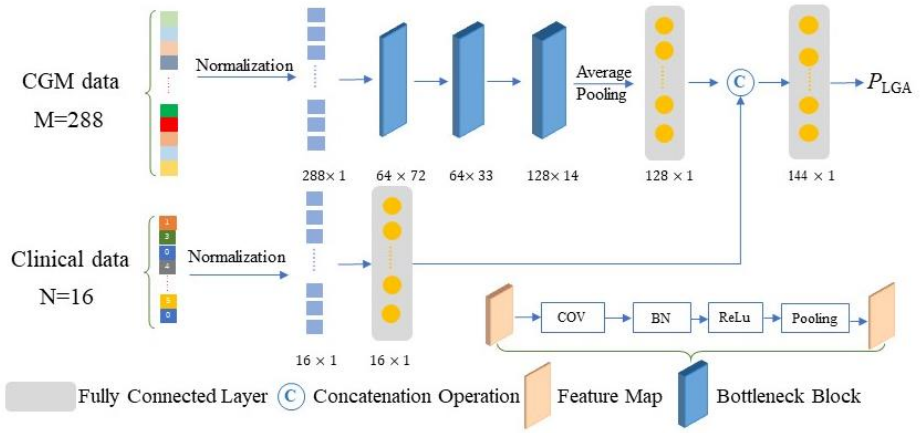
Supplementary

Supplementary methods

Figure S1-The overall network architecture for our fusion model proposed.



The entire fusion network is composed of two branches (CNN-based model and MLP-based model) to extract two streams of features. (1) The CGM data of each case with size of 288*1 is normalized into the feature tensor (vector). We further extract features by 1d convolution layer with kernel size of 7 and obtain feature tensor with size of 64×72. Next, the feature tensor is fed into two convolution layer with kernel size of 5 and obtain feature tensor 128×14. The max pooling operation is utilized to generate CGM data feature with size of 128×1. (2) The clinical data is also first processed using normalization for each case. We extract feature using MLP to obtain feature tensor with size of 16×1. Two branches of feature tensors are fused via the concatenate operation to generate the fusion feature with size of 144× 1. The fusion feature tensor is utilized to predict the LGA with a final convolution layer.

Loss function

In our study, the positive samples account for about 20% of the entire dataset. Generally, the unbalanced sample distribution biases the network model towards negative samples, which results in the low prediction accuracy for positive samples. We have conducted experimented with the standard Binary Cross-Entropy loss[1] to optimize the prediction model:

$$L = -(y\log(p(x) + (1-y)\log(1-p(x))) \quad (1)$$

Experiment Setup

We adopted the Adam solver to optimize the model for 20 epochs with a batch size of 16. The initial learning rate is set to 0.004, which decays with a factor of 10 for every 5 epochs. Our method is implemented using PyTorch, and all the experiments were conducted on a NVIDIA Geforce RTX 3090 GPU. We also performed a series of ablation studies to verify the contribution of each module. Specifically, the two branches are individually trained to predict the LGA under the same experimental setup. We evaluated and compared the performance of several methods, including random forest[2], decision tree[3], and logistic regression, using AUC_{ROC} and accuracy.

Visualization analysis

To visualize further the effectiveness of the proposed model, we calculate the weighted heatmap[4] from convolutional layers output in first branch via implementing the multiplication of features with 128×14 dimensions and fully connected layer weights. We obtain the 232 feature-tensor with non-LGA and 62 features with LGA from the training dataset based on the trained CGM model. The two groups of heatmaps are presented in **Fig.S4 (A)** and **Fig.S4 (B)**, respectively. Comparing two cohorts, the feature values have significant differences via using color bar values. Additionally, we calculate the mean values of all feature tensors from two cohorts, and obtain **Fig.S4(C)**. The two curves show that the prediction model proposed have the ability to distinguish the LGA and non-LGA. Finally, we normalize the single feature tensor from training dataset, and obtain the mean curves of feature tensors as shown in **Fig.S4(D)**. The values of the feature tensor after the normalization operation represent the importance of corresponding feature values. This figure shows that the GDM data from 24h have different importance for predicting LGA cohorts and non-LGA ones.

