Bozok J Eng Arch, 2024; 3(2) 78-86
DOI: 10.70700/bjea.1581791

Bozok
Journal of Engineering and Architecture

e-ISSN: 3023-4298
Research Article

Simulation of infant suction vacuum pressure using WGAN

Fatih Furkan ARSLAN[1] *, Önder DİNCEL[1]

[1] *Yozgat Bozok University, Faculty of Engineering and Architecture, Department of Electrical and Electronics Engineering, Yozgat, Türkiye*

ARTICLE INFO

Article history:

Received
08.11.2024
Accepted
18.11.2024
Published
31.12.2024

Keywords:

Infant Sucking Vacuum
WGAN
Simulation
Sucking Physiology

ABSTRACT

In this study, it is aimed to simulate infant suction vacuums by replicating infant suction vacuum pressures with the Wasserstein Generative Adversarial Network (WGAN) model. To understand the physiology of infants' feeding behaviour, it is very important to understand the physiology of sucking. One of the most important elements of feeding physiology is the infant suction vacuum. In this context, the infant suction vacuum pressure produced by the WGAN model was successfully simulated. Data preprocessing and hyperparameter optimization were performed to ensure the most accurate simulation of the WGAN model used in the study. As a result, the low RMSE and high R2 value between the real data and the generated data show that the model works successfully. This method is promising in terms of better understanding the sucking behaviour of infants and contributing to solving problems in the field of infant feeding. In the future, infant sucking simulation will be more successful if the model is supported with larger data and other elements of infant sucking physiology such as swallowing and respiration are considered. This kind of simulation will pave the way for new applications in the fields of infant feeding and breastfeeding.

WGAN ile bebek emme vakum basıncının simülasyonu

MAKALE BİLGİSİ

Makale Tarihleri:

Geliş tarihi
08.11.2024
Kabul tarihi
18.11.2024
Yayın tarihi
31.12.2024

Anahtar Kelimeler:

Bebek Emme Vakumu
WGAN
Benzetim
Emme Fizyolojisi

ÖZET

Bu çalışmada, bebeklerin emme vakum basınçlarının Wasserstein Generative Adversarial Network (WGAN) modeli ile çoğaltılarak bebek emme vakumlarının benzetiminin yapılması amaçlanmıştır. Bebeklerin beslenme davranışlarını fizyolojisini anlamak için emme fizyolojisini kavramak oldukça önemlidir. Beslenme fizyolojisinin en önemli unsurlarından bir tanesi bebek emme vakumudur. Bu bağlamda, WGAN modelinin ile üretilen bebek emme vakum basıncı başarıyla taklit edilmiştir. Çalışmada kullanılan WGAN modelinin en doğru benzetimi yapabilmesi için veri ön işleme yapılmış ve hiperparametre optimizasyonu yapılmıştır. Sonuç olarak, gerçek verilerle üretilen veriler arasındaki düşük RMSE ve yüksek R2 değeri modelin başarılı çalıştığını göstermektedir. Bu yöntem ile bebeklerin emme davranışlarının daha iyi anlaşılması ve bebek beslenmesi alanındaki sorunların çözülmesine katkı sağlaması açısından umut vadetmektedir. Gelecekte, modelin daha geniş verilerle desteklenmesi ve bebek emme fizyolojisinin diğer unsurları olan yutma ve solunum gibi unsurlarında ele alınmasıyla bebek emme benzetim daha başarılı bir şekilde yapılabilecektir. Bu tür bir benzetim ile bebek beslenmesi ve anne sütü sağımı alanlarında yeni uygulamalara zemin hazırlayacaktır.

1. INTRODUCTION

The feeding behaviour exhibited by infants constitutes a multifaceted activity that needs various aspects of physiologic functions, such as sucking, swallowing, and breathing, to be harmonized [1]. In this regard, a better understanding of the stretch and patterns of infant sucking behaviour is critical in the identification and management of feeding problems and also in the designing of appropriate feeding dynamics [2]. This subject has been the focus of several previous works. For instance, one such study analysed the suction pressure, milk expression, and the period of lactation deprivation in infants [3]. Similar to this study, the pattern of

ORCID ID: Fatih Furkan ARSLAN: 0000-0001-7887-7951, Önder DİNCEL: 0000-0002-5373-8338

***Corresponding author:** Yozgat Bozok Üniversitesi Mühendislik-Mimarlık Fakültesi Elektrik-Elektronik Mühendisliği Bölümü, Yozgat
Phone:+90 354 2421001.
Fax: +90 354 2421005.
E-mail:f.furkan.arslan@yobu.edu.tr

To cite this article: Arslan F.F., Dincel Ö., “Simulation of infant suction vacuum pressure using WGAN”, Bozok Journal of Engineering and Architecture, vol. 3, no. 2, pp. 78-86, 2024.

