

DiabEye-Q: AI-driven Longitudinal Analysis of Ophthalmoscopic Images for Early Diabetes Prediction in Qatari Adults

Dr. Zubair Shah

Associate Professor, College of Science and Engineering
Hamad Bin Khalifa University (HBKU), Qatar

Sulaiman Khan, Md. Rafiul Biswas (Co-authors)

Background

- Diabetes is a chronic condition that reflects body's inability to produce or respond to insulin and is characterized by abnormal glucose levels in the blood.
- WHO reports that it is the fifth most common reason for mortalities worldwide [1].
- The IDF statistics show that after every six seconds one person dies of diabetes [2].
- In 2019, approximately 425 million adults were living with diabetes; and it is projected to be 629 million by 2045 [3].

Current AI Focus

- Numerous AI/ML techniques are proposed to classify diabetic vs. non-diabetic cohorts using:
 - Electronic Health Records (EHRs)
 - Genetic data
 - Imaging data (e.g., retinal fundoscopy)
 - Other clinical/omics modalities
- Several detection models exist that integrate retinal images with other modalities to capture vascular/metabolic signals.

Clinical Reality of Diabetes Diagnosis Using AI

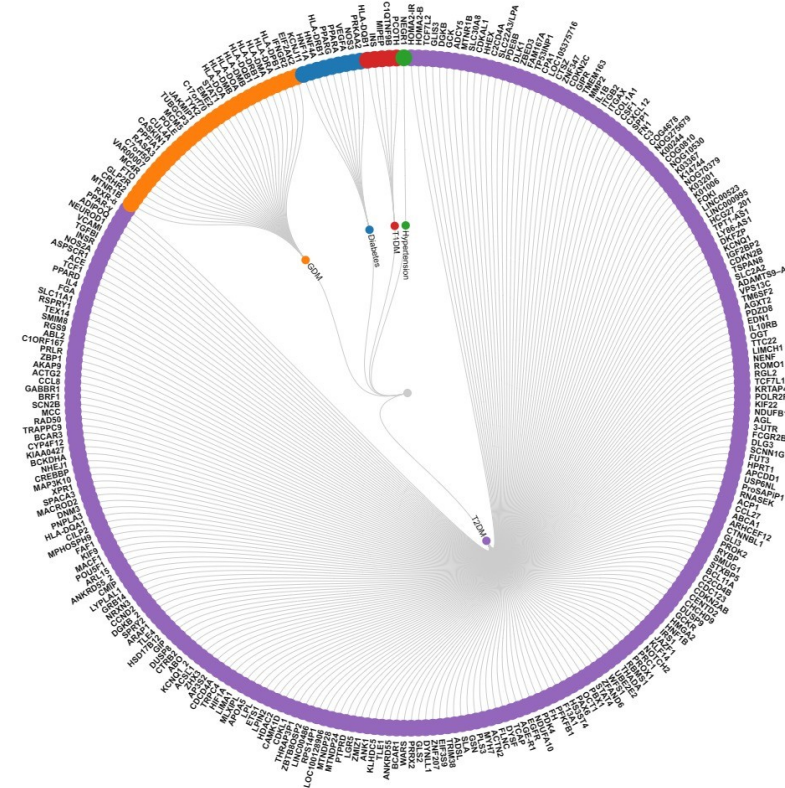
- **Diabetes Diagnosis is not the real problem.**
- Diabetes can already be detected cheaply and reliably using:
 - HbA1c
 - Fasting Plasma Glucose (FPG)
 - Oral Glucose Tolerance Test (OGTT)
- **Replacing these with AI is unlikely to happen and AI-based detection adds little clinical value in diabetes diagnosis,** since these gold-standard methods are already effective and inexpensive.
- The true challenge is **identifying who will become diabetic in the future.**

