

Multimodal Deep Learning Model for Diabetes Hyperglycemic Prediction

Vitale Ciro

Deep Learning

Università degli Studi di Salerno

Agosto 28, 2025

<https://github.com/cirovitale/diabetes-multimodal>

Contents

1	Introduzione	3
2	Research Question	3
3	Dataset	4
4	Data Preprocessing	6
5	Progettazione Modello Multimodale	8
5.1	Input Layers	9
5.2	Branch CGM	9
5.3	Branch Multi-Sensor Fitbit	9
5.4	Branch Dati Terapeutici	11
5.5	Fusion Layer	12
5.6	Output Layer	12
5.7	Metriche di Monitoraggio e Valutazione	12
5.8	Tecniche di Ottimizzazione e Regolarizzazione	12
6	Risultati Training	13
7	Testing e Validazione	14
7.1	Prestazioni del Modello Multimodale	14
7.2	Analisi delle Research Questions	14
7.2.1	Risposta a RQ1: Efficacia dell'Approccio Multimodale	14
7.2.2	Risposta a RQ2: Generalizzazione e Robustezza	15
8	Sviluppi Futuri	15
9	Conclusioni	15

1 Introduzione

Il diabete rappresenta una delle maggiori sfide sanitarie a livello globale, secondo il **2025 IDF Diabetes Atlas**, l'undicesima edizione del testo di riferimento curato dalla **International Diabetes Federation (IDF)**, circa **589 milioni di adulti** (tra i 20 e i 79 anni) convivono attualmente con il diabete, un valore destinato a salire a **853 milioni entro il 2050** [1]. Nel 2024 ci sono state circa 3.4 milioni di morti per diabete, 1 ogni 9 secondi [1]. Tutto ciò ha contribuito negativamente nella spesa sanitaria, con almeno **mille miliardi di dollari spesa** [1]. Questi dati particolarmente allarmanti rendono necessario lo sviluppo di nuove tecniche avanzate per migliorare la diagnosi e la **gestione** del diabete.

In questo ambito i modelli di **Machine Learning (ML)** e **Deep Learning (DL)** hanno mostrato grande potenziale nel migliorare la diagnosi e la gestione del diabete. I **modelli tradizionali** hanno fornito buoni risultati basandosi esclusivamente su singole *feature*, quindi su **input unimodali** [2]. Negli ultimi anni il problema è stato affrontato considerando l'applicazione di **modelli di deep learning multimodali**, che come suggerisce il nome, riescono a sfruttare simultaneamente fonti eterogenee di informazioni, come immagini di radiografie, dati testuali e serie storiche di rilevazioni mediche di vario genere, fornendo una **visione olistica del paziente**. Le informazioni comunemente utilizzate negli approcci unimodali:

- **Immagini mediche**, come le radiografie.
- **Serie temporali**, come i valori continui da **CGM (Continuous Glucose Monitoring)**.
- **Dati testuali**, come i referti medici.
- **Dati clinici strutturati**, come i parametri da **EHR (Electronic Health Records)**.

Queste strategie unimodali, sebbene abbiano fornito spunti utili, mancano spesso di una visione integrata che consideri informazioni di varia natura.

In letteratura è stata affrontata sebbene in maniera non approfondita rispetto all'approccio unimodale, la combinazione di diverse fonti dati per i modelli multimodali [2, 3, 4, 5], evidenziando l'efficacia della combinazione di informazioni eterogenee, superando nettamente i risultati ottenuti dai modelli basati su singole fonti di dati.

2 Research Question

- **RQ1:** L'integrazione di informazioni eterogenee, provenienti da fonti diverse, in un unico modello di DL migliora le prestazioni predittive nel rilevare o predire esiti legati al diabete rispetto a modelli basati su una singola modalità?
- **RQ2:** Il modello multimodale proposto generalizza bene a differenti pazienti e scenari clinici? Al fine di evidenziare la **robustezza** dell'approccio proposto in un contesto delicato come quello sanitario.

3 Dataset

In letteratura sono stati identificati i seguenti dataset:

- **DiaTrend** [6]: Ampio dataset con oltre 27 mila giorni di dati CGM e 8 mila giorni di dati da pompe insuliniche, raccolti da 54 pazienti diabetici tramite dispositivi medici wearable (vedi Figura 1) [6]. Questo dataset si focalizza su dati sensoriali e non include componenti testuali o informazioni sullo stile di vita come note o diari del paziente.
- **HUPA-UCM Diabetes Dataset** [7]: Dati raccolti in uno studio medico su **25 pazienti con diabete mellito di tipo 1 (T1DM)**, quindi **dati reali** raccolti in condizioni di vita quotidiana (**free-living**) e *non simulati*, il che aumenta il valore del dataset per sviluppare modelli con utilità pratica. Ogni paziente è stato monitorato per un periodo di almeno 14 giorni, durante il quale sono stati acquisiti diversi tipi di dati:
 - *Tracciati glicemici continui (CGM)*: misurazioni della glicemia tramite sensori Freestyle Libre 2, registrate ogni 15 minuti (vedi Figura 1).
 - *Somministrazione di insulina*: dose di insulina misurata mediante dispositivi di rilevazione *Continuous Insulin Infusion System (CIIS)* (vedi Figura 1).
 - *Diario alimentare*: quantità di carboidrati ingeriti ad ogni pasto (grammi di *CHO*).
 - *Dati da dispositivi wearable (Fitbit Ionic)*: passi effettuati, calorie bruciate, frequenza cardiaca.
- **mBRSET (Mobile Brazilian Retinal Dataset)** [8]: Dataset multimodale contenente **5,164 immagini retiniche** da **1,291 pazienti diabetici** di origine brasiliana ma con diversità etnica [8]. Le caratteristiche multimodali includono:
 - **Immagini retiniche**: immagini ad alta risoluzione ($1,600 \times 1,600$ pixel) catturate con fotocamera portatile Phelcom Eyer, integrata in un Samsung Galaxy S10.
 - **Dati demografici strutturati**: età, sesso.
 - **Dati clinici strutturati**: durata del diabete, terapia insulinica, ipertensione sistemica, livello educativo, consumo di alcol, fumo, obesità, malattie vascolari correlate al diabete, infarto del miocardio, nefropatia, neuropatia.
 - **Classificazioni cliniche**: grading della retinopatia diabetica secondo la classificazione **International Clinical Diabetic Retinopathy (ICDR)** con 5 livelli (da 0 *normale* a 4 *proliferativo*), presenza di edema maculare, valutazione della qualità delle immagini e presenza di artefatti.

