



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



ESCUELA TÉCNICA
SUPERIOR INGENIEROS
INDUSTRIALES VALENCIA

TRABAJO FIN DE MÁSTER EN INGENIERÍA BIOMÉDICA

DISEÑO, IMPLEMENTACIÓN Y VALIDACIÓN DE UNA APLICACIÓN PARA EL SCREENING MASIVO DE LA ENFERMEDAD CELIACA

AUTOR: GONTZAL ELORZA EXEA

TUTOR: ANTONIO MARTÍNEZ MILLANA

COTUTOR: ÁLVARO FIDES VALERO

Curso Académico: 2018-19



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

RESUMEN

Introducción y objetivo: Las preocupantes cifras sobre la epidemiología y la dificultad de diagnóstico de la enfermedad celiaca justifican la necesidad de un sistema para el cribado masivo de la enfermedad celiaca. Para poder llevar a cabo esta acción, se ha propuesto diseñar y desarrollar un sistema que contenga una aplicación móvil capaz de almacenar datos referentes a pacientes, y construir un modelo predictivo para la clasificación de las imágenes tomadas de las tiras reactivas.

Materiales y métodos: Se ha diseñado una aplicación para Android, que almacene de manera local los datos de los usuarios registrados y los pacientes. Por otro lado, se ha propuesto construir un modelo basado en redes neuronales convoluciones para la clasificación automática de imágenes tomadas a las tiras reactivas empleadas por sujetos reales. Se ha dispuesto de un total de un total de 38 tiras válidas. Se han tomado 5 fotos de cada una de las tiras, formando un total de 190 fotos para la creación del modelo. Para la verificación de la aplicación móvil se propone emplear test en base a seis requisitos funcionales y para la validación del modelo de clasificación se proponen métricas clásicas sobre el subconjunto de imágenes de test.

Resultados: Se ha desarrollado y validado la aplicación móvil Android funcional en los teléfonos con versión Android 9.0. Se han construido y validado dos modelos predictivos empleando una red neuronal convolucional capaces de clasificar los resultados de las tiras reactivas con una exactitud del 78%.

Conclusiones: Se ha llevado a cabo un análisis de la enfermedad celiaca y su diagnóstico justificando así este proyecto. Además, se ha diseñado, desarrollado y validado una aplicación móvil capaz de almacenar datos referentes a pacientes con enfermedad celiaca. Asimismo, se ha construido y validado un modelo capaz de clasificar de manera correcta el resultado de diversos tipos de tiras serológicas para la enfermedad celiaca. Con todo esto, se ha creado un sistema para el cribado poblacional sistematizado para esta enfermedad.

Palabras clave: Salud móvil, enfermedad celiaca, procesado de imagen, cribado.



RESUM

Introducció i objectiu: Les preocupants xifres sobre l'epidemiologia i la dificultat de diagnòstic de la malaltia celíaca justifiquen la necessitat d'un sistema per al cribratge massiu. Per poder dur a terme aquesta acció, s'ha proposat dissenyar i desenvolupar un sistema basat en una aplicació mòbil capaç d'emmagatzemar dades referents a pacients, i construir un model predictiu per a la classificació de les imatges preses de les tires reactives.

Materials i mètodes: S'ha dut a terme el disseny d'una aplicació per a Android que guarde de manera local les dades dels usuaris registrats i dels pacients. D'altra banda, s'ha proposat construir un model basat en xarxes neuronals convolucions per a la classificació automàtica de les imatges preses a més a les tires reactives preses emprades per subjectes reals. S'ha disposat d'un total d'un total de 38 tires vàlides. S'han pres 5 fotos de cadascuna de les tires, formant un total de 190 fotos per a la creació del model. Per a la verificació de l'aplicació mòbil es proposa emprar test sobre la base a sis requisits funcionals i per a la validació del model de classificació es proposen mètriques clàssiques sobre el subconjunt d'imatges de test.

Resultats: S'ha desenvolupat i validat l'aplicació mòbil Android funcional en els telèfons amb versió Android 9.0. S'han construït i validat dos models predictius emprant una xarxa neuronal convolucional capaç de classificar els resultats de les tires reactives amb una exactitud del 78%.

Conclusions: S'ha dut a terme una anàlisi de la malaltia celíaca i el seu diagnòstic justificant així aquest projecte. A més, s'ha dissenyat, desenvolupat i validat una aplicació mòbil capaç d'emmagatzemar dades referents a pacients amb malaltia celíaca. Així mateix, s'ha construït i validat un model capaç de classificar de manera correcta el resultat de diversos tipus de tires serològiques per a la malaltia celíaca. Amb tot això, s'ha creat un sistema per al cribratge poblacional sistematitzat per a aquesta malaltia.

Paraules clau: Salut mòbil, malaltia celíaca, processat d'imatge, cribratge.



ABSTRACT

Introduction and objective: The worrying numbers on the epidemiology and the difficulty of diagnosing celiac disease justify the need for a system for mass screening for celiac disease. In order to carry out this action, it has been proposed to design and develop a system containing a mobile application capable of storing patient data, and to construct a predictive model for the classification of images taken from test strips.

Materials and methods: An application has been designed for Android, which stores locally the data of registered users and patients. On the other hand, it has been proposed to build a model based on convolutional neural networks for the automatic classification of images taken from test strips used by real subjects. A total of 38 valid strips were available. Five photos were taken of each of the strips, forming a total of 190 photos for the creation of the model. For the verification of the mobile application it is proposed to use a test based on six functional requirements and for the validation of the classification model classical metrics are proposed on the subset of test images.

Results: A functional Android mobile application has been developed and validated for phones with Android version 9.0. Two predictive models have been constructed and validated using a convolutional neural network capable of classifying the results of the test strips with an accuracy of 78%.

Conclusion: An analysis of celiac disease and its diagnosis has been carried out, thus justifying this project. In addition, a mobile application capable of storing data on patients with celiac disease has been designed, developed and validated. A model capable of correctly classifying the result of different types of serological strips for celiac disease has also been built and validated. With all this, a system has been created for the systematized population screening for this disease.

Key words: mHealth, coeliac disease, image processing, screening.



DOCUMENTOS

1. MEMORIA
2. PRESUPUESTO

ÍNDICE DE LA MEMORIA

CAPÍTULO 1: INTRODUCCIÓN	1
1.1. MOTIVACIÓN	1
1.2. OBJETIVOS	2
1.3. ENFERMEDAD CELIACA	3
1.3.1. PERSPECTIVA HISTÓRICA	3
1.3.2. PATOGÉNESIS	3
1.3.3. EPIDEMIOLOGÍA	4
1.3.4. DIAGNÓSTICO	6
1.3.5. TRATAMIENTO	8
1.4. EL PARADIGMA DE SALUD MÓVIL (MHEALTH)	9
1.5. MACHINE LEARNING	9
1.6. CONTEXTO DE ESTE TRABAJO	11
CAPÍTULO 2: MATERIALES Y MÉTODOS	13
2.1. SISTEMA OPERATIVO Y ENTORNO DE PROGRAMACIÓN INTEGRADO	14
2.1.1. ANDROID	15
2.1.2. ANDROID STUDIO	17
2.2. TIRAS REACTIVAS EMPLEADAS	17
2.3. MODELO DE CLASIFICACIÓN DE IMAGEN	19
2.3.1. ALGORITMO DE CLASIFICACIÓN	20
2.3.2. EVALUACIÓN DEL MODELO	22
2.4. ESPECIFICACIONES INICIALES	23
2.4.1. BASE DE DATOS	23
2.4.2. CUESTIONARIO	25
2.4.3. FLUJO DE TRABAJO DE LOS PROFESIONALES DE LA SALUD	25
2.4.4. FLUJO DE LA APLICACIÓN MÓVIL	26
2.4.5. REQUISITOS FUNCIONALES DE LA APLICACIÓN MÓVIL	29
CAPÍTULO 3: RESULTADOS	33
3.1. APLICACIÓN MÓVIL	33
3.1.1. BASE DE DATOS	33



