Cardiotocography Modeling via Transfer Entropy Bottleneck to Predict Intrapartum Fetal Deterioration

Mahdi Shamsi[1,4], Michael Kuzniewicz[2], Cornet Marie-Coralie[3], Yvonne Wu[3], Lawrence Gerstley[2], Robert Kearney[1], Philip Warrick[1,4]

[1] McGill University, Montreal, Canada
[2] Kaiser Permanente Northern California, Oakland, USA
[3] University of California, San Francisco, USA
[4] PeriGen Inc., Montreal, Canada

Abstract

Fetal acidosis is a key clinical precursor to hypoxic-ischemic encephalopathy (HIE) and an important marker of impaired fetal oxygenation during labor. Fetal monitoring via cardiotocography (CTG), which records uterine pressure (UP) and fetal heart rate (FHR), is a standard of care, yet its interpretation is subjective. In this work, we propose the Sequential Variational Autoencoder with Transfer Entropy Bottleneck (SeqVAE-TEB), a novel VAE model designed to capture the directed and lagged influence of UP contractions on subsequent FHR responses. SeqVAE-TEB is trained in a predictive setting, forecasting future FHR to emphasize clinically meaningful coupling between UP and FHR. For the classification task, we train the model on acidosis cases and healthy cases with blood gas, and test and evaluate its generalization on unseen HIE cases and other healthy subgroups. Our model maintains robust specificity across different healthy subgroups and provides early sensitivity in detecting acidosis and cases at risk of developing HIE.

1. Introduction

Hypoxic ischemic encephalopathy (HIE) is a severe neonatal condition caused by insufficient oxygen and blood flow to the infant's brain during the perinatal period. It is one of the leading cause of neonatal mortality and long term neurodevelopmental impairment. A key precursor to HIE is fetal acidosis, defined by low umbilical cord pH or high base deficit, which directly reflects impaired fetal oxygenation and metabolic stress [1]. Acidosis is more common than HIE and provides a clinically meaningful window for early detection and intervention before neurological injury becomes irreversible. HIE, in turn, can be viewed as the most severe progression of acidosis, associated with more profound hypoxia and poorer outcomes. As such, accurate and early detection of acidosis is very important in intrapartum monitoring. [2]. Cardiotocography (CTG), which records maternal uterine pressure (UP) and fetal heart rate (FHR), is the standard non-invasive method for intrapartum fetal surveillance. Certain FHR patterns have been associated with hypoxia in clinical guidelines, particularly late or prolonged decelerations; bradycardia or tachycardia; and reduced variability [3]. However, the visual interpretation of CTG remains subjective and inconsistent, with poor inter- and intra-observer agreement [1].

Machine learning approaches have recently emerged as promising tool for analyzing CTG signals. These methods can identify subtle temporal patterns and physiologic relationships that are difficult to detect visually. Among them, unsupervised representation learning has shown particular utility in modeling biomedical time series, enabling the extraction of compact, informative features for downstream classification and risk prediction tasks. Variational Autoencoders (VAEs) offer a principled probabilistic framework for such representation learning [4]. Despite these advances, most prior work has primarily focused on the FHR signal, while clinical interpretation of CTG routinely incorporates both UP and FHR. Uterine contractions are frequently followed, with a physiological delay, by FHR decelerations. This directed and lagged relationship is clinically important, as it reflects the fetus's adaptive response to maternal uterine activity and provides an early indicator of hypoxic stress. To explicitly model this coupling, we leverage the concept of transfer entropy, an information-theoretic measure that quantifies directed information flow between time series [5].