I dati clinici presentano un'ampia diversità:

- genere maschile 34.93 %, genere femminile 65.06 %,
- età media 61.44 ± 11.63 anni,
- durata media del diabete 9.53 ± 8.64 anni.

- **AI-READI (Artificial Intelligence Ready and Equitable Atlas for Diabetes Insights)** [9]: Dataset multimodale flagship finanziato dal *National Institutes of Health (NIH)* attraverso il programma *Bridge2AI*, progettato specificamente per abilitare la ricerca di tecniche di AI nel diabete di tipo 2. Lo studio trasversale mira a raccogliere dati da 4,000 entro il 2026, bilanciati per etnia, genere e stadio del diabete. L'ultima versione include 1,067 partecipanti da tre siti di raccolta dati con 15 modalità diverse per un totale di 165,051 file (2.01 TB) [9]. Le modalità multimodali includono:
 - **Dati cardiovascolari:** elettrocardiogrammi a 12 derivazioni, monitoraggio Holter, dati da smartwatch.
 - **Imaging medico:** immagini *DICOM* oculari, cardiache, e di altra natura.
 - **Dati clinici strutturati:** seguendo il modello *OMOP Common Data Model (CDM)*.
 - **Sensori wearable:** dati di monitoraggio continuo seguendo standard *Open mHealth*.
 - **Dati ambientali e comportamentali:** fattori di stile di vita e ambientali.

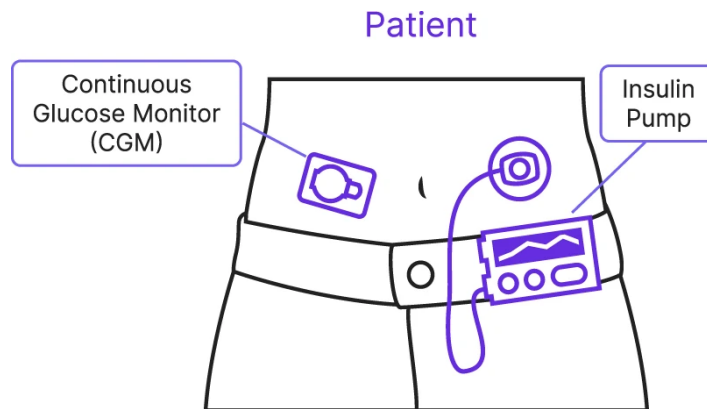


Figure 1: Metodologia di raccolta dati [6]

Dopo aver analizzato i dataset descritti precedentemente, si è proceduto alla **valutazione dell'accessibilità e disponibilità per il progetto**. Purtroppo, i dataset multimodali identificati (**mBRSET** [8] e **AI-READI** [9]) presentavano significative **barriere di accesso** che ne hanno impedito l'utilizzo nel presente progetto, non si è ricevuta alcuna risposta alle richieste di accesso effettuate. In particolare, sono state inviate **richieste formali di accesso** per entrambi i dataset alle rispettive piattaforme, tuttavia, **non sono state ricevute risposte**. Queste limitazioni, seppur comprensibili dal punto di vista etico e di protezione della privacy dei partecipanti, hanno reso necessario orientarsi verso alternative immediatamente disponibili.

Di conseguenza, è stato selezionato il dataset **HUPA-UCM** [7] come soluzione pragmatica per lo sviluppo di un modello di DL **multimodale**. Sebbene questo dataset non presenti la multimodalità classica (ad esempio, immagini e testo), offre comunque caratteristiche vantaggiose per il progetto: combina **segnali eterogenei da dispositivi diversi** (CGM, smartwatch Fitbit, input manuali) con **significati fisiologici distinti**,

permettendo lo sviluppo di modelli che integrano informazioni complementari ed eterogenee. Il dataset HUPA-UCM presenta inoltre **dati reali** raccolti in condizioni di vita quotidiana da 25 pazienti con diabete di tipo 1, è **già preprocessato e ben documentato**, facilitando l'implementazione diretta. La sua **disponibilità immediata** e la **documentazione completa** lo rendono una scelta strategica per dimostrare l'efficacia degli approcci di integrazione di segnali multimodali nel contesto diabetologico, aprendo la strada a future estensioni verso dataset comprensivi di immagini e altre tipologie di dati quando il loro accesso sarà disponibile.

4 Data Preprocessing

Il dataset HUPA-UCM [7] (vedi Sezione 3), sebbene fornito in formato preprocessato (vedi Figura 2), ha richiesto ulteriori interventi per ottimizzare l'addestramento del modello multimodale.

