TELKOMNIKA, Vol.17, No.1, February 2019, pp.268~274
ISSN: 1693-6930, accredited First Grade by Kemenristekdikti, Decree No: 21/E/KPT/2018
DOI: 10.12928/TELKOMNIKA.v17i1.11612

Cognitive artificial-intelligence for doernenburg dissolved gas analysis interpretation

Karel Octavianus Bachri*[1], Umar Khayam[2], Bambang Anggoro Soedjarno[3], Arwin Datumaya Wahyudi Sumari[4], Adang Suwandi Ahmad[5]
[1,2,3,4,5]Bandung Institute of Technology, Ganesha St. 10, Bandung, Indonesia, 40132, +62-22-2502260
*Corresponding author, e-mail: karel232@students.itb.ac.id[1], umar@hv.ee.itb.ac.id[2], b.anggoro55@yahoo.com[3], arwin.sumari@yahoo.com[4], adangsahmad@yahoo.com[5]

Abstract

This paper proposes Cognitive Artificial Intelligence (CAI) method for Dissolved Gas Analysis (DGA) interpretation adopting Doernenburg Ratio method. CAI works based on Knowledge Growing System (KGS) principle and is capable of growing its own knowledge. Data are collected from sensors, but they are not the information itself, and thus, data needs to be processed to extract information. Multiple information are then fused in order to obtain new information with Degree of Certainty (DoC). The new information is used to identify faults occurred at a single observation. The proposed method is tested using the previously published dataset and compared with Fuzzy Inference System (FIS) and Artificial Neural Network (ANN). Experiment shows CAI implementation on Doernenburg Ratio performs 115 out of 117 accurate identification, followed by Fuzzy Inference System 94.02% and ANN 78.6%. CAI works well even with small amount of data and does not require trainings.

Keywords*: cognitive artificial-intelligence, DGA interpretaion, information fusion, knowledge growing system*

1. Introduction

Dissolved Gas Analysis (DGA) interpretation is the most reliable method in transformer fault diagnosis, as it is capable of detecting incipient faults before they becomes catastrophic [1–3]. Transformer oil degrades over time during operation [4–6]. At the presence of stresses, radices will be released from the oil as shown in Table 1 [7]. The radices will then form combustible gases as shown in Table 2 [5], [8].

Table 1. Released Radices [7]

Energy Level (KJ/mol)	Fault	Radix Released
< 338	Partial Discharge	H^*, CH_3^*
338	Low-Energy	CH_2^*
607	Arcing	CH^*
720	Thermal-Low	C^*
> 960	Thermal-High	C^*

Table 2. Gases Formed by Radices [5], [8]

Radices	H^*	CH_3^*	CH_2^*	CH^*
H^*	H_2	CH_4	CH_4	CH_4
CH_3^*	CH_4	C_2H_6	-	-
CH_2^*	CH_4	-	C_2H_4	-
CH^*	CH_4	-	-	C_2H_2

The more faults occurs, the more combustible gases are produced in the oil. These gases accelerate the degradation process. There are two kinds of transformer paper degradation processes, Hydrolysis and Pyrolysis [8], [9]. Pyrolysis is related to heat, while Hydrolysis is related to water. Pyrolysis produces Oxygen, which will then oxidize oil and paper insulator. Hydrolysis causes depolymerization of the oil insulator and later on produces Carbon Monoxide and Carbon Dioxide, which are acidic oxides that accelerate Hydrolysis even more.

There are several conventional methods commonly used in DGA interpretation, one of which is Doernenburg Ratio Method (DRM). Conventional methods have limitation, they have low accuracy [10–12], moreover, they have high failure rate [1]. There have been several researches working on DGA interpretation. They started from labeled dataset to test their proposed methods [10]. All of them used Artificial Intelligence (AI) to imitate human experts,

Received July 28, 2018; Revised October 10, 2018; Accepted November 6, 2018

some used FIS [13], [14], some used ANN [1], [12], while some others used modification of standard-AI method [11]. Ratio DGA interpreting methods have been developed to identify faults [1], [10–14] as seen in Table 3.

