An Android Business Card Reader Based on Google Vision: Design and Evaluation

Nguyen Hoang Thuan[1(✉)], Dinh Thanh Nhan[1], Lam Thanh Toan[1], Nguyen Xuan Ha Giang[1], and Quoc Bao Truong[2]

[1] Can Tho University of Technology, 256 Nguyen Van Cu Street, Can Tho City, Vietnam
{nhthuan, dtnhan, lttoan, nxhgiang}@ctuet.edu.vn
[2] College of Engineering Technology, Can Tho University, 3/2 Street, Can Tho City, Vietnam
tqbao@ctu.edu.vn

Abstract. Business cards have been widely used to greet business professionals and exchange contact information. However, the current paper-based way to manage business cards impedes their effective usage, leading to a need for digitalising and extracting business card information. This paper aims to design a business card reader (BCR) application for Android devices. Based on Google vision library, the application digitalises and extracts business card information. We evaluate the application on a dataset of 170 business cards. The results show that the application can digitalise business cards and extract contact information with 88.4% of accuracy. We then further conduct a comparative analysis of our application and other commercial BCR applications. Based on the results, the paper suggests several recommendations for future research.

Keywords: Business card reader · Android · Design science · Experiment

1 Introduction

Today, the exchange of business cards become a norm to start new business connections. Business cards help self introduction and provide contact information for future correspondence. In the current practice, business card receivers manage business cards manually, e.g. manually input contact information into a cell phone, which will take little effort to utilise business card information [1]. Further, managing the contact information manually leads to a risk of input errors. Thus, it will be more convenient to extract business card information automatically without any typing activities, which leads to the need of digitalising and extracting business card information [1–3].

From a user perspective, digitalising and extracting information from business cards brings the following benefits. First, it saves time and efforts to transfer paper-based business cards into contact information [2, 4]. Second, it supports business card receivers to manage and search particular contacts since information are stored in digitalised form. The searching capability is especially important in cases where too many business cards are received. Third, it is easier to access and back up business card information in digitalised form than a paper-based form. Finally, when users are on

P. C. Vinh and A. Rakib (Eds.): ICCASA 2019/ICTCC 2019, LNICST 298, pp. 223–236, 2019.
https://doi.org/10.1007/978-3-030-34365-1_18

movement (e.g. business travel), digitalised business cards provide flexibility and convenience in comparison to a bunch of paper-based business cards.

These benefits have motivated an increased research interest in business card reader (BCR) applications. By and large, the literature related to BCR applications can be classified into two research streams, which focus on problems related to computer science (CS) and information systems (IS). In the computer science domain, researchers focus on optical character recognition (OCR) to identify business cards' texts at the technical levels [3, 5, 6]. They proposed techniques and algorithms to process business card images [7], find text blocks [8], and extract texts [3, 4]. These algorithms and techniques provide a foundation supporting the development of BCR applications.

From a different perspective, the IS domain focuses on the design, adoption, use of BCR applications at the business levels. Examples of IS problems include design BCR applications on certain mobile platforms [1, 2, 6], management of semantics, and evaluation of the designed BCR applications [9, 10]. Consequently, several BCR applications have been designed and evaluated. They are based on several computer vision libraries, including Mobile Vision, OpenCV and Tesseract [2, 9, 10]. However, other computer vision libraries like Google Vision exist, which have not been used to design BCR applications. Further, the evaluation so far is limited in terms of sample size, e.g. experimenting on less than 55 business cards [2, 6, 10].

This paper adopting an IS perspective concerns the design and evaluation of BCR applications. In particular, we are interested in designing an Android BCR application based on Google Vision [11] and Image Cropper [12], developing algorithms to extract business card information, and then evaluating the application on a sample of 170 business cards. This problem is challenging: while Google Vision can detect and extract texts from business card images with high accuracy, the extracted texts have no semantics. Thus, our design needs to add meaning to the extracted texts and transfer them to useful contact information (i.e. name, phone number, and email address). This problem is relevant because the proposed solution creates opportunities for further evaluating BCR applications.