3.1.2.	DISEÑO DE LA APLICACIÓN	34
3.1.3.	IMPLEMENTACIÓN DE LA APLICACIÓN	42
3.2.	VALIDACIÓN DE LA APLICACIÓN MÓVIL.....	46
3.3.	MODELO DE CLASIFICACIÓN	46
3.3.1.	PROCESO DE CREACIÓN DEL MODELO.....	46
3.3.2.	RESULTADOS DE LOS MODELOS.....	48
CAPÍTULO 4: DISCUSIÓN		55
4.1.	LIMITACIONES	57
4.2.	LÍNEAS FUTURAS	58
CAPÍTULO 5: CONCLUSIONES		59
CAPÍTULO 6: BIBLIOGRAFÍA.....		61

ÍNDICE DEL PRESUPUESTO

1.	INTRODUCCIÓN.....	1
2.	PRESUPUESTOS PARCIALES	1
2.1.	COSTE DE PERSONAL.....	1
2.2.	COSTE DE MATERIAL	2
3.	PRESUPUESTO GENERAL	2

ÍNDICE DE ILUSTRACIONES

Ilustración 1 Patogénesis de la enfermedad celiaca	4
Ilustración 2 Epidemiología en forma de iceberg de la enfermedad celiaca	5
Ilustración 3 Aproximación general al diagnóstico de la enfermedad celíaca.....	7
Ilustración 4 Tipos de machine learning	10
Ilustración 5 Logo del proyecto EURO-glutenscreen	11
Ilustración 6 Distribución de los sistemas operativos empleados en teléfonos en Europa.....	15
Ilustración 7 Ilustración Pila de Android	16
Ilustración 8 Posibles resultados de las tiras Simple CD1WB.....	17
Ilustración 9 Posibles resultados de las tiras Simple CD2WB.....	18
Ilustración 10 Posibles resultados de las tiras Simple CD2WB.....	19
Ilustración 11 Ejemplo del proceso de convolución	21
Ilustración 12 Esquema funcional general de una CNN	21
Ilustración 13 Matriz de confusión.....	22
Ilustración 14 Curvas ROC con diferentes AUC	22
Ilustración 15 Diagrama de entidad - relación de la base de datos	24
Ilustración 16 Diagrama de secuencia de componentes físicos.....	26
Ilustración 17 Diagrama de secuencia del flujo de la aplicación.....	28
Ilustración 18 Diagrama entidad - relación final de la base de datos	33
Ilustración 19 Pantalla de inicio de sesión	34
Ilustración 20 Aviso de identificación de usuario errónea.....	35
Ilustración 21 Pantalla de creación de nuevo usuario	35
Ilustración 22 Aviso de nombre de usuario existente.....	36
Ilustración 23 Aviso de creación de usuario satisfactoria	36
Ilustración 24 Pantalla del menú principal.....	36
Ilustración 25 Pantalla de creación de nuevo paciente	37
Ilustración 26 Aviso de campos nulos en creación del paciente.....	38
Ilustración 27 Pantalla de cuestionario con disposición ScrollView	39
Ilustración 28 Pantalla de la cámara antes y después de tomar una foto	40
Ilustración 29 Solicitud de permisos para acceder a los archivos del dispositivo y tomar fotos	41
Ilustración 30 Aviso de guardado de foto satisfactoria y finalización de cribado.....	41
Ilustración 31 Ventana emergente de eliminación de paciente	41
Ilustración 32 Aviso de eliminación de paciente.....	41
Ilustración 33 Pantalla de consulta de base de datos con datos referentes a un paciente.....	42
Ilustración 34 Parte del código necesario para comunicar fragmentos	44
Ilustración 35 Código necesario para realizar acciones con boton back en fragmentos.....	45
Ilustración 36 Herramienta LabelImg utilizada para delimitar el área de interes de una tira	47
Ilustración 37 Pérdida durante el entrenamiento del modelo 2 en Tensorboard	48
Ilustración 38 Foto de tira reactiva CD2 y la predicción realizada por el modelo 2	49
Ilustración 39 Foto de tira reactiva SimtomaX y la predicción realizada por el modelo 1	49
Ilustración 40 Matriz de confusión y curva ROC del modelo 1	52
Ilustración 41 Matriz de confusión y curva ROC del modelo 2	53



ÍNDICE DE TABLAS

Tabla 1 Análisis de las diferentes pruebas serológicas específicas para la EC	8
Tabla 2 Tiras SimtomaX Glutencheck disponibles.....	19
Tabla 3 Tiras del fabricante Operon disponibles.....	19
Tabla 4 Requisito funcional 1: Registrar usuario.....	29
Tabla 5 Requisito funcional 2: Iniciar sesión	29
Tabla 6 Requisito funcional 3: Registrar paciente.....	30
Tabla 7 Requisito funcional 4: Rellenar cuestionario	30
Tabla 8 Requisito funcional 5: Tomar foto	31
Tabla 9 Requisito funcional 6: Consulta de base de datos.....	31
Tabla 10 Cumplimiento de los requisitos en diferentes modelos de teléfono	46
Tabla 11 Resultados de la validación del modelo 1	50
Tabla 12 Resultados de la validación del modelo 2	51
Tabla 13 Comparación entre los indicadores de los modelos	52





1. MEMORIA



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

CAPÍTULO 1: INTRODUCCIÓN

1.1. MOTIVACIÓN

La enfermedad celiaca (EC) es una enfermedad autoinmune que afecta al sistema gástrico. La consecuencia de esta enfermedad es una respuesta autoinmune contra el gluten, presente en la gran mayoría de los alimentos como el pan de trigo.

La EC afecta al menos al 1% de la población europea. Con 5 millones de personas potencialmente afectadas, es una de las hipersensibilidades alimentaria más común en Europa y, por lo tanto, es un problema de salud pública importante (Mustalahti et al., 2010). A pesar de los avances en medicina y los esfuerzos ocasionales en algunas regiones para detectar activamente la enfermedad, la mayoría de los casos aún no se han diagnosticado o se han diagnosticado con retrasos importantes (Riznik et al., 2019).

La EC ocurre en individuos genéticamente susceptibles, comenzando en la primera infancia y afectando al doble de mujeres que de hombres. La EC es causada por una reacción inflamatoria anormal al gluten, una proteína que se encuentra en los componentes básicos de la dieta europea normal, como el trigo, la cebada y el centeno. Una vez diagnosticado, la EC se trata de manera efectiva con una dieta sin gluten (GFD, por sus siglas en inglés). Los pacientes no diagnosticados y, por lo tanto, no tratados están expuestos a complicaciones graves a largo plazo, debido a que la mayoría de los alimentos contienen gluten. Estas complicaciones, algunas de las cuales son irreversibles, incluyen pérdida de peso, retraso en el crecimiento, pubertad retrasada, osteoporosis, anemia por deficiencia de hierro, trastornos autoinmunes, infertilidad y desarrollo de tumores malignos.

Un diagnóstico temprano de la EC en la infancia es esencial para asegurar un crecimiento óptimo, desarrollo, manejo de síntomas y evitar complicaciones. Sin embargo, la mayoría de los niños con EC no son diagnosticados y son portadores de la enfermedad con sus complicaciones hasta la edad adulta. Además, como la EC puede manifestarse clínicamente en la edad adulta, las complicaciones irreversibles debidas a la tardía o la falta de diagnóstico son inevitables.

1.2. OBJETIVOS

Dadas las preocupantes cifras de la EC y la dificultad de su diagnóstico a nivel poblacional, el objetivo de este trabajo es desarrollar e implementar una aplicación móvil que permita realizar un *screening* masivo y sistematizado de la EC. Para poder cumplir con este objetivo principal se han definido una serie de hitos (objetivos secundarios) que se han abordado de manera secuencial.

En primer lugar, se debe estudiar el contexto sobre el que nace este proyecto, desde el contexto de la propia enfermedad celiaca y su diagnóstico hasta el contexto relacionado con el propio cribado y diagnóstico de esta.

Más adelante, se debe definir la estructura del sistema que se debe tener, las herramientas y plataformas que se deben utilizar para su desarrollo y su flujo de uso previsto, todo esto teniendo en cuenta todas las partes interesadas en este proyecto.

Por ello, el objetivo principal de este trabajo se ha definido como **“diseñar y desarrollar un sistema que facilite el proceso de cribado para la enfermedad celiaca en pacientes en edad pediátrica en diferentes centros de salud o centros escolares”**.

