

University Guide using Speech Recognition and Computer Vision

¹Anthony Kosasi, ²Nur Hadisukmana, ³Rosalina, ⁴R.B. Wahyu, ⁵Budi Sulisty, ⁶Yuyu Wahyu

Fakultas Komputer
President University
Bekasi, Indonesia

¹anthonykosasi777@yahoo.com, ²nhadisukmana@gmail.com, ³rosalina@president.ac.id, ⁴rbw0101@gmail.com,
⁵budi241@yahoo.com, ⁶yuyuwahyusr@gmail.com

Abstract- Education is an important factor inside mankind's existence. However, choosing a great and promising place to learn is a very hard thing to do, especially for senior high school student who want to continue their education into the university. There are a lot of universities around us and a lot of major or concentration that we must choose before entering a university. Considering this issue, Some application will be developed and the application will help people in understanding some majors those are available and help people to know about some universities in Indonesia whether about vision, mission, curriculum, etc. This application will not be a boring information application that provides information in text and let user read it. This is rather an attractive application that will tell some information about each major in speech and also give information about each university when user scan the university logo with computer vision. Not only that, after user scans the university logo, video profile of the university will be played and then user can pause, stops or plays the video also by using speech. Some additional feature like migration into major field and website link will also be provided.

Keywords: *university guide; speech recognition; computer vision.*

I. INTRODUCTION

Nowadays, technology evolution is growing rapidly. People tend to like something more attractive and unique. One of the examples is about finding information using computer because it can gather all the important data that people want and it is more flexible and easier to use. However, people cannot deny that even by using computer, the procedure of gaining information compare to using books is the same. People must read it. Considering this issue, the thing that can be improved in here is the way to gaining the information. To make the way of gaining information more attractive, an application that use speech as input and also the output will be developed.

In order to make it more interesting, this application also uses computer vision feature for some university logo to improve the attractiveness of this application and make people choose this kind of application in gaining information. This application will use speech, click action and university logo as

input so that even blind people can use it, although it is not 100% effective. The main problem that is happening in here is finding information about universities and any major that they provide, also the description about the major itself. This application will make it easier for senior high school student and their parent to understand more about the future path that students from high school are going to take. This application differ the finding criteria into 2 topics. First, it is about each major's information and second is about the information about any university. For people that already have the university that they will attend to in mind, they can just search for major that interest them or vice versa. What make this application more interesting is about finding the university's description. To do it, first the user must scan the university logo first and then the application will show the image of the university with computer vision. Not only that, we can gain some information about the scanned university through speech or click action. Any available command will be displayed. Besides that, some features like opening website of the university and showing video about the university will also be available.

II. LITERATURE STUDY

Speech Recognition (SR) is the ability of a machine or program to identify words and phrases in spoken language and convert them to a machine-readable format [2]. It is also can be concluded that Speech recognition is the process of capturing spoken words using a microphone or telephone and converting them into a digitally stored set of words. The quality of a speech recognition systems are assessed according to two factors: its accuracy (error rate in converting spoken words to digital data) and speed (how well the software can keep up with a human speaker). Speech recognition technology has endless applications. Commonly, such software is used for automatic translations, dictation, hands-free computing, medical transcription, robotics, automated customer service, and much more [3].

Some SR systems use "training" (also called "enrolment") where an individual speaker reads text or isolated vocabulary into the system because a critical component in the pattern matching approach to speech recognition is the training algorithm, which aims at producing typical (reference) patterns or models for accurate pattern comparison [11]. The performance of speech recognition systems is usually evaluated

in terms of accuracy and speed. Accuracy is usually rated with word error rate (WER), whereas speed is measured with the real time factor. Other measures of accuracy include Single Word Error Rate (SWER) and Command Success Rate (CSR). However, speech recognition (by a machine) is a very complex problem. Vocalizations vary in terms of accent, pronunciation, articulation, roughness, nasality, pitch, volume, and speed. Speech is distorted by a background noise and echoes, electrical characteristics [9].

In speech analysis, the voiced-unvoiced decision is usually performed in conjunction with pitch analysis. The linking of voiced-unvoiced (V-UV) decision to pitch analysis not only results in unnecessary complexity, but makes it difficult to classify short speech segments which are less than a few pitch periods in duration [10].

III. SYSTEM OVERVIEW

This application will accept input in a form of speech, provided university logo and button/label click. After receiving any kind of the input, System will process it and give the output. This application output is also varies and they are in a form of speech, printed out information, video, opening Microsoft word or any university website link. Figure 3 will show this application system overview figure of its input and output.

