Towards Ubiquitous Learning of Balinese-to-Latin Script Transliteration as Part of Balinese Language Education

Gede Indrawan*, I Gede Aris Gunadi
Dept. of Computer Science
Universitas Pendidikan Ganesha
Singaraja, Indonesia
*gindrawan@undiksha.ac.id,
igedearisgunadi@undiksha.ac.id

I Ketut Paramarta
Dept. of Balinese Language
Universitas Pendidikan Ganesha
Singaraja, Indonesia
ketut.paramarta@undiksha.ac.id

Abstract—**As part of the Balinese culture in Indonesia, the Balinese script has rarely been used today. Its Balinese-to-Latin script transliteration knowledge is also affected by that condition and has caused concern over the threat of extinction. This study is aimed to preserve that knowledge through a technological approach by collaboration between Information Technology and Balinese Language discipline. This research analyzed the ubiquitous learning aspect of Balinese-to-Latin script transliteration as part of Balinese Language education, which is a mandatory local subject from basic to high school in Bali Province. This analysis was conducted on the Balinese Glyph Recognition (BalineseGR) web application that was developed as a technology product of Universitas Pendidikan Ganesha (Undiksha), Indonesia. Balinese is a sort of Optical Character Recognition (OCR) web application that receives a Balinese script image and outputs Latin text. It was considered as the first Balinese-to-Latin script transliteration web application using JavaScript for the algorithm. Recently, its initial development has covered transliteration knowledge of eighteen Balinese script's basic syllables (*Akśara Wreşāstra*) only. Since ubiquitous learning is supported by mobile computer through a wireless network, and it is aimed to provide learners with content and interaction anytime and anywhere, BalineseGR web application has several aspects to concern about (1) its counterpart mobile application development where image input not only from a file but also more useful directly from a camera, and (2) richer knowledge content, including a transliteration of the vowels (*Akśara Suara*), additional basic syllables (*Akśara Şwalalita*), sound killers (*Pangangge Tengenan*), numerals, punctuations, ligatures, and miscellaneous glyph.**

Keywords: Balinese, Latin, transliteration, styling, ubiquitous learning

I. INTRODUCTION

Balinese script has rarely been used today and has caused concern over the threat of the extinction [1,2]. Its transliteration knowledge [3-5], is also affected by that condition. Preservation through political approach was conducted by releasing: (1) Bali Governor Regulation No. 80 Year 2018 about protection and usage of Balinese Language, Script, and Literature, and organizing the Balinese Language Month, and (2) Bali Governor Circular Letter No. 3172 the Year 2019 about the recommendation of the usage of Balinese Scripts on a national and international event hosted in Bali.

However, the political approach is not enough. Multi-side approaches were needed to strengthen the preservation effort, including the technological approach.

Balinese Language preservation effort covers the transliteration knowledge between Latin and Balinese script, and vice versa. There are previous researches by the author related to the Latin-to-Balinese script transliteration that was related to the string-pattern processing field from the technological approach [6-9]. All of those use rule-based methods and different types of font, i.e. Noto Sans Balinese [6], Bali Simbar [8], and Bali Simbar Dwijendra [9]. Its learning mobile application [7] was also developed based on [8]. Comprehensive testing with all possible rules and their examples [10] gave accuracy result up to 91% [6] and 98% [8][9]. Additional 47 testing cases were added in [9] related to the sound killers that have no example in [10]. The next milestone in this kind of transliteration is to enrich the repository consists of words belong to seventeen kinds of special words from the existing Balinese script dictionary [11].

On the other side, not so many references related to the Balinese-to-Latin script transliteration that was related to the image-processing field from the technological approach. This study enriches research in this area by analyzing the Balinese Glyph Recognition (BalineseGR) web application [12], which is adopted from GNU Optical Character Recognition (OCR) algorithm of Latin character [13]. Its ubiquitous learning aspect of Balinese-to-Latin script transliteration, as part of Balinese Language education, was also been observed. Ubiquitous learning improves learning by providing learners with content and interaction anytime and anywhere [14], by mobile and embedded computers and wireless networks in everyday life [15].

II. METHODS

This analysis was conducted on BalineseGR web application that was developed as a technology product of Universitas Pendidikan Ganesha (Undiksha), Indonesia. BalineseGR is a sort of Optical Character Recognition (OCR) web application that receives Balinese script image and output Latin text. It was considered as the first Balinese-to-Latin script transliteration web application using JavaScript for the algorithm. It was developed by using an ad hoc method specific to each character based on Bali Simbar font [8]. Recently, its initial development has covered transliteration knowledge of eighteen Balinese script's basic syllables (*Akśara Wreṣāstra*) only.