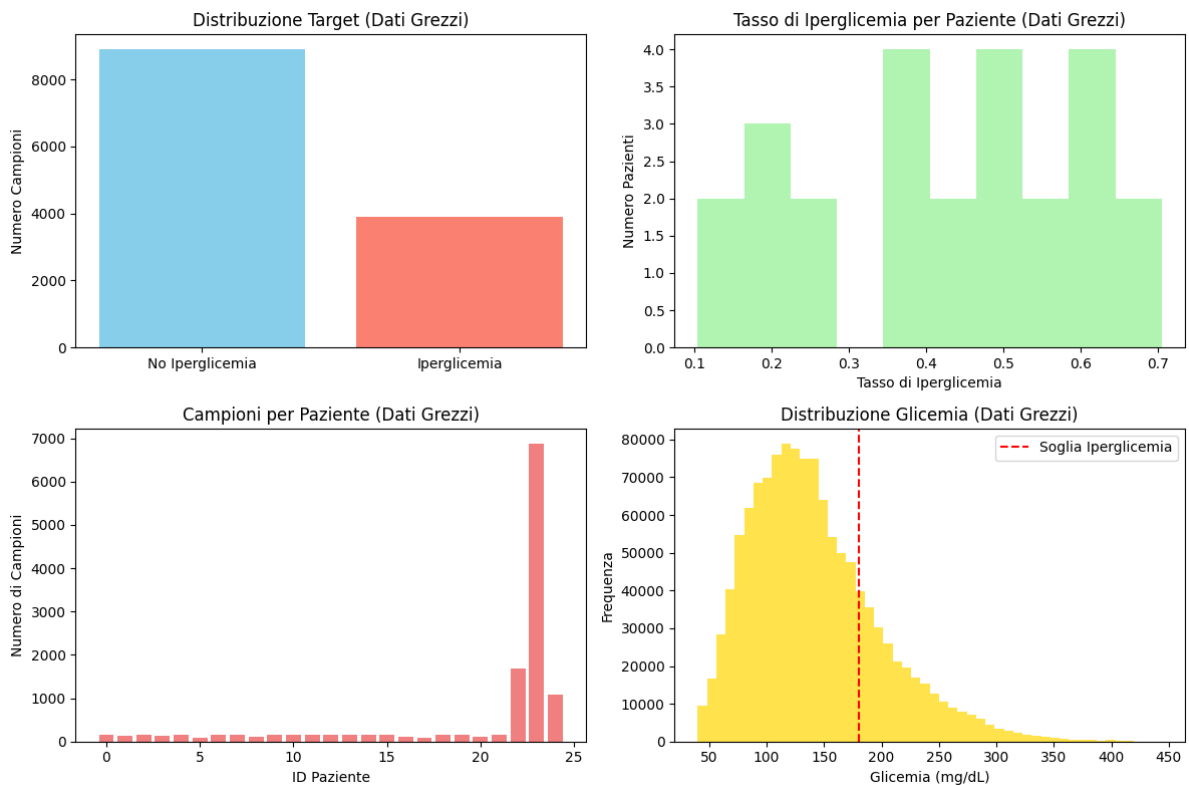


Figure 2: Analisi distribuzione del dataset HUPA-UCM prima di essere processato.

Di seguito sono illustrate le caratteristiche principali del dataset processato (vedi Figura 3).

