Biopotential Signals and their Applicability to Cybersecurity Problems

by

Caterina Fuster-Barceló

A dissertation submitted by in partial fulfillment of the requirements for the degree of Doctor of Philosophy in

Computer Science and Technology

Universidad Carlos III de Madrid

Advisor:
Carmen Cámara Núñez

Tutor:
Pedro Peris López

December 2022

ACKNOWLEDGEMENTS

HASTA los dieciocho años mi vida la hacía completamente en catalán. No fue a hasta esa edad cuando me mudé a Barcelona y empecé a usar el castellano en mi día a día. Así que, como podréis entender, me veo en la obligación de repartir mi cariño y agradecimiento a cada persona en su lengua y, también, en la mía. Así que, a todas estas personas que me habéis ayudado en este camino, això va per voltros:

Nunca en mi vida, hasta los 18 años, me planteé hacer una ingeniería y mucho menos un doctorado. De hecho, como muchas personas sabrán, yo empecé derecho porque pensaba que era la mejor forma en la que podía poner mi granito de arena y hacer de este mundo un lugar mejor. Después de un año muy difícil, me di cuenta de que ese no era mi camino. Al ser esto unos agradecimientos, voy a empezar a dar las gracias a la primera mujer que fue mi referente a la hora de ser ingeniera (o en ese caso, estudiar una ingeniería), a Arantza. Gracias a ella vi algo que nunca me había planteado. Para mí, cuando era pequeña, quién hacía ingeniería siempre era un hombre mucho más listo que yo. Por lo que jamás me imaginé ese camino. Fue gracias a Arantza, y a su apoyo incondicional durante un año, que pude seguir adelante y elegir lo que haría que mi vida cambiara por completo.

Evidentment, tampoc hauria arribat fins aquí sense haver passat per sa Universitat de les Illes Balears on vaig estudiar el grau en enginyeria telemàtica. Van ser quatre anys que record amb molta estima. Es cafès, es vespres a sa biblioteca de Sa Riera, es treballs en grup, ses matrícules extraordinaires, ses tutories... tot té es seus bons i es seus mals moments. Allò que no dubt és que vaig tenir molt de suport per part des meus companys i ses meves companyes a qui sempre hi estaré agraïda per fer aquests quatre anys tan èpics, així que gràcies a en Toni, Anas, Gabriel, Javi, Medina, Juli, Ferran i més.

Con este afán de coleccionar títulos que hay ahora, yo no podía ser menos y tenía que hacer un máster. Así que, con dos maletas y un coche, nos mudamos a Madrid para lo que iba a ser un año. Cuando me encontré, después del máster, teniendo que decidir si hacer o no hacer un doctorado ya con el tema empezado y en mitad de una pandemia mundial, no tuve muchas dudas. Ahora, ya llevo más de tres años en Madrid y ya lo siento mi casa. Esta ciudad, y la gente que la habita, me ha sabido acoger de manera que me sienta muy a gusto.

Para quién me pregunte, le puedo adelantar que hacer un doctorado no es tarea fácil. Necesitarás una serie de ingredientes para hacerlo más llevadero. Estos son: amistad, apoyo, mucho café, cerveza, terapia grupal, que te repitan una y otra vez que no eres la única sufriendo esto y mucho más. En mi caso, tengo que admitir que todas estas necesidades las he tenido más que cubiertas gracias a ciertas personas que me han acompañado durante estos últimos tres años.

A Ramiro y a Maru les tengo que agradecer infinitamente haberme aguantado -que tampoco es tarea fácil- con todos mis cambios de humor y mis idas de olla. Habéis sido pieza fundamental para pasar todo esto con vuestros consejos y vuestro apoyo. Os quiero muchísimo.

Si hay una persona que ha aguantado 8 horas al día a mi lado con mis quejas, mis lloros e incapacidad de gestionar los problemas sin volverme loca es Josemi. Los desayunos en la cafetería del Sabatini cada mañana me han dado, en muchas ocasiones, las ganas de aparecer por el Laboratorio y ponerme a trabajar fingiendo que todo va bien y que todo sigue según lo planeado. De verdad, tienes un sitio ganado en el cielo por tu paciencia y apoyo.

También tengo que agradecer todas estas comidas que siempre se han alargado más de la cuenta con medio departamento de TSC. Me habéis hecho ver que no estaba sola y que me esperaban muchas más cosas fuera de la universidad, que no todo era escribir papers y lanzar experimentos. No habría sido lo mismo sin Emese, Dani, David, Peis, Lorena, Jesús, Santi y José Carlos. A todos os deseo muchísima suerte en todo y muchísimos ánimos porque esto, aunque no lo creáis, ya se acaba.

Esta tesis no habría salido adelante sin el esfuerzo incansable de Diana, quién ha estado ayudándome y dándome consejos de incalculable valor desde el minuto uno. Conocerla ha significado un antes y un después en el desarrollo de mi tesis y en mi vida. Gracias, gracias y mil gracias.

Me gustaría agradecer también a mi directora de tesis Dra. Carmen Cámara Núñez y mi tutor de tesis Dr. Pedro Peris López, su apoyo, sus útiles consejos, sus amables aportaciones y su gran autonomía en esta tesis.

No podria faltar de cap manera l'agraïment fort i sincer a sa meva família. A mon pare i ma mare per donar-me suport i per sempre ajudar-me quan els he necessitat. Per sempre haver-me ajudat amb tot el que han pogut i més i haver-me empès a fer tot el que jo me proposava encara que no ho entenguessin. A sa meva germana Francina per ser sa millor amiga, companya i germana que algú podria haver demanat mai. Es seu suport sempre ha estat incondicional. Encara que estiguem enfora, sé que sempre puc comptar amb voltros per tot allò que necessiti i que sempre hi sou per jo. També he de mencionar que no hi ha suport més incondicional que aquell que se t'aporta fins i tot quan no saps què és allò que estàs recolzant. Per això he d'agrair a sa padrina, es seu amor i devoció cap a ses seves nétes que m'ha fet que sigui més fàcil seguir. Mai vos podré agrair lo suficient lo molt que m'heu ajudat a treure això endavant. Vos estim.

Para terminar, no tengo ni la más mínima duda de que esta tesis no solo me pertenece a mí, sino que también es parte tuya. El apoyo incondicional, los ánimos, el amor, las noches de vino y queso, las tardes programando, las

noches de consuelo, las de borrachera, las de peli y manta y todo lo que me has dado durante estos últimos años es lo que ha hecho que esta tesis haya salido adelante. Me has acompañado siempre en todo lo bueno y en lo malo, has hecho que mis días sean más llevaderos y que estos años hayan sido un poco más felices, todo gracias a ti. No me basta ni todo el oro del mundo, ni todas las palabras del diccionario, ni todo el tiempo que dispongo para agradecerte todo lo que has hecho por mí. Te quiero y eres lo mejor que me ha pasado. Gracias por estar ahí siempre.

PUBLISHED AND SUBMITTED CONTENT

The following works and contributions are the result of this doctoral thesis. The content of the journal *peer-reviewed* publications and *preprints* is detailed during each of the chapters where it is indicated.

Type:	Journal Publication
Title:	ELEKTRA: ELEKTRokardiomatrix application to biometric identification with convolutional neural networks
Authors:	Caterina Fuster-Barceló, Pedro Peris-López, Carmen Cámara
Submitted to:	Neurocomputing
DOI:	`https://doi.org/10.1016/j.neucom.2022.07.059`
Note:	Submitted in April 2021 and Published in September 2022. Included mainly in Chapter 3 and part of Chapter 4. Material from this source included in the thesis is not marked by typographical means or references.

Type:	Journal Publication
Title:	An in-depth study of ELEKTRA as a biometric system
Authors:	Caterina Fuster-Barceló, Carmen Cámara, Pedro Peris-López
Submitted to:	Engineering Applications of Artificial Intelligence
Status:	Under review
Note:	The publication was Submitted on July 2022 and we are currently waiting for an answer. The results from this publication are shown in Chapter 4 and 5. Material from this source included in the thesis is not marked by typographical means or references.

OTHER RESEARCH MERITS

Other contributions and merits done during thie doctoral thesis.

Type:	Talk
Title:	Protegiendo el oro con tu corazón
Conference:	Woman Techmakers Madrid
URL:	`http://wtmgdgmadrid.github.io/agenda/caterina_fuster_2022.html`
Note:	Presentation of a non-technical view of the thesis.

Type:	Talk
Title:	Protegiendo el oro con tu corazón
Conference:	PyConES Granada
URL:	`https://charlas.2022.es.pycon.org/pycones2022/talk/review/BPRDBMLFMH9YW7RBALCHHS7THN7PPPKM`
Note:	Presentation of a non-technical view of the thesis but focusing on the Python implementation.

ABSTRACT

Biometric systems are an uprising technique of identification in today's world. Many different biometric systems have been used in everyone's daily life in the past years, such as fingerprint, face scan, ECG, and others. More than 20 years evince that the Elektrokardiogramm (EKG) or Electrocardiogram (ECG) is a feasible method to perform user identification as each person has their unique and inherent Elektrokardiogramm (EKG). A biometric system is based on something that every human being *is* and cannot lose or possess as it is an eye, the DNA, palm print, vein patterns, iris, retina, etc. For this reason, during the last decade, biometric identification or authentication has gained ground between the classic authentication systems as it was a PIN or a physical key. All biometric systems, to be accepted, must fulfill a set of requirements including universality, uniqueness, permanence, and collectability. The EKG is a biometric trait that not only fulfills those requirements but also has some advantages over other biometric traits. To use an EKG as the biometric trait for identification is motivated by four key points: 1) the collection of an EKG is a non-invasive technique so may contribute to the acceptability among the population; 2) a human being can only be identified if they are alive as their heart must be beating; 3) all living beings have their EKG so it is inclusive; 4) an EKG not only provides identification but also provides a medical and even emotional diagnose.

There exist many works regarding user identification with EKGs in the current state-of-the-art. Biometric identification with EKGs has been deployed using many different techniques. Some works use the fiducial points of the EKG signal (T-peak, R-peak, P-onset, QRS-offset, ...) to perform the user identification and others use feature extraction performed by a Neural Network as the classification or identification method. As the EKG is a signal which is expressed in time and frequency, many different Neural Network models can exploit the dissimilarity between each EKG signal from each user to perform user identification such as Recurrent Neural Networks, Convolutional Neural Networks, Long-Short Term Memory, Principal Component Analysis, among others offering very competitive results.

Focusing on user identification, depending on the user condition in each case, as has been commented before, the EKG not only contributes as an identification method but also offers a diagnosis as it is a person's condition from a medical point of view or a person's status regarding their emotional state. Some research has studied certain conditions such as anxiety over EKG identification showing that higher heart rates might be more complex to iden-

tify individuals.

Nevertheless, there are some drawbacks in the current state-of-the-art regarding identification with EKG. Many systems use very complexly Deep Learning architectures or, as commented, extract the features by a fiducial analysis making the biometric system too complex and computationally costly. One important flaw, not only in biometric systems but in science, is the lack of publicly available datasets and the use of private ones to perform different studies. Using a private database for any research makes the experiments and results irreproducible and it could be considered a drawback in any science field. Furthermore, many of these works use the EKG signal in a sense that it can be recovered from the identification system so there is no privacy protection for the user as anyone could retrieve their EKG signal.

Owing to the many drawbacks of a biometric system based on ECG signals, ELEKTRA is presented in this thesis as a new identification system whose aim is to overcome all the inconveniences of the current proposals. ELEKTRA is a biometric system that performs user identification by using EKGs converted into a heatmap of a set of aligned R-peaks (heartbeats), forming a matrix called an Elektrokardiomatrix (EKM).

ELEKTRA is based on past work where the EKM was already created for medical purposes. As far as the literature covers up to this date, all the existing research regarding the use of the EKM is focused on the diagnosis of different Cardiovascular Disease (CVD) such as Congestive Heart Failure, Atrial Fibrillation, and Heart Rate Variability, among others. Therefore, the work presented in this thesis, presumably, is the first one to use the EKM as a valid identification method.

In aim to offer reproducible results, four different public databases are taken to show the model feasibility and adaptability: i) the Normal Sinus Rhythm Database (NSRDB), ii) the MIT-BIH Arrhythmia Database (MIT-BIHDB), iii) the Physikalisch-Technische Bundesanstalt (PTBDB), and iv) the Glasgow University Database (GUDB). The first three of them (*i*, *ii* and *iii*) are taken from Physionet a freely-available repository with medical research data, managed by the MIT Laboratory. However, the fourth database (*iv*) is also freely available by petition to Glasgow University.

Furthermore, to test ELEKTRA's adaptability and feasibility of the biometric system presented, four different datasets are built from the databases where the EKG signals are segmented into windows to create several Elektrokardiomatrix (EKM)s. The number of EKMs built for each dataset will depend on the length of the records. For example, for the Normal Sinus Rhythm Database (NSRDB) as the EKG records are very extensive, 3000 EKMs or images per user will be obtained. However, for the three other databases, the highest possible number of EKM images is obtained until the signal is lost. It is important to take into account that depending on the number of heartbeats taken to be represented in each EKM, a different number of EKMs is obtained for the three databases in which EKG recordings are shorter. As higher the number of heartbeats o R-peaks taken (i.e., 7bpf), the fewer images will be obtained.

Once the datasets of EKMs are constructed, a simple yet effective Convolutional Neural Network (CNN) is built by one 2D Convolution with ReLU activation, a max-pooling operation followed by a dropout to include regularisation and, and finally, a layer with flattened and dense operations with a softmax or sigmoid function depending if the classification task is categorical o binary to achieve the final classification. With this simple CNN, the feasibility and adaptability of ELEKTRA are demonstrated during all the experiments.

The four databases are tested during chapters 3, 4, and 5 where the experimentation takes place. In Chapter 3, the NSRDB is studied as the baseline of identification with control users. Different experiments are conducted with aim of studying ELEKTRA's behavior. In the first experiments, how many heartbeats are needed to identify a user and the costs of convergence of the model depending on the time computing and the number of heartbeats taken to be represented in the EKM are studied. In this case, similar results are achieved in all the experiments as results close to 100% of accuracy are obtained. In the classification of a non-seen user a user, from a different database that has not been seen in any other experiment, is processed and tested against the network. The result obtained is that a non-seen user or an impersonator would only bypass the system one in ten times which can be considered a low ratio when many systems are blocked after three to five attempts. The classification of a user is tested to have a closer situation in which a low-cost sensor is used. For this experiment, an EKG signal is modified by adding Gaussian noise and then processed as any other signal. As a demonstration of our robust system, an accuracy of 99% is obtained indicating that a noisy signal can be processed too. The last experiment over the NSRDB is where this database is used to test the feasibility of ELEKTRA by testing how many images or EKM are enough to identify a user. Even though there is a decrease in accuracy when the number of images used to train the network is decreased too, a 97% of accuracy is obtained when training the network with only 300 EKMs per user. This chapter concludes that, as shown in all the experiments, ELEKTRA is a valid and feasible identification method for control users.

The MIT-BIH Arrhythmia Database (MIT-BIHDB) is a database comprising patients with Arrhythmia and random users, and the Physikalisch-Technische Bundesanstalt (PTBDB) comprises patients with different CVD together with healthy users. Hence, the main goal in Chapter 4 is to study the identification system proposed over users with CVD showing ELEKTRA's adaptability. First of all, the MIT-BIHDB is tested achieving outperforming results and showing how ELEKTRokardiomatrix Application to biometric identification with Convolutional Neural Networks (ELEKTRA) is capable to identify a pool of users with and without arrhythmia with just a slight decrease of the network's accuracy as a 97% of accuracy is obtained. Secondly, the whole PTBDB is taken to test the biometric system. The result obtained in this experiment is lower than in the other ones (a 93% of accuracy) as the number of images used to train the network has suffered a great decrease compared to the other experiments and 232 users are being studied. Lastly, ELEKTRA has tested over 162 users from the PTBDB with specific CVD which, namely, are Bundle branch block, Cardiomyopathy, Dysrhythmia, Myocardial infarction, Myocarditis, and Valvular heart disease. Through this experiment, the aim is to see ELEKTRA's behaviour when only users with CVD are included. Better results are obtained compared to the last experiment. It can be owed that the number of users has decreased and that a CVD makes more unique each EKG as many researchers use the EKM for diagnosis purposes. The conclusion extracted from all the experiments from this chapter is that ELEKTRA is capable to identify users with and without CVD approaching a real-life scenario.

In Chapter 5 the Glasgow University Database (GUDB) is tested to evaluate the performance of user identification when the users are performing different activities. The GUDB comprises 25 users performing five different activities with different levels of cardiovascular effort: sitting, walking on a treadmill, doing a maths exam, using a handbike, and running on a treadmill. The proposed biometric system is tested with each of these activities for 3 and 5 bpf achieving different results in each case. For the experiments performed where an activity requiring lower cardiovascular effort such as sitting or walking, the accuracy obtained is close to 100% as it is 99.19% for *sitting* and 98.59% for *walking*. Then for the scenarios where higher heartbeat rates are supposed the experiment results in lower accuracies as it is *jogging* with an 82.63% and *biking* with a 95.51%. For the *maths* scenario, its outcome is different; the heartbeat rate for each user could be different depending on how nervous each user is. Hence, a 94.0% is obtained with this activity. The conclusion extracted from these first experiments is that it is more complex to identify users when they are performing an activity that requires a higher cardiovascular effort and, consequently have a higher heart rate. For the following experiment, all scenarios have been merged to study the behaviour of a system that has been trained with users performing different activities. In this case, the results obtained seemed to be close to the mean of the results obtained before as the general accuracy for all the scenarios with 3bpf is 91.32%. For the subsequent experiments, some of the scenarios have been merged into two different categories. On the one hand, the more calmed activities (*sitting* and *walking*) have been merged in the so-called Low Cardiovascular Activity

(LCA) scenario. The accuracy obtained by training and testing with these two activities together is 97.74% and an EER of 1.01%. On the other hand, the High Cardiovascular Activity (HCA) scenario is composed by activities that require a higher cardiovascular effort (*jogging* and *biking*). In this case, the results obtained have decreased compared to the last ones as the accuracy is 85.71%. It can be noticed that what has suffered a considerable increase is the False Rejection Rate (FRR) which is 14.17% without implying an increase in the False Acceptance Rate (FAR) which is still very low as it is 0.6%. The last experiments have been called *fight of scenarios* as there is a confrontation between scenarios by merging some of them and training with some activities or scenarios and predicting with different ones. The first experiments that can be found in this section are training with the LCA group and testing with the HCA group and vice versa. The results here show a great decrease in the performance as accuracies are 37.24% and 46.42%, respectively. This fact implies that it is more complex to identify users that have been registered with a different heartbeat rate. Last but not least, there are a set of experiments where the activities have been confronted such as training the network with the *sitting* scenario and testing with the *jogging* scenario. These experiments confirm the hypothesis for higher heart rates, are more complex to identify users, and even more when the network has been trained over calmed users. Even though, one of the main advantages of the presented model is that, even for low accuracies, the False Acceptance Rate has not increased compared to the other experiments meaning that an impostor could not achieve bypassing the system.

Lastly, in Chapter 6 conclusions and discussions are offered. A comparison between ELEKTRA and other biometric systems based on EKGs from the current state-of-the-art is offered. These researches from the literature are examined to show how ELEKTRA outperforms all of them in regards to some of the aspects such as efficiency, complexity, accuracy, error rates, and reproducibility among others. It is important to remark that, compared to the other works, in all experiments performed in this doctoral thesis, really high performances with high accuracies and low error rates are achieved. In fact, what is remarkable is that this performance is obtained using a very simple CNN conformed by just one convolutional layer. By achieving outstanding results with a simple neural network, the solidity of ELEKTRA is proven.

By this, ELEKTRA contributes to the state-of-the-art by providing a new method for user identification with EKGs with many benefits. Outstanding results in terms of high accuracy and low error rates in the experiments assure the efficiency of ELEKTRA. The fact that the databases used to perform the experimentation in this doctoral thesis are publicly available, makes this work reproducible in contrast to many other works in the literature. In fact, as the databases used are different depending on the users' nature conforming to each database, it is established that the identification method proposed is inclusive as all living beings have their own EKG and high accuracies are also obtained when testing the model over users with different CVD. Moreover, as it has been proven that users with CVD can also be identified without having major drawbacks, ELEKTRA offers an identification system that can also offer a diagnosis of the user who is being identified in terms of their medical health. In addition, thanks to the GUDB, ELEKTRA can determine for the first time, as far as the literature reaches, that performing user identification with EKGs over users performing activities requiring a higher cardiovascular effort and consequently having higher heartbeat rates, is more complex.

In conclusion, by the studies and experiments performed in this doctoral thesis, it can be assumed that ELEKTRA is a feasible and efficient identification method for biometrics with EKG and outperforms the current state-of-the-art proposals in user identification with EKG.

ABSTRACT (NON-TECHNICAL)

BIOMETRIC systems are an uprising technique of identification in today's world. Many different systems have been used in everyone's daily life in the past years, such as fingerprint, face scan, ECG, and others. In fact, more than 20 years evidence that the Elektrokardiogramm (EKG) or Electrocardiogram (ECG) is a feasible method to perform user identification. In this thesis, a new identification method called ELEKTRA is proposed. Nevertheless, there are some drawbacks in the current state-of-the-art regarding identification with EKG. Many systems use very complex Deep Learning architectures or extract the features by a fiducial analysis making the biometric system too complex. One important flaw, not only in biometric systems, is the lack of publicly available datasets and the use of private ones to perform different studies. Using a private database for any research makes the experiments and results irreproducible and it could be considered a drawback in any science field.

In this doctoral thesis, ELEKTRA is developed as a biometric identification system by using EKGs converted into a heatmap of a set of aligned R-peaks (heartbeats), forming a matrix called an Elektrokardiomatrix (EKM). In aim to offer reproducible results, four different public databases are taken to show the model feasibility and adaptability: the Normal Sinus Rhythm Database (NSRDB), the MIT-BIH Arrhythmia Database (MIT-BIHDB), the Physikalisch-Technische Bundesanstalt (PTBDB) and the Glasgow University Database (GUDB). New datasets of EKMs are created from each of the aforementioned databases. Furthermore, to test ELEKTRA's adaptability and feasibility of the biometric system presented a simple yet effective CNN with only one Convolutional Layer constructed.

The four databases are tested during Chapters 3, 4, and 5 where the experimentation takes place. In Chapter 3, the NSRDB is studied as the baseline of identification with control users. Different experiments are conducted with aim of studying ELEKTRA's behaviour. The areas studied through this database are how many heartbeats are needed to identify a user; the costs of convergence of the presented model; the classification of a non-seen user obtained from a different database; the classification of a user whose EKG signal has been modified by adding Gaussian noise; and the feasibility of ELEKTRA by testing how many images or EKM are enough to identify a user.

Regarding databases that comprise users with CVD, the MIT-BIHDB is a database comprising patients with Arrhythmia and random users and the

PTBDB comprises patients with different CVD together with healthy users. These two databases are studied in Chapter 4 where the adaptability of ELEKTRA to different patients with CVD is studied. First of all, the MIT-BIHDB is tested achieving promising results and showing how ELEKTRA is capable to identify users with and without arrhythmia in the same pool. Secondly, the whole PTBDB is taken to test the biometric system obtaining high accuracies. And lastly, ELEKTRA is tested over some users with specific CVD from the PTBDB to see its behaviour when only users with CVD are included. The result from these experiments show how ELEKTRA is capable to identify users with and without CVD approaching a real-life scenario.

Lastly, in Chapter 5 the GUDB is tested to evaluate the performance of user identification when the users are performing different activities. The GUDB comprises 25 users performing five different activities with different levels of cardiovascular effort (sitting, walking, doing a maths exam, using a handbike, and running on a treadmill). The proposed biometric system is tested with each of these activities to show how it is more complex to identify users when they are performing an activity that requires a higher cardiovascular effort and, consequently, has a higher heart rate. The following experiments consist of different activities merged to study the differences between heartbeat rates and how user identification is related to the heartbeat rate. The more representative experiment is performed by training the model with the scenario where a user is sitting and performing the user classification with the scenario where the user is running. This way, a low accuracy is obtained proving that for higher heartbeat rates it is more complex to identify a user. In fact, one of the main advantages of the presented model is that, even with low accuracy, the False Acceptance Rate has not increased compared to the other experiments meaning that an impostor could not achieve bypassing the system. Even though, if the database is launched over all the activities merged, accurate results are shown to offer an inclusive model to train and test users performing different activities.

By this, ELEKTRA contributes to the state-of-the-art by providing a new method for user identification with EKGs with many benefits. Outstanding results in terms of high accuracy and low error rates in the experiments assure the efficiency of ELEKTRA. The fact that the databases used to perform the experimentation in this doctoral thesis are publicly available, makes this work reproducible. In fact, as the databases used are different depending on the users conforming to each database, it is established that the identification method proposed is inclusive as all living beings have their own EKG and high accuracies are also obtained when testing the model over users with different CVD. In addition, thanks to the GUDB, ELEKTRA can determine that performing user identification with EKGs over users performing activities requiring a higher cardiovascular effort is more complex.

In conclusion, by the studies performed in this doctoral thesis, it can be assumed that ELEKTRA is a feasible and efficient identification method for biometrics with EKG.

RESUMEN (NO TÉCNICO)

Los sistemas biométricos son una técnica de identificación en auge en la actualidad. En los últimos años se han utilizado muchos sistemas diferentes en la vida cotidiana, como la huella dactilar, el escáner facial, o el ECG, entre otros. De hecho, son más de 20 años los que avalan que el Elektrokardiogramm (EKG) o el Electrocardiogram (ECG) es un método fiable para realizar identificación de usuarios. En esta tesis se propone un nuevo método de identificación biométrica denominado ELEKTRA. Por otro lado, existen algunos inconvenientes en el estado del arte actual respecto a la identificación con EKG. Muchos sistemas utilizan arquitecturas muy complejas de Deep Learning o extraen las características importantes mediante un análisis fiduciario, haciendo que el sistema biométrico sea demasiado complejo o costoso. Un fallo importante, no solo en los sistemas biométricos, es la falta de bases de datos públicas y el uso de bases de datos privadas para la investigación. El uso de bases de datos privadas en cualquier estudio hace que los experimentos y los resultados sean irreproducibles y son un inconveniente en cualquier campo de la ciencia.

En esta tesis doctoral se ha desarrollado ELEKTRA, un sistema de identificación biométrica, mediante el uso de imagénes llamadas Elektrokardiomatrix (EKM). Estas imágenes se construyen a partir de realizar un mapa de calor de un conjunto de picos R (latidos) alineados, formando una matriz. Con el fin de ofrecer resultados reproducibles, se usan cuatro diferentes bases de datos públicas para demostrar la viabilidad y adaptabilidad del modelo: la Normal Sinus Rhythm Database (NSRDB), la MIT-BIH Arrhythmia Database (MIT-BIHDB), la Physikalisch-Technische Bundesanstalt (PTBDB) y la Glasgow University Database (GUDB). Se han creado nuevas sub-bases de datos de EKMs a partir de cada una de las bases de datos mencionadas. Además, para testear la adaptabilidad y viabilidad de ELEKTRA como sistema biométrico se construye una CNN sencilla, pero eficaz, con una sola capa Convolucional.

Las cuatro bases de datos anteriormente mencionadas se han testeado en los Capítulos 3, 4 y 5. En el Capítulo 3 se estudia la NSRDB como prueba de concepto de identificación en usuarios control. Se realizan diferentes experimentos con el objetivo de estudiar el comportamiento de ELEKTRA. Las características estudiadas con esta base de datos son: cuántos latidos son necesarios para identificar a un usuario; los costes de convergencia del modelo presentado; la clasificación de un usuario jamás visto proveniente de una base de datos diferente; la clasificación de un usuario cuya señal EKG ha sido modificada añadiendo ruido Gaussiano; y la viabilidad de ELEKTRA probando cuántas imágenes o EKM son suficientes para identificar a un usuario.

En cuanto a las bases de datos que contienen usuarios con CVD, la MIT-BIHDB contiene pacientes con Arritmia y usuarios sanos, y la PTBDB contiene pacientes con diferentes CVD junto a usuarios sanos. Estas dos bases de datos se estudian en el Capítulo 4, donde se estudia la adaptabilidad de ELEKTRA a distintas CVDs. En primer lugar, se testea la MIT-BIHDB logrando resultados prometedores y mostrando cómo ELEKTRA es capaz de identificar usuarios con y sin arritmia en el mismo grupo. En segundo lugar, se toma la PTBDB completa obteniendo porcentajes altos de acierto y bajos en cuanto a tasas de error concierne. Y por último, se prueba ELEKTRA sobre algunos usuarios con CVD específicos de la PTBDB para ver su comportamiento cuando sólo se incluyen usuarios con CVD. El resultado de estos experimentos muestra cómo ELEKTRA es capaz de identificar a los usuarios con y sin CVD acercándose a un escenario real.

Por último, en el capítulo 5 se prueba ELEKTRA sobre la GUDB para evaluar el rendimiento de la identificación de usuarios cuando éstos realizan diferentes actividades cardiovasculares. La GUDB consta de 25 usuarios que realizan cinco actividades diferentes con distintos niveles de esfuerzo cardiovascular (sentarse, caminar, hacer un examen de matemáticas, usar una bicicleta de mano y correr en una cinta). El sistema biométrico propuesto se prueba con cada una de estas actividades para mostrar que es más complejo identificar a los usuarios cuando realizan una actividad que requiere un mayor esfuerzo cardiovascular y, en consecuencia, tienen una mayor frecuencia cardíaca. Los experimentos realizados consisten en fusionar diferentes actividades para estudiar las diferencias entre las frecuencias cardíacas y cómo la identificación del usuario está relacionada la misma. El experimento más representativo se realiza entrenando el modelo con el escenario en el que el usuario está sentado y realizando la clasificación ciega de usuarios del escenario en el cual están corriendo. En este experimento, se obtiene una precisión realmente baja demostrando que para frecuencias de latidos más altas es más complejo identificar a un usuario. De hecho, una de las principales ventajas del modelo presentado es que, incluso con una precisión baja, la Tasa de Falsa Aceptación no ha aumentado en comparación con los otros experimentos, lo que significa que un impostor no podría conseguir eludir el sistema. Sin embargo, si la base de datos se lanza sobre todas las actividades fusionadas, se muestran resultados precisos que ofrecen un modelo inclusivo para entrenar y probar sobre usuarios que realizan diferentes actividades.

De este modo, ELEKTRA contribuye al estado del arte proporcionando un nuevo método de identificación de usuarios con EKGs con muchas ventajas. Los excelentes resultados en términos de alta precisión y bajas tasas de error en los experimentos, aseguran la eficiencia de ELEKTRA. El hecho de que las bases de datos utilizadas para realizar la experimentación en esta tesis doctoral estén disponibles públicamente, hace que este trabajo sea reproducible. De hecho, como las bases de datos utilizadas son diferentes en función de los usuarios que conforman cada una, se establece que el método de identificación propuesto es inclusivo ya que todos los seres vivos tienen su propio EKG. También, se obtienen altas precisiones al probar el modelo sobre usuarios con diferentes CVD. Además, gracias a la GUDB, ELEKTRA determina que identificar usuarios en base a sus EKGs mientras hacen actividades cardiovasculares, que requieren un mayor esfuerzo, es más complejo.

En conclusión, por los estudios realizados en esta tesis doctoral, se puede asumir que ELEKTRA es un método de identificación factible y eficiente para la biometría con EKG.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

NOWADAYS, biometrics techniques are widely used in everyday life. Examples range from the start of the day when an individual unlocks their phone with facial recognition or fingerprint, to enter their workspace with again their fingerprint, to access specific resources with an iris or retina scanner, or even by the police department to gather pieces of evidence from a crime scene with latent fingerprints [4, 5, 6]. Particularly, identification systems based on Elektrokardiogramm (EKG) or Electrocardiogram (ECG) have increased their use due to their uniqueness and inherent characteristics that may introduce new features to human identification.

Still, in the current state-of-the-art regarding the use of the EKG as the base for a biometric system, some drawbacks can be found. Through time, identification systems with different biometric traits have moved to use different Deep Learning (DL) techniques such as Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM) and others [7, 8, 9, 10]. But, unfortunately, many of the existing works base their study on very complex DL architectures making it difficult to reproduce those systems in a real-life situation. Moreover, computationally speaking, many of the studies in the current literature perform a fiducial analysis, which is very costly despite having good performances [11].

A general flaw of any biometric system and, even more over EKG, is the lack of publicly available databases and how many researchers use their own private databases to perform different studies and demonstrate their proposal feasibility making it irreproducible [12].

Regarding the integrity and inclusivity of an identification system from a human perspective, a weakness of many of them is that there exists a risk in which the EKG can be recovered from the biometric system endangering the user's privacy. And finally, not many of the works presented in the current literature cover the study of cardiovascular diseases over identification systems or even changes in the heart rate rhythm [12].

Owing to the many drawbacks of a biometric system based on EKG signals, ELEKTRA is presented in this thesis as a new identification system whose aim is to overcome all the inconveniences of the current proposals.

ELEKTRA is based on the work presented in [13]. It is a biometric system where the EKG is converted into an EKM by aligning a set of R_{peaks} forming a matrix and then plotting them as a heatmap. Hence, instead of implement user identification with the EKG signal, a CNN performs the classification of users through EKM images (see Chapter 2).

1.1 The EKG as a biometric trait

In cybersecurity, there are three types of identification means: something you know (a password, a PIN, or any personal information); something you have (i.e., a card key or a token); or something you are, a biometric trait [14]. In this thesis, we are going to see an in-depth study of the third one, what you are, the EKG.

1.1.1 Biometrics

Biometry Introduction

The term *biometry* comes from *bio-* as *life* and *-metry* as the process of measuring. There is not much consensus on when it was first seen. Still, it is thought that it was first introduced by the philosopher and historian of science William Whewell in 1831 meaning the calculation of life expectancy. Then, both terms *biometry* and *biometrics* became expressions of common use after the foundation of the journal *Biometrika* in 1901 by Francis Galton and Karl Pearson [15].

In addition, a biometric system is essentially a pattern-recognition system that recognises a person based on a feature vector derived from a specific physiological or behavioural characteristic that the person possesses [16]. Basically, in biometrics, you are your own password. There are many different types of biometric traits that can be used to access a system; these range from facial characteristics and hand geometry to vein patterns, voice analysis, or, obviously, an EKG [17].

Characteristics and Requirements of a Biometric System

All different types of biometrics are based on specific characteristics that all of them should have in common [18, 19]. These are:

- *Universality.* Every person in the target population should be able to access the system. Universality refers that each (live) individual must possess the biometric trait that is being used to authenticate that individual [20]. There might be some identification systems that are not universal for all the population as fingerprint recognition where a user may have their fingers burnt or no fingers at all.

- *Uniqueness.* The biometric trait should be sufficiently distinguishable across individuals. There must be enough evidence for a biometric trait to be different for each individual in the targeted population. For example, in the case of EKG, there exist more than twenty years of peer-reviewed scientific publications proving how each person has a unique EKG [21, 22, 23, 24, 25, 26].

- *Permanence.* The biometric trait should be sufficiently invariant over a long period. Some biometric traits may have difficulties coping with this characteristic as it is face recognition, which may need to re-enroll the participants through the passing of time.

- *Collectability.* The system should be capable to collect the biometric trait by employing suitable devices that can be later used to authenticate the user. An important factor of a biometric trait is how costly the sensors or devices used to collect the biometric trait itself are. Resistance to low sensors is important when developing new identification systems.

In addition, a practical biometric trait is also required to address some additional requirements [18, 19] such as the following:

- *Performance.* Achievable recognition accuracy. It is required for a system to not only achieve high accuracies in identification but also achieve low error rates to avoid impersonation.

- *Acceptability.* The biometric identifier should have broad public acceptance among the target population. Some biometric traits could have difficulties with the acceptance of the targeted population as could be a retina scanner or a DNA test. Others, like fingerprint or face recognition, are widely accepted among the general population.

- *Resistance to circumvention.* The biometric identifier should be challenging to bypass or counterfeit. Having a system that offers low error rates together with difficulties to copy or duplicate a person's biometric trait, makes it more complex to pursue an impersonation attack.

- *Privacy preservation.* Protect users' private information in the biometric templates. The biometric system proposed should protect the user's private data and should not allow replicating the user's biometric trait that is being used to authenticate the user.

Types of Biometrics

Generally, there are two main groups of biometric identifiers depending on their nature: behavioural and physical [27]. Firstly, in the behavioural biometric identifiers it can be found signature recognition [28, 29, 30], voice recognition [31, 32, 33], keystroke dynamics [34, 35, 35] and even touch-dynamics on mobile devices [36, 37]. These behavioural recognition modalities are based on the actions of human beings. Secondly, the physiological biometrics are the well-known iris [38, 39, 40], face [41, 42, 43] and fingerprint [44, 6, 45] and some others less popular such as finger veins [46, 47], ear [48, 49] and footprint [50, 51] which are based on physical characteristics of a human being.

Then, depending on the application, biometric modalities can be divided into the following three main groups [27]:

- *Commercial applications.* Such as e-commerce, the Internet, credit cards, smartphones, or wearable devices.
- *Government applications.* Such as national ID cards, driver's licenses, social security, or border control.
- *Forensic applications.* Such as corpse identification, criminal investigation, terrorist identification, and others.

Biometric System Types

In biometrics, there are two types of systems; the ones that perform user identification and the ones performing user verification.

First of all, before performing any identification or verification, enrollment of the user is needed. Enrollment is a phase that implies that the system needs to capture a biometric trait from the user which will be processed by this biometric system and stored for later comparison. Regarding the process of user identification with EKG in Section 1.2.1, this phase would include the data acquisition, signal preprocessing and, in some cases, feature extraction [52].

Once the enrollment is completed, a system can either identify or verify those enrolled users. The identification type (also called recognition) is a one-to-many matching, it occurs when the user's identity is unknown and the biometric system identifies a person from the whole enrolled population by searching in a database. Regarding verification, the biometric system authenticates a person who claims to be enrolled in the system in a one-to-one matching [52, 53].

1.1.2 EKGs

Regarding the EKG collection, the heart's electricity is detected by adhesive electrodes attached to the skin. The resulting measurements are referred to as leads. Over the past century, various lead systems have been developed, such as the ones from Einthoven, Goldberger, and Wilson [1].

In fact, the electrocardiograph was first recorded by Dutch physiologist Willem Einthoven in 1902 and gave physicians a powerful tool to help them diagnose various forms of heart diseases [54]. As in the XX century, adhesive

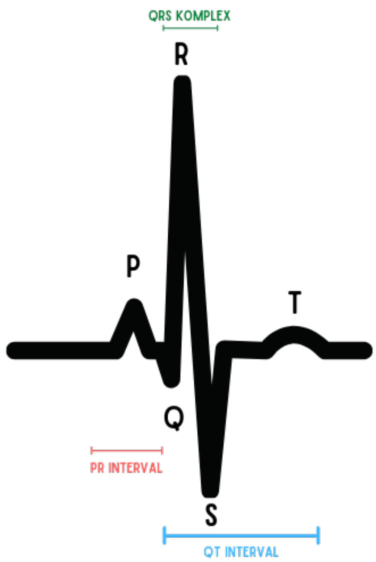


Figure 1.1: Electrocardiogram with its fiducial medical points

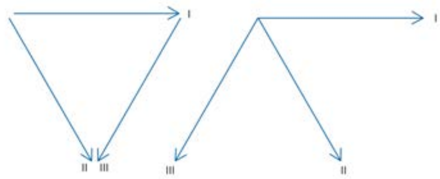
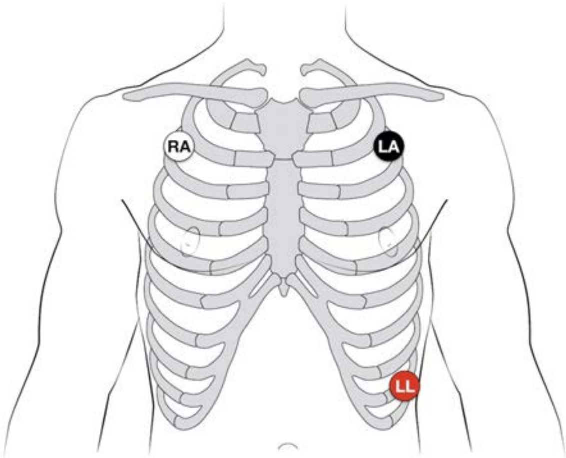


Figure 1.2: Einthoven's triangle. Position of leads to record an EKG [1].