Table 3. Related Works

Author, year	Problem	Solution	Result
M Duval, 2001	conventional methods have Low accuracy	Dataset creation	In-depth description of the five main type of fault [10].
F Zakaria, 2012	DGA is reliable, but conventional methods have low accuracy	Muti-layer feed forward back preopagation ANN	Increase accuracy, use three layer as the optimal number of layer [12]
S. A. Khan, 2014	conventional methods have Low accuracy	ANFIS on IEC 599	Fault identification and location [11].
M Beykverdi, 2014	IEC Ratio method has high failure rate	Combination of ANN Nero-ICA	Failure rate has been successfully decreased[1]
F D Samirmi, 2015	Basic ontology model requires improvement	Fuzzy Ontology Reasoning	Multi-agent oriented System improves accuracy of basic ontology model[13]
Trianto, 2016	AI and conventional methods have different characteristics	Combining AI and conventional method	Accuracy 93.7% [14]

Normal AI methods have some disadvantages. ANN requires a large number of samples in training [15], while FIS requires complex computing resources and the lacks of design techniques [16]. Adaptive Neural Fuzzy Inferece System (ANFIS) improves the accuracy, but increase complexity of the system.

The most recent development of AI is Cognitive Artificial Intelligence (CAI) [17]. CAI uses a method knowns as Knowledge Growing System (KGS) that is able to solve multi-input and multi-output problems occured in the environment. This method is developed from the combination of Bayesian Inference Method (BIM), Maximum a Posteriori (MAP) theorem, and Linear Opinion Pool (LOP) to obtain decision options [17–19]. In the area of DGA interpretation, DGA problem requires multi-input processing in order to make decision in a multi-output situation, and hence, this method is suitable for this situation.

2. Proposed Method

CAI method solves problems based on Knowledge Growing System (KGS) principle and can be described using Figure 1.

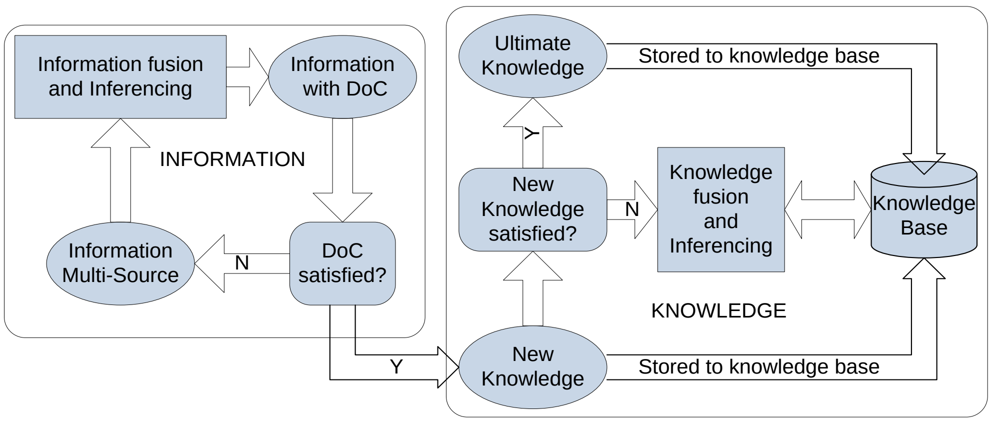

Figure 1. Diagram of KGS [17]

The diagram consists of two parts, the Information Part and the Knowledge Part. In the Information Part, information from multi sources are fused to extract new information with Degree of Certainty (DoC). If the DoC satisfies certain value, the information will then be sent to the knowledge part as the Current Knowledge for further processes.

In the Knowledge Part, the Current Knowledge will be fused with the Existing Knowledge in order to obtain the New Knowledge with DoC. If the DoC satisfies certain value, the knowledge will then become the Ultimate Knowledge.

The main component of CAI is Information Fusion [17]. Information Fusion processes information by imitating the way human brain processes information. Multi-source information were processed in order to obtain new information.