Además, se han definido una serie de objetivos secundarios para cumplir con el objetivo principal:

- Estudiar la EC, epidemiología de la enfermedad y sus métodos de diagnóstico.
- Diseñar una aplicación móvil capaz de realizar las funciones necesarias para poder llevar a cabo el cribado.
- Implementar y validar la aplicación móvil de acuerdo con los requisitos propuestos en el diseño.
- Implementar y validar un modelo predictivo para automatizar el reconocimiento del resultado de las diferentes pruebas serológicas empleadas para el cribado.

1.3. ENFERMEDAD CELIACA

La EC o celiacía es una enfermedad, crónica, autoinmune que afecta principalmente al intestino delgado. Esta enfermedad tiene como causa una reacción autoinmune hacia las gliadinas y gluteninas, que son proteínas que se encuentran en algunos cereales como el trigo, avena, cebada o centeno. Algunos de estos, como el trigo, son fundamentales en la dieta occidental.

1.3.1. PERSPECTIVA HISTÓRICA

La enfermedad celiaca fue descrita por primera vez hace poco menos de 2000 años por el físico griego Areteo de Capadocia. Dio el nombre de “koiliakos” a la enfermedad descrita en honor a la palabra griega “koelia” (abdomen), para dar a entender que estos pacientes estaban “sufriendo en su abdomen”. A finales del siglo XIX, el doctor Samuel Gee redescubrió la enfermedad celiaca y abrió la puerta a futuras investigaciones modernas, sugiriendo que, en caso de haber un tratamiento, sería a través de una dieta (Guandalini, 2008).

Más adelante, en 1944, se logró identificar la causa de la enfermedad celiaca. Durante esta época, la segunda guerra mundial estaba en marcha y, por lo tanto, había una escasez de cereales en todo Europa. Karel Dicke, un pediatra holandés evidenció una mejora en los niños que padecían la enfermedad celiaca. Una vez superada la escasez de cereales, los síntomas volvieron a aparecer en los pacientes que previamente los padecían. Esto llevo a Dicke a deducir la relación entre la ingesta de diferentes cereales o harinas y la enfermedad celiaca (Guandalini, 2008).

Fue en 1950 cuando se identificó al trigo, a la avena y a el centeno como los cereales causantes de la reacción, y demostró cómo una dieta en la que se excluyan estos cereales producía una mejora drástica en los pacientes con EC. De esta manera, se introdujo la GFD (dieta sin gluten por sus siglas en inglés) como tratamiento efectivo para esta enfermedad. Durante esta misma década se introdujo la biopsia intestinal como método de diagnóstico (Guandalini, 2008).

1.3.2. PATOGÉNESIS

La patogénesis de la enfermedad celíaca implica una compleja interacción entre factores ambientales, genéticos e inmunológicos. El gluten y otras proteínas relacionadas provocan respuestas inmunitarias adaptativas en el intestino delgado que conducen al daño de la mucosa. Los genes que codifican antígenos leucocitarios humanos de clase II HLA-DQ2 y –DQ8 están estrechamente relacionados con la enfermedad y se encuentran en casi todos los pacientes celíacos.

El término gluten se refiere a las principales proteínas de almacenamiento del trigo e incluye más de 100 moléculas diferentes. Las proteínas del gluten se dividen dependiendo de la base de la solubilidad en gliadinas y gluteninas, ambas implicadas en la enfermedad celíaca. El centeno y la cebada también contienen proteínas que pueden desencadenar la enfermedad. Algunos péptidos incompletamente digeridos del gluten de trigo y las proteínas relacionadas del centeno y la cebada pueden atravesar el epitelio y entrar en la lámina propia del intestino delgado bajo ciertas condiciones.

Mientras tanto, los residuos neutros de glutamina en los péptidos de gluten pueden convertirse en ácidos glutámicos con carga negativa a través de la deamidación por tTG. Las

células que presentan antígenos que expresan las moléculas HLA-DQ2 y HLA-DQ8 tienen una mayor afinidad con estos péptidos deamidados. La unión posterior de los péptidos inmunogénicos generados a las moléculas HLA da como resultado complejos peptídicos que pueden activar las células T CD4+ específicas del gluten en la lámina propia. La activación de estas células viene acompañada con la producción de diferentes citoquinas que generan un proceso inflamatorio y atrofia en la vellosidad intestinal (Green, Jabri, & Kasarda, 2005) (Ilustración 1).

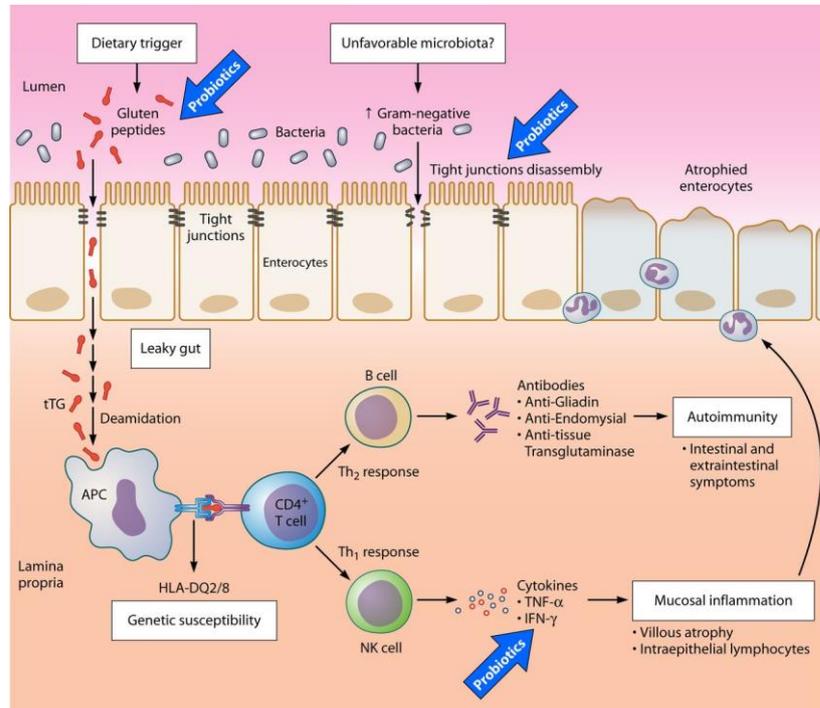


Ilustración 1 Patogénesis de la enfermedad celíaca (Fuente: (de Sousa Moraes, Grzeskowiak, de Sales Teixeira, & Gouveia Peluzio, 2014))

1.3.3. EPIDEMIOLOGÍA

La EC ha sido considerada durante varias décadas una enfermedad rara que afecta a las personas durante los primeros años de vida. En esta época, el diagnóstico de la enfermedad se centraba principalmente en la realización de una endoscopia intestinal. Hoy en día, con la utilización de diferentes pruebas serológicas, se ha podido evaluar la prevalencia real de la EC. De esta forma, se han podido detectar una gran cantidad de casos en los que la EC no podía ser diagnosticada previamente, bien porque eran casos atípicos de la enfermedad o bien porque eran asintomáticos. Se considera que aproximadamente el 1% de la población occidental padece esta enfermedad y que la prevalencia es el doble en mujeres que en hombres (Briani, Samaroo, & Alaedini, 2008).

Un estudio realizado a 17.201 estudiantes entre 6-15 años fue el primero en realizar un *screening* masivo de la EC utilizando pruebas serológicas. Cabe destacar que 1 de cada 7 casos diagnosticados en este estudio no había sido previamente diagnosticado (Catassi et al., 1996). Durante las últimas décadas se han realizado estudios de este tipo en varios países europeos que evidencian la necesidad de pruebas serológicas activas para el diagnóstico de los “casos silenciosos” de la EC.

La prevalencia en Europa varía mucho dependiendo del país, siendo más alta en los países escandinavos y menor en Alemania (Mustalahti et al., 2010). La distribución de la EC está relacionada con las propias causas de la enfermedad. Es más común en zonas donde los genotipos HLA (DQ2 y DQ8) están presentes en mayor proporción, y en zonas donde el consumo de cereales que contienen gluten es mayor.