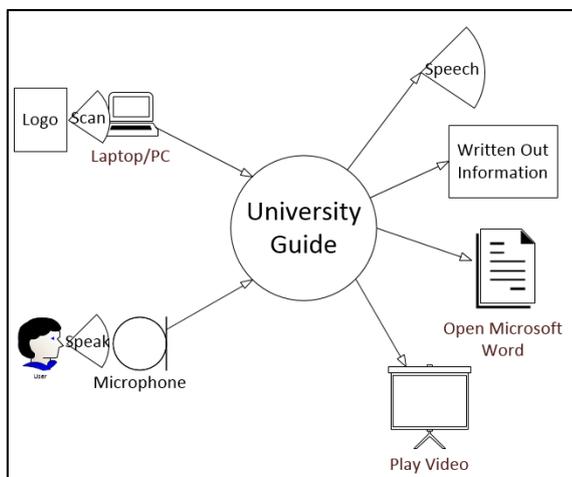


Figure 3. Application System Overview

IV. APPLICATION PREVIEW

This section is giving the application preview of its implementation. This section will include the use case and some application screenshots. Figure 4 will shows the use case diagram of the application. The use case includes “Download the University Logo”, “Choose Majors”, “Search Information about the Chosen Major”, “Upload Information about Certain Major”, “Scanning the University Logo”, “Open the Scanned University Website Link”, “Search Information about Certain University” and “go into Available Major in Scanned University”.

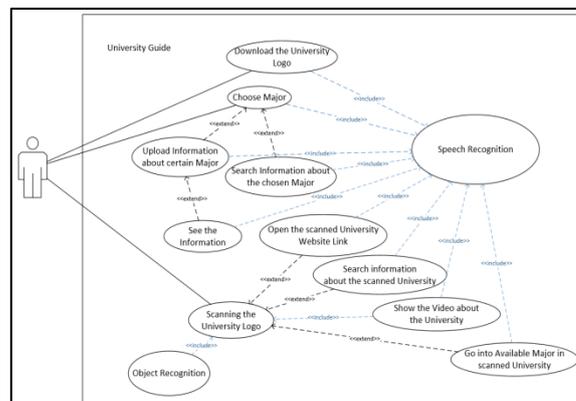


Figure 4. Use Case Diagram

Figure 5 will shows the Main Form of the application. In here, 3 menus will be displayed and they are Logo, Majors and Universities. If user chooses Logo, the will go into Form Logo (Figure 6) and in here user can download university logos. Figure 7 will shows the form to choose a major. After choosing any major, user will be redirected into major description form, for example, Information Technology Form can be seen in Figure 8.

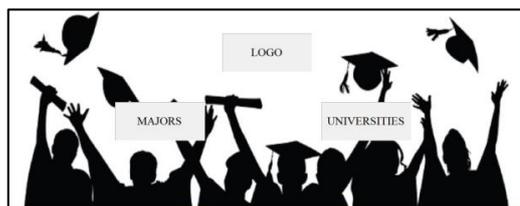


Figure 5. Main Form

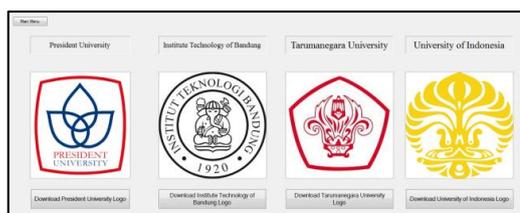


Figure 6. Download Logo

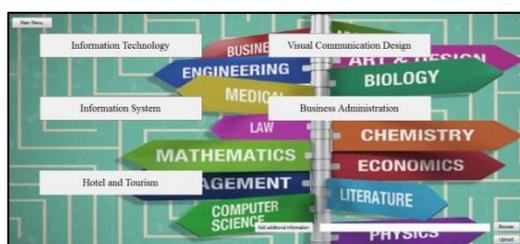


Figure 7. Choose Major

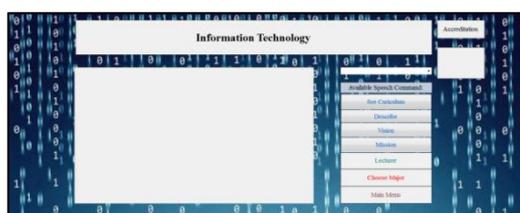


Figure 8. IT Description Form

There is additional feature if user wants to upload some more information for the admin, user can browse and upload it in choose major form and the data can be seen in figure 9.

	major	university name	description	vision	mission
▶	Industrial Engine...	President Univers...	School of Engine...	To become a ce...	To create and to ...
*					

Figure 9. Uploaded Data

If user clicks or speaks lecturer in any Major Description form, lecturer of the related major and university will be shown (Figure 10). The last part is Universities. If user goes into University Recognition form, user will first scans any provided university logo and then the application will provide some more information. Figure 11 will shows the University Recognition default form.



Figure 10. Lecturer Form



Figure 11. University Recognition Form

Figure 12 shows the Swimlane Diagram of use case Download the University Logo. User speaks download after choose the University logo and then The System will download the Logo automatically into the computer/laptop.