Supplementary results

Table S1- Patients' characteristics of training, validation and test datasets

Characteristics	Total	Training Dataset	Validation Dataset	Test Dataset	*P*-value
N	371	293 (79.0%)	33 (8.9%)	45 (12.1%)	
Age (years)	31.8±4.5	31.7±4.5	33.2±4.1	31.6±5.2	0.172
Pregestational BMI (kg/m^2)	23.4±3.6	23.4±3.7	23.0±2.7	23.6±3.6	0.765
BMI < 25	266(71.7)	212(72.4)	24(72.7)	30(66.7)	0.726
BMI ≥ 25	105(28.3)	33(27.6)	33(27.3)	72(33.3)	
Height(cm)	160±5	160±5	160±4	161±5	0.954
GDM History (%)	31(8.4)	23(7.8)	3(6.1)	6(13.3)	0.378
Macrosomia History* (%)	12(3.3)	7(2.4)	1(3.0)	4(8.9)	0.068
Parity [n (%)]					0.954
0	181(48.8)	144(49.1)	16(48.5)	21(46.7)	
≥1	190(51.2)	149(50.9)	17(51.5)	24(53.3)	
Measurements in second trimester					
Systolic Pressure (mmHg)	116±10	116±10	114±10	115±10	0.334
Diastolic Pressure (mmHg)	71±9	71±9	70±10	70±7	0.611

Maternal Weight Gain (Kg)	7.4±4.6	7.4±4.6	8.2±4.3	6.8±4.9	0.410
FPG (mmol/L)	5.0±0.6	5.0±0.6	4.8±0.6	5.0±0.7	0.253
OGTT 1h Glucose (mmol/L)	10.2±1.7	10.2±1.8	10.2±1.2	10.1±1.5	0.904
OGTT 2h Glucose (mmol/L)	8.5±1.6	8.5±1.6	8.4±1.8	8.3±1.6	0.665
HbA_{1c} (%)	5.2±0.4	5.2±0.4	5.2±0.5	5.1±0.4	0.662
Fasting Insulin (μIU/L)	62.0(45.4,87.5)	60.7(45.8,88.6)	66.0(37.4,81.0)	67.4(44.1,90.8)	0.175
Cholesterol (mmol/L)	5.8±1.1	5.8±1.1	5.9±0.9	5.9±1.1	0.596
Triglycerides (mmol/L)	2.6±1.0	2.6±1.1	2.4±0.7	2.7±1.1	0.589
HDL (mmol/L)	1.9±0.7	1.9±0.8	1.8±0.3	1.9±0.4	0.968
LDL (mmol/L)	3.0±0.9	3.0±0.9	3.1±0.9	3.0±0.9	0.660
Glycated Albumin (%)	12.9(11.9, 14.0)	12.9(11.8,14.0)	13.1(12.4,14.2)	13.0(11.8,14.0)	0.982
1h Insulin (μIU/L)	425.8(298.2,630.0)	430.0(305.6,644.4)	421.7(278.8,577.5)	390.6(302.0,739.0)	0.929
2h Insulin (μIU/L)	415.5(267.9,662.7)	434.0(274.0,688.7)	350.6(177.7,670.1)	368.8(260.8,520.1)	0.048
CGM examine time (weeks)	26.6(25.4-28.3)	26.5(26.1-26.7)	26.1(25.3-28.0)	26.7(26.2-28.3)	0.974
Treatments after CGM					0.561
Medical treatment	77(20.8)	59(20.1)	6(18.2)	12(26.7)	
Diet-exercise treatment	294(79.2)	234(79.9)	27(81.8)	33(73.3)	
Maternal and neonatal outcomes					
Delivery gestation time (weeks)*	38(37,39)	38(37,39)	39(37,39)	39(38,39)	0.420

Birth Weight(g)	3217.8±527.7	3225.6±530.0	3109.4±571.6	3246.6±479.2	0.453
LGA (%)					0.991
Yes	76 (20.5)	60 (20.5)	7 (21.2)	9 (20.0)	
No	295 (79.5)	233 (79.5)	26 (78.8)	36 (80.0)	

Abbreviations: BMI, body mass index; CGM, continuous glucose monitoring; GDM, gestational diabetes mellitus; LGA, large for gestational age; OGTT, oral glucose tolerance test; SD, standard deviation.