Diabetes Risk Prediction Using AI

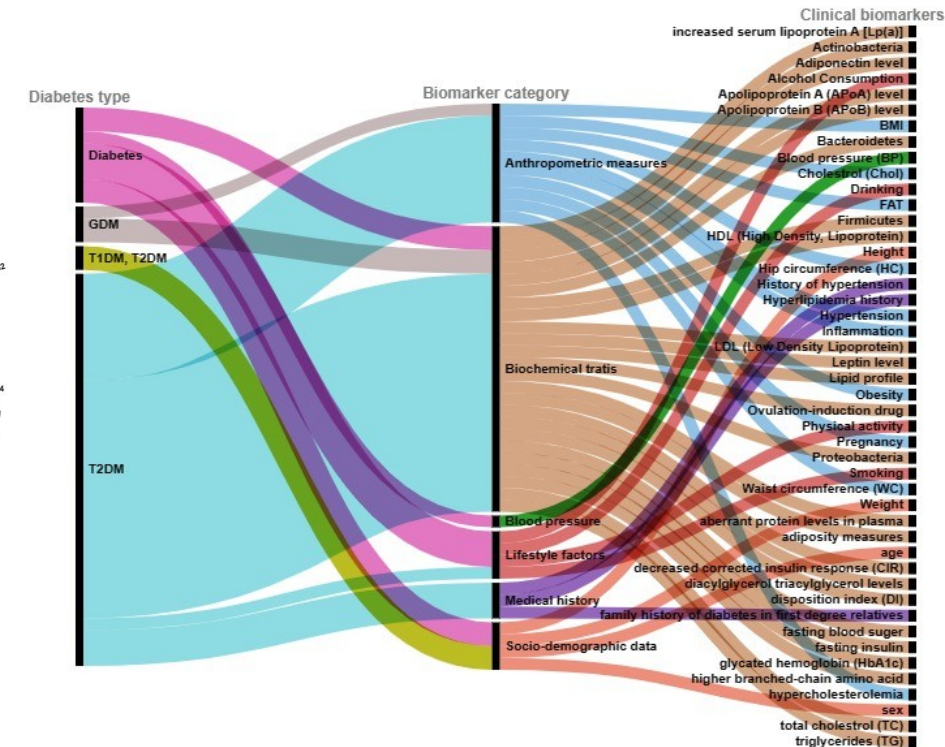
- Current **clinical tools** for **predicting who will develop diabetes** are limited.
- Value of AI lies not in **classification of diabetes**.
- Instead, the value of AI lies in identifying **risk factors** using longitudinal analysis of patients, enabling **early prediction**.
- Uncovering **longitudinal risk factors** is crucial for prevention and early intervention.
- **AI offers unique potential** to model these risk trajectories by fusing multimodal data, such as (EHR, genetics, imaging, lifestyle).
- True clinical value lies in **shifting from diagnosis to early prediction**.

Clinical and Genetic Risk Factors

- Studies have identified diabetes risk factors, such as:
 - Genetic risk factors
 - Clinical risk factors
- However, most of these studies are **unimodal**.



Genetic factors [4]



Clinical factors [4]

Gaps in the Literature

- Many studies are **retrospective and cross-sectional**; few address prediction before onset using longitudinal data.
- **Unimodal approaches** dominate (EHR-only, genetics-only, imaging-only).
- Very limited **multimodal approaches** exist, especially those fusing **retinal fundoscopy, EHR, genetic, and clinical measures**.
- Question: Can ML models capture **longitudinal risk trajectories** of diabetes using **multimodal fusion**?

Research Aim

- Addressing this gap, we:
Develop a pipelined architecture that combines XGBoost with a Vision Transformer (ViT) to predict diabetes onset in the Qatari population and identify longitudinal risk factors through multimodal data fusion.

Objectives



Task 1 – Explore longitudinal associations between fundoscopy features (e.g., vessel morphology) and T2DM onset in the Qatari cohort.



Task 2 – Develop an interpretable AI pipeline by fusing XGBoost with ViT for longitudinal diabetes risk prediction.



Task 3 – Assess how age and gender influence diabetes prediction accuracy by analyzing model performance across distinct age cohorts and quantifying gender-specific susceptibility patterns for diabetes onset