tongue movement during sucking was assessed using an ultrasound, and it was correlated to the flow of milk and the pressure inside the mouth [4]. Similarly, this study emphasizes the importance of coordination of sucking, swallowing and breathing during the breastfeeding process and argues for more research by pointing out the gaps in existing research and the limitations of clinical management strategies. It also provides an important source of information for care providers by describing the sucking and breathing abilities of infants[5]. Another study analysed normal feeder sucking behaviour in detail, demonstrating how sucking frequency, pressure and consumption rate vary according to the age of the infants. Thanks to the objective measurement devices used in the study, the data obtained provide a basis for the clinical evaluation of abnormal sucking behaviour in preterm infants and other infants at risk of neuromuscular disorders or swallowing difficulties[6]. Newborn sucking patterns have been studied using various techniques using devices that allow for the analysis of sucking in terms of complete and partial closure of the suction system [7]. Another study seeks to analyse and detail the sucking dynamics in preterm breastfeeding infants through the use of advanced techniques, including synchronized ultrasound imaging and intra-oral vacuum measurements. This method aims to offer a more comprehensive understanding of the sucking patterns in preterm infants, and it shows how they may differ from those of term infants[8].

With the objective of ascertaining feeding abilities and evaluating infant health, new devices have been designed to assess non-nutritive sucking by measuring intra-oral and lip expression pressures of infants [9]. Likewise, the sucking-feeding coupling of the infant's non-nutritive sucking response has also been done using specific described devices, which take into account the supply of nutrients as per the performance of the components [9][10]. Similarly, another article introduces a non-nutritive suckling (NNS) system that measures intraoral vacuum profiles in real time in newborns, providing an analysis of objective sucking characteristics. By presenting quantitative data of sucking behaviors, the study provides a framework for future research that will contribute to the assessment of oromotor dysfunctions[11]. Ultrasound imaging was also employed to assess the dynamics of sucking of infants during breastfeeding, showing the presence of nutrition and non-nutrition sucking [12]. In another study, he developed an innovative tool that simulates breastfeeding mechanics in a laboratory setting, providing reliable data and contributing to the understanding of breastfeeding dynamics with measurements consistent with clinical data. This study has elucidated breastfeeding biomechanics with advanced measurement techniques and laid the foundation for the development of future breastfeeding devices[13].

