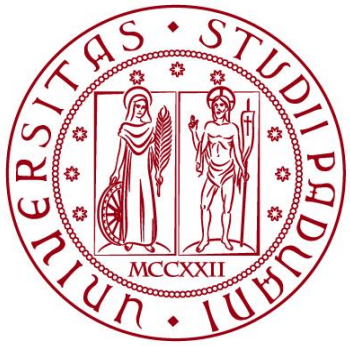

UNIVERSITY OF PADOVA

DEPARTMENT OF INFORMATION ENGINEERING

Master's Degree in Bioengineering

HYPOGLYCEMIA DETECTION IN PATIENTS WITH TYPE 1 DIABETES USING EEG SIGNALS

Supervisor:
Prof. Giovanni SPARACINO

Co-supervisor:
Fabio SCARPA
Maria RUBEGA

Master candidate:
Debora TEODORI

April 8, 2019

Abstract

The main risk for patients with type 1 diabetes (T1D) is to fall into hypoglycemia, an event which leads to both short and long-terms consequences and can be life-threatening, especially when it occurs at night without subject awareness. Avoiding hypoglycemia is important since cognitive tests conducted on T1D adults have assessed that hypoglycemia results in altered cerebral activity, most likely due to the complete dependence of the brain on glucose supply. Indeed, the first organ influenced by this fall of glucose in the blood is the brain. Many studies investigated the mirroring of cognitive dysfunction due to hypoglycemia in the spectral power of the electroencephalogram (EEG) signal. In particular, the increase of the power in low frequency EEG bands is a well-known effect during hypoglycemia that seems more pronounced in the EEG recording in the posterior areas of the brain. Recent studies have shown these variations also in the evaluation of signal complexity. Indeed, in neural signals, complexity is high for systems that contain specialized and organized elements, while it is lower in uniform or random systems and during hypoglycemia, the EEG signal results more regular and more uniform, less complex respect to normal EEG. Pilot studies about the real-time processing of the EEG signal to detect hypoglycemia have indicated that it might be possible to alert the patients by means of EEG analysis. The main purpose of this work is to broaden the quantitative and qualitative detection of hypoglycemia from the altered EEG signal in T1D patients and consequently to identify potential features sensitive to hypoglycemia. In particular, the analyses are extended to different domains, i.e., fractal domain and entropy domain, to deepen the knowledge about the changes due to hypoglycemia in the EEG signal. This dissertation is structured in 6 chapters, briefly presented below. Chapter 1 will start with a brief overview about the impact of T1D in patients' daily life. Furthermore, the results on the effects of hypoglycemia on brain activity will be described by reporting the main findings in literature. Chapter 2 will describe the Database, composed by the EEG signal acquired by 34 T1D patients and the respective cognitive test for each of them. Patients underwent an 8-h hyperinsulinemic - hypogycemic clamp experiment with the general goal of having a more complete overview on the effects of hypoglycemia in the brain. Subsequently Chapter 3 will consider the effects of hypoglycemia on EEG complexity. The entropy indices applied to this dataset and the comparison with fractal features will be enounced. The aim will be to reduce the computational cost compared to traditional algorithms. Chapter 4 will examine the possibility of selecting a limited number of EEG channels that are more sensitive to hypoglycaemia by studying the EEG signal complexity in all channels. In Chapter 5, the classification method of Neural Networks will be applied to select the minimum number of EEG channels able to identify with adequate accuracy the presence of the disease. Eventually, Chapter 6 will close the dissertation by interpreting the ensemble of the results from both the clinical and the engineering point of view and presenting the possible future developments of this work.

Sommario

Il principale rischio per i pazienti affetti da diabete di tipo 1 (T1D) è cadere in ipoglicemia, un abbassamento della concentrazione di glucosio nel sangue che può provocare conseguenze sia a breve sia a lungo termine, mettendo a rischio la vita del paziente soprattutto quando si verifica durante la notte senza averne coscienza. Evitare questa condizione glicemica è essenziale poiché test cognitivi hanno dimostrato un'alterata condizione cerebrale durante l'ipoglicemia, molto probabilmente causata dalla completa dipendenza del cervello dalla fornitura di glucosio. Si trovano vari studi in letteratura che provano come la riduzione della concentrazione di glucosio nel sangue, oltre a cambiamenti nelle funzioni cognitive, si rifletta in cambiamenti della potenza spettrale del segnale elettroencefalografico (EEG). In particolare, la crescita della potenza delle basse frequenze nel segnale EEG, soprattutto nelle aree posteriori del cervello, è un effetto ben noto in letteratura durante l'ipoglicemia. Studi recenti hanno mostrato queste variazioni anche nel livello di regolarità del segnale. Infatti, durante gli episodi l'ipoglicemici, il segnale EEG risulta più regolare e uniforme rispetto al segnale EEG acquisito in condizioni normali. Studi pilota hanno dimostrato che potrebbe essere possibile utilizzare anche il segnale EEG per segnalare l'entrata in ipoglicemia. Il vantaggio potrebbe essere dato dal fatto che i cambiamenti nel segnale EEG avvengano prima che la condizione ipoglicemica sia grave a tal punto da causare conseguenti disfunzioni cerebrali. Il principale obiettivo di questa tesi è quello di ampliare la rilevazione dell'ipoglicemia dal segnale EEG nei pazienti T1D e di identificare potenziali indici sensibili all'ipoglicemia. Le analisi sono state estese a diversi domini, il dominio frattale e dominio entropico, per approfondire la conoscenza dei cambiamenti dovuti all'ipoglicemia nel segnale EEG. Finora, gli studi in letteratura hanno principalmente valutato queste variazioni a livello di singolo canale EEG nel dominio della frequenza e limitate informazioni sono disponibili sull'analisi della regolarità del segnale EEG. La tesi è stata strutturata in 6 capitoli brevemente descritti in seguito per affrontare i temi appena anticipati. Il Capitolo 1 presenta una panoramica sulle conseguenze e sull'impatto nella vita di tutti i giorni del paziente T1D. Inoltre, si descrivono brevemente i risultati sugli effetti dell'ipoglicemia sull'attività cerebrale riportati in letteratura. Nel Capitolo 2 sono descritti il database dei soggetti affetti da diabete T1D e i rispettivi test cognitivi. Il segnale EEG e la concentrazione di glucosio nel sangue sono stati raccolti in parallelo per circa 8 ore in 31 pazienti ospedalizzati affetti da T1D indotti in ipoglicemia attraverso un clamp ipoglicemico iperinsulinemico. Successivamente gli indici di entropia applicati a questo set di dati e il confronto con gli indicatori frattali sono enunciati nel Capitolo 3. L'obiettivo è ridurre il costo computazionale delle operazioni da eseguire in contrasto con gli algoritmi tradizionali. Il Capitolo 4 esaminerà la possibilità di selezionare un numero limitato di canali EEG più sensibili all'ipoglicemia studiando la complessità del segnale EEG in tutti i canali. Nel Capitolo 5 verrà applicato il metodo di classificazione delle Reti Neurali per selezionare il numero minimo di canali EEG in grado di identificare con adeguata accuratezza la presenza di patologie. Alla fine, il Capitolo 6 chiuderà la tesi interpretando l'insieme dei risultati dal punto di vista sia clinico che ingegneristico e presentando i possibili sviluppi futuri di questo lavoro.

Contents

Abstract iii

Sommario v

1 Hypoglycemic induced EEG alterations in type 1 diabetes 1

1.1 Type 1 diabetes and hypoglycemia 1

1.2 Hypoglycemia and the brain 4

1.2.1 The role of glucose in brain function 4

1.2.2 Cognition dysfunction during hypoglycemia 5

1.2.3 Hypoglycemia-related EEG changes 5

1.3 Aim of the thesis 10

2 Database 11

2.1 Description of the data 11

2.2 Experimental protocol 11

2.3 Behavioural tests 13

2.3.1 Trail Making Test 13

2.3.2 Stroop Test 13

3 Analysis of complexity of EEG signals during hypoglycemia 15

3.1 State of Art 15

3.2 Description of the chosen methodologies 15

3.2.1 Empirical Permutation Entropy 16

3.2.2 Empirical Conditional Entropy 20

3.2.3 Features extraction based on fractal dimension 22

3.3 Implementation parameters and details 23

3.4 Results 25

3.5 Conclusions 29

4 Relationship between EEG quantitative features and awareness of hypoglycemia 31

4.1 The starting point 31

4.2 Selection of EEG channels sensitive to hypoglycemia: single-channel evaluation 32

4.3 Multi-channel evaluation 34

4.4 Conclusions 39

5 Classification of hypoglycemia through Neural Networks 41

5.1 The use of Neural Networks to classify hypoglycemia 41

5.2 Dataset creation 42

5.3 Design the network 45

5.4 Results 46
5.5 Conclusions 48
6 Conclusions 49
6.1 Main achievements and discussion 49
6.2 Future challenges 50
A Fundamentals of Neural Networks 53
A.1 The concept of artificial neuron 53
A.2 Networks of artificial neuron 53
A.2.1 Mathematical modelling of the artificial neuron 54
A.2.2 Activation function 56
A.3 Network architecture: Feedforward neural networks 58
A.4 Supervised training of the network 59
A.4.1 Back-propagation: Levenberg Marquardt algorithm 60
Bibliography 63
Acknowledgment 69

List of figures

Figure 1.1: Diabetes Atlas Eighth edition.....**Errore. Il segnalibro non è definito.**
Figure 1.2: Hierarchy of response to hypoglycemia.....**Errore. Il segnalibro non è definito.**
Figure 1.3: Schematic diagram of HAAF in diabetes.....**Errore. Il segnalibro non è definito.**
Figure 1.4: The role of glucose in brain function .....**Errore. Il segnalibro non è definito.**
Figure 1.5: 10/20 International System. .....**Errore. Il segnalibro non è definito.**
Figure 1.6: EEG signals in euglycemic and hypoglycemic state. 7
Figure 1.7: Hyposafe device.....8

Figure 2.1: EEG recordings of subject during the hyperinsulinemic-hypoglycemic clamp. 13

Figure 3.1: An example of ordinal patterns 17
Figure 3.2: Empirical Permutation Entropy in two successive overlapped sliding windows 19
Figure 3.3: Algorithm of fast computing empirical permutation entropy in sliding windows 19
Figure 3.4: Algorithm of computing empirical conditional entropy of ordinal patterns in sliding windows. ..22
Figure 3.5: The ePE values in a random subject.....25
Figure 3.6: The ePE in all subjects 25
Figure 3.7: The eCE values in a random subject 26
Figure 3.8: The eCE in all subjects.....26
Figure 3.9: Fractal dimension in a representative subject 27
Figure 3.10: Fractal dimension in alla subjects 27
Figure 3.11: Results of the ROC analysis.....28

Figure 4.1: Awareness vs unawareness in cognitive tests 31
Figure 4.2: Empirical Permutation Entropy indices in a random subjects.....33
Figure 4.3: Awareness influence on P3 channel.....34
Figure 4.4: EEG in euglycemia and hypoglycemia 35
Figure 4.5: Topographies of the highest frequency density in EU and HYPO.....36
Figure 4.6: The boxplot for values of complexity indices in all EEG channels by averaging all subjects. 38
Figure 4.7: Awareness influence on all EEG channels.....39

Figure 5.1: The brain as a Neural Network. .....**Errore. Il segnalibro non è definito.**
Figure 5.2: From Biological neuron to Artificial Neuron.....**Errore. Il segnalibro non è definito.**
Figure 5.3: Neural Network.....54
Figure 5.4: Non-linear model of the neuron 55
Figure 5.5: Reformulation of the neuron model.. 56
Figure 5.6: Activation function 57
Figure 5.7: Single layer and multi layer Neural Network.....59
Figure 5.8: A representative FFNN with back-propagation learning algorithm. 60
Figure 5.9: Higuchi matrix in hypoglycemia of Channel 13 43
Figure 5.10: Representative vector necessary for creating the dataset. 43
Figure 5.11: Graphic representation of FFNN implemented from MATLAB.....45
Figure 5.12: Exploratory analysis of the features.....48

List of Tables

Table 1: A progressive reduction of intellectual functions of the subject involved in patient with T1D 5
Table 2: Database. 12
Table 3: The ordinal patterns. 17
Table 4: Efficiency of computing the complexity's indices. 29
Table 5: Final representation of the dataset constructed for this classification. 44
Table 6: The obtained results with 28 features and the FFNN described above. 47
Table 7: The obtained results with 21 features and the FFNN described above. 48

List of abbreviations

ANN	**A**rtificial **N**eural **N**etwork
ApEn	**Ap**proximate **En**tropy
AUC	**A**rea **U**nder the **C**urve
BBB	**B**lood-**B**rain **B**arrier
BG	**B**lood **G**lucose
C	**C**entral lobe
CGM	**C**ontinuous **G**lucose **M**onitoring
eCE	**E**mpirical **C**onditional **E**ntropy
EEG	**E**lectro**E**ncephalo**G**ram
ePE	**E**mpirical **P**ermutation **E**ntropy
ERS/ERD	**E**vents of **S**ynchronization / **D**e**S**ynchronization
EU	**EU**glycemia
F	**F**rontal lobe
FFNN	**F**eed**F**orward **N**eural **N**etwork
FN	**F**alse **N**egative
FP	**F**alse **P**ositive
GD	**G**radient **D**escend
GDM	**G**radient **D**escent with **M**omentum
HAAF	**H**ypoglycemia-**A**ssociated **A**utomatic **F**ailure
HU	**H**ypoglycemia **U**nawareness
HYPO	**HYPO**glycemia
ICA	**I**ndependent **C**omponent **A**nalysis
ID	**ID**entifier
IDF	**I**nternational **D**iabetes **F**ederation
LOOCV	**L**eave **O**ne **O**ut **C**ross **V**alidation
MSE	**M**ulti **S**cale **En**tropy
MSE	**M**ean **S**quare **E**rror
NN	**N**eural **N**etwork
O	**O**ccipital lobe
OP	**O**rdinal **P**attern
P	**P**arietal
PCA	**P**rincipal **C**omponent **A**nalysis
R	**R**egression
RBFNN	**N**eural **N**etworks **R**adial-**B**asis **F**unction
rePE	**R**obust **E**mpirical **P**ermutation **E**ntropy
RNN	**R**ecurrent **N**eural **N**etwork
ROC	**R**eceiver **O**perating **C**haracteristic
RP	**R**esilient **B**ackpropagation
SampEn	**Samp**le **En**tropy
T	**T**emporal lobe
T1D	**T**ype **1** **D**iabetes
TN	**T**rue **N**egative
TP	**T**rue **P**ositive
TMT	**T**rail **M**aking **T**est
YSI	**Y**ellow **S**pring **I**nc

Chapter 1

Hypoglycemic induced EEG alterations in type 1 diabetes

1.1 Type 1 diabetes and hypoglycemia

Type 1 diabetes is a chronic disease in which pancreatic cells are no longer able or nearly able to produce insulin. It is configured as an autoimmune disease [1], characterized by the distraction of β-cells located in the pancreas [2], which involves association with insulin-deficiency. Insulin is the key hormone that allows the metabolism of nutrients and in particular sugar, in human cells: insulin promotes the absorption of nutrients (in particular glucose) in the cells. In particular, the glucose has problems in entering into the muscles and adipose tissue, consequently its accumulation in the blood causes, in the diagnosis, is defined as hyperglycemia. The main preponderant factor that contributes to T1D is genetics, but also environment and immunology. Indeed, T1D is one of the most frequent chronic disease in childhood and adolescence and it rarely develops in adults, it accounts for almost 85% of all pediatric diabetes cases. Despite active research, T1D has no cure yet and its only possible therapy is the exogenous administration of what the organism cannot produce autonomously: insulin. It is for this reason that the patient suffering from T1D must do a lifelong insulin therapy: this treatment is focused on the management of blood sugar levels with exogenous insulin, diet and lifestyle both aimed to prevent complications. Nowadays, a non-negligible share of people in the world suffers from T1D. Diabetes has reached epidemic proportions throughout the whole world and will continue to grow and remains the greatest global health challenge the world has ever known [3]: 415 million people have diabetes (1 in 11 adults), and the number of diagnosed people is predicted to rise beyond 642 million in less than 25 years. Fig. 1 describes the incidence rate of diabetes in the world, no world area is immune to this deadly disease. In absolute numbers, the greatest increase of individuals with diabetes will be seen in the Southeast Asia region (61.9 million), followed closely by the Western Pacific region (61.6 million). For instance, about 300,000 people live with T1D in Italy, about 10% of the population and the Italian rate of incidence of T1D in children is 12 out of 100,000. The Italian island of Sardinia is an anomaly to the rest of the country: the second-highest rate of incidence of T1D in the world, e.g., rate of incidence in children is 45 out of 100,000. As previously described, the challenge faced daily by people suffering from diabetes is to keep blood sugar, i.e. the amount of sugar in the blood, under control and the role played by glucose and insulin is a key factor in this metabolic disease. In the diabetic patient, the tendency is to have a too high blood sugar (hyperglycemia), but the opposite condition, having too low blood sugar (hypoglycemia), is either likewise if not even more dangerous, that is. There are many factors that influence the counter-regulatory responses of glucose and insulin and their metabolic balance in order to avoid hypoglycemia, which is considered to be one of the most dangerous complications in diabetes treatment. Its definition is internationally approved: hypoglycemia is an event during which typical neurogenic symptoms, such as shakiness, anxiety, nervousness, palpitations, sweating, dry mouth, pallor, and pupil dilation, are accompanied by a measure plasma glucose concentration lower than 70 mg/dl (3.9 mmol l^{-1}). The huge range of physiological, hormonal, symptomatic and cognitive

responses triggered by the fall in blood glucose are visible in Fig. 1.2, [1]. Higher levels of epinephrine cause all the neurogenic signs, i.e., the primary fast-acting hormone in the defence against acute hypoglycemia, besides glucagon.

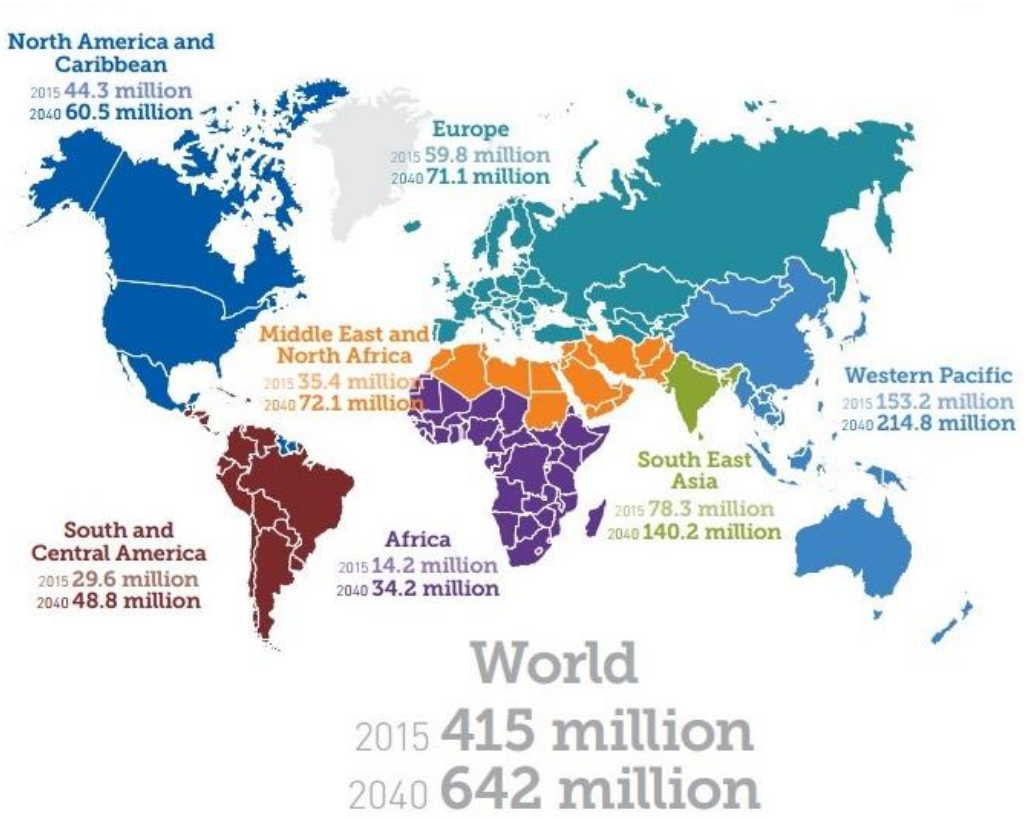

Figure 1.1: International Diabetes Federation (IDF), Diabetes Atlas Eighth edition 2017, estimated number of people with diabetes worldwide and per region in 2015 and 2040 (20-79 years) [3].

However, epinephrine responses to hypoglycemia could become impaired in T1D [2], leading to hypoglycemia unawareness. Hypoglycemia unawareness (HU) is defined as the failure to sense a significant fall of blood glucose below normal levels [4]. In patients with T1D, recurrent hypoglycemia has been shown to reduce the glucose level that precipitates the counter-regulatory response necessary to restore euglycemia during a subsequent episode of hypoglycemia. Aberrant glucose counter-regulation (as a result of a failure in the reduction of insulin production and an increase in glucagon release), and HU (as the result of an attenuated increase in sympathoadrenal activity) are the components of hypoglycemia-associated autonomic failure (HAAF) in diabetic patients. HAAF is most often caused by recent/recurrent iatrogenic hypoglycemia, and indeed it is maintained by recurrent hypoglycemia Fig 1.3, [4]. Nevertheless, hypoglycemia awareness can be reversed by avoiding new hypoglycemia episodes, thanks to a strict glycemic control [5]. In a healthy subject, the blood glucose concentration should remain in the range [70 - 180] mg / dl ([3.9 - 10] mmol l-1), which is defined as “normal” blood glucose, also called euglycemia [6]. Although there have been many improvements in the last years, current methods of blood glucose detection and insulin replacement e.g. through continuous glucose monitoring (CGM) devices and smart pumps [7], are not still perfect and the hypoglycemia risk is high in T1D patients. Iatrogenic hypoglycemia could even causes recurrent physical and psychological morbidity and in some cases mortality, it impairs defences against subsequent hypoglycemia, and precludes the maintenance of euglycemia. Hypoglycemia can have both short-term dangerous effects, such as cognitive dysfunction,

behavioural abnormalities, confusional state, coma, seizures, transient ischaemic attack and transient hemiplegia, as well as long-term ones, such as cerebrovascular events, hemiparesis, focal neurological deficits, ataxia, choreoathetosis, cognitive impairment with behavioural and psychosocial problems and rarely, epilepsy and vegetative state. Nocturnal hypoglycemia can have these drastic effects. Indeed, during night, the usual hypoglycemia warning symptoms and the sympatho-adrenal response are blunted [8] and moreover, the symptoms' intensity decreases when lying down. It must be mentioned that hypoglycemia has also social and economic effects in both industrialized and developing countries. Indeed, hospitalization of diabetic people after a hypoglycemic episode leads to a significant increase of treatment costs due to an increased length of hospital stay and a higher risk of mortality [9]. In conclusion, it is important, both medically and economically, to avoid hypoglycemia in T1D.

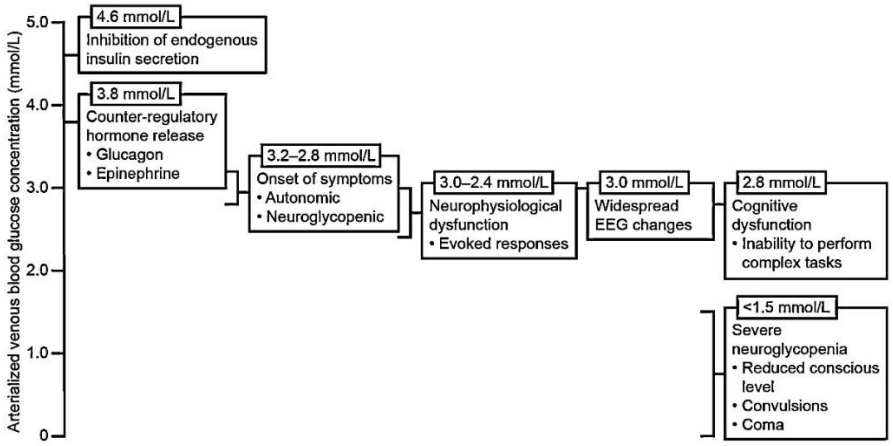

Figure 1.2: Physiological, symptomatic and cognitive response to hypoglycemia. reproduced from [4].

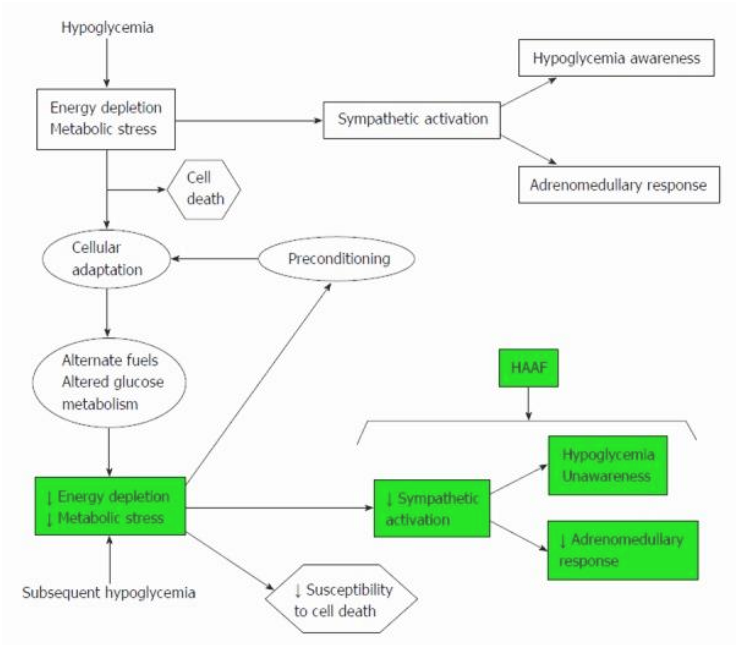

Figure 1.3: Recurrent hypoglycemia leads to cellular adaptation and hypoglycemia-associated autonomic failure. HAAF: hypoglycemia-associated autonomic failure. Reproduced from [6].

1.2 Hypoglycemia and the brain

1.2.1 The role of glucose in brain function

Glucose is the main organism's energy substrate from which it is possible to obtain ready-to-use energy for the proper functioning of the main organs. The glucose-dependent organs like brain, medulla of the adrenal gland, marrow and retina constantly require glucose to carry forward their energy metabolism. Despite the brain accounts only for 2% of body weight, it is the principal consumer of glucose, i.e., 20% of glucose-derived energy [10]. In Fig. 1.4, the metabolism of the glucose in the human brain is highlighted. Specialized centres in the brain, including neurons in the hypothalamus, sense central and peripheral glucose levels and regulate glucose metabolism (panel A in Fig. 1.4). Glucose supply to the brain is regulated by neurovascular coupling and may be modulated by metabolism-dependent and -independent mechanisms (panel B in Fig. 1.4). The glucose enters the brain trough the blood by crossing the blood-brain barrier (BBB) through glucose transporter 1, while glucose and other metabolites (e.g. lactate, Lac) are rapidly distributed through a highly coupled metabolic network of brain cells (panel C in Fig. 1.4). Glucose provides the energy for neurotransmission (panel D in Fig. 1.4), and several glucose metabolizing enzymes control cellular survival (panel E in Fig. 1.4) [11]. Indeed, the glucose is required for both neurotransmitters' synthesis and several glucose-metabolizing enzymes control cellular survival, but the largest amount of glucose is consumed for neural computation and signal processing, since the brain increases its glucose consumption upon activation [12]. In particular, in adult brains neurons have the highest energy demand and the BBB that separates the circulating blood from the brain extracellular fluid is highly selective and permeable for glucose. Thus, glucose is a fundamental and not-replaceable energy source for the brain.