Qatar BioBank (QBB) Data

Longitudinal data

- Diabetic patients
- Pre-diabetic patients
- Normal/controlled people

Demographics

- Age
- Gender
- Nationality

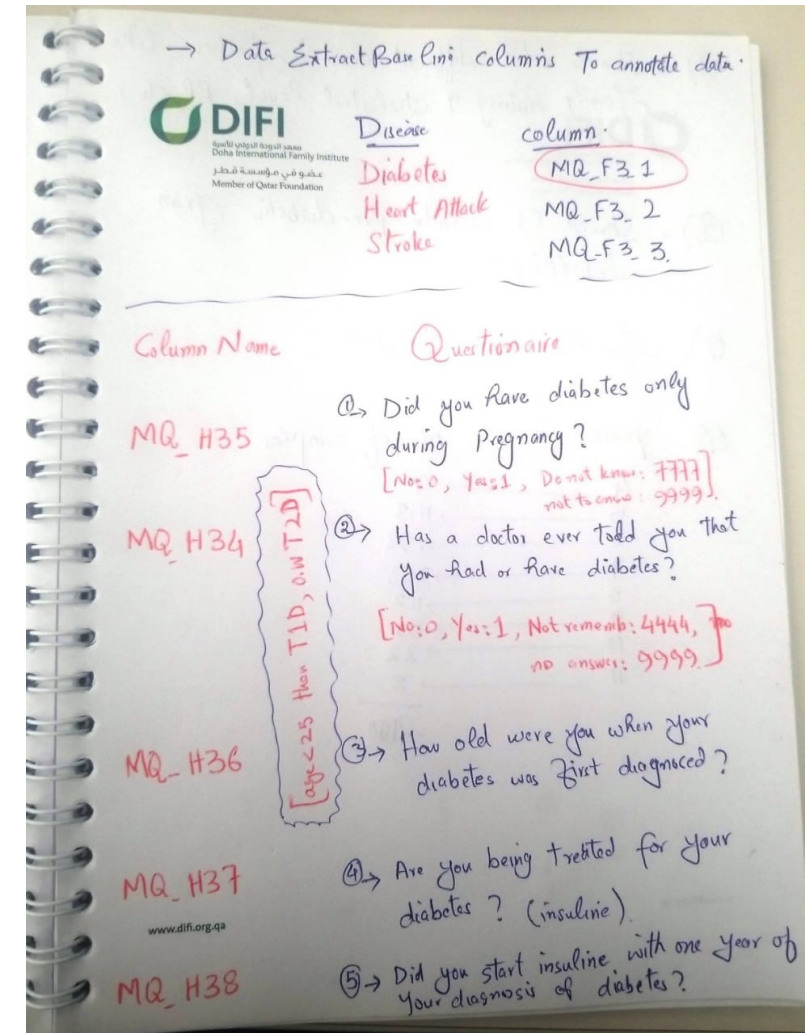
Type of data

- Retinal funduscopy images
- Clinical data

Longitudinal timeframe (years)	# of Samples
4	4
5	1118
6	798
7	293
8	113
9	50
10	22
11	2

Data Annotation

- The QBB data contains baseline and follow-up records
- We filtered the **Baseline Data** by removing pre-diabetic and diabetic participants using the following criteria:
 - (Pre-diab: IF $HbA1c \geq 5.7$ and ≤ 6.4)
 - (Diabetic: $HbA1c \geq 6.5$)
 - Self-reported pre-diabetes
 - Those who are using tablets for diabetes care (Metformine etc.)
 - start insulin within one year of your diagnosis of diabetes?
- We used only baseline data from healthy participants