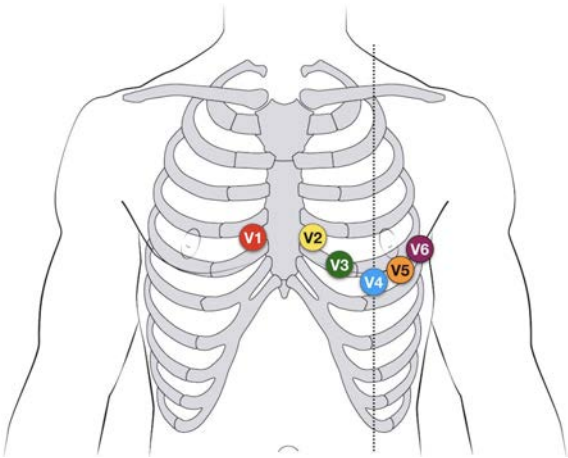
electrodes were not invented, Einthoven could measure the tension between the right and left arm (lead I), the right arm and left leg (lead II), and the left arm and left leg (lead III) by placing the extremities in a bucket of a salty solution (see Figure 1.2).

Then, the EKG was introduced by A. Goldberg in 1981 in [55] as a book for nurses, medical students, paramedical assistants, or any health care provider who read EKGs. The main goal of this book was to introduce how to properly read an EKG for health purposes. Before, in 1942, Goldberg, as the lead configuration that Einthoven had designed years ago had the leads too apart from each other, changed the angles where the leads were positioned by cutting them in half to improve diagnosis.

The current positioning for the EKG electrodes follows the image in Figure 1.3, Figure 1.3a for three electrodes and Figure 1.3b for ten electrodes producing 12 leads. A 1-lead EKG configuration is widely common for long recordings of 14 hours. The 3-lead EKG configuration is most often used for even longer recordings of 24 hours to, for example, diagnose heart problems.



(a) 3-lead EKG configuration.



(b) 12-lead EKG configuration.

Figure 1.3: Lead configuration for recording EKGs [2].

And the 12-lead EKG configuration is a system that combines Einthoven, Goldberger, and Wilson configurations. It is the most used configuration for recording both resting and stress EKG recordings [1].

The first work presenting the EKG as a biometric trait is done by L. Biel et al in 2001 [21]. Their experiments showed that it is possible to identify a person by features extracted from one lead only. Besides, one year later, T.W. Shen et al confirmed that it is possible to identify a person among a different group of candidates with 1-lead EKG [56]. The authors of this last work can identify 20 different subjects from the MIT-BIHDB with two different techniques: Decision Based Neural Network (DBNN) and Template Matching (TM).

Nowadays, the study of the EKG as a biometric trait is motivated by four key points [57]: 1) the collection of an EKG is a non-invasive technique, whereas an iris or retina scanner would be [58, 59]; 2) It is very rough against counterfeiting or spoofing practices because one can only be identified through its EKG if it is alive; 3) It is inclusive since all living beings have their own EKG; and 4) An EKG also provides additional information related to psychological states and physiological status, which can be interesting for some applications [60].

Since that, many works have considered the EKG in cybersecurity as a biometric trait [61, 62, 63] or even in other areas of cybersecurity application such as in a key generator [64] or a Random Number Generator [65]. As a result, there are more than twenty years of peer-reviewed studies confirming the feasibility of user recognition and identification with EKGs.

1.1.3 Metrics

To test and evaluate the different biometric models presented in this thesis, some metrics are going to be studied in this section.

One of the well-known metrics evaluated on biometric modalities is the accuracy of a model or an identification system, which is calculated as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{1.1}$$

being True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN)).

The quality of a biometric system is measured by evaluating its ability to accurately make accept/reject decisions over a user population. False Acceptance Rate (FAR) can be defined as the rate at which a non-authorised person is authorised as genuine:

$$FAR = \frac{FP}{FP + TN} \tag{1.2}$$

The False Rejection Rate (FRR) is defined as the rate of a genuine person getting rejected (Eq. 1.3) [66]:

$$FRR = \frac{FN}{FN + TP} \tag{1.3}$$

The security aspect of biometrics is associated with the ability to prevent false acceptance [67, 68]. False acceptance happens if FAR of the system is very high and the system is exposed to possible attacks. FRR depends on the number of users and the biometric system. High rates on the FRR on a valid identification model can be due to poor quality of image or signals or non-proper placement of the biometric sensor [66].

FAR and FRR are considered to be interdependant [14]. Hence, by plotting both metrics we obtain the Receiver Operating Characteristic (ROC) curve which allows determining another important performance metric referred to as a Crossover Error Rate (CER) or Equal Error Rate (EER):

$$EER = \frac{FAR + FRR}{2} \tag{1.4}$$

The EER is the point where the FAR is equal the FRR. In Figure 1.4 it can be found the graphic representing the EER point, where the FAR and FRR meet. The lower the EER is, the more accurate the biometric system [19].

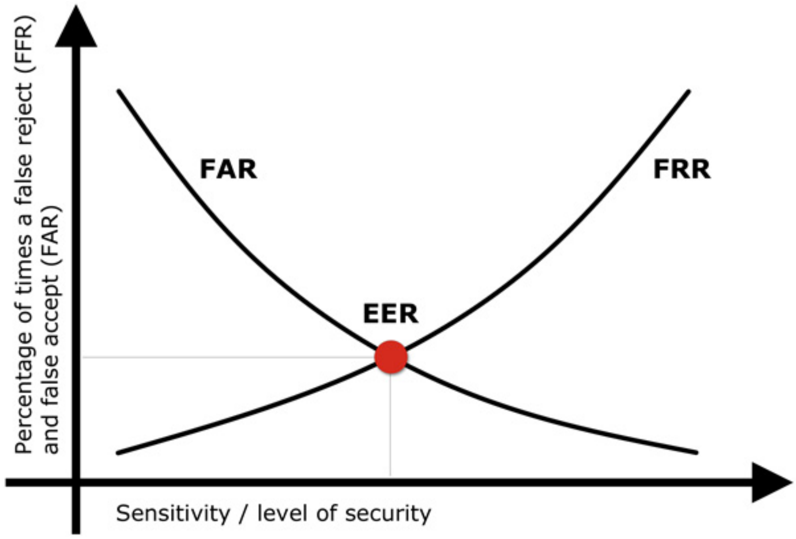


Figure 1.4: ROC Curve, where the EER is represented by the FAR and FRR meeting point [3].

Even though, from a cybersecurity point of view, in this thesis, the focus will be set over the FAR as the most important (in this area) factor is to prevent an impostor to acccess the system. Otherwise, a FRR can be prevented by blocking the system after certain number of attempts as many identification systems does. Thus, this thesis will not try to locate this EER point in a balance between FAR and FRR having almost the same values for each of these error rates.

Lastly, the Identification Rate (IR) is a very common metric used to evaluate a biometric system, is the rate at which a biometric subject in a database is correctly identified:

$$IR = 1 - \frac{FP + FN}{TP + FN + TN + FP} \tag{1.5}$$

1.2 Overview of Models and Contributions

In this section, a literature review of the current identification systems is offered. Different perspectives are going to be studied as it is: which is the process followed by an identification method step by step; different systems based on whether they use a fiducial or a non-fiducial analysis; a study of the subject through the identification with EKG based on intra-subject and inter-subject variability; a general view of the previous works with the EKG as an identification system; and, finally, all the preceding works with EKM as a system whether it is for identification of humans or medical diagnosis.

1.2.1 The process of EKG identification

To achieve a final classification result with a biometric system, the EKG signal must be subject to a certain process from the moment that is retrieved until a result or classification is obtained. This process is almost the same for all identification models, which is the one that can be seen in Figure 1.5.

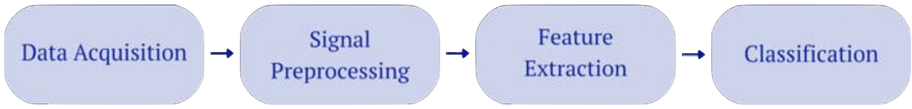


Figure 1.5: Conventional Identification Process of a Biometric System

The process consists of four different steps. The first one is to retrieve the data that will be necessary to perform the user identification. This data can be obtained by public or private means. In fact, one of the main obstacles of identifying users through their EKG is that many proposals used private databases in their experiments (e.g., [60, 22, 69]). Supporting this reasoning, in 2020, a comprehensive review of EKG-based solutions is presented in [70]. In this work, the author shows a list of the commonly used EKG databases, including information on whether they are public or private. From the twenty-one databases listed, six (30%) are public, but fifteen (70%) are private. There are just a few available databases on sites such as Physionet [71], one of the most used repositories of EKG records. Unfortunately, many works use private databases created ad-hoc for that research, which hinders the result's reproducibility. In this thesis, it is argued that in research all data used to perform experimentation of any kind should be public and available for everyone in the field in order to reproduce the results and test the veracity of the experimentation.

The second step of this identification process consists of de-noising, filtering the signal by removing the most common noises that can appear, and delivering it in a pure shape to extract features from it [72]. As in [18] is stated, "the noise sources and artifacts are characterized as interfering signals that emanate from anything that does not belong to the electrical activity generated by the heart". Hence, in this preprocessing phase, the main goal is to remove such noise and artifacts from the EKG signal to, in some cases, identify the following fiducial points: P_{onset}, P_{peak}, P_{offset}, QRS_{onset}, R_{peak}, QRS_{offset}, T_{peak}, T_{offset}, U_{peak} and U_{offset} [73]. In the works presented in [69] and [74] the fiducial points are used to perform user identification, therefore, preprocessing methods including the analysis and detection of fiducial points are needed. One of the algorithms that include signal processing with noise removal and the detection of the fiducial points is the well-known algorithm of Pan-and-Tompkins [75, 76, 77]. Other preprocessing techniques for EKG signals include wavelet transform such as [78] or even use the wavelet transform to process the signal and to detect the R_{peak} as in [79, 61, 80]. Other methods for noise removal might be Independent Component Analysis (ICA) [81, 82, 83], Adaptive Filtering [84, 85] or Singular Value Decomposition (SVD) [86, 87] among others.

The third step of the process of user identification shown in Figure 1.5 is feature extraction. There are two possibilities to extract the features of an EKG signal which can be based on handcrafted or non-handcrafted feature-based algorithms [18]. In both of these two methods, it also has to be taken into account if a fiducial or non-fiducial analysis is performed (Fiducial and non-fiducial analysis is explained in Section 1.2.2).
On the one hand, there are the works presented in [88, 89, 90] which have the QRS complex of the EKG signal extracted as a fiducial point with a handcrafted feature-based algorithm to perform user identification over the extracted features. Another example of feature extraction with handcrafted methods over fiducial points could be the work presented in [91] where they demonstrate that P, Q, R, S, and T peak wave also holds significant information for user classification. Then, handcrafted feature-based algorithms can also be used over non-fiducial analysis of the signal through, for example, Wavelet Transform such as [92].

On the other hand, there are non-handcrafted features-based algorithms most of which are based on deep feature extraction. Extraction of features with deep learning is grounded on letting the network or model decide which are the important features that represent the signal, in this case, the EKG signal. For example, a CNN performs feature extraction by exploiting the spatial correlation existing in the input data (the EKG signal in 2D) as it is done in [93] and [94]. An RNN performs feature extraction similar to a CNN but instead of focusing on the spatial correlation, the temporal correlation is exploited [83, 95]. Then, regarding statistical feature extraction based on non-handcrafted features, there is the Principle Component Analysis (PCA). A PCA performs dimensionality reduction by applying SVD (as in noise removal) to the input signal [96, 97]. Following the data science methods for feature extraction with a non-fiducial analysis, a data mining technique is possible as shown in [98] where the authors present a Continuous Authentication technique considering the EKG as continuous data streams.

Finally, the last step should be the classification or user identification itself. Many works use TM techniques which consist of comparing a set of features extracted from the signal, a vector, or what is being used to classify to pick a possible matching candidate such as in [56, 99, 100, 101]. Many proposals can also be found that use Machine Learning (ML) or DL algorithms to perform user identification and classification. Some of the most promising algorithms tested for user identification could be some of the ML algorithms such as Support Vectors Machine (SVM) [72, 102, 103], a Random Forest [104, 105] or Logistic Regression [106, 107], other deep learning algorithms as a k-Nearest Neighbour (k-NN) [108, 109, 110] or any Neural Network (NN) [93, 111, 112, 113, 114].

Despite the process shown in Figure 1.5 being the most common process for an identification biometric system, some works do not follow this structure. This is the case of the work performed in [23] or [115] where, thanks to deep learning techniques, a feature extraction step is not required and the classification is directly performed over the signal.

1.2.2 Fiducial and non-fiducial analysis of EKG

Regarding the analysis of the signal, fiducial and non-fiducial approaches are the two main alternatives:

A fiducial analysis of a signal consists of using the signal with markers. Concerning the EKG, these markers or points are the well-known waves or phases of one heartbeat as shown in Figure 1.1 and other relevant characteristics or points of the signal (e.g., PQ interval, RT interval, PQ amplitude or RS amplitude). An excellent example of a fiducial analysis is the study conducted by a medical professional when inspections on an EKG trace for heart diseases such as Atrial Fibrillation, Tachycardia, and others [116]. Fiducial analysis can also be found in studies where user identification is their primary goal, as in [74, 69, 117]. These studies use the fiducial points as the features to perform user identification instead of the whole signal or part of it. It can be found that many works achieve promising results by using fiducial points to identify users. Nevertheless, identifying and extracting the fiducial points from an EKG signal is known to be computationally more costly than performing a non-fiducial analysis of the signal.

On the other hand, the non-fiducial analysis does not focus on the signal's specific points or markers but on the whole signal. It could be considered a general analysis of the whole EKG signal or part of it. As the entire signal is taken and not some points of it, the EKG can be processed either in time, frequency, time-frequency, or even as an image. Thanks to that, the identification over the signal can be performed through k-NN, SVM or a CNN between others [118, 111, 25, 99, 62] resulting in systems less costly than the systems that need to extract each fiducial point of the signal. It is relevant to remark that extracting features with a NN is not a fiducial analysis. Those features extracted from a NN are the ones that explain the data itself, not the fiducial medical points represented in Figure 1.1.

1.2.3 Intra-subject and inter-subject variability of EKG

There exist many factors that can alter intra-subject and inter-subject variability of an EKG signal. In the case of inter-subject variability, one of the main factors is any Cardiovascular Disease (CVD) a subject can suffer. As it can be seen, there are many works that perform not user, but CVD identification through the EKG signal as the ones in [116, 119, 120, 121, 122, 123, 124]. In fact, in the particular case of the EKM, as it is commented in Section 1.2.5, all the previous work performed with the EKM is done to study the classification and diagnostics of CVD [125, 126, 127, 128]. With all these works it is manifested that CVD affects the EKG signal to make it distinguishable enough to identify users.

Whereas there exists plenty of evidence of inter-subject variability, there are few pieces of research studying intra-subject variability. Regarding intra-subject variability, many different factors can affect the EKG signal by hindering user classification. Any fluctuation in the heartbeat rhythm can alter the final classification of a user as it is a crucial characteristic of the EKG signal.

In the work presented in [69], the authors perform a study of user classification with a private database containing multiple states with different levels of anxiety for each user. Even though anxiety can affect the heartbeat rhythm the authors claim that the differences between subjects (inter-subjects) are more significant than the differences between the levels of anxiety of each subject (intra-subject), therefore, a classification of users can be performed regardless of the level of anxiety.

However, K. Kim et al. study the possibility of using normalized time-domain features of EKG to improve human identification through EKG [22]. They use a private database of 10 users with two states: resting and active. Hence, after some testing, they conclude that identifying a user which is active and, consequently, with a higher heartbeat rate, is more complex. Therefore, the authors reveal how intra-subject variability based on performing an activity affects the EKG identification.

Besides, another important factor altering intra-subject variability is how each heart ages. This fact is considered in longitudinal studies as [109, 60, 26]. All three works use databases that comprise more than one recording session. In the case of [60], the authors use a private database with three sessions recorded on different days. The best-offered approach of a longitudinal study for user identification is when training with two sessions and testing with the third one. Even though, the best results in terms of accuracy are obtained when training and testing with sessions on the same day. The same happens for the work presented in [26], the best result is achieved when the training and testing set is from the same session. And, of course, the work [109] happens the same. With these results, it can be concluded that recording the EKG in different sessions establishes differences in the intra-subject variability of the EKG. Therefore, the best approach would be to hold a database that comprises recordings from more than two days or sessions to train for two days and test over the third one to achieve a result as similar as we can to a real situation.

1.2.4 Previous works over EKG identification

In this section, a look over the past works studied and presented over different biometric systems of user identification through the EKG are commented. This section is focused on comparing different works which have performed user identification either with the databases studied in this thesis or with an approach with certain similarities to the presented identification method.

In the literature, a wide variety of approaches for identification using EKG records can be found. Some proposals use the raw signal after the signal preprocessing [23], and other works transform the EKG into an image or a spectrogram and process it as a picture [129, 93, 26].

In the case of [26], the authors convert the raw EKG signal into a spectrogram which is the visual representation of the energy of the signal expressed as a function of frequency and time. They tested their identification method over two databases, the University of Toronto Database (UofTDB) and the Check Your Biosignals Here Initiative (CYBHi). Their results are presented through the EER (see Eq. 1.4) when testing with both databases which achieve a 19.57% for the CYBHi and a 19.37% for the UofTDB.

Other possible solution to handle the EKG signal is to extract its fiducial points (as commented in Section 1.2.2) and use them directly to classify between different users as in the DETECT method presented in [130]. The authors tested DETECT using the database of the challenge titled "You Snooze You Win" [131], the offered error rates are low (FRR and FAR of 6.4×10^{-2} and 3.3×10^{-5}), and the accuracy is moderately high (i.e., 92%).

In the case of the work presented in [117] the authors combine a fiducial analysis with a non-fiducial one to perform user identification with EKGs. A 2D CNN is fed by the features extracted from a fiducial analysis combined with one beat of the EKG. They have tested their proposal over four databases. One of them is private and is not going to be commented on due to the lack of reproducibility and the other three are public databases that obtained accurate results with an accuracy of 99.69% with the MIT-BIHDB and an EER of 1.63% and 4.47% with the PTBDB and CYBHi, respectively.

Yefei Zhang et al. present an EKG identification method through a recurrence plot to avoid the time-frequency domain as they claim that EKG signals show substantial interclass similarity in the 1D time dimension. Therefore, user classification through 2D plots makes easier the recognition between users [132]. The authors have tested their proposal with a 2D CNN plus transfer learning. They have achieved a 99.17% of accuracy over the PTBDB and a 100% of accuracy over the CYBHi.

Another presented work that performs user identification through the transformation (S-transformation, for this particular case) of the EKG signal into an image to set it as the input of a CNN is the one presented in [133]. The authors capture the EKG trajectory at each time point in the frequency domain to obtain a 2D image. Their proposal is tested over two different databases. When testing over 50 healthy users of the ECG-ID database they achieve an accuracy of 96.63%. Then, they take 100 users of the 2017 PhysioNet/CinC Challenge database to perform a test on 50 users with atrial fibrillation (obtaining an accuracy of 96.23%) and 50 users with a noisy signal (achieving an accuracy of 96.18%). This result shows us, among other possible conclusions, that they can classify with the same results a healthy user, a user with a cardiac condition, and a user with a noisy signal demonstrating the capacity of adaptability of their proposal.

Given the outlook of identification of EKG using NN, there exist very different methods. One interesting work is [23] where Salloum and Kuo design an identification system based on cardiac signals and over different types of RNN. As the input of their NN, the authors use a vector of several consecutive heartbeats waveforms. The particular number of waveforms is a hyperparameter in the proposal. For experimentation, they use traditional RNNs, NN with LSTM units and with Gated Recurrent Unit (GRU) and test each identification proposal with two public databases: ECG-ID database [134] and the MIT-BIHDB [135]. The higher result achieved with the ECG-ID database is 100% of accuracy with an LSTM network and an input vector containing nine heartbeats. Next, the second-best result doing the same experiment but using only three heartbeats is 98.2% of accuracy. Similar results are achieved with the MIT-BIHDB.

Another striking example of identification through NN is the proposal in [93] where the identification is performed over a CNN. This CNN's input is an image containing a concatenation of several QRS complexes. This picture only contains a line (time series of several QRS segments), and the resting image is a blank space. Despite this, the performance, measured with database PTBDB [71], is comparable to the existing works as claimed by the authors as they achieve a 1.63% of EER.

Then, very similar to this last paper [93], the work presented in [94] can be found. They also implement feature extraction with a CNN and use the Euclidean Distance to calculate the difference between feature vectors. There also exist similarities with [113] because, in both studies, they use different databases to train and test the network. The PTBDB and MIT-BIHDB are the databases chosen by the authors of [94] to test the network. They achieve a 0.59% of EER over the PTBDB and a 4.74% of EER over the MIT-BIHDB.

A different approach of classification with NN is the one in [136]. Their used method to treat with the EKGs is the Second Order Difference Plot (SODP), a non-linear time-series analysis method that allows extracting features. The authors propose a Logarithmic Grid Analysis (LGA) as a quantification method of SODP. These techniques are applied to finally classify EKGs with a k-NN. Their method is tested over three different databases: the ECG-ID database, the NSRDB, and the MIT-BIHDB. Their best results in terms of accuracy are 91.96%, 99.86%, and 95.12% for each database, respectively.

Concerning the use of different databases to test the identification performance, in [63] Sidek et al. tested their proposal over databases with users having abnormal cardiac conditions. Those databases are the well-known MIT-BIHDB, the MIT-BIH supraventricular arrhythmia database (SVDB) and the Charles Sturt Diabates Complication Screening Initiative (DiSciRiDB). To perform the identification with EKGs the authors use a Normalize-Convolute-Normalize (NCN) technique over the EKG segments and then Bayes Network (BN), Multilayer Perceptron (MLP) and k-NN are used as the classification algorithms. Metrics such as Sensitivity (Se) and Specificity (Sp) or Positive Predictive Value (PPV) are used to assess the performance. In terms of accuracy the results are significantly high (i.e., 96.7%, 96.4% and 99.3% for MIT-BIHDB, SVDB and DiSciRiDB databases, respectively).

An interesting proposal using public databases is the one found in [115] in which Kim and Pyun propose an identification solution based on bidirectional LSTM and tested with two public databases (NSRDB and MIT-BIHDB). The authors use as input the LSTMs the EKG itself after filtering and normalising the EKG signal. They conduct experimentation analyses by varying the number of hidden units used in the LSTM (from 125 to 250) and the number of heartbeats (3, 6, or 9) used in the input. Outstanding results are achieved with the bidirectional LSTM with both databases. For example, with the proposed LSTM and the MIT-BIHDB, a 99.8% of accuracy is achieved when using nine heartbeats and three layers of hidden units.

Regarding research over the PTBDB as a public database, the work presented in [99] can be found. In this work, the authors perform user identification with a non-fiducial feature extraction algorithm. The authors implement both identification (with template matching) and authentication (with a random forest) of users. They tested their method over 50 healthy users and 50 patients with myocardial infarction. They achieve promising results for both classifications (for healthy and unhealthy users, separately). In fact, 49 of 50 healthy users were correctly identified and all of the non-healthy users were accurately identified achieving a 98% of accuracy over healthy users and a 100% of accuracy over non-healthy users.

The work presented in [113] performs user identification with a CNN.

Specifically, a CNN is formed by four uni-dimensional convolutional layers, one max pooling layer, and a fully-connected layer. The special interest of this paper is that the authors use different databases to train and test over their network. They train their network over the CYBHi and then test it over the PTBDB. When testing the network directly with the PTBDB not very good results are offered, but, after fine-tuning with the testing databases, their performance is improved as they achieve a 9.06% of EER.

Following the line of experimentation for user identification over the PTBDB, the work done in [129] can be found. A pre-trained CNN is used as the feature extractor method together with some more layers to perform user classification. To confirm the validity of their method, they compare themselves with a PCA with a k-NN and they achieve better results than this last method by finally achieving a 98.1% of IR.

In the work presented by Q. Zhang et al. in [24] the EKG segments are transformed to the wavelet domain to set them as the input of a 1D CNN. Eight different databases are compared with different CVD achieving high values of IR such as a 91.1% over the MIT-BIHDB and a 95.1% over the NSRDB.

Cascaded-CNN is the NN technique applied to the work done in [111]. In the presented work two CNNs are used: a first CNN for the feature extraction and a second CNN for User Identification. Their work is tested over five public datasets FANTASIA, CEBSDB, NSRDB, STDB, and AFDB achieving, for example, with the NSRDB an IR of 93.1%.

M. Hammad et al. presented in [137] a different approach of multimodal biometric identification combining EKGs together with fingerprint claiming that the EKG signal is better at rejecting impostors and the fingerprint is better at accepting genuine users. Hence, they have a three-step process where they first reject impostors with EKG authentication, then accept the genuine users with fingerprints, and, in the end, they use a fusion between the two inputs to finally authenticate each person. For the authentication part with the EKGs the authors transform the EKG signal into a two-dimensional EKG image by plotting each heartbeat as an individual greyscale image. The EKG databases that they use to test their method are the PTBDB and CYBHi achieving a 98.66% of accuracy with the PTBDB and a 98.97% of accuracy with the CYBHi.

Another work exploring multimodal biometric identification combining different metrics together with the EKG is the one presented in [138]. The EKG signal with the Galvanic Skin Response and Airflow are combined in an Identification System which offers outperforming results even in a real environment.

Recently, some works over attention models have been presented for user identification through EKGs, which is the case of [139]. The authors propose an end-to-end CNN with attention mechanisms without the need for any handcrafted processing, feature extraction, and classification stages. They achieve promising results as 98.85% and 99.27% of accuracy with the PTBDB and CYBHi, respectively.

1.2.5 Previous works over the EKM

The Electrocardiomatrix (ECM) or Elektrokardiomatrix (EKM) is a set of several aligned R-peaks of an EKG trace composing a matrix and then transformed into an image through a heatmap representation. It appeared for the first time in the work presented in [126] in 2015. In this presented work, the researchers use the EKM to manually inspect and detect different CVD. Therefore, this first work is far from using deep learning techniques to perform user identification. On 2018, some researches from the same lab that performed the first study with the EKM [126], worked on a new investigation studying the detection of Atrial Fibrillation (AFIB) and Atrial Flutter (AFL) with the EKM [125]. They achieved good results by detecting those CVD over the EKG heatmap and conclude that EKM is a reliable method for accurate identification for AFIB and AFL.

Later on the same year, a team of researchers in which Gang Xu was part of[1] presenting the work in [127] where the EKM is used to manually inspect the EKG validating and analysing the Heart Rate Variability (HRV) over two databases of users. They conclude that the EKM streamlines the identification of depressed HRV over a manual inspection of the EKG. Gang Xu et. al continue their research over the EKM in [140]. In this new study, they use the EKM to detect AFIB in stroke patients. They can demonstrate that superior accuracies are obtained when the EKM is applied to detect AFIB over traditional monitor analysis. Another research that studies AFIB with the EKM is the one in [141] where the authors grabbed three different databases to convert the EKG signals into EKMs by segmenting the signal every ten beats plus three seconds and feed it to a CNN. They conclude that the rhythm and morphological characteristics of an EKG signal can be learnt from the EKM when processed through a CNN.

To the best of our knowledge, the last work published in the study of the EKM (despite ELEKTRA [13]) is the one presented in [128]. In this case, the EKM is used to manually study patients with Congestive Heart Failure (CHF). The authors confirm that the EKM preserves all the features of cardiac electrical signals that can be decoded from raw EKG data in a simplified

[1]Gang Xu is one of the authors of the works presented in [126, 125] and [127].

way.

Therefore, all the previous works where the EKM is presented are based on studying different CVD. The work presented in this thesis, presumably, is the first one to use the EKM as a valid identification method.

1.3 Thesis Organization

The present doctoral manuscript is divided into six chapters that are shortly reviewed in the following paragraphs for a better comprehension of the document and the organisation of this thesis.

1.3.1 Chapter 2: Materials and Methods

The base of the whole thesis is shown in Chapter 2 where the materials and methods are presented.

ELEKTRA as a biometric system is tested over four different databases which are presented in this chapter. These four databases have been chosen due to their users' characteristics as there is one database with control users (NSRDB), two databases comprising users with CVD (PTBDB and MIT-BIHDB) and another database which its 25 users are performing five different activities requiring different cardiovascular efforts on each of the activities (GUDB).

Then, to create the EKM, the whole process is explained through the signal processing and the creation of the EKM per se. The formation of the EKM follows a set of steps from obtaining a list of R_{peaks}, segmenting the signal depending on the number of beats per frame (bpf) chosen aligning those peaks and segments, and, finally, plot the matrix formed into a heatmap to obtain the EKM image.

Once the process of the EKM is explained in detail, the establishment of new datasets conforming to all the EKM images for all users and databases is also described.

Finally, the architecture of the CNN used for user classification is detailed. All the hyperparameters are set and defined.

1.3.2 Chapter 3: Experimentation over Control Users

There are three chapters for the experimentation and demonstration of ELEKTRA's feasibility. Each of these chapters studies a different type of user. In this Chapter 3, the control or healthy users are exposed. The main goal of this chapter is to show how ELEKTRA is a capable and efficient method to work as a biometric system over a pool of control users. Hence, different aspects of ELEKTRA and these control users are studied to explore its efficiency and feasibility through different experiments such as the influence of bpf over the classification of users, the time costs of convergence or others. To do so, the NSRDB is studied.

1.3.3 Chapter 4: Experimentation over Users with CVD

In Chapter 4 the second chapter of experimentation is exposed. The MIT-BIHDB and PTBDB are the databases analysed in this chapter as healthy users together with non-healthy users (having a CVD) are mixed. The content presented in this part of the thesis allows studying how ELEKTRA behaviours in a scenario closer to a real-life situation where healthy and non-healthy users are present.

1.3.4 Chapter 5: Experimentation over users performing activities

The last chapter including experimentation is the Chapter 5 where the GUDB is studied. As commented before, the GUDB is a database combining 25 healthy users developing five different activities (sitting, walking, doing a maths exam, biking, and jogging). Each of these activities is supposed to involve a different heartbeat rhythm, hence, different results in terms of accuracy and error rates are obtained depending on the heartbeat rate. In this chapter, it is demonstrated how it is more complex to identify a user when his or her heart beats faster. This fact is studied during the whole chapter in different types of experiments by analysing each of the activities separately, merging different scenarios, or merging all of the scenarios.

1.3.5 Chapter 6: Related Work

In Chapter 6 a survey of the current literature regarding user identification with EKG compared to the proposal presented in this thesis is offered. Through this, ELEKTRA's performance is studied and contrasted against other different research regarding user identification. In fact, it is settled how [13] outperforms all research that has been studied for this thesis.

1.3.6 Chapter 7: Conclusions and Future work

The thesis is concluded in Chapter 7 where results and conclusions are summarised to show the reader how ELEKTRA contributes to the state-of-the-art by providing a new method for user identification with EKGs which is efficient, inclusive, reproducible, reliable and offers outperforming results in terms of accuracy and error rates. In addition, a discussion over different aspects of this doctoral thesis is offered. Finally, additional ideas for future development and application are provided to follow the research over ELEKTRA as the future identification system with EKG.

1.3.7 Appendix: EKM and CNN

In addition with the presented document, an Appendix is added to detail the two principal axis of this project: the creation of the EKM in Section A.1 and how the CNN is constructed in Section A.2. In those sections, different algorithms are showed to demonstrate how the creation of these two aspects is made. Mainly, in this appendix, it is shown the important aspects that have to be taken into account to reproduce this thesis.

1.4 Thesis Contributions

As a result of the aforementioned objectives, two main contributions have been accomplished during the development of this doctoral thesis.

C1. The first contribution to this thesis was presented in *Neurocomputing* journal, sent in 2021 and published in 2022. In it, ELEKTRokardiomatrix Application to biometric identification with Convolutional Neural Networks (ELEKTRA) is presented as a new identification method based on Elektrokardiomatrix (EKM). The presented biometric system is tested over three different databases: NSRDB, MIT-BIHDB and PTBDB. An in-depth study of the NSRDB is conducted as a way to show how ELEKTRA is a feasible identification method over control users. This research and demonstration is available in Chapter 2 where the Materials and Methods are presented and in Chapter 3 where the experimentation with control users is shown.

C2. The second contribution is an in depth study of ELEKTRA where our proposal is tested against different types of population as it is users with CVD (through the MIT-BIHDB and the PTBDB) and users performing different activities with the GUDB. The results and experimentation of these contribution can be seen in Chapter 4 where the experimentation of ELEKTRA over users with CVD are shown and in Chapter 5 where the results for studying ELEKTRA with users performing different activities are detailed.

CHAPTER 2

MATERIALS AND METHODS

To obtain an Elektrokardiomatrix there is a full process consisting of various steps from the acquisition of the database to the final plot of the image to then classify those images through the use of a CNN. All these steps can be seen summarised in Figure 2.1. To perform all these steps a set of materials and methods are needed and are presented in this chapter.

Hence, this chapter presents how and which databases of EKGs are processed, which are the steps of signal processing for noise removal and detection of R_{peaks} over the EKG, the process to create the EKM and which is the architecture used to perform user classification with a CNN.

2.1 Data

The experimentation of this article has been performed over four different databases. All of these databases are detailed in Table 2.2. These four databases have been chosen due to users' attributes[1]. In addition, it is important to remark that these four databases are publicly available (three of them by a free download in Physionet[2] and the other one by petition) to follow the principle in which the data used in research should be public to allow the reproduction of results.

The experiments were intended to focus on different aspects of different subjects to study the behaviour of the presented model from different perspectives. For this reason, a database with control users with no significant cardiopathies (NSRDB) has been chosen, as well as one database with patients with different CVD (PTBDB), two databases with both healthy users and non-healthy patients (PTBDB and MIT-BIHDB), and a database of healthy users in different scenarios (GUDB). This way, the approach presented in this thesis is studied and analysed in different conditions that simulate a real-world application.

The first database that is chosen and tested is the NSRDB, publicly available in Physionet [142]. The database contains 18 different healthy subjects

[1]See Section 7.1.4 where a discussion on why this databases are chosen and not others is proposed.

[2]Physionet [71] is a repository of freely-available medical research data managed by the MIT Laboratory for Computational Physiology.

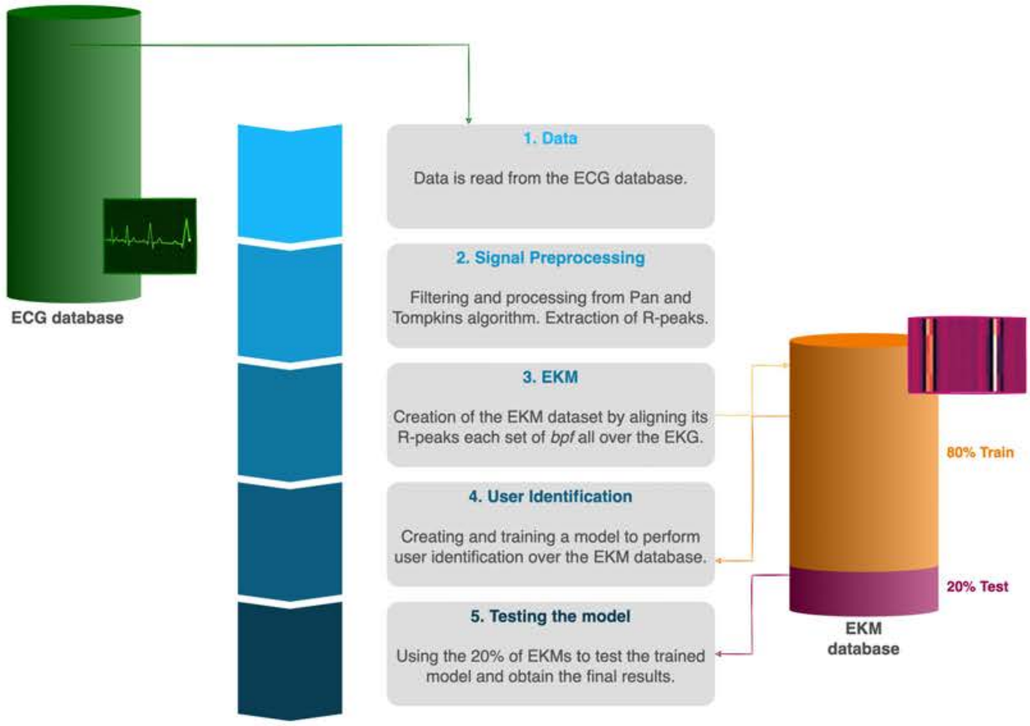


Figure 2.1: Pipeline showing the process of ELEKTRA including the steps for the creation of the EKM and the final classification of users.

(with no significant arrhythmias). Its records were obtained in the Arrhythmia Laboratory at Boston's Beth Israel Hospital. This database contains one long-term record (24h) per user. The NSRDB is the database comprising the largest recordings from the four databases.

The next database is the widely used MIT-BIHDB which is also publicly available in Physionet [135]. This database consists of 48 recordings from 47 different subjects. Of the pool of 47 subjects, there are 23 are random users, and the other 24 patients are considered to have a significant arrhythmia. Consequently, what is remarkable from this database is that there is a combination of healthy people with patients with CVD as happens in real-world situations.

The PTBDB is also a publicly available database from Physionet with 549 records from 290 subjects [143]. The important aspect of this database is that is composed of patients with different CVD and healthy subjects. In the original database, there are 52 healthy control users and 238 users with different CVD. This database is considered to be a strongly unbalanced dataset, having more than one EKG per user in some cases. It can be observed from Table 2.1 that, from the 290 subjects comprised in this database, there are 74 subjects with two or more EKGs, and there is even one subject with 20 EKGs. Furthermore, no description of when and how the different records were acquired is provided. For this reason, one random EKG per user has been considered in all the experiments.

# ECGs	2	3	4	5	6	7	10	13	20
# subj	44	16	5	3	1	2	1	1	1

Table 2.1: Number of users per ECG recordings for users that have two or more recordings from the PTBDB.

Some experiments of this presented work over the PTBDB will study the entire database and other experiments will study only some specific CVD without healthy subjects to show how ELEKTRA performs over users with those cardiac conditions. The CVD considered in the commented experiments are Bundle branch block, Cardiomyopathy, Dysrhythmia, Myocardial infarction, Myocarditis, and Valvular heart disease. Therefore, there will be two approaches regarding this database: processing all the data and processing the subjects with the commented CVD.

	Subject Information					Records	Sampling Frequency (Hz)
	Number	Male	Women	Ages	Pathologies		
NSRDB	18	5	13	20-50	Healthy	18	1000
MIT-BIHDB	47	26	22	23-89	23 random subjects 24 subjects with arrhythmias	48	360
PTBDB	290	209	81	17-87	Different Cardiac Conditions and healthy subjects	549	1000
GUDB	25	9	15	18-X	Healthy	125	250

Table 2.2: Databases information and characteristics.

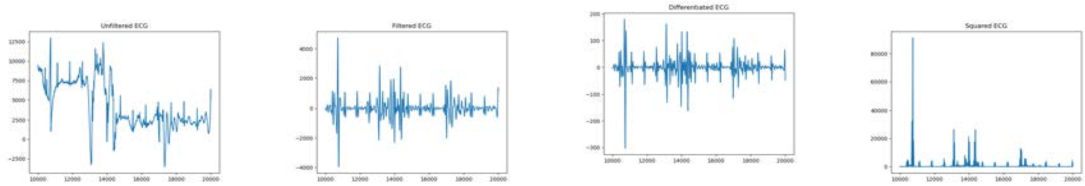
Lastly, referring to the studied databases, the GUDB has been considered. This database is available by petition to Glasgow University [144]. The special trait of this database is that all 25 users have been recorded in five different scenarios, which are: sitting, doing a maths test on a tablet, walking on a treadmill, running on a treadmill, and using a hand bike. Consequently, there will be five EKG recordings for each user. All participants from this database have not any significant CVD or health issues, they are considered to be healthy subjects.

2.2 Signal Preprocessing

The main goal of signal preprocessing is to obtain a clean ECG signal and a list of its R_{peaks} (see Figure 1.1) for each user. To do so, some steps are required.

To generate the EKM, the first step that has to be performed is the R_{peaks} detection procedure[3]. The *Biosignals* library, which provides a set of functions (e.g., filter or event detection) commonly used in the pipeline of biosignals

[3]Some of this steps are covered and detailed in Appendix A.1



(a) Unfiltered ECG (b) Filtered ECG (c) Differentiated ECG (d) Squared ECG

Figure 2.2: Preprocess of the ECG signal to obtain its R-peaks.