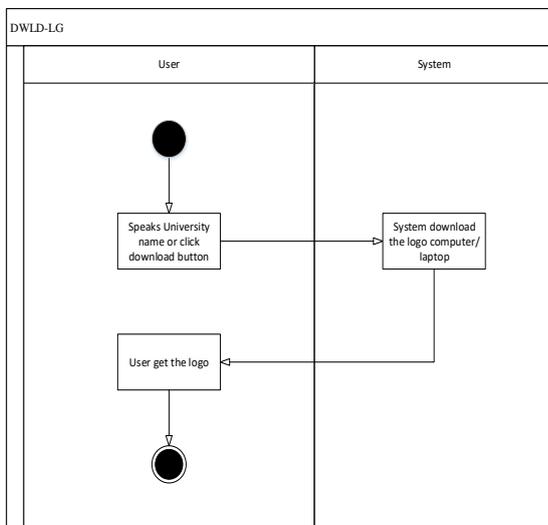


Figure 12.. Swimlane Diagram – Download the University Logo

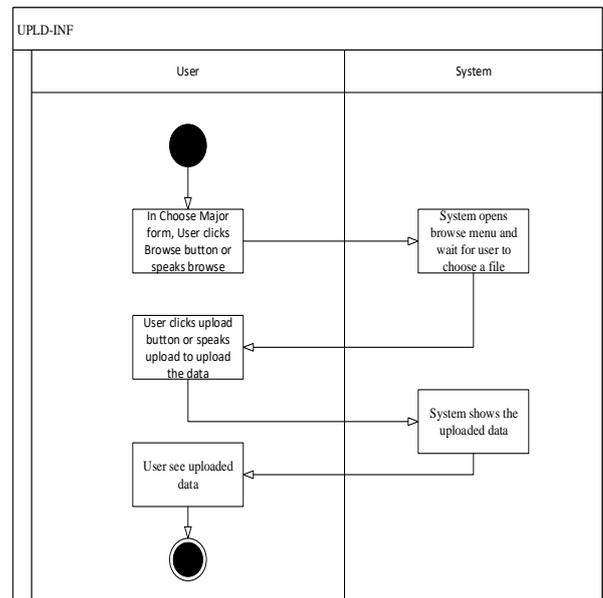


Figure 13. Swimlane Diagram – Upload Information about Certain Major

Figure 13 shows the Swimlane Diagram of use case Upload Information about Certain Major. This Swimlane describes about condition after user speaks browse or clicks the button manually in choose major form, user will browse the file that he/she want to upload. Later on, User clicks upload and then user will see all the data that he/she already inputted.

V. EXPERIMENTAL RESULTS

In order to check and be sure that application runs properly, some simulations are needed to check it. Simulations will be divided based on the activities of the program executed. Testing scenario describes the activities which have to be checked in order to prove that the system works properly. The testing scenario will be divided into two sections, which are Speech Recognition Section and Object Recognition Section. Speech Recognition Section explains about all input using speech recognition or button/label click. Object Recognition Section explains about any input using image and the recognition feature.

This section contains Table 1 which shows the testing scenarios, expected result and evaluation for input that use Speech Recognition or button/label click.

TABLE I SPEECH RECOGNITION SECTION

No.	Scenario	Expected Result	Evaluation
1.	Speaks Logo in Main Menu Or Clicks Logo Label in Main Menu	Go to Brand Form	Success
2.	Speaks Majors in Main Menu Or Clicks Logo Major in Main Menu	Go to Choose Major Form	Success
3.	Speaks Universities in Main Menu or Clicks Logo	Go to University	Success

No.	Scenario	Expected Result	Evaluation
	Universities in Main Menu	Recognition Form	
4.	Speaks Main Menu in Form Brand or Clicks Main Menu Button in Form Brand	Go back into Main Menu	Success
5.	Speaks (University Name) in Brand Form or Clicks Download (University Name) Logo Button	Download the spoken University Name Logo	Success
6.	Speaks Main Menu in Choose Major Form or Clicks Main Menu Button in Choose Major Form	Go back into Main Menu	Success
7.	Speaks Exit Application in Any Form	Exit the Application	Success
8.	Speaks Browse in Choose Major Form or Clicks Browse Button in Choose Major Form	Open the Browse file window	Success
9.	Speaks Upload in Choose Major Form or Clicks Upload Button in Choose Major Form	Go into Upload Form to see the data	Success
10.	Speaks (major name) in Choose Major Form or Clicks any Major Name Label	Go into (major name) Form	Success
11.	Speaks box (University Name) in any Major Description Form	Combo Box in the Form will change into spoken University Name	Success
12.	Use Combo Box in any Major Description Form to choose University	Combo Box Text will change according to the chosen University	Success
13.	Speaks See Curriculum in any Major Description Form or clicks See Curriculum Button	Open Microsoft Word with curriculum data of related Major and University	Success
14.	Speaks Describe or Clicks Describe button in any Major Form when comboBox1.Text is not equal to "" in any Major Description Form	System will change label1.text text into description of current University Major text and read it.	Success
15.	Speaks Vision or Clicks Vision button in any Major Form when comboBox1.Text is not equal to "" in any	System will change label1.text text into vision of current University	Success

No.	Scenario	Expected Result	Evaluation
	Major Description Form	Major text and read it.	