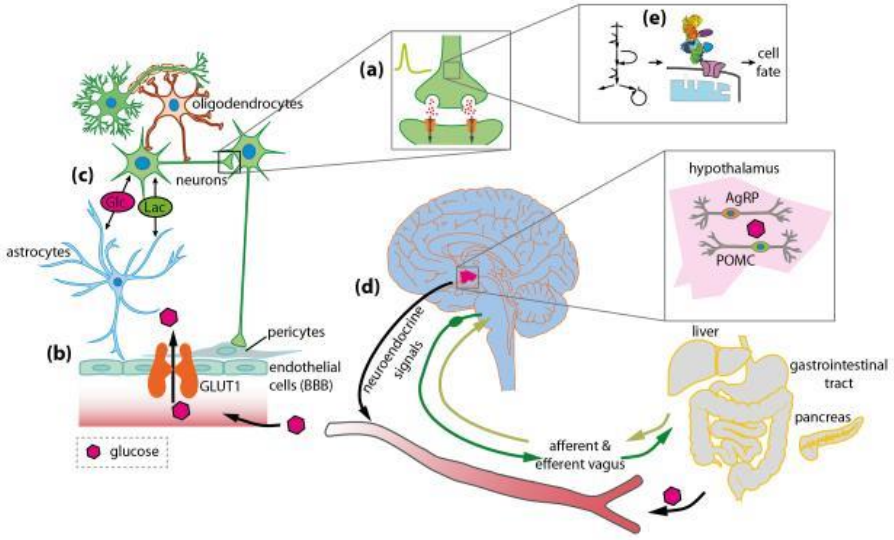

Figure 1.4: (A) Specialized centres in the brain sense central and peripheral glucose levels and regulate glucose metabolism. (B) Glucose supply to the brain is regulated by neurovascular coupling. (C) Glucose and other metabolites are rapidly distributed through a highly coupled metabolic network of brain cells. (D) glucose provides the energy for neurotransmission, and (E) several glucose-metabolizing enzymes control cellular survival. reproduced from [13].

1.2.2 Cognition dysfunction during hypoglycemia

Several studies have shown how the brain, in a state of prolonged fasting, is able to derive energy from ketones and other substances [13]. Despite this, the human brain depends on the continuous supply of glucose and therefore it is vulnerable for any glucose deprivation. It is for this reason, that the lack of glucose can lead to both-short term and long-term cognitive dysfunctions. In particular, a recurrent moderate hypoglycemia state may already lead to irreversible synaptic dysfunction, even in absence of necrosis [14]. The subjects affected by hypoglycemia are also likely to have a less favourable development in cognitive function during follow-up if compared with control subjects and especially in memory (assessed by Rey Auditory Verbal Learning Test), problem solving and verbal function (assessed by Wechsler Adult Intelligence Scale III), and psychomotor efficiency (assessed by finger tapping). In [15], where cognitive tests like the California Computerized Assessment Package, the Stroop test and the Trail Making Test B (TMT B) were performed to assess selective attention, visual attention, cognitive alterations, reaction time and speed of information processing in T1D adults during hypoglycemia compared to euglycemia, the cognitive function deteriorated as expected during hypoglycemia. The above-mentioned studies, using cognitive tests, assessed that hypoglycemia results in an altered cerebral activity, likely due to the entirely dependence of the brain from glucose supply [16]. The most important consequence is the progressive decline of the IQ, with a persistent reduction of cognitive levels and, more generally, a progressive decay of the higher brain functions [table 1]. Therefore, the brain is the first organ to be affected by hypoglycemia and the decrease of cerebral functioning consequent to this condition may be investigated through suitable analysis of the electroencephalogram (EEG) signal.

T1D
Attention
Memory
Learning
Psychomotor efficiency
Problem Solving
Vocabulary
General intelligence
Vasoconstrictive abilities
Visual perception
Mental flexibility
Executive functions

Table 1: A progressive reduction of previously acquired intellectual functions and alterations in the behaviour of the subject involved in patient with T1D, [16].

1.2.3 Hypoglycemia-related EEG changes

The electroencephalogram (EEG) is a record of the electrical activity of the brain from the scalp. The recorded waveforms reflect the cortical electrical activity by pairs of conductive electrodes. Fig, 1.5

shows a schematic representation of the EEG system placement: the location of electrodes on the scalp follows the 10–20 International System, which provides the positioning of the electrodes according to ideal lines (anterior-posterior sagittal line, medial and lateral, coronal frontal, central and parietal lines) drawn from fixed points of reference, i.e. inion, nasion and preauricular points. The distance between one electrode and another is always 10% or 20% of the total length of the line, hence the name of the system. In this system, each electrode site is labelled with a letter and a number referring to the area of the brain e.g. letters F, T, C, P and O respectively stand for frontal area (blue area in Fig. 1.5), temporal (green area), "central"[1], parietal area (yellow area), and occipital lobes (red area). Additionally, even numbers denote the right side of the hemisphere while the odd numbers denote the left one. Overall, therefore, the 10-20 system includes 19 sites of which 8 on the left side, 8 on the right side and 3 on the center, and on each of them is located an electrode. The difference in voltage between the electrodes is measured, and the changes over time reflect neuronal activity. Since the acquired signal is weak (30-100µV), it must be amplified. The EEG signal is usually described in terms of rhythmic activity and transients. The rhythmic activity is divided into frequency bands indeed, the EEG signal can be seen as the sum of simpler parts characterized by different frequency components each with a specific biological meaning [17].

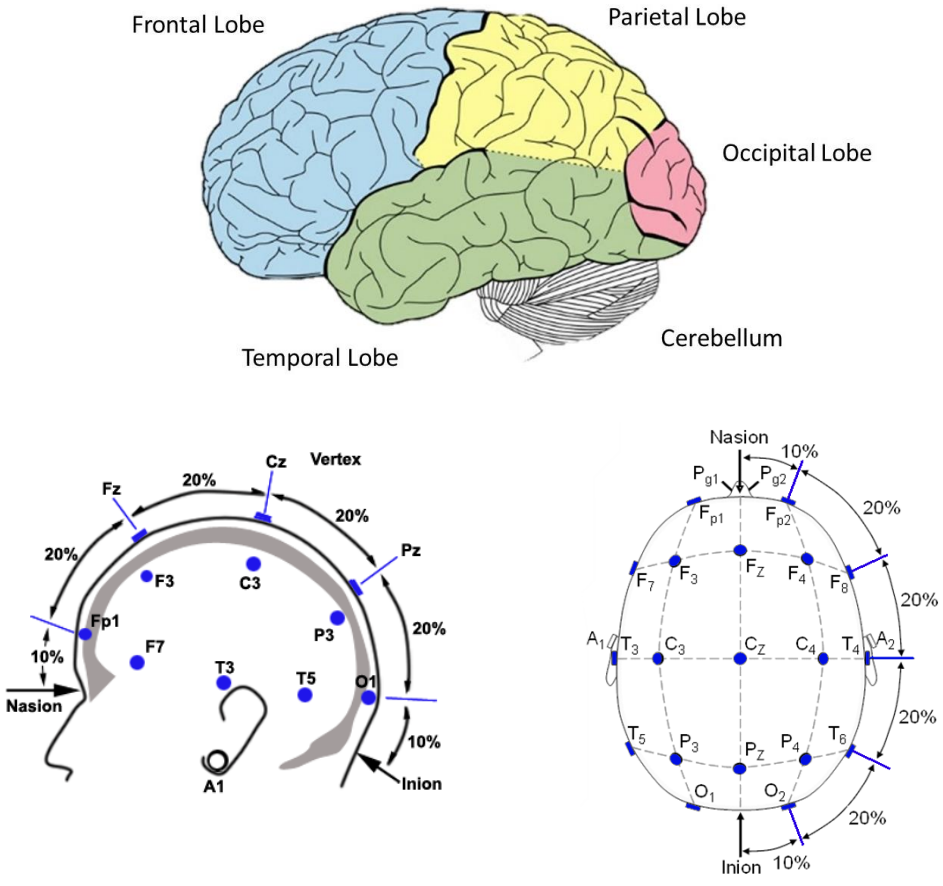

Figure 1.5: Brain lobes (upper panel) and standardize nomenclature and location of electrodes in the 10/20 International System (lower panel).

[1] Notation indicating the central part of the brain given by the division into lobes.

Frequency bands are usually extracted using spectral methods and waveforms are subdivided into bandwidths known as: delta waves, from 1 to 4 Hz, are characterized by high amplitude waves; theta waves, from 4 to 8 Hz, appear during meditative and relaxed state in adults; alpha waves, from 8 to 13 Hz, appear with closing eyes and decrease during mental efforts or opening eyes; and beta waves, from 13 to 30 Hz, which represent thinking, active concentration and movements. Being a non-invasive recording, the EEG represents a basic examination in neurology. It assumes its greatest diagnostic value in the diagnosis of epilepsy and inflammatory diseases such as meningitis and encephalitis or in metabolic encephalopathies. An association between EEG changes and hypoglycemia has also been proved in many papers in the literature. Indeed, the EEG signal reflects the metabolic state of the brain during the critical state of hypoglycemia. In [18], authors provided a review of the current literature regarding changes in the EEG during episodes of low blood glucose. Hypoglycemia causes an extensive range of clinical effects and highly affects the brain functions that are associated with measurable EEG changes. The first studies aiming to see if cognitive dysfunction was mirrored by altered EEG activity during hypoglycemia date back to 40's: in [19], authors observed abnormal EEG activity defined as slow potentials, high amplitudes fast potentials and irregular or disorganized patterns, during hypoglycemia. In [20], authors also emphasized the importance of blood sugar concentration on cerebral function, observing EEG dysrhythmia, i.e., distorted EEG waves with higher amplitude, during hypoglycemia. In Fig. 1.6 it is possible to see an example of EEG dysrhythmia in a T1D patient. Hypoglycemia-effect on EEG was also quantified by the increase of low-frequency slow-wave activity. This increase in the power of low frequency bands during hypoglycemia in T1D patients is a well-established effect in literature.

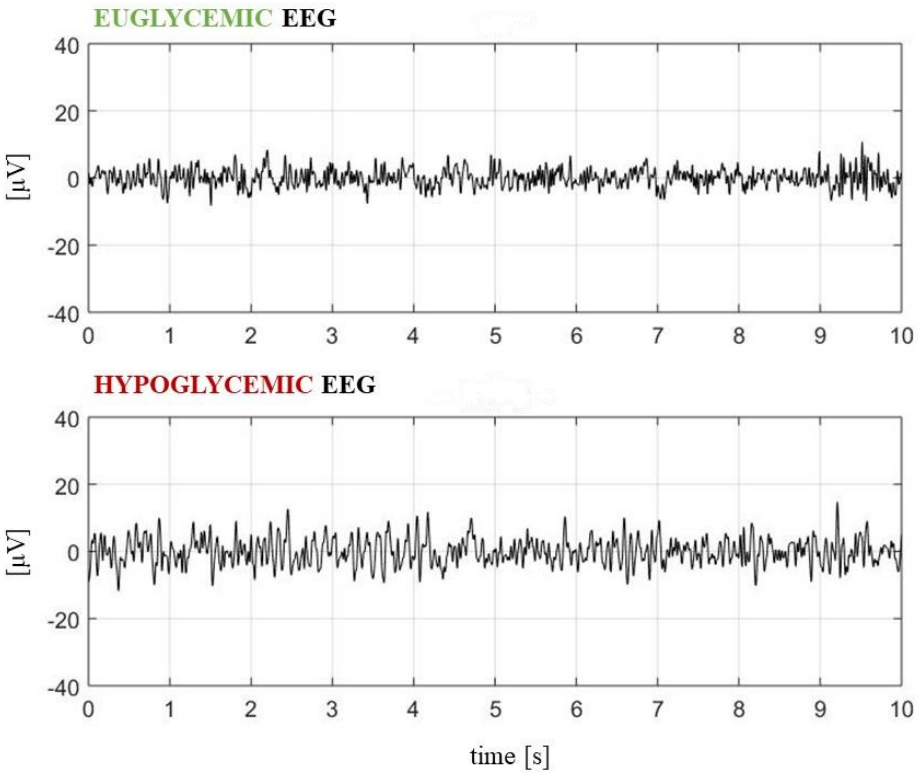

Figure 1.6: Comparison between 10 s of EEG signals in euglycemic and hypoglycemic state.

Indeed, in [21], authors observed an increase in theta activity set in abruptly at a mean blood glucose concentration close to 2 mmol l^{-1}. These changes were symmetrically and diffusely spreaded over the cortex though slightly more pronounced in the recordings from the temporal and parieto-occipital areas. In [22], the EEG signal during hypoglycemia was analysed in nineteen T1D patients. The general slowing in the lower EEG frequency bands induced by hypoglycemia influenced also the results of the EEG functional connectivity. Information partial directed coherence, used to evaluate the mutual relation between the EEG channels at scalp level, tended to decrease when passing from euglycemia to hypoglycemia in both theta and alpha bands in all combinations of the considered EEG channels. After the hypoglycemic period, EEG activity recovered to the baseline. Also non-linear features were applied in the EEG analysis during hypoglycemia proving a generalized slowing of cerebral activity [22]. Nonlinear indices originally developed for the representation of chaotic systems, were then implemented to provide information complementary to the ones obtained by linear methodologies with a computational improvement.

1.2.3.1 Prototype device EEG-based for real-time hypoglycemia detection

Hyposafe proposed a research prototype of portable device for the detection of hypoglycemia events, able to measure the EEG signal by subcutaneous electrodes and then analyse and classify it in real-time [18]. This device, showed in Fig. 1.7, was designed to detect and alert T1D patients of hypoglycemia before the situation would become critical by continuous measurement of the electrical activity within the brain by EEG [23].

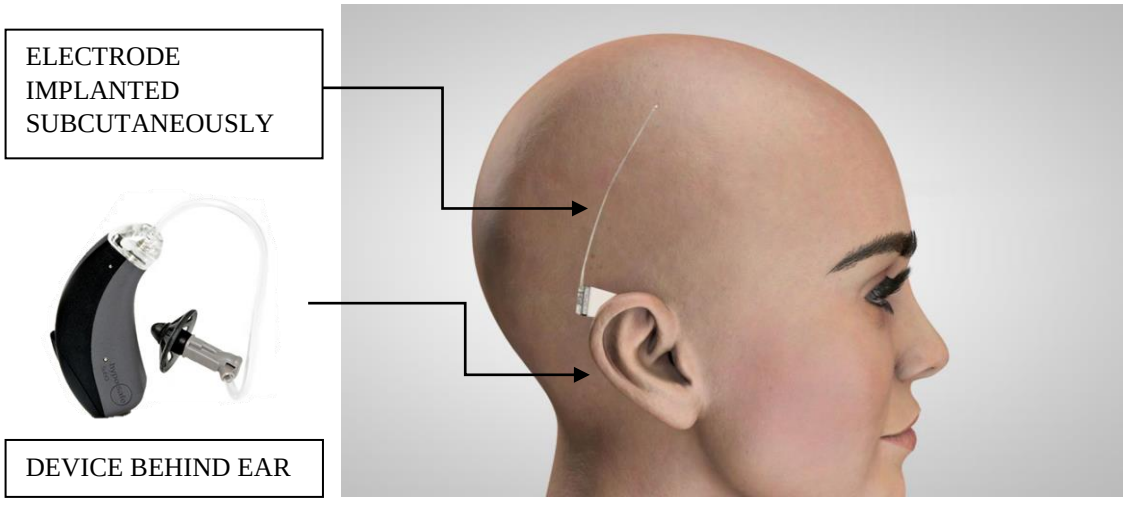

Figure 1.7: ideal device for hypoglycemia detection from Hyposafe. images available online.

This technology, patented in [24], is made of a 5 cm long, 0.5 mm diameter electrode implanted only few millimetres below the skin, and a non-implanted device which stores and processes the EEG signal as well as contains a power source (18 h), a sound generator and a light indicator to inform the user in case of critical events. Thus, Hyposafe idea is based on using the brain directly as a biosensor, measuring changes in EEG signals and monitoring signs of hypoglycemia. The Hyposafe team proposed to detect specific changes in the EEG time-series, by processing the signals through a

general mathematical algorithm based on power estimations and noise indicators, with the aim of detect hypoglycaemia before the development of cognitive failure. The EEG was recorded within a pilot-experiment. In particular, the hypoglycaemia was induced in 15 T1D patients not allowed to sleep Frequency, amplitude and power of the EEG signal recorded were extracted as features from the pilot experiment to realize the final model. The final model consists of a curve of integrated events: a window in time domain weights past hypoglycaemia epochs by integrating events weighted by a kernel. The threshold value to detect hypoglycaemia is a steep increase above 20 as shown in Fig 1.8.

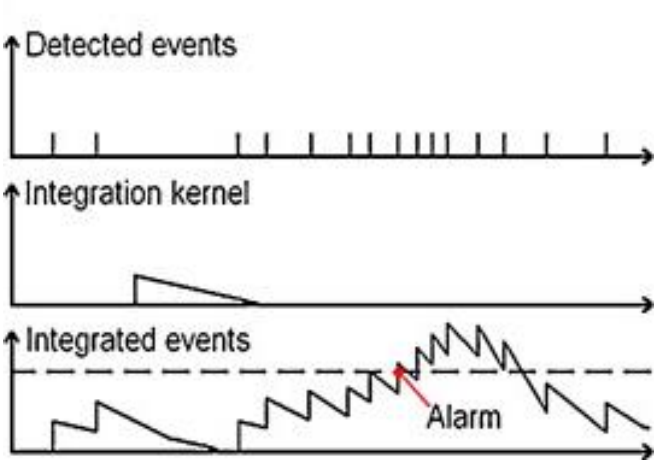

Figure 1.8: It is possible see the detected events in the first chart. These events were integrated applying the integration kernel (in the second chart) and obtaining the integration curve (in the third chart). You can notice that the integration curve overcomes the threshold, thus it's generated a warning to the patient.

In other studies, the same authors of the previous experiment also studied if severe hypoglycaemia is preceded by changes in both ECG (electrocardiography) and EEG features by extraction of delta, theta and alpha bands power and QT[2] interval features. They noted that EEG theta energy increases when plasma glucose detects hypoglycemia. Results, obtained by these works, support the use of the brain as a biosensor for hypoglycaemia detection. Analysis of EEG changes, as a device to detect and predict hypoglycaemia events, was proposed in some other papers.The device can remain implanted for more than 10 years without creating inconvenience to the user. Moreover, there is no risk of infection or allergic reactions, since there are not physical elements passing through the skin and the use of bandage/plaster is not required [23]. The clinical studies of continuous EEG recording and real-time data processing during insulin-induced hypoglycemia in T1D patients that have been conducted so far indicate that it might be possible to predict incidents of severe hypoglycemia before the patients become severely cognitive impaired both during daytime and sleep, [23] [10] [25]. Despite several contributions, a systematic evaluation of the effects of hypoglycemia in the brain to identify solid indicators sensitive to hypoglycemia has not been reported yet.

[2] In cardiology, the QT interval is a measure of the time between the start of the Q wave and the end of the T wave in the heart's electrical cycle.

1.3 Aim of the thesis

Starting from the findings in literature, the aim of this work is to bring important improvements in the EEG signals analysis, investigating new features and algorithms of hypoglycemia classification and detection in subjects with T1D. The use of the EEG signal was suggested to detect the entering in hypoglycemia as given that symptom perception as an estimation of the blood glucose level can be unreliable, as alternative or, more likely, complement to blood glucose sensors. This work can help to make EEG portable device for real-time hypoglycemia detection (e.g. Hyposafe) usable in clinical practice and in monitoring patients during day life activity. So far, many previous studies detected change in EEG signals, mostly in frequency domain, from euglycemic state respect to hypoglycemic, obtaining important information but not enough for real-time application able to correctly detect hypoglycemia onset. Recently, the EEG analysis carried out through fractal analysis opened the way to new valuation and insights. This type of analysis is further investigated in this thesis and applied to a database consisting of EEG recordings collected in 34 T1D patients undertaken to a hyperinsulinemic– hypoglycemic clamp.

Chapter 2

Database

2.1 Description of the data

Thirty-four patients[3] with T1D underwent an hyperinsulinemic– hypoglycemic clamp with a target of 2.0 – 2.5 mmol l^{-1}. Nineteen patients out of them (58 % males; mean age, 55 ± 2.5 years; diabetes duration, 28.5 ± 2.5 years) were recruited from the diabetes outpatient clinics in the Nordsjællands Hospital Hillerød and the Steno Diabetes Center, Denmark. These data were anonymised and saved with an ID consisting of a number. Fifteen patients (67 % males; mean age, 37.5 ± 3.5 years; diabetes duration, 15 ± 3 years) were recruited from the diabetes outpatient clinics in the Hvidovre Hospital, Denmark, these data were anonymised and saved with an ID consisting of the letter "D" followed by a number. All the details are reported in Table 2.1. The database was kindly provided by Dr Claus B. Juhl, Dr Anne-Sophie Sejling, and Dr Christian Frandsen. Hyposafe A/S, Lynge, Denmark. The inclusion criteria were the following: subjects affected by T1D for more than 3 years, with an age of more than 18 years old, and being either hypoglycemia aware or unaware. Exclusion criteria included pregnancy; breastfeeding; any brain disorder; use of antiepileptic drugs, β-blocking drugs, or neuroleptic drugs; use of benzodiazepines within the last month; cardiovascular disease; alcohol or drug abuse. The hypoglycemia awareness status was classified by the Pedersen-Bjergaard method [26], the Gold score [27], and the Clarke method [28]. Of the 34 participants, 13 patients were classified as hypoglycemia unaware and 15 patients as hypoglycemia aware according to all three methods, while 3 patients as hypoglycemia unaware according to two methods (labelled with reduced awareness in Tab. 2.1). There are 3 patients who are not qualify as either hypoglycemia aware or unaware. were excluded from the analyses. Their EEG signal was however considered in the complexity analysis study. Finally, database is composed of 34 patients.

2.2 Experimental protocol

For the clamp procedure, insulin (Actrapid; Novo Nordisk, Ballerup, Denmark) mixed with heparinized plasma from the patient and isotonic saline was administered intravenously at a rate of 1 mU insulin/kg/min for the progressive induction of hypoglycemia. A variable 20% glucose infusion was administered to keep the plasma glucose at the desired levels. During the 8 h experiment, blood glucose (BG) samples were monitored (Yellow Springs Inc (YSI), USA) every 5 min to identify at least 1-h of hypoglycemia (HYPO) and 1-h of euglycemia (EU), and 19 EEG channels were simultaneously recorded (Cadwell, Easy II, Kennewick, USA) using standard cap electrodes placed on the scalp according to the international 10/20 system, while the patients were sitting on a chair with eyes open. EEG was analogically low-pass filtered offline, to avoid aliasing, then digitally acquired and finally, down-sampled at 200 Hz. The dynamic range of the EEG was ± 4620 μV with an amplitude resolution of 0.14 μV. The internal noise level in the analogue data acquisition system

[3] The study was approved by the local ethical committee and the patients gave written informed consent.

was estimated to be 1.3 μV RMS. In Fig. 2.1, P3-A1A2 EEG channel recording and simultaneous BG data from a representative subject are visible. 1-h of EU and 1-h of HYPO and 1-h of recovery after induced-hypoglycemia were identify by visual inspection of the BG time-series before and after it has crossed the hypoglycemic threshold of 70 mg/dL. Furthermore, two different cognitive tests were performed in EU (pre and post induced-hypoglycemia) and HYPO: Trail 5 B of the Comprehensive Trail Making Test (TMT B) and the Stroop tests.

#	Patient ID	Awareness statuts	Gender	Age(y)	Diabetes duration (y)
1)	1	Unaware	F	68	33
2)	2	Unaware	M	65	38
3)	3	Aware	M	45	15
4)	6	Aware	M	56	28
5)	D01	–	–	–	–
6)	D06	Aware	F	42	19
7)	D09	Aware	M	51	14
8)	D12	Aware	M	38	8
9)	D13	–	–	–	–
10)	D15	Aware	M	29	15
11)	D18	Reduced awareness	M	40	20
12)	D20	Aware	F	36	22
13)	D21	Aware	M	22	7
14)	D26	Aware	M	66	37
15)	D28	Aware	M	28	16
16)	D33	Aware	F	26	20
17)	D34	Reduced awareness	M	29	15
18)	D37	–	–	–	–
19)	D40	Reduced awareness	F	43	3
20)	8	Unaware	M	60	15
21)	9	Aware	F	66	33
22)	10	Unaware	M	62	37
23)	11	Unaware	M	55	35
24)	12	Aware	F	45	11
25)	13	Unaware	F	69	45
26)	15	Aware	M	42	16
27)	16	Unaware	M	59	20
28)	19	Unaware	F	52	34
29)	20	Unaware	M	67	23
30)	21	Aware	M	36	18
31)	23	Unaware	F	53	33
32)	24	Unaware	F	45	27
33)	26	Unaware	M	66	54
34)	27	Unaware	F	40	27

Table 2: Database.

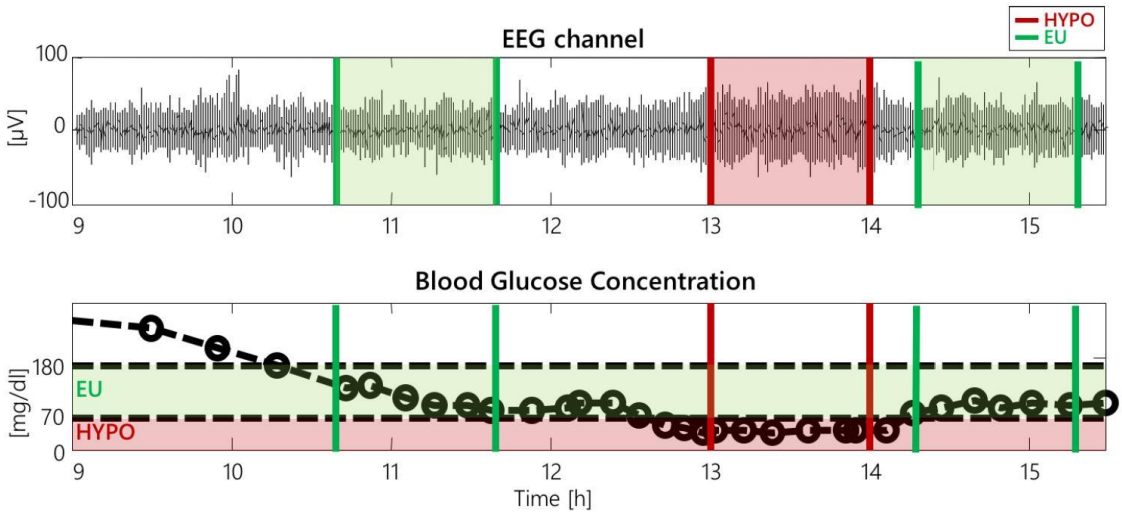

Figure 2.1: Data from a representative subject during the hyperinsulinemic-hypoglycemic clamp. on the upper panel, P3-A1A2 electroencephalography (EEG) recordings and on the lower panel, simultaneous blood glucose concentration (open circles denote reference YSI samples and the dashed line denotes their spline interpolation). vertical green and red lines refer to 1-h eu- and hypoglycemic and recovery intervals. Reproduced from [23]

2.3 Behavioural tests

2.3.1 Trail Making Test

The Trail Making Test is a neuropsychological test of motor speed and visual attention composed of two parts. In part A, i.e. TMT A, the subject's task is to quickly draw lines on a page connecting 25 consecutive numbers without lifting the pen from the paper. If the patient makes an error, it is pointed out immediately and the patient is allowed to correct it. The performance is assessed by the time taken to complete the trial correctly. In part B, i.e. TMT B, it is tested how fast the participant can connect numbers and letters in alternating increasing sequence (i.e., 1-A-2-B, etc.). TMT B is 56 cm longer and has more visually interfering stimuli than TMT A. Part B is more difficult than Part A not only because it is a more difficult cognitive task, but also because of its increased demands in motor speed and visual search [29]. Thus, TMT B evaluates visual attention, motor speed, and cognitive alternation.