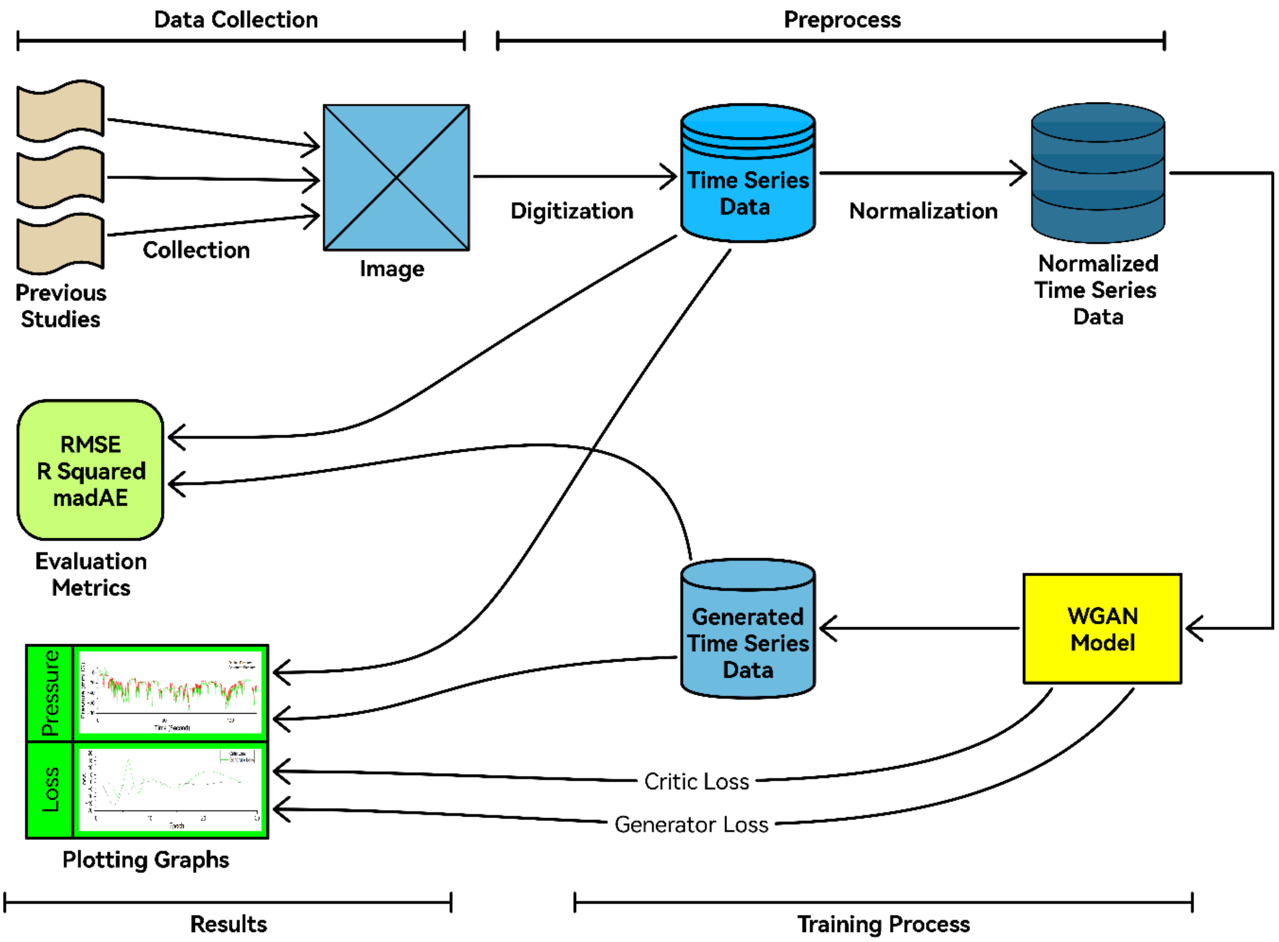

Figure 1. Flow chart of the process.

Our goal is to make the mechanical baby sucking simulation mentioned above with artificial intelligence a more natural and realistic milking process for the mother. To achieve this a novel approach to mimicking the suction derived from an infant's mouth using a Wasserstein Generative Adversarial Network (WGAN). In generative modelling techniques, the WGAN is viewed as a sophisticated one since it can comprehend and internalize the data characteristics of suction vacuum pressure samples obtained from infants. In contrast to traditional GAN frameworks, the WGAN paradigm incorporates the Wasserstein distance as a cost, allowing for more consistent output and broader usability [14][15].

To this end, the accrual of data on infant suction vacuum pressure from existing literature and data washing will be discussed first. Next, we present the WGAN model and its implementation, the training process, and the model evaluation metrics. Eventually, the findings of the simulation are presented along with the generated suction vacuum pressure data, the performance of the model with regard to generated data and actual data, and the trends of generator and critic loss over the training iterations, as illustrated in Figure (3). The flowchart of the whole process is given in Figure (1).

2. MATERIALS and METHODS

2.1. Data Collection and Preprocessing

This study utilized data stemming from the lactational practice of nutritional sucking. This included time series data on the suction vacuum pressure of babies during suckling. The data is a mean vacuum pressure plot concerning time for six mother-infant pairs [7]. The following measures were taken during data processing;

- OpenCV transforms the image from a graphics form to a greyscale [16].
- The plot data points are then extracted from the greyscale image after binarization is performed using a threshold value [16].
- The subsequent step involves digitizing the obtained data points to numerically represent the suction vacuum pressure and the related time intervals.
- In addition, the data is further adjusted for use in the WGAN training purpose.

2.2. WGAN

According to [17], two neural networks make up Generative Adversarial Networks (GANs) – Generator (G) and Discriminator (D). The generator, G, accepts an input of random noise, which enables it to learn the features of actual data samples. Conversely, the discriminator, D, takes as an input a mixture of actual data as well as generated data and attempts to tell the difference between the two and to indicate what data is actual and what is generated out of G. The loss functions that enable G and D to operate within the framework of GANs are given in the Equation (1) below:

$$\begin{cases} Loss_G = E_{z\subset P_z}\left[D\big(G(z)\big)\right] \\ Loss_D = -E_{x\subset P_x}[D(x)] + E_{z\subset P_z}\left[D\big(G(z)\big)\right] \end{cases} \quad (1)$$

In this setting, it is the case that $Loss_G$ and $Loss_D$ refer to the loss function in turns for the Generator and the Discriminator, respectively. The purpose of the generator is to generate samples $G(z)$ that when z is in the latent space, closely approximates the image data x. On the contrary, the task of the discriminator is to tell the difference between the actual data x and the generated image data $G(z)$.

In the case of WGAN, the similarity between two distributions is quantified using the concept of Wasserstein distance. According to [15], in contrast with Jensen-Shannon and Kullback-Leibler divergence measurements, the Wasserstein distance has a continuous and valuable gradient. This property helps to alleviate the mode collapse problem that is usually associated with GAN training relying on JS divergence. The Wasserstein distance is represented in the Equation (2) below:

$$W(P_x, P_z) = sup_{L\leq 1}\left(E_{x\subset P_x}[D(x)] - E_{z\subset P_z}\left[D\big(G(z)\big)\right]\right) \quad (2)$$

In this context, the notation $W(P_x, P_z)$ is reserved for the Wasserstein distance between two probability distributions. The first distribution P_x refers to the true data distribution while the second distribution P_z denotes the distribution of the data generated in the model. The constant L (where $L \leq 1$) acts as a bound on the difference in expectation that can be achieved under the 1-Lipschitz assumption to guarantee safe gradient optimization in the network.

