Data-Driven Anesthesia: An Ensemble Model for Propofol and Remifentanil Dosage Control During Medical Surgery

Abas Almaayofi[1], Farah Almulla[2], Mohammed Almulla[1]

[1]Computer Science Department – College of Science - Kuwait University - Kuwait

[2]Pediatrics Department – Al-Adan Hospital – Ministry of Health – Kuwait

ABSTRACT

Anesthesia is a critical medical intervention used to ensure patients remain unconscious, pain-free, or immobile during surgical and diagnostic procedures. The choice of anesthetics is influenced by the type of surgery, the patient's medical history, and the preferences and expertise of the anesthesiologist. Anesthetics is usually administered through inhalation, intravenous injection, or a combination of both. Administering anesthesia during medical procedures is vital to patient care, requiring precision, flexibility, and real-time adaptability. In this work, we propose a new machine learning model that relies on LSTM and a fully connected neural network to control the patient's anesthetic state during surgery for all stages including induction, maintenance, and emergence, using a synergy of Propofol and Remifentanil. Propofol is used primarily for sedation, whereas Remifentanil is mainly used for pain relief. Since the duration of their effect is very short, constant infusion of both drugs is necessary to maintain the patient's sedation state. Experience indicates that the synergistic effect of both drugs yields better control of the anesthesia level. This model is meant to elevate the burden that comes with the task of anesthesia control in real-time but shouldn't take complete control as the presence of anesthesiologists remains vital to monitor its performance.

Keywords: Anesthesia, Machine Learning Predictive Model, LSTM, Fully Connected Neural Network, Personalized Medicine.

INTRODUCTION

Anesthesia is a critical medical intervention used to ensure patients remain unconscious, pain-free, or immobile during surgical and diagnostic procedures. It involves administering specific anesthetics that are carefully selected based on the type of surgery, the patient's age, medical history, and other factors. Moreover, the choice of anesthetics is also influenced by the preferences and expertise of anesthesiologists. Anesthetics can be general which means they tend for all aspects of anesthesia, like inducing unconsciousness, pain relief, and muscle relaxation. There are two types of anesthetics: (i) inhalation anesthetics (administered via the raspatory system) like gases or vapors that are inhaled using a mask or endotracheal tube, and (ii) intravenous anesthetics (administered via an injection through the veins) are categorized based on their physiological effects into hypnotics, analgesics, and neuromuscular blocking (NMB) drugs [1]. Hypnotics, such as propofol, induce unconsciousness during surgery, with propofol being widely preferred due to its rapid metabolism and lower risk of side effects. Analgesics, like opioid-based remifentanil, alleviate pain sensations. NMB drugs cause skeletal muscle paralysis by blocking nerve signals at the neuromuscular junction, facilitating procedures like endotracheal intubation and mechanical ventilation. These three drug types (shown in Fig. 1) collectively achieve the main goals of general anesthesia: unconsciousness, pain relief, and muscle relaxation.

Fig. 1 The three main goals of general anesthesia.

Anesthesia is administered and managed in three distinct stages: (1) Induction: This initial phase involves transitioning the patient from a conscious state to unconsciousness or from full sensation to partial or complete sensory loss. (2) Maintenance: this stage is dedicated to monitoring the patient's vital signs, such as blood pressure, oxygen saturation, and heart rate, to adjust the dosage or type of anesthetic agents. This ensures the patient remains stable, unconscious, pain-free, and appropriately relaxed throughout the surgery. (3) Emergence: The final phase involves reducing or discontinuing the anesthetic agents, allowing the patient to regain consciousness and normal bodily functions. This process must be carefully managed to ensure smooth and safe recovery from the anesthetic state. Anesthesia is a highly specialized field that requires precise monitoring and expertise to tailor the treatment to the patient's individual needs, ensuring safety and comfort throughout the surgical process [2]. Approximately 266 million surgeries are performed worldwide each year, many of which require general anesthesia. Effective management of anesthesia plays a key role in operating room efficiency and the length of procedures. This is influenced by factors such as the time needed to achieve target anesthetic concentrations and the efficiency of drug delivery, which affect both the induction of anesthesia and the timing of patient emergence [3].

