UNIVERSITÀ DEGLI STUDI DI PADOVA

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN BIOINGEGNERIA

Analisi e sviluppo di algoritmi in tempo reale per identificazione automatica dei pasti nel diabete di tipo 1

Relatore: Prof. Andrea Facchinetti

Correlatore: Dott. Luca Cossu

Laureanda: Elena Diurno

ANNO ACCADEMICO 2024– 2025

Data di laurea 08-10-2025

A chi crede che l'istruzione sia il primo mattone della pace.
Alle donne e agli uomini che con la conoscenza aprono strade,
costruiscono ponti e spengono i fuochi dell'odio.
Perché un mondo senza guerra nascerà solo
dove ogni mente avrà avuto il diritto di imparare.

ABSTRACT

Il diabete mellito di tipo 1 (T1DM) è una malattia autoimmune caratterizzata dalla distruzione delle cellule β pancreatiche, responsabili della produzione di insulina. In assenza di questa funzione, le persone affette necessitano di una terapia insulinica esogena, basata sulla somministrazione di insulina basale e boli ai pasti. Nonostante i notevoli progressi tecnologici, tale trattamento non riesce a replicare in modo ottimale la regolazione fisiologica della glicemia, esponendo i pazienti al rischio di episodi di ipoglicemia e iperglicemia, che a lungo termine possono favorire lo sviluppo di complicanze croniche. Per supportare la gestione della terapia, negli ultimi anni sono stati sviluppati sistemi di supporto alla decisione che sfruttano i dati forniti dai sensori per il monitoraggio continuo della glicemia (Continuous Glucose Monitoring, CGM). Questi sistemi, tuttavia, richiedono spesso l'inserimento manuale di informazioni fondamentali, come l'orario e il contenuto dei pasti. Tale procedura risulta impegnativa, soggetta a dimenticanze o stime imprecise, con il rischio di ridurne l'efficacia.

In questo contesto, l'identificazione automatica e in tempo reale dei pasti rappresenta un obiettivo strategico. Automatizzare questo processo è fondamentale perché riduce l'onere del monitoraggio manuale, limita gli errori dovuti a dimenticanze o stime imprecise e permette un intervento tempestivo, migliorando il controllo glicemico e la qualità della vita dei pazienti.

Questa tesi si propone di analizzare e confrontare i principali algoritmi presenti in letteratura per il rilevamento automatico dei pasti, valutandone l'efficacia sia quando si basano esclusivamente sull'informazione glicemica fornita da sensori CGM, sia quando integrano dati esterni, come i diari terapeutici e il segnale della frequenza cardiaca.
Tutti gli algoritmi considerati sono stati testati su dati reali provenienti dallo studio Type 1 Diabetes Exercise Initiative, che ha monitorato per circa quattro settimane giovani adulti con T1DM impegnati in attività fisica regolare. Il confronto ha permesso di evidenziare punti di forza e i limiti di ciascun approccio, offrendo una panoramica utile per orientare futuri sviluppi in questo ambito. Inoltre, l'analisi ha incluso anche un'indagine critica delle performance ottenute, valutando possibili cause della diversità delle performance tra soggetti e metodologie diverse, con particolare attenzione alla presenza di segnali rumorosi o anomalie nei dati. Questo ha permesso di riflettere sulla robustezza e sulla generalizzabilità degli algoritmi, valutando se il loro comportamento sia uniforme o influenzato da caratteristiche individuali.

INDICE

Abstract *4*

1.Introduzione *8*

1.1 Il diabete mellito di tipo 1: caratteristiche e terapie **8**

1.2 Tecnologie di supporto per il monitoraggio del diabete **10**

1.3 L' importanza dell' identificazione dei pasti nella terapia: problematica aperta **13**

1.4 Scopo e struttura della tesi **14**

2. Algoritmi per l' identificazione dei pasti: stato dell' arte *16*

2.1 Approcci esistenti: offline vs real-time **16**

2.2 Metodologie analizzate per rilevazione pasti da segnali CGM e fisiologici ..**17**

2.2.1 Algoritmo multi variabile per pancreas artificiale di Samadi et al. 17

2.2.2. Algoritmo basato sul tasso di incremento glicemico di Harvey et al. 20

2.2.4 Algoritmo fisiologico con simulazione predittiva di Zheng et al. 28

2.2.5 Algoritmo HR+CGM di Cossu et al 32

2.2.6 Il metodo RAP sviluppato da Mosquera-Lopez et al. 36

3. Dataset e fase di preprocessing *39*

3.1 Dataset T1DEXI: descrizione, soggetti e raccolta **39**

3.2 Analisi e preprocessing dei dati **40**

4. Valutazione sperimentale e risultati *45*

4.1 Metriche di valutazione **45**

4.2 Risultati ottenuti **48**

4.2.1. Valutazione su singolo paziente 48

4.2.3 Validazione dell'analisi con estensione del margine temporale di rilevamento 53

4.2.4 Valutazione dell'algoritmo HR+CGM con l'ausilio della derivata seconda del glucosio 56

5. Discussione *58*

5.1 Analisi delle metriche di rumore e correlazione con le performance **58**

5.2 Interpretazione della correlazione tra variabili **61**

5.3 Analisi locale del rumore nei singoli eventi di detection **64**

5.3.1 Rappresentazione grafica delle distribuzioni per categoria di evento 65

5.4 Ranking delle performance algoritmiche.....68

5.5 Sintesi conclusiva degli algoritmi valutati 71

6.Conclusioni e sviluppi futuri..... 73

6.1 Limiti dello studio.....73

6.2 Sviluppi futuri.....74

Bibliografia 78

Ringraziamenti..... 81

CAPITOLO 1

1.Introduzione

1.1 Il diabete mellito di tipo 1: caratteristiche e terapie

Il diabete mellito (DM) è un disturbo metabolico caratterizzato dalla presenza di iperglicemia cronica accompagnata da una maggiore o minore compromissione del metabolismo di carboidrati, lipidi e proteine. Il DM è probabilmente una delle malattie più antiche conosciute dall'uomo.[1]

È causato da una carenza nella produzione di insulina da parte delle cellule del pancreas, da una inefficacia dell'insulina prodotta, o da entrambe le condizioni.

L'insulina è un ormone prodotto dalle cellule β del pancreas che permette al glucosio presente nel sangue di entrare nelle cellule, dove viene utilizzato come fonte di energia. Inoltre, regola il metabolismo dei carboidrati, dei grassi e delle proteine e favorisce la crescita e la divisione cellulare. L'assenza o l'inefficacia dell'insulina è la causa principale del T1DM.[2]

Esistono differenti tipologie di diabete mellito e ad oggi la comunità medica ha formulato una classificazione completa e di facile consultazione. L'organizzazione mondiale della sanità (OMS) e l'associazione americane del diabete (ADA) hanno evidenziato tre principali tipologie di diabete.[3]

1. Diabete mellito di tipo 1(*Type 1 Diabetes Mellitus, T1DM*), malattia autoimmune causata dalla distruzione delle cellule β del pancreas, che porta solitamente a una carenza assoluta di insulina. È una patologia tipica degli adolescenti.[3]
2. Diabete mellito di tipo 2*(Type 2 Diabetes Mellitus, T2DM)*, condizione che inizia a causa di un'insensibilità dei tessuti all'insulina e si consolida con una ridotta produzione di quest'ultima. Si sviluppa nella popolazione adulta. [3]

3. Diabete gestazionale (*Gestational Diabetes, GD*), stato di iperglicemia diagnosticato nella paziente durante la gravidanza. Si presenta a causa dell'azione di alcuni ormoni della placenta che generano insulino-resistenza. [3]

Nel T1DM, l'assenza di insulina determina tre possibili condizioni glicemiche che ne caratterizzano la gestione quotidiana:

- Iperglicemia (>180 mg/dL): è la manifestazione clinica tipica del T1DM non adeguatamente controllato. Si verifica quando si instaura uno squilibrio tra l'assunzione di carboidrati, l'attività fisica e la terapia insulinica sostitutiva. Altri fattori scatenanti includono stress fisico o psicologico, infezioni acute o patologie intercorrenti, che aumentano il fabbisogno insulinico e compromettono il controllo glicemico [2,4].
- Ipoglicemia (<70 mg/dL): rappresenta la complicanza iatrogena più frequente e temuta nel trattamento del T1DM. Può insorgere a seguito di una somministrazione eccessiva di insulina rispetto ai carboidrati assunti, attività fisica intensa non compensata, errori nel dosaggio o nella tempistica della terapia insulinica, oppure ritardi significativi nei pasti programmati [5].
- Euglicemia (70-180 mg/dL)**:** è l'obiettivo terapeutico principale nella gestione del T1DM. Si definisce come il mantenimento di valori glicemici il più possibile simili a quelli di soggetti non diabetici. Raggiungere e sostenere l'euglicemia richiede un equilibrio dinamico tra terapia insulinica (basale e prandiale), apporto nutrizionale, attività fisica e gestione dello stress. Il mantenimento dell'euglicemia è fondamentale per prevenire complicanze microvascolari e macrovascolari e per garantire una buona qualità di vita [6],[7].

Dal punto di vista epidemiologico, in Italia il 4,9% della popolazione è affetto da diabete mellito, di cui circa il 10% presenta T1DM, per un totale stimato di circa 300.000 casi. A livello globale, il T1DM rappresenta il 10-15% di tutti i casi di diabete, mentre il diabete di tipo 2 è la forma più comune. Tuttavia, il T1DM è la forma predominante nei bambini sotto i 15 anni, con oltre mezzo milione di bambini attualmente affetti a livello mondiale [8].

Questi dati evidenziano l'impatto significativo del T1DM sulla sanità pubblica e sottolineano l'urgenza di sviluppare soluzioni tecnologiche innovative per ottimizzare la gestione terapeutica e migliorare la qualità della vita delle persone con questa patologia cronica [9].

1.2 Tecnologie di supporto per il monitoraggio del diabete

Negli ultimi decenni, abbiamo assistito a progressi straordinari nelle tecnologie per la gestione del T1DM, con l'obiettivo di sviluppare approcci personalizzati e automatizzati che possano migliorare la vita dei pazienti.

I principali strumenti per la gestione del diabete rientrano in due categorie: strumenti per la misurazione della glicemia e strumenti per la somministrazione dell'insulina. [6] Nel corso del tempo, si è verificata una significativa evoluzione nel monitoraggio della glicemia. Infatti, negli anni '50, non esistevano dispositivi in grado di misurare i livelli di glucosio nel sangue in tempo reale; iniezioni di insulina, diete equilibrate e attività fisica regolare erano i pilastri del trattamento.

La situazione è migliorata negli anni '80 con l'introduzione di dispositivi noti come Self Monitoring of Blood Glucose (SMBG). Questi strumenti utilizzano un misuratore interno che consente la visualizzazione del valore glicemico, espresso in mg/dL, e un pungidito per prelevare una piccola goccia di sangue capillare dal polpastrello. Grazie all'SMBG, è possibile misurare la glicemia 3-4 volte al giorno, soprattutto prima dei pasti, per calcolare il bolo di insulina corretto per compensare l'escursione glicemica. [6]

Tuttavia, questa procedura è invasiva a causa della necessità di prelevare sangue più volte al giorno. Inoltre, l'SMBG (Figura 1) non riesce sempre a individuare eventi di iperglicemia o ipoglicemia, a causa delle ampie e repentine variazioni glicemiche che si verificano durante la giornata in seguito alla dieta, all'attività fisica e alla terapia farmacologica.

Figura 1: Dispositivo pungi-dito per l'SMBG

Per questo motivo, per consentire un monitoraggio più frequente, solitamente ogni 1-5 minuti, sono stati sviluppati i dispositivi Continuous Glucose Monitoring (CGM).

Questi sensori sono minimamente invasivi perché non misurano la glicemia nel sangue, ma nel fluido interstiziale, cioè nello spazio tra cellule e capillari dove avvengono gli scambi di sostanze, incluso il glucosio.[10]

Il monitoraggio avviene con sensori sottocutanei, fornendo misurazioni frequenti e meno invasive tramite un sistema composto da sensore, monitor e dispositivo di elaborazione, spesso con comunicazione wireless. [11]

Figura 2 : Esempio di dispositivo per il monitoraggio in continua del glucosio

Inoltre, la maggior parte dei sistemi CGM è dotata di allarmi che avvertono il paziente in caso di variazioni glicemiche imminenti, consentendo una gestione più efficace della malattia e aumentando il livello di sicurezza.

Per quanto riguarda la somministrazione dell'insulina nei pazienti diabetici si prevedeva l'uso di siringhe con aghi grandi, ingombranti e poco pratici. *(Figura 3a*)

Con l'introduzione di penne insuliniche precaricate, la procedura è diventata più semplice e precisa; infatti, può essere eseguita in modo autonomo e sicuro. *(Figura 3b*)

La svolta nella terapia insulinica è avvenuta con l'introduzione della pompa per insulina, un piccolo dispositivo computerizzato in grado di somministrare insulina in modo continuo, sia durante il giorno che durante la notte. Analogamente al pancreas, essa rilascia piccole dosi di insulina ad azione rapida per mantenere stabile il livello glicemico tra un pasto e l'altro e durante il sonno. *(Figura 3c*)[12]

Con l'evoluzione della biomedicina sono in via di sviluppo degli analoghi dell'insulina grazie ai quali può essere migliorata la stabilità dell'azione insulinica e la flessibilità adattiva del trattamento.[7]

Figura 3: Evoluzione della somministrazione dell'insulina: (a) siringhe tradizionali; (b) penne insuliniche precaricate; (c) pompe per insulina, dispositivi computerizzati che erogano insulina in modo continuo.

Oggi le combinazioni di infusori e sensori glicemici hanno raggiunto un nuovo livello, creando sistemi adattivi che cambiano automaticamente la quantità di dosi somministrate.

1.3 L'importanza dell'identificazione dei pasti nella terapia: problematica aperta

Nel trattamento del diabete, e in particolare in T1DM**,** il momento dell'assunzione dei pasti riveste un ruolo fondamentale per il controllo glicemico. Dopo un pasto, infatti, la glicemia tende ad aumentare, rendendo necessaria nei soggetti con T1DM una somministrazione insulinica tempestiva, così da prevenire scompensi metabolici.[13]

La segnalazione manuale dei pasti da parte del paziente avviene tramite l'inserimento dell'orario di assunzione del pasto o il conteggio dei carboidrati nei dispositivi di monitoraggio, però risulta spesso imprecisa o del tutto assente.

Questo limita l'efficacia della terapia insulinica, compromettendo il controllo della glicemia postprandiale e incrementando la variabilità glicemica. Per questo motivo, i sistemi moderni di identificazione dei pasti stanno cercando di integrare la rilevazione automatica del pasto attraverso l'analisi continua dei trend glicemici, sfruttando algoritmi predittivi e tecnologie di monitoraggio in tempo reale.

Questi sistemi, però, presentano ancora numerose limitazioni significative, tra cui l'utilizzo di soglie glicemiche fisse, l'alta incidenza del tasso di eventi falsi positivi, la ridotta capacità di adattarsi ai principali fattori che impattano sulla vita quotidiana quali stress e attività e in particolare il ritardo nel rilevare l'assunzione del pasto.[14],[15]
Queste criticità spiegano perché l'identificazione automatica dei pasti rappresenti una delle principali sfide ancora aperte nella moderna gestione del diabete.
Inoltre, rispetto ad altri segnali fisiologici, la risposta glicemica ai pasti nei soggetti con T1DM non sempre si manifesta con variazioni marcate e facilmente distinguibili. L'aumento post-prandiale può risultare attenuato dall'azione dell'insulina somministrata, smussato dal ritardo dei sensori CGM o mascherato da fattori concomitanti quali attività fisica, stress o variazioni ormonali. Questa mancanza di risposte nette rende ancora più complesso il compito degli algoritmi di rilevazione automatica.
Senza una chiara individuazione dell'evento alimentare, l'erogazione di insulina può essere ritardata o inadeguata, aumentando il rischio sia di iperglicemia che di ipoglicemia.
Il rischio di ipoglicemia aumenta ulteriormente quando i pasti vengono posticipati o saltati, nei soggetti trattati con insulina o farmaci ipoglicemizzanti. Questo fenomeno può verificarsi anche durante la notte (ipoglicemia notturna) e passare inosservato nei pazienti con ridotta percezione dei sintomi (ipoglicemia inconsapevole) [16], con possibili implicazioni pericolose, soprattutto durante l'attività fisica o la guida.
Queste difficoltà spiegano perché l'identificazione automatica dei pasti rappresenti una delle principali sfide ancora aperte nella moderna gestione del diabete.

1.4 Scopo e struttura della tesi

Come evidenziato nelle sezioni precedenti, l'identificazione tempestiva dei pasti rappresenta una delle principali sfide nella gestione del T1DM, influenzando direttamente il rischio di iper- e ipoglicemia e la qualità del controllo glicemico.
In questo contesto, automatizzare il processo di monitoraggio dei pasti è un obiettivo chiave per migliorare il controllo glicemico e ridurre il carico di gestione dei pazienti.
Lo scopo di questa tesi è analizzare e confrontare i principali algoritmi presenti in letteratura per il rilevamento automatico e in tempo reale dei pasti, sfruttando segnali provenienti da dispositivi indossabili per il monitoraggio continuo della glicemia, informazioni fisiologiche aggiuntive come la frequenza cardiaca, e dati terapeutici come la somministrazione di insulina. L'obiettivo è valutare l'efficacia e la robustezza di tali approcci su dati reali, al fine di identificare le soluzioni più promettenti per un'integrazione futura nei sistemi di gestione automatizzata del diabete.

Partendo da una revisione sistematica delle metodologie esistenti, con attenzione agli approcci real-time, e alla varietà di segnali utilizzati, si descrive poi il dataset utilizzato, che comprende dati glicemici, frequenza cardiaca e informazioni terapeutiche raccolti su soggetti con T1DM in condizioni reali e durante attività fisica regolare.
In questo contesto, vengono testati e confrontati diversi algoritmi, analizzandone le prestazioni in termini di accuratezza, tempestività, robustezza al rumore e generalizzabilità tra soggetti.
La tesi si conclude con una discussione critica dei risultati ottenuti, mettendo in luce i principali punti di forza e debolezza delle soluzioni analizzate. Vengono inoltre proposte riflessioni e indicazioni per lo sviluppo di futuri algoritmi più adattivi, precisi e resilienti, in grado di integrarsi efficacemente nei sistemi di supporto decisionale per la gestione del diabete.

CAPITOLO 2

2. Algoritmi per l'identificazione dei pasti: stato dell'arte

2.1 Approcci esistenti: offline vs real-time

Nel contesto degli algoritmi per l'identificazione automatica dei pasti tramite monitoraggio continuo del glucosio, si distinguono principalmente due strategie opposte: approcci offline e approcci real-time.

L'approccio offline prevede l'elaborazione dei dati glicemici a posteriori. Questa modalità consente l'impiego di modelli complessi, inclusi metodi di machine learning avanzati, e permette una valutazione più accurata delle prestazioni. Tuttavia, la natura retrospettiva limita la possibilità di fornire interventi immediati durante o subito dopo l'assunzione del pasto, rendendo l'approccio offline poco adatto a scenari che richiedono feedback tempestivi o supporto decisionale in tempo reale.[17]

L'approccio real-time, invece, mira a rilevare il pasto quasi istantaneamente, elaborando i dati man mano che vengono acquisiti [18]. Questo paradigma è particolarmente promettente per la gestione dinamica e personalizzata del diabete, poiché può abilitare suggerimenti terapeutici immediati. Tuttavia, la latenza fisiologica intrinseca del glucosio interstiziale introduce un ritardo inevitabile nella rilevazione, che può variare tipicamente tra 9 e 45 minuti.

Per raggiungere tali prestazioni, gli algoritmi real-time devono essere semplificati e ottimizzati, riducendo al minimo la complessità computazionale e privilegiando strategie leggere.

Nonostante queste sfide, l'approccio real-time offre il vantaggio di fornire feedback immediati, abilitando un controllo glicemico più reattivo. La variabilità dei tempi di rilevamento e la necessità di validazione estesa in contesti reali restano però ostacoli critici che ne limitano, ad oggi, l'adozione su larga scala.

Alla luce di queste considerazioni, la sezione successiva analizzerà in dettaglio i principali algoritmi proposti in letteratura per la rilevazione dei pasti, con l'obiettivo di evidenziare caratteristiche, prestazioni e limiti applicativi.

2.2 Metodologie analizzate per rilevazione pasti da segnali CGM e fisiologici

2.2.1 Algoritmo multi variabile per pancreas artificiale di Samadi et al.

Descrizione dettagliata del metodo originale

L'algoritmo proposto da Samadi et al. [19] introduce un modulo di meal detection basato su una rappresentazione qualitativa del segnale glicemico e su un sistema fuzzy per la stima dei carboidrati ingeriti.

Il rilevatore è progettato per minimizzare l'interazione del paziente e rendere il sistema il più possibile autonomo, in grado di interpretare l'andamento del glucosio monitorato in continuo. Il rilevatore si basa su un sistema che analizza in tempo reale la serie temporale dei valori di glucosio sottocutaneo, con campionamento ogni 5 minuti, provenienti dal sensore CGM. La logica fuzzy consente di descrivere in maniera qualitativa la variazione della glicemia, evitando l'impiego diretto di modelli matematici della dinamica glucosio-insulina, la cui parametrizzazione risulta spesso complessa e soggetta a elevata variabilità inter-individuale e intra-individuale.

L'algoritmo opera analizzando l'andamento del glucosio a crescere, stabilizzarsi o decrescere, a partire dalle prime e seconde derivate numeriche del segnale. A ciascuna combinazione di queste derivate viene assegnato un valore qualitativo, che rappresenta un determinato tipo di variazione glicemica (ad esempio, un rapido aumento o una lenta discesa), permettendo di codificare in modo semplice e interpretabile l'andamento del glucosio nel tempo.

Il processo di identificazione del pasto prevede che ad ogni campionamento vengano calcolate la prima derivata discreta $dG(n)$, indicativa della velocità di variazione del glucosio, e la seconda derivata $d^2G(n)$ che misura l'accelerazione o la decelerazione del trend glicemico. In base ai segni di queste derivate, il sistema assegna una delle sette "forme qualitative" predefinite (A–G), ognuna caratterizzata da un valore simbolico $I_q(n)$ che sintetizza l'intensità e la direzione del cambiamento (da rapido calo a rapida crescita).

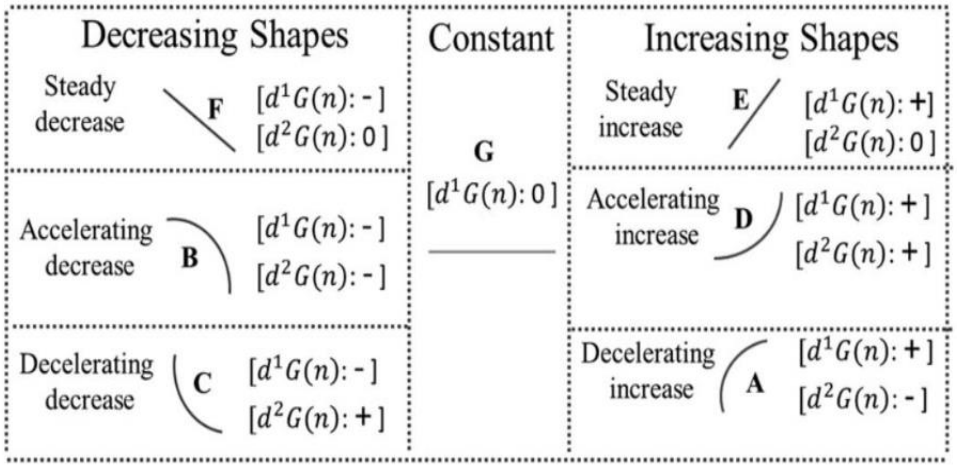

Figura 4 : Forme qualitative predefinite presa da [19]

La variabile di rilevamento $I(n)$ viene quindi ricavata come media ponderata fuzzy dei valori $I_q(n)$ all'interno di una finestra temporale mobile, producendo un indice continuo compreso tra –3 e +3.