2.3. Model Components

Two networks make up the WGAN model, which competes in two components – the generator and the critic:

The Generator tries to produce samples that correspond closely with the actual ones from a given random vector. During training, the generator, which is built as a stack of fully connected layers, receives not only a latent vector (random noise) but also an extra Gaussian-distributed noise, as temporal information. As a result, the model is capable of generating data samples that are reasonably and temporally dynamic in nature, changing with time. The first layer inputs the latent dimension. The following layers consist of 512, 1024, and 2048 neurons, thus endowing the model with the ability to comprehend and produce intricate patterns. The last layer presents the sequence length as an output, the same as the arrangement of the target data. Each layer uses the Leaky ReLU activation function [18], which assists in avoiding information loss in negative ranges, leading to a more realistic distribution. The output layer employs Tanh [19] which constricts the output values to the range of [-1, 1]. Furthermore, the layers employ dropout regularization to mitigate overfitting, with a dropout rate corresponding to the fixed value named dropout rate, enhancing generalization and performance [20].

The Critic is a model that is capable of telling the difference between actual data and fake data. In contrast to a Discriminator in the classical sense, the function of the loss for the Critic is more straightforward and, therefore, leads to more equal and stable training of the model. The Critic network includes four layers of fully connected networks: the first one has 1024 neurons, the second 512, the third 256, and the last one produces a single scalar output with one neuron. Input data is sequential. Therefore, input tensors are unrolled in a manner of a view, preferably before the Critic. The first three layers are composed of ReLU activation [21]and a wok returns a single scalar. ReLU function is utilized due to its positive range linearity, which minimizes information loss and fastens training. As illustrated in Figure (2), the model architecture is presented in detail.

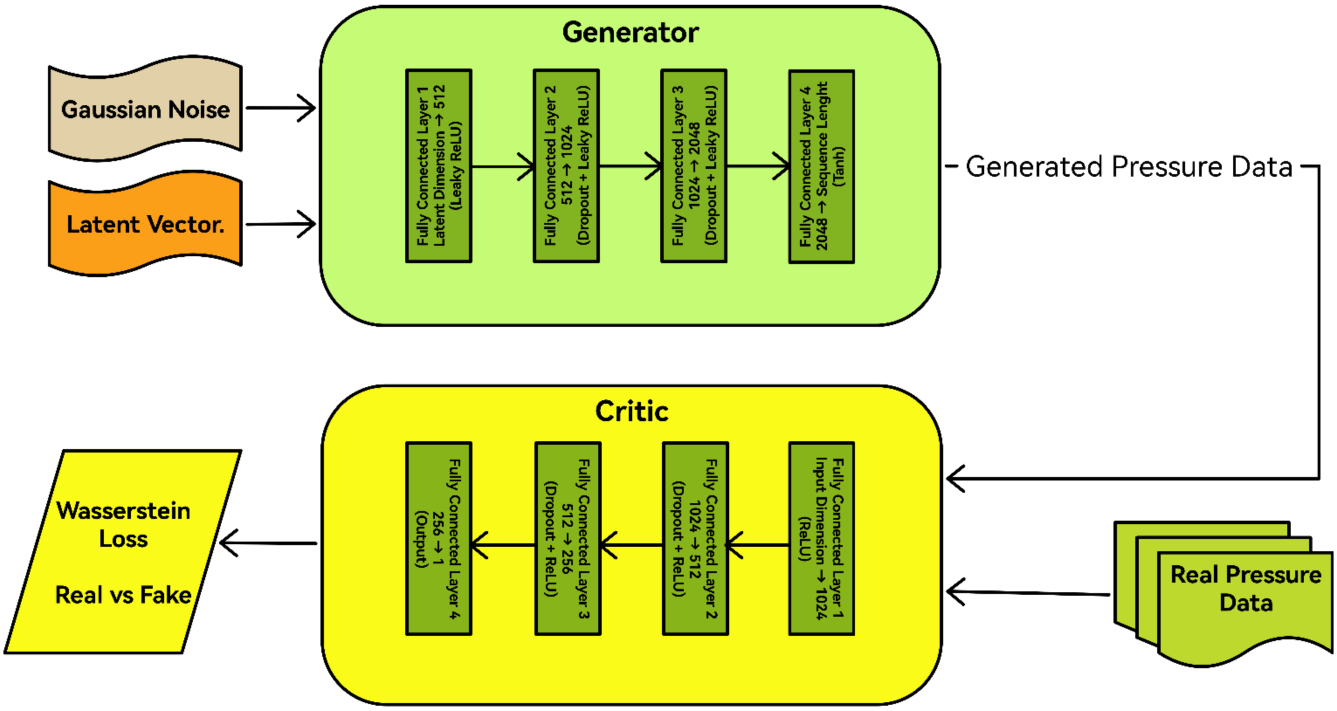

Figure 2. Block diagram of the model.

2.4. Training Process and Hyperparameters

The process of training the WGAN model was methodically structured so that learning could proceed in a stable manner, and the quality of the generated data was good. By emphasizing the adjustment of hyperparameters and the use of special training regimes, the model managed to perform well and precisely reproduced realistic patterns of baby-sucking vacuum.