ASSA2010 is the new name given by the method's inventors, to the original method's name namely A3S (Arwin-Adang-Aciek-Sembiring) [17]. It is the algorithm used to fuse information that that is derived from BIM, MAP, and LOP (BIM + MAP + LOP) [17–19] as shown in (1). The probability of Hypothesis B given A is the probability of Indication A given B times the probability of B per all the possible events.

$$P(B_j|A_i) = \sum_{i=1}^{n} \left(\frac{P(A_i|B_j)P(B_j)}{\sum_{k=1}^{m} P(A_i|B_k)P(B_k)} \right) \quad (1)$$

where $P(B_j|A_i)$ is the probability of Hypothesis B_j given A_i, $P(A_i|B_j)$ is the probability of Indication A_i given B_j. $P(B_j)$ is the probability of Hypothesis B_j itself, and $\Sigma(P(A_i|B_k) P(B_k)$ is the combination of all possible events. They are put into arrangement in the Indication-Hypothesis matrix as fused information ($P(\upsilon_i^j)$). BIM + MAP is unable to do decision making in multi-input multi-ouput situation as faced in DGA situation. ASSA2010 improves BIM + MAP by combining the concept of LOP so it has capability in making decision in multi-input multi-ouput situation based on the New Knowledge Probability Distribution (NKPD) it produces after performing computation on Indication-Hypothesis data.

NKPD comprises of DoC for each hypothesis, which indicates how a Hypothesis can be believed regarding to the presence of the combination of Indications. NKPD can be calculated using (2):

$$P(\psi_i^j) = \frac{\sum_{i=1}^{\delta} P(v_i^j)}{\delta} \quad (2)$$

where δ is the number of sensors and $P(\psi_i^j)$ is the inferenced fused-information. The system keeps receiving information from sensors in the form of $P(\psi_i^j)$ according to (3).

$$P(\phi_\gamma^j) = \begin{cases} 1, & P(\psi_\gamma^j) > \frac{P(\psi_\gamma^j)}{\lambda} \\ 0, & P(\psi_\gamma^j) \le \frac{P(\psi_\gamma^j)}{\lambda} \end{cases} \quad (3)$$

where λ is the number of fused information. The New Information is fused with the previous information to produce NKPD over Time (NKPDT), which is as shown in (4).

$$P(\theta_j) = \frac{\sum_{\gamma=1}^{\Gamma} P(\phi_\gamma^j)}{\Gamma} \quad (4)$$

where Γ is the number of observation and $P(\phi_\gamma^j)$ is the NKPD of each observation. Decision will be made based on the highest DoC in the NKPDT matrix as shown in (5):

$$DoC = \begin{cases} max[P(\psi_\gamma^j)], & single\ observation \\ max[P(\theta_j)], & time\ series\ observation \end{cases} \quad (5)$$

where j = 1, 2, 3, …, λ. The proposed method can be described in pseudocode as follows:
1. Read data [gas].
2. Calculate Ratio using (6) to (9).
3. Localize Ratio using Table 4.
4. Create observation matrix as Table 5.
5. Obtain NKPD using (2).
6. determine fault using (5)

7. if more data available then goto step 1

The Doernenberg Ratio Method uses gas concentration Ratios that is calculated using (6)-(9) and works based on Figure 2.