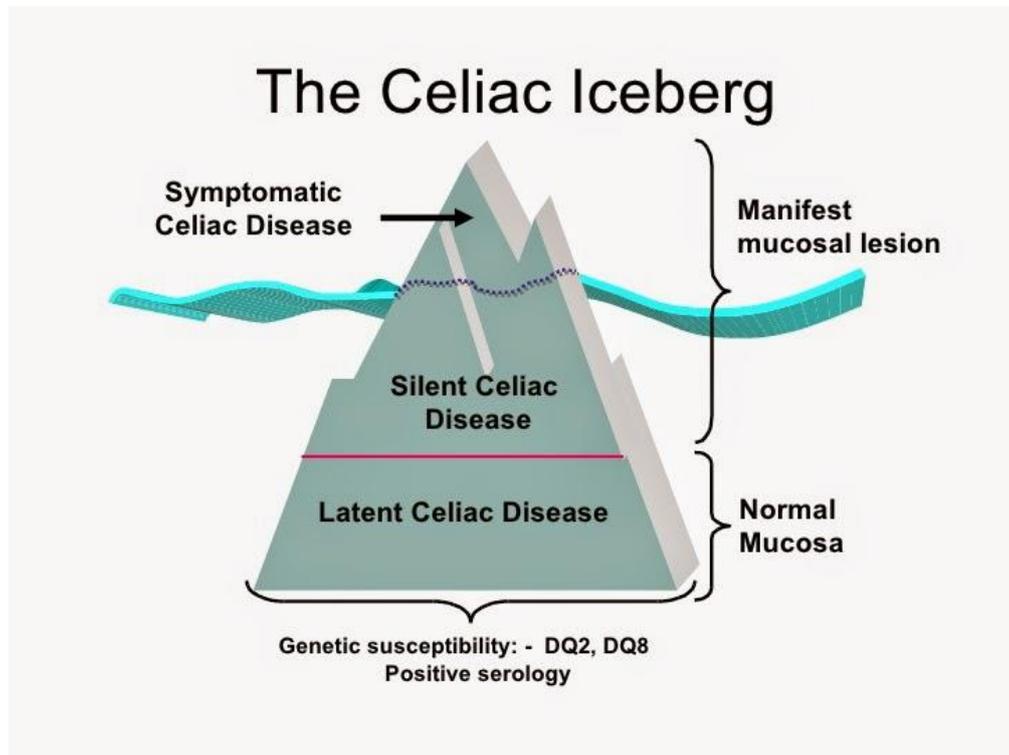


Ilustración 2 Epidemiología en forma de iceberg de la enfermedad celiaca (Fuente: (Fasano, 2005))

La estructura de la epidemiología de la EC se ha denominado como iceberg (Ilustración 2). Por un lado, están los casos diagnosticados (casos por encima del agua), por otro lado, los casos no diagnosticados (casos por debajo del agua) que corresponden a la gran mayoría de los casos. La mayoría de los casos diagnosticados vienen de pacientes con síntomas visibles. Sin embargo, en los países occidentales, por cada sujeto diagnosticado, 5 sujetos permanecen sin diagnosticar. Estos sujetos no diagnosticados están en riesgo de padecer complicaciones a largo plazo que están relacionadas con la EC como infertilidad, osteoporosis o cáncer. Este grupo de pacientes sin síntomas visibles se divide en dos grupos, los que simplemente no presentan los síntomas comunes, pero tienen la mucosa intestinal afectada, y los que no presentan los síntomas comunes, y además no tienen la mucosa intestinal afectada. Sin embargo, todos ellos tienen en común que en el caso de que se les realice una prueba serológica de anticuerpos el resultado sería positivo.

Los casos no diagnosticados en la enfermedad celiaca forman la gran mayoría de casos (Riznik et al., 2019). Este hecho ha abierto un debate entre los expertos sobre los métodos de diagnóstico de la EC. En algunos casos los expertos opinan que el diagnóstico a pacientes en riesgo es una mejor opción como puede ser el caso del actual Colegio Americano de Gastroenterología, la Sociedad Británica de Gastroenterología y las directrices NICE, que recomiendan realizar pruebas a adultos de alto riesgo con serología celíaca (Green, & Cellier, 2007) (Butterworth & Los, 2019). Sin embargo, que tantos sujetos permanezcan sin ser

diagnosticados durante su vida y la mejora en los métodos y en el proceso de diagnóstico puede sugerir que se deba realizar un cribado global a la población.

1.3.4. DIAGNÓSTICO

La incidencia de la enfermedad celíaca está aumentando en todo el mundo debido, en parte, a una mayor apreciación de la variabilidad en la presentación clínica. Hasta la década de los 50, el diagnóstico de la enfermedad celíaca se basaba en observaciones clínicas centradas en las características relacionadas con los síntomas principales. El desarrollo de la biopsia intestinal durante esta misma década produjo un cambio sustancial en el diagnóstico. Desde entonces, la enteropatía dependiente del gluten, basada en la evaluación histológica de la mucosa intestinal, ha sido el estándar para el diagnóstico. En la década de los 80, se desarrollaron pruebas serológicas sensibles y específicas para la enfermedad celíaca. Estas se utilizan ahora como el primer paso cuando hay una sospecha de enfermedad celíaca, para identificar a los pacientes que deben someterse a un análisis de biopsia intestinal. Además, durante las últimas décadas se ha observado que la presentación clínica de la enfermedad celiaca es mucho más heterogénea de lo que se pensaba. Además de la forma gastrointestinal clásica, se han descrito una serie de manifestaciones clínicas adicionales de la enfermedad gracias a la aparición de estas pruebas de cribado serológico (Fasano & Catassi, 2001).

Hoy en día, se recomienda realizar un cribado poblacional a diferentes grupos de pacientes en riesgo de padecer la EC. Estos grupos son familiares de primer y segundo grado de enfermos de EC, personas con diabetes tipo 1 y personas con síndrome de Down, además de los pacientes con síntomas relacionados. A pesar de esto, se considera que la enfermedad celiaca es diagnosticada sólo en 1 de cada 5 personas que la padecen (Green, & Cellier, 2007). Teniendo en cuenta este último dato y las complicaciones que puede suponer el consumo de gluten prolongado en pacientes no diagnosticados, se ha planteado llevar a cabo un cribado poblacional global en la población, pero nunca se ha llegado a realizar.

Varias organizaciones relacionadas con enfermedades gastrointestinales han diseñado diferentes procedimientos para el diagnóstico de la EC. Sin embargo, todas estas utilizan los test serológicos como primer paso para realizar un cribado (Ilustración 3) (Kelly, Bai, Liu, & Leffler, 2015). Estos test serológicos han mejorado su sensibilidad y especificidad a niveles superiores al 90% por lo que han resultado ser una primera aproximación más que suficiente para realizar un cribado eficaz. Una vez se han obtenido resultados positivos en las pruebas serológicas se puede proceder a realizar una biopsia para corroborar que haya daño en el tejido intestinal. Además, esta aproximación mediante pruebas serológicas sirve para no realizar biopsias a todos los pacientes con sospecha de padecer la EC, ya que este procedimiento es más invasivo y por lo tanto no tan ético.

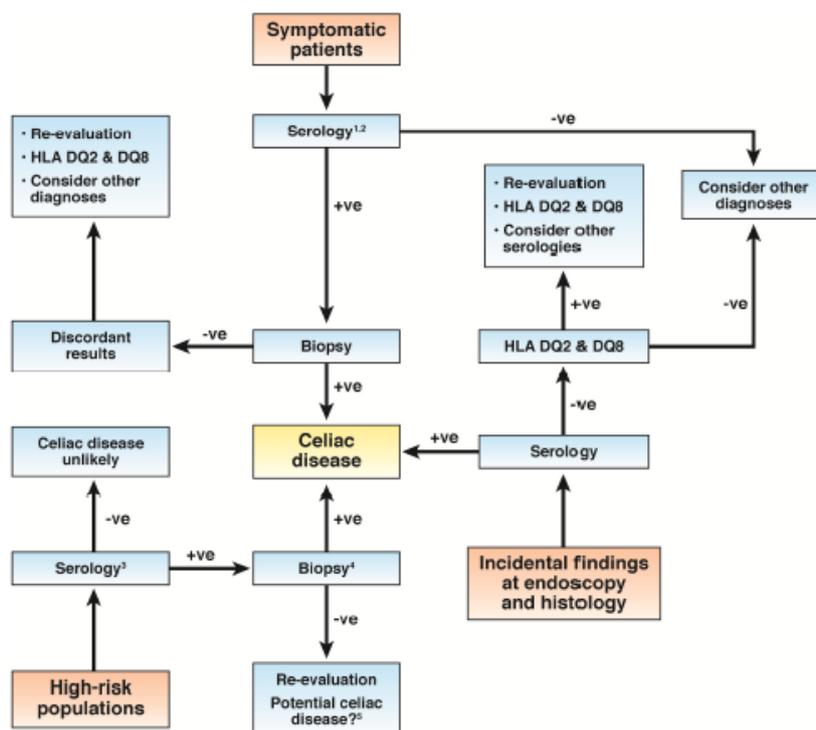


Ilustración 3 Aproximación general al diagnóstico de la enfermedad celíaca (Fuente: (Kelly et al., 2015))

1.3.4.1. PRUEBAS SEROLÓGICAS

Las pruebas serológicas sensibles y específicas a la EC han ido mejorando su especificidad y sensibilidad progresivamente desde su primer desarrollo. El precio de estas también ha bajado. Han permitido que se pueda realizar un cribado entre los pacientes en grupos de riesgo permitiendo evitar realizar una gran cantidad de biopsias intestinales.