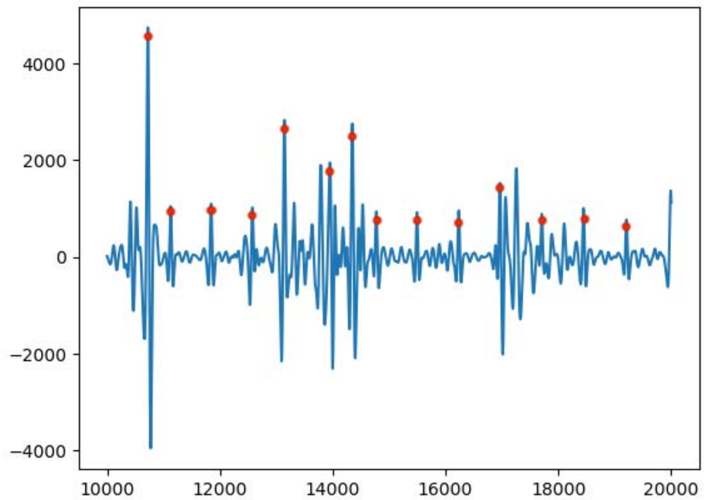


Figure 2.3: Filtered ECG with its detected R peaks.

processing[4] has been used.

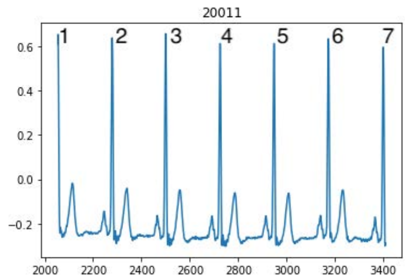
With the Biosignals Python library, the steps that have been carried out are the following: a) Bandpass filtering of the signal between 5 and 15 Hz; b) differentiation; c) rectification (squaring), and d.) moving-window integration. The application of all of these steps into an EKG can be seen in Figure 2.2. Figure 2.3 shows the result of the list of R_{peaks} obtained plotted over the EKG filtered signal.

Particularly, the well-known Pan and Tompkins (PT) algorithm has been used to generate the definitive list of R peaks [75][5]. This algorithm employs a dynamic threshold (a) Buffer initialisation; b) Detection of possible and provable R-peaks; and c) Exclusion of peaks ([75] criteria) and lag correction).

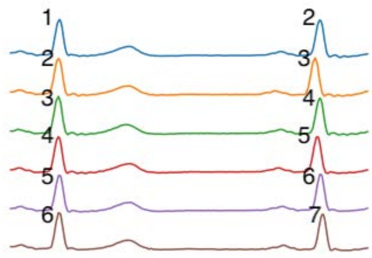
Using these steps with the EKG records in the database, a dictionary is built with N entries (being N the number of users of each database), and each entry contains a list of all its R_{peaks}.

[4] `https://github.com/biosignalsplux/biosignalsnotebooks`

[5] The PT algorithm is implemented in the Biosignals library.



(a) EKM I: Seven peaks plotted from an EKG



(b) EKM II: Alignment of seven peaks from an EKG or EKM

Figure 2.4: Steps I and II of the process of creating the EKM

2.3 EKM

As mentioned in Chapter 1, the EKM is a set of several aligned R-peaks of an EKG trace composing a matrix and then transformed into an image through a heatmap representation.

The EKM has already been proposed in several works for medical purposes [126, 127, 140, 128, 145]. In this thesis, the main goal is to show the feasibility of EKMs for identification (cybersecurity purposes). As far as the literature review reaches, it is the first time that this approach is proposed.

To create the EKM[6], first of all, a *window*, which is an EKG segment containing the number of beats declared to be on the EKM (see Figure 2.4 as illustration), is defined. To refer to this number of heartbeats (R-peaks), the metric used is called beats per frame (bpf), which is the number of different beats that there appear on each window, which is then taken to construct the EKM. The value of bpf is a hyperparameter that depends on the problem to address. In the presented experiments, the tests are performed over three bpf values: 7, 5, and 3[7].

Once the EKG is divided into windows, each window needs to be processed to obtain the matrix[8]. Precisely, the window is split into smaller segments containing two R_{peaks} forming each row of the matrix. As shown in Table 2.3, two values delimit each segment, where the μ parameter represents the average distance between beats for the entire EKG record of a subject[9]:

- Init Segment: Given by $p_x - \alpha_i\, \mu$ where p_x is the sample position of

[6]The implementation to generate EKMs is available at: `https://github.com/cfusterbarcelo/ELEKTRA-approach` and detailed in Appendix A.1.

[7]A discussion about how these values have been chosen to construct the EKM and not others is offered in Section 7.1.3.

[8]One EKM is obtained for each window processed. See Section 7.1.2 to see an evaluation of what could happen if an EKM was obtained for every heartbeat.

[9]The creation of the EKM segments has been discussed in Section 7.1.1 and detailed in Appendix A.1.

the $peak_x$ and α_i is a free hyperparameter indicating the percentage of samples that are taken before p_x. For all our experiments we set α_i to 0.2.

- End Segment: Given by $p_x + \alpha_e\, \mu$ where p_x is the sample position of the $peak_x$ and α_e is a free hyperparameter indicating the percentage of samples that are taken after p_x. We set $\alpha_e = 1.3$ in our experiments.

Init Segment	$p_x - \alpha_i\mu$
End Segment	$p_{x+1} + \alpha_e\mu$

Table 2.3: Segment Separation of the window

Each of the processed segments represents a row of the EKM. Note that all the peaks are aligned since the α_i, and α_e values are fixed for all the rows. The size of the EKM matrix is $bpf - 1$ rows with two peaks per row. To obtain the final image of the EKM representation, a heatmap needs to be computed. As illustration, Figure 2.5 is the representation of the plot seen in Figure 2.4b.

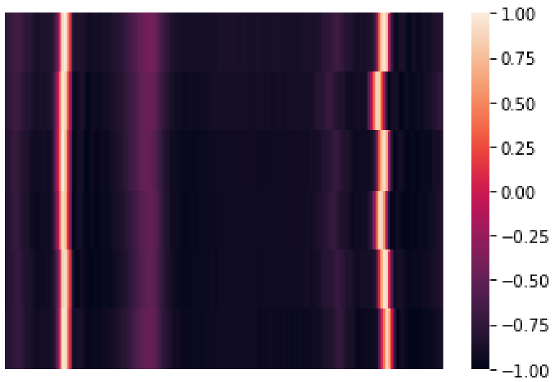


Figure 2.5: EKM III: Heatmap of the EKM

In Figure 2.6 the three representations of an EKM depending on each of the bpf parameters chosen for this work ($bpf = (3, 5, 7)$), can be observed. It can be noticed how, for each of the bpf chosen, the number of R_{peaks} seen in each picture, changes. Due to that, for example, for the EKM with 3bpf, the first peak appears once, the second peak appears twice and the third peak appears once too. This means that there is more representation for the first and third peaks than for the second. In the EKMs with 5 and 7 bpf the number of appearances per peak, is more balanced.

However, a CNN focus on the features that seem important for each of the classification problems (see Section 2.4). In this case, when classifying users through their EKM, it is not possible to know what considers the CNN that

(a) EKM to fed into the CNN with 3bpf

(b) EKM to fed into the CNN with 5bpf

(c) EKM to fed into the CNN with 7bpf

Figure 2.6: EKMs fed into the CNN depending on the bpf.

is more important. Consequently, it might or might not be important if the appearance of each peak is balanced or not.

2.4 Classification or User Identification

2.4.1 EKMs databases

Following the procedure described in Section 2.3, several EKM datasets are built, one per each bpf parameter and for the databases mentioned in section 2.1[10]. Considering the creation of the EKM datasets, there are two different types of databases depending on how large the EKG recordings are. For large EKG recordings, such as NSRDB, the EKM images are obtained until an upper threshold equal for each user in the database (i.e., 3000 images per user in our experiments) is reached. Otherwise, for shorter EKG recordings, such as the MIT-BIHDB or PTBDB, the highest possible number of EKM images are obtained until the signal is lost (see Table 2.4). Once the datasets are generated and users are identified (categorical classification problem), a CNN is trained as it will be explained later in Section 2.4. As conventional, the dataset is divided into several sets:

- Train dataset: From the whole dataset of EKMs, an 80% of the NSRDB is used to train the CNN and a 90% of the MIT-BIHDB, PTBDB and GUDB. Besides, from this training dataset another split is done in order to cross-validate the CNN parameters, creating:
  - Train dataset: (x_{train}). A 70% (for the NSRDB) or a 90% (for the MIT-BIHDB, PTBDB and GUDB) of EKM samples for each user.
  - Validation data-set: (x_{val}). A 30% (for the NSRDB) or a 10% (for the MIT-BIHDB and PTBDB) of EKM samples for each user.
- Test data-set: (x_{test}). It is a 20% of the NSRDB and a 10% for the MIT-BIHDB, PTBDB and GUDB. The number of images in this dataset split is the resting from the training ones. It is noteworthy that the trained

[10]The process of creating EKM datasets to be feeded into a CNN is detailed in Appendix A between sections A.1 and A.2.

model never sees the testing images. Testing images are only used to test and predict once the model is trained.

The length of the EKG recordings for the MIT-BIHDB, PTBDB and GUDB also influence the datasets created for each of the bpf values analyzed. As it can be seen in Table 2.4, the number of EKM extracted for each database depends on how large the EKG signal for each user is. Therefore, in the first place, the larger the EKG recording, the most number of images can be extracted (as for the NSRDB). Secondly, the higher the value of the parameter bpf fewer images are going to be obtained for each user and dataset as it happens for the other three databases.

Datasets	**#EKM**		
	3bpf	5bpf	7bpf
NSRDB	54,000	54,000	54,000
MIT-BIHDB	35,949	21,149	15,119
PTBDB	9,854	5,891	4,180
GUDB	8,099	4,816	-

Table 2.4: Number of EKM extracted per database depending on how large each EKG signal from each user is.

Finally, after computing the EKM datasets for the three bpf values, all the EKM heatmaps are resized before introducing them to the CNN to reduce computational costs. The raw image extracted from the EKG is 57x108 pixels, including all its blank space and the margins of the picture indicating the axis values. At this step, all images are resized into 29x54 pixels. Then, as it will be commented in Section 2.4, all axis and blank spaces are removed in the first layer of the CNN to make the model more efficient by not introducing useless information.

2.4.2 CNN Architecture

Nowadays, CNNs are the most used NNs for feature extraction in computer vision [146, 147, 148]. Furthermore, CNNs are widely used for feature extraction, specifically, for biomedical images such as in [149, 150, 151, 152]. Therefore, there exist many works that use a CNN to perform user identification over EKG images with a CNN as mentioned in Section 1.2.4.

The first intuition about a CNN was presented in [153] where it was discovered that a convolution function was able to extract features from spatial information. The first time that a CNN was presented was by LeCun in [154]. LeCun et al. used a CNN to extract spatial information about the well-known database MNIST [155] to perform classification of images.

A CNN is a cascaded NN which each layer is composed by convolutional filters. These filters or kernels are intended to find spatial structure in the image such as borders, edges, and contours. By setting a deep convolutional structure, a CNN can extract latent information regarding the spatial distribution of the data. This network provides a latent feature vector for each processed image. Finally, this latent feature vector is fed to a simple classifier composed of a fully-connected layer and a Sigmoid or Softmax activation (depending if the problem faced is a binary or categorical classification) to classify each image.

As a proof-of-concept solution to show the feasibility of using EKMs (extracted from EKG traces) for identification purposes, a simple CNN architecture that offers high performance in terms of accuracy and low error rates is offered[11]. In detail, the architecture of the CNN consists of just one convolutional layer plus a fully connected layer to classify users. It has been chosen to test all the work with a very simple CNN in comparison to the other state-of-the-art architectures to show that this approach can identify or classify users even with an elementary architecture (see Table 2.5). Thus, the focus of this approach is to introduce the EKM as a valid identification method, and for that purpose, in-depth and rigorous experimentation has been conducted. Figure 2.7 depicts the used architecture, specifying all dimensions and layers.

Model	**Year**	**# Parameters**
VGG-16 [156]	2014	138 Million
ResNet-50 [157]	2015	25 Million
Inception v3 [158]	2015	24 Million
EfficientNetB0 [159]	2019	5.3 Million
EfficientNetB7 [159]	2019	66 Million
ELEKTRA [13]	**2022**	**1.1 Million**

Table 2.5: Comparison of parameters needed to train in different CNNs of the state-of-the-art.

The first step that the network performs can be considered a preprocessing step. All input images are cropped to reduce the number of parameters and eliminate the image axis and white spaces, discarding irrelevant information. Again, the images that are fed into the CNN after this cropping step, can be observed in Figure 2.6.

Next, the second layer of the CNN is structured as follows:

- Conv2D: A 2D Convolution occurs where the kernel size is set to 3×3

[11] The construction of the CNN, which is the same for all experiments of this thesis, can be found in Appendix A.2.

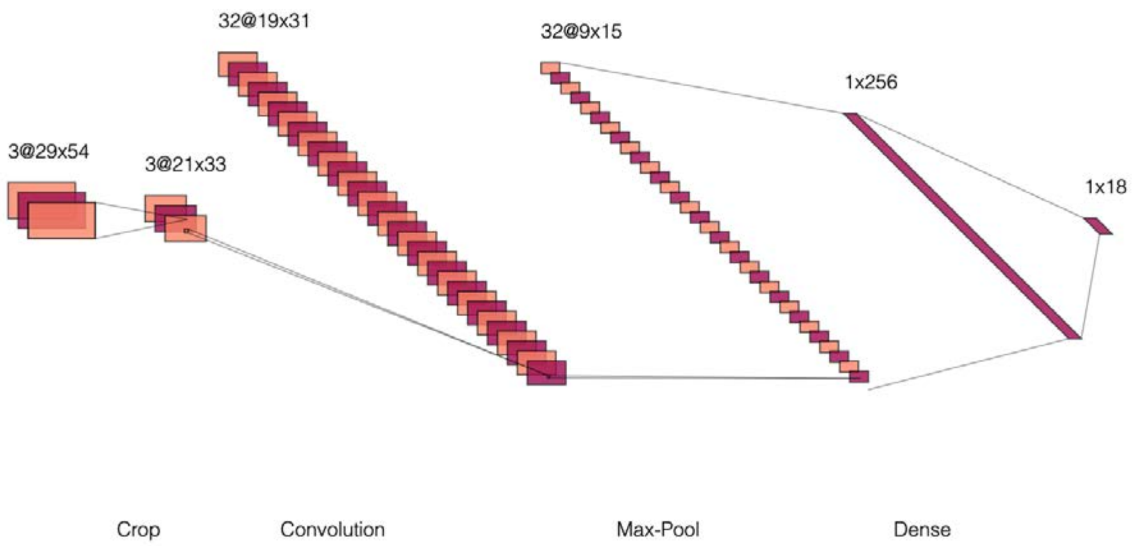


Figure 2.7: CNN Architecture

kernel, this means that a 3×3 filter is going to slide through the image with a stride of 1 extracting high-level spatial features.

- Relu: A Rectified Linear Unit (ReLU) function is going to create non-linear activation. That way, the ReLU selects which features extracted by the Conv2D are important for the next layer.
- MaxPooling: The Pooling layer reduces the size of the samples. A kernel of size 2×2 inspects the whole image and gets the maximum value of that kernel, downsampling the features extracted from the convolution to half of their input size.
- Dropout: A dropout of 70% takes place as a regularisation technique to prevent over-fitting after the MaxPooling operation.

If the CNN had another layer in order to make it more complex, it would follow the same intuition as the first layer. Nevertheless, the focus here is to show how the presented approach can identify users with satisfactory results with only one Conv2D layer showing that an uncomplicated architecture is enough to obtain those good results with the ELEKTRA proposal.

The second layer of the network is the Fully-Connected Layer (FCL) or classification layer. This layer aims to reduce the number of features and group them into the number of classes that have to be identified (18, 47, 290, or 25, depending on the database). Since this approach deals with a categorical classification problem, a *softmax* activation function is used to get the probability that an image belongs to a class[12]. The FCL consists on three steps:

[12] In some experiments a binary classification is performed, and some adjustments are needed, as the substitution of Softmax by a Sigmoid function and the Categorical Cross-Entropy (CCE) for the Binary Cross-Entropy (BCE)

- Flatten: Once the previous layers extract features, they are flattened into a column vector.
- Dense: Two dense layers are used to reduce the number of features extracted to the number of existing classes.
- Softmax: The last dense layer's output is activated by a *softmax* function to get the probability of belonging to each class.

A NN updates its parameters at each layer by optimising a cost or loss function during the training of the model. The CNN trains the model while a CCE cost function is optimised as a categorical problem is faced. For optimisation purposes, Adam optimiser is chosen since it is a widely used approach when training CNNs [160]. Due to the Graphical Process Unit (GPU) availability[13] the network has to be trained in batches, meaning that all images cannot be fed into the GPU at the same time, they have to be delivered in packages called *batches*. Once all the batches have been processed by the model, an epoch is completed. Hence, the experiments train the models with different numbers of epochs and steps per epoch depending on the chosen database. In the beginning, for some of the experiments, 50 epochs and 70 steps per epoch were set, lately, 50 epochs were not enough to train properly all the databases such as the MIT-BIHDB and PTBDB. Hence, databases have been trained for 150, 200, or even 250 or 300 epochs due to the small EKM datasets. Also, steps per epoch (see Equation 2.1) for these two commented datasets are changed depending on the number of training images:

$$Steps\,per\,epoch = \frac{length(train\,dataset)}{batch\,size} \tag{2.1}$$

where the batch size is set to 32.

After training each of the models used in each experiment using the training and validation datasets, the models only have to be tested with the testing images that the CNN has never seen. All the metrics specified in Section 1.1.3 are obtained for each experiment using only the testing images.

[13]GPU Tesla Pascal P100 PCIe by Artemisa [161]

CHAPTER 3

EXPERIMENTATION OVER CONTROL USERS

MOST of the studies in the literature perform user identification over control user to demonstrate their method efficiency through private databases or different public databases such as the NSRDB, CYBHi, *You Snooze you win* database, the ECG-ID DB or the UofTDB [115, 111, 132, 26, 137, 115]. All users from these databases are considered to be healthy. Specifically, they are considered to not have any significant Cardiovascular Disease such as Arrhythmia among others.

In this chapter, the Normal Sinus Rhythm Database (NSRDB) is going to be studied to perform user identification over control users. This database has been chosen due to its records' dimensions. Thanks to each record length, a large number of EKMs per user can be created and a more accurate result can be obtained. In fact, for each user, the extensive recordings allow us to obtain 3000 EKMs for each user regardless of the number of bpfs chosen. This means that there are 54000 EKMs or images to train and evaluate the model for each one of the experiments where the whole database is studied. In fact, this is the most extensive database that is held in this thesis.

All the experiments performed in this chapter are going to be conducted following the CNN architecture explained in Section 2.4 unless another architecture or parameter details are specified.

3.1 Influence of bpf over the database

In this first experiment, the focus of the study is the influence of how many R peaks are taken to identify a user over this database modifying the bpf parameter (i.e., $bpf = \{3, 5, 7\}$).

The CNN is tested with the produced datasets commented in Section 2.3. This experiment's main goal is to determine how many bpf (R_{peaks} in an EKM) are enough to identify an individual. Table 3.1 summarises the results of these experiments. Four common metrics are used to compare the proposals. Regarding the correct classification of instances, Loss and Accuracy metrics are studied. Concerning errors, False Acceptance Rate (FAR) and False Rejection Rate (FRR) commonly used in biometrics [162] are studied (all metrics have been studied in Section 1.1.3).

The results are very similar in the three tests with minor differences (see Table 3.1). The correct classification of instances is almost perfect in the presented system. An accuracy higher than 99.0% and loss in the range of 10^{-2} confirms this desirable property. Concerning errors, the FAR is below the FRR. That is, authorised access is K times more costly than locking out a legitimate user (i.e. $K \times FAR = FRR$). For example, the parameter K is equal to 15.5 for $bpf = \{3\}$, 18 for $bpf = \{5\}$ and 15 for $bpf = \{7\}$.

Since the differences are minimal between the three alternatives examined, as a trade-off between usability and performance, it is considered setting the bpf parameter to 5 as a fair value, and this is the value assumed in the subsequent experiments. This is due to that for $bpf = \{3\}$ (focusing on Figure 2.6a) it can be seen how the first and third R peak only appears once while the second R peak appears twice which may not seem fair to proper classification (see Section 2.3). In addition, for $bpf = \{7\}$ the problem that one may have is provided by the number of images obtained on a database depending on the number of bpf chosen. As manybpf are taken to form an EKM to obtain a dataset, fewer images are going to be obtained in a signal that is not as large as the ones that are being used in this experiment (the ones from the NSRDB).

3.2 Time Costs of Convergence

In this experiment, the focus is set on evaluating the time costs of convergence. Meaning that how to find a balance between the performance results and the time needed to achieve those results will be evaluated.

Hence, by focusing on the costs of convergence of this model together with this database (the NSRDB), the results by epochs for each of the bpf experiments in Table 3.1 can be seen. There is a tendency in which the 150 epochs are the best option. This is because training the model with 100 epochs tends to be under fitted and, with 200 epochs, overfits. For example, looking at the outcomes from the experiment done with 5bpf, we see that a 99.78% of accuracy is obtained with 100 epochs. Then, with a bit more training, with 150 epochs, a slightly better result is reached, which is a 99.82% of accuracy. Finally, by training the network during 200 epochs, the result suffers a slight decrease up to 99.69% indicating that the network is incapable of keeping on learning features about the model and suffers an overfit of the data.

Even though for other experiments the expected behaviour is to achieve better results once the network trains during more epochs, the lower result obtained in this experiment is for 3 bpf and 150 epochs. The cost function that is being used to optimise the network is the CCE which is a non-convex optimisation problem. Thus, local minima can be achieved. This fact is a possible explanation for why this result is obtained.

Otherwise, the best-obtained result is with 7 bpf and 150 epochs where an Accuracy of 99.84% is achieved. It is also remarkable that all values for FAR and FRR are shallow, indicating outperforming results with ELEKTRA.

bpf	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
3	100	0.0205	99.5	0.03	0.5
	150	0.0307	99.04	0.06	0.96
	200	**0.0115**	**99.69**	**0.02**	**0.31**
5	100	0.0104	99.78	0.01	0.22
	150	**0.0072**	**99.82**	**0.01**	**0.18**
	200	0.0123	99.69	0.02	0.31
7	100	0.0138	99.67	0.02	0.33
	150	**0.0057**	**99.84**	**0.01**	**0.15**
	200	0.0078	99.84	0.01	0.16

Table 3.1: Results obtained from the experiments done for the **NSRDB**.

On Figure 3.1 there are the Loss function optimisation plotted of training and validation with the CNN for 100 and 200 epochs and for $bpf = \{3, 5, 7\}$. It can be observed that all six images are very similar due to very similar results. Focusing on each result for the same bpf, it can be seen that training for 100 or 200 epochs makes no relevant change in the loss optimisation.

To summarise, very similar results are obtained regardless of the number of epochs chosen to train the network. That means that the number of epochs and, consequently, the time used to train the network is a parameter that can be modified depending on the requirements of the system. In other words, if having slightly high FRR (legitimate users may attempt two times to get access to the system) can be afforded, the use of fewer epochs is appropriate. If that FRR is excessively high for the target application, it is better to use more epochs in which both FAR and FRR are at a low level and a more robust model is offered.

3.3 One-vs-the-rest (OvR) Classifier

As an additional experiment, the system has been tested using the OvR approach. For instance, this solution can build the identification system of a smartphone. To make the system more resistant to attacks (e.g., replay or impersonation attacks), two classes are used (i.e., the samples of the target user and the rest).

A new dataset is computed following the approach described below. N EKM images of a user-q are randomly chosen for building the target class being $q \in \{0, 1, \ldots 17\}$. Next, the same number of randomly sampled images for the resting users are chosen (i.e., where $q \notin p$ being $p \subset \{0, 1, \ldots 17\}$). For the training phase, 2400 EKMs are used and 600 EKMs for the testing datasets ($N = 3000$ in our experiments), as in other experiments.

Concerning the CNN model, many of the model's parameters (e.g., epochs or dropout) introduced in Section 2.4.2 are maintained. However, some modifications are applied due to the problem's nature. The changes consist of using a Sigmoid as the activation function and a BCE for the loss function as a binary problem is being solved instead[1].

In Table 3.2 results of this binary classification are compared with the results from a multi-class classifier with 5bpf and 50 epochs [2]. In terms of accuracy, the differences are meaningless (a decrease of less than 0.2%). A worsening of the results concerning errors is observed although the system works in the EER point (i.e., $FRR = FAR$). This result is still appropriate for a biometrics system since it is still below a 1% [162]. In Table 3.2, we observe the normalised confusion table of this experiment, which clearly shows a tiny misclassification rate.

Classification	Loss	Accuracy (%)	FAR (%)	FRR(%)
Categorical	0.027	99.47	0.03	0.06
Binary	0.023	99.33	0.67	0.67

Table 3.2: Metrics obtained with the OvR experiment with user *14*.

0	**0.987**	**0.013**
1	**0.000**	**1.000**
	0	1

Table 3.3: Normalised Confusion Matrix from the OvR experiment.

3.4 Classification of a non-seen user

In this experiment, the performance of ELEKTRA against illegitimate samples is tested. In cybersecurity, this kind of attack represents an impersonation attack.

Similarly than in Section 3.3, a OvR classifier for each of the eighteen users in the NSRDB is built. For testing these models against impersonation, a user of the Brno University of Technology ECG Quality Database

[1] The differences between a binary and a categorical classification are explained in detail in Appendix A.2.

[2] Both classifications, binary and categorical, are trained for 50 epochs to simplify the experiment.

(BUTQDB) [163] has been randomly chosen. Three-thousand EKM images are generated for this user, using the procedure explained in Section 2.3. Note that there is no bias between users from the same database or between both databases since no significant heart diseases were discovered in any subjects.

After that, each of the models is tested with a set of illegitimate EKMs (i.e., 3000 EKMs of user 111001). The measurement performed is how many times each of the images are misclassified as a legitimate user. Next, the average value of these eighteen experiments is computed, representing the probability of success for an adversary to conduct an impersonation attack. Mathematically, assuming a target *user-p* that belongs to the *class-p* and an EKM^q sample of a *user-q* with $q \neq p$:

$$p(\mathcal{A}_I) = p(y = class\text{-}p|x = EKM^q) < \alpha \tag{3.1}$$

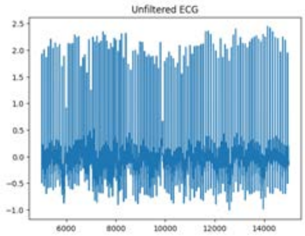
In the experiments conducted, the average α value is 0.077. It is a reasonable value for many applications. Note that it implies that the adversary, on average, only bypasses the system one in ten times. A reader can consider that this value is high but likely, many systems are blocked after a much lower number of attempts (e.g., ATM cards are blocked after three attempts[3]).

3.5 Classification of a Noisy User

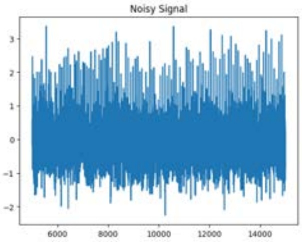
As a final experiment, the evaluation of the system performance when dealing with noisy samples takes place. With this experiment, the main goal is to demonstrate that not only the CNN used is a simple yet effective one, not only the images or EKMs offer a good approach for user identification but also the preprocess of the signal presents outperforming results as it can remove noise from a signal achieving great performance.

Hence, a random user from the NSRDB is chosen (e.g., the user-12) and Gaussian Noise ($\mathcal{N}(0, 0.05)$) is added to its EKG record. In Figure 3.2 it can be seen the process where, at first, there is the unfiltered EKG from a user, then, the noise is added and, finally, the EKG is filtered and processed to obtain a list of its R_{peaks} (see Section 2.2) to compute the EKM image (see Section 2.3).

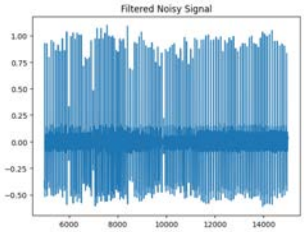
[3] e.g., `https://www.barclays.co.uk/help/cards/debit-card/wrong-pin/`



(a) Unfiltered EKG from user 12 of the NSRDB.



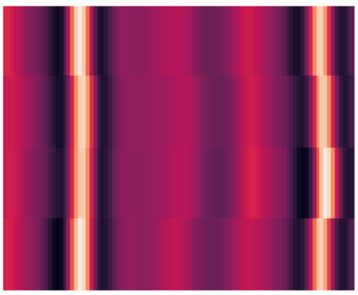
(b) Noisy EKG from user 12 of the NSRDB.



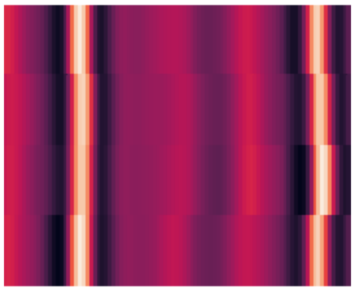
(c) Filtered noisy EKG from user 12 of the NSRDB.

Figure 3.2: Process of a noisy EKG: from an unfiltered random EKG to a noisy and finally filtered EKG.

To illustrate that there are minimum distances between the normal EKG and the same noisy EKG, there are the images in Figure 3.3. On the left side, Figure 3.3a shows the initial EKM from this random user without any added noise. On the right side, Figure 3.3b represents the same EKM as the one on the left but after adding the noise and processing it. The differences between these two images can hypothesize that there are not many dissimilarities between the two signals meaning that the proposed system (including the preprocess and obtaining the EKM) offers a strong performance against noisy signals.



(a) EKM num. 60 of user 12 from the NSRDB.



(b) EKM num. 60 of user 12 from the NSRDB after adding Noisy to its signal.

Figure 3.3: Differences between an EKM and its same EKM after adding noise to a signal.

As a result, these new instances are evaluated with the CNN network explained in Section 2.4. After testing 2400 EKMs of the randomly chosen user, the classification accuracy is 99.083%. The above result implies that the proposed system maintains its workability with noisy signals. Note that noisy signals are widespread when the EKG records are captured with low-cost EKG sensors [164]. In addition, Gaussian Noise can in fact be used to simulate noise from a real-world application as it is stated by the Central Limit Theorem in Section 7.1.5.

3.6 Feasibility test over the NSRDB

The last experiment performed over control or healthy users is a feasibility test in which different sizes of datasets are analysed to study how ELEKTRA performs when the different amounts of EKMs are used to train and test the model.

In these experiments, the split for training and testing datasets has been 80-20%. Moreover, the network has been trained during 100 epochs in each of the experiments and the batch size has been adapted to the different dataset sizes. The rest of the network has been implemented as it is explained in Section 2.4.

EKMs/user	Total EKMs	Acc (%)	Loss	FAR (%)	FRR (%)
3000	54000	99.65	0.0173	0.000207	0.003519
2500	45000	99.50	0.0203	0.000295	0.005020
2000	36000	99.53	0.0207	0.000276	0.004698
1500	27000	99.56	0.0194	0.000261	0.004444
900	16200	99.17	0.0479	0.000490	0.008333
600	10800	99.26	0.0452	0.000435	0.007407
300	5400	97.69	0.1564	0.001361	0.023148

Table 3.4: Results from the feasibility test where different sizes of datasets are analysed.

By taking a look at Table 3.4 the different results obtained and which datasets are used can be observed. On the first column there are the number of EKM per each user and on the second column the total number of EKMs[4]. Then, in the coming columns, the results for each of these experiments are shown in terms of Accuracy, Loss, FAR and FRR.

If the focus is set on Table 3.4 in general, there exists a tendency in which the fewer images that are used to train, the lower the results. The results in terms of accuracy are shown in Figure 3.4 where this tendency is plotted. Even though, all results obtained with all datasets are above a 97.0% of accuracy showing a really good performance. Even with a dataset with only 5400 EKMs and training during 100 epochs, ELEKTRA can perform user identification over control users with promising results.

[4] $Total\, EKMs = Num.of\, EKMs \cdot 18$

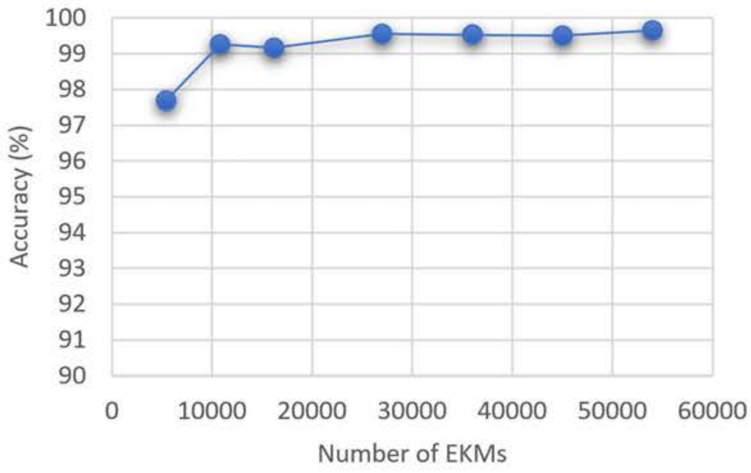


Figure 3.4: Graphic showing the accuracy results for the feasibility test where different sizes of datasets are analysed.

3.7 Discussion

After seeing all the experiments performed in this chapter over the NSRDB comprising 18 healthy users considered control users, it can be assured how ELEKTRA is a system capable of performing user identification with EKMs achieving promising results.

In this chapter, identification through control users has been studied and analysed in many different approaches. First, it has been demonstrated how the proposed system can identify users achieving an accuracy close to 100.0% and really low error rates when testing with three different parameters of bpf being $bpf = \{3, 5, 7\}$ shown in Table 3.1. Then, on the same table, it can be seen that less than 150 epochs are needed to train the network properly. Again, setting the training epochs on different values to study the time costs of convergence ($epochs = \{100, 150, 200\}$) the accuracy obtained in all the experiments is above a 99.0% claiming that with only 100 epochs, the model can train properly and can identify users with a good performance. In fact, these results allow playing between the costs of time and the results and error rates in a way that one can choose between the system that needs to develop depending on which factor of performance is more important: less training time or higher results.

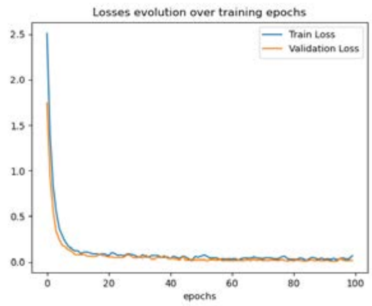
The experiment seen in Section 3.3 studied a one-vs-the-rest classifier showing a different approach where a binary classification takes place. The achieved results in Table 3.2 demonstrate that a real-life application where this kind of classification is applied is feasible.

Following a real-life situation experiment, a study of how ELEKTRA performs against impersonation attacks is performed in Section 3.4. A user from a different database was chosen to test the network against a non-seen user. The result offered claims that an impersonation attack would take place in 1 for every 10 attempts, which is not a high rate considering that many systems block their identifying attempts every 3 failed attempts.

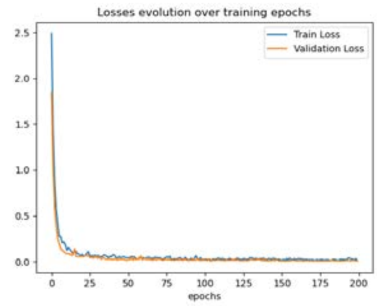
Between these situational experiments simulating a real-life scenario, there is the experiment presented in Section 3.5 where again, really good results are obtained when using a noisy signal to create the EKMs and classify the user.

Lastly, ELEKTRA has been tested to show its feasibility by changing the data size test that has been used in all other experiments for this chapter. The model has been tested against different sizes of datasets showing that even with very few images (300 EKMs per user) the presented model is capable of classifying users obtaining accuracy results above a 97.0% and really low error rates below 0.01% of EER.

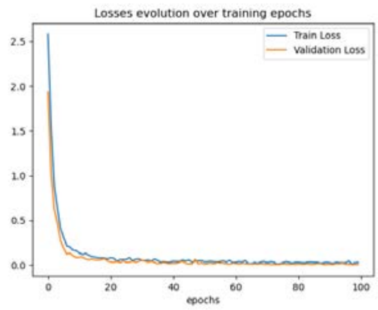
In conclusion, this chapter has demonstrated how ELEKTRA works and performs over a pool of control users without a CVD or activity performed in special. This chapter is the baseline to establish ELEKTRA's feasibility to implement user identification using EKG signals converted in to EKMs.



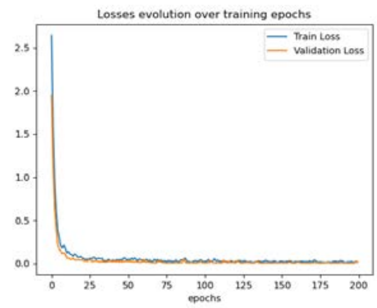
(a) Loss function optimisation during training and validation for **100** epochs with **3bpf** over the NSRDB.



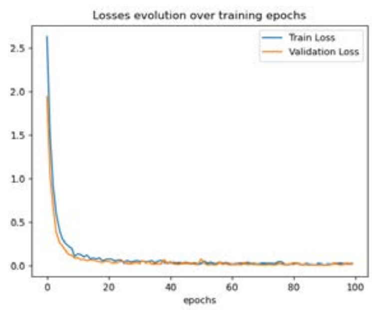
(b) Loss function optimisation during training and validation for **200** epochs with **3bpf** over the NSRDB.



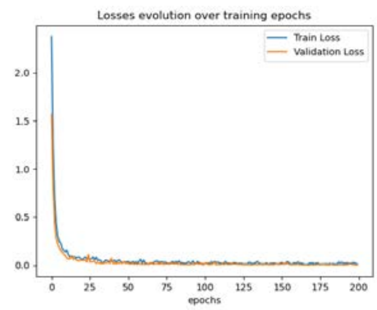
(c) Loss function optimisation during training and validation for **100** epochs with **5bpf** over the NSRDB.



(d) Loss function optimisation during training and validation for **200** epochs with **5bpf** over the NSRDB.



(e) Loss function optimisation during training and validation for **100** epochs with **7bpf** over the NSRDB.



(f) Loss function optimisation during training and validation for **200** epochs with **7bpf** over the NSRDB.

Figure 3.1: Loss function optimisation during training and validation for 100 and 200 epochs for EKMs with 3, 5 or 7 bpf over the NSRDB.

CHAPTER 4

EXPERIMENTATION OVER USERS WITH CVD

REGARDING identification of users with different Cardiovascular Disease (CVD), not many studies are performed where the focus lies on whether the user is healthy or not. In the work presented in [63], abnormal cardiac conditions over user identification are studied. However, in other works such as [133, 56, 63, 113] users with different CVD are included in their studies but it is not the main focus. In other words, there exist many studies including databases comprising users with CVD but it is not the main goal of the study, those databases are used without commenting on which CVD the users have or even taking care of the CVD itself.

In Chapter 3, the feasibility of ELEKTRA is tested over control users or users considered to be healthy. This chapter intended to go further by showing how ELEKTRA behaviours over non-healthy users or users with different CVD. To do so, two different databases, both containing users with CVD, were studied: the MIT-BIH Arrhythmia Database and the Physikalisch-Technische Bundesanstalt.

4.1 Experiments over the MIT-BIHDB

As it has been commented in Section 2.1, the MIT-BIHDB is composed of 24 patients with significant Arrhythmia and 21 random subjects. Hence, this database has been chosen to study the behaviour of ELEKTRA over a pool of users where patients and healthy users are mixed. The main goal of this experiment is to see if a CVD, as it is Arrhythmia, may interfere in the identification results where both users with and without cardiopathies are included.

First, the number of images or EKMs taken for this experiment must be considered. The EKG recordings for the MIT-BIHDB are not as large as the ones in the NSRDB. Consequently, as many EKMs as the ones obtained for the NSRDB cannot be obtained for this database. As it has been hypothesized, the more bpf that are taken, the fewer the number of EKMs will be obtained. This is shown in Table 4.1 on its first and second columns. For 3bpf there are 35949 images to train and evaluate the model. When the number of bpf taken to construct each EKM is increased, the number of images decreases as it happens with 5bpf and 7bpf where there are 21149 and 15119 images, respectively. Therefore, there is more or less half the number of images for 7bpf compared with the ones for 3bpf.

What is remarkable about this database is that, despite significantly decreasing the number of images regarding the NSRDB, promising results are achieved. As it has been commented in the previous experiment in Section 3.1, for this database which results are shown in Table 4.1, with 200 epochs, there is a slight overfit of the network but not in all experiments with a different number of bpf. It might be that, due to the smaller amount of images, for seven bpf, 200 epochs is the training that gives the best result. Not only for all EKMs with 7bpf, but for all the databases. Generally, results with 5bpf are the best for this database, but the best particular result is the one for 7bpf and a 200 epochs-training with a 97.89% of accuracy.