Table S2 Comparisons of variables values with missing data before and after multiple imputation

Variable	Missing size	Before	After	*P*-value
		Mean (SD)	Mean (SD)	
SBP	1	116.10(10.57)	116.13(10.33)	0.9681
DBP	1	70.49(9.13)	70.72(8.73)	0.7418
Cholesterol	35	5.81(1.07)	5.83(1.04)	0.8050
TG	35	2.61(1.06)	2.59(1.01)	0.7594
OGTT 1h	16	10.15(1.70)	10.19(1.71)	0.7901
HbA1c	26	5.16(0.44)	5.17(0.43)	0.8070

Abbreviations: DBP, diastolic blood pressure; OGTT, oral glucose tolerance test; SBP, systolic blood pressure; SD, standard deviation; TG, triglyceride.

Figure S2- 24-h mean glucose profiles in LGA and non-LGA GDM pregnancies.
(A) GDM women with BMI < 25 kg/m^2; (B) GDM women with BMI ≥25 kg/m^2; (C) Diet-treated GDM; (D) Medically-treated GDM; (E) in 2nd trimester; (F) in 3rd trimester.

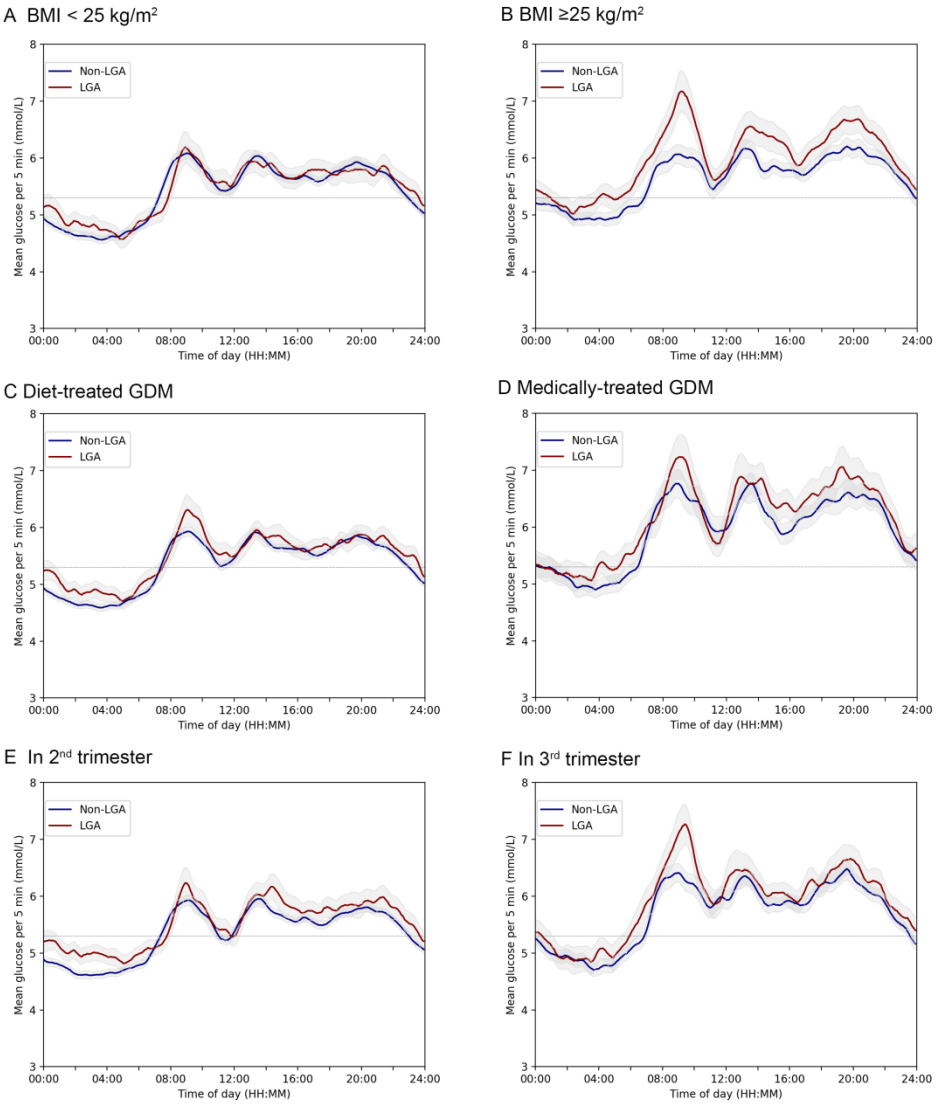


Figure S3 Mean glucose difference profile of GDM with and without LGA infants. (A) GDM women with BMI < 25 kg/m^2; (B) GDM women with BMI ≥25 kg/m^2; (C) Diet-treated GDM; (D) Medically-treated GDM; (E) In 2nd trimester; (F) In 3rd trimester.