Las pruebas actuales sobre la precisión de las pruebas utilizadas para diagnosticar la enfermedad celíaca respaldan la excelente sensibilidad y especificidad de las pruebas de anti-transglutaminasa tisular de tipo inmunoglobulina A (IgA anti-tTg, por sus siglas en inglés). Las pruebas de IgA de anticuerpos endomisiales (EmA, por sus siglas en inglés) tienen una sensibilidad más baja pero una especificidad similar a las pruebas de IgA anti-tTG. Además, las pruebas de IgA anti-tTG son más sensibles que las pruebas de IgA anticuerpos frente al péptido desamidado de la gliadina (DGP, por sus siglas en inglés) y de IgG DGP en adultos que no tienen deficiencia de IgA.

Por otro lado, existen otro tipo de pruebas serológicas que no responden ante anticuerpos, sino que son sensibles a marcadores genéticos. Estas pruebas, detectan la presencia de antígenos leucocitarios humanos (HLA-DQ2 o HLA-DQ8, por sus siglas en inglés) y tienen una sensibilidad y un valor predictivo negativo excelentes. Son muy útiles para descartar pero no para diagnosticar la enfermedad celíaca, ya que todas las personas con enfermedad celíaca tienen presentes estos antígenos, pero no todas las personas con estos antígenos tienen la enfermedad celíaca (Tabla 1) (John M. Eisenberg Center for Clinical Decisions and Communications Science, 2007).

Tabla 1 Análisis de las diferentes pruebas serológicas específicas para la EC (Fuente: (John M. Eisenberg Center for Clinical Decisions and Communications Science, 2007)

Test	Descripción	Detalles
IgA, anti-tTG	tTG es una enzima implicada en la reticulación de ciertas proteínas.	Esta prueba está incluida en los algoritmos de todas las directrices recientes. Se pueden emplear otras pruebas para personas con deficiencia de IgA.
IgA, anti-EmA	El endomisio es la capa delgada de tejido conectivo que cubre las fibras musculares individuales. Los EmA se desarrollan cuando el revestimiento intestinal está dañado.	La mayoría de los pacientes con enfermedad celíaca activa tienen IgA anti-EmA.
IgA e IgG, DGP	Las gliadinas son uno de los principales grupos de proteínas que componen el gluten.	Puede dar un resultado positivo en algunos individuos con enfermedad celíaca que son anti-tTG negativos.
Tipificación de HLA	Se cree que las moléculas HLA presentan antígenos de gluten a las células T, que a su vez inducen daño tisular.	Aproximadamente el 95 por ciento de los pacientes de enfermedad celíaca tienen el heterodímero HLA-DQ2, mientras que el 5 por ciento restante tiene el heterodímero HLA-DQ8.

1.3.4.2. BIOPSIA INTESTINAL

Un resultado en la prueba serológica positivo debe ser sucedido por una biopsia intestinal. Además, también puede realizarse una biopsia si los resultados serológicos son negativos, pero hay evidencias significativas de que haya EC. La biopsia intestinal tiene como objetivo identificar la afección histológica de la vellosidad del intestino. Si se confirma la presencia de tejido dañado, al paciente se le dará la instrucción de seguir una dieta libre de gluten. Solo podrá realizarse un diagnóstico definitivo de la EC al poder apreciarse una mejora tras seguir durante un tiempo la dieta libre de gluten. No es necesaria una segunda biopsia para poder confirmar la mejora de los tejidos.

1.3.5. TRATAMIENTO

No hay tratamiento para curar la EC y la Dieta Libre de Gluten (GFD, por sus siglas en inglés) es la única alternativa aceptada internacionalmente para mitigar los daños producidos por la EC. La GFD implica la eliminación de por vida de trigo, centeno y cebada de la dieta.

Las organizaciones de apoyo a los pacientes son una valiosa fuente de información sobre la enfermedad y la dieta. La mayoría de los países tienen grupos de apoyo nacionales que son de fácil acceso en Internet. El costo de los productos sin gluten varía según el país, pero los productos sin gluten suelen ser más caros que los convencionales, lo que hace que el tratamiento suponga un problema para los pacientes con dificultades financieras. Los productos sin gluten son especialmente costosos y difíciles de encontrar en países en desarrollo. En la

mayoría de los países desarrollados estos productos suelen estar subsidiados (Green & Cellier, 2007). Por ejemplo, en el caso de España, los panes y harinas sin gluten tienen un IVA del 4% respecto al 10% de los productos convencionales.

1.4. EL PARADIGMA DE SALUD MÓVIL (MHEALTH)

El término mHealth o mobile health se utiliza comúnmente para referirse al uso de dispositivos de comunicación móvil, tales como teléfonos móviles, tabletas y dispositivos portátiles, para servicios de asistencia, información y recolección de datos en salud. El campo de la mHealth ha surgido como una parte de la salud digital. Las aplicaciones de mHealth incluyen el uso de dispositivos móviles en la recogida de datos de salud comunitarios y clínicos, la entrega de información sanitaria a los médicos, investigadores y pacientes, la monitorización en tiempo real de los signos vitales de los pacientes, la prestación directa de asistencia (a través de la telemedicina móvil), así como la formación y la colaboración de los trabajadores de la salud (Martinez-Millana, Jarones, Fernandez-Llatas, Hartvigsen, & Traver, 2018).

El crecimiento de la salud digital al igual que el de la mHealth está siendo exponencial. Cada vez son más los procesos en el entorno sanitario que se apoyan en dispositivos móviles. Hoy en día se pueden encontrar más de 97.000 aplicaciones móviles relacionadas con la salud y el bienestar. El 70% de estas están destinadas para su uso en el deporte y el bienestar, el resto están destinadas para profesionales de la salud (Comisión europea, 2014).

Según la consultora PricewaterhouseCoopers las apps destinadas para profesionales de la salud se pueden dividir en 4 grupos principales: respuesta ante emergencias, apoyo a profesionales de la salud, supervisión sanitaria y administración sanitaria.

La categoría de respuesta ante emergencias está formada por aplicaciones que facilitan la respuesta ante desastres o emergencias. La categoría de apoyo a los profesionales compone la gran mayoría de apps destinadas para profesionales y se centran principalmente en proporcionar información, manuales o guías sobre procedimientos, procesos o productos sanitarios. La categoría de supervisión sanitaria incluye aplicaciones que ayudan al personal sanitario en entornos externos al ámbito hospitalario a recopilar información de los pacientes y comunicarse con los médicos. Esta categoría también incluye aplicaciones que permiten el acceso móvil al registro de salud electrónico. Por último, la parte de aplicaciones para la administración sanitaria incluye aplicaciones que facilitan el proceso administrativo de los hospitales como pueden ser recordatorios o cancelaciones de última hora de citas médicas (Singh & Landman, 2017).

Teniendo en cuenta los objetivos propuestos para este proyecto, la aplicación que se plantea desarrollar entra dentro de dos de estos grupos. Por un lado, tiene como objetivo facilitar el proceso de recolección de datos de los pacientes (apoyo a profesionales). Por otro lado, el de almacenar los datos recolectados (supervisión sanitaria).

1.5. MACHINE LEARNING

El machine learning es el estudio científico de los algoritmos y modelos estadísticos que los sistemas informáticos utilizan para realizar una tarea específica de forma eficaz sin utilizar instrucciones explícitas, basándose en patrones e inferencias.