In particular, we aim to address the following goals:

– Enable users to digitalise business cards and extract contact information from business card pictures. For the purpose, we develop an Android BCR application that extracts business card texts, adds semantics to the extracted data and transfers them to useful information;
– Evaluate the developed application to measure the accuracy of extracted information. Here, we aim to evaluate the application using a reasonable sample size, which fulfils the current research gap where the related research is limited in its sample size in evaluating BCR applications [2, 6, 10].

The current study adopts a design science research (DSR) paradigm to design and evaluate an Android BCR application. DSR focuses on identification of a relevant problem, design artefacts to address the problem, evaluation to assess the artefact's utility [13–15]. Translating into the current research, we address the problem of digitalising business cards, design an Android BCR application, and evaluate the application through the sample size of 170 business cards. We further conduct a comparative analysis to assess the capabilities of our application against other commercial BCR applications (e.g. ABBYY and CamCard).

This research contributes to explore the design and adoption of BCR applications. It contributes to fulfill the need to digitalise and extract business card information. The research designs a BCR application that allows end users to extract and manage business card information. The design extends the research stream of BCR development [2, 6, 10] by using alternative computer vision libraries, i.e. Google Vision and Image Cropper. Further, the research also addresses the problem of the evaluation using a reasonable sample size of 170 business cards and using comparative analysis, which illustrates the usefulness of the BCR application.

2 Literature Review

Business card readers (BCR) aim to transfer business card pictures into useful information using optical character recognition (OCR). At the early time, BCR were researched as long as the development of OCR technology, and thus belong to the research stream of computer science. Saiga et al. [16] proposed an OCR system for reading business cards based on line segmentation and linguistic processing. In the same vein, Luo et al. [17] developed an OCR method based on multi-resolution analysis of document images to implement a BCR application. To improve the accuracy of optical character recognition, several algorithms have been developed. For instance, Lin [3] used neural network to extract texts from business card pictures while Mollah et al. [7] used fuzzy-based algorithms to improve OCR accuracy. All in all, this research stream contributes to the improvement of OCR capabilities to recognise texts from business card pictures, which, to some extent, founds the development of BCR applications.

Recently, the BCR research area has gained momentum with the availability of several computer vision libraries. Examples of qualified computer vision libraries include Tesseract OCR, Mobile Vision, OpenCV, and Google Vision. This availability has shifted the BCR research focus from CS domain into the IS domain, where research concerns how to utilise the computer vision libraries to develop BCR applications and assess their usage in practice.

In the IS domain, researchers have developed and evaluated many BCR applications based on available computer vision libraries. Dangiwa and Kumar [2] based on Tesseract library built a BCR application that extracted names, phone numbers, and email addresses for iOS devices. They evaluated the application using a sample of 55 business cards with the accuracy of 74%. Also using available libraries, Dat [10] combined OpenCV and Tesseract to develop an Android BCR application. The application was tested on a sample of 50 business cards with the accuracy of 81.1%. Other BCR applications have been developed based on Tesseract and Mobile Vision [6, 9].

Table 1 summarises the related research using available computer vision libraries for developing BCR applications.

Table 1. Summary of related research

Related research	Library	Experimental sample size	Accuracy
[9]	Mobile Vision	4	Name: 100% Phone number: 100% Email address: 50%
[10]	OpenCV & Tesseract	50	Name: 80.32% Phone number: 83% Email address: 80.04%
[2]	Tesseract	55	Name: 53.8% Phone number: 100% Email address: 83.3%
Current research	Google Vision & Image Cropper	>100	To improve accuracy level

We further position the current research into Table 1 in order to show how it extends the related literature. In the current research, we adopt an alternative library using a combination of Google Vision and Image Cropper to construct an Android BCR application, namely CTUT_NameCard. We further note that the related research is limited in the sample size of evaluating BCR applications (less than 55 business cards), which impedes an overall assessment of the BCR applications. The current research addresses this research gap by increasing the sample size for evaluating the CTUT_NameCard application. Further, our research also aims to improve the accuracy level of extracting business card information.

The paper proceeds as follows. In Sect. 3, we propose an architecture backing the development of CTUT_NameCard application. Section 4 introduces the design of CTUT_NameCard application, while Sect. 5 experiments the application on a dataset of 170 business cards and then compares it with other commercial BCR applications. In Sect. 6, we discuss the findings, their implications, and outline future work.