In this paper, we propose the Sequential Variational Autoencoder with Transfer Entropy Bottleneck (SeqVAE-TEB), a novel framework that integrates the predictive power of VAEs with an explicit bottleneck designed to capture the directed influence of UP on FHR. The Transfer Entropy Bottleneck (TEB) is an information theoretic ex-

tension of the variational information bottleneck that explicitly constrains the latent space to retain only the predictive information flowing from a source process to a target process [6] Unlike traditional reconstruction-based approaches, our model is trained in a predictive setting, focusing on forecasting FHR conditioned on uterine activity. This predictive formulation emphasizes the causal role of uterine contractions in shaping FHR dynamics, aligns with clinical practice. Furthermore, we evaluate model generalization on HIE cases as the most severe manifestation of acidosis, providing additional validation of our approach.

2. Methods

In our prior work [7], we introduced the Sequential Variational Autoencoder (SeqVAE), a probabilistic deep learning framework inspired by variational recurrent neural networks (VRNNs), to model FHR signals for the early detection of HIE. The SeqVAE was designed to capture temporal dynamics while encouraging smoother and more informative latent representations. We demonstrated that SeqVAE produced physiologically meaningful latent features and showed promising discriminative performance in distinguishing between healthy and HIE outcomes. In this paper, we have considered the UP signal as a second source of information to make the latent space more informative. Moreover, to emphasize UP-FHR coupling during pretraining, we used a prediction task rather than reconstruction. We used a large-scale cardiotocography (CTG) dataset consisting of over 250,000 singleton births sampled at $4\,\mathrm{Hz}$ [8]. We defined two groups of clinical outcomes: (i) healthy deliveries with no neurological symptoms, and (ii) adverse outcomes including acidosis (umbilical artery pH< 7.0 or base deficit$\geq 10\,\mathrm{mmol\,L^{-1}}$) and HIE which is defined as acidosis plus neurological signs. It is important to note that acidosis is identified exclusively through cord blood gas measurements, which are generally only available after delivery. Thus, while acidosis provides the most direct evidence of impaired fetal oxygenation, it cannot be observed prospectively during labor and serves here as a post hoc marker for evaluating model performance. For model pre-training, we selected 20,968 healthy vaginal deliveries with blood gas. For classification, we trained and tested with data not seen during pre-training. We used 2,988 healthy vaginal deliveries with blood gas and 1,981 acidosis vaginal deliveries for training, and reserved for testing all 391 HIE cesarean section (CS) and vaginal deliveries and 1217 acidosis cases with CS. Moreover, we performed auxiliary tests on several clinically relevant subclasses that were not included during training as well. These included healthy deliveries without blood gas (500 vaginal and 500 CS records), and 500 healthy deliveries with blood gas but delivered via CS. The no blood gas subgroup represents the majority of deliveries in the full dataset. Consequently, model performance on this group has substantial clinical impact specifically reduced false positive rates in this set translate to fewer unnecessary interventions across a large number of deliveries. These analyses assessed the generalization capability of the SeqVAE-TEB model beyond the training distribution and provided insight into its robustness across different clinical scenarios. The UP and FHR signals were segmented into $20\,\mathrm{min}$ epochs, each corresponding to 4,800 samples.

Preprocessing Step: In TEB context, the UP signal is regarded as the source time series denoted as X, while the FHR signal is treated as the target time series denoted as Y. We applied the wavelet scattering transform to both UP and FHR signals. The scattering transform extracts multiscale, translation-invariant representations of the signals while preserving physiologically relevant frequency dynamics. The resulting scattering coefficients were subsampled by a factor of 16, which reduces redundancy and improves computational efficiency while retaining relevant temporal structures. These transformed time series were used as input to the SeqVAE-TEB model.

