, thereby increasing the accuracy of the recognition process.
- *Sharpen*: The sharpen convolution technique increases the accuracy of recognition by giving a sharper image. The kernel matrix of this method, according to Abraham (2020), is
$$\begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix}.$$
- *Image binarization*: Binarization is a process that transforms a pixel in grayscale to a pixel that has only two values: black (1) and white (0). The formula for the conversion is as follows:

$$g(x,y) = \begin{cases} 1 & \text{if } f(x,y) \geq T \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where $f(x, y)$ is a function that represents the value at the position (x, y) of the image, and T is the threshold that has values from 0 to 255. After experimenting with our application, we determined that a T value of 250 is suitable for clarifying pixels when the students make fuzzy marks or small strokes when filling in answers with pencils. This is the default value of our program. The user can change this parameter as desired when using the program.

3.3. Calibration mark recognition

According to Cip and Horak (2011), calibration marks are points used to locate position on the answer sheets. The calibration marks are usually placed at the corners and are a circle or square shape. Finding these points is the first step in the recognition process. From these points an application locates the position of the sheet, from which rows, columns, and cells can be determined and cut. This action is the basis for taking image areas, analyzing pixels, and recognizing data from the image pixels.

When observing two survey answer sheets of Dalat University, we saw that calibration marks have not been included, so it is difficult to determine the positioning points to cut and detect the blocks. To solve this problem, we proposed a method that determines location by using an edge detection technique. This is a convolution technique that applies a boundary detection matrix to define a horizontal or vertical line. Following Sinha (n.d), the kernel matrices for vertical and horizontal lines are

$\begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix}$ and $\begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$, respectively. These matrices are used in the convolution

method, which determines the nearest horizontal or vertical line of the scanned image from the top and the left side. The lines form a basis to determine the area of the image to be cropped for the next steps of the OMR process. When using the boundary detection technique, all the calibration points on the front and back side are determined at this time, so the image area can be cropped on both sides of the answer sheet.

3.4. Image cropping process

The image of a scanned survey answer sheet consists of three blocks: The first block includes personal information and instructions. The second block is the handwriting area consisting of questions and boxes for marking answers, and the final block is the area for the students' opinions. After determining the calibration point, the scanned image will be cut based on these three blocks (Figure 4).

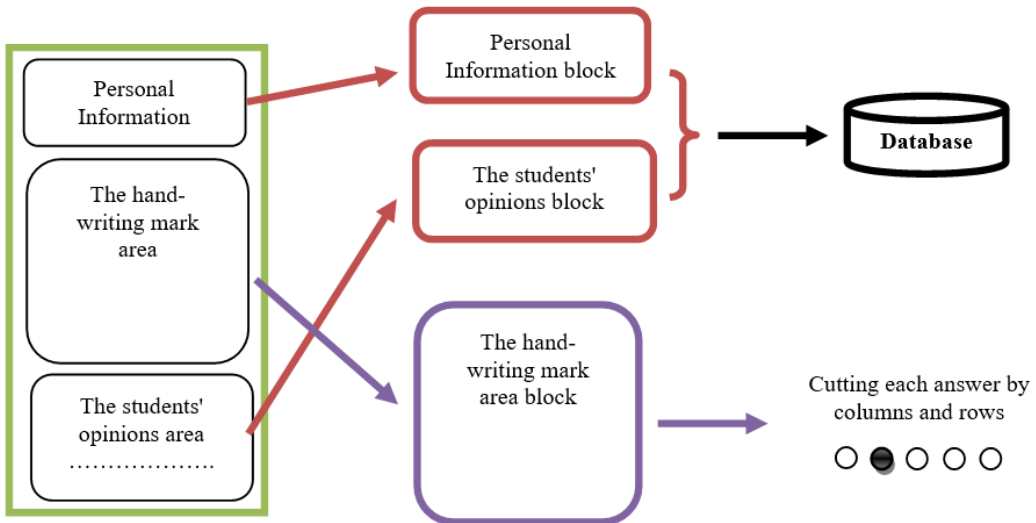


Figure 4. Cutting the three blocks of the scanned image

In some cases, the block is too small after cropping, so the software will zoom in to an appropriate size for more accuracy in the next steps. The blocks are cut by our application as follows:

- *Information and student's opinion blocks:* These blocks are cut according to the position determined by the calibration points and saved to the system. When the software displays the results of each image, the student's opinion block can be deleted if it is blank.
- *Handwriting block:* The handwriting block is also cut by positioning the image based on the calibration points. Both survey answer sheets have questions on both sides of the image. The first step is to cut the block that contains the

answers on each side of each image, then the application will cut each question and answer box by column and row. The student survey answer sheet on teaching quality has 18 questions on the front and 5 questions on the back, while the student survey answer sheet on the administration and departments has 16 questions on the front and 15 questions on the back. Each question on the two answer sheets has five answer options.