Hoy en día, la digitalización de la sociedad está posibilitando la recolección de una cantidad de datos exponencialmente al alza. Esta gran cantidad de datos es llamada “Big data”. Estos datos se caracterizan por la gran velocidad a la que se generan, el volumen que representan, la inmensa variedad de tipologías que engloban, y el grado de veracidad que poseen.

Tal y como se ha mencionado anteriormente, el machine learning se basa en datos, para poder aprender de ellos y tomar decisiones. El surgimiento del Big data y de otras tendencias relacionadas con la digitalización como Internet of Things junto con el procesamiento computacional más rápido y económico han facilitado que el machine learning haya ido obteniendo cada vez más protagonismo durante los últimos años.

El aprendizaje automático se divide en dos ramas principales, el aprendizaje supervisado y el no supervisado (Ilustración 4). Los algoritmos de clasificación y los algoritmos de regresión son tipos de aprendizaje supervisado. Los algoritmos de clasificación se utilizan cuando las salidas se limitan a un conjunto limitado de valores. Para un algoritmo de clasificación que filtra los correos electrónicos, la entrada sería un correo electrónico entrante, y la salida sería el nombre de la carpeta en la que se archivará el correo electrónico. Para un algoritmo que identifica los correos electrónicos de spam, el resultado sería la predicción de “spam” o “no spam”, representada por los valores booleanos verdadero y falso. Los algoritmos de regresión se denominan así por sus salidas continuas, lo que significa que pueden tener cualquier valor dentro de un rango. Ejemplos de un valor continuo son la temperatura, la longitud o el precio de un objeto (Bishop, 2006).

En el aprendizaje no supervisado, el algoritmo construye un modelo matemático a partir de un conjunto de datos que contiene sólo entradas y etiquetas de salida no deseadas. Los algoritmos de aprendizaje no supervisados se utilizan para encontrar la estructura de los datos, como la agrupación o agrupación de puntos de datos. El aprendizaje no supervisado puede descubrir patrones en los datos, y puede agrupar las entradas en categorías, como en el aprendizaje de características (Bishop, 2006).

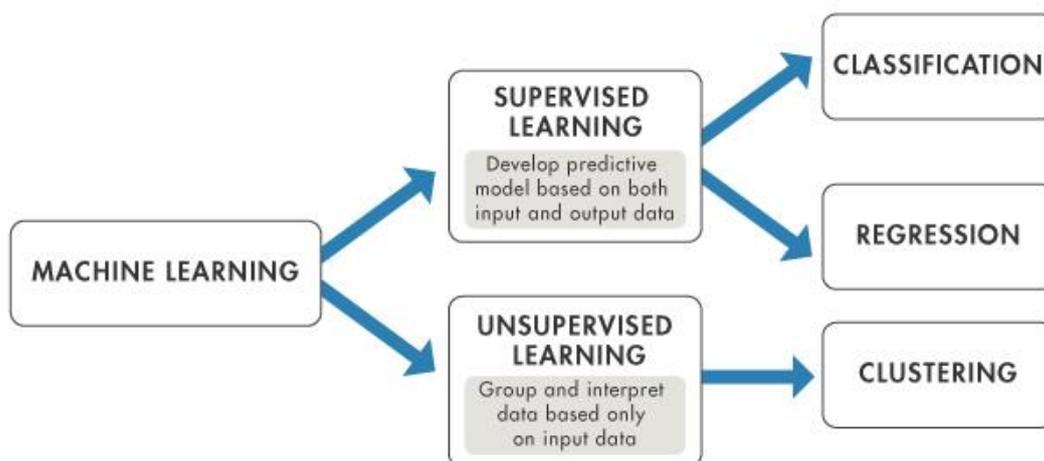


Ilustración 4 Tipos de machine learning (Fuente: <https://www.mathworks.com/help/stats/machine-learning-in-matlab.html>)

Teniendo en cuenta los objetivos propuestos para este proyecto se deberá realizar un modelo de aprendizaje supervisado para clasificación ya que se ha planteado clasificar las imágenes de los resultados de las pruebas serológicas para automatizar en la medida de lo posible la recolección de los datos.

1.6. CONTEXTO DE ESTE TRABAJO

El presente trabajo fin de master se engloba en la colaboración del grupo de Gastroenterología Pediátrica del Hospital Universitari i Politènic La Fe, dirigido por la Dra. Carmen Ribes y el grupo de investigación ITACA-SABIEN de la Universitat Politècnica de València dirigido por el Dr. Vicente Traver. Ambos grupos han colaborado estrechamente en la investigación de nuevo métodos basados en mHealth para el apoyo a enfermedades crónicas como la Fibrosis Quística en el proyecto H2020 MyCyFAPP, cuyos resultados serán explotados por la farmacéutica Mylan a partir del año 2020 en 6 países de la Unión Europea.

En lo que respecta a este TFM, la necesidad de realizar *screening* de EC a nivel poblacional y de manera sistematizada fue detectada por el grupo que dirige la Dra. Ribes y junto a un consorcio de más de 10 socios entre los cuales destacan 5 universidades, 2 centros de innovación y 2 asociaciones de pacientes de EC, se ha redactado una propuesta de proyecto titulada “*Effective European screening programmes to achieve prevention of coeliac disease by early diagnosis and treatment: EURO-glutenscreen*” (Ilustración 5) enviada al tema **SC1-BHC-30-2019** de la convocatoria en dos fases de Octubre 2018-Abril 2019.15361828J



Ilustración 5 Logo del proyecto EURO-glutenscreen (Fuente: ITACA-SABIEN)

CAPÍTULO 2: MATERIALES Y MÉTODOS

La realización de este proyecto requiere el empleo de diferentes programas informáticos, herramientas de aprendizaje automático y diferentes tiras reactivas para el diagnóstico de la enfermedad celiaca. Este apartado describe el material empleado para el desarrollo de este proyecto además de justificar su elección. En la última sección de este capítulo, se describe el diseño inicial del sistema planteado para el *screening* masivo de la EC.

A la hora de definir el alcance del proyecto, se definieron varias opciones para abordar los objetivos establecidos. Algunos de los elementos del sistema propuesto han recibido diferentes enfoques.

Por un lado, se han planteado diferentes formas para almacenar los datos obtenidos de los pacientes y visualizarlos. En un principio, había dos opciones, almacenar los datos de manera local en la aplicación móvil, o almacenar los datos en un servidor y visualizarlos mediante una aplicación web. Obviamente, la segunda opción requiere más tiempo y una serie de conocimientos más amplios (programación web, envío de datos a un servidor...) mientras que la primera opción sólo requiere conocimientos en bases de datos en aplicaciones móviles. **Se ha decidido que se empleará una base de datos local en la propia aplicación** sin necesidad de recurrir a comunicarse con la nube.

Por otro lado, ha habido dos opciones para abordar la integración del modelo al sistema. La primera opción es realizar un modelo integrado en la aplicación móvil que reconoce directamente las imágenes obtenidas. La segunda opción es realizar un modelo sin integrarlo en la aplicación, de manera que funcione en un PC o servidor. La integración del modelo en la aplicación móvil es más difícil de realizar debido a la falta de documentación, información y tutoriales que hay al respecto. Por ello, **se ha decidido no integrar el modelo en la aplicación y construirlo de manera que funcione en un PC.**

Por lo tanto, la idea inicial del proyecto contará con tres partes principales, dos de software y una física. Por un lado, **se deberá diseñar una aplicación móvil en el sistema operativo Android que sea capaz de cumplir con los objetivos propuestos para este proyecto**, es decir, deberá poder mostrar una interfaz mediante la cual se pueden almacenar los datos obtenidos de los pacientes (cuestionario, datos demográficos y foto de la tira serológica), deberá posibilitar el registro e identificación del personal sanitario que empleará la aplicación. Además, se puede considerar la posibilidad de visualizar los datos obtenidos.

Por otro lado, **se deberá construir un modelo a partir de las imágenes tomadas a las tiras reactivas para poder realizar la predicción del resultado.** Se deberán analizar las diferentes opciones que se disponen a la hora de definir las clases de salida del modelo en base a sus resultados. Hay tres tipos de tiras reactivas disponibles con resultados diferentes en cada una de ellas. La opción más simple es realizar un modelo que clasifique entre los resultados positivos y negativos. La segunda opción es que el modelo clasifique por un lado el modelo de la tira, y por otro el resultado positivo o negativo, por ejemplo, *SimtomaX Glutencheck* positivo. La opción más compleja es la de realizar un modelo que clasifique por un lado el tipo de tira que es

y por otro lado el resultado exacto del test, por ejemplo, *SimtomaX Glutencheck* EC positivo IgA positivo.

