ICIC Express Letters
Part B: Applications
Volume **12**, Number **3**, March **2021**

ICIC International ©2021 ISSN 2185-2766
pp. **275–281**

A NOVEL MODEL AND IMPLEMENTATION OF HUMANOID ROBOT WITH FACIAL EXPRESSION AND NATURAL LANGUAGE PROCESSING (NLP)

Widodo Budiharto[1,*], Vincent Andreas[1]
and Alexander Agung Santoso Gunawan[2]

[1]Computer Science Department
[2]Mathematics Department
School of Computer Science
Bina Nusantara University
JL. K. H. Syahdan No. 9, Kemanggisan, Palmerah, Jakarta 11480, Indonesia
[*]Corresponding author: wbudiharto@binus.edu

Received March 2020; accepted June 2020

Abstract. *Rapid development of humanoid robot focuses on self-learning and knowledge retrieval to make an interactive conversation and natural interaction with human in Indonesian language is very rare. In this research, we would like to propose a model of humanoid robot with the self-learning capability for accepting questions and giving a response with facial expression and Natural Language Processing (NLP). The prototype of our humanoid robot can recognize object and show facial expression for showing happy, angry, sad and natural expression. This kind of robot should be able to be used widely in universities and public services for industry 4.0 era. In our scenario, the knowledge from users is based on speech, text and showing images in front of the robot. We use Google translator technology for speech recognition systems. If we compared the result with our previous research, the accuracy is increasing about 30% (60% accuracy from previous research to 90% accuracy). There is an improvement with previous research where the robot is able to detect an object. The comparative experiments using prototype of humanoid robot with expression with samples from users were presented and analyzed.*
Keywords: Humanoid robot, Facial expression, Speech recognition systems, Indonesian language, NLP, Big data, Object detection

1. **Introduction.** Intelligent humanoid robots with facial expression and interactive conversation are our hope for the future. For example, one of the newest expressive humanoid robots and commercially produced is Reeti. Reeti is provided with user friendly interface, and an iPhone or iPad applications. For another example, Bruce et al. [1] explored the effects of the expression of emotion and indication of attention on the robot's success at initiating interaction. Without the face or the ability to move, the robot relies solely on verbal cues to attempt to engage people in interaction. Understanding speech of human is a difficult and long task for the humanoid robot with natural interaction, as shown in our previous research in developing humanoid robot for education [2] and robot that is able to get basic knowledge and answer questions [3].

Google Speech to Text [4] is usually used in humanoid robot and capable of more general speech recognition but requires a stable Internet connection and incurs additional processing time. Perera and Veloso [2] proposed a novel approach to enable a mobile service robot to understand questions about the history of tasks it has executed. We frame the problem of understanding such questions as grounding an input sentence to a query that can be executed on the logs recorded by the robot during its runs. We define a query as an operation followed by a set of filters. Robotics' engineer Angelo Cangelosi of the

DOI: 10.24507/icicelb.12.03.275

University of Plymouth in England and Linda B. Smith, a developmental psychologist at Indiana University Bloomington, have demonstrated how crucial the body is for procuring knowledge. The shape of the robot's body, and the kinds of things it can do, influence the experiences it has and it can learn from [3]. Learning from demonstration approaches focuses on the development of algorithms that are generic in their representation of the skills and in the way, they are generated. One of the most promising approaches is those that encapsulate the dynamics of the movement into the encoding [4].

The contribution of this research is how to develop a novel model of humanoid robot with facial expression that is able to self-learn based on the knowledge from the speech and images provided by human or user. If we have this model, then everyone can replicate this model to develop many humanoid robots for general application. Our experiment shows that the model is satisfied for humanoid robot with interactive conversation with facial expression. This is answering the research question of this study. In summary, our model is very simple, reliable and can be used as a novel model for developing intelligent humanoid robot.

In this paper, our objective is to make a humanoid robot that can talk as a receptionist and answer questions from the dataset using deep learning. We propose a simple model of self-learning for the humanoid robot using deep learning, Part 1 is an introduction, Part 2 shows deep learning and speech recognition systems, we propose a method in Part 3, experimental results are analyzed in Part 4 and Part 5 concludes the paper. Figure 1 shows our prototype of the humanoid robot with a servo and servo controller.