Su $I(n)$ si basa la gestione della cosiddetta "meal flag", una variabile discreta che può trovarsi in tre stati: inattiva (nessun pasto in corso), attiva (pasto in corso) o in pausa (rilevamento sospeso per questioni di sicurezza, ad esempio per rischio di ipoglicemia o attività fisica). Quando $I(n)$ supera una soglia superiore, la meal flag si attiva; se scende sotto una soglia inferiore, si disattiva; se resta in un intervallo intermedio mentre permangono condizioni di rischio, entra in pausa. Durante lo stato attivo, il sistema stima in modo incrementale i carboidrati assunti: a ogni passo temporale calcola un contributo $\Delta CHO(n)$ che viene sommato alle stime precedenti per ottenere il totale. Questa stima si basa su due input principali: il rapporto $R(n) = \frac{D(n)}{D_0}$ dove $D(n)$ è l'aumento del glucosio negli ultimi cinque minuti e D_0l'incremento iniziale al momento dell'attivazione.

Per garantire la sicurezza dell'utente, l'algoritmo include regole che sospendono temporaneamente la stima dei CHO e la proposta di bolo in caso di ipoglicemia nascente o rilevazione di movimento, limitano la dose totale consigliata per evitare episodi notturni di ipoglicemia e consentono stime "multi-step" aggiornate qualora la meal flag si riattivi più volte nel corso dello stesso pasto. Inoltre, i parametri soggettivi quali rapporto insulina/CHO e sensibilità insulinica possono essere rivisti periodicamente per migliorare ulteriormente la precisione.

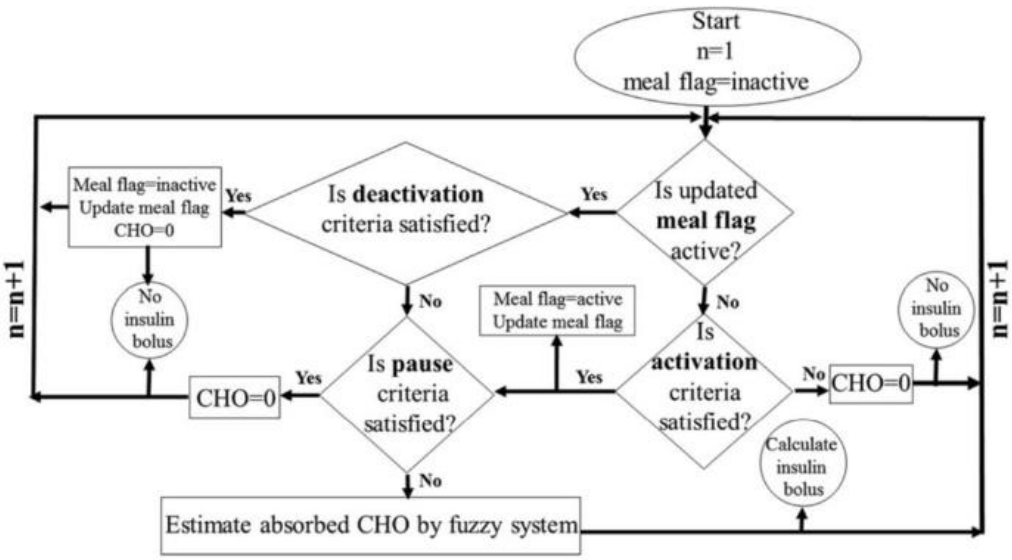

Figura 5:Logica fuzzy presa da [19]

Risultati del metodo originale

Le prestazioni dell'algoritmo sono state valutate utilizzando il simulatore di pazienti virtuali sviluppato dall'Università della Virginia e dall'Università di Padova.

I pazienti simulati presentavano eterogeneità nella risposta glicemica individuale e l'algoritmo è riuscito a adattarsi a quest'ultima, dimostrando un buon livello di precisione nell'analisi dei dati.

In totale, l'algoritmo è stato testato su 117 episodi tra pasti e spuntini; i risultati hanno mostrato un tasso di falsi positivi pari al 20,8%, il che significa che in alcuni casi il sistema ha segnalato la presenza di un pasto anche quando in realtà non si era verificato alcun evento.

Si è inoltre evidenziato che la sensibilità dell'algoritmo nel riconoscere gli spuntini è risultata inferiore rispetto ai pasti principali, a conferma del fatto che il sistema fa fatica a indentificare le variazioni glicemiche ridotte.

L'algoritmo di Samadi et al. [19] riesce a identificare autonomamente l'inizio di un pasto non annunciato, a stimarne la dimensione in carboidrati e a guidare la somministrazione del bolo insulinico senza ricorrere a modelli fisiologici espliciti: questo approccio si è dimostrato robusto in simulazioni e applicazioni cliniche preliminari, con elevata sensibilità, bassa frequenza di falsi positivi e stime sufficientemente accurate per supportare in sicurezza un pancreas artificiale completamente automatizzato.

Modifiche apportate all'algoritmo di Samadi et al. in fase di implementazione

L'algoritmo di meal detection basato su logica fuzzy è stato reimplementato, permettendo così di validarne le prestazioni in un contesto applicativo concreto e di adattarlo alle caratteristiche specifiche dei pazienti. Le modifiche principali riguardano:

- Personalizzazione dei parametri: sono stati ottimizzati i parametri relativi alle soglie di attivazione, pausa e disattivazione dell'automa a stati finiti, mediante Bayesian Optimization, considerando anche indicatori di rischio derivati da frequenza cardiaca, passi e ipoglicemia.
- Dataset e validazione: i dati sono stati suddivisi temporalmente in training, validazione e test, per permettere l'ottimizzazione dei parametri su un sottoinsieme e la valutazione su dati indipendenti.
- Funzione di costo personalizzata: è stata definita una funzione che valuta la qualità della rilevazione dei pasti in base a precisione, recall e F1-score, applicando un filtro temporale per evitare rilevazioni troppo ravvicinate.

Altri aspetti dell'algoritmo, come il calcolo delle derivate del segnale glicemico e la codifica fuzzy delle variazioni glicemiche, non sono stati modificati rispetto alla versione originale.

2.2.2. Algoritmo basato sul tasso di incremento glicemico di Harvey et al.

Descrizione dettagliata del metodo originale

L'algoritmo Glucose Rate Increase Detector (GRID) [20] è stato sviluppato per rilevare automaticamente i pasti nei sistemi di pancreas artificiale e per attivare in modo tempestivo una risposta insulinica adeguata, contribuendo così alla gestione efficace della glicemia postprandiale.

Funziona in parallelo al controller glicemico principale e utilizza esclusivamente i dati provenienti dal sensore continuo di glucosio senza necessità di input esterni.

L'algoritmo viene strutturato in tre fasi principali: preprocessing del segnale, stima del tasso di variazione del glucosio e logica di rilevamento. Inizialmente, i dati glicemici vengono filtrati per eliminare picchi di rumore o variazioni improvvise non fisiologiche.

Questo avviene tramite un filtro noise-spike che limita la variazione massima accettabile tra due campioni consecutivi (tipicamente ±3 mg/dL/min):

$$G_{F_{NS}}(k) = \begin{cases} G_{F_{NS}}(k-1) + \Delta G \quad se\ G_m(k) + G_{F_{NS}}(k-1) > \Delta G \\ G_{F_{NS}}(k-1) - \Delta G \quad se\ G_m(k) - G_{F_{NS}}(k-1) < -\Delta G \\ G_m(k)\ altrimenti \end{cases}$$

dove:

- $G_m(k)$ rappresenta la misura del glucosio al tempo k
- $G_m(k)$ rappresenta il massimo ROC accettabile (es. 3 mg/dL)
- $G_{F_{NS}}(k)$ rappresenta il valore filtrato

Questo è seguito da un filtro passa-basso di primo ordine che attenua le oscillazioni ad alta frequenza mantenendo la tendenza generale della curva glicemica:

$$G_F(k) = \frac{\Delta t}{\tau_F + \Delta t} \cdot G_{F_{NS}}(k) + \frac{\tau_F}{\tau_F + \Delta t} \cdot G_F(k-1)$$

dove:

Δt rappresenta il tempo di campionamento del segnale

τ_F rappresenta la costante di tempo del filtro passa basso

Una volta pulito il segnale, GRID calcola la derivata prima della glicemia, stimando quindi il suo tasso di crescita nel tempo.

Questa operazione viene effettuata tramite interpolazione polinomiale lagrangiana su tre punti successivi:

$$G'(k) = \frac{G_F(k-2) - 4\,G_F(k-1) + 3G_F(k)}{2\Delta t}$$

La fase successiva consiste nella valutazione logica di tali informazioni: un pasto viene rilevato solo se la glicemia supera una soglia minima (130 mg/dL) e se la sua velocità di crescita è sufficientemente elevata per almeno due o tre campioni consecutivi.

$$GRID^{+}(k) = \begin{cases} 1 \;\; se\; G_F(k) > G_{min} \;\wedge (\; G'(k),\; G'(k-1),\; G'(k-2) > G'_{min,3} \\ \qquad oppure \;\; G'(k),\; G'(k-1) \;>\; G'_{min,2}) \\ \qquad\qquad 0 \;\; altrimenti \end{cases}$$

$G_F(k)$ rappresenta valore glicemico filtrato

$G'(k)$rappresenta il tasso di variazione del glucosio

G_{min} rappresenta la soglia glicemica minima (es. 130 mg/dL)

$G'_{min,3}$ = 1.5 mg/dL/min: soglia ROC per 3 punti consecutivi

$G'_{min,2}$= 1.6 mg/dL/min: soglia ROC per 2 punti consecutivi

Il sistema segnala un evento positivo solo se queste condizioni sono soddisfatte, garantendo un buon compromesso tra sensibilità e specificità. Una volta identificato il pasto, GRID può agire in modalità utente o automatica.

Nel primo caso, il sistema genera un allarme che invita l'utente a inserire le informazioni sul pasto, in modo da poter somministrare un bolo completo o parziale in base alla risposta. Nella modalità automatica, invece, il sistema può erogare direttamente un bolo predefinito – ad esempio equivalente a un pasto da 75 grammi di carboidrati – oppure calcolare un bolo correttivo per riportare la glicemia verso un valore target, come 80 o 140 mg/dL.

In entrambi i casi, l'intervento può essere modulato attraverso una funzione opzionale detta Recent History Correction (RHC), che tiene conto sia dell'insulina recentemente somministrata (evitando sovradosaggi), sia dei valori glicemici minimi delle ultime ore (limitando il rischio di ipoglicemia tardiva).

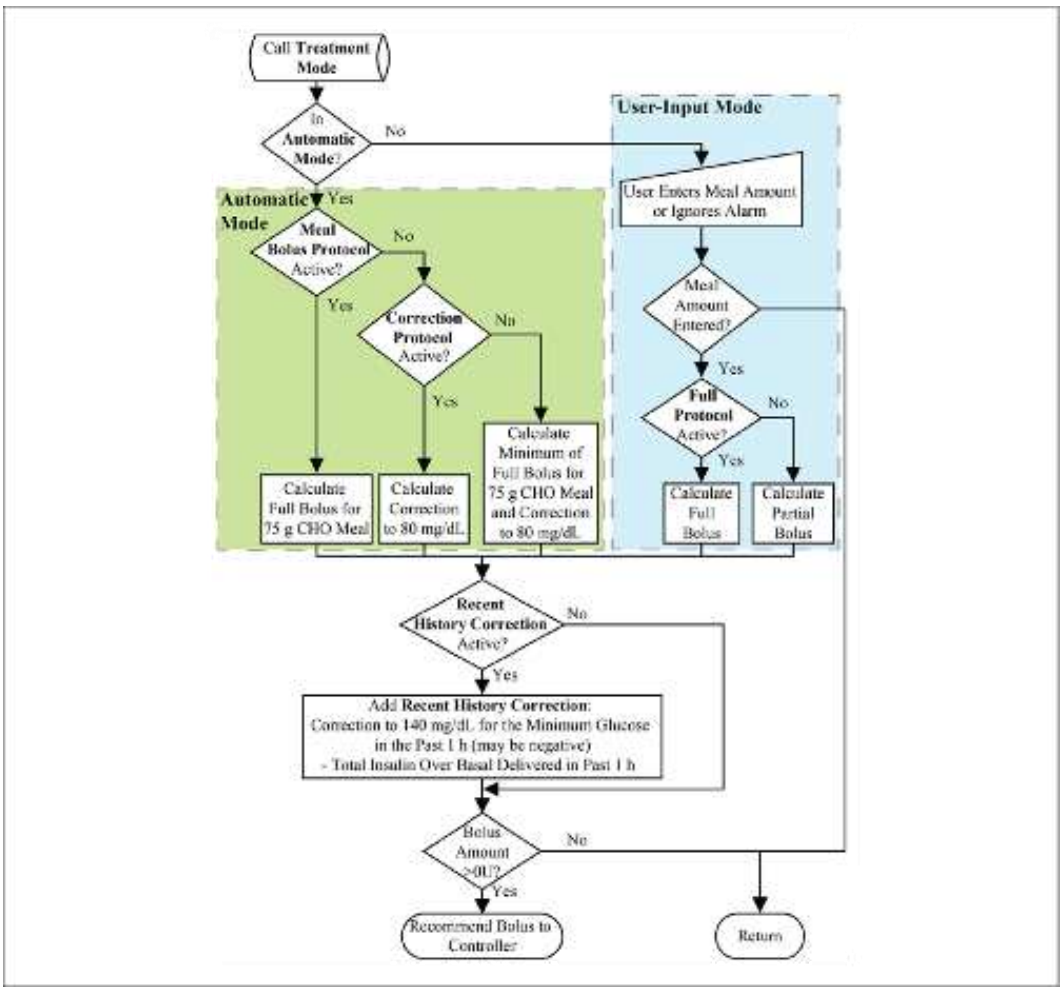

Figura 6: Schema a blocchi dell'algoritmo GRID presa da [20]

Risultati del metodo originale

L'algoritmo originale è stato ottimizzato su dati clinici ottenuti in modalità closed-loop e validato sia su soggetti differenti sia tramite simulazioni in silico, utilizzando il simulatore UVA/Padova approvato dalla FDA. I parametri selezionati manualmente includevano una costante di tempo del filtro passa basso pari a 6 minuti, una soglia glicemica minima di 130 mg/dL e soglie di tasso di variazione della glicemia pari a 1,5 e 1,6 mg/dL/min per rispettivamente tre o due punti consecutivi. I risultati riportati in letteratura indicano un tempo medio di rilevamento dei pasti pari a 42 minuti, con un tasso di falsi positivi inferiore a uno al giorno (0.58), accompagnato da un miglioramento significativo del tempo trascorso nel range glicemico ottimale (80–180 mg/dL) rispetto a scenari senza rilevamento automatico.

Modifiche apportate all'algoritmo di Harvey et al. in fase di implementazione

L'algoritmo GRID[20] implementato in questo lavoro si ispira fedelmente alla struttura presentata dagli autori, articolata in tre fasi principali: filtraggio del segnale glicemico, stima del tasso di variazione e logica di rilevamento basata su soglie.

Tuttavia, sono state apportate diverse modifiche significative per adattarlo al contesto applicativo specifico e migliorare la robustezza e la personalizzazione dell'algoritmo su dati reali.
Le modifiche principali apportate riguardano:

- Personalizzazione dei parametri: ottimizzazione tramite Bayesian Optimization delle soglie principali (G_{min} $G'_{min,3}$,$G'_{min,2}$), della costante del filtro passa basso (τ_F) e del limite di variazione glicemica, per migliorare la precisione del rilevamento dei pasti su base individuale.
- Correzione retrospettiva temporale**:** implementata per stimare con maggiore accuratezza l'orario effettivo del pasto, individuando il minimo glicemico all'interno di una finestra antecedente al rilevamento e compensando il ritardo fisiologico della risposta postprandiale.
- Intervallo di inattività (cooldown)**:** introdotto per evitare sovrapposizioni tra pasti ravvicinati e il doppio rilevamento dello stesso evento.

L'algoritmo si concentra esclusivamente sul rilevamento dei pasti dalla dinamica glicemica, senza includere funzioni di somministrazione automatica di insulina, gestione di input utente o correzioni basate sulla storia recente dell'insulina, lasciando tali estensioni a sviluppi futuri.

2.2.3 Algoritmo Data-driven sui pattern glicemici di De Carvalho et al.

Descrizione dettagliata del metodo originale

L'algoritmo proposto adotta un approccio data-driven[21] per il rilevamento dei pasti, sfruttando esclusivamente i dati glicemici ottenuti da sensori CGM e costruendo pattern personalizzati della risposta post-prandiale per ciascun individuo.

Il processo parte dall'analisi dei dati del singolo soggetto, rappresentati dalla tupla:

$$X = (T, E, P, D)$$

dove:

- $T = \{t_1, t_2, \dots, t_m\}$ rappresenta la serie temporale di valori glicemici,
- $E = \{e_1, e_2, \dots, e_k\}$ è l'insieme di timestamp associati a pasti registrati,
- $P = \{p_1, p_2, \dots, p_i\}$ è l'insieme di pattern glicemici estratti dalle risposte post-pasto,

- $D = \{d_1, d_2, \dots, d_i$ è l'insieme dei profili di distanza tra ogni pattern e il segnale glicemico completo.

A partire dai pasti noti E, viene estratta per ciascun evento e_k una risposta glicemica di durata fissa Δresponse di 2 ore che costituisce il pattern:

$$T_{i,n=}\{t_i, t_{i+1}, \dots, t_{i+n-1}\}$$

Tutti i pattern così ottenuti vengono utilizzati per identificare le risposte glicemiche tipiche del soggetto. Successivamente, per ciascun pattern $p_j \in P$, si effettua una ricerca per similarità su tutto il segnale glicemico T , utilizzando tecniche efficienti come MASS (Mueen's Algorithm for Similarity Search).

La distanza tra pattern e sottosequenze di T è calcolata tramite distanza z-normalizzata**,** definita come:

$$dist(Q, C) = \sqrt{2m(1 - corr(Q, C)}$$

dove:

- Q è il pattern di riferimento,
- C è una sottosequenza di TT di lunghezza m,
- corr(Q,C) è la cross-correlazione normalizzata tra Q e C.

Questa distanza viene calcolata in ogni punto del segnale, generando un profilo di distanza $d_j(t)$, che rappresenta la dissimilarità tra il pattern e ogni finestra scorrevole di T.

Le occorrenze candidate vengono identificate come i minimi locali all'interno del profilo $d_j(t)$, che risultano inferiori a una soglia predefinita dcutoff, e che rispettano una distanza temporale minima Δvalley tra due occorrenze successive. Questi istanti vengono considerati potenziali rilevamenti di pasto.

Per ciascun candidato identificato, si estrae una finestra temporale del segnale glicemico centrata sull'evento, e da essa si calcolano diverse feature descrittive, tra cui:

- valore glicemico iniziale,
- derivata iniziale stimata,
- valore minimo e massimo,

- tempo per raggiungere min e max,
- pendenza media verso minimo/massimo,
- orario della giornata (espresso come minuti da mezzanotte).

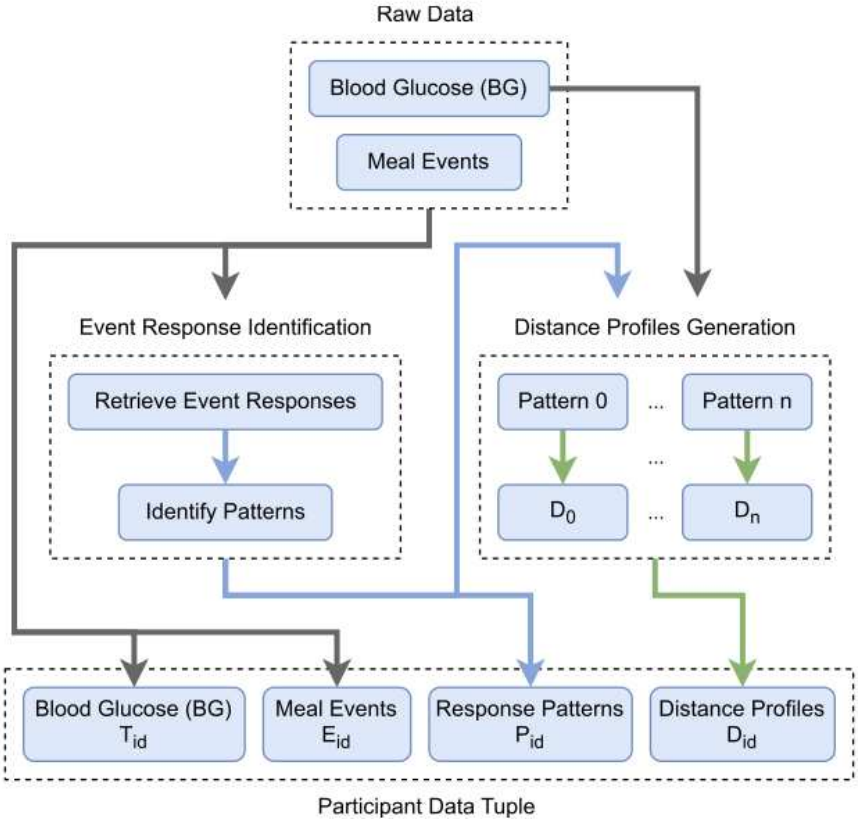

Figura 7 : Tupla dei candidati presa da [21]

Tali feature vengono utilizzate come input per un classificatore binario, che determina se l'evento candidato corrisponde a un pasto oppure no.

Il classificatore viene selezionato tra diversi modelli di machine learning, tra cui AdaBoost, Random Forest, Gradient Boosting, MLP, Gaussian Naive Bayes e alberi decisionali,scegliendo, per ciascun soggetto, il modello che ottiene la migliore prestazione sul set di validazione in termini di $F_{\beta-score}$

La metrica utilizzata per la selezione del classificatore è l'$F_{\beta-score}$definita come:

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{Precision} \cdot \text{Recall}}{(\beta^2 \cdot \text{Precision}) + \text{Recall}}$$

Nel lavoro, si utilizza β=2 per dare maggiore peso al recall e quindi ridurre i falsi negativi, favorendo la sensibilità del rilevamento.

Infine, per costruire il set di training supervisionato, ogni candidato viene etichettato come positivo se si verifica entro un intervallo simmetrico di ±1 ora da un pasto noto, ovvero:

$$t_c \in [t_{\text{meal}} - \delta,\ t_{\text{meal}} + \delta], \quad \text{con } \delta = 60\,\text{min}$$

Al di fuori di questo intervallo, l'evento candidato viene etichettato come negativo (non pasto).

L'algoritmo è in grado di adattarsi ai pattern glicemici specifici di ciascun individuo, migliorando significativamente la capacità di riconoscere pasti reali in assenza di input espliciti da parte dell'utente.

È importante precisare che l'algoritmo descritto da de Carvalho et al.[21] non è concepito per un'esecuzione in tempo reale, bensì opera in modalità offline.