Based on the GNU OCR algorithm of Latin character [13]., BalineseGR accommodate basic syllables transliteration by only comparing the shape differences that are relevant to choose between two character categories, mostly like a binary search.

No	Latin	Regular Form	Appended Form	No	Latin	Regular Form	Appended Form
1	*ha / a*			10	*la*		
2	*na*			11	*ma*		
3	*ca*			12	*ga*		
4	*ra*			13	*ba*		
5	*ka*			14	*nga*		
6	*da*			15	*pa*		
7	*ta*			16	*ja*		
8	*sa*			17	*ya*		
9	*wa*			18	*nya*		

Fig. 1. Balinese script basic syllables.

On basic syllables regular form, the transliteration process involves one line of layouts and differentiates character shape based on possession of the number of holes, i.e.:

- No hole, for Balinese basic syllables: "ha", "ra", "ta", "wa", "la", "ma", "ga", "ba", "nga", "pa", "ja", "ya", and "nya". This category adopts recognition of Latin characters: "1", "2", "3", "5", "7", "C", "E", "F", "G", "H", "I", "J", "K", "L", "M", "N", "S", "T", "U", "V", "W", "X", "Y", "Z", "c", "f", "g", "h", "i", "j", "k", "l", "m", "n", "r", "s", "t", "u", "v", "w", "x", "y", and "z".

- One hole, for Balinese basic syllables: "ca" and "da". This category adopts recognition of Latin characters: "0", "4", "6", "9", "A", "D", "O", "P", "Q", "R", "a", "b", "d", "e", "o", "p", and "q".

- Two holes, for Balinese basic syllable: "na" and "ka". This category adopts recognition of Latin characters: "8"and "B ".

III. RESULTS AND DISCUSSION

All of the good quality of input images of a regular form of basic syllables has been successfully transliterating into their related text. So do 400x400 pixel input image consists of all regular forms of basic syllables. Figure 2 shows BalineseGR image input samples with their certain positions All of those input images can be downloaded from the BalineseGR web application [12], and after that can be uploaded individually for the testing.
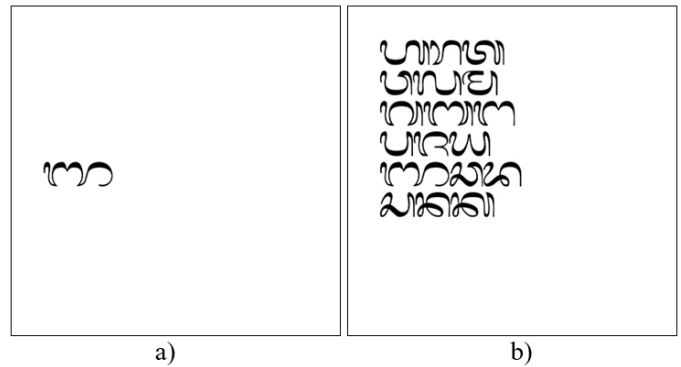

a)   b)

Fig. 2. BalineseGR image input samples: a) regular form "nya"; b) all regular forms of basic syllables.

Figure 3 shows BalineseGR was introduced among other learning applications [7], to twenty-six high-school teachers at Undiksha's community service event at SMAN 1 Bali Mandara, Indonesia.

The feedback from that event, related to the ubiquitous learning aspect of Balinese-to-Latin script transliteration, confirmed several aspects need to be taken care of: (1) mobile technology development, and (2) knowledge functionality accommodation which also used by its counterpart Latin-to-Balinese script transliteration application as another technology product of Undiksha [8,16].

Fig. 3. BalineseGR introduction to the high school teachers at community service event at SMAN 1 Bali Mandara, Indonesia.