3 Architecture

In this section, we propose an architecture for building CTUT_NameCard application. The architecture is based on the work by Holsapple [18], which seems to clearly separate the major application components [19]. Figure 1 shows the architecture which relies on three components: GUI, library, and semantic.

The GUI component manages the interaction between the CTUT_NameCard application and its users, i.e. business card receivers who want to digitalise business cards. This component accepts business card pictures as its inputs through two means. The component allows users to take pictures of business cards by cell phone cameras. Alternatively, users may want to load business card pictures that are stored in the cell phone library. These inputs are processed by the library component.

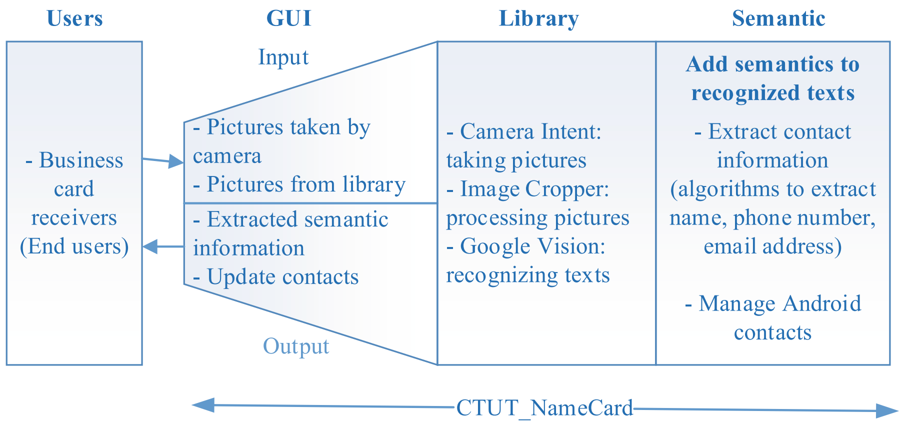


Fig. 1. Architecture for CTUT_NameCard (adapted from [18])

The Library component consists of three libraries founding the development of the CTUT_NameCard application: Android camera intent, Image Cropper, and Google Vision. Android camera intent provides functionalities for cell phones to capture pictures. In CTUT_NameCard, the component takes pictures of business cards and thus uses mainly back camera. The component then calls Image Cropper library to process the captured pictures. Image Cropper allows users to choose, flip, rotate, and crop business card images for eliminating noise background and focusing the areas that include business card information. Then, Google Vision is used to recognise texts from the business card images. This library returns a bunch of texts, which are normally high accuracy yet have no semantics. This bunch of texts feed the semantic component.

The Semantic component adds meaning to the extracted texts. In CTUT_NameCard, this component develops algorithms to classify business card information (e.g. name, phone number, and email address). After classifying business card information, this component enables users to update their contacts. As a result, it provides semantics for the extracted texts and manage Android contacts, which are presented as GUI's outputs.

4 Design and Implementation

This section presents the design and implementation of CTUT_NameCard application. We start with an overview on the workflow of the application that involves five main activities. Figure 2 shows the application workflow.

First, business card receivers (users) access the application and select how to load business card images. They can either use cell phone's camera to take a picture of business cards or load an existing business card image from Gallery. As a result, the business card image is loaded. Second, users can crop the loaded image for focusing on the main content of the picture, which removes noise background. If the image is vertical, users can choose to rotate it in order to support the text recognition in the next activity.

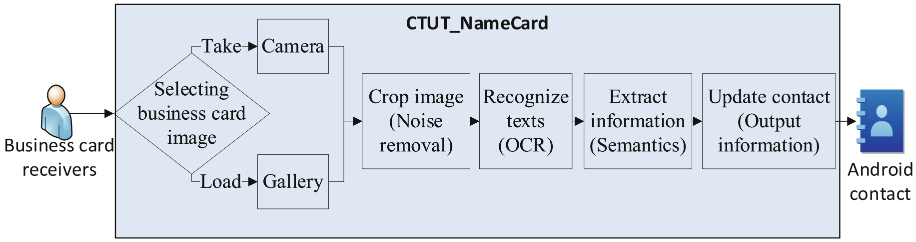


Fig. 2. Workflow of CTUT_NameCard

Third, CTUT_NameCard recognises texts on the pictures. It calls Google Vision for identifying and recognising texts, which return a bunch of texts without semantics. Fourth, we add semantics to the recognised texts. More precisely, we develop three algorithms to extract contact information regarding name, phone number, and email address from the recognised texts. Finally, the extracted information is used to update the users' contact list.