Problem Statement

The main issue with anesthesia control delivery is the human factor. The task of sustaining a stable hypnotic state is complex, as the anesthesiologist needs to monitor, accurately assess, and adapt to the patient's vital signs variability to deliver the precise drug dosage [4]. This means that the anesthesiologist must be an expert and well-rounded in his field to the point that he/she can predict the patient's response to certain drug infusions, and this is quite challenging, given that every patient responds differently to different anesthetic agents. There's also the issue of human fatigue, as some surgeries might last for days, hence, the anesthesiologist can't maintain the same level of attention and readiness to make constant decisions for drug infusion during the surgery. Even if there are shifts between anesthesiologists, where another anesthesiologist is ready to take over if one feels tired, there's the issue of transition delay. Also, some hospitals might lack staff in the anesthesiology department. Lastly, anesthesiologists rely on different monitoring tools, some are expensive. Some devices are advanced, but they are also limited in providing a complete picture of the patient's physiological signs, and this can lead anesthesiologists to imprecise administration of anesthetic agents, causing fluctuation in the hypnotic state (patient might inadvertently wake up).

RELATED WORK

The safe and personalized administration of anesthetics during surgery is a major concern in clinical practice, necessitating precision, adaptability, and a personalized approach [5]. Integrating machine learning into anesthesia control can improve the accuracy and efficiency of its delivery, as well as ensure patient safety by reducing human errors such as over or under-dosing episodes [6]. Anesthesiologists have historically been at the forefront of developing closed-loop devices. As early as the 1950s, Bickford and colleagues introduced automated volatile anesthetic delivery systems guided by electroencephalogram (EEG) data [7]. Modeling plays a critical role in both feed-forward and feed-backward control systems. In automated anesthesia, these approaches are commonly known as target-controlled infusion (TCI) and closed-loop drug delivery, respectively. Both aim to regulate a system where the anesthetic drug infusion rate serves as the input, and the (measured) clinical effect acts as the output [8].

In 2023 the authors of [9] proposed an interoperative EEG model that uses gradient boosting to accurately detect loss and recovery of consciousness during anesthesia, showcasing high precision. Whereas in [10], the authors suggested a framework that predicts continuous depth of anesthesia using SWT and fractal features. This model achieved 97.1% classification accuracy and superior regression performance. Also in 2023, the authors of [11] introduced machine learning models and artificial intelligence (AI) to allow objective and personalized nociception-antinociception prediction in the patient safety era for the design and evaluation of closed-loop analgesia controllers.

Earlier in 2021 the authors of [12] published a study that considered machine learning for anesthesiologist decisions on remifentanil, with LSTM showing promising performance, enhancing anesthesia decision-making potential. In a different attempt, Liu et. al (2019) analyzed EEG signals using a convolution neural network to

model the patient's consciousness level based on the anesthesiologist's experience. The model achieved 93.50% accuracy [13]. Similarly, Asai et. al [14] proposed an anesthetics dose prediction model to avoid post-induction hypotension using electronic anesthesia records. Their model used ridge regression on electronic records with promising results.