2.3.2 Stroop Test

The Stroop tests evaluate processing speed and selective attention. Specifically, in the Danish version of the Stroop Colour and Word Test by Golden [30], the participant must name as many items as possible in 45 s at three different conditions: colour names printed in black (Stroop word), blocks printed in different colours (Stroop colour), and colour names printed in non-matching colours (Stroop combi) [31]. The Stroop test proved how the brain slows down if it must deal with a conflict of information. For instance, to complete the last task is more challenging than the first ones because of the brain processing delay caused by the interfering information of colour names printed in non-

matching colour. Moreover, the more the capacity to direct attention is fatigued, the harder this last stage becomes.

Chapter 3

Analysis of complexity of EEG signals during hypoglycemia

3.1 State of Art

In T1D patients, a considerable association between hypoglycemic patient state and EEG signals change has been identified. In [18], authors provide a review of current literature regarding EEG changes during episodes of low glycemia. As previously introduces, an increase of signal power and general symmetrically slowdown frequency appeared in all cerebral regions, independently from age, diabetes duration, and awareness or not of episodes antecedent to hypoglycemia. So far, many studies have been carried out in frequency domain, many of which are presented in section 1.2.3, trying to describe the difference between euglycemic EEG signals with compared to hypoglycemic, but also the influence of the pathology on functional brain connectivity. An increment of power has been found in theta band and an occipital and parietal increase of power has been detected in alpha band, compensated by a parallel overall decrease of power in beta band during HYPO respect to EU. At this point, to improve the analysis of this kind of data, we have to consider the nature of the EEG signal itself: it is a non-stationary biological temporal series, whose statistical properties are variable over time. Indeed, the frequency domain parameters provide an evaluation of the autonomic function, but the reliability of the spectral power decreases with the decrease of the power signal and the signal-to-noise ratio. Many studies, e.g., [32], evinced a decrease of EEG signals complexity in compromised pathophysiological and conscience conditions, showing the usefulness of complexity tools to increase the knowledge of this altered states. In [22], non-linear methods are applied to try and locate important dynamic proprieties hidden in hypoglycemic signals. In particular, the application of non-linear complexity indices in analysis of changes in EEG signals in HYPO can provide more information on the functioning brain during the disease and its effect. The indices used so far, ApEn and SampEn, include only the results reported in [33], where a decrease of EEG complexity, understood as irregularity, has been demonstrated to be induced by HYPO, by calculating Sample Entropy (SampEn) at various factors scale, that is the multi-scale Entropy algorithm (MSE), [34]. The entropy algorithms are often used in the analysis of biomedical signals but, their high computational cost, in particular of ApEn and SampEn, make them not suitable for a real-time application aimed at the detection of hypoglycemia event.

3.2 Description of the chosen methodologies

In this chapter, we introduce different non-linear indices with intent to detect HYPO from EEG. The first, Empirical Permutation Entropy and Empirical Conditional Entropy, are innovative entropy-based methods, widely used to evaluate the complexity of biological signals [35]. The following ones are indices belonging to base on the fractal dimension, the first is the traditional Higuchi index [36],

while Residuals and Tortuosity are two new indices which reflect both the amplitude of signals as well as frequency properties [37] already used in [22] for our same purpose, but on a less wide database and on longer temporal series (60 min vs. 5 min). By applying the above-mentioned methods, we can obtain new evaluation of EEG signals complexity, with a reduced computational cost compared to previous algorithms. Furthermore, we will evaluate which of the illustrated methodologies is the most performing, not only in terms of capability to detect hypoglycemia, but also in terms of computational cost. The methods proposed in this section could be of practical relevance, since they can be used in real-time application and also because they can allow the simultaneous computation of more features and the grouping of multiple assessments temporally consecutive.

3.2.1 Empirical Permutation Entropy

Entropies and entropy-like quantities are playing an increasing role in data analysis, in the contexts of dynamical systems and of stochastic processes. Various applications of entropies are used, for example, in the analysis of physiological time series. Permutation entropies and variants are based on the ordinal structure of time series and systems behind them. In order to quantify complexity on the base of real-world data, Bandt and Pompe introduced Permutation Entropy [38]. It is based on the distributions of ordinal patterns (OPs), which describe order relations between the values of a time series [39] [40]. The definition of the ordinal pattern is needed to introduce the empirical permutation entropy. Let $\prod_d be$ the set of permutations of the set $\{0,1, \ldots, d\}$ for $d \in \mathbb{N}$. Then, the real vector $(x_0, x_1, \ldots \ldots, x_d) \in \mathbb{R}^{d+1}$ has the ordinal pattern $\pi = (r_0, r_1, \ldots \ldots, r_d) \in \prod_d of$ order $d \in \mathbb{N}$ if

$$x_{r0} \geq x_{r1} \geq \cdots \geq x_{rd} \text{ , where} \quad (3.1)$$
$$r_{l-1} > r_l \text{ in the case } x_{r-l} = x_{rl}.$$

Originally, ordinal patterns of order d were defined as permutations of the set $\{0,1, \ldots, d\}$. Subsequently, it was affirmed the following equivalent definition, which provides simpler enumeration of ordinal patterns [41]. A delay vector $(x_t, x_{t-\tau}, \ldots \ldots, x_{t-d\tau})$is said to have an *ordinal pattern* $\left(i_1^\tau(t), i_2^\tau(t), \ldots, i_d^\tau(t)\right)$ of order $d \in \mathbb{N}$ and delay $\tau \in \mathbb{N}$, where for $l = 1,2, \ldots, d$

$$i_l^\tau(t) = \#\{r \in \{0,1, \ldots \ldots, l-1\} | \, x_{t-l\tau} \geq x_{t-r\tau}\}. \quad (3.2)$$

Note that the occurrence of equal values (tied ranks) in a time series is assumed to be quite rare; for this reason, the relation “equal to" is combined with the relation “greater than". Briefly, each $i_l^\tau(t)$ codes how many points from $\left(x_t, x_{t-\tau}, \ldots \ldots, x_{t-(l-1)\tau}\right)$ are not larger than $x_{t-l\tau}$. Note that in Definition 3.1 the distance $\tau \in \mathbb{N}$ between the values in ordinal patterns is fixed to $\tau = 1$, so Definition 3.2 is more flexible for application to real-world data. There are $(d + 1)!$ ordinal patterns of order d, and it is assigned to each of them a number from in a $\{0,1, \ldots, (d+1)! - 1\}$ one-to-one way, by [41]

$$n_d^\tau(t) = n_d^\tau\left(\left(i_1^\tau(t), i_2^\tau(t), \ldots, i_d^\tau(t)\right)\right) = \sum_{l=1}^{d} i_l^\tau(t) \frac{(d+1)!}{(l+1)!} \quad (3.3)$$

For example, all OPs of order d = 2 are given in Table 3.1

Ordinal pattern						
(i_1,i_2)	(0,0)	(0,1)	(0,2)	(1,0)	(1,1)	(1,2)
$n_2(i_1,i_2)=3i_1+i_2$	0	1	2	3	4	5

Table 3: The ordinal patterns of order d=2.

In Figure 3.1 we present the OPs of the time series $(x_t)_{t=1}^{10}$, the delay $\tau = 2$ indicates a distance between points in OPs, the order $d = 2$ indicates number of points $(d + 1) = 3$ in ordinal patterns.

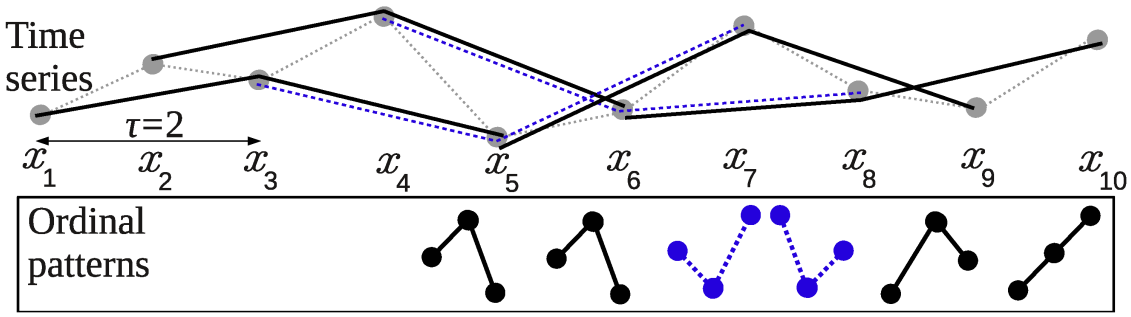

Figure 3.1: Illustration of computing the ordinal patterns of order d=2. Reproduced from [37]

It is possible to see that the blue OPs (dashed line) “overlap" the previous black OPs in $d = 2$ points, as well as the black OPs “overlap" the previous blue OPs. This maximal “overlapping" between OPs is usually used in order to obtain the maximal information from a time series [41]. Due to the “overlapping" between OPs, the successive OP $i_d^\tau(t+\tau) = (i_1', i_2', \ldots, i_d')$ is obtained from the previous one $i_d^\tau(t) = (i_1, i_2, \ldots, i_d)$ by:

$$i_{l+1}' = \begin{cases} i_l & if\ x_{t-l\tau} < x_{t+\tau} \\ i_l + 1 & otherwise \end{cases} \tag{3.4}$$

for $l = 0, \ldots \ldots, d$ with $i_0 = 0$. In this way the original time series is converted to a sequence of ordinal patterns of an order d, each of them describing order relations between $(d + 1)$ successive points of the time series. The more complex the underlying dynamical system is, the more diverse the ordinal patterns occurring for the time-series are. This diversity is just what the permutation entropy measures. According to (3.4), only d comparisons and at most d incrementation operations are needed to obtain the successive OP when the current OP is given, which provides a relatively fast computing of OPs [41]. When counting OPs in their number representation (3.3), which is more convenient than in the representation provided by (3.2), one needs d multiplications more. In order to compute the complexity of EEG time series, is defined the empirical permutation entropy (ePE) [40] of order $d \in \mathbb{N}$ and of delay $\tau \in \mathbb{N}$ of a time series $(x_i)_{i=1}^N$ with $N \in \mathbb{N}$ we understand the quantity:

$$\text{ePE}(d, \tau, (x_i)_{i=1}^{N}) = -\frac{1}{d}\sum_{j=O}^{(d+1)!-1} p_j ln p_j \text{ , where}$$

$$p_j = \frac{\#\{\, i=d\tau+1,d\tau+2,\ldots..,N \mid (x_i,x_{i-\tau},\ldots..,x_{i-d\tau}) \text{ has the ordinal pattern } j\}}{N - d\tau}$$

(with 0 ln0:= 0). (3.5)

The empirical permutation entropy is a well-interpretable measure of complexity. Indeed, the higher the diversity of ordinal patterns of order d in a time series $(x_i)_{i=1}^{N}$ is, the larger the value of ePE(d, τ, N) is. It holds for d, N, $\tau \in \mathbb{N}$:

$$0 \leq \text{ePE}(d, \tau, N) \leq \frac{\ln((d+1)!)}{d}, \tag{3.6}$$

which is restrictive for estimating a large complexity. The choice of order d is rather simple. The larger d is, the better the estimate of complexity underlying a system by the empirical permutation entropy is. On the other hand, excessively high d value leads to an underestimation of the complexity because by the bounded length of a time series not all ordinal patterns representing the system can occur. In [42],

$$5(d+1)! < N \tag{3.7}$$

It is recommended. The choice of the delay τ is slightly complicated; in many applications $\tau = 1$ is used, however larger delays τ can provide additional information. For this reason, it was tried to use value of τ different by 1, but the results are clearly worse than those previously obtained. Note also that increasing the delay τ can lead to increasing the values of ePE. In order to reflect complexity changes in a time series in the course of time, the ePE values are usually computed in sliding time-windows of a fixed size, see also the previous definition:

By the empirical permutation entropy (ePE) of order d and of delay τ of a time-window $(x_t, x_{t-1}, \ldots\ldots, x_{t-M-d\tau+1}), t, M \in \mathbb{N}$ one understands the quantity

$$\text{ePE}(d, \tau, M, t) = -\frac{1}{d}\sum_{j=O}^{(d+1)!-1} \frac{q_j(t)}{M} \ln \frac{q_j(t)}{M} = \ln M - \frac{1}{M}\sum_{j=O}^{(d+1)!-1} q_j ln q_j(t),$$

$$\text{where } q_j(t) = \#\{i = t, t-1, \ldots., t-M+1 \mid n_d^{\tau}(i) = j\}$$

(with 0 ln0: = 0). (3.8)

We use a time-window $(x_t, x_{t-1}, \ldots\ldots, x_{t-M-d\tau+1})$ because it contains exactly M ordinal patterns which is convenient for computations and further explanations, [35]. In Figure 3.2, we illustrate computing ePE in the two maximally overlapped windows $(x_t)_{t=1}^{9}$ and $(x_t)_{t=2}^{10}$ containing M = 5 OPs of order d = 2 for a delay $\tau = 2$.

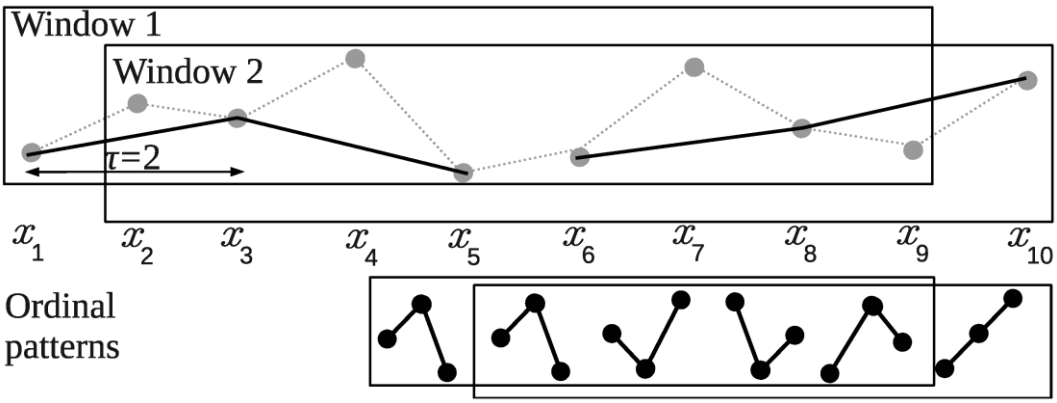

Figure 3.2: Illustration of computing the empirical permutation entropy in two successive and maximally overlapped sliding windows. Reproduced from [45].

It is truly important mention that the main feature of ePE, is that it should satisfy the weak stationarity requirement. This means that in each sliding windows the distribution of ordinal patterns does not change. In Fig. 3.3, it was summarized the diagram of the calculation algorithm, where the size of a sliding window is denoted by M, the order of ordinal patterns by d, the delay by τ and the length of a time series by W:

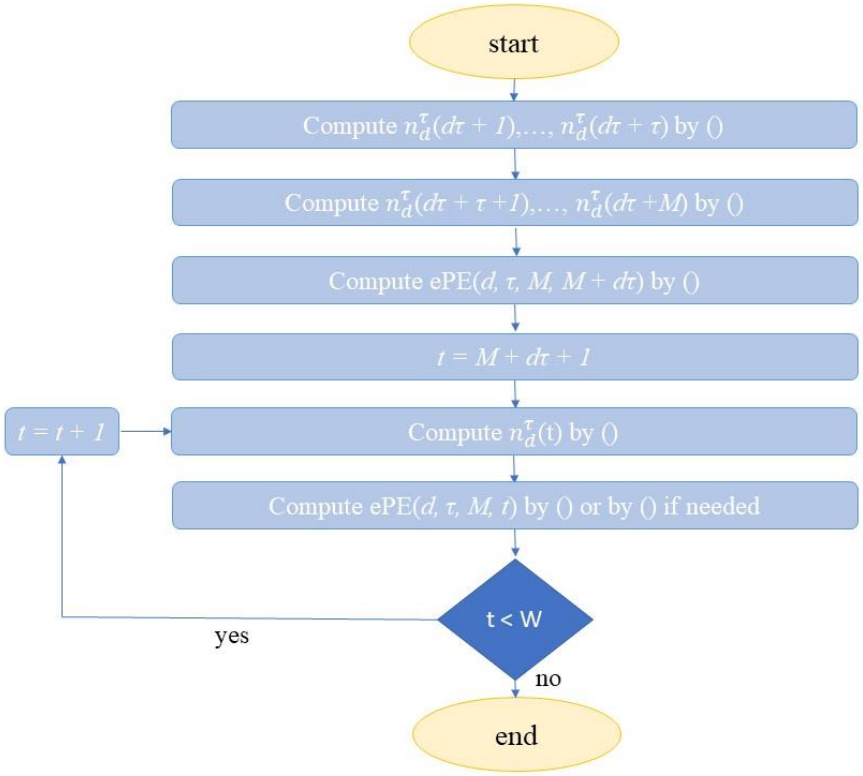

Figure 3.3: Algorithm of fast computing empirical permutation entropy in sliding windows [45].

Empirical permutation entropy is robust to noise [43], computationally simple and fast [41]. It has been applied to detect and visualize EEG changes related to epileptic seizures (e.g., [44]), to distinguish brain states related to anaesthesia [45], to discriminate sleep stages in EEG data [40], to analyse and classify heart rate variability data [46], and for financial, physical and statistical time series analysis (see [47] for a review of applications). Moreover, we considered the robustness of the

ePE with respect to observational noise, since there were situations in which it is not as robust to noise as usually reported (e.g., [48], [49]).
Given a time series $(x_i)_{i=1}^N$, the observational noise $(\xi_i)_{i=1}^N$adds an error ξ_i to each value x_i:

$$y_i = x_i + \xi_i. \tag{3.9}$$

The robust empirical permutation entropy (rePE) is introduced [40], which is based on counting robust ordinal patterns defined in the following way:

For positive $\eta \in \mathbb{R}$, it is call an ordinal pattern of the vector $(x_t, x_{t-\tau}, \dots\dots, x_{t-d\tau}) \in \mathbb{R}^{d+1}$ η-robust if

$$\#\{(i,j): 0 \le i < j \le d, |x_{t-i\tau} - x_{t-j\tau}| < \eta\} < \frac{(d+1)d}{8}. \tag{3.10}$$

The threshold $\frac{(d+1)d}{8}$ is chosen as a quarter of the amount of the ordered pairs of entries from the vector $(x_t, x_{t-\tau}, \dots\dots, x_{t-d\tau})$.

$$\text{rePE}(d, \tau, (x_i)_{i=1}^N, \eta) = -\frac{1}{d}\sum_{j=O}^{(d+1)!-1} p_j ln p_j \text{ , where}$$

$$p_j = \frac{\#\{\, i = d\tau+1,\ d\tau+2,\dots..,N \mid (x_i, x_{i-\tau},\dots.., x_{i-d\tau}) \text{ has the } \eta\text{−ordinal pattern } j\}}{\#\{\, i = d\tau+1,\ d\tau+2,\dots..,N \mid (x_i, x_{i-\tau},\dots.., x_{i-d\tau}) \text{ has an } \eta\text{−ordinal pattern}\}}$$

(with 0 ln0: = 0 and 0/0: = 0). (3.11)

RePE is applied in EEG data in order to verify if the signal is compromise by the error. The final valuation is widely satisfactory since rePE and ePE lead back to the same result, and the detection of hypoglycemia turns out to be the same for all subject. This means that the data is cleaned good from artefacts. However, the rePE has two drawbacks. Firstly, it is necessary to set the parameter η in order to compute rePE, which is ambiguous. Secondly, rePE has a slower computational algorithm than ePE [35].

3.2.2 Empirical Conditional Entropy

The quantity called Conditional Entropy of ordinal patterns (eCE) was investigated, akin to the permutation entropy. The theoretical conditional entropy is introduced in [50], and it seems that this approach estimates the Kolmogorov-Sinai entropy better than ePE for several cases, in particular for a finite order *d*. It can be also used in several applications as demonstrated in [50], as, for example, the prediction of epileptic seizures. In this section we describe the algorithm, about the average diversity of the ordinal patterns succeeding a given ordinal pattern.
By the empirical conditional entropy of ordinal patterns $\text{eCE}(d, \tau, (x_i)_{i=1}^N)$ of order $d \in \mathbb{N}$ and of delay $\tau \in \mathbb{N}$ of a time series $(x_i)_{i=1}^N$ with $N \in \mathbb{N}$ we understand the quantity [40]:

$\mathrm{eCE}(d, \tau, (x_i)_{i=1}^{N}) = -\frac{1}{d} \sum_{j=O}^{(d+1)!-1} \sum_{j=O}^{(d+1)!-1} p_j\, q_{j,l}\, ln q_{j,l}$, where

$$p_j = \frac{\#\{\, i \in I \mid (x_i, x_{i-\tau}, \ldots\ldots, x_{i-d\tau}) \text{ has the ordinal pattern } j\}}{N - d\tau - \tau}$$

$$q_{j,l} = \frac{\#\{i \in I \mid (x_i, x_{i-\tau}, \ldots\ldots, x_{i-d\tau}) \text{ and } (x_{i+\tau}, x_i, \ldots\ldots, x_{i-(d-1)\tau}) \text{ has the ordinal pattern } j \text{ and } l\}}{\#\{\, i \in I \mid (x_i, x_{i-\tau}, \ldots\ldots, x_{i-d\tau}) \text{ has ordinal pattern } j\}}$$

$$\text{for } I = \{d\tau + 1, d\tau + 2, \ldots, N - \tau\} \tag{3.12}$$

(with 0 ln0 := 0 and 0/0 := 0).

It holds for $d, N, \tau \in \mathbb{N}$:

$$0 \leq \mathrm{eCE}(d, \tau, N) \leq \ln(d + 1). \tag{3.13}$$

The choice of the order d of the empirical conditional entropy is similar to that of the empirical permutation entropy. As in (3.7), order d and length N of a time series should satisfy $5(d + 1)!(d + 1) < N$, otherwise eCE can be underestimated. Note that to compute eCE, one has to estimate frequencies of all possible pairs of ordinal patterns, and there are $(d + 1)!(d + 1)$ such couples. However, according to authors' experience, in many cases it is enough to have:

$$(d + 1)!(d + 1) < N. \tag{3.14}$$

In this section, we used the method we adapted to ePE for the efficient calculation of ordinal patterns and was introduced by [40]. In contrast to the ePE, the eCE is computed not only from the distribution of OPs but also from the distribution of (2, d)-words that are pairs of successive OPs of order d, in particular:

We say that a real vector $(x_0, x_1, \ldots\ldots, x_{d+n-1}) \in \mathbb{R}^{d+n}$ has the (n, d)-word $\pi_1, \pi_2, \ldots, \pi_n$ if $(x_i, x_{i+1}, \ldots, x_{i+d})$ has the ordinal pattern $\pi_{i+1} \in \prod_d$ for $i = 0,1, \ldots\ldots, n - 1$. It is possible to define the delay vector $(x_t, x_{t-\tau}, \ldots, x_{t-(d+1)\tau})$which have an (2,d)-word $(i_d^\tau(t), i_d^\tau(t - \tau))$ when the OPs $i_d^\tau(t)$ and $i_d^\tau(t - \tau)$ are given by (3.2). There are (d+1)!(d+1) possible (2, d)-words, we assign to each of them a number from $\{0,1, \ldots\ldots, (d + 1)!\,(d + 1) - 1\}$ in a one-to-one way by:

$$w_d^\tau(t) = w_d^\tau(i_d^\tau(t), i_d^\tau(t - \tau)) = (d + 1) n_d^\tau(t - \tau) + l, \tag{3.15}$$

$$\text{where } l = \#\{r \in \{1,2, \ldots, d\} \mid x_{t-r\tau} \geq x_t\}. \tag{3.16}$$

In [35], it is illustrated that only two more arithmetical operations are needed to compute $w_d^\tau(t)$ together with $n_d^\tau(t)$ instead of computing $n_d^\tau(t)$ only, because the same l from (3.16) is used for both of them. Now we define, like for ePE, empirical conditional entropy of ordinal patterns introduced in [35] computed in sliding time-windows of a fixed size. By the empirical conditional entropy of ordinal patterns (eCE) of order $d \in \mathbb{N}$ and of delay $\tau \in \mathbb{N}$ of a time-window $(x_t, x_{t-1}, x_{t-2}, \ldots\ldots\ldots, x_{t-M-(d+1)\tau+1})$, it is possible to understands the quantity:

$\mathrm{eCE}(d, \tau, M, t) = \sum_{j=O}^{(d+1)!-1} \frac{q_j(t-1)}{M} \ln \frac{q_j(t-1)}{M} - \sum_{j=O}^{(d+1)!(d+1)-1} \frac{p_j(t)}{M} ln \frac{p_j(t)}{M}$, where

$$p_j(t) = \#\{i = t, t-1, \ldots., t-M+1 \mid w_d^\tau(i) = j\}$$
$$q_j(t-1) = \#\{i = t-1, t-2 \ldots., t-M+2 \mid n_d^\tau(i) = j\}$$

(with 0 ln0: = 0). (3.17)

More details of compute in [35], while in Fig. 3.4, it was summarized the diagram of the calculation algorithm, where the size of a sliding window is denoted by *M*, the order of ordinal patterns by *d*, the delay by τ and the length of a time series by *W*.

The conditional entropy of ordinal patterns has rather good properties. Theoretical results and numerical experiments show how in many cases the conditional entropy provides a reliable estimation of the KS entropy. In this regard it is important to note that the eCE is computationally simple: it has the same computational complexity as the permutation entropy. Meanwhile, some questions concerning the conditional entropy of ordinal patterns remain open, see [50] for more details.