SeqVAE-TEB Model: The proposed Sequential Variational Autoencoder with Transfer Entropy Bottleneck (SeqVAE-TEB) extends our earlier SeqVAE framework by reformulating the learning objective from signal reconstruction to time series prediction. Unlike reconstruction, prediction presents a substantially more difficult challenge; the model must capture the underlying temporal dynamics and anticipate delayed dependencies between the source and target signals. The prediction task is defined as forecasting the next two minutes of FHR activity conditioned on the UP and FHR histories. To address this task, SeqVAE-TEB integrates the TEB framework, which explicitly quantifies the directed influence of the source signal (UP) on the target signal (FHR). The encoder, $q_\phi(z \mid X_{\text{past}}, Y_{\text{past}})$, learns a latent distribution z conditioned jointly on the source and target histories, while the conditional prior, $r_\psi(z \mid Y_{\text{past}})$, captures the distribution conditioned only on the target history. The Kullback–Leibler (KL) divergence between these distributions constrains the latent space to retain only the additional predictive information contributed by the source, therefore serving as a variational upper bound on transfer entropy. The decoder, $p_\theta(Y_{\text{future}} \mid z, Y_{\text{past}})$, then predicts the future trajectory of the FHR signal from the latent representation.

Training Objective: The model is trained by optimizing a loss function that balances prediction fidelity with the information bottleneck constraint:

$$\mathcal{L} = \mathbb{E}_{q_\phi}\left[-\log_{p_\theta}(Y_{\text{future}} \mid z, Y_{\text{past}})\right] + \beta\,\mathrm{KL}\big(q_\phi(z \mid X_{\text{past}}, Y_{\text{past}}) \parallel r_\psi(z \mid Y_{\text{past}})\big), \quad (1)$$

where the first term encourages accurate prediction of the

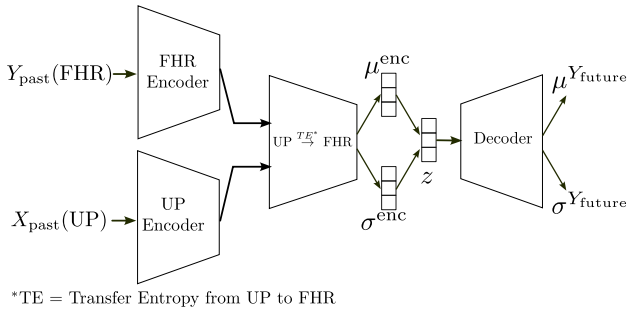

Figure 1: SeqVAE-TEB model block diagram.

future FHR signal, while the second term enforces the TEB constraint by ensuring that the latent space captures only the minimal predictive information from UP about FHR based on the minimum necessary information principle [6]. The hyperparameter β governs the trade-off between predictive accuracy and the TEB regularization term and regulates the strength of the transfer entropy bottleneck. A large β can overly suppress source (UP) to target (FHR) information and discard useful UP contributions, while a small β may fail to enforce the minimum necessary information (MNI) principle and allow irrelevant or redundant UP features to leak into the latent space. Figure 1 illustrates the block diagram of SeqVAE-TEB model.

Classification Extension: We use a LSTM-based classification head for the task of distinguishing healthy and abnormal cases. Specifically, the model leverages the learned TEB-informed latent representations, which capture both the directed dependencies of UP on FHR and the FHR characteristics that are not predictable from UP. After pre-training on the prediction objective, we froze the VAE backbone, and trained the lightweight multi-layer LSTM head. This two-stage training strategy ensures that the discriminative task benefits from the physiologically meaningful latent features learned during prediction and benefits from large healthy datasets to pre-train the mode.

Latent Representation and Classification Protocol: After pretraining SeqVAE-TEB in the predictive setting, we froze the model and extracted latent representations for each epoch independently. These latent vectors were then passed to the classification head, which outputs the probability of the epoch being healthy vs. acidosis. For each record, the predictions were aggregated between epochs as follows: if an epoch was classified as acidosis, all subsequent epochs up to delivery were marked as acidosis. In contrast, a record was considered healthy only if no epoch before delivery was flagged as acidosis. Such aggregation reflects the clinical principle that once fetal compromise is detected, the pregnancy is considered at risk until delivery. To control the false positive rate (FPR), we adopted a validation-based calibration strategy. Specifically, we tuned the decision threshold on the validation set to achieve a desired FPR within one hour of delivery. This calibrated threshold was then applied to the test set, and all reported figures and performance metrics are based on the test evaluation. While the CS rate in the United States is approximately 32% based on the National Center for Health Statistics (NCHS), we adopted a more restrictive operating point and set the allowable FPR to 0.2 in order to minimize unnecessary interventions while maintaining clinically useful sensitivity.