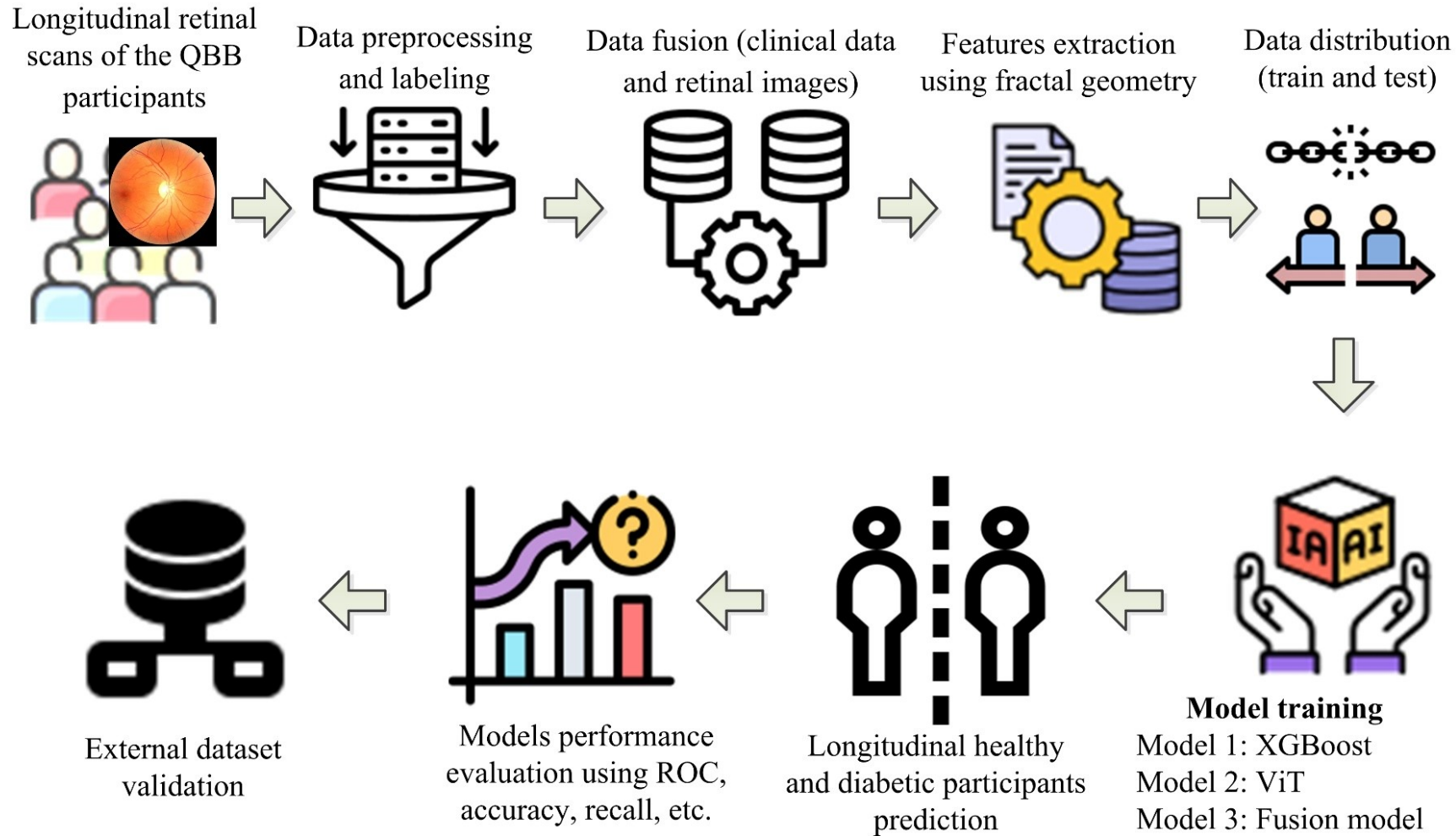


Data Annotation cont...

- A total of 2,043 participants have their longitudinal follow-up retinal imaging data with the following specifications:
 - **Diabetic**: 430 and **Healthy**: 1,633
 - **Male**: 1,252 participants (diabetic: 246, healthy: 1,006)
 - **Female**: 791 participants (diabetic: 184, healthy: 607).
- Longitudinal Period

S. No	Longitudinal period	Number of participants
1.	4 years	12
2.	5 years	949
3.	6 years	920
4.	7 years	137
5.	8 years	23

Methodology



Prediction of Future Risk of Diabetes

● Step 1 – Labeling Strategy

- Baseline data (EHR, genetics, retinal images, etc.) is linked with **follow-up outcomes**.
- If a participant develops diabetes at follow-up → their **baseline record is labeled “future diabetic.”**
- If not → their **baseline record is labeled “non-diabetic.”**

● Step 2 – Data Preparation

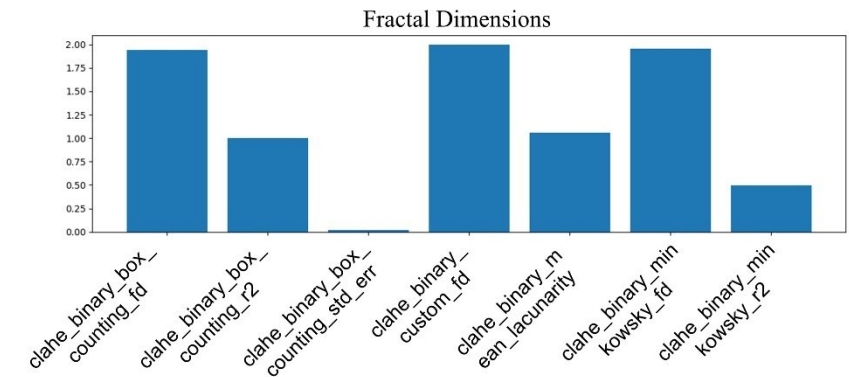
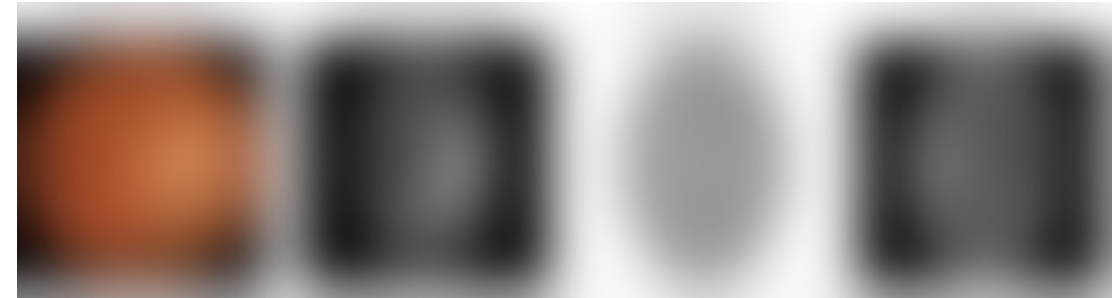
- Baseline data is now paired with future outcomes.
- Dataset is split into training and testing cohorts.