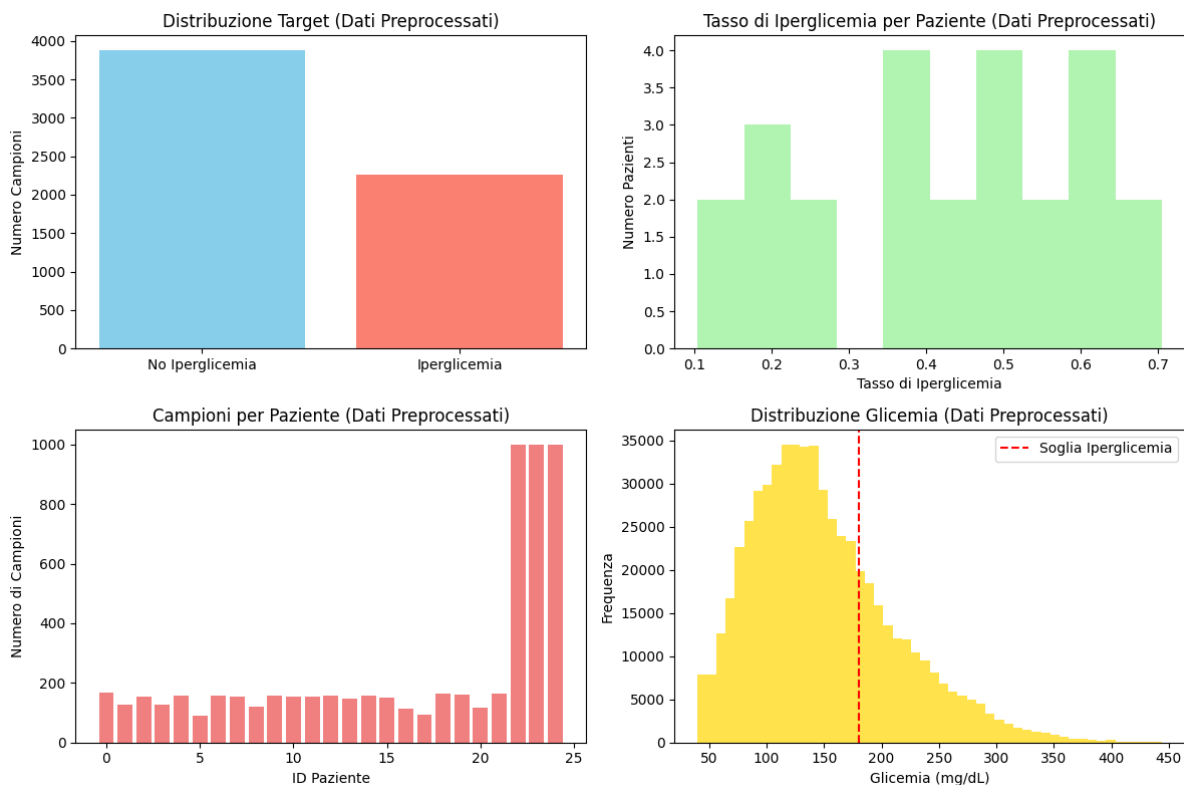


Figure 3: Analisi distribuzione del dataset HUPA-UCM processato.

Gli interventi di preprocessing attuati includono:

- **Definizione Target Iperglicemico:** Creazione delle etichette binarie basate sulla soglia clinica di **180 mg/dL** [10] per identificare episodi iperglicemici nelle successive 4 ore di monitoraggio.
- **Definizione Periodi Temporal:** Ogni campione è costruito considerando una **finestra di input di 96 timesteps** (pari a 24 ore di monitoraggio continuo) e un **orizzonte di predizione di 16 timesteps** (4 ore future). Questa scelta permette al modello di apprendere i pattern glicemici recenti e stimare con anticipo la probabilità di un episodio iperglicemico nelle ore successive.
- **Creazione Sequenze Temporal:** Generazione di *sliding window* di 96 timesteps (24 ore) con overlap del 75% per aumentare la robustezza del training e catturare pattern temporali continui, garantendo continuità temporale.
- **Bilanciamento:** Implementazione di campionamento per limitare lo sbilanciamento nel numero di sequenze per paziente (massimo 1,000 campioni per paziente). Questa tecnica previene la dominanza di pazienti con maggiore quantità di dati (fino a 7.000 campioni) (vedi Figura 2) rispetto a quelli con dati limitati (200 campioni) (vedi Figura 3), garantendo rappresentazione equa di tutti i profili metabolici nel training set.

- **Normalizzazione Multi-Modalità:** Implementazione di `StandardScaler` separati per ogni modalità di dati (CGM, Fitbit, Insulin).
- **Gestione Valori Mancanti:** Per garantire la completezza delle sequenze ed evitare interruzioni nel flusso temporale, sono state applicate diverse strategie di imputazione, specifiche per ciascuna modalità:
 - **CGM:** *Interpolazione lineare* (`interpolate(method='linear')`) per preservare la continuità temporale delle curve glicemiche. Questa scelta è motivata dal fatto che i sensori CGM forniscono dati pressoché regolari e i valori mancanti possono essere stimati in modo affidabile in base ai vicini.
 - **Fitbit:** *Imputazione con media* (`fillna(df['colonna'].mean())`) per calorie e frequenza cardiaca, in quanto queste variabili presentano distribuzioni relativamente stabili. Per i **passi**, invece, è stato applicato *zero-filling* (`fillna(0)`), poiché l'assenza di rilevazioni dovrebbe coincidere con assenza di movimento.
 - **Dati terapeutici:** *Zero-filling* (`fillna(0)`), in quanto l'assenza di somministrazioni di insulina o di assunzioni di carboidrati equivale a valori nulli.
- **Gestione Outlier:** Applicazione di soglie clinicamente appropriate per i valori glicemici (*40-500 mg/dL*) e frequenza cardiaca (*40-200 bpm*).
- **Dataset Splitting:** Suddivisione del dataset: 60% *training set*, 20% *validation set*, 20% *test set*.

5 Progettazione Modello Multimodale

L'architettura del modello (vedi Figura 4) proposto sfrutta un **approccio multi-sensor** che integra **informazioni eterogenee provenienti da diversi dispositivi di monitoraggio** per **monitorare** in maniera **proattiva** i pazienti. La progettazione segue il paradigma della **Functional API** di Keras, permettendo la creazione di un'architettura flessibile con branch specializzati per ogni tipologia di segnale.

Il modello è strutturato in quattro componenti principali:

1. **Branch Temporale CGM:** Processa le serie temporali dei valori glicemici continui utilizzando soprattutto layer LSTM bidirezionali per catturare le dipendenze temporali nei pattern glicemici.
2. **Branch Multi-Sensor Fitbit:** Gestisce simultaneamente i segnali eterogenei del dispositivo wearable (passi, calorie, frequenza cardiaca) attraverso layer LSTM bidirezionali che processano le correlazioni temporali multivariate.
3. **Branch Dati Statici:** Elabora le informazioni non temporali (dosaggio insulina, carboidrati) mediante layer densi.
4. **Fusion Layer:** Combina le rappresentazioni estratte dai branch precedenti attraverso concatenazione, proiezioni e layer densi per la predizione finale.