The design of the model and the settings of the hyperparameters were customized to suit the baby-sucking vacuum data. Some of the primary hyperparameters were as follows:

- Latent Dimension: The latent dimension, which is fixed at 200, indicates the length of the noise vector fed into the Generator. This dimension was chosen in order to strike a balance on the degree of complexity required to model variations in the data while ensuring the computational costs are kept at a reasonable level.
- Sequence Length: Set to 992, the sequence length is half the length of the complete data set. This technique helps the model’s ability to create plausible incomplete sequences, which, when put together, form a coherent whole of sucking vacuum data.
- Batch Size: The batch size of 128 was employed as it enhances the stability of the gradients and helps in faster convergence. The application of multiple samples at once allowed for effective parameter updates.
- Critic Training Frequency: The Critic was subjected to 10 training sessions for every Generator update, which was more effective in learning the nuanced textures of the real and fake samples. This update frequency allows for better feedback from the generator and, therefore, improves the quality and convincement of the generated data.
- Weight Clipping: A clipping threshold of 0.01 was imposed on the weights of the Critic, which is useful in restraining the excessive growth of weights as well as stabilizing the Wasserstein loss, which is important for training purposes.
- Weight Decay: Given a value of 0.0001, weight decay is one of the techniques used to prevent overfitting by discouraging large weights in the Generator and Critic networks, thereby encouraging less complex and more general learning.
- Patience for Early Stopping: We found it suitable to apply a patience setting of 15 epochs in early stopping in order to cease the training whenever no improvement was recorded in the loss of the Generator in order to prevent overfitting.
- Rate of Dropout: In both models, a dropout rate of 0.3 was used, reducing overfitting through random shutdown of a certain percentage of neurons during each training step.
- MSE Threshold: An MSE threshold of 25 was established as the acceptable level for the fidelity of generated data. Training would be stopped if this level was achieved or surpassed. They were indicating that the model was able to reproduce the real data closely.

To enhance both the efficiency and stability of the training, early stopping and an adaptive learning rate scheduler were also incorporated into the training regime. Early stopping was applied, considering there is no improvement in the Generator’s loss for 15 epochs, thus minimizing the unnecessary computational burden and overfitting windows. Furthermore, since the performance was observed to have plateaued, the learning rate was tailored down by a ReduceLROnPlateau. In cases of aggressive training, when performance did not improve, the learning rate was decreased by a factor of 0.5. This ensured that the model avoided retraining in the local minima, enabling optimally updating parameters during the model's convergence.

The training regime was implemented to provide progressive and focused updates to both the generator and the critic. The Critic received ten updates for each update of the Generator, allowing enough training to learn the difference between actual samples and generated samples. In every Practicum update, real and generated samples were tested and paired to measure the Wasserstein loss metric, corresponding to the difference between actual and generated samples in mean scores the Practicum gave. This way, the Critic learned to approximate the Wasserstein distance between two distributions, producing better gradients for optimizing the generator.

The generator was then updated by decreasing the Critic’s score for the artificial samples. This enhanced the Generator's ability to produce realistic-looking data by shrinking the Wasserstein distance between the actual images and their generated counterparts. This sort of feedback loop enabled the model to create more and more realistic pictures over time.

2.5. Evaluation Metrics

The Critic and Generator losses were supervised throughout the training process. These losses also illuminated the convergence and balance of the model. Moreover, extreme differences between the Critique and Generator losses were considered signs of instability, and thus, such parameters and training strategies were adjusted accordingly, as depicted in Figure (4). Furthermore, evaluation of the quality of data generated through the network against the original data also took place at the end of every epoch, focusing on metrics such as the mean squared error (MSE), the coefficient of determination (R^2), and mean absolute deviation (madAE). The value of MSE, which was set at the threshold of 25, was considered the target objective. When this cue was reached, training was stopped, and indicating that the model had produced realistic data. Through this approach and with the appropriate selection of the hyperparameters, the WGAN model was able to create realistic sequential data relevant to the simulation of baby-sucking pressure patterns.

3. RESULTS and DISCUSSION

Figure (3) shows the infant sucking vacuum pressure data in both raw and filtered forms for a duration of 120 seconds. The red line illustrates the original pressure data, while the data after filtering is presented as a green line. A close observation of the image shows that the filtered line is very close to the initial one while preserving the significant shifts and minimizing the noise; thus, it proves the adequacy of the filtering method employed.

Throughout the period, the pressure exhibits oscillation-like behavior, with sharp increases and decreases in values, typical of an infant sucking. This compatibility of the initial pressure vacuum signal with the filtered one illustrates the efficiency of the filtered data in replicating the pressure vacuum profile. Such precision is essential in ensuring that accurate models are developed as well as analyzing infant sucking patterns, which is one of the objectives of this research concerning the representation of data.