Il rilevamento degli eventi pasto si basa su un'analisi retrospettiva dell'intera serie glicemica, già acquisita per intero, mediante la ricerca di sottosequenze simili a pattern glicemici estratti a partire da pasti noti. Questo approccio richiede la disponibilità dell'intero segnale temporale per poter calcolare i profili di distanza tra i pattern di riferimento e tutte le possibili finestre del segnale glicemico continuo.

Di conseguenza, il metodo non è immediatamente applicabile in ambienti operativi che richiedono una rilevazione tempestiva e in continua evoluzione, come nei sistemi di somministrazione insulinica automatica o nei moduli di allerta in tempo reale. Sebbene l'algoritmo presenti ottime prestazioni in termini di personalizzazione e accuratezza, la sua architettura non consente l'elaborazione dei dati in modo incrementale né la generazione di decisioni istantanee a partire da input parziali. Pertanto, la sua utilità risiede principalmente in contesti offline, come l'analisi retrospettiva dei dati, la validazione di altri modelli, o il supporto nella calibrazione di algoritmi predittivi destinati a essere eseguiti in real-time

Risultati del metodo originale

Nel lavoro originale di De Carvalho et al. [26], l'algoritmo di rilevamento dei pasti è stato validato sul dataset OhioT1DM, sfruttando esclusivamente i dati glicemici da CGM e gli eventi pasto auto-riportati. Il metodo, di natura interamente offline, è stato testato su 11 partecipanti (escludendo uno privo di annotazioni). Le prestazioni ottenute variano sensibilmente tra i partecipanti, riflettendo la diversa quantità e qualità dei pasti annotati. Il

punteggio F_2-score sul set di test risulta compreso in un intervallo che va da 0.06 a 0.87, a seconda del soggetto e del classificatore selezionato.

I modelli più performanti si sono rivelati essere Random Forest, AdaBoost e MLP, in funzione della variabilità individuale della risposta glicemica.

La tabella seguente riassume le principali metriche per ciascun paziente:

ID	Modello	TP	FP	FN	Precisione	Recall	F1	F2
540	**Decision Tree**	1	8	19	0.11	0.05	0.07	0.06
544	**Random Forest**	29	9	3	0.76	0.91	0.83	0.87
552	**MLP**	6	12	8	0.33	0.43	0.38	0.41
559	**Gaussian NB**	11	20	12	0.35	0.48	0.41	0.45
563	**Decision Tree**	10	18	13	0.36	0.43	0.39	0.42
570	**Random Forest**	24	13	7	0.65	0.77	0.71	0.75
575	**MLP**	24	18	13	0.57	0.65	0.61	0.63
584	**Decision Tree**	11	23	9	0.32	0.55	0.41	0.48
588	**AdaBoost**	25	25	9	0.50	0.74	0.60	0.67
591	**AdaBoost**	26	38	11	0.41	0.70	0.51	0.61
596	**Gradient Boosting**	32	27	13	0.54	0.71	0.62	0.67

Tabella 1:Risultati sul dataset OhioT1DM

Modifiche apportate all'algoritmo di De Carvalho et al. in fase di implementazione

Nel presente lavoro, l'algoritmo originale è stato modificato per consentire un'elaborazione in tempo reale del segnale glicemico, superando la principale limitazione della versione offline. Le modifiche principali riguardano:

- Selezione dei candidati in tempo reale: la fase di identificazione degli eventi pasto è stata ristrutturata per funzionare su dati acquisiti progressivamente, aggiornando

dinamicamente la lista dei potenziali eventi man mano che nuove osservazioni diventano disponibili.

- Adattamento dell'algoritmo MASS: il calcolo del profilo di distanza tra pattern glicemici medi e finestre scorrevoli del segnale è stato adattato per garantire compatibilità con dati parziali, rumorosi o mancanti.
- Soglia e filtro di soppressione temporale: la rilevazione dei candidati utilizza una soglia fissa (dcutoff) e un filtro basato sulla finestra "Δvalley", che mantiene solo i minimi locali sufficientemente distanziati nel tempo, riducendo falsi positivi e ridondanza.
- Efficienza computazionale: la logica è stata ottimizzata per una pipeline online, in cui la selezione dei candidati dipende esclusivamente dalle informazioni disponibili fino all'istante corrente.

Queste modifiche rendono l'algoritmo applicabile a scenari reali, permettendo decisioni progressive e affidabili senza necessità di dati futuri.

2.2.4 Algoritmo fisiologico con simulazione predittiva di Zheng et al.

Descrizione dettagliata del metodo originale

Il metodo proposto da Zheng et al.[22], è un approccio innovativo per la rilevazione automatica dei pasti basato sulla simulazione della risposta glicemica mediante un modello fisiologico e l'uso integrato di dati provenienti da sensori indossabili. A partire dalle serie temporali di glucosio osservato $G(t)$,insulina basale $I(t)$, in bolo $B(t)$ e frequenza cardiaca $HR(t)$ l'algoritmo simula l'andamento glicemico atteso $G'(t)$ utilizzando il Glucose-Insulin Model (GIM), un modello validato che descrive il metabolismo del glucosio tenendo conto di molteplici componenti fisiologiche.

Il consumo di glucosio viene rappresentato come la somma di due contributi: uno insulino indipendente $U_{ii}(\mathrm{t})$, e uno insulino-dipendente $U_{id}(\mathrm{t})$ secondo la relazione complessiva:

$$U(t) = U_{ii}(\mathrm{t}) + U_{id}(t)$$

Un elemento distintivo del metodo è l'integrazione dell'effetto dell'attività fisica sulla sensibilità insulinica. Questo è incluso tramite il contributo della frequenza cardiaca, modellato secondo una dinamica differenziale che stima il consumo energetico $Y(t)$ e l'efficacia insulinica $Z(t)$.

La dinamica del consumo energetico è descritta da:

$$\frac{dY(t)}{dt} = \frac{1}{T_{HR}}\, Y(t)\, (HR(t) - HR_b)\, , Y(0) = 0$$

mentre l'evoluzione della sensibilità insulinica è governata da:

$$\frac{dZ(t)}{dt} = \frac{f(Y(t))}{T_{in}} \left(\frac{1}{T_{ex}} Z(t) + f\big(Y(t)\big)\right), Z(0) = 0$$

dove la funzione $f\big(Y(t)\big)$ è una funzione sigmoidea che modula l'effetto dell'attività in base al livello di intensità percepito e assume la forma:

$$f\big(Y(t)\big) = \frac{(\frac{Y}{aHR_b})^n}{\left(1 + \left(\frac{Y}{aHR_b}\right)^n\right)}$$

In aggiunta, viene calcolata un'integrazione cumulativa del surplus di attività, espressa dal termine $W(t)$, che rappresenta l'effetto prolungato dell'attività fisica sulla dinamica del glucosio:

$$W(t) = \begin{cases} \int_0^t (HR(\tau) - HR_b) d\tau, & se\ t < t_z \\ 0\ , & altrimenti \end{cases}$$

Il modello così costruito è in grado di generare una previsione del profilo glicemico atteso $G'(t)$, per ogni istante t, tenendo conto dell'insulina somministrata e del livello di attività fisica, ma in assenza di pasti

Una volta ottenuta la curva glicemica prevista $G'(t)$, questa viene confrontata con quella osservata $G(t)$, e viene identificato un punto di divergenza quando la differenza tra le medie in una finestra f supera una soglia φ secondo:

$$\left|\overline{G'(t-f:t)} - \overline{G(t-f:t)}\right| > \varphi$$

dove $\overline{G'(t-f:t)}$ e $\overline{G(t-f:t)}$ rappresentano, rispettivamente, la media della glicemia simulata e osservata negli ultimi minuti.

Dopo aver rilevato il punto di divergenza, l'algoritmo avvia la generazione di pasti candidati definiti come triple $m = [m_{st}, m_c, m_{du}]$, dove:

- m_{st} è l'orario di inizio pasto (in minuti),
- m_c rappresenta la quantità di carboidrati assunti (in grammi),
- m_{du} la durata del pasto, ovvero il tempo durante il quale si assume l'intera quantità di carboidrati.

Si assume che l'ingestione avvenga in modo costante nel tempo: ciò significa che i carboidrati vengono distribuiti uniformemente durante la durata del pasto. Il tasso di ingestione, quindi, è definito come:

$$dd(t) = \frac{mc}{mdu}, \quad t \in [mst, mst + mdu]$$

dove $dd(t)$ rappresenta la velocità di assunzione dei carboidrati al tempo t.

Ogni pasto candidato viene iniettato nel modello per generare una nuova simulazione $G'_m(t)$.

Per valutare la bontà della simulazione rispetto al segnale reale, si calcola la distanza euclidea media tra $G'_m(t)$ e $G(t)$ sull'intervallo corrispondente:

$$d_{\text{Euc}}(m) = \frac{1}{t - mst} \sqrt{\sum_{i=mst}^{t} \left(G'_m(i) - G(i)\right)^2}$$

Il pasto viene accettato se $d_{\text{Euc}}(m) < \varepsilon$, dove ε è una soglia predefinita (5 mg/dL).

Tra tutti i pasti candidati accettabili, l'algoritmo seleziona quello che minimizza la distanza euclidea, ovvero quello che spiega meglio la discrepanza osservata.

Questo processo viene ripetuto ciclicamente lungo la serie temporale, permettendo di rilevare i pasti in modo continuo e in tempo quasi reale. L'utilizzo di un modello fisiologico consente di distinguere con maggiore affidabilità le variazioni dovute ai pasti da quelle causate da altri fattori, come l'attività fisica o la variabilità endogena del soggetto.

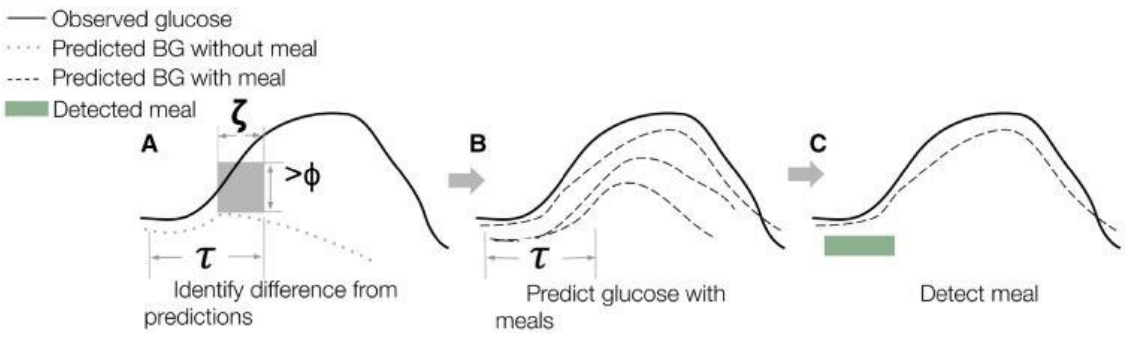

Figura 8 : Predizioni del tracciato glicemico con pasto e senza pasto presa da[22]

Risultati del metodo originale

I risultati ottenuti su dati simulati mostrano un ritardo medio nella rilevazione dei pasti di 25,7 minuti e un errore medio sulla stima della quantità di carboidrati di 1,2 g.

La validazione su dati reali, mediante criteri operativi basati sull'aumento glicemico postprandiale, ha confermato la maggiore capacità del metodo GIM nel rilevare pasti plausibili in presenza di rumore, attività fisica e assenza di annotazioni manuali, con il 76,7% dei pasti rilevati considerati validi.

La combinazione di precisione, robustezza e integrazione fisiologica rende GIM un candidato altamente promettente per l'integrazione in sistemi automatizzati per il supporto alla terapia del T1DM.

Modifiche apportate all'algoritmo di Zheng et al. in fase di implementazione

Il metodo originale è stato rielaborato, semplificando l'approccio e omettendo il calcolo dei carboidrati.

Le modifiche principali includono:

- Simulazione individualizzata: sviluppo di un modulo basato sul modello GIM, calibrato per ciascun soggetto in funzione di peso corporeo e frequenza cardiaca basale, per confrontare il glucosio simulato con quello osservato.
- Rilevamento basato sulla soglia φ: attivazione del meccanismo di rilevamento quando la glicemia osservata si discosta significativamente da quella simulata, con soglia ottimizzata per ciascun individuo.
- Analisi retrospettiva della derivata prima: valutata su una finestra temporale definita per stimare con maggiore precisione l'inizio effettivo del pasto, riducendo falsi positivi dovuti a variazioni glicemiche non correlate all'alimentazione.

Queste modifiche consentono una rilevazione tempestiva e robusta dell'inizio del pasto, migliorando la corrispondenza temporale tra eventi rilevati e pasti effettivamente consumati, senza stimare la quantità di carboidrati ingeriti.

2.2.5 Algoritmo HR+CGM di Cossu et al

Descrizione del metodo originale

Il metodo implementato da Cossu et al. [23] rappresenta un algoritmo innovativo per la rilevazione automatica dei pasti, pensato specificamente per soggetti affetti da ipoglicemia post-bariatrica (PBH).

A differenza degli approcci tradizionali che si basano esclusivamente sull'andamento del glucosio, il metodo proposto sfrutta anche il segnale della frequenza cardiaca (HR), con l'obiettivo di catturare in modo più tempestivo e robusto le dinamiche fisiologiche associate all'assunzione di cibo.

L'idea alla base dell'algoritmo è che il pasto rappresenti un evento multisistemico, che coinvolge non solo il metabolismo glucidico ma anche il sistema nervoso autonomo. In particolare, nei soggetti PBH è stato osservato che l'ingestione di un pasto si accompagna a un aumento della frequenza cardiaca, riconducibile a una risposta autonomica gastrointestinale. Questo fenomeno precede spesso l'innalzamento glicemico rilevabile dai sensori CGM, rendendo il segnale HR un potenziale indicatore anticipatorio.[24]

L'algoritmo agisce utilizzando una catena di regole decisionali, progettata per operare in tempo reale. Il processo si attiva quando il ROC del glucosio supera una certa soglia, indicando una possibile salita glicemica. A questo punto, viene analizzato l'andamento della frequenza cardiaca nei minuti precedenti alla salita: se viene rilevato un picco di HR superiore a una soglia prefissata, e il glucosio soddisfa contemporaneamente criteri minimi di valore assoluto e variazione, l'evento viene classificato come pasto.

Tutti i parametri decisionali sono personalizzabili per ciascun soggetto. La personalizzazione avviene attraverso una fase di calibrazione iniziale, durante la quale i parametri vengono ottimizzati in funzione di una metrica di performance, al fine di adattare la sensibilità del sistema alle specifiche risposte fisiologiche individuali.

L'algoritmo è stato sviluppato originariamente su dati raccolti in ambiente controllato da soggetti PBH, utilizzando sensori indossabili per la misura continua di glucosio e frequenza

cardiaca. La sua struttura modulare e generalizzabile ha poi consentito di testarlo anche su altre coorti, attraverso un porting diretto, mantenendo inalterata l'architettura decisionale e verificandone la robustezza in contesti diversi.

Presentava un confronto con un algoritmo basato sul tasso di variazione del glucosio, che è stato mantenuto e impiegato come metodologia aggiuntiva di identificazione dei pasti.

L'algoritmo utilizzato come riferimento, da ora in avanti denominato ROC, rileva i pasti monitorando il ROC del glucosio. Se l'incremento del glucosio risulta sostenuto per più letture e superiore a una soglia prestabilita, il pasto viene considerato confermato.

In caso contrario, il processo viene resettato, garantendo maggiore accuratezza e riducendo i falsi positivi.

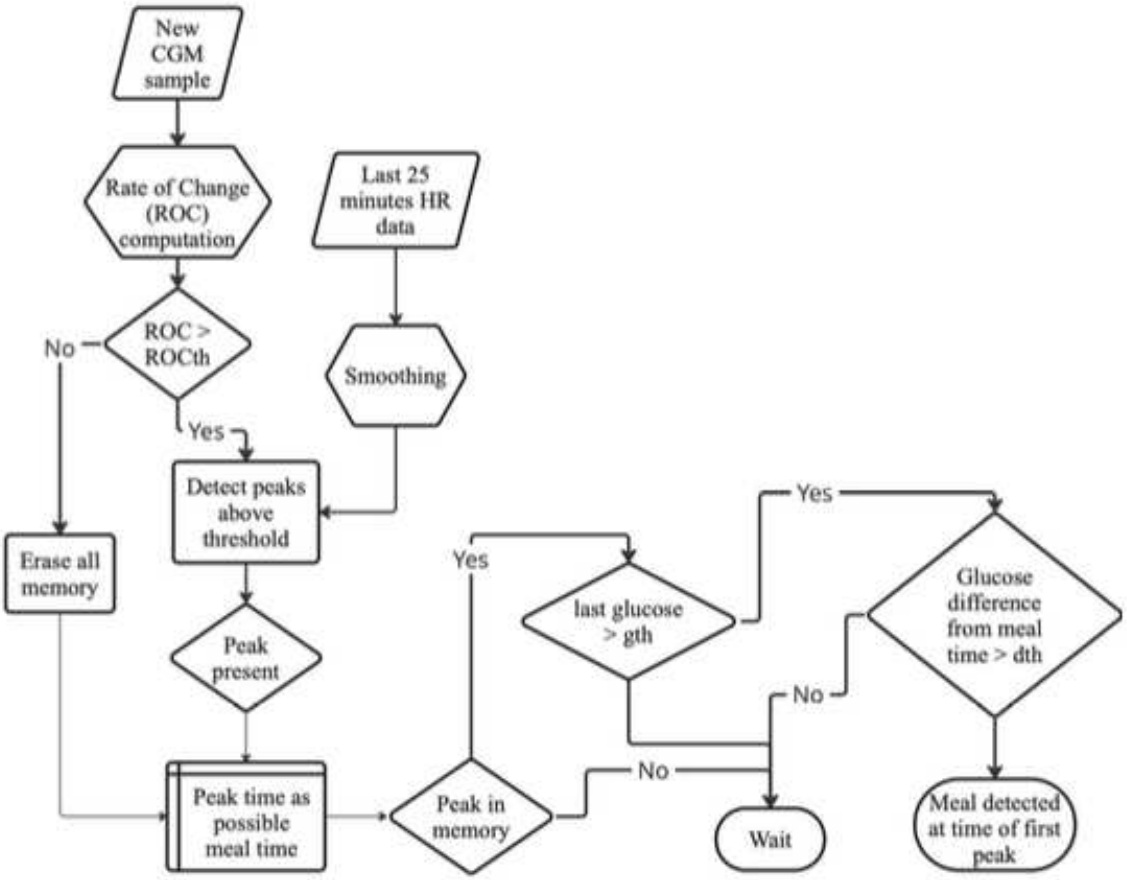

Figura 9 : Schema a blocchi del metodo HR+CGM presa da [23]

Risultati sul metodo originale

I risultati ottenuti dal metodo originale confermano che l'algoritmo risulta essere robusto ed efficace presentando un recall del 100 % ,una precisione superiore al 90% e un ritardo

mediano di rilevamento di appena 1 minuto rispetto all'effettivo inizio del pasto. Tali performance sottolineano l'accuratezza temporale del sistema, con una finestra di rilevamento aderente all'evento reale.

Anche in condizioni di vita reale, il metodo ha comunque mantenuto prestazioni elevate, nonostante la maggiore variabilità del contesto. La precisione risulta essere dell'85%, con un recall del 78% e un tempo mediano di rilevamento pari a 28 minuti. Particolarmente rilevante è il tasso contenuto di falsi positivi, pari a uno ogni due giorni circa, inferiore rispetto ai metodi di rilevamento pasti basati solo sul CGM.

L'utilizzo di una strategia di personalizzazione dei parametri ha fatto si che l'algoritmo risultasse adattabile alle caratteristiche fisiologiche dell'individuo, migliorando sensibilmente la capacità di rilevamento in soggetti con frequenza cardiaca di base più bassa o con risposte post-prandiali attenuate.

I risultati confermano come l'integrazione della frequenza cardiaca consenta di migliorare significativamente l'affidabilità e la specificità del rilevamento dei pasti nei soggetti PBH, offrendo una base solida per l'estensione del metodo anche ad altre popolazioni cliniche. Proprio l'efficacia dimostrata in questo contesto ha motivato l'esplorazione di una versione adattata per i soggetti con T1DM, con l'obiettivo di verificarne la trasferibilità e il potenziale beneficio anche in presenza di dinamiche fisiologiche più complesse e variabili.

Modifiche apportate all'algoritmo di Cossu et al. in fase di implementazione

L'algoritmo HR+CGM[23] utilizzato in questo lavoro è stato originariamente sviluppato per soggetti con ipoglicemia post-bariatrica (PBH). In tale contesto, la risposta fisiologica all'assunzione di un pasto è caratterizzata da dinamiche particolarmente rapide e marcate: la glicemia tende ad aumentare in modo repentino, così come la frequenza cardiaca, che mostra un picco immediatamente successivo all'ingestione.

Questo comportamento rende il pattern HR+CGM facilmente identificabile e ripetibile, facilitando la progettazione di regole threshold-based.

Nel caso dei soggetti con T1DM, tuttavia, la situazione è sensibilmente diversa. Le differenze principali che giustificano un adattamento dell'algoritmo sono le seguenti:

- Risposta glicemica più variabile: nei soggetti T1DM, l'andamento post-prandiale del glucosio dipende fortemente dalla quantità e dal tipo di insulina somministrata, dal

tempo di somministrazione rispetto al pasto e dalla variabilità interindividuale. Di conseguenza, la salita glicemica può essere più lenta o attenuata rispetto a quanto osservato nei soggetti PBH.

- Risposta della frequenza cardiaca meno prevedibile: nei soggetti diabetici la risposta autonomica al pasto può essere meno marcata, o talvolta assente, a causa della neuropatia autonomica diabetica. Questo rende meno affidabile l'utilizzo di soglie assolute di HR definite a priori.
- Rumore e variabilità del segnale HR: nella vita reale, i segnali HR nei soggetti T1DM sono spesso affetti da rumore, artefatti o picchi causati da attività fisica, stress o movimenti improvvisi, il che può aumentare i falsi positivi se non si adotta una logica più cauta.

Per questi motivi, pur mantenendo invariata la struttura dell'algoritmo, è stato necessario adattare i parametri decisionali. Le principali modifiche introdotte sono state:

- Ottimizzazione soggetto-specifica delle soglie: tutti i parametri sono stati calibrati individualmente su ciascun soggetto, utilizzando dati di training e ottimizzando una funzione obiettivo basata su F1-score. Questo ha permesso di compensare la maggiore variabilità interindividuale nella risposta ai pasti.
- Gestione conservativa degli eventi HR: è stata mantenuta una soglia HR elevata, ma integrata con controlli addizionali sulla traiettoria glicemica (es. differenza minima di glucosio), al fine di ridurre la sensibilità a picchi HR non legati al pasto.
- Introduzione post-implementativa di un criterio ispirato all'algoritmo di Colmegna: successivamente all'implementazione dell'algoritmo HR+CGM,[23] al fine di ridurre i falsi positivi legati a variazioni glicemiche non prandiali, è stato introdotto un criterio derivato dal lavoro di Colmegna et al. relativo alla rilevazione automatica dei pasti in soggetti con T1DM. [25]

  In tale studio, gli autori propongono un metodo per identificare pasti (anche non annunciati) analizzando il segnale glicemico tramite le sue derivate prima e seconda. In particolare, viene calcolato l'indice $\mu = ROC \cdot ROC2$, dove entrambe le derivate sono saturate a zero per eliminare la componente negativa, e i picchi di μ vengono usati per stimare l'insorgenza del pasto.