METHODOLOGY

The traditional method for anesthesia control relies on the anesthesiologist's experience and clinical judgment, such methods are prone to subjective judgment and human error. To address this issue, we propose an ensemble machine learning model that uses LSTM and a fully connected neural network (Dense) with the help of a dropout layer to control and predict the drug dosages of Propofol and Remifentanil. This model will utilize real-time patient data for patients during the administration of anesthesia. ML's exponential growth in medicine is made possible by the availability of large datasets and improvement in computing power, as it is a computer-controlled technique that automates analytical model building [15]. The data was obtained from a large clinical database called VitalDB (https://vitaldb.net/dataset/) [16]. The data in this set consists of intraoperative bio signals and clinical information related to 6388 surgical patients, the data includes many features, but we only focus on seven to avoid overfitting, and to encourage generalization to make informed decisions. Both drugs, Propofol and Remifentanil, are characterized by a rapid onset and offset, meaning they take effect in a short period, and they wear out also in a fast manner [17]. So, to sustain a stable unconsciousness and analgesia state, a constant infusion of a combination of these drugs is necessary. This might seem flawed, but in contrast, the behaviour of the combination of these drugs is perfect for all three stages of anesthesia: induction, maintenance, and emergency. This might be obvious for induction and emergency because both require a fast change of anesthetic state, otherwise, it won't be convenient for physicians to wait long before they start operating on the patient (induction), same for emergencies, where a prolonged recovery can result in complications. Maintenance on the other hand, can't be achieved with drugs that have prolonged effects, as the effect of overdosing is not recoverable and might result in prolonged unconsciousness, raspatory depression, and delayed emergency. In other words, it's hard to fine-tune drugs with prolonged effects to sustain a certain range of anesthetic states. The value that we seek to maintain is the BIS (Bispectral index), which is a measurement of brain activity. Its values range from 0 to 100, 0 means there's no brain activity at all, and 100 indicates fully awake. For surgery purposes, the BIS value has to be within the range of 40 to 60 to sustain an unconscious state with minimum side effects (See Fig. 2). It relies on the concertation of Propofol and Remifentanil in the effect-site, so to control it, we need to control these two drugs. Finally, this model will elevate the burden that comes with the task of anesthesia control but shouldn't take complete control as the presence of anesthesiologists remains vital to monitor its performance.
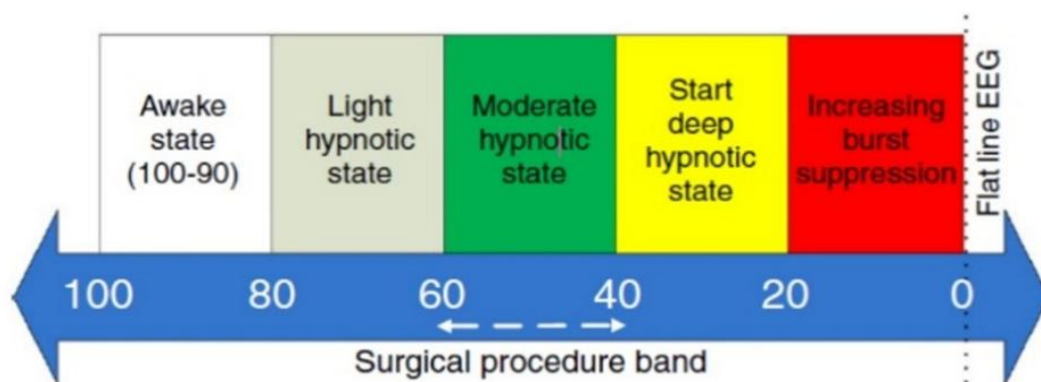

Fig. 2 Bispectral index range interpretation.

System Description

In healthcare, machine learning algorithms helped in disease diagnosis, treatment recommendation, and patient recovery prediction, laying the framework for personalized medicine. Machine learning algorithms analyze extensive datasets comprising patients' information, surgical details, and medical histories. The insights gained will be used to develop a customized anesthetic prediction model that optimizes medication administration based

on individual patient characteristics. Our system utilizes Long-Short Time Memory (LSTM) for the Propofol and Remifentanil dosage prediction since LSTMs are perfect for sequential data where the order is crucial; given that this characteristic aligns with the behavior exhibited by this discipline of anesthesia control, where previous dosages determine the current required dosage. The model seeks to mitigate the risks posed by unpredictable patient responses. Leveraging trained algorithms to analyze individual characteristics, it enables safer and more precise anesthetic protocols, potentially improving patient outcomes and recovery. During surgery, the system will continuously monitor data, updating predictions in real time to adapt to evolving conditions [5].