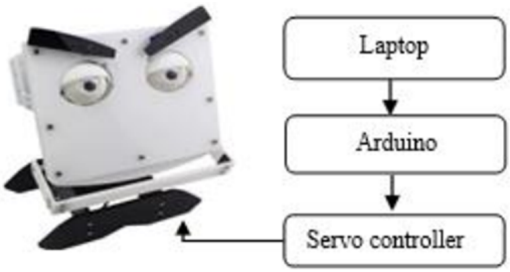

FIGURE 1. HappyBot, our prototype of intelligent humanoid robot with facial expression and deep learning [5]

2. **Deep Learning and Speech Recognition Systems.** Deep learning is a specific subset of machine learning, which is a specific subset of artificial intelligence. Computer vision and natural language processing are a great example of a task that deep learning has transformed into something realistic for robot applications. Using deep learning to classify and label images and text will be better than actual humans. Deep learning methods are proving very good at text classification, achieving state-of-the-art results on a suite of standard academic benchmark problems. Feng et al. [6] tried to propose an answer selection using some methods. First is by converting question and candidate answers to word vector, and then count the cosine similarity of the answer. The highest cosine value means the answer to the question. The second method is by using an information retrieval baseline, then using cosine similarity to find the highest value. For the deep learning method, they use CNN to handle it. They use java to program their model. The result shows that a deep learning based model shows a better result than another method. As a source of big data for our robot, we obtain knowledge for our robot using our campus's website. Yin et al. [7] did the comparative study of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) for NLP. They do some task (sentiment classification, relation classification, textual entailment, answer selection, question relation match, path query answering, part-of-speech tagging) to test the performances. To get a fair result, they train the models from scratch, using basic setup,

and search optimal hyperparameters for each task and model separately. They found that RNN performs better than CNN in most of the task, except in keyphrase recognition task and question-answer match setting. A novel deep learning optimization algorithm for human motions anomaly detection proposed by Omae et al. [8] showed good result of the implementation of deep learning.

Question Answering (QA) is a general task in humanoid robots for answering questions written in natural language with adequate sentences. Komiya et al. [9] tried to improve QA system performance by using two methods. First do a modification for query expansion. Query expansion is a method to extend query words by adding a new word to improve search answer results. They use mutual information based on two words in each question and a word in its answer, and count degree of co-occurrence between them. The more co-occurrence in a corpus, the greater their mutual information becomes equal. When they do the experiments, the accuracy and Mean Reciprocal Rank (MRR) score are increased. The second is to evaluate the module for candidate answers. There are two methods to do the candidate answer evaluation. First, count topic relevance by feedback from the web. The second is to treat questions and answer as translation. The system will calculate translation probabilities. The two evaluating candidate answers methods perform better results than without the methods. In this research, we use Google API Speech Recognition and Python language to recognize voice [10]. To understand and find the answer, we use deep learning technology developed using Python. After finding the answer, we will use Google Text to Speech to speak the answer to the user. Natural language processing is a general tool for processing the Indonesian language to be understood by the robot.

3. **Proposed Method.** We propose a method of our research as described in Figure 2. This research is commenced from designing a model of intelligent humanoid robot named as HappyBot.

For the knowledge base, we can input with 3 methods. First, user directly inputs the knowledge, by saying the knowledge to robot. The second method is by summarizing the data that we get from crawled website. The third is by showing an object in front of the camera of the robot and storing the information in the database of the robot. Our robot uses camera for real-time face detection using Haar cascade from Viola and Jones [11], because for the task of face detection, the initial rectangle features selected by AdaBoost are meaningful and easily interpreted based on OpenCV. When user comes in front of the robot, the algorithm for accepting command from user is started. The architecture of our knowledge base is to use an array for saving knowledge base, expression and store of id_object for knowledge of an object.