  Ispirandomi a questa strategia, anche nel presente lavoro è stato adottato un indice μ analogo, con soglia empirica adattata al contesto applicativo. L'obiettivo non

è la ridefinizione del timestamp del pasto, come nel lavoro originale, ma piuttosto la validazione o il rafforzamento del segnale CGM al fine di ridurre i falsi positivi. Questo approccio ha mostrato maggiore robustezza nella discriminazione degli eventi reali, soprattutto in situazioni caratterizzate da elevata variabilità glicemica o rumore .

In sintesi, l'algoritmo è stato portato direttamente nel contesto diabetico senza modificarne la logica, ma adattandone l'operatività attraverso la calibrazione dei parametri e una gestione più conservativa delle condizioni di attivazione.

Questo approccio ha consentito di esplorare l'utilizzo della frequenza cardiaca come segnale ausiliario anche in soggetti con T1DM, ampliando il potenziale applicativo dell'algoritmo originario.

2.2.6 Il metodo RAP sviluppato da Mosquera-Lopez et al.

Descrizione dettagliata del metodo originale

Il metodo proposto da Mosquera-Lopez et al.[26] si basa sull'integrazione tra un sistema di somministrazione automatica dell'insulina e un modello di intelligenza artificiale in grado di rilevare autonomamente l'assunzione di un pasto e stimarne il contenuto di carboidrati. Questo sistema, indicato come Robust Artificial Pancreas (RAP), mira a ridurre la necessità di intervento da parte della persona con diabete, rendendo più e sicura la gestione postprandiale della glicemia.

L'algoritmo è stato implementato all'interno della piattaforma iPancreas, che permette l'integrazione tra sensori CGM, pompa insulinica, smartwatch e smartphone.

Questa rappresenta, inoltre, lo strumento utilizzato per la raccolta dati e il monitoraggio remoto.

Il metodo si focalizza sullo sviluppo di una rete neurale artificiale multi-output. La rete è stata inizialmente addestrata su dati simulati, ottenuti da due simulatori clinicamente validati per il T1DM relativi a 199 soggetti virtuali osservati per 14 giorni .Successivamente, l'algoritmo è stato valutato in uno studio clinico crossover condotto su 15 adulti con diabete di tipo 1, di cui 13 hanno completato il protocollo, al fine di testarne la sicurezza e l'efficacia in condizioni reali.

Il modello prende in input 32 caratteristiche calcolate da due ore di dati storici di glucosio e somministrazione di insulina. Tra queste caratteristiche sono presenti la glicemia media, il tasso di variazione del glucosio, la disponibilità di insulina e l'ora del giorno.

La rete neurale è composta da tre livelli condivisi (512, 32 e 16 nodi) e due rami distinti: uno per classificare la presenza o meno di un pasto e uno per stimarne la dimensione.

Per l'ottimizzazione dei pesi della rete è stato utilizzato l'algoritmo Adam [27] e il modello è stato addestrato minimizzando contemporaneamente due funzioni di perdita: l'entropia incrociata binaria per la rilevazione del pasto e l'entropia incrociata categoriale per la stima della dimensione. È stato applicato un meccanismo di interruzione anticipata dell'addestramento quando le prestazioni sul set di validazione smettono di migliorare, così da evitare l'overfitting.

Il sistema RAP integra questa rete neurale all'interno di un'applicazione di controllo automatizzato dell'insulina basata su Model Predictive Control (MPC). Quando la probabilità di un pasto supera una soglia prestabilita (PTH = 0.86), il sistema invia una notifica all'utente, proponendo una dose di insulina calcolata automaticamente sulla base del contenuto stimato di carboidrati e del rapporto insulina/carboidrati individuale. Poiché l'insulina viene somministrata dopo l'ingestione del pasto, la dose proposta è proporzionalmente ridotta, in funzione del tempo trascorso dal pasto stesso, con una riduzione dell'1% per ogni minuto di ritardo stimato.

Risultati del metodo originale

L'algoritmo di rilevamento pasti e stima della dimensione del pasto ha dimostrato ottime prestazioni sia in fase preclinica che in studio clinico. La sensibilità nella rilevazione dei pasti è risultata dell'83,3%, con un basso tasso di falsi positivi (16,6%) e un tempo medio di rilevazione pari a circa 26 minuti.

Tali risultati sono in linea con le simulazioni effettuate precedentemente in silico dagli autori, confermando l'affidabilità del sistema anche in contesti reali.

Il sistema di somministrazione automatizzata dell'insulina che integra l'algoritmo (RAP) ha mostrato un significativo miglioramento del controllo postprandiale rispetto al sistema al MPC: il tempo trascorso in iperglicemia si è ridotto in media del 10,8% (P = 0.04), mentre il

tempo in range è aumentato del 9,1%, pur senza raggiungere significatività statistica (P = 0.09).

Questi benefici sono stati ottenuti senza aumentare il rischio di ipoglicemia, con valori di tempo sotto i 70 mg/dL comparabili tra RAP e MPC.

Modifiche apportate all'algoritmo di Mosquera et al. in fase di implementazione

Per quanto riguarda l'algoritmo RAP descritto da Mosquera et al. [26], nel presente lavoro non sono state apportate modifiche sostanziali alla logica di base. L'unico intervento effettuato è stato un tentativo di personalizzare la soglia PTH per ciascun paziente, con l'obiettivo di adattarla alle caratteristiche individuali. Tuttavia, questa personalizzazione non ha determinato un miglioramento significativo delle prestazioni dell'algoritmo.

CAPITOLO 3

3. Dataset e fase di preprocessing

3.1 Dataset T1DEXI: descrizione, soggetti e raccolta

Il set di dati T1DEXI è stato generato nel quadro della sperimentazione clinica Type 1 Diabetes Exercise Initiative, che aveva lo scopo di studiare l'impatto dell'attività fisica sul controllo glicemico nei giovani adulti con diabete di tipo 1.[28]
Lo studio ha seguito ogni partecipante per circa quattro settimane, richiedendo un minimo di 150 minuti di esercizio alla settimana come criterio di inclusione. Durante il periodo di osservazione sono stati raccolti dati multidimensionali ad alta risoluzione, tra cui misurazioni continue della glicemia con i sensori Dexcom G6 CGM, informazioni dettagliate sull'assunzione di insulina (basale e bolo), la registrazione dei pasti e degli spuntini tramite l'applicazione mobile T1-DEXI e i dati relativi all'attività fisica (tipo, intensità percepita e durata dell'esercizio).

Oltre alle componenti comportamentali e terapeutiche, lo studio ha incluso anche parametri fisiologici ottenuti da dispositivi indossabili. In particolare, sono stati utilizzati il Verily Study Watch e una fascia toracica per la rilevazione continua della frequenza cardiaca (HR), fornendo così un quadro fisiologico utile alla comprensione delle risposte cardiache in condizioni reali.
I dati grezzi raccolti sono stati originariamente salvati in più file per ciascun paziente, successivamente riorganizzati tramite una pipeline di elaborazione sviluppata in ambiente Matlab.
La pipeline comprendeva diversi passaggi, tra cui la conversione dei formati originali in strutture compatibili per l'analisi, l'estrazione e l'allineamento delle variabili di interesse, la ristrutturazione dei dati per soggetto e la loro visualizzazione sintetica.

Questo processo ha permesso l'integrazione di informazioni eterogenee in una struttura di dati unica e unificata, consentendo un'analisi quantitativa dell'interazione tra esercizio fisico, somministrazione di insulina e risposta metabolica.
Tuttavia, la struttura dei dati presenta problemi critici che devono essere risolti nella fase di preprocessing.
Tra questi si evidenziano la presenza di pasti registrati in rapida successione che possono generare ambiguità temporali, e la mancanza di una stima accurata della quantità di carboidrati in alcuni episodi alimentari, spesso riportata in modo parziale o approssimato.
Tali aspetti richiedono talvolta interventi manuali di correzione o imputazione per evitare distorsioni nell'analisi.
Lo studio ha coinvolto complessivamente 496 soggetti, di cui 495 sono stati inclusi nell'analisi preliminare dopo le operazioni di controllo qualità.
Il dataset è composto da 133 uomini e 362 donne, con un'età compresa tra i 18 e i 70 anni, tutti residenti negli Stati Uniti.
Ogni soggetto è identificato da un codice univoco ed è associato a un insieme di variabili demografiche e cliniche, tra cui età, sesso, informazioni sulla gestione terapeutica del diabete, tracciati glicemici da CGM, registrazioni dei pasti, e segnali fisiologici continui come la frequenza cardiaca.
I dati di frequenza cardiaca, in particolare, sono stati acquisiti con una risoluzione temporale elevata ovvero 1 misura ogni 10 secondi, mentre i tracciati glicemici presentano un intervallo di campionamento regolare di 5 minuti. Questa densità di misurazione consente di eseguire analisi dettagliate e sincronizzate della risposta metabolica all'attività fisica e alla terapia insulinica.
Complessivamente il dataset T1DEXI rappresenta uno strumento utile per lo studio del comportamento glicemico in tempo reale.
Offre una visualizzazione realistica della condizione diabetica poiché presenta i dati dei pazienti per un periodo prolungato, raccoglie misurazioni molto frequenti e include diversi tipi di informazioni.

3.2 Analisi e preprocessing dei dati

Nel contesto dell'analisi dei dati del dataset T1DEXI, è stata implementata una procedura di preprocessing articolata, con l'obiettivo di selezionare un sottoinsieme di pazienti che garantisse un'elevata qualità e completezza delle informazioni.
La valutazione iniziale ha riguardato l'identificazione della percentuale di valori mancanti nelle variabili principali. In particolare, i dati glicemici presentano una percentuale media di

valori NaN pari al 4,20%, mentre per la frequenza cardiaca tale valore è più elevato, raggiungendo in media il 16,85%.
Per tali ragioni, è stato introdotto un filtro che ha permesso l'inclusione esclusivamente dei pazienti con meno del 10% di dati mancanti in entrambe le variabili.
In aggiunta alla qualità del segnale, sono stati considerati anche aspetti legati alla completezza informativa, tra cui la disponibilità di dati di insulina basale e bolo per ciascun giorno di osservazione e la presenza di almeno tre pasti giornalieri mediamente rilevati.
Questo ha permesso di garantire una copertura temporale continua e di disporre di un contesto nutrizionale e terapeutico sufficiente per interpretare correttamente l'andamento glicemico. Il rispetto simultaneo di tutti questi criteri ha portato alla selezione finale di 120 soggetti, considerati "validi" ai fini dell'analisi.
Per ciascun paziente incluso, sono state quindi calcolate statistiche descrittive riferite alle serie temporali della glicemia, tra cui la media, la mediana, i percentili 25° e 75°, la deviazione standard e il coefficiente di variazione.
Tali misure hanno permesso di descrivere centralità, dispersione e simmetria della distribuzione glicemica individuale. Oltre alle statistiche classiche, sono stati considerati anche indicatori clinicamente rilevanti come il Time In Range (TIR), ovvero la percentuale di tempo in cui la glicemia risulta compresa tra 70 e 180 mg/dL, il Time Below Range (TBR), che rappresenta la quota di tempo trascorsa in ipoglicemia, e il Time Above Range (TAR), relativo alle condizioni di iperglicemia.
Per una stima temporale, è stato inoltre calcolato il tempo totale (in minuti) trascorso in ciascuna fascia critica, tenendo conto della frequenza di campionamento di 5 minuti.
Parallelamente, è stata condotta un'analisi sulla frequenza cardiaca rilevata dai dispositivi indossabili. Anche in questo caso, oltre alle misure descrittive, è stato calcolato il tempo totale trascorso in condizioni di bradicardia (T Brady), definite da una frequenza inferiore a 60 bpm, e tachicardia (T Tachy), superiore a 100 bpm. Le stime temporali sono state ricavate calcolando gli intervalli di tempo associati ai valori validi di HR, sfruttando la risoluzione temporale elevata dei dati (campionamento ogni 10 secondi).
Durante il preprocessing è stata inoltre verificata la regolarità temporale delle misurazioni, con particolare attenzione alla presenza di discontinuità nei segnali e alla sincronizzazione tra gli eventi alimentari e le misure glicemiche e fisiologiche.
È stata adottata una rappresentazione temporale unificata tramite oggetti timetable, che ha permesso di gestire in maniera efficiente l'allineamento tra le serie, l'aggregazione dei dati per giorno e l'applicazione di funzioni di raggruppamento e sintesi.

Infine, per facilitare l'interpretazione e la sintesi dei risultati, sono stati realizzati grafici riepilogativi che mostrano la distribuzione delle principali metriche calcolate tra i pazienti selezionati. In particolare, sono stati prodotti istogrammi per visualizzare la dispersione delle medie glicemiche, del TIR, TAR e TBR, e boxplot per rappresentare la variabilità intra-soggetto attraverso la deviazione standard e il coefficiente di variazione. Sono stati inoltre creati diagrammi a barre per illustrare la durata media dell'esposizione a condizioni di ipo e iperglicemia, nonché a stati di bradicardia e tachicardia. (Figura 10 e Figura 11)
La Figura 10 mostra un insieme di grafici descrittivi relativi alle statistiche glicemiche calcolate per il gruppo di pazienti validati. Viene riportata la distribuzione della media e della mediana del glucosio (pannelli in alto a sinistra e centro), entrambe tendenzialmente comprese tra 130 e 170 mg/dL, con alcune code verso valori elevati. Questo suggerisce un controllo glicemico complessivamente nella norma, sebbene in alcuni casi si osservino valori medi superiori ai limiti raccomandati.
Nel pannello in alto a destra è illustrata la distribuzione della percentuale di dati mancanti (NaN), che risulta generalmente bassa: la maggior parte dei pazienti presenta meno del 3% di valori mancanti, indicando una buona qualità del dataset.
La variabilità glicemica è rappresentata nei boxplot centrali tramite la deviazione standard (SD) e il coefficiente di variazione (CV). I valori medi di SD si aggirano intorno ai 40–50 mg/dL, mentre il CV è prevalentemente compreso tra 0.25 e 0.35.
Questi dati indicano una moderata variabilità glicemica tra i pazienti, con alcuni casi di maggiore instabilità.

Per quanto riguarda l'analisi del tempo trascorso nei diversi intervalli glicemici:

- Il TIR (70–180 mg/dL) risulta elevato nella maggioranza dei pazienti, con valori tra il 60% e il 90%
- Il TBR (<70 mg/dL) è generalmente inferiore al 5%, suggerendo una bassa incidenza di ipoglicemie.
- Il TAR (>180 mg/dL) è più variabile e in alcuni pazienti raggiunge il 60% del tempo segnalando una prevalenza dell'iperglicemia rispetto all'ipoglicemia.

Infine, l'ultimo grafico del boxplot riassume il tempo medio trascorso in ipo e iperglicemia, espresso in minuti: il tempo in iperglicemia risulta nettamente superiore, confermando che questa rappresenta la condizione fuori target prevalente nel campione analizzato.

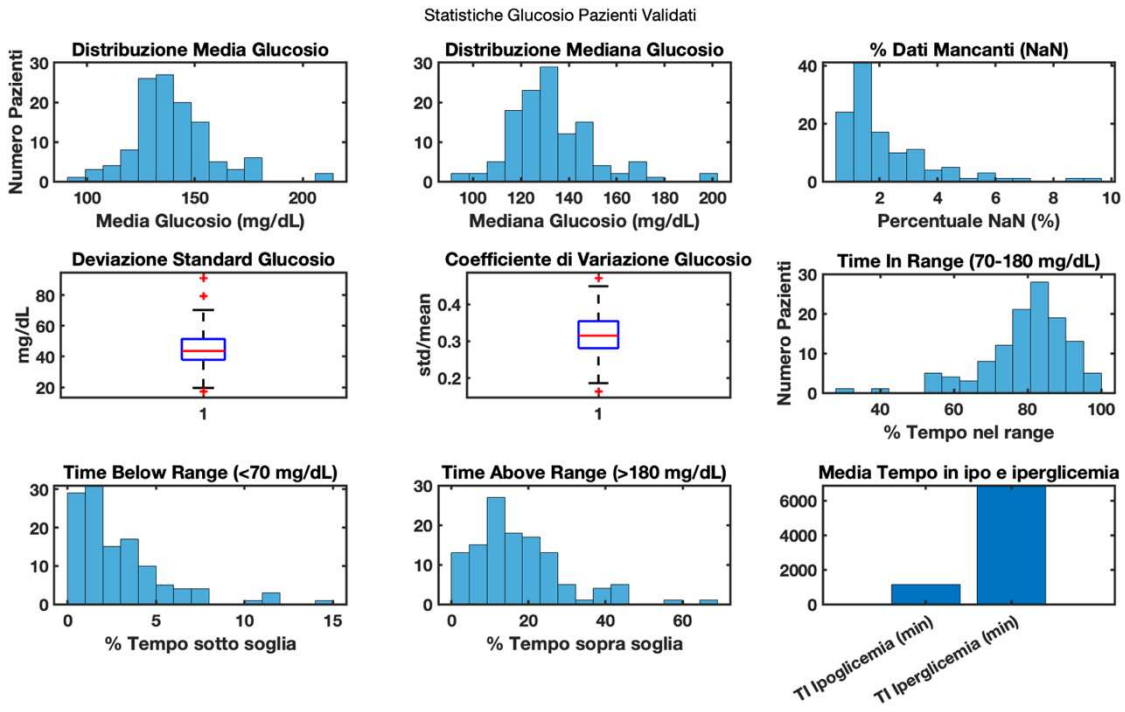

Figura 10: Distribuzione delle principali metriche glicemiche derivate dal monitoraggio continuo della glicemia (CGM) in pazienti con diabete di tipo 1, incluse media e mediana del glucosio, deviazione standard e coefficiente di variazione, percentuale di dati mancanti, Time in Range (70–180 mg/dL), Time Below Range (<70 mg/dL), Time Above Range (>180 mg/dL) e durata media degli episodi di ipo- e iperglicemia.

La Figura 11 riporta un'analisi statistica della frequenza cardiaca nei pazienti validati, considerando valori medi, variabilità, dati mancanti e tempi trascorsi in bradicardia e tachicardia.

Nei pannelli superiori vengono mostrate:

- La distribuzione della frequenza cardiaca media, che evidenzia una concentrazione tra i 70 e gli 85 bpm, con un picco intorno agli 80 bpm.
- La distribuzione della mediana della frequenza cardiaca, che rispecchia quella della media, indicando una buona simmetria nei dati individuali.
- La percentuale di dati mancanti relativi alla HR, con una distribuzione più dispersa: molti pazienti presentano tra il 4% e il 10% di dati mancanti, a differenza della glicemia, per cui i dati risultavano più completi.

I boxplot centrali illustrano la variabilità della frequenza cardiaca:

- La SD è mediamente compresa tra 10 e 25 bpm, con alcune osservazioni oltre i 30 bpm.
- Il CV si distribuisce tra 0.15 e 0.3, con valori più contenuti rispetto a quelli osservati per il glucosio, a conferma di una minore variabilità relativa.

Infine, l'ultimo pannello a destra mostra la durata media (in minuti) trascorsa in condizioni di bradicardia e tachicardia. I due valori risultano simili, indicando che i pazienti validati trascorrono in media tempi comparabili sia in frequenza cardiaca bassa che elevata.

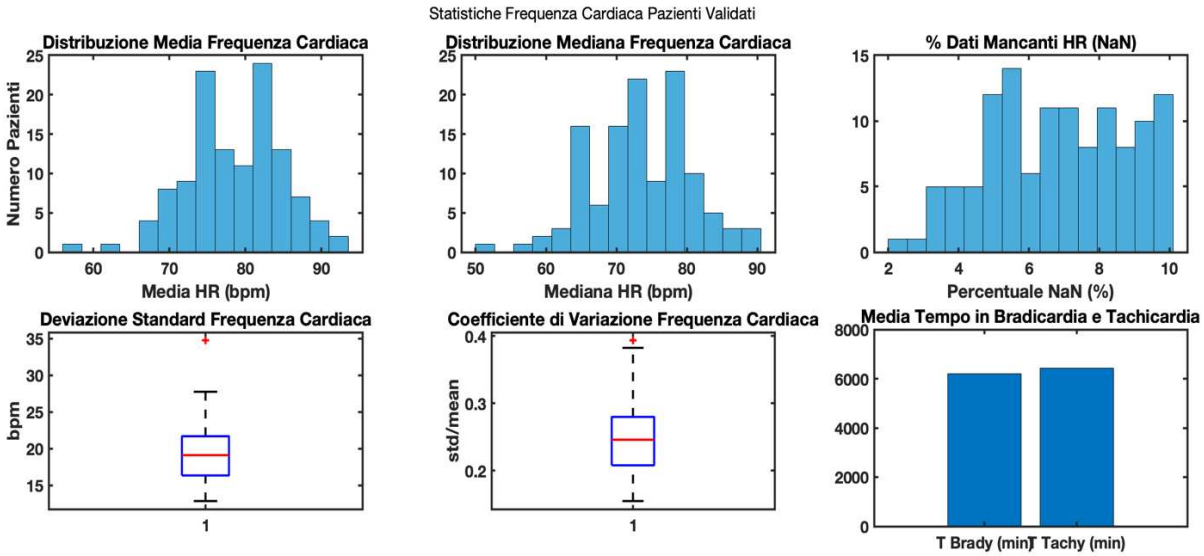

Figura 11: Statistiche descrittive della frequenza cardiaca in pazienti validati. La figura mostra: distribuzione della media e della mediana della frequenza cardiaca (HR, bpm), percentuale di dati mancanti, deviazione standard e coefficiente di variazione della frequenza cardiaca, e la media del tempo trascorso in bradicardia ***e*** *tachicardia(minuti)*

Questa fase di analisi e selezione ha permesso di derivare un dataset raffinato, strutturato e omogeneo, caratterizzato da una copertura temporale completa, su cui si basano le analisi successive.

L'approccio adottato ha garantito un equilibrio tra ampiezza del campione e robustezza delle informazioni, assicurando coerenza tra le diverse fonti dati e una base solida per l'esplorazione di correlazioni complesse tra attività fisica, risposta glicemica e parametri fisiologici.

CAPITOLO 4

In questo capitolo vengono presentati i risultati ottenuti dall'applicazione degli algoritmi descritti nel Capitolo 2 al dataset di pazienti analizzato. Inizialmente vengono definite le principali metriche utilizzate per la valutazione delle performance degli algoritmi di rilevamento dei pasti, sia per i dati di monitoraggio continuo della glicemia (CGM) che per i parametri di frequenza cardiaca. Successivamente, vengono illustrati i risultati sperimentali ottenuti, includendo l'analisi statistica descrittiva del dataset e le performance di detection raggiunte dai diversi approcci metodologici implementati.

4. Valutazione sperimentale e risultati

4.1 Metriche di valutazione

Nel contesto di questo lavoro, il rilevamento di un pasto è considerato corretto se avviene entro ± 30 minuti dall'orario effettivo.

Sulla base di tale criterio, i veri positivi (True Positives, TP), i falsi negativi (False Negatives, FN), i falsi positivi (False Positives, FP) e i veri negativi (True Negatives, TN) vengono definiti secondo le seguenti modalità:

- Sono considerati TP i pasti rilevati correttamente all'interno dell'arco temporale di mezz'ora dalla loro assunzione.
- Sono considerati FN i pasti non rilevati all'interno dell'arco temporale di mezz'ora dalla loro assunzione.
- Sono considerati FP i pasti rilevati in modo errato.
- I TN non vengono considerati, in quanto non si hanno gli strumenti per verificare che un pasto venga correttamente non rilevato.