Il modello è progettato per predire eventi di **iperglicemia** nelle successive 4 ore, definiti secondo la **soglia clinica** di **180 mg/dL** stabilita dalle linee guida dell'**American Diabetes Association (ADA)** per pazienti con diabete di tipo 1 [10]. Questa soglia rappresenta il punto critico oltre il quale si raccomanda intervento terapeutico immediato. Il sistema fornisce quindi un **early warning**, permettendo ai pazienti e ai clinici di implementare strategie preventive prima del raggiungimento di stati iperglicemici critici. Infine, l'implementazione dell'architettura multimodale comprende **204,789 parametri addestrabili** e **12 parametri non addestrabili**, per un totale di **204,801 parametri**.

5.1 Input Layers

Il modello utilizza **tre input distinti**, corrispondenti alle diverse sorgenti di dati:

- **CGM Input:** Sequenza glicemica di 96 timestep, rappresentata come vettore unidimensionale di **shape (96)**.
- **Fitbit Input:** Sequenza multi-sensoriale di 96 timestep con 3 feature parallele (passi, calorie bruciate, frequenza cardiaca), con **shape (96, 3)**.
- **Insulin Input:** Tre valori scalari aggregati su finestra di 24 ore (dosaggio insulinico totale, tasso medio, carboidrati totali), con **shape (3)**.

5.2 Branch CGM

Il **branch CGM** elabora sequenze glicemiche di *96 timesteps* (corrispondenti a 24 ore di monitoraggio continuo con frequenza di 15 minuti). L'architettura (77,088 parametri) è composta da:

- **Reshape Layer:** Conversione dell'input unidimensionale in formato (batch_size, 96, 1) per compatibilità con i layer LSTM (0 parametri).
- **Bidirectional LSTM (64 unità):** Primo layer ricorrente con configurazione `return_sequences=True` per preservare la dimensione temporale, regolarizzato con `dropout=0.2` e `recurrent_dropout=0.1` (33,792 parametri).
- **Self-Attention Layer:** Meccanismo di attenzione che permette al modello di focalizzarsi sui pattern glicemici più rilevanti della sequenza temporale (0 parametri).
- **Bidirectional LSTM (32 unità):** Secondo layer ricorrente, che sintetizza le informazioni temporali in un vettore di feature ad alto livello, con regolarizzazione `dropout=0.2` (41,216 parametri).
- **Dense Layer (32 unità):** Layer fully-connected con attivazione ReLU, responsabile della proiezione finale in una rappresentazione compatta del branch (2,080 parametri).

5.3 Branch Multi-Sensor Fitbit

Il **branch Fitbit** elabora *sequenze multivariate* di *96 timesteps* (corrispondenti a 24 ore di monitoraggio continuo con frequenza di 15 minuti) con tre segnali fisiologici: *passi*, *calorie bruciate* e *frequenza cardiaca*. L'architettura (86,380 parametri) è composta da:

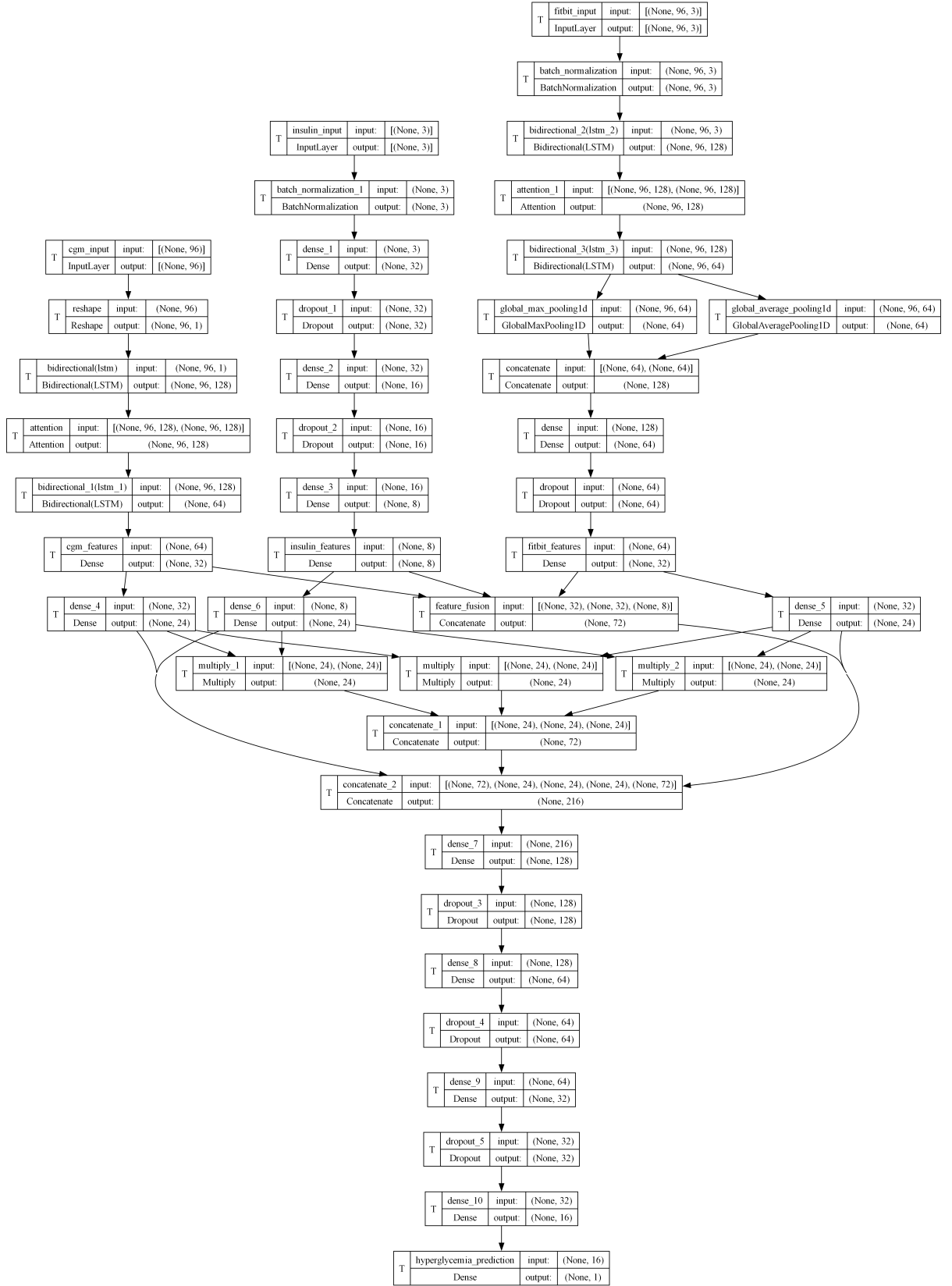


Figure 4: Architettura modello multimodale

- **Batch Normalization:** Normalizzazione degli input (`batch_size`, 96, 3) per stabilizzare il training e migliorare la convergenza (12 parametri).
- **Bidirectional LSTM (64 unità):** Primo layer ricorrente bidirezionale, configurato con `return_sequences=True`, `dropout=0.3` e `recurrent_dropout=0.1`, utile a catturare le correlazioni temporali (34,816 parametri).
- **Self-Attention Layer:** Meccanismo di attenzione per focalizzarsi sui pattern più rilevanti nella sequenza (0 parametri).
- **Bidirectional LSTM (32 unità):** Secondo layer ricorrente bidirezionale, con `return_sequences=True`, `dropout=0.3` e `recurrent_dropout=0.1`, per estrazione di feature di più alto livello (41,216 parametri).
- **Global Average Pooling 1D:** Aggrega l'informazione temporale calcolando la media lungo la sequenza (0 parametri).
- **Global Max Pooling 1D:** Complementare al pooling medio, cattura le attivazioni massime (0 parametri).
- **Concatenate:** Concatenazione delle rappresentazioni provenienti da average e max pooling, 128 dimensioni (0 parametri).
- **Dense Layer (64 unità):** Primo layer fully-connected con attivazione ReLU, usato per proiettare le feature combinate (8,256 parametri).
- **Dropout (40%):** Regularizzazione per prevenire overfitting (0 parametri).
- **Dense Layer (32 unità):** Layer fully-connected finale con attivazione ReLU, responsabile della rappresentazione compatta del branch (2,080 parametri).

5.4 Branch Dati Terapeutici

Il **branch terapeutico** elabora tre feature scalari aggregate su una finestra temporale di 24 ore (*dosaggio insulinico totale*, *tasso medio*, *carboidrati totali*). L'architettura (876 parametri) è composta da:

- **Batch Normalization:** Normalizzazione dei tre input (`batch_size`, 3) per stabilizzare l'addestramento (12 parametri).
- **Dense Layer (32 unità):** Primo layer fully-connected con attivazione ReLU per espandere la rappresentazione (128 parametri).
- **Dropout (40%):** Regularizzazione per prevenire overfitting (0 parametri).
- **Dense Layer (16 unità):** Secondo layer fully-connected intermedio con attivazione ReLU (528 parametri).
- **Dropout (30%):** Ulteriore regularizzazione per ridurre la complessità e il rischio di overfitting (0 parametri).
- **Dense Layer (8 unità):** Layer intermedio compatto con attivazione ReLU (136 parametri).

- **Dense Layer (8 unità):** Layer finale del branch con attivazione ReLU, che produce la rappresentazione compatta (72 parametri).

5.5 Fusion Layer

Il **Fusion Layer** integra le rappresentazioni provenienti dai tre branch descritti precedentemente (*CGM*, *Fitbit* e *Dati Terapeutici*). L'architettura complessiva per la fusione (40,440 parametri) è composta da:

- **Dense Layer (24 unità):** Le feature risultanti dai branch precedenti vengono proiettate in spazi latenti da 24 unità, permettendo una rappresentazione comparabile tra le diverse modalità.
- **Interazioni Cross-Modali:** Vengono introdotte moltiplicazioni element-wise tra coppie di modalità, evidenziando pattern condivisi e dipendenze tra sorgenti eterogenee.
- **Concatenation:** Le feature dei tre branch vengono concatenate.
- **Fusion Finale:** Tutte le feature (“originali”, “proiettate” e “interazioni”) vengono ricombinate in un vettore unico e passate attraverso una gerarchia di layer fully-connected:
 - Dense (128 unità, ReLU) + Dropout (50%).
 - Dense (64 unità, ReLU) + Dropout (40%).
 - Dense (32 unità, ReLU) + Dropout (30%).
 - Dense (16 unità, ReLU).

5.6 Output Layer

L'**Output Layer** costituisce il layer finale del modello multimodale ed è implementato come una **unità singola densa** con attivazione **sigmoid** (17 parametri), che restituisce la probabilità di un episodio di iperglicemia nelle **4 ore successive**. Infine, la scelta architetturale privilegia la **specializzazione** di ogni branch per il tipo di dato specifico, seguita da una **fusione** che preserva le caratteristiche distintive apprese nei branch e layer precedenti.