Both LSTM networks are considered within a single layer, their outputs are passed to the next layer which is a *fully connected neural network*. This neural network aims at predicting the BIS value, but before that, we need to add a third layer: the *dropout* to avoid overfitting. The dropout layer randomly disables neurons in the fully connected neural network to ensure a model with generalized behavior. The last layer is a single sigmoid neuron that outputs the value for the BIS in the range 0-1 (will be multiplied by 100 to accurately represent the BIS value). In this section, we will illustrate the working principle of the model and how it's constructed.

Data Cleaning and Preprocessing

The most vital aspect for the model to work correctly is the kind of data used to train it. For medicine-related research and applications, a well-known database named VitalDB contains rich and valuable datasets for physiological and demographic data from patients that undergone surgeries. Out of these physiological and demographic data, we are interested in a few that we believe are enough for building an effective machine-learning agent that can control the infusion rate at an expert level. The seven parameters are: Propofol Volume, Remifentanil Volume, Age, Sex, Height, Weight, and BIS level.

- Removing records with an initial BIS level lower than 80; the usual initial BIS level is between 90 and 100, but due to the huge drop that would result in the number of records if we were to go with 90 instead of 80 (which can lower the model's accuracy), we think 80 is a good threshold for indicating patient's awareness and also maintaining a reasonable number of records for the training and testing phases. Also, eliminating records with the last BIS level that is less than 80, as the last BIS level should indicate the patient's recovery status.
- Removing records where patients are under the age of 20, and the reason behind such selection criteria is to minimize variability and focus on a specific population.
- Removing records where patients are overweight or underweight, our thresholds are: 130 KG and 40 KG respectively. These thresholds are guessed, and the logic is similar to the previous selection criteria, which is to ignore outliers and account only for a specific population.
- Removing records with anesthesia type that are not general.
- Fill in records that have *zero* or *no* values with the last non-zero values. Also, fill in the records that have *no* values in their initial readings with zeros.
- Change the values in the *Propofol volume* and *Remifentanil volume* columns to rates by subtracting the immediate last reading from the current.
- Adding additional records at the top of the dataset with zero values. These zeroed records are used for the LSTM training as this type of neural network takes inputs as sequence, and so the idea is to feed the first sequences that are zeros so that the last value of the sequence is a real value, and the LSTM will have to predict the following output to this real value.

Example: Assuming we want to pass the first sequence to the LSTM and assume that the number of zeroed records added at the beginning of the dataset is 119, then, the length of the sequence is 120 where the 120th value is the value that belongs to the first record before adding the additional zeroed records. The sequence is then passed to the LSTM as a single input. After implementing these selection criteria on the dataset, the input should be ready to be fed into the model for training.

Model Design

An ensemble machine learning model was developed for the BIS level prediction and Propofol and Remifentanil dosage control. The first layer of the ensemble consists of two LSTM models, with the following specifications

in Table 1.

Table 1: LSTM Models Specifications

Hyperparameters	LSTM1 (Propofol)	LSTM2 (Remifentanil)
Nodes	8	8
Time Steps	120	120
Activation Function	Leaky ReLU	Leaky ReLU