Come si potrà vedere nel capitolo successivo verrà anche affrontato il caso di allungamento di finestra di rilevazione per confrontare l'analisi.

Durante l'analisi delle performance, è stata inoltre adottata una procedura specifica per la gestione di alcuni falsi positivi associati a eventi glicemici estremi come ipoglicemia imminente o ipoglicemia/iperglicemia in corso.

Questi falsi positivi sono stati identificati attraverso un'analisi del profilo glicemico nelle vicinanze temporali dell'evento rilevato:

- Si è verificata la presenza di una rapida diminuzione del ROC nei 15 minuti precedenti l'evento, segnalando una possibile ipoglicemia imminente.

- Sono stati identificati valori glicemici al di sotto della soglia di 70 mg/dL entro ±15 minuti dall'evento o valori superiori a 180 mg/dL nell'intervallo da 5 minuti prima a 60 minuti dopo l'evento.

Gli eventi falsi positivi che soddisfacevano uno di questi criteri sono stati considerati giustificati e quindi esclusi dal calcolo delle metriche di performance, al fine di ottenere una valutazione più accurata e clinicamente significativa dell'algoritmo. Tale approccio ha permesso di ridurre l'impatto di possibili errori di etichettatura manuale e di valorizzare la rilevanza clinica degli eventi identificati.

Le performance dell'algoritmo vengono valutate utilizzando le seguenti metriche di valutazione:

1. La precisione, detta anche positive predictive value, misura la capacità di un algoritmo di classificazione di limitare i falsi positivi. Più precisamente, rappresenta la proporzione di predizioni positive che corrispondono effettivamente a casi reali positivi. Un'alta precisione indica che, tra tutte le istanze segnalate come positive dal modello, la maggior parte è correttamente individuata. Questa metrica è particolarmente importante nei contesti in cui è prioritario evitare segnalazioni errate o allarmi ingiustificati, che potrebbero generare interventi non necessari o falsi allarmi.

$$precisione = \frac{TP}{TP + FP} * 100$$

2. La sensitività, nota anche come recall o true positive rate, rappresenta la capacità di un algoritmo di identificare correttamente tutti i casi positivi presenti nel dataset. Dal punto di vista formale, essa esprime la proporzione di veri positivi rispetto al totale dei casi realmente positivi, ovvero la probabilità che un evento positivo venga correttamente rilevato dal modello. Un'elevata sensibilità indica che l'algoritmo commette pochi falsi negativi, risultando particolarmente utile in contesti in cui è cruciale non trascurare condizioni cliniche rilevanti.

$$sensitività = \frac{TP}{TP + FN} * 100$$

3. L'F1-score è una misura di performance che sintetizza in un singolo valore la capacità di un algoritmo di classificazione di identificare correttamente i casi positivi, bilanciando sensitività (recall) e precisione. Si calcola come la media armonica tra precisione e sensibilità.

$$F1_{score} = \frac{2 * precisione * sensitività}{precisione + sensitività}$$

L'F1-score penalizza in modo più severo i modelli che ottengono un valore molto basso in una delle due metriche. In questo modo, fornisce una valutazione più equilibrata delle prestazioni complessive, risultando particolarmente utile in situazioni in cui è importante contenere sia il numero di falsi positivi che di falsi negativi.

4. Il tempo di rilevazione del pasto da parte dell'algoritmo, noto anche come detection time, è definito come il ritardo, espresso in minuti, tra il momento in cui l'algoritmo riconosce internamente la presenza di un pasto e l'istante in cui viene emessa la conferma ufficiale della rilevazione. In questo contesto, il detection time non è calcolato rispetto all'orario di riferimento del pasto reale, ma piuttosto riflette la latenza interna dell'algoritmo nel consolidare la decisione di rilevazione.

5. L'errore di rilevazione misura la differenza temporale (in minuti) tra l'orario effettivo del pasto (ground truth) e il momento in cui l'algoritmo rileva tale evento.
   Viene calcolato solo sui pasti correttamente identificati e fornisce una misura diretta dell'accuratezza temporale del rilevamento.
   È particolarmente utile per valutare quanto tempestivamente l'algoritmo riesce a rilevare un pasto, aspetto cruciale in contesti clinici in cui il tempismo dell'intervento è fondamentale.
   Nei risultati presentati sarà indicato come difference time.

6. L'F1-score da solo non tiene conto dei ritardi nella rilevazione: un pasto rilevato con precisione, ma con ritardo significativo, può essere poco utile in applicazioni real-time. Per superare questa limitazione, si introduce l'F1-delay, una metrica estesa che penalizza i ritardi temporali nella rilevazione rispetto all'evento reale. Questa metrica integra la correttezza dell'F1-score con la tempestività, permettendo una valutazione più completa dell'efficacia del metodo. In formula:

$$F_{1-delay} = \left(\frac{F_{1-score}}{1+\left(\frac{Errore\ di\ rilevazione}{Finestra\ di\ tolleranza}\right)}\right)$$

dove l'errore di rilevazione rappresenta la differenza in minuti tra la rilevazione e l'evento reale, mentre la finestra di tolleranza definisce l'intervallo massimo accettabile per considerare corretta la rilevazione. Questa formulazione fornisce una

valutazione più rigorosa dell'efficacia temporale del metodo, integrando sia accuratezza che tempestività.

7. In aggiunta alle metriche sopra descritte, una misura di performance importante utilizzata in questo lavoro è la frequenza dei falsi positivi normalizzata al giorno (FP/day). Questa metrica rappresenta il numero medio di falsi positivi rilevati dall'algoritmo per ogni giorno di monitoraggio, fornendo una stima diretta del carico di segnalazioni errate nell'arco temporale dell'analisi. Un valore basso di FP/day è particolarmente desiderabile, poiché indica un ridotto numero di allarmi non giustificati, migliorando l'usabilità e la praticità del sistema di rilevazione in condizioni reali.

4.2 Risultati ottenuti

4.2.1. Valutazione su singolo paziente

Per una prima analisi puntuale delle prestazioni, gli algoritmi descritti nel Capitolo 2 sono stati applicati al paziente 110, scelto come caso rappresentativo. Questa valutazione ha permesso di esaminare in dettaglio la capacità di ciascun metodo di rilevare correttamente gli eventi (pasti) e la loro accuratezza temporale.

METODO	PASTI	TP	FP	FN	Recall	Precision	F1-score	Delta Detection Time	Difference Time
FUZZY LOGIC	21	15	14	6	0.71	0.51	0.6	15.5 min	5.6 min
GRID	21	11	9	10	0.550	0.524	0.537	25.67 min	13 min
ROC	21	18	33	3	0.85	0.34	0.48	20 min	13.8 min
HR+CGM	21	14	11	7	0.66	0.56	0.6	38 min	11.4 min
GIM	21	16	2	5	0.76	0.88	0.82	15 min	11.76 min
DATA-DRIVEN	21	12	11	9	0.57	0.55	0.55	60 min	7.5 min
MOSQUERA	21	4	0	17	0.19	1	0.31	50 min	21.25 min

Tabella 2 : Risultati sul paziente 110

Dal confronto emerge che l'algoritmo di GIM [22] mostra il miglior bilanciamento tra sensibilità e precisione, raggiungendo il valore più elevato di F1-score (0.82), con un numero molto contenuto di falsi positivi. Anche Fuzzy Logic [19] e HR+CGM [23] evidenziano buone performance complessive, pur con livelli di precisione leggermente inferiori.

L'algoritmo ROC, pur presentando un'elevata sensibilità (recall = 0.85), soffre di una bassa precisione (0.34) a causa dell'elevato numero di falsi positivi, suggerendo una tendenza a sovrastimare la presenza di pasti.

Al contrario, l'algoritmo Mosquera [26] risulta estremamente conservativo: rileva soltanto 4 eventi su 21 reali (recall = 0.19), ma con una precisione perfetta (1.00), indicando che ogni rilevazione è corretta ma accompagnata da una forte sottostima.

Infine, l'algoritmo Data-driven [21] mostra tempi di rilevazione notevolmente più lenti rispetto agli altri metodi (fino a 60 minuti di ritardo medio), pur mantenendo un equilibrio accettabile tra recall e precisione.

Dal confronto dei valori di Difference Time, che quantifica l'accuratezza temporale della rilevazione, si osserva come gli algoritmi differiscano non solo nella capacità di rilevare correttamente un pasto, ma anche nella precisione temporale con cui lo stimano.

Algoritmi come GIM [22] e Fuzzy Logic[19] si dimostrano più affidabili nel posizionamento temporale degli eventi, aspetto fondamentale per l'impiego in contesti clinici reattivi, dove anche scostamenti di pochi minuti possono incidere sull'efficacia dell'intervento terapeutico.

Al contrario, algoritmi come Mosquera [26] e ROC evidenziano una maggiore incertezza temporale, che ne limita la prontezza d'uso.

Questa valutazione, applicata a un singolo paziente, ha permesso di evidenziare le caratteristiche comportamentali di ciascun metodo, fornendo un primo riscontro qualitativo sul tipo di errore prevalente e sulla reattività temporale.

Di seguito è riportata una figura rappresentativa che mostra l'andamento del glucosio e le rilevazioni dei pasti effettuate dai diversi algoritmi sul test set. I grafici illustrano, per tale soggetto, la dinamica glicemica osservata, i pasti reali (ground truth) e quelli rilevati automaticamente, permettendo un confronto visivo diretto tra la risposta fisiologica e le predizioni algoritmiche.

Paziente 110

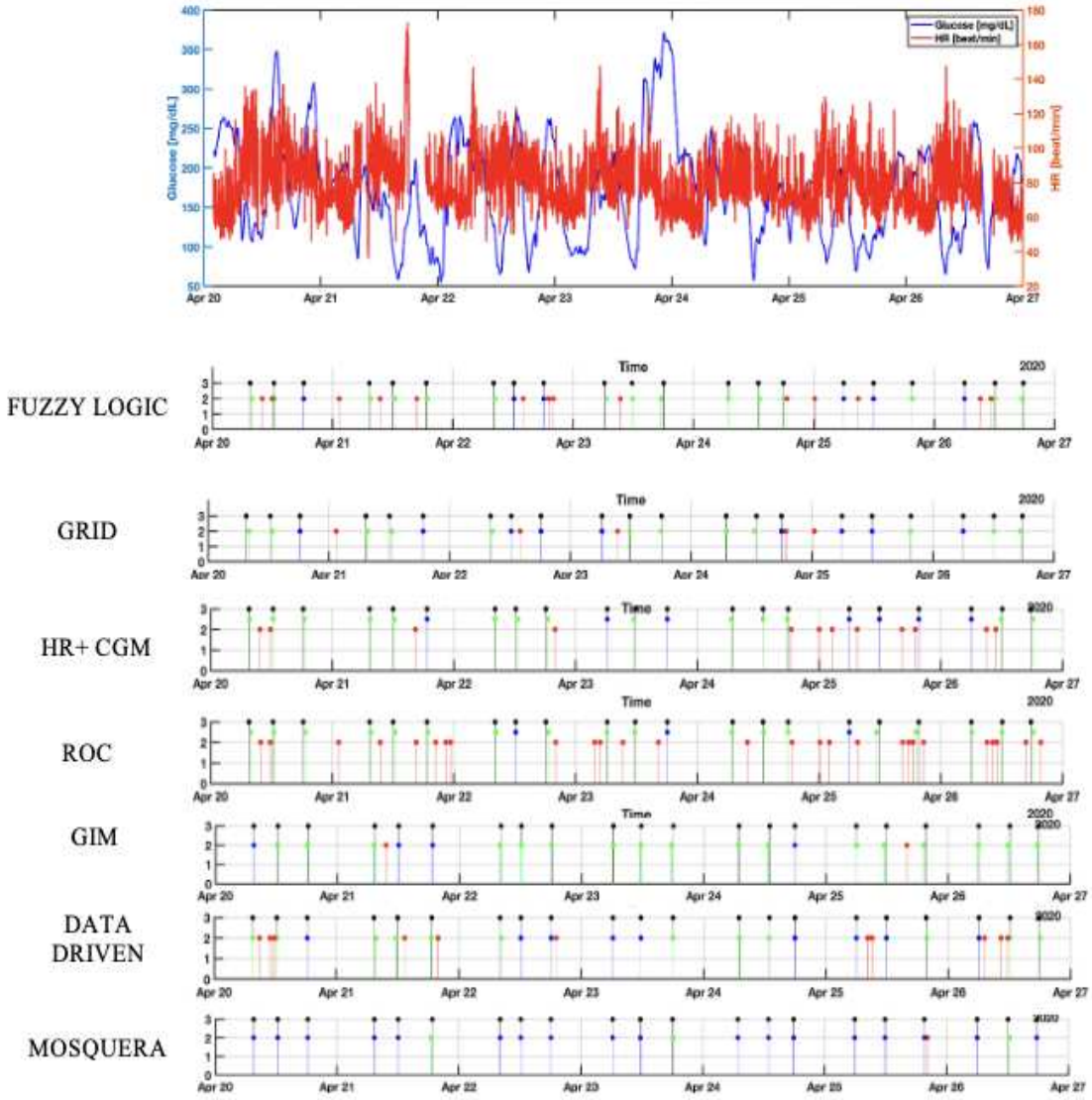

Figura 12: Performance dell'algoritmo di meal detection nel paziente 110. I marker rappresentano: meal flag di riferimento (nero), pasti correttamente identificati - True Positives (verde), pasti non rilevati - False Negatives (blu), e false identificazioni - False Positives (rosso).

La Figura 12 mostra il confronto tra diversi algoritmi di rilevamento pasti applicati al paziente 110, utilizzando dati continui di glicemia (Glucose, linea blu, mg/dL) e frequenza cardiaca (HR, linea rossa, bpm) registrati nel periodo 20–27 aprile 2020. Nella parte superiore sono riportate le due serie temporali, con l'asse sinistro riferito alla glicemia e l'asse destro alla frequenza cardiaca, evidenziando incrementi post-prandiali e variazioni dovute a fattori fisiologici o comportamentali. Nella parte inferiore, ciascun pannello corrisponde a uno specifico algoritmo) e presenta uno stem plot in cui i marker neri indicano i pasti annotati, i marker verdi i TP, i marker rossi i FP e l'assenza di marker in corrispondenza di pasti reali rappresenta FN. L'analisi visiva evidenzia che Fuzzy Logic[19] e GRID [20] ottengono un'elevata sensibilità con numerosi TP ma generano anche un maggior numero di FP, mentre

HR+CGM[23] e GIM [22]risultano più specifici, riducendo i FP ma aumentando i FN. Data-driven [21] mantiene un compromesso tra sensibilità e specificità, pur mostrando talvolta offset temporali nei rilevamenti rispetto al pasto reale. L'approccio Mosquera [26], basato su logiche soglia, presenta invece bassa sensibilità con prevalenza di FN.

In alcuni algoritmi si osservano rilevamenti multipli in prossimità dello stesso pasto, verosimilmente dovuti a variazioni prolungate o fluttuanti del segnale che portano a riattivazioni dell'algoritmo entro il periodo di osservazione, per tale motivo si è utilizzata l'applicazione di un periodo di cooldown. La rappresentazione tramite stem plot consente così di valutare in modo immediato e ad alta risoluzione temporale il compromesso precisione–recall e l'allineamento tra eventi reali e predetti.

Si è proseguito con l'estensione dell'analisi a tutto il dataset

4.2.2 Valutazione su intero dataset

Per fornire una valutazione comparativa più robusta delle prestazioni degli algoritmi, è stata effettuata un'analisi aggregata su tutti i pazienti del test set. La Figura 13 mostra i boxplot delle sei metriche principali: precisione, recall, F1-score, FP/day, tempo medio di rilevazione (Detection Time) ed errore assoluto medio (Difference Time). Questa rappresentazione consente di apprezzare non solo le medie delle prestazioni, ma anche la loro variabilità inter-soggetto.

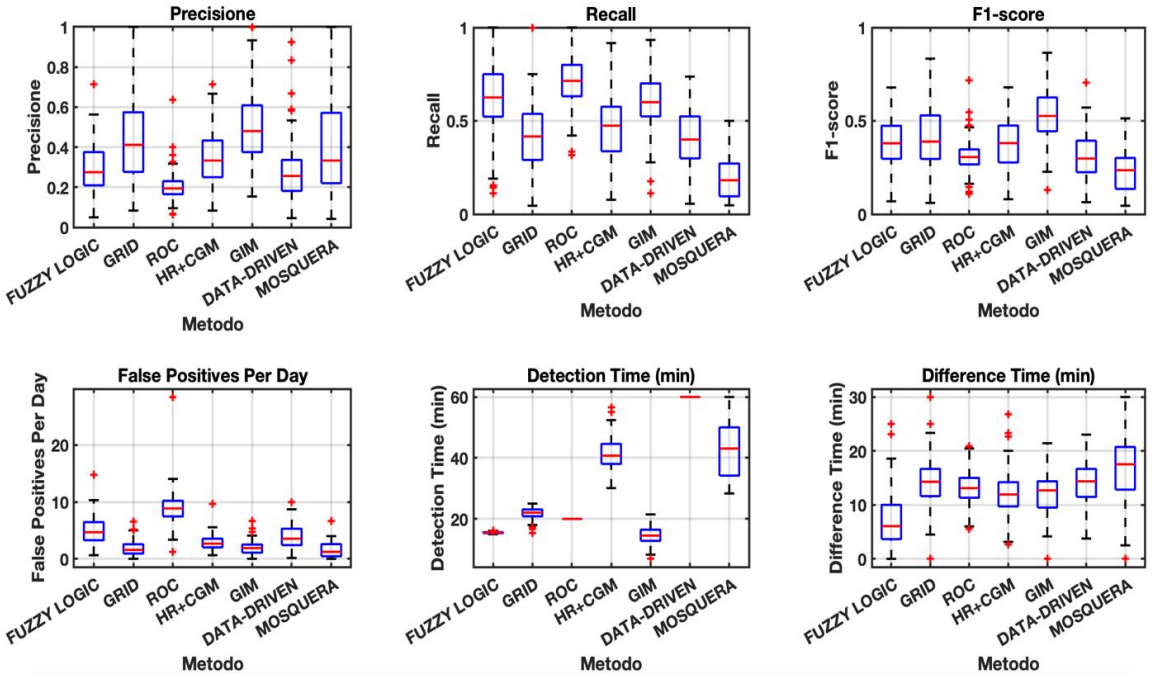

Figura 13: Boxplot relativo all'analisi su tutti i pazienti.

L'algoritmo Fuzzy Logic[19] mostra una buona stabilità complessiva, con recall elevato e uno dei tempi medi di rilevazione più bassi. Tuttavia, la precisione risulta moderata, indicando una certa incidenza di falsi positivi. Il metodo GRID[20] si distingue per l'alta precisione e un basso numero di FP/day, ma con recall più basso, segno di una tendenza conservativa che può portare a mancata rilevazione di alcuni pasti. Al contrario, l'approccio ROC raggiunge il recall più elevato, ma con una precisione molto bassa e una frequenza di falsi positivi significativamente maggiore, penalizzando l'F1-score. Il metodo HR+CGM [23]mostra prestazioni più bilanciate, con valori centrali medi e una variabilità contenuta, pur evidenziando un Detection Time relativamente alto.

Il metodo GIM [22], coerentemente con l'analisi su singolo paziente, si conferma tra i metodi più affidabili, con un buon compromesso tra precisione, recall e tempestività. Il metodo Data-driven[21], pur mostrando un basso numero di FP/day, presenta un Detection Time costantemente elevato (valore fisso a 60 minuti), compromettendone l'utilità pratica in tempo reale. Infine, il metodo Mosquera[26] si colloca tra quelli con le performance più conservative, caratterizzate da alta precisione, ma bassa sensibilità e tempi medi di rilevazione tra i più alti.

Nel complesso, l'analisi aggregata evidenzia che nessun algoritmo domina su tutte le metriche, confermando l'importanza di un approccio di valutazione multidimensionale, che consideri tanto la correttezza delle rilevazioni quanto la loro tempestività e affidabilità operativa in contesti reali.

Tali risultati sono in linea con quanto ci si può aspettare da un contesto realistico: il dataset utilizzato è infatti caratterizzato da dati clinici reali, eterogenei e affetti da rumore, discontinuità e variabilità interindividuale, elementi che rendono il compito di rilevazione dei pasti particolarmente sfidante.
Questi costituiscono la base per la successiva discussione, dove verranno analizzati in dettaglio i boxplot di F1-score e introdotte le metriche di qualità del segnale , al fine di comprendere le possibili cause delle prestazioni limitate e valutare il loro impatto sull'efficacia dei metodi di rilevamento dei pasti.

4.2.3 Validazione dell'analisi con estensione del margine temporale di rilevamento

Per valutare la robustezza degli algoritmi e la sensibilità delle metriche di performance rispetto alla precisione temporale, è stata condotta un'analisi supplementare estendendo la tolleranza temporale per la corrispondenza tra pasto rilevato e pasto reale da ±30 minuti a ±45 minuti. L'ampliamento di questa finestra tiene conto della possibile variabilità fisiologica tra

il momento effettivo dell'assunzione del pasto e l'effetto osservabile nei segnali (glucosio, frequenza cardiaca), oltre a eventuali ritardi sistematici introdotti dagli algoritmi. Inoltre, l'incremento a ± 45 minuti permette di valutare il comportamento degli algoritmi in scenari più realistici, dove un certo grado di imprecisione temporale è accettabile.

Di seguito sono mostrati i box plot delle principali metriche di performance, evidenziando le differenze emerse:

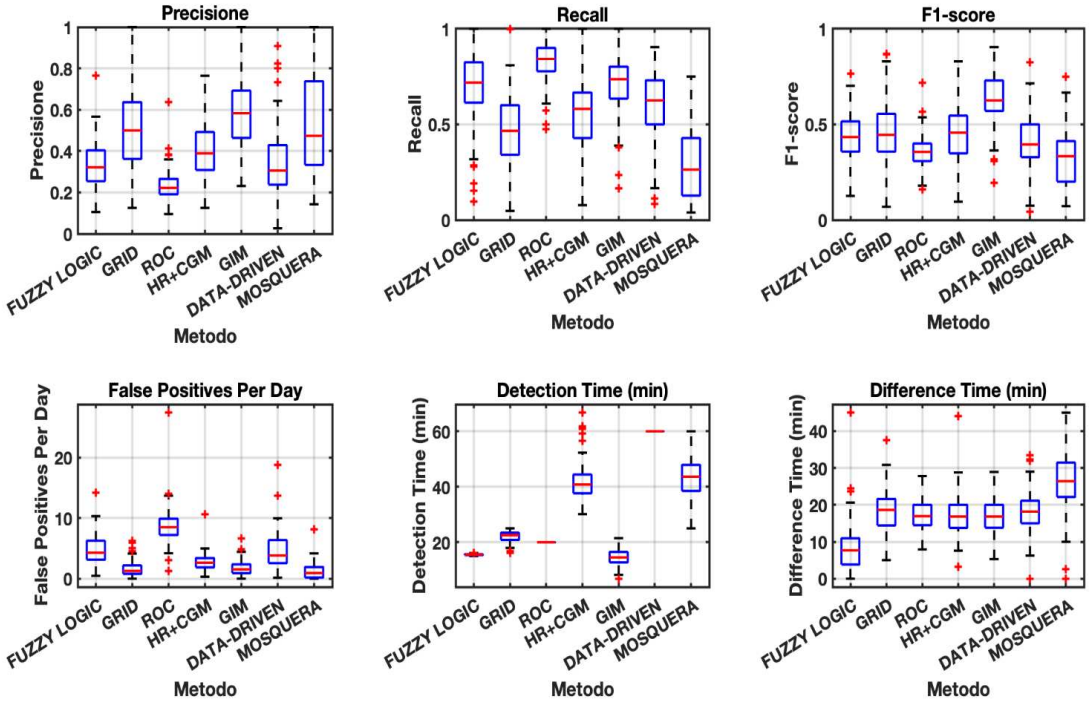

Figura 14 : Boxplot relativo all'analisi su tutti i pazienti con aumento di margine di tolleranza temporale

All'aumentare della tolleranza temporale si osservano i seguenti effetti generali:

- Recall in aumento per quasi tutti gli algoritmi: un margine più ampio consente di classificare come corretti anche rilevamenti temporalmente imprecisi.
- Leggero miglioramento dell'F1-score in alcuni casi, non sempre proporzionale all'aumento del recall.
- Precisione invariata o leggermente ridotta, poiché i falsi positivi rimangono numerosi.
- Riduzione dei valori medi di Difference Time, segno che un maggior numero di rilevamenti cade entro il margine considerato.