● Step 3 – Model Training & Evaluation

- AI model is trained on the **training set** to learn patterns in baseline data that predict future diabetes.
- Model performance is evaluated on the **testing set** to assess predictive accuracy.

Feature Extraction

- We applied fractal geometry to retinal images to calculate patterns from these images.
- Fractal geometry studies objects or patterns that exhibit self-similarity at different scales. These patterns can be described using fractals, which often have complex structures, but are generated by simple recursive processes
- We applied the box-counting method to funduscopy images, where the goal was to estimate the fractal dimension of the image, which can capture texture, roughness, and other complex features.



Fractal features

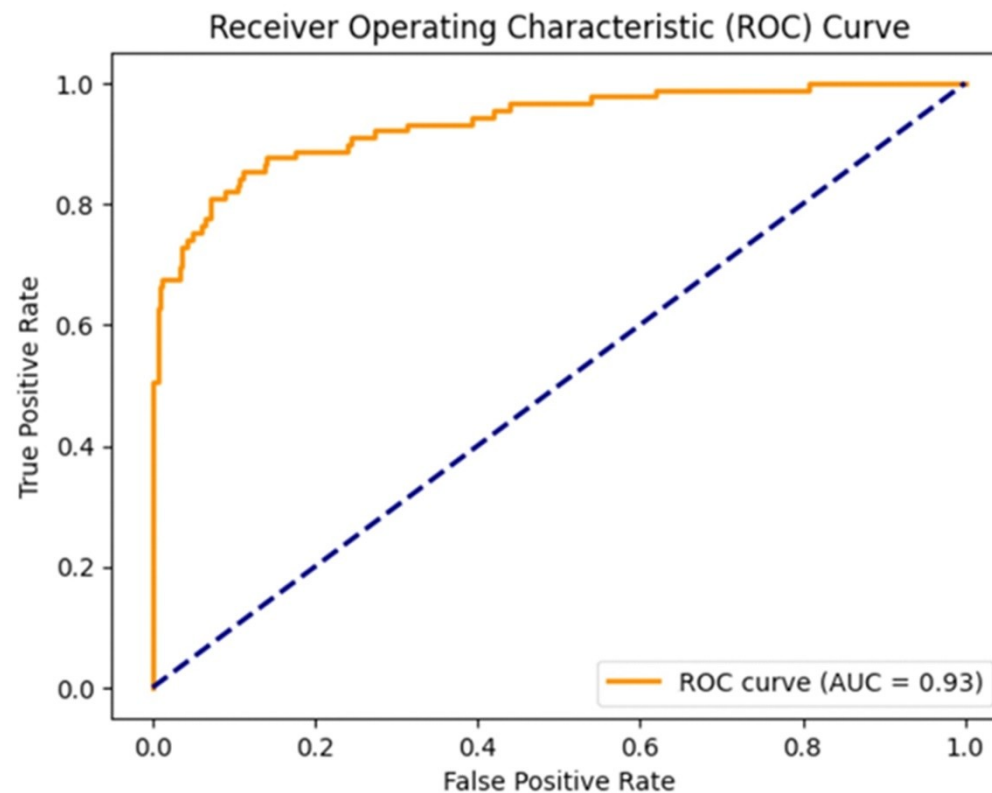
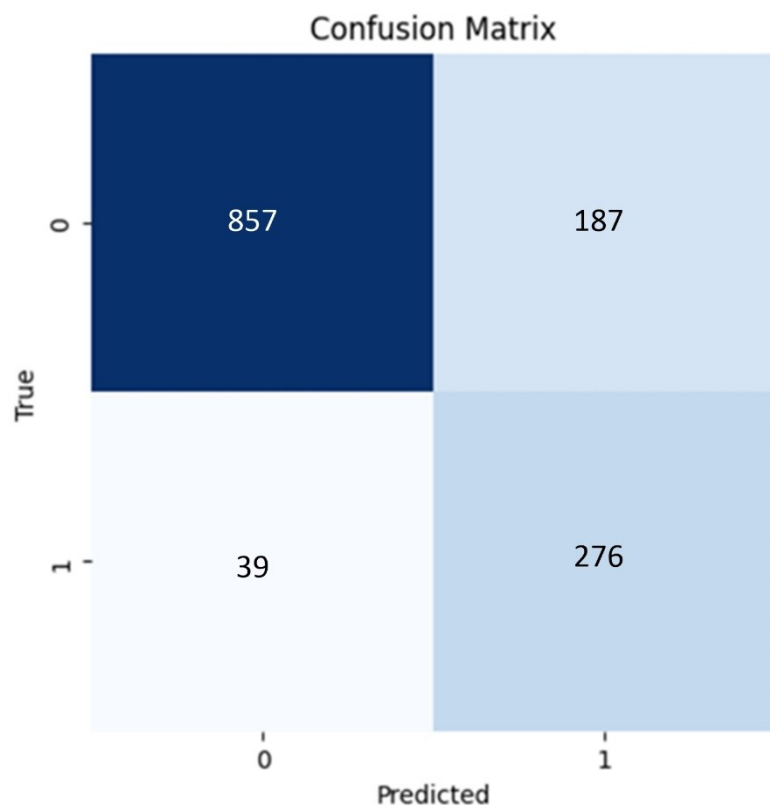
Hardware and Software Specs