4.1 Implementation

The CTUT_NameCard application was developed using Java and Android Studio. The graphical user interface of the application is shown in Figs. 3, 4, and 5.

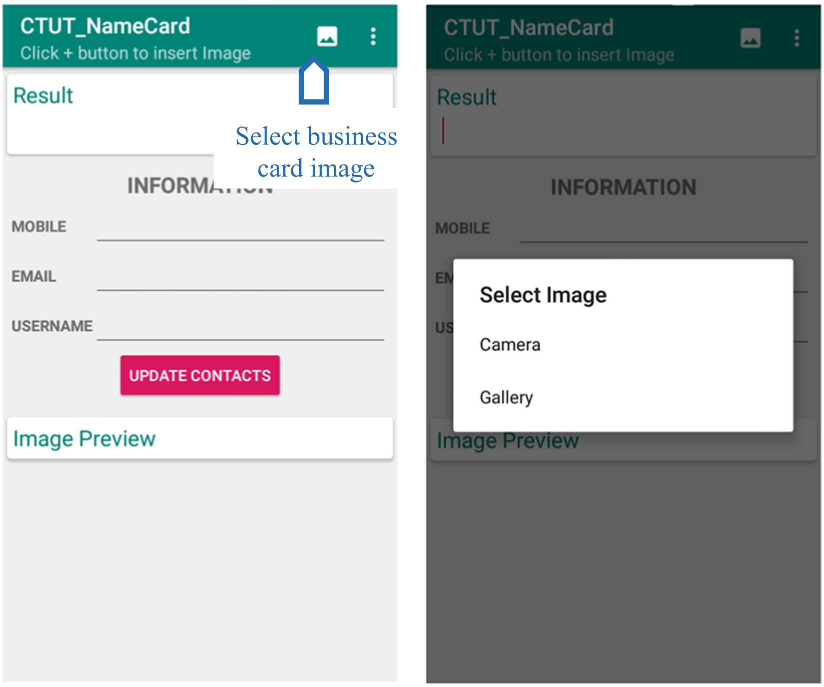


Fig. 3. CTUT_NameCard: Home screen (left) and Select image (right)

In particular, Fig. 3 (left-hand side) shows the home screen of the application, where users can select business card images in order to digitalise and extract information. In the right-hand side of Fig. 3, users can choose taking pictures of business cards or loading pictures from gallery. As a result, business card images are loaded into the application, ready for the next processing functions.