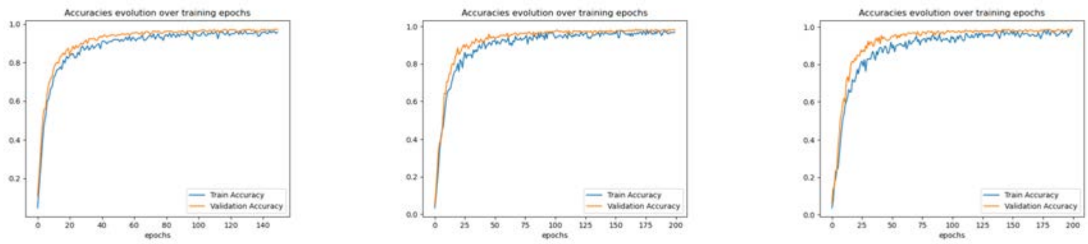
bpf	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
3	35949	100	0.1268	96.88	0.07	3.12
		150	**0.1160**	**96.80**	**0.07**	**3.20**
		200	0.1136	96.38	0.71	3.17
5	21149	100	0.1762	96.12	0.08	3.88
		150	0.1071	97.63	0.05	2.37
		200	**0.1004**	**97.89**	**0.07**	**3.27**
7	15119	100	0.1551	95.92	0.09	4.08
		150	0.1277	96.12	0.08	3.88
		200	**0.0934**	**97.89**	**0.04**	**2.11**

Table 4.1: Results obtained from the experiments done over the MIT-BIHDB.

After seeing the results for each experiment, if the focus is set on Figure 4.1 the performance evolution for each of the bests results commented on Table 4.1 are shown. It can be observed how, mainly, the three experiments with three different values of bpf, are very similar. There are not many differences between the training and validation evolution as well as with the results commented before, it can be affirmed that training the network during more epochs would not offer much higher results in terms of accuracy as it can be seen how the three training curves plateau since epoch 150 approximately.

Moreover, the three Confusion Matrix (CM) for the same three experiments are also offered in Figures 4.3, 4.4 and 4.5. These CM are the best results obtained when testing the model trained over the MIT-BIHDB with 3, 5, and 7 bpf. As it can be seen, mostly all users (as commented in Section 2.1, there are 47 users) are correctly classified obtaining an almost perfect CM. If the focus is set on the black part of any of the three CM, there are some misclassifications but, regardless of those users that were not able to be classified, the performance is still very accurate as seen on each image.

Seeing that all the results from the MIT-BIHDB where healthy users are mixed with patients with different CVD, are above 95% and we even obtain a



(a) Accuracy evolution during training and validation for the MIT-BIHDB with 3bpf and 150 epochs.

(b) Accuracy evolution during training and validation for the MIT-BIHDB with 5bpf and 200 epochs.

(c) Accuracy evolution during training and validation for the MIT-BIHDB with 7bpf and 200 epochs.

Figure 4.1: Accuracies for the best results obtained during training and validation with the MIT-BIHDB for 3, 5 and 7bpf.

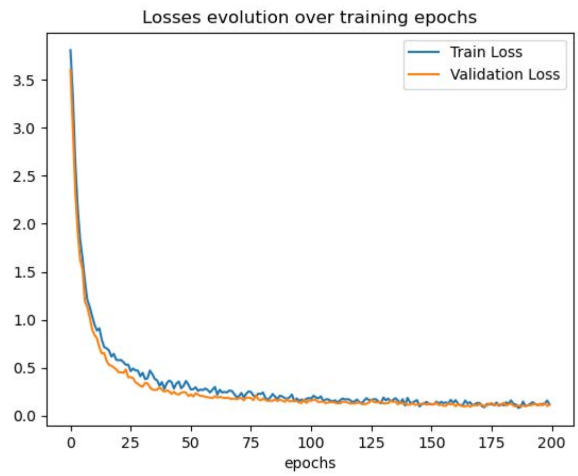


Figure 4.2: Loss evolution during training and validation for the experiment with 3bpf and 200 epochs from the MIT-BIHDB where a slightly overfit of the network is shown.

97.89% in terms of accuracy, it can be presumed that this model can classify and identify users with ELEKTRA regardless their cardiac health condition. Furthermore, in terms of FAR and FRR, the results obtained show that it is feasible to use ELEKTRA to identify users obtaining a very high performance with meagre error rates.

4.2 Experiments over the PTBDB

The central aspect of the PTBDB, as it is commented in Section 2.1, is that, similar to the MIT-BIHDB, some users from it have a CVD, and some have not. This database is built so that some users from it can be excluded or not, depending on the CVD that the user has. This fact allows the study of how

ELEKTRA works over a dataset where everyone has a CVD and compares it to other databases with different users.

Thus, in this section, two different approaches are studied. One approach is offered by processing the whole database and the other processed only some CVD to study the difference between performing user identification over a pool of users which all of them have a CVD and a pool of users where healthy and non-healthy users are mixed.

4.2.1 PTBDB Study

In Table 4.2, the experiments launched over the whole database can be seen. This includes different CVD such as Myocarditis or Dysrhythmia and healthy subjects. This might be similar to the experiments performed over the MIT-BIHDB where healthy users and patients are mixed.

Two hundred thirty-two users comprise this database, but with small EKG recordings; for this reason, the amount of EKMs that are held in these experiments is even smaller than the previous ones, which might affect the results.

Regarding the results obtained with this experiment (found in Table 4.2), it can be observed that with less than half of the images as the ones from the MIT-BIHDB, a really good performance achieving an accuracy up to 93.96% for 3bpf is obtained. Hence, in terms of accuracy, it has only decreased as much as 6 points with, remarking again, less than half of the images. This means that the proposed method is feasible enough that, having reduced drastically the number of EKMs, is able to perform user identification achieving promising results.

Focusing on FAR and FRR, the results for the FAR are very low, resulting in a high performance when a user wants to be identified and it is whom it claims to be. However, the FRR is not as good as the FAR. This means that when one user wants to be identified, it might be possible that this user is going to be rejected 1 of every 10 times that they may try, which may be considered a low ratio.

bpf	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
3	9854	150	0.4454	88.55	0.05	11.45
		200	0.4214	89.8	0.05	10.2
		250	**0.2831**	**93.96**	**0.03**	**6.04**
5	5891	150	0.7463	82.85	0.08	17.15
		200	0.512	87.27	0.06	12.73
		250	**0.4594**	**87.44**	**0.06**	**12.56**
7	4180	150	1.2036	71.98	0.15	28.02
		200	0.9239	79.12	0.11	20.88
		250	**0.5128**	**87.64**	**0.07**	**12.36**

Table 4.2: Results obtained over the experiments done with the **PTBDB** with 3, 5 and 7bpf.

In Figure 4.6 there are the three plots for the best results in terms of accuracy obtained for 3, 5 and 7 bpf over the whole PTBDB. By taking a look at the three figures as a process, it can be noted that as higher the number of bpf chosen is, the more spiky the plot is. This might be due to the number of images. As higher the number of bpf chosen, the fewer images are held to train and test the network. Therefore, the network can have more difficulties to generalise a proper model to perform user classification.

Based on the results achieved by processing the whole PTBDB, it can be concluded that ELEKTRA is totally capable of identifying users when different types of CVD are mixed with healthy users with even not a significantly large database.

4.2.2 Study over users with CVD

When studying the PTBDB with only users with CVD 162 users are held. Consequently, the number of images obtained to train and test the model decreases even more.

Generally, focusing on Table 4.3 where all these results are shown, better results with this database segmentation are achieved than with the whole database. This might be because it could be easier to identify users that all have CVD due to the dissimilarity of EKG recordings. As it is shown in other studies, it is possible to identify different cardiac conditions through the study and classification of EKGs or EKMs [165, 166, 125, 128, 140, 116]. That implies that each EKM changes depending on the user if they have any CVD. In conclusion, it may be easier to identify a pool of users where everyone has a different CVD due to the inter-subject variability of users with CVD.

Then, regarding the experimentation, one more experiment has been added.

As shown, for this dataset, only 250 epochs to train the model was not enough to train it properly and achieve the best results. This way, the higher results for 3 and 5 bpf were achieved with 300 epochs. For the experimentation with 7bpf, 300 epochs were too much, and the model suffered a slightly overfit of the network. Thus, the best results are achieved with 250 epochs. The results with 7bpf are not as good as the ones with 3bpf, but this might be due to, again, the number of EKMs that are disposed to train and test the network as it is shown in the Table 4.3.

bpf	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
3	7266	150	0.4445	89.61	0.07	10.39
		200	0.2703	93.91	0.04	6.09
		250	0.2302	96.4	0.02	3.6
		300	**0.192**	**97.09**	**0.02**	**2.91**
5	4350	150	0.6338	84.13	0.1	15.87
		200	0.4225	88.91	0.01	11.09
		250	0.3125	93.04	0.05	6.96
		300	**0.2124**	**95.0**	**0.03**	**5.0**
7	3066	150	1.4065	65.1	0.25	34.9
		200	0.856	81.21	0.13	18.79
		250	**0.5645**	**88.26**	**0.09**	**11.74**
		300	0.5114	86.58	0.1	13.42

Table 4.3: Results obtained when testing over the **PTBDB** for patients with CVD.

In Figure 4.7 there are the accuracy plots for those best results commented above. In figure 4.7a it can be seen the best result for 3bpf, the accuracy evolution during training and validation of the network is observed. The same plot is shown for 5bpf (Figure 4.6b) and 7bpf (Figure 4.6c). As it is commented in the above section, as more bpf and fewer images are used to train and test the network, the plot appears to be spikier. Comparing the images with the ones in Figure 4.6 where all the database is studied, the plots in this section are spikier than the ones before. Owing to having fewer images because not the whole database is processed in this experiment and only some CVD are studied.

Based on the experiments performed in this section where the PTBDB is studied, it can be affirmed that ELEKTRA can identify a pool of people with and without different CVD. Moreover, based on the experiments where only the users with CVD were taken, it is even easier to identify those users if they all have any CVD.

4.3 Discussion

As a result of all of the experiments done in this chapter based on the two databases comprising healthy and non-healthy users, it can be affirmed that ELEKTRA can perform user identification with EKM achieving promising results. Moreover, the MIT-BIHDB is one of the most used databases to test biometric systems with EKGs, by achieving an accuracy of 97.89% it demonstrated how ELEKTRA is a compatible and efficient method compared to the literature (see Chapter 6).

As explained before, studying the PTBDB with the whole database permits us to compare this study with the rest of the proposals using this same database. And then, studying it with only some of the CVD shows how ELEKTRA can obtain very good results when performing the classification of users with cardiopathies which may interfere in the creation of the EKM.

To sum up, this chapter shows ELEKTRA's feasibility for the classification of users with CVD and a mixture of healthy and non-healthy users. This approach is poorly studied in the literature and this chapter offers an in-depth study of how user identification is affected by a CVD.

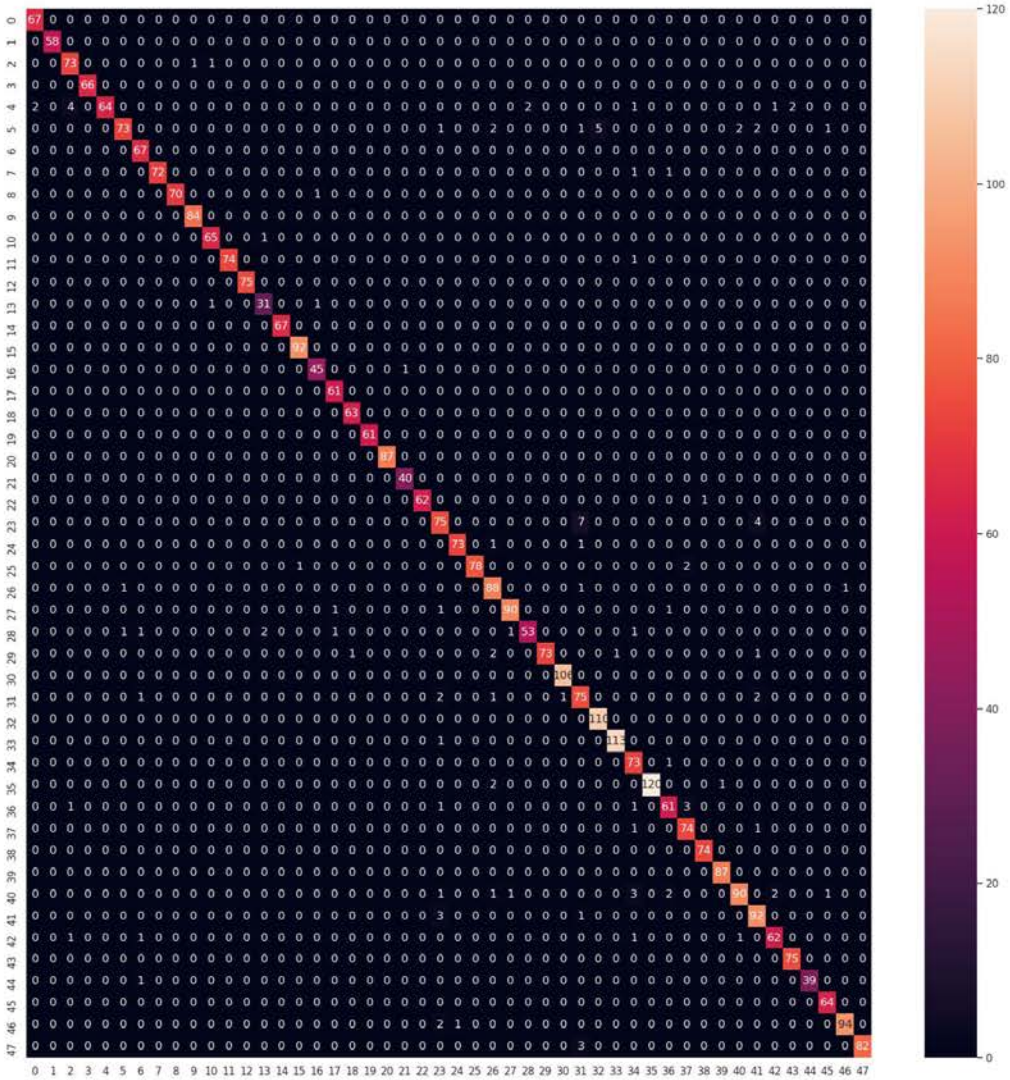


Figure 4.3: Confusion Matrix obtained when testing the trained model with the MIT-BIHDB for 3bpf and 150 epochs.

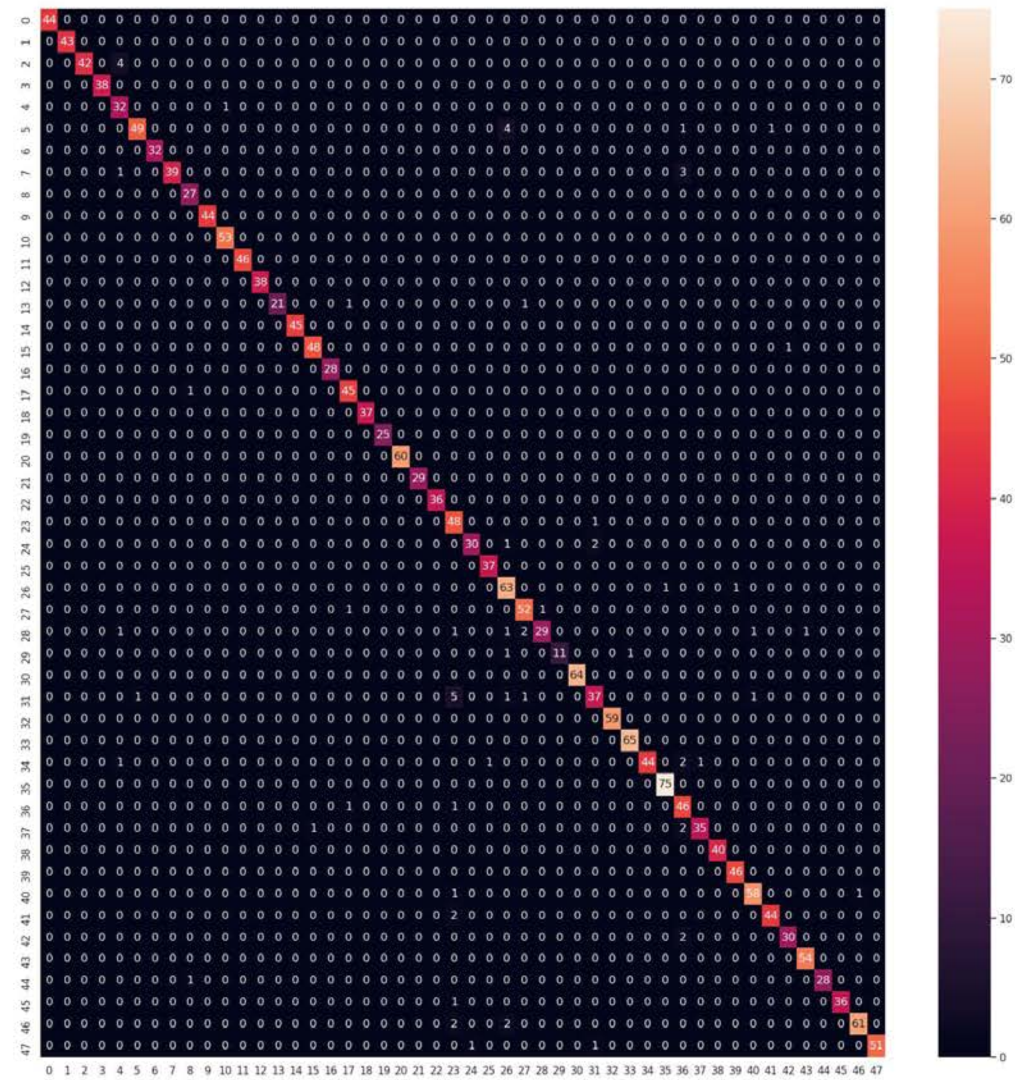


Figure 4.4: Confusion Matrix obtained when testing the trained model with the MIT-BIHDB for 5bpf and 200 epochs.

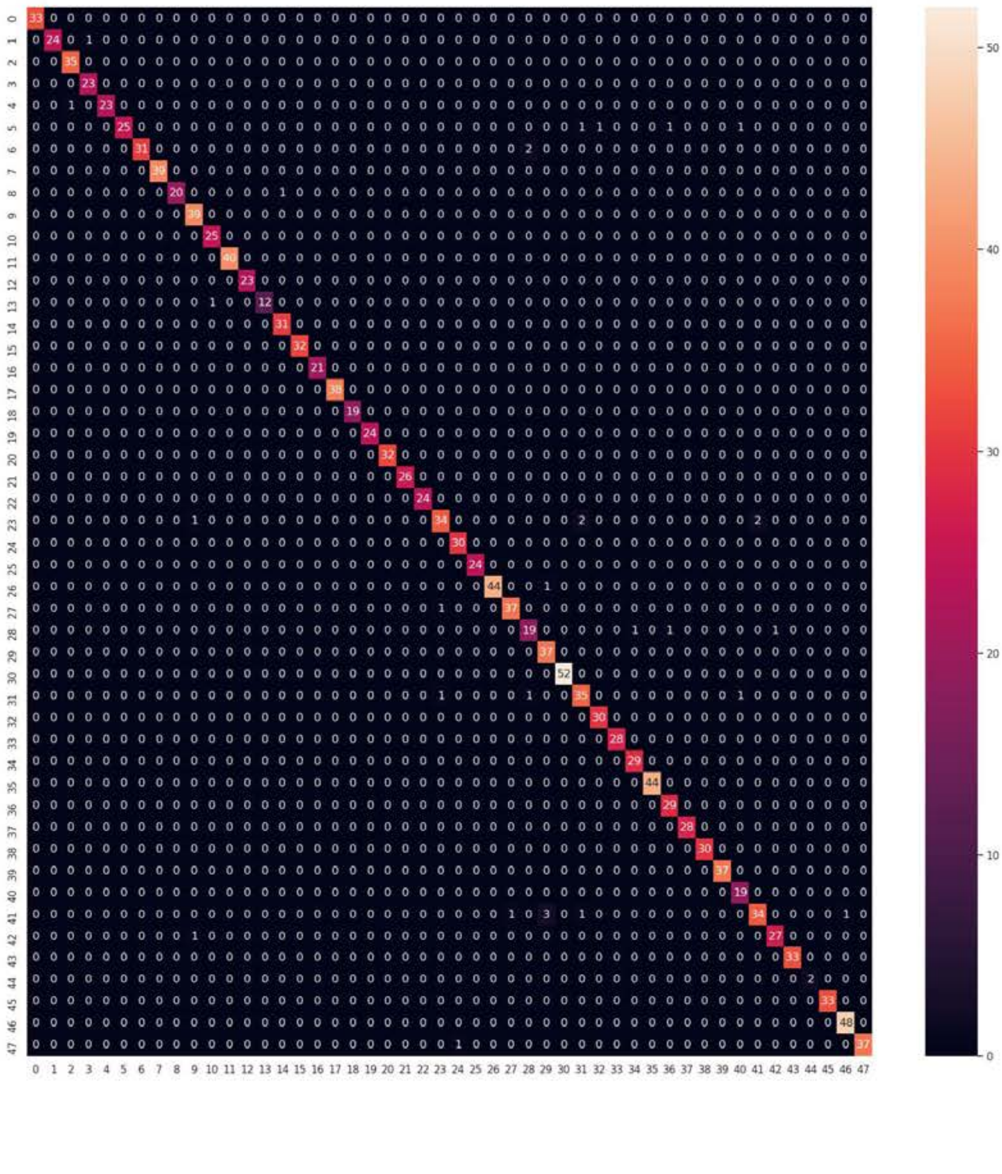
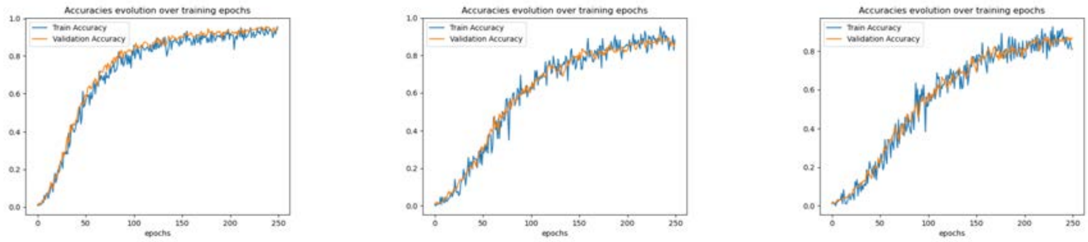


Figure 4.5: Confusion Matrix obtained when testing the trained model with the MIT-BIHDB for 7bpf and 200 epochs.

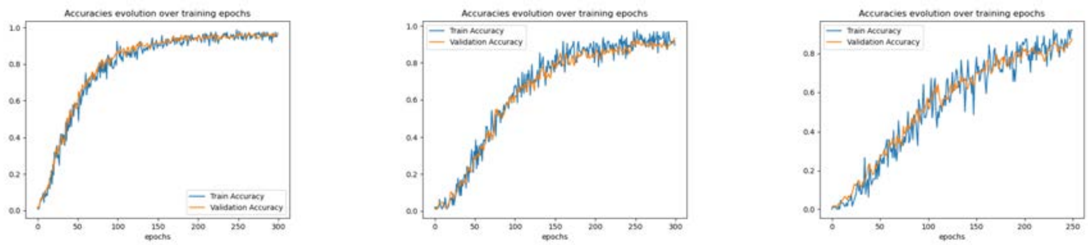


(a) Accuracy evolution during training and validation for the PTBDB with 3bpf and 250 epochs.

(b) Accuracy evolution during training and validation for the PTBDB with 5bpf and 250 epochs.

(c) Accuracy evolution during training and validation for the PTBDB with 7bpf and 250 epochs.

Figure 4.6: Accuracies for the best results obtained during training and validation with the PTBDB for 3, 5 and 7bpf.



(a) Accuracy evolution during training and validation for the PTBDB for CVD with 3bpf and 300 epochs.

(b) Accuracy evolution during training and validation for the PTBDB for CVD with 5bpf and 300 epochs.

(c) Accuracy evolution during training and validation for the PTBDB for CVD with 7bpf and 250 epochs.

Figure 4.7: Accuracies for the best results obtained during training and validation with the PTBDB over users with CVD for 3, 5 and 7bpf.

CHAPTER 5

EXPERIMENTATION OVER USERS PERFORMING ACTIVITIES

Differences in the heartbeat rate when performing user identification with EKG have been studied several times in the literature such as in [69] or [22]. Unfortunately, in the literature, there is no work performing user identification with the Glasgow University Database (GUDB) where various states of 25 users are presented (as commented in Section 2.1).

In this chapter, the GUDB helps to understand how ELEKTRA works in different scenarios and, consequently, how cardiovascular activity or differences in the heartbeat rate can affect user identification with EKGs.

It is important to remark that the number of EKMs obtained to perform the user identification is very low. This can affect the identification because the CNN has insufficient data to properly train the model with high performance for all the scenarios as for the other databases. Thus, the main goal of this chapter is to study how those different activities affect user classification and compare the experiments between them.

As it has been commented in Section 2.1 the GUDB works with 25 healthy users in five different scenarios. The EKG of each participant was recorded when they were performing these five activities: sitting, a maths test on a tablet, walking on a treadmill, running on a treadmill and using a hand bike. It is understood that there is a different heartbeat rhythm and behaviour in each of these scenarios. For example, *sitting* is going to be considered an activity with a low heart rate and running a cardiovascular activity with a higher heart rate. Therefore, as commented, the main goal of using this database is to compare the experimentation between scenarios to evaluate the performance of ELEKTRA when the user performs different activities and, consequently, has different heartbeat rates.

5.1 First approach to the GUDB

In this first experiment, each of the possible scenarios is going to be studied separately. This way, an approach of how ELEKTRA works over each scenario and, consequently, over different heartbeat rates, is offered.

Focusing on Table 5.1 where the experimentation of the GUDB with 3bpf is presented, it can be seen that all results in terms of accuracy are above

74.58%. Considering the number of images used in this experiment, it can be considered a good result. Even a 99.19% of accuracy for one of the experiments is achieved.

Hence, if the results obtained per scenario are analysed, it can be noted how the best result is obtained for the *sitting* scenario, where the heart rate is supposed to be calmer (a 99.19% of accuracy). In contrast, the worst result is obtained in the *jogging* scenario, which can be considered cardiovascular training with a higher heart rate (82.63% of accuracy). The difference in accuracy between *sitting* and *jogging* is about 17 points, which is quite a difference if it is compared to past experiments and considering that there are more images for *jogging* than for *sitting*. This result presumably means that it might be easier to identify a user resting with a lower heart rate than when the heart beats faster.

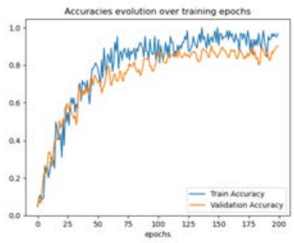
Better results are obtained when the user's heart rate is slower could also explain the rest of the results from the same Table 5.1. The *walking* scenario, which could be considered calm if not walking too fast, also has high accuracy and low error rates. Then, for the scenario where a user is *taking a Maths exam*, one could think that a user may or may not be nervous depending on each experience of the user resulting in different error rates and accuracy for each user. And then, the hand-bike activity could be considered between a calm scenario and a cardiovascular activity due to the results obtained.

Scenario	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
Hand-bike	1529	150	0.2721	94.87	0.21	5.13
		200	**0.2504**	**95.51**	**0.19**	**4.49**
		250	0.1918	94.23	0.24	5.77
Jogging	2205	150	0.8128	74.58	1.06	25.42
		200	**0.6302**	**82.63**	**0.72**	**17.37**
		250	0.5329	82.63	0.72	17.37
Sitting	1335	150	0.1358	94.31	0.24	5.69
		200	**0.0798**	**99.19**	**0.03**	**0.81**
		250	0.0907	94.56	0.1	2.44
Walking	1556	150	0.2008	93.66	0.26	6.34
		200	0.0867	97.89	0.09	2.11
		250	**0.0568**	**98.59**	**0.06**	**1.41**
Maths	1474	150	0.2001	93.33	0.28	6.67
		200	**0.1803**	**94.0**	**0.25**	**6**
		250	0.1254	94.0	0.25	6

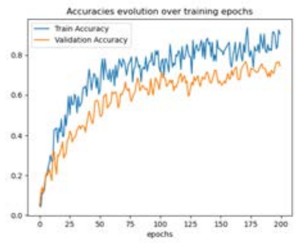
Table 5.1: Results obtained from the experiments done over the **GUDB** with **3bpf** and over the five different scenarios.

The evolution during training and validation for all these experiments over the GUDB with 3bpf can be seen in Figure 5.1. The best result for each of the experiments shown in Table 5.1 is shown. If the focus is set on Figure 5.1a it can be seen how from epoch 50 approximately there is an overfit of the network as the training curve is getting to a higher accuracy each time while the validation curve is not capable to keep on learning and get stuck. The same happens for the *jogging* scenario shown in Figure 5.1b but even earlier, almost at the beginning of the training. At epoch 200, it can be seen how the network suffers clear overfitting as the validation curve is below the training one. This figure indicates that the problem of classification of users with higher heart rates (as in those scenarios) is too complex and the network may learn it is “noise” or irrelevant information within the dataset. Hence, the model is unable to generalize well to new data as it is the data from validation.

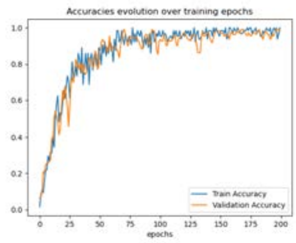
For the other three Figures (Figure 5.1c, 5.1d and 5.1e) where the scenarios in which the users are supposed to have lower heartbeat rates, the training of the model is not overfitted resulting in higher accuracies as the model is capable to generalize for non-seen data.



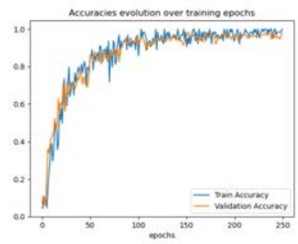
(a) Accuracy evolution during training and validation for the GUDB for *Hand bike* scenario with 3bpf and 200 epochs.



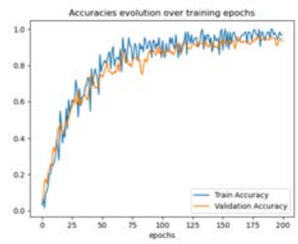
(b) Accuracy evolution during training and validation for the GUDB for *Jogging* scenario with 3bpf and 200 epochs.



(c) Accuracy evolution during training and validation for the GUDB for *Sitting* scenario with 3bpf and 200 epochs.



(d) Accuracy evolution during training and validation for the GUDB for *Walking* scenario with 3bpf and 250 epochs.



(e) Accuracy evolution during training and validation for the GUDB for *Maths* scenario with 3bpf and 200 epochs.

Figure 5.1: Accuracies for the best results obtained during training and validation with the GUDB over users performing different tasks for 3bpf.

Lastly, for this section, in Figure 5.2 it can be seen the CM for the best result from the *biking* scenario obtained after testing the model with the GUDB and 3bpf which was the one with 200 epochs as seen in Table 5.1. The same is obtained for the other four scenarios (Figures 5.3, 5.4, 5.5 and 5.6). By comparing the five of them, it can be observed how the two CM which are from scenarios which are supposed to need a higher cardiovascular effort, show a lower performance of the model when it is tested. For instance, by simply looking at the figures and the legend to its right, it can be seen how the parts which are supposed to be all black if good performance is obtained, clearer colours are shown. In fact, the matrix that shows more variety of colours is the one from the *jogging* scenario, which, as commented before, is the one presumably with higher heart rates. For the rest of the CM, normal results are obtained compared to the other CM presented in previous chapters.

Therefore, almost all scenarios conclude with the hypothesis that it is more difficult to identify a user with higher heart rates. Consequently, more data should be needed to identify users with higher accuracy and lower error rates.

In the table 5.2 where the experimentation for the GUDB and 5bpf are shown, it can be seen almost the same results as the ones in the Table 5.1. As shown here, the bpf is a hyperparameter that has to be cross-validated for every dataset as it seems to be an important parameter that should be chosen depending on each database studied and on how large the EKG records are.

Scenario	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
Hand-bike	914	150	0.4945	83.33	0.69	16.67
		200	**0.359**	**90.74**	**0.39**	**9.26**
		250	0.385	87.96	0.5	12.04
Jogging	1317	150	0.7453	79.10	0.91	20.9
		200	0.6248	83.58	0.71	16.42
		250	**0.6068**	**85.82**	**0.62**	**14.18**
Sitting	792	150	0.2813	95.52	0.19	4.48
		200	0.1683	97.01	0.13	2.99
		250	**0.1221**	**98.51**	**0.06**	**1.49**
Walking	912	150	0.3204	93.10	0.29	6.9
		200	**0.2077**	**95.4**	**0.19**	**4.6**
		250	0.2821	90.80	0.38	9.2
Maths	881	150	0.3690	90.32	0.4	9.68
		200	0.1891	94.62	0.22	5.38
		250	**0.1316**	**97.85**	**0.09**	**2.15**

Table 5.2: Results obtained from the experiments done over the **GUDB** with **5bpf** and over the five different scenarios.

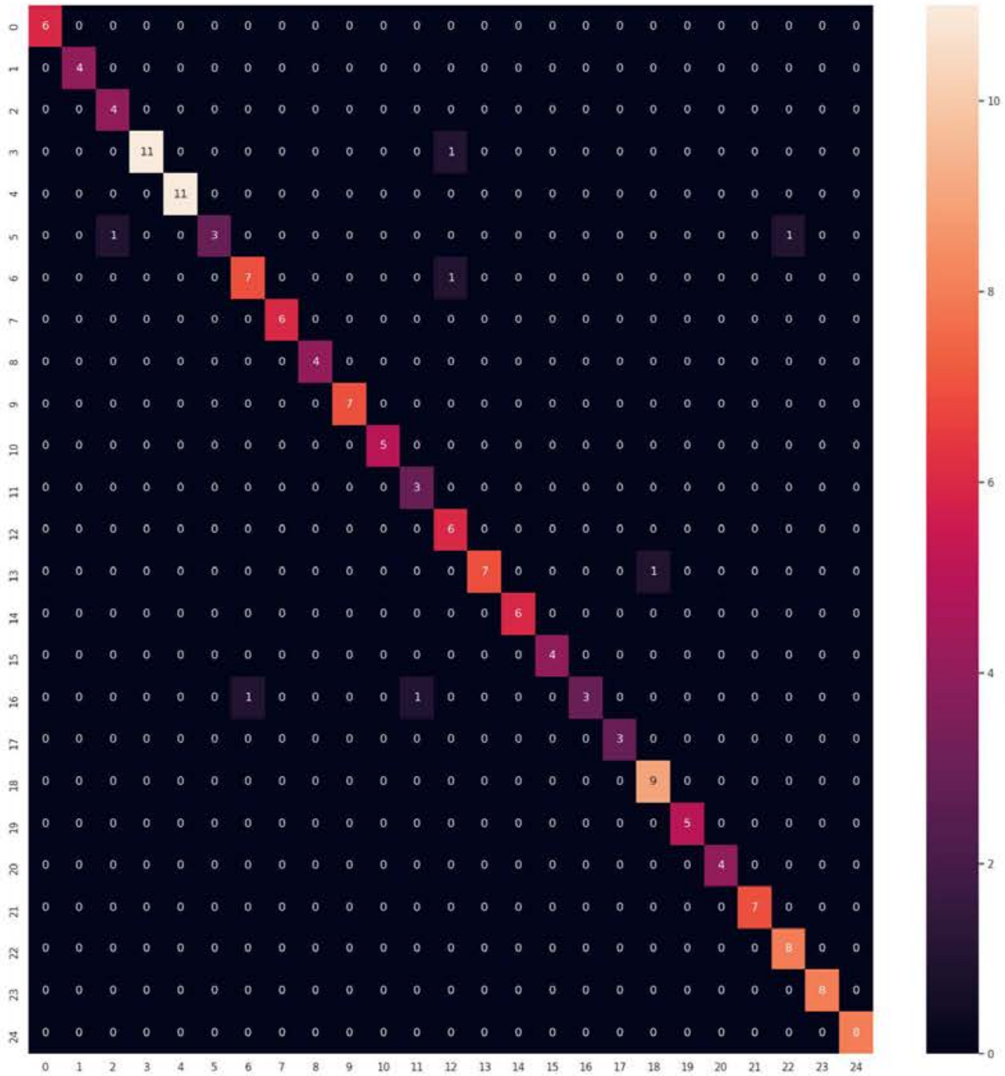


Figure 5.2: Confusion Matrix obtained when testing the *biking* scenario from the GUDB for 3bpf and 200 epochs.

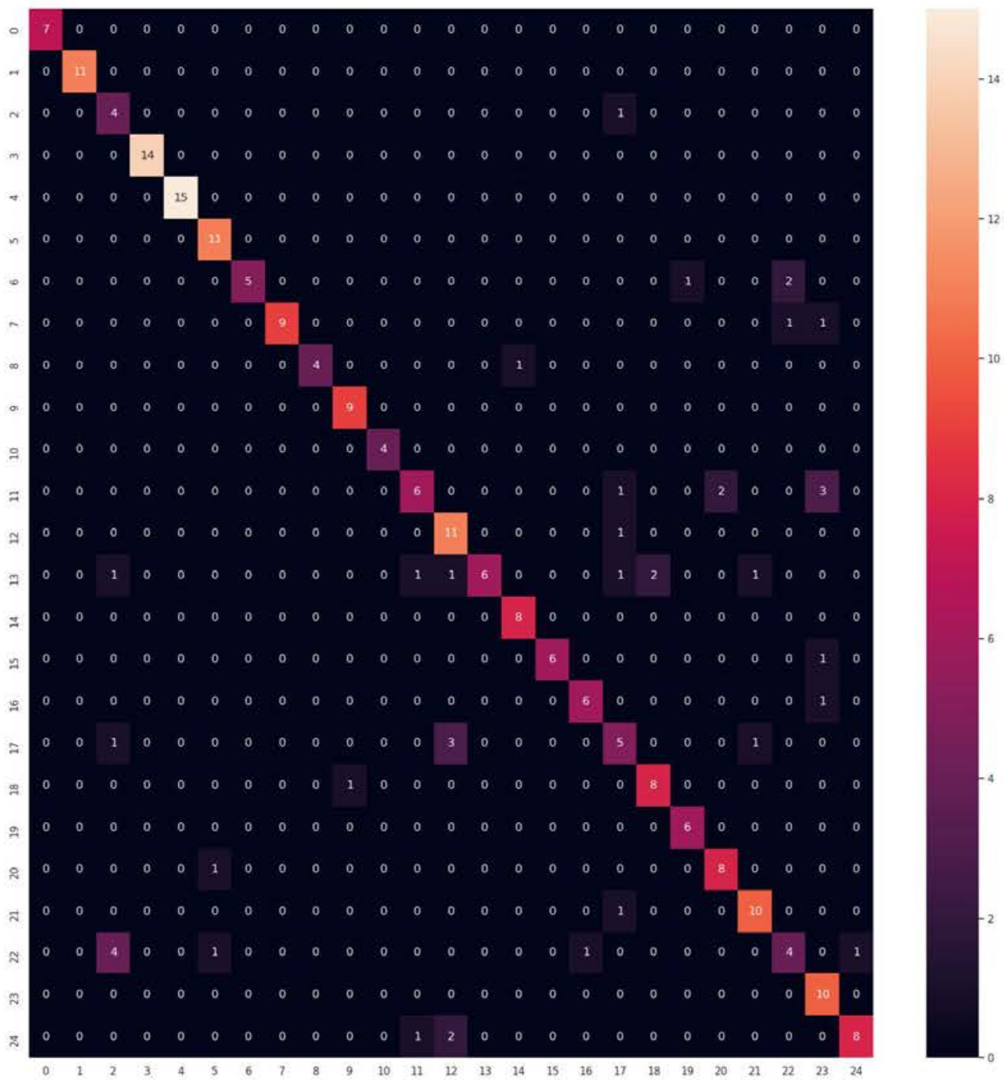


Figure 5.3: Confusion Matrix obtained when testing the *jogging* scenario from the GUDB for 3bpf and 200 epochs.