- The experiments were performed on Google Coalab with A100 40GB SXM, 83 GB of RAM, and an NVIDIA A100 Tensor Core GPU.
- Scikit-learn was used for shallow models, while TensorFlow-Keras was employed for ViT.

Table. Different parameters selected for ViT to train and evaluate

Parameters	Description	Parameters	Description
Number of layers	16	Embedding dimensions	768
Input projection	$768 \rightarrow 3072$	Output projection	$3072 \rightarrow 768$
Learning rate	$1e^{-4}$	Weight decay	$1e^{-5}$
Epoch size	500	Training and test split (%)	70 : 30
Activation function	GeLU	Patch size	16×16
Performance metrics	Accuracy, recall, AUROC, etc.	Parameters	$\approx 86M$
Optimizer	AdamW	Attention heads per layer	12

Results



Comparative Analysis

- To compare our model with other benchmark models and datasets,
 - We selected the work of Mohsen et al. [5], that also utilized the same Qatari Cohort for research analysis.
 - The **Indian Diabetic Retinopathy dataset** used by Saxena et al. [6].
 - **Kaggle datasets**: Messidor-1, Messidor-2, and EyePACS

Research studies	Test dataset	F-score	AUC
Mohsen et al., [12]	QBB	82.5%	94.76%
Our model	QBB	88%	93%
saxena et al., [15]	Messidor-2		92%
Our model	Messidor-2	92%	95.30%

Age-group vs Model Performance

- We stratified test data into three distinct age cohorts as shown in [Table](#).
- For younger participants (22 — 40 years), the models achieved higher prediction accuracy and recall values for healthy participants. Prediction metrics (precision values) were lower for diabetic cases within this age group.
- For older cohorts (41 — 55 years and 56 — 69 years), precision and recall values increased for diabetic participants and decreased for healthy participants in the QBB dataset.

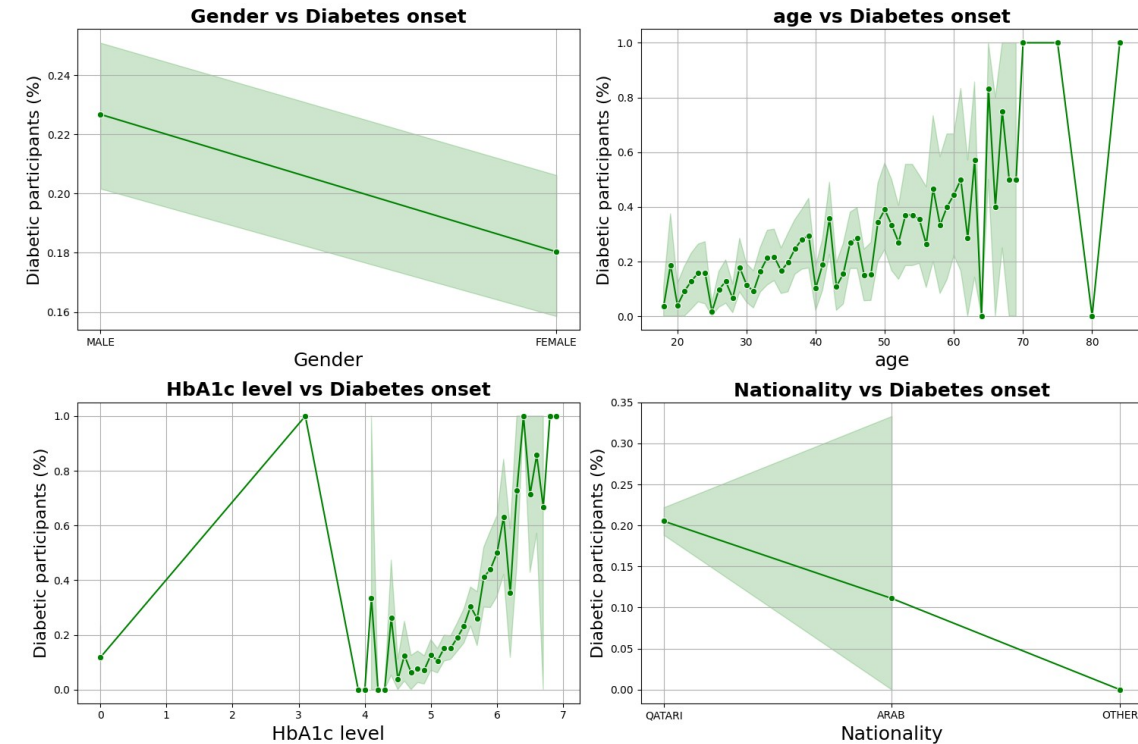
Age groups	No. of participants	Accuracy (%)		Precision (%)		Recall (%)	
		<i>cont</i>	<i>cas</i>	<i>cont</i>	<i>cas</i>	<i>cont</i>	<i>cas</i>
$22 \leq age < 40$	613 (cont: ^a 485, cas: ^b 128)	87.03	86.41	93.95	66.62	87.05	86.47
$40 \leq age < 56$	293 (cont: ^a 205, cas: ^b 88)	86.05	80.75	91.89	74.93	83	80.88
$\geq 56 \text{ years}$	69 (cont: ^a 27, cas: ^b 42)	86.89	90.35	86.90	90.35	86.89	92.35