Example 1: "You like to drink milk" expression = "Happy"

Knowledge 1: Array	[0]	[1]	[2]	[3]	[4]
	["You"]	["like"]	["drink"]	["milk"]	["Happy"]

Example 2: "This is a rose"

Knowledge 2: Array	[0]	[1]	[2]	[3]	[4]
	["This"]	["is"]	["a"]	["rose"]	[id_object=2]

We propose a model for storing information with the emotion using low cost humanoid robot. We define the information that each knowledge has feature of expression as shown in Figure 3. For example,

My name is HappyBot.	Expression: Happy
I am angry.	Expression: Angry
I am hungry.	Expression: Sad
I am robot from BINUS Univ.	Expression: Natural
This is a rose.	Expression: Happy

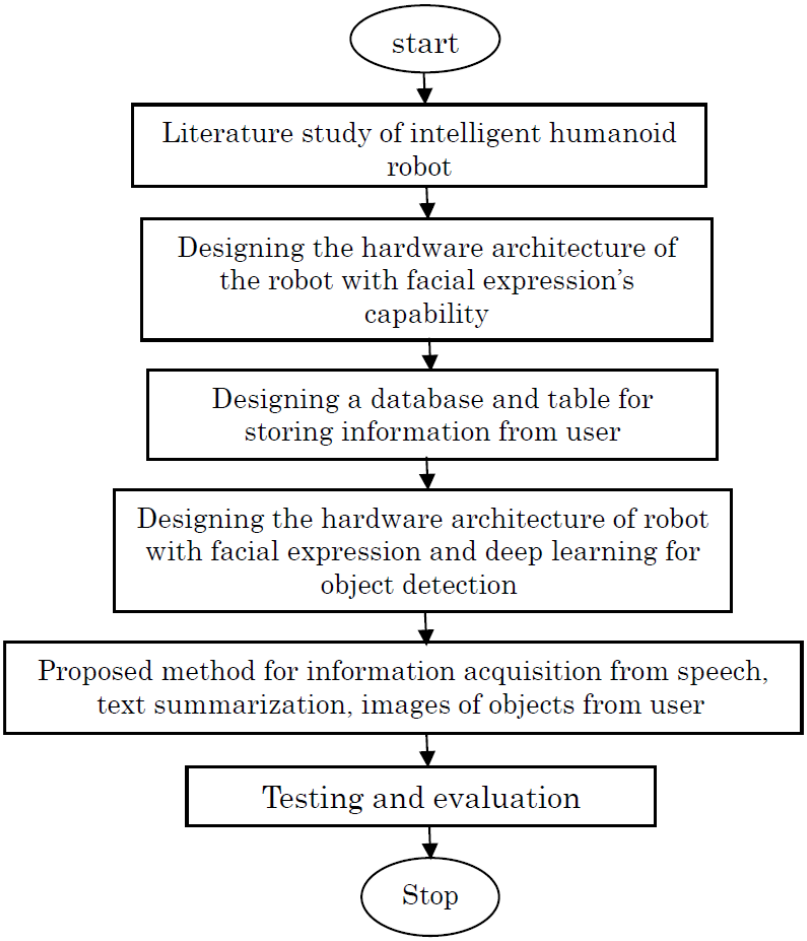

FIGURE 2. Method of our research

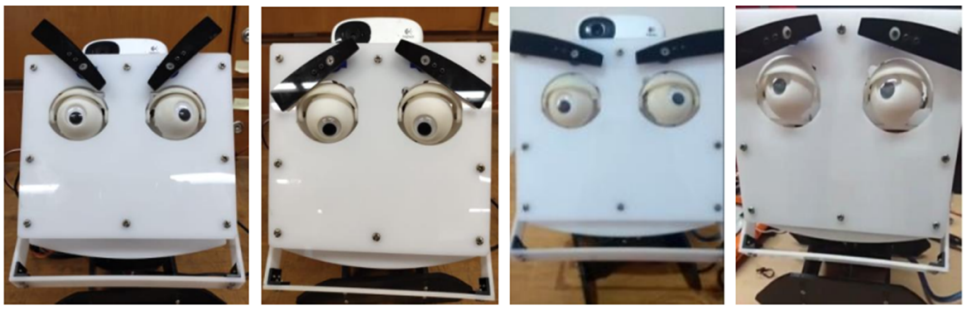

FIGURE 3. Proposed facial expressions for humanoid robot

In our previous work, we successfully proposed the face recognition and speech recognition system using stemming and tokenization for humanoid robot for education by using various NLP principles and basic self-learning capability [2,12,13]. To improve the previous research, we want to make robot with the ability for self-learning and able to detect an object. The source of knowledge can be from texts, images or speech. To achieve a complex behavior of humanoid robot, it would be necessary to have inclusive and comprehensive repertoires of skills especially in response to the questions [11]. Our model is using deep learning to do the question-answering task based on the dataset for improving the accuracy. For the model training, we will use SQuAD Dataset that contains 100,000+ questions as shown in Figure 4.