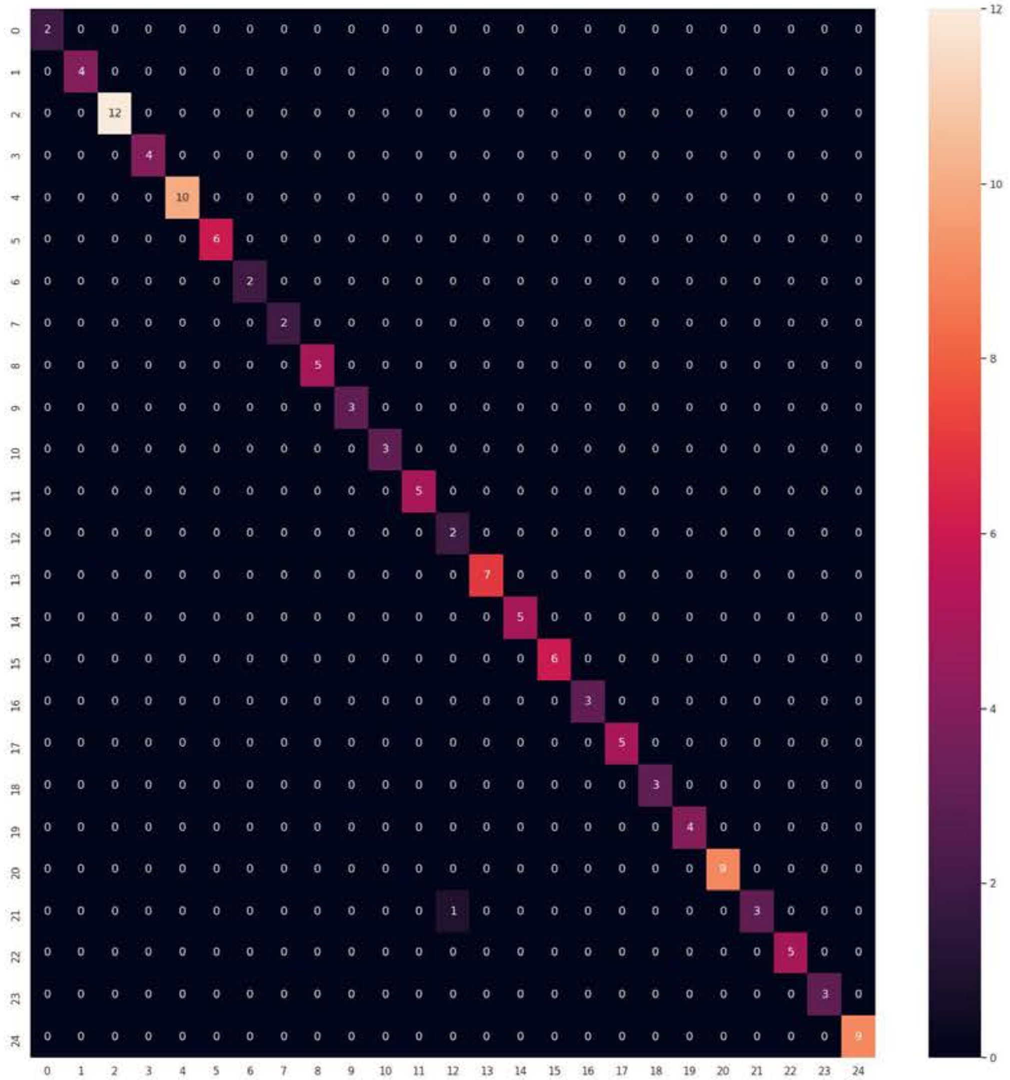


Figure 5.4: Confusion Matrix obtained when testing the *sitting* scenario from the GUDB for 3bpf and 200 epochs.

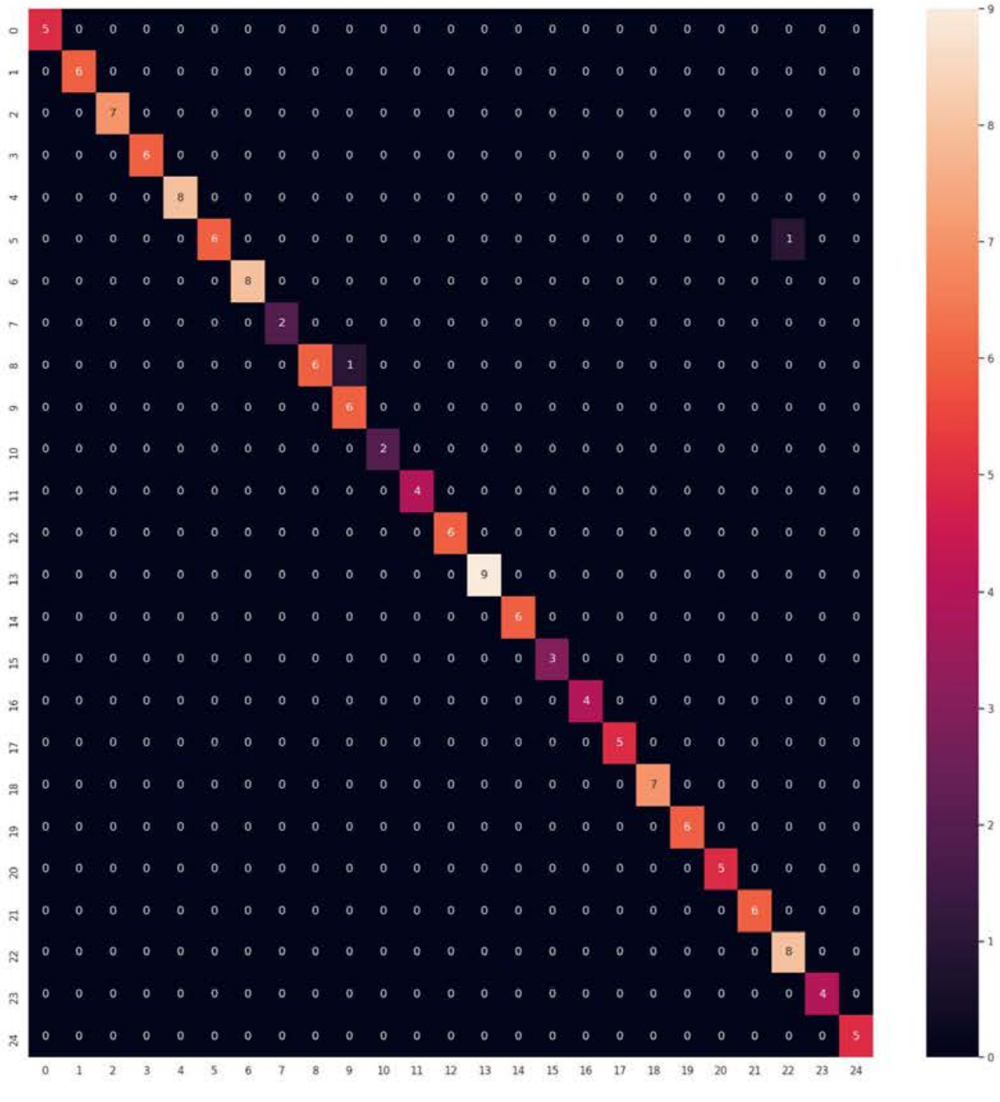


Figure 5.5: Confusion Matrix obtained when testing the *walking* scenario from the GUDB for 3bpf and 250 epochs.

For this experiment with 5bpf results are expected to be a bit lower due to the significant decrease in the number of images that the model holds to properly train and test the network (see Table 5.2). As it was seen in the previous experiment and also can be seen in Figure 5.7b, the training and validation performance for the *jogging* scenario from the GUDB with 5bpf suffers an overfit of the network as happened before. In this figure, it can be seen that, from the very beginning, the network is unable to generalize with the data that it holds and, consequently, it suffers an overfit indicated by the low performance shown in the validation curve. The overfitting might be caused, as commented, due to the higher heart rates when *jogging* as the network might learn the "noise" or other features instead of the important ones that make a user unique.

The lack of data for this experiment also explain why all the training and validation curves on the rest of figures for this experiment (Figures 5.7a, 5.7d, 5.7e and 5.7c) are so spiky. As seen in the experimentation, the data hold is not enough to offer the same results as the ones in Chapter 3 where control users are presented and thousands of users are held to train, validate and test the model. Even though, results obtained with the GUDB and 5bpf for the different activities, are above an 85.82% of accuracy and even a 97.85% of accuracy is achieved with the *Maths scenario* indicating that ELEKTRA can cope with these different scenarios.

CM for this experiment can be seen on Figures 5.8, 5.9, 5.12, 5.10 and 5.11. As happened in the previous experiment, this CM follows more or less the same lines. The difference is, again, the number of images. As there are fewer images held for these experiments than in the previous ones, results have decreased and, consequently, the CM are affected. For example, the one for the *jogging* scenario in Figure 5.9 has even more variation of colours on the "black side " of the matrix indicating that some users have been misclassified. For instance, user called *21* has been misclassified into user *6*, as it is indicated in the CM. The rest of the CM, are more or less similar to the ones before.

The experiments for the GUDB and 7bpf have not been included due to the small number of images available. Less than 700 or 800 images to identify 25 users with the CNN is not enough to obtain conclusive results based on the previous experiments.

5.2 Combination of users performing different activities

Following the line of the tests for the GUDB, this database has also been tested where the whole dataset with all the scenarios is merged. The results of these experiments can be seen in Table 5.3. The highest result (91.32% of accuracy) was achieved in the experiments with 3bpf, 250 epochs and 8099 images. If it is compared with the results obtained in the experiment where the scenarios are detached in Table 5.1 it can be observed that this new result is better than the ones obtained in the scenarios where a cardiac activity occurred (such as *jogging* with an 85.82%). However, it is worse than resting scenarios (such as *sitting* with a 98.51%). It means that this approach offers almost an average of the past results. In fact, when looking to Figures 5.13a, 5.13b and 5.13c it can be seen the loss evolution for training and testing for each of the experiments shown in Table 5.3 for 3bpf. As it is noted, there is overfitting of the network as the validation curve is positioned above the training curve in all of the experiments. This may be because the data held in the experiment is not enough to train or that, as the EKMs are so different in each of the merged scenarios, the model is not capable to generalize for one set of features that makes each user unique.

Then, concerning the results with 5bpf, the approach offers lower accuracies and error rates. It is important to remark that there are nearly half of the number of images compared to the experiment with 3bpf. Seeing Figures 5.14a, 5.14b and 5.14c the training and validation evolution for the loss function can be observed. Compared to the experiment with 3bpf, the network with 5bpf does not suffer as much overfit as the model with 3bpf. Even though lower results are obtained and it can be observed how the network has problems conveying as both curves are very spiky indicating difficulties to generalize and learn the proper features of the network.

This result shows that more data is needed to train the network correctly. For this reason, more data will be needed to identify a user when working with scenarios where cardiovascular activity occurs.

bpf	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
3	8099	150	0.3444	89.06	0.45	11.62
		200	0.3441	88.81	0.47	11.49
		250	**0.2963**	**91.32**	**0.36**	**9.06**
5	4816	150	0.7209	79.19	0.87	20.81
		200	0.6205	81.89	0.75	18.11
		250	**0.5708**	**82.47**	**0.73**	**17.53**

Table 5.3: Results obtained from the experiments done over the GUDB with 3 and 5bpf joining all scenarios by doing an All vs. All.

5.3 HCA or LCA

Following the lines of the experiments done in this chapter over the GUDB, as it has been already noticed, for scenarios where a cardiovascular effort is required, the identification of users becomes more complex. To study this hypothesis, four of five scenarios have been merged into two categories: LCA and HCA.

The LCA scenario is where the user is supposed to perform an activity that does not demand great effort and a low heart rate is supposed. It has been named as LCA to have a reference. Hence, the LCA scenario comprises the *walking* and *sitting* activities presented in Section 2.1.

The other scenario comprises the activities from the GUDB which are considered to need a higher effort and, consequently, have a higher heart rate, called HCA. These are *jogging* and *biking* scenarios.

The missing scenario is where users perform *a maths exam*. This scenario has not been included because a user can have a lower or higher heart rate depending on whether the person is nervous or not. Hence, this scenario is not included in this experiment to have a clear separation between heart rates. All results from this experiment can be found on Table 5.4[1].

The 90% of the heatmaps from all the images obtained are used to train the network while the 10% resting is used to test the network. The network has been trained and tested over 25 users with a categorical classification but feeding two scenarios simultaneously in each test.

So, as it can be seen on Table 5.4, despite having less images for the LCA scenario than for the HCA scenario, better results are obtained for LCA one.

In Figure 5.15a, it can be seen the training and validation evolution for the LCA scenario with 3bpf and the best result obtained for it, which is 150 epochs. And, in Figure 5.15b the same plot can be seen but for the HCA scenario and 250 epochs, which was the best result. Both plots show spiky curves indicating some difficulties in learning features to generalise in a proper model. As commented before, despite having fewer images for the HCA, the model needs more epochs to achieve proper results with this scenario.

[1]The bpf hyperparameter chosen for all these experiments is 3bpf. It has been chosen to represent this experiment with 3bpf because more images can be obtained and used to test our network.

Scenario	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
LCA	2891	100	0.1797	96.6	0.14	3.08
		150	**0.1245**	**97.74**	**0.09**	**1.93**
		200	0.1091	96.98	0.13	2.61
		250	0.1118	95.47	0.19	4.86
HCA	3734	100	0.7642	75.77	1.01	23.83
		150	0.5784	83.16	0.7	17.4
		200	0.518	84.18	0.66	15.99
		250	**0.4595**	**85.71**	**0.6**	**14.17**

Table 5.4: Results obtained when testing over the GUDB distinguishing between two scenarios: LCA or HCA.

This trait fully confirms the hypothesis that higher results are obtained when a user rests than when performing a cardiovascular activity. In fact, the best result for the LCA scenario is a 97.74% of accuracy; meanwhile, cardiovascular activity obtains 12 points less with an 85.71% of accuracy. There is an essential difference between the results of both scenarios.

To keep on studying how different heartbeat rates are obtained as a result of performing different activities, training and testing scenarios have been mixed to follow investigations over the GUDB. It means that, by seeing Table 5.5 for the first experiment on this table, the network has been trained with the LCA images and tested with the images from the HCA scenario. Furthermore, for the second experiment, the opposite has been done: training with the HCA scenario and testing with the LCA one.

Thus, focusing on Table 5.5, once again, what has been hypothesized can be confirmed: it is more complex to identify users with higher heart rates. In the first row, where the tested users are the ones performing cardiovascular activities, there are low accuracies, and in the second row, where the users tested are the resting ones, improved results are obtained.

By setting the focus on Figures 5.16 and 5.17 it can be seen how these CM are not very similar to the ones from the previous experiments. The main goal of this experiment is to show that it is not possible to train a model when the user is in one *state* and perform the testing on a different *state* with a different heartbeat rhythm with this amount of images, the CM were expected to show it by showing less accurate results.

Results from this experiment were not expected to be very high compared to the others. These results were expected to show the difference between training and testing with different scenarios so they can be compared. These two databases used for this experiment are more balanced, even though the resting one the one with fewer images, is the one obtaining higher performances.

Scenario	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
Train	3018	100	4.2492	32.91	2.81	68.99
LCA		150	4.8969	33.16	2.8	69.17
Test		200	**5.187**	**37.24**	**2.62**	**65.42**
HCA		250	6.263	35.71	2.69	66.84
Train	3607	100	2.0277	41.16	2.47	57.91
HCA		150	1.8747	45.28	2.28	54.05
Test		200	**2.1209**	**46.42**	**2.25**	**52.83**
LCA		250	2.5796	43.77	2.35	55.85

Table 5.5: Results obtained when testing over the GUDB training and testing with different scenarios. Testing with the LCA images and training with HCA ones on the first row and the other way round on the second row.

5.4 Fight of scenarios

The last experiment launched of this project to study the GUDB and the differences between heartbeat rates over user identification is the one that can be seen in Table 5.6. In this table, different scenarios have been mixed in their training and testing phases. Like the experiment commented above in Table 5.5, an already trained model has been used for a scenario or activity and then, another scenario has been tested over this trained model.

The first two rows of the Table 5.6 refers to two scenarios: *walking* and *biking*. On the first row, there appear the results shown for training the model with the *walking* scenario and testing with *biking* and, on the second row, there is the opposite case. What is interesting about these two scenarios is that both of them achieve the same higher accuracy, which is 50%. This might be because, depending on the user's speed when *walking* and *biking*, the effort of each one, might be similar, resulting in similar results.

Then, on the following rows of the same table, the *sitting* scenario is the one used for all the experiments (the one with better results when trained and tested by itself) and the testing phase has been performed with all other scenarios to study the difference among them.

The third row is the one in which the lower results were obtained. This was expected because the *sitting* scenario was used for training and the testing scenario was *jogging*. Hence, in this particular experiment, the model was trained with the scenario with the lower heartbeat rate, and the users used to predict were having higher heartbeat rates. Consequently, the lowest result in Table 5.6 is for this experiment. If the error rates are studied, it is also obtained the lowest FRR in all the experiments which is 89.05%. This means that every nine of each ten times a user attempts to access the secure resource after our biometric system will be rejected. On the contrary, the FAR is high but not as high as the FRR because it is 3.67% (the highest). This means that from every 100 attempts of impersonating the proposed biometric system, only almost 4 of them would succeed. This result tells that the proposed approach will not accept an impostor in most cases, even with low performances over users with higher heartbeat rates.

The left experiments on Table 5.6 confirm that with higher heartbeat rates (such as *biking*) the identification has worst results than with lower heartbeat rates (such as *walking* or *performing a maths exam*) as demonstrated during all our experiments.

Scenario	EKMs	Epochs	Loss	Accuracy (%)	FAR (%)	FRR (%)
Train		100	2.4924	46.15	2.24	58.29
Walk	1570	150	3.2895	48.08	2.16	51.56
Test		200	**3.4784**	**50**	**2.08**	**52.23**
Bike		250	3.2747	48.72	2.14	53.13
Train		100	2.4393	42.96	2.38	53.49
Bike	1515	150	2.261	49.3	2.12	46.73
Test		200	3.1415	47.89	2.18	47.38
Walk		250	**2.9027**	**50**	**2.09**	**45.37**
Train		100	6.2369	13.14	3.65	88.08
Sitting	1448	150	**7.7108**	**16.53**	**3.51**	**86.4**
Test		200	8.169	16.1	3.52	86.42
Jogging		250	7.7617	12.71	3.67	89.05
Train		100	3.5188	43.33	2.37	58.02
Sitting	1362	150	4.3421	47.33	2.19	54.86
Test		200	4.413	47.33	2.19	54.65
Maths		250	**4.2452**	**50**	**2.08**	**50.73**
Train		100	2.9046	35.21	2.7	66.97
Sitting	1556	150	3.581	33.8	2.76	68.36
Test		200	3.441	36.62	2.64	66.78
Walking		250	**3.7029**	**41.55**	**2.44**	**61.61**
Train		100	3.9325	26.28	3.08	77.25
Sitting	1368	150	4.4912	30.13	2.91	74.15
Test		200	**4.5356**	**33.97**	**2.75**	**72.02**
Bike		250	5.1502	28.21	2.99	76.52

Table 5.6: Results obtained when testing over the GUDB training and testing with different scenarios.

As there is no training in these experiments, the CM of these results are offered. The perfect CM would be the one that would have a perfect diagonal line coloured and the rest of the CM would be all black. Hence, by checking the CM for these experiments, very different results can be seen.

As hypothesized the CM closer to the perfect one are the ones where the scenarios used to train and test which are supposed to be closer in heartbeat rates such as in Figures 5.18 and 5.19 which are for training with the *walking* scenario and predicting with the *biking* scenario and vice versa.

Then, it can be seen that the worst CM is obtained when training with the scenario where the lowest heartbeat rate is supposed and testing with the higher heartbeat rate, which is training with *sitting* and testing with *jogging* seen in Figure 5.23. Instead of having the diagonal highlighted, there are two columns indicating a clear misclassification of users. In fact, most of the users have been misclassified into the same users which are, namely, users 3, 4, 13 and 15.

5.5 Discussion

An in-depth study of the GUDB over user classification with EKM has been performed in this chapter. First, a first approach seen in Section 5.1 has demonstrated how the ELEKTRA methodology is feasible to identify users from the GUDB to a lesser or greater extent depending on the scenario or activity that the user was performing.

All scenarios have been merged in Section 5.2 to show how ELEKTRA can classify a pool of users where the users were performing different activities to show the need for more data in this situation. This means that when in an identification system the main goal is to identify users performing any activity or having different heartbeat rates, more data is needed to achieve better performance as the network does not have enough data to generalize in a proper model. Then, the differences between scenarios depending on the heartbeat rhythm where different activities were merged are shown in Section 5.3. In this section, the first demonstration of how for higher heart rates it is more complex to identify those users takes place.

Finally, some of the scenarios had been put against others in Section 5.4 to follow the lead investigation over differences in heartbeat rhythms resulting in having lower accuracies for scenarios where more differences in heartbeat rates were shown.

In conclusion, thanks to this extensive study over the GUDB where the users are performing five different activities with different heartbeat rhythms, support the put forward theory in this work about the differences in heartbeat rates. With this work, it is supported that it is more complex to identify users and classify them for higher heart rates. However, this does not apply when a user is being impersonated. It can be seen in the error rates obtained in the experiments, as there are high values for the FRR and lower values for the

FAR.

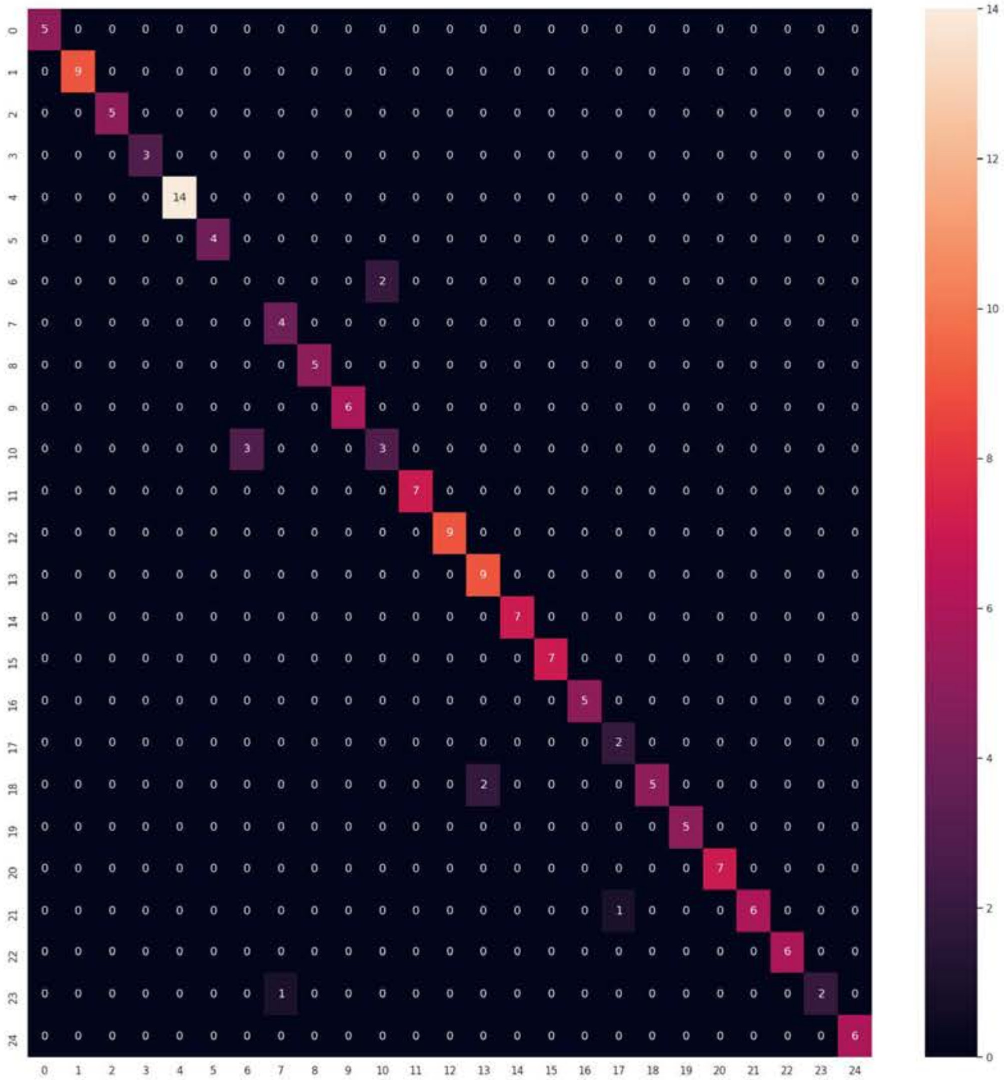
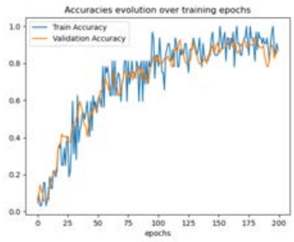
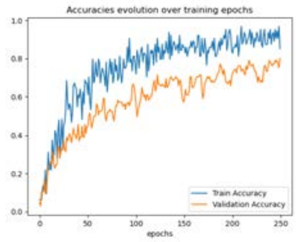


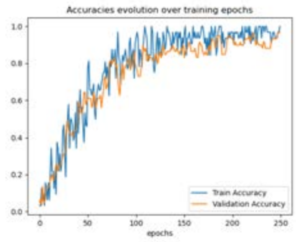
Figure 5.6: Confusion Matrix obtained when testing the *maths* scenario from the GUDB for 3bpf and 200 epochs.



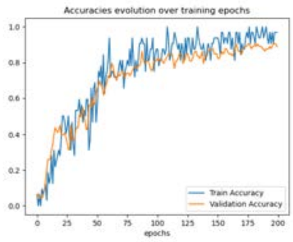
(a) Accuracy evolution during training and validation for the GUDB for *Hand bike* scenario with 5bpf and 200 epochs.



(b) Accuracy evolution during training and validation for the GUDB for *Jogging* scenario with 5bpf and 250 epochs.



(c) Accuracy evolution during training and validation for the GUDB for *Sitting* scenario with 5bpf and 250 epochs.



(d) Accuracy evolution during training and validation for the GUDB for *Walking* scenario with 5bpf and 200 epochs.

(e) Accuracy evolution during training and validation for the GUDB for *Maths* scenario with 3bpf and 250 epochs.

Figure 5.7: Accuracies for the best results obtained during training and validation with the GUDB over users performing different tasks for 5bpf.

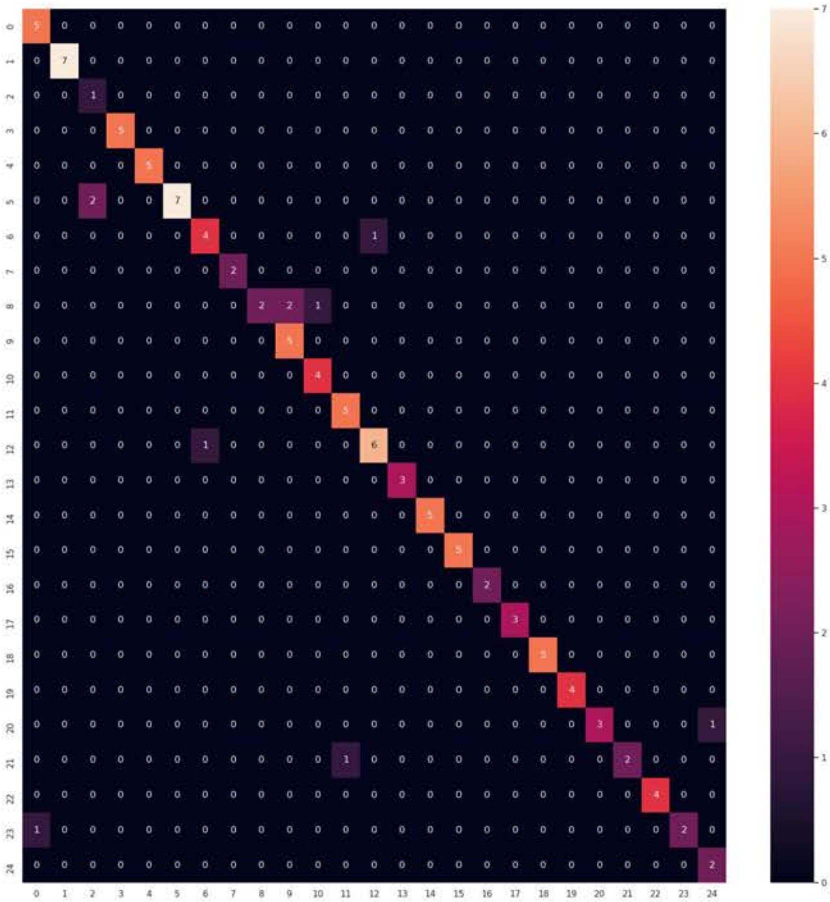


Figure 5.8: Confusion Matrix obtained when testing the *biking* scenario from the GUDB for 5bpf and 200 epochs.

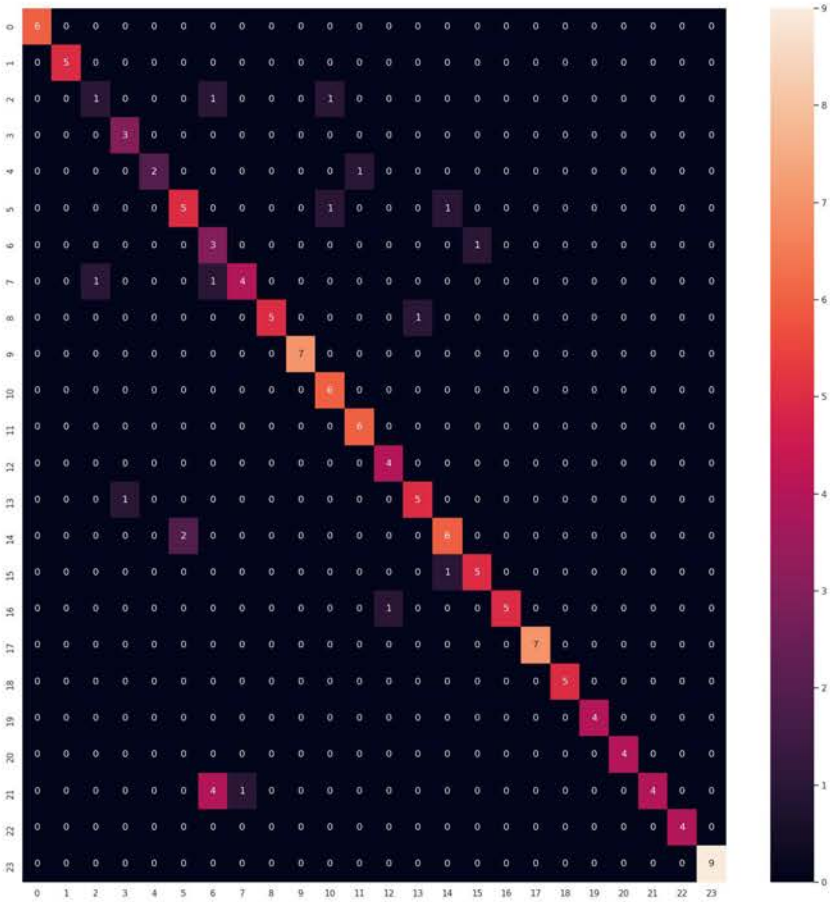


Figure 5.9: Confusion Matrix obtained when testing the *jogging* scenario from the GUDB for 5bpf and 250 epochs.

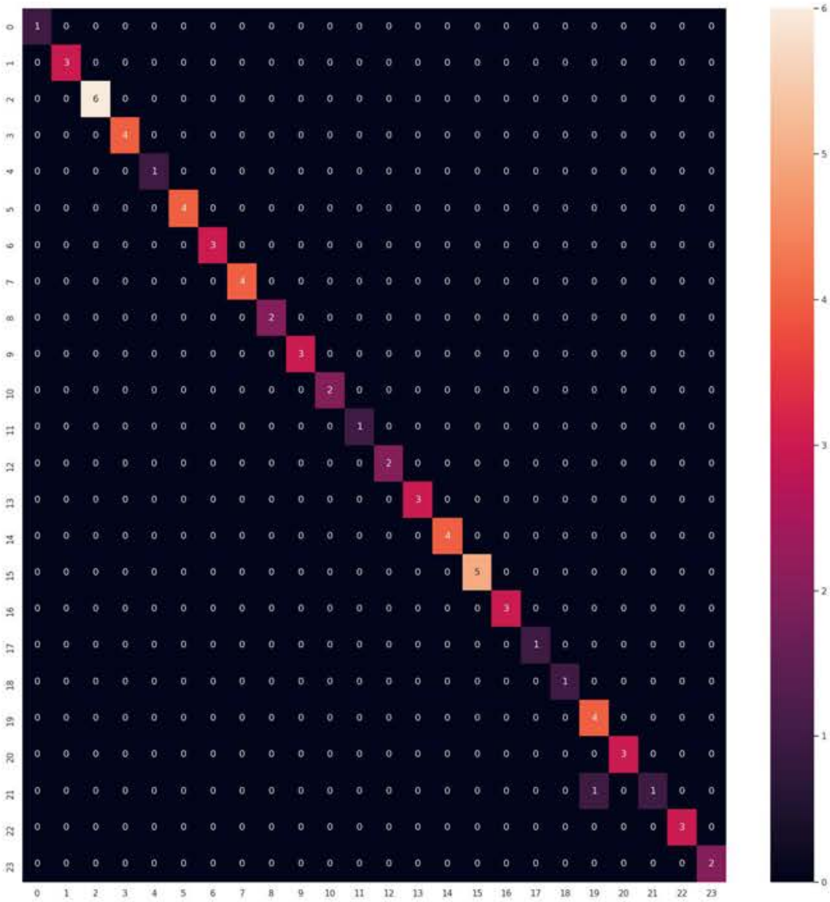


Figure 5.10: Confusion Matrix obtained when testing the *sitting* scenario from the GUDB for 5bpf and 250 epochs.

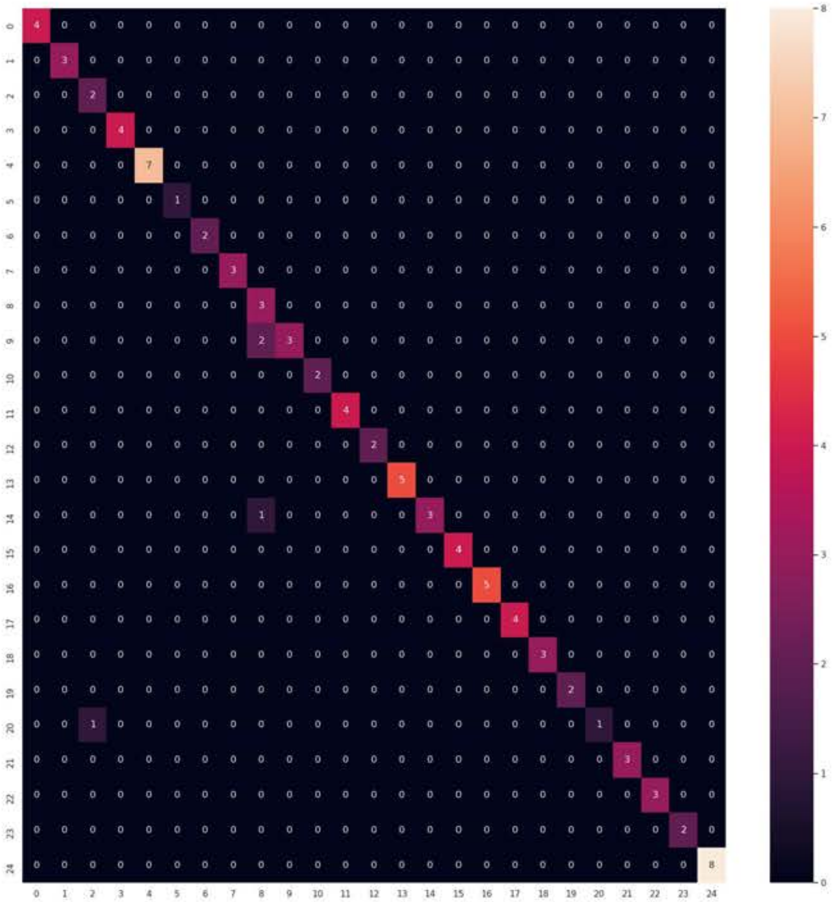


Figure 5.11: Confusion Matrix obtained when testing the *walking* scenario from the GUDB for 5bpf and 200 epochs.

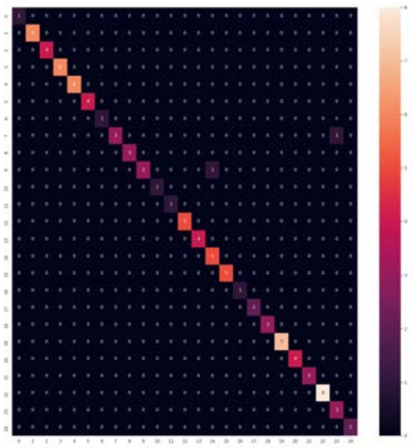
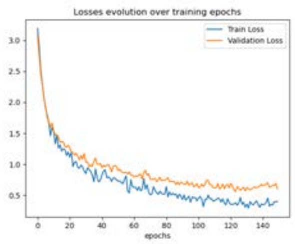
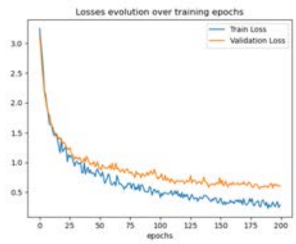


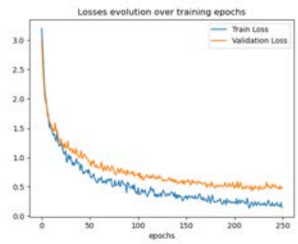
Figure 5.12: Confusion Matrix obtained when testing the *maths* scenario from the GUDB for 5bpf and 250 epochs.



(a) Loss evolution during training and validation for the GUDB where all the scenarios are merged with 3bpf and 150 epochs.

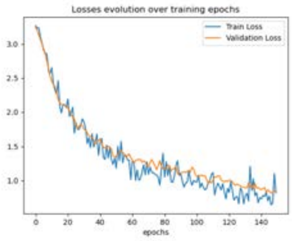


(b) Loss evolution during training and validation for the GUDB where all the scenarios are merged with 3bpf and 200 epochs.



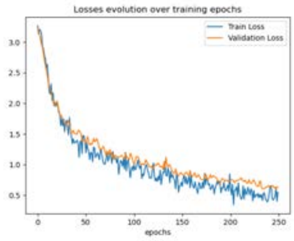
(c) Loss evolution during training and validation for the GUDB where all the scenarios are merged with 3bpf and 250 epochs.

Figure 5.13: Loss evolution during training and validation from Section 5.2 with the GUDB for 3bpf.



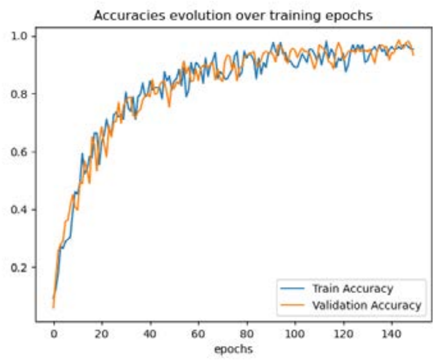
(a) Loss evolution during training and validation for the GUDB where all the scenarios are merged with 5bpf and 150 epochs.

(b) Loss evolution during training and validation for the GUDB where all the scenarios are merged with 5bpf and 200 epochs.

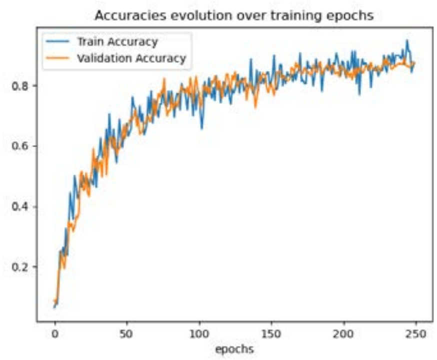


(c) Loss evolution during training and validation for the GUDB where all the scenarios are merged with 5bpf and 250 epochs.

Figure 5.14: Loss evolution during training and validation from Section 5.2 with the GUDB for 5bpf.



(a) Accuracy evolution during training and validation for the GUDB for LCA scenario with 3bpf and 150 epochs.



(b) Accuracy evolution during training and validation for the GUDB for HCA scenario with 3bpf and 250 epochs.

Figure 5.15: Accuracy evolution for training and testing from Section 5.3 with the GUDB and 3bpf.