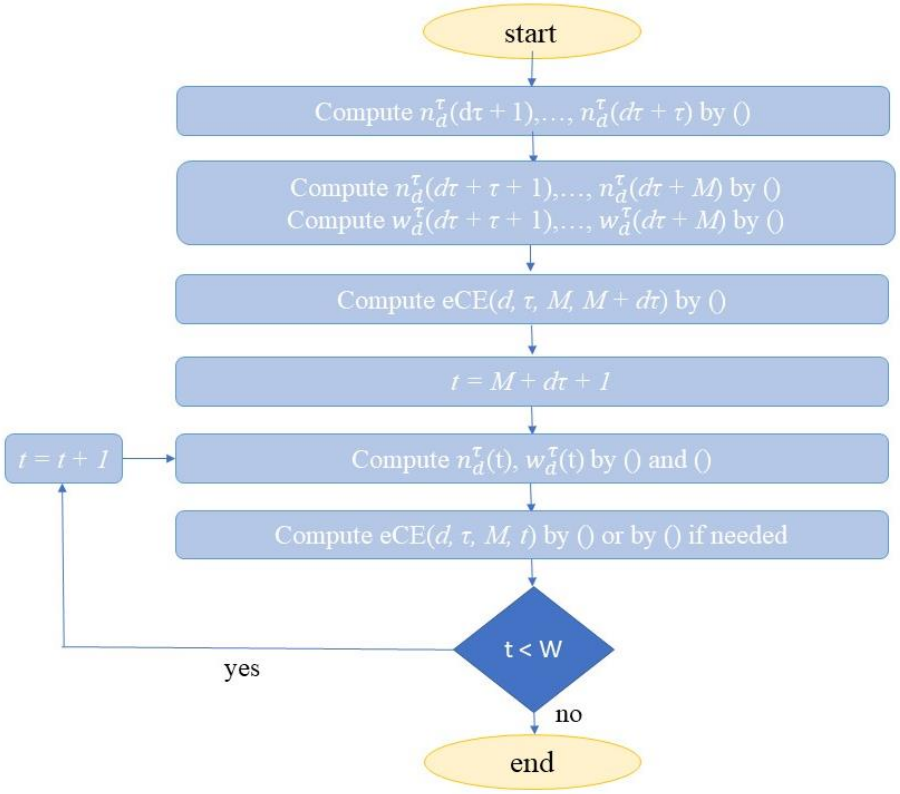

Figure 3.4: Algorithm of computing empirical conditional entropy of ordinal patterns in sliding windows.

3.2.3 Features extraction based on fractal dimension

The fractal dimension provides a complexity index that describes how the measure of the length of a curve L(k) changes depending on the scale k used as unit of measurement. Higuchi's algorithm [51] is often employed in EEG analysis to estimate the fractal dimension D [36] [52]. The Higuchi's

standard approach calculates fractal dimension of time series in the time domain, and is based on the log(L(k)) vs log(k) curve computed as follow:

- For each sample *i* of the EEG epoch S, absolute differences between the values S(i) and S(i-k) (i.e. samples at distance k) are computed, considering k = 1,….., k_{max};

- Each absolute difference is multiplied by a normalization coefficient that takes into account the different number of samples available for each value of k. The computation of this coefficient is based on the starting point m = 1,…,k and on the total number (N) of samples of an epoch;

- L(k) is computed by summing the obtained values and dividing by k;

$$L(k) = \frac{1}{k}\sum_{m=1}^{k}\left[\frac{1}{k}\left(\sum_{i=1}^{q}|S(m+ik) - S(m+(i-1)k)|\right)\frac{N-1}{qk}\right],$$

$$\text{where } q = int\left[\frac{(N-m)}{k}\right]; \tag{3.18}$$

- The log[L(k)] vs log(k) curve, referred in the following as l_k, is finally derived.

By definition, if L(k) is proportional to k^{-D} (i.e. log(k) and log[L(k)] have a linear relationship) for k = 1,2,…., k_{lin}, then the curve is fractal with dimension D. Consequently, k_{lin} is the maximum k for which L(k) is proportional to k^{-D} and D is estimated by ordinary least squares as the linear coefficient of the regression line of the l_k curve for k = 1,2,….,k_{lin}. Therefore, the Higuchi's fractal dimension increases as the signal irregularity increases. The non-linear part of the l_k curve ($k > k_{lin}$) presents an oscillatory behaviour whose characteristics depend on the periodicity of the signal itself [37]. In addition to Higuchi's fractal dimension, we also extract two additional features [37], which consider a wider domain of the l_k curve with respect to the Higuchi fractal dimension. The first additional feature evaluates the deviation (i.e. the sum of squares of the residuals) of the l_k curve from the regression line computed on its linear region. The second additional feature is a tortuosity measures τ of l_k. It consists in a measure of the rate at which the curve is changing with respect to its coordinates changes (x = log(k) and y = log(L(k))), by using their first and second partial differences:

$$\tau_{lk} = \sum_{n=3}^{k_{max}}\left|\frac{\Delta x(n)\Delta^2 y(n) - \Delta^2 x(n)\Delta y(n)}{[(\Delta x(n))^2 + (\Delta y(n))^2]^{3/2}}\right| \tag{3.19}$$

Where $\Delta x(n) = x(n) - x(n-1)$, $\Delta^2 x(n) = \Delta x(n) - \Delta x(n-1)$, $\Delta y(n) = y(n) - y(n-1)$, $\Delta^2 y(n) = \Delta y(n) - \Delta y(n-1)$. A complete description of this tortuosity measures can be found in [53]. These two features increase as the signal regularity and periodicity increase. The linear region is here defined considering $k_{lin} = 6$, according to [36], while the two other features are computed considering also the non-linear region up to $k_{max} = 30$ according to [37].

3.3 Implementation parameters and details

This subsection is about the influence of the choice of EEG channel, of the delay τ, of the order d and of the window size N when applying the empirical permutation entropy ePE(d,τ,N) and the

empirical conditional entropy eCE(d,τ,N). In fact, being very similar to each other in the implementation of the algorithms, the choices made for the parameters described above will be applied for both methods. In particular, for each subject, from the two 5-m intervals corresponding to EU and HYPO conditions, which presented the highest SNR and the least number of artifacts displayed through a visual inspection, the P3 channel was chosen since the EEG signal from this scalp position was highly affected by hypoglycemic events [54] [55]. This choice was done in order to compare our results to fractal dimension indicators obtained in [22], SampEn obtained in [33] considering that our analysis is carried out on the same database of [22] and [33]. The ePE and the eCE values strongly depend on the chosen EEG channel and for this reason, authors recommended computing them from all the EEG channels when analysing EEG data. For this reason, for a more in-depth study, the analysis was effectively extended on all the channels. ePE and eCE need to satisfy the weak stationarity requirement. Therefore, the size of the sliding window should be chosen in a way that the distribution of ordinal patterns does not change in the window. The size of a sliding window for EEG data could also be chosen equal to 2 s, as in [56]. However, we are going to use the sliding windows of 4 s size in order to compare this analysis with the one carried out in the previous studies using the feature domain. We compute all entropies in maximally overlapping sliding windows of 4 s size. About the parameters, the order d is usually set to 6, i.e., the maximum for the chosen window size, according to the recommendation in equation (3.7). Different values of τ were tested [1 : 3], finally we set 1, because it provides the best hypoglycemia detection. However, in [35], authors recommend combining the measures computed for different delays τ and order d since it allows revealing different features of underlying system and it is helpful for discrimination of time series with different underlying complexity. An example in [35], shows how an increase of delay τ can lead to an increase of the ePE(d,τ,N) values so that they almost attain the upper bound, described in the following relation:

$$\text{ePE}(d,\tau,N) \leq \frac{\ln((d+1)!)}{d} \tag{3.20}$$

Concerning the fractal analysis, it is applied to 5-min intervals corresponding to EU and HYPO at theP3-EEG channel in order to compare fractal dimension indicators to SampEn/ePE/eCE results obtained still now. The 5 min chosen are the same as described above and the P3 channel in EU and HYPO was split into epochs of 4 s length, with 2 s overlap. Epoch length and overlap were chosen as a trade-off between the amount of available data samples required to calculate the features for each epoch and the readiness required for the prompt identification of a physiological change. EEG data can be affected by several sources of noise, such as body movements and propagation of bioelectrical muscles potentials, which are likely to sum up into the electroencephalographic signal generated by the brain. Consequently, recorded EEG signals may show trends, low frequency oscillations and baseline changes due to these components. In order to remove these artefacts a detrending step to each EEG epoch was applied to correct drifts and linear trends. Briefly, a first order model consisting of an offset plus a linear trend was fitted to each epoch, estimated and then subtracted from the epoch itself. Furthermore, if the value of one or more of its samples was greater than 100 μV or lower than -100 μV, the epoch was considered to be affected by artefacts and thus not considered for the analysis (on average, 0.4% epochs rejected per subject).

3.4 Results

In this work, several indices have been calculated: entropy-based, i.e. ePE, rePE and eCE, and fractals, i.e Higuchi, Residuals and Tortuosity. All the indices have been computed using 4-s EEG epochs in EU and HYPO and finally compared to each others. Furthermore, since a similar analysis was conducted in [22] with a limited database of 19 subjects and longer EEG time series (almost 1 h long), we compared the results of the fractal analysis with those resulting from this study. First, we will show the empirical permutation entropy outcomes that following different trend in EU and HYPO. The entropic methods consist of two steps: segmentation of time series into pseudo-stationary segments and the calculation of OPs distribution. Given the study conducted in [22] and the research in literature, the EEG signal epoch can be considered pseudo – stationary if its length is between 2 and 4 s. We chose a length of 4 s in order to enable comparisons with previous studies and because it provides good results in the experiments. Finally, the distribution of OPs was computed using the algorithm implemented in [57]. Remembering from the literature the property for which the ePE values vary from a minimum of 0 to a maximum of $\frac{\ln((d+1)!)}{d}$, the same is applies to rePE and eCE, the greater the complexity of the signal, much more the entropic value will be close to the maximum assumable and vice versa. Observing Fig 3.5, we can confirm a decrement of ePE amplitude, in particular of hypoglycemic state (blue) compared to euglycemic (green) in a random subject of database. This result is extended to all subjects, as shown in Fig 3.6, in which the distribution of ePE values was computed for all database subjects in both euglycemic states.

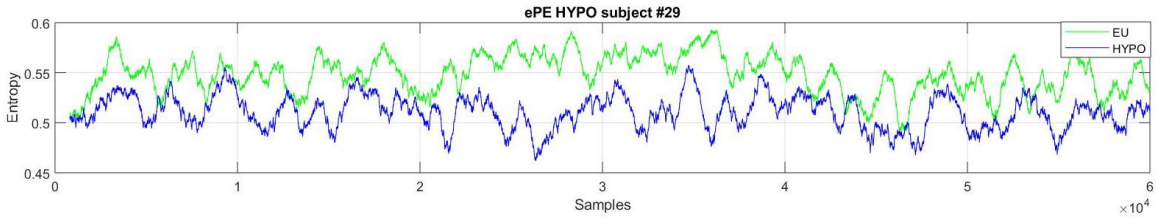

Figure 3.5: The ePE values of a random subject are represented in its EU status (green) and HYPO (blue). This representation refers to the entire available EEG signal, which is 5 min long. Notice how hypoglycemic values are lower along the whole path compared to euglycemic ones.

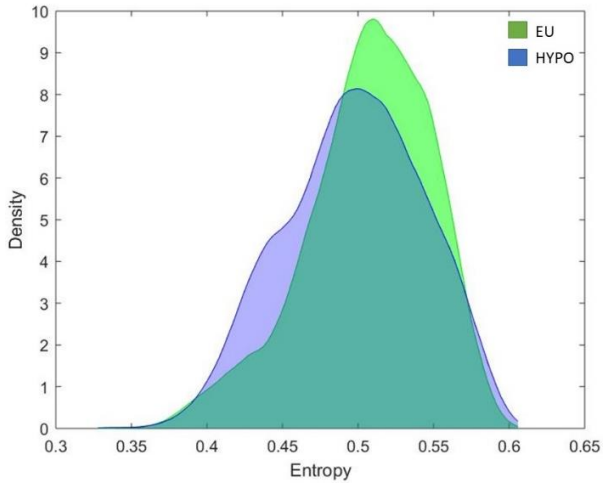

Figure 3.6: Distributions of the values for 4-s epochs in EU (green distribution) and HYPO (blue distribution) of ePE for all subjects.

Subsequently, the rePE index was computed in order to verify if the signal is compromised by the error. The final valuation is widely satisfactory since rePE and ePE lead back to the same result: the detection of hypoglycemia turns out to be the same for all subject. Finally, the last entropy index was also evaluated: eCE. As in the case of ePE, the eCE values are lower in the case of hypoglycemic than in euglycemic one, but in a less evident way as can be seen from Fig 3.7. These results are confirmed on all subjects as shown in Fig 3.8, in which the distribution of the eCE values for all the subjects in database in each glycemic state was computed. The indices obtained from both the eu- and the hypoglycemic state were compared by a two-sided Wilcoxon rank sum test to identify significant changes in the signal properties in two 5-m intervals corresponding to EU and HYPO. A statistically significant decrease of the standard ePE was found from eu- to hypoglycemic state in 28 over 34 subjects (p-value always < 0.01, except for subjects with id 1, D13, D18, D28, 19, 23 for which p-value > 0.05). Same results were found for rePE. A statistically significant decrease was found in 30 subjects over 34 for eCE (p-value < 0.01 while for subjects with id 1, D15, 23 and 27, p-value > 0.05, respectively).

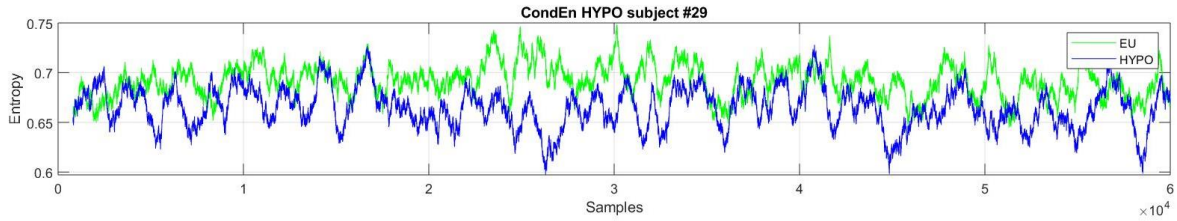

Figure 3.7: The eCE values of a random subject are represented in its EU status (green) and HYPO (blue). This representation refers to the entire available EEG signal, which is 5 min long. Notice how hypoglycemic values are slightly lower along the whole path compared to euglycemic.

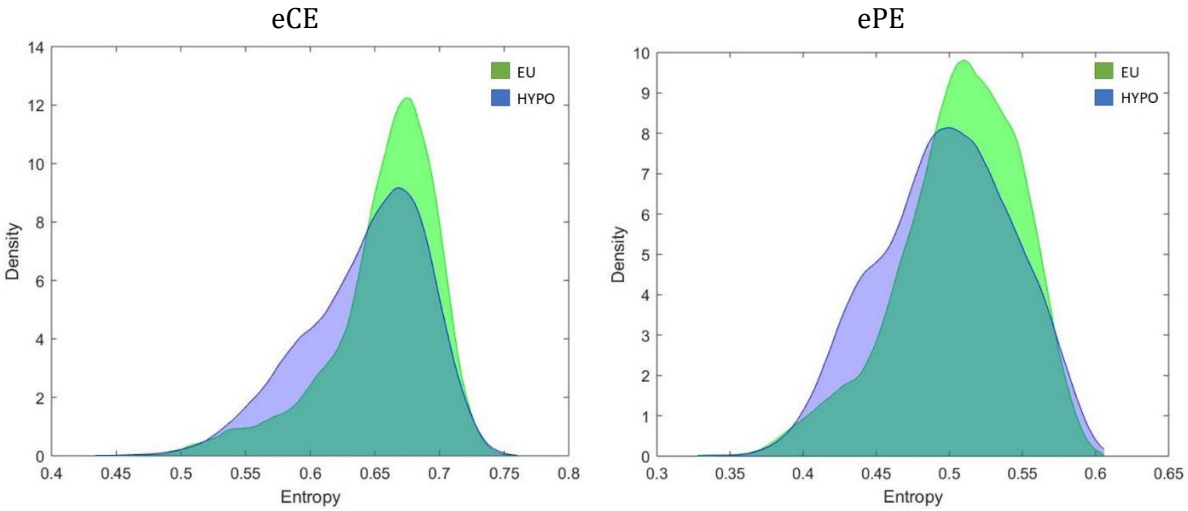

Figure 3.8: Distributions of the values for 4-s epochs in EU (green distribution) and HYPO (blue distribution) of eCE (left) and ePE (right) for all subjects.

For each 4-s EEG epoch in EU and HYPO, the fractal dimension features described in the previous section were computed on the 19 subjects in common with the study of [22]. This comparison is necessary to evaluate the efficiency of features on a temporal series much shorter, i.e. 5 min of EEG signal. The obtained results are comparable to the previous ones: we can confirm a decrease in Higuchi and an increase in Tortuosity, while the increase in Residuals is less evident. Subsequently a new evaluation was carried out on the entire database composed of 34 subjects, and it shows

different distributions in the two glycemic states, denoting a decrease in the EEG signal complexity in all the subjects during HYPO. Indeed, as visible in Fig. 3.9 for a representative subject where the distributions of the features in EU (green distribution) and HYPO (blue distribution) are reported, the Higuchi's fractal dimension tends to decrease shifting from EU to HYPO, while the fractal dimension Residuals and Tortuosity tend to increase. These results are confirmed across all subjects, in Fig. 3.10, except for Residuals that is almost coincident in both situations.

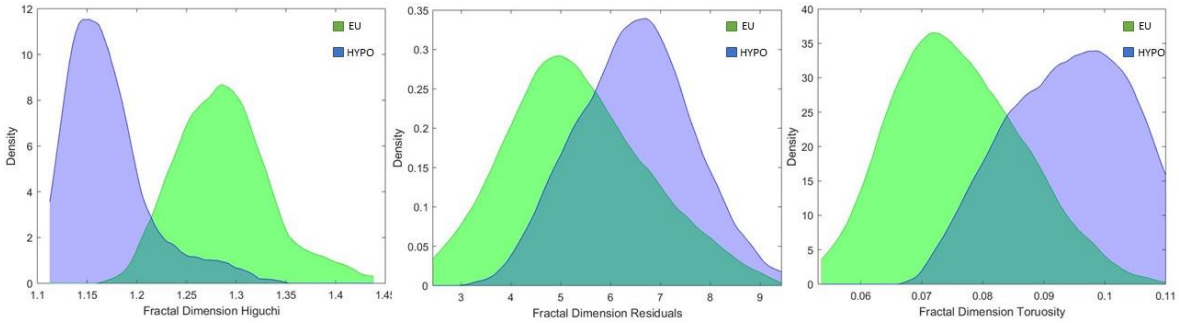

Figure 3.9: Distributions of the values for 4-s epochs in EU (green distribution) and HYPO (blue distribution) of Higuchi's fractal dimension (left), the Residuals (middle) and Tortuosity (right) for the representative subject id 16.

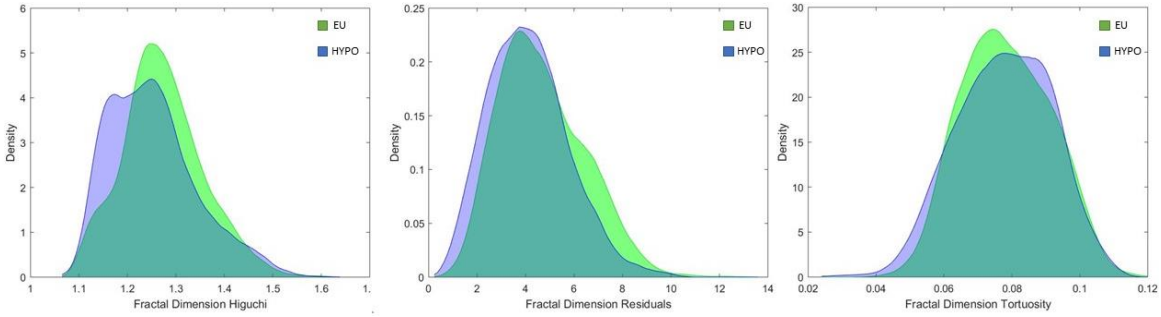

Figure 3.10: Distributions of the values for 4-s epochs in EU (green distribution) and HYPO (blue distribution) of Higuchi's fractal dimension (left), the Residuals (middle) and Tortuosity (right) for all subjects.

These features were compared by a two-sided Wilcoxon rank sum test in order to identify significant changes in two 5-m signals intervals corresponding to EU and HYPO. A statistically significant decrease of the standard Higuchi fractal dimension and a statistically significant increase for Residuals and Tortuosity features were found from eu- to hypoglycemic state in 33 over 34 subjects (p-value always < 1e-07, except for subject with id D15, for which p-value > 0.05). The decrease of Higuchi fractal dimension, the increase of the two additional features and the decrease of ePE and eCE in HYPO, highlight a reduction of the EEG signal complexity in this physiological condition.

Receiver operating characteristic (ROC) analysis was performed to quantitatively measure the capability of each indices to identify changes in the signal properties related to the different glycemic state. For each subject, the ROC curve was computed by plotting the true positive rate (sensitivity) versus the false positive rate (1- specificity), and the area under the curve (AUC) was measured. The AUC is a measure of how well a parameter can distinguish between two groups: it can vary from 0 to 1 and if equal to 1 indicates maximal sensitivity and specificity, and thus perfect discrimination between the two glycemic states. AUCs were calculated for Higuchi fractal dimension, Residuals

and Tortuosity features, ePE and eCE. In Fig. 3.11, the results of ROC analysis for all subjects are depicted. The AUC scores obtained by the features based on fractal dimension on many of the subjects denote their effectiveness in extracting useful information with discriminatory power from the EEG signals.

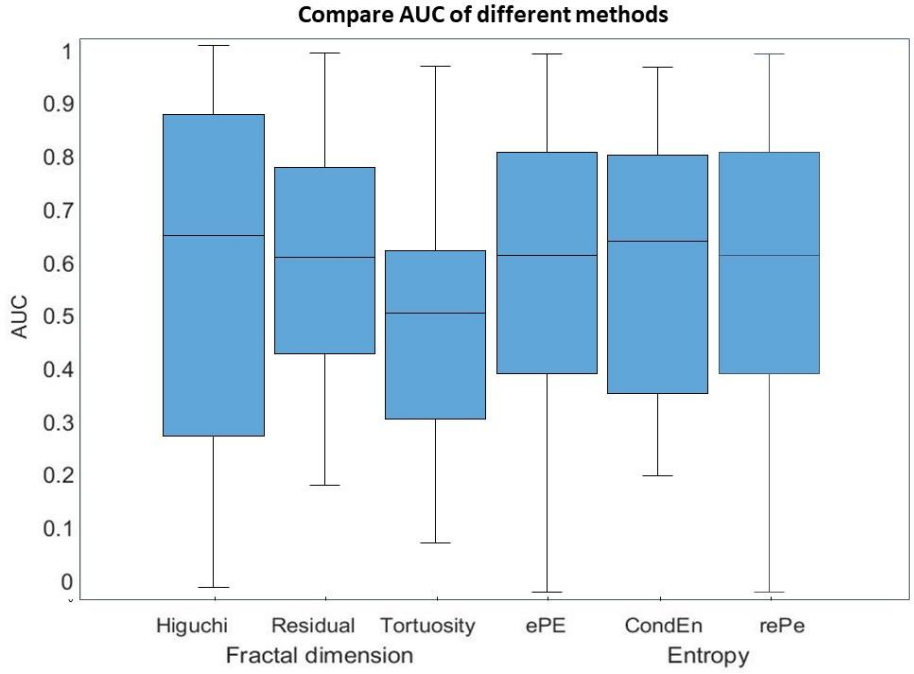

Figure 3.11: Results of the ROC analysis. Area Under the Curve (AUC) obtained by each feature for all subjects is reported. In each box, the central thick line is the median value, the edges of the box are the 75th and 25th percentiles, the whiskers extend to the most extreme data points that are not outliers. Each AUC is also plotted individually.

In order to make this analysis more complete, we compared the obtained results with previous studies about valuation and detection of hypoglycemia, in particular with [33], since in this research two well-known algorithms were used in the analysis of complexity: ApEn and SampEn. The purpose of this comparison is the evaluation and subsequent selection of the algorithm with the best computational cost – accuracy ratio for a deepening of possible future applications in real time. Computing ApEn and SampEn on epochs of 4-s epochs (N = 800) length could be applied without a scale factor, but their computational cost, $O(N^2)$, would be too high for a real-time application [58]. Instead, the computational cost of EEG complexity estimation assessed by fractal analysis from 4-s epochs (N = 800 samples) is O(KN). Indeed, the computation of L(k), see Eq. 3.18, requires the execution of N mathematical operations for each $k = 1 : K$, with K smaller than N (in this work, $K = klin = 6$ for Higuchi's fractal dimension, $K = kmax = 30$ for Residuals and Tortuosity, and $K = kmax = 30$ for all of them simultaneously). Finally, to understand the efficiently computing ordinal patterns we remained to [57]; in this work, we consider the ePE, computed in sliding windows of a size M with maximal overlapping (the same is for eCE), i.e., the first point of the successive window is the second point of the previous one. The successive windows $(x_{t-1}, x_{t-2}, \ldots\ldots, x_{t-M-d\tau})$ and $(x_t, x_{t-1}, x_{t-2}, \ldots\ldots, x_{t-M-d\tau+1})$ differ in points x_t and $x_{t-M-d\tau}$, therefore the ordinal distributions in the windows differ in the frequencies of occurrence of the ordinal patterns $n_d^\tau(t)$ and $n_d^\tau(t-M)$, see Eq 3.8. In order to obtain the ordinal distribution of the successive window given the current one, the frequency of the "outcoming" ordinal pattern and the frequency of the "incoming" ordinal pattern need to be substituted, see details in [57]. In this way, the total number of samples (representative

5min of signals), and then only 4 s are used to set the two algorithms ePE and eCE, as shown in section 3.2, otherwise the samples would not be enough to accurately compute them. In the following Table 4, what has been described so far is summarized.

Quantity	Seconds	Computational time	Storage
ApEn	4s	$O(N^2)$	$O(N)$
SampEn	4s	$O(N^2)$	$O(N)$
ePE	60s	$O(N)$	$O((d+1)!(d+1))$
eCE	60s	$O(N)$	$O((d+1)!(d+1))$
Fractal features	4s	$O(kN)$	

Table 4: Efficiency of computing the complexity's indices.

The short epochs length and the low computational cost allow fractal features and empirical permutation entropy to provide frequent assessment of the physiological state, which is essential for a useful real-time detection of HYPO. However, a more in-depth analysis of the performance of all the algorithms applied is necessary, considering that both the analyses of complexity yield comparable results. Starting from features of fractal domain, the most relevant property is their validity in non-stationary signals, which can be applied directly on real EEG signals. Furthermore, we can consider this algorithm for real – time application since for each 4-s of EEG signal (necessary limited storage in memory) it is possible to compute all three fractals features, Higuchi, Residuals and Tortuosity. We have highlighted this point in order to accurately evaluate the complexity of the EEG through the fractal domain, all three fractal features must be calculated. Regarding the analysis of entropy, we have discussed the applicability and performance of three empirical measures based on the distribution of ordinal patterns in time series, the computational cost of which is the same for all proposed entropic algorithms. Whereas the ePE is well established and eCE is recently introduced [50], the rePE is a new concept proposed for deal with noisy data. However, these algorithms provide some choices to make before implementing the distribution of OPs, in particular, we must set 3 parameters that depend on the length of the EEG time series: the order (d), τ and windows size. Additionally, rePE forecast the assignment of further parameter, η, which is still ambiguous in the literature, since its setting is not clearly defined. The rePE is a promising new quantity yet requiring further investigation, while eCE is an index recently introduce whose value and accuracy should be further deepened, see [50]. Moreover, unlike fractal dimension, the empirical permutation entropy requests two significant assumptions: the pseudo-stationarity of signal and the continuous storage of at least 60-s of analysed signal. In return, the calculation speed of ePE and eCE is undisputed, the calculation of rePE much slower and heavier. It is important to note that these three empirical measures are computationally simple, especially with very long time series.

3.5 Conclusions

We can conclude this comparison highlighting the easier implementation of fractal features, considering the fact that the results of all the algorithms applied, in terms of computational cost, are comparable. Fractal features can be used on non-stationary signals and do not require the choice of any parameter for the analysis, nor the storage of a signal quantity of more than 4 seconds. Moreover, the information content of these indices leads to results in line with those achieved by nonlinear

entropy-based indicators like SampEn e ApEn. Indeed, the decrease of the Higuchi fractal dimension as well as the increase of the residuals and the tortuosity features highlight the same decrease of the EEG signal complexity in HYPO condition. The next step involves the calculation of these complexity indices on several EEG channels in order to acquire as much information as possible from the signal.

