International Journal of
Applied Sciences and Smart Technologies

Volume 03, Issue 01, June 2021

p-ISSN 2655-8564 & e-ISSN 2685-9432

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, page i

p-ISSN 2655-8564, e-ISSN 2685-9432

CONTENTS

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, page ii

p-ISSN 2655-8564, e-ISSN 2685-9432

EDITORIAL BOARD

Contact us
International Journal of Applied Sciences and Smart Technologies
Faculty of Science and Technology
Sanata Dharma University
Kampus III Paingan, Maguwoharjo, Depok, Sleman
Yogyakarta, 55282
Phone : +62 274883037 ext. 523110, 52320
Fax : +62 272886529
Email : editorial.ijasst@usd.ac.id
Website : http://e-journal.usd.ac.id/index.php/IJASST

IJASST is an open-access peer-reviewed journal that mediates the dissemination of research and studies conducted by academicians, researchers, and practitioners in science, engineering, and technology.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages iii-iv

p-ISSN 2655-8564, e-ISSN 2685-9432

PREFACE

Volume 3 Issue 1 of *International Journal of Applied Sciences and Smart Technologies* (IJASST) contains 12 research articles. In the present issue, the authors are from three countries, namely, the United States of America, India and Indonesia. It covers research results in the areas of Mathematics, Science, Engineering and Technology.

Submitted papers were peer-reviewed using the open journal system (OJS). After the review process, accepted papers of the journal are publicly available at the website of IJASST. Readers can access the published papers free of charge.

Currently, we obtain increasing interests of researchers submitting their results to be published in IJASST. With this fact, we are sure that IJASST has gained more attention from the scientific community. Please, enjoy reading our IJASST papers!

Dr. I Made Wicaksana Ekaputra
Editor in Chief
IJASST

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages iii-iv

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

Frequency Distribution Fitting for Electronic Documents

Arockia David Roy Kulandai[1,2,*]

[1]*Department of Computer Science, Klingler College of Arts and Sciences, Marquette University, Milwaukee, Wisconsin, U.S.A.*
[2]*St. Xavier's College (Autonomous), Ahmedabad, Gujarat, India*
[*]*Corresponding Author: david.roy@marquette.edu*

(Received 19-09-2020; Revised 21-09-2020; Accepted 21-09-2020)

Abstract

Studies of frequency distributions of natural language elements have identified some distributions that offer a good fit. Using electronic documents, we show that some of these distributions cannot be used to model the frequency of bytes in electronic documents even if these documents represent natural language documents.

Keywords: Frequency fitting, quantitative linguistics, phase change memories

1 Introduction

Mathematical linguistics has studied the frequency of phonemes, words, and graphemes in natural languages. At its best, this work is used to obtain linguistical insights or even applications. For example, the Flesch Reading Ease Index [1] uses a combination of average word length and average sentence length. Best [2] still upholds its usefulness but notices that word length and sentence length are only indirectly related to readability. Our own motivation does not stem from linguistics but from the study of new non-volatile memory devices and their integration into future systems. We

Volume 3, Issue 1, pages 1–10
p-ISSN 2655-8564, e-ISSN 2685-9432

are interested in researching how to minimize bitflips in Phase Change Memories (PCM) [3]. PCM are a new non-volatile memory technology that offer bye-addressability, very high density, non-volatility, high retention, and high capacity. Unfortunately, PCM exhibit limited endurance. They use energy only while reading and writing, and usually writing consumes most of the energy. The number of bitflips caused by overwriting electronic documents of one kind by documents of the same kind depends on the encoding. For example, the web-browser cache contains HTML documents which could be placed in the same area of a PCM. To find good encodings, we want to model the frequency of graphemes in these documents [3]. The most frequent encoding for internet documents is UTF-8 so that our graphemes are bytes.

Here, we apply the methods of mathematical linguistics to modelling the frequency of bytes. Linguists are interested in language and graphemes are important as carriers of information on phonems. Unlike linguistics, we are interested in the effects of storing graphemes instead of using them. This makes for important differences. For instance, a linguist is not likely to make a distinction between capital letters and non-capital letters. Similarly, a linguist might conflate equivalent spellings, for example, the English and the US English versions of “tre” and “ter”, the recent abolition of the German letter “ß” in favor of “ss”, or even remove accents in Spanish.

Linguistics has shown that the frequency distribution of graphemes can be modelled by one or two parameter distributions successfully. Our results show that distribution fitting is less successful for bytes than for letters and phonems. Our research has convinced us that modelling a broad category such as text documents using distributions and parameters fitted to one corpus does not translate to another corpus. Evaluation of byte overwrites using these models are dangerous. Fortunately, we did find an encoding strategy that leads to energy savings for a broad class of electronic documents [3].

2 Research Methodology

We observed that encoding, e.g. utf-8, utf-16, ASCII, has a strong impact on the number of bits over-written when string text based electronic documents. This translates immediately into energy savings because each bit over-write costs energy. Also, each

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

bit-write is potentially destructive of the cell. We, therefore concentrated on HTML files stored, for example, in a browser's cache and to a lesser extent on text files. For comparison, with results in linguistics [1], [4], [5], [6] we also extracted pure text content from HTML files by gathering long text between the paragraph meta tag if the text was at least 50 bytes long. This excludes instances where the webpage used a paragraph meta tag only as a structural element. We also only processed letters and did not include punctuations or space. We collected corpora from Internet newspaper articles, Wikipedia, and the Project Gutenberg library of books in four European languages namely, English, German, Spanish and French. Each corpus contained at least 10 MB of raw data. We gathered ten corpora for English and five each for the other languages. For each corpus, we then calculated the frequency of each letter in the language or the frequency of each possible byte. We then fitted various distributions proposed in the linguistics literature to the frequency tables we obtained. For fitting we used Python's SciPy module. We minimized the relative sum of squared differences between the ordered relative frequency of the letters or bytes and the prediction by the distribution. Since the distribution has one, two, or three parameters, this means minimizing a function of one, two, or three variables. For each distribution, and for each of the 25 corpora, we tabulated the best fitting parameters and the goodness of fit for bytes.

3 Distributions

Zipf is an ancestor of modern quantitative linguistics, but the distribution named after him is also used almost as a default when modelling uneven usage of resources or uneven sizes in Computer Science. He ranked words in descending order of frequency of occurrence and observed that the frequency of the i^{th} word is proportional to $1/i$. Thus, we fit an ordered array of n descending frequencies with an array:

$$[\alpha/1, \alpha/2, \alpha/3, \ldots, \alpha/n]$$

where α is chosen so that the array sums up to one, which means that α is the inverse of the n^{th} harmonic number. Over time, many other distributions have been proposed to

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

model frequency of elements in natural languages. In his later works, Zipf generalized his distribution, matching the ordered array of n descending frequencies with

$$\left[{}^{\alpha}/_{1^{\beta}}, {}^{\alpha}/_{2^{\beta}}, {}^{\alpha}/_{3^{\beta}}, \dots, {}^{\alpha}/_{n^{\beta}}\right],$$

where β is a parameter of the text and α is calculated from β and the length n of the frequency array because the Probability Density Function (PDF) needs to sum up to 1. With other words, the frequency of the i^{th} most frequent item, denoted by f_i, is

$$f_i \sim {}^{1}/_{i^{\beta}},$$

where $\sim$ denotes proportionality. This distribution is also known as the Power Law distribution. Mandelbrot generalized the Zipf distribution by adding a second independent parameter γ so that

$$f_i \sim {}^{1}/_{(i+\gamma)^{\beta}}.$$

The Good distribution [7] is a parameter-less distribution where

$$f_i \sim \sum_{j=i}^{n} \frac{1}{j}.$$

We parameterize the Good distribution by setting

$$f_i \sim \sum_{j=i}^{n} \frac{1}{j^{\alpha}}.$$

In addition, we went through a list of distributions given by Li and Miramontes [5].

Exponential:	$f_i \sim \exp(-\alpha i)$
Logarithmic:	$f_i \sim 1 - \alpha \log(i)$
Quadratic Logarithmic:	$f_i \sim 1 - \alpha \log(i) - \beta (\log(i))^2$
Weibull:	$f_i \sim \log((n+1)/i)^{\alpha}$
Cocho – Beta:	$f_i \sim (n+1-i)^{\beta}/i^{\alpha}$
Frappat:	$f_i \sim \beta i + \exp(-\alpha i)$
Yule:	$f_i \sim \beta^i / i^{\alpha}$
Menzerath-Altmann:	$f_i \sim \exp\left(-\frac{\beta}{i}\right)/i^{\alpha}$

The actual value of the PDF of a distribution with $f_i \sim \psi(i, \alpha, \beta)$ is $c\big(\psi(i, \alpha, \beta)\big)$, where $1/c$ is equal to $\sum_{i=1}^{n} \psi(i, \alpha, \beta)$. The purpose of c is to ensure that the PDF sums up to 1.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

n is the number of symbols obtained, namely, $n = 256$ for bytes and $n =$ the number of symbols in a language. We are following the notation of Li and Miramontes [5], which has idiosyncrasies. For some values of β, Yule and Menzerath-Altmann are virtually indistinguishable. What Li and Miramontes call the Yule distribution is in fact not the well-known Yule-Simon distribution. The Yule-Simon distribution would have $f_i \sim \alpha B(i, \alpha + 1)$, where B is the beta function, but is not suited for frequency matching.

4 Results

There are two criteria for a distribution fit for modelling. Most importantly, the distribution should predict the frequency well. We measure this by calculating the sum of the differences squared and dividing it by the number of symbols. The number of symbols n is equal to 256 when we process raw documents, consisting of bytes. For text, it is just the total number of letters that can appear. To allow comparisons between text and raw data, we divide by n.

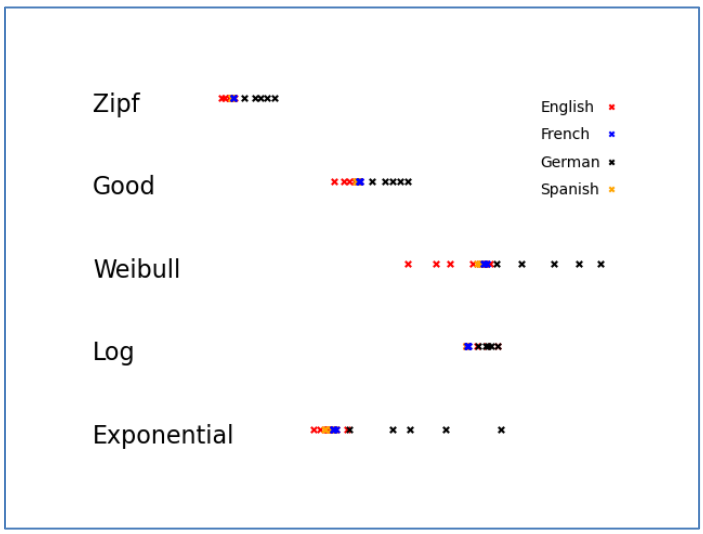


Figure 1. Distribution of the parameter for fitted one parameter distributions.

The second criterion is good clustering of the parameters. If two different corpora can be fitted well to the same distribution but with widely different parameters, then either we have too many parameters or the parameters are specific to one corpus. In the first case, we are better off with a distribution with less parameters and in the second case

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

the distribution with these parameters does not generalize and is not suitable for modelling.

For one parameter distributions, the fitted parameters lie close together and often in bands determined by the language, Figure 1. Only the parameters for German raw documents are more spread out in the case of the Weibull distribution and the Exponential distribution. In Figure 1, we plotted the sole parameter along the x-axis multiplying the parameter for the Logarithmic distribution by 10 and the parameter for the Exponential distribution by 20. Because the best fitting parameters in general appear in small ranges with sometimes differences between the languages, we conclude that modeling byte distribution with a single parameter will apply across a broad spectrum of corpora as long as they are in the same language.

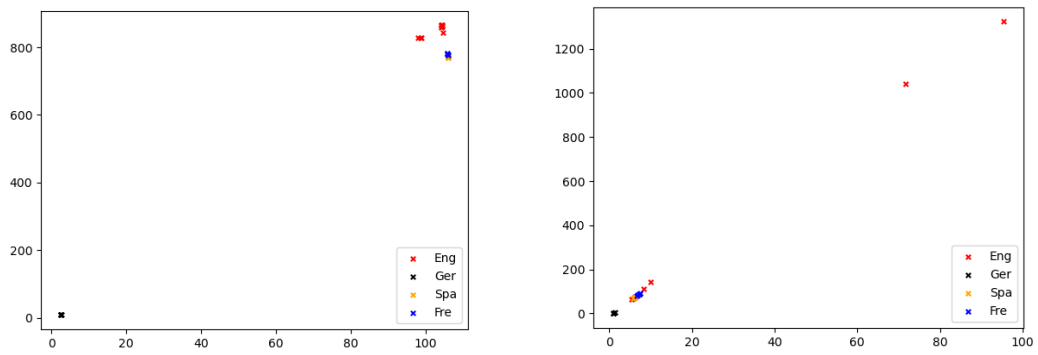


Figure 2. Parameters for Zipf Mandelbrot for text (left) and raw HTML (right) corpora.

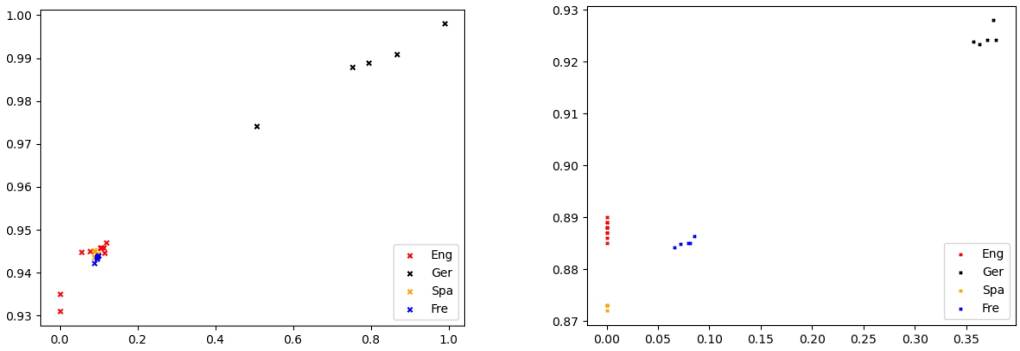


Figure 3. Parameters for Yule for text (left) and raw HTML (right) corpora.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

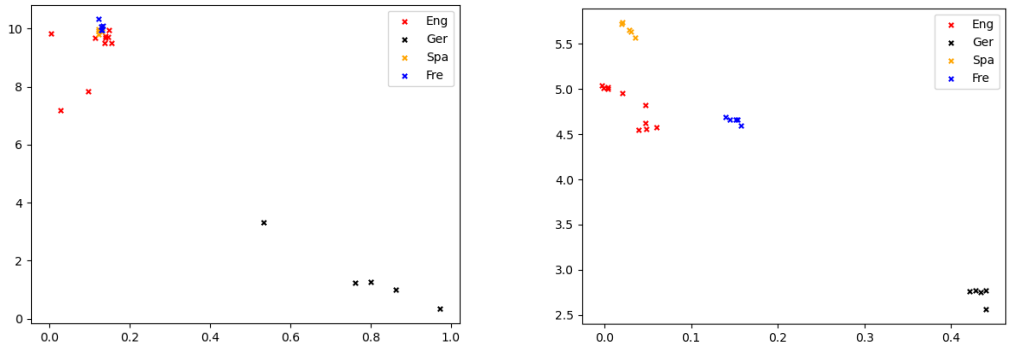


Figure 4. Parameters for Cocho-Beta for text (left) and raw HTML (right) corpora.

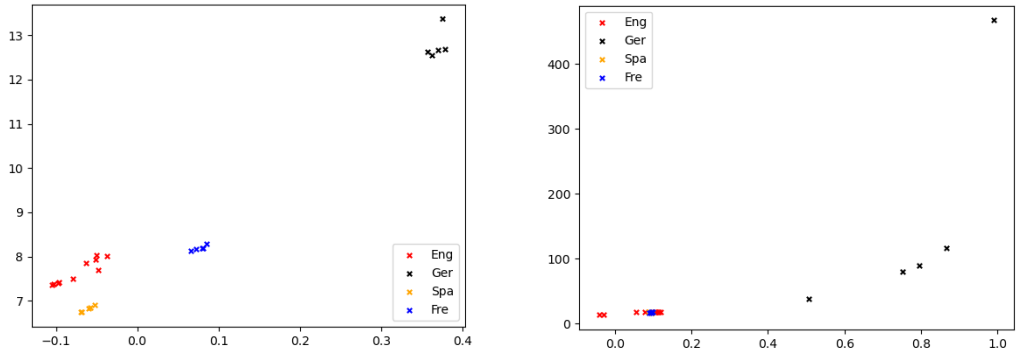


Figure 5. Parameters for Menzerath-Altmann for text (left) and raw HTML (right) corpora.

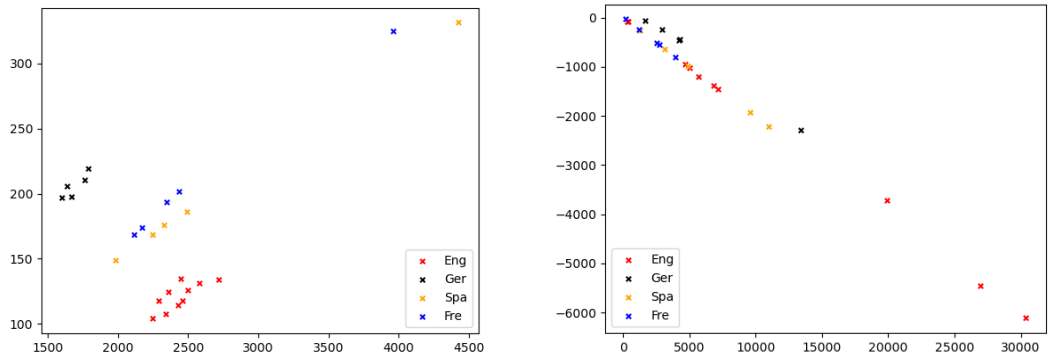


Figure 6. Parameters for Quadratic logarithmic for text (left) and raw HTML (right) corpora.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

Table 1. Range and average of goodness of fits for distributions and language corpora.

	Method	Range of Fits (Text)	Text Avg	Range of fits (Raw)	Raw Avg
English	Zipf	0.008573 - 0.009917	0.009379	0.005736 - 0.008065	0.006442
	Good	0.004025 - 0.004987	0.004601	0.003446 - 0.005398	0.004305
	Logarithmic	0.002510 - 0.002940	0.002745	0.002630 - 0.005295	0.004440
	Weibull	0.002129 - 0.002758	0.002481	0.001313 - 0.002383	0.001534
	Exponential	0.000649 - 0.000908	**0.000795**	0.000109 - 0.000258	**0.000208**
	Zipf-Mandelbrot	0.000664 - 0.000901	0.000801	0.000115 - 0.000175	0.000143
	Yule	0.000649 - 0.000908	0.000795	0.000109 - 0.000167	0.000134
	Cocho-Beta	0.000507 - 0.000620	**0.000568**	0.000109 - 0.000173	0.000146
	Quadratic log	0.060180 - 0.064961	0.062800	0.015871 - 0.023709	0.019974
	Menzerath-Altmann	0.000627 - 0.000838	0.009440	0.000102 - 0.000160	**0.000132**
	Frappat	0.000651 - 0.000814	0.009440	0.000109 - 0.016701	0.004980

	Method	Range of Fits (Text)	Text Avg	Range of fits (Raw)	Raw Avg
German	Zipf	0.003868 - 0.004299	0.004149	0.001605 - 0.002684	0.002095
	Good	0.001315 - 0.001661	0.001509	0.000785 - 0.001904	**0.001209**
	Logarithmic	0.004030 - 0.004748	0.004310	0.005799 - 0.021800	0.013460
	Weibull	0.000444 - 0.000653	**0.000521**	0.000511 - 0.006085	0.003124
	Exponential	0.001871 - 0.002235	0.002009	0.003151 - 0.016770	0.009970
	Zipf-Mandelbrot	0.001061 - 0.001414	0.001194	0.001236 - 0.002684	0.001882
	Yule	0.000431 - 0.000649	0.000515	0.000401 - 0.002671	**0.001493**
	Cocho-Beta	0.000323 - 0.000500	**0.000392**	0.000344 - 0.002625	0.001416
	Quadratic log	0.067545 - 0.071953	0.069897	0.028288 - 0.059766	0.044443
	Menzerath-Altmann	0.000431 - 0.000649	0.009389	0.000401 - 0.002671	**0.001493**
	Frappat	0.001353 - 0.001763	0.009389	0.002006 - 0.004427	0.003596

	Method	Range of Fits (Text)	Text Avg	Range of fits (Raw)	Raw Avg
Spanish	Zipf	0.011111 - 0.011337	0.011283	0.006152 - 0.006370	0.006254
	Good	0.006357 - 0.006546	0.006494	0.00422 - 0.004456	0.004333
	Logarithmic	0.004410 - 0.004520	0.004474	0.004648 - 0.005131	0.004877
	Weibull	0.003246 - 0.003374	0.003314	0.001400- 0.001472	0.001435
	Exponential	0.001116 - 0.001281	**0.001193**	0.000163 - 0.000208	**0.000184**
	Zipf-Mandelbrot	0.001123 - 0.001297	0.001195	0.000063 - 0.000114	0.000085
	Yule	0.001116 - 0.001281	**0.001193**	0.000095 - 0.000126	**0.000111**
	Cocho-Beta	0.000817 - 0.000972	0.000891	0.000135 - 0.000159	0.000147
	Quadratic log	0.074113 - 0.074588	0.074396	0.018778 - 0.020212	0.019447
	Menzerath-Altmann	0.001068 - 0.001248	0.011604	0.000095 - 0.000126	**0.000110**
	Frappat	0.001021 - 0.001183	0.011604	0.000094 - 0.000149	0.000117

	Method	Range of Fits (Text)	Text Avg	Range of fits (Raw)	Raw Avg
French	Zipf	0.010303 - 0.010482	0.010381	0.006304 - 0.006504	0.006358
	Good	0.006042 - 0.006157	0.006093	0.004383 - 0.004586	0.004442
	Logarithmic	0.005219 - 0.005289	0.005244	0.005038 - 0.005300	0.005148
	Weibull	0.003713 - 0.003848	0.003778	0.001451 - 0.001534	0.001473
	Exponential	0.002739 - 0.002869	**0.002795**	0.000165 - 0.000203	**0.000181**
	Zipf-Mandelbrot	0.002720 - 0.002860	0.002776	0.000104 - 0.000128	0.000113
	Yule	0.002677 - 0.002797	0.002736	0.000093 - 0.000120	**0.000102**
	Cocho-Beta	0.002216 - 0.002328	**0.002270**	0.000109 - 0.000141	0.000120
	Quadratic log	0.074867 - 0.075727	0.075307	0.020078 - 0.020616	0.020275
	Menzerath-Altmann	0.002677 - 0.002797	0.012376	0.000093 - 0.000120	**0.000102**
	Frappat	0.002656 - 0.002783	0.012376	0.000129 - 0.016198	0.003352

Volume 3, Issue 1, pages 1–10
p-ISSN 2655-8564, e-ISSN 2685-9432

For two parameter distributions, the situation is more difficult. In some cases, such as the Zipf-Mandelbrot distribution, Figure 2, language specific parameters are nicely clustered by language if we only look at text. If, however, we look at raw text, then the English cluster dissolves. For the five German corpora, the parameters are too widely distributed for text and raw files. We attribute this to over-fitting, a phenomenon well known from machine learning. Fitting Zipf-Mandelbrot "learns" the corpus but not the general category. In addition, we observe that the parameters for raw HTML lie along a line, indicating a linear relationship between the two parameters. This indicates that the distribution should be made into a one-parameter distribution. In fact, as can be seen from Table 1, the goodness of fits for Zipf-Mandelbrot is better than for the Zipf distribution but still at the worst range of two parameter distributions. Similarly, the parameters for Menzerath-Altmann are nicely clustered for text but lie on a one-dimensional curve for the raw corpus. For the Quadratic logarithmic distribution, again the results differ between text and raw corpora. For this reason alone, a number of distributions suitable in linguistics are not suitable to model byte frequencies. We refer readers to see Figures 2-6 for the details of our illustration results.

5 Discussion

Our interest is not in linguistics but modelling the overwriting of non-volatile memory. Therefore, our frequency tables make a distinction between capital and non-capital letters. For a linguist, this distinction is probably artificial. Also, unlike for example, Li and Miramontes [5], we do not conflate the letters that differ only in an accent or *umlaut* because they are encoded differently even though they can be considered the same letter. We gave results for texts as a comparison point for raw data. For example, we learned that some distributions such as Zipf-Mandelbrot overfit for raw data and are therefore probably useless for analytics while this does not happen for text. Overall, just as in the work of Li and Miramontes, the Cocho-Beta distribution and the Yule distribution allow best fits without the overfitting phenomenon. Among single parameter distributions the Zipf or Power Law distribution does not fare so well as it is outperformed by the Exponential distribution and by the parametrized Good distribution.

Volume 3, Issue 1, pages 1–10

p-ISSN 2655-8564, e-ISSN 2685-9432

6 Conclusion

Frequency modelling of bytes in electronic documents can be done with the Exponential distribution. While a better fit can be achieved with the Menzerath Altmann distribution or the Cocho-Beta distribution, their parameter range is not only language but also corpus specific. It is hard to see how scientific conclusions can be obtained with such variety. When restricted to text, our observation is not valid.

References

[1] R. Flesch. "A new readability yardstick." *Journal of Applied Psychology,* **32** (3), 221, 1948.

[2] K. H. Best. "*Sind Wort-und Satzlänge brauchbare Kriterien zur Bestimmung der Lesbarkeit von Texten? In: Wichter, Sigurd/Busch, Albert (eds.) Wissenstransfer Erfolgskontrolle und Rückmeldungen aus der Praxis*." Peter Lang Verl, Frankfurt, 2006.

[3] A. D. R. Kulandai and T. Schwarz. "Content-Aware Reduction of Bit Flips in Phase Change Memory." *IEEE Letters of the Computer Society*, 2020.

[4] B. Krevitt and B. Griffith. "A Comparison of Several Zipf-Type Distributions in Their Goodness of Fit to Language Data." J*ournal of the American Society for Information Science*, **23** (3), 220, 1972.

[5] W. Li and P. Miramontes. "Fitting Ranked English and Spanish Letter Distribution in U.S and Mexican Presidential Speeches." *Journal of Quantitative Linguistics,* **18** (4), 359–380, 2011.

[6] C. Manning and H. Schütze. "*Foundations of Statistical Natural Language Processing.*" MIT Press, 2003.

[7] H. Pande and H.S. Dhami. "Mathematical Modelling of Occurrence of Letters and Word's Initials in Texts of Hindi Language." *SKASE Journal of Theoretical Linguistics*, **7** (2), 2010.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

Artificial Generation of Realistic Voices

Dhruva Mahajan[1,*], Ashish Gapat[1], Lalita Moharkar[1],

Prathamesh Sawant[1], Kapil Dongardive[1]

[1]*Department of Electronics and Telecommunication Engineering, Xavier Institute of Engineering, Mahim, Mumbai, Maharashtra, India*

[*]*Corresponding Author: dhruvam17@gmail.com*

(Received 17-07-2020; Revised 27-08-2020; Accepted 15-09-2020)

Abstract

In this paper, we propose an end-to-end text-to-speech system deployment wherein a user feeds input text data which gets synthesized, variated, and altered into artificial voice at the output end. To create a text-to-speech model, that is, a model capable of generating speech with the help of trained datasets. It follows a process which organizes the entire function to present the output sequence in three parts. These three parts are Speaker Encoder, Synthesizer, and Vocoder. Subsequently, using datasets, the model accomplishes generation of voice with prior training and maintains the naturalness of speech throughout. For naturalness of speech we implement a zero-shot adaption technique. The primary capability of the model is to provide the ability of regeneration of voice, which has a variety of applications in the advancement of the domain of speech synthesis. With the help of speaker encoder, our model synthesizes user generated voice if the user wants the output trained on his/her voice which is feeded through the mic, present in GUI. Regeneration capabilities lie within the domain Voice Regeneration which generates similar voice waveforms for any text.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

Keywords: Speech synthesis, speaker encoder, synthesizer, Text-to-Speech, vocoder

1 Introduction

In the near future, everything around us will be voice operated. With the growing trends of Alexa and Google home, advancements are being made to create an environment of Artificial Intelligence and its operating medium would be Voice. With the advent of signal processing, voice signals have extensive upgradation in terms of global standards and kept on challenging for better platform in all its forms of uses. Right from voice screeching out of airhorn to voice search engines in smart phones, voice applications have always been vital in any or all of their feats. Artificial generation itself is a form of voice cloning that is implemented with the help of neural nets that help with generating the Mel-spectrograms. This process intensifies on the input, where the characters present are synthesized into signal waveforms which are digitally spectograms. These spectograms are then coupled and then linked through a vocoder, which generates voice corresponding to the characters that are given as input.

The goal of the this paper is to build a TTS system which can generate natural speech for a wide variety of speakers which are absent throughout the process. Our model can run in real time by implementing the mic function. This is possible by achieving a powerful form of voice cloning. The output must be aligned in such a manner that runs on correct lines to provide a clone of the dataset that is trained within the system. The output speech must allign with the exact speaker voice picked from the dataset. This voice gets matched with the help of RNN, which cycles the implementation process wherein the dataset voice gets linked with input text. This implementation works throughout for all the speakers whether present or absent. When a speaker implements mic function, the speaker encoder does not operate.

Uncertainties do arise as we train the model with the speaker's reference speech (trained dataset):

- The output utterance is slightly composed.
- Dataset voices could be identical, depending on the dataset picked (VCTK and LibriSpeech).

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

- Voice output obtained, when implementing through the mic tends to be rough. This is because while recording in real time, it is very important to check the environment around. Mic integrated with system is highly sensitive and tends to pick smallest utterance.
- Recording a large amount of high-quality data for many speakers seems rather impractical.

The approach which we deploy, decouples speaker modeling from speech synthesis by independently training a speaker effective embedding network that traces and sequentially captures the character space of speaker characteristics and training a high-quality TTS [1], [2]. Taking findings from a 2017 research paper, Google- Tacotron, where we figure out the synthesis and voice generation process [3]. We took implementation steps from a deep learning architecture that we researched and coupled it with WaveNet [3], [4]. WaveNet is a neural network which acts as a vocoder, to convert mel-spectograms into corresponding voice signals. We now had to check the process and operation for which we used two public datasets that are Librispeech and VCTK. After the respective implementation we got our process confirmed for synthesis and started the model. To run with time and technology, we implement a neural network architecture of our research pertaining to synthesis of voice. This research model is a deep learning text-to-speech model. Text-to-Speech describes that a string of text would be converted to speech output. There are three major operations.

- Speaker Encoder- Where the embedded text feeded in is sent to a convolution bank highway network. This convolution bank is nothing but a deep learning tool which breaks our text in separate characters. Breaking the string of characters allows the network to work on each character modulation, making it more in tune with voice rather sounding modulated. i.e. For example- GOOD MORNING in the convolution bank will be presented as G-O-O-D M-O-R-N-I-N-G.
- Synthesis- We implement trained datasets in our paper. These trained datasets arefeeded and loaded before operating the model. During the synthesis our character embedding which is disintegrated character by character is sent to attention, where the character is coupled with trained dataset voice signal.

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

- Vocoder- This is the final part of our model where the signals which are bindedas spectrogram is converted into voice. This operation uses WaveNet neural network [5]. (WaveNet was developed in deepmind labs, which is a AI research wing of google.) WaveNet as a vocoder acts as a binder, where all end results (melspectrograms obtained) are combined. These results are the voice signal modulation that take place along with each character given as input. Explaining mathematically, in theory, our input text gets converted into an algebraic equation (which consists are input characters), this equation is then simplified by taking K constant. (K constant here is the dataset selected). This simplified equation is then collected and generated into series function equation (This is the desired output). With spectrograms for these character strings, this waveform passes through the WaveNet which perform natural language processing, giving voice as the end result.

2 Methodology

Text-to-Speech implementation for a speaker model to pick up character embeddings, the model needs to be well disintegrated in three subsets as clarified. These subsets work in a sequential manner providing speech in a uniform and desired manner. Putting light on each model now will magnify each role clearly with technical overviews and details. See Figure 1 for the block diagram.

Figure 1. Block diagram

Speaker encoder illustration is shown in Figure 2. Speaker encoder plays an important role by estimating the embedding from various text characters that is feeded as input. While implementing Text-to-Speech model, it's up to the developer by what means it should be operated. To cut the unnecessary word-word translation by using a conventional Text-to-Speech system, where the text is sent in a string and the output obtained has passive speech, we use deep learning architecture, where the input text is

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

feeded as the prenet (pre-feeded information/local information). This prenet is then sent to 1D-CBHG. This is one-dimensional convolution bank highway network. This block is the place where the string is broken in separate character by character formation [5]. Example: GOOD G-O-O-D.

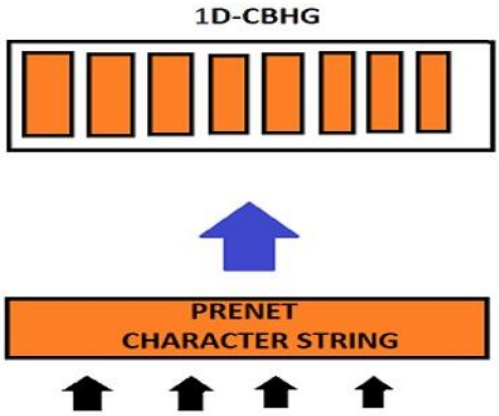


Figure 2. Speaker encoder

Synthesizer is illustrated in Figure 3. Functions of daily text-to-speech is evident in our voice search engines, smart electronics and various voice modulating devices that inherit the use of synthesizer. Text-to-speech control system employs an easy-to-use transport medium from which a user can control most text-to-speech variating functions without any prior training. Now, the synthesizer basically works in the areas of translating a plurality of discontinuous user-selected part of text in an independent form of target application into an audio output that resembles the sound of human speech. To generalize, the synthesizer is the core of the Text-to-Speech engine, where it mainly focuses on tracing and adapting the text input in a sequence and delivers audio sample of it [5].