^a cont — controls/healthy participants.

^b cas — cases/diabetic participants.

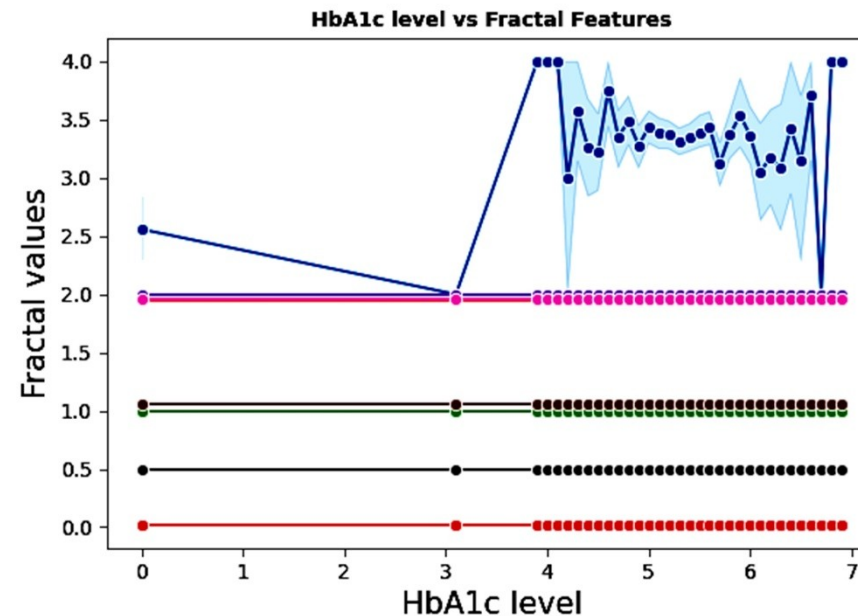
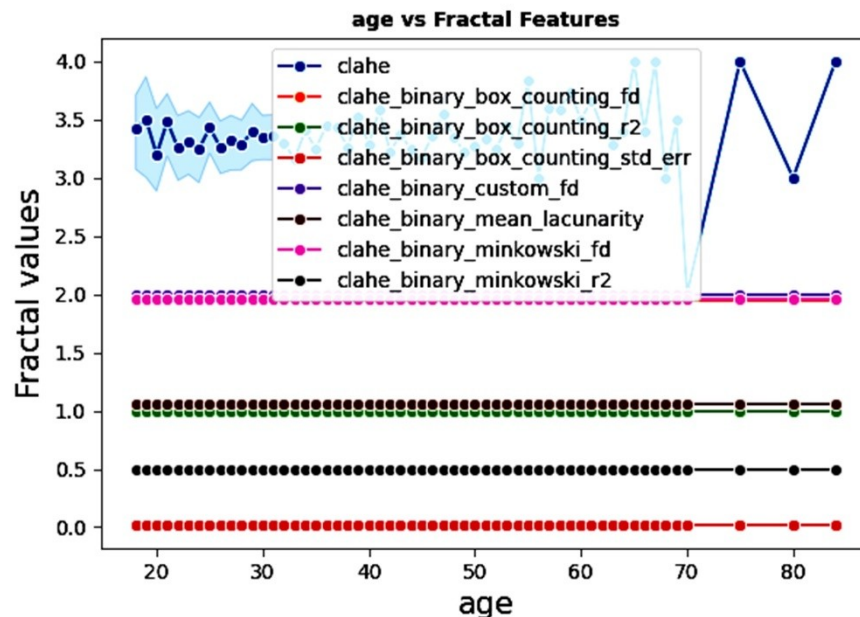
Clinical factors vs Diabetes Onset