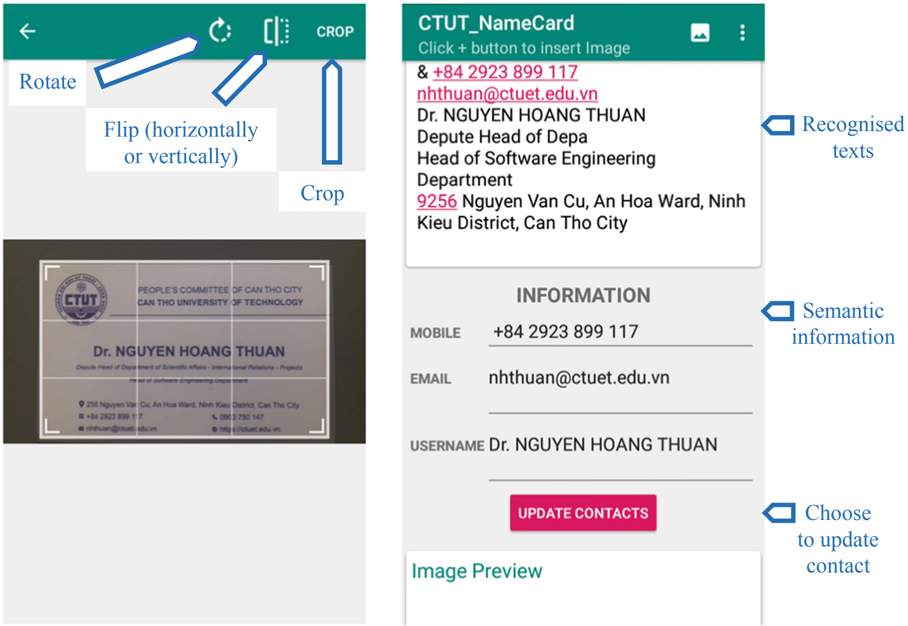


Fig. 4. CTUT_NameCard: Crop (left) and OCR and Sematic screen (right)

Figure 4 (left-hand side) shows the Crop screen. The application auto-detects the focus of the business card image and shows a cropping grid surrounding the image. Users may adjust this grid to remove noise background. They may also rotate, flip, and crop images, which will increase the accuracy of recognising texts.

With the cropped business card image, the OCR and semantic functions are operated, and their outcomes are displayed on the right-hand side of Fig. 4. The top part shows the OCR outcome that recognises texts from the business card image. The recognised texts are presented line by line without semantics. The bottom part shows the semantic outcome that adds meaning to the recognised texts. More precisely, the recognised texts are extracted and classified into useful information, i.e. phone number, email address, and name, which are ready for updating Android contacts.

Figure 5 shows the update contact screen where the extracted information is transferred into Android contacts. Here, the user can check and modify the extracted contact information before updating (left-hand side of Fig. 5). The updated outcome is shown in the right-hand side of Fig. 5.

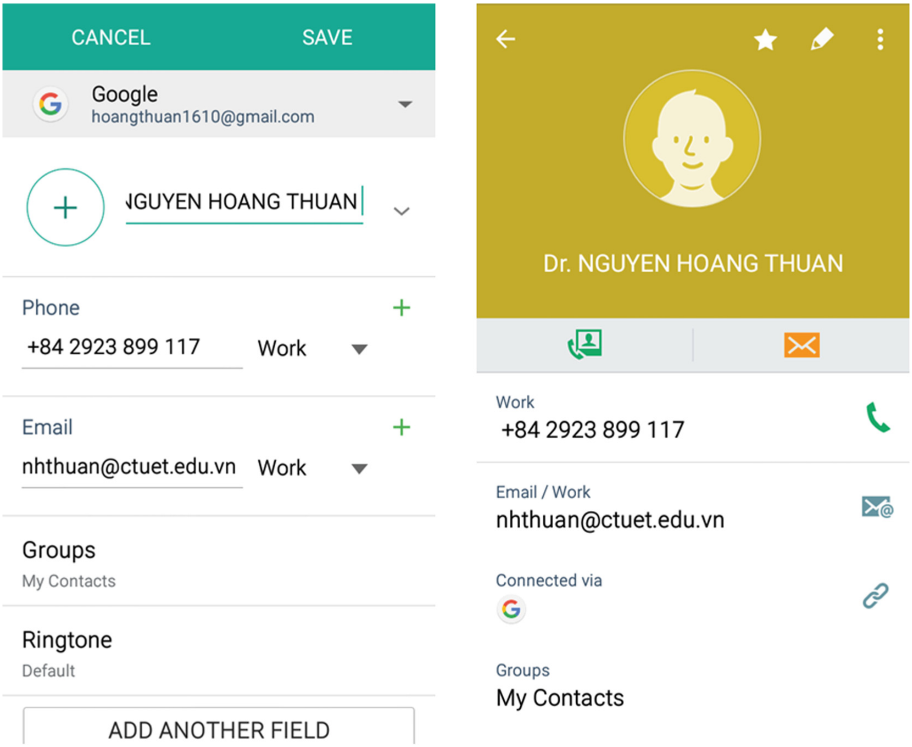


Fig. 5. CTUT_NameCard: Update contact

5 Evaluation

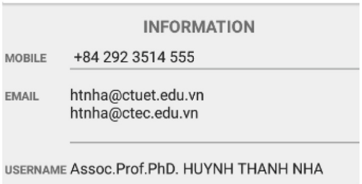
Experiments were arranged to evaluate the CTUT_NameCard application for its accuracy. The evaluation was conducted in two different phases: analytical assessment and comparative assessment.