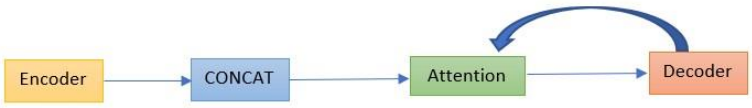


Figure 3. Synthesizer block

In the block diagram represented above, the speaker encoder gets linked to the Concat block. In this Concat block, the character broken down in the CBHG, gets

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

forwarded as a string with intervals between each character to the attention. Attention is a deep learning tool where each character is concentrated and constant is removed, and binded with the trained dataset. This trained dataset is forwarded from the Decoder, which acts as a gate for audio signal and couples it with the data in Concat. The constant removed is called 'k-constant', which is used for signal processing.

Vocoder: Simplicity in the output should be in a form such that all the speech is synthesized in the manner that the corresponding input is heard in tone that is tuned to match the tempo of the unseen speakers' voice. For this purpose, a vocoder is used where it captures the characteristic elements of the audio signal and then uses this characteristic signal to affect other audio signals. It is also dubbed as a "talking synthesizer" for its ability to fine-tune the synthesized signal in accordance to vocal frequency. See Figure 4 for Log Mel-waveforms.

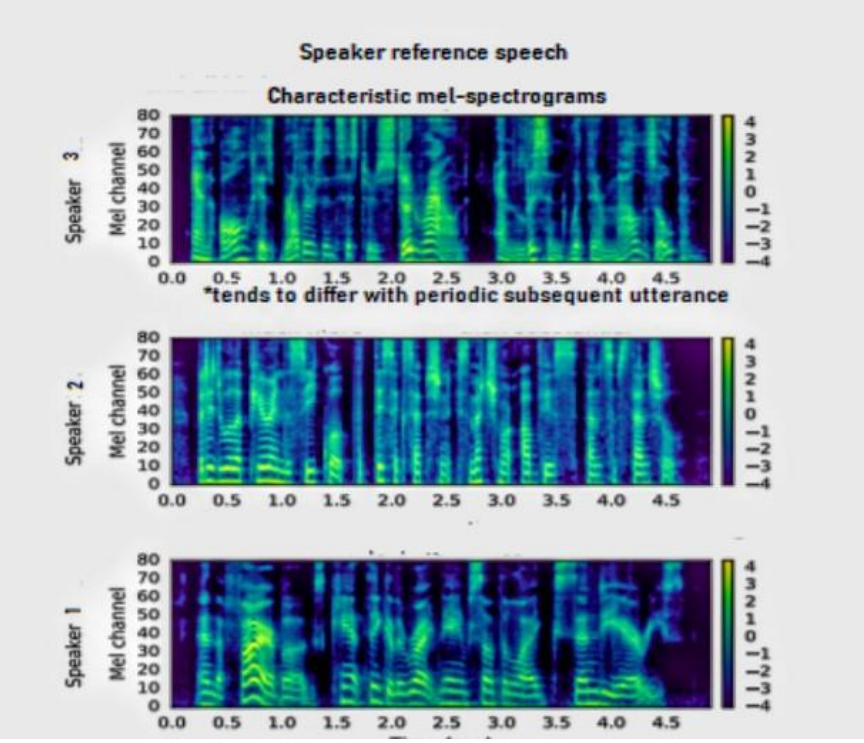


Figure 4. Log Mel-waveforms

The vocoder provides a bank of multiple bandpass filters which dissociate the input signal into narrow spectral slices. Consider, we excite channel 'k' of vocoder with the input signal $a(nT)\cos(wknT)$ for $n = 0,1,2,3,4,5, \ldots$ where wk is the centre frequency

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

of the channel in radians per second, T is the sampling interval in seconds and bandwidth of $a(nT)$ is smaller than channel bandwidth. We regard this input signal as an amplitude modulated sinusoid. The component $\cos(wknT)$ can be called the carrier wave, while $a(nT) > 0$ is the amplitude envelope.

If the phase of each channel filter is linear in frequency within the passband (or at least across the width of the spectrum) and if each channel filter has a flat amplitude response in its passband, then the filter output will be, by the analysis of the previous section. $Yk(n) \sim a[nT - D(wk)]\cos(wk[nT - P(wk)])$

Creating a GUI is illustrated in Figure 5 as follows. With the reach of applications with software, the main over layer and the visible interaction is the Graphical User Interface (GUI) which allows dynamic ability to the user for its functioning [6]. For our model we intended the working on an interface to allow the user to interact with the model. Creation of GUI also implements the ease of functions laid out at disposal of the user within a fixed framework. GUI requires the interaction to be in a flow that does not hamper the user and neither causes any sort of imbalance within the process. Taking for consideration, our model is heavily based on synthesis, making it completely oriented to user interaction (input/character embedding). So, with this, the GUI should be in a manner that allows easy flow of task within the same framework. To create the GUI for our model, we implement tkinter library of python. Within the tkinter package, there are many functions that are used to make things more organized and presentable. Tkinter allows us to make various frameworks, buttons and organizes the functions systematically. Major properties to include in the GUI is illustrated in Figure 5.

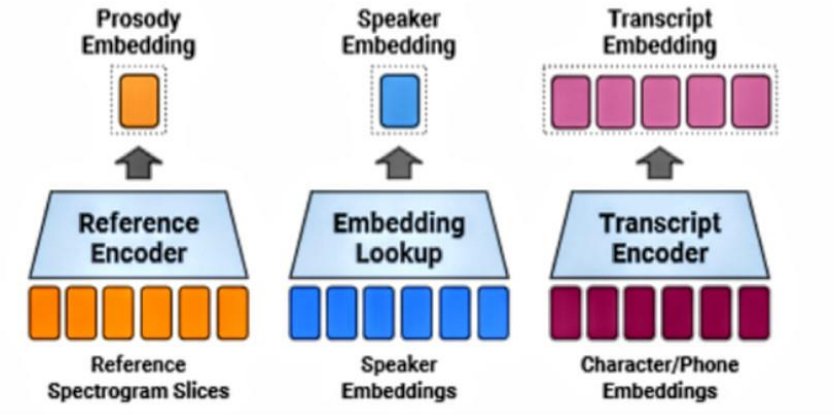


Figure 5. Main GUI components

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

3 Implementation

Real time Voice Cloning (Using SV2TTS): The entire approach to real time voice cloning is adapted on the basis of Transfer learning from speaker. i.e. dubbed as prosody transfer (voice styling implementation). It is a speaker verification to multi-speaker text-to-speech synthesis. It essentially defines the framework for voice cloning that barely requires 4-6 seconds of reference speech. It is majorly dependent on three early works from Google: the GE2E loss (Wan et al., 2017), Tacotron (Wang et al., 2017) and WaveNet (van den Oord et al., 2016). This proposed model is three-stage pipeline, as listed above in order. The google cloud services, google search engine, or google assistant make use of these same models.

Model Architecture of Speaker encoder, Synthesizer and Vocoder are shown in Figure 6.

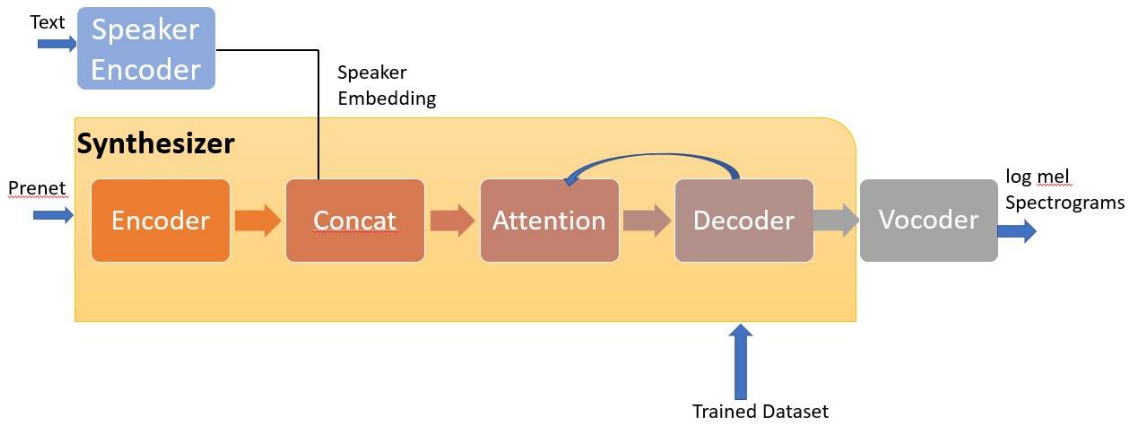


Figure 6. Three-stage pipeline for Text-to-Speech

Model Architecture of Speaker Encoder: It is a three-layer LSTM with 768 hidden nodes which is followed up by a projection layer comprising of 256 units. Although there is no reference in any of references present, as to what the projection layer defines. Hence, we round up to consider the overall function of the projection layer as a 256-output fully connected layer (per LTSM) which is repeatedly applied to each and every output of the LTSM. The inputs to this model are 40-channels log-mel spectrograms

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

with 25ms window width and a 10ms step. The desired output is the L2 normalized hidden state of the last layer, which works as a vector of 256 elements.

Model Architecture of Synthesizer: In the synthesizer implementation, the target Mel spectrograms present for the synthesizer provide more features than those used for speaker encoder which are computed from a 50ms window with a 12.5ms step and have 80 channels. We use a python implementation of LogMMSE algorithm, which is used for filtering the audio speech by erasing the noise in the early frameworks. Consequently, we train the synthesizer for 150k steps, comprising of a batch size of 144. The outputs per second is set to 2 for the decoder. While implementing, the architecture tends to provide speech synthesis to attain identical cloning to the unseen speaker. During this process there are losses that are accounted in the verge predicted and ground truth mel spectrograms. This is the L2 loss function. While training, the model is set to ground truth aligned (GTA). The reason for this is that if we do not set the model to GTA, then the synthesizer would produce different variations of the same utterance (text or embedding). Implementation of neural nets are shown in Figure 7.

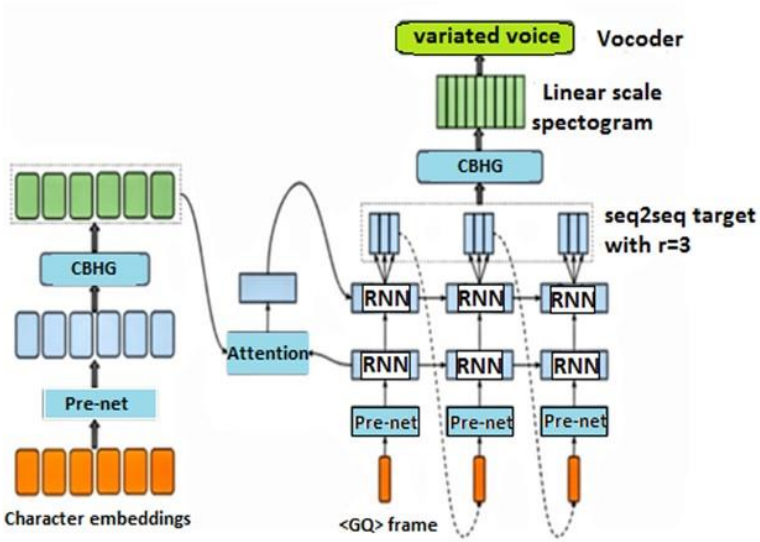


Figure 7. Implementation of neural nets

Model Architecture of vocoder The vocoder implemented in the model is WaveNet [1, 4]. WaveNet produces naturalness in TTS. This is the primary reason it is fairly used in Tacotron and SV2TTS. However, the efficiency of this neural net is too

Volume 3, Issue 1, pages 11–26
p-ISSN 2655-8564, e-ISSN 2685-9432

good coming at the cost of the speed. It is very slow and the slowest deep learning architecture at interference time. Even if this is a matter of stress on the implementation part, there are improvements that can be initiated on it. Googles own vocoder works at the rate of output of 8000 samples per second, which is by far not bad for a neural net. The model implemented is an open-source Pytorch implementation that is based on an RNN model deployed by Github user fatchord.

Zero-shot speaker adaption the speech characteristics to be synthesized are picked up from the audio signal. Zero-shot adaption is the ease with which the model gets linked with the training data from the unseen speaker which is not present during the training. It just requires time of 4-6 seconds for the model to generate new speech by synthesizing the speaker characteristics. Interference can cause the speaker information to get synthesized without knowledge of the input fed through. However, in our model, interference occurs with arbitrary untranscribed speech audio which does need the text to match with the synthesizer, thereby making the implementation hassle free and comparatively quick.

Dataset Used Two datasets which are public by nature for synthesis and speech training for vocoder network is used for implementation. VCTK comprises of 43 hours of clean speech from 109 speakers, among which most have British accents. We down sampled the audio to 25 kHz, trimmed leading and trailing silence (reducing the median duration from 3.3 seconds to 1.8 seconds). It's been split into three subsets: train, validation which has the same speakers as the train set and test which has 11 speakers held out from the train and validation sets. LibriSpeech comprises of the union of the two "clean" training sets, consisting 436 hours of speech from 1,172 speakers, sampled at 16 kHz. The speech is US English majorly, however since it is sourced from audio books, the vocals and style of speech can differ significantly between utterances from the same speaker. We reassembled the data into shorter utterances by force aligning the audio to the transcript using an automatic speech recognition (ASR) model and breaking segments on silence, reducing the median duration from 14 to 5 seconds.

Naturalness of Speech Clearer the person's voice is, crisper is the audibility factor that follows. This might tend to differ in cases of speech synthesis, where natural speech gets encoded and decoded with help of speech synthesizers and in our case deep

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

learning networks too. So, to arrange it cordially, there is prenet data, also termed as local information that gets coupled with character embeddings from the input. There is interference that plays a major role and thereby noise gets added on while processing through the highway networks. With subsequent processing and ability to trace character by character (using Fourier transform) we get the log-mel spectrogram through vocoder. After attaining the desired output, the major question is how much of the part is natural speech. As we implement two public datasets, we tried comparing the VCTK and LibriSpeech datasets to check the speech synthesis through synthesizers and vocoder. We tried comparing by using 11 unseen and seen speakers for VCTK and 10 unseen and seen speakers for LibriSpeech. Each of the comparison was conducted independently [2]. Comparison of time taken for synthesis is listed in Table 1.

Table 1. Comparison of time taken for synthesis

System	Speaker Information	VCTK	LibriSpeech
Ground truth	Same Speaker	4.67 ± 0.04	4.33 ± 0.08
Ground truth	Same Gender	2.25 ± 0.07	1.83 ± 0.07
Ground truth	Different Gender	1.15 ± 0.04B	1.04 ± 0.03
Embedding Table	Seen	4.17 ± 0.06	3.70 ± 0.08
Proposed Model	Seen	4.22 ± 0.06	3.28 ± 0.08
Proposed Model	Unseen	3.28 ± 0.07	3.03 ± 0.09

Speaker Similarity and Verification To check the speech having cleaner detail, we check the results of the above comparison of the two datasets. This is done to check whether the desired output is identical to the input given. From the comparison table, we get to know that the values delivered by VCTK tend to be an edge above the LibriSpeech dataset. The speech from the VCTK dataset is cleaner. This can be also understood by higher ground truth baselines in the VCTK. That makes the VCTK better, but on using LibriSpeech on VCTK model, it was visible that the output was better than that of VCTK model. This means that depending on the dataset and type of groundtruth, the similarity can be accomplished. For the part of speaker verification, on LibriSpeech, the synthesized speech is at most similar to the ground truth voices. The

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

LibriSpeech synthesizer obtains similar EERs of 5-6 % using reference speakers from both datasets, whereas the one trained on VCTK performs much worse, especially on out-of-domain LibriSpeech speakers.

4 Results and Discussion

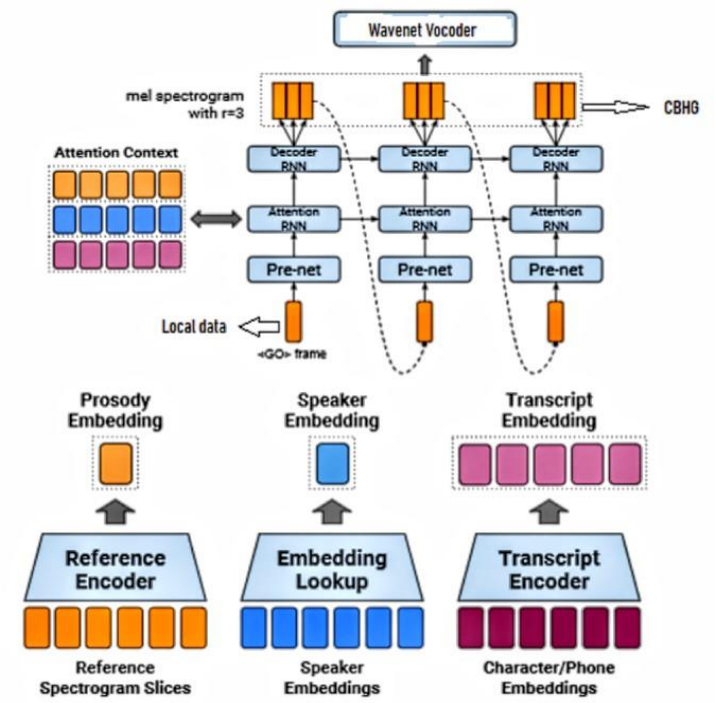


Figure 8. Proposed-method

Having all the details and implementation incorporated, we got our proposed design (see Figure 8) that would be implemented by us. We had to take in consideration every part as sequential support that will provide proportional synthesis and deliver the desired output. We had two procedures followed to obtain results.

Method 1: In the first method we used the train datasets for speech synthesis, which were LibriSpeech and VCTK respectively. The output obtained was clear and audible with the help of headphones. In open environment, the voice felt a little light, and required a speaker with an equalizer.

Method 2: This method is an advancement to our paper which could be extended further successfully with correct implementations. We tried feeding in our voice through

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

the mic, and tried cloning. The speech did get synthesized but the output obtained wasn't as clear as the previous case and was little gibberish in nature.

Audio Quality Analysis: The quality audio can be better heard by using speakers with better bass and equalizer in an open environment. There is a 2-3 seconds lag in setting the audio sample from the dataset which, again can be improved by trimming the audio samples. Within all testing periods, the model was reliable and highly efficient in terms of user interface. Further extension is possible by implementing user generated voice (refined), and miscellaneous dataset implementation of various accents and dialects.

Now, in accordance to the above model, we had to implement a GUI (see Figure 9), that can provide subsequent functioning of the above model and also should be user friendly to operate. We implemented Python tkinter package in accordance to the system model and created the model that equips and provides most of the proposed system through easy flow of framework design.

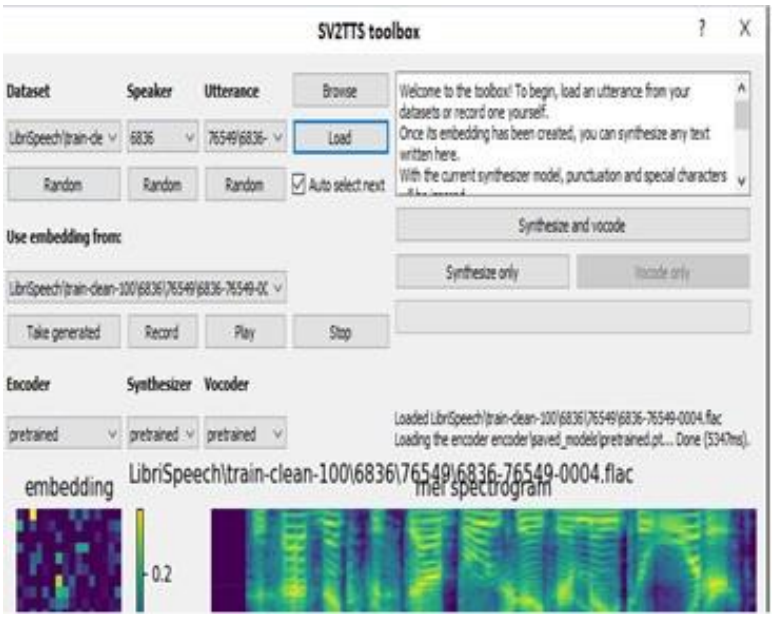


Figure 9. Output GUI

5 Conclusion

The entire concept of having voice cloning has always had some advancements tending to always grow for the better. Incorporating all the factors, we try instill the

Volume 3, Issue 1, pages 11–26
p-ISSN 2655-8564, e-ISSN 2685-9432

model with all datasets and pretrained data through which the model infers and tries to pullout the desired output claimed. To attain the desired output, we need to verify the speech to be identical to the trained unseen speech, so as to be sure. If the speech synthesis is not identical, we cannot term the output as the desired output. Such models have implications and applications mostly in the armed forces for providing stealth mode a new edge. Voice cloning can help make communication between borders, across seas and even on telephone a rounded-up mystery, which is the reason it benefits the armed forces greatly. The application also lies with media. People tend to use voice cloning for media concentration applications or entertainment media such as Dubsmash. This model also helps with regeneration of voice wherein, a certain person can communicate if he/she is disabled or lost ability of speech under certain circumstances. The voice cloning works for the future, wherein with on-time upgradation, and correct use we can use regeneration factor for taking over music and pop culture with a wave. With a sytem having a multi-area application domain. It is very important to place certain regulations and boundaries bounded by legal clauses for implementation. Major part of the model lies within the use of the Text-to-Speech system, which is the backbone of the model. Many systems integrated in artificial intelligence pick Text-to-Speech system as prime domain because of its broad area of implementation, right from voice assist, voice detection to voice cloning. Our model works with voice cloning and moves in direction of voice regeneration, which will be a major breakthrough in the near future. The proposed model does not attain human-like naturalness, despite the use of a WaveNet vocoder (along with its very high inference cost), in contrast to the single speaker results. This is a consequence of the additional difficulty of generating speech for a variety of speakers given significantly less data per speaker, as well as the use of datasets with lower data quality.

Acknowledgements

The corresponding author acknowledges all the co-authors and group members for co-operating and working with an optimistic mindset. Every work related to the paper required dedicated and devoted attention from the department in association, and

personal guidance from the project guide, Prof. Lalita Moharkar who stood by all along the buildup of the research model. Trying to walk the entire path from scratch required detailed reference help which acted as a walking stick, which were journal papers and technical papers present in international journals and publications.

References

[1] L. Wan, Q. Wang, A. Papir and I. L. Moreno, "Generalized end-to-end loss for speaker verification." *Acoustics, Speech and Signal Processing (ICASSP), IEEE International Conference*, 2018.

[2] Y. Jia, Y. Zhang, R. J. Weiss, Q. Wang, J. Shen, F. Ren, Z. Chen, P. Nguyen, R. Pang, I. L. Moreno and Y. Wu, "Transfer Learning from Speaker Verification to Multispeaker Text-To-Speech Synthesis." *Advances in neural information processing systems*, **31**, 4485–4495, 2018.

[3] Artificial Intelligence at Google – Our Principles. https://ai.google/principles/, 2018.

[4] A.V.D. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior and K. Kavukcuoglu. "Wavenet: A generative model for raw audio." *arXiv preprint 1609.03499*, (2016).

[5] J. Shen, R. Pang, Ron J. Weiss, M. Schuster, N. Jaitly, Z. Yang, Z. Chen, Y. Zhang, Y. Wang, R. J. Skerry-Ryan, R. A. Saurous, Y. Agiomyrgiannakis and Y. Wu. "Natural TTS synthesis by conditioning WaveNet on mel spectrogram predictions." *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2018.

[6] M. Grechanik, Q. Xie and C. Fu, "Creating GUI testing tools using accessibility technologies." *Proceedings of the IEEE International Conference on Software Testing, Verification, and Validation Workshops*, 243–250, 2009.

Volume 3, Issue 1, pages 11–26

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 27–34

p-ISSN 2655-8564, e-ISSN 2685-9432

Human Detection in Video Surveillance

Sushama Khanvilkar[1], Santosh Gupta[1], Hinal Rane[1],
Calvin Galbaw[1,*]

[1]*Department of Computer Engineering,*
Xavier Institute of Engineering, Mahim, Mumbai, Maharashtra, India
[*]*Corresponding Author: calving2012@gmail.com*

(Received 17-07-2020; Revised 31-07-2020; Accepted 01-08-2020)

Abstract

Recognition of the human activities in videos has gathered numerous demands in various applications of computer vision like Ambient Assisted Living, intelligent surveillance, Human-Computer interaction. One of the most pioneering techniques for Human Detection in Video Surveillance based on deep learning and this project mainly focuses on various approaches based on that. This paper provides an idea of solution to use video surveillance more effectively, by detecting any humans present and notifying the concerned people. The deep learning model, preferred for fast computation, Convolution Neural Network is used by stacking 3 blocks of layers on fully connected layers. This provided an identification of humans and naïve approach to eliminate inanimate human like objects such as mannequins.

Keywords: deep learning, CNN, human detection

1 Introduction

Human Activity detection is a major problem in smart videos surveillance. It is an elementary drawback in computer vision, i.e. to notice the activity of human in

Volume 3, Issue 1, pages 27–34
p-ISSN 2655-8564, e-ISSN 2685-9432

surveillance videos. These applicants need real time detection performance, but it is generally very time consuming to detect the actual activity. Since the use of CCTV, the cases of forced entries and robberies have decreased drastically. But the delay in response to such cases can cause problems. If the owner can get the notification of such events, the culprit can be caught red handed. It becomes important to alert the user by detecting what activity is been performed by the subjects [1-3].

2 Research Methodology

This prospective implementation was carried out using simple programming tools and cloud resources. The Convolutional Neural Network (CNN) is the most promising network to work with images and videos. Hence, developing an architecture using CNN was an optimal and efficient choice.

Implementation Design. In order to implement the system Modified AlexNet design which is trained on frames of video has been used.

Dataset size. 8 videos have been used as a dataset.

Sample size calculation. The sample size was chosen from multiple videos which satisfies the needs of the required datasets. Each video chosen have average of 8000 frames from which about 10% are taken into consideration. This is to reduce redundancy of the data.

Subjects and selection method. The dataset is formed of videos which are taken using CCTV cameras. These videos all include people trying to break into the shops and houses. Some videos also include mannequins and are taken mostly at night. The dataset are labeled according to visibility of humanity.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 27–34

p-ISSN 2655-8564, e-ISSN 2685-9432

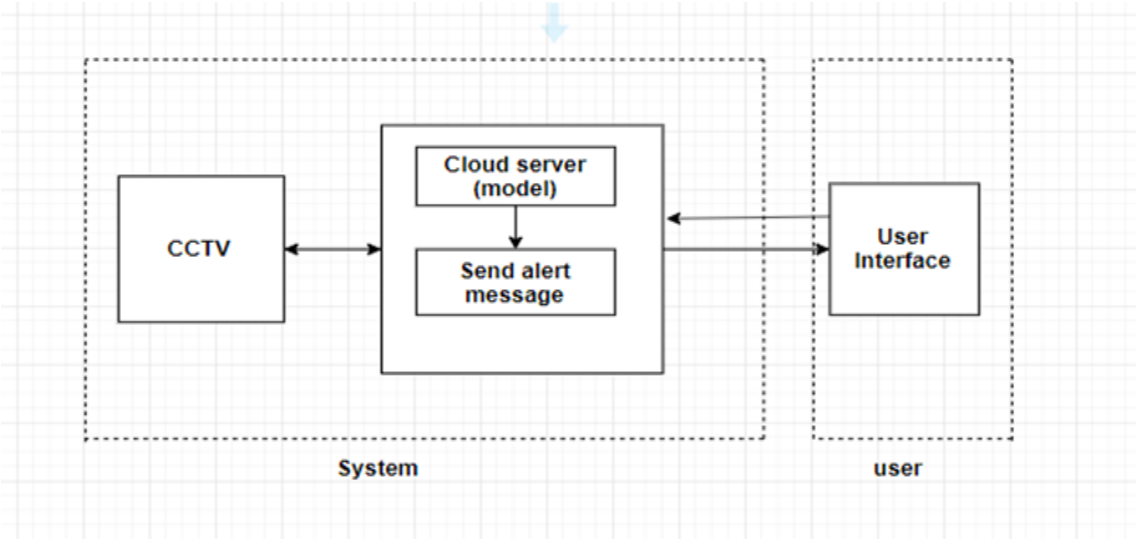


Figure 1. Block diagram

Preprocessing methodology. The main source is a raw video recorded by the CCTV, as in Figure 1. Such videos have very high fps and using such videos requires a lot of processing power by the system. To reduce the processing power, we reduce the fps i.e. dividing current fps by 10. These frames are further sent to be processed [4].

The frames extracted from the videos are RGB images. Processing of images begins with resizing the image into 227×227 and then converting RGB images into Grayscale images. This method converts or compresses the three channels of RGB to a single channel. This single channel contains the values of luminance. Luminance can also be described as brightness or intensity, which can be measured on a scale from black (zero intensity) to white (full intensity). Therefore, the output will have the monochromatic range of black and white.

Most of the theft and break-ins occur at night; hence the images will be dark and will be not clear. To brighten up the image techniques like histogram equalization, alpha and beta transformation can be used. We choose histogram equalization to brighten up the images. Histogram equalization improves the contrast of the image by spreading out the most frequent intensity values.

To remove the noise from the images, blurs are used. This reduces the sharpness of the image and smoothen it. Gaussian blur is the most popular blur and is used for processing. Blur also helps in detection of edges and for thresholding. Thresholding

Volume 3, Issue 1, pages 27–34

p-ISSN 2655-8564, e-ISSN 2685-9432

converts the image to have only two intensities or values. Thresholding using OTSU is used in the project. Finding edges using Canny edge detection is the last pre-processing of the image. Edges help reduce the processing done by the neural model. It reveals the important parts of the image discarding others and helps in extraction of features by CNN.

Neural Network Model. Human detection module will make use of a convolutional neural network to detect and recognize human in the video surveillance. For creating the network, CNN are regularized versions of multilayer perceptron.

Multilayer perceptron's usually means fully connected networks, that is, each neuron in one layer is connected to all neurons in the next layer. CNN uses relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand engineered. The structure of network consists of different components: Input layer, Hidden layer and Output layer. Input Layer is reflecting the potential descriptive factors that may help in prediction. Hidden Layer is defined number of layers with a specified number of neurons in each layer. Output Layer is reflecting the thing is a human present or not.

The CNN architecture used is a modified AlexNet. The input is a series of 3 continuous frames to help whether the entity is a human or a human like mannequin. Due to this, each frame in the input stack is has its own CNN layers. The features extracted or output of the CNN layers are concatenated and given to the fully connected network. The classification of the images is done by using the softmax activation layer. Figure 2 depicts the CNN block for each frame [5], [6]. The fully connected network is illustrated in Figure 3.

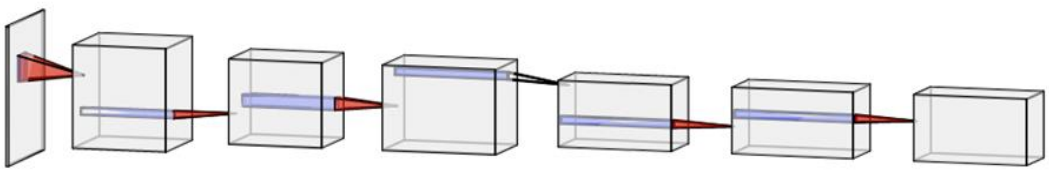


Figure 2. CNN block

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 27–34

p-ISSN 2655-8564, e-ISSN 2685-9432

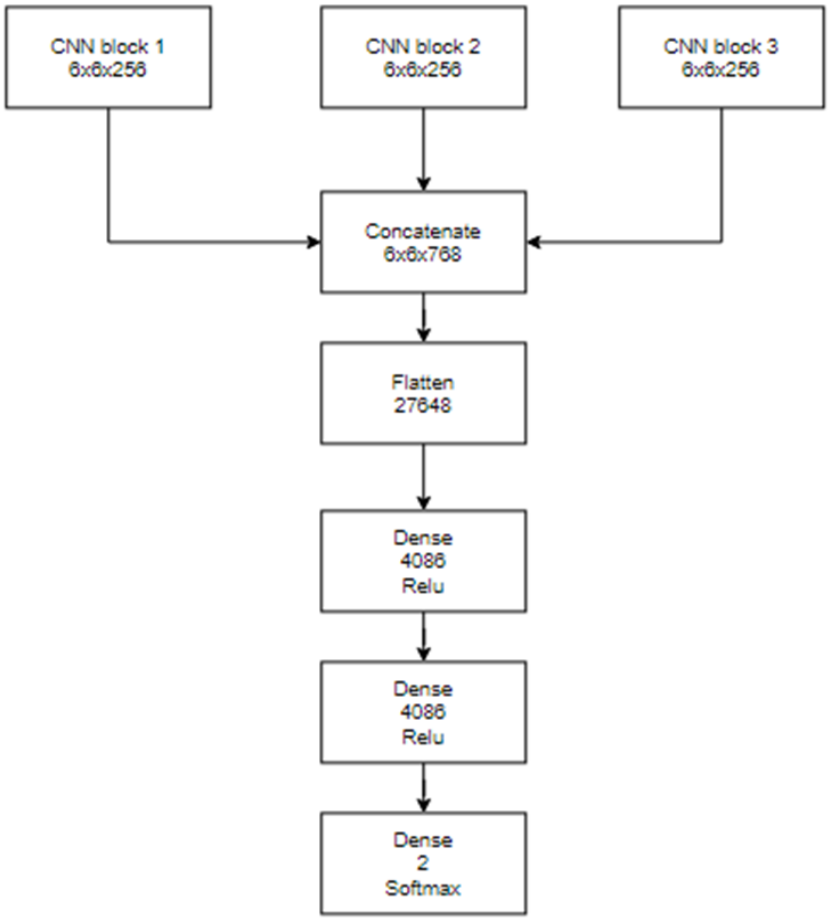


Figure 3. Fully connected network

3 Results and Discussion

The model classifies the data properly at the accuracy rate of 87%. This accuracy is measured by feeding the test data containing both positive and negative labelled images. From the predicted labels, the number of correctly labelled data, positive and negative both, is divided by the total number of data gives the accuracy of the model. The model is trained also in the way that it does not detect mannequins. The model implementation uses android GUI to alert the user of the CCTV and system. This will help the damage done due to the robbery or catch the intruder.