L'obiettivo era verificare se un aumento del margine di rilevamento potesse migliorare le prestazioni, in particolare l'F1-score, mantenendo al contempo significatività clinica.

Tutti gli algoritmi analizzati mostrano un incremento dell'F1-score con l'allargamento della finestra temporale, indicando che una parte delle predizioni corrette si verifica con un piccolo ritardo o anticipo rispetto ai 30 minuti.

Algoritmo	**F1-score (± 30 min)**	**F1-score (± 45 min)**
FUZZY LOGIC	0.386 ± 0.130	0.438±0.124
GRID	0.401 ± 0.169	0.463 ± 0.169
ROC	0.309 ± 0.078	0.361 ±0.083
HR+CGM	0.383 ± 0.131	0.452 ±0.138
GIM	0.540 ± 0.14	0.635 ±0.133
DATA-DRIVEN	0.311 ± 0.118	0.413 ± 0.134
MOSQUERA	0.228 ± 0.105	0.313 ± 0.146

Tabella 3 : Confronto performance con tempi di rilevamento differenti

In particolare:

- GIM [22] e Data-driven [21] sono tra i metodi che beneficiano maggiormente dell'estensione del margine, con incrementi rispettivamente di +0.095 e +0.102 punti F1.
- Mosquera [26] e HR+CGM[23] migliorano in modo marcato, anche se le performance assolute restano inferiori rispetto ai metodi più efficaci.
- Anche ROC, pur con precisione limitata, guadagna +0.052 di F1-score, evidenziando un buon recupero nei casi in cui il rilevamento è leggermente sfasato.

Questi risultati confermano l'importanza di includere un'analisi della tolleranza temporale nella valutazione delle performance, specialmente per algoritmi destinati all'uso in tempo

reale. Aumentare il margine di tolleranza può migliorare sensibilmente la rilevanza pratica del sistema, purché non si ecceda al punto da includere falsi positivi non correlati all'evento reale.

Come anticipato nel Capitolo 2, per approfondire ulteriormente l'accuratezza del rilevamento, nella sezione successiva verrà esaminato l'algoritmo HR+CGM [23] integrato con la derivata seconda del glucosio, al fine di valutare come questo approccio possa influenzare la tempestività e l'affidabilità delle rilevazioni.

4.2.4 Valutazione dell'algoritmo HR+CGM con l'ausilio della derivata seconda del glucosio

In questa sezione vengono presentati i risultati ottenuti dall'applicazione dell'algoritmo HR+CGM[23]integrato con l'indice μ, come descritto nella sezione metodologica.

L'indice μ, combinando la derivata prima e seconda del segnale glicemico, consente di rafforzare la discriminazione tra eventi prandiali e variazioni glicemiche non correlate ai pasti. L'analisi dei risultati mira a valutare l'impatto di questa integrazione sulla riduzione dei falsi positivi e sulla robustezza complessiva dell'algoritmo in condizioni di variabilità glicemica elevata o rumore di segnale.

Metodo	Precision	Recall	F1-score	FP/day	Delta Detection Time(min)	Difference Time(min)
HR+CGM	0.340 ± 0.128	0.462 ± 0.163	0.383 ± 0.131	2.881 ± 1.217	41.245 ± 5.398	12.230 ± 3.782
HRCGM(μ)	0.520 ± 0.232	0.310 ± 0.176	0.355 ± 0.175	1.039 ± 0.846	28.192 ± 2.994	12.433 ± 5.266

Tabella 4 : Valutazione performance con introduzione dell'indice μ

I risultati ottenuti confrontano l'algoritmo HR+CGM[23] standard con la versione integrata dall'indice μ.

L'integrazione dell'indice ha comportato un miglioramento significativo della precisione , indicando una maggiore affidabilità nel confermare gli eventi prandiali rilevati. Questo

suggerisce che l'indice μ consente di ridurre l'influenza delle fluttuazioni glicemiche non correlate ai pasti.

Al contrario, il recall ha mostrato una diminuzione evidenziando che alcuni eventi prandiali reali non sono stati rilevati. Questo comportamento riflette un compromesso tipico tra selettività e sensibilità: l'integrazione dell'indice migliora la confidenza sugli eventi rilevati, a scapito della capacità di intercettare tutti i pasti.

L'F1-score è rimasto sostanzialmente invariato, confermando che l'adozione dell'indice μ sposta l'algoritmo verso una maggiore selettività senza peggiorare drasticamente la performance complessiva.

Un risultato particolarmente rilevante riguarda i falsi positivi per giorno, che si riducono significativamente passando da 2,881 ± 1,217 a 1,039 ± 0,846. Questo evidenzia come l'indice μ contribuisca a discriminare meglio le variazioni glicemiche genuine dai picchi prandiali, migliorando la robustezza dell'algoritmo in condizioni di rumore o elevata variabilità glicemica.

La differenza temporale rispetto al pasto reale risulta sostanzialmente invariata tra le due versioni dell'algoritmo, indicando che l'integrazione dell'indice μ non compromette la precisione della stima temporale dei pasti.

Tuttavia, come negli altri casi, anche qui le performance complessive dei metodi non risultano ottimali; per questo motivo, il capitolo successivo sarà dedicato a un'analisi approfondita di tutti i metodi per comprenderne le cause.

CAPITOLO 5

5. Discussione

L'analisi dei risultati evidenzia come, in diversi scenari, le prestazioni degli algoritmi testati siano inferiori alle aspettative, mostrando valori di precisione e recall spesso ridotti.

In questa sezione si indagano le possibili cause di tali criticità, considerando fattori come la qualità del segnale acquisito, la presenza di rumore o artefatti nei dati e la variabilità inter e intra-soggetto nelle risposte glicemiche e cardiache.

Per raggiungere questo scopo, l'analisi tiene conto di indicatori quantitativi, come le metriche di performance standard, e di osservazioni qualitative derivanti dall'esame dei tracciati temporali, allo scopo di identificare pattern di errore ricorrenti e correlazioni significative.

L'obiettivo principale è comprendere se le prestazioni limitate siano imputabili a limiti degli algoritmi, a caratteristiche del dataset o a una combinazione di entrambi. Questa valutazione approfondita consente di individuare criticità specifiche e suggerire possibili strategie di miglioramento nella progettazione degli algoritmi.

5.1 Analisi delle metriche di rumore e correlazione con le performance

Per ogni metodo di rilevamento pasti, sono state calcolate specifiche metriche di rumore volte a caratterizzare la qualità e la variabilità dei segnali glicemico.

Le metriche considerate includono:

- Entropia del segnale: misura quantitativa della complessità e della variabilità irregolare del segnale, indicativa di dinamiche non prevedibili e potenziali fonti di disturbo. In ambito biomedico, l'analisi dell' entropia è stata applicata sia a segnali ECG per valutarne la qualità [29], sia a serie temporali fisiologiche complesse, incluse quelle glicemiche [30]. L'impiego di misure entropiche consente di stimare il grado di irregolarità che può ostacolare il rilevamento accurato degli eventi alimentari.
  Ciò la rende una scelta appropriata per quantificare il grado di complessità che può ostacolare il rilevamento accurato degli eventi alimentari.
- RMSSD (Root Mean Square of Successive Differences): indicatore della variabilità a breve termine. Studi clinici hanno evidenziato correlazioni significative tra RMSSD e

parametri glicemici nonché una riduzione del RMSSD in soggetti con peggior controllo glicemico o maggiore durata del diabete [31].
L'estensione di RMSSD all'analisi del segnale glicemico consente di rilevare fluttuazioni rapide tra misurazioni successive, potenzialmente riconducibili a rumore o a variazioni fisiologiche rilevanti per la rilevazione pasti.

- Potenza in alta frequenza (HF): componente spettrale del segnale nella banda 0,15–0,40 Hz, generalmente associata alla modulazione parasimpatica dell'attività cardiaca e fortemente correlata con RMSSD. In ambito diabetologico, valori ridotti di HF sono stati associati a disfunzione autonomica e peggioramento del controllo glicemico [31],[32].
  La valutazione della potenza HF nel dominio della frequenza offre una prospettiva complementare rispetto a RMSSD, permettendo di distinguere tra fluttuazioni fisiologiche e rumore ad alta frequenza.
- Outlier ratio e percentili: il rilevamento di valori anomali è stato effettuato tramite il criterio dell'intervallo interquartile (IQR). L'outlier ratio misura la proporzione di valori al di fuori dell'intervallo accettabile, fornendo un indice della robustezza del segnale rispetto a deviazioni estreme. Contestualmente, l'analisi dei percentili (25°, 50° e 75°) consente di descrivere la distribuzione centrale e la variabilità intrinseca della glicemia.
- CV: definito come il rapporto tra la deviazione standard e la media del segnale, espresso in percentuale. Questo indice sintetizza la variabilità relativa del segnale, distinguendo soggetti con fluttuazioni glicemiche fisiologiche da quelli con segnali caratterizzati da maggiore instabilità o rumorosità.
- Rapporto segnale-rumore (SNR): misura la proporzione tra la potenza del segnale utile e quella del rumore residuo. Valori più elevati indicano una migliore qualità del segnale e una minore contaminazione da rumore. In contesto glicemico, un basso SNR riflette una predominanza di oscillazioni ad alta frequenza o artefatti che possono compromettere l'affidabilità della meal detection.[33]

Tale approccio è fondamentale per interpretare in maniera più robusta le prestazioni degli algoritmi di meal detection, evidenziando i casi in cui le ridotte performance possano essere attribuite a una bassa qualità del segnale piuttosto che a limiti intrinseci del metodo.

I boxplot riportati in figura confrontano le prestazioni dei diversi metodi di identificazione automatica dei pasti valutate tramite F1 score e F1 delay.

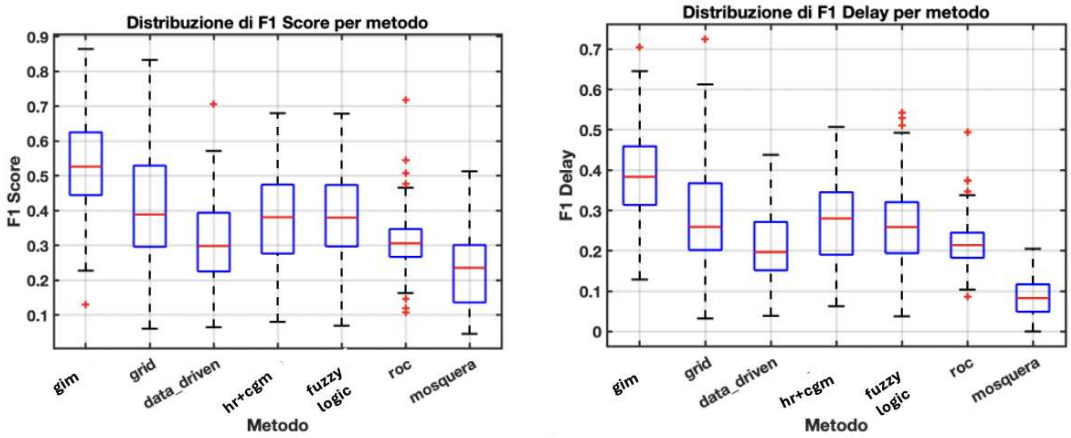

Figura 15. Distribuzione delle metriche di performance per metodo

Il confronto tra F1-score e F1-delay evidenzia quanto l'introduzione della componente temporale nella valutazione sia cruciale per una stima realistica delle prestazioni degli algoritmi di meal detection. In particolare, alcuni metodi, come Fuzzy Logic[19], GRID[20] e HR+CGM[23], mostrano inizialmente valori discreti di F1-score, indicativi di una buona capacità di identificare correttamente i pasti. Tuttavia, quando si considera anche la tempestività della rilevazione (F1-delay), si osserva un ridimensionamento significativo, spesso superiore al 15–20%, confermando che l'accuratezza da sola non riflette l'efficacia reale dell'algoritmo in contesti real-time.

Il calo risulta ancora più marcato per gli algoritmi ROC e Data-Driven[21], che già partono da prestazioni moderate: la mancata tempestività accentua le loro limitazioni, evidenziando che i pasti rilevati fuori finestra temporale non contribuiscono positivamente alla valutazione.

L'algoritmo Mosquera[26] si conferma come il meno performante, dove la rapidità nella rilevazione non compensa la scarsa accuratezza complessiva, portando a valori molto bassi sia di F1-score sia di F1-delay.

Al contrario, il metodo GIM[22] si distingue per la sua robustezza complessiva: mantiene il miglior punteggio sia di F1-score che di F1-delay, dimostrando di combinare efficacemente precisione e tempestività. Questo conferma la sua affidabilità anche in scenari real-time, dove non è sufficiente individuare correttamente i pasti, ma è fondamentale farlo entro una finestra temporale strettamente definita.

Nel paragrafo successivo valuteremo la correlazione tra questi due indici di performance e il rumore.

5.2 Interpretazione della correlazione tra variabili

In questa sezione vengono esaminati i risultati dell'analisi di correlazione tra le metriche di qualità dei segnali glicemici e le prestazioni dei diversi algoritmi di rilevamento pasti.
L'obiettivo è identificare quali caratteristiche del segnale possano influenzare significativamente l'accuratezza e la tempestività del rilevamento.
A tal fine, è stato condotto uno studio sistematico sui segnali CGM del test set, seguendo un approccio in più fasi. In primo luogo, per ciascun soggetto sono state estratte le ultime porzioni temporali dei segnali, corrispondenti al 25% finale della registrazione, e sono stati associati i punteggi F1 ottenuti dai diversi algoritmi.
Successivamente, sono state calcolate diverse metriche di rumore e variabilità del segnale che permettono di caratterizzare sia le fluttuazioni fisiologiche sia le eventuali anomalie o artefatti presenti nei dati, elementi che possono compromettere la rilevazione dei pasti.
Per ciascun algoritmo, le metriche di qualità del segnale sono state correlate con i punteggi F1 mediante coefficienti di Pearson e Spearman, e le relazioni più significative sono state visualizzate tramite scatter plot[34]. Inoltre, per i metodi maggiormente sensibili al rumore, sono stati generati boxplot dei quartili delle metriche e analizzate le tracce con il maggiore impatto negativo sulle performance.
Infine, per quantificare globalmente quanto queste variabili spiegano la variabilità nelle performance, è stata effettuata una regressione multipla, il cui coefficiente di determinazione adjusted (R^2 adj) indica la frazione della variabilità spiegata dal modello.[35]
Questo approccio consente di quantificare sistematicamente l'impatto delle caratteristiche del segnale CGM sulla performance dei vari algoritmi di rilevamento pasti, fornendo un quadro completo della qualità dei dati e della loro influenza sull'accuratezza e tempestività delle rilevazioni.
Nella figura sottostante (Figura 16) viene riportato il boxplot che mostra la distribuzione delle prestazioni (F1-score) di alcuni metodi di meal detection in relazione alle principali metriche di rumore del segnale CGM.

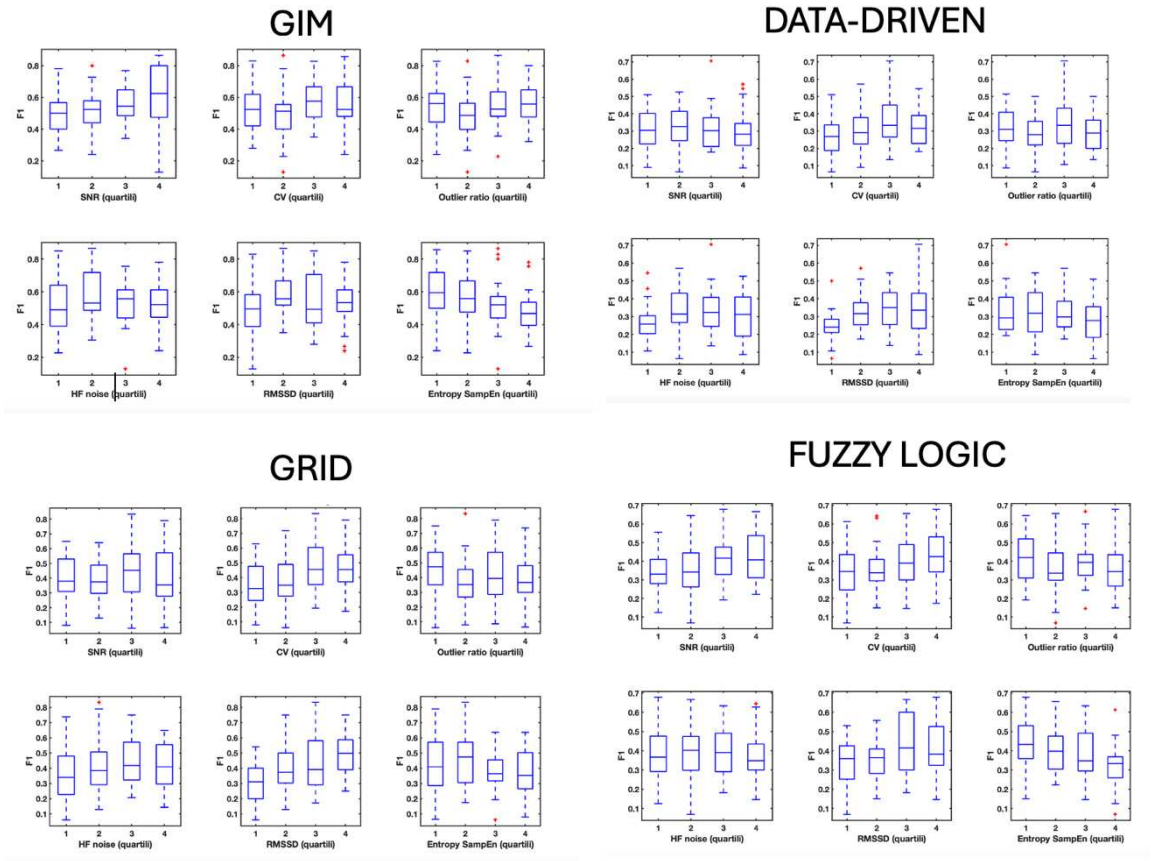

Figura 16. Distribuzione delle prestazioni dei metodi di meal detection in relazione al rumore del segnale CGM utilizzando come metrica di performance l'F1 score

Analizzando i singoli metodi, emergono differenze marcate nella loro robustezza al rumore. Il metodo GRID[20] mostra un comportamento bilanciato, con un miglioramento moderato ma costante delle prestazioni all'aumentare del rapporto segnale-rumore (SNR) e una sensibilità contenuta rispetto alla presenza di outlier e al rumore ad alta frequenza (HF). Il metodo Data-driven[21], al contrario, si conferma il più sensibile: l'F1-score cala in modo marcato in presenza di segnali degradati, in particolare con alti valori di coefficiente di variazione (CV), outlier o rumore HF, indicando una forte dipendenza dalla qualità del segnale e una conseguente ridotta affidabilità in scenari reali. Il metodo Fuzzy Logic[19] si distingue per l'elevata robustezza: le sue prestazioni rimangono relativamente stabili anche in condizioni di rumore, con variazioni dell'F1-score minime. Tuttavia, questa stabilità si accompagna a una limitata capacità di sfruttare segnali di alta qualità, come suggerito dai miglioramenti modesti in presenza di SNR elevato. Infine, il metodo GIM [22] rappresenta una soluzione intermedia, con performance che migliorano moderatamente in condizioni favorevoli, ma che risultano penalizzate dalla presenza di outlier e da alti livelli di entropia, segnalando una sensibilità di livello medio.

Viene introdotta anche un'analisi di correlazione con l'F1 delay e nella figura qui riportata si evincono le principali relazioni tra performance e metriche di rumore.

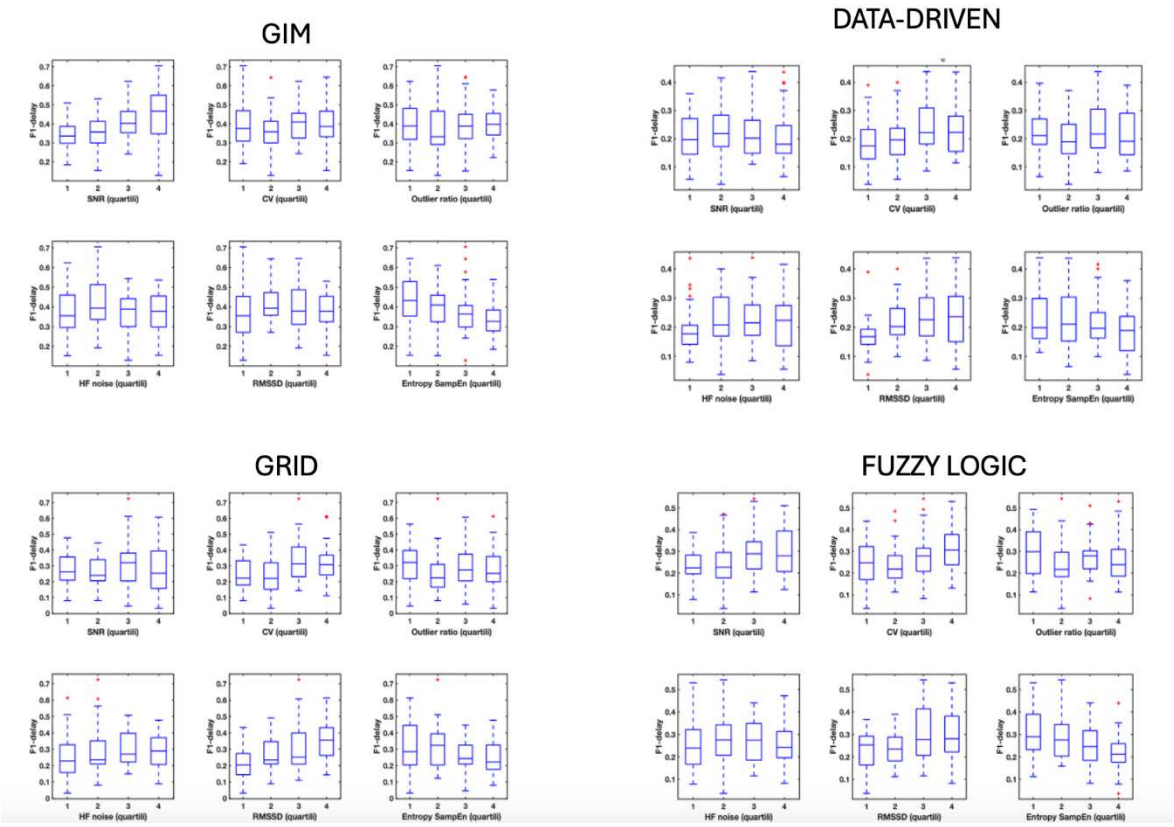

Figura 17. Distribuzione delle prestazioni dei metodi di meal detection in relazione al rumore del segnale CGM utilizzando come metrica di performance l'F1 delay

In generale, l'F1-delay ha mostrato una maggiore sensibilità alla qualità del segnale e alla presenza di rumore. Algoritmi come GIM[22] e Fuzzy Logic[19], ad esempio, hanno beneficiato di segnali ad alto SNR, ma hanno mostrato un degrado più evidente in presenza di entropia elevata o rumore ad alta frequenza. Al contrario, gli approcci Grid[20] e Data-Driven [21] hanno mostrato una maggiore robustezza all'outlier ratio e alla variabilità (CV), ma con performance più stabili anche in condizioni non ottimali, specialmente se valutati con il solo F1-score.