3. Results and Discussion

For robust evaluation, we performed 10-fold cross-validation at the record level, with 90% training, 5% validation, and 5% testing partitions. We restricted the pre-training of SeqVAE-TEB to the last 12 hours prior to birth, enabling the model to learn predictive representations when clinically relevant intrapartum dynamics should be most present. For training the classifier, only the last 6 hours before delivery were used to minimize the effect of weak labeling. The labels are weak since initially healthy fetuses may deteriorate over time and transition to abnormal states, even though the entire record is labeled according to the final outcome. During evaluation, all test results are reported on the last 12 hours prior to delivery.

Auxiliary Subclass Analysis In addition to the main classification task, Figure 2a reports specificity stratified by the blood-gas criteria, while Figure 2b illustrates the specificity across all healthy records when stratified by CS status. In all figures, the solid curves represent the average specificity over 10-fold cross-validation, and the shaded regions denote the range between the minimum and maximum values observed across folds. The results demonstrate that the model achieves comparable performance across all subgroups. Importantly, the specificity reduces gradually to near 0.8 during the final hour before delivery. Figure 2c presents the specificity for healthy records without blood gas, stratified by CS status. The model demonstrates consistently slightly better performance between 8 and 2 hours before delivery, particularly for the no CS subgroup. However, both subgroups converge to the desired operating point, with specificity near 0.8 during the final hour before birth. Figures 3a and 3b present the sensitivity results for acidosis and HIE subgroups, stratified by CS status. The model was trained exclusively on acidosis cases with vaginal deliveries; acidosis with CS and HIE cases were excluded during training. Despite this, the sensitivity profiles across subgroups show broadly similar trends. For acidosis without CS, the sensitivity was near 0.4 in the final hour before delivery, while for acidosis with CS and HIE cases, it was between 0.3 and 0.4. Notably, the HIE subgroup exhibits narrower confidence bounds across the 10 folds, suggesting more consistent

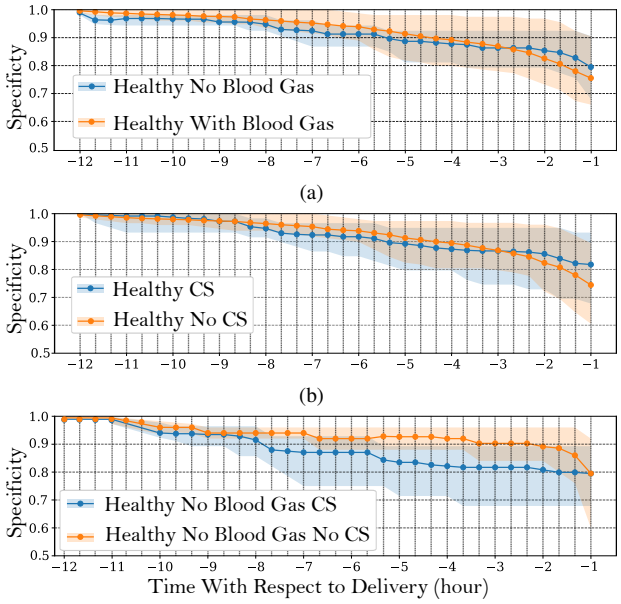

Figure 2: Test specificity for healthy records (a) stratified by blood-gas criteria (b) stratified by CS criteria and (c) no blood gas alone, stratified by CS criteria.

performance. This is particularly noteworthy for the HIE no CS subgroup, as these cases were by definition missed by clinicians during labour, which indicates the potential advantage of the model in enabling earlier detection.