The model is fast and efficient but the delay due to cloud and pre-processing hamper the performance a little bit. This can be neglected by using faster network speed and faster hardware.

Volume 3, Issue 1, pages 27–34

p-ISSN 2655-8564, e-ISSN 2685-9432

Example. Figure 4 shows the correct prediction on the GUI of the system. This depicts the notification and alert used in the system.

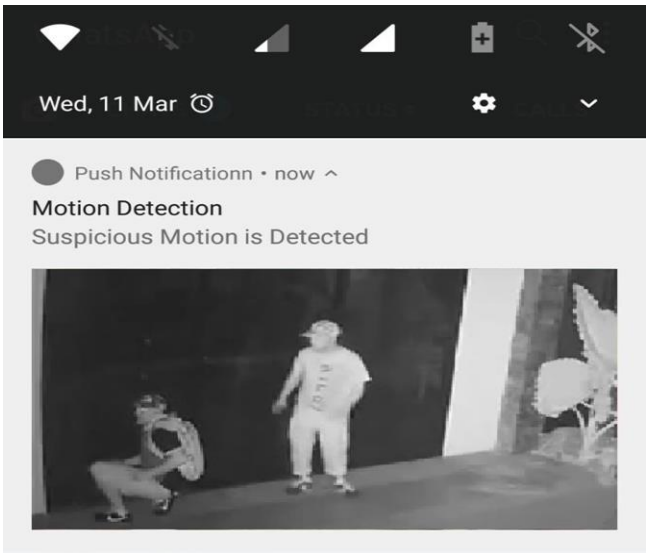


Figure 4. Prediction on Android GUI

Figure 5 shows the incorrect prediction which is also called False Positive. This depicts that the model used is not perfect having accuracy of 87%.

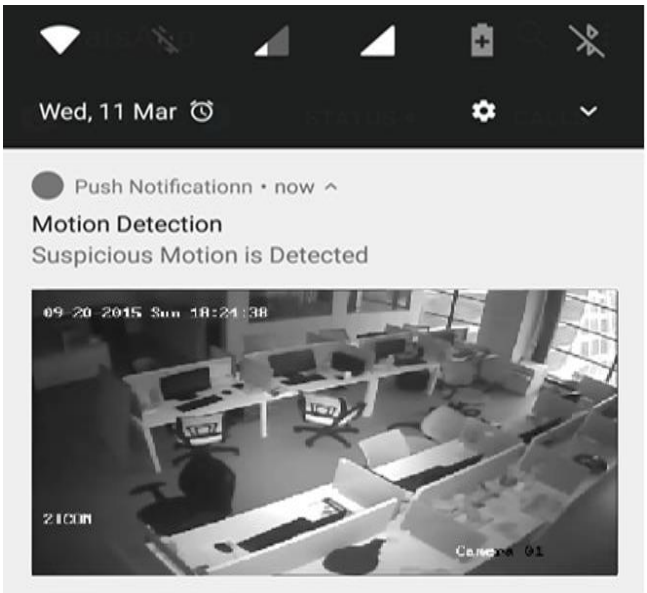


Figure 5. False positive prediction

Volume 3, Issue 1, pages 27–34

p-ISSN 2655-8564, e-ISSN 2685-9432

Discussion. Human Detection in video surveillance using deep learning techniques is the growing area in the field of computer vision. In general, Human Detection is the process of automatically finding the action in the sequence of videos. In this project, we are making use of Convolutional Neural Network. A convolutional neural network (CNN) is an artificial neural network architecture targeted at pattern recognition. In CNN is the methods require labels which are difficult to attain due to the video high dimension information. On this, human activity is detected using this algorithm and after that user can be alert through android application about the subjects.

Environment sensing is the process of detecting a change in the position of an object relative to its surroundings or a change in the surroundings relative to an object. The performance of the system can be enhanced by detecting the changes in its surrounding and it can adapt to the change at the same time. For example, if there is any moment in the shop after closing then the system will alert the user about suspicious activity by send alert message that can user take action on it. Our main goal was to detect human at low visibility due to night time.

The deconstruction of Implementation is as follows:

1. Initially, the input video is taken from video surveillance.
2. This video is processed by the video processing which is used to detect the human activity in the video by frame by frame.
3. The output video is provided to the network to identify the human detection using CNN model.
4. The output of model is sent to user to alert about human activity to take action via application.

4 Conclusion

The accuracy of actually catching a robbery is not calculated in the study but this will reduce the success rate. The purpose of project is to achieve goal to find techniques for behavioural identification. Various techniques for motion recognition based on deep learning. Today, human activity detection in video is the most popular hot space. For security purposes, behaviour recognition can be use in shop or mall. For example, the

use of CCTV, the cases of forced entries and robberies have decreased drastically. But the delay in response to such cases can cause problems. If the owner can get the notification of such events, the culprit can be caught red handed. It becomes important to alert the user by detecting what activity is been performed by the subjects.

Acknowledgements

We (the authors) thank Mrs. Sushama Khanvilkar who gave valuable suggestions and ideas when we were in need of them. She encouraged us to work on this project. We are also grateful to our college for giving us the opportunity to work with them and providing us the necessary resources for the project. Working on this project also helped us to do lots of research and we came to know about so many new things.

References

[1] R. Khurana and A. Kushwaha, "Deep Learning Approaches for Human Activity Recognition in Video Surveillance - A Survey." *First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 542–544, 2018.

[2] A. Khaleghi and M. Moin, "Improved anomaly detection in surveillance videos based on a deep learning method." *8th Conference of AI & Robotics and 10th RoboCup Iranopen International Symposium (IRANOPEN)*, 73–81, 2018.

[3] L. Anishchenko, "Machine learning in video surveillance for fall detection." *Ural Symposium on Biomedical Engineering, Radio electronics and Information Technology (USBEREIT)*, 99–102, 2018.

[4] https://www.pyimagesearch.com/2019/07/15/video-classification-with-keras-and-deep-learning/

[5] https://towardsdatascience.com/introduction-to-video-classification-6c6acbc57356

[6] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5469670/

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

IoT Based Smart Classroom

Prajas Kadepurkar[1], Prim Dsouza[1,*], Nivya Jomichan[1]

[1]*Department of Computer Engineering,*
Xavier Institute of Engineering, Mahim, Mumbai, Maharashtra, India
[*]*Corresponding Author: prim.prd@gmail.com*

(Received 22-07-2020; Revised 09-08-2020; Accepted 23-08-2020)

Abstract

A classroom is a place where there is always room for development; therefore just as the development for a student results in the ease of living, similarly a smart classroom focuses on the structural development leading to effective time and energy utilization. This project offers three major upgradation in classroom; the first being able to book a classroom dynamically using Raspberry Pi and toggling lights and fans using NodeMCU and a mobile application which also helps notify a student about the subject of the lecture, the time and the venue related to the commencement of a period. The second section of this project inputs the attendance of a student check in, using a portable real-time biometric system whose data can further be used to calculate the attendance statistics of each student which can be viewed by the respective student or the teachers can keep a track on their own assigned class using the mobile application. And the last part of the project focuses on keeping track of all the lights and fans, which are on after the lecture. After the lecture is done, it will check the status of the room, whether some other teacher for some other subject is using the lecture hall. Also when the teacher ends the class, the lights and fans will be switched off after 5 minutes buffer provided for the students to

Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

checkout. It will wait for 15 minutes and if it still does not receive any request then it will switch off the lights and fans of that particular lecture hall.

Keywords: Raspberry Pi, NodeMCU, portable, real-time biometric system

1 Introduction

In most of the Colleges and Universities, we have witnessed scenarios where the lights and fans of a classroom are turned on even if there is no person in the class or a similar scenario where a small group of people are sitting in one corner of a classroom where the lights and the fans in the whole classroom are switched on. These scenarios account to a great deal of electricity wastage. As mentioned in [1], without adequate electricity, it becomes challenging for a person towards concentrating on their professional work or study and hence, current scenario insist towards highly efficient and effective usage of any form of power in educational institutes. Suppose the working hour of an institution is 7 hours, let us assume that the lights and fans are switched on for most of it. This implies that a total of 4830W is being consumed in 7 hours. This means that the institution consumes 1110.9 kW per month whereas, with a smarter system it can be reduced to 634.8 kW per month (Detailed calculation shown further under Results). The traditional attendance system in an educational institute makes the teacher shouts out the roll numbers in the class and mark the attendance of a student on a paper upon acknowledgement of the student. This system proves to be inefficient in a lot of ways. Firstly, calling out the roll numbers in front of a class is labor expensive. Secondly, marking the attendance on a paper wastes a lot of paper and losing this piece of paper essentially means there is no record of the student's attendance. Consider 10 subjects per year; 4 attendance sheet per class per academic year, i.e. 40 sheets per year per class. Let's say there are 12 classes in a given institution, this implies that with a smarter system we can save 480 sheets per year. Lastly, a student can fraudulently acknowledge for a student not present in the class. In essence, the traditional way of marking attendance can prove to be inefficient in terms of labor work, time, paper and security. Considering the above two scenarios, we developed a system where one part of

Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

the system focuses on automated control over the electrical components in a classroom where these components of a particular classroom or lecture hall will be turned on only when a lecture is taking place in the aforementioned classroom or lecture hall which helps in saving power. This solves one of the two previously mentioned problems. The second part of our system helps reduce the difficulties encountered in the traditional attendance system. In this portable attendance system, attendance of a student is marked using a biometric sensor that prevents fraudulent entries and saves labor work. Further, the records are stored into a database, which reduces the risk of losing the records. Also, the system being portable can be circulated among the students as the lecture is taking place that helps in saving time required after the teaching session as in traditional attendance system.

2 Literature Survey

In this section we elaborate the automatic lighting and control system for classroom, a classroom scheduling service for smart classes, design and development of portable classroom attendance system based on arduino and fingerprint biometric, web-based student attendance system using RFID technology.

2.1 Automatic Lighting and Control System for Classroom.

Current scenario insists towards highly efficient and effective usage of any form of power in educational institutions like Colleges and universities where we use power for our teaching in classroom or labs. It is common practice that most of us leave the classrooms or labs with Air conditioner, Fan and lighting on even if no students or Faculty members present. These amount to unnecessary wastage of power, contributing to country's energy resource. So accordingly, an automatic lighting and control using Arduino for the efficient use of energy in Classroom condition where we have divided the classroom intro grids has been developed. The system developed will control lighting in a particular area of classroom based on the presence of human using relay control compared to the one placed in ceiling which would switch on or off based on presence of human in room irrespective of position. In addition to relay control, we have

Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

also provided mobility and remote command execution to system using Android mobile App via Bluetooth to control lighting based on voice command [1].

2.2 A Classroom Scheduling Service for Smart Classes.

A typical case study demonstrates that SmartClass provides a new efficient paradigm to the traditional classroom scheduling problem, which could achieve high flexibility by software services reuse and ease the burden of educational programmers. Evaluation results on efficiency, overheads and scheduling performance demonstrate the SmartClass has lower scheduling over heads with higher efficiency [2].

2.3 Design and Development of Portable Classroom Attendance System based on Arduino and Fingerprint Biometric.

The objective of this paper is to design and develop a portable student attendance system used in educational institutions as well as to design a user friendly attendance mechanism especially for the lecturer which incorporates security criteria for the stored data. The design and development of a portable classroom attendance system based on fingerprint biometric is presented. This paper introduces a portable fingerprint based biometric attendance system which addresses the weaknesses of the existing paper based attendance method or long time queuing. The system helped to reduce many issues such as, denying the possibilities of cheating in recording the attendance, helps to ease the lecturers to keep track of students attendance, the encryption technique adds more security so there will be no anonymous fingerprint which is able to tamper with the recorded data, and the portability saves time in taking attendance instead of queuing in a line [3].

2.4 Web-based Student Attendance System using RFID Technology

The existing conventional attendance system requires students to manually sign the attendance sheet every time they attend a class. As common as it seems, such system lacks of automation, where a number of problems may arise. This include the time unnecessarily consumed by the students to find and sign their name on the attendance sheet, some students may mistakenly or purposely signed another student’s name and

Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

the attendance sheet may got lost. Having a system that can automatically capture student's attendance by flashing their student card at the RFID reader can really save all the mentioned troubles. The main idea behind the system is to capture student attendance in a semi-automated way where the students are required to flash their student card at the RFID reader upon entering the classroom. This way, the student ID is instantly captured by the reader, after which the data is sent to the online server for recording purpose [4].

The aforementioned papers have been taken as a reference and a guide for our project as a collective idea consisting each feature presented by the individual papers. As presented in the paper on Automatic light and control system for classroom [1], our IoT Based smart classroom has a system which would switch on or off the lights and fans of a classroom based on the lecture status as and when initiated by a lecturer. Using the mobile application, the lecturer decides the class and lecture to be conducted and the lights and fans will be switched on if the class has started and switches off automatically when the lecturer ends the class. All the students are also notified regarding the lecture status, using the web application. This leads us to the need of smart class scheduling system to avoid collision and chaos. As highlighted in the paper on A Classroom Scheduling Service for Smart Classes [2], the issues concerning classroom scheduling are vital and our system ensures minimal collisions. The lecturer can only select from a set of available lecture halls for them to initiate a notification regarding their lecture. Such measures lead to a more dynamic and more convenient scheduling methodology. The IoT based smart classroom provides the main functionality of integrating intelligent toggling of the physical electrical aspects of a class along with a portable biometric attendance system on a dynamically scheduled lecture hall. As mentioned in the paper on Web-based Student Attendance System using RFID Technology [4], a more secure, efficient way than a pen and paper system for attendance is using technology to do the required work than humans, but as present in this paper, the use of RFID has many drawbacks, some of them including false entry or non-portability in terms of scanning for attendance, which is why our project uses a fingerprint sensor as used in the paper on Design and Development of Portable Classroom Attendance System based on

Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

Arduino and Fingerprint Biometric [3], but rather making it portable using a raspberry Pi, such that it can be circulated in the lecture hall. In our research, enhancements have been made in terms of security, cost, and performance, therefore making a classroom smart as the project suggests. In this project we are making use of NodeMCU Microcontroller board, which consist of ESP8266 Wi-Fi enabled chip and Raspberry Pi single board computer.

3 System Design

The IoT based smart classroom provides the main functionality of integrating intelligent toggling of the physical electrical aspects of a class along with a portable biometric attendance system on a dynamically scheduled lecture hall. In this project we are making use of NodeMCU Microcontroller board, which consist of ESP8266 Wi-Fi enabled chip and Raspberry Pi which is a single board computer.

3.1 Hardware Processing Unit

The Hardware Processing Unit is further divided into two major subunits. First subunit focuses on controlling the electrical components of the classroom and the second subunit is used for marking student attendance. The attendance handling subunit uses NodeMCU Development Board / kit v1.0 (Version 2) microcontroller whereas the electricity handling subunit use Raspberry Pi Single Board Computer. The NodeMCU microcontrollers act as a client to the local server which is described further. The Raspberry Pi acts as an admin to the Firebase Server and can directly communicate with the database.

International Journal of Applied Sciences and Smart Technologies
Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

3.1.1 Electricity Handling Unit

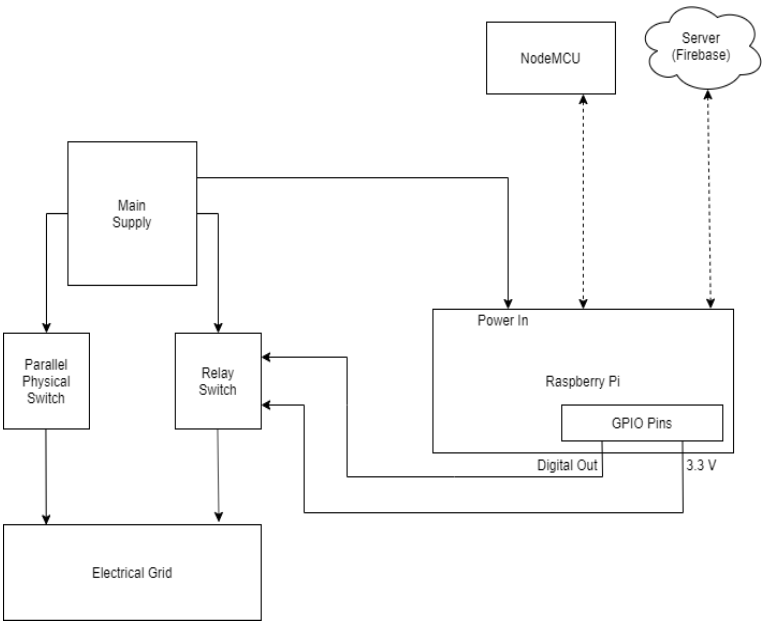


Figure 1. Block diagram of electricity handling subunit

Figure 1 shows the block diagram for the working of Electricity Handling unit. Each lecture hall will contain its own electrical processing unit with a unique ID assigned to it. A parallel physical switch is provided to the electricity grid to manually control the switches as well. When the teacher starts or ends a lecture, the Pi toggles the relay switch and controls the electricity grid. The Raspberry Pi is provided with an external power supply. Along with the handling of electricity of the particular lecture hall, the Raspberry Pi also updates its local file and the database according to the attendance marked. The server in turn, also keeps validating the data on the database with the local file (.json extension) and updates it's contents.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

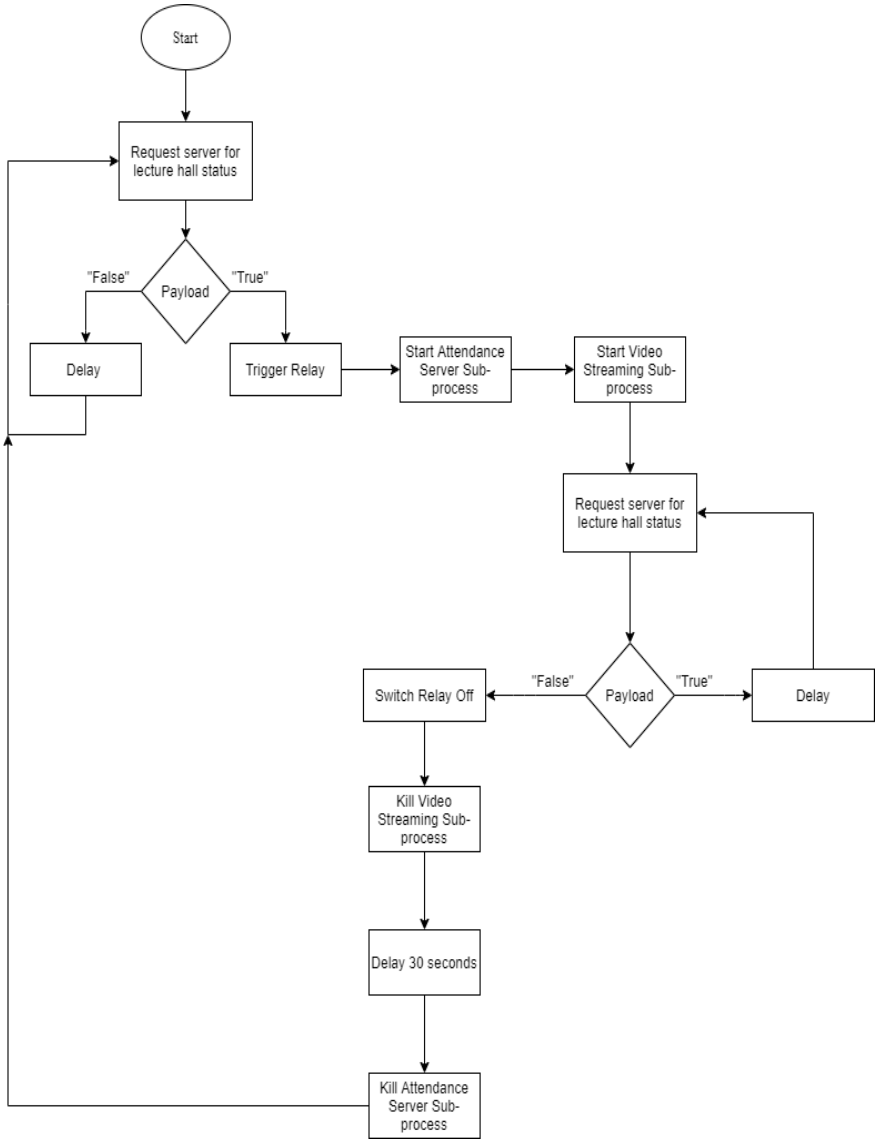


Figure 2. Flowchart of electricity handling subunit

A relay is an electrically controlled electromagnetic switch. It can be controlled with low level voltages like 5V from Raspberry Pi GPIOs or 3.3V from NodeMCU pins. Their main use is controlling circuits by a low power signal or when several circuits must be controlled by one signal. The input to a relay consists of VCC, usually 5V or 3.3V in some cases, GND normally connected to negative supply, and INn pins where n is the number of channels the relay consists. The output consists of three pins which provide 2 different configurations to control circuits. One of the output pins is COM which is the common pin. The other two output pins are NO which stands for Normally

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

Open and keeps the circuit open before the relay is triggered and NC which stands for Normally Closed which keeps the circuit closed before the relay is triggered. The flowchart of electricity handling subunit is shown in Figure 2, and a two channel relay module is illustrated in Figure 3.

Figure 3. Two channel relay module [5]

3.1.2 Attendance Handling Unit

Here we describe our Attendance Handling Unit. This includes Block diagram of attendance processing subunit, flowchart of attendance processing subunit, etc.

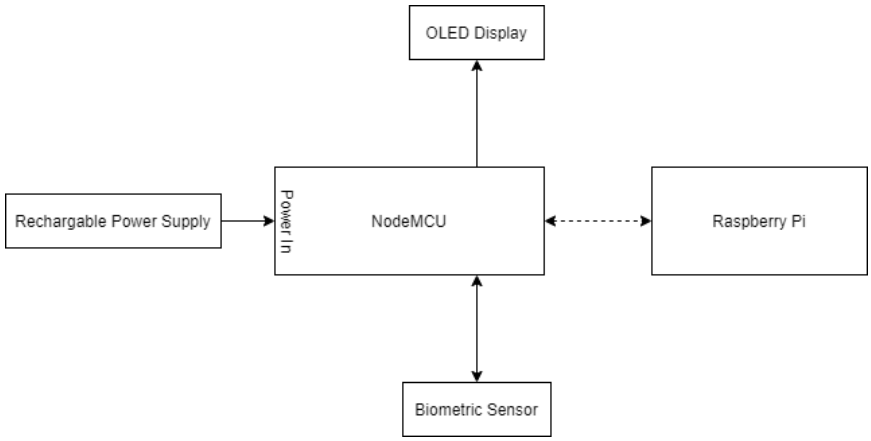


Figure 4. Block diagram of attendance processing subunit

Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

Figure 4 shows the block diagram for Attendance Processing system which uses a biometric sensor to verify each registered student's fingerprint with its respective ID and sends it to the Raspberry Pi to update in the local file and the database. The OLED screen displays the information of the student once his/her attendance is marked. NodeMCU is provided with a rechargeable power supply.

When the device is circulated amongst the students, each student scans his / her fingerprint on the system. This fingerprint is then matched with database stored on the sensor. In reference to the matched ID, a local JSON file is updated to mark the respective student as present and the matched ID is sent wirelessly to the server. The name of the student, depending upon the ID matched, is then displayed on the screen which acts as a validation for the student that his / her attendance has been recorded. The screen displays the message 'Try Again' if something goes wrong. Figure 5 shows a flowchart of attendance processing subunit. Figures 6 and 7 depicts R307 optical fingerprint scanner [6] and R307 optical fingerprint scanner [7], respectively.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

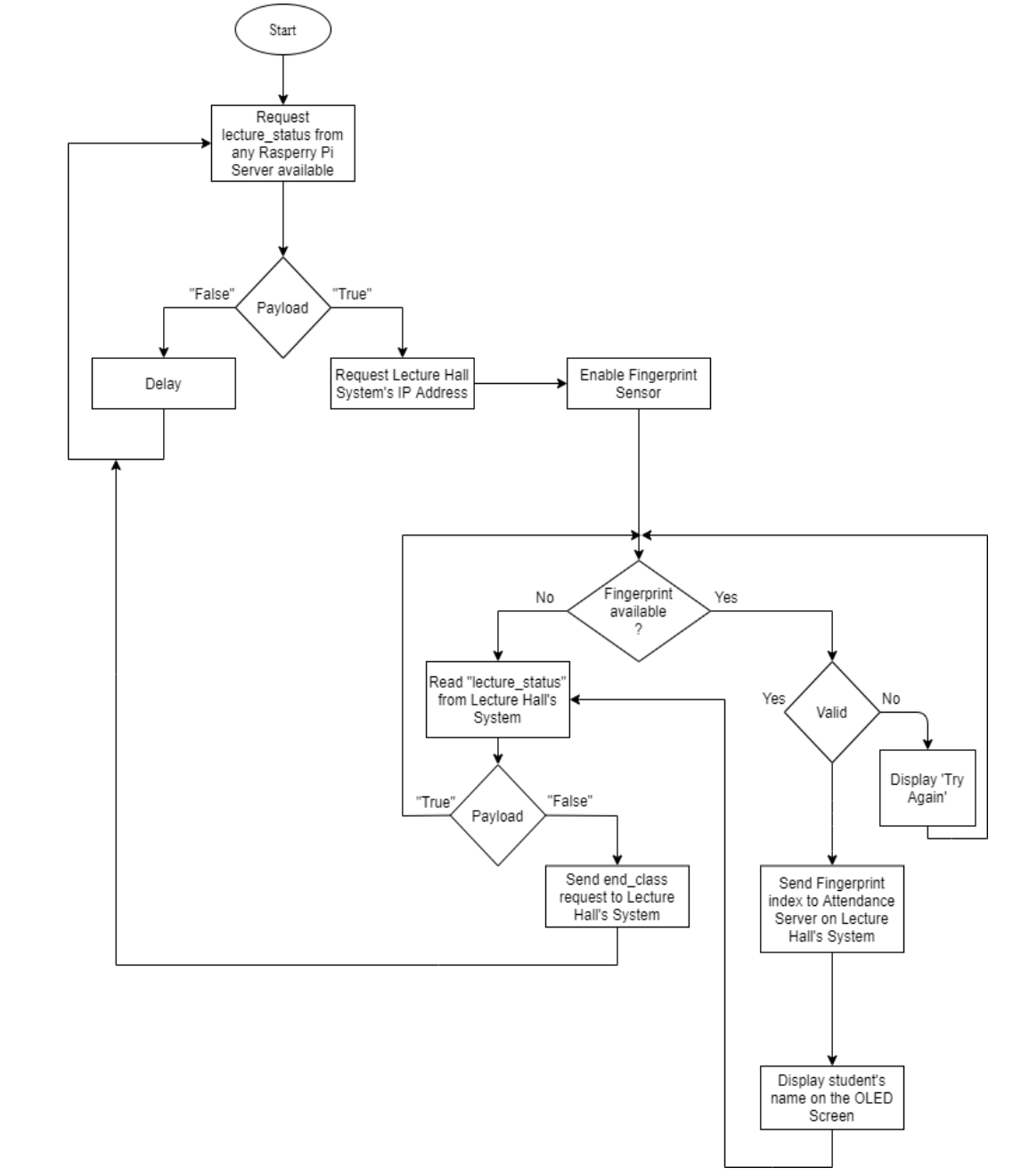


Figure 5. Flowchart of attendance processing subunit

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

Figure 6. R307 optical fingerprint scanner [6]

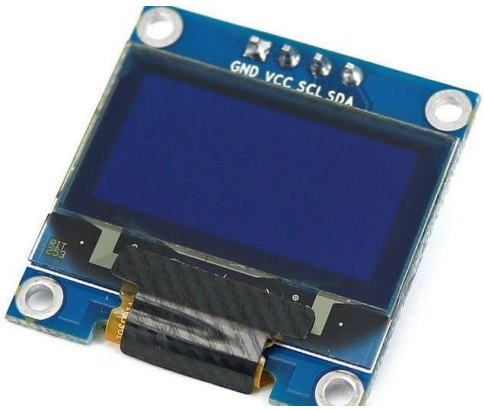


Figure 7. R307 optical fingerprint scanner [7]

Figure 8 shows the flowchart of the Attendance Server which runs on Raspberry Pi which keeps on checking for incoming requests and updates the database accordingly.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

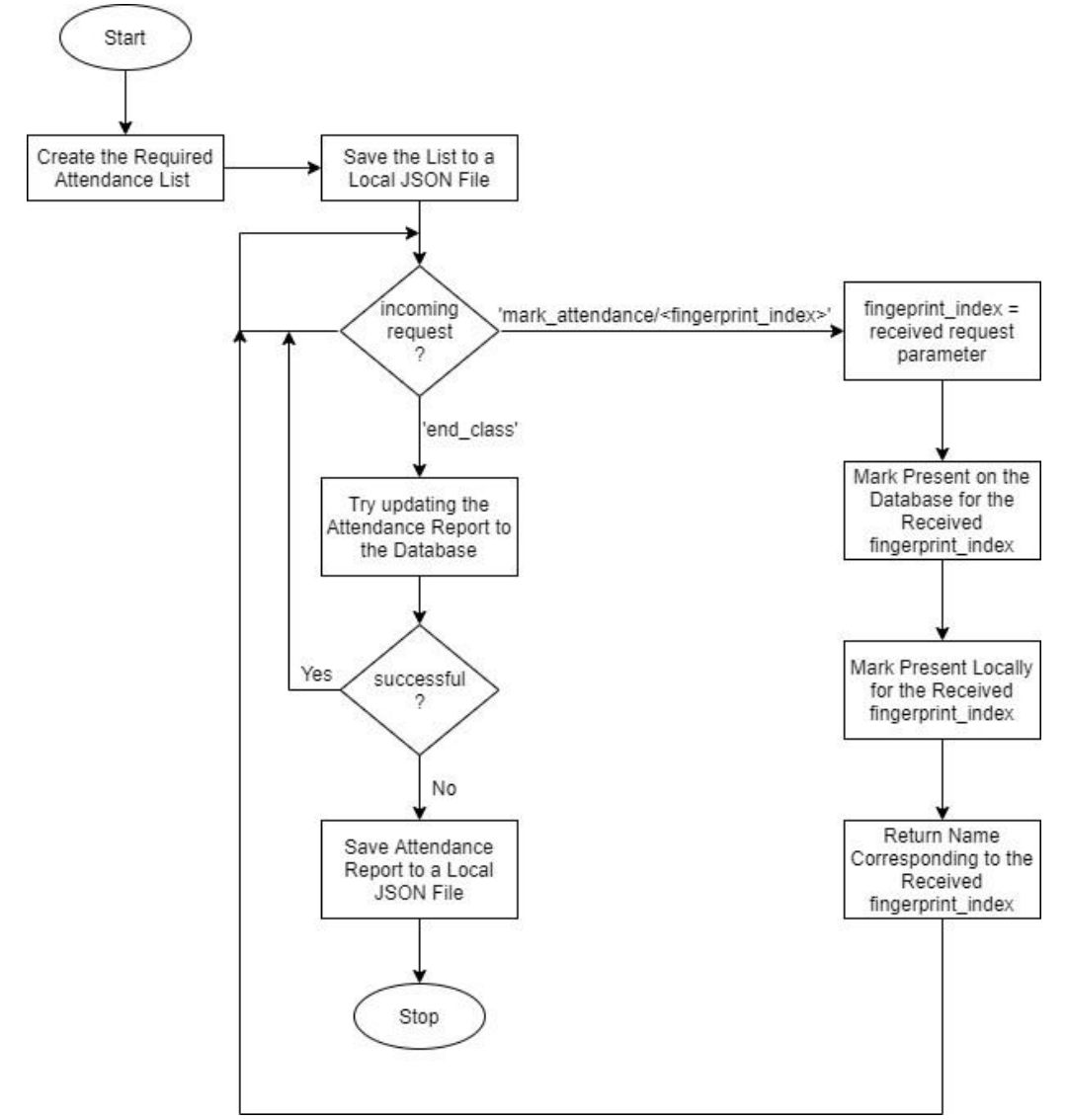


Figure 8. Flowchart of attendance server on Raspberry Pi

3.2 Software Processing Unit

The software processing unit consists of a cross platform native mobile application. The application is developed using Flutter SDK for Dart programming language developed by Google. Flutter enables the development of cross platform mobile applications i.e. Android and iOS from a single codebase. It means that we only have to write a single piece of that, which further will be converted to native machine languages by the Flutter SDK, to work on both the platforms.