3.5. Recognizing image blocks

To recognize the handwritten marks in the answer blocks, we apply the histogram to the image of each answer box. This diagram depends on two colors: black and white. The main color used for comparison is black. We analyze the number of black pixels per answer block and compare it with the given threshold. Variable sbp is the total number of black pixels, and T is the threshold value to distinguish marked cells. If $sbp \geq T$, then the cell is read as marked by the student; otherwise the cell is read as not marked. Experimentation with our software determined that $T = 960$ is a suitable value to guarantee the accuracy of the recognition process (Figure 5).





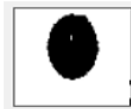
Hand-writing images					
Sbp value	521	524	520	506	1355
Results	False	False	False	False	True

Figure 5. Example of a filled-in answer mark by a student

4. EXPERIMENTATION RESULTS

The program was built in the C# language based on the Aforge foundation. The program allows inputting image files for two types of survey questionnaires, the student survey of teaching quality and the student survey of the administration and departments. The output of the program is displayed by batch number, according to the answer sheets and questions. Although most answers from the survey sheets are recognized accurately, there are still some recognition errors due to bad input, which the program will indicate with a red x. For instance, these problems may come from noisy scanned images as the filled-in answers may be too fuzzy or double-marked on one answer. After the results are displayed, the program allows the user to review, remove, or edit incorrect questions. This action only allows correcting the data; the picture of the scanned images is unchanged. The program also allows users to delete blocks of user opinions if they are blank. The last function of our program is to provide statistics of the results in the form of graphs or files for export to Microsoft Excel (Figure 6).

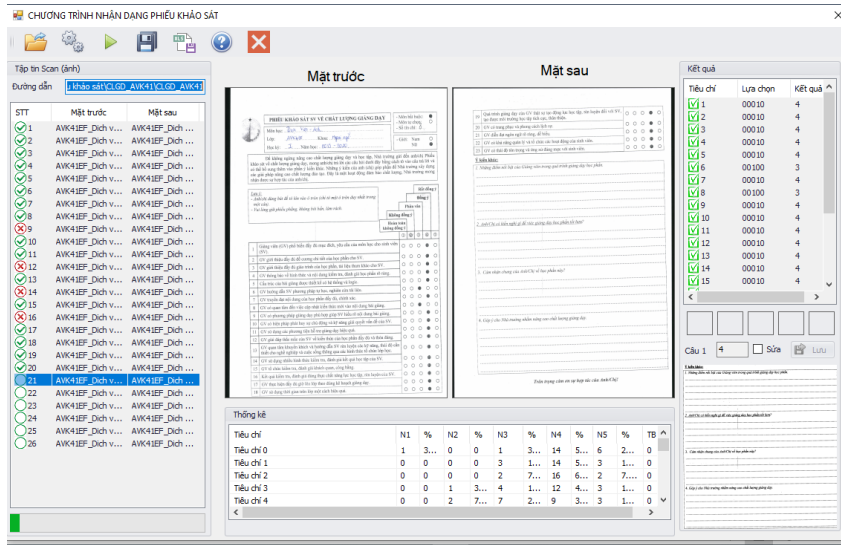


Figure 6. Experiment program

We used 677 survey answer sheets provided by the Quality Assurance and Testing department for the second term of the 2019-2020 school year. The sheets are classified and grouped by class and faculty. Due to security reasons, we used the concept of Lot instead of the class name. Survey files were scanned and the size of each image was 2,550 x 3,300 pixels. The experimental results showed that 98.9% of the images were correctly recognized. There was some incorrect recognition because of noise in the scanning process (Figure 7a) or because of a large image angle. In addition, there were many questionnaires that were invalid because students did not fill in an answer or filled in more than one answer per question (Figure 7b). The results of the program are given in Tables 1 and 2.