5.1 Analytical Assessment

In the analytical assessment, we analysed how the application meets its designed target of extracting business card information. For the purpose, we tested the application using a sample of 170 business cards. The sample was collected through a convenient approach, where the business cards were collected through the authors' relationships. For each business card, we used the application to digitalise it and then consider whether the application correctly extracts its contact information: name, phone number, and email address.

Table 2 presents an example of the assessment results. In this table, the first column shows a picture of the business card, digitalised by the application. The second column shows the application's screen that extracts business card information. The third column shows the evaluation results of extracting name, phone number, and email address. If the extracted information is correct, the column value is ‘T’. Otherwise, it is ‘F’.

Table 2. An example of the assessment results

Business cards	**Information digitalised and extracted by CTUT_NameCard**	**Evaluation results (True/False)**		
		Name	Phone number	Email address
	CANTHO UNIVERSITY OF TECHNOLOGY 256 Nguyen Van Cu Street, An Hoa Ward, Ninh Kieu District, Can Tho City, Viet Nam ms CTUT Tel: +84 292 3514 555 Mobile: +84 919 209 555 Fax:+84 292 3894 103 Meritorious Teacher, Senior Lecturer Assoc.Prof.PhD. HUYNH THANH NHA Rector Website: w.ctuet.edu.vn E-mail: htnha@ctuet.edu.vn htnha@ctec.edu.vn 	T	T	T

After testing the application on the sample of 170 business cards, we calculated the positive rate and negative rate. The positive rate is defined as proportion of correctly digitalising and extracting contact information. The negative rate is defined as proportion of falsely digitalising and extracting contact information. Table 3 presents the evaluation results of CTUT_NameCard.

Table 3. Evaluation results of CTUT_NameCard

Assessment		Number of business card (n = 170)	Percentage
Name	Positive rate	131	77.1%
	Negative rate	39	22.9%
Phone number	Positive rate	159	93.5%
	Negative rate	11	6.5%
Email address	Positive rate	161	94.7%
	Negative rate	9	5.3%

As seen via Table 3, CTUT_NameCard identified and extracted business cards' name, phone number, and email address with the positive rate of 77.1%, 93.5%, and 94.7% respectively. On average, the accuracy of digitalising and extracting contact information is 88.4%. We note that the positive rates of extracting phone numbers and email addresses are higher than the positive rate of extracting names. This is because phone numbers and email addresses have clear semantic patterns. For instance, if recognised texts include the character '@', they are mostly identified as email addresses. On the other hand, names do not have clear semantic patterns, which can be used to differentiate them and other information on the business cards like company's names and addresses.

To further interpret the evaluation results, Table 4 presents the evaluation results of the current study with the evaluation of the related studies. Table 4 also adds the average levels of accuracy regarding each research, which shows that the current research yields highly reasonable level of accuracy in digitalising and extracting contact information (88.4%).

Table 4. Evaluation results of CTUT_NameCard and related research

Related research	Library	Experimental sample size	Accuracy
[9]	Mobile Vision	4	Name: 100% Phone number: 100% Email address: 50% **Average: 83.0%**
[10]	OpenCV & Tesseract	50	Name: 80.3% Phone number: 83.0% Email address: 80.0% **Average: 81.1%**
[2]	Tesseract	55	Name: 53.8% Phone number: 100% Email address: 83.3% **Average: 79.0%**
CTUT_NameCard	Google Vision & Image Cropper	170	Name: 77.1% Phone number: 93.5% Email address: 94.7% **Average: 88.4%**

To provide another view on the evaluation results, Table 5 shows the number of business cards that CTUT_NameCard correctly identified three, two, one, zero fields of contact information. Out of 170 business cards, the application correctly identified all three information fields in 122 cards (72%), two information fields in 38 cards (22%), one information field in 10 cards (6%). There is no case that the application could not identify any information field.