International Journal of Applied Sciences and Smart Technologies
Volume 3, Issue 1, pages 35–54
p-ISSN 2655-8564, e-ISSN 2685-9432

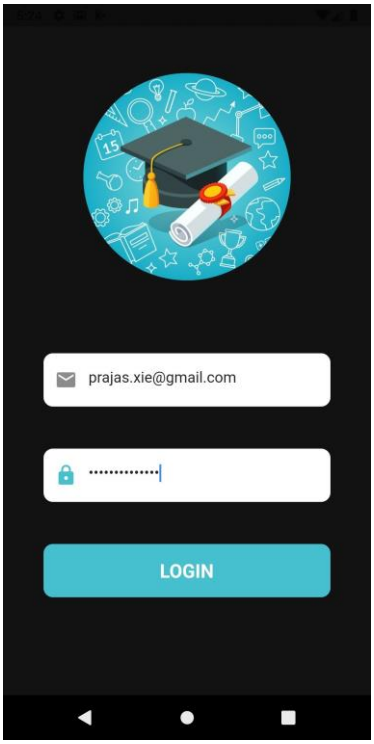


Figure 9. Mobile application login screen

The above mentioned application is a role based application which consists of the following two roles:

3.2.1 Professor

A professor after logging in, can select the class and a lecture hall, given the status of both is free, to start the lecture. The above mentioned professor can also view the attendance statistics of a particular class in the application. The professor can view the list of students who are defaulters subject wise and the Class teacher can see the defaulter list of the whole class with all the subjects.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

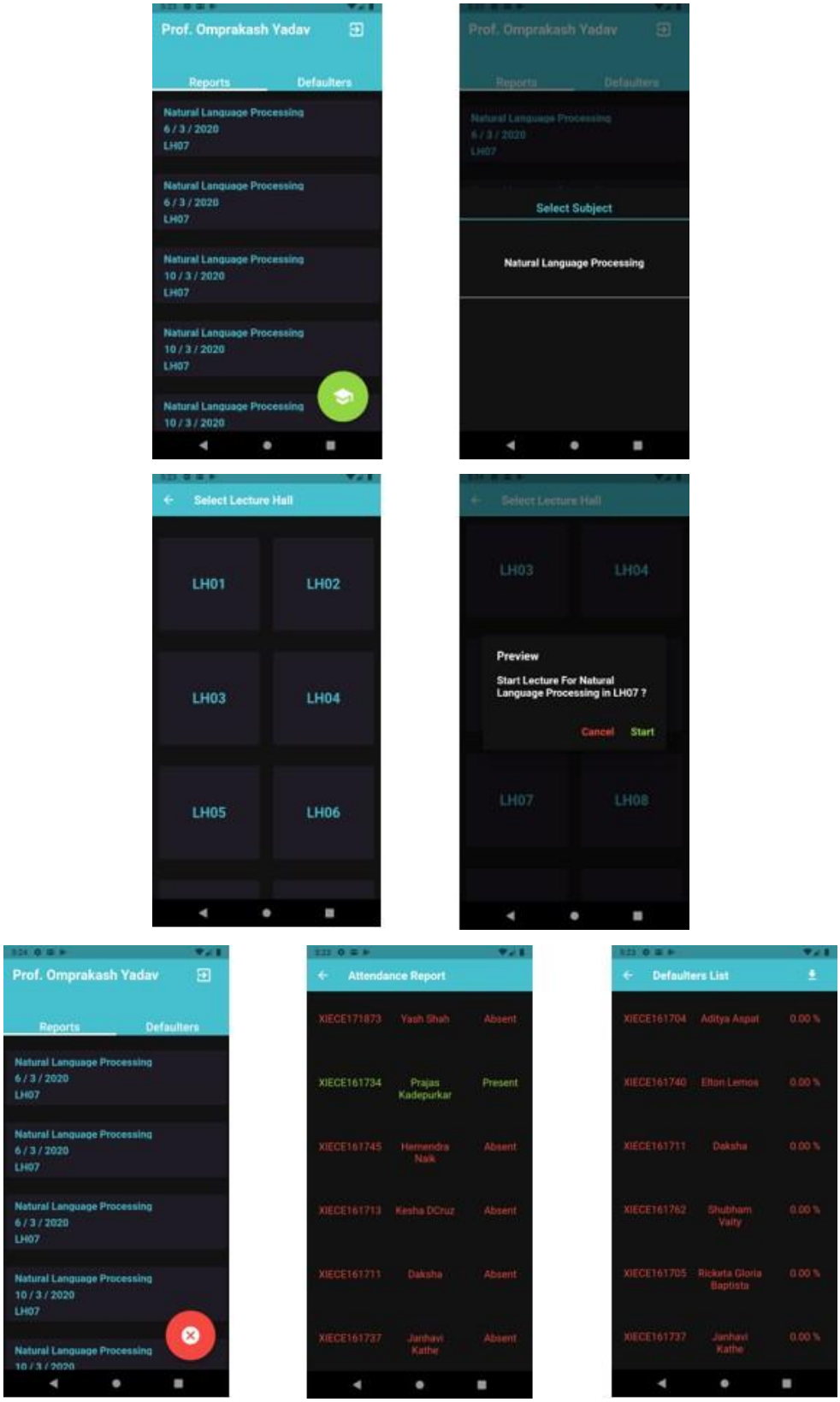


Figure 10. Professor's mobile application interface

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

3.2.2 Student

As a student, the user can login to the application by entering his / her credentials. When the professor starts the lecture, each registered student is notified about the details of the lecture to be conducted. After logging in, the user can view his / her attendance statistics for each class he has registered for.

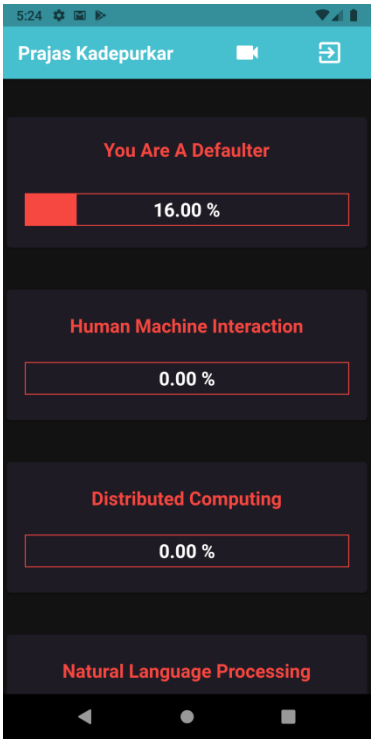


Figure 11. Student's mobile application interface

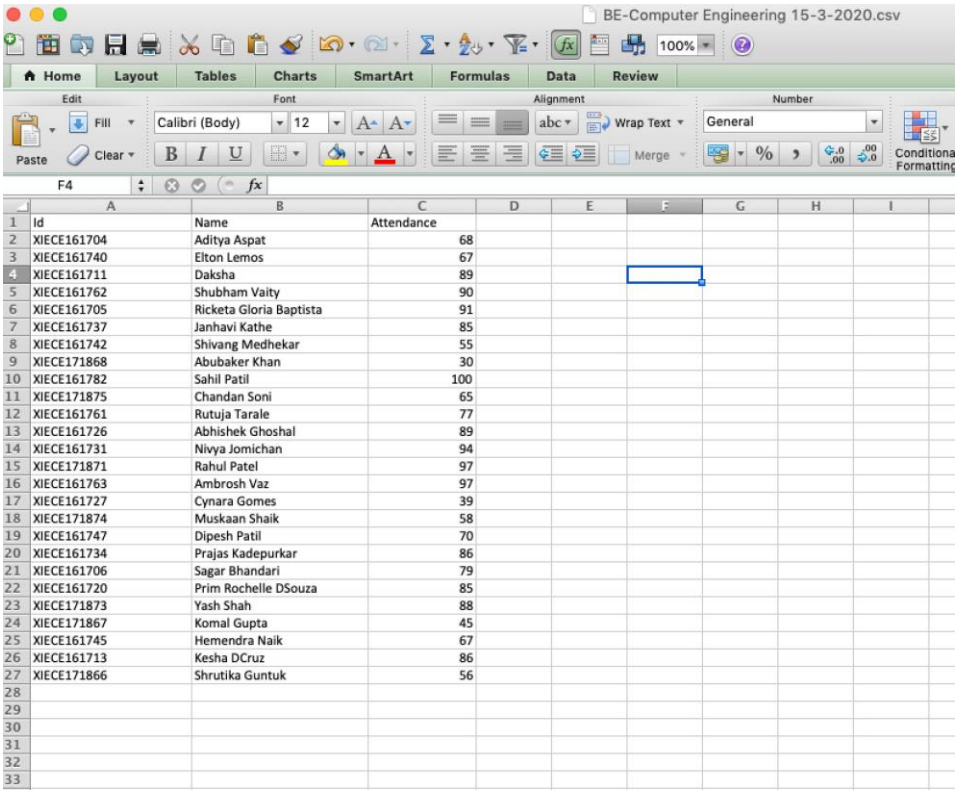
4 Results

The user interface is shown in Figure 9, 10 and 11. The registered professors and students can successfully log in to their accounts. Professors are able to select the classroom, the subject and the batch of student registered for the given course, he or she may start or end the lecture and view the overall attendance of the students and the defaulter list for the class they teach. Class teachers can also view the attendance of the entire class for all the courses the students have enrolled for. Professors can also download the attendance statistical list as a csv file as shown in Figure 12.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432



	A	B	C
1	Id	Name	Attendance
2	XIECE161704	Aditya Aspat	68
3	XIECE161740	Elton Lemos	67
4	XIECE161711	Daksha	89
5	XIECE161762	Shubham Vaity	90
6	XIECE161705	Ricketa Gloria Baptista	91
7	XIECE161737	Janhavi Kathe	85
8	XIECE161742	Shivang Medhekar	55
9	XIECE171868	Abubaker Khan	30
10	XIECE161782	Sahil Patil	100
11	XIECE171875	Chandan Soni	65
12	XIECE161761	Rutuja Tarale	77
13	XIECE161726	Abhishek Ghoshal	89
14	XIECE161731	Nivya Jomichan	94
15	XIECE171871	Rahul Patel	97
16	XIECE161763	Ambrosh Vaz	97
17	XIECE161727	Cynara Gomes	39
18	XIECE171874	Muskaan Shaik	58
19	XIECE161747	Dipesh Patil	70
20	XIECE161734	Prajas Kadepurkar	86
21	XIECE161706	Sagar Bhandari	79
22	XIECE161720	Prim Rochelle DSouza	85
23	XIECE171873	Yash Shah	88
24	XIECE171867	Komal Gupta	45
25	XIECE161745	Hemendra Naik	67
26	XIECE161713	Kesha DCruz	86
27	XIECE171866	Shrutika Guntuk	56

Figure 12. CSV attendance file Sample

This results in the cut down of paper usage for keeping the record of attendance for each subject of each academic year as shown below

- Let’s consider that there are 10 subjects per year.
- Each class will require 4 attendance sheets per Year.
- Therefore, 10 subjects × 4 sheets = 40 sheets per year per class.
- For each batch from First year to the Fourth year let’s say there are 12 batches in total, this implies that 40 sheets per year per class × 12 such batches = 480 sheets per year.

Hence, with our current system, there is no wastage of sheets for marking attendance. As for the electricity module, consider the current scenario, without the IoT Based Smart classroom system, let’s say that the average college working hour is seven hours and since there is no system to toggle the electricity according to the occurrence of a lecture or any human presence, the following is the estimated electricity consumption:

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

- 4830W in 7 hours (average time considering all the events when there is no one in the class as well).
- So for 10 such classes in a college 4830W × 10 = 48300W in one day.
- For one month, the electricity consumption will be 1110900W i.e. 1110.9 kW for 10 classes in 1 month.
- The cost of electricity for this month will be Rs. 5165 (per month).

Now, consider the following case where our system is being used in the institution, the cost of electricity for a month will be calculated as follows:

- 2760W in 4 hours; since the system automatically switches of the fans and lights of a lecture hall, unnecessary wastage of energy is minimized. Therefore the average time considered for a lecture hall in a day is 4 hours.
- So for 10 such classes in a college 2760W × 10 = 27600W in one day.
- For one month, the electricity consumption will be 634800W i.e. 634.8 kW for 10 classes in 1 month.
- The cost of electricity for this month will be Rs. 2915 (per month).

Clearly, the energy consumption and the cost reduces drastically for just one month (see Figure 13), therefore our system helps conserve a lot of electricity in the long run with a decrease in the expense.

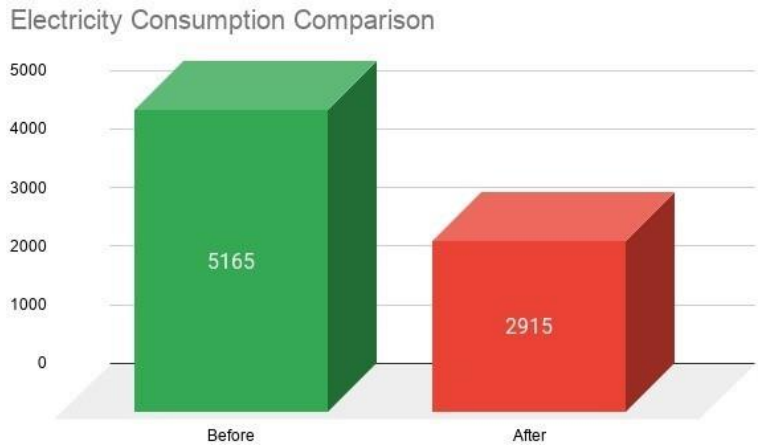


Figure 13. Electricity consumption comparison

5 Conclusion

Thus, we have proposed Smart Classroom System, which introduces automated control over electrical components in a classroom for energy saving, and a Portable

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

Electronic Attendance System based on fingerprint identification for efficient use of time and paper, reduce labor work, and even prevent fraudulent attempts to mark the attendance. The initial results are promising but the work is still ongoing as there is a lot of scope for further development. Use of facial recognition to mark the attendance of the students, dividing each lecture hall into sections called grid with motor sensor to detect moving objects in order to toggle the sectional electricity module according to the presence of a person, developing student teacher portal on the app for further usability etc are some of the future scope of this project we are working on.

References

[1] S. Suresh, H.N.S. Anusha, T. Rajath, P. Soundarya and S.V. P. Vudatha, "Automatic lighting and Control System for Classroom." *International Conference on ICT in Business Industry Government (ICTBIG),* 1–6, 2016.

[2] C. Wang, X. Li, A. Wang and X. Zhou, "A Classroom Scheduling Service for Smart Classes." *IEEE Transactions on Services Computing **X** (X),* 1–11, 2015.

[3] N. I. Zainal, K. A. Sidek, T. S. Gunawan, H. Manser and M. Kartiwi, "Design and development of portable classroom attendance system based on Arduino and fingerprint biometric." *The 5th International Conference on Information and Communication Technology for The Muslim World (*ICT4M), 2015.

[4] M. Kassim, H. Mazlan, N. Zaini and M. K. Salleh, "Web based student attendance system using RFID technology." *IEEE Control and System Graduate Research Colloquium (ICSGRC 2012),* 213–218, 2012.

[5] 2 Channel 5V Relay Module-Wiki, http://wiki.sunfounder.cc/index.php?title=2_Channel_5V_Relay_Module

[6] Interfacing Fingerprint Sensor (R307) with evive – Fingerprint Matching, https://thestempedia.com/tutorials/interfacing-fingerprint-sensor-r307-evive-fingerprint-matching/

[7] How to Use an OLED Screen (128 per 64)-Hackster.io, https://www.hackster.io/MisterBotBreak/how-to-use-an-oled-screen-128-per-64-cb6e4d

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 35–54

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 55–64

p-ISSN 2655-8564, e-ISSN 2685-9432

Stone, Paper, Scissors Mini-Game for AI Pet Robot

Aditya Aspat[1,*], Elton Lemos[1], Abhishek Ghoshal[1]

[1]*Department of Computer Engineering,*
Xavier Institute of Engineering, Mahim, Mumbai, Maharashtra, India
[*]*Corresponding Author: aditya.dir@gmail.com*

(Received 20-07-2020; Revised 19-12-2020; Accepted 19-12-2020)

Abstract

The Artificial Intelligence (AI) Pet Robot is a combination of various fields of computer science. This paper showcases the various functionalities of our AI Pet. Most of the functionalities showcased use the immage processing modules made available through OpenCV. The pet robot has various features such as emotion recognition, follow routine, mini-game etc. This paper discusses the mini-game aspect of the robot. The game has been developed by using VGG16 convolutional network for identification of the action performed by the user. To improve the accuracy we have made use of background subtraction which gives removes all the unwanted objects from the background and gives a simple cutout of the users hand.

Keywords: Pet Robot, VGG16, background subtraction

1 Introduction

Owning a pet comes with a zillion benefits for our physical as well as mental health [1]. Pets give us increased opportunities to go outside, exercise and socialize with people. In today's time such as the pandemic, having a pet by your side gives us

Volume 3, Issue 1, pages 55–64
p-ISSN 2655-8564, e-ISSN 2685-9432

company when we are required to social distance from everyone else. However, with all the benefits of owning a pet, there are some drawbacks as well [2]. Pets can be a source for transmission of various diseases if they are not properly taken care of. Many societies have rules which don't allow pets inside the buildings. They may start barking in the middle of the night which might cause nuisance in some cases to the neighbors. For the reasons given above, we propose the development of an inexpensive pet using modern technologies like artificial intelligence (emotion recognition and face recognition), internet of things (for communicating with the bot and it's movement), image processing (building a mini game) which will seamlessly interact with humans and comes with all the benefits while avoiding the drawbacks of owning a pet.

2 Existing Systems

There are a few pet robots that are already available in the market. Learning about these systems will help better understand how our system compares to them. This section also looks into the different algorithms that have previously been implemented regarding our robot and its functionalities.

There have been multiple attempts at making realistic AI pets with some dating back to 1990s. Tamagotchis were handheld digital pets created in Japan in 1997 [3]. This representation of pets of pets quickly became popular. The colourfully designed creatures would grow differently based on the level of care provided by the user. In more recent times companies like Sony have come up with realistic pets such as the Aibo [4]. It was a slew of features such as facial recognition, emotion detection, automatic battery charging and many more. Pibo is another such AI pet created by Circulus [5]. Pibo's features include weather reports, alarms, notification, taking photos and other interactive functions. There is however one common problem with these products. Their cost is upwards of $1000. One of our goals is to create a fully functioning AI pet by using equipment that is cheap and commercially available.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 55–64

p-ISSN 2655-8564, e-ISSN 2685-9432

3 Implementation Methodology

The architecture of our pet closely resembles a living pet. Our pet has 4 major parts:

- The Brain: The computer acts as the brain and helps with all the computation.
- The Spine: All the messages flow through it, the Pi4B acts as a kernel.
- The Eyes: The NOIR camera are the eyes of our bot. They help in capturing images.
- The Limbs: The Arduino UNO and its motors acts as the limbs and perform various movements as per the commands given to it.

Hardware Design

- Vision Unit:

  This unit is responsible for what the bot sees. It helps in capturing the images and sending them to the computer for Image Processing. It comprises of the Raspberry 4 and its camera module. We have also used a NOIR camera which helps the bot in capturing images in low light. The camera's video stream is hosted using the RPi-Cam-Web-Interface. The computer captures this stream and uses it for various processing features

- Motor Unit:

  This unit is responsible for the movement of our bot. It comprises of various actuators like DC Motor and Servos which are controlled by the Arduino UNO. The power for the motor unit is provided by a 12V battery. The Arduino oversees the motor operations but the final decision as to what movement should be performed is taken by the Computer.

- Communication Unit:

  This Unit is the tunnel for information transfer from the Computer to the Pi4B and to the Arduino UNO and vice versa. This Unit consists of a Serial Interface between the Pi4B and the Arduino UNO which is used by the Pi4B to instruct the Arduino or forward instructions given by the Computer. There is also a Socket connection between the Pi4B and the Computer for wireless communication using TCP.

Volume 3, Issue 1, pages 55–64
p-ISSN 2655-8564, e-ISSN 2685-9432

- Interfacing with Robot:

  The robot has a joystick for an arm. If the user shakes hands with the robot, the Robot goes into listening mode and hears any commands given by the user. If the robot cannot understand, the robot will inform the user that it couldn't understand and then continue what it was doing previously. The list of commands that the robot can obey are:

- Stop:

  On receiving a STOP command, the robot will stop whatever it is doing and will enter the face detection mode.

- Tell me the time:

  The UTC time is grabbed from the servers and is displayed on the screen as well as spoken out loud by the bot.

- Read "book name":

  The robot makes use of the Google Text-to-Speech module to read out books to the user.

- Follow me:

  On receiving the "Follow Me" command, the bot starts tracking the user's movement and follows him or her while maintaining a safe distance.

- Play a game:

  Saying this will start the game of stone, paper and scissors. The action performed by the bot is displayed on the screen attached to it.

4 Implementation

This section elaborates on the algorithms we have used for implementation of the stone, paper and scissors game. How they work, and how integrating all of them makes the game easy to play for anyone.

4.1 VGG16

We have used the VGG16 convolution network model proposed by K. Simonyan and

Volume 3, Issue 1, pages 55–64

p-ISSN 2655-8564, e-ISSN 2685-9432

A. Zisserman from the University of Oxford in the paper "Very Deep Convolutional Networks for Large-Scale Image Recognition" [6]. The model achieves 92.7% top-5 test accuracy in ImageNet, which is a dataset of over 14 million images belonging to 1000 classes. The architecture [7] of VGG16 is as in Figure 1.

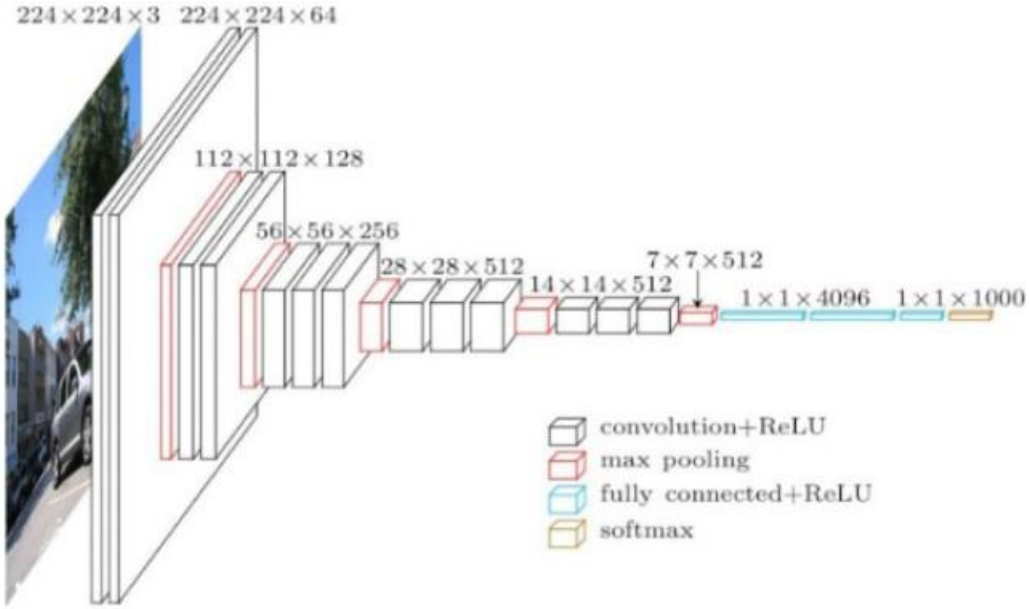


Figure 1. VGG16 architecture

4.2 Training

Deep convolutional neural network models may take days or even weeks to train on very large datasets. A way to short-cut this process is to re-use the model weights from pre-trained models that were developed for standard computer vision benchmark datasets, such as the ImageNet image recognition tasks. Top performing models can be downloaded and used directly, or integrated into a new model for your own computer vision problems. The way to do this is Transfer Learning where we apply pre-trained models on our own dataset to get a state of the art model. We used 550 images with variation for all the three cases and achieved an accuracy of 98% in the validation set. This provides a very good platform to test the model in a real time scenario. As we have already seen in previous implementations, the real time accuracy of models is affected due to the deviations from ideal test conditions. Training images are shown in Figure 2.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 55–64

p-ISSN 2655-8564, e-ISSN 2685-9432

Figure 2. Training images

```
acc: 0.5817 - val_loss: 0.9739 - val_acc: 0.6077

acc: 0.6691 - val_loss: 0.7196 - val_acc: 0.6731

acc: 0.8196 - val_loss: 0.3430 - val_acc: 0.8692

acc: 0.9337 - val_loss: 0.1310 - val_acc: 0.9588

acc: 0.9733 - val_loss: 0.0670 - val_acc: 0.9661

acc: 0.9871 - val_loss: 0.0881 - val_acc: 0.9782

acc: 0.9814 - val_loss: 0.1783 - val_acc: 0.9564
```

Figure 3. Model accuracy

In Figure 3, metrics for some of the epochs are shown. acc is the accuracy of the model while training the model, that's is the total number of accurately classified images divided by the number of images in the training set. Val loss is the value of the cost function for the validation data. Val acc is the accuracy of the validation set, that is the total number of accurately classified images divided by the total number of images in the validation set. Some of the barriers we faced while performing real time testing were related to both hardware and software. These have been mentioned in the real time testing explanation below.

5 Results and Discussion

In this section, we provide our research results about Real Time Results for Mini-Game. As shown in the fowchart (see Figure 4), right after we turn on the camera, we take a snapshot of an empty background. We make use of background subtraction [8] to take a cutout of the hand. We need to make sure no object in the background does not move, especially within the boundaries of the blue square. In the image below, the left side is the normal image and the right side shows a continuous background subtracted

Volume 3, Issue 1, pages 55–64
p-ISSN 2655-8564, e-ISSN 2685-9432

image. Since the person is the only thing in the image that is moving, an outline comes around the person.

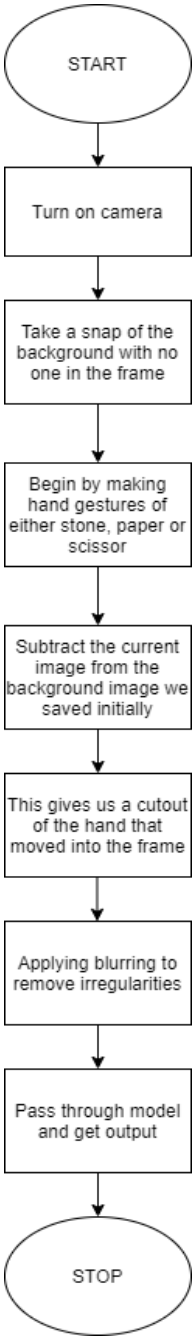


Figure 4. Flowchart for realtime testing

What we do next, is that we use the empty frame we captured to detect any new objects moving into the blue square, i.e. our hand. We cut out this part of the image and

Volume 3, Issue 1, pages 55–64

p-ISSN 2655-8564, e-ISSN 2685-9432

apply blurring so that we reduce the noise that we encounter due to the shortcomings of the camera. As it gets darker, the graininess of the image will increase giving rise to even more noise.

1. The program takes the image of the still background. This will be used later to subtract from the real time image (see Figure 5).

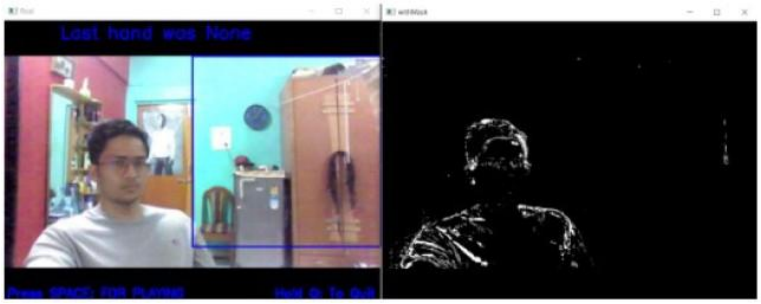


Figure 5. Continuous background subtraction

2. We perform the gesture we wish to do and the program captures this image (see Figure 6).

Figure 6. Gesture made by user

3. We subtract the current image with the saved image. This gives use the required format of the image (see Figure 7).

Figure 7. Background subtraction with hand

Volume 3, Issue 1, pages 55–64

p-ISSN 2655-8564, e-ISSN 2685-9432

4. We cut out the part of the image we need (hand) which is then fed to the model (see Figure 8).

Figure 8. Cut out of hand

The previously encountered software issues involved difficulty in extracting just the hand from the image in such a way that it resembles the images that the model was trained on. This is necessary to reduce any real time errors that we may encounter. All these steps had to be performed in order to get our real time images to be fed to the model, as close to the data set as possible. Only then would we be able successfully use our trained model.

6 Conclusion

The proposed algorithms show the functionality of our AI Pet Robot. We have used VGG16 and background subtraction to implement a simple game of stone, paper and scissors and integrated it in out AI Pet Robot with various other features. The use of background subtraction has helped us solve the issues we faced while taking a cut out of the action performed by the user. This game can be played by children to pass time and have some kind of entertainment without being exposed to the internet/mobile devices at a young age.

Volume 3, Issue 1, pages 55–64

p-ISSN 2655-8564, e-ISSN 2685-9432

References

[1] Healthy Pets, Healthy People. U.S. Department of Health & Human Services. https://www.cdc.gov/healthypets/healthbenefits/index.html (Accessed: 02/03/2020).

[2] E. Paul Cherniack, MD and Ariella R. Cherniack, "Assessing the benefits and risks ofowning a pet." *Canadian Medical Association Journal*, 2015.

[3] The life and death of Tamagotchi and the virtual pet, https://wellcomecollection.org/articles/WsT4Ex8AAHruGfWb (Accessed: 13/08/2020).

[4] S. Aibo: The Dog and Personal Assistant of the Future, https://www.forbes.com/sites/moorinsights/2019/05/01/sony-aibo-the-dog-and-personal-assistant-of-the-future (Accessed: 05/03/2020).

[5] pibo, https://pibo.circul.us (Accessed: 05/03/2020).

[6] Simonyan, K. Zisserman and Andrew. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv 1409*.1556, 2014.

[7] Step by step VGG16 implementation in Keras for beginners, https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c (Accessed: 08/03/2020).

[8] Using Background Subtraction, https://docs.opencv.org/master/d1/dc5/tutorial background subtraction.html (Accessed: 10/03/2020).

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

OpenCV Image Processing for AI Pet Robot

Abhishek Ghoshal[1,*], Aditya Aspat[1], Elton Lemos[1]

[1]*Department of Computer Engineering,*
Xavier Institute of Engineering, Mahim, Mumbai, Maharashtra, India
[*]*Corresponding Author: abhighosh98@gmail.com*

(Received 23-07-2020; Revised 21-12-2020; Accepted 21-12-2020)

Abstract

The Artificial Intelligence (AI) Pet Robot is a culmination of multiple fields of computer science. This paper showcases the capabilities of our robot. Most of the functionalities stem from image processing made available through OpenCV. The functions of the robot discussed in this paper are face tracking, emotion recognition and a colour-based follow routine. Face tracking allows the robot to keep the face of the user constantly in the frame to allow capturing of facial data. Using this data, emotion recognition achieved an accuracy of 66% on the FER-2013 dataset. The colour-based follow routine enables the robot to follow the user as they walk based on the presence of a specific colour.

Keywords: OpenCV, image processing, AI Pet robot

1 Introduction

There are many health benefits of owing a pet [1]. They can increase opportunities to exercise, get outside and socialize. There are also a slew of health benefits such as decreased blood pressure, decreased cholesterol levels and better mental health. However, there are also a lot of drawbacks of having to own a pet [2]. There can be issues of disease transmission and infections from pets. There also other problems like

Volume 3, Issue 1, pages 65–82
p-ISSN 2655-8564, e-ISSN 2685-9432

restrictions from housing societies, training, cleaning up, extra expenses and spending time with the pet to ensure its well being. There have been numerous cases of cruelties and animals across the world. We propose the development of a cheap "AI Pet" which can help provide the benefits of having a pet. This machine will be able to seamless interact with humans and indulge in activities like real pets can. As this pet will be a machine, a lot of the drawbacks of owing a pet will be nullified such as the transmission of diseases and infections, requirement to spend time, cleaning up.

2 Existing Systems

As part of the research conducted before starting the implementation of the pet robot, we looked through some of the commercially available products. We have also mentioned below the algorithms that we have looked at before creating the different functionalities of the pet robot.

2.1. Commercial Products

Some of the earliest products that resembled an artificial pet can be pinpointed back to the late 1990s. Tamagotchis first went on sale in Japan in 1997. They were small, handheld devices with a screen and buttons. The 'pet' on the screen would evolve in various ways depending on user inputs. Over the years there have been many more artificial pets created but pet robots are much more recent. One of the latest products which has been hugely acclaimed is Sony's Aibo. It boasts of numerous artificial intelligence based features including but not limited to facial recognition, emotion recognition, automatic battery charging and others. However it also comes with a price tag to match it; $1800. There are other pet robots with similar features but all of them cost more than $1000. It is an important consideration for us to be able to limit the development cost of our robot and to use cheap and commercially available parts.

2.2. Face Detection using Haar Cascade

Haar Cascade is a machine learning object detection algorithm proposed by Paul Viola and Michael Jones in their paper "Rapid Object Detection using a Boosted

Volume 3, Issue 1, pages 65–82
p-ISSN 2655-8564, e-ISSN 2685-9432

Cascade of Simple Features" in 2001 [3]. It is a machine learning based approach where a cascade function is trained from a lot of positive and negative images. It is then used to detect objects in other images. OpenCV offers pre-trained Haar cascade algorithms, organized into categories (faces, eyes and so forth), depending on the images they have been trained on.

We will be using the Haar Cascade algorithm for facial recognition from a live feed. For this we will have to incorporate the "haarcascade frontalface default.xml" into our project. From an input video feed of say 640 X 480 pixels, the xml file applies the algorithm on individual frames. It then returns the X and Y coordinates that help in forming a rectangle around the face of the user. These X and Y coordinates can then be used to cut out the face from the image for further processing.

2.3. Facial Expression Recognition

This implementation is using a Convolutional Neural Network to classify the FER-2013 [4] dataset for the task of emotion recognition. There is also a creation of a real time system for validation of the model. FER-2013 is a large dataset with over 30,000 monochrome images of size 48 X 48 pixels. The initial model achieved an accuracy of 56%. After making changes based on the results, we improved the accuracy to 66%. We will try to further improve this result. Two models were compared in this paper with the aim to get good results with simpler architectures. Initial model that was proposed was a standard fully-convolutional neural network. This model achieved an accuracy of 66% with 35,887 images in the dataset. For real-time images, the input images were pre-processed before being passed to the model. This including conversion of images to grayscale and resizing the images to 48 X 48 pixels as this was the format for the FER-2013 dataset.

3 Implementation Methodology

Our Robot has an architecture closely related to that of a living pet. The Computer works as the brain of the pet, the Pi4B acts as the Spine, the camera as the eyes while the Arduino UNO and its motors act as Limbs.

Volume 3, Issue 1, pages 65–82
p-ISSN 2655-8564, e-ISSN 2685-9432

In this subsection, we describe our hardware design, which includes vision unit, motor unit, communication unit, and interfacing with robot.

- **Vision Unit**

The Vision Unit is responsible for providing vision to the computer for performing Image processing. The Vision Unit comprises of the Pi4B and the Raspberry Pi camera module. The Raspberry Pi camera module captures a video stream and that stream is hosted by the Pi4B by using the RPi-Cam-Web-Interface. The Computer picks up the stream from the hosted site and can now use the data for its programs.