The inputs to these LSTM neural networks are sequences of the length of 119. Each entry in the sequences is organized in an orderly manner, corresponding to each time step, and the next step that we wish to predict will be the 120th in the sequence. The outputs for both drugs, along with the covariates (age, gender, height, and weight) are concatenated and flattened to be processed by the fully connected neural network, which consists of a single hidden layer with neurons (activated using ReLU activation function) and an output neuron activated using the sigmoid activation function. To maintain a model of smooth and generalized nature we add a dropout layer that regularly disables neurons in the hidden layer. The percentage set for neurons to be disabled in the hidden layer is 20%. So, we get an output that represents the BIS level in the range of 0-1, to better resemble the range we can multiply the output by 100. The other two outputs are simply the control dosages taken from the last entry (which is predicted) output from the LSTM networks. You may reference Fig. 3 for better comprehension of the model's architecture.
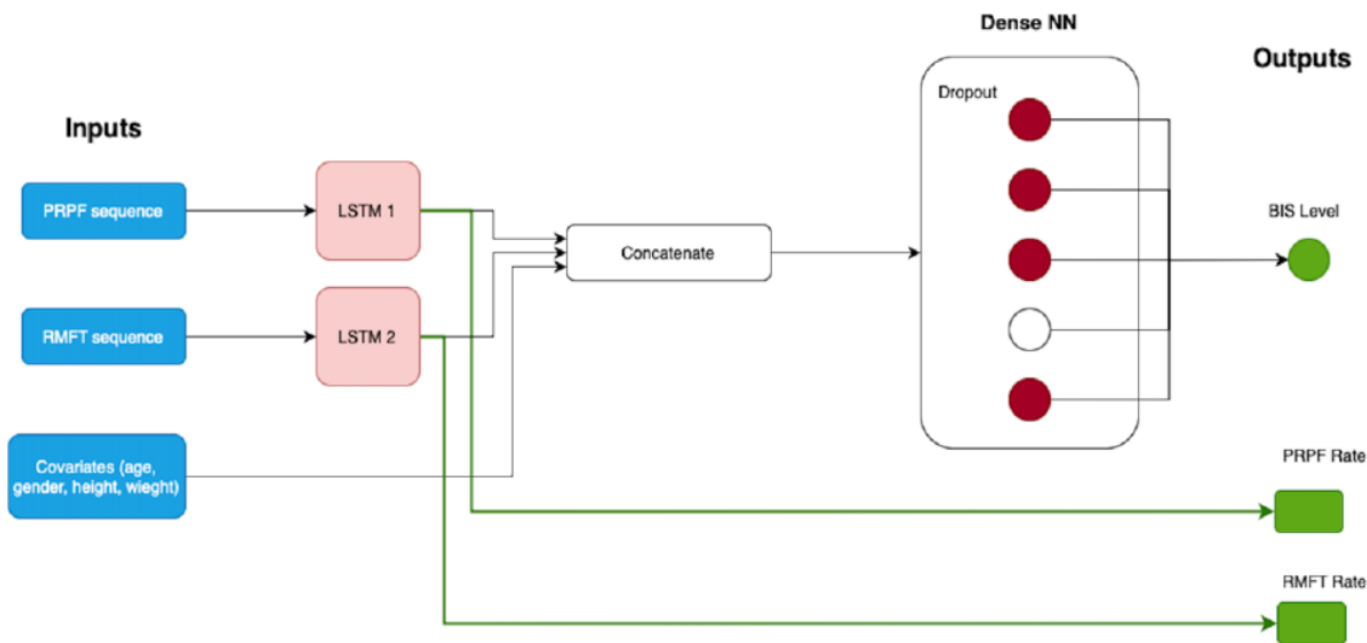