$$R_1 = [CH_4]/[H_2] \quad (6)$$

$$R_2 = [C_2H_2]/[C_2H_4] \quad (7)$$

$$R_3 = [C_2H_2]/[CH_4] \quad (8)$$

$$R_4 = [C_2H_6]/[C_2H_2] \quad (9)$$

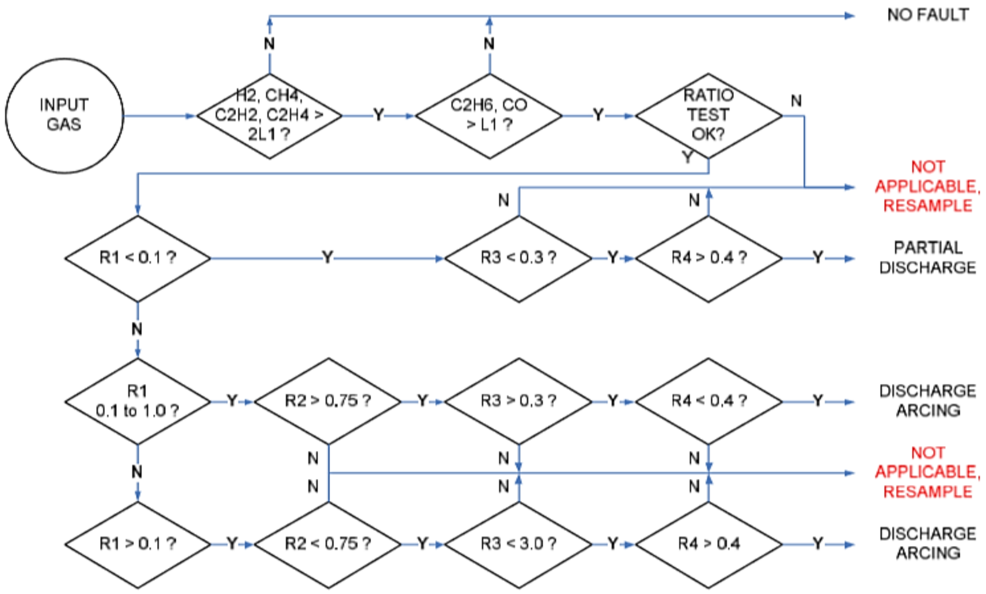

Figure 2. Doernenberg decision system flowchart [7]

Localization of Ratios is shown in Table 4. The input of Table 4 is ratio R_i, and the output is score A_i, if R_1 exceeds 1, A_1 will be given score '2', if R_1 is between 0.1 and 1, A_1 will be given score '1', and if R_1 is below 0.1, A_1 will be given score '0'. If R_2 is below 0.75, A_2 will be given score '0', and if R_2 exceeds 0.75, R_2 will be given score '1'. If R_3 exceeds o.3, A_3 will be given '1', otherwise A_3 will be given '0'. If R_4 exceeds 0.4, A_3 will be given '1', otherwise A_4 will be given '0'.Observation matrix is shown in Table 5, while NKPD is calculated using (2) and is shown in Table 6. Since the test data is not time-series, NKPDT does not need to be calculated, but for time-series data, NKPDT can be calculated using (4) to show trends in a certain time interval. Fault is identified using (5) for single observation. It is the max of $P(\psi_i^j)$ for each sample.

Table 4. Ratios Localization [7]

	TD	PD	A
R_1	> 1	<0.1	0.1 < <1
R_2	<0.75	ns	>0.75
R_3	<0.3	<0.3	>0.3
R_4	>0.4	>0.4	<0.4

Table 5. Observation Matrix [17]

R	Score	Hypotheses B_1	B_2	B_3
R1	A_1	$P(B_1\|A_1)$	$P(B_2\|A_1)$	$P(B_3\|A_1)$
R2	A_2	$P(B_1\|A_2)$	$P(B_2\|A_2)$	$P(B_3\|A_2)$
R3	A_3	$P(B_1\|A_3)$	$P(B_2\|A_3)$	$P(B_3\|A_3)$
R4	A_4	$P(B_1\|A_4)$	$P(B_2\|A_4)$	$P(B_3\|A_4)$

Table 6. NKPD Matrix [17]

Sample	Hypotheses H_TD	H_PD	H_A
1	$P(\psi_1^1)$	$P(\psi_2^1)$	$P(\psi_3^1)$
2	$P(\psi_1^2)$	$P(\psi_2^2)$	$P(\psi_3^2)$
…	…	…	…
j	$P(\psi_1^j)$	$P(\psi_2^j)$	$P(\psi_3^j)$
…	…	…	…
n	$P(\psi_n^1)$	$P(\psi_n^1)$	$P(\psi_n^1)$