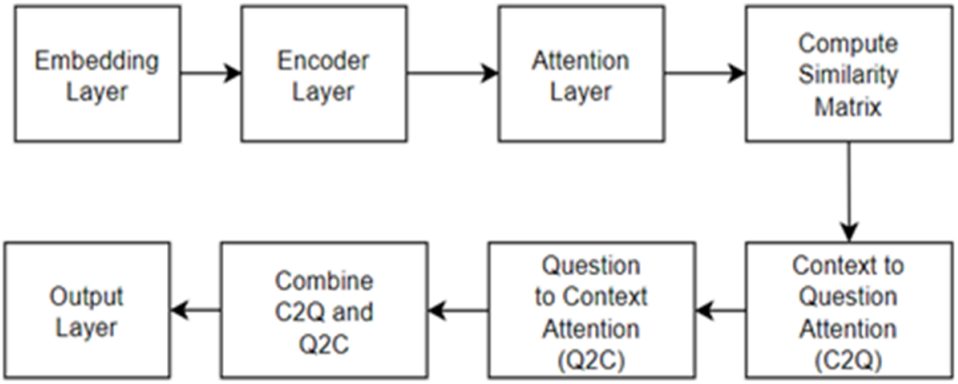

FIGURE 4. Our learning model using deep learning with attention layer and combining C2Q and Q2C will improve the accuracy.

4. **Experimental Results.** We conducted an experiment in our labs with the distance between the robot and the user being about 40cm. Before we test the new algorithm, we give the robot basic knowledge. First, we test the Indonesian language, using the given datasheet:

Context:
School of computer science mempunyai jurusan teknik informatika, jurusan game application and technology, jurusan mobile application and technology, jurusan cyber security. school of computer science memiliki akreditasi sangat baik. Ruang LSC atau ruang AOC atau ruang dosen ada di lantai dua. kantin terdapat di lantai basement. Starbuck terdapat di lantai satu. prof Haryanto Prabowo adalah rektor binus. prof Haryanto Prabowo adalah pemimpin binus. Universitas Bina Nusantara dimulai pada '21 oktober 1974'.

School of computer science has major in informatics engineering, game application and technology, mobile application and technology, and cyber security. school of computer science has very good accreditation. The LSC room or the AOC room or the lecturer room is on the second floor. the canteen is in the basement. Starbuck is on the first floor. Prof. Haryanto Prabowo is the rector of Binus. Prof. Haryanto Prabowo is the leader of Binus. Bina Nusantara University established in 'October 21, 1974'.

Untuk menguasai Teknik Informatika, paling tidak kamu memahami logika dan matematika. Mahasiswa Teknik Informatika akan menghabiskan banyak waktunya dengan pemrograman, dan pemrograman itu sendiri adalah perwujudan dari algoritma (matematika dan logika). Tapi, kalau kamu tertarik dengan dunia teknologi dan memiliki minat untuk mengembangkan sebuah perangkat lunak maka program studi ini tentu bisa menjadi pilihan kamu. orang yang dapat masuk ke jurusan teknik informatika adalah yang dulunya SMA IPA ataupun SMK RPL.

To mastering Informatics Engineering, at least you understand logic and mathematics. Informatics Engineering students will spend a lot of time with programming and programming itself is an embodiment of the algorithm (mathematics and logic). But if you are interested in the world of technology and have an interest in developing a software program then this study program can certainly be your choice. people who can enter the department of informatics engineering are those who used to be science high school or computer vocational high school.

Then, we try to ask the robot and the result is shown in Table 1.

Our proposed model and system are also able to recognize object learned before using deep learning, as shown in Figure 5, and the robot says that it is a rose.