Chapter 4

Relationship between EEG quantitative features and awareness of hypoglycemia

4.1 The starting point

Recent studies have investigated the relationship between resting state slow wave power, healthy aging and cognitive performance [59]. Similar study is made in this work to investigate the cognitive performance during HYPO for T1D patients and the relationship with the decrease in signal complexity. In order to evaluate the cognitive performance, all subjects performed the Stroop test and the TMT B during three glycemic periods, i.e., EU, HYPO and the recovery after the period in HYPO. The outcomes of the test in [22] shown that during the recovery period the performances for each of the tests are comparable with the performances during the period of EU before entering in HYPO. Thus, the visible patients drop in performance during HYPO was not due to tiredness (the HYPO interval is about 4 h after the start of the experiment in the hospital). Secondly, after applying Wilcoxon rank-sign tests, a significant worsening during HYPO compared to EU was found in the performance of all the tasks ($p < 0{,}001$). In particular, during TMT B performances deteriorated up to 30% and during the Stroop test up to 15%, revealing the drop in visual attention, processing speed and selective attention during HYPO. We also replicated the analyses by dividing the subjects with and without awareness of HYPO, but no significant difference in the performance for the two groups were found. In Fig. 4.1, we reported the results for each of the test for both groups in EU, HYPO and Recovery.

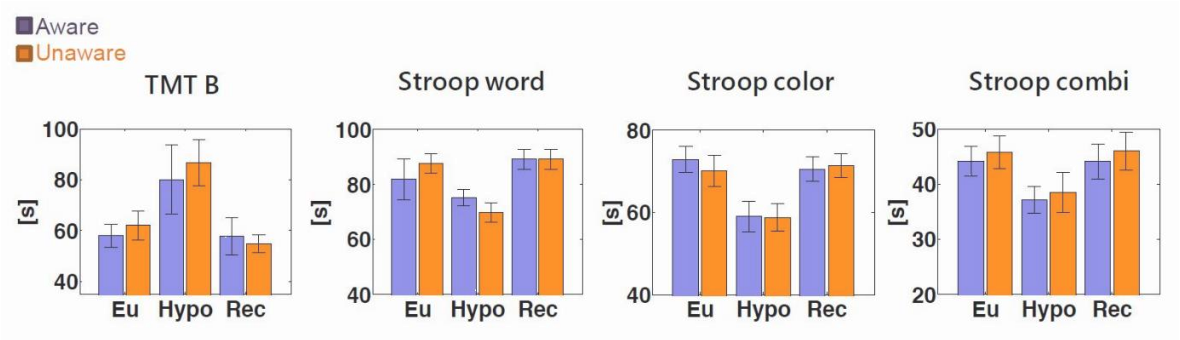

Figure 4.1: From left to right, results of TMT B and Stroop test in aware (violet) and unaware (orange) patients in EU, HYPO and Recovery. In all conditions, the performance for awareness patients is comparable with the one of unawareness subjects. There are not significant differences.

The second step is to better understand the changes in cognitive performance in relation to the EEG signal complexity decrease during hypoglycemia. One of the aims of this chapter is to evaluate the relationship between the awareness status of each subject and the results of the complexity analysis, in order to investigate if the accuracy of hypoglycemia detection is associated to different awareness status of subjects in database. Hypoglycemia awareness status was classified by 3 methods described in Section 2.1, and the result is that, in our database, there are subjects in Aware, Unaware and

Reduced awareness status, and three who did not qualify as either hypoglycemia aware or unaware, which were excluded from the analysis. Therefore, in this chapter when we refer to the database, we will consider the 31 patients qualified through the inclusion criterion of chapter 2, section 2.1. Remembering that HU is defined as the failure to sense a significant fall in blood glucose below normal levels, we want to verify if this condition has repercussions on the EEG signal and on the detection of hypoglycemia. It is interesting, for a better understanding of brain functioning in hypoglycemia, to investigate further approaches which consider several EEG channels at the same time. For this reason, we tried to apply some signal pre-processing techniques. Firstly, applying PCA and ICA that led to no viable road and then, trying to study on EEG signal the events of Desynchronization / Synchronization (ERD/ERS). This index consists of the temporary modulation of the underlying EEG oscillatory activity and reflects the desynchronization / synchronization of the activity of a population of neurons, but our data are not adapted to carry out this type of analysis.

Not having achieved good results working on the pre-processing front, we decided to investigate further the important change in the EEG signal during HYPO state compared to the EU one. The algorithms applied on the P3 EEG channel appear to be informative, in fact, differences in terms of signal complexity have been also evidenced (Chapter 3), but not such accuracy as to allow real-time application on a device such as the one devised by Hyposafe. For this reason, our analysis will not be limited to a single channel, but the same calculation made in Chapter 3 are performed on all channels.

4.2 Selection of EEG channels sensitive to hypoglycemia: single-channel evaluation

Until this moment, the analysis of complexity has been carried out only on EEG P3 channel, following the same indications found in [22]. Below, in Fig 4.2, we show the comparison of the results obtained on a random subject of the database by applying entropy measures. An evident difference between the EEG EU and HYPO signals is given by ePE, while rePE and eCE do not introduce particular relevant information. In fact, we can see how the graph representing the rePE trend (fig 4.2, (c)) coincides with that of ePE (fig 4.2, (b)), and considering the higher computational cost of rePE as introduced in section 3.4, it makes sense to delete it. Regarding eCE instead (fig 4.2, (d)), it is very difficult to make a clear distinction between the trend in the case of an euglycemic state and that in the hypoglycemic case. Instead, it is possible to notice that often they assume the same value. Taking into account two important aspects, these results are not so far from our expectations. Firstly, the available EEG signal is devoid of artifacts having been previously filtered, then it was not expected to obtain significant differences between ePE and rePE. Secondly, from the literature [50] we know that the distribution of the OPs calculated in the eCE algorithm, quantifies the average difference between an OP and the next one. For this reason, knowing the difficulty in detecting different oscillation and hidden patterns in EEG signals and in particular, the distinction between EEG epochs in EU and HYPO, we did not expect significant results through this calculation of OPs. These assessments lead us to exclude rePE and eCE from the possible real time application and therefore from the following steps of this work.

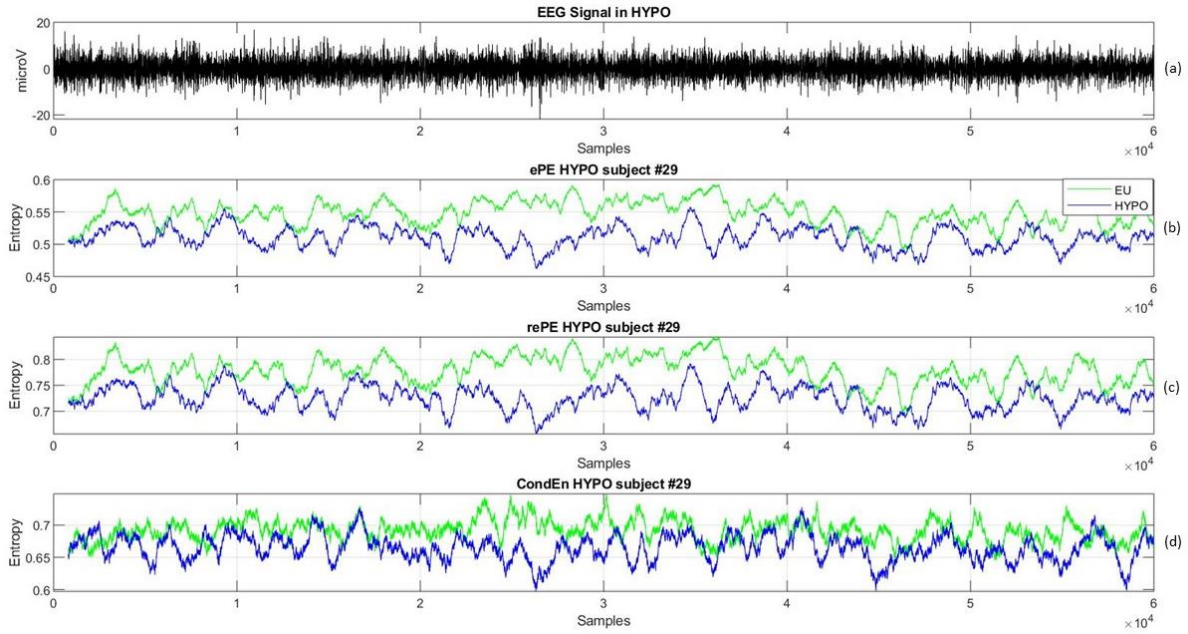

Figure 4.2: In the first panel (a) the EEG signal of random subject in HYPO state is represented. Following, the trend of the same subject in EU (green) and HYPO (blue) state, calculated by ePE (panel b), rePE (panel c) ed eCE (panel d). This representation refers to all the available EEG signal of a subject, 5 min long. It is also possible to note with a visual inspection the difference between the applied algorithms.

In the previous chapter, we also showed the results obtained from the fractal dimension, which is currently the first choice for future applications on the device, as described in section 3.4. Before extending the study of this algorithm and the ePE across all EEG channels, we have tried to study whether there were differences in outcomes between hypo-aware and unaware patients in the most significant results of complexity. As depicted in Fig. 4.3, the difference between the results of each index in EU and HYPO state computed so far has been computed, dividing the subjects in hypo-aware and unaware. It is possible to notice that in Higuchi case (panel a.1) the difference is 9.8% in hypo-aware subjects and 0.09% in hypo-unaware, in Residuals (pannel a.2), 82.4% in hypo-aware patients and 15% in hypo-unaware, finally, in Tortuosity (panel a.3) there is no difference in percentage in hypo-aware subjects and 3.1% in hypo-unaware. Looking at the following panel, in the case of ePE (panel b.1), the discrepancy is 20.7% in the first case and 0.2% in the second, instead in the eCE panel (b.2), there is a widely difference between the two conditions: 23,5% against 0,9%. Then, we can conclude this section by saying that the outcomes, obtained by applying complexity algorithms only on the P3 EEG channel, detect an important difference between hypo-aware and unaware patients mostly in entropic indices. Fig 4.3 shows that the decrease in complexity of the hypo-aware EEG signal is more evident than hypo-unaware, but not to the point of causing repercussions on the non-linear analysis of the EEG and on the detection of hypoglycemia.

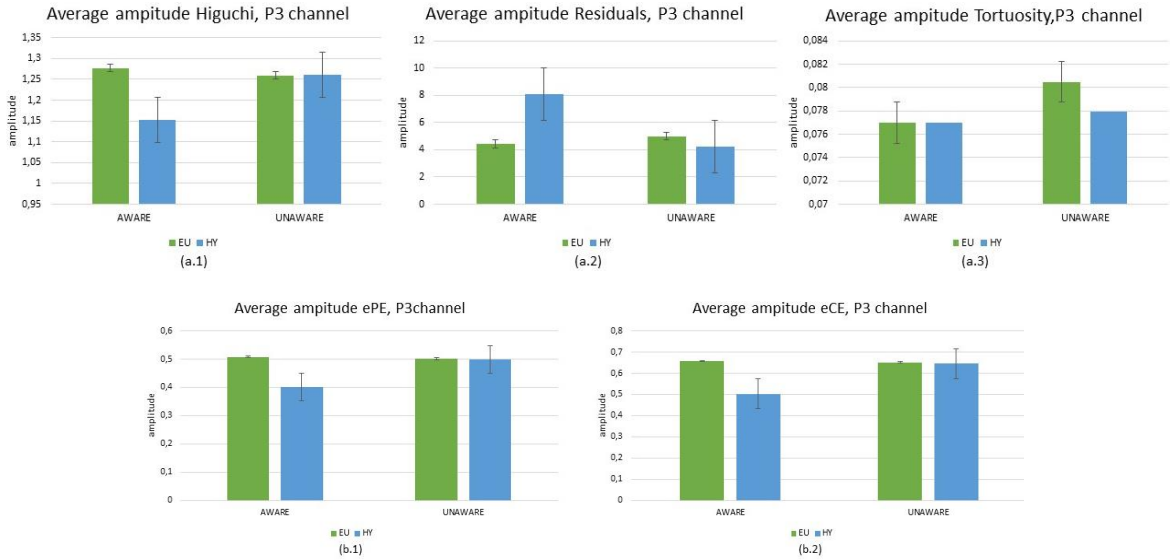

Figure 4.3: In panel (a), average ampitude on P3 channel of Higuchi, Residuals and Tortuosity in EU and HYPO in the patients with hypo-awareness versus the patients with hypo-unawareness. In panel (b), results of ePE and eCE in EU and HYPO in the patients with hypo-awareness versus the patients with hypo-unawareness.

4.3 Multi-channel evaluation

The studies carried out so far, allow us to detect hypoglycemia from an electroencephalographic signal with almost 60% of accuracy. In order to try to raise this percentage, our work has been focused on a wider evaluation of the available signal, not limited to one channel only. In fact, representative randomly chosen 2-s EEG sweeps are depicted in Fig. 4.4 to appreciate in all channels, albeit qualitatively, differences of the EEG signal in EU (in panel (a)) and HYPO (in panel (b)). As in Fig. 1.6, the EEG in HYPO shows higher amplitudes in the low frequency bands and greater regularity than the one in EU. This first visual ispection, seems to show a variation in the regularity of the signal in all EEG channels to such a point that it makes sense to carry out a new analysis of the complexity on all 19 channels. We expect not only information consistent with those obtained from the P3 channel but also more exhaustive, with the aim of being able to make the detection of HYPO as precise as possible. Indeed, as we can see in the topographies below in Fig 4.5, the result is very interesting. Complexity analysis is performed on all the 19 EEG channels. The topographies represent the highest frequency density of all subjects on each channel. The eCE analysis is not present for the reasons described in previous section. In Fig 4.5, panel a, the topography of Higuchi values is represented. It shows a statistically significant decrease in its amplitude during hypoglycemia compared to euglycemia in all electrodes (p-value < 1.5e-06) and in a more specific way, in the brain area covered by T5, P3, Pz, P4, O1 and O2 channel (in parietal and occipital lobes). In the topography showing the values of the features Residuals (panel b), as already anticipated by the analysis performed on the P3 channel, the results are not so clear as to show a significant difference between euglycemic and hypoglycemic states, as it was also confirmed by the statistical tests (p-value > 0.52). The highest frequency density in Tortuosity is shown in panel c, in which it is possible to see results consistent with what has been achieved up to now, remembering that, as for Residuals, the amplitude of the tortuosity increases with decreasing signal complexity. The result is

a statistically significant increase of amplitude feature when passing from EU to HYPO in parietal and occipital lobes (p-value < 4e-05) and in a more specific way in Cz, T5, P3, Pz, P4, O1 and O2 channels. Also the last topography, relative to the highest frequency density of ePE (panel d), shows a general decrease in the complexity of the EEG signal in HYPO, statistically significant (p-value <0.0082) on all EEG channels and with particular attention to the parietal and occipital lobe. As we hoped, a visual inspection by topography of all EEG channels returned very relevant information: we can easily note, except for the Residuals feature, that the lower channels (temporal lobe, parietal and occipital) show a greater discrepancy between the EU and HYPO signal compared to the frontal lobe channels.

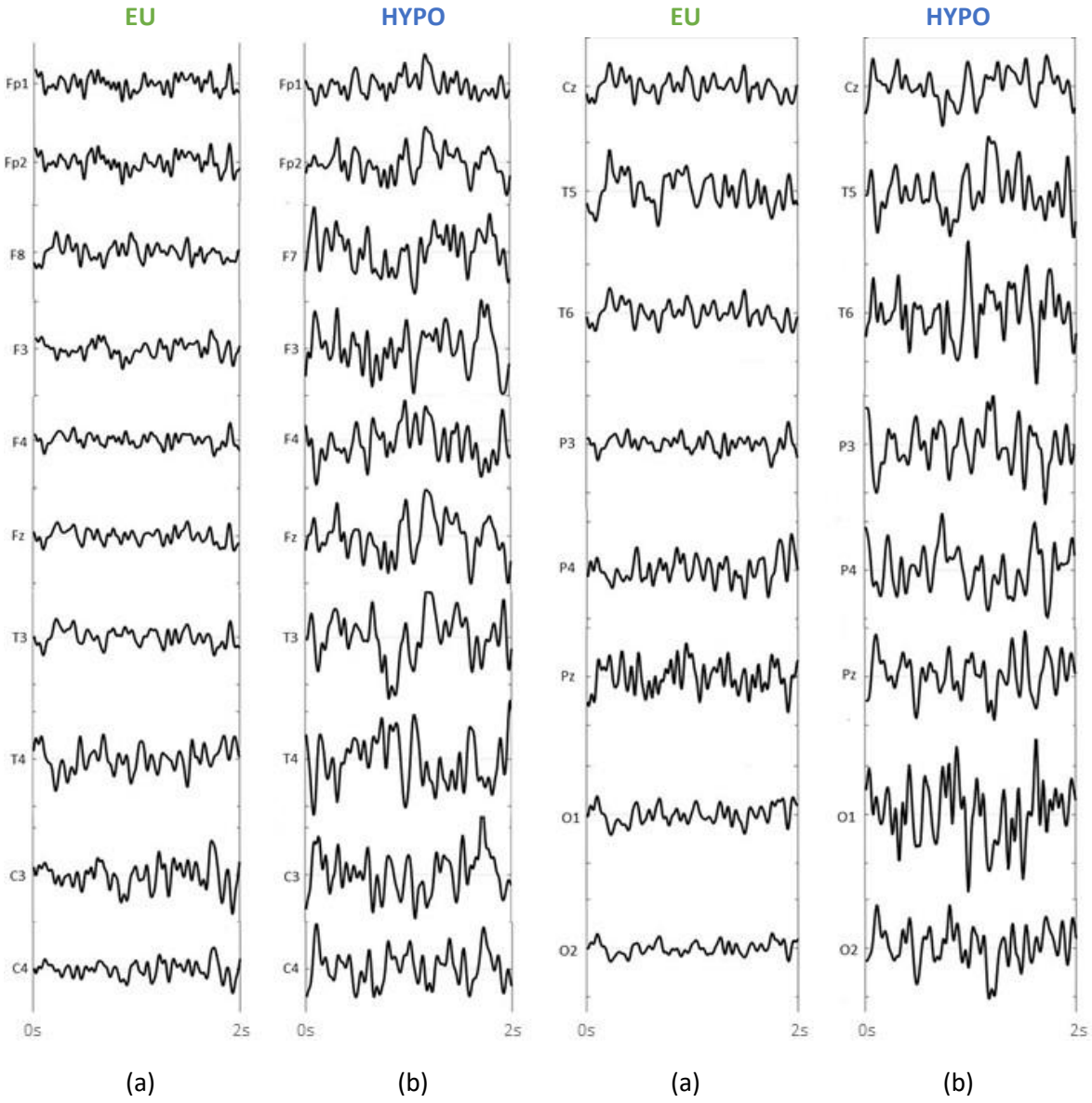

Figure 4.4: 2-s EEG signals in time domain during (a) EU and (b) HYPO in a representative subject.

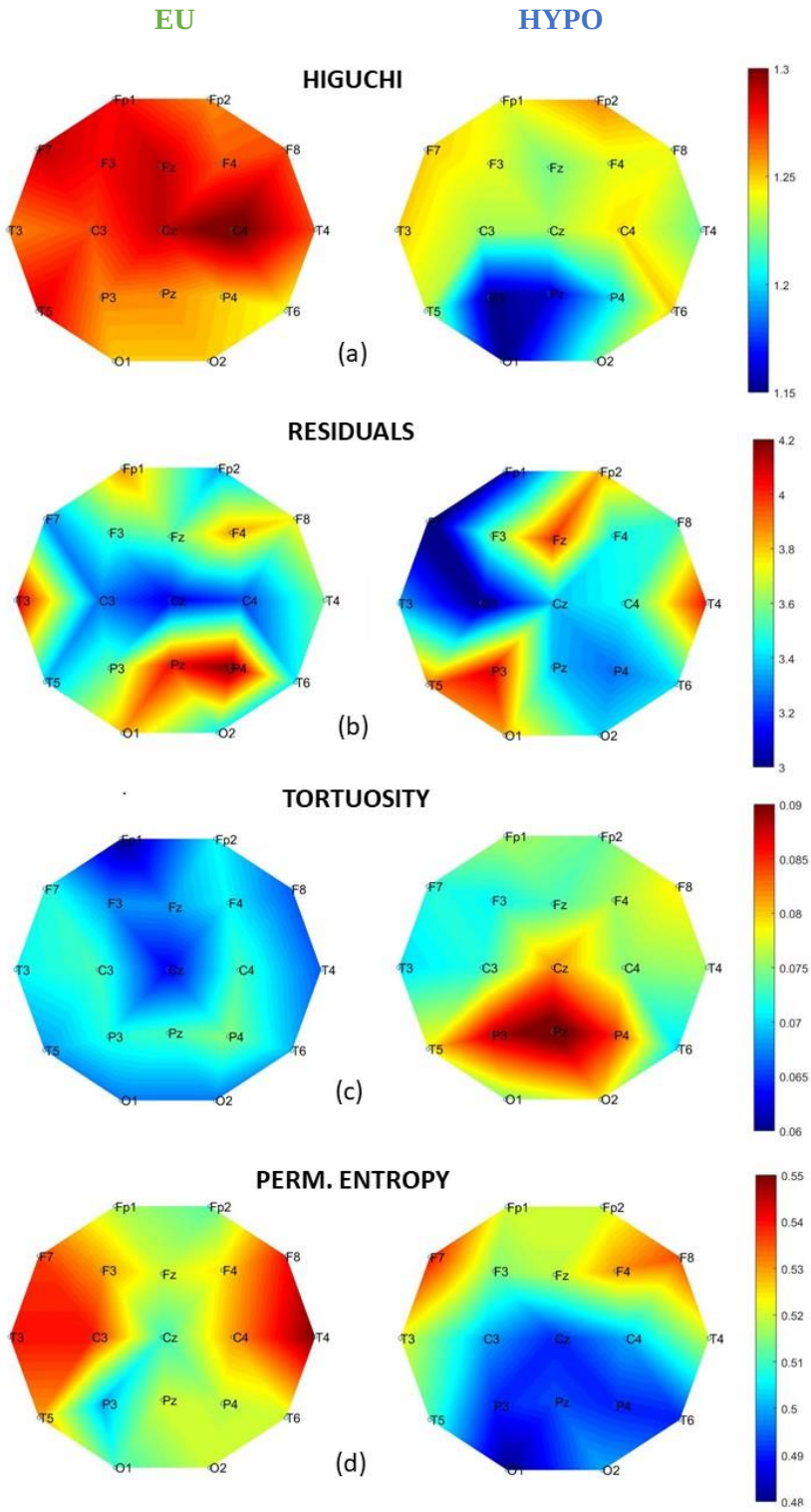

Figure 4.5: Topographies of the highest frequency density in EU and HYPO were calculated on average across all T1D patients in the most relevant algorithms. During HYPO, there is: (a) significant difference in amplitudes with respect to EU in Higuchi's feature (p -value < 1.5e-06); (b) not a significant increase in Residuals feature in all electrodes; (c) a significant increase in the amplitude of the Tortuosity in parietal and occipital lobes (p -value < 4e-05); (d) a significant decrease in ePE values in occipital lobes (p-value < 0.0082).

In topographies, we have compared the results between eu- and hypoglycemic states obtained from each complexity algorithms, computed on each single channel. Now we will use the visual results obtained through topographies and try to identify the most significant channels. In Fig 4.6, each boxplot represents the average signal calculated on all the subjects, for each EEG channel. The boxplot graphs are derived for each complexity indices, in all 19 EEG channels. We will analyse the discrepancy between EU and HYPO in the different channels. The difference between the EEG channels in the frontal lobe and the channels in the temporal and parietal lobes is immediately evident. In a and c panels, representing the Higuchi's feature and the ePE index, the first channels show a slight difference between the two glycemic conditions, while the remaining channels show clearly lower values in HYPO state than in the EU. We notice that starting from the Fp1 channel, that is located on the frontal lobe, up to the T6, positioned on the temporal lobe, the boxplots that represent the amplitudes in EU and HYPO are very close, even overlapped in some channels, to the point of not distinguishing the two glycemic states; from the T6 channel instead it is possible to highlight a decrease in the hypoglycemic boxplot values up to the minimum in the Pz channel. In b and d panels, which represent the Residuals and Tortuosity features, we have noticed in general that the disparity between EU and HYPO status is not as clear as in the previous ones, also remembering that for these two features the value of the amplitude increases with the decrease of complexity. In detail, the Residuals feature returns non-informative values, as shown by previous topographies, it is not possible to clearly distinguish the eu- from hypoglycemic state. In panel d, the results obtained by Tortuosity reflect those obtained by Higuchi and ePE. In fact, it is possible to notice an increase in the values starting from the T6 channel until obtaining the P4 matrix, a channel placed on the parietal lobe. Quantitatively, the values between the EU and HYPO signals in the channels affecting the parietal, occipital and temporal lobes differ by 5% compared to the others. In particular, the most significant channels are: Cz, T5, T6, P3, P4, Pz, O1 and O2. From this evaluation, we can confirm that more than one channel is potentially informative for our study, that is, more precisely chosen channels can provide better detection of the HYPO status in a T1D subject.

As in single-channel case, also in multi-channels evaluation we computed investigated if there were differences in the results between hypo-aware and unaware patients in signal complexity. In fig 4.7 the difference between the hypo-aware and unaware subjects in the same way of the previous section was calculated. It is possible to notice that in Higuchi case (panel a.1) the difference is 15.8% in hypo-aware subjects and 4% in hypo-unaware, in Residuals (pannel a.2), 8.6% in hypo-aware patients and 3.3% in hypo-unaware, finally, in Tortuosity (panel a.3) there is no difference in percentage in hypo-aware subjects and 7.2% in hypo-unaware. Looking at the following panel, in the case of ePE (panel b.1), the discrepancy is 40.5% in the first case and 2.6% in the second, instead in the eCE panel (b.2), the difference between the two conditions are 47.2% and 1.2%. In conclusion, the outcomes obtained by applying complexity algorithms on all EEG channels, confirm the difference between hypo-aware and unaware patients, mostly in entropic indices. Fig 4.7 demonstrates that the decrease in complexity of the hypo-aware EEG signal is more evident than hypo-unaware, but not to the point of causing repercussions on the non-linear analysis of the EEG and on the detection of hypoglycemia.