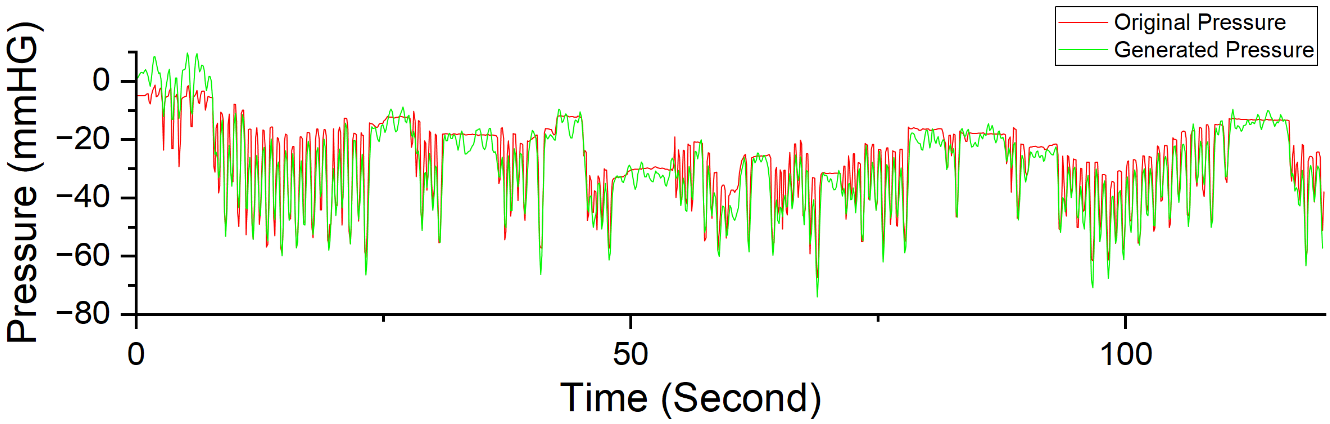

Figure 3. Original and generated sucking vacuum pressure.

Figure (4) presents the losses for both the generator and the critic throughout 30 epochs. In this context, the generator loss is illustrated with a green line, while the critical loss is shown in gray. In the first few epochs, both losses are exceptionally high and vary greatly, suggesting that the model is still in the process of 'fine-tuning' itself to the data.

With increased training, these fluctuations also decrease, and the losses begin to steady out, indicating that the training is reaching its airport. The trends of the loss curve for the generator in the later epochs show relatively smoother trends than in the earlier cycles, meaning that the device is progressively learning how to generate target data. On the other hand, the losses of the critic show the same trend; hence, the output of his network is adequate enough for the generator. This relationship between the generator and the critic plays a vital role in the WGAN replicating the infant-sucking vacuum pattern.

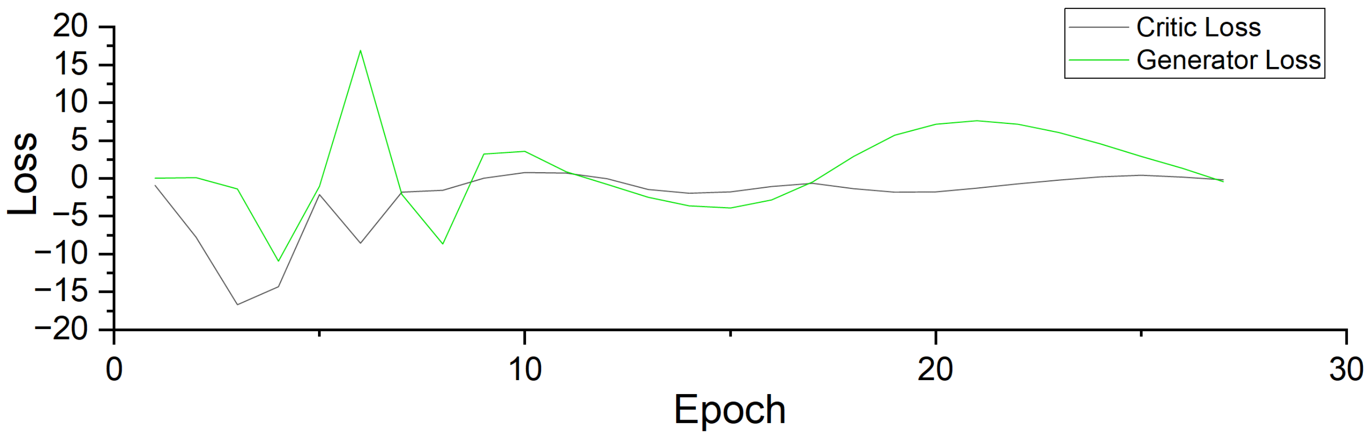

Figure 4. Generated and critic loss per epoch.

The evaluation metrics for the model's performance in simulating infant sucking vacuum pressure are illustrated in Table 1:

- Root Mean Squared Error (RMSE): 4.86. This is the model's mean square error for pressure values. The low MSE suggests that the model accurately captures the pressure variations without significant error.
- R – Squared Score (R^2): 0.93, which implies that the application of the model accounts for 93% of the variability in the data. This very high value of R^2 indicates that the predicted values from the model are closely aligned with the observed values from the experiment.
- Mean Absolute Deviation of Absolute Error (madAE):3.92, which provides an average absolute error deviation close to 4. This low madAE value is advantageous because it shows how close the model's predictions are to the actual values, proving its credibility.