The proposed model successfully makes our intelligent humanoid robot accept questions and respond the user with appropriate answers. Based on experiments we have done

TABLE 1. Result of experiments for robot with expression using Indonesian language

No	Questions	True answer	Predicted answer
1	Apakah kamu sedang marah? *(Are you angry?)*	Ya, saya sedang marah *(Yes, I am angry)*	Ya, saya sedang marah *(Yes, I am angry)* Expression: Angry
2	Apakah kamu sedang sedih? *(Are you sad?)*	Ya, saya sedang sedih *(Yes, I am sad)*	Ya, saya sedang sedih *(Yes, I am sad)* Expression: Sad
3	Hai, siapa namamu? *(Hi, what is your name?)*	Nama saya HappyBot *(My name is HappyBot)*	Nama saya HappyBot *(My name is HappyBot)* Expression: Happy
4	Hai, apa kabar? *(Hi, how are you?)*	Saya lapar *(I'm hungry)*	Saya lapar *(I'm hungry)* Expression: Hungry
5	What is it? *(showing a rose)*	Itu adalah bunga mawar *(That is a rose)*	Itu adalah bunga mawar *(That is a rose)*

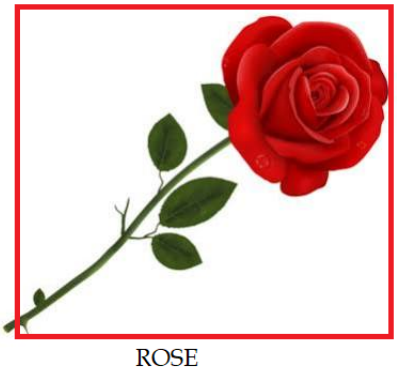

FIGURE 5. Robot is able to recognize and says that the object is a rose.

many times; our system has proven to be quite realistic and feasible to be used for real applications.

If we compared the result with our previous research, the accuracy is increasing about 30% (60% accuracy from previous research to 90% accuracy) because when we use deep learning, the pattern of questions and answers is already known, then we can use it to find the answers [3]. There is an improvement with previous research where the robot is able to detect an object.

5. **Conclusion.** The outstanding feature of the humanoid robot is the probability to communicate with it, to teach and to interact with a human. Our system is successfully able to obtain knowledge using speech and deep learning. Our architecture of knowledge base uses array for saving knowledge base, expression and storage of id_object for knowledge of an object. For future development, we will use the database to save knowledge, so the knowledge can store more data and manage easily. We also will improve our algorithm to make better results to find the answer and improve the face of the robot that is able to show the best emotion.

REFERENCES

[1] A. Bruce, I. Nourbakshs and R. Simmons, The role of expressiveness and attention in human-robot interaction, *Proc. of the 2001 AAAI Fall Symposium*, Cape Cod, MA, USA, 2001.

[2] V. Perera and M. Veloso, Learning to understand questions on the task history of a service robot, *The 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, Lisbon, Portugal, DOI: 10.1109/ROMAN.2017.8172318, 2017.
[3] V. Andreas, A. A. S. Guwanan and W. Budiharto, Anita: Intelligent humanoid robot with self-learning capability using Indonesian language, *2019 the 4th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*, Tokyo, 2019.
[4] *Google Speech to Text*, https://cloud.google.com/speech-to-text, 2019.
[5] *HappyBot: Intelligent Humanoid Robot*, https://www.youtube.com/watch?v=-LrF-wO7UkQ, 2020.
[6] M. Feng, B. Xiang, M. R. Glass, L. Wang and B. Zhou, Applying deep learning to answer selection: A study and an open task, *IEEE Workshop on Automatic Speech Recognition and Understanding*, pp.813-820, 2015.
[7] W. Yin, K. Kaan, M. Yu and H. Schütze, Comparative study of CNN and RNN for natural language processing, *arXiv Preprint arXiv:1702.01923v1*, 2017.
[8] Y. Omae, M. Mori, T. Akiduki and H. Takahashi, A novel deep learning optimization algorithm for human motions anomaly detection, *International Journal of Innovative Computing, Information and Control*, vol.15, no.1, pp.199-208, 2019.
[9] K. Komiya, Y. Abe, H. Morita and Y. Kotani, Question answering system using Q&A site corpus Query expansion and answer candidate evaluation, *SpringerPlus*, vol.2, 2013.
[10] *Python Programming Language*, http://python.org, 2019.
[11] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, *International Conference on Computer Vision and Pattern Recognition*, 2001.
[12] W. Budiharto and A. D. Cahyani, Behavior-based humanoid robot for teaching basic mathematics, *Internetwork Indonesia Journal*, vol.9, no.1, pp.33-37, 2017.
[13] W. Budiharto and P. C. B. Rumondor, Modeling of natural interaction and algorithm of intelligent humanoid robot for education, *2017 IEEE International Conference on Mechatronics and Automation (ICMA)*, Takamatsu, Japan, pp.172-176, 2017.