Table 1. Results of the student survey on teaching quality

Lot	Quantity	Recognition results		Timing (seconds)
		Invalid sheets	Invalid responses	
L1	26	3	3	156
L2	46	6	10	276
L3	29	3	5	174
L4	76	11	11	456
L5	113	27	29	678
L6	75	9	15	450
L7	23	3	5	138
L8	28	2	2	168
L9	27	2	4	162
L10	32	5	7	192
Total	475	71	91	2,850

Table 2. Results of the student survey on the administration and departments

Lot	Quantity	Recognition results		Timing (seconds)
		Invalid sheets	Invalid responses	
L1	26	3	8	156
L2	46	6	10	276
L3	29	3	9	174
L4	76	11	11	456
L5	25	27	30	150
Total	202	50	68	1,212

**Figure 7. Examples of invalid responses**

Notes: a) Image has noise; b) Invalid answer.

Tables 1 and 2 show that the total time for processing 677 survey answer sheets was 4,062 seconds. When added to the time to process the incorrect results (assuming each incorrect result takes 3 seconds), the total processing time is 4,539 seconds. The total input time for the staff, assuming that each form takes 60 seconds, is 40,620 seconds. Thus, using the software will be about 9 times faster, not including the time for sorting the survey answer sheets and calculating the statistics.

5. CONCLUSION

In this article, we have handled the recognition of survey answer sheets by applying a number of convolution image processing techniques, such as edge detection, noise filtering, and image sharpening. Combining the convolution operations with an optical mark reader, we have built a recognition program and have automatically read two types of answer sheets used at Dalat University. The program reads faster than manual input, gives accurate results, and allows erroneous results to be corrected quickly.

In the future, we will improve the program in a general way to read more types of forms. This improvement helps increase work efficiency for university staff. We also propose to redesign the survey sheets with calibration points at the four corners for easier and more convenient reading.

REFERENCES

Abraham, S. (2020). *Image manipulation: Filter and convolutions*. <https://www.cs.utexas.edu/~theshark/courses/cs324e/lectures/cs324e-6.pdf>.

- Bergeron, B. P. (1998). Optical mark recognition. *Postgraduate Medicine*, 104(2), 23-25.
- Cip, P., & Horak, K. (2011). *Concept for optical mark processing*. Paper presented at the 22nd International DAAAM Symposium, Austria.
- de Elias, E. M., Tasinaflfo, P. M., & Junio, R. H. (2019). *Alignment, scale and skew correction for optical mark recognition documents based*. Paper presented at the 2019 XV Workshop de Visão Computacional (WVC), Brazil.
- Đỗ, N. T., & Phạm, V. B. (2007). *Xử lý ảnh*. Trường Đại học Thái Nguyên.
- Kim, U. (2016). *Phép tích chập trong xử lý ảnh (convolution)*. [https:// www.stdio.vn /computer-vision/phep-tich-chap-trong-xu-ly-anh-convolution-r1vHu1](https://www.stdio.vn/computer-vision/phep-tich-chap-trong-xu-ly-anh-convolution-r1vHu1)
- Kumar, S. (2015). A study on optical mark readers. *International Interdisciplinary Research Journal*, 3(11), 40-44.
- Mai, H. A. (2014). Nghiên cứu ứng dụng kỹ thuật xử lý ảnh vào xử lý phiếu đánh giá môn học Trường Đại học Lâm nghiệp. *Tạp chí Khoa học và Công nghệ Lâm nghiệp*, (1), 141-146.
- Ngô, Q. T., & Đỗ, N. T. (2000). Một số phương pháp nâng cao hiệu quả nhận dạng phiếu điều tra dạng dấu phục vụ cho thiết kế hệ nhập liệu tự động MarkRead. *Tạp chí Tin học và Điều khiển học*, 16(3), 65-73.
- Phan, T. T. N., Nguyen, T. H. T., Nguyen, V. P., Thai, D. Q., & Vo, P. B. (2017). Vietnamese text extraction from book covers. *Dalat University Journal of Science*, 7(2), 142-152.
- Popli, H., Parekh, H., & Sanghvi, J. (2014). *Optical mark recognition*. <https://www.slideshare.net/HimanshuPopli/optical-mark-recognition-40292822>
- Sinha, U. (n.d). *Image convolution examples*. <https://aishack.in/tutorials/image-convolution-examples>.
- Surbhi, G., Geetila, S., & Parvinder, S. S. (2012). *A generalized approach to optical mark recognition*. Paper presented at the International Conference on Computer and Communication Technologies (ICCCT'2012), Thailand.
- Yang, Y. (2006). *Image filtering: Noise removal, sharpening, and deblurring*. http://eeweb.poly.edu/~yao/EE3414/image_filtering.pdf
- Yunxia, J., Xichang, W., & Xichang, C. (2019). *Research on OMR recognition based on convolutional neural network Tensorflow platform*. Paper presented at the International Conference on Measuring Technology and Mechatronics Automation (ICMTMA), China.