Fig. 3: Model Architecture

EXPERIMENTATION & DISCUSSION

Model Performance & Comparison

The model exhibits promising performance as seen in Fig.4, where the predicted values are smooth and stable with a Mean Absolute Percentage Error of less than 17% (it would've been less than 10%, but due to fluctuation and added noises during device readings, the MAPE showed an increase). It was observed that the percentage of predicted BIS values that are maintained within the 40-60 range, is more than 81%, the rest of the predictions are either outliers or related to the induction and emergence stages.
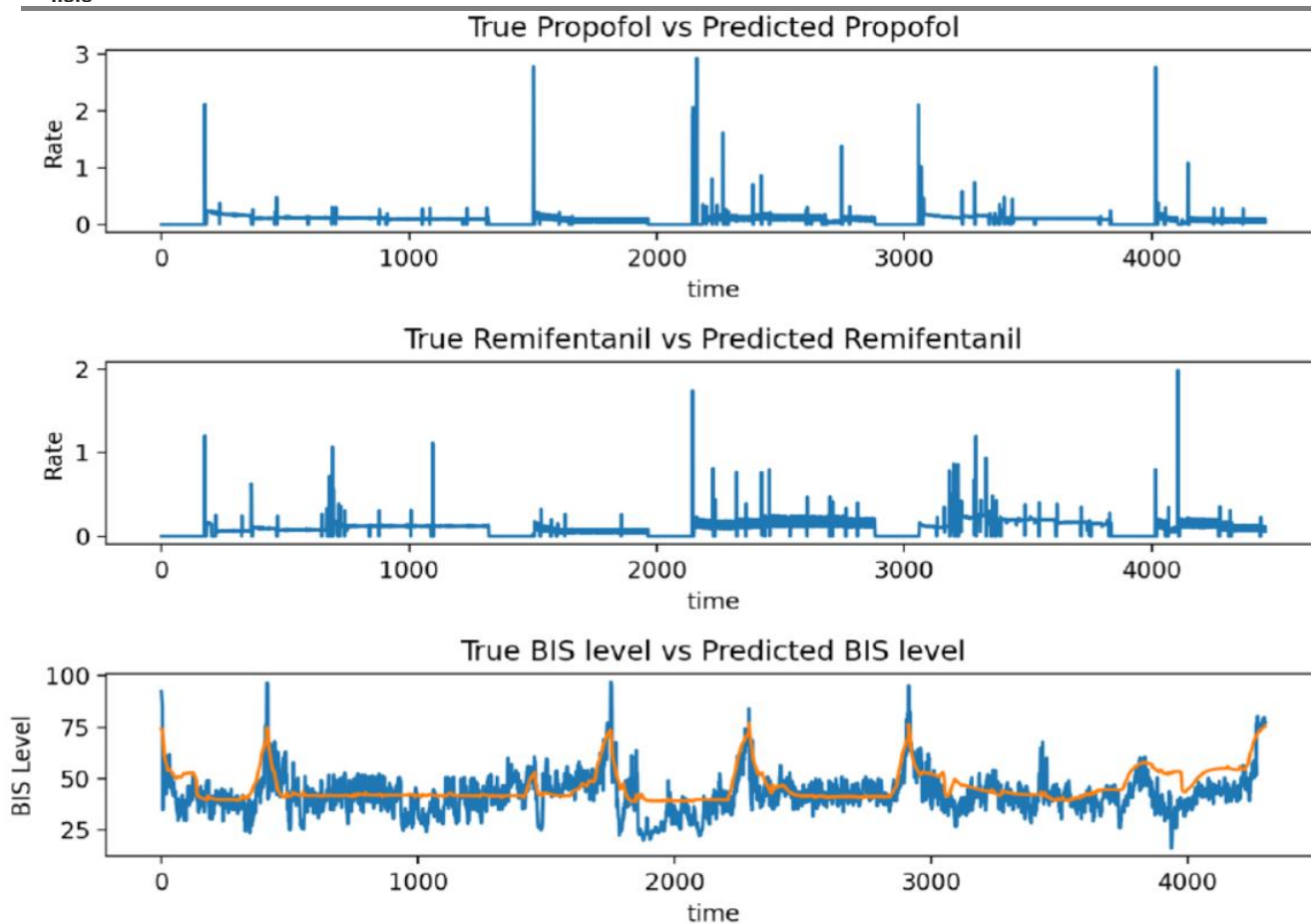
Fig. 4 Actual vs predicted BIS, Propofol & Remifentanil.

Unfortunately, the goal was to also do an extended comparison and analysis with another model that performs the same task with relatively the same type of machine learning, but the literature lacked any ensemble-based machine learning model that predicts BIS and attempts to control Propofol and Remifentanil, which indicates the novelty of our model. Moreover, testing proportional hazards assumption results showed that the model fitted reasonably well to the data and met the proportional hazards assumption [18].

Advantages and Limitations

The model excels in three aspects: (1) the usage of real-time data during actual operations. There's nothing more valuable for training a model than the quality of the dataset. Not only does the data resemble real anesthetic administration on real patients, but also the abundance of recorded readings holds significance to the model's accuracy. (2) the model training time is less than 5 minutes, with an impressive performance. (3) the model predicts the administration of not only a single anesthetic agent but two, to give a synergetic effect by combining both Propofol and Remifentanil which has the prospect of yielding more control on the BIS level with relatively fewer dosages.