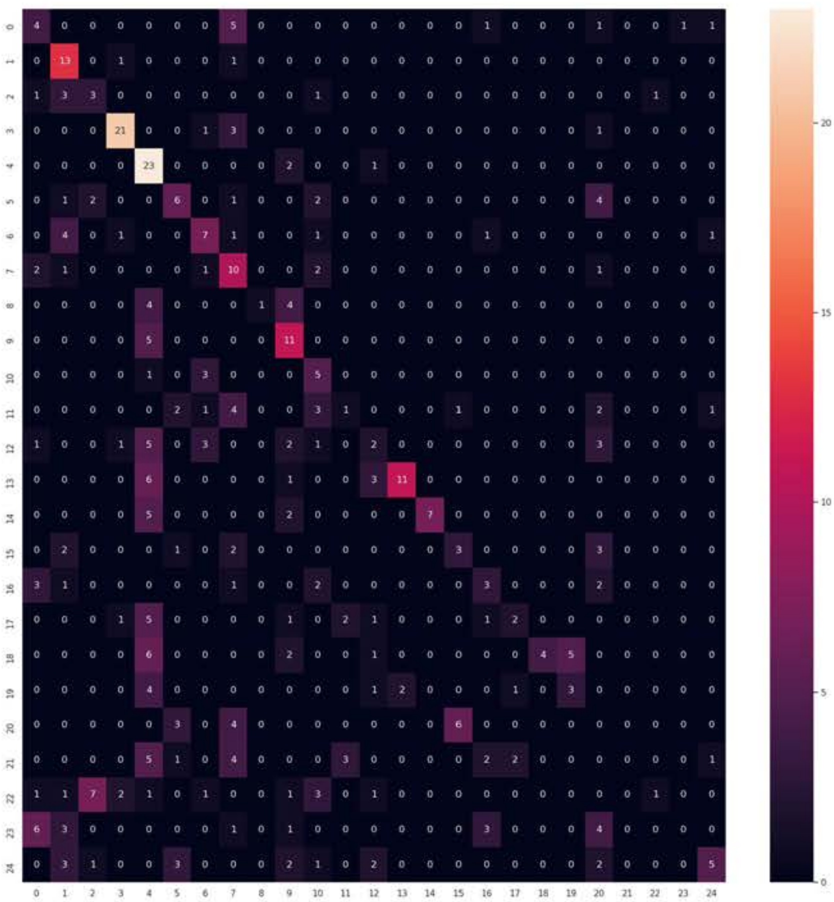


Figure 5.16: Confusion Matrix obtained when training the model with the LCA scenario and predicting with the HCA scenario from the GUDB with 3bpf and 200 epochs.

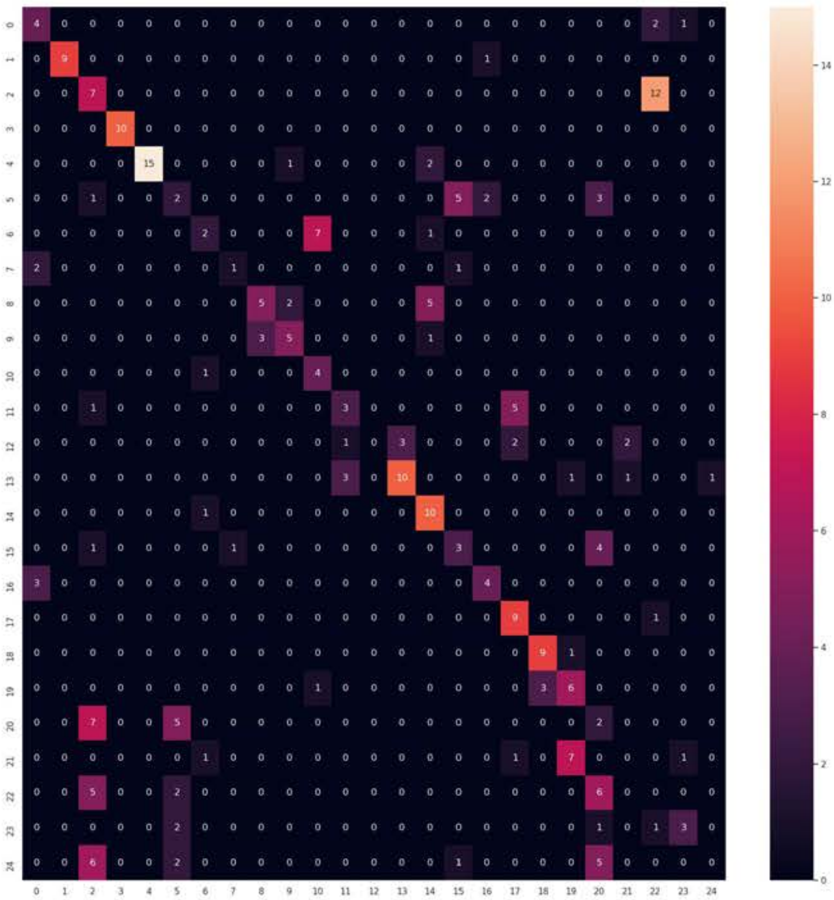


Figure 5.17: Confusion Matrix obtained when training the model with the HCA scenario and predicting with the LCA scenario from the GUDB with 3bpf and 200 epochs.

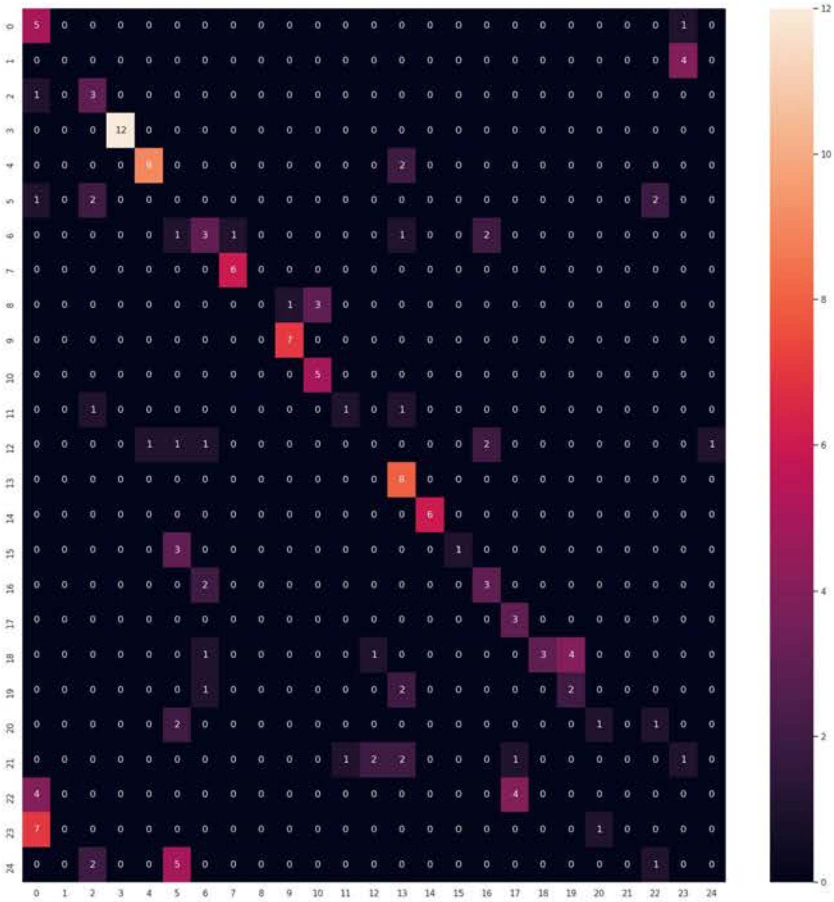


Figure 5.18: Confusion Matrix obtained when training the model with the *walking* scenario and predicting with the *biking* scenario from the GUDB with 3bpf and 200 epochs.

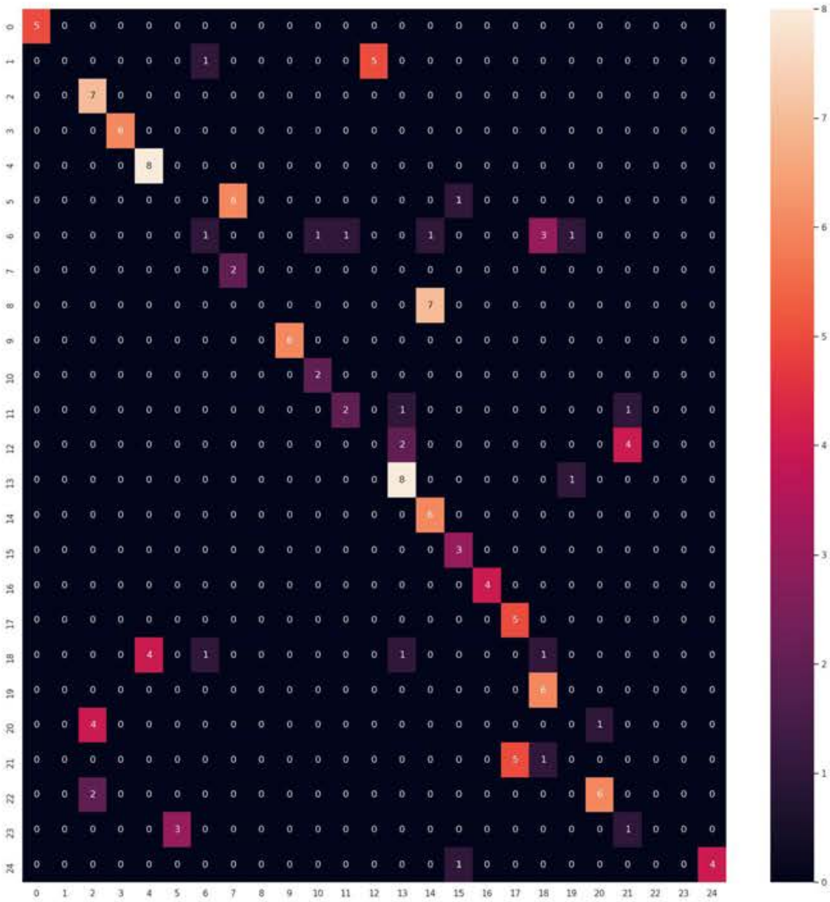


Figure 5.19: Confusion Matrix obtained when training the model with the *biking* scenario and predicting with the *walking* scenario from the GUDB with 3bpf and 250 epochs.

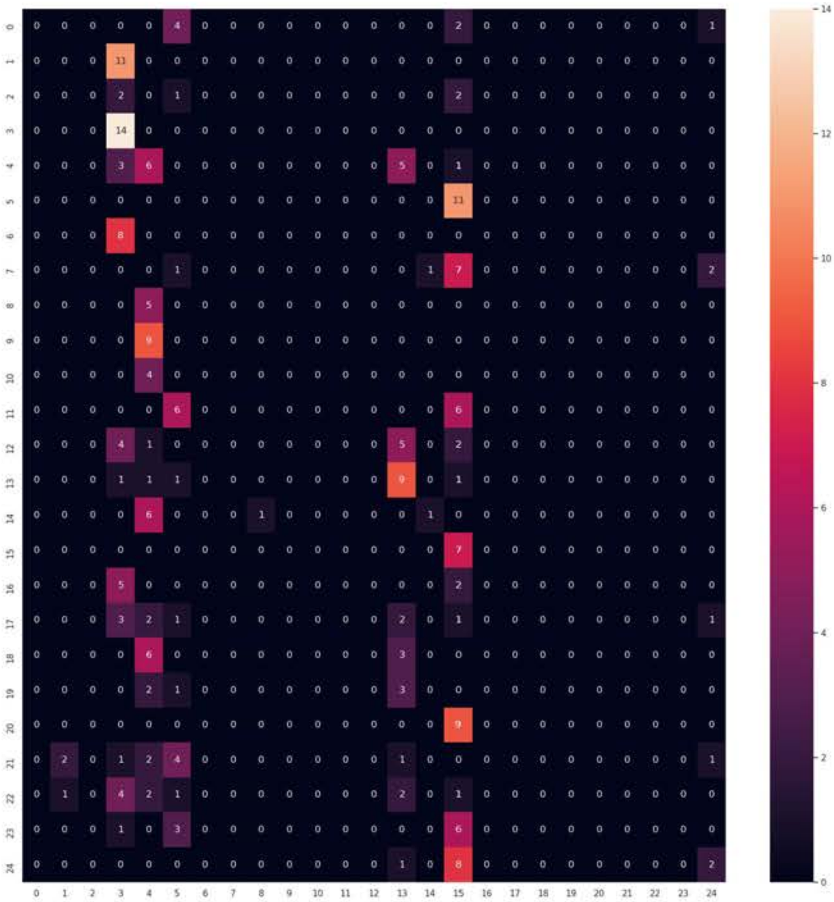


Figure 5.20: Confusion Matrix obtained when training the model with the *sitting* scenario and predicting with the *jogging* scenario from the GUDB with 3bpf and 150 epochs.

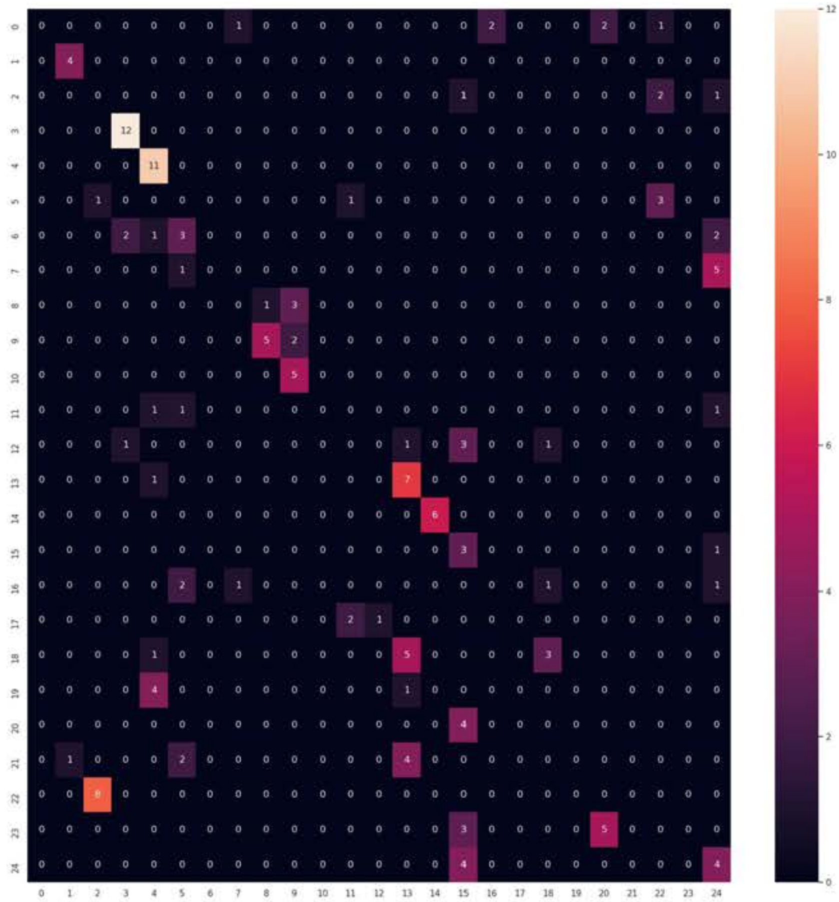


Figure 5.21: Confusion Matrix obtained when training the model with the *sitting* scenario and predicting with the *biking* scenario from the GUDB with 3bpf and 200 epochs.

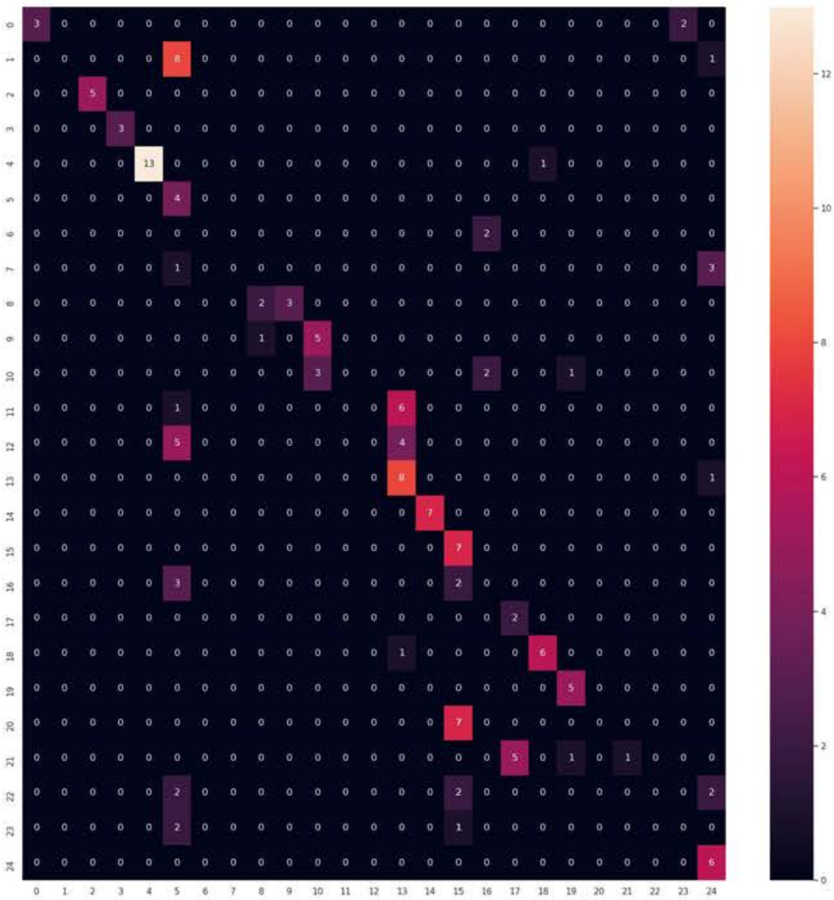


Figure 5.22: Confusion Matrix obtained when training the model with the *sitting* scenario and predicting with the *maths* scenario from the GUDB with 3bpf and 250 epochs.

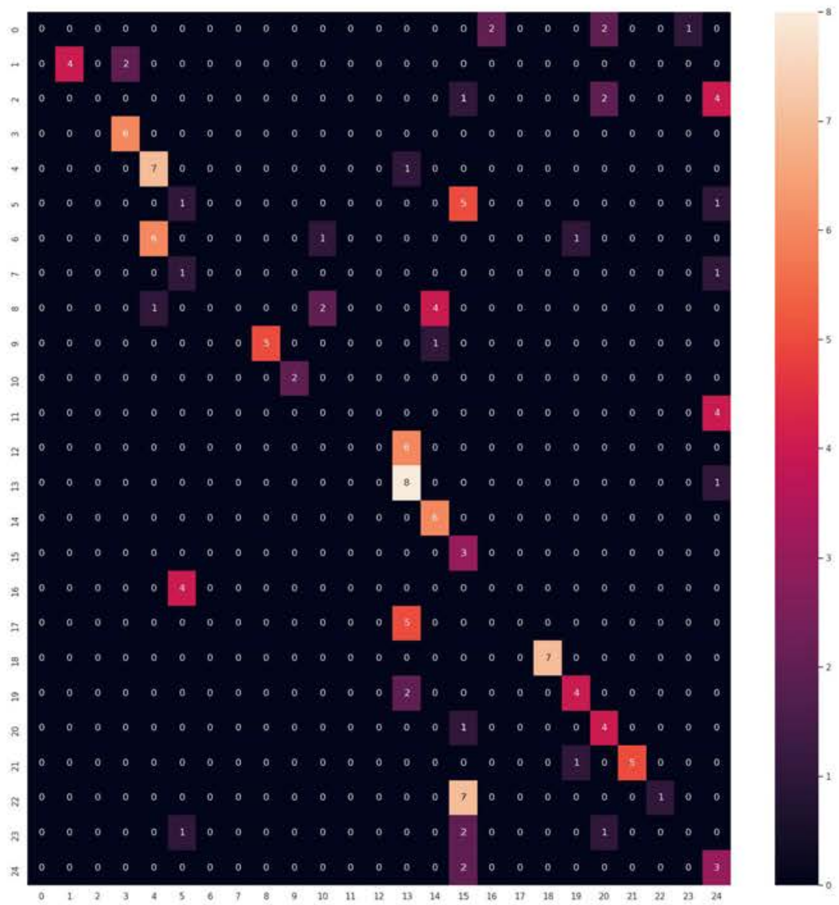


Figure 5.23: Confusion Matrix obtained when training the model with the *sitting* scenario and predicting with the *walking* scenario from the GUDB with 3bpf and 250 epochs.

CHAPTER 6

RELATED WORK

In this chapter a comparison between ELEKTRA's performance and the state-of-the-art models for human identification through their EKG is offered. The works presented in Table 6.1 have been chosen due to similarities with the presented model in this thesis or due to studying the same databases as the ones in the previous chapters. This way, an accurate comparison between similar methods can be performed and an evaluation of ELEKTRA is presented.

6.1 Comparative Analysis

Jikui Liu et al., 2018 [99]. In the first work exposed on Table 6.1 a non-fiducial feature extraction based on AutoRegressive Model (ARM) is offered to perform user identification through EKG. To test their efficiency the authors perform two different kinds of tests over the same database, the PTBDB. The authors grab 50 healthy users from this database and another 50 users but non-healthy ones with a CVD achieving a 98% and a 100% of accuracy, respectively. Their work explained that these accuracy metrics are obtained using one data vector per user. This means that, of 50 healthy users, only one was misclassified. The same happens for the non-healthy users: they use one data vector per user, and all were correctly classified; this is how they obtain a 98% and 100% of accuracy. However, ELEKTRA proposes a better generalization method by testing over hundreds of images for each user while simultaneously achieving similar results with the PTBDB as it is a 98.89% of accuracy. We also offer lower error rates, as shown with an EER of 1.075%. Therefore, ELEKTRA outperforms [99].

Joao Ribeiro Pinto et al., 2020 [113]. The next work seen on Table 6.1 is the one presented in [113] also testing their proposal over the PTBDB. The authors use a CNN to train and test not over segmentation of data points or data vectors but users, meaning that they train the model with the UofTDB, fine-tune it with the PTBDB and test it over the PTBDB. Their proposal offers a good point of view but, unfortunately, worse results than ELEKTRA are offered as the EER obtained with the PTBDB is a 9.06% while with ELEKTRA a 1.075 of EER is achieved.

Ruggero Donida Labati et al., 2019, *Deep-ECG* [93]. Following the line of testing identification models through EKG over the PTBDB the work presented in [93] is offered. In this work, the authors use 6 layers CNN, three max-pooling layers, three Local Response Normalisation layers, one dropout layer and a fully connected plus a softmax layer as the feature extractor. ELEKTRA is able to perform feature extraction and classification with a one-layer CNN achieving similar results. It has to be taken into account that the user classification with the work in [93] achieves a 1.63% of EER over 52 healthy users, not the whole database. Therefore, with a simpler CNN and more users, ELEKTRA outperforms [93].

Yifan Chu et al., 2019 [94]. In the work presented in [94] exposed in Table 6.1 the authors also used two separate methods to perform feature extraction and classification instead of one, as is proposed with ELEKTRA. A CNN is used as the identification method but the Euclidean Distance is needed to calculate the difference between vectors of features representing the users. As explained in Section 1.2.1, there exist a process for user identification with EKG that may vary between proposals. In this case, four steps are needed: data acquisition; signal preprocessing performing a concatenation of heartbeats; feature extraction with a CNN and; classification with Euclidean Distance. What is remarkable about the project presented in this thesis is that similar results are achieved in many of the proposals with one step less as the feature extraction and classification is performed in just one step. Therefore, ELEKTRA is more efficient than [94] in terms of processing steps. In addition, with this work commented, when testing with the MIT-BIHDB the authors obtain a 4.74% of EER while with ELEKTRA a 1.075% is achieved (see Table 6.2).

Pei-Lun Hong et al., 2019 [129]. In the case of [129], the user identification is also performed through a CNN but, in this case, the NN works as the feature extractor and the classifier. There are three types of images fed into the CNN which are based on converting EKG data time series into these images by a) spatial correlation-based, b) temporal correlation-based and c) raw beat bundle-based methods. They use the PTBDB to train the Inception-v3 model to finally achieve an IR of 98.1% over 200 users from the database while over the whole database, ELEKTRA can achieve a 99.94% of accuracy (see Table 6.2) showing higher results.

Beom-Hun Kim et al., 2020 [115]. Classification of users through their EKG by using bidirectional LSTMs and RNNs tested over the MIT-BIHDB and NSRDB is the work proposed in [115]. Their highest result is obtained by a network composed of three bilateral LSTM, a late fusion and one softmax layer achieving accuracies of 99.8% for the MIT-BIHDB and a 100% for the NSRDB. Meanwhile, with ELEKTRA very similar results are achieved by using just a one-layer CNN which is much simpler than the network used by the authors of the presented paper.

Yazhao Li et al., 2020 [111]. In chapter 3 the NSRDB was taken as the baseline to demonstrate the feasibility of ELEKTRA over control users. In the work presented in [111], the NSRDB is also the database used to evaluate their identification method through EKGs with CNN. In this case, instead of using a simple yet effective one-layer CNN, the authors use a Cascaded-CNN where three CNNs are concatenated and trained progressively. Unfortunately, despite using a more complex CNN, their achieved accuracy with the NSRDB is a 99.0% which is below the result obtained with ELEKTRA over the same database (99.84% of accuracy).

Eduardo José da Silva Luz et al., 2017 [26]. In the work presented in [26] the authors use two strategies to perform user identification, one is using the raw ECG signal and the other is using a heartbeat spectrogram representation of the EKG to finally fuse them to perform user classification. The spectrogram they use is the visual representation of the energy of the signal expressed as a function of frequency and time, which can be considered similar to the last step of the construction of the EKM where the matrix is plotted into a heatmap. Despite offering an interesting approach, their results have pretty high error rates. Over a database of healthy users (CYBHi), they obtain almost 20% of EER while with ELEKTRA over the database with control users a 0.08% is obtained. The same happens for the UofTDB where healthy users are mixed with non-healthy ones (as with the MIT-BIHDB), a 19.37% of EER is obtained while ELEKTRA obtains a 1.075% of EER over the MIT-BIHDB.

Mohamed Hammad et al., 2018 [137]. A different proposal shown in Table 6.1 is the one presented in [137] where a multi-modal biometric system is developed by fusing EKG together with fingerprint. The authors claim that an EKG is better at rejecting impostors and the fingerprint is more suitable to accept genuine users. To perform the final identification of users the EKG has to be converted into an image where each EKG beat is plotted as an individual grey-scale image. Consequently, the image used to feed into a CNN is a white image with a grey line providing irrelevant information (white spaces) to the CNN. Through the whole process of creating the EKM, the proposed model in this thesis avoids adding irrelevant information for the CNN to process by plotting the frequency on each heatmap. Hence, all the information provided in the process of user identification is useful and both works (the one presented in [137] and the one presented in this thesis) achieve very similar results as with the PTBDB in the work done in [137] a 98.66% of accuracy is obtained with 100 users while with ELEKTRA and 162 users from the same database, a 97.09% of accuracy is obtained.

T.W. Shen et al., 2002 [56]. The work proposed in [56] was one of the first works to implement user identification through EKG. As expected, new techniques and investigations have to lead to better results by achieving higher accuracies and lower error rates than 80.0% or 95.0% of accuracy with the MIT-BIHDB. Mostly all EKG studies achieve better performances.

Alex Barros et al., 2020 [167]. The work presented in [167] proposes an EKG identification system based on fiducial analysis and tests it over the Physionet database challenge called "You snooze you win" which is conformed by all healthy users, as the NSRDB. It could be considered that the main drawback of this work is to base their proposal on a fiducial analysis per se. As explained in Section 1.2.2 a fiducial analysis consists of taking as a reference the fiducial points of the EKG signal. Hence, it is computationally more costly than a non-fiducial analysis as all the fiducial points have to be computed and studied. ELEKTRA do not perform a fiducial analysis over the whole signal to perform user classification and achieve similar results. While [167] achieves an EER of 0.32% over the database "You snooze you win", ELEKTRA achieves an EER of 0.08% over the NSRDB, which is also a database comprising all healthy users.

Khairul Sidek et al., 2014 [63]. The authors of the work presented in [63] and shown in Table 6.1 test their proposal over databases with users having abnormal cardiac conditions. They use a NCN technique over the EKG segments and then use different classification algorithms to perform user identification. The authors can achieve good results in terms of accuracy over their tested databases but, for example, for the MIT-BIHDB, the authors of this presented work achieve 96.7% of accuracy while with ELEKTRA over the same database a 97.89% of accuracy is achieved.

Ronald Salloum et al., 2017 [23]. Promising results are achieved in the work done in [23] where the identification of EKG takes place through RNN and NN with LSTM and GRU. The input to their NN is a vector of several consecutive heartbeat waveforms. To achieve the results shown in Table 6.1 the authors use 9 heartbeats while with ELEKTRA 3, 5 o 7 heartbeats are used. When they set the number of heartbeats used to three, the result changes to 98.2% of accuracy while with ELEKTRA a 96.8% is obtained which is a very similar result achieved with a much simpler system. All in all, their good results are achieved by using complex NN and a high number of heartbeats per instance.

Gokhan Altan et al., 2019 [136]. A Second Order Difference Plot is a non-linear time-series analysis method that allows extracting features. The authors of [136] have based their identification work over the SODP together with a k-NN to perform user classification. Their method is tested over three different databases as shown in Table 6.1. The accuracy of their model is evaluated achieving 99.86% and 95.12% for the NSRDB and the MIT-BIHDB, respectively. While, for the same databases, ELEKTRA achieves 99.84% and

97.89%, respectively.

Qingxue Zhang et al., 2017 *HeartID* [24]. In line with the user identification using CNNs, the work presented in [24] presents their proposal based on EKG segments that are transformed to the wavelet domain to set them as the input of a 1-D CNN. The authors test it over eight different databases with control users and users with CVD, but their results in terms of the IR are not as accurate as the ones presented with ELEKTRA. As it can be seen in Table 6.1 their IR over the NSRDB is a 95.1% while the IR obtained in this thesis with the same database a 99.98%. The same happens with the MIT-BIHDB where the authors obtain a 91.1% of IR while in Chapter 4 an IR of 99.91% is offered.

Joao Ribeiro Pinto et al., 2020 [113]. The next work in Table 6.1 is the one found in [113] where this proposal is based on a 1-D CNN, where the training is set over one database and the testing of the identification of users, is done with another database. It is a very interesting proposal but it may not offer very good performance as seen in their results. Their highest result over the PTBDB is an EER of 9.06% while, over the same database, the result achieved with ELEKTRA is a 3.04% of EER.

Mohamed Hammad et al., 2019 [113]. In the work presented in [117] very good results are offered by achieving really low error rates. That might be due to the that their proposal is based on fusing features extracted from a fiducial analysis of the EKG signal together with one beat of the signal to set them as the input of a 6-layer CNN. Hence, there is evidence that the authors offer very good results but to do so, they have to use a very complex model with a complex architecture which includes a fiducial analysis which may be considered computationally costly compared to ELEKTRA. Hence, while in [117] an EER of 1.63% is obtained over the PTBDB by fusing the fiducial analysis of the signal with one beat of it, with ELEKTRA and a one-layer of a 2D CNN a 3.04% of EER is achieved.

Zhidong Zhao et al., 2018 [133]. The last work in Table 6.1 is the one found in [133] where the authors convert the EKG signal into a 2D image through the S-transformation (by plotting the EKG trajectory image). To use a transformation of an EKG into an image to feed a 2D CNN can be a delicate task as the whole image has to provide useful and relevant information to the CNN to be as efficient as possible. Therefore, some works have commented that they do different transformations to the EKG signal to obtain an image so the authors can use it as the classification method. As is the case of some commented works such as [26], [137] and the work just commented in [133]. The main drawback of some of them ([26] and [132]) is that the authors create the image with lots of blank spaces providing irrelevant information instead of plotting different features (such as the frequency) as [137] or ELEKTRA does. As a consequence, despite having good results, a proposal which considers images to provide useless information is considered to be less efficient than other works in the state-of-the-art.

6.2 Summary of ELEKTRA's results

In this Section, Table 6.2 is presented as a summary of the results obtained during the experimentation with four databases over ELEKTRA as a biometric system for user identification with EKG. The best results are shown for each case regardless of the number of bpf chosen to extract the EKMs. In the case of the GUDB the result obtained from the *sitting* scenario is the best approach with this database.

Proposal	DDBB	Subjects	Condition	Methodology	Metric	Result (%)
Jikui Liu et al., 2018 [99]	PTBDB	50 50	healthy non-healthy	AutoRegressive Models	Acc	98.0 100
Joao Ribeiro Pinto et al., 2020 [113]	PTBDB	290	healthy and non-healthy	IT-CNN	EER	9.06
Ruggero Donida Labati et al., 2019 *Deep-ECG* [93]	PTBDB	52	healthy	CNN	EER	1.63
Yifan Chu et al., 2019 [94]	PTBDB MIT-BIHDB	52 47	healthy healthy and non-healthy	CNN + Euclidian Distance	EER	0.59 4.74
Pei-Lun Hong et al., 2019 [129]	PTBDB	200	healthy and non-healthy	Pretrained CNN	IR	98.1
Beom-Hun Kim et al., 2020 [115]	MIT-BIHDB NSRDB	47 18	healthy and non-healthy healthy	biLSTM	Acc	99.8 100
Yazhao Li et al., 2020 [111]	NSRDB	18	healthy	Cascaded-CNN	Acc	99.0
Eduardo José da Silva Luz et al., 2017 [26]	CYBHi UofTDB	65 1020	healthy healthy and non-healthy	CNN Spectrogram	EER	19.57 19.37
Mohamed Hammad et al., 2018 [137]	CYBHi PTBDB	63 100	healthy healthy and non-healthy	CNN-2D Plot	Acc	98.97 98.66
T.W. Shen et al., 2002 [56]	MIT-BIHDB	47	healthy and non-healthy	DBNN Template Matching	Acc	80.0 95.0
Alex Barros et al., 2020 [167]	You Snooze You Win	1985	healthy	Fiducial Analysis	EER	0.32
Khairul Sidek et al., 2014 [63]	DiSciRiDB SVDB MIT-BIHDB	51 67 47	non-healthy non-healthy healthy and non-healthy	NCN+ MLP/kNN/RBF	Acc	99.3 96.4 96.7
Ronald Salloum et al., 2017 [23]	ECG-ID MIT-BIHDB	90 47	healthy healthy and non-healthy	LSTM	Acc	100.0 100.0
Gokhan Altan et al., 2019 [136]	ECG-ID NSRDB MIT-BIHDB	90 18 47	healthy healthy healthy and non-healthy	SODP + kNN	Acc	91.96 99.86 95.12
Qingxue Zhang et al., 2017 *HeartID* [24]	NSRDB MIT-BIHDB	18 47	healthy healthy and non-healthy	1D CNN	IR	95.1 91.1
Joao Ribeiro Pinto et al., 2020 [113]	PTBDB	232	healthy and non-healthy	1D CNN	EER	9.06
Mohamed Hammad et al., 2019 [117]	CYBHi PTBDB	65 232	healthy healthy and non-healthy	CNN + Fiducial analysis	EER	4.47 1.63
Zhidong Zhao et al., 2018 [133]	ECG-ID 2017 Physionet	50 50 50	healthy non-healthy noisy	S-transformation + CNN	Acc	96.63 96.23 96.18

Table 6.1: Comparative analysis of EKG identification among the state-of-the-art.

†Notation: Acc (Accuracy); EER (Equal Error Rate); IR (Identification Rate).

DDBB	**Subjects**	**Condition**	**Metric**	**Result (%)**
NSRDB	18	healthy	Acc	99.84
			IR	99.98
			EER	0.08
MIT-BIHDB	47	healthy and non-healthy	Acc	97.89
			IR	99.91
			EER	1.075
PTBDB	232	healthy and non-healthy	Acc	93.96
			IR	99.94
			EER	3.04
	162	non-healthy	Acc	97.09
			IR	99.96
			EER	1.465
GUDB	25	healthy sitting	Acc	99.19
			IR	99.93
			EER	0.42

Table 6.2: Best results obtained with ELEKTRA from every database .

[†]Notation: Acc (Accuracy); IR (Identification Rate);
EER (Equal Error Rate)

CHAPTER 7

DISCUSSION AND CONCLUSIONS

The preceding chapters developed a detailed analysis of ELEKTRA as an identification system and an in-depth study of four public EKG databases for user identification comprising control users, users with CVD and users performing different activities. All this presented information and chapters showed the proposed system's feasibility through its accuracy and error rate results. In this last chapter, a set of discussions that had to be faced during the elaboration of this project are shown together with the summarised conclusions and achievements and the further work suggested as the next step with ELEKTRA.

7.1 Discussion

In this section of this last chapter, different discussions, choices and problems that have been faced during the elaboration of the thesis, or the elaboration of some methods (seen in Chapter 2) and launching of the experiments are seen.

7.1.1 Creation of the EKM segments

The work presented in [126] shows the first development and application of the EKM. Despite having a fascinating approach, the process of creating the EKM is not explained in detail. For this reason, we asked the research authors if there was any available implementation of how to create each EKM, but there was not. Hence, we had to create our own implementation which is available in Github[1] and in Appendix A.1. As it has been commented in Section 2.3, where the creation of the EKM is explained, each EKG recording has to be split into windows[2] to then be split again into smaller segments containing two peaks each. To define begin and end of each segment of the window, the following equations are followed:

$$InitSegment = p_x - \alpha_i \mu \tag{7.1}$$

$$EndSegment = p_x + \alpha_i \mu \tag{7.2}$$

where p_x is each peak from the current window, α_i is a hyperparameter set for how much of the signal is going to be taken before and after the peak

[1] Github available at: `https://github.com/cfusterbarcelo/ELEKTRA-approach`

[2] How large the windows are will depend on the number of parameters chosen to appear on each EKM (see discussion on Section 7.1.3)

(expressed as a percentage) and μ is the mean distance between peaks. This equation is used and expressed on Algorithm 2.

Hence, the free parameters α_i and α_e are two parameters that do not depend on the database. They have to be fixed for all the users in the same experiment. In detail, if we are comparing N users, the values of α_i and α_e have to be the same for all the users to have the same matrix construction. As shown in Figure 7.1, and as it has been commented, these two parameters represent a percentage of how much we want to take before and after the first peak (in a segment). Meaning that α_i will have to be a value under a 50% of the signal to take just some samples before the R_{peak} and α_e will be a value above a 100% of the signal to have the representation of the next peak on this segment. As a result, α_i and α_e are two hyperparameters that can be chosen depending on the database to adjust how the EKM is plotted.

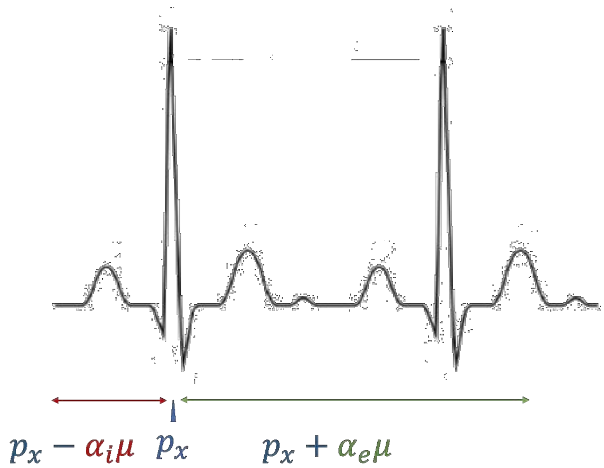


Figure 7.1: EKG segment to create the EKM, detailed in Algorithm 2

In the case of the μ value, which is expressing the mean distance between peaks, it is calculated for each user of the database after the normalisation of the signal. By calculating the mean distance between peaks, we can construct the segments always taking two peaks for each segment. As each peak of the window is being taken as a reference to construct each segment, the process is repeated for each peak from the same window until all the peaks from that window (3, 5 or 7 bps are taken in each window) are all segmented. The only peak that is not taken is the last peak from the window as, if it was taken, we would have to use the next peak of the next window which is not included in the window that is being treated.

The discussion faced here, regarding the construction of each segment of the EKM, is offered by the possibility of using either the minimum, the mean or the maximum distance between peaks. The main goal is to plot always two R-peaks in each segment. Thus, each of the possibilities has been studied and considered to construct the EKM segment:

Minimum. If the minimum distance between peaks is taken, probably and frequently, not two peaks would be taken to be drawn in each segment as the minimum distance could not be enough. As a possible case, there exists a physical phenomenon which is called Heart Rate Variability (HRV)[3] which consists of the variation in time between consecutive heartbeats or R-peaks, that is, the time that happens between consecutive heartbeats can vary. This phenomenon can happen almost undetected but, when evaluating the distances between peaks, it can affect the final result of it. Therefore, if the minimum distance between peaks was the one used to construct the segments, taking the HRV into account, it may not be enough to show always two peaks on the same segment.