- We investigated multiple clinical factors such as HbA1c, gender, and age that influencing diabetes onset in Qatari adults.
- Diabetes incidence was higher in **males during the follow-up period**. This gender disparity was high in the Qatari population.
- Demographic analysis revealed that **approximately one in nine participants from other Arab nationalities** (Yemen, Lebanon, Jordan, Oman, etc.) developed diabetes during follow-up, while the majority of diabetic cases occurred among Qatari adults, predominantly males.
- Age and HbA1c levels emerged as critical risk factors. **Diabetes susceptibility increased with age, with notably higher risk for individuals over 40 years** and a marked increase for those 70 and older.



Clinical Factors vs Fractals

- After analyzing fractal values in relation to demographic and clinical factors, a steady change in fractal values was observed for HbA1c level ≥ 4 .
- In contrast, an abrupt change was noted in participants older than 55 years, suggesting that older individuals in the QBB dataset may be more susceptible to diabetes onset.



Research Implications

- From a clinical perspective,
 - It will assist both the patients and practitioners to decide about the development of diabetes in future using his/her current clinical examination results
 - To suggest medications and lifestyle guidelines based on his/her future health status
- From a computer science community perspective
 - To fuse different data modalities and achieve higher diabetes prediction capabilities
 - To perform longitudinal analysis of fundoscopy images and clinical data modalities to perform diabetes prediction

Limitation

- The dataset is sourced exclusively from QBB, containing only Qatari participants, which limits the generalizability to other populations.
- The sample size is relatively small (2,043 participants) and highly imbalanced, increasing the risk of model overfitting.
- The follow-up period is short, with only 23 participants having eight years of medical history, limiting insights into long-term diabetes risk factors.
- The exclusion of prediabetic individuals, and type 1 diabetes patients may affect the study's comprehensiveness.

Conclusion

- This research work presents an approach for diabetes onset prediction using artificial intelligence and multi-modal data.
- After fusing ophthalmoscopic imaging features with EHR, we employed both conventional and transformer-based model to perform predictions.
- During experimental analysis, it was observed that integrated model (DiabEye-Q) outperformed both standalone ViT and XGBoost, achieving 88.08% accuracy, 93% AUROC, and 88% recall values.
- Our analysis revealed distinct diabetes prediction patterns across demographic segments. Among younger participants (22—40 years), the model demonstrated elevated false positive rates when classifying healthy individuals. In contrast, prediction accuracy significantly improved for diabetic participants in older cohorts, with peak performance observed in the 56 — 69 years age period.
- Additionally, analysis of the QBB dataset indicated that male participants exhibited markedly higher susceptibility to diabetes onset compared to their female counterparts.

Acknowledgements

- We thank the Qatar Biobank laboratory and spectrometry teams for providing clinical and retinal scanning data and acknowledge Qatar Biobank (Project QF-QBB-RES-ACC-00181) for data access.
- We are grateful to all biobank participants for their contributions to scientific advancement.

Thank you!

zshah@hbku.edu.qa

References

- [1]. P. Balakumar, K. Maung-U, and G. Jagadeesh, "Prevalence and prevention of cardiovascular disease and diabetes mellitus," Pharmacol. Res., vol. 113, pp. 600 - 609, Nov. 2016.
- [2]. S. Pugazhenth, L. Qin, and P. H. Reddy, "Common neurodegenerative pathways in obesity, diabetes, and Alzheimer's disease," Biochim. Biophys. Acta-Mol. Basis Disease, vol. 1863, no. 5, pp. 1037-1045, 2017.
- [3]. N. Cho, J. Shaw, S. Karuranga, et al., "IDF diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045," Diabetes Res. Clin. Pract., vol. 138, pp. 271281, Apr. 2018.
- [4]. Khan, S., Mohsen, F., & Shah, Z. (2024). Genetic biomarkers and machine learning techniques for predicting diabetes: systematic review. Artificial Intelligence Review, 58(2), 41.
- [5]. Mohsen, F., Shah, U., Islam, A., Shah, Z., Belhaouari, S.B.: Multimodal deep learning for diabetic retinopathy grading: Integrating linear-radon sinograms and retinal fundus images. In: 2024 IEEE 26th International Workshop on Multimedia Signal Processing (MMSP). pp. 1-5. IEEE (2024)
- [6]. Saxena, G., Verma, D.K., Paraye, A., Rajan, A., Rawat, A.: Improved and robust deep learning agent for preliminary detection of diabetic retinopathy using public datasets. Intelligence-Based Medicine 3, 100022 (2020)