The limitation can be summarized as the lack of wider population coverage, in the preprocessing step, we've removed many records that don't comply with our selection criteria, like age, weight, type of anesthesia, and more. Restrictions were put on the dataset to get a cleaner and population-specific dataset. Thus, the model is promising with a portion of the dataset, but it lacks consideration to rare cases, which indicates that the model architecture lacks improvements to include predictions for all cases. The improvements to the model might be as simple as changing the hyperparameters and implementing try-and-error to derive the optimum.

CONCLUSION & FUTURE WORK

Using a synergy of Propofol and Remifentanil, a new ensemble machine learning model that relies on LSTM and fully-connected neural networks to control a patient's anesthetic state during surgery in all stages including

induction, maintenance, and emergence has been proposed. Although ML-guided anesthesia has a great impact on the quality and the cost-effectiveness of the patient's access to healthcare, ML applications for clinical anesthesia might raise ethical challenges and safety concerns [19]. For example, a patient's life may depend on an anesthesiologist's ability to regain control from the machine learning model if the latter fails to deliver the right amount of anesthesia, hence maintaining some clinical and cognitive skills will be necessary.

Possible future extensions may include considering an additional relaxant drug like Succinylcholine (usually used to render patients paralyzed during surgeries, particularly for those who suffer from involuntary muscle movements). Another enhancement could be to create a human dynamics simulator, in the sense: that the simulator is the environment where we can generate and observe physiological signs during drug infusion as if we are administering the agent on a real patient, this can allow us to test on a new and reliable data. A special problem in modern medicine is the diagnosis and monitoring of the condition of children who have had critical surgery [17]. The causes of the onset and development of diseases are individual. Currently, we are testing the proposed model on child anesthesia.

ACKNOWLEDGMENT

REFERENCES

1. Donohue, C., Hobson, B., & Stephens, RC. (2013). An introduction to anesthesia. British Journal of Hospital Medicine, 74(5): C71-75. https://doi.org/10.12968/hmed.2013.74.Sup5.C71
2. Toma, A. & Sahib, M.A. (2023). A Comprehensive Review on Automated Control of Anesthesia: Recent Methods, Challenges, and Future Trends, Wasit Journal of Pure Science 2(2): 291-315, https://doi.org/10.31185/wjps.160.
3. Beard, J.W., Yacoubian, S. Luchetti, M. Yapici, H.O. & Kennedy, R.R. (2024). Anesthesia delivery via manual control versus end-tidal control: A scoping review, Trends in Anesthesia and Critical Care, 58: (101501), https://doi.org/10.1016/j.tacc.2024.101501.
4. Hashemi, S., Yousefzadeh, Z., Abin, A., Ejmalian, A., Nabavi, Sh. & Dabbagh, A. (2024). Machine Learning-Guided Anesthesiology: A Review of Recent Advances and Clinical Applications, Journal of Cell Mol. Anesth., 9(1): https://doi.org/10.5812/jcma-145369.
5. Vinoth, M.M., Lordson, S.R.B. & Ramana, M. (2024). Anesthesia prediction using machine learning, International Research Journal of Engineering and Technology, 11(5): 2039-2046.
6. Milanesi, M., Paolino, N., Schiavo, M., Padula, F. & Visioli, A. (2024). PIDA control of depth of hypnosis in total intravenous anesthesia, IFAC-Papers Online, 58(7): 192-197, https://doi.org/10.1016/j.ifacol.2024.08.033.
7. Wingert, Th., Lee, Ch., & Cannesson, M. (2021). Machine Learning, Deep Learning, and Closed Loop Devices Anesthesia Delivery, Anesthesiol Clin. 39(3): 565–581. https://doi.org/10.1016/j.anclin.2021.03.012.
8. Soltesz, K., Van Heusden, K. & Dumont, G.A. (2020). 5 - Models for control of intravenous anesthesia. In Automated drug delivery in anesthesia, pp. 119-166. Elsevier. https://doi.org/10.1016/B978-0-12-815975-0.00010-2.
9. Saint Aubin, O., Khemir, I., Perdeeau, J.C., Touchard, C., Vallée, F. & Cartailler, J. (2023). Repurposing electroencephalographic signal for automatic segmentation of intra-operative periods under general anesthesia, 20th International Conference on Smart Technologies. Torino, Italy, pp. 286-290, https://doi.org/10.1109/EUROCON56442.2023.10199018.
10. Dutt, M. I. & Saadeh, W. (2023). Monitoring Level of Hypnosis Using Stationary Wavelet Transform and Singular Value Decomposition Entropy with Feedforward Neural Network, in IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31: 1963-1973, https://doi.org/10.1109/TNSRE.2023.3264797.