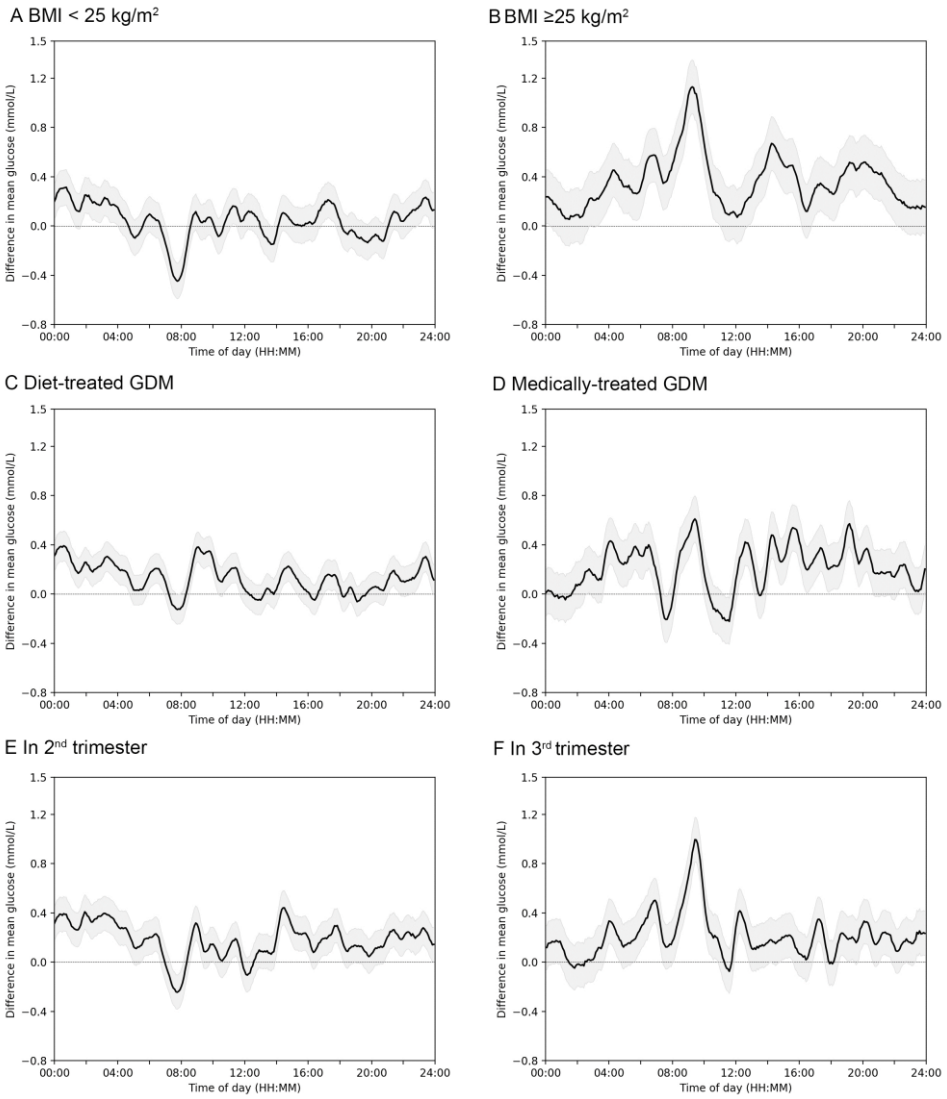
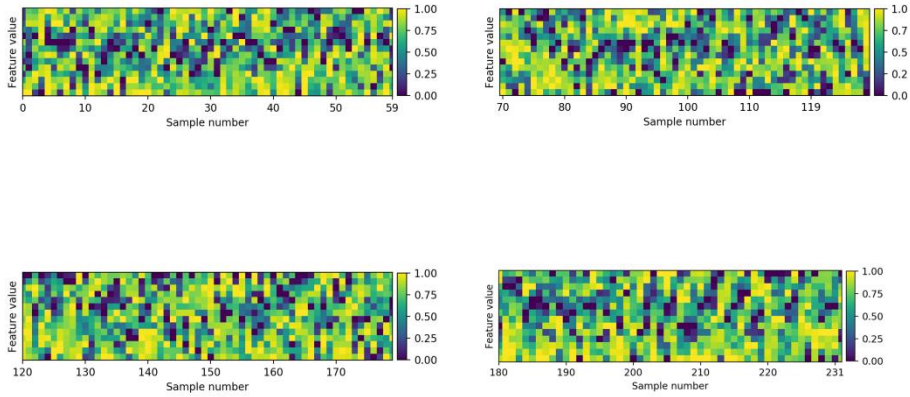
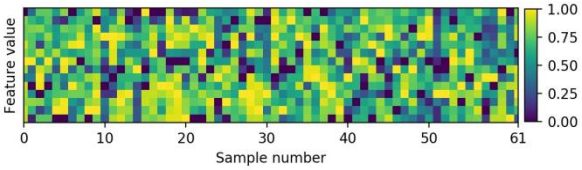