Maximum. Again, we could have that HRV could affect the distances between peaks but instead of reducing this time or distance between the consecutive peaks, it could increase it resulting in some of the heartbeats of the EKG signal being significantly apart from each other. Consequently, if the maximum distance between peaks is taken, it might be possible that, in some cases, three peaks were shown in the EKM segment instead of two.

Mean. Therefore, the most exact option to perform this step is to take the mean distance between peaks. By taking the mean distance it does not mean that in some of the EKMs there will appear only one peak or even three. But the possibility of this happening is significantly lower than with taking the minimum or maximum distances between peaks.

In fact, if for a database, it is discovered that even though we have taken the mean distance to calculate and construct each segment, only one peak is appearing for some of the segments, we could force it by changing the $alpha_e$ value. If this is a general problem for the database, the possible solution taken can be that, instead of, for example, taking the 130% of the signal multiplied by the mean distance after the first peak, take the 150% of the signal making sure that two peaks are taken and shown in the EKM segment.

7.1.2 Repeating heartbeats when creating the EKM

When creating the EKM, each one is created and extracted for each window segmented from the EKG signal. This means that for each window, we will obtain one EKM. This implies that if the EKG signal is not too large, not many EKMs can be extracted. Or that, if the number of bpfs chosen to create the EKM is 7 (this will be discussed in the next Section 7.1.3) the user that is being identified has to wait for more time because more beats are needed to perform user identification.

A possible solution for this, from a cybersecurity point of view, could be to create one EKM for each beat processed instead of for each window. This means that, if we have a window of 7bpf, instead of creating one single EKM for that window, we would create seven EKMs. This way, the windows will

[3] `https://www.firstbeat.com/en/blog/what-is-heart-rate-variability-hrv/`

only be used to plot each EKM but an EKM will be extracted from each beat of the signal. By this, the number of EKMs extracted would be, in this case, seven times more. With this, two improvements would be made:

1. More EKM images will be extracted for each EKG recording implying that there would be more images for either training and testing with the CNN and it might be an improvement of the results obtained.

2. Less time would be needed to identify a person as the EKMs would be extracted more frequently.

From a cybersecurity point of view, some advantages are offered. But, from an information theory point of view, we have to take into account that we would be repeating a lot of information and a lot of beats would be represented more than one time. Seeing that, some problems would be derived from using each heartbeat to construct an EKM:

1. Extracting an EKM for each beat will make that if for example 7bpfs are taken, each heartbeat or R_{peak} would be shown in seven different EKMs. This means that what is shown in an EKM would also be contained in the other six EKMs. This implies that all EKMs would be very similar among them and it could affect the classification.

2. In this work, the training and testing subsets of images have been randomly chosen. This means that, once all the EKMs are created, they are randomly sorted into training and testing subsets depending on the number of images that we want for each of these subsets. With this new approach, this would not be possible. As the information contained in one EKM is also shown in the other six EKMs, to split randomly the images in train and test would make that the information would be repeated in train and test and the classification with a CNN would not be fair. In any Neural Network, when the classification is performed, there has to be a clear partition between the training and testing information because if the information that is going to be used in the test has already been seen in the train, the model will be already trained with the supposedly non-seen information and the final classification will not be fair.

A possible solution and use of this methodology in which each heartbeat extracts an EKM could be to use it in a real-world application. This means that the EKMs that would be used for training the model, would be extracted as it has been done in this thesis. However, once we have the model already trained, the testing images to identify a person in a real-world scenario could be obtained by constructing each EKM for each heartbeat to take the advantages that we have already commented on in the testing phase. This may be possible as it is not important if the information is repeated when identifying a user.

7.1.3 Using 3, 5 and 7bpf

In the work presented in [126], which is the base work for this doctoral thesis regarding the EKM, the authors segment each EKG signal into windows containing 7 beats each. Other subsequent works such as [125], have reproduced the work done in [126] and used the same method to construct the EKM with 7 beats in each window.

When this project was started (in 2020) there was not any work using a CNN to process the EKM. Hence, we had to face how many beats were going to be taken to construct the EKM to then process the images to perform user identification.

Even though, in 2021, Salinas et al., used the EKM together with a CNN to detect Atrial Fibrillation in [141]. To do so, they used windows of 10 consecutive beats plus 3 seconds to construct the EKMs and feed them into a CNN.

By checking the literature regarding window segmentation to perform user identification, many different approaches can be found where the segmentation of the EKG signal is performed either depending on a certain number of heartbeats or in time intervals. Concerning time intervals, many works use 3s as it is known that a typical heart rate ranges from 40 to 208 beats per minute. Therefore, 3s must be enough to have at least one heartbeat on each window if the starting point of the window is set to the P wave from that heartbeat or the T wave from the previous one. This is the case of the works found in [132] and [167]. Other research is more conservative and, to make sure there is at least one heartbeat per signal, take a window of 5s as it is the case of [118] or [136]. And other do not give that important if there is an entire cardiac cycle or not on their window, so they use 2s windows of the EKG signal to perform user identification.

Concerning how many heartbeats have been chosen in the literature to perform user identification with EKGs, again, many values can be found. In the works presented in [94] and [132], the window segmentation is performed after each heartbeat. The opposite case can be found in [25] where the authors use segments of 10 heartbeats. Then, there is the special case of [23] where they test their proposal by segmenting the window every 3 or 9 heartbeats. Their best result is achieved by using 9 heartbeats.

Owing to the variety of approaches regarding window segmentation, we had to face the decision of how many heartbeats were needed or enough to construct the EKM and identify users. There is one sure thing: to construct an EKM an odd number of heartbeats are needed because of the nature of the EKM. For example, for the work performed in [141] where the EKM was created using 10 heartbeats plus 3 seconds, they use an even number of heartbeats but, this is the reason that they had to add 3 more seconds, to add one more heartbeat. Consequently, by adding one more heartbeat that does not belong in the segmented window, we would have this extra heartbeat would appear in two different EKMs. Therefore, the information regarding that extra heartbeat would be represented and repeated in two different EKMs. This problem is similar to the one faced in Section 7.1.2 where there was an information repetition. Since that, these are the number of bpf that could be considered:

1bpf. As the EKM is, in fact, a coloured matrix, if we consider that the R_{peak} defines the number of columns and rows of the matrix, the minimum number of R_{peak} that has to be taken is 3. As we can see in Figure 7.2a where a matrix with 3 R_{peak} is plotted, obtaining an EKM with fewer beats would not be possible. If we had only 1 data point (one R_{peak}) a matrix could not be constructed because we would have just that: one data point, not a matrix.

3bpf. As we have just seen, the minimum number of beats that can be taken to construct the EKM is 3bpf. It is important to remark that as the second R_{peak} appears twice in this EKM (see Figure 7.2a) while the first and third peaks appear only once, we could consider that the information regarding the second peak is duplicated compared to the other two peaks. Even though, from a Cybersecurity point of view, as we introduced in Section 7.1.1 it would be interesting to offer user identification with windows of only 3 heartbeats as more images can be extracted per minute and, consequently, less time would be needed to identify one person in a real-world application.

5bpf. As commented before, from the second peak to the penultimate one are repeated twice in the creation of an EKM. In this case, for 5bpf, the repetition of information is fairer than for the 3bpf case. In Figure 7.2b we can see how the peaks number 2, 3 and 4 would be the ones repeated. The approach with 5bpf would be the more fair regarding two important points: i) the number of heartbeats that appear repeated and, ii) the number of EKMs that can be extracted from the EKG signal. Therefore, using 5bpf is one of the best approaches for this particular case.

7bpf. Taking 7bpf was the value that was chosen by the authors who first implemented the EKM. Of course, their main goal was not user identification but inspection of cardiac signals for medical diagnosis. Hence, for them, it might not be as important as for user classification how many EKMs can be extracted from one EKG recording. If we focus on the information repeated from the second to the penultimate R_{peak} we can see in Figure 7.2c that the peaks number 2, 3, 4, 5 and 6 are repeated twice and the first and last peak are not repeated. Hence, we could consider that there is not any peak which is more represented than others. However, the problem that could be faced by having 7bpf to create the

EKM is that not that much number of EKMs can be extracted from the EKG signal as with 3bpf, for example. It is the case for the GUDB, the EKG recordings were not long enough to extract sufficient EKM images with 7bpf and perform user identification.

9bpf. As we have seen, other works in the literature, consider 9bpf as the number of heartbeats used to perform user identification. Even though, we considered that 9bpf would decrease even more the number of EKMs that can be extracted from the EKG signal and, consequently, decrease the performance of our method because the CNN would not have enough images to properly train.

(a) Random EKM from PTBDB 3bpf

(b) Random EKM from PTBDB 5bpf

(c) Random EKM from PTBDB 7bpf

Figure 7.2: Random EKMs from the PTBDB for 3, 5 and 7bpf

In conclusion, to have a fair representation of each of the heartbeats shown in the EKM and to obtain as many EKMs as we could, we considered that the best approaches have been obtained with 3, 5 and 7bpf being the 5bpf the fairest option.

7.1.4 Chosen Databases

In our experiments, we use three well-known databases (NSRDB, MIT-BIHDB and PTBDB) used widely and previously in the literature together with the GUDB [142, 143, 135, 144]. Reviewing the Table 6.1 in Chapter 6, we can see that numerous previous studies that we have used to compare our model to have also been validated with these same databases [23, 129, 93, 63, 115, 136, 24, 111, 113].

The first database that we chose to perform user identification with ELEKTRA was the NSRDB. We had a set of requirements for this database to be tested with our work. These are: i) a long-term database with long EKG recordings to be able to test with a huge number of images, ii) to be freely available, iii) with control or healthy users to have no biases between users and, iv) with `.dat` and `.hea` files to be processed as it was what we had processed before to do some first tests. The NSRDB fulfilled all of our requirements. Hence, there were not many doubts about using it. Even more, this database has been used for user identification in the literature thus we were validated by the peer-reviewed journals that already used it.

To follow the study over ELEKTRA but over users with CVD, we required two different databases comprising different types of users (healthy and non-healthy) and users with different types of CVD. The first database that was proposed was the MIT-BIHDB as it is one of the most famous databases for user identification with EKGs and even for the classification of EKGs for other purposes such as medical diagnosis. This database is also available in Physionet and contains the same type of files as the previous database making the preprocessing steps simple. In fact, this database has also been used in many research before making us confident of using it another time. On the contrary, we had some difficulties finding another database to perform user identification over users with CVD that fulfilled all our requirements. What we needed, in this case, was a database that comprised healthy users and users with different CVD as in the MIT-BIHDB. It might include Arrhythmia patients but we did not want all of them to have Arrhythmia as it was already studied in the previous database. By going through almost all databases from Physionet and classifying them on how many users each database have, the condition of the users (healthy, non-healthy or with which CVD) and, of course, how large the EKG recordings where we discovered some database that fulfilled some of our requirements:

MIT-BIH Malignant Ventricular Ectopy Database[4]. This database is comprised of twenty-two recordings of thirty minutes each from users with ventricular tachycardia, ventricular flutter, and ventricular fibrillation. It also has the type of files that we already knew how to process. The problem with this database is that it does not contain healthy users and the number of subjects contained is not sufficient enough. Therefore, we could not offer a classification where healthy and non-healthy users were mixed.

Sudden Cardiac Death Holter Database[5] This database is very similar to the previous one but the number of recordings and subjects is different. The 18 subjects comprised in this database were recorded to obtain 23 EKG recordings. These subjects suffer mostly from ventricular tachyarrhythmia and cardiac arrest. Hence, again, not enough users are presented in this database and all of them have a CVD.

Physikalisch-Technische Bundesanstalt[143] The database used was finally the PTBDB as it fulfills all the requirements. As it has been explained before, it holds 290 subjects with and without CVD. In fact, in the `.hea` file where all the information regarding each subject is contained, there is a line which specifies which is the CVD that the users have (i.e., `# Reason for admission: Myocardial infarction`). Thanks to that,

[4] The MIT-BIH Malignant Ventricular Ectopy Database is available at `https://www.physionet.org/content/vfdb/1.0.0/`

[5] The Sudden Cardiac Death Holter Database is available at `https://www.physionet.org/content/sddb/1.0.0/`.

before processing the signal, we can classify the users depending on the CVD that they have and include them or not depending on what we want to study. Moreover, the recordings from this database are between 38 to 104 seconds and there are 290 subjects which is more than sufficient to perform user classification and to show ELEKTRA's adaptability and reliability.

For the last experiments, we thought that a good proposal would be to test our biometric system over users performing different activities to offer a different approach that has not been widely studied. When performing the first steps of the research regarding user identification, we discovered the GUDB[144]. Hence, as this type of database comprising users in different scenarios or situations is not very common, we rapidly opted for this database as it had each of the users performing each of the activities. This way, a fair comparison between activities without biases between users, would be offered.

7.1.5 Gaussian Noise

In the experiment performed in Chapter 3.5, Gaussian Noise has been added to the EKG signal to simulate real noise in a real-wold scenario. Hence, it can be discussed if Gaussian Noise could be similar to a real noise obtained from the environment.

The Central Limit Theorem (CLT) [168] is a well-known theorem in probability and statistics because it states that, under certain conditions, when independent random variables are added will tend towards a normal distribution. Thus, this means that many random variables from many different areas such as shoe size, birth weight, IQ or even noise coming from low-resolution sensors, will tend to follow a normal distribution. Noise is the "byproduct of interference from potentially many different sources[6]". Hence, if many of these products coming from different sources are joined in one distribution, it can be modeled as a Gaussian distribution.

In conclusion, as the CLT confirms, we could state that Gaussian Noise can behave as noise obtained from a real-life situation

7.1.6 Ad-hoc or Pre-trained CNN

A Convolutional Neural Network is a classification algorithm capable of exploiting the spatial correlation between pixels in images. Hence, this type of Neural Network is needed to perform user classification using images. In the literature, it is widely common to use pre-trained CNNs for such a job. For this reason, we tried to perform user classification with an ad-hoc CNN as well as with a pre-trained one.

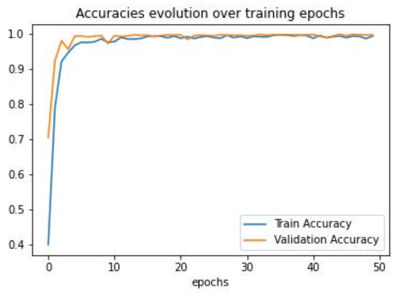
We first tried to launch a tailored CNN which we built for this classification problem. After launching some experiments, we saw that just one

[6]Random Noise and the Central Limit Theorem https://gregorygundersen.com/blog/2019/02/01/clt/

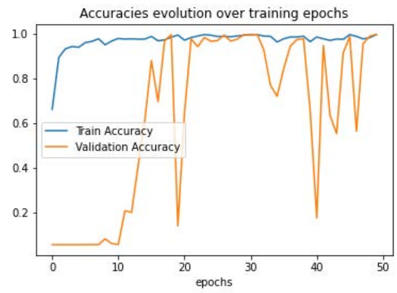
convolutional layer was enough to extract the necessary information from the images to perform user classification.

To study how a pre-trained CNN would cope with user identification with EKMs, we studied the ResNet50 [169]. This CNN is a 50-layer CNN designed to solve computer vision problems such as object detection or image recognition. To use this network, we added two classification layers to the pre-trained CNN: a fine-tuning of the last layer of the network and a final classification with a softmax activation (remember that we are facing a categorical classification).

To offer a comparison between the two approaches (a tailored CNN and a pre-trained one) we trained both networks with the same database (the NSRDB) and the same images, during 50 epochs and with a 70% of dropout to offer regularisation and to prevent overfitting. The result of these two training phases can be seen in Figure 7.3



(a) Accuracy evolution for training and validation with the two-layer CNN for 50 epochs.



(b) Accuracy evolution for training and validation with the ResNet50 for 50 epochs.

Figure 7.3: Differences between using a particular CNN or a pre-trained one on training.

It can be observed that many differences are noted between the two training. As a result, a smooth evolution of the training and validation accuracy for the ad-hoc CNN is obtained and a spiky one for the ResNet50. A very complex CNN such as the ResNet50 is designed to find latent information that is somehow hidden in real data. As in our problem we have found data that perfectly identifies each user (the EKM), a complex network suffers from overfitting as it is extracting too complex information from really intuitive data.

Consequently, we discarded the idea of using a pre-trained network as the ones offered in the literature are too complex inducing overfit. The CNN finally used is detailed in Appendix A.2.

7.2 Conclusions

Identification of users with Electrocardiogram (ECG) or Elektrokardiogramm (EKG) is a widely used technique since 2002 [21] and many proposals including many different approaches exist since then as seen in Section 1.2.

This thesis proposes a cutting-edge biometric identification technique based on the novel representation of an EKG trace named Electrocardiomatrix (ECM or EKM). Recall that EKMs are a heatmap representation of several aligned peaks of an EKG composed within a matrix as explained in Chapter 2. To show the identification proposal's feasibility, a simple CNN is used, showing that the approach presented in [13] is enough to identify each user with just one convolutional layer.

It is important to remark that, apart from achieving good results in terms of accuracies or error rates, the main focus of this work is to show ELEKTRA's feasibility by using a very simple CNN which shows that is capable to achieve those higher results.

To show ELEKTRA's performance, different experiments have been conducted during chapters 3, 4 and 5 of this thesis. A summary of the achieved results in the preceding chapters can be found in Table 6.2.

In Chapter 3, ELEKTRA has been tested over healthy users (control users) with the NSRDB. Really low error rates and high accuracy have been obtained showing promising results with 3, 5 or 7bpf. A test to simulate a real-world scenario with a One-vs-the-rest classifier is offered to show the adaptability of the method proposed. ELEKTRA has even been tested over the classification of a non-seen user and a user whose signal has been added noise achieving always outperforming results. To test ELEKTRA's feasibility over control users, the size of the dataset has been increased until only 300 images per user having just a slight decrease in the results.

Following the tests performed to show this thesis proposal, it has been tested over users with different Cardiovascular Disease through the MIT-BIHDB and PTBDB in Chapter 4. As seen in the experimentation with both databases, the results obtained are really good in terms of accuracy and error rates and competitive with the current state-of-the-art. Not many works are focused on studying how a CVD can affect the identification of users through their EKG as it is done in this thesis. However, studies of classification techniques of different CVD such as Myocardial Infarction, Arrhythmia, Ventricular Ectopic Beats (VEB) and Supraventricular Ectopic Beats (SVEB) among others are very common (see [170, 171, 172, 173, 174]).

To sum up, the extensive study of the two databases commented how ELEKTRA can adapt to databases closer to a real-life situation where control users and users with cardiac diseases are mixed.

Finally, the last experiments presented in this work in Chapter 5 are the ones over the GUDB showing how variability over the heartbeat rate does affect user identification. By performing experiments with over 25 users doing five different activities (sitting, walking, doing a maths test, biking and jogging) it is confirmed that when a user is doing an activity that requires a cardiovascular effort, it is more complex to identify. On the contrary, when a user is sitting or doing a more calmly activity such as walking, higher accuracies and lower error rates are obtained.

The fact that identifying a user with higher heart rates is more complex might be due to that a NN may focus on the possible noise of the EKG signal instead of the user's characteristics that make them unique.

As seen in all experiments, ELEKTRA is suggested as a biometric system for accessing high-profile areas or assets such as a bank, a locker, a government area or others as it has been demonstrated that it is a suitable identification method and it might be more difficult to be identified while being nervous or overwhelmed.

As a matter of fact, the work presented in this thesis, as far as the literature review covers, is the first work to show that identifying a user with EKGs when this user is nervous or anxious with higher heart rates, is more difficult.

With ELEKTRA a lot of statements are made as seen during all the chapters. One of the most important affirmations of this thesis is the lack of publicly available datasets. Many works perform user identification over EKG or even other biometric traits with their private database. On the contrary, in this thesis, reproducibility is assured by using exclusively public datasets.

Also, we can forecast the system acceptability due to the possibility of acquiring an EKG through non-invasive devices such as a smartwatch and the widespread adoption of these devices [175]. Furthermore, system usability is guaranteed thanks that 3, 5 or 7 heartbeats are enough to generate the input images (EKMs). Collectibility is also satisfied even via low-cost EKG sensors.

Another benefit that is acquired in a biometric system based on EKG and, even more with the EKM is the possibility to acquire a diagnosis from the EKG signal used to identify the user. As the previous works with the EKM are all based on CVD diagnosis, it can be assured that ELEKTRA's system enables a possible diagnosis of the user who is being identified.

In conclusion, by the studies in this thesis, it can be assumed that ELEKTRA is a feasible and efficient identification method for biometrics with Elektrokardiograms.

7.3 Suggestions for Future Research and Application

In this last section of the thesis, future lines to follow the investigation or how to apply ELEKTRA in the real world are studied. Different approaches are discussed and investigated to have different perspectives on how ELEKTRA can be continued, as it has been demonstrated to be a feasible method for user identification.

7.3.1 Longitudinal Study

Despite having an in-depth study over ELEKTRA as an identification method, one of the possible approaches to focus on future research is a longitudinal perspective studying the time permanence of identification of users with EKG with the presented method.

A few works or proposals exist studying the stability over time for EKG biometrics, many of them over private databases. It is the example of [60] in which Odinaka et al. measured ECG stability using three user sessions on different days from a private database, achieving promising results (93.5% of Accuracy and 4.75% of EER). Also, the work presented in [176] feeds a 2D representation of an EKG to a CNN obtained from a private database comprising 400 users. In their work, the authors conclude that the experiments where one single session of enrollment is used to train effects negatively by making the accuracy degrade as time pass. Nevertheless, when more EKG records from different sessions are used, the results in terms of accuracy improve.

Then, the works found in [177], and [26] confirm an evident decay over time when training with one session and testing with a session from another day using the E-HOL-03-0202-003 and CYBHi databases, respectively. In brief, even though the use of the EKG guarantees the universality (everyone alive is beating) and permanence (cardiac signals are stable during at least five years [178]) of the system, presumably, there exists a need of training with more than one day to achieve good performance over time stability with an EKG identification method. For this reason, future work comprising a balanced database with sessions from different records from different days would be needed to achieve a final evaluation over time permanence with ELEKTRA.

In fact, in a real-world scenario, to overcome the possible obstacles with permanence over the EKG signal, a good approach could be to retrain the last layer of the CNN (performing transfer learning as it is suggested in [177]) each time a user is being identified into the system to re-enrol each subject each time and adapt the model and the biometric system to possible changes over time.

7.3.2 ELEKTRA implementation over a wearable

Regarding a real-scenario approach, a future line of work is widely open. To test ELEKTRA in the wild, first of all, a wearable is needed to record real user EKGs. Current literature is developing low-cost wearable systems that record your EKGs in real-time, such as [179], where they claim that their EKGs follow the gold standard of Physionet (same databases as used in this thesis). Other authors, such as [180], create an EKG recorder wearable system and train a Neural Network to identify cardiac arrhythmia. Moreover, research has been done recording EKGs in the wild and using, then, a CNN to identify arrhythmias [181]. As seen, recording EKGs with a wearable is feasible. Therefore, as future work, we could test the feasibility of ELEKTRA over a real scenario.

Then, in collaboration with other researchers, we could develop a wearable system to be tested directly with ELEKTRA. First, we could record EKGs over a certain time for each user to train the ELEKTRA model. Once trained, ELEKTRA would be tested over the same user to identify them every time they connect the wearable via Bluetooth to the smartphone (similar to the OnevsRest classifier seen in Section 3.3). Finally, we could create an enrollment system which, after a specific amount of EKGs are collected, retrains ELEKTRA making the model dynamic and keeping it actualised.

7.3.3 ELEKTRA's feasibility over low-resolution sensors

All databases used in the studies of this thesis have been medical ones, i.e., with quality sensors recording the EKGs. Following the previously exposed in Section 7.3.2, a proof-of-concept over low-quality sensors could be done as future work.

Current research is done recording EKGs with low-resolutions sensors, such as [182, 183, 179, 180, 181] . However, any of that works explicitly says if those databases are public to the research community. Nevertheless, other works such as [184] present a public and open database of EKGs recorded with low-resolution sensors.

Hence, a future line of work could be to test ELEKTRA's feasibility over these databases coming from low-resolution sensors and see its performance in an environment similar to the real one.

7.3.4 Photoplethysmomatrix

A Photoplethysmogram (PPG) is an optical technique applied to the monitoring of the cardiac signal similar to an EKG. One of its main characteristics is that it is capable of measuring HRV. The PPG measures changes in blood volumes through tissues by the emission of light rays [185]. In figure 7.4 there is shown how a PPG signal looks like. The PPG signal can be captured from a small and low-cost sensor totally feasible to be connected into a wearable.

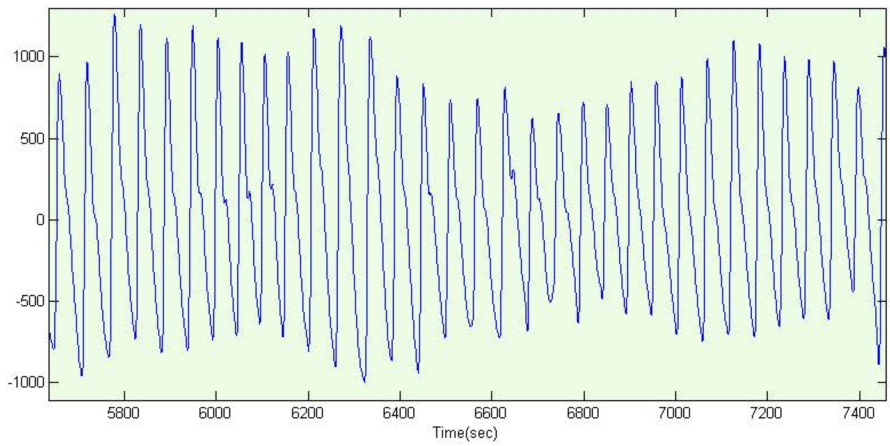


Figure 7.4: PPG Image

For example, in the work presented in [186] the authors developed a peak detection algorithm for PPG signals registered from a wearable.

Thus, a future line of experimentation could be to construct the Elektrokardiomatrix (EKM) but instead of using an EKG signal, using a PPG to finally have a Photoplethysmomatrix (PPM). As we have seen for the construction of the EKM in Sections 2.3 and A.1, the R_{peaks} of the EKG signal need to be detected to construct each EKM. As the PPG signal also records those peaks, it might be possible to build a similar matrix with the PPG(the PPM).

Once it is empirically proven that the PPM can be constructed and have meaningful information, this thesis could be reproduced over this type of data.

7.3.5 User impersonation with a GAN

A future research line is open regarding the security of our approach and AI. Nowadays, Generative Adversarial Networks (GAN)s are used to generate fake images (such as faces), and bypass identification controls [187]. In this field, literature has been creating powerful models to bypass face identification [188] and real face generation [189]. Hence, a future line of work could be a feasibility test for ELEKTRA to know if it can be fooled.

The procedure of this test would be the following. First, we would create a GAN to generate realistic EKMs of a given user. A GAN is a network that consists of two parts: a generator and a discriminator. The generator usually is an AutoEncoder (AE), or a Variational AutoEncoder [190], which given random noise can generate an image by minimising a reconstruction loss and a Kullback Leibler (KL) divergence over the latent space. In this case, we would train the reconstruction loss to reconstruct the EKMs of a given user. Then, a discriminator consisting of an MLP binary classifier is trained to distinguish between the generator's fake EKMs and the real user's real EKMs. Then, both networks are trained simultaneously. Hence, the generator learns how to create images that can fool the discriminator, i.e., realistic EKMs indistinguishable from real ones.

In conclusion, a future study could use these realistic EKMs to test the feasibility of our biometric system to a fake image generator attack.

APPENDIX A

AN IN-DEPTH STUDY OF EKM AND CNN

Throughout the document, the processes of how an EKM is created and how our CNN is constructed are presented. This chapter will detail an in-depth study of the processes needed to create the EKM database and the CNN. Algorithms with pseudo-code are offered to illustrate those processes and fully understand the steps, parameters and hyperparameters required in each case. The detailed Python code is available in a Github project called ELEKTRA-approach.

A.1 Creation of an EKM dataset

The process of creating an EKM has already been presented in Section 2.3. Some other aspects regarding the construction of the EKM or its databases are studied. In Chapter 7 a discussion of these aspects referring to how and why we have chosen those parameters to construct the EKM is offered. The mean distance used to calculate each segment of the EKM is discussed in Section 7.1.1 and showed in detail in Figure 7.1. How many beats per frame or EKM have been used to create the EKM is specified and discussed in Section 7.1.3. And lastly, the possibility of creating an EKM for each of the heartbeats instead of for each window segmented from the EKG signal is studied in Section 7.1.2.

Once all these aspects are studied, in this section, we will see in detail the necessary steps, parameters and hyperparameters to create each EKM and EKM database.

First of all, to create the EKM dataset, some specific and specialised Python libraries have been used:

- **SciPy**. This library provides different algorithms as an extension to NumPy, providing additional tools for array computing for specialised data. In this case, it is used for signal processing.

- **Biosignalsnotebooks**. It is a Python library containing a set of notebooks and tools to perform signal processing. The algorithm Pan and Tompkins is the one implemented and used through this library.

- **wfdb**. The Python package waveform-database (wfdb) is a library of tools for reading, writing, and processing WFDB signals and annotations.

- **Seaborn**. It is a Python library for data visualisation based on matplotlib. With it, we can plot the EKM matrix into a heatmap.

In Algorithm 1 the process of creating the EKM dataset is presented. The first step to perform user identification or classification with EKGs is to obtain the EKG recordings from the databases chosen. In our work, as we have commented, we have downloaded three of our databases from Physionet[1] with an `wget` command. As databases from Physionet are not standardised, each of the databases, regardless of the type files (`.hea` and `.dat`), comes in a different directory format. For example, for the NSRDB, each file is contained in one single directory. Otherwise, for the PTBDB, each file is included inside a directory for each user. Hence, the first step, which is to read all `.hea` and `.dat` files, must be adapted to each database.

Algorithm 1 Creation of an EKM database

Define path to .hea and .dat files
Initialise: $Initial\,window = 0$, sf
Define Hyperparameters: bpf, α_i, α_e, train percentage, #EKM total
R peaks list, filtered EKG ← **PanTompkins**(unfiltered EKG, sf)
detrend EKG ← **Detrend**(filtered EKG)
norm EKG ← **Normalise**(detrend EKG)
μ = mean peak distance(R peaks list)
while (train and test EKM < #EKM total) **do**
    Create EKM matrix(μ, R peaks list, norm EKG, Initial window, α_i, α_e)
    Standarise EKM$(-1, 1)$
    Generate EKM image
    Save EKM in train or test
    $Initial\,window += bpf$
end while

Once the EKG is read, some parameters are needed to be initialised. The Sampling Frequency (sf) is given by each database. For example, for the NSRDB will be 128 and 360 for the MIT-BIHDB. In the beginning, the *initial window* is initialised and will be updated after each EKM obtained as we need to move the beginning of that window the number of *bpf* that have been chosen as a hyperparameter ($bpf = (3, 5, 7)$). Hence, the next step is defining those hyperparameters such as the *bpf*. As it has been explained in Section 2.3 and 7.1.1, the values for α_i and α_e are the percentage of the signal that is going to be taken before and after each peak of a window to construct each segment of the EKM. Thus, it can be defined by the user who who will construct the EKM under some thresholds ($\alpha_i < 100\% > \alpha_e$). Also, the total number of EKMs has to be specified if we are coping with an extensive database as it is the NSRDB as well as the percentage of the subsets

[1]Physionet: `https://www.physionet.org`

for training and testing that are going to be obtained once the EKM datasets are constructed.

Two lists are needed to construct each EKM: i) a clean and filtered EKG signal and ii) a list of the R peaks of this EKG signal. Therefore, those lists are obtained through the Pan and Tompkins algorithm [75] as it has been specified in Section 2.2. Hence, the unfiltered EKG signal and the sampling frequency are given to this algorithm to obtain those two lists finally. Once the filtered EKG is obtained, the signal must be detrended to remove the linear trend along its axis and normalise. Before creating the EKM, we need to obtain the mean distance between R peaks for each signal. To do so, it is calculated with the list of the R peaks that were obtained before.

The while condition will be performed while the number of training and testing EKM images saved is less than the total number of EKMs (specified at the beginning). This way, we obtain images until a total threshold is obtained. If we had a shorter EKG recording, we would run through the whole signal until we ran out of it. So, once the window is defined by the number of *bpf* and the initial point of each window, an EKM matrix for this window can be created. To see how each EKM is made for each window, we have to look at Algorithm 2 as a function designed to construct each EKM for each window.

Hence, the requirements to create the EKM specified in 2 are the mean distance between peaks (μ), the list of the R peaks that were obtained before through the PanTompkins algorithm, the normalised EKG, the initial or starting point of the window, and the values for α_i and α_e. Each segment of the EKM (see Figure 7.1) is obtained for each window peak. Hence, the `for` condition will be repeated for each of the peaks of the window except for the last peak. Each segment will be obtained from the normalised EKM and defined by the α_i, α_e and μ values. To obtain the segment, we will set the starting point of that segment to a certain number of samples ($\alpha_i\mu$) before the current peak (p_x) and then, the segment will end a certain number of samples ($\alpha_e\mu$) after that same peak ((p_x)). This way, a segment containing the current peak and the peak after that will be constructed. We will store this segment and then concatenate all the segments from this window (i.e. six segments if $bpf = 7$) by aligning them to form the EKG matrix.

Once the matrix is obtained, we return to the main Algorithm 1. Before plotting the EKM, it will be standardised between -1 and 1 to have similar colours when the matrix is plotted. After that, the EKM can be plotted with the Seaborn library. As commented at the beginning of the algorithm, at the end of it, the `Initial window` needs to be actualised to move to the next window.

Hence, this process will be repeated for each window of the EKG signal

Algorithm 2 Creation of **one** EKM

Require: μ, R peaks list, norm EKG, Initial window, α_i, α_e
  for each peak (p_x) of the current window **do**
    segment = norm EKG$[p_x - \alpha_i\mu : p_x + \alpha_e\mu]$
    All segments $\leftarrow$ Append(segment)
    EKM = Concatenate(All segments, 1)
  end for
  return EKM

until we reach a certain threshold or run out of the signal. Again, this process has to be repeated for each user of the database to obtain an EKM dataset for the database chosen.

A.2 Classification with a CNN

In this thesis, the development of a tailored CNN has been needed to perform user classification over the EKM datasets. Other possibilities were faced regarding a CNN, as seen in Section 7.1.6 where it is discussed if it would be more efficient to use a tailored ad-hoc CNN or a pre-trained one. A tailored CNN with only one convolutional layer is the best approach as it offers more accurate results.

A.2.1 Objective Optimisation

How this CNN has been constructed is explained in Section 2.4 together with all hyperparameters of the network. This section will show an in-depth analysis of this chosen neural network. It is important to remark that the structure followed by Algorithms 3 and 4 has been followed for all the experiments performed in this doctoral thesis. The only differences that can be found are if the experimentation that is being performed is categorical or binary as follows:

Classification	**#classes**	**Loss Function**	**Activation**
Categorical	18, 47, 290, 162, 25	CCE	Softmax
Binary	2	BCE	Sigmoid

Table A.1: Comparison between a Categorical and Binary classification when training a CNN

As can be observed in Table A.1 there are some differences between a categorical and a binary classification when training a CNN. By its nature, a binary classification will only have two classes. Otherwise, a categorical classification can have many different classes; it will be defined by the number of users identified in each database. A loss function or cost function calculates the distance between the current and the expected output of the algorithm. It is a method which will evaluate how the algorithm models the data. The Categorical Cross-Entropy (CCE) loss function is used when dealing with a categorical classification. It is defined by:

$$\text{CCE} = -\frac{1}{N}\left(\sum_{u=1}^{N}\sum_{c=1}^{C}\left(t_{u,c}\log(p_{u,c})\right)\right) \tag{A.1}$$

where u is a specific user of the current database, N is the total number of users in our database, c is the current user we are testing of the C total users of the database, and $t_{u,c}$ is the true identifier of that user, and $p_{u,c}$ is the Softmax output probability estimation from our CNN.

Then, in the case of binary classification, we use the Binary Cross-Entropy (BCE) which is a specific case of the previous formula where $C = 2$, then, the formula is developed as follows:

$$\text{BCE} = -\frac{1}{N}\left(\sum_{u=1}^{N}\left(t_u\log(p_u)(1-t_u)\log(1-p_u)\right)\right) \tag{A.2}$$

where u is a specific user of the current database, N is the total number of users in our database, t_u is the true identifier of that user, and p_u is the Sigmoid output probability estimation from our CNN.

Ultimately, the BCE or the CCE, depending on the problem faced, is optimised to train the parameters of the network.

A.2.2 Code Implementation

To perform user classification, the code implementation seen in Algorithm 3 has to be followed. The first steps needed are always to define the paths in which the EKM datasets obtained in the previous steps (Section A.1) are stored and define the paths where the model, results and outputs will be saved. Then, a set of hyperparameters that have already been commented on in Section 2.4 are defined.

It is important to remark that the image size specified is not given by the image loaded to train the model but by the user, as it is a hyperparameter. When the EKMs are extracted from the EKG signal, the size of the generated image is $432x288$ pixels. In our experiments, we reduced the image size up to $54x29$ to reduce computational costs. Different experiments were launched to validate this parameter. This means that we reduced the size of the image as much as we could until the classification was affected. Consequently, the size of each image has been reduced eight times by width and almost ten times by height.

Once the parameters of the network have been specified, the train and validation dataset is obtained through an *ImageDataGenerator* to be able to feed the images into the CNN model. Next, a model is constructed by the addition of each of the layers forming this model.

The first layer is cropping the image to remove the axis and blank spaces; how many pixels can be removed from each frame can be specified. Once we have the clean EKM, the first and only convolutional layer takes place. The channel's output and kernel size must be set (we used a $3x3$ kernel). Then, the ReLu activation also takes place. After the convolution, the Max Pooling operation is defined with a pool size ($2x2$ in all of our experiments) to finally apply regularisation through a dropout defined to 70% of the parameters.

For the second layer of the CNN, a flatten is applied to construct a vector with all the parameters. Then, with a Dense layer and another ReLu activation, reduce the number of neurons. Finally, a Dense layer is again applied to have at the output the number of specified classes (depending on the database) together with a Softmax or Sigmoid activation (depending on the problem faced).

Once all layers are constructed, the last layer specifies how the model will be compiled by selecting the Loss function that will be used (BCE or CCE), an optimiser (Adam optimiser) and the metrics that will be optimised. The final step will be to store the trained model and results from this training and validation phase.

Once we have a properly trained model, we can test it with the testing dataset which is the subset of images that have not been seen in the training phase as done in Algorithm 4. Once the model is trained, the only steps that need to be performed are to define the test dataset, load the model and execute the evaluation with the imported images and model. Once the evaluation is done, all metrics, results and plots can be obtained from this evaluation.

Algorithm 3 Training the CNN

Define paths: EKM train dataset, results, output
Define hyperparameters: #epochs, #classes, batch size, validation percentage, image size, kernel size, pool size, dropout
train dataset, test dataset ← ImageDataGenerator(EKM train dataset, image size, validation percentage)
for each layer in model **do**
  Add(**Cropping**((top,bottom),(left,right), input shape))
  Add(**Conv2D**(channel output, kernel size, activation))
  Add(**MaxPooling2D**(pool size))
  Add(**Dropout**(dropout))
  Add(**Flatten**())
  Add(**Dense**(neuron numbers, activation))
  Add(**Dense**(#classes, activation))
  Compile(CE, optimiser, metrics)
end for
trained model ← **Fit**(train and validation datasets, model, batch size, epochs)
Save model and results

Algorithm 4 User Classification - testing

Define paths: EKM test dataset, trained model
Define hyperparameters: image size
test dataset ← ImageDataGenerator(EKM test dataset, image size)
model ← **Load**(trained model)
FP, TP, FN, TN ← Model **Evaluation**(model, test dataset)

BIBLIOGRAPHY

[1] "Ecg lead systems." `https://www.cardiosecur.com/magazine/specialist-articles-on-the-heart/lead-systems-how-an-ecg-works`. Accessed: 2022-08-17.

[2] "Ecg lead positioning." `https://litfl.com/ecg-lead-positioning/`. Accessed: 2022-08-17.

[3] "Far and frr: security level versus user convenience." `https://www.recogtech.com/en/knowledge-base/security-level-versus-user-convenience`. Accessed: 2022-08-17.