Le differenze tra F1-score e F1-delay sono particolarmente marcate nell'approccio Data-Driven, dove l'F1-score tende a mascherare errori temporali che invece vengono penalizzati dal F1-delay. Questo suggerisce che la metrica classica può sovrastimare le prestazioni in scenari dove la precisione temporale è rilevante.

Come passo successivo , per ciascun metodo di rilevamento del pasto, è stata effettuata una regressione lineare multipla per modellare l'F1-score in funzione delle principali metriche di rumore del segnale.
Le metriche sono state standardizzate tramite z-score per rendere comparabili i coefficienti stimati.
I risultati mostrano che i metodi Fuzzy Logic[19] e GRID[20] presentano i valori più elevati di R^2 adj (rispettivamente 0.351 e 0.313), indicando una maggiore sensibilità delle loro prestazioni alle metriche di rumore. Al contrario, metodi come Mosquera[26] e HR+CGM[23] evidenziano valori di R^2 adj molto più bassi (0.031 e 0.108), suggerendo una maggiore robustezza al rumore. (Figura 18)

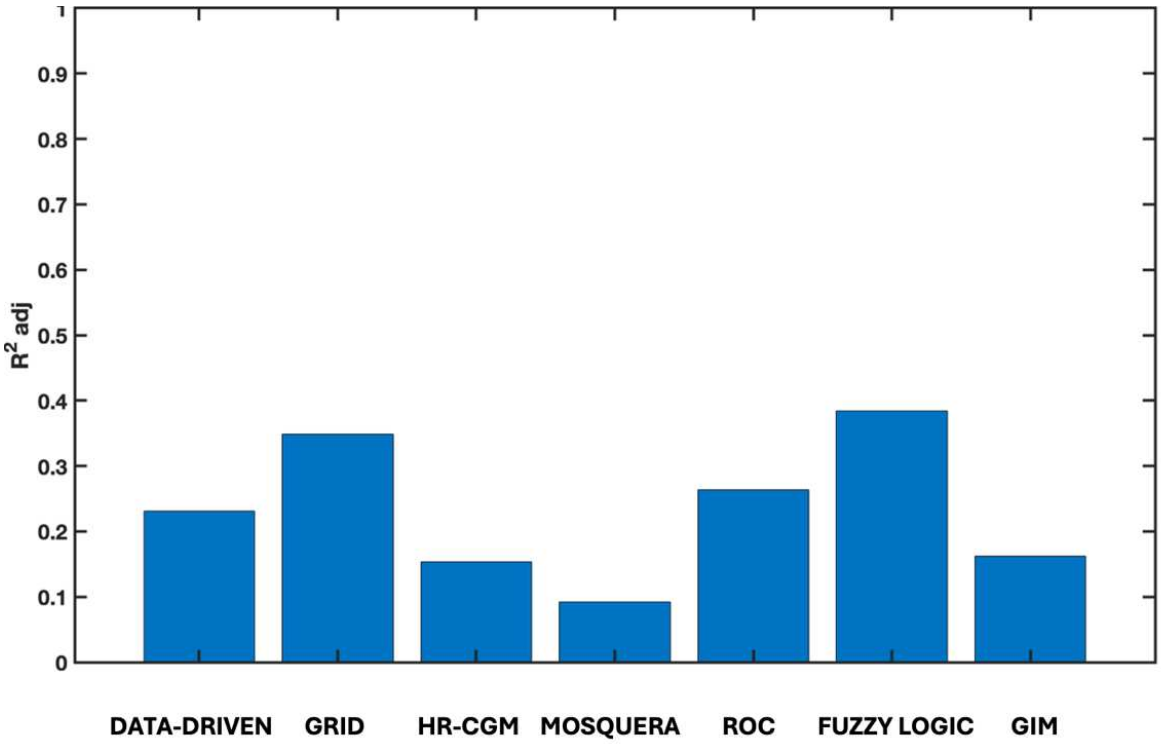

Figura 18. Qualità della regressione multipla dei diversi metodi (Adjusted R^2)

Nel complesso, questi risultati indicano che le prestazioni limitate non dipendono esclusivamente da limiti intrinseci degli algoritmi, ma anche dalle condizioni di qualità del segnale.
Come ulteriore analisi ci si è focalizzati sulla tipologia di evento (TP, FP, FN), così da valutare in che misura tali metriche possano spiegare le condizioni specifiche che favoriscono o ostacolano la corretta identificazione dei pasti.

5.3 Analisi locale del rumore nei singoli eventi di detection

Dopo aver condotto un'analisi globale tramite regressione multipla, finalizzata a valutare l'effetto complessivo delle metriche di rumore sulla variabilità delle prestazioni dei metodi, è

stata effettuata un'analisi complementare a livello locale, focalizzandosi direttamente sugli eventi rilevati.

A tal proposito, per ogni evento identificato dai vari algoritmi, è stato estratto un segmento temporale del segnale glicemico centrato sull'istante dell'evento stesso. La durata della finestra di estrazione è stata adattata in base al metodo, per riflettere le diverse dinamiche temporali di rilevazione.

All'interno di ogni segmento così ottenuto sono state calcolate diverse metriche descrittive del segnale, finalizzate a quantificare il livello di rumore, variabilità e complessità del segnale nel contesto specifico in cui si è verificato l'evento. Le metriche considerate includono in parte gli stessi identificatori utilizzati per l'analisi globale con l'aggiunta di due indicatori specifici calcolati localmente: la ΔG, che rappresenta l'ampiezza della variazione glicemica nel segmento temporale analizzato, e il ROC.

Queste due metriche aggiuntive risultano particolarmente informative per caratterizzare la dinamica del segnale in prossimità degli eventi rilevati.

Poiché non tutte le metriche hanno mostrato un contributo informativo significativo o differenze apprezzabili tra i diversi tipi di evento, nel seguito vengono riportati e discussi solo i risultati relativi agli indici che si sono rivelati maggiormente discriminanti e con un impatto più diretto sulle performance.

5.3.1 Rappresentazione grafica delle distribuzioni per categoria di evento

Per valutare oggettivamente l'impatto del rumore locale sulle prestazioni di detection, sono state analizzate le distribuzioni delle metriche calcolate per ciascuna categoria di evento attraverso rappresentazioni box plot. L'analisi del SNR (Figura 18) ha mostrato che gli algoritmi più robusti (Data- Driven[21], HR+CGM[23] e GIM[22]) associano valori più elevati di rapporto segnale-rumore ai casi veri positivi, indicando che la qualità del segnale gioca un ruolo determinante nel corretto riconoscimento degli eventi. Al contrario, altri approcci (Mosquera[26], Fuzzy Logic[19]) evidenziano distribuzioni simili tra TP, FP e FN, suggerendo una minore capacità di discriminare in base alla pulizia del segnale.

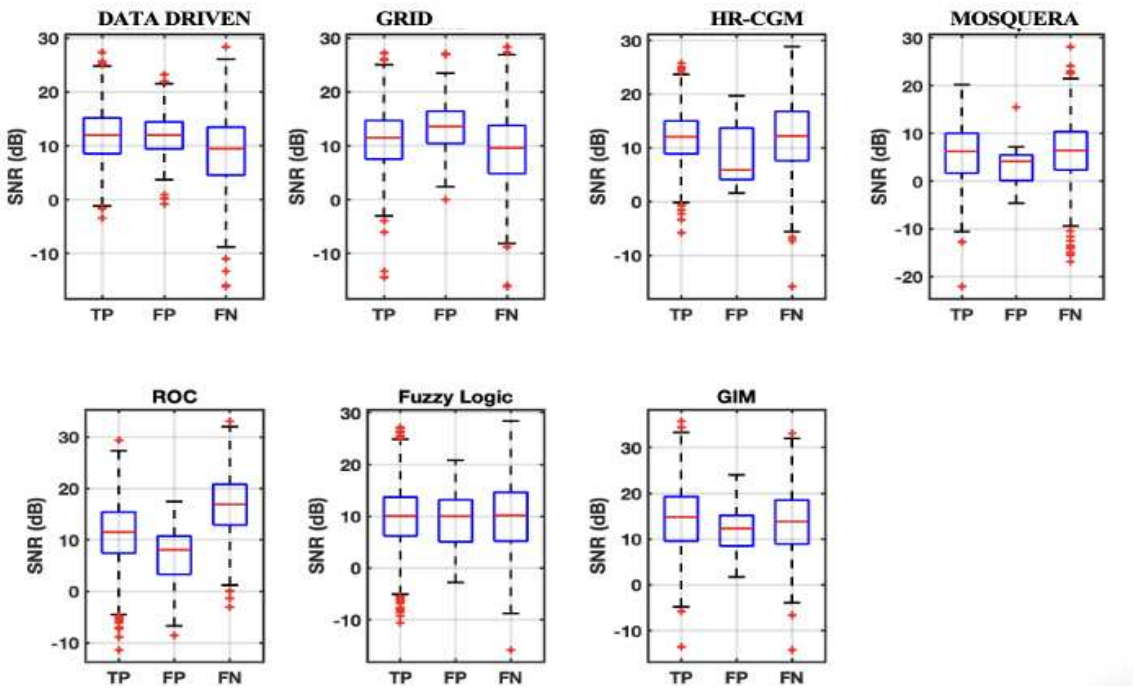

Figura 19: Confronto delle prestazioni dei diversi metodi in termini di SNR per TP, FP e FN

Anche la Sample Entropy è stata analizzata e nella figura 20 sono riportati i risultati relativi a questa.

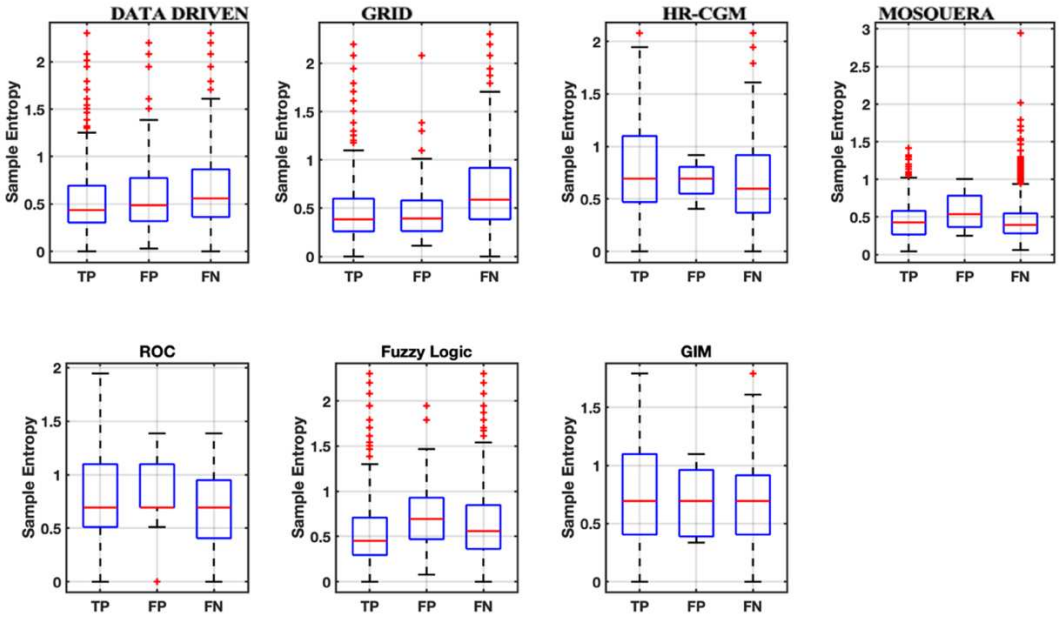

Figura 20. Confronto delle prestazioni dei diversi metodi in termini di SE per TP, FP e FN

In particolare, tale analisi ha evidenziato come alcuni metodi tendano a produrre falsi negativi in corrispondenza di segnali caratterizzati da maggiore complessità, mentre altri

presentano valori mediamente più bassi di entropia ma con un'elevata dispersione, segno di instabilità locale.

Infine, la Figura 21 relativa a ΔG ha rivelato differenze sostanziali nella rilevanza clinica degli errori: alcuni algoritmi (Data-driven[21], Mosquera[26]) associano i falsi negativi a variazioni glicemiche molto ampie, rendendo tali errori particolarmente critici; al contrario, ROC si è mostrato più conservativo, con ΔG ridotti nei TP e FP ma elevati nei FN. L'algoritmo GIM[22], invece, ha fatto registrare distribuzioni più compatte e consistenti, a conferma di una maggiore stabilità locale.

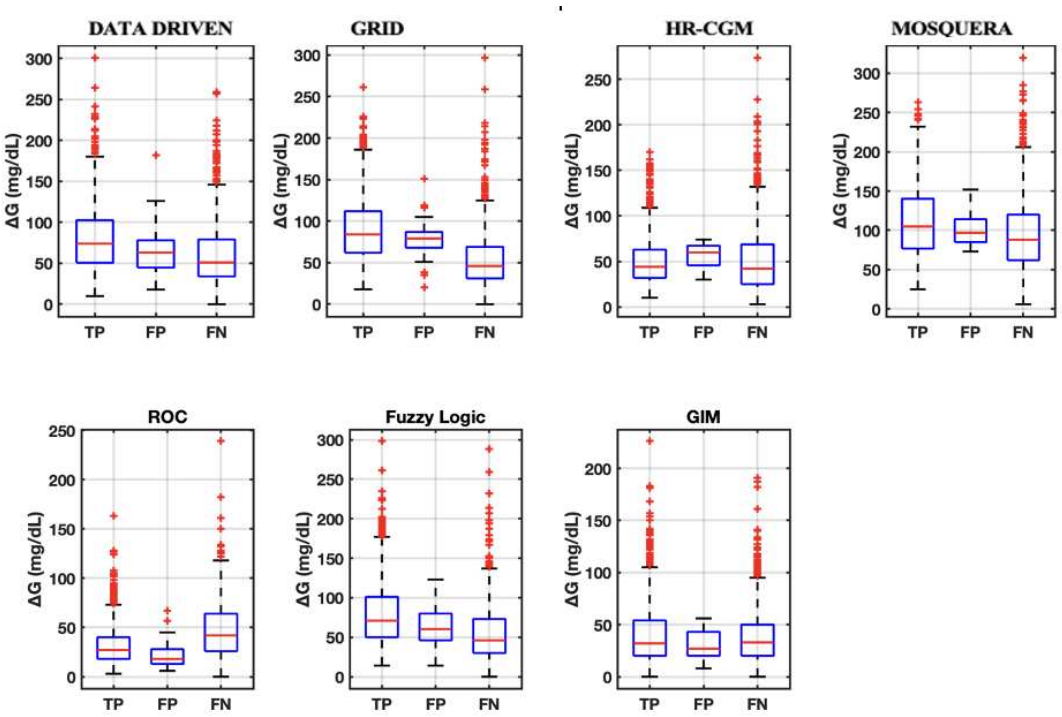

Figura 21. Confronto delle prestazioni dei diversi metodi in termini di ΔG per TP, FP e FN

L'analisi locale mette in evidenza che le prestazioni dei diversi algoritmi non dipendono esclusivamente dal numero complessivo di TP, FP e FN, ma sono strettamente influenzate dalle caratteristiche del segnale nei singoli casi. In particolare, segnali complessi o rumorosi aumentano la probabilità di errore, e in alcuni algoritmi i falsi negativi coincidono con variazioni glicemiche più significative, con implicazioni cliniche rilevanti. Tali osservazioni sottolineano l'importanza di un'analisi locale complementare a quella globale, in grado di spiegare non solo quanto un algoritmo sbaglia, ma anche perché e in quali condizioni lo fa.

5.4 Ranking delle performance algoritmiche

Le evidenze raccolte nelle sezioni precedenti hanno mostrato come le prestazioni dei metodi siano influenzate sia da fattori globali di qualità del segnale, sia da condizioni locali specifiche legate al tipo di evento. Un aspetto complementare, tuttavia, consiste nel valutare non solo perché gli algoritmi falliscono o riescono, ma anche quali tra essi si distinguono sistematicamente come più affidabili. Per questo motivo, è stata introdotta un'analisi di ranking, finalizzata a confrontare le performance relative dei diversi metodi e a individuare quelli caratterizzati da maggiore robustezza o, al contrario, da fragilità costante.

Per ogni soggetto e per ogni traccia, gli algoritmi sono stati ordinati in base alla metrica di valutazione scelta (F1-score e F1 delay). All'algoritmo con la performance migliore è stato assegnato un rank pari a 1, mentre al peggiore un rank pari a 7.

Verrà introdotta un'analisi ranking mediante l'utilizzo dei boxplot che conferma la presenza di algoritmi sistematicamente robusti algoritmi costantemente deboli e altri più sensibili al contesto i cui rank mostrano una variabilità marcata. Questa evidenza supporta la necessità di valutare i metodi non solo sulla performance media, ma anche sulla stabilità relativa nelle diverse condizioni sperimentali.

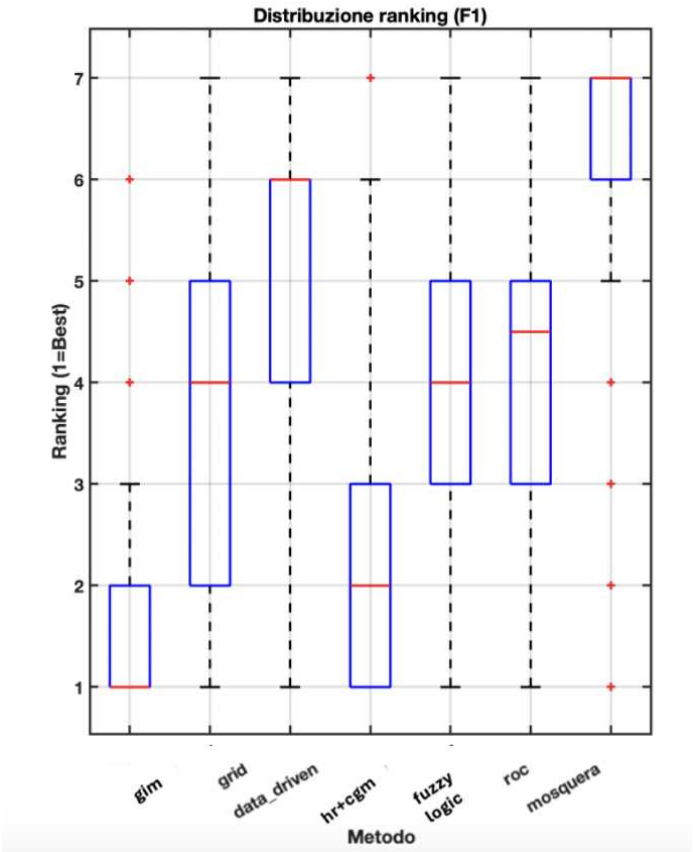

Figura 22. Ranking dei metodi in base al punteggio F1.

Dalla figura 22 si osserva chiaramente come GIM[22] ottenga costantemente rank molto bassi (quindi migliori), con mediana vicina a 1 e variabilità ridotta, a conferma della sua robustezza. Al contrario, il metodo Mosquera[26] si colloca sistematicamente tra le ultime posizioni, con scarsa variabilità, indicando performance costantemente inferiori. Gli altri algoritmi mostrano andamenti intermedi: ad esempio GRID[20] e Fuzzy Logic[19] ottengono risultati discreti ma con una variabilità maggiore, suggerendo una dipendenza dal soggetto o dalla traccia.

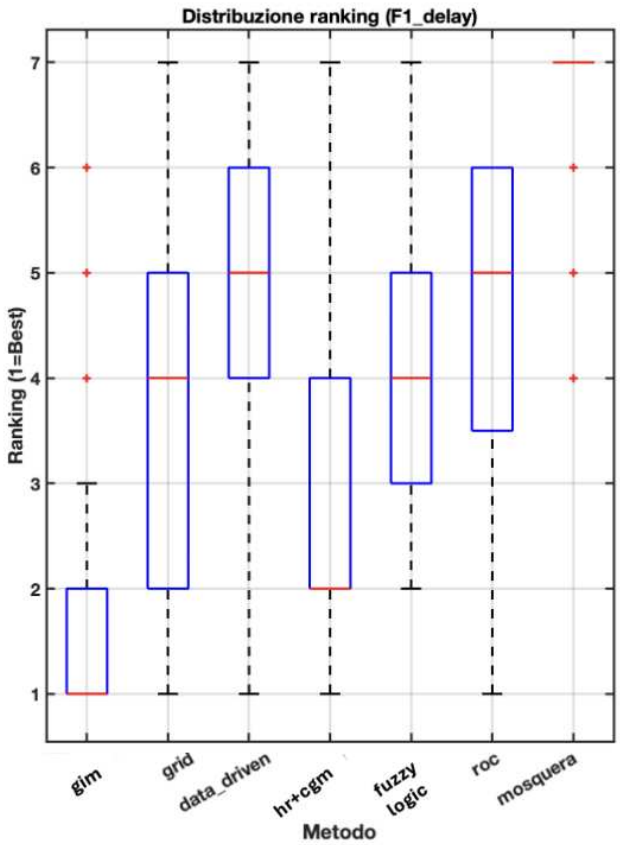

Figura 23. Ranking dei metodi in base al F1-Delay.

Anche la figura 23 mostra che il metodo GIM[22] risulta nettamente il migliore, con mediana pari a 1 e distribuzione molto concentrata, mentre il metodo Mosquera[26] mostra un comportamento opposto, posizionandosi costantemente all'ultimo posto. Gli altri algoritmi, come GRID[20], ROC e Data-driven[21], presentano distribuzioni più ampie dei rank, indicando una forte variabilità delle performance a seconda delle condizioni

Oltre all'analisi della distribuzione dei rank, è stata valutata la frequenza con cui ciascun algoritmo risulta il migliore per singolo paziente (Figura 24).

Questo approccio fornisce una misura complementare: mentre i boxplot mostrano la posizione relativa media, l'istogramma evidenzia la capacità di un metodo di adattarsi ai singoli casi clinici.