- **Motor Unit**

The Motor Unit is responsible for the movement of the robot. It consists of actuators like DC Motor and Servos that are controlled by the Arduino UNO. The Motor Unit is powered by a 12V Battery. The Arduino is responsible for overseeing the motor operations. However the decision as to what should the Arduino UNO do is taken by the Computer.

- **Communication Unit**

While the robot receives most of its inputs through the mic, camera and other sensors, most of the processing happens on a computer. To be able to send data from the robot to the computer a socket connection for communication has been set up using the TCP protocol. This helps to get information to and fro from the computer and the Pi4B and Arduino Uno on the robot.

- **Interfacing with Robot**

There are two ways to interact with the robot. The user can “shake the hand” of the robot by moving a joystick mounted on the robot and proceed by saying something voice commands. Alternatively commands can also be input in the CLI present on the computer.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

4 Implementation

In this section we elaborate on the specific algorithms we have implemented. As our major processing tasks are dominated by image processing, there is extensive use of OpenCV and Deep Learning algorithms in this section.

4.1. Face Recognition

The Robot is trained to remember its creators. We have trained a model to remember our faces using Convolution Neural Network. The Computer takes the data from the camera and compares it with the model. If it does not find a match, it assumes that it does not know that person and hence calls it an unknown face. However if it does find a match, it is able to print out the name associated with the matched face. This Module works in conjunction with the Emotion Recognition Module.

4.2. Emotion Recognition

The Robot is also trained to remember our expressions when we feel a certain emotion like Happiness, Sadness, Fear, Anger and Surprise. For this we have used Convolutional Neural Network. The computer takes the data from the camera and compares it with the model. It will then print out the emotion that matches the most features of the tested data. This module is explained below further in detail.

4.2.1. Dataset

The process began with acquiring the right dataset for our project. This was the FER-2013 Dataset that was also used in [5] and can be acquired from kaggle [4]. This dataset contains 32,557 labeled images of size 48 X 48 pixels. The given labels are Happy, Sad, Angry, Fear, Surprise and Neutral. We performed some basic exploratory data analysis on this dataset and understood the distribution of the data is show in Figure 1.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

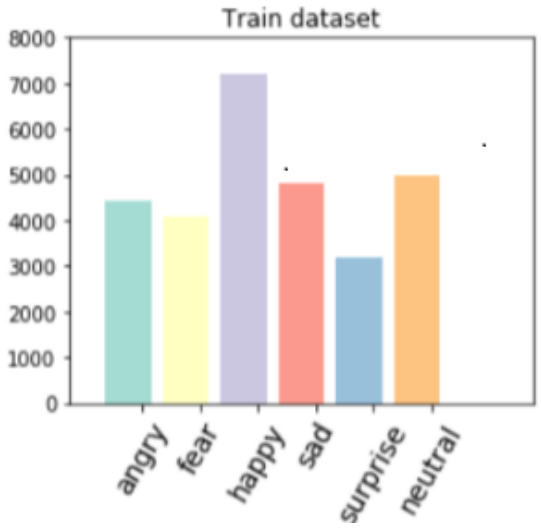


Figure 1. Dataset bar graph

Figure 2. Example of images in dataset

To try to improve the performance of the model, we took a preemptive measure and added 255 images of ourselves spread into the different emotion categories. We took images of our upper bodies and to match our own data to the data in FER-2013, some processing was needed. Figure 2 shows an example of images in dataset. Figure 3 illustrates types of images added. The steps taken to achieve are explained in Figure 4.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

Figure 3. Types of images added

Figure 4. Pre-processing of images before adding to dataset

Before we could proceed to create the model, the target variable that is the “emotion” column was converted in to a encoded column. This is done because Machine Learning models are only able to perceive numerical values. Words like “happy” or “sad” hold no

Volume 3, Issue 1, pages 65–82
p-ISSN 2655-8564, e-ISSN 2685-9432

meaning for it. Hence, we convert such columns using encoding that replaces the possible words with numbers. For example, Happy = 1, Sad = 2, Fear = 3 and so on.

4.2.2. Training

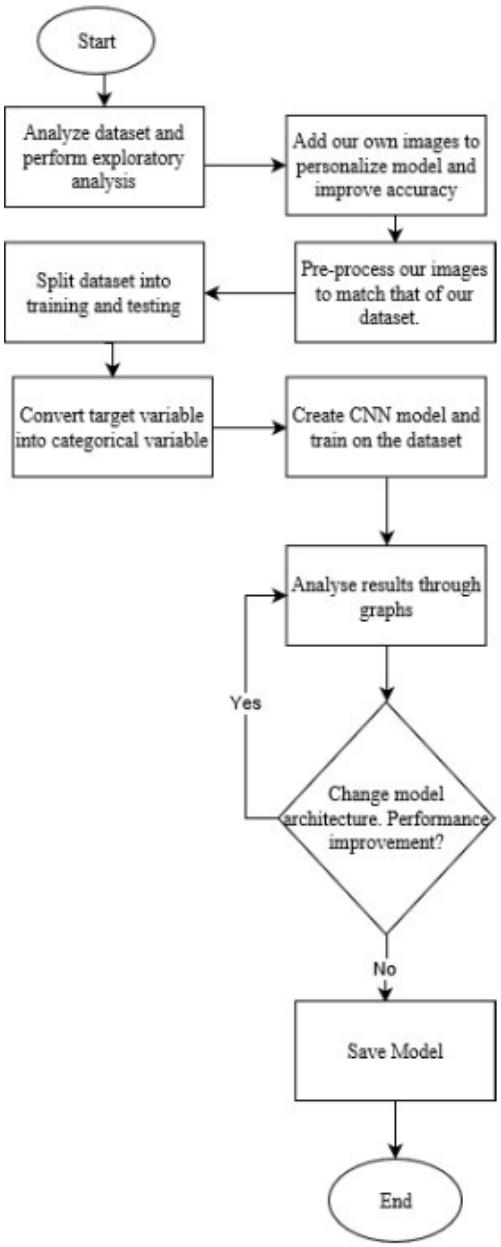


Figure 5. Facial expression recognition flow chart

Figure 5 shows the basic steps taken to perform the training process. These steps have been further elaborated below. We referred to [5] to help get an idea of what type

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

of model needs to be created and what the architecture of this model should be. We created a CNN that takes an input shape array of (48,48) and the final output is an array of probabilities which shows which of the classes the given image most closely belongs to. Figure 6 shows the model architecture.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 48, 48, 32)	320
conv2d_2 (Conv2D)	(None, 48, 48, 32)	9248
conv2d_3 (Conv2D)	(None, 48, 48, 32)	9248
max_pooling2d_1 (MaxPooling2	(None, 24, 24, 32)	0
conv2d_4 (Conv2D)	(None, 24, 24, 64)	18496
conv2d_5 (Conv2D)	(None, 24, 24, 64)	36928
conv2d_6 (Conv2D)	(None, 24, 24, 64)	36928
max_pooling2d_2 (MaxPooling2	(None, 12, 12, 64)	0
conv2d_7 (Conv2D)	(None, 12, 12, 128)	73856
conv2d_8 (Conv2D)	(None, 12, 12, 128)	147584
conv2d_9 (Conv2D)	(None, 12, 12, 128)	147584
max_pooling2d_3 (MaxPooling2	(None, 6, 6, 128)	0
flatten_1 (Flatten)	(None, 4608)	0
dense_1 (Dense)	(None, 64)	294976
dense_2 (Dense)	(None, 64)	4160
dense_3 (Dense)	(None, 6)	390

Total params: 779,718
Trainable params: 779,718
Non-trainable params: 0

Figure 6. Model architecture

Looking at the model we can see the input shape of the first layer is 48 x 48 since the size of the images in the dataset in 48 pixels by 48 pixels. The goal of the MaxPooling

Volume 3, Issue 1, pages 65–82
p-ISSN 2655-8564, e-ISSN 2685-9432

layers is to reduce the dimensionality of the image so that we can extract meaningful features from it.

4.3. Colour-based Follow Function

The Robot is trained to follow colour. We used the tools available in OpenCV to extract information like color from an image. We set a color range matching a very unique color of a particular shoe we have. The idea is that the robot should not receive presence of multiple areas of the same colour through its camera. Hence we have choosen a relatively unique colour for the follow mechanism; fluorescent yellow.

As shown in [6], we are able to create a circle around the pixels with the required color. This in turn gives us the radius of that circle. Using the size of this radius we then determine the approximate distance of the robot from the user. We have used simple if-condition statements based on the radius of this circle. Along with this we have defined two thresholds experimentally. These thresholds are used with the if-condition statements to give inputs to the motors of the robot.

- If the radius is smaller than lower threshold, move forward (Robot is too far from the user).
- If radius is greater than the upper threshold, move back (Robot is to close to the user).
- If radius is between lower and upper threshold, stay.

Using the location of the circle on the screen we have also defined conditions for the robot to turn left or right.

5 Analysis

We provide analysis in this section. This includes emotion recognition and colour-based follow function.

5.1. Emotion Recognition

Following is the performance of our model. We achieved an accuracy of 56% in the testing set however we find that we have greater accuracy in real time testing conditions

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

of specific emotions like happy, angry, neutral and surprise. Figure 7 shows the performance of this model through the epochs.

```
- acc: 0.2508 - val_loss: 1.7626 - val_acc: 0.2494
 acc: 0.2513 - val_loss: 1.7618 - val_acc: 0.2494
 acc: 0.2926 - val_loss: 1.5196 - val_acc: 0.3834
 acc: 0.4149 - val_loss: 1.4220 - val_acc: 0.4366
 acc: 0.4602 - val_loss: 1.3196 - val_acc: 0.4804
 acc: 0.4994 - val_loss: 1.2611 - val_acc: 0.4965
 acc: 0.5396 - val_loss: 1.2407 - val_acc: 0.5191
 acc: 0.5688 - val_loss: 1.2082 - val_acc: 0.5383
 acc: 0.6006 - val_loss: 1.1544 - val_acc: 0.5600
 acc: 0.6315 - val_loss: 1.1971 - val_acc: 0.5628
```

Figure 7. Model performance per epoch

In the figure above we have shown the performance of the model for 10 epochs. We can see that the model's performance on the training set starts surpassing its performance on the validation set. This gap widens as we train for more epochs as shown below in Figure 8.

```
acc: 0.9475 - val_loss: 2.9265 - val_acc: 0.5327
acc: 0.9490 - val_loss: 3.0391 - val_acc: 0.5464
acc: 0.9591 - val_loss: 3.0722 - val_acc: 0.5461
acc: 0.9616 - val_loss: 3.1316 - val_acc: 0.5456
acc: 0.9605 - val_loss: 3.3258 - val_acc: 0.5430
acc: 0.9640 - val_loss: 3.2917 - val_acc: 0.5375
acc: 0.9596 - val_loss: 3.0188 - val_acc: 0.5361
```

Figure 8. Overfitting of model

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

As we see that the training accuracy of the model crosses 90% whereas the validation accuracy actually starts dropping from 56%. This suggests that our model is overfitting the dataset. Figure 9 shows the accuracy and loss representation of our results.

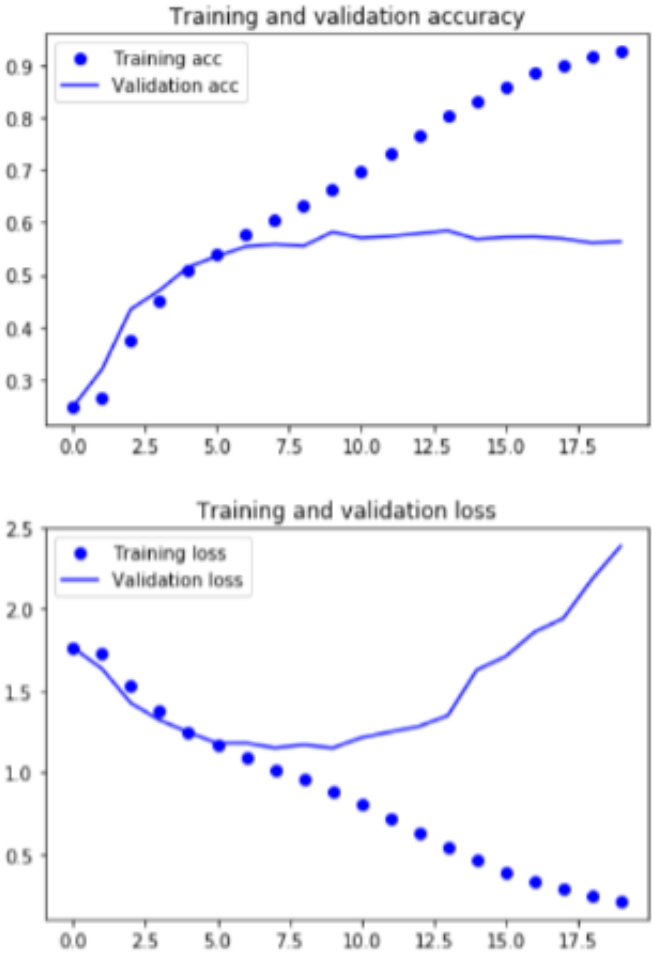


Figure 9. Accuracy and loss representation

We can see from both these graphs that after a certain point, training accuracy keeps on increasing but validation accuracy plateaus out. This further shows us that the model is overfitting the data and certain measures need to be taken. Reference [7] suggests a few steps we can try to stop our CNN from overfitting a dataset. The steps we have used are:

1. Use a simpler model.
2. Add Regularization.
3. Use batch normalization.

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

4. Use dropout.

Let us understand what these do to our model. References [7,8] explains that regularization works on the assumption that using smaller weights simplifies the model and prevents overfitting. L2 Regularization or Ridge Regularization term is the sum of square of all feature weights as shown in the equation below. L2 forces weights to be small but does not make then zero. Batch Normalization is a relatively new concept that was introduced after the VGG model. It is recommended to do this process for every model. This adds a normalization layer which helps the model to converge much faster in training. This in turn allows us to use higher learning rates. Another common measure used to stop overfitting is to use dropout. This randomly sets the activations of neurons to 0.

After taking into account the aforementioned steps, a new model was created which uses L2 Regularization, Batch Normalization, and Dropout layers. These help the model break past the previous highest accuracy by a large margin. We were able to achieve and accuracy of 66%. Figure 10 shows the performance of our model in different epochs.

```
acc: 0.7703 - val_loss: 1.0172 - val_acc: 0.6585

acc: 0.7743 - val_loss: 1.0385 - val_acc: 0.6601

acc: 0.7822 - val_loss: 1.0468 - val_acc: 0.6653

acc: 0.7823 - val_loss: 1.0442 - val_acc: 0.6625

acc: 0.7890 - val_loss: 1.0586 - val_acc: 0.6594

acc: 0.7919 - val_loss: 1.0655 - val_acc: 0.6663

acc: 0.7950 - val_loss: 1.0328 - val_acc: 0.6582

acc: 0.7980 - val_loss: 1.0267 - val_acc: 0.6672

acc: 0.8029 - val_loss: 1.0731 - val_acc: 0.6635
```

Figure 10. New model performance per epoch

Volume 3, Issue 1, pages 65–82
p-ISSN 2655-8564, e-ISSN 2685-9432

We can see by comparing Figure 8 and Figure 10 (previous model and current model) that the amount of overfitting is less in the new model. The accuracy has also increased significantly.

5.2. Colour-Based Follow Function

This follow function was tested in two different environments; indoors and outdoors. Due to the lack of proximity sensors in our robot we were worried about it crashing into other objects. However due to the colour-based following, these issues were not encountered. The robot simply did not move towards an object if it was not the required color. Due to our choosen color being rarely encountered in the everyday environment, we did not experience any unwanted behaviour during our testing. In Figure 11 we can see the robot's view of the follow routine. We have stuck a fluorescent yellow coloured paper to the shoe to emulate a sticker or an entire shoe that can be an accessory for the robot. In this example, we observe that the size of the circle is relatively small. This would prompt the robot to move towards the user. Depending on further movements from the user, the robots movements will also chance.

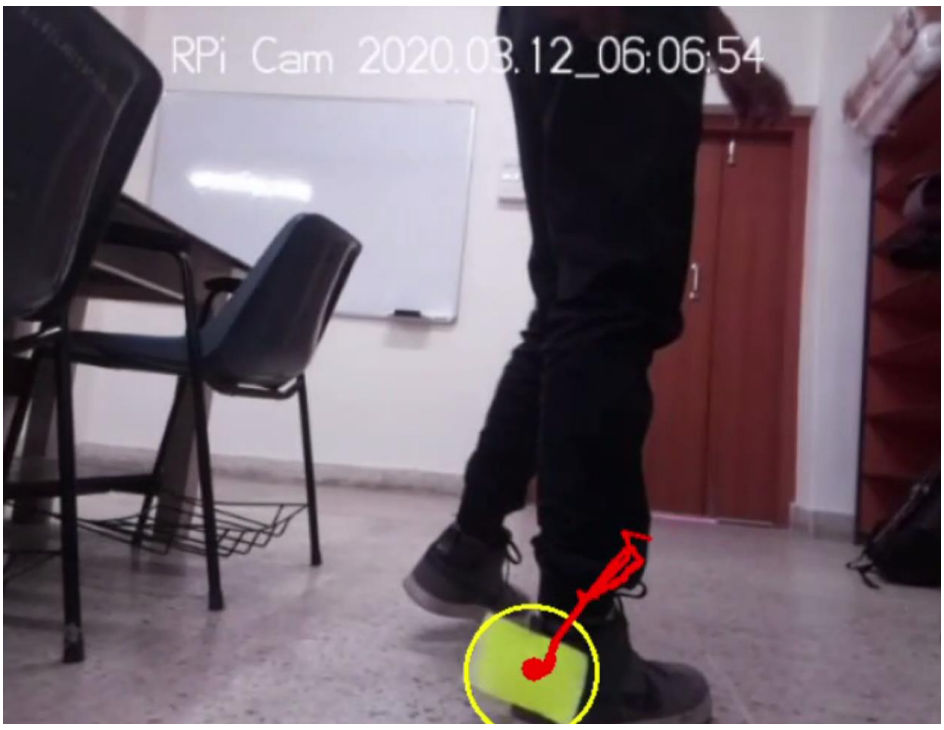


Figure 11. Real time follow functionality

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

6 Results and Discussion

Our results are provided in this section consisting of colour-based follow function and real time results for emotion recognition.

6.1. Colour-Based Follow Function

Using OpenCV and a normal camera we have created a basic distance measurement system. This enables the robot to judge how far or close it is to our desired colour in the frame. Our system is able to perform the required function in both indoor and outdoor environments provided there is ample illumination.

One of the inherent issues with the system is that when there is presence of more than one object with the desired colour, the robot cannot distinguish between the two. We would recommend using another kind of tagging mechanism which the robot can follow. For example, using RFID tags would prevent the robot from following any unwanted objects.

6.2. Real Time results for Emotion Recognition

After we have saved the model, we need to run this model for real time data. The real time module begins with the computer receiving the video feed from the Pi Camera mounted on the Raspberry Pi 4. Haar Cascade is used for finding a face in the individual frames of the video feed. If a face is not found in this step, Haar Cascade is run again on the subsequent frames till the face is found. In the meanwhile, the camera that is mounted on servos begin to sweep its 1800 field of view in a systematic way. If a face is found in this sweep, the motion stops and a snapshot of the frame is taken for further processing. The face is cut out from the entire frame and the resulting image is converted to monochrome and then resized to 48 X 48 pixels before it can be fed to the model for making a prediction:

1. A picture is captured from the camera (see Figure 12).
2. The Face is extracted using Haar-Cascade and converted to gray scale (see Figure 13).
3. This is then fed to the model and the model returns the predicted emotion (see Figure 14).

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

4. The final prediction can be seen on the screen (see Figure 15).

Figure 12. Emotion recognition input image

Figure 13. Grayscaling and face cutout

Figure 14. Emotion recognition output

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

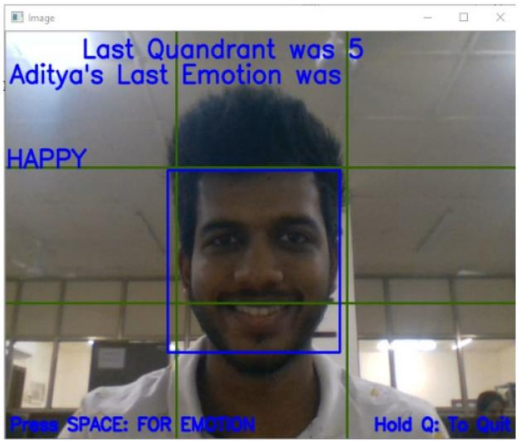


Figure 15. Detecting happy emotion with face recognition

7 Conclusion

The proposed algorithms act as some of the functions that our AI Pet Robot can perform. It is capable of differentiating between a known person and an unknown person. This will allow implementation of further functions such as "Sentry Mode" where the bot could sound an alarm if there is presence of unknown people. The robot is also capable of recognizing the emotions of the user. This could be a useful feature for someone undergoing therapy or counselling. Finally the follow functionality can allow the user to take the AI Pet Robot on walks like he would take any other pet, for example a dog.

References

[1] Center for Disease Control and Prevention, Healthy Pets, Healthy People, https://www.cdc.gov/healthypets/health-benefits/index.html, Accessed: 12/01/2020.

[2] E. Paul Cherniack, M. D and Ariella R. Cherniack, “Assessing the benefits and risks of owning a pet.” *Canadian Medical Association Journal*, 2015.

[3] P. Viola and M. Jones, “Rapid Object Detection using a Boosted Cascade of Simple Features.” *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, **1**, 2001.

Volume 3, Issue 1, pages 65–82

p-ISSN 2655-8564, e-ISSN 2685-9432

[4] Challenges in Representation Learning: Facial Expression Recognition Challenge, https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data, 2013.

[5] N. Poornadithya C., P.C. Chengappa, T. Raman, S. Pandey and G. K. Shyam, "Emotion Identification and Classification using Convolutional Neural Networks." *International Journal of Advanced Research in Computer Science*, 2018.

[6] A. Rosebrock, *Ball Tracking with OpenCV*, https://www.pyimagesearch.com/2015/09/14/ball-tracking-with-opencv/, 2015, Accessed 02/02/2020.

[7] R. Ruizendaal, *Deep Learning #3: More on CNNs Handling Overfitting*, https://towardsdatascience.com/deep-learning-3-more-on-cnns-handling-overfitting-2bd5d99abe5d, 2017, Accessed 23/02/2020.

[8] R. Khandelwal, *L1 and L2 Regularization*, https://medium.com/datadriveninvestor/l1-l2-regularization-7f1b4fe948f2, Nov. 4, 2018, Accessed 23/02/2020.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 83–92

p-ISSN 2655-8564, e-ISSN 2685-9432

Effects of the Existence of Fan in the Wood Drying Room and the Performance of the Electric Energy Wood Dryer

Wibowo Kusbandono[1,*], Petrus Kanisius Purwadi[1]

[1]*Department of Mechanical Engineering, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*
[*]*Corresponding Author: bowo@usd.ac.id*

(Received 24-01-2021; Revised 05-02-2021; Accepted 21-03-2021)

Abstract

The purpose of this study is to determine the effect of the presence of a fan in the wood drying room in the drying time of wood. In addition, it is also to determine the performance of the steam compression cycle engine used in wood drying machines and the conditions of air entering and leaving the wood drying room. Wood drying machines work on a source of electrical energy. The research was conducted experimentally. Variations in the study were carried out on the presence of fans in the drying room: (a) there were no fans and (b) there were 2 fans. The dried object is a sengon wood board, which has a length of 2 m, a width of 20 cm, and a thickness of 2 cm. The number of wooden planks is 70 wooden planks of uniform size. The wooden planks before drying have a moisture content of 29.6%, and when dry, have a moisture content of 10%. The research gave the following results: (a) if there are 2 fans in the drying room, the time needed to dry the sengon wood planks is around 42.6 hours, whereas if there is no fan around 49.9 hours (b) the average Coefficient of Performance (COP) of the steam compression cycle engine is 10.65 (c) The air condition enters the drying room when

Volume 3, Issue 1, pages 83–92
p-ISSN 2655-8564, e-ISSN 2685-9432

there are 2 fans, has a dry ball air temperature of 40^{o}C with a relative humidity of 32% RH and the air condition when it comes out, has a dry ball air temperature of 28^{o}C with a relative humidity of 73% RH.

Keywords: Wood drying machine, steam compression cycle, electricity, COP

1 Introduction

Wood is the main material in the furniture industry. Before being processed into finished goods, the wood must be dry. If it is not dry, the products can change shape into unwanted shapes. One of the problems that exist in the wood industry is how to dry wood that does not interfere with production. At this time, generally the wood drying process is carried out with the help of solar energy or with fuel energy from wood that is no longer used. For energy from wood fuel, besides being complicated, it is also not environmentally friendly. In addition, the time used for the drying process is long. So it needs careful calculations so that production is not disturbed by the absence of dry wood. The use of solar energy is more practical, more environmentally friendly, cheaper, with unlimited energy sources, free and available everywhere, but this method cannot be relied on when the rainy season arrives. Even though the rainy season lasts quite a long time. It is necessary to look for other alternatives, which are more practical, environmentally friendly and can be done anytime and anywhere. The solution is to use electrical energy.

Many studies related to the drying process with electrical energy have been carried out. A suitable solution for the use of electrical energy is to use a steam compression cycle machine. As has been done by these researchers, in the drying process. Steam compression cycle machine, used to produce dry and hot air. With dry and hot air, the object drying process can run well. In the research that has been done, the dried objects are different. Mitsunori [1] conducted research with the cloth object. Balioglu [2] and Bison [3] conducted research with the object of laundry. Kusbandono and Purwadi [4,5] conducted research with the object to be chilled.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 83–92

p-ISSN 2655-8564, e-ISSN 2685-9432

From the research that has been conducted by these researchers, a question arises for the authors, can the wood drying process be carried out using electrical energy involving a steam compression cycle machine. As it is known, that wood is a solid, dense, hard object, with a relatively large size, with a very low moisture content in the wood. Starting from this problem, the research was carried out.

1.1. Steam Compression Cycle

The main components of a steam compression machine include: compressor, condensor, capillary tube and evaporator. The working fluid used in the vapor compression cycle is a refrigerant. Figure 1 presents a series of main components of a steam compression engine and Figure 2 presents a simple vapor compression cycle shown on the p-h diagram.

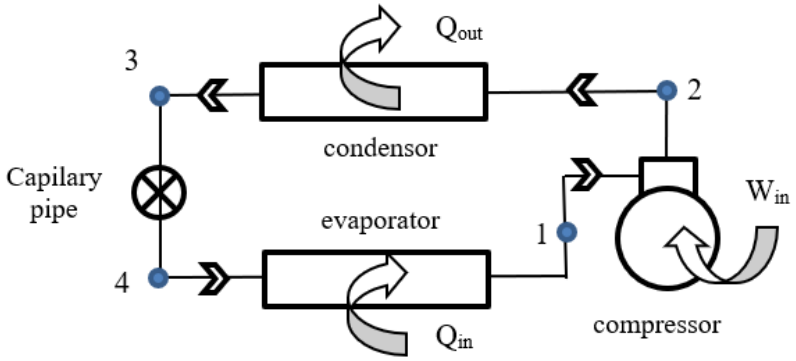


Figure 1. Circulation of the main components of a steam compression cycle engine

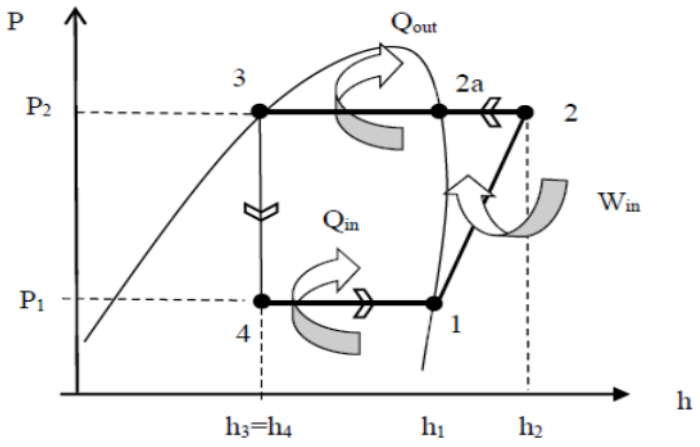


Figure 2. The vapor compression cycle in the p-h diagram

Volume 3, Issue 1, pages 83–92

p-ISSN 2655-8564, e-ISSN 2685-9432

The process of the steam compression cycle engine includes: (1) process 1-2: isoentropic compression (2) process 2-2a: desuperheating (3) process 2a-3: condensation or condensation (4) process 3- 4: throttling or lowering pressure isoenthalpy (fixed enthalpy) (5) 4-1 process: evaporation or refrigerant evaporation process. During the evaporation process, heat flows from the ambient air around the evaporator to the refrigerant that flows in the evaporator pipe. The amount of heat absorbed by the mass-refrigerant unity evaporator is Q_{in}. The amount of heat released by the mass-refrigerant unity condensor is Q_{out}. During the compression process, the compressor works to increase the refrigerant pressure, from low pressure to high pressure or from the evaporator working pressure to the condensor working pressure. Work done by the compressor During the desuperheating and condensation process, heat flow occurs from the condensor out into the environment around the condensor. The mass-refrigerant unit is equal to W_{in}. The amount of work done by the compressor is the power supplied by a power source or from electrical energy to the mass-refrigerant unity compressor. In this steam compression cycle, the subcooling and superheating processes are eliminated.

The amount of heat absorbed by the mass-refrigerant unity evaporator (Q_{in}) can be expressed by

$$Q_{\text{in}} = h_1 - h_4. \quad (1)$$

The amount of heat released by the mass-refrigerant unity condensor (Q_{out}) can be expressed by

$$Q_{\text{out}} = h_2 - h_{3.} \quad (2)$$

The amount of work performed by the mass-refrigerant unity compressor (W_{in}) can be expressed by

$$W_{\text{in}} = h_2 - h_1. \quad (3)$$

The values of h_1, h_2, h_3 and h_4 are the enthalpy value of the refrigerant when it enters the compressor, the enthalpy of the refrigerant when it leaves the compressor, the enthalpy of the refrigerant when it leaves the condensor and the enthalpy of the refrigerant when it enters the evaporator.

Volume 3, Issue 1, pages 83–92
p-ISSN 2655-8564, e-ISSN 2685-9432

The performance (Coefficient of Performance or COP) of the steam compression cycle machine used in wood drying machines with electrical energy is the ratio between the useful energy for mass-refrigerant unity with the amount of energy required for mass-refrigerant unity. The Coefficient of Performance or COP of the steam compression cycle of this electric energy wood drying machine can be expressed by

$$\mathrm{COP} = \frac{(Q_{\mathrm{in}} - Q_{\mathrm{out}})}{W_{\mathrm{in}}}. \quad (4)$$

The heat absorbed by the evaporator causes the ambient air that passes through the evaporator to dry, and the heat released by the condensor causes the temperature of the air that passes through the condensor to increase to become hot. In Equation (4), the energy used to drive the fan has not been calculated: the evaporator fan, the condensor fan and the fan in the drying room.

1.2. Open Air System Wood Drying Machine

The wood drying machine used in this study uses an open air system. In this open air system, the air used as a wood drying medium is taken from the outside air. The outside air is introduced into the drying room, by first passing through the evaporator and condensor. Outside air can flow in because of the evaporator fan and condensor fan. After the air is put into the drying room, it performs the wood drying process. The air that has been used to dry wood is then flown out of the dryer, and is no longer used.

2 Research Methodology

In this section we describe about how this research was carried out among others the research methods and research variations, research flow, equipment used, dried object, and how to collect data.

2.1 Research Methods and Research Variations

The research was conducted experimentally. Variations in the study were carried out on the presence of fans in the wood drying room: (a) there was no fan and (b) there were two fans in the wood drying room.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 83–92

p-ISSN 2655-8564, e-ISSN 2685-9432

2.2 Research Flow

The research flow follows the flow as presented in Figure 3.

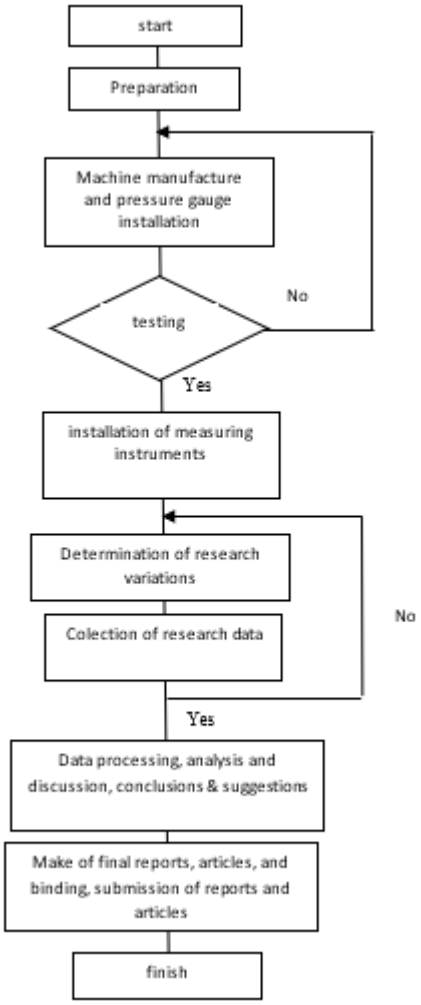


Figure 3. Research flow

2.3 Equipments

The rotary compressor power used in the steam-compression cycle engine is 1 HP. The other main component, its size adjusts to the amount of compressor power. The compression cycle uses R134a refrigerant. The power per fan used in the wood drying room is 30 watts. The evaporator fan has 17.5 watts of power and the condensor fan has a power of 20 watts. The wood drying machine has a total length of 350 cm, a total height of 180 cm and a total width of 140 cm.

Volume 3, Issue 1, pages 83–92
p-ISSN 2655-8564, e-ISSN 2685-9432

2.4 Dried object

The object that was dried was a wet sengon wood board with an initial moisture content of 29.6%. The wet wooden planks are dried until they have a moisture content in the wood of 10%. The size of the wooden boards to be dried is uniform, has a length of 2 m, a width of 20 cm and a thickness of 2 cm. Number of board to dry: 70 boards of wood, for one drying. Wooden boards are arranged on racks, shelves are arranged from bottom to top, one shelf contains 5 wooden boards (horizontal direction), there are 14 shelves arranged.