El último componente necesario para la realización del proyecto son las tiras reactivas descritas en el apartado 2.2.

2.1. SISTEMA OPERATIVO Y ENTORNO DE PROGRAMACIÓN INTEGRADO

Una de las piezas claves para el desarrollo de este proyecto es la propia aplicación móvil. A la hora de poder diseñar y desarrollar una aplicación móvil, uno de los primeros pasos es decidir el sistema operativo para el cual estará creada la aplicación. Hoy en día, la amplia mayoría de teléfonos móviles inteligentes se dividen en dos sistemas operativos: iOS (Apple) y Android.

Uno de los parámetros más importantes a la hora de poder seleccionar el sistema operativo es la cantidad de dispositivos vendidos en los países donde se desea aplicar este proyecto. La razón por la cual se ha decidido crear este proyecto es un proyecto europeo para realizar un cribado masivo poblacional. Por ello, se ha analizado la cantidad de dispositivos de cada tipo en los países europeos.

Aunque no haya datos oficiales sobre la distribución de los sistemas operativos de los móviles, se ha podido observar que, en los datos disponibles, a nivel europeo la diferencia es muy significativa (Ilustración 6). Se puede observar que como se ha mencionado anteriormente, los dos sistemas operativos principales son iOS, con un 26,64% las ventas y Android, con un 72,48% en junio de 2019. Además, se puede observar, que la tendencia de Android ha sido ascendente durante los últimos años. De esta manera, se puede considerar que Android se ha establecido durante este periodo como el sistema operativo de dispositivos móviles de facto en Europa. En el caso de España, la diferencia entre los sistemas operativos empleados es aún más significativa, ya que en junio de 2019 los dispositivos Android componían en 77,97% del mercado mientras que los iOS solo el 21,71%¹.

A todo esto, se le debe sumar el hecho de que Android es un sistema operativo libre, disponible para dispositivos de una infinidad de fabricantes mientras que iOS solo sirve para teléfonos móviles o dispositivos fabricados por Apple.

Asimismo, el software necesario para poder desarrollar aplicaciones en cada uno de estos sistemas operativos es diferente. Por un lado, el entorno de desarrollo integrado para Android es gratuito, y se puede probar la aplicación diseñada en cualquier dispositivo Android sin coste alguno. No obstante, en iOS se debe pagar una cuota de desarrollador iOS para poder realizar pruebas en un dispositivo.

Teniendo en cuenta todas estas razones, se ha decidido utilizar el sistema operativo Android para el desarrollo de la aplicación móvil. El entorno de desarrollo integrado que se ha decidido utilizar es el propio de Android, llamado Android Studio.

¹ <http://gs.statcounter.com/os-market-share/mobile/europe/#monthly-201406-201906>

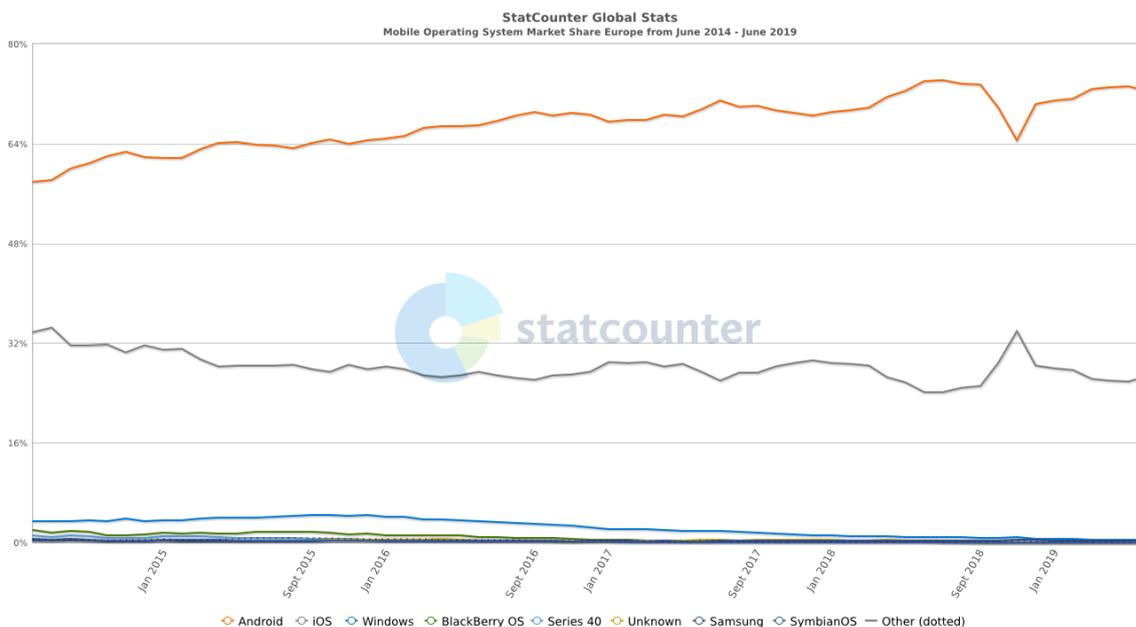


Ilustración 6 Distribución de los sistemas operativos empleados en teléfonos en Europa (Fuente: <http://gs.statcounter.com/os-market-share/mobile/europe/#monthly-201406-201906>)

2.1.1. ANDROID

Android es un sistema operativo para dispositivos móviles desarrollado por Google. Está basado en una versión modificada del kernel de Linux y otro software de código abierto, y está diseñado principalmente para dispositivos móviles con pantalla táctil como *smartphones* y tabletas.

Desarrollado inicialmente por Android Inc. Que Google compró en 2005, Android fue presentado en 2007, con el primer dispositivo comercial Android lanzado en septiembre de 2008. La versión estable actual es Android 9 “Pie”, lanzada en 2018. Google lanzó la primera beta de la próxima versión, Android Q, en teléfonos Pixel en marzo de 2019. El código fuente central de Android se conoce como Android Open Source Project (AOSP), que está licenciado principalmente bajo la licencia Apache, que es una licencia de software libre.

Android está basado en capas con una pila fragmentada (Ilustración 7). A continuación, se explicará cada una de las capas para poder entender mejor la forma en la que funciona este sistema operativo.

La capa base de la pila de Android es el kernel basado en Linux. El uso de un kernel de Linux permite a Android aprovechar las principales funciones de seguridad y permite a los fabricantes de dispositivos desarrollar controladores de hardware para un núcleo ampliamente conocido.

La capa de abstracción de hardware consiste en varios módulos que implementan una interfaz para un tipo específico de componente de hardware, como puede ser el módulo de la cámara o el del altavoz. Cuando el *framework* de una aplicación realiza una llamada para acceder a hardware del dispositivo, el sistema carga el módulo de biblioteca para el componente de hardware seleccionado.

El Android Runtime está escrito para ejecutar varias máquinas virtuales en dispositivos de memoria baja y compila fuentes de Java de manera que puedan ser ejecutadas en la plataforma Android.

Hay componentes que se basan en código nativo que requiere bibliotecas desarrolladas en C y C++. Android proporciona la posibilidad de acceder a estas bibliotecas, es decir, da la posibilidad de utilizar librerías escritas en C y C++.

El conjunto completo de funciones del sistema operativo Android está disponible a través de APIs escritas en lenguaje Java. Estas APIs forman los bloques de construcción necesarios para crear aplicaciones Android al simplificar la reutilización de los componentes y servicios centrales y modulares del sistema.

Por último, Android viene con una serie de aplicaciones preinstaladas como puede ser el navegador. Las aplicaciones del sistema funcionan tanto como aplicaciones destinadas a usuarios como para proporcionar funcionalidades clave para los desarrolladores. Es necesario crear esa funcionalidad, sino que se puede invocar cualquier aplicación que tome fotos que ya esté instalada (Google, 2019b).