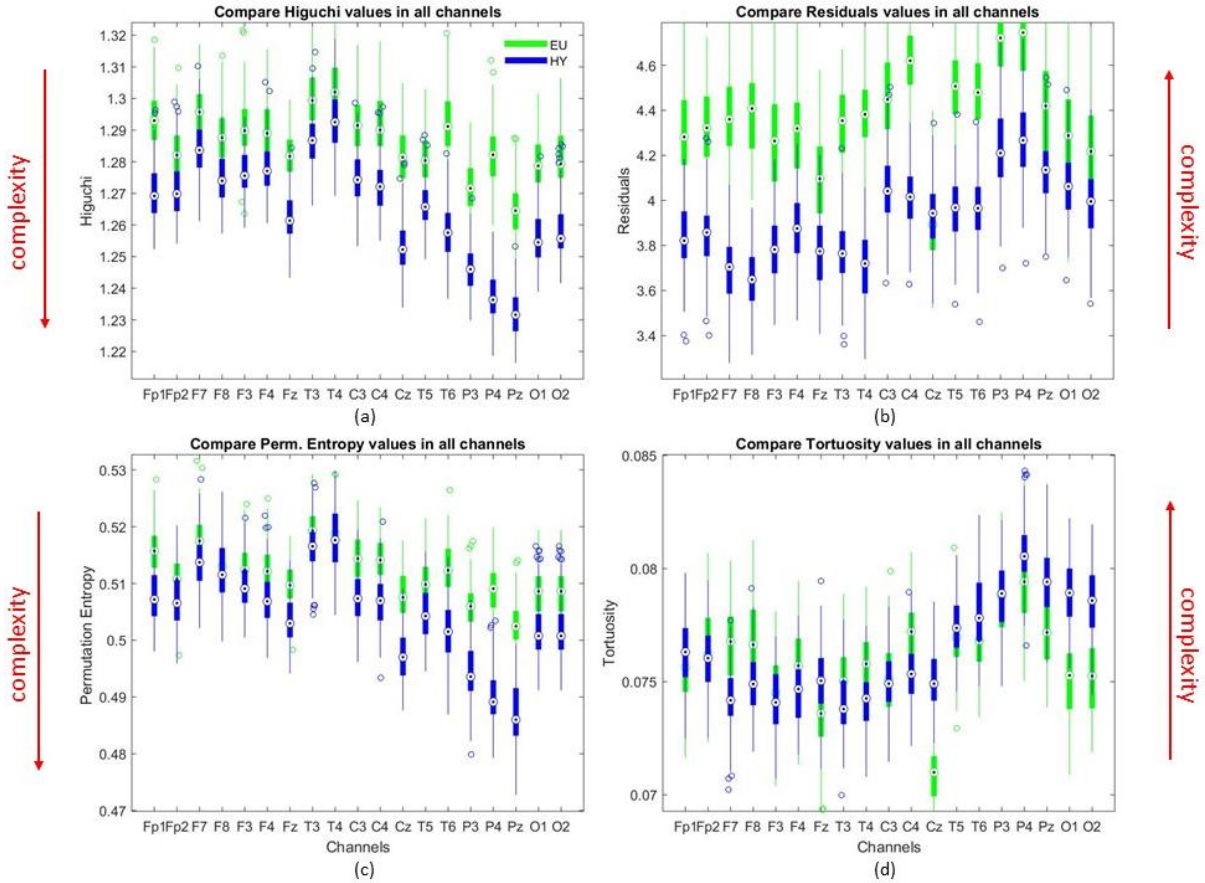

Figure 4.6: Representing the boxplot for values of complexity indices in all EEG channels by averaging all subjects. In Higuchi, panel a, and ePE, panel c, a decrease in the regularity of the signal already starting from the Cz channel is evident. In Residuals, panel b, whose amplitude value increases with the decrease of complexity, there are no significant results, while in Tortuosity, panel d, an increase in the width of the features can be seen starting from the T6 channel. Boxplots represent the average of all subjects calculated on the signal coming out of each channel. Data are mean ± sd values. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. Circle on the top of the bars indicate statistically significant differences (p-value < 0,05) between entropy indices values at the considered channel.

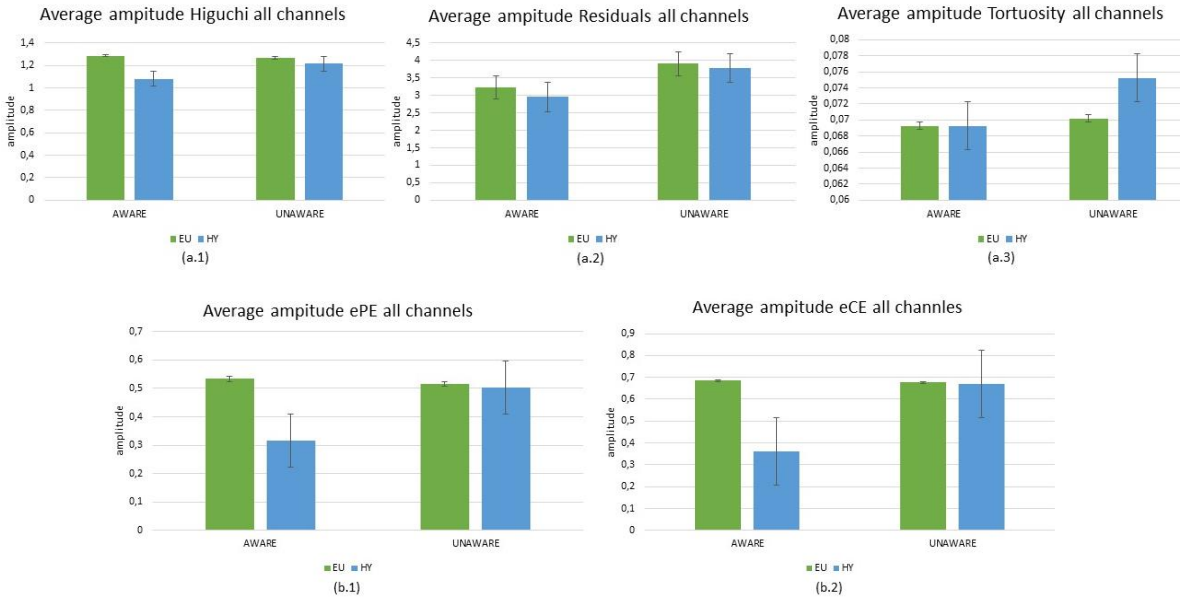

Figure 4.7: In panel (a), average ampitude on all channels of Higuchi, Residuals and Tortuosity in EU and HYPO in the patients with hypo-awareness versus the patients with hypo-unawareness. In panel (b), results of ePE and eCE in EU and HYPO in the patients with hypo-awareness versus the patients with hypo-unawareness. All the analysis was camputed on all EEG channels

4.4 Conclusions

The present chapter investigated if algorithms based on Higuchi's fractal dimension and empirical permutation entropy can reveal changes in the complexity of the EEG signal induced by HYPO. We checked the interaction of the results presented previously with hypo-aware or not of subjects. Using the P3 EEG channel data, the proposed algorithms could already be used to assess changes in EEG complexity which occur during HYPO, likely related to progressive loss of cognitive function and altered cerebral activity. rePE and eCE do not return additional information with respect to ePE. Analysing the P3 channel, we emphasize a distinction in the evaluation of the HYPO between subjects that are in an aware or unaware state, with an average difference of about 10%. Repeating the same analysis, using all the available EEG channels, the difference in the assessment among subjects in awareness status turned out to be even more pronounced, reaching about 15% for the Higuchi features and 40% for the ePE algorithm. We can not explain exactly why there is so much difference between the resulting percentage of the algorithms. This could be an interesting starting point for future investigations. Furthermore, we showed that multi-channel analysis leads to a classification of hypoglycemia with greater accuracy than that obtained by analysing only the P3 channel. Indeed, a decrease in the complexity of the EEG signal in HYPO condition was significantly found in 7 EEG channels: T5, T6, P3, P4, Pz, O1 and O2. Subsequently, we may consider narrowing the analysis to the 7 channels mentioned above and classifying hypoglycemia from information provided by complexity indices in these revealing channels.

Chapter 5

Classification of hypoglycemia through Neural Networks

5.1 The use of Neural Networks to classify hypoglycemia

The work carried out so far has expanded the knowledge on the relationship between hypoglycemia and electroencephalographic signal. The effect of hypoglycemia is shown, through brain topographies, on the outgoing signal, in particular from 7 EEG channels (T5, T6, P3, P4, Pz, O1 and O2). The analysis was performed using two types of indices, fractal features and empirical permutation entropy. All indices revealed analogue results, and the usage of these non-linear methods is promising for possible future applications of real-time EEG analysis, given the non-stationary nature of the signal. We still need to investigate if one of the two types of complexity index is more informative than the other, and if it is possible to classify with good accuracy HYPO and EU. For these reasons we have introduced in our work an analysis based on Neural Networks (NN).

A NN exploits a dense network of simple units that are strongly interconnected, also known as neurons Fig 5.1, with the aim of emulating the ways in which the brain performs certain functions. The feedforward NN are the most used, both thanks to their simplicity and their versatility. They are often used for voice and optical recognition, for classification, for identification and control of systems, for prediction and for the approximation of functions [61]. This is not the first time they have been used in the biomedical field and applied to data from diabetic patients. For example, in [62], NN were used to diagnose Alzheimer's disease and the network was able to distinguish between male and female Alzheimer's. In [63], they attempted to identify type 2 diabetes from data collected from patients' daily habits. One of the NN main fields of applications is the prediction of time series, for example, in [64], [65] and [66] they were used for the prediction of diabetes with sensors for CGM. What makes them so attractive and versatile is their ability to learn for examples the relationship that connects entry and output, as well as the fact that they do not need to define a model that describes the data, which requires a priori assumptions on the process to be analysed. Another appreciable property is their ability to generalize and produce correct results even if they have data that have never been seen before and partially corrupted by noise. Finally, NN are, by their nature, non-linear (just use non-linear activation functions) and, therefore, particularly suitable for describing non-linear components of the time series to be predicted. These just described are the non-negligible qualities of the NN, but they also have limits to be considered. One of the biggest problems is the absence of theoretical results on the optimal network structure to be used. Consequently, it is necessary to make attempts, testing various configurations and choosing the best one. In this chapter we will apply the NN to the results obtained using the fractal domain and ePE algorithms with the aims of detecting HYPO with a high percentage of accuracy, and to evaluate if between the two algorithms used so far it is recognized by the network which is the most informative. We are opening the doors to a new study to classify Euglycemia and Hypoglycemia with a high accuracy and with as few EEG channels as possible.

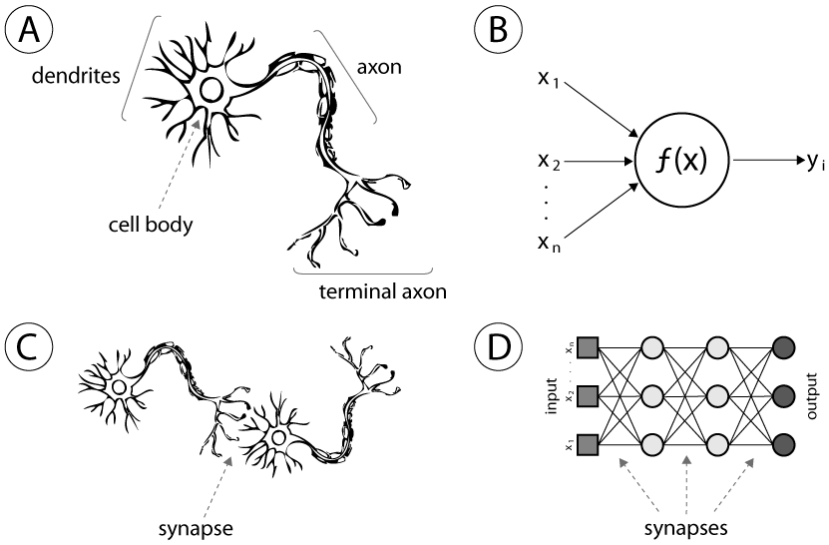

Figure 5.1: Biological neuron on the left (A) and biological connection, called synapse (C) and artificial neuron on the right (B) and artificial connection of neurons bottom right (D). Representing the architecture of artificial neuron inspired from biological in way that is function is to calculate an output value based on inputs and interconnection weights; adjust interconnection weights specify the flow of information.

In particular, a FFNN network was applied on a dataset composed by the complexity indices carried out in Chapter 3 and in Chapter 4. In general, it can be considered that the parameters to be determined in defining the network architecture are:

- the subdivision of the database;
- the number of hidden layers and neurons to be inserted in each layer;
- the connective mechanisms between the different layers;
- the activation function;
- the learning rules;
- updating the connection weights of neurons.

Some of these points have been described in the previous section, in the following we will explain the choice of features for the creation of the dataset and the subsequent definition of the neural network architecture for this first classification.

5.2 Dataset creation

For the creation of the dataset, the results obtained from the Multi-channel evaluation in paragraph 4.3 were used. In particular, we considered the complexity analysis performed with the ePE indices and with the fractal dimension on the 7 most significant EEG channels. For each channel, we derived a matrix that contains the values of each complexity index relating to EU and HYPO, for each 4-s segment, as shown in Fig. 5.2. Consequently, a matrix is derived for each channel and for each complexity index (Higuchi, Residuals, Tortuosity, ePE and eCE). Subsequently, each matrix was

transformed into a vector 1x (148x34), to obtain a resulting vector containing in sequence the values of the respective index of all 34 subjects. For example, the Higuchi vector, related to EU values of channel 13, is built as shown in Fig 5.3.

HIGUCHI Matrix, Hypoglycemia, Channel 13 (T5):

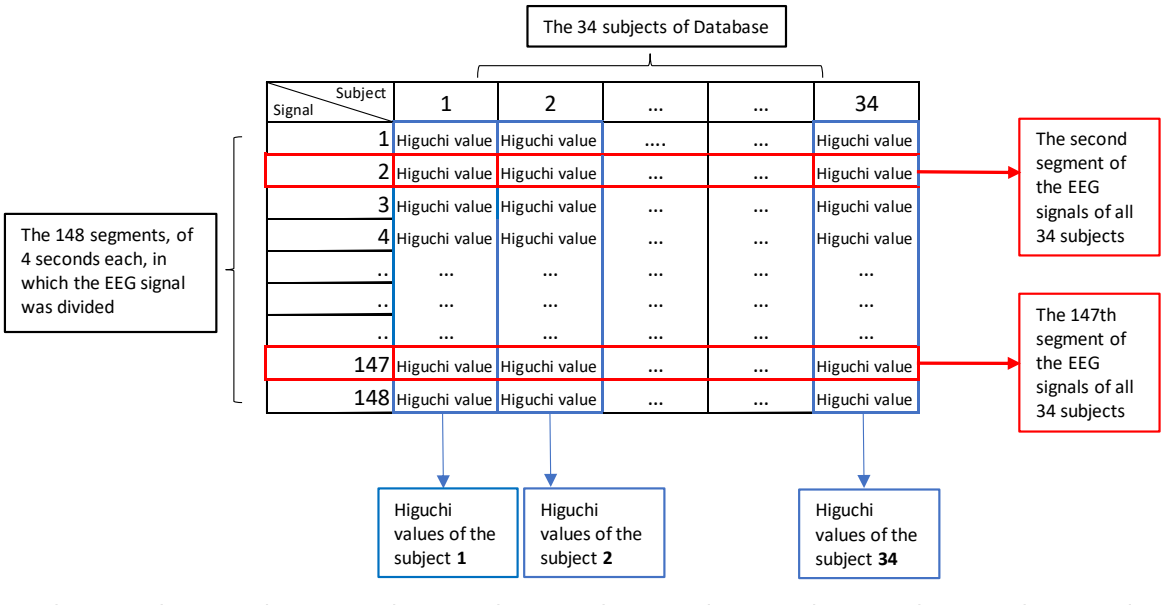

Figure 5.2: Representative scheme of the outgoing matrix from the analysis of complexity. In this example a Higuchi matrix in hypoglycemia of Channel 13 is shown, but the other indices have the same configuration. It is a 148x34 matrix and it contains on each column the Higuchi values related to a single EEG signal of a patient, and on each row the Higuchi values related to the same segment of the EEG signals of all 34 patients. The Higuchi matrix relating to the euglycemic EEG signal values of channel 13 (T5) has the same identical setting.

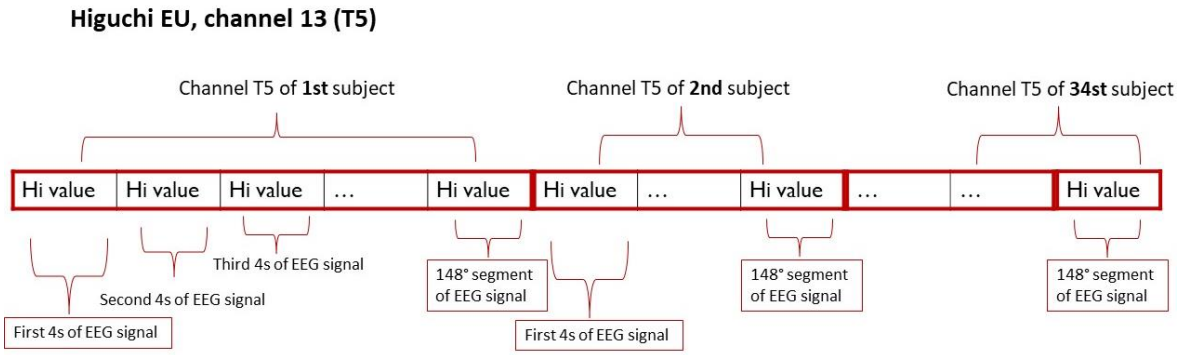

Figure 5.3: Representation of a vector necessary to create the dataset. It is a 1x5032 vector that contains in sequence the Higuchi values related to the euglycemic EEG signal of a channel, of all 34 subjects. The matrices of the remaining complexity indices have undergone the same setting variation, becoming vectors1x5032.

Finally we can set the features matrix that describe our dataset. The resulting matrix has dimensions 28 x 10.064, each row contains the values of an index of complexity of a given channel, first relating

to the EU signal and then to the HY signal. Since the multi-channel chapter showed that the most informative channels for all applied complexity indexes are 7, the features matrix shown in the following table contains the complexity values related to the selected channels: T5, T6, P3, P4, Pz, O1 e O2.

Higuchi EU channel 13 sub #1	Higuchi EU channel 13 sub #34	Higuchi HY channel 13 sub #1	Higuchi HY channel 13 sub #34
Residuals EU channel 13 sub #1	Residuals EU channel 13 sub #34	Residuals HY channel 13 sub #1	Residuals HY channel 13 sub #34
Tortuosity EU channel 13 sub #1	Tortuosity EU channel 13 sub #34	Tortuosity HY channel 13 sub #1	Tortuosity HY channel 13 sub #34
P. Entropy EU channel 13 sub #1	P. Entropy EU channel 13 sub #34	P. Entropy HY channel 13 sub #1	P. Entropy HY channel 13 sub #34
Higuchi EU channel 14 sub #1	Higuchi EU channel 14 sub #34	Higuchi HY channel 14 sub #1	Higuchi HY channel 14 sub #34
⋮	⋮	⋮	⋮
P. Entropy EU channel 19 sub #1	P. Entropy EU channel 19 sub #34	P. Entropy HY channel 19 sub #1	P. Entropy HY channel 19 sub #34

Table 5: Final representation of the dataset constructed for this classification.

Once the contents of the dataset are defined, the desired output must be specified. The purpose of this analysis is to classify EU and HYPO from the values obtained from the complexity analysis. For this reason, we have defined a "target" vector. Its size is 1x10.064 and it identifies with a value of 1 the output "EU" and with –1 the output "HYPO". The choice of output values is a consequence of the setting of the FFNN which presents a single layer of hidden neurons with sigmoid activation function with computed values between –1 and 1. Once the form of the output and the content of the dataset from which to extract the input variables are defined, the dataset is divided into sub-periods, which delimit the learning (training set) and evaluation (validation set and test set). The network learns by trying to recognize the dynamics of the training set, verifies how it fits on the validation set and then applies it to a set of data, test sets, which it has never seen. Given the stochastic nature of neural networks, a subdivision method called cross validation was used. In this method all available data are used, in fixed size groups, alternatively as test and training set. Therefore, each pattern is both tested at least once, and used for training. The criteria for the extraction of the subperiods, called N-folds, are very varied and in our case, it is random, i.e. a percentage proportion of the two N-folds was defined based on what was found in the literature in the various applications of NN. In detail, the previously constructed dataset was randomly divided into two N-folds, one representing the train set and the other the test set. The training of the NN and the respective evaluation was repeated 1000 times, and each time 75% of the dataset was selected as a train and the remaining 25% as a test set. In this way the neural network has been trained on a train set, and subsequently classified on a test set; those sets are always different for each analysis performed. Performance as an average was evaluated between the performances obtained on each of the resulting groups. The performances obtained depend however on the division into groups. Therefore, the cross validation has been repeated several times changing the train and test percentages, so as to become independent of the particular division into groups. A further attempt was made by applying the Leave One Out Cross Validation (LOOCV) procedure. In this case the test set is composed of the values relating to a single subject, while the training of the network is performed on the remaining subjects, in our case, on 33 subjects. This procedure is repeated as many times as the subjects making up the dataset, each time choosing a different subject for the test set. The LOOCV has some advantages over other validation procedures but in the case of our dataset, it was excluded from the study because the results obtained are unsatisfactory. Probably the 34 subjects available are not enough to carry out a good training of the network, and to model the high inter-subjects variability typical of EEG signal.

5.3 Design the network

NN design is very important, as it influences the ability of the neural network to validate output. A two-layer FFNN was created using the neural network toolbox from MATLAB software. For a simple feedforward neural network, trained with the back-propagation algorithm, the possible variables are the number of hidden layers and the number of neurons in each layer. In [65], authors confirmed that a neural network with only one hidden layer can accurately approximate any function, if it has enough hidden neurons with a nonlinear activation function. For some functions, however, the number of needed neurons could be very large or even infinite and a NN with two hidden layers is usually much more adequate. Generally, the number of layers and the number of neurons in each layer are set according to a trial and error procedure, i.e. by testing various possible configurations and choosing the one that produces the best result. This procedure undoubtedly leads to a sub-optimal solution, since it is not possible to test all the imaginable architectures. On the other hand, since only empirical rules are present for the construction of a FFNN, this is the only way forward.

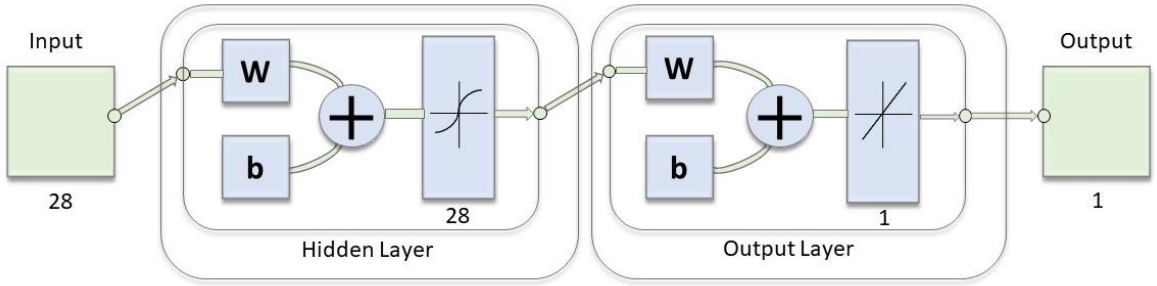

Figure 5.4: Graphic representation of FFNN implemented using the neural network toolbox from MATLAB software. It is a single hidden layer with a hyperbolic tangent sigmoid transfer function for the hidden and linear layer for output neurons.

The NN architecture used in this study is a single hidden layer, with a hyperbolic tangent sigmoid transfer function for the hidden layer and a linear one for output neurons. Various techniques exist to determine the number of units appropriate for a NN, among them the most used methods are the "growing" and "pruning". They themselves, respectively, increase and decrease the number of neurons and connections dynamically, during network training. The method used in this paper is a statistical technique instead. It consists in the subsequent training of networks with an increasing number of neurons and in the recording of the performances of each one. The chosen network was the smaller one that provided a significant improvement in performance compared to simpler architectures. In detail, it has several neurons for the input layer equal to the number of features, i.e. 28 (4 complexity indices for each of the 7 EEG channels), while for the output layer the number of neurons is equal to the number of classes. The network was trained with Levenberg-Marquardt back propagation algorithm. Training automatically stops when validation stops improving, as indicated by an increase in the mean square error of the validation samples. The magnitude of the gradient and the number of validation checks are used to terminate the training. The gradient will become very small as the training reaches a minimum of the performance. If the magnitude of the gradient is less than a minimum value imposed in the software, by default $1e^{-7}$, the training will stop. The number of validation checks represents the number of successive iterations that the validation performance fails to decrease. If this number reaches 6 (default value), the training will stop. We have changed

this criterion by setting the parameter of validation checks at 20. This change allowed our optimization not to terminate in the case is hits a small "plateau".

5.4 Results

The results of applying the ANN methodology to our database showed very good abilities of the network to classify the correct glycemic status of the patients from the EEG signal complexity indices. The network was trained and tested 1000 times, each with a different testing set (i.e. cases the network has not seen before), as specified in section 5.2. The NN was able to classify each 4-s segment with high accuracy, indeed it correctly classify, each time, about 90% of the testing set. The network trained with Levenberg-Marquardt algorithm stops when validation stops improving, as indicated by an increase in the mean square error of the validation samples. Indeed, the mean square error (MSE), that is the average squared difference between outputs and targets, reaches its best validation performance at a value of 0.41601 at epoch 23. For this parameter, lower values are better and zero means no error. Another parameter that allows to evaluate if the network has been well trained is Regression (R) Values measure which represents the correlation between outputs and targets. An R value of 1 means a close relationship, 0 a random relationship. The results of regression are reported for training, validation, testing and all of them together, whose values are: 0.83019 for training, 0.76895 for validation, 0.74625 for test and finally 0.80759 for all together. In an ideal situation it is expected that a classification method will be able to perfectly discriminate two populations (in our case, EU and HYPO) that cannot be superimposed (mutually exclusive), where the ‘cut off’ represents the threshold value of the method. It usually happens that the two populations overlap in part, and the method necessarily identifies as positive some values that are negative (False Positive) and as negative some values that are positive (False Negatives). In order to evaluate the classification provided by the neural network, we defined as positive the hypoglycemic state, and as negative the euglycemic state. Consequently, we defined True Positives, True Negatives, False Positives and False Negatives as followed:

- True Positive (TP), a 4-s segment related to hypoglycemic state, classified by the NN as HYPO;
- True Negative (TN), a 4-s segment related to euglycemic state, classified by the NN as EU;
- False Positive (FP), a 4-s segment related to euglycemic state, classified by the NN as HYPO;
- False Negative (FN), a 4-s segment related to hypoglycemic state, classified by the NN as EU;

The final classification of each 4-s segment is achieved through a threshold. If the threshold used to define a segment positive or negative is shifted to the right or left (i.e. raised or lowered) it will have a greater or lesser risk than FP and FN. This will affect the sensitivity and specificity of the method itself, where sensitivity refers to the proportion of values with a positive condition that were correctly classified and it is calculated with equation (5.1), and specificity refers to the proportion of values with a negative condition that were correctly classified, equation (5.2).

$$\text{Sensitivity} = TP/(TP+FN); \tag{5.1}$$

$$\text{Specificity} = TN/(TN+FP); \tag{5.2}$$

A high cut off will allow to correctly identify the majority of EU segments, giving the method a high specificity (therefore few FP), but it will underestimate the proportion of HYPOs, giving the method a low sensitivity (therefore many FN). A low cut off on the contrary will allow to correctly identify the majority of HYPOs, giving the method a high sensitivity (few FN), but it will underestimate the proportion of EU people, giving the method a low specificity (many FP). In our work, we have varied the cut off between the values [0, 1] and selected the one that returns a better accuracy. This parameter calculates the average performance of the FFNN based on the difference between the classification calculated and the real classification, as shown in equation (5.3).

$$\text{Accuracy} = (TP+TN) /N; \tag{5.3}$$

$$\text{Precision} = TP/(TP+FP); \tag{5.4}$$

It does not optimize the network, but it returns this value in the output data, offering the possibility of selecting the architecture that has the best performance [71]. Instead, precision indicates how close or how repeatable the results are, as shown in equation (5.4). It is possible to say that it is a measure of reliability and consistency. The results are summarized in the following Table 6.