Based on these parameters, the model delivers high accuracy in reproducing beating patterns of specific vacuum pressures relevant to infant-sucking simulations.

Table 1. Performance Metrics

RMSE	**R^2**	**madAE**
4.86	0.93	3.92

4. CONCLUSIONS

In this research, we illustrated how a Wasserstein Generative Adversarial Network (WGAN) was successfully employed in synthesizing infant suction vacuum pressure, a multi-dimensional complex behaviour associated with an infant's feeding ability. The WGAN paradigm provides a novel means of creating consistent time series data on the weights of physiologically relevant images by searching for the rhythm and oscillations seen during the pressure changes in sucking. The beauty of this model is not just limited to how well the real data features are reproduced but also includes the ability to create realistic-looking samples of the infant sucking that is evidenced by low MSE and high R2 values.

In this case, the main benefit of using the WGAN model comes from the fact that it does not suffer from the weaknesses of ordinary GANs, especially mode collapse, because of employing the Wasserstein metric as a loss. This is very important for generating physiological data as there is a lot of variation that should be preserved within each generated sequence. The stable and consistent performance of the model is evidenced by the convergence behaviours of the generator and critic loss trends, which suggests that training was successful and there was meaningful interaction between the generator and critic networks. Therefore, the WGAN model is useful for appropriately simulating physiology processes that change over time, such as the infant sucking vacuum pressure.

The ability of the model to produce pressure data corresponding to infant feeding could offer considerable advantages in clinical and developmental studies. This method, for example, could be useful in designing devices for non-invasive observation of a subject's feeding activity and diagnosing abnormalities in this activity. As the model continues to be perfected, this data could become the gold standard for what normal sucking looks like versus deviations from this normality, which is helpful in diagnosing feeding problems and is useful for clinicians.

Even so, it is important to highlight some areas for improvement in this research. The corpus of data employed for the model's training had a relatively low sample size, which may affect the ability to generalize the findings to a bigger sample of infants. Furthermore, the model considers only the suction pressure and neglects to assess vital parameters of feeding behaviour, such as swallowing or breathing, which all take place during the process of feeding. Future work seeking to build on this analysis should focus on improving the model by enriching its database and using different types of data.

To sum up, the current research highlights the efficacy of WGAN in creating artificial time-sequenced data for various physiological simulations focusing on infant feeding. This method provides a basis for reliable and cost-effective modelling of vacuum suction pressure, which positively impacts neonatal healthcare by introducing new ways of understanding and treating infant-feeding dynamics. In the long run, such a WGAN model could embody significant advancements in the knowledge of heightened interest in the feeding capabilities of infants and the improvement of the health systems for children.

LIMITATION

The availability of infant suction data for this study is perhaps the most limiting factor. This is primarily because the process of feeding a mother and her child is highly sensitive therefore data collecting becomes difficult. In addition to this, the field has little interest in the past research, which means there are fewer datasets that have been made available to the public or recorded scientifically. Generalizability of the findings is hindered by insufficient data and present simulation has limited realism. To overcome this problem, it will be necessary for the current research to find new, and responsible ways of gathering fuller information on how infants feed, possibly by observing but not disturbing those being studied.

AUTHOR CONTRIBUTIONS

The authors contributed equally at every stage of the article.

CONFLICT OF INTEREST

There is no conflict of interest.

ETHICS

There is no ethical problem in the publication of this article.

ACKNOWLEDGMENT

A significant part of this paper includes the Master Thesis data of Fatih Furkan ARSLAN.

REFERENCES

[1] M. Martell, G. Martínez, M. González, and J. L. D. Rosselló, “Suction patterns in preterm infants,” *J Perinat Med*, vol. 21, no. 5, pp. 363–369, 1993, doi: https://doi.org/10.1515/jpme.1993.21.5.363.

[2] A. M. Cannon, V. S. Sakalidis, C. T. Lai, S. L. Perrella, and D. T. Geddes, “Vacuum characteristics of the sucking cycle and relationships with milk removal from the breast in term infants,” *Early Hum Dev*, vol. 96, pp. 1–6, 2016, doi: https://doi.org/10.1016/j.earlhumdev.2016.02.003.

[3] C. Prieto, H. Cardenas, A. Salvatierra, C. Boza, C. Montes, and H. Croxatto, “Sucking pressure and its relationship to milk transfer during breastfeeding in humans,” *Bioscientifica Ltd*, vol. 108, no. 1, pp. 69–74, Nov. 1996, doi: https://doi.org/10.1530/jrf.0.1080069.