3. Results and Analysis

The proposed method is verified using previously-published data consisting 117 samples with 9 Partial Discharge (PD), 26 Low-Energy Discharge (LE), 48 High-Energy Discharge (A), 16 Thermal-Low (TL), and 18 Thermal-High (TH) that is put into groups based on the fault types. For example, data #1 of PD, R_1=0.07. R_2=0.00, R_3=0, and R_4=inf. Using Table 3, Ratios are then localized and given score, A_1=0, A_2=0, A_3=0, A_4=1. These scores are arranged in observation matrix in Table 5. Table 3 is also used to give scores to each Hypothesis, for example R1='0', column H_TD, H_PD, and H_A for row R1 is given '0', '1', and '0' respectively to be arranged in Table 7. Other types of faults can be analyzed using the same method. NKPD is calculated using (2) and is shown in Table 8.

Table 7. Observation Matrix PD

Data	A	score	Hypotheses H_TD	H_PD	H_A
1	A_1	0	0	1	0
	A_2	0	1	1	0
	A_3	0	1	1	0
	R_4	1	1	1	0
…	…	…	…	…	…
9	A_1	1	0	0	1
	A_2	0	1	1	0
	A_3	0	1	1	0
	A_4	1	1	1	0

Table 8. NKPD for PD.

No.	H_TD	H_PD	H_A
1	0.3750	0.6250	0.0000
2	0.2500	0.6250	0.1250
3	0.3750	0.3750	0.2500
4	0.3750	0.3750	0.2500
5	0.2500	0.3750	0.3750
6	0.2500	0.3750	0.3750
7	0.3750	0.6250	0.0000
8	0.2500	0.6250	0.1250
9	0.3750	0.3750	0.2500

Sample number 1, 2, 7, and 8 in dataset PD indicates the major faults were PD, some with the minor fault TD, while others with A. meanwhile, in sample number 3, 4, 5, 6, and 9, no major faults occurred, however PD are still identified as the DoCs are among the highest. All samples in the dataset LE indicate the major faults are Arcing, while minor faults are mostly PD, this is due to the limitation in Doernenburg method, which identifies only verylow Energy Discharge (PD) and High-Energy Discharge (A). All samples in the dataset HE indicate the major faults are Arcing, while some others are identified as TD and also some others as PD. Dataset TL and TH give TD, with two samples in the dataset TL are identified as PD.

The dataset are also tested using Fuzzy Inference System (FIS) and Artificial Neural Network (ANN). FIS and ANN are designed to be implemented on the same conventional DRM using the same dataset. FIS has four input membership functions (R_1, R_2, R_3, and R_4) and three output membership functions (TD, PD, and A). The result is shown in Table 9. For dataset PD, FIS has 88.9% accuracy, while ANN has 55.6%. For dataset LE, FIS has 92.3% accuracy, while ANN has 53.9%. For dataset HE, FIS has 100% accuracy, while ANN has 95.8%. For dataset TL, FIS has 75% accuracy, while ANN has 81.3%. For dataset TH, FIS has 100% accuracy, while ANN has 77.8%. The accuracy of FIS in the testing is confirmed by various authors, which lies between 93% to 96% with various datasets and conditions [13], [14], while the accuracy of ANN varies due to the small numbers of samples [1], [11, 12].

The proposed method performs 87.5% accuracy for dataset TL and 100% accuracy for other dataset. The overall performance of CAI is 98.3% accuracy, with 115 correct identifications out of 117 samples. The accuracy of ANN is proportional to the number of samples.

Table 9. Comparison between CAI and other Methods

Fault Type	#Data	#Accurate			%Accuracy		
		FIS	ANN	CAI	FIS	ANN	CAI
PD	9	8	5	9	88.9	55.6	100
LE	26	24	14	26	92.3	53.9	100
HE	48	48	46	48	100	95.8	100
TL	16	12	13	14	75	81.3	87.5
TH	18	18	14	18	100	77.8	100
Overall	117	110	92	115	94.02	78.6	98.3

4. Conclusion and Future Works

In this research, a novel method of DGA interpretation is proposed. The proposed method is based on CAI with KGS as the core. In this paper CAI adopts Doernenburg Ratio method. There are some important things that can be noted from experiments.