Figure S4 Heatmaps and feature maps

A Non-LGA

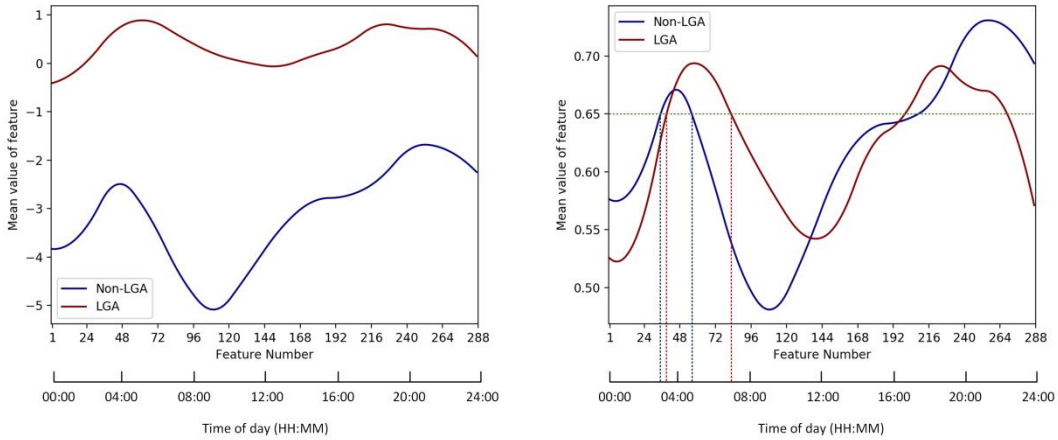


B LGA



C Mean value of heatmap

D Mean value of normalized heatmap



Notes: Figure S4.A presented feature values for a single case using the Class Activation Mapping (CAM) method from the final convolution layer of the Convolutional Neural Network (CNN) in the Non-LGA group. Similarly, the LGA group was illustrated in Figure S4.B. To evaluate the model's effectiveness, we plotted the mean values of feature values from all cases in the training dataset within 24 hours in Figure S4.C. The distinct difference between the two groups highlighted the model's proficiency in distinguishing between Non-LGA and LGA. In Figure S4.D, we analyzed the feature value distribution from another perspective by applying normalization and mean operations. A blue dashed line was introduced as a reference threshold to clearly illustrate the shift phenomenon.

References

[1] Mao Anqi, Mehryar Mohri, Yutao Zhong. Cross-entropy loss functions: Theoretical analysis and applications. In International conference on Machine learning, pp. 23803-23828. PMLR, 2023.

[2] Parmar Aakash, Rakesh Katariya, Vatsal Patel. A review on random forest: An ensemble classifier. In International conference on intelligent data communication technologies and internet of things (ICICI) 2018, pp. 758-763. Springer International Publishing, 2019.

[3] Priyanka Kumar D. Decision tree classifier: a detailed survey. International Journal of Information and Decision Sciences. 2020;**12**(3):246-69

[4] Zhou Bolei, Aditya Khosla, Agata Lapedriza, Aude Oliva, Antonio Torralba. Learning deep features for discriminative localization. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2921-2929. 2016