5.7 Metriche di Monitoraggio e Valutazione

La **valutazione delle prestazioni** del modello utilizza il seguente set di metriche per il task di classificazione binaria: **accuracy**, **precision**, **recall** e **AUC (Area Under ROC Curve)**

5.8 Tecniche di Ottimizzazione e Regularizzazione

L'architettura implementa molteplici strategie per prevenire overfitting, migliorare la generalizzazione e ottimizzare il processo di addestramento:

- **Batch Size Ottimizzata:** Dimensione batch di 32 campioni per bilanciare stabilità del gradiente ed efficienza computazionale.

- **Early Stopping:** Monitoraggio della *validation loss* con `patience=15` (epoche) e ripristino dei migliori pesi, mediante `restore_best_weights=True`, per prevenire overfitting e ridurre i tempi di training.
- **Learning Rate Reducing:** Riduzione automatica del learning rate con `factor=0.5` quando la *validation loss* stagna per 8 epoche consecutive, con soglia minima impostata tramite `min_lr=1e-6`.
- **Model Checkpointing:** Salvataggio automatico del modello migliore per garantire conservazione delle prestazioni ottimali.
- **Dropout:** Applicazione di vari layer di dropout di diversa entità, come precedentemente descritto nel dettaglio, per regolarizzazione adattiva.

6 Risultati Training

I risultati del training del **modello multimodale** (vedi Figura 5) mostrano un andamento **complessivamente positivo** delle metriche principali durante le 56 epoche di addestramento, interrotte da *early stopping*. Il modello evidenzia una **convergenza stabile**, con curve di training e validation che rimangono vicine tra loro, indicando una buona capacità di generalizzazione e assenza di fenomeni di overfitting marcato.

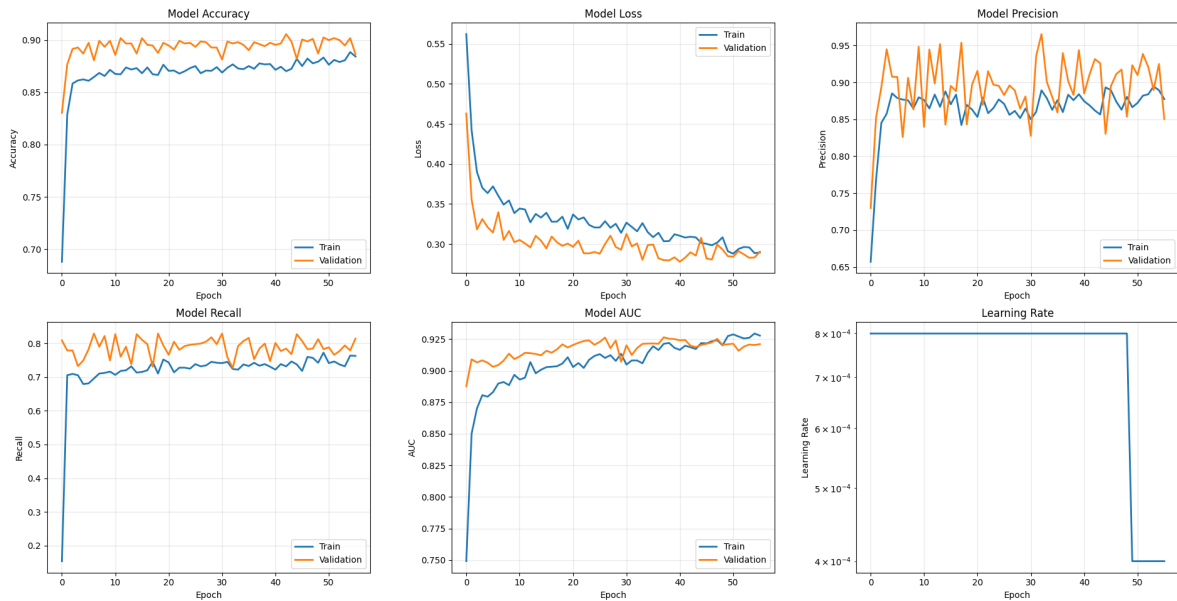


Figure 5: Risultati training

Alla fine dell'addestramento (epoca 56), il modello ha ottenuto le seguenti performance:

- **Training:** Loss = 0.2895, Accuracy = 0.8841, AUC = 0.9278, Precision = 0.8774, Recall = 0.7626

- **Validation:** Loss = 0.2905, Accuracy = 0.8863, AUC = 0.9211, Precision = 0.8505, Recall = 0.8141

Le **migliori prestazioni** sono state osservate all'epoca 41.

7 Testing e Validazione

I risultati sperimentali sono stati ottenuti attraverso una valutazione sistematica su un **test set di 1,519 campioni** provenienti da pazienti mai visti durante training e validation, garantendo una stima realistica delle prestazioni di generalizzazione.

7.1 Prestazioni del Modello Multimodale

La **valutazione comparativa rispetto ai modelli baseline unimodali** evidenzia pattern chiari nelle prestazioni: il **modello multimodale** mostra un **vantaggio consistente** in tutte le metriche considerate, come riportato nella Tabella 1, dove i migliori valori sono evidenziati in grassetto.

È interessante osservare come i modelli *Fitbit-only* e *Insulin-only* presentino performance sensibilmente inferiori, segnalando una scarsa capacità predittiva se utilizzati isolatamente. Al contrario, la baseline *CGM-only* ottiene risultati comparabili al multimodale, confermando il ruolo centrale dei segnali glicemici continui come principale fonte informativa. Infine, l'integrazione multimodale, pur partendo da una forte base legata al CGM, consente un ulteriore incremento in termini di metriche garantendo maggiore robustezza predittiva.