Definition	Train	Test
Specificity	0.9214	0.8852
Sensitivity	0.9081	0.8685
Accuracy	0.9147	0.8768
Precision	0.9071	0.8755

Table 6: The obtained results with 28 features and the FFNN described above.

The results immediately showed excellent network capabilities. With 7 EEG channels out of 19 it is possible to classify with an accuracy of about 88% EU and HYPO starting from the values resulting from the complexity analysis. This result leads us to a more careful evaluation of the complexity indices calculated in the previous chapters. In particular, we investigated if their information content can be considered complementary and therefore if each index contributes differently to the achievement of such a high percentage of accuracy. Indeed, it is important to avoid the overfitting of data and consequently to eliminate the features, in this case the indices of entropy, not indispensable for achieving the result. First of all, by taking each index separately, Higuchi, Residuals, Tortuosity and ePE, the NN achieved an accuracy of about 55%. This result confirms what found in the previous chapters: the indices taken individually denote differences between the two glycemic states, but there is a remarkable overlap. In a second attempt, it turned out that the features of the fractal domain must be all three included in the NN training for the classification to be accurate. The result obtained by considering the three fractal features but not the features related to ePE is quite interesting. Indeed, ePE was found to offer a negligible contribution on the accuracy obtained in Table 6. Using the same network, with the same architecture and simply reducing the number of features from 28 to 21, due to the elimination of the ePE value of each of the 7 channels, the result varied less than 1%. The obtained results with 21 features and the FFNN described above are summarized in the following table.

Definition	Train	Test
Specificity	0.9070	0.8753
Sensitivity	0.8911	0.8580
Accuracy	0.8990	0.8666
Precision	0.9064	0.8729

Table 7: The obtained results with 21 features and the FFNN described above.

5.5 Conclusions

We can conclude by saying that although the indices of the fractal domain calculate the complexity of the signal with an implementation different from ePE and the calculation of the OPs, the results are almost comparable. We found analogies between fractal and entropy indices in the computational cost, in the ability to distinguish between EU and HYPO in all EEG channels and finally in the classification performed by the NN. Carrying out further tests, we have deepened the evaluations regarding these complexity indices. Finally, we can confirm that the contribution of ePE with respect to the indices of the fractal domain is relatively low. The results obtained also suggest and encourage the use of NN for classification EU and HYPO starting from the values obtained from the complexity analysis. A further in-depth analysis can start from the exploratory analysis of the data used in this analysis to try to further reduce the features needed for classification. Fig 5.5 shows a first analysis carried out on the 21 used features, highlighting the importance of every of them to the achievement of that 90% accuracy previously obtained. It is immediately noticeable that not all features are indispensable for achieving good accuracy.

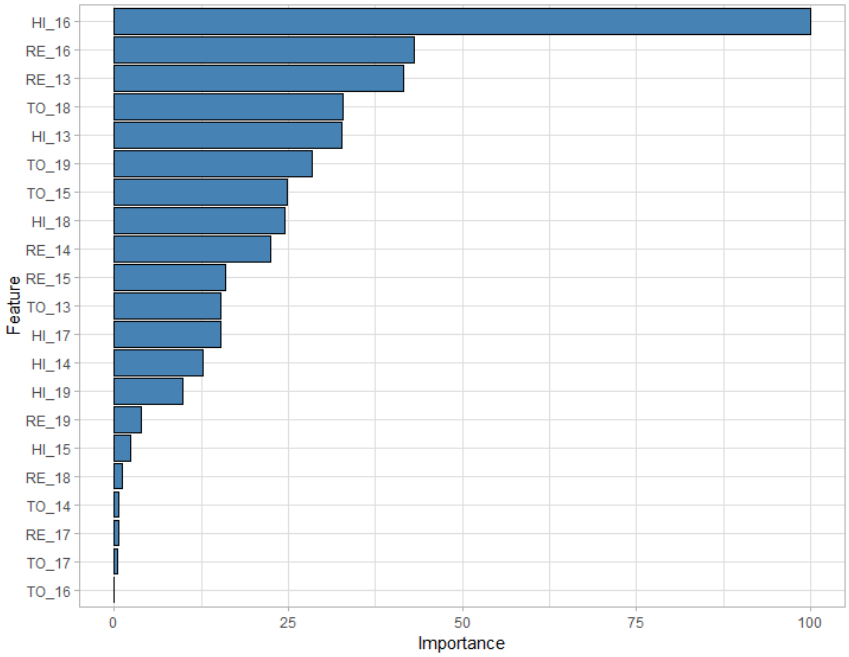

Figure 5.5: Exploratory analysis of the 21 features used for NN fillings. On the ordinate axis the features used are listed, on the horizontal axis the percentage of importance.

Chapter 6

Conclusions

6.1 Main achievements and discussion

Hypoglycemic events in patients affected by T1D were proved to be associated with measurable EEG changes. It is also qualitatively observable by visual inspection that the EEG signal in the hypoglycemic condition originates from a process of higher amplitudes in the low frequency bands and reveals greater regularity (Chapter 1). In this dissertation, we proposed an overview on the effects of hypoglycemia in the brain by broaden the EEG analysis already reported in literature, in particular deepening the analysis of EEG complexity during hypoglycemia. We applied our investigation on a database of 34 hospitalized patients who underwent a hypernsulinemic-hypoglycemic clamp for about 8 h. The EEG signals were measured in parallel for all the duration of the experiment (Chapter 2). We quantitatively analysed the EEG changes due to HYPO exploiting single and multi-channels EEG algorithms. Our work was structured with the following goals:

1. Analysis of complexity by fractal features to identify the effects of HYPO in the EEG signal, confirming in our dataset what demonstrated in the literature on smaller dataset;

2. Analysis of complexity by Empirical Permutation Entropy and comparison with fractal features;

3. Multi-channel analysis of complexity and relationship between EEG quantitative features and awareness of hypoglycemia, i.e., widening of the study on all the EEG channels in order to identify the most informative and subsequent comparison of the complexity values calculated on subjects with different states of awareness;

4. Classification by Neural Networks starting from the analysis of complexity values of EU and HYPO.

To deepen on the analysis of the EEG signal on the brain HYPO conditions, we evaluated the degree of complexity taking in account only P3 channel, which is in the posterior area of the brain. In neural signals, complexity is high for systems that contain specialized and organized elements, while it is lower in a uniform or random system. Indeed, during HYPO, the EEG signal results more regular, more uniform, less complex respect to EU. In previous studies, they were able to detect hypoglycemia on a 60-minute electroencephalographic signal, which turns out to be too long for a possible future real-time application. In our work, this decrease in complexity highlighted by a significant decrease of Higuchi's fractal dimension and an increase of both the new fractal features Residuals and Tortuosity, confirms the loss in cognitive function during HYPO in 5 min of signals. Subsequently, three new entropic indices were applied to our data, ePE, rePE and eCE. Only ePE was found to be potentially applicable, showing a decrease in the regularity of the HYPO signal compared to the EU one. From a first comparison between the various applied indices, the features of the fractal dimension are more easily applicable in a real-time context (Chapter 3). In our data, this loss of cognitive function during the pathological state of HYPO is proved by the results of TMT B and Stroop tests, where patients in HYPO have a drop in the performance of all the tests. Moreover, how

the alteration of cerebral activity, which is mainly due to entering in HYPO and not to tiredness, is demonstrated by the equal performance in TMT B and Stroop test in EU before and after the hypoglycemic period. We investigated the relationship between EEG quantitative features and hypoglycemic awareness by performing a multi-channel complexity analysis. Applying the above-mentioned complexity indices on all EEG channels, we demonstrate that the decrease in complexity of the hypo-aware EEG signal is more evident than hypo-unaware of about 15%, but not to the point of causing repercussions on the non-linear analysis of the EEG and on the detection of hypoglycemia (Chapter 4). We mainly showed that multi-channel analysis leads to a classification of hypoglycemia with greater accuracy than that obtained by analysing only the P3 channel. Indeed, a decrease in the complexity of the EEG signal in HYPO condition was significantly found in 7 EEG channels: T5, T6, P3, P4, Pz, O1 and O2. Finally, considering these important results, we wanted to clarify which of the two complexity indices was more informative than the other and then if it was possible to classify HYPO and EU with the data obtained so far with good accuracy. Applying a FFNN on a dataset created for this work, we have succeeded in classifying EU and HYPO with about 90% accuracy (Chapter 5). This outcome was achieved using 21 features for neural network training, which contain the values resulting from the analysis performed through the features of the fractal dimension without ePE.

6.2 Future challenges

Analyses and results of the thesis can be further developed and interpreted applying them in extended databases. The implementation of different indices of complexity leads to an overall increase in the information content, which could be further increased by exploring new indices. Results obtained from the features of the fractal dimension by means of multi-channel analysis seem to be the most suitable for a clinically useful real-time application. In particular, the many potentials of NN emerge from the presented overview. However, the use of NN also involves problems and difficulties, for example its intrinsically stochastic nature prevents the users from completely controlling the NN behaviour and thoroughly understanding the reasons that led to a certain configuration of weights instead of another. The results obtained also suggest and encourage the use of NN for classification EU and HYPO starting from the values obtained from the complexity analysis. It could therefore be useful, first of all, to study alternative methods for NN training and to deepen the study of individual features. In such a way it is possible to achieve high accuracy with fewer data.

Previous studies reported that EEG changes in HYPO suggest a possible use of the brain as a biosensor to detect hypoglycemia in T1D by EEG monitoring through subcutaneous electrodes (with a potential lifetime of several years) and real-time data processing by means of a multi-parameter algorithm [55]. Regarding this aspect, the accuracy achieved in this study may seem far from ideal, but a further investigation of the potential usefulness of EEG in detecting hypoglycemia is still warranted. The benefit of brain as a biomarker is that the EEG changes associated to hypoglycemia are not blunted during hypoglycemia in patients hypo-unaware, and in contrast to the symptom of hypoglycemia, the counterregulatory hormonal response and skin temperature change. By using EEG signals as the measuring principle there is a direct correlation between oncoming hypoglycemia and the signal that is measured. In this way, the measurement principle will not be affected by delays and transition functions. Furthermore, a frequent calibration as in blood glucose sensors is not required to identify the characteristic hypoglycemia pattern. Authors in [55] focused on the possibility to

construct a hypoglycemia alarm system based on continuous EEG monitoring and real-time data processing by means of a multi-parameter algorithm. At the first level, there is the feature extraction. The second level consists of three blocks, each of which analyses the features to determine if there is evidence of impending hypoglycemia, deep sleep patterns, or noise contamination, respectively. At the third level, EEG epoch is labelled as consistent or not-consistent with hypoglycemia. Lastly, taking the recent history into account it is determined whether a sufficient amount of hypoglycemia evidence is present to constitute an alarm. By carrying out the assessments and making the deductions described above, our study could contribute to a possible implementation of the device. An EEG-based hypoglycemia detection device, however, still needs to be tested in a clinical setting. Certainly, fundamental issues, such as sensitivity, specificity, reliability, resolution, and influence of activity, must be carefully addressed before demonstrating a real-time usability of such device.

Appendix A

Fundamentals of Neural Networks

A.1 The concept of artificial neuron

It is a system of parallel processing with a highly distributed character, born from an analogy with the human brain, see Fig A.1.

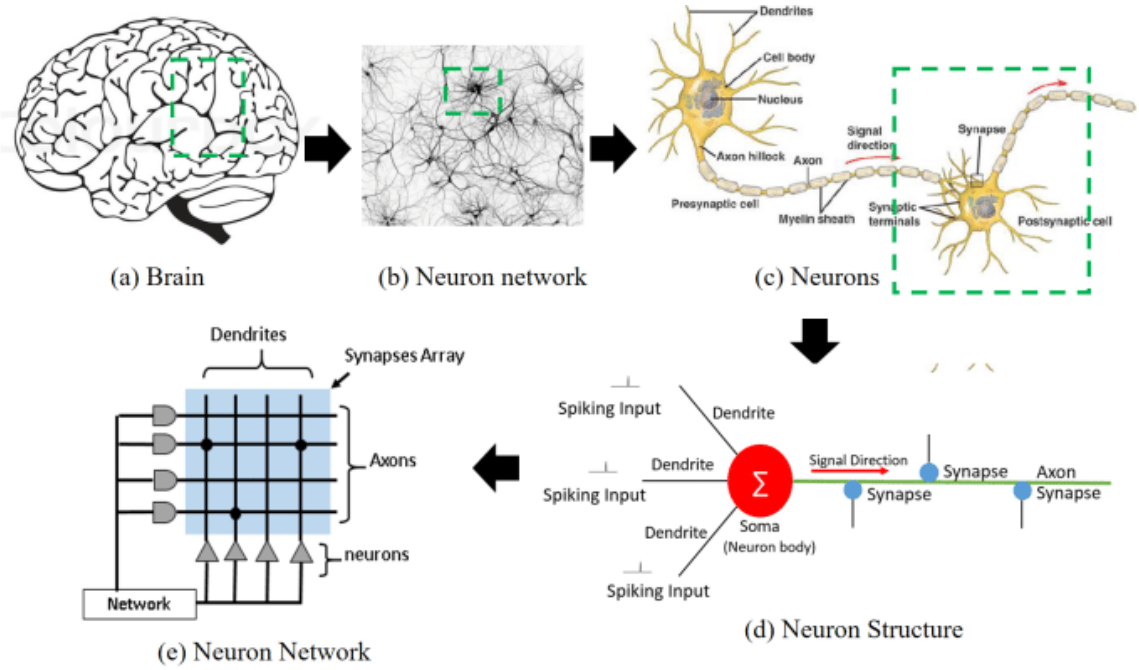

Figure A.1: The brain as a Neural Network. (a) brain; (b) biological neuron; (c) neuron connecting structure; (d) artificial neuron; (e) Neural Network architecture. Image reproduced from [60].

A.2 Networks of artificial neuron

Neural Network is a parallel and distributed calculation system with the natural propensity to store knowledge based on experience, making it available for use. The knowledge:

- is acquired from the external environment through learning.

- is stored in the forces that bind neurons, also known as synaptic weights [61].

The process through which learning takes place is called a learning or training algorithm. Its function is to neatly modify the synaptic weights of the network to complete a predetermined objective. This adjustment of the synaptic weights is the standard method by which the nets are trained. The NN history has seen in the last decades the birth of innovative models that have been employed in various fields, from the preliminary idea proposed by Hebb in 1949 to the present days. Despite the

computational speed of modern processors (≈ 10^{-9} s) is sharply higher than the processing speed of a biological neuron (≈ 10^{-3} s), the total number of operations per second performed by the brain greatly exceeds the number of calculations carried out by a computer. This superiority is due to the nature of the brain: many elementary units, called neurons, process simple information in parallel, rather than sequentially, as in a computer, and communicate with each other through synapses [67].

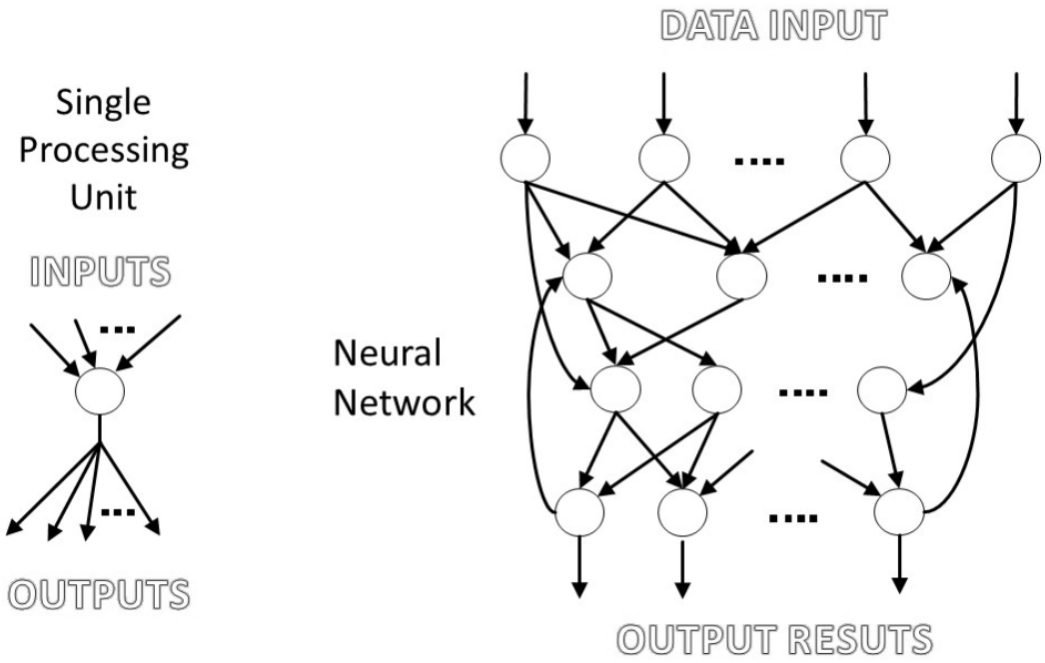

Figure A.2: Single processing unit, neuron, on the left and a NN on the right. Starting from the input units, to which the data of the problems to solve are supplied, the computation is propagated in parallel in the network up to the output units, which provide the result. Reproduced from [61].

An Artificial Neural Network (ANN) is a computing paradigm inspired by the structure and functioning of the brain described above. Usually a NN changes its structure based on the external and internal information presented to it during the learning process. It uses an elaborative architecture that presents a series of analogies with natural intelligence in information processing: strong interconnection, rapid communication, simple processing units, large amount of memory (in connections) and unplanned but trained by machine learning. As anticipated, the NN is a parallel computational model, made up of numerous homogeneous processing units that are strongly interconnected by links of varying intensity. The activity of the single unit is simple, guided by the transfer function, and the power of the model lies in the configuration of the connections. The data of the problems to be solved are supplied to the input units, then the computation is propagated in parallel in the network up to the output units, which supply the result, as shown in Fig A.2. A NN is not programmed to perform a certain activity but using an automatic learning algorithm is trained through a series of examples to be modelled.

A.2.1 Mathematical modelling of the artificial neuron

The artificial neuron represents the fundamental processing unit for NN. Figure A.3 shows a model of the artificial neuron described with a block diagram, in which it is possible to identify different elements:

1. A series of synapses or connections, each of them characterized by a weight. Specifically, the input x_j signal to the neuron k is multiplied by its weight $w_{k,j}$ which can take both positive and negative values, unlike the biological case. In this way it has the possibility to express both excitatory and inhibitory synapses;
2. An adder, which linearly combines the various inputs according to the respective synaptic weights;
3. An activation function $\varphi((\cdot))$, which limits the amplitude of the neuron output in a given range of defined values, typically in [0; 1] or in [–1; 1]);
4. A bias term, indicated with b_k, which can increase or decrease the input signal to the activation function depending on whether it assumes a positive or negative value.

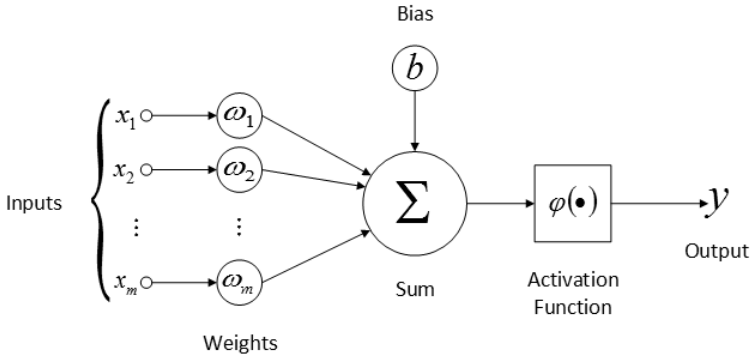

Figure A.3: Non-linear model of the neuron k. Reproduced from [68].

The mathematical terms that describe the functioning of the neuron are:

$$v_k = \sum_{j=1}^{m} w_{kj} x_j \tag{A.1}$$

$$y_k = \varphi(v_k + b_k) \tag{A.2}$$

Where $x_1, x_2, \ldots\ldots, x_m$ are the network input; $w_{k1}, w_{k2}, \ldots\ldots, w_{km}$ are the synaptic weights related to the neuron k; v_k is the result of the linear combination of input signals; b_k is the bias; $\varphi((\cdot))$ is the activation function; and y_k is the neuron output. The bias effect consists in applying a similar transformation to the output v_k of the linear adder of the neuron model of Fig. A.4. Since bias b_k is an external parameter of neuron k, as are weights, we can reformulate the equation (5.1) as:

$$v_k = \sum_{j=0}^{m} w_{kj} x_j \tag{A.3}$$

And the (A.2) becomes:

$$y_k = \varphi(v_k) \tag{A.4}$$

This result is obtained simply by adding a new synapse whose input is fixed and is worth:

$$x_0 = +1 \tag{A.5}$$

The neuron model thus becomes the one illustrated in Fig. A.4. We can see the contribution of the bias taken into consideration by adding a new entry with a constant value of +1 and adding a new synaptic weight equal to b_k. Although the two models presented may seem different, from the mathematical point of view they are equivalent.

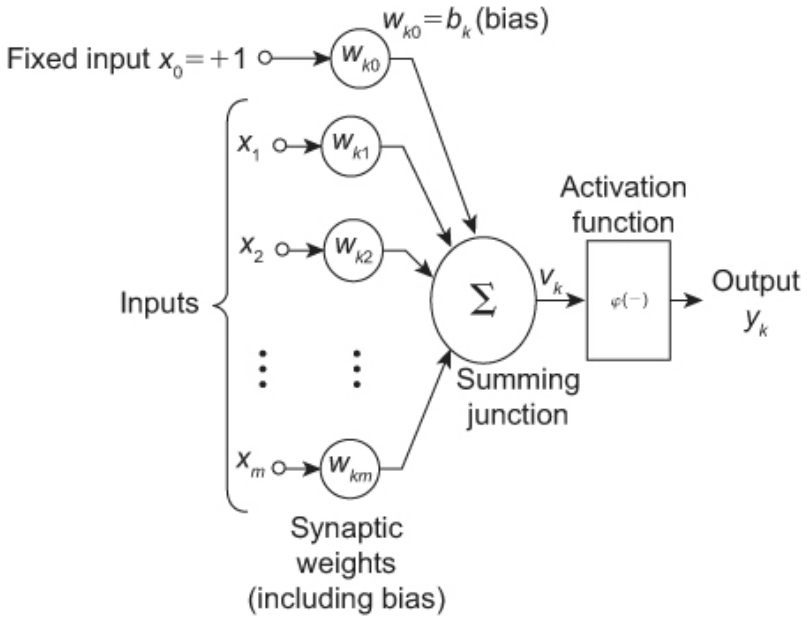

Figure A.4: Reformulation of the neuron model. The bias has been included among the weights thanks to the addition of a new input whose value is always + 1, [61].

A.2.2 Activation function

The activation function, indicated with $\varphi((\cdot))$, calculates the output of the neuron based on the activity level at its entry. The main types of activation functions are:

1. *Threshold function*. This function, represented in panel 2 of Figure A.5, is defined by the expression

$$\varphi(v) = \begin{cases} 1 & se\ v \geq 0 \\ 0 & se\ v < 0 \end{cases} \tag{A.6}$$

This neuron follows all or nothing behaviour, since its output can take only two values, 1 if its internal activity exceeds a certain threshold, 0 otherwise. This function is particularly used in the last neuron of networks used for classification.

2. *Piecewise linear function.* This function is defined by

$$\varphi(v) = \begin{cases} 0 & se\ v \leq -\frac{1}{2} \\ v & se\ -\frac{1}{2} < v < \frac{1}{2} \\ 1 & se\ v \geq \frac{1}{2} \end{cases} \tag{A.7}$$

where the amplification factor, in the linear region, is unitary. this function tends to the threshold function if the amplification factor in the linear zone is made infinitely large.

3. *Linear function*. Represented in panel 1 of Figure A.5, it is simply defined as

$$\big(\varphi(v)\big) = v \tag{A.8}$$

and is widely used as a function of activation of neurons in the output layer, especially if the target output to be reproduced varies in a certain range of values, not necessarily finite.

4. *Sigmoid function*. It is a strictly increasing, continuous, derivable and limited function between 0 and 1. Represented in panel 3 of Figure A.5

$$\varphi(v) = \frac{1}{1 + e^{-cv}} \tag{A.9}$$

where *c* is a parameter that influences the slope.

5. *Hyperbolic tangent function*. Together with the sigmoid function and the most used among the activation functions, it is represented in panel 4 of Figure A.5. It has similar properties to the sigmoid function, but also assumes negative values and is antisymmetric with respect to the origin. It is defined by

$$\varphi(v) = \frac{1 - e^{-2cv}}{1 + e^{-2cv}} \tag{A.10}$$

This function is particularly useful when the neuron output has to be zero if the weighted sum of its inputs is zero, and if it is desired it has an antisymmetric behaviour with respect to the internal activation threshold reached. Another very important feature that this function shares with the sigmoid function is the non-linear behaviour that allows obtaining outputs that depend not linearly on the inputs of the neuron. There are also other types of activation functions and for a more in-depth view, see [68].

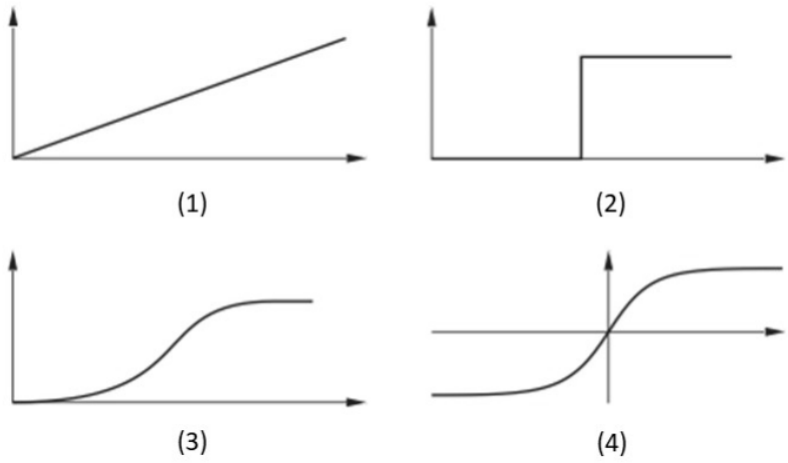

Figure A.5: Some possible activation functions; (1) Linear activation function; (2) Threshold activation function; (3) Sigmoid activation function with computed values between 0 and 1; (4) Sigmoid activation function with computed values between -1 and 1. Reproduced from [68].

A.3 Network architecture: Feedforward neural networks

The architecture of the neural network is defined by the way in which neurons are connected and structured and it is closely connected to the training algorithm used to train the network. For this reason, there are many different types of neural networks and a common choice is to classify them according to their structure:

- The Feedforward Neural Network (FFNN) contains only "forward" connections, i.e. the information flows orderly and sequentially from the entrance to the output; the structure remained more widespread and used.
- The Recurrent Neural Network (RNN), differ from the FFNN due to the presence of at least one link to feedback loop. It has a hidden layer in which, in addition to the classic inputs outside the network, there are self feedback loops, that are synapses that allow hidden neurons to "self-feed" with their outputs. The presence of the feedback loop has a profound impact on the performance of the network and its ability to learn.
- Neural networks Radial-Basis Function (RBFNN). This type of network differs substantially from the networks considered so far. Network learning corresponds to finding a surface capable of adhering well to the data provided in the training phase, statistically evaluating the quality of the result.
- Competitive networks and self-organizing maps: they are networks that learn by themselves to identify certain regularities and correlations between their inputs and adapt their response accordingly. The neurons of these networks, particularly suitable for classification, are able to recognize groups of similar inputs and belonging with high probability to the same class [69].