Table 5. Evaluation results: three, two, one fields of contact information

Assessment	Number of business card (n = 170)	Percentage
All three fields of contact information	122	72%
Two fields of contact information	38	22%
One field of contact information	10	6%
No field of contact information	0	0%

From these results, we suggest the application has met its designed goals for digitalising and extracting business card information with high level of accuracy.

5.2 Comparative Assessment

While the above assessment has showed utility of the CTUT_NameCard, we note that the evaluation results may be biased, given that it used a convenient sample approach. To overcome this limitation, we now evaluate the application using a sample provided by a third party.

For the purpose, we conducted a comparative assessment. We used a dataset provided by a third party in order to compare our application with two commerce BCR applications, i.e. CamCard and ABBYY. We installed CamCard, ABBYY, and CTUT_NameCard in the same smartphone, and used them to digitalised 50 business cards from the Stanford mobile visual search dataset [20]. We checked the accuracy of digitalising and extracting contact information regarding each application on the given dataset. Table 6 shows the comparative assessment results.

Table 6. Comparative results: CamCard, ABBYY, and CTUT_NameCard

BCR applications	Information field	Number of business card (n = 50)	Positive rate
CamCard	Name:	36	72%
	Phone number:	41	82%
	Email address:	36	72%
	Average:		**75.3%**
ABBYY	Name:	38	76%
	Phone number:	42	84%
	Email address:	35	70%
	Average:		**76.7%**
CTUT_NameCard	Name:	26	52%
	Phone number:	43	86%
	Email address:	42	84%
	Average:		**74%**

From this table, we suggest that on average CTUT_NameCard performs as well as other commerce BCR applications. In particular, it achieves the positive rate of 74%, while CamCard and ABBYY achieve a little higher positive rate of 75.3% and 76.7%

respectively. The results show that CTUT_NameCard extracts phone numbers and email addresses with higher levels of accuracy than the other applications. However, it extracts names with lower levels of accuracy, which suggests us to improve this feature in the future.

6 Conclusion

BCR applications are developed to allow the business card receivers digitalising and extracting contact information from the paper-based business cards. This study designed and evaluated CTUT_NameCard to digitalise and extract contact information with ease of use and accuracy. We designed the application based on Google Vision and Image Cropper. Inheriting the strength of Google Vision to recognise texts, we integrate several semantic algorithms to extract useful information, i.e. phone number, email address, and name. The application was evaluated in two experiments. The first experiment used the application to digitalise 170 business cards, while the second experiment analytically compared the applications with CamCard and ABBYY. The evaluation results have showed that the application met its design goals of digitalising and extracting contact information with high level of accuracy.

This study fulfils the need to digitalise and extract business card information. As seen via Table 1, the related research has addressed this need by designing BCR applications based on Mobile Vision, OpenCV and Tesseract [2, 9, 10]. We extend this stream of research by designing an BCR application based on Google Vision. Our work suggests that Google Vision is an important library for digitalising and extracting business card information.

Another contribution of the current study is its empirical evaluation. The evaluation on the dataset of 170 business cards has showed that the application has achieved high level of accuracy in digitalising and extracting contact information (88.4%). When positioning the empirical results within the related research, we note that the domain has a shortage of sample size in evaluating BCR applications (e.g. evaluating with less than 55 business cards [2, 6, 10]). This research addresses the above shortage and shows the evaluation results on a larger sample of 170 business cards.

From a practical viewpoint, the paper provides an Android BCR application for digitalising and extracting business card information. This BCR application has been starting used in practice. The comparative assessment has showed that the application performs as well as other commercial BCR applications (Table 6).

Our work presents several opportunities for future work. First, future work can improve our semantic algorithms in order to increase the levels of accuracy for extracting contact information, e.g. extracting names. Second, during our work, we note that many business cards use symbols for presenting information. For instance, the symbol ☎ indicates information of phone number. Consequently, our next version of the application targets to identify these symbols for increasing the ability of recognising and extracting information. Third, future work should also develop BCR application for multiple languages, like both English and Vietnamese [6]. Finally, while two experiments were conducted to test the application, they were conducted in the lab environment. Thus, future work should bring actual users into the experiments. Their feedbacks are really important to improve the application.