4.1. Conclusion

The proposed method has successfully identify fault based on DGA with the overall accuracy of 98.3% from five types of fault with 87.5% accuracy on Thermal-Low, and 100% accuracy on Partial Discharge, Low-Energy Discharge, High-Energy Discharge, and Thermal High, which results in higher accuracy than FIS (94.02%) and ANN (78.6%). This CAI-based proposed method does not require a large amount of data and training as ANN does and does not require a complex computation as FIS does.

The more the sample number, ANN tends to have more accuracy. The accuracy of FIS is the proportional to the membership functions and rule base. The more detail the membership functions, the more accuracy FIS will have, resulting more complexity.

The accuracy of the proposed method is independent to the number of samples. It is shown in Table 8, where various number of samples resulting 100% accuracy. The complexity of the proposed method is only proportional on the number of hypotheses and indications. Another advantage of the proposed method is CAI works well even with small amount of data and does not require trainings. Meaning that the proposed method is able to capture the ability of the human brain in performing fast learning to obtain decision options for multi-input multi-output problems.

4.2. Future Works

The proposed method is a novel method in computing, and AI area that is developed to solve multi-input and multi-output problems, such as in Biomedic [20], DPA countermeasure [21], Intrusion Detection System [22], emotion modeling [23], and many other area requiring a Multi-input-Multi-Output processing such as in big data [24]. The proposed method can be implemented in software [22] and hardware [21], [25]. For the time being, a cognitive processor is being developed [26] and is going to be used in an embedded system to perform such tasks with more power efficiency and higher speed.

Acknowledgment

The first author would like to show his gratitude to Mr. Harry Gumilang for his support through discussions.

References

[1] M. Beykverdi, F. Faghihi, and others. A new approach for transformer incipient fault diagnosis based on dissolved gas analysis (DGA). *Nova Journal of Engineering and Applied Sciences*. 3(2). 2016.
[2] D. Bhalla, R. K. Bansal, and H. O. Gupta. Transformer incipient fault diagnosis based on DGA using fuzzy logic. in *2010 India International Conference on Power Electronics (IICPE)*. 2011: 1–5.
[3] H. Ma, Z. Li, P. Ju, J. Han, and L. Zhang. Diagnosis of power transformer faults on fuzzy three-ratio method. in *The 7th International Power Engineering Conference IPEC 2005*. 2005: 1–456.
[4] V. Arakellian and I. Fofana. Physicochemical aspects of gassing of insulating liquids under electrical stress. *IEEE Electrical Insulation Magazine*. 2009; 25(3): 43–51.
[5] R. Sanghi. Chemistry behind the life of a transformer. *Resonance*. 2003; 8(6): 17–23.

[6] J. B. Taylor. Discussion on 'Corona Losses between Wires at High Voltages' (Harding), 'The Law of Corona and the Dielectric Strength of Air-II' (Peek), and 'The Electric Strength of Air-III' (Whitehead). Boston, Mass., June 25, 1912. *Transactions of the American Institute of Electrical Engineers*. 1912; XXXI(1): 1119–1130.

[7] "IEEE Guide for the Interpretation of Gases Generated in Oil-Immersed Transformers - Redline. *IEEE Std C57.104-2008 (Revision of IEEE Std C57.104-1991) - Redline*. 2009; Feb: 1–45.

[8] H. Gumilang. Hydrolysis process in PLN P3BJB transformers as an effect of oil insulation oxidation. in *2012 International Conference on Condition Monitoring and Diagnosis (CMD)*. 2012: 1147–1150.

[9] B. CIGRE Working Group 12.18. 227 guidelines for life management techniques for power transformers. *CIGRE WG*. 2003: 12.

[10] M. Duval and others. Interpretation of gas-in-oil analysis using new IEC publication 60599 and IEC TC 10 databases. *IEEE Electrical Insulation Magazine*. 2001. 17(2): 31–41.