From that public dissemination of the research, Figure 4 shows the next BalineseGR evolution planning to support ubiquitous learning on the mobile computer where image input not only from the file but also more useful directly from the camera. Figure 5 shows BalineseGR functionality accommodation planning.
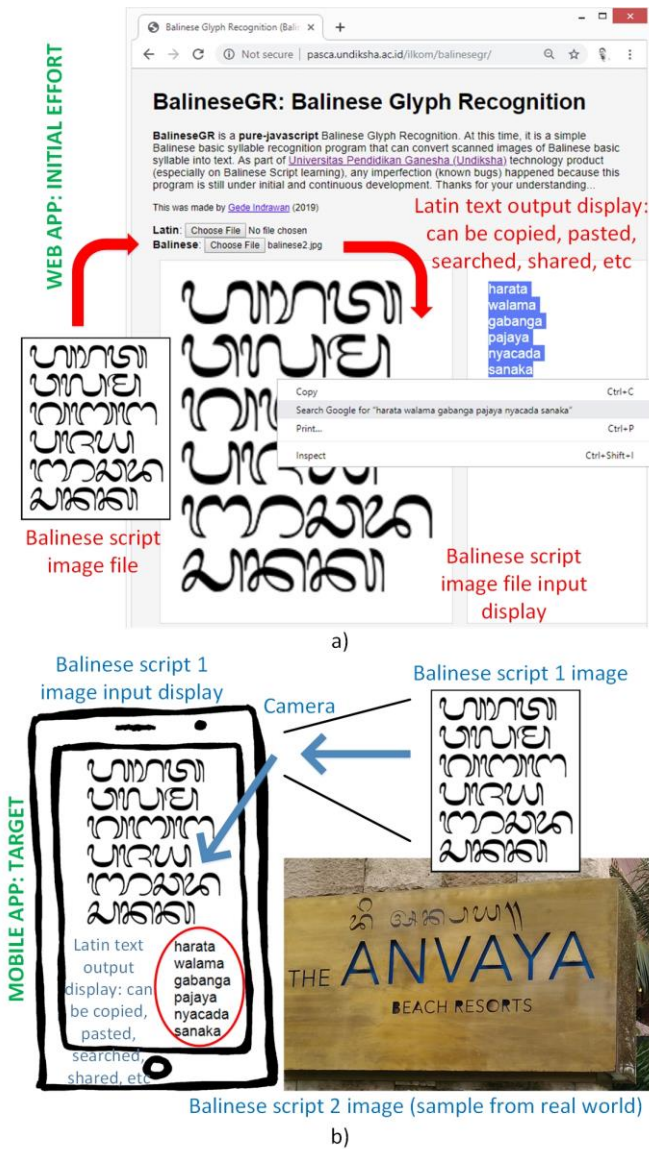

Fig. 4. BalineseGR application: a) initial web app; b) targeted mobile app.
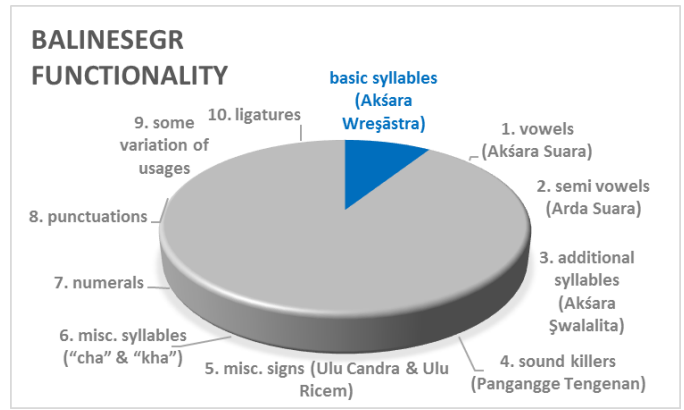

Fig. 5. BalineseGR functionality: a) initial web app; b) targeted mobile app.

A. Mobile Technology Development

BalineseGR runs only by two files, i.e.: Hypertext Markup Language (HTML) and JavaScript file. HTML file provides the User Interface (UI) functionality, while its counterpart JavaScript file provides transliteration core functionality. For supporting ubiquitous learning, BalineseGR needs to run on a mobile platform. Since HTML and JavaScript technology is widely used on many platforms, the existing BalineseGR functional code could be reused on mobile platforms.

B. Knowledge Functionality Accommodation

Figure 5 shows that BalineseGR needs major enhancement on functionality since recently it accommodate Balinese-to-Latin script transliteration only on Balinese basic syllables. This functionality also needs to be more robust for any case of character shape rotation, stretching, and cut. To make it worse, the input image is also not always in good quality.

For near-future work on a limited case of ideal good quality of input image, there are ten others knowledge category functionality need to be accommodated by BalineseGR, as shown by Figure 5, i.e. (1) vowels (*Akśara Suara*), (2) semivowels (*Arda Suara*), (3) additional syllables (*Akśara Şwalalita*), (4) sound killers (*Pangangge Tengenan*), (5) miscellaneous signs (*Ulu Candra* and *Ulu Ricem*), (6) miscellaneous syllables ("cha" and "kha") (7) numerals, (8) punctuations, (9) some variation of usages, and (10) ligatures.

Based on Figure 1, more complexity processing comes from basic syllables appended form (and also appended form from other categories) where the transliteration process involves two lines of layouting of character shape. Appended form from other categories also can have three lines of layouting of character shape, as shown by Figure 6. These multiple lines of layouting need to be addressed on the next improvement of BalineseGR or other Balinese-to-Latin script transliteration application.

No	Latin	Balinese
1	*Om Swastiastu* (May God blesses you)	
2	*Om Şanti, Şanti, Şanti, Om* (May peace be everywhere)	

Fig. 6. Balinese Script Miscellaneous Signs (Ulu Candra & Ulu Ricem) Word Examples.