Metrica	Multimodale	CGM-only	Fitbit-only	Insulin-only
Accuracy	0.8749	0.8713	0.6545	0.6545
Precision	0.8998	0.8891	-	-
Recall	0.8167	0.7788	-	-
AUC	0.9228	0.9007	0.5659	0.4651

Table 1: Confronto modelli

7.2 Analisi delle Research Questions

7.2.1 Risposta a RQ1: Efficacia dell'Approccio Multimodale

Il **modello multimodale supera sistematicamente le baseline unimodali** in tutte le metriche utilizzate (accuracy, precision, recall, AUC), confermando che la fusione di segnali eterogenei consente una migliore capacità predittiva rispetto all'utilizzo di singole sorgenti informative. Tuttavia, è interessante osservare come la *baseline CGM-only* si avvicini maggiormente alle prestazioni del modello multimodale, confermando il ruolo centrale dei segnali glicemici continui come principale fonte informativa. L'integrazione di ulteriori sorgenti informative contribuisce a raffinare, seppur di poco, le previsioni, garantendo una maggiore robustezza rispetto a modelli basati su una singola modalità.

7.2.2 Risposta a RQ2: Generalizzazione e Robustezza

L'impiego di un **test set composto esclusivamente da pazienti non visti in fase di training** e i risultati positivi ottenuti evidenziano la capacità del modello multimodale di **generalizzare a nuovi individui**.

8 Sviluppi Futuri

Il lavoro ha dimostrato l'efficacia dell'integrazione multimodale di dati glicemici, terapeutici e da sensori wearable per la predizione di episodi iperglicemici. Tuttavia, sono stati individuati i seguenti **sviluppi futuri**:

- **Espansione dei dataset**: Ulteriore addestramento e validazione con i dataset multimodali individuati: *AI-READI* e *mBRSET* (vedi Sezione 3).
- **Integrazione di nuove modalità**: Aggiungere dati testuali (referti clinici, diari alimentari) e immagini (retiniche, radiologiche) per estendere la capacità multimodale del modello.

9 Conclusioni

In questo lavoro è stato sviluppato e validato un modello multimodale di deep learning capace di integrare **dati eterogenei** da **CGM**, dispositivi **Fitbit** e **informazioni terapeutiche**. I risultati hanno evidenziato che l'**approccio multimodale supera sistematicamente i modelli unimodali**, confermando il valore aggiunto derivante dalla fusione di diverse sorgenti informative.

Bibliografia

- [1] D. J. Magliano, E. J. Boyko, I. Genitsaridi, L. Piemonte, P. Riley, and P. Salpea, Eds., *IDF Diabetes Atlas, 11th Edition*. International Diabetes Federation, 2025. [Online]. Available: <https://diabetesatlas.org>
- [2] S. Gundapaneni, Z. Zhi, and M. Rodrigues, "Deep learning-based noninvasive screening of type 2 diabetes with chest x-ray images and electronic health records," 12 2024.
- [3] K. Wardhani, S. Kasim, A. Erianda, and R. Hassan, "Deep learning-based method in multimodal data for diabetic retinopathy detection," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 14, pp. 1602–1608, 10 2024.
- [4] J.-E. Ding, P. N. M. Thao, W.-C. Peng, J.-Z. Wang, C.-C. Chug, M.-C. Hsieh, Y.-C. Tseng, L. Chen, D. Luo, C. Wu *et al.*, "Large language multimodal models for new-onset type 2 diabetes prediction using five-year cohort electronic health records," *Scientific reports*, vol. 14, no. 1, p. 20774, 2024.
- [5] M. S. Haleem, D. Katsarou, E. I. Georga, G. E. Dafoulas, A. Bargiota, L. Lopez-Perez, M. Rujas, G. Fico, L. Pecchia, D. Fotiadis *et al.*, "A multimodal deep learning architecture for predicting interstitial glucose for effective type 2 diabetes management," *Scientific Reports*, vol. 15, no. 1, p. 27625, 2025.

- [6] T. Prioleau, A. Bartolome, R. Comi, and C. Stanger, “Diatrend: A dataset from advanced diabetes technology to enable development of novel analytic solutions,” 04 2023.
- [7] J. I. Hidalgo, J. Alvarado, M. Botella, A. Aramendi, J. M. Velasco, and O. Garnica, “Hupa-ucm diabetes dataset,” *Data in Brief*, vol. 55, p. 110559, 2024.
- [8] L. F. Nakayama, L. Zago Ribeiro, D. Restrepo *et al.*, “mbrset, a mobile brazilian retinal dataset,” *PhysioNet* <https://doi.org/10.13026/QXPD-1Y65>, 2024.
- [9] P. M. <http://orcid.org/0000-0001-6343-2140> Drolet Caroline 4 <http://orcid.org/0000-0003-2287-4190> Lucero Abigail 8 Matthies Dawn 7 <http://orcid.org/0009-0003-4909-6058> Pittock Hanna 3 Watkins Kate 3 York Brittany 1 and N. P. S. W. X. 11, “Ai-readi: rethinking ai data collection, preparation and sharing in diabetes research and beyond,” *Nature metabolism*, vol. 6, no. 12, pp. 2210–2212, 2024.
- [10] A. D. A. P. P. Committee, “6. glycemic goals and hypoglycemia: Standards of care in diabetes—2025,” *Diabetes Care*, vol. 48, no. Supplement_1, pp. S128–S145, 01 2025. [Online]. Available: <https://doi.org/10.2337/dc25-S006>