[11] S. A. Khan, M. D. Equbal, and T. Islam. ANFIS based identification and location of paper insulation faults of an oil immersed transformer. in *2014 6th IEEE Power India International Conference (PIICON)*. 2014: 1–6.

[12] F. Zakaria, D. Johari, and I. Musirin. Artificial neural network (ANN) application in dissolved gas analysis (DGA) methods for the detection of incipient faults in oil-filled power transformer. in *2012 IEEE international conference on Control system, computing and engineering (ICCSCE).* 2012: 328–332.

[13] F. D. Samirmi, W. Tang, and Q. Wu. Fuzzy Ontology Reasoning for Power Transformer Fault Diagnosis. *Adv. Electr. Comput. Eng*. 2015; 15: 107–114.

[14] Trianto, Suwarno, Y. Li, and G.-J. Zhang. Combining conventional and artificial intelligence DGA interpretation methods using optimized weighting factor. in *2016 International Seminar on Intelligent Technology and Its Applications (ISITIA)*. 2016: 37–42.

[15] M.-Y. Chow and P. Goode. The advantages and challenges of machine fault detection using artificial neural network and fuzzy logic technologies. *in Proceedings of the 36th Midwest Symposium on Circuits and Systems, 1993.* 1993: 708–711.

[16] I. Bousserhane, A. Hazzab, M. Rahli, M. Kamli, and B. Mazari. Adaptive PI controller using fuzzy system optimized by genetic algorithm for induction motor control. *in International Power Electronics Congress, 10th IEEE*. 2006: 1–8.

[17] A. D. W. Sumari, A. S. Ahmad, A. I. Wuryandari, and J. Sembiring. Constructing brain-inspired knowledge-growing system: A review and a design concept. *in, 2010 International Conference on Distributed Framework and Applications (DFmA).* 2010: 1–7.

[18] K. O. Bachri, B. Anggoro, and A. S. Ahmad. Transformer performance calculation using information fusion based on DGA interpretation. *in International Symposium on Electronics and Smart Devices (ISESD).* 2016: 11–15.

[19] A. D. W. Sumari and A. S. Ahmad. Knowledge-growing system: The origin of the cognitive artificial intelligence. in *Electrical Engineering and Informatics (ICEEI), 2017 6th International Conference on* 2017: 1–7.

[20] L. Goeirmanto, R. Mengko, and T. L. Rajab. Direction of ventricle contraction based on precordial lead ECG signal. in *Cyber and IT Service Management, International Conference on Electronics and Smart Devices (ISESD).* 2016: 1–3.

[21] S. D. Putra, A. S. Ahmad, and S. Sutikno. DPA-countermeasure with knowledge growing system. in *International Symposium on Electronics and Smart Devices (ISESD)*. 2016: 16–20.

[22] H. R. Talompo, A. S. Ahmad, Y. S. Gondokaryono, and S. Sutikno. NAIDS design using ChiMIC-KGS. *in 2017 International Symposium on Electronics and Smart Devices (ISESD).* 2017: 346–351.

[23] T. Sutikno, M. Facta, and G. A. Markadeh. Progress in Artificial Intelligence Techniques: from Brain to Emotion. *TELKOMNIKA Telecommunication Computing Electronics and Control.* 2013; 9(2): 201–202.

[24] T. Sutikno, D. Stiawan, and I. M. I. Subroto. Fortifying big data infrastructures to face security and privacy issues. *TELKOMNIKA Telecommunication Computing Electronics and Control*. 2014; 12(4): 751–752.

[25] T. Sutikno, N. R. N. Idris, and A. Z. Jidin. Overview on Strategies and Approaches for FPGA Programming. *TELKOMNIKA Telecommunication Computing Electronics and Control.* 2014; 12(2): 273–282.

[26] C. O. Sereati, A. D. W. Sumari, T. Adiono, and A. S. Ahmad. Implementation Knowledge Growing System Algorithm using VHDL. in *International Symposium on Electronics and Smart Devices (ISESD).* 2016: 7–10.