[4] M. Taskiran, N. Kahraman, and C. E. Erdem, "Face recognition: Past, present and future (a review)," *Digital Signal Processing*, vol. 106, p. 102809, 2020.

[5] N. Senthilkumar *et al.*, "Comprehensive review of fingerprint based biometric systems," *Journal of Critical Reviews*, vol. 7, no. 4, 2020.

[6] N. Singla, M. Kaur, and S. Sofat, "Automated latent fingerprint identification system: A review," *Forensic Science International*, vol. 309, p. 110187, 2020.

[7] J. M. Ackerson, R. Dave, and N. Seliya, "Applications of recurrent neural network for biometric authentication & anomaly detection," *Information*, vol. 12, no. 7, 2021.

[8] R. Salih Kuzu, E. Maiorana, and P. Campisi, "Loss functions for cnn-based biometric vein recognition," in *2020 28th European Signal Processing Conference (EUSIPCO)*, pp. 750–754, 2021.

[9] C. Q. Lai, H. Ibrahim, M. Z. Abdullah, and S. A. Suandi, "Eeg-based biometric close-set identification using cnn-ecoc-svm," in *International Conference on Artificial Intelligence and Big Data in Digital Era*, pp. 723–732, Springer, 2022.

[10] D. Jyotishi and S. Dandapat, "An ecg biometric system using hierarchical lstm with attention mechanism," *IEEE Sensors Journal*, vol. 22, no. 6, pp. 6052–6061, 2022.

[11] T.-W. Bae, K.-K. Kwon, and K.-H. Kim, "Electrocardiogram fiducial point detector using a bilateral filter and symmetrical point-filter structure," *International Journal of Environmental Research and Public Health*, vol. 18, no. 20, p. 10792, 2021.

[12] M. Ingale, R. Cordeiro, S. Thentu, Y. Park, and N. Karimian, "Ecg biometric authentication: A comparative analysis," *IEEE Access*, vol. 8, pp. 117853–117866, 2020.

[13] C. Fuster-Barceló, P. Peris-Lopez, and C. Camara, "Elektra: Elektrokardiomatrix application to biometric identification with convolutional neural networks," *Neurocomputing*, vol. 506, pp. 37–49, 2022.

[14] S. Liu and M. Silverman, "A practical guide to biometric security technology," *IT Professional*, vol. 3, no. 1, pp. 27–32, 2001.

[15] A. Albrizio, "Biometry and anthropometry: from galton to constitutional medicine," *Journal of Anthropological Sciences*, vol. 85, pp. 101–123, 2007.

[16] S. Prabhakar, S. Pankanti, and A. Jain, "Biometric recognition: security and privacy concerns," *IEEE Security & Privacy*, vol. 1, no. 2, pp. 33–42, 2003.

[17] A. Eng and L. A. Wahsheh, "Look into my eyes: A survey of biometric security," in *2013 10th International Conference on Information Technology: New Generations*, pp. 422–427, 2013.

[18] A. N. Uwaechia and D. A. Ramli, "A comprehensive survey on ecg signals as new biometric modality for human authentication: Recent advances and future challenges," *IEEE Access*, vol. 9, pp. 97760–97802, 2021.

[19] M. S. Obaidat, I. Traore, and I. Woungang, *Biometric-based physical and cybersecurity systems*. Springer, 2019.

[20] Y. N. Singh, S. K. Singh, and A. K. Ray, "Bioelectrical signals as emerging biometrics: Issues and challenges," *International Scholarly Research Notices*, vol. 2012, 2012.

[21] L. Biel, O. Pettersson, L. Philipson, and P. Wide, "Ecg analysis: a new approach in human identification," *IEEE Transactions on Instrumentation and Measurement*, vol. 50, no. 3, pp. 808–812, 2001.

[22] K.-S. Kim, T.-H. Yoon, J.-W. Lee, D.-J. Kim, and H.-S. Koo, "A robust human identification by normalized time-domain features of electrocardiogram," in *2005 IEEE Engineering in Medicine and Biology 27th Annual Conference*, pp. 1114–1117, 2005.

[23] R. Salloum and C.-C. J. Kuo, "Ecg-based biometrics using recurrent neural networks," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2062–2066, 2017.

[24] Q. Zhang, D. Zhou, and X. Zeng, "Heartid: A multiresolution convolutional neural network for ecg-based biometric human identification in smart health applications," *IEEE Access*, vol. 5, pp. 11805–11816, 2017.

[25] B. Pourbabaee, M. Howe-Patterson, E. Reiher, and F. Benard, "Deep convolutional neural network for ecg-based human identification," *CMBES Proceedings*, vol. 41, May 2018.

[26] E. J. da Silva Luz, G. J. P. Moreira, L. S. Oliveira, W. R. Schwartz, and D. Menotti, "Learning deep off-the-person heart biometrics representations," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 5, pp. 1258–1270, 2018.

[27] M. Al Rousan and B. Intrigila, "A Comparative Analysis of Biometrics Types: Literature Review," *Journal of Computer Science*, vol. 16, no. 12, pp. 1778–1788, 2020.

[28] E. Alajrami, B. A. M. Ashqar, B. S. Abu-Nasser, A. J. Khalil, M. M. Musleh, A. M. Barhoom, and S. S. Abu-Naser, "Handwritten signature verification using deep learning," *International Journal of Academic Multidisciplinary Research (IJAMR)*, vol. 3, no. 12, pp. 39–44, 2020.

[29] K. Bibi, S. Naz, and A. Rehman, "Biometric signature authentication using machine learning techniques: Current trends, challenges and opportunities," *Multimedia Tools and Applications*, vol. 79, no. 1, pp. 289–340, 2020.

[30] A. N. Azmi, D. Nasien, and F. S. Omar, "Biometric signature verification system based on freeman chain code and k-nearest neighbor," *Multimedia Tools and Applications*, vol. 76, no. 14, pp. 15341–15355, 2017.

[31] K. Aizat, O. Mohamed, M. Orken, A. Ainur, and B. Zhumazhanov, "Identification and authentication of user voice using dnn features and i-vector," *Cogent Engineering*, vol. 7, no. 1, p. 1751557, 2020.

[32] H. N. M. Shah, M. Z. Ab Rashid, M. F. Abdollah, M. N. Kamarudin, C. K. Lin, and Z. Kamis, "Biometric voice recognition in security system," *Indian journal of Science and Technology*, vol. 7, no. 2, p. 104, 2014.

[33] N. Singh, A. Agrawal, and R. Khan, "Voice biometric: A technology for voice based authentication," *Advanced Science, Engineering and Medicine*, vol. 10, no. 7-8, pp. 754–759, 2018.

[34] B. Bhana and S. Flowerday, "Passphrase and keystroke dynamics authentication: Usable security," *Computers & Security*, vol. 96, p. 101925, 2020.

[35] N. Raul, R. Shankarmani, and P. Joshi, "A comprehensive review of keystroke dynamics-based authentication mechanism," in *International Conference on Innovative Computing and Communications*, pp. 149–162, Springer, 2020.

[36] E. Ellavarason, R. Guest, F. Deravi, R. Sanchez-Riello, and B. Corsetti, "Touch-dynamics based behavioural biometrics on mobile devices – a review from a usability and performance perspective," *ACM Comput. Surv.*, vol. 53, Dec. 2020.

[37] E. Ellavarason, R. Guest, F. Deravi, R. Sanchez-Riello, and B. Corsetti, "Touch-dynamics based behavioural biometrics on mobile devices–a review from a usability and performance perspective," *ACM Computing Surveys (CSUR)*, vol. 53, no. 6, pp. 1–36, 2020.

[38] S. Adamović, V. Miškovic, N. Maček, M. Milosavljević, M. Šarac, M. Saračević, and M. Gnjatović, "An efficient novel approach for iris recognition based on stylometric features and machine learning techniques," *Future Generation Computer Systems*, vol. 107, pp. 144–157, 2020.

[39] M. Dua, R. Gupta, M. Khari, and R. G. Crespo, "Biometric iris recognition using radial basis function neural network," *Soft Computing*, vol. 23, no. 22, pp. 11801–11815, 2019.

[40] A. S. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi, and T. A. Nagem, "A multi-biometric iris recognition system based on a deep learning approach," *Pattern Analysis and Applications*, vol. 21, no. 3, pp. 783–802, 2018.

[41] P. Kaur, K. Krishan, S. K. Sharma, and T. Kanchan, "Facial-recognition algorithms: A literature review," *Medicine, Science and the Law*, vol. 60, no. 2, pp. 131–139, 2020. PMID: 31964224.

[42] J. Hernandez-Ortega, J. Galbally, J. Fiérrez, and L. Beslay, "Biometric quality: Review and application to face recognition with faceqnet," *arXiv preprint arXiv:2006.03298*, 2020.

[43] M. Zulfiqar, F. Syed, M. J. Khan, and K. Khurshid, "Deep face recognition for biometric authentication," in *2019 international conference on electrical, communication, and computer engineering (ICECCE)*, pp. 1–6, IEEE, 2019.

[44] L. T. Nguyen, H. T. Nguyen, A. D. Afanasiev, and T. V. Nguyen, "Automatic Identification Fingerprint Based on Machine Learning Method," *Journal of the Operations Research Society of China*, 2021.

[45] W. Yang, S. Wang, J. Hu, G. Zheng, and C. Valli, "Security and accuracy of fingerprint-based biometrics: A review," *Symmetry*, vol. 11, no. 2, p. 141, 2019.

[46] D. P. Wagh, H. S. Fadewar, and G. N. Shinde, "Biometric finger vein recognition methods for authentication," in *Computing in Engineering and Technology* (B. Iyer, P. S. Deshpande, S. C. Sharma, and U. Shiurkar, eds.), (Singapore), pp. 45–53, Springer Singapore, 2020.

[47] K. Shaheed, A. Mao, I. Qureshi, M. Kumar, S. Hussain, and X. Zhang, "Recent advancements in finger vein recognition technology: methodology, challenges and opportunities," *Information Fusion*, vol. 79, pp. 84–109, 2022.

[48] L. Olanrewaju, O. Oyebiyi, S. Misra, R. Maskeliunas, and R. Damasevicius, "Secure ear biometrics using circular kernel principal component analysis, Chebyshev transform hashing and Bose–Chaudhuri–Hocquenghem error-correcting codes," *Signal, Image and Video Processing*, vol. 14, no. 5, pp. 847–855, 2020.

[49] A. Kamboj, R. Rani, and A. Nigam, "A comprehensive survey and deep learning-based approach for human recognition using ear biometric," *The Visual Computer*, pp. 1–34, 2021.

[50] R. Kushwaha and N. Nain, "Person identification using footprint minutiae," in *Proceedings of 3rd International Conference on Computer Vision and Image Processing* (B. B. Chaudhuri, M. Nakagawa, P. Khanna, and S. Kumar, eds.), (Singapore), pp. 285–299, Springer Singapore, 2020.

[51] R. Khokher and R. C. Singh, "Footprint identification: Review of an emerging biometric trait," in *Macromolecular Symposia*, vol. 397, p. 2000246, Wiley Online Library, 2021.

[52] F. L. Podio, J. S. Dunn, *et al.*, "Biometric authentication technology: From the movies to your desktop," *ITL Bulletin, May*, pp. 1–8, 2001.

[53] V. Matyas and Z. Riha, "Toward reliable user authentication through biometrics," *IEEE Security & Privacy*, vol. 1, no. 3, pp. 45–49, 2003.

[54] W. Fye, "A history of the origin, evolution, and impact of electrocardiography," *The American Journal of Cardiology*, vol. 73, no. 13, pp. 937–949, 1994.

[55] A. L. Goldberger and E. Gold-berger, "Clinical electrocardiography, a simplified approach," *Critical Care Medicine*, vol. 9, no. 12, pp. 891–892, 1981.

[56] T. Shen, W. Tompkins, and Y. Hu, "One-lead ecg for identity verification," in *Proceedings of the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] [Engineering in Medicine and Biology*, vol. 1, pp. 62–63 vol.1, 2002.

[57] S. Khan, S. Parkinson, L. Grant, N. Liu, and S. Mcguire, "Biometric systems utilising health data from wearable devices: applications and future challenges in computer security," *ACM Computing Surveys (CSUR)*, vol. 53, no. 4, pp. 1–29, 2020.

[58] K. Nguyen, C. Fookes, R. Jillela, S. Sridharan, and A. Ross, "Long range iris recognition: A survey," *Pattern Recognition*, vol. 72, pp. 123–143, 2017.

[59] V. Conti, L. Rundo, C. Militello, V. M. Salerno, S. Vitabile, and S. M. Siniscalchi, "A multimodal retina-iris biometric system using the levenshtein distance for spatial feature comparison," *IET Biometrics*, vol. 10, no. 1, pp. 44–64, 2021.

[60] I. Odinaka, P. Lai, A. D. Kaplan, J. A. O'Sullivan, E. J. Sirevaag, S. D. Kristjansson, A. K. Sheffield, and J. W. Rohrbaugh, "Ecg biometrics: A robust short-time frequency analysis," in *2010 IEEE International Workshop on Information Forensics and Security*, pp. 1–6, 2010.

[61] A. Pal and Y. N. Singh, "Ecg biometric recognition," in *International Conference on Mathematics and Computing*, pp. 61–73, Springer, 2018.

[62] K. N. Plataniotis, D. Hatzinakos, and J. K. M. Lee, "Ecg biometric recognition without fiducial detection," in *2006 Biometrics Symposium: Special Session on Research at the Biometric Consortium Conference*, pp. 1–6, 2006.

[63] K. A. Sidek, I. Khalil, and H. F. Jelinek, "Ecg biometric with abnormal cardiac conditions in remote monitoring system," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 44, no. 11, pp. 1498–1509, 2014.

[64] L. González-Manzano, J. M. de Fuentes, P. Peris-Lopez, and C. Camara, "Encryption by heart (ebh)—using ecg for time-invariant symmetric key generation," *Future Generation Computer Systems*, vol. 77, pp. 136–148, 2017.

[65] C. Camara, P. Peris-Lopez, H. Martín, and M. Aldalaien, "Ecg-rng: A random number generator based on ecg signals and suitable for securing wireless sensor networks," *Sensors*, vol. 18, no. 9, p. 2747, 2018.

[66] J. Malik, D. Girdhar, R. Dahiya, and G. Sainarayanan, "Reference threshold calculation for biometric authentication," *IJ Image, Graphics and Signal Processing*, vol. 2, pp. 46–53, 2014.

[67] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, "Person re-identification by symmetry-driven accumulation of local features," in *2010 IEEE computer society conference on computer vision and pattern recognition*, pp. 2360–2367, IEEE, 2010.

[68] D. O. Gorodnichy and R. Hoshino, "Calibrated confidence scoring for biometric identification," in *NIST International Biometric Performance Conference (IBPC 2010)*, pp. 2–4, 2010.

[69] S. A. Israel, J. M. Irvine, A. Cheng, M. D. Wiederhold, and B. K. Wiederhold, "Ecg to identify individuals," *Pattern Recognition*, vol. 38, no. 1, pp. 133–142, 2005.

[70] M. O. Diab, A. Seif, M. Sabbah, M. El-Abed, and N. Aloulou, *A Review on ECG-Based Biometric Authentication Systems*, pp. 17–44. Singapore: Springer Singapore, 2020.

[71] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000 (June 13). Circulation Electronic Pages: http://circ.ahajournals.org/content/101/23/e215.full PMID:1085218; doi: 10.1161/01.CIR.101.23.e215.

[72] M. M. Bassiouni, E.-S. A. El-Dahshan, W. Khalefa, and A. M. Salem, "Intelligent hybrid approaches for human ecg signals identification," *Signal, Image and Video Processing*, vol. 12, no. 5, pp. 941–949, 2018.

[73] M. A. Awal, S. S. Mostafa, M. Ahmad, and M. A. Rashid, "An adaptive level dependent wavelet thresholding for ecg denoising," *Biocybernetics and biomedical engineering*, vol. 34, no. 4, pp. 238–249, 2014.

[74] F. Gargiulo, A. Fratini, M. Sansone, and C. Sansone, "Subject identification via ecg fiducial-based systems: Influence of the type of qt interval correction," *Computer Methods and Programs in Biomedicine*, vol. 121, no. 3, pp. 127–136, 2015.

[75] J. Pan and W. J. Tompkins, "A real-time qrs detection algorithm," *IEEE Transactions on Biomedical Engineering*, vol. BME-32, no. 3, pp. 230–236, 1985.

[76] S. A. El_Rahman, "Biometric human recognition system based on ecg," *Multimedia Tools and Applications*, vol. 78, no. 13, pp. 17555–17572, 2019.

[77] M. Hammad, G. Luo, and K. Wang, "Cancelable biometric authentication system based on ecg," *Multimedia Tools and Applications*, vol. 78, no. 2, pp. 1857–1887, 2019.

[78] H. M. Rai, A. Trivedi, and S. Shukla, "Ecg signal processing for abnormalities detection using multi-resolution wavelet transform and artificial neural network classifier," *Measurement*, vol. 46, no. 9, pp. 3238–3246, 2013.

[79] M. J. Rooijakkers, C. Rabotti, S. G. Oei, and M. Mischi, "Low-complexity r-peak detection for ambulatory fetal monitoring," *Physiological measurement*, vol. 33, no. 7, p. 1135, 2012.

[80] D. Alduwaile and M. S. Islam, "Single heartbeat ecg biometric recognition using convolutional neural network," in *2020 International Conference on Advanced Science and Engineering (ICOASE)*, pp. 145–150, 2020.

[81] J. Rodrigues, D. Belo, and H. Gamboa, "Noise detection on ecg based on agglomerative clustering of morphological features," *Computers in biology and medicine*, vol. 87, pp. 322–334, 2017.

[82] W. Yang, Y. Si, D. Wang, and G. Zhang, "A novel method for identifying electrocardiograms using an independent component analysis and principal component analysis network," *Measurement*, vol. 152, p. 107363, 2020.

[83] D. Belo, N. Bento, H. Silva, A. Fred, and H. Gamboa, "Ecg biometrics using deep learning and relative score threshold classification," *Sensors*, vol. 20, no. 15, p. 4078, 2020.

[84] B. Widrow, J. R. Glover, J. M. McCool, J. Kaunitz, C. S. Williams, R. H. Hearn, J. R. Zeidler, J. E. Dong, and R. C. Goodlin, "Adaptive noise cancelling: Principles and applications," *Proceedings of the IEEE*, vol. 63, no. 12, pp. 1692–1716, 1975.

[85] R. K. Mahendran and P. Velusamy, "A secure fuzzy extractor based biometric key authentication scheme for body sensor network in internet of medical things," *Computer Communications*, vol. 153, pp. 545–552, 2020.

[86] J. S. Paul, M. Reddy, and V. J. Kumar, "Data processing of stress ecgs using discrete cosine transform," *Computers in biology and medicine*, vol. 28, no. 6, pp. 639–658, 1998.

[87] F. Liu, Y. Si, T. Luo, D. Wang, Q. Liu, and L. Li, "The ecg identification based on grnn," in *2018 IEEE International Conference on Communication Systems (ICCS)*, pp. 418–422, IEEE, 2018.

[88] M. Komeili, W. Louis, N. Armanfard, and D. Hatzinakos, "Feature selection for nonstationary data: Application to human recognition using medical biometrics," *IEEE transactions on cybernetics*, vol. 48, no. 5, pp. 1446–1459, 2017.

[89] E. K. Zaghouani, A. Benzina, and R. Attia, "Ecg based authentication for e-healthcare systems: Towards a secured ecg features transmission," in *2017 13th international wireless communications and mobile computing conference (IWCMC)*, pp. 1777–1783, IEEE, 2017.

[90] H. Kim and S. Y. Chun, "Cancelable ecg biometrics using compressive sensing-generalized likelihood ratio test," *IEEE Access*, vol. 7, pp. 9232–9242, 2019.

[91] E. Ercelebi, "Electrocardiogram signals de-noising using lifting-based discrete wavelet transform," *Computers in Biology and Medicine*, vol. 34, no. 6, pp. 479–493, 2004.

[92] R. Tan and M. Perkowski, "Toward improving electrocardiogram (ecg) biometric verification using mobile sensors: A two-stage classifier approach," *Sensors*, vol. 17, no. 2, p. 410, 2017.

[93] R. Donida Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ecg: Convolutional neural networks for ecg biometric recognition," *Pattern Recognition Letters*, vol. 126, pp. 78–85, 2019. Robustness, Security and Regulation Aspects in Current Biometric Systems.

[94] Y. Chu, H. Shen, and K. Huang, "Ecg authentication method based on parallel multi-scale one-dimensional residual network with center and margin loss," *IEEE Access*, vol. 7, pp. 51598–51607, 2019.

[95] H. M. Lynn, S. B. Pan, and P. Kim, "A deep bidirectional gru network model for biometric electrocardiogram classification based on recurrent neural networks," *IEEE Access*, vol. 7, pp. 145395–145405, 2019.

[96] M. Hejazi, S. Al-Haddad, Y. P. Singh, S. J. Hashim, and A. F. Abdul Aziz, "Ecg biometric authentication based on non-fiducial approach using kernel methods," *Digital Signal Processing*, vol. 52, pp. 72–86, 2016.

[97] D. Wang, Y. Si, W. Yang, G. Zhang, and T. Liu, "A novel heart rate robust method for short-term electrocardiogram biometric identification," *Applied Sciences*, vol. 9, no. 1, p. 201, 2019.

[98] C. Camara, P. Peris-Lopez, L. Gonzalez-Manzano, and J. Tapiador, "Real-time electrocardiogram streams for continuous authentication," *Applied Soft Computing*, vol. 68, pp. 784–794, 2018.

[99] J. Liu, L. Yin, C. He, B. Wen, X. Hong, and Y. Li, "A multiscale autoregressive model-based electrocardiogram identification method," *IEEE Access*, vol. 6, pp. 18251–18263, 2018.

[100] A. Y. Shdefat, M.-I. Joo, S.-H. Choi, and K. Hee-Cheol, "Utilizing ecg waveform features as new biometric authentication method," *International Journal of Electrical and Computer Engineering*, vol. 8, no. 2, p. 658, 2018.

[101] V. Krasteva, I. Jekova, and R. Schmid, "Perspectives of human verification via binary qrs template matching of single-lead and 12-lead electrocardiogram," *PloS one*, vol. 13, no. 5, p. e0197240, 2018.

[102] S. Aziz, M. U. Khan, Z. A. Choudhry, A. Aymin, and A. Usman, "Ecg-based biometric authentication using empirical mode decomposition and support vector machines," in *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)*, pp. 0906–0912, IEEE, 2019.

[103] M. U. Khan, S. Aziz, K. Iqtidar, A. Saud, and Z. Azhar, "Biometric authentication system based on electrocardiogram (ecg)," in *2019 13th International Conference on Mathematics, Actuarial Science, Computer Science and Statistics (MACS)*, pp. 1–6, IEEE, 2019.

[104] R. Tan and M. Perkowski, "Ecg biometric identification using wavelet analysis coupled with probabilistic random forest," in *2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA)*, pp. 182–187, 2016.

[105] T. N. Alotaiby, S. R. Alrshoud, S. A. Alshebeili, and L. M. Aljafar, "Ecg-based subject identification using statistical features and random forest," *Journal of Sensors*, vol. 2019, 2019.

[106] M. N. Sahadat, E. L. Jacobs, and B. I. Morshed, "Hardware-efficient robust biometric identification from 0.58 second template and 12 features of limb (lead i) ecg signal using logistic regression classifier," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1440–1443, 2014.

[107] V. Chandrashekhar, P. Singh, M. Paralkar, and O. K. Tonguz, "Pulse id: the case for robustness of ecg as a biometric identifier," in *2020 IEEE 30th International Workshop on Machine Learning for Signal Processing (MLSP)*, pp. 1–6, IEEE, 2020.

[108] S. A. A. Yusuf and R. Hidayat, "Mfcc feature extraction and knn classification in ecg signals," in *2019 6th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE)*, pp. 1–5, IEEE, 2019.

[109] G. Wübbeler, M. Stavridis, D. Kreiseler, R.-D. Bousseljot, and C. Elster, "Verification of humans using the electrocardiogram," *Pattern Recognition Letters*, vol. 28, no. 10, pp. 1172–1175, 2007.

[110] M. Pelc, Y. Khoma, and V. Khoma, "Ecg signal as robust and reliable biometric marker: Datasets and algorithms comparison," *Sensors*, vol. 19, no. 10, 2019.

[111] Y. Li, Y. Pang, K. Wang, and X. Li, "Toward improving ecg biometric identification using cascaded convolutional neural networks," *Neurocomputing*, vol. 391, pp. 83 – 95, 2020.

[112] N. Bento, D. Belo, and H. Gamboa, "Ecg biometrics using spectrograms and deep neural networks," *Int. J. Mach. Learn. Comput*, vol. 10, no. 2, pp. 259–264, 2019.

[113] J. R. Pinto and J. S. Cardoso, "An end-to-end convolutional neural network for ecg-based biometric authentication," in *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pp. 1–8, 2019.

[114] J. R. Pinto, J. S. Cardoso, and A. Lourenço, "Deep neural networks for biometric identification based on non-intrusive ecg acquisitions," in *The Biometric Computing*, pp. 217–234, Chapman and Hall/CRC, 2019.

[115] B.-H. Kim and J.-Y. Pyun, "Ecg identification for personal authentication using lstm-based deep recurrent neural networks," *Sensors*, vol. 20, no. 11, 2020.

[116] Z. Ebrahimi, M. Loni, M. Daneshtalab, and A. Gharehbaghi, "A review on deep learning methods for ecg arrhythmia classification," *Expert Systems with Applications: X*, vol. 7, p. 100033, 2020.

[117] M. Hammad, S. Zhang, and K. Wang, "A novel two-dimensional ecg feature extraction and classification algorithm based on convolution neural network for human authentication," *Future Generation Computer Systems*, vol. 101, pp. 180–196, 2019.

[118] W.-H. Jung and S.-G. Lee, "Ecg identification based on non-fiducial feature extraction using window removal method," *Applied Sciences*, vol. 7, p. 1205, Nov 2017.

[119] R. He, K. Wang, N. Zhao, Y. Liu, Y. Yuan, Q. Li, and H. Zhang, "Automatic detection of atrial fibrillation based on continuous wavelet transform and 2d convolutional neural networks," *Frontiers in Physiology*, vol. 9, 2018.

[120] A. Ullah, S. u. Rehman, S. Tu, R. M. Mehmood, and M. Ehatisham-Ul-Haq, "A hybrid deep cnn model for abnormal arrhythmia detection based on cardiac ecg signal," *Sensors*, vol. 21, no. 3, p. 951, 2021.

[121] M. S. Haleem, R. Castaldo, S. M. Pagliara, M. Petretta, M. Salvatore, M. Franzese, and L. Pecchia, "Time adaptive ecg driven cardiovascular disease detector," *Biomedical Signal Processing and Control*, vol. 70, p. 102968, 2021.

[122] J. D. Pollard, K. T. Haq, K. J. Lutz, N. M. Rogovoy, K. A. Paternostro, E. Z. Soliman, J. Maher, J. A. Lima, S. K. Musani, and L. G. Tereshchenko, "Electrocardiogram machine learning for detection of cardiovascular disease in african americans: the jackson heart study," *European Heart Journal-Digital Health*, vol. 2, no. 1, pp. 137–151, 2021.

[123] T. Wang, C. Lu, Y. Sun, M. Yang, C. Liu, and C. Ou, "Automatic ecg classification using continuous wavelet transform and convolutional neural network," *Entropy*, vol. 23, no. 1, p. 119, 2021.

[124] S. P. Jillella, C. Rohith, S. Shameem, and P. S. S. Babu, "Ecg classification for arrhythmias using cnn & heart disease prediction using web application," in *2022 First International Conference on Electrical, Electronics, Information and Communication Technologies (ICEEICT)*, pp. 1–8, IEEE, 2022.

[125] V. Lee, G. Xu, V. Liu, P. Farrehi, and J. Borjigin, "Accurate detection of atrial fibrillation and atrial flutter using the electrocardiomatrix technique," *Journal of Electrocardiology*, vol. 51, no. 6, Supplement, pp. S121–S125, 2018.

[126] D. Li, F. Tian, S. Rengifo, G. Xu, M. M. Wang, and J. Borjigin, "Electrocardiomatrix: A new method for beat-by-beat visualization and inspection of cardiac signals," *J Integr Cardiol*, vol. 1, no. 5, pp. 124–128, 2015.

[127] G. Xu, S. Dodaballapur, T. Mihaylova, and J. Borjigin, "Electrocardiomatrix facilitates qualitative identification of diminished heart rate variability in critically ill patients shortly before cardiac arrest," *Journal of Electrocardiology*, vol. 51, no. 6, pp. 955–961, 2018.

[128] K. Sharma, M. Rao, P. Marwaha, and A. Kumar, "Accurate detection of congestive heart failure using electrocardiomatrix technique," *Multimedia Tools and Applications*, 04 2022.

[129] P.-L. Hong, J.-Y. Hsiao, C.-H. Chung, Y.-M. Feng, and S.-C. Wu, "Ecg biometric recognition: Template-free approaches based on deep learning," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2633–2636, 2019.

[130] A. Barros, P. Resque, J. Almeida, R. Mota, H. Oliveira, D. Rosário, and E. Cerqueira, "Data improvement model based on ecg biometric for user authentication and identification," *Sensors*, vol. 20, no. 10, p. 2920, 2020.

[131] M. M. Ghassemi, B. E. Moody, L.-W. H. Lehman, C. Song, Q. Li, H. Sun, R. G. Mark, M. B. Westover, and G. D. Clifford, "You snooze, you win: the physionet/computing in cardiology challenge 2018," in *2018 Computing in Cardiology Conference (CinC)*, vol. 45, pp. 1–4, 2018.

[132] Y. Zhang, Z. Zhao, Y. Deng, X. Zhang, and Y. Zhang, "Human identification driven by deep cnn and transfer learning based on multiview feature representations of ecg," *Biomedical Signal Processing and Control*, vol. 68, p. 102689, 2021.

[133] Z. Zhao, Y. Zhang, Y. Deng, and X. Zhang, "Ecg authentication system design incorporating a convolutional neural network and generalized s-transformation," *Computers in Biology and Medicine*, vol. 102, pp. 168–179, 2018.

[134] A. Goldberg, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, March 2014.

[135] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[136] G. Altan, Y. Kutlu, and M. Yeniad, "Ecg based human identification using second order difference plots," *Computer Methods and Programs in Biomedicine*, vol. 170, pp. 81–93, 2019.

[137] M. Hammad, Y. Liu, and K. Wang, "Multimodal biometric authentication systems using convolution neural network based on different level fusion of ecg and fingerprint," *IEEE Access*, vol. 7, pp. 26527–26542, 2019.

[138] C. Camara, P. Peris-Lopez, J. E. Tapiador, and G. Suarez-Tangil, "Non-invasive multi-modal human identification system combining ecg, gsr, and airflow biosignals," *Journal of Medical and Biological Engineering*, vol. 35, no. 6, pp. 735–748, 2015.

[139] M. Hammad, P. Pławiak, K. Wang, and U. R. Acharya, "Resnet-attention model for human authentication using ecg signals," *Expert Systems*, vol. 38, no. 6, p. e12547, 2021.

[140] D. L. Brown, G. Xu, A. M. B. Krzyske, N. C. Buhay, M. Blaha, M. M. Wang, P. Farrehi, and J. Borjigin, "Electrocardiomatrix facilitates accurate detection of atrial fibrillation in stroke patients," *Stroke*, vol. 50, no. 7, pp. 1676–1681, 2019.

[141] R. Salinas-Martínez, J. De Bie, N. Marzocchi, and F. Sandberg, "Detection of brief episodes of atrial fibrillation based on electrocardiomatrix and convolutional neural network," *Frontiers in physiology*, vol. 12, 2021.

[142] A. Goldberg, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals. circulation [online]. 101 (23), pp. e215–e220," Aug 1999.

[143] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der ekg-signaldatenbank cardiodat der ptb über das internet," *Biomedizinische Technik / Biomedical Engineering*, vol. 40, pp. 317–318, 1 1995.

[144] L. Howell and B. Porr, "High precision ecg database with annotated r peaks, recorded and filmed under realistic conditions," 2018.

[145] R. Salinas-Martínez, J. De Bie, N. Marzocchi, and F. Sandberg, "Automatic detection of atrial fibrillation using electrocardiomatrix and convolutional neural network," in *2020 Computing in Cardiology*, pp. 1–4, 2020.

[146] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, pp. 1097–1105, 2012.

[147] Y. LeCun, B. E. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. E. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," in *Advances in neural information processing systems*, pp. 396–404, 1990.

[148] A. Krizhevsky and G. Hinton, "Convolutional deep belief networks on cifar-10," *Unpublished manuscript*, vol. 40, no. 7, pp. 1–9, 2010.

[149] D. Varshni, K. Thakral, L. Agarwal, R. Nijhawan, and A. Mittal, "Pneumonia detection using cnn based feature extraction," in *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pp. 1–7, 2019.

[150] H. Sharma, J. S. Jain, P. Bansal, and S. Gupta, "Feature extraction and classification of chest x-ray images using cnn to detect pneumonia," in *2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 227–231, 2020.

[151] I. Rizwan I Haque and J. Neubert, "Deep learning approaches to biomedical image segmentation," *Informatics in Medicine Unlocked*, vol. 18, p. 100297, 2020.

[152] L. D. Nguyen, R. Gao, D. Lin, and Z. Lin, "Biomedical image classification based on a feature concatenation and ensemble of deep cnns," *Journal of Ambient Intelligence and Humanized Computing*, pp. 1–13, 2019.

[153] K. Fukushima and S. Miyake, "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition," in *Competition and cooperation in neural nets*, pp. 267–285, Springer, 1982.

[154] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[155] L. Deng, "The mnist database of handwritten digit images for machine learning research," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 141–142, 2012.

[156] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[157] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, 2016.

[158] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 2818–2826, 2016.

[159] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, pp. 6105–6114, PMLR, 2019.

[160] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[161] 2018.

[162] R. Srivastva, A. Singh, and Y. N. Singh, "Plexnet: A fast and robust ecg biometric system for human recognition," *Information Sciences*, vol. 558, pp. 208–228, 2021.

[163] A. Nemcova, R. Smisek, K. Opravilová, M. Vitek, L. Smital, and L. Maršánová, "Brno university of technology ecg quality database (but qdb)," *PhysioNet*, 2020.

[164] V. Ponciano, I. M. Pires, F. R. Ribeiro, and N. M. Garcia, "Data acquisition of timed-up and go test with older adults: accelerometer, magnetometer, electrocardiography and electroencephalography sensors' data," *Data in Brief*, vol. 32, p. 106306, 2020.

[165] I. Monedero, "A novel ecg diagnostic system for the detection of 13 different diseases," *Engineering Applications of Artificial Intelligence*, vol. 107, p. 104536, 2022.

[166] E. D. Übeyli, "Support vector machines for detection of electrocardiographic changes in partial epileptic patients," *Engineering Applications of Artificial Intelligence*, vol. 21, no. 8, pp. 1196–1203, 2008.

[167] A. Barros, P. Resque, J. Almeida, R. Mota, H. Oliveira, D. Rosário, and E. Cerqueira, "Data improvement model based on ecg biometric for user authentication and identification," *Sensors*, vol. 20, no. 10, 2020.

[168] J. W. Lindeberg, "Eine neue herleitung des exponentialgesetzes in der wahrscheinlichkeitsrechnung," *Mathematische Zeitschrift*, vol. 15, no. 1, pp. 211–225, 1922.

[169] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition. corr abs/1512.03385 (2015)," 2015.

[170] S. W. Chen, S. L. Wang, X. Z. Qi, S. M. Samuri, and C. Yang, "Review of ecg detection and classification based on deep learning: Coherent taxonomy, motivation, open challenges and recommendations," *Biomedical Signal Processing and Control*, vol. 74, p. 103493, 2022.

[171] G. Sannino and G. De Pietro, "A deep learning approach for ecg-based heartbeat classification for arrhythmia detection," *Future Generation Computer Systems*, vol. 86, pp. 446–455, 2018.

[172] W. Liu, Q. Huang, S. Chang, H. Wang, and J. He, "Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram," *Biomedical Signal Processing and Control*, vol. 45, pp. 22–32, 2018.

[173] F. Andreotti, O. Carr, M. A. F. Pimentel, A. Mahdi, and M. De Vos, "Comparing feature-based classifiers and convolutional neural networks to detect arrhythmia from short segments of ecg," in *2017 Computing in Cardiology (CinC)*, pp. 1–4, 2017.

[174] M. M. Al Rahhal, Y. Bazi, M. Al Zuair, E. Othman, and B. BenJdira, "Convolutional neural networks for electrocardiogram classification," *Journal of Medical and Biological Engineering*, vol. 38, no. 6, pp. 1014–1025, 2018.

[175] B. Reeder and A. David, "Health at hand: A systematic review of smart watch uses for health and wellness," *Journal of Biomedical Informatics*, vol. 63, pp. 269–276, 2016.

[176] A. Ranjan, "Permanence of ecg biometric: Experiments using convolutional neural networks," in *2019 International Conference on Biometrics (ICB)*, pp. 1–6, 2019.

[177] G. Lopes and J. S. Pinto, João Ribeiroand Cardoso, "Don't you forget about me: A study on long-term performance in ecg biometrics," in *Pattern Recognition and Image Analysis* (A. Morales, J. Fierrez, J. S. Sánchez, and B. Ribeiro, eds.), (Cham), pp. 38–49, Springer International Publishing, 2019.

[178] C. Camara, P. Peris-Lopez, and J. E. Tapiador, "Human identification using compressed ecg signals," *Journal of Medical Systems*, vol. 39, p. 148, Sep 2015.

[179] E. S. Winokur, M. K. Delano, and C. G. Sodini, "A wearable cardiac monitor for long-term data acquisition and analysis," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 1, pp. 189–192, 2013.

[180] Y. Xia, H. Zhang, L. Xu, Z. Gao, H. Zhang, H. Liu, and S. Li, "An automatic cardiac arrhythmia classification system with wearable electrocardiogram," *IEEE Access*, vol. 6, pp. 16529–16538, 2018.

[181] Y. Xia and Y. Xie, "A novel wearable electrocardiogram classification system using convolutional neural networks and active learning," *IEEE Access*, vol. 7, pp. 7989–8001, 2019.

[182] S. Y. Chun, J.-H. Kang, H. Kim, C. Lee, I. Oakley, and S.-P. Kim, "Ecg based user authentication for wearable devices using short time fourier transform," in *2016 39th international conference on telecommunications and signal processing (tsp)*, pp. 656–659, IEEE, 2016.

[183] A. H. Sodhro, A. K. Sangaiah, G. H. Sodhro, S. Lohano, and S. Pirbhulal, "An energy-efficient algorithm for wearable electrocardiogram signal processing in ubiquitous healthcare applications," *Sensors*, vol. 18, p. 923, 2018.

[184] Z. Cai, C. Liu, H. Gao, X. Wang, L. Zhao, Q. Shen, E. Ng, and J. Li, "An open-access long-term wearable ecg database for premature ventricular contractions and supraventricular premature beat detection," *Journal of Medical Imaging and Health Informatics*, vol. 10, pp. 2663–2667, 2020.

[185] J. L. Moraes, M. X. Rocha, G. G. Vasconcelos, J. E. Vasconcelos Filho, V. H. C. De Albuquerque, and A. R. Alexandria, "Advances in photoplethysmography signal analysis for biomedical applications," *Sensors*, vol. 18, no. 6, p. 1894, 2018.

[186] D. Han, S. K. Bashar, J. Lázaro, F. Mohagheghian, A. Peitzsch, N. Nishita, E. Ding, E. L. Dickson, D. DiMezza, J. Scott, *et al.*, "A real-time ppg peak detection method for accurate determination of heart rate during sinus rhythm and cardiac arrhythmia," *Biosensors*, vol. 12, no. 2, p. 82, 2022.

[187] A. Piplai, S. S. L. Chukkapalli, and A. Joshi, "Nattack! adversarial attacks to bypass a gan based classifier trained to detect network intrusion," in *2020 IEEE 6th Intl Conference on Big Data Security on Cloud (BigDataSecurity), IEEE Intl Conference on High Performance and Smart Computing, (HPSC) and IEEE Intl Conference on Intelligent Data and Security (IDS)*, pp. 49–54, 2020.

[188] Y. Liu, F. Wei, J. Shao, L. Sheng, J. Yan, and X. Wang, "Exploring disentangled feature representation beyond face identification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2018.

[189] N. Zaltron, L. Zurlo, and S. Risi, "Cg-gan: An interactive evolutionary gan-based approach for facial composite generation," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, pp. 2544–2551, 2020.

[190] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.