We analyse in detail the FFNN, that is the most used network architecture. FFNN is a network in which neurons are organized in layers. In Figure A.6, panel (a), its simplest form is illustrated, in which there is only one entry layer that directs the information to the layer of output neurons, where it is processed. This network is defined as a “single-layer network”, since the input layer is not counted, since no data operation takes place in it and the inputs, that are properly weighed, are processed directly by the neurons of the output layer. Figure A.6, panel (b), shows a multi-layered neural network. It is characterized by one or more layers of hidden neurons, whose function is to intervene between the external input and the network output, processing the information provided to the network and extracting higher-order statistics with respect to a single-layer network [61]. Usually the activation function is a sigmoid or hyperbolic tangent type for the neurons of the hidden layers, while for the neurons of the output it can be both sigmoid or linear, based on the task that the neural network has to solve [65]. Each of the nodes in the first layer receives an element from the input vector and transfers it to the neurons of the second layer, i.e. the first hidden layer, which processes the information received and passes it to the neurons of the immediately following layer. This process continues thus to the neurons of the last layer of the network, whose output corresponds to the response of the network to the input vector. Typically, the neurons of each layer of the network only have the output signals of the previous layer in input and direct their output exclusively to the neurons of the next layer. The hidden neurons are able to learn the pattern in data during the training phase and mapping the relationship between input and output pairs [67]. The FFNN are widely and successfully used for forecasting and problem solving, thanks to their simplicity and their versatility.

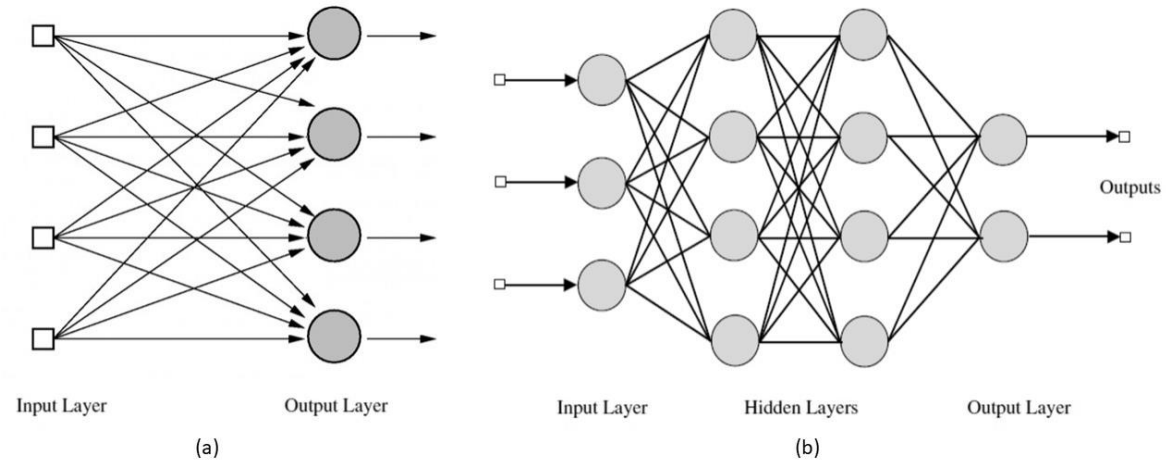

Figure A.6: The two types of neural networks. On the left, (a), the NN with a single layer of neurons is represented. On the right, (b), an FFNN is represented with a hidden layer consisting of 4 neurons and 2 output neurons. Reproduced from [61].

A.4 Supervised training of the network

There are various ways to train a neural network, that is to determine the weights so that it solves the desired task. Typically, a distinction is made between:

- Supervised learning, during which the desired output vector is known for each input vector. Consequently, in this learning method the aim is to minimize the error, that is the difference between the desired output and the actual network output, called also back-propagation; Models that use this learning: Perceptron, Multi-layer perceptron and radial basis function;
- Weakly guided learning, in which an error signal cannot be generated for each input vector presented to the network. Also in this case feedback is given to the network, but not after each time point, but more occasionally and usually at the end of a long evaluation;
- Unsupervised learning, during which the network learns on its own to recognize important statistical properties of the inputs presented to it and responds appropriately to sets of similar incoming vectors. This technique is usually used for classification, when the classes to which the different vectors belong to not known [67]. Models that use this learning: Kohonen and Hopfield.

The back-propagation algorithm is probably the most used methods for training FFNN. It is based on the Gradient Descent (GD) method that attempts to minimize the error of the network by moving down the gradient of the error curve. The back-propagation algorithm was developed by Paul Werbos in 1974 and rediscovered independently by Rumelhart and Parker [70].

The training algorithm consists of two basic parts: a forward pass (forward pass) and a backward pass (backward pass).

- In the first phase the input is presented to the network and the corresponding output is calculated. The weights remain unchanged. The output of the network is compared with the desired output and the error committed by each neuron of the output layer is calculated.

- In the second phase the synaptic weights are modified by applying the "error-correction" rule. The error is propagated backwards through the network, from the output to the entrance and an attempt is made to understand the error made by each hidden neuron, of which neither the target output nor the real one is known, so as to adequately modify the relative weights.

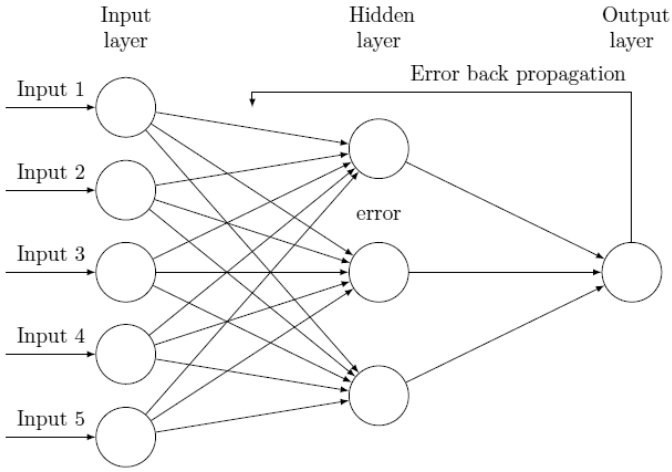

Figure A.7: A representative FFNN with back-propagation learning algorithm.

A NN whose neurons are equipped with a differentiable activation function produces an output that is also differentiable, both according to the inputs and according to the weights and biases. Thus, defining an error function derivable on the basis of the output of the network itself, it will also be derivable with respect to synaptic weights. Thanks to this property, the partial derivatives can be used in order to identify a law for updating the weights such as to minimize the error function by means of the gradient-based back-propagation algorithm (or another algorithm). There are many variations of this algorithm that optimize the speed and the ability to identify a global minimum of the error function, without getting stuck in a local minimum, the most used is the Levenberg - Marquardt algorithm.

A.4.1 Back-propagation: Levenberg Marquardt algorithm

As already described above, back-propagation training algorithm is a widely used technique in ANN and it is also a very popular optimization task in finding an optimal weights sets during the training process. However, traditional back-propagation algorithms have some drawbacks such as getting stuck in local minimum and slow speed of convergence. Some possible variations are proposed from MATLAB® software: the back-propagation training functions that use Jacobian derivatives are Levenberg-Marquardt algorithm and Bayesian Regulation back-propagation, while others that use gradient derivatives could be Gradient descent with momentum, RPROP backpropagation and BFGS quasi-Newton backpropagation. Below is a brief description of what distinguishes each of these back-propagation algorithms.

1. *Batch Gradient Descent with Momentum training algorithm (GDM):* this training algorithm updates the network weights in the direction of the negative gradient of the performance function by a factor determined by a parameter known as the learning rate. This algorithm makes use of momentum, which allows a network to respond not only to the local gradient, but also to recent trends in the error surface, allowing networks to avoid getting stuck in shallow minima [69].

2. *Resilient Backpropagation training algorithm (RP):* the algorithms that rely on gradient descent can get stuck in local minima or slow down significantly when the magnitude of the gradient is small. The resilient back-propagation training algorithm avoids this problem by using the sign of the gradient to determine the direction of the weight change. The amplitude of the weight change is obtained by a value that is sensitive to the behaviour of this sign. This method of changing magnitudes allows the resilient back-propagation algorithm to converge very rapidly [69].
3. *The BFGS training algorithm* belongs to a class of training algorithms known as Quasi-Newton algorithms. These algorithms approximate Newton's method [69]. Quasi-Newton algorithms approximate the complex and computationally expensive calculation of the Hessian matrix by using a function of the gradient instead of calculating the second derivative.
4. *Levenberg-Marquardt training algorithm:* This training algorithm is another algorithm that approximates Newton's method. The Levenberg–Marquardt algorithm blends the steepest descent method and the Newton's algorithm. Fortunately, it inherits the speed advantage of the Newton's algorithm and the stability of the steepest descent method. It's more robust than the previous ones, because in many cases it can converge well even if the error surface is much more complex than the quadratic situation.
5. *Bayesian Regulation backpropagation*: This training algorithm updates the weight and bias values according to Levenberg-Marquardt optimization. It minimizes a combination of squared errors and weights and then determines the correct combination to produce a network that generalizes well. The process is called Bayesian regularization.

In this work, we have applied the Levenberg – Marquardt back-propagation algorithm which will be described in detail below. It is specifically designed to minimize sum-of-square error functions [69], of the form

$$E = \frac{1}{2}\sum k\,(e_k)^2 = \frac{1}{2}\,\|e\|^2 \tag{A.11}$$

Where e_k is the error in the k^{th} exemplar or pattern and e is a vector with element e_k. If the difference between the pervious weight vector and the new weight vector is small, the error vector can be expanded to first order by means of a Taylor series.

$$e(j+i) = e(j) + \frac{\partial e_k}{\partial w_i}(w(j+i) - w(j)) \tag{A.12}$$

Consequently, the error function can be expressed as

$$E = \frac{1}{2}\left\| e(j) + \frac{\partial e_k}{\partial w_i}\,(w(j+1) - w(j))\right\|^2 \tag{A.13}$$

Minimizing the error function with respect to the new weight vector, it gives

$$w(j+1) = w(j) - (Z^T Z)^{-1} Z^T e(j)$$
$$\text{Where } (Z)_{ki} \equiv \frac{\partial e_k}{\partial w_i} \tag{A.14}$$

Since the Hessian for the sum-of-square error function is

$$(H)_{ij} = \frac{\partial^2 E}{\partial w_i}\partial w_j = \sum\left\{\left(\frac{\partial e_k}{\partial w_i}\right)\left(\frac{\partial e_k}{\partial w_i}\right) + e_k\frac{\partial^2 e_k}{\partial w_i}\partial w_j\right\} \qquad \text{(A.15)}$$

Neglecting the second term, the Hessian can be written as $H = Z\hat{}T$.

The weight's update therefore involves the inverse Hessian or an approximation thereof for nonlinear networks. The Hessian is relatively easy to compute, since it is based on first order derivatives with respect to the network weights that are easily accommodated by back-propagation. Although the updating formula could be applied iteratively to minimize the error function, this may result in a large step size, which would invalidate the linear approximation on which the formula is based. In the Levenberg-Marquardt algorithm, the error function is minimized, while the step size is kept small in order to ensure the validity of the linear approximation. This is accomplished through the use of a modified error function of the form.

$$E = \frac{1}{2}\left\|e(j) + \frac{\partial e_k}{\partial w_i}\left(w(j+1) - w(j)\right)\right\|^2 + \lambda\|w(j+1) - w(j)\|^2 \qquad \text{(A.16)}$$

In which λ is a parameter governing the step size. Minimizing the modified error with respect to $w(j+1)$ gives

$$w(j+1) = w(j) - (Z^T Z + \lambda I)^{-1} Z^T e(j) \qquad \text{(A.17)}$$

When the scalar λ is zero, this is basically Newton's method, using the approximate Hessian matrix. When λ is large, this becomes descent gradient with a small step size. Newton's method is faster and more accurate near an error minimum, so the aim is to shift toward Newton's method as quickly as possible. Thus, λ is decreased after each successful step (reduction in performance function) and it is increased only when a tentative step would increase the performance function. In this way, the performance function is always reduced at each iteration of the algorithm. Various more or less heuristic arguments have been put forward for the best choice for the damping parameter λ. Theoretical arguments exist showing why some of these choices guarantee local convergence of the algorithm; however, these choices can make the global convergence of the algorithm suffer from the undesirable properties of steepest descent, in particular, very slow convergence close to the optimum. The absolute values of any choice depend on how well-scaled the initial problem is. Marquardt recommended starting with a value λ_0. This algorithm appears to be the fastest method for training moderate-sized feedforward neural networks. It also has an efficient implementation in MATLAB® software, because the solution of the matrix equation is a built-in function, so its attributes become even more pronounced in a MATLAB environment.

Bibliography

[1] K. Colin, «When hypoglycemia is not obvious: Diagnosis and treating underecognized and undisclosed hypoglycemia,» *Primary Care Diabetes,* vol. 8, n. 1, pp. 3-11, 2013.

[2] V. a. S. N. Davis, «Hypoglycemia in type 1 and type 2 diabetes: Physiology pathophysiology and management,» *Clinical Diabetes,* vol. 24, n. 3, pp. 115-121, 2006.

[3] I. D. Atlas, «http://www.diabetesatlas.org/,» [Online].

[4] d. C.-G. F. Martin-Timon I, «Mechanisms of hypoglycemia unawareness and implications in diabetic patients,» *World J Diabetes,* vol. 6, n. 7, pp. 912-926, 2015.

[5] A. R. S. a. C. Kahn, «Insulin signaling and the regulation of glucose and lipid metabolism,» *Nature,* vol. 414, n. 6865, pp. 799-806, 2001.

[6] P. Cryer, «The barrier of hypoglycemia in diabetes,» *Diabetes,* vol. 57, n. 12, pp. 2169-3176, 2008.

[7] C. A. J. v. B. a. J. H. DeVries, «Analysis: The accuracy and efficacy of the dexcom g4 platinum continuous glucose monitoring system,» *Journal of Diabetes Science and Technology,* vol. 9, n. 5, pp. 1027-1029, 2015.

[8] P. L. F. a. W. T. T.W.Jones, «Decreased epinephrine responses to hypoglycemia during sleep,» *New England Journal of Medicine,* vol. 338, n. 23, pp. 1657-1662, 1998.

[9] B. T. M. W. J. a. M. P.McEwan, «Healtcare resource implications of hypoglycemia-related hospital admissions and inpatient hypoglycemia: Retrospective record-linked cohort studies in england,» *BMJ Open Diabetes REsearch and Care,* vol. 3, n. 1, 2015.

[10] A. B. a. H. H. F. Erbsloh, «The glucose consumption of the brain & its dependence of the liver,» *Archiv fur Psychiatrie und Nervenkrankheiten, vereinigt mit Zeitschrift fur die gesamte Neurologie und Psychiatrie,* vol. 196, n. 6, pp. 611-626, 1957.

[11] U. L. G. D. a. A. M. P. Merhenthaler, «Sugar for the brain: The role of glucose in physiological and pathological brain function,» *Trends in neurosciences,* vol. 36, n. 10, pp. 587-597, 2013.

[12] L. Sokoloff, «Energetics of functional activation in neural tissues,» *Neurochemical research,* vol. 24, n. 2, pp. 321-329, 1999.

[13] L. K. G. S. e. a. Wessels AM, «Diabetes and cognitive decline in elderly African Americans: a 15year followuup study,» *Alzheimers Dement,* vol. 7, pp. 418-424, 2011.

[14] C. D. Marco Di Mauro, «www.diabete-rivistamedia.it,» 2017. [Online].

[15] T. U.-B. A.S. Sejling, «Hypoglycemia-associated changes in the electroencephalogram in patients with type 1 diabetes and normal hypoglycemia awareness or unawareness,» *Diabetes,* vol. 64, n. 5, pp. 1760-1769, 2015.

[16] K. M. R. C. e. a. Feinkohl I, «Clinical and subclinical macrovascular disease as predictors of cognitive decline in older patients with type 2 diabetes: the Edinburgh Type 2 Diabetes Study,» *Diabetes Care,* vol. 36, n. 9, pp. 2779-86, 2013.

[17] E. A. Larsen, «Classification of EEG Signals in a Brain-Computer interface system,» 2011.

[18] L. B. a. C. Juhl, «Hypoglycemia-induced changes in the electroencephalogram: An overview,» *Journal of diabetes science and technology,* 2016.

[19] M. F. a. B. L. Pacella, «Association of spontaneous hypoglycemia with hypocalcemia and electrocerebral dysfunction,» *Archives of Internal Medicine,* vol. 81, n. 2, pp. 184-202, 1948.

[20] I. R. a. L. Loeser, «Electroencephalographic findings ini essential hypoglycemia,» *Electroencephalography and Clinical Neurophysiology, 2, Elsevier,* vol. 3, pp. 141-148.

[21] B. T. B. S. a. C. B. S. Pramming, «Glycaemic threshold for changes in electroencephalograms during hypoglycaemia in patients with insulin dependent diabetes,» *British Medical Journal (Clin Res Ed),* vol. 296, n. 66623, pp. 665-667, 1988.

[22] M. Rubega, «Quantitative analysis of hypoglycemia-induced EEG alterations in type 1 diabetes [Ph.D thesis],» 2017.

[23] Hyposafe, «www.hyposafe.com,» [Online].

[24] R. M. a. H. Beck-Nielsen, «Implantable electronic devices for detecting hypoglycaemia using eeg singals,» *US Patent* , 2014.

[25] L. F. R. E. L. S. R. H. B.-N. B. T. P. M. G. a. C. B. J. L.S. Snogdal, «Detection of hypoglycemia associated eeg changes during sleep in type 1 diabetes mellitus,» *Diabetes Research and Clinical Practice,* vol. 98, n. 1, pp. 91-97, 2012.

[26] S. a. B. T. U. Pedersen-Bjergaard, «Recall of severe hypoglycaemia and self-estimated state of awareness in type 1 diabetes,» *Diabetes Metabolism Research Reviews,* vol. 19, n. 3, pp. 232-240, 2003.

[27] K. M. M. a. B. M. F. A. E. Gold, «Frequency of severe hypoglycemia in patients with type 1 diabetes with imparied awareness of hypoglycemia,» *Diabetes Care,* vol. 17, n. 7, pp. 697-703, 1994.

[28] D. J. C. L. G.-F. D. J. D. S. a. W. P. W. L. Clarke, «Reduced awareness of hypoglycemia in adults with iddm. a prospective study of hypoglycemic frequency and associated symptoms,» *Diabetes Care,* vol. 18, pp. 517 - 522, 1995.

[29] K. A. &. J. Frank, «Trail Making Test, Part B as a Measure of Executive Control: Validation Using a Set-Switching Paradigm,» *Journal of clinical and experimental neuropsychology,* vol. 22, n. 4, pp. 518-528, 2000.

[30] C. Golden, «A group version of the stroop color and word test,» *Journal of personality assessment,* vol. 39, n. 4, pp. 386-388, 1975.

[31] U. P.-B. T. K. N. O. F. D. J. H. J. F. L. T. a. B. T. P.L. Kristensen, «Influence of erythropoietin on cognitive performance during experimental hypoglycemia in patients with type 1 diabetes mellitus: A randomized cross-over trial,» *PloSon*, vol. 8, n. 4, 2013.

[32] D. A. R. H. P. E. a. M. L. J. Escudero, «Analysis of electroencephalograms in alzheimer's disease patients with multiscale entropy,» *Physiological measurement*, vol. 27, n. 11, p. 1091, 2006.

[33] G. S. A. S. A. G. J. D. L. S. R. C. B. J. a. C. C. C. Fabris, «Hypoglycemia related electroencephalogram changes assessed by multiscale entropy,» *Diabetes technology & therapeutics*, vol. 16, n. 10, pp. 688-694, 2014.

[34] M. A. G. a. C.-K. Peng, «Multiscale entropy analysis of complex physiologic time series,» *Physical review letters*, vol. 89, n. 6.

[35] V. Unakafova, «Investigating measures of complexity for dynamical systems and for time series,» 2015.

[36] M. A. M. C. a. F. B. A. Accardo, «Use of the fractal dimension for the analysis of electroencephalographic time series,» *Biological cybernetics*, vol. 77, n. 5, 1997.

[37] F. S. a. M. Z. F. Finotello, «EEG signal features extraction based on fractal dimension,» *IEEE*, pp. 4154 - 4157, 2015.

[38] B. P. C. Bandt, «Permutation Entropy - A natural complexity measure for time series,» *Phys Rev*, 2002.

[39] M. S. Keller K., «Ordinal analysis of time series,» *Phys*, vol. 356, pp. 114-120, 2005.

[40] U. A. M. a. U. V. K. Keller, «Ordinal patterns, entropy and EEG,» *Entropy*, vol. 16, n. 12, pp. 6212-6239, 2014.

[41] M. S. J. E. K. Keller, «Time series from the ordinal viewpoint,» *Stochastics and Dynamics*, vol. 07, n. 02, pp. 247-272, 2007.

[42] J. M. Z. S. S. M. Amigò, «Combinatorial detection of determinism in noisy time serires,» *Europhys*, vol. 83, 2008.

[43] J. Amigo, Permutation complexity in dynamical systems. Ordinal patterns, permutation entropy and all that, Springer, 2010.

[44] J. B. G. Y. Cao, «Detecting dynamical changes in time series using the permutation entropy,» *Phys Rev*, 2004.

[45] X. L. Z. L. L. V. J. S. D. Li, «Multiscale permutation entropy analysis of EEg recordings during sevourane anesthesia,» *Journal of neural engineering*, vol. 7, n. 4, 2010.

[46] C. Q. Q. M. C. Bian, «Modi ed permutation-entropy analysis of heartbeat dynamics,» *Physical Review*, 2012.

[47] K. K. J.M. Amigo, «Permutation Entropy: one concept, two approaches,» *The European Physical Journal Special Topics*, 2013.

[48] M. Zanin, L. Zunino, O. Rosso e D. Papo, «Permutation entropy and its main biomedical and econophysics applications:,» *Entropy,* vol. 14, p. 1553–1577, 2012.

[49] F. Morabito, D. Labate, F. La Foresta, A. Bramanti, G. Morabito e I. Palamara, «Multivariate multi-scale permutation entropy for complexity analysis of Alzheimer's disease EEG,» *Entropy,* vol. 14, p. 1186–1202, 2012.

[50] K. K. Anton M. Unakafov, «Conditional entropy of ordinal patterns,» *Physica D: Nonlinear Phenomena,* vol. 269, pp. 94-102, 2014.

[51] T. Higuchi, «Approach to an irregular time series on the basis of the fractal theory,» *Physica D: Nonlinear Phenomena,* vol. 31, n. 2, pp. 277-283, 1988.

[52] P. K. S. A. S. G. Pokychronaki, «Comparison of fractal dimension estimation algorithms for epileptic seizure,» *Journal of neural engineering,* vol. 7, n. 4, 2010.

[53] X. Z. Y. O. a. A. R. F. Scarpa, «Automatic evaluation of corneal nerve tortuosity in images from in vivo confocal microscopy,» *Investigative ophthalmology e visual science,* vol. 52, n. 9, pp. 6401-6408, 2011.

[54] K. H. G. H. P. A. H. T. a. Z. G Tribl, «Eeg topography during insulin-induced hypoglycemia in patients with insulin-dependent diabetes mellitus,» *European Neurology,* vol. 36, p. 303–309, 1996.

[55] L. R. H. B.-N. a. C. J. R. Elsborg, «Detecting hypoglycemia by using the brain as a biosensor,» *InTech,* 2011.

[56] B. D. B.E. Brodsky, «Non-parametric statistical diagnosis. problems and methods,» *Dordrecht: Kluwer Academic Publishers,* 2000.

[57] V. A. U. a. K. Keller, «Efficiently Measuring Complexity on the Basis of Real-World Data,» *Entropy,* vol. 15, n. 10, pp. 4392-4415, 2013.

[58] C.-W. W. S.-G. L. K.-Y. S.-D. Wu, «Analysis of complex time series using refined composite multiscale entropy,» *Physics Letters A,* vol. 378, n. 20, pp. 1369 - 1374, 2014.

[59] F. T. I.-T. K. a. W. S. E.L.Vlahou, «Resting state slow wave power, healthy aging and cognitive performance,» *Scientific Reports,* vol. 4, p. 5101, 2014.

[60] H. Z. Z. &. Y. Y. An, «Opportunities and challenges on nanoscale 3D neuromorphic computing system,» *IEE International Symposium on Electromagnetic Compatibility,* pp. 416-421, 2017.

[61] S. Haykin, «Neural Networks: a Comprehensive Foundation,» *Third Avenue,* vol. 866, 1994.

[62] M. M, «Alzheimer. Come diagnosticarlo precocemente con le reti neurali artificiali.,» 2010.

[63] L. Z.-C. V. Alcalà-Rmz, «Identification of Diabetic Patients through Clinical and Para-Clinical Features in Mexico: An Approach Using Deep Neural Networks,» *Int. J. Environ.Res.Public Health,* vol. 3, n. 16, p. 381, 2019.

[64] Q. Q. K. L. Y. Y. D. J. Y. &. T. H. Zou, «Predicting diabetes mellitus with machine learning techniques,» *Frontiers in genetics,* 2018.

[65] Z. C, «Reti neurali per la predizione della glicemia futura mediante sensori per il continuous glucose monitoring.,» 2010.

[66] C. L., «Valutazione di nuove reti neurali per la predizione della glicemia futura mediante sensori CGM e modelli dell'assorbimento del glucosio dopo pasto.,» 2012.

[67] M. Wahde, «Biologically Inspired Optimization Methods. An Introduction.,» in *WIT Press*, 2008.

[68] S. Haykin, «Neural Network: a Comprehensive Foundation.,» 1999.

[69] M. H. a. H. D. M. Hudson Beale, «Neural Network toolbox. User's guide,» [Online]. Available: http://www.mathworks.com/help/pdf_doc/nnet/nnet.pdf..

[70] K. L. P. a. P. E.Keller, «Artificial Neural Networks An Introduction,» *The international Society of Optical Engineering,* 2005.

[71] S. A. Nye M, «Are efficient deep representations learnable?,» *ICLR,* 2018.

[72] G. G. D. E. G. A. G. a. G. S. C. C. Patterson, «Incidence trends for childhood type 1 diabetes in europe during 1989-2003 and predicted neq cases 2005-20: A multicentre prospective registration study,» *Lancet,* vol. 373, n. 9680, pp. 2027-2033, 2009.

[73] D.Danelea, «The accelerating epidemic of childhood diabetes,» *Lancet,* vol. 373, n. 9680, pp. 1999-2000, 2009.

[74] K. H. R. E. M. P. P. S. J. J. H. C. C. a. H. B.-N. C.B. Juhl, «Automated detection of hypoglycemia-induced eeg changes recorded by subcutaneous electrodes in subjects with type 1 diabetes-the brain as a biosensor,» *Diabetes Research Clinical Practice,* vol. 88, n. 1, pp. 22-28, 2010.

Acknowledgment

I would like to thank Prof. Giovanni Sparacino for giving me the opportunity to work on a similar study. An interesting job that allowed me to grow and interface with scientific research by developing initiative and motivation. Thanks also to Fabio Scarpa, an available and professional researcher. It was a guide showing enthusiasm, dispensing advice and patience during all phases of my research work. Finally, I would like to thank Maria Rubega for making her advice available on this work very close to her personally and professionally.

I dedicate this thesis to my family