IV. CONCLUSION AND FUTURE WORK

The possibility of ubiquitous learning of Balinese-to-Latin script transliteration, as part of Balinese Language education, was analyzed by the possible use of BalineseGR web application, as an initial technology product of Universitas Pendidikan Ganesha (Undiksha), Indonesia.

Since ubiquitous learning is supported by mobile computer through the wireless network, and it is aimed to provide learners with content and interaction anytime and anywhere, BalineseGR has several aspects to enhance: (1) its counterpart mobile application where image input not only from the file but also more useful directly from a camera, and (2) richer knowledge content, including a transliteration of the vowels (*Akśara Suara*), *Akśara Şwalalita*, sound killers (*Pangangge Tengenan*), numerals, punctuations, ligatures, and miscellaneous glyph.

ACKNOWLEDGMENT

REFERENCES

[1] I.N. Jampel, G. Indrawan, and I.W. Widiana, "Accuracy Analysis of Latin-to-Balinese Script Transliteration Method," International Journal of Electrical and Computer Engineering, vol. 8, no. 3, 2018.

[2] G. Indrawan, I.K. Paramarta, K. Agustini, and Sariyasa, "Latin-to-Balinese Script Transliteration Method on Mobile Application: A Comparison," Indonesian Journal of Electrical Engineering and Computer Science, vol. 10, no. 3, 2018.

[3] S. Karimi, F. Scholer, and A. Turpin, "Machine transliteration survey," ACM Computing Surveys, 2011.

[4] K. Kaur, "Review of Machine Transliteration Techniques," International Journal of Computer Applications, vol. 107, no. 20, 2014.

[5] N.S. Kharusi and A. Salman, "The English Transliteration of Place Names in Oman," Journal of Academic and Applied Studies, vol. 1, no. 3, pp. 1–27, 2011.

[6] G. Indrawan, N.N.H. Puspita, I.K. Paramarta, and Sariyasa, "LBtrans-Bot: A Latin-to-Balinese Script Transliteration Robotic System based on Noto Sans Balinese Font," Indonesian Journal of Electrical Engineering and Computer Science, vol. 12, no. 3, 2018.

[7] G. Indrawan and I.K. Paramarta, "The Development of Learning Mobile Application of Latin-to-Balinese Script Transliteration [Pengembangan Aplikasi Mobile Pembelajaran Transliterasi Tulisan Latin ke Aksara Bali]," Pekommas, vol. 4, no. 2, pp. 123–130, 2019.

[8] G. Indrawan, Sariyasa, and I.K. Paramarta, "A New Method of Latin-To-Balinese Script Transliteration based on Bali Simbar Font," in The 4th International Conference on Informatics and Computing (ICIC), 2019.

[9] G. Indrawan, I.P.E. Swastika, Sariyasa, and I.K. Paramarta, "An Improved Algorithm and Accuracy Analysis Testing Cases of Latin-to-Balinese Script Transliteration Method based on Bali Simbar Dwijendra Font," International Journal of Recent Technology and Engineering, forthcoming 2020.

[10] Ida Bagus Adi Sudewa, "The Balinese Alphabet," 2003. [Online]. Retrieved from: http://www.babadbali.com/aksarabali/alphabet.htm.

[11] I.G.K. Anom et al., Balinese - Indonesian Dictionary with Its Writing in Latin and Balinese Script [Kamus Bali - Indonesia Beraksara Latin dan Bali]. Denpasar: Language, Script, and Literature Development Agency, Culture Office, Bali Province, 2009.

[12] G. Indrawan, "BalineseGR | Balinese Glyph Recognition," 2019. [Online]. Available: http://pasca.undiksha.ac.id/ilkom/balinesegr/.

[13] Antonio Diaz, "Ocrad - The GNU OCR," 2017. [Online]. Retrieved from: https://www.gnu.org/software/ocrad/.

[14] G.J. Hwang, C.C. Tsai, and S.J.H. Yang, "Criteria, strategies and research issues of context-aware ubiquitous learning," Journal of Educational Technology and Society, vol. 11, no. 2, 2008.

[15] H. Ogata, Y. Matsuka, M.M. El Bishouty, and Y. Yano, "LORAMS: Linking Physical Objects and Videos for Capturing and Sharing Learning Experiences Towards Ubiquitous Learning," International Journal of Mobile Learning and Organisation, 2009.

[16] G. Indrawan and I.K. Paramarta, "Computerized Transliteration of Latin Text to Balinese Script [Komputerisasi Transliterasi Teks Latin ke Aksara Bali]," 2019. [Online]. Retrieved from: https://ebooks.gramedia.com/books/komputerisasi-transliterasi-teks-latin-ke-aksara-bali-seri-teknologi-informasi-untuk-budaya.