[4] D. T. Geddes, J. C. Kent, L. R. Mitoulas, and P. E. Hartmann, “Tongue movement and intra-oral vacuum in breastfeeding infants,” *Early Hum Dev*, vol. 84, no. 7, pp. 471–477, 2008, doi: https://doi.org/10.1016/j.earlhumdev.2007.12.008.

[5] D. T. Geddes and V. S. Sakalidis, “Breastfeeding: how do they do it? Infant sucking, swallowing and breathing,” *Infant*, vol. 11, no. 5, pp. 146–150, 2015.

[6] J. S. McGowan, R. R. Marsh, S. M. Fowler, S. E. Levy, and V. A. Stallings, “DEVELOPMENTAL PATTERNS OF NORMAL NUTRITIVE SUCKING IN INFANTS,” *Dev Med Child Neurol*, vol. 33, no. 10, pp. 891–897, 1991, doi: https://doi.org/10.1111/j.1469-8749.1991.tb14798.x.

[7] L. Chen, R. F. Lucas, and B. Feng, “A Novel System to Measure Infants’ Nutritive Sucking During Breastfeeding: the Breastfeeding Diagnostic Device (BDD),” *IEEE J Transl Eng Health Med*, vol. 6, pp. 1–8, 2018, doi: 10.1109/JTEHM.2018.2838139.

[8] D. T. Geddes, K. Chooi, K. Nancarrow, A. R. Hepworth, H. Gardner, and K. Simmer, “Characterisation of sucking dynamics of breastfeeding preterm infants: a cross sectional study,” *BMC Pregnancy Childbirth*, vol. 17, no. 1, Nov. 2017, doi: https://doi.org/10.1186/s12884-017-1574-3.

[9] S. Akbarzadeh *et al.*, “Evaluation of Apgar Scores and Non-Nutritive Sucking Skills in Infants Using a Novel Sensitized Non-Nutritive Sucking System,” *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pp. 4282–4285, 2020, doi: 10.1109/EMBC44109.2020.9176146.

[10] A. J. Sameroff, “An apparatus for recording sucking and controlling feeding in the first days of life,” *Psychon Sci*, vol. 2, no. 1, pp. 355–356, 1965.

[11] P. Truong, E. Walsh, V. P. Scott, T. Coleman, G. Tilvawala, and J. Friend, “Non-Nutritive Suckling System for Real-Time Characterization of Intraoral Vacuum Profile in Full Term Neonates,” *IEEE J Transl Eng Health Med*, vol. 11, pp. 107–115, 2023, doi: 10.1109/JTEHM.2022.3231788.

[12] V. S. Sakalidis *et al.*, “Ultrasound Imaging of Infant Sucking Dynamics during the Establishment of Lactation,” *Journal of Human Lactation*, vol. 29, no. 2, pp. 205–213, 2013, doi: 10.1177/0890334412452933.

[13] L. Jiang and F. Hassanipour, “Bio-Inspired Breastfeeding Simulator (BIBS): A Tool for Studying the Infant Feeding Mechanism,” *IEEE Trans Biomed Eng*, vol. 67, no. 11, pp. 3242–3252, 2020, doi: 10.1109/TBME.2020.2980545.

[14] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, “Improved Training of Wasserstein GANs,” in *Advances in Neural Information Processing Systems*, I. Guyon, U. Von Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/892c3b1c6dccd52936e27cbd0ff683d6-Paper.pdf

[15] M. Arjovsky, S. Chintala, and L. Bottou, “Wasserstein Generative Adversarial Networks,” in *Proceedings of the 34th International Conference on Machine Learning*, D. Precup and Y. W. Teh, Eds., in Proceedings of Machine Learning Research, vol. 70. PMLR, Nov. 2017, pp. 214–223. [Online]. Available: https://proceedings.mlr.press/v70/arjovsky17a.html

[16] S. Brahmbhatt, *Practical OpenCV*, no. 2. Apress, 2013.

[17] I. Goodfellow *et al.*, “Generative adversarial networks,” *Commun ACM*, vol. 63, no. 11, pp. 139–144, 2020.

[18] J. Xu, Z. Li, B. Du, M. Zhang, and J. Liu, “Reluplex made more practical: Leaky ReLU,” in *2020 IEEE Symposium on Computers and Communications (ISCC)*, 2020, pp. 1–7. doi: 10.1109/ISCC50000.2020.9219587.

[19] M. M. Lau and K. Hann Lim, “Review of Adaptive Activation Function in Deep Neural Network,” in *2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)*, 2018, pp. 686–690. doi: 10.1109/IECBES.2018.8626714.

[20] P. Baldi and P. J. Sadowski, “Understanding Dropout,” in *Advances in Neural Information Processing Systems*, C. J. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, Eds., Curran Associates, Inc., 2013. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2013/file/71f6278d140af599e06ad9bf1ba03cb0-Paper.pdf

[21] A. D. Rasamoelina, F. Adjailia, and P. Sinčák, “A Review of Activation Function for Artificial Neural Network,” in *2020 IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI)*, 2020, pp. 281–286. doi: 10.1109/SAMI48414.2020.9108717.