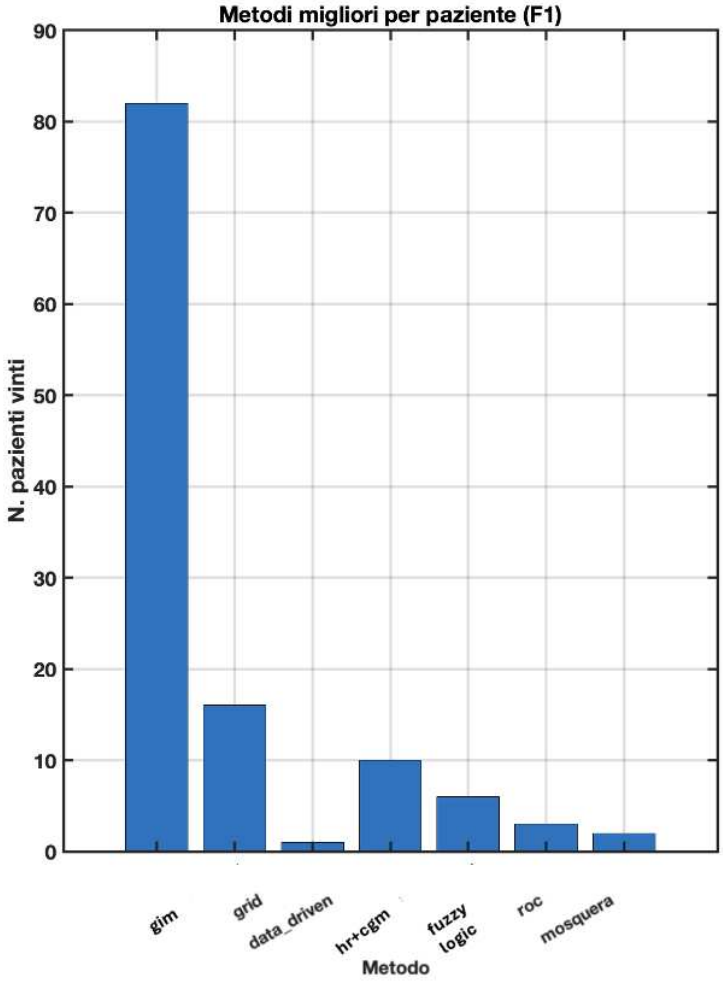

Figura 24. Numero di pazienti in cui ciascun metodo ha ottenuto la miglior performance in termini di F1-score

Dai risultati emerge chiaramente che un algoritmo in particolare domina la classifica, ottenendo il miglior risultato nella grande maggioranza dei pazienti. Gli altri metodi si distribuiscono su valori molto più bassi, con vittorie solo occasionali. Questo suggerisce che, oltre a distinguersi in termini di rank medio, l'algoritmo dominante, ovvero il metodo GIM[22], possiede anche una maggiore robustezza intra-paziente, risultando sistematicamente competitivo indipendentemente dal soggetto.

Si è riportato per semplicità solo l'istogramma del F1-score poiché per la misura del F1 delay presenta dei risultati concordi.

Prima di passare alle conclusioni, si è ritenuto opportuno inserire un ultimo paragrafo dedicato a un'analisi complessiva degli algoritmi, così da fornire una visione d'insieme più chiara e bilanciata delle prestazioni osservate.

5.5 Sintesi conclusiva degli algoritmi valutati

Per quanto riguarda gli algoritmi analizzati nello studio, emergono ulteriori limitazioni specifiche.

L'algoritmo Fuzzy Logic [19] mostra una forte dipendenza esclusiva dal segnale CGM, rendendolo vulnerabile a rumore, ritardi e valori mancanti, con il rischio di confondere variazioni glicemiche dovute ad altri fattori (ad esempio attività fisica o stress) con pasti reali. La definizione di soglie fisse per l'attivazione della meal flag limita inoltre la generalizzabilità e introduce possibili errori di rilevamento.

Analogamente, l'algoritmo Data-driven [21], pur essendo più flessibile grazie alla personalizzazione dei pattern glicemici individuali, risente della qualità dei pasti dichiarati, che condiziona la costruzione dei pattern di riferimento. Anche in questo caso, la sensibilità ai parametri di soglia e al rumore può compromettere l'affidabilità del rilevamento, mentre la necessità di addestrare classificatori specifici per ogni soggetto riduce la scalabilità del metodo in scenari clinici più ampi.

Il metodo GIM [22], pur più sofisticato grazie all'uso di un modello fisiologico integrato con la frequenza cardiaca, richiede una corretta calibrazione dei parametri individuali, talvolta non disponibile, ed è influenzato da rumore sensoriale e da fenomeni non modellizzati, come stress o ormoni contro-regolatori; inoltre, l'ipotesi di un'assimilazione uniforme dei carboidrati semplifica eccessivamente la dinamica reale dei pasti.

D'altro canto, l'algoritmo HR+CGM [23], sebbene sensibile grazie all'integrazione della frequenza cardiaca e alla personalizzazione dei parametri, presenta limiti simili: rimane vulnerabile a rumore o dati mancanti, l'effetto di stress o attività fisica può confondere il rilevamento, la calibrazione individuale riduce la scalabilità e l'assunzione uniforme dei carboidrati semplifica la dinamica reale dei pasti.

Il metodo RAP di Mosquera [26], che combina CGM, intelligenza artificiale e automazione dell'insulina, presenta limitazioni legate all'addestramento su dati simulati, che potrebbe ridurre la robustezza in contesti reali. La soglia fissa per la rilevazione dei pasti e la stima dei carboidrati non si adatta perfettamente a tutti i soggetti, mentre eventuali malfunzionamenti dei sensori o ritardi nell'integrazione dei dispositivi possono compromettere l'affidabilità. Infine, la riduzione post-prandiale della dose di insulina, pur volta a limitare il rischio

ipoglicemico, non corregge errori nella stima dei pasti, lasciando margini di discrepanza nella gestione glicemica.

Infine, l'algoritmo GRID [20], pur essendo progettato per rilevare rapidamente i pasti e attivare una risposta insulinica automatica, si basa esclusivamente sul CGM, risultando vulnerabile a rumore, ritardi o valori mancanti. Le soglie fisse per la glicemia e per il tasso di variazione possono non adattarsi perfettamente a tutti i soggetti, limitando sensibilità e specificità. L'assenza di informazioni esterne, come la frequenza cardiaca, riduce la capacità di anticipare in modo robusto gli eventi postprandiali, mentre la modulazione automatica dei boli dipende dalla correttezza della stima del pasto, lasciando margini di errore in caso di ingestione non convenzionale o ritardata.
Sebbene gli algoritmi di rilevamento dei pasti basati su HR+CGM rappresentino un passo significativo verso l'automazione nella gestione del diabete, le loro applicazioni pratiche sono ostacolate da sfide tecniche e metodologiche. La dipendenza da dati incompleti o imprecisi, la mancanza di adattamento individuale e la difficoltà nell'integrazione di informazioni contestuali limitano la loro efficacia. Pertanto, è fondamentale proseguire nella ricerca e nello sviluppo di sistemi più robusti e personalizzati, che possano affrontare queste sfide e migliorare la gestione del diabete in scenari reali.

CAPITOLO 6

6.Conclusioni e sviluppi futuri

In questo lavoro di tesi si è cercato di fornire un contributo alla tematica, ancora aperta e di grande interesse, dell'identificazione automatica dei pasti, con particolare riferimento all'ambito clinico, dove il monitoraggio accurato dell'alimentazione riveste un ruolo cruciale per la gestione di diverse condizioni di salute.

Il percorso seguito si è articolato in due fasi principali, complementari tra loro.

Nella prima fase, è stata condotta una revisione approfondita della letteratura scientifica al fine di identificare e selezionare gli algoritmi più significativi e promettenti disponibili allo stato attuale. L'analisi si è concentrata sia su approcci classici sia su metodi più innovativi, valutandone la capacità di rilevare in modo affidabile i pasti sulla base dei dati disponibili.

Successivamente, gli algoritmi selezionati sono stati implementati, opportunamente adattati al contesto specifico del dataset clinico utilizzato e sottoposti a test sperimentali. In questa fase, sono stati analizzati in dettaglio i risultati ottenuti, evidenziando sia le performance raggiunte sia le principali criticità e limitazioni emerse, in termini di accuratezza, sensibilità e robustezza rispetto alle variazioni presenti nei dati.

Nella seconda fase, l'attenzione si è concentrata sulle possibili cause alla base delle performance non sempre soddisfacenti osservate nella fase sperimentale. In particolare, è stata effettuata un'analisi approfondita del rumore presente nei dati, al fine di comprendere come esso influenzasse l'output degli algoritmi e di individuare potenziali strategie per mitigarne l'impatto. Questo approccio ha permesso non solo di identificare specifiche fonti di errore, ma anche di delineare le condizioni in cui gli algoritmi risultano più o meno affidabili, fornendo così indicazioni utili per sviluppi futuri e miglioramenti delle metodologie.

In sintesi, il lavoro ha cercato di combinare un approccio teorico, basato sull'analisi della letteratura, con un'indagine sperimentale rigorosa, finalizzata a comprendere le dinamiche e le difficoltà reali legate all'identificazione dei pasti, offrendo spunti di riflessione sia sulle metodologie esistenti sia sulle possibili direzioni per studi futuri.

6.1 Limiti dello studio

Questo studio presenta diverse limitazioni, la prima delle quali riguarda il dataset.

Il dataset T1DEXI costituisce una risorsa di grande interesse per lo studio della risposta glicemica e dei parametri fisiologici in contesti reali; tuttavia, presenta alcune limitazioni che ne devono essere considerate nell'interpretazione. La raccolta dei dati è avvenuta in un periodo relativamente breve (circa quattro settimane) e ha incluso solo soggetti con un livello minimo di attività fisica, limitando la generalizzabilità dei risultati a popolazioni con differenti abitudini motorie. Alcune informazioni critiche, come l'assunzione di carboidrati, sono state registrate manualmente dai soggetti tramite applicazione mobile e presentano frequenti stime parziali o approssimative, con la conseguente necessità di interventi manuali di correzione o imputazione. I dati continui di glicemia e frequenza cardiaca, sebbene ad alta risoluzione, contengono valori mancanti rispettivamente del 4,2% e fino al 16,85%, con implicazioni sulla qualità complessiva del segnale e sull'analisi delle dinamiche temporali. La differente risoluzione temporale tra le serie (glicemia ogni 5 minuti, frequenza cardiaca ogni 10 secondi) e la complessità di allineamento tra variabili eterogenee rappresentano ulteriori fonti di incertezza.

La fase di selezione dei pazienti, volta a garantire completezza e coerenza dei dati, ha comportato la riduzione del campione da 495 a 120 soggetti, introducendo possibili bias di selezione e limitando la variabilità interindividuale. La necessità di interventi di preprocessing, la gestione di artefatti nei dispositivi indossabili e la discrepanza tra dati auto-riportati e misurazioni oggettive costituiscono ulteriori fattori di potenziale distorsione.

Queste limitazioni sottolineano l'importanza di interpretare i risultati con cautela e di considerare la necessità di studi futuri con campioni più ampi e osservazioni più prolungate.

6.2 Sviluppi futuri

Lo studio condotto ha evidenziato il potenziale e al tempo stesso i limiti degli attuali algoritmi di rilevamento dei pasti, sottolineando la complessità intrinseca del problema e la difficoltà di ottenere soluzioni robuste in contesti clinici reali.

Alla luce di queste considerazioni, diversi sviluppi futuri possono essere individuati per progredire verso sistemi più efficaci e affidabili.

Un primo aspetto riguarda la qualità e la varietà dei dati disponibili. Il dataset T1DEXI, pur rappresentando una risorsa preziosa e unica nel suo genere, presenta limitazioni temporali, demografiche e metodologiche che riducono la generalizzabilità dei risultati.

Inoltre, la raccolta di informazioni alimentari non dovrebbe basarsi solo su auto-report, notoriamente soggetti a errori sistematici e omissioni, ma integrare strumenti più oggettivi come sistemi di image-based meal recognition, dispositivi indossabili dedicati o tecniche di tracciamento automatizzato dell'assunzione di cibo [36]. Parallelamente, l'acquisizione di ulteriori segnali fisiologici, ad esempio parametri cardiovascolari, indicatori di stress o biomarcatori ormonali, potrebbe arricchire la descrizione dei processi metabolici postprandiali, riducendo l'ambiguità degli algoritmi che si basano esclusivamente sulla glicemia [37].

Un secondo fronte di ricerca è di natura metodologica. Gli approcci attuali mostrano forti dipendenze da soglie prefissate o da modelli relativamente rigidi, che faticano a adattarsi alle variazioni interindividuali e intraindividuali. L'adozione di tecniche avanzate di machine learning e deep learning rappresenta una prospettiva promettente: modelli basati su reti neurali ricorrenti hanno già dimostrato di poter catturare pattern complessi nei segnali glicemici e nei dati multimodali [38],[39]. Tuttavia, l'uso di tali tecniche non è privo di criticità. La necessità di grandi quantità di dati per l'addestramento, il rischio di overfitting su campioni limitati e la scarsa trasparenza decisionale (il cosiddetto black-box effect) rappresentano ostacoli significativi per un impiego clinico. Una direzione promettente consiste quindi nello sviluppo di modelli ibridi, che integrino conoscenze fisiologiche con metodi di apprendimento automatico, combinando interpretabilità e capacità predittiva [40].

Un ulteriore elemento chiave riguarda la personalizzazione e l'adattività degli algoritmi. L'evidenza sperimentale mostra che non esiste un metodo universalmente valido: algoritmi che si comportano bene su un sottoinsieme di pazienti possono risultare inefficaci su altri, a causa di differenze individuali nei profili glicemici, nelle abitudini alimentari e nelle risposte fisiologiche. Per questo, lo sviluppo di sistemi adattivi, capaci di aggiornarsi dinamicamente sui dati del singolo paziente, rappresenta una priorità.

Particolare attenzione merita l'integrazione multimodale dei dati. Affidarsi esclusivamente al segnale CGM comporta inevitabili vulnerabilità: il rumore, i ritardi di misura e gli artefatti possono generare falsi positivi o falsi negativi. L'inclusione di segnali complementari, come frequenza cardiaca, attività fisica da accelerometri, qualità del sonno o livelli di stress, può migliorare la discriminazione tra pasti reali e altre variazioni glicemiche [41],[37]. Tuttavia, l'integrazione multimodale porta con sé nuove sfide, tra cui la sincronizzazione temporale dei segnali, la gestione dei dati mancanti e la necessità di garantire una bassa complessità computazionale per implementazioni in tempo reale.

Infine, non va trascurato l'aspetto della trasferibilità clinica. Lo sviluppo di algoritmi efficaci in ambiente sperimentale rappresenta solo un primo passo: per una reale applicazione nei sistemi di closed-loop insulin delivery, è necessario che i metodi siano affidabili, interpretabili e conformi agli standard di sicurezza dei dispositivi medici. Questo implica non solo un miglioramento delle metriche di performance ,ma anche la capacità di fornire spiegazioni comprensibili al clinico e al paziente, riducendo il rischio di interventi automatizzati inappropriati. Inoltre, la validazione deve necessariamente passare attraverso trial clinici prospettici e controllati, che permettano di valutare l'impatto reale degli algoritmi sulla gestione del diabete e sugli outcome clinici [42],[43].

In sintesi, i futuri sviluppi dovranno affrontare la sfida di conciliare tre dimensioni:

1. Robustezza dei dati, tramite raccolte più accurate, complete e multimodali.
2. Avanzamento metodologico, con algoritmi adattivi, interpretabili e basati su modelli ibridi.
3. Trasferibilità clinica, che richiede validazioni rigorose e integrazione sicura nei sistemi di automazione terapeutica.

Solo la convergenza di questi tre aspetti potrà portare a sistemi di rilevamento dei pasti realmente utili, personalizzati e clinicamente applicabili, contribuendo a un passo avanti significativo verso una gestione sempre più automatizzata e sicura del diabete.

Bibliografia

1. Baynes HW. *J Diabetes Metab.*2015;6:541. doi:10.4172/2155-6156.1000541.
2. Wilcox G. Insulin and insulin resistance. *Clin Biochem Rev.* 2005;26(2):19–39. PMID:16278749.
3. Tassone F. Nuove proposte in letteratura sulla classificazione del diabete mellito. *JAMD.* 2019;21(4):142–147.
4. Du M, Li S, Zhang Y, Li J, Zhou X. Advances in the pathogenesis and treatment strategies for type 1 diabetes mellitus. *Int Immunopharmacol.* 2025;127:110560. doi:10.1016/j.intimp.2024.110560.
5. Brutsaert EF. Ipoglicemia. In: *Manuale MSD, edizione professionisti.* 2023. Available from: https://www.msdmanuals.com/it-it/professionale/endocrinologia-e-malattie-metaboliche/diabete-mellito-e-disturbi-del-metabolismo-glucidico/ipoglicemia
6. Heinemann L, Freckmann G. Continuous glucose monitoring: A perspective on its past, present, and future applications for diabetes management. *Diabetes Metab Syndr Obes.* 2015;8:239–248. doi:10.2147/DMSO.S84765.
7. Vettoretti M, Facchinetti A. Combining continuous glucose monitoring and insulin pumps to automatically tune the basal insulin infusion in diabetes therapy: A review. *Biomed Eng Online.* 2019;18:65. doi:10.1186/s12938-019-0658-x.
8. Istituto Superiore di Sanità (Epicentro). Report to Parliament 2021: prevalenza diabete in Italia e nel mondo.
9. Vecchio I, Tornali C, Bragazzi NL, Martini M. The discovery of insulin: An important milestone in the history of medicine. *Front Endocrinol (Lausanne).* 2018;9:613.
10.Bender C, Vestergaard P, Cichosz SL. The history, evolution and future of continuous glucose monitoring (CGM). *Diabetology.* 2025;6:17. doi:10.3390/diabetology6030017.
11.Kvist PH, Jensen HE. Recent advances in continuous glucose monitoring: Biocompatibility of glucose sensors for implantation in subcutis. *J Diabetes Sci Technol.* 2007;1(5):746–752. doi:10.1177/193229680700100520. PMID:19885143; PMCID:PMC2769647.
12. Movi-Elab. La terapia con pompe per insulina [Internet]. Available from: https://movi-elab.it/utilizzatori/il-diabete/la-terapia-con-pompe-per-insulina/
13. American Diabetes Association. Standards of medical care in diabetes—2023. *Diabetes Care.* 2023;46(Suppl. 1):S1–S154.
14. Xie J, Wang X, Zhao M, et al. Automated meal detection and estimation algorithms for Type 1 diabetes: A systematic review. *BMC Med Inform Decis Mak.* 2022;22:100. doi:10.1186/s12911-022-01805-3.
15. Patel N, Bansi A, Hillson R. The impact of meal timing on glycemic control in Type 1 diabetes: A review. *Diabetes Therapy.* 2021;12:617–629. doi:10.1007/s13300-021-01010.
16. Moser EG, Phillip M, Battelino T. Management of hypoglycemia unawareness in Type 1 diabetes. *Diabetes Care.*2021;44(5):e91–e92. doi:10.2337/dc20-3017.
17. Eyal D, Wayne B. Detection of a meal using continuous glucose monitoring: Implications for an artificial β-cell. *Diabetes Care.* 2008;31(2):295–300. doi:10.2337/dc07-1293.
18. Brummer J, Glasbrenner C. Continuous glucose monitoring for automatic real-time assessment of eating events and nutrition: A scoping review. *Front Nutr.* 2024;10:1308348. doi:10.3389/fnut.2023.1308348.

19. Samadi S, et al. Meal detection and carbohydrate estimation using continuous glucose sensor data. *IEEE J Biomed Health Inform.* 2017;21(3):619–627.
20. Harvey RA, et al. Design of the Glucose Rate Increase Detector: A meal detection module for the health monitoring system. *J Diabetes Sci Technol.* 2014;8(4).
21. de Carvalho F, et al. Data driven meal events detection using blood glucose response patterns. *BMC Med Inform Decis Mak.* 2023;23:1.
22. Zheng M, et al. Automated meal detection from continuous glucose monitor data through simulation and explanation. *J Am Med Inform Assoc.* 2019.
23. Cossu L, Cappon G, Wortmann F, et al. A personalized and automated real-time meal detection algorithm based on continuous glucose monitoring and heart rate data for individuals with post-bariatric hypoglycemia. 2024.
24. Hotta S, Mori T, et al. Eating moment recognition using heart rate responses. In: *Proc. ACM Int Joint Conf Pervasive Ubiquitous Comput.* 2017;69–72. doi:10.1145/3123024.3123093.
25. Colmegna P, Bisio A, McFadden R, et al. Evaluation of a web-based simulation tool for self-management support in Type 1 diabetes: A pilot study. *IEEE J Biomed Health Inform.* 2023;27(1):515–525. doi:10.1109/JBHI.2022.3209090.
26. Mosquera Lopez C, et al. Enabling fully automated insulin delivery through meal detection and size estimation using artificial intelligence. *npj Digit Med.* 2023;6:1.
27. Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980. 2014.
28. Riddell MC, Li Z; T1DEXI Study Group. Examining the acute glycemic effects of different types of structured exercise sessions in Type 1 diabetes in a real-world setting. *Diabetes Care.* 2023;46(4):704–713. doi:10.2337/dc22-1721.
29. Zhang Y, Wei S, Long Y, Liu C. Performance analysis of multiscale entropy for the assessment of ECG signal quality. *J Electr Comput Eng.* 2015;Art. ID 563915. doi:10.1155/2015/563915.
30. Costa M, Goldberger AL, Peng CK. Multiscale entropy analysis of complex physiologic time series. *Phys Rev Lett.*2002;89(6):068102. doi:10.1103/PhysRevLett.89.068102.
31. Benichou T, et al. Heart rate variability in Type 2 diabetes mellitus: A systematic review and meta-analysis. *PLoS ONE.* 2018;13(4):e0195166. doi:10.1371/journal.pone.0195166.
32. Nickel A, et al. The relationship between heart rate variability and glucose clearance in healthy men and women. *PLoS ONE.* 2024;19(6):e0303346. doi:10.1371/journal.pone.0303346.
33. Rodbard D, et al. Interpretation of continuous glucose monitoring data: glycemic variability and quality of glycemic control. *Diabetes Technol Ther.* 2009;11(S1):S55–S67. doi:10.1089/dia.2008.0132.
34. Rovetta A. Raiders of the lost correlation: A guide on using Pearson and Spearman coefficients to detect hidden correlations in medical sciences. *Cureus.* 2020;12(11):e11794. doi:10.7759/cureus.11794.
35. Matabuena M, Sartini J. Multilevel functional data analysis modeling of human glucose response to meal intake. arXiv preprint arXiv:2405.14690v1. 2024. PMID:38827463; PMCID:PMC11142320.
36. Leal Y, et al. Meal detection algorithms for artificial pancreas systems: challenges and opportunities. *Sensors.* 2022.

37. Sun Q, et al. Multimodal physiological signals for glucose metabolism analysis. *IEEE J Biomed Health Inform.*2021.
38. Xie J, et al. Deep learning for automated meal detection using CGM. *Diabetes Technol Ther.* 2020.
39. Zhao Y, et al. Machine learning approaches in diabetes management. *JMIR Med Inform.* 2021.
40. Contreras I, Vehi J. Artificial intelligence for diabetes management and decision support: A review. *J Diabetes Sci Technol.* 2021.
41. Li K, et al. Multisensor data fusion for personalized diabetes management. *Sensors.* 2020.
42. Weisman A, et al. Closed-loop insulin delivery systems: current status and future perspectives. *Diabetes Care.*2020.
43. Boughton CK, Hovorka R. Advances in artificial pancreas systems. *Nat Rev Endocrinol.* 2021.

Ringraziamenti

Ringrazio il Prof. Andrea Facchinetti per la guida, i suggerimenti e il supporto continuo che hanno reso possibile il completamento di questo lavoro di ricerca.
Un ringraziamento speciale va anche al Dott. Luca Cossu per il contributo significativo e il tempo dedicatomi, che hanno arricchito in maniera considerevole il mio lavoro.
La loro competenza, disponibilità e impegno hanno reso questa esperienza di ricerca estremamente stimolante e formativa.