4. Conclusion

In this work, we introduced SeqVAE-TEB, a predictive VAE based framework that incorporates the transfer entropy bottleneck to explicitly model the directed influence of UP on FHR dynamics. By reformulating the learning objective from reconstruction to prediction, the model learns to predict next two minutes of FHR from the latent respresentation which encodes the history of FHR conditioned on UP. Evaluation on the CTG dataset demonstrated stable specificity across diverse healthy subgroups and consistent sensitivity for acidosis detection, with generalizable performance on unseen HIE cases.

Acknowledgments

This study was supported by funding provided by the Bill & Melinda Gates Foundation and the National Institutes of Health.

References

[1] Freeman RK, Garite TJ, Nageotte MP. Fetal heart rate monitoring. Lippincott Williams & Wilkins, 2003.

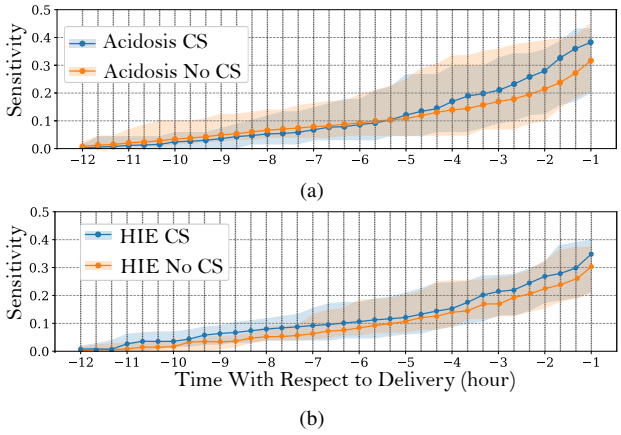

Figure 3: Test sensitivity stratified by CS criteria for (a) Acidosis records and (b) HIE records.

[2] Larma JD, Silva AM, Holcroft CJ, Thompson RE, Donohue PK, Graham EM. Intrapartum electronic fetal heart rate monitoring and the identification of metabolic acidosis and hypoxic-ischemic encephalopathy. American journal of obstetrics and gynecology 2007;197(3):301–e1.

[3] Macones GA, Hankins GD, Spong CY, Hauth J, Moore T. The 2008 national institute of child health and human development workshop report on electronic fetal monitoring: update on definitions, interpretation, and research guidelines. Journal of Obstetric Gynecologic Neonatal Nursing 2008; 37(5):510–515.

[4] Kingma DP, Welling M. Auto-encoding variational bayes. arXiv preprint arXiv13126114 2013;.

[5] Schreiber T. Measuring information transfer. Physical review letters 2000;85(2):461.

[6] Kalajdzievski D, Mao X, Fortier-Poisson P, Lajoie G, Richards B. Transfer entropy bottleneck: Learning sequence to sequence information transfer. arXiv preprint arXiv221116607 2022;.

[7] Shamsi M, Lahiri A, Wu Y, Gerstley L, Kuzniewicz M, Cornet M, Hamilton E, Kearney R, Warrick P. Analyzing fetal heart rate patterns via latent representations with sequential variational autoencoder (seqvae). In Proceedings of the 51st Computing in Cardiology Conference (CinC 2024). Karlsruhe, Germany, 2024; .

[8] Kearney RE, Wu YW, Vargas-Calixto J, Kuzniewicz MW, Cornet MC, Forquer H, Gerstley L, Hamilton E, Warrick PA. Construction of a comprehensive fetal monitoring database for the study of perinatal hypoxic ischemic encephalopathy. MethodsX 2024;12:102664.

Address for correspondence:

Mahdi Shamsi

Address: 3775 University - Montreal - QC - H3A 2B4

E-mail: mahdi.shamsi@affiliate.mcgill.ca