Object Recognition Section

This section contains Table 2 which shows the testing scenarios, expected result and evaluation for Object Recognition.

TABLE II. OBJECT RECOGNITION SECTION

No.	Scenario	Expected Result	Evaluation
1.	Choose President University in the combo box then scan President University Logo	System will shows President University appearance, website link and profile video and there will be a blue square surrounding the recognized area	Success
2.	Choose Institute Technology of Bandung in the combo box then scan Institute Technology of Bandung Logo	System will shows Institute Technology of Bandung appearance, website link and profile video and there will be a blue square surrounding the recognized area	Success
3.	Choose Tarumanegara University in the combo box then scan Tarumanegara University Logo	System will shows Tarumanegara University appearance, website link and profile video and there will be a blue square surrounding the recognized area	Success
4.	Choose University of Indonesia in the combo box then scan University of Indonesia Logo	System will shows University of Indonesia appearance, website link and profile video and there will be a blue square surrounding the recognized area	Success

Figure 14 will shows the example of successful President University Recognition.

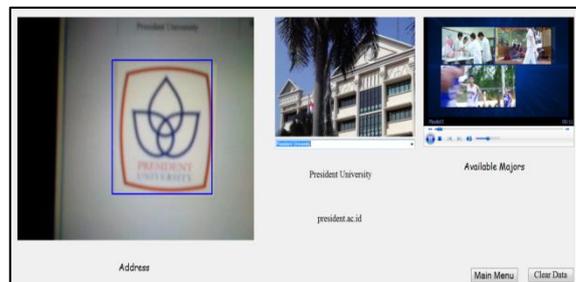


Figure 14. President University Logo Recognition

VI. CONCLUSION AND FUTURE WORKS

This research has been successfully achieved all the objectives which are becoming a source of information about some majors and universities that is very useful for high school

student who want to continue their study into the university and become a very attractive desktop application because of its special features which are speech and object recognition.

This system could be used to search information about all universities in Indonesia as long as the admin has already inputted all the detail needed in the application. This application can be installed in any highschool who wants to give their students information about some majors which are available in any universities listed in the application and help their students to know about its vision, mission, curriculum, etc. The listed university could be added by the admin of the highschool.

For future works this application could be easily shared if it is build under android platform.

REFERENCES

- [1] Atal, B.S. , Hill, Murray, Rabiner, L. *A pattern recognition approach to voiced-unvoiced-silence classification with applications to speech recognition*. Bell Laboratories. 29 January 2003.
- [2] *Computer Vision*. <http://research.microsoft.com/en-us/research-areas/computer-vision.aspx>.
- [3] Emgu CV. http://www.emgu.com/wiki/index.php/Main_Page.
- [4] Hwang Juang, B., Huo,W., Hill,M., *Minimum classification error rate methods for speech recognition*. Bell Laboratories. 06 August 2002.
- [5] *How Speech Recognition Works*. <https://msdn.microsoft.com/en-us/library/hh378337%28v=office.14%29.aspx>.
- [6] *Introducing Visual Studio*. <https://msdn.microsoft.com/en-us/library/fx6bk1f4%28v=vs.90%29.aspx>.
- [7] Ji, Qiang. *3dvision_intro*. Department of Electrical, Computer, and Systems Engineering. Rensselaer Polytechnic Institute. 2 February 2015.
- [8] Kumar, Dileep. *Creating a haar cascade classifier aka haar training*. 18 August 2011.
- [9] Microsoft Visual Studio. https://en.wikipedia.org/wiki/Microsoft_Visual_Studio
- [10] Seo, Naotoshi. *Tutorial: OpenCV haartraining (Rapid Object Detection With A Cascade of Boosted Classifiers Based on Haar-like Features)*. 16 October 2008.
- [11] *Speech Recognition Definition*. <http://searchcrm.techtarget.com/definition/speech-recognition>.
- [12] *Speech Recognition*. https://en.wikipedia.org/wiki/Speech_recognition.
- [13] Thompson, Lawrence. *Key Differences Between Speech Recognition and Voice Recognition*.
- [13] *What is Computer Vision?*. <http://www.bmva.org/visionoverview>
- [14] *Working with Images*. http://www.emgu.com/wiki/index.php/Working_with_Images.