2.5 Data collection

Before entering wet wood, the engine is started first. Data collection starts after all the wet wood is in the drying room, and the door to the wood drying room has been closed. Data collection was stopped after all sengon wood was dry (having moisture content in the wood <10%). Measurement of moisture content in wood using a digital moisture meter. Temperature and humidity data were measured using a thermocouple and hygrometer. Meanwhile, for low pressure (evaporator working pressure) and high pressure (condensor working pressure) in the steam compression cycle using a pressure gauge. The enthalpy data is taken from the P-h diagram.

3 Results and Discussion

Figure 4 presents the drying process of sengon wood boards from time to time, from the initial condition of the wet wood boards with an average moisture content of 29.6% to an average moisture content of 10%. It appears that the presence of a fan in the wood drying room affects the drying time of the wood. The existence of a fan, causes the airflow to be faster and the drying time is faster. In other words, the drying time of wood is affected by the speed of air flow in the drying room. The greater the speed of air flow across the wood surface, the faster the drying time. If there is no fan in the drying room, it means that the air flow that occurs in the drying process is only caused by the evaporator fan and condensor fan.

From Table 1 and Figure 4, information is obtained, if there is no fan in the drying room, the wood drying time is 49.9 hours, and if there is a fan, the wood drying time is only 42.6 hours. The use of 2 fans can shorten the drying time by about 7.3 hours or

Volume 3, Issue 1, pages 83–92

p-ISSN 2655-8564, e-ISSN 2685-9432

speed up the drying time by 14.6%. Even though it is able to shorten the wood drying time, the wood drying process requires additional electrical power which is used to drive 2 fans. The air condition when there are 2 fans in the drying room is presented in Table 2. Further research is needed, regarding the effect of adding a fan on the optimal work of a wood drying machine.

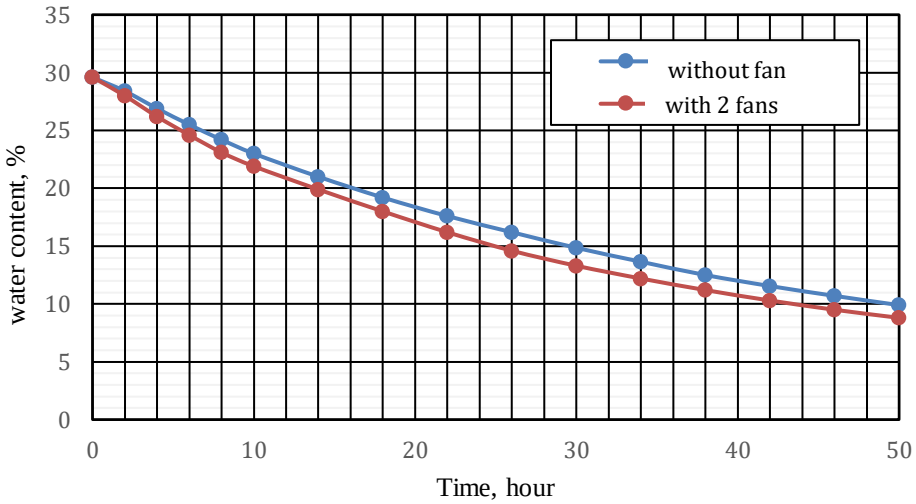


Figure 4. Moisture content in sengon wood from time to time

Table 1. Drying time of sengon wood

No	Research Variations	Drying time t
		t (hour)
1	There is No fan in wood dryer room	49,9
2	There are 2 fans in the wood dryer room	42,6

Table 2. Air conditions in and out of the sengon wood drying room

No	Conditions in the wood drying room	The air condition exits the wood drying room		The air condition enters the wood drying room	
		T_{db} (°C)	RH (%)	T_{db} (°C)	RH (%)
1	There is not fan	30	68%	40,5	31
2	There are 2 fans	28	73%	40	32

With increasing air flow velocity, the greater the air flow across the wood. The greater the air flow, the greater the ability of air to take water from the wood. Because the air system is made open, the air that has passed through the wood is immediately flown out of the drying chamber, and the air is replaced with new air coming from the condensor.

Volume 3, Issue 1, pages 83–92

p-ISSN 2655-8564, e-ISSN 2685-9432

The new air used to dry the wood is dry and the temperature is high enough. The component of the vapor compression cycle which functions to make the air dry is the evaporator and the component that makes the air a high enough temperature is the condensor.

Having 2 fans in the wood drying chamber does not have much impact on the characteristics of the steam compression cycle machine. The air fan has more impact on the evenness of the air flow velocity that occurs in the drying room. More even and easier to get out of the drying room. Because the air flow system is open, the airflow flowing through the evaporator and condensor is more dominant due to the evaporator fan and condensor fan only. However, the work of the evaporator fan and the condensor fan work becomes lighter with the fan in the drying room. Average characteristics of the steam compression cycle machine on a wood drying machine are presented in Table 3.

Table 3. Performance of steam compression cycle machine on wood drying machine (Sengon wood drying)

No	$T_{evap,}$ °C	$T_{kond,}$ °C	Q_{in} (kJ/kg)	Q_{out} (kJ/kg)	W_{in} (kJ/kg)	COP
1	6	50	130,3	157,3	27	10,65

4 Conclusion

The results of the research for the electric energy wood drying machine steam compression cycle of this closed air system are: (a) if there are 2 fans in the drying room, the time required to dry the sengon wood planks is 42.6 hours, whereas if there is no fan for 49.9 hours; (b) The Coefficient of Performance or COP of the wood drying machine with an average electrical energy of 10.65; (c) the air enters the drying room when there are 2 fans, has a dry ball air temperature of 40^{o}C with 32% RH and the dry ball air temperature out from 28^{o}C drying chamber with 73% RH. This research can be developed to determine the relationship between the condensor working temperature and the COP and the time required for the drying process or it can be developed by looking for the relationship between the working temperature of the evaporator and the performance of the dryer and the time required for the wood drying process.

Volume 3, Issue 1, pages 83–92

p-ISSN 2655-8564, e-ISSN 2685-9432

References

[1] T. Mitsunori, "Dehumidifying and heating apparatus and clothes drying machine using the same", *European Patent Specification*, EP 2 468 948 B1, 27.11.2013, 2013.

[2] Balioglu, "Heat Pump Laundry Dryer Machine", *Patent Application Publication*, Pub. No: US 2013/0047456 A1, April 2013.

[3] Bison, "Heat Pump Laundry Dryer and a Method for Operating a Heat Pump Laundry Dryer", *Patent Application Publication*, Pub. No: US 2012/0210597 A1, 2012.

[4] W. Kusbandono and P.K. Purwadi, "Pengaruh Adanya Kipas yang Mengalirkan Udara Melintasi Kondensor terhadap COP dan Efisiensi Mesin Pendingin Showcase." *Prosiding Seminar Nasional XI Rekayasa Teknologi Industri dan Informasi*, 313-317, 2016.

[5] W. Kusbandono and P.K. Purwadi, "COP Mesin Pendingin Refrigeran Sekunder." *Jurnal Penelitian*, **19** (1), 79-86, 2015.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 93–100

p-ISSN 2655-8564, e-ISSN 2685-9432

Effects of Shock Wave Phenomenon on Different Convergent Lengths in the Mixing Chamber of the Steam Ejector

Stefan Mardikus[1,*]

[1]*Department of Mechanical Engineering, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*
[*]*Corresponding Author: stefan@usd.ac.id*

(Received 26-03-2021; Revised 24-04-2021; Accepted 05-05-2021)

Abstract

The shock wave phenomenon is a phenomenon in a steam ejector that caused when the working fluid has high pressure, and suddenly it turns into low pressure and high speed. The shock wave effect will be investigated to the different convergent length in the mixing chamber to find the highest entrainment ratio as the performance of steam ejector. Operating pressure in the primary flow was in the range 0.68 MPa - 1.39 MPa, and the secondary flow was set 0.38 MPa to 0.65 MPa. The result of this study demonstrated that the highest entrainment ratio occurred in the convergent length of 69 mm.

Keywords: Steam ejector, shock wave, convergent length

1 Introduction

The Steam Ejector is a utilization that is used to recover waste energy from low pressure fluid (secondary pressure) and low thermal energy to high-pressure fluid (primary pressure) without the use of the moving part and electrical energy

Volume 3, Issue 1, pages 93–100
p-ISSN 2655-8564, e-ISSN 2685-9432

resources [1]. On the other hand, the steam ejector can be used to mixing fluid in the chemical industry when the fluid has different physical properties through operating condition by pressure or temperatures of fluid. The steam ejector system was applied in many industries as power plant, refrigeration, food processing technology, etc. For example, the difficult problem to handle the liquid or gas corrosively in process industries, the ejector can pump this fluid through different pressure when the primary fluid passed part of the nozzle; thus; the secondary fluid can be moved to the mixing chamber area [2]. The primary fluid that has high pressure and temperature when passed the nozzle decreased the pressure and temperature. Moreover, the fluid flow of the primary fluid shaped expansion angle in the mixing chamber; therefore, the low-pressure fluid as secondary fluid can entrain the ejector's system. In the chemical industry, the steam ejector was used to mix the substance with different properties of physics. Furthermore, the steam ejector model geometry consists of the nozzle and mixing chamber area like convergent, throat and divergent is an essential part to enormously improve the steam ejector performance [3].

In many investigations of steam ejector performance, Navid and Majid optimized the nozzle to reduce the steam ejector system's energy across the nozzle geometries. The investigation proved that the nozzle geometries' improvement could augment the entrainment ratio of steam ejector [4]. Meanwhile, the shockwave phenomenon in the convergent and divergent nozzle was evaluated by Yinhai that the increase of shockwave length can decrease characteristic of the steam ejector performance [5]. The shockwave phenomenon appears when the working fluid has high pressure, and suddenly it turns into low pressure and high speed. Furthermore, the expansion angle of fluid flow will show. One of the strategies to improve the steam ejector's performance will be investigated based on the shockwave phenomenon in the mixing chamber area. This paper will be experimentally evaluated by the different convergent length that can provide shock wave phenomenon with several operating pressure conditions in primary and secondary flow to find the highest entrainment ratio at the different of convergent length.

International Journal of Applied Sciences and Smart Technologies
Volume 3, Issue 1, pages 93–100
p-ISSN 2655-8564, e-ISSN 2685-9432

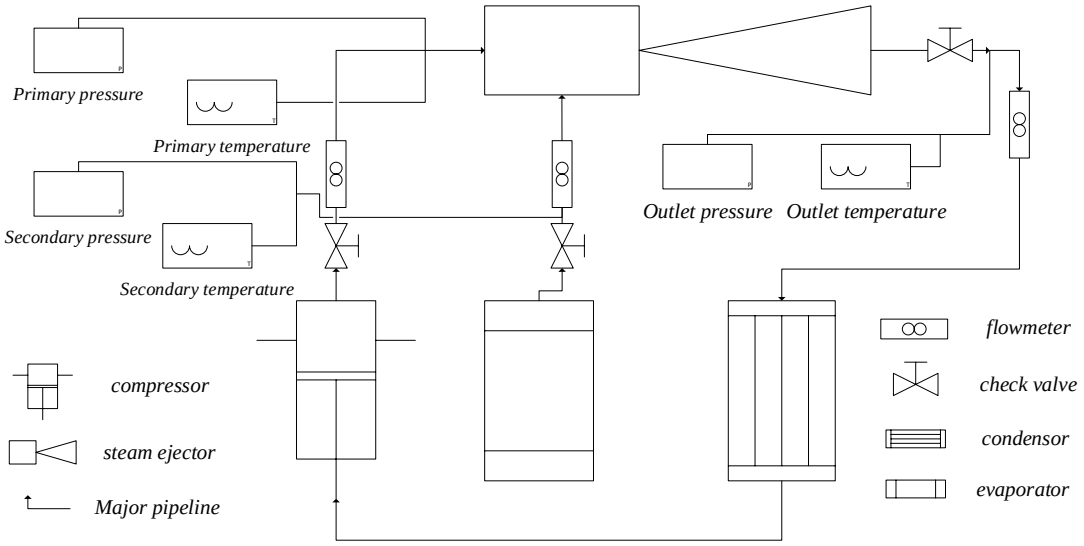


Figure 1. The schematic of steam ejector experimental.

2 Research Methodology

The schematic of the experimental steam ejector can be seen in Figure 1. The part of the main experimental setup consists of a steam ejector (1), compressor (2), condenser (3), an evaporator (4). In this research, 1 PK compressor used to compress the primary fluid. The characteristic of the primary fluid, secondary fluid, and discharge fluid was measured by thermocouple type K and pressure gauge. Primary pressure and secondary pressure can be set using regulator valve. The primary pressure set to operate 0.68 MPa to 1.39 MPa and the secondary pressure 0.38 MPa to 0.65 MPa. This study uses an R-22 as the working fluid. The condenser was an air cooled plate heat exchanger constructed from a 3/8inch diameter cooper pipe. The steam ejector geometry model can be seen in Figure 2, where the convergent length will be investigated with the different model. Three types of models of convergent length are 51 mm, 69 mm, and 75 mm.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 93–100

p-ISSN 2655-8564, e-ISSN 2685-9432

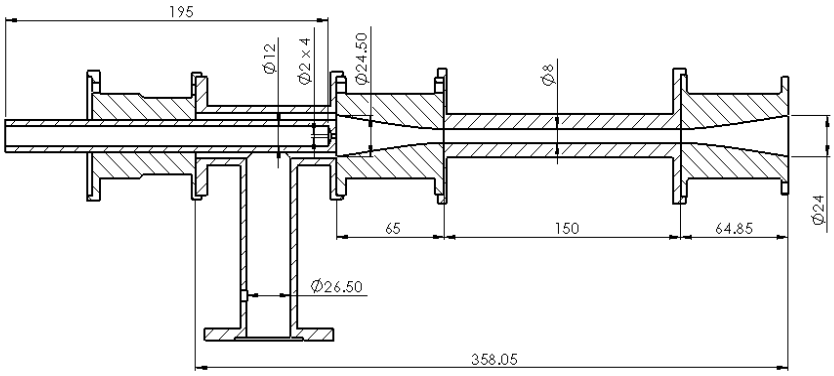


Figure 2. Geometry model of the steam ejector component.

3 Results and Discussion

As shown in Figures 3-6, our experiment results showed that the optimum entrainment ratio with different convergent length was on the convergent length 69 mm for all secondary pressure due to the shock wave phenomenon that appeared when the working fluid through the nozzle part decreased pressure and increasing velocity suddenly occurred. Increasing the primary pressure can augment the mass flow rate of working fluid into the mixing chamber.

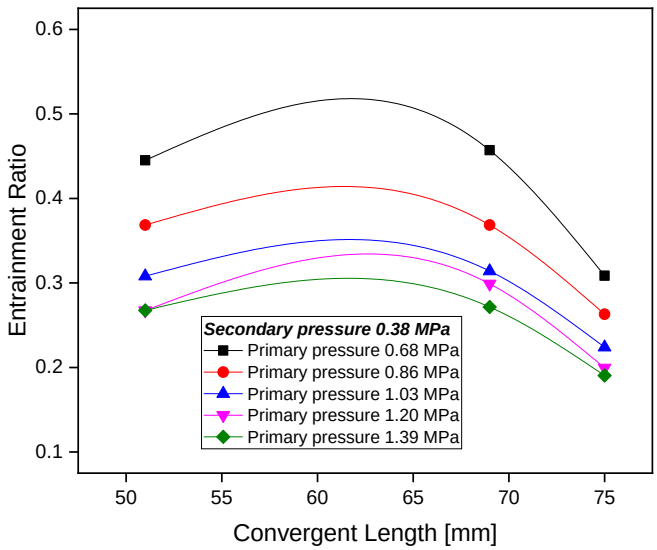


Figure 3. Effect of entrainment ratio versus the convergent length variations on secondary pressure 0.38 MPa in case operations of primary pressure

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 93–100

p-ISSN 2655-8564, e-ISSN 2685-9432

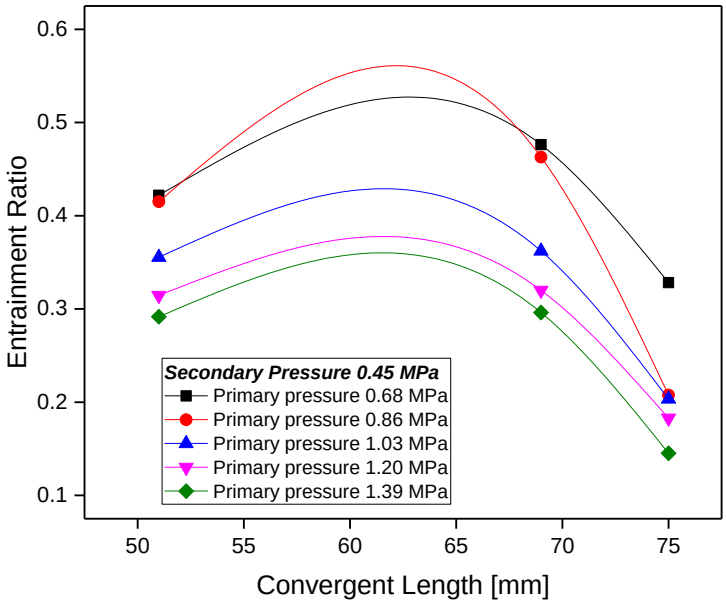


Figure 4. Effect of entrainment ratio versus the convergent length variations on secondary pressure 0.45 MPa in case operations of primary pressure

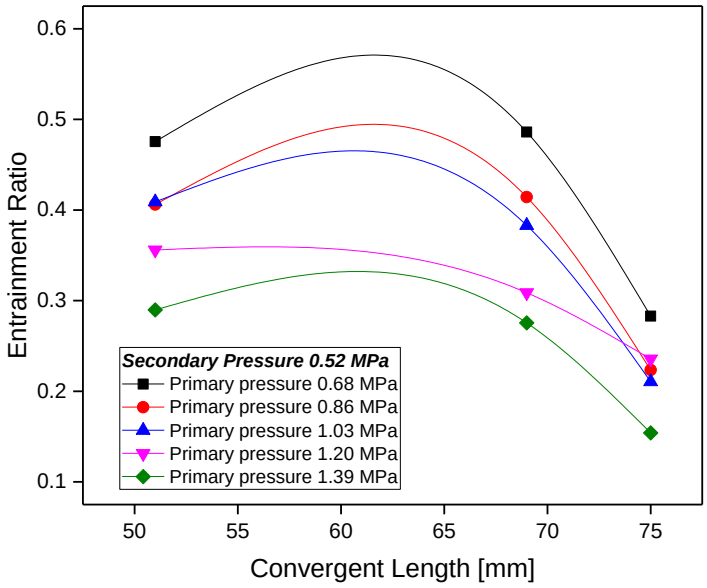


Figure 5. Effect of entrainment ratio versus the convergent length variations on secondary pressure 0.52 MPa in case operations of primary pressure

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 93–100

p-ISSN 2655-8564, e-ISSN 2685-9432

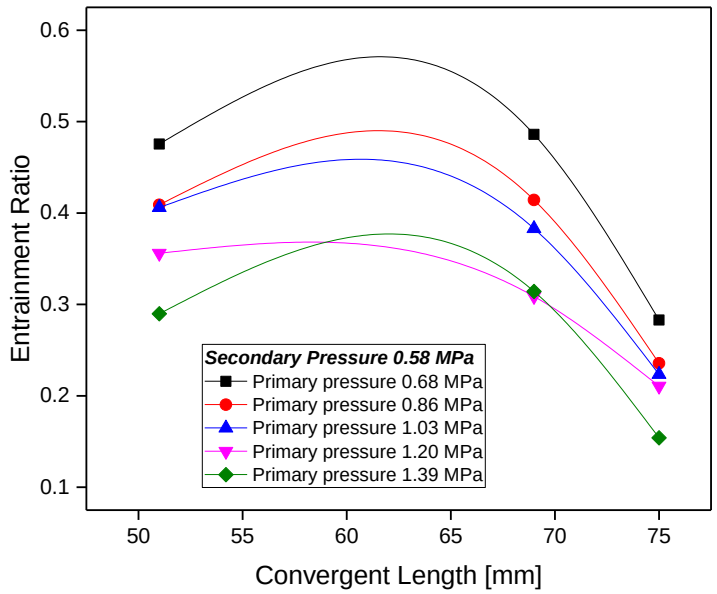


Figure 6. Effect of entrainment ratio versus the convergent length variations on secondary pressure 0.58 MPa in case operations of primary pressure

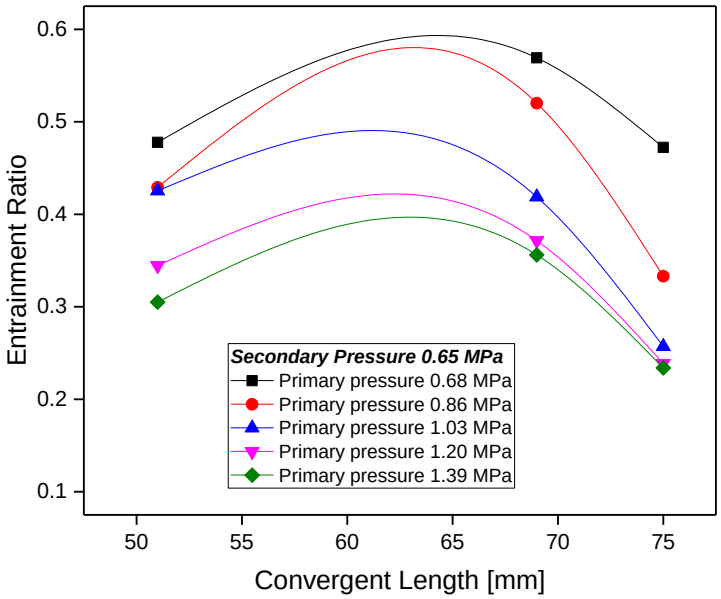


Figure 7. Effect of entrainment ratio versus the convergent length variations on secondary pressure 0.65 MPa in case operations of primary pressure

Volume 3, Issue 1, pages 93–100
p-ISSN 2655-8564, e-ISSN 2685-9432

In Figure 7, the highest entrainment ratio occurred at the primary pressure of 0.68 MPa. This experiment was found that operating of primary pressure could effect the entrainment ratio as the performance of steam ejector. At the same of the primary pressure, when the convergent length more extensive, the entrainment ratio slightly decreased at all of operating pressure in secondary because there was a small expansion angle in convergent length that more extensive [6]. When it happened, the mass flow rate of secondary pressure slightly reduced to move into the steam ejector's mixing chamber [5]. This phenomenon was caused by entraining duct shorter in the suction chamber area [2].

4 Conclusion

Based on experimentally evaluated results, the different convergent length in the mixing chamber could influence the steam ejector performance like an entrainment ratio. The shock wave phenomenon occurred when working fluid passed the nozzle area, and it could provide the diamond shock wave in the mixing chamber. The diamond shock wave phenomenon became one of the important things that caused the entrainment of steam ejector was augmented; thus; on this investigation, the highest entrainment ratio was occurred in the convergent length of 69 mm.

References

[1] J. Dong, X. Chen, W. Wang, C. Kang, and H. Ma, “An experimental investigation of steam ejector refrigeration system powered by extra low temperature heat source.” *Int. Commun. Heat Mass Transf.*, 2017.

[2] V. V. Chandra and M. R. Ahmed, “Experimental and computational studies on a steam jet refrigeration system with constant area and variable area ejectors.” *Energy Convers. Manag.*, 2014.

[3] N. Ruangtrakoon, S. Aphornratana, and T. Sriveerakul, “Experimental studies of a steam jet refrigeration cycle: Effect of the primary nozzle geometries to system performance.” *Exp. Therm. Fluid Sci.*, 2011.

[4] N. Sharifi and M. Sharifi, “Reducing energy consumption of a steam ejector through experimental optimization of the nozzle geometry.” *Energy*, 2014.

Volume 3, Issue 1, pages 93–100

p-ISSN 2655-8564, e-ISSN 2685-9432

[5] Y. Zhu and P. Jiang, "Experimental and analytical studies on the shock wave length in convergent and convergent-divergent nozzle ejectors." *Energy Convers. Manag.*, 2014.

[6] Y. Wu, H. Zhao, C. Zhang, L. Wang, and J. Han, "Optimization analysis of structure parameters of steam ejector based on CFD and orthogonal test." *Energy*, 2018.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

Identity Graph of Finite Cyclic Groups

Maria Vianney Any Herawati[1,*], Priscila Septinina Henryanti[1], Ricky Aditya[1]

[1]*Department of Mathematics, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*

[*]*Corresponding Author: any@usd.ac.id*

(Received 26-03-2021; Revised 24-04-2021; Accepted 05-05-2021)

Abstract

This paper discusses how to express a finite group as a graph, specifically about the identity graph of a cyclic group. The term chosen for the graph is an identity graph, because it is the identity element of the group that holds the key in forming the identity graph. Through the identity graph, it can be seen which elements are inverse of themselves and other properties of the group. We will look for the characteristics of identity graph of the finite cyclic group, for both cases of odd and even order.

Keywords: Graph, identity graph, group, identity element.

1 Introduction

Mathematics as a science has several branches including abstract algebra and graph theory [1-3]. The phrase of abstract algebra has been used since the early 20th century to distinguish them from what is now more commonly referred to as elementary algebra, which is the study of the rules of manipulation of algebraic formulas and expressions involving real or complex variables and numbers [4-6]. Abstract algebra is a field of

Volume 3, Issue 1, pages 101–110
p-ISSN 2655-8564, e-ISSN 2685-9432

mathematics that studies algebraic structures, such as monoids, groups, rings, fields, modules, etc. [3, 4].

Students often find it difficult to learn abstract structure such as group. Therefore, some writers are looking for a way to represent a group with a diagram called a graph. Graph theory is a branch of mathematics that has been studied and developed by researchers. In its development, the application of graph theory is often found both in mathematics itself and in other fields such as computer science, biology, chemistry and in problems in human life such as transportation problems, installation of public facilities, and traffic light management.

In this paper, graph theory will be applied in abstract algebra, especially to represent groups in the form of a graph so that it can be visualized diagrammatically and studied its properties through the graph of the group. The group discussed here is a finite group. There are several previous articles that examine graphs formed from groups including Cayley graphs, G-graphs, coprime graphs, and identity graphs of dihedral groups.

2 Methodology: Notations and Definitions

The method used is literature study with the initial step of forming an identity graph of several cyclic groups then looking for general patterns of their properties, making conjectures and proving them. Before going into those steps, in this section we will discuss some basic concepts and definitions in group theory and graph theory.

Group Theory

These are some definitions in group theory [4] which will be used in the next section:

1. Group is a set with one binary operation on the set which fulfills associative properties, has an identity element, and each member of the group has an inverse.
2. Order of a group is the number of its elements. A finite group is a group of finite order. Let e is the identity element of a finite group G. Order of an element a in G is defined as the smallest natural number n such that $a^n = e$.
3. Let G be a group. A non-empty subset $H \subseteq G$ is called a subgroup of G if and only if H is also a group with the same operation defined in G [4].

Volume 3, Issue 1, pages 101–110
p-ISSN 2655-8564, e-ISSN 2685-9432

4. If G is a group and $a \in G$, then the set $\langle a \rangle = \{a^n : n \in Z\}$ is a subgroup of G, and $\langle a \rangle$ is called the cyclic subgroup generated by a. Group G is called a cyclic group if and only if $a \in G$ exists, such that $G = \langle a \rangle$ [4].

Related to order of groups and order of elements, we have these two important theorems in group theory [4]:

1. (Cauchy's Theorem) Let G be a finite group and p be a prime number. If p divides the order of G, then G has an element of order p.
2. (Lagrange's Theorem) If H is a subgroup of a finite group G, then order of H divides order of G.

Graph Theory

Graph G is a pair of finite sets (V, E), written with the notation $G\ (V, E)$, in which case V is a non-empty set of vertices and E is a non-empty set of edges connecting a pair of vertices or connect a vertex with the vertex itself. A graph G can be represented by a diagram, each vertex of G is represented by a dot or small circle while an edge connecting two vertices is represented by a curve connecting the corresponding vertices in the diagram.

3 Results and Discussion

In this section, we write our research results in terms of theorems and their proofs. Some illustrations of graphs are also presented. First, we need to understand the concept of identity graph of a group.

Definition [2]

Let G be a group. The identity graph of group G is a graph with the elements of group G as its vertices which satisfies these properties:

a) Two elements x, y in group G are connected by an edge if $xy = e$, with e is the identity element for group G.

b) Each element of G is connected by an edge with the identity element e.

To develop the previous research, we shall examine the identity graph of finite cyclic groups. There are two possibilities for the order of a finite cyclic group: it is an odd

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

natural number, or it is an even natural number. The order may also be a prime number. We will examine the case of odd prime order first.

Theorem 1 [2]

If $G = \langle g \mid g^p = 1, p \neq 2 \rangle$ is a cyclic group of the pth order, where p is prime, then the identity graph formed by G consists of $(p-1)/2$ triangles.

Proof:

Let $G = \langle g \mid g\text{^}p = 1, p \neq 2 \rangle$ be a cyclic group of the pth order, where p is prime. Then G does not have a proper subgroup, according to Lagrange theorem. Therefore, there is no element in G having inverse which is itself; in other words, there is no $g^i \in G$ such that $\left(g^i\right)^2 = 1$. Suppose there is g^i in G such that $(g^i)\text{^}2 = 1$. Then, G has a subgroup $H = \{1, g^i\}$. This contradicts with the fact that G does not have a proper subgroup. As a result, G does not have element of the 2nd order. Using Cauchy theorem with $p \neq 2$ then G does not have element of the 2nd order.

For every(g^i) in G there is exactly one inverse of g^i that is g^j such that $g^i\, g^j = 1$ with i and j are positive integers and $i \neq j$. Because $g^i\, g^j = 1$ then $g^i\, g^j = g^{i+j} = 1 = g^p$ such that $p = i + j$ which is equivalent to $j = p - i$. Consequently, we can form the identity graph of G as illustrated by Figure 1.

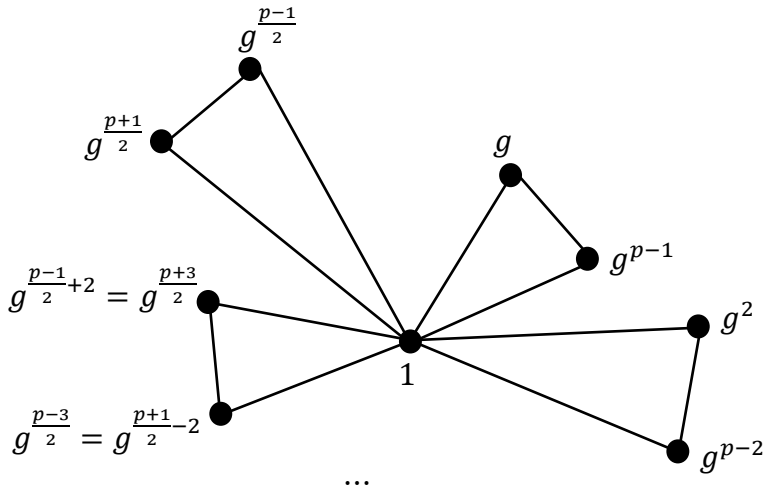


Figure 1. Identity graph of $G = \langle g \mid g^p = 1, p \neq 2 \rangle$

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

We have seen a unique pattern of the identity graph of a finite cyclic group with odd prime order. We skip the case of even prime order since the only even prime number is 2, and the identity graph of group of order 2 looks too trivial and not so interesting. Now we look at more general case of finite cyclic groups of odd order.

Theorem 2 [2]

If G is a cyclic group with an odd order, then G has the identity graph G_i which can be formed by triangles without a unique edge.

Proof:

Let $G = \langle g \,|g^n = 1\rangle$ with n is an odd integer is a cyclic group with multiplication operation and is of the nth order. The elements of G are $\{1, g, g^2, g^3, g^4, \cdots, g^{n-1}\}$. We shall use the Cayley table given by Table 1 to show results of the operation for each element of G.

Table 1. Cayley table for Theorem 2.

	1	g	g^2	g^3	$\cdots$	g^{n-3}	g^{n-2}	g^{n-1}
1	1	g	g^2	g^3	$\cdots$	g^{n-3}	g^{n-2}	g^{n-1}
g	g	g^2	g^3	g^4	$\cdots$	g^{n-2}	g^{n-1}	1
g^2	g^2	g^3	g^4	g^5	$\cdots$	g^{n-1}	1	g
g^3	g^3	g^4	g^5	g^6	$\cdots$	1	g	g^2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
g^{n-3}	g^{n-3}	g^{n-2}	g^{n-1}	1	$\cdots$	g^{n-6}	g^{n-5}	g^{n-4}
g^{n-2}	g^{n-2}	g^{n-1}	1	g	$\cdots$	g^{n-5}	g^{n-4}	g^{n-3}
g^{n-1}	g^{n-1}	1	g	g^2	$\cdots$	g^{n-4}	g^{n-3}	g^{n-2}

From the Cayley table above (Table 1) we can see that $gg^{n-1} = 1, g^2\, g^{n-2} = 1, \ldots, g^{n-1}g = 1$. Thus, for any non-identity element g, we have $g^k\, g^{n-k} = 1, for\ k \in$

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

Z^+ and $k < n$. From the definition of identity graph part (iii), for any elements $a, b \in G, a \neq b, a \neq e, b \neq e$, there exists an edge which connects a to b if and only if $ab = ba = e$. Therefore, there exists an edge which connects $g^k to\ g^{n-k}$, and since for every $a \in G, a \neq e$ there exists an edge connecting a to e, then from the definition of identity graph part (ii), there exists an edge from g^k to 1. These will form a triangle connecting 1, g^k and g^{n-k}. Moreover, since n is an odd number, then $n = 2x + 1$ for any $x \in Z^+$, and the number of non-identity elements in G is even. So, those non-identity elements can be partitioned into two sets with same cardinality, where the inverse of an element in one set is in the other set and vice versa. Therefore, there is no single edge in identity graph of G.