Acknowledgement. We would like to thank Phi Thi Ngoc Minh for helping us taking business card pictures and compare business card data and data extracted by the application.

References

1. Hing, V., Khoo, H.K.: Business card reader with augmented reality engine integration. In: Ibrahim, H., Iqbal, S., Teoh, S.S., Mustaffa, M.T. (eds.) 9th International Conference on Robotic, Vision, Signal Processing and Power Applications. LNEE, vol. 398, pp. 219–227. Springer, Singapore (2017). https://doi.org/10.1007/978-981-10-1721-6_24
2. Dangiwa, B.A., Kumar, S.S.: A business card reader application for iOS devices based on Tesseract. In: 2018 International Conference on Signal Processing and Information Security (ICSPIS). IEEE (2018)
3. Lin, L., Tan, C.L.: Text extraction from name cards using neural network. In: Proceedings of 2005 IEEE International Joint Conference on Neural Networks. IEEE (2005)
4. Madan Kumar, C., Brindha, M.: Text extraction from business cards and classification of extracted text into predefined classes. Int. J. Comput. Intell. IoT **2**(3), 595–602 (2019)
5. Mollah, A.F., et al.: Text/graphics separation for business card images for mobile devices. In: IAPR International Workshop on Graphics Recognition, pp. 263–270 (2009)
6. Hung, P.D., Linh, D.Q.: Implementing an android application for automatic Vietnamese business card recognition. Pattern Recogn. Image Anal. **29**(1), 156–166 (2019)
7. Mollah, A.F., Basu, S., Nasipuri, M.: Text detection from camera captured images using a novel fuzzy-based technique. In: 2012 Third International Conference on Emerging Applications of Information Technology. IEEE (2012)
8. Mollah, A.F., et al.: Text region extraction from business card images for mobile devices. In: Information Technology and Business Intelligence (2009)
9. Duy, N.: Xây dựng ứng dụng nhật diện danh thiếp trên Android. Can Tho University: in Vietnamese, Bachelor thesis (2017)
10. Đat, N.T.: Ứng dụng nhật diện danh thiếp tiếng Việt trên Android. Da Lat University: in Vietnamese (2018)
11. Google. Google Vision, 2017 Jun 2019. https://developers.google.com/android/reference/com/google/android/gms/vision/text/package-summary
12. Arthur. Android Image Cropper, 2016 June 2019. https://github.com/ArthurHub/Android-Image-Cropper
13. Hevner, A., Chatterjee, S.: Design Research in Information Systems: Theory and Practice. Integrated Series in Information Systems, vol. 22. Springer, Heidelberg (2010). https://doi.org/10.1007/978-1-4419-5653-8. Ed. R. Sharda and S. Voß
14. Hevner, A., et al.: Design science in information systems research. MIS Q. **28**(1), 75–105 (2004)
15. Thuan, N.H., Drechsler, A., Antunes, P.: Construction of design science research questions. Commun. Assoc. Inf. Syst. **44**(1), 332–363 (2019)
16. Saiga, H., et al.: An OCR system for business cards. In: Proceedings of 2nd International Conference on Document Analysis and Recognition (ICDAR 1993). IEEE (1993)
17. Luo, X.-P., Li, J., Zhen, L.-X.: Design and implementation of a card reader based on build-in camera. In: Proceedings of the 17th International Conference on Pattern Recognition (ICPR). IEEE (2004)

18. Holsapple, C.W.: DSS architecture and types. In: Burstein, F., Holsapple, C.W. (eds.) Handbook on Decision Support Systems 1, pp. 163–189. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-48713-5_9
19. Thuan, N.H.: Business Process Crowdsourcing: Concept, Ontology and Decision Support. Progress in IS. Springer, Heidelberg (2019). https://doi.org/10.1007/978-3-319-91391-9. Ed. C. Rauscher
20. Chandrasekhar, V., et al.: Dataset: Stanford Mobile Visual Search Dataset, V. Research Datasets for Image, and Multimedia Systems Group at Stanford, Editor. Stanford (2013)