11. Ghita, M., Birs, I.R., Copot, D., Muresan, C.I., Neckebroek, M. & Ionescu. C.M. (2023). Parametric Modeling and Deep Learning for Enhancing Pain Assessment in Postanesthesia. IEEE Trans Biomed Eng, 70(10):2991-3002. https://doi.org/10.1109/TBME.2023.3274541.

12. Miyaguchi, N., Takeuchi, K., Kashima, H., Morita M. & Morimatsu H. (2021). Predicting anesthetic infusion events using machine learning. Scientific Reports, 11, 23648 https://doi.org/10.1038/s41598-021-03112-2.

13. Liu, Q., Cai, J., Fan, Sh., Abbod, M.F., Shieh, J. & Kung, Y. (2019). Spectrum Analysis of EEG Signals Using CNN to Model Patient's Consciousness Level Based on Anesthesiologists' Experience, in IEEE Access, 7: 53731-53742, https://doi.org/10.1109/ACCESS.2019.2912273.

14. Asai, N., Doi, Ch., Iwai, K., Ideno, S., Seki, H. & Kato, J. (2019). Proposal of Anesthetic Dose Prediction Model to Avoid Post-induction Hypotension Using Electronic Anesthesia Records, 2019 Twelfth International Conference on Mobile Computing and Ubiquitous Network (ICMU), Kathmandu, Nepal, pp. 1-4, https://doi.org/10.23919/ICMU48249.2019.9006672.

15. Rellum, S.R., Schuurmans, J., Van der Ven, W.H. Eberl, S., Driessen, A.H.G., Vlaar, A.P.J. & Veelo. D.P. (2021). Machine learning methods for perioperative anesthetic management in cardiac surgery patients: a scoping review. J Thorac Dis. 13(12): 6976-6993. https://doi.org/10.21037/jtd-21-765.

16. Lee, H.C., Park, Y., Yoon, S.B., Yang, S.M., Park, D. & Jung. C.W. (2022). VitalDB, a high-fidelity multi-parameter vital signs database in surgical patients. Sci Data; 9(1): 279. https://doi.org/10.1038/s41597-022-01411-5.

17. Mashevskiy G. & Dubrovina, P. (2020). System for Predicting the Origin of Neurological Symptomatics in Children after an Ischemic Stroke, IEEE International Conference on Electrical Engineering and Photonics (EExPolytech), St. Petersburg, Russia, 136-139, https://doi.org/10.1109/EExPolytech50912.2020.9243867.

18. Liu, L. Wang, K. Yang, Y. Hu, M. M. Chen, M. Liu, X. Yan, P. Wu, N. & Xiang, X. (2024). Population pharmacokinetic/pharmacodynamic modeling and exposure-response analysis of ciprofol in the induction and maintenance of general anesthesia in patients undergoing elective surgery: A prospective dose optimization study, Journal of Clinical Anesthesia, 92(111317), https://doi.org/10.1016/j.jclinane.2023.111317.

19. Danton C. & Burgart, A. (2020). Machine-Learning Implementation in Clinical Anesthesia: Opportunities and Challenges, Anesthesia & Analgesia 130(6): 1709-1712, https://doi.org/10.1213/ANE.0000000000004656.