Then the identity graph which corresponds with G is shown in Figure 2.

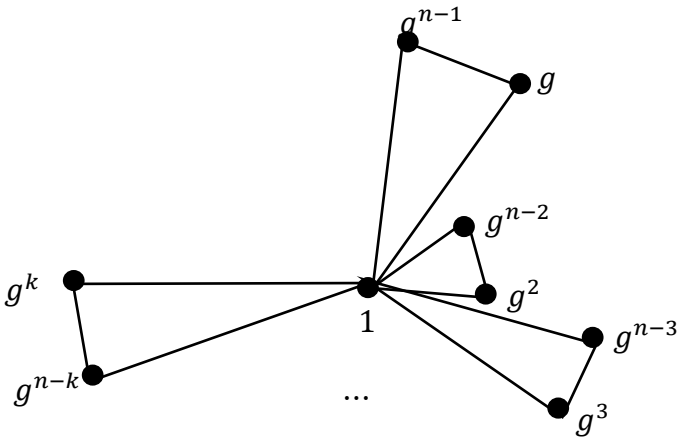


Figure 2. Identity graph of group $G = \{g \mid g^n = 1\}$, n is an odd number.

Based on previous theorems, we can conclude a characterization of the identity graph of cyclic group of odd order in the following Theorem 3.

Theorem 3 [2]

If $G = \langle g | g^n = 1 \rangle$ is a cyclic group of order n, where n is an odd number, then the identity graph G_i of G is formed by $(n-1)/2$ triangles.

Proof:

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

This can be proved using Theorem 1 and Figure 1. Since the number of non-identity elements is even and those elements forms (n-1)/2 couples, where each couple is inverse to each other, then there will be (n-1)/2 triangles.

Similar with the result in the odd order case, we can obtain a characterization of the identity graph of cyclic group of even order case in the following Theorem 4. The proof is also using similar principle with the odd order case.

Theorem 4 [2]

If $G = \langle g | g^m = 1 \rangle$ is a cyclic group of order m where m is an even number, then its identity graph G_i has $(m-2)/2$ triangle and a single edge.

Proof:

Table 2. Cayley table for Theorem 4.

	1	g	g^2	g^3	$\cdots$	$g^{\frac{m}{2}}$	$\cdots$	g^{m-3}	g^{m-2}	g^{m-1}
1	1	g	g^2	g^3	$\cdots$	$g^{\frac{m}{2}}$	$\cdots$	g^{m-3}	g^{m-2}	g^{m-1}
g	g	g^2	g^3	g^4	$\cdots$	$g^{\frac{m}{2}+1}$	$\cdots$	g^{m-2}	g^{m-1}	1
g^2	g^2	g^3	g^4	g^5	$\cdots$	$g^{\frac{m}{2}+2}$	$\cdots$	g^{m-1}	1	g
g^3	g^3	g^4	g^5	g^6	$\cdots$	$g^{\frac{m}{2}+3}$	$\cdots$	1	g	g^2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
$g^{\frac{m}{2}}$	$g^{\frac{m}{2}}$	$g^{\frac{m}{2}+1}$	$g^{\frac{m}{2}+2}$	$g^{\frac{m}{2}+3}$	$\cdots$	1	$\cdots$	$g^{\frac{3m}{2}-3}$	$g^{\frac{3m}{2}-2}$	$g^{\frac{3m}{2}-1}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
g^{m-3}	g^{m-3}	g^{m-2}	g^{m-1}	1	$\cdots$	$g^{\frac{3m}{2}-3}$	$\cdots$	g^{m-6}	g^{m-5}	g^{m-4}
g^{m-2}	g^{m-2}	g^{m-1}	1	g	$\cdots$	$g^{\frac{3m}{2}-2}$	$\cdots$	g^{m-5}	g^{m-4}	g^{m-3}
g^{m-1}	g^{m-1}	1	g	g^2	$\cdots$	$g^{\frac{3m}{2}-1}$	$\cdots$	g^{m-4}	g^{m-3}	g^{m-2}

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

Let $G = \langle g \,|g^m = 1\rangle$, m is an even number, be a cyclic multiplicative group and its order is m. Elements of G are $\{1, g, g^2, g^3, g^4, \cdots, g^{m-1}\}$. We will use the Cayley table given by Table 2 to show the operations between elements in G.

From the Cayley table above (Table 2) we can see that $g\,g^{m-1} = 1, g^2\,g^{m-2} = 1, \ldots, g^{m-1}\,g = 1$. However, for $g^{\frac{m}{2}}$ we have something different, that is $g^{\frac{m}{2}}\,g^{\frac{m}{2}} = 1$. Thus, for any non-identity elements other than $g^{\frac{m}{2}}$ we have $g^k\,g^{m-k} = 1$, for $k \in Z^+$ and $k < m$. From the definition of identity graph part (iii), for any elements $a, b \in G, a \neq b, a \neq e, b \neq e$, vertices a and b are adjacent if and only if $ab = ba = e$. This means the vertex $g\text{^}k$ is adjacent with vertex g^{m-k} and since for each $a \in G, a \neq e$ vertex a is adjacent with e based on definition of identity graph part (ii), then there exists an edge from g^k to 1. These form a triangle connecting 1, g^k dan $g\text{^}(m-k)$. In other side, for $g^{\frac{m}{2}}$ there will be only one edge connecting it, that is the edge which connects $g^{\frac{m}{2}}$ with 1. Moreover, since m is even, then $m = 2x$ for $x \in Z^+$ and the number of elements which is not the identity and not $g^{\frac{m}{2}}$ in G is even. Therefore, there will be $(m-2)/2$ triangles and a single edge. The identity graph G_i which corresponds with G is given in Figure 3.

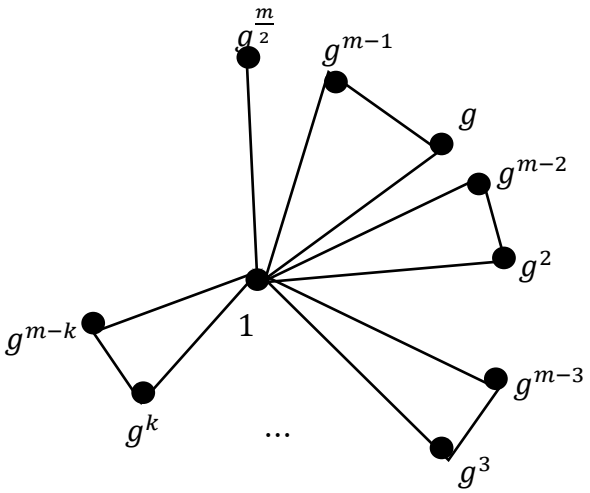


Figure 3. Identity graph of group $G = \langle g|\, g^m = 1\rangle$, n is an even number.

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

We see that in the even order case, there exists exactly one single edge that does not form a triangle. This happens because in a finite cyclic group of even order, there exists exactly one element of order 2.

4 Conclusion

From what we have discussed, there are some conclusions about identity graph of finite cyclic groups as following:

a. Identity graph of a group is a way to represent the relations between elements of a group in a graph. In identity graph of a group, the identity element of a group is connected with any other elements and each non-identity element is connected with its inverse.

b. For cyclic group of odd order n, its identity graph consists of $(n-1)/2$ triangles. There is no single edge in this case since in such group there is no element of order 2.

c. For cyclic group of even order m, its identity graph consists of $(m-2)/2$ triangles and a single edge. The single edge connects the identity element to the only element of such group that has order 2.

References

[1] A. Bretto, A. Faisant, L. Gilibert, G-graphs: A new representation of group, *Journal of Symbolic Computation*, **42**, 549-560, 2007.

[2] W. B. V. Kandasamy, F. Smarandache, *Groups as Graphs*. Slantina: Editura CuArt. 2009.

[3] S. Lovett, *Abstract Algebra*. Boca Raton: CRC Press. 2016.

[4] C. C. Miller, *Essentials of Modern Algebra*. Dulles, VA: Mercury Learning and Information, 2013.

[5] R. Rajkumar, P. Devi, *Coprime Graph of Subgroups of a Group*, https://www.semanticscholar.org

[6] M. U. Sherman-Bennett, *On Groups and Their Graphs*, MA: Bard College, 2016.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 101–110

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

Obtaining the Efficiency and Effectiveness of Fin in Unsteady State Conditions Using Explicit Finite Difference Method

Petrus Kanisius Purwadi[1,*], Budi Setyahandana[1], R.B.P. Harsilo[1]

[1]*Department of Mechanical Engineering, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*
[*]*Corresponding Author: pkpurwadi1966@gmail.com*

(Received 10-04-2021; Revised 10-05-2021; Accepted 17-05-2021)

Abstract

This paper discusses the search for fin efficiency and effectiveness in unsteady state conditions using numerical computation methods. The straight fin under review has a cross-sectional area that changes with the position x. The cross section of the fin is rectangular. The fins are composed of two different metal materials. The computation method used is the explicit finite difference methods. The properties of the fin material are assumed to be fixed, or do not change with changes in temperature. When the stability requirements are met, the use of the explicit finite difference methods yields satisfactory results. The use of the explicit finite difference methods can be developed for various other fin shapes, which are composed of two or more different materials, time-varying convection heat transfer coefficient, and the properties of the fin material that change with temperature.

Keywords: fin, efficiency, effectiveness, finite-difference, unsteady state

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

1 Introduction

In the design of fins, the important thing to know is the efficiency and effectiveness of the fins. There are many ways to know the value of fin efficiency and fin effectiveness. For certain fin shapes, the value of the efficiency and effectiveness of the fins can be found by using existing charts. For fins that do not yet exist, other ways are needed to get them. For cases in unstable state, the solution becomes more complicated. One way that can be done is by using numerical computation methods.

Several articles [1-7] related to the efficiency and fin effectiveness of fins in the unsteady state have helped in in solving this problem. In this case, the case discussed is a fin which is composed of two different materials and is in an unsteady state. The shape and cross-section of the fins are different from those that have been used as the object of the previous discussion. In this case, the cross-sectional area of the fin changes with the change in position x. The shape of the fin chosen is a truncated rectangular pyramid. Figure 1 presents an image of the fin shape to be discussed. With the total length of the fin L, where $L = L_1 + L_2$ and $L_1 = L_2$

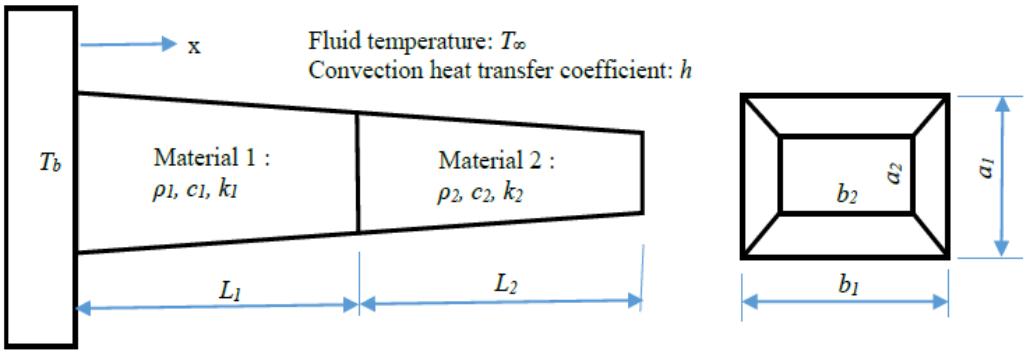


Figure 1. Straight fin with a truncated rectangular pyramid shape

The mathematical model for this problem is expressed by equation (1):

$$\frac{\partial^2 T(x,t)}{\partial x^2} + \frac{hP}{kA_p}(T(x,t) - T_\infty) = \frac{1}{\propto}\frac{\partial T(x,t)}{\partial t} \qquad 0 < x < L, x \neq L_1, t > 0 \qquad (1)$$

$$\alpha = \frac{k}{\rho c} \qquad (2)$$

with initial condition:

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

$$T(x,0) = T_i = T_b \qquad 0 \le x \le L, t = 0 \tag{3}$$

and boundary condition:

$$T(0,t) = T_b \qquad x = 0, t > 0 \tag{4}$$

$$k_2 \frac{\partial T(x,t)}{\partial x} = h \frac{As}{Ap_1}(T(x,t) - T_\infty) \qquad x = L, t > 0 \tag{5}$$

$$k_1 \frac{\partial T(x,t)}{\partial x} = k_2 \frac{Ap_2}{A_{p1}} \frac{\partial T(x,t)}{\partial x} + hA_s \frac{As}{Ap_1}(T_\infty - T(x,t)) \qquad x = L_1, t > 0 \tag{6}$$

In equations (1)-(6) we have used the notations:

x	:	Stated the position on fin, m
t	:	Stated the time, seconds
$T(x,t)$	:	Temperature at the x position, at the t time, ^oC
T_∞	:	Fluid temperature, ^oC.
T_i	:	Initial temperature of the fin, ^oC
T_b	:	Temperature at the base of the fin, ^oC
L_1	:	The length of the fin with material 1or the length of the fin with material 1, m
L_2	:	The length of the fin with material 1or the length of the fin with material 2, m
L	:	The length of the fin, $L = L_1+L_2$, m
A_p	:	Cross-sectional of the fin, m^2
P	:	Around the cross section, m
k	:	Conduction heat transfer coefficient of fin material 1 or material 2, W/(m.^oC)
k_1	:	Conduction heat transfer coefficient of fin material 1, W/(m.^oC)
k_2	:	Conduction heat transfer coefficient of fin material 2, W/(m.^oC)
h	:	Convection heat transfer coefficient, W/(m^2 ^oC)
ρ	:	Density of the material, kg/m^3
c	:	Specific heat of the material, kJ/(kg.^oC)
c_1	:	Specific heat of the material 1, J/(kg.^oC)
c_2	:	Specific heat of the material 2, J/(kg.^oC)
α	:	Thermal diffusivity of the material, m^2/s

1.1. Calculation Steps

The steps used to calculate the efficiency and effectiveness of the fin in an unsteady state condition using the numerical method are as follows:

1. Determine fixed variable variables, such as: $h, \rho_1, \rho_2, c_1, c_2, k_1, k_2, a_1, b_1$ $a_2, b_2, T_\infty, T_i, T_b, L_1, L_2, L, P, As_i, Ap_i, m, \Delta x, \Delta t$.

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

2. Perform temperature calculations for each control volume from time to time from volume control 1 to control up to volume control to m. Volume control 1 is on the bottom of the fin and volume control m is on the end of the fin. The total volume of the control is m.
3. Calculate the actual heat flow rate (q_{act}) released by the fins, the ideal heat flow rate released by the fins (q_{ideal}), and the heat flow rate released if there are no fins ($q_{no,fin}$). Temperature calculations are carried out from time to time.
4. Perform calculations of fin efficiency (η) from time to time
5. Perform calculations of fin effectiveness (ϵ) from time to time

Figure 2 presents a flow chart to get the efficiency and effectiveness of the fins.

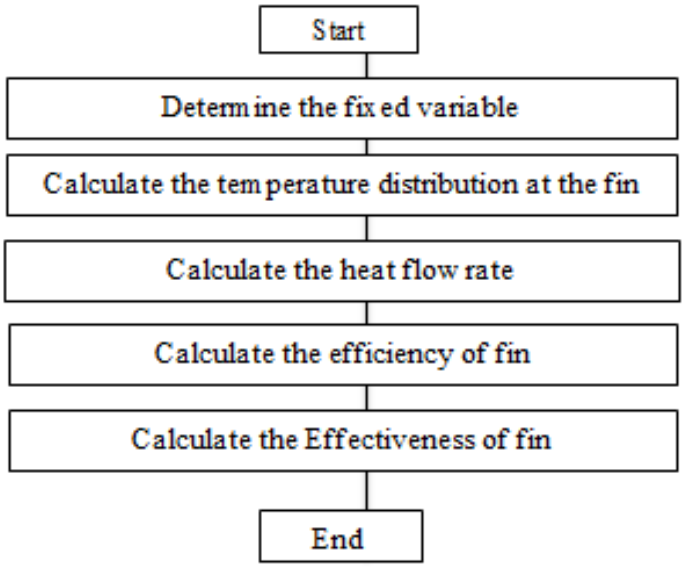


Figure 2. Flow chart to get the efficiency and effectiveness of the fins

Some of the assumptions used in this calculation are:

1. The fluid temperature and the convection heat transfer coefficient are uniform and constant.
2. Material properties (density, specific heat, thermal conductivity of the material) are uniform and constant.
3. The shape of the fins is fixed and does not change in volume during the process.

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

4. The transfer of heat by radiation is negligible.
5. The connection on the two fin materials is assumed to be perfectly connected.
6. No energy generated in the fins.
7. The heat transfer by conduction in the fins is assumed to take place in only one direction, the x direction.
8. The entire surface of the fins is in contact with the fluid around the fins.

1.2 Control volume and energy balance on the control volume

Figure 3 shows a fin image divided into many controls. The total control volume is m. Volume control 1 is at the base of the fin, and volume control m is at the end of the fin. The volume control p is located at the border of the two fin materials. In the control volume p, the half control volume is made of material 1, and the half control volume is made of material 2. The numbering of the control volume starts from the bottom of the fin to the end of the fin in sequence, with the names of the volume controls $1, 2, 3, \dots, m$. The distance between control volumes is Δx. At the control volume $i = 1, 2, 3, \dots, m - 1$, the control volume performs a convection heat transfer process through its blanket surface. At the control volume m, the process of convection heat transfer outside the blanket surface, also carries out the process of convection heat transfer through the cross-sectional area of the fins. The thickness of the volume control at the base of the fin and at the tip of the fin is $0{,}5\,\Delta x$. The control volume thickness of $i = 1, 2, 3, \dots, m - 1$ is Δx.

In Figure 4, Figure 5 and Figure 6, the convection heat flow is represented by q_c. The conduction heat flow from the control volume in position $i - 1$ to the control volume at i, is represented by q_{i-1}. The conduction heat flow from the control volume in position $i + 1$ to the control volume in position i, expressed as q_{i+1}.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

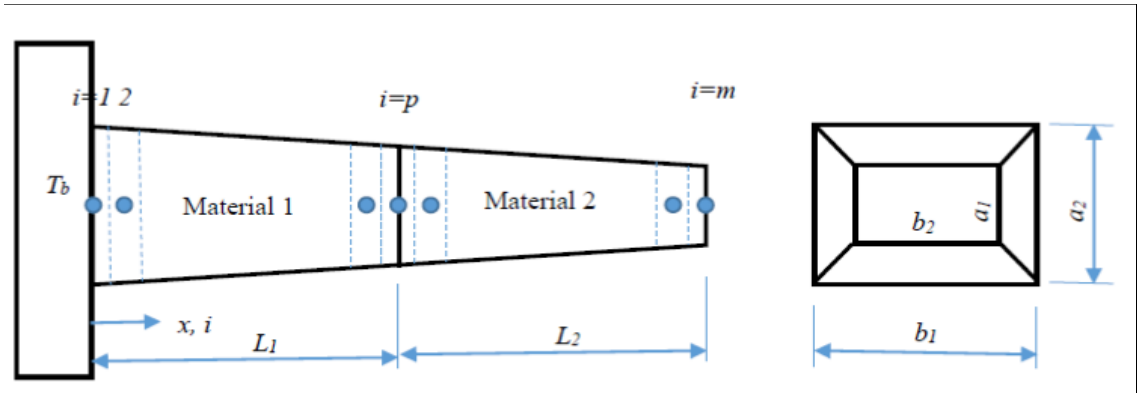


Figure 3. Distribution of volume control on the fins

The energy balance of the control volume on the fins, without energy generation, in the unsteady state condition can be stated by the following statement:

$$\begin{bmatrix}\text{all energy entering}\\ \text{the control volume during}\\ \text{the time interval } \Delta t\end{bmatrix} = \begin{bmatrix}\text{the energy change}\\ \text{in the control volume}\\ \text{during in the time interval } \Delta t\end{bmatrix}$$

It can also be expressed by equation (7):

$$\sum_{i=1}^{i=k} q_i = \rho c V \frac{\Delta T}{\Delta t} \tag{7}$$

Figure 4 presents the energy balance at the control volume inside the fin body, or at the control volume at position $i = 2, 3, 4, \dots, p-3, p-2, p-1$ (between the base of the fin with the border of the two fin materials) and in position $i = p+1, p+2, p+3, \dots, m-2, m-1$ (between the border of the two fin materials with the tip of the fin). The energy balance in this control volume can be expressed by equation (8):

$$q_{i-1} + q_{i+1} + q_c = \rho c V \frac{T_i^{n+1} - T_i^n}{\Delta t}. \tag{8}$$

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

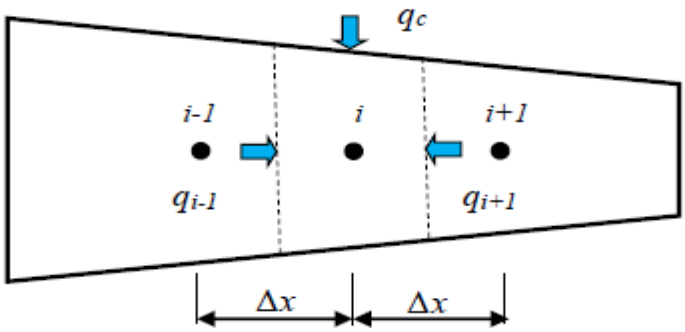


Figure 4. The energy balance at the control volume at $i = 2, 3, 4, \dots, p-2, p-1$ and at the control volume at $i = p+1, p+2, p+3, \dots, m-2, m-1$

In equations (8)-(10), T_i^n is the temperature in control volume i at time $t = n$ or at time t, and T_i^{n+1} is at control volume i at time $t = n+1$ or when $t = t + \Delta t$. Assume that the control volume has a uniform temperature. The variable V represents the volume, Δt and represents the time interval.

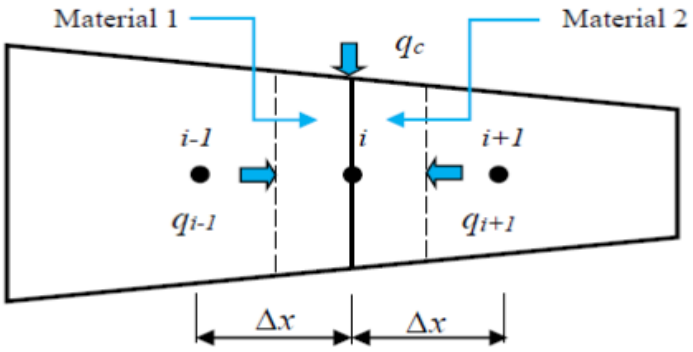


Figure 5. Energy balance on volume control at $i = p$

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

Figure 5 presents the energy balance at the control volume at the boundary of the two materials or at $i = p$. At the control volume at $i = p$, the control volume is composed of two different materials. The thickness of the control volume for material 1 is $0.5\,\Delta x$, the thickness of the control volume for material 2 is $0.5\,\Delta x$. The energy balance at the control volume in position $i = p$ can be expressed by equation (9):

$$q_{i-1} + q_{i+1} + q_c = \rho_1 c_1 V_i . \frac{T_i^{n+1} - T_i^n}{\Delta t} + \rho_2 c_2 V_2 \frac{T_i^{n+1} - T_i^n}{\Delta t}. \quad (9)$$

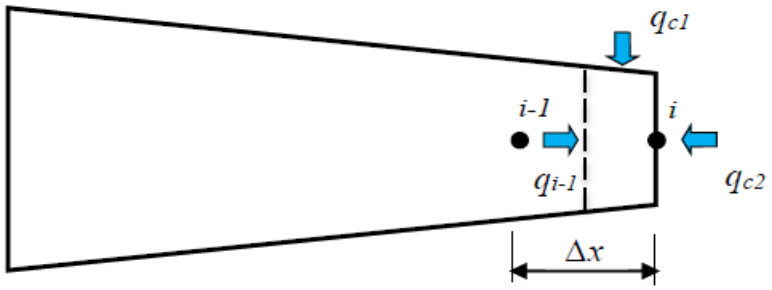


Figure 6. Energy balance on the control volume di $i = m$

Figure 6 presents the energy balance at the control volume at the tip of the fin or at $i = m$. In the control volume at the fin tip, the process of convection heat transfer through the blanket surface and the cross-sectional surface of the fin tip. The thickness of the volume control at the tip of the fin is $0.5\,\Delta x$. The energy balance at the control volume at position $i = m$ can be expressed by equation (10):

$$q_{i-1} + q_{c1} + q_{c2} = \rho_2 c_2 V_2 \frac{T_i^{n+1} - T_i^n}{\Delta t} \quad (10)$$

1.3 Temperature Distribution at the Fin

By using equation (7), the equation used to calculate the temperature at the control volume $i = 2, 3, 4, \ldots, m-2, m-1, m$, when $t > 0$ can be found. The equation for calculating the temperature at the control volume at $i = 2, 3, 4, \ldots, p-2, p-1$, when $t > 0$ is expressed by equation (11):

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

$$T_i^{n+1} = \frac{\Delta t}{\propto_1 \Delta x^2}[(T_{i-1}^n - 2T_i^n + T_{i+1}^n) + Bi_1(T_\infty - T_i^n)] + T_i^n \,. \quad (11)$$

The stability requirements for equation (11) are expressed by equation (12):

$$\Delta t \leq \frac{\rho_1 c_1 \Delta x^2}{k_1\left(2 + Bi_1 \frac{\Delta x As_i}{A_p}\right)}. \quad (12)$$

The equation for calculating the temperature at the control volume at $i = p$ when $t > 0$, can be expressed by equation (13):

$$T_i^{n+1} = \frac{\Delta t}{(\rho_1 c_1 V_1 + \rho_2 c_2 V_2)}\left[\left(k_1 A_p \frac{T_{i-1}^n - T_i^n}{\Delta x} + k_2 A_p \frac{T_{i+1}^n - T_i^n}{\Delta x}\right) + hA_s(T_\infty - T_i^n)\right] + T_i^n \,. \quad (13)$$

The stability requirements for equation (13) are expressed by equation (14):

$$\Delta t \leq \frac{(\rho_1 c_1 V_1 + \rho_2 c_2 V_2)}{\left(\frac{k_1 A_p}{\Delta x} + \frac{k_2 A_p}{\Delta x} + hAs\right)}. \quad (14)$$

The equation for calculating the temperature at the control volume at $i = p$ when $t > 0$, can be expressed by equation (15):

$$T_i^{n+1} = \frac{\Delta t}{\propto_2 \Delta x^2}[(T_{i-1}^n - 2T_i^n + T_{i+1}^n) + Bi_2(T_\infty - T_i^n)] + T_i^n \,. \quad (15)$$

The stability requirements for Equation (15) are expressed by equation (16):

$$\Delta t \leq \frac{\rho_2 c_2 \Delta x^2}{k_2\left(2 + Bi_2 \frac{\Delta x As_i}{A_p}\right)}. \quad (16)$$

The equation for calculating the temperature at the control volume at $i = m$ when $t > 0$, can be expressed by equation (15):

$$T_i^{n+1} = \frac{\Delta t}{0.5 \propto_2 \Delta x^2}\left[(T_{i-1}^n - T_i^n) + Bi_2(T_\infty - T_i^n) + \left(Bi_2 \frac{As}{A_p}(T_\infty - T_i^n)\right)\right] + T_i^n \,. \quad (17)$$

The stability requirements for Equation (17) are expressed by equation (18):

$$\Delta t \leq \frac{0.5\, \rho_2 c_2 \Delta x^2}{k_2\left(1 + Bi_2 + Bi_2 \frac{As}{A_p}\right)}. \quad (18)$$

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

1.4. Heat flow rate, efficiency and effectiveness

The amount of actual heat released by the fins in the unsteady state condition can be calculated by equation (19):

$$q_{\text{act}}^{n} = \sum_{i=1}^{m} hAs_i(T_i^n - T_\infty) \tag{19}$$

The amount of ideal heat released by the fins can be calculated by equation (20):

$$q_{\text{ideal}} = \sum_{i=1}^{m} hAs_i(T_b - T_\infty). \tag{20}$$

The amount of heat released by the bottom of the fin, if the length of the fin is considered zero, can be calculated by equation (21):

$$q_{\text{no fin}} = \sum_{i=1}^{m} hA_b(T_b - T_\infty). \tag{21}$$

The efficiency of the fin at an unsteady state condition can be calculated by equation (22):

$$\eta^n = \frac{q_{\text{act}}^{n}}{q_{\text{ideal}}} = \frac{\sum_{i=1}^{m} hAs_i(T_i^n - T_\infty)}{\sum_{i=1}^{m} hAs_i(T_b - T_\infty)}. \tag{22}$$

The effectiveness of the fin at an unsteady state condition can be calculated by equation (23):

$$\in^n = \frac{q_{\text{act}}^{n}}{q_{\text{no fin}}} = \frac{\sum_{i=1}^{m} hAs_i(T_i^n - T_\infty)}{\sum_{i=1}^{m} hA_b(T_b - T_\infty)}. \tag{23}$$

2 Research Methodology

The search for efficiency and effectiveness is carried out using numerical methods. The numerical method used is the explicit finite difference methods. The selected fin shape is a truncated rectangular pyramid (Figure 1). Fins are composed of two different materials. The properties of the fin material are presented in Table 1. The total fin length is L, where $L = 10\ cm$. The length of the fin with material 1 is the same as length p with material 2, where $L_1 = L_2 = 5\ cm$. The section of the fin at the base of the fin, has a_1 width and b_1 height, where $a_1 = 1\ cm$ and $b_1 = 0.5\ cm$. The section of the fin at the end of the fin, has a_2 width and b_2 height, where $a_2 = 0.5\ cm$ and $b_2 = 0.25\ cm$. The sum of the control volume on the fin is m, where $m = 25$. The distance

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

between the control volume is Δx, and $\Delta x = 0.004167\ cm$. The time interval is Δt, where $\Delta t = 0.05$ second. The temperature of the fluid around the fin is T_∞, where $T_\infty = 30℃$ The base temperature of the fin remains equal to T_b, with $T_b = 100℃$. The initial temperature of the fin is T_i, where $T_i = T_b = 100℃$. The convection heat transfer coefficient is h, where $h = 100 W/m^2.℃$.

Table 1. Properties of the material
(Y.A. Cengel, Heat Transfer A Practical Approach, pp 868-870, see [8])

Material	Density (ρ) (kg/m^3)	Thermal Conductivity (k) ($W/m.^oC$)	Specific Heat (c) ($J/kg.^oC$)
Copper (Cu)	8933	401	385
Aluminum (Al)	2702	237	903
Zinc (Zn)	7140	116	389
Nickel (Ni)	8900	90.7	444
Iron (Fe)	7870	80.2	447

3 Results and Discussion

The results of the calculation of temperature distribution, actual heat flow rate, efficiency and effectiveness of fins are presented in Figure 7, Figure 8, Figure 9 and Figure 10. The solution using explicit finite difference methods gives satisfactory results. As long as the stability requirements are met, the results are realistic. The truncated rectangular pyramid shape is an example of the shape of the fin. Thus, in the same way, the use of the explicit finite difference methods can be developed for the calculation of efficiency and effectiveness with other fin shapes.

Figure 7 presents the temperature distribution that occurs in the fins with various variations in the composition of the material. At the boundary of the two materials, the temperature distribution looks broken. The fins with an iron-copper composition were more likely to be broken than those with an iron-nickel composition. The results of the temperature distribution in the unsteady state are influenced by the properties of the material such as mass density, thermal conductivity and specific heat. Figure 8, Figure 9 and Figure 10 present the actual heat flow rate of fin removal, fin efficiency and fin effectiveness. The results of this calculation, it all depends on the temperature distribution that occurs in the fin.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

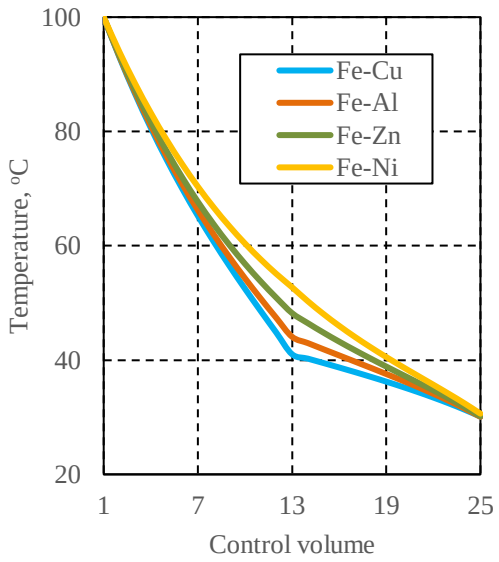


Figure 7. The temperature of the fin at time t= 100 seconds

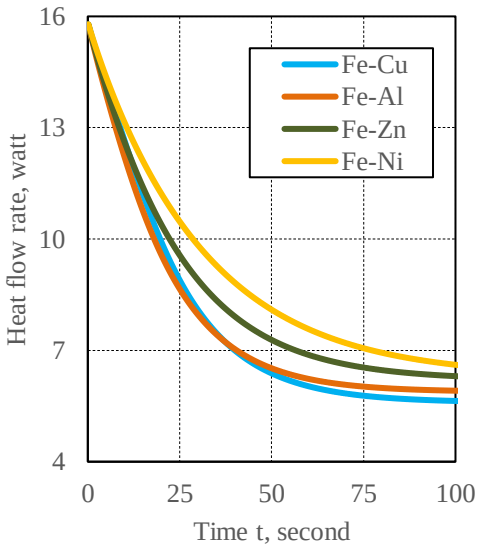


Figure 8. The heat flow rate of the fin

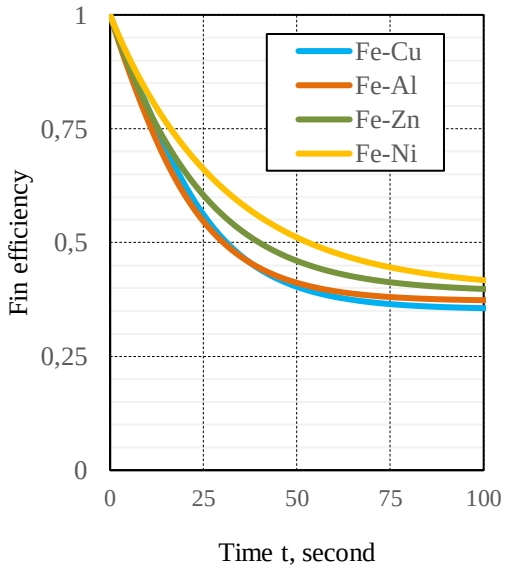


Figure 9. Fin efficiency

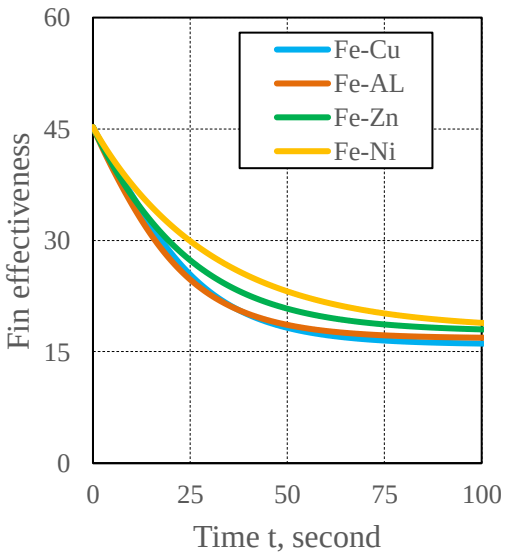


Figure 10. Fin Effectiveness

The use of the explicit finite different methods can also be developed for the case of conduction heat flow which is not only for the one-dimensional case, but also the two-dimensional unsteady state case. The two-dimensional case is when the conduction heat

flow in the object takes place in two directions: the x direction and the y direction. The two-dimensional case occurs for fins with a greater length and width than the thickness of the fin. It can also be developed for the three-dimensional case, with three-directions of conduction heat flow: x, y, and z directions.

The fin object discussed in this issue is for fins which are composed of two different materials. If the fins are composed of three or more different fin materials, the use of the explicit finite difference methods can also be used. In this case, the fin material used does not change with changes in temperature. The use the explicit finite difference methods can also be used for fins with temperature changing material properties. Material properties that change due to temperature, such as density, specific heat and thermal conductivity of the fin material.

4 Conclusion

As long as the stability requirements are met explicit finite difference methods can be used well to calculate the efficiency and effectiveness of fins in the unsteady state condition. The use of the explicit finite difference methods can be developed for various other fin shapes, which are composed of two or more different materials, the time-changing convection heat transfer coefficient value, and the temperature-changing properties of the fin material.

Acknowledgements

This research was conducted at Sanata Dharma University. The authors thank Sanata Dharma University for supporting this research.

References

[1] A.W. Vidjabhakti, P.K. Purwadi, S. Mungkasi, “Efficiency and effectiveness of a fin having pentagonal cross section dependent on the one dimensional position”, *Proceedings of the 1st International Conference on Science and Technology for an Internet of Things*, Yogyakarta, Indonesia, 20 October 2018, doi:10.4108/eai.19-10-2018.2282540.

Volume 3, Issue 1, pages 111–124

p-ISSN 2655-8564, e-ISSN 2685-9432

[2] K.S. Ginting, P.K, Purwadi, S. Mungkasi, "Efficiency and effectiveness of a fin having capsule shaped cross section dependent on the one dimensional position", *Proceedings of the 1st International Conference on Science and Technology for an Internet of Things*, Yogyakarta, Indonesia, 20 October 2018, doi:10.4108/eai.19-10-2018.2282543.

[3] P.K. Purwadi and Bramantyo Yudha Pratama, "Efficiency and effectiveness of a truncated cone-shaped fin consisting of two different materials in the steady-state", *AIP Conference Proceedings 2202*, **020091** (2019), doi:10.1063/1.5141704.

[4] P.K. Purwadi and Michael Seen, "Efficiency and effectiveness of a fin having the capsule-shaped cross section in the unsteady state", Cite as: *AIP Conference Proceedings* **2202**, 020092 (2019), doi:10.1063/1.5141705.

[5] P.K. Purwadi and Michael Seen, "The efficiency and effectiveness of fins made from two different materials in unsteady-state", *Journal of Physics: Conference Series*, *Volume 1511*, *International Conference on Science Education and Technology (ICOSETH) 2019*, 23 November 2019, Surakarta, Indonesia, doi:10.1088/1742-6596/1511/1/012082.

[6] P.K. Purwadi, Yunus Angga Vantosa, Sudi Mungkasi, "Efficiency and effectiveness of a rotation-shaped fin having the cross-section area dependent on the one dimensional position", *Proceedings of the 1st International Conference on Science and Technology for an Internet of Things*, Yogyakarta, Indonesia, 20 October 2018, doi:10.4108/eai.20-9-2019.2292097.

[7] T.D. Nugroho and P.K. Purwadi, "Fins effectiveness and efficiency with position function of rhombus sectional area in unsteady condition", *AIP Conference Proceedings* **1788**, 030034 (2017), doi:10.1063/1.4968287.

[8] Y.A. Çengel, *Heat Transfer: A Practical Approach* (University of Nevada, Reno: The McGraw-Hill Companies, Inc., United States of America, 2008), pp. 163-164.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 125–130

p-ISSN 2655-8564, e-ISSN 2685-9432

Conceptual Design of Modular Chassis Jig of Student Competition Car

Heryoga Winarbawa[1,*]

[1]*Department of Mechanical Engineering, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*
[*]*Corresponding Author: winarbawa@gmail.com*

(Received 28-05-2021; Revised 16-06-2021; Accepted 16-06-2021)

Abstract

Chassis jig is needed to ensure that the welded chassis components does not warp or deform during welding process. Through concept screening and concept scoring, multiple design of chassis jigs is narrowed down to next development process. This study aims to design a chassis jig for the fabrication of student car competition chassis. The desired result of this design process is chassis jig with the ability to manufacture a wide variety of student competition car chassis.

Keywords: Car, chassis, jig, fixture

1 Introduction

Students of mechanical engineering programs are often challenged to study implementations. Student competition is one of study implementations. Those competitions drive students to build something based on what they learn in class. Formula Student (FS), Shell Eco Marathon (SEM), Kompetisi Mobil Hemat Energi (KMHE), and Kompetisi Mobil Listrik Indonesia (KMLI) are examples where student need to build cars to compete.

Volume 3, Issue 1, pages 125–130
p-ISSN 2655-8564, e-ISSN 2685-9432

Car chassis is the important thing to ensure vehicle stability and driver/passenger safety. There are several types of chassis, ladder frame, space frame, and monocoque. A lot of student teams apply space frame chassis for their cars. Space frame chassis is chosen due to simpler manufacturability and cost effectiveness.

Space frame chassis are manufactured by welding some of tubing. To ensure accuracy, space frame chassis materials are welded on chassis jig and several fixtures. Welding fixtures are typically the most common devices used to align and retain the various pieces for welding [1]. Some chassis jig on the market are using steel as their table top. Steel table top are usually coated with other material to prevent rust.

The scope of this study was limited to accommodate competition car chassis welding process. This study summarizes the design of chassis jig that accommodate several kind of chassis that suite for each competition.

2 Research Methodology

This study applied three stages of process as shown at Figure 1. The first stage is to collect all requirement measurement, such as, length, width, and height of car chassis based on technical rules for each competition. As the second stage, we recall that concept selection is the process of evaluating concepts with respect to customer needs and other criteria, comparing the relative strengths and weaknesses of the concepts, and selecting one or more concepts for further investigation, testing, or development [2]. Material selection criteria are limited to ease of manufacturing and assembly process, and rust resistance properties. The third stage is when chassis jig was design and some feature were added.

Figure 1. Study process

Volume 3, Issue 1, pages 125–130

p-ISSN 2655-8564, e-ISSN 2685-9432

Collect Data. Car dimension can be found at technical data of each competition [3-5]. The KMHE competition adopts SEM environment, so basically the technical rules for car dimension are the same. Maximum dimension requirement are selected so the chassis that need to be manufactured doesn't break the competition rules. The closest car chassis CAD design for each type of competition are also collected to simulate availability on chassis jig top.

Concept Selection. A chassis jig is constructed by table top and some legs. Steel tube, hollow aluminum tube, and aluminum extrusion are the most common materials to build such thing. Concept A is bulky, heavy duty chassis jig constructed by square hollow steel as frame and thick steel sheet as table top. Concept B is knock-down slatted steel sheet as top, frame and some legs. Concept C is constructed by all aluminum extrusion. Concept D has steel sheet with welded ribs as table top, and bent steel sheet that constructed as frame. Concept E is assembled by welded square hollow steel and aluminum extrusion as table top. All of those concepts are narrowed down through concept screening which tabulated in Table 1.

Table 1. Concept screening

	Concepts				
Selection Criteria	A	B	C	D	E
Weight	-	-	0	+	-
Rust Resistant	-	0	0	0	0
Easy to Assemble	-	0	0	+	+
Manufacturing Cost	-	-	0	+	+
Fixture Modularity	+	+	0	0	0
Rigidity	+	+	0	-	+
Easy to Move	-	+	0	+	-
Sum +'s	2	3	0	4	3
Sum 0's	0	2	7	2	2
Sum –'s	5	2	0	1	2
Net Score	-3	1	0	3	1
Rank	5	3	4	1	2
Continue?	No	Yes	Combine	Yes	Combine

Through concept screening, concept variation has been reduced to three concepts, Concept B, Concept D, and combination of Concept C and E, thus called Concept CE. Further investigated issue needs to be clarified before chosen final concept. Selection criteria are weighted to increase the sensitivity of concept determination as shown in Table 2.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 125–130

p-ISSN 2655-8564, e-ISSN 2685-9432

Table 2. Concept scoring

		Concepts					
		B		CE		D	
	Weight	Rating	Weighted Score	Rating	Weighted Score	Rating	Weighted Score
Selection Criteria							
Weight	10%	2	0.2	4	0.4	5	0.5
Rust Resistant	15%	3	0.45	4	0.6	3	0.45
Easy to Assemble	10%	3	0.3	2	0.2	3	0.3
Manufacturing Cost	25%	1	0.25	3	0.75	4	1
Fixture Modularity	15%	4	0.6	4	0.6	3	0.45
Rigidity	20%	4	0.8	4	0.8	1	0.2
Easy to Move	5%	4	0.2	4	0.2	4	0.2
Total Score		2.8		3.55		3.1	
Rank		3		1		2	
Continue		No		Develop		No	

The total score for each concept is the sum of the weighted scores, and formulated as:

$$S_j = \sum_{i=1}^{n} r_{ij}\, w_i. \quad (1)$$

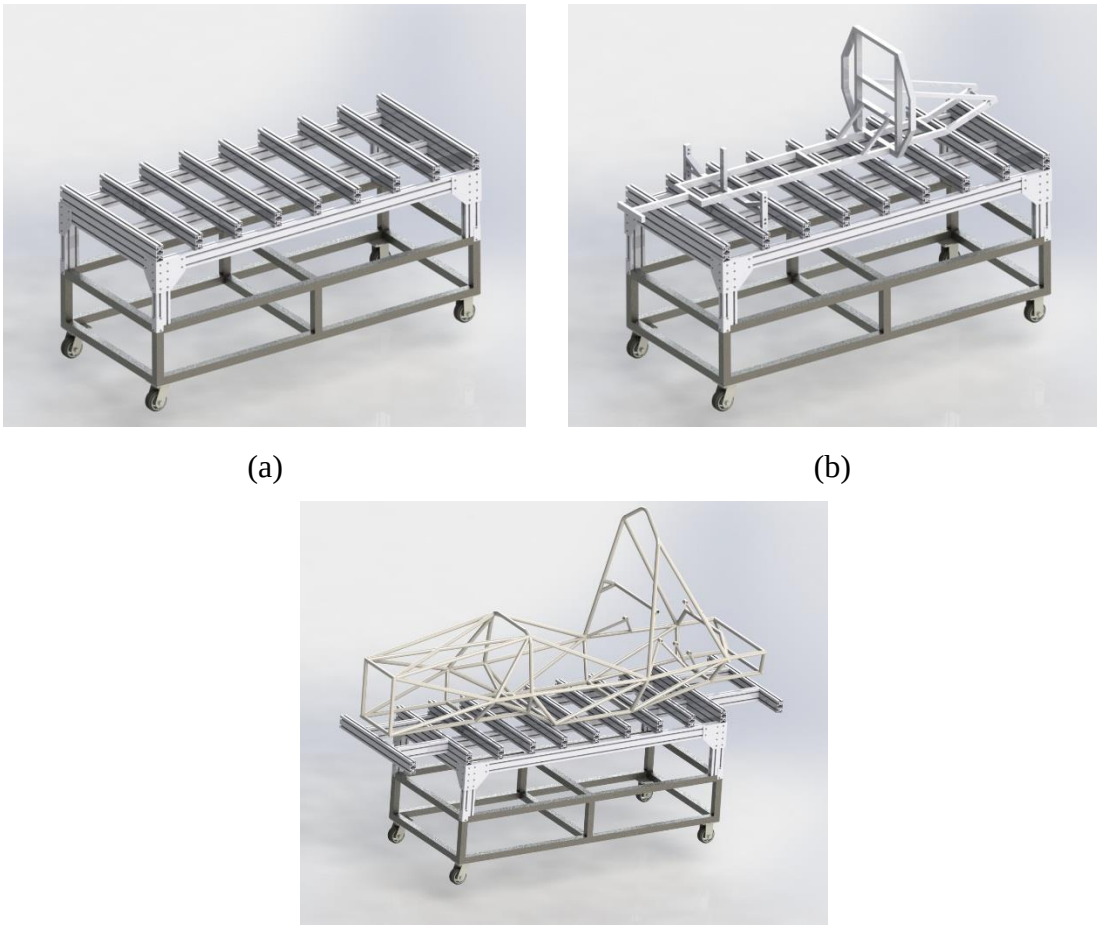
Here, r_{ij} is the raw rating of concept j for the ith criterion, w_i represents the weighting for ith criterion, n is the number of criteria, and S_j denotes the total score of concept j.

3 Results and Discussion

Chosen concept is combination of concept C and E. Concept CE is constructed by welded square hollow steel as bottom part of frame and slatted arrangement of aluminum extrusion as table top as shown in Figure 2a. From bottom part of frame to table top is raised by another aluminum extrusion. That configuration provide better stability as center gravity is lower. The lower part also useful for storing jigs/fixtures, welding equipment, and power tools that required during chassis fabrication. Figure 2b shows chassis jig with SEM prototype car chassis on top of it. For FS car chassis, some additional jig members were added to support front and back section of car chassis as shown in Figure 2c. Fixtures for holding chassis materials will be design as future work.

Volume 3, Issue 1, pages 125–130

p-ISSN 2655-8564, e-ISSN 2685-9432



(a) (b)

(c)

Figure 2. Some illustrations of chassis jigs: (a) Chassis jig, (b) Chassis jig with SEM prototype car chassis, (c) Chassis jig with FS car chassis

4 Conclusion

Concept CE has been concluded as best concept after several steps. The excellence of Concept CE amongst others are:

- **Lightweight**: using aluminum extrusion as the most component are reducing overall weight
- **Rust resistant**: aluminum is already rust resistant since it already provides oxide film all over the surface.
- **Easy to assemble**: aluminum extrusion provides tracks that has standardized dimensions and their own accessories such as nuts, bolts and connectors.

Volume 3, Issue 1, pages 125–130

p-ISSN 2655-8564, e-ISSN 2685-9432

- **Low manufacturing cost**: since it already has tracks and accessories, fabricating process and assembly is just simply as cut and connect. It does not need complex manufacturing process on many machines.
- **Modular fixture**: lower fixture design are mainly based on aluminum extrusion track design, and the top one are following work piece design, that is, aluminum hollow tube for chassis.
- **Rigid enough**: aluminum extrusion cross-section is already designed to maintain rigidity along its length. Bigger cross-sectional area provides minimum deflection and maximum rigidity when they are made into a structure.
- **Easy to move**: since it is already lightweight because half part is using aluminum material, the movement also help by castor wheel.

Note that because this concept using aluminum as top, so it does not provide impact protection. Top surface is only used as planar reference and placing fixtures.

References

[1] E. G. Hoffman, *Jig and Fixture*, 5th Edition, Clifton Park, NY: Delmar, 2004.

[2] K.T. Urlich and S.D Epingger, *Product Design and Development*, New York, NY: McGraw-Hill Education, 2016.

[3] "Formula Student," [Online]. Available: https://www.imeche.org/docs/default-source/1-oscar/formula-student/2021/forms/fs2020-rules.zip?sfvrsn=2. [Accessed 24 March 2021].

[4] "Shell Eco Marathon," [Online]. Available: https://base.makethefuture.shell/en_gb/service/api/home/shell-eco-marathon/global-rules/_jcr_content/root/content/document_listing/items/download_1733386233.stream/1598973488368/0e057b48fe5e2adc044ac860f83ebf26f8ccd5f9/shell-eco-marathon-2020-official-rul. [Accessed 24 March 2021].

[5] "Kompetisi Mobil Listrik Indonesia," [Online]. Available: http://kmli.polban.ac.id/component/phocadownload/category/1-panduan-lomba.html?download=6:panduan-kmli-xi-2019. [Accessed 24 March 2021].

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

Classification of Toddler Nutrition Using C4.5 Decision Tree Method

Kartono Pinaryanto[1,*], Robertus Adi Nugroho[1],

Yanuarius Basilius[1]

[1]*Department of Informatics, Faculty of Science and Technology, Sanata Dharma University, Yogyakarta, Indonesia*

[*]*Corresponding Author: kartono@usd.ac.id*

(Received 07-05-2021; Revised 15-06-2021; Accepted 16-06-2021)

Abstract

Nutrition is very much needed in the growth of toddlers. It is very important to give babies a balanced nutritional intake at the right stage so that the baby grows healthy and is accustomed to a healthy lifestyle in the future. Children under five years of age are a group that is vulnerable to health and nutrition problems. In determining the nutritional status, it can be done in a system manner using the C4.5 decision tree classification method and entering several variables or attributes. The dataset tested was 853 toddlers. Classification is carried out to determine the nutritional status based on the weight/age (BB/U), height/age (TB/U) and weight/height (BB/TB) categories. The attributes used for the classification of BB/U are gender, weight and age. The attributes used for TB/U are gender, body length or height, and age. The attributes used for BB/TB are gender, weight, body length or height, and age. The average accuracy of the BB/U category is 90.16%, the average accuracy of the TB/U category is 76.64%, and the average accuracy of the BB/TB category is 83.83%.

Keywords: Classification, decision tree, C4.5, nutrition for toddlers

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

1 Introduction

Nutrients are organic substances required for normal functioning of the body's systems, growth and health maintenance. It is very important to give babies a balanced nutritional intake at the right stage so that the baby grows healthy and is accustomed to a healthy lifestyle in the future. Children under five years of age are a group that is vulnerable to health and nutrition problems, so that the toddler years are an important period of growth and need serious attention [1]. Based on the results of the 2018 Ministry of Health's Basic Health Research, 17.7% of infants under 5 years of age (toddlers) still experience nutritional problems. This figure consisted of Under-fives who suffered from malnutrition by 3.9% and those suffering from malnutrition by 13.8% [2]. The nutritional status of toddlers can be measured anthropometry, anthropometric indices are often used, namely: body weight for age (BB/U), height for age (TB/U), body weight for height (BB/TB). The weight index based on age (BB/U) is the most commonly used indicator because it has the advantage of being easy and quicker to understand by the general public. The reference standard used for determining nutritional status by anthropometry is based on the Decree of the Minister of Health No. 920/Menkes/SK/VIII/2002, to use the reference book of the "World Health Organization-National Center for Health Statistics" (WHO-NCHS) by looking at the Z-score.

In determining the nutritional status, it has been done manually by the Community health centers, so patients have to come physically to the Community health centers. This is of course very troublesome especially in the current pandemic situation and conditions. Determining nutritional status can be done automatically using a classification approach. One approach that can be taken is to use the C4.5 decision tree method. The C4.5 method is an algorithm that works by applying the concept of a decision tree. A decision tree is a predictive model using a tree structure or hierarchical structure. The concept of a decision tree is to transform data into a decision tree with decision rules.

In previous research on the comparison of the performance of the C4.5 and Naive Bayes algorithms for the classification of scholarship recipients by Choirul Anam and

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

Harry Budi Santoso, stated that the C4.5 algorithm has better performance than Naive Bayes with the level of accuracy obtained using the C4.5 algorithm of 96.4%, while the accuracy rate of Naïve Bayes is 95.11% [3].

Based on research [4] on the classification of typhoid fever (TF) and dengue hemorrhagic fever (DHF) by applying the C4.5 decision tree algorithm. It can be concluded that by using the k-folds cross validation test, the highest average accuracy value is 91.875% using 32 test data and 128 training data.

From the description above, a study was conducted using the C4.5 decision tree method in determining the nutritional status of children under five. It is hoped that applying the C4.5 decision tree method can help classify the nutritional status of toddlers to determine the growth of children under five.

2 Methodology

The methodology used in this study is as follows (Figure 1).

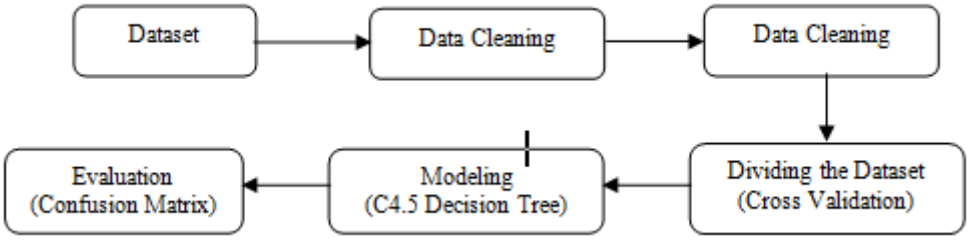


Figure 1. Decision tree classification research methodology C4.5

The research began to prepare the dataset, then the dataset went through the cleaning process and continued with data selection. The next stage, the data will be divided into testing data and training data. Training data will be used to form a decision tree, while testing data will be used to evaluate the system being created. In the next sub-section, it will be explained in detail about the stages that are passed.

2.1. Dataset

The dataset used in this study is the monitoring data on the nutritional status of toddlers, obtained from the Kebong Health Center, Kelam Permai District, Sintang

Volume 3, Issue 1, pages 131–142
p-ISSN 2655-8564, e-ISSN 2685-9432

District, West Borneo in 2017 with a total of 853 toddlers. Monitoring data on the nutritional status of toddlers has three categories, namely the category of body weight according to age (BB/U), height for age (TB/U), and body weight for height (BB/TB). The BB/U category has 4 classification labels namely Best, Good, Bad and Worst. The TB/U category has 4 classification labels namely High, Normal, Short, and Very Short. While the BB/TB category has 4 classification labels namely Fat, Normal, Thin, and Very Thin (Table 1).

Table 1. Categories and Labels

No	Category	Label
1	BB/U	Best, Good, Bad, Worst
2	TB/U	High, Normal, Short, Very Short
3	BB/TB	Fat, Normal, Thin, Very Thin

2.2. Data Cleaning

Data cleaning is a process for cleaning unused data [5]. In this study, some data were deleted because were incomplete. An example of deleted data is that it does not have a BB/TB label, has no PB/TB value, and does not have a TB/PB conversion value.

2.3. Data Selection

In the dataset, there are 19 attributes, including name, date of birth, gender M/F, body weight, PB/TB, measured position, age, age calculation process, conversion of TB/PB, age family, code, code1, code2, nutritional standards Poor BB/U, Nutritional Standards Good BB/U, Short PB/U or TB/U Standards, Normal PB/U or TB/U Standards, Weight Standards BB/TB or BB/TB, and Normal Standards of BB/TB or BB/TB. At the data selection stage, the attributes used for the classification were determined (feature selection). In the selection of attributes, the attributes of Gender M/F, Body Weight, PB/TB and Age were selected. These attributes were selected based on recommendations from the health center. The results of the attribute selection are shown in Table 2.

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

Table 2. Attributes used by each category

No	Category	Attribute	Label
1	BB/U	Gender M/F, Body Weight, Age	Best, Good, Bad, Worst
2	TB/U	Gender M/F, PB/TB, Age	High, Normal, Short, Very Short
3	BB/TB	Gender M/F, Body Weight, PB/TB, Age	Fat, Normal, Thin, Very Thin

2.4. Dividing the Dataset

The dataset is divided into testing data and training data using k-folds validation. The number of k is chosen by the user where the values of k are 3, 5, 7 and 9 folds. If the value of $k = 3$, then the data is divided into 3 parts, 2 parts used for training data and 1 part for testing data, and likewise for dividing the value of 5, 7 and 9 folds.

2.5. Modeling C4.5 Decision Tree

Every fold is modeled using the C4.5 decision tree method, so that there are n models for each n folds. The C4.5 decision tree method classifies the data by looking for the value of Entropy, Information Gain, Split Info and Gain Ratio. Tree formation begins with finding the highest Gain Ratio value to become the root node, then for leaf nodes it is carried out recursively until a decision tree is formed [6].

The following is an example of a tree formation step:

1. Prepare the data that will be used for the formation of the C4.5 decision tree model. In this example, 9 data on children under five are used for the classification of the BB / U category with the attributes used according to Table 3.
2. Separating data into training data such as Table 4 and testing data as in Table 5 with a total of 3 folds.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

Table 3. Dataset

Gender M/F	Body Weight	Age	BB/U
1	8	9	Good
1	7.8	8	Good
1	10.1	8	Good
2	6.1	6	Good
2	4.6	6	Worst
2	10	44	Worst
2	7.3	27	Worst
2	8.9	17	Worst
1	8.1	26	Worst

Table 4. Data Training

Gender M/F	Body Weight	Age	BB/U
1	8	9	Good
1	7.8	8	Good
1	10.1	8	Good
2	6.1	6	Good
2	4.6	6	Worst
2	10	44	Worst

Table 5. Data Testing

Gender M/F	Body Weight	Age	BB/U
2	7.3	27	Worst
2	8.9	17	Worst
1	8.1	26	Worst

3. Calculating entropy using formula (1), information gain using formula (2), split info using formula (3), and calculating the gain ratio using formula (4) for each attribute. The entropy is formulated as

$$\text{Entropy}(S) = \sum_{i=1}^{n} -p_i * \log_2 p_i. \quad (1)$$

Description of formula (1) follows: S is the set of cases, n is the number of partitions S and p_i is the proportion of S_i to S. The gain is formulated as

$$\text{Gain}\,(S, A) = \text{Entropy}(S) - \sum_{i=1}^{n} \frac{|S_i|}{|S|} * \text{Entropy}(S_i). \quad (2)$$

Description of formula (2) follows: S is Sample, A is attribute, n is the number of partitions of the attribute set A, $|S_i|$ is the number of samples on the partition, and $|S|$ is the number of samples in S. Now we formulate the Split Info as

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

$$\text{SplitInfo}(S, A) = -\sum_{i=1}^{v} \frac{|S_i|}{|S|} \times \log_2\left(\frac{|S_i|}{|S|}\right). \quad (3)$$

Description of formula (3) follows: v is the subset resulting from solving using attribute A which has as many as v values. Then, we have the Gain Ratio as

$$\text{GainRatio}(S, A) = \frac{\text{Gain}(S, A)}{\text{SplitInfo}(S, A)}. \quad (4)$$

Next, look for the root node candidates by looking for the highest information gain value for each attribute. Determine the root node by finding the highest gain ratio value for each candidate. The highest gain ratio value is found in the weight attribute with a variable value of 4.6, thus the root node of the tree is Weight B. with a value of 4.6. The decision tree formed from the calculation is shown in Figure 2.

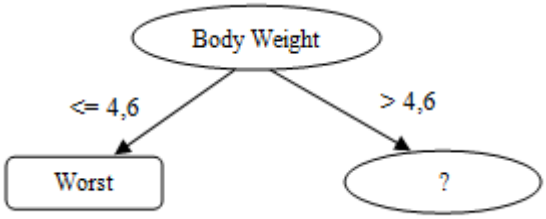


Figure 2. Root Node

4. After getting the root node, then we do a leaf node search. Data with a weight value of 4.6 are deleted / removed from the dataset before searching for leaf nodes (Table 6).

Table 6. The dataset table at node 2

Gender M/F	Body Weight	Age	BB/U
1	8	9	Good
1	7.8	8	Good
1	10.1	8	Good
2	6.1	6	Good
2	4.6	6	Worst
2	10	44	Worst

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

After it has been removed, it is followed by looking for leaf nodes, and searching for the highest information gain value. The highest information gain value is in the Age attribute with a value of 9, thus the leaf node is Age, if the age is below 9 then the classification label is Good and if it is above 9 then the classification label is Worst. The resulting tree is shown in Figure 3.

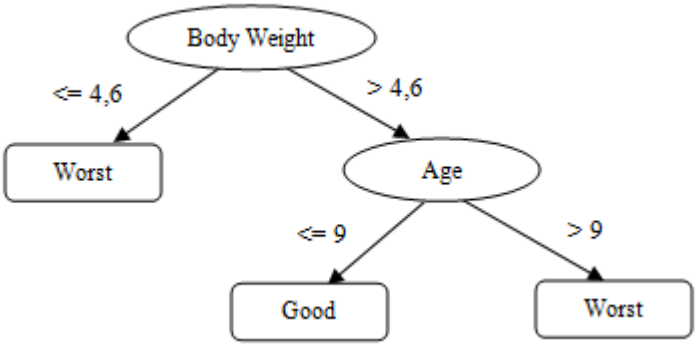


Figure 3. Leaf Node

2.6. Evaluation

Several experiments were carried out to evaluate this system. Each experiment was carried out by dividing the data into 3, 5, 7 and 9 folds. Each experiment was carried out for each category, namely the categories BB/U, TB/U and BB/TB. The experiments are shown in Table 7.

Table 7. C4.5 decision tree experiment

Experiment	Number of Folds
1st	*3-folds*
2nd	*5-folds*
3rd	*7-folds*
4th	*9-folds*

3 Results and Discussion

Based on the experiments, the system is able to classify the nutritional status of the toddler based on BB/U, TB/U and BB/TB. The test results for the BB/U category showed when the number of 3 folds the measured accuracy was 89.52%, when the

Volume 3, Issue 1, pages 131–142
p-ISSN 2655-8564, e-ISSN 2685-9432

number of 5 folds the measured accuracy was 90.93%, when the number of 7 folds the measured accuracy was 90.10% and when the number was 9 folds measured accuracy was 90.10%. These results indicate the average level of accuracy is 90.16%. Where the greatest accuracy occurs when using 5 folds (Table 8). This shows that the system can classify the BB/U category well.

Table 8. Results of the BB / U experiment

BB/U		
Experiment	**Number of Folds**	**Average accuracy (%)**
1	3	89.52
2	5	90.93
3	7	90.10
4	9	90.10

While the TB/U category trial showed the average accuracy rate was 76.64% and the highest accuracy occurred at folds 7 (Table 9).

Table 9. Results of the TB/U experiment

TB/U		
Experiment	**Number of Folds**	**Average accuracy (%)**
1	3	75.27
2	5	75.96
3	7	78.32
4	9	77.03

While the BB/TB category trial showed the average accuracy rate was 83.83% and the highest accuracy occurred at folds 7 (Table 10).

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

Table 10. Results of the BB/TB experiment

BB/TB		
Experiment	**Number of Folds**	**Average accuracy (%)**
1	3	83.27
2	5	83.27
3	7	84.45
4	9	84.34

Based on the test results, we observe that the C4.5 decision tree works well for classifying the categories of BB/U, TB/U and BB/TB using the selected attributes. Although a minority of cases cannot be classified properly.

4 Conclusion

Based on the results of the nutritional classification of children under five using the C4.5 decision tree method, the following conclusions can be drawn:

1. The C4.5 decision tree classification method can be used to classify the nutrition of toddlers quite well.
2. The average accuracy for each category is as follows:
   a. The BB/U category classification has an average accuracy of 90.16%.
   b. The TB/U category classification has an average accuracy of 76.64%.
   c. The BB/TB category classification has an average accuracy of 83.83%.

References

[1] P.T. Juniman. “4 Ancaman Bahaya yang Dialami Balita dengan Gizi Buruk” [Online]. Available: https://www.cnnindonesia.com/gaya-hidup/20180125110614-255-271456/4-ancaman-bahaya-yang-dialami-balita-dengan-gizi-buruk, 2008

[2] Kemenkes. *Hasil Utama Riset Kesehatan Dasar Kementerian Kesehatan 2018* [Online]. Available: https://www.depkes.go.id/resources/download/info-terkini/materi_rakorpop_2018/Hasil%20Riskesdas%202018.pdf. 2018

[3] C. Anam and H.B. Santoso. "Perbandingan Kinerja Algoritma C4.5 dan Naive Bayes untuk Klasifikasi Penerima Beasiswa," *Jurnal ENERGY*, **8** (1), 13–19, 2018. [Online]. Available: https://ejournal.upm.ac.id/index.php/energy/article/view/111

[4] U. Febriana, M.T. Furqon, and B. Rahayudi. (2017). "Klasifikasi Penyakit Typhoid Fever (TF) dan Dengue Haemorhagic Fever (DHF) dengan Menerapkan Algoritma Decision Tree C4.5 (Studi Kasus : Rumah Sakit Wilujeng Kediri)," *Jurnal Pengembangan Teknlogi Informasi dan Ilmu Komputer*, **2** (3), 1275–1282, 2017. [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/1124.

[5] J. Han and M. Kamber. *Data Mining: Concept and Techniques*, Second Edition, Morgan Kaufmann Publishers, 2006.

[6] D.T. Larose. *Discovering Knowledge in Data: An Introduction to Data Mining*, John Willey & Sons, Inc., 2005.

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 131–142

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 143–144

p-ISSN 2655-8564, e-ISSN 2685-9432

AUTHOR GUIDELINES

Author guidelines are available at the journal website:
http://e-journal.usd.ac.id/index.php/IJASST/about/submissions#authorGuidelines

International Journal of Applied Sciences and Smart Technologies

Volume 3, Issue 1, pages 143–144

p-ISSN 2655-8564, e-ISSN 2685-9432

This page intentionally left blank