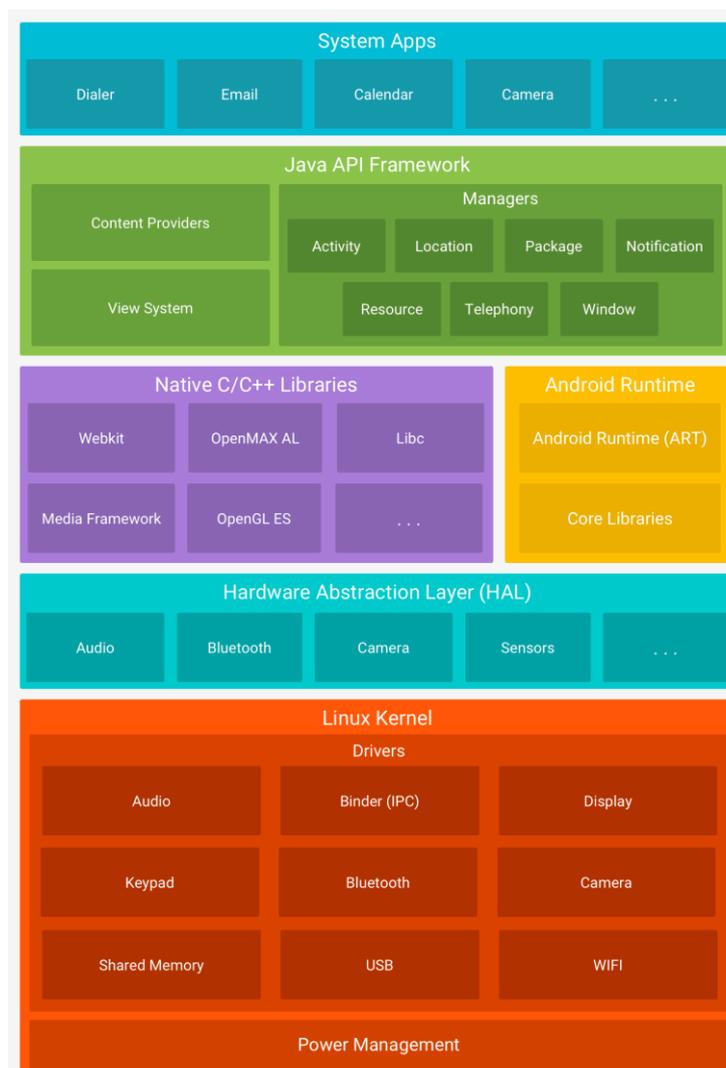


Ilustración 7 Ilustración Pila de Android (Fuente: (Google, 2019b))

2.1.2. ANDROID STUDIO

Android Studio es el entorno de desarrollo integrado oficial para el desarrollo de aplicaciones Android, basado en IntelliJ IDEA. Además de las potentes herramientas de desarrollo y edición de código de IntelliJ, Android Studio ofrece aún más funciones que mejoran su productividad al crear aplicaciones para Android.

Cada proyecto de Android Studio contiene como mínimo un módulo, donde estarán almacenados los datos referentes al proyecto. Cada módulo contiene las carpetas de manifest, java y res. La carpeta manifest contiene el archivo AndroidManifest.xml que es un archivo de configuración donde se pueden aplicar las configuraciones básicas de la app. La carpeta java contiene los archivos de código fuente y la carpeta res contiene los recursos que no sean código, como las imágenes empleadas para el diseño de la app o los archivos XML que forman la interfaz de la app.

2.2. TIRAS REACTIVAS EMPLEADAS

Otro de los componentes necesarios para poder realizar este proyecto son las tiras reactivas para el cribado de la enfermedad celiaca. En este caso, las tiras reactivas que se han empleado son las recogidas por el personal sanitario durante los últimos 10 años en el Hospital Universitari i Politècnic La Fe.

Dentro de estas tiras reactivas, hay tres tipos que se dividen en dos fabricantes. Por un lado, están las tiras reactivas del fabricante *Augurix*, cuyo nombre de producto es *SimtomaX GlutenCheck*. Por otro lado, están las tiras reactivas del fabricante *Operon*, que se dividen en tiras *Simple CD1WB* y *Simple CD2WB*.

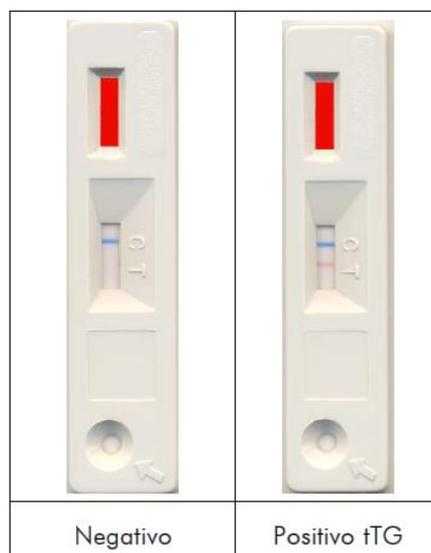


Ilustración 8 Posibles resultados de las tiras Simple CD1WB (Fuente: <https://operon.es/es/ceeliac-disease/>)

Simple CD1WB es un test inmunocromatográfico diseñado para la detección en sangre humana de anticuerpos (IgA, IgG e IgM) dirigidos contra transglutaminasa tisular humana, principal antígeno reconocido por los anticuerpos antiendomisiales. Este tipo de tiras solo tiene dos resultados posibles, positivo o negativo. En el caso negativo sólo aparece una línea transversal Azul en la zona central del dispositivo de reacción alineada con la letra “C” marcada en la carcasa. Esta línea siempre debe aparecer, sino, el test se considerará inválido. El resultado

será positivo si además de la línea azul de control aparece una línea transversal roja en la zona alineada con la letra “T” de la carcasa (Ilustración 8). La sensibilidad y especificad de este tipo de test están en torno al 90 %².

Simple CD2WB es un test parecido al *CD1WB*. Detecta los anticuerpos de tipo IgA dirigidos contra transglutaminasa tisular humana y contra gliadinas en la sangre humana. Los resultados posibles de este tipo de test son 5. Este test tiene tres bandas de colores. La banda azul constituye la banda de control que indica el correcto funcionamiento del test, la banda roja central indica la presencia de IgAs anti-gliadina y la banda roja inferior indica la presencia de IgAs anti-tTG. De esta manera, se considerará el resultado negativo cuando sólo aparezca la línea azul alineada con la letra “C” de la carcasa. El resultado será inválido si esta tira no aparece. El resto de resultados son positivos y hay 3 posibles opciones. Si se detectan IgAs anti-tTG aparecerá una tira roja debajo de la de control alineada con la letra “T1” de la carcasa. Si se detectan IgAs anti-gliadina aparecerá una tira roja alineada con la letra “T2” de la carcasa (Ilustración 9). Si se detectas IgAs anti-gliadina y anti-tTG aparecerán todas las tiras. La sensibilidad y especificidad de este test rondan también el 90% siendo mejor para IgAs anti-tTG que para anti-gliadina.

Tira 1	Tira 2	Tira 3	Tira 4	Tira 5
				
Negativo	Positivo IgAs anti-tTG	Positivo IgAs anti-gliadinas	Positivo IgAs anti-tTG y anti-gliadinas	Inválido

Ilustración 9 Posibles resultados de las tiras Simple CD2WB (Fuente: <https://operon.es/es/ceeliac-disease/>)

Por último, las tiras *SimtomaX glutencheck* combinan la detección de anticuerpos IgA frente a péptidos de gliadina desamidados con la detección de anticuerpos IgA anti-tTG. Este tipo de test tiene 4 resultados válidos posibles. El resultado será negativo si una barra transversal azul aparece alineada con las letras “CT” de la carcasa. En el caso de que esta línea no aparezca, el resultado del test no será válido. En el caso de que aparezca una barra transversal roja alineada con la letra “B” de la carcasa, el resultado será negativo para la EC, pero habrá presencia de IgAs. En el caso de que la barra transversal que aparezca esté alineada con la letra “A” de la carcasa, el resultado para la EC será positivo, pero habrá ausencia de IgAs. En el caso de que ambas barras transversales rojas aparezcan, el resultado será positivo para la EC y habrá

² <https://operon.es/es/ceeliac-disease/>

presencia de IgAs (Ilustración 10). La sensibilidad para este tipo de tiras es de 93,1% y la especificidad es de 95%. El fabricante de estas tiras presume de tener un valor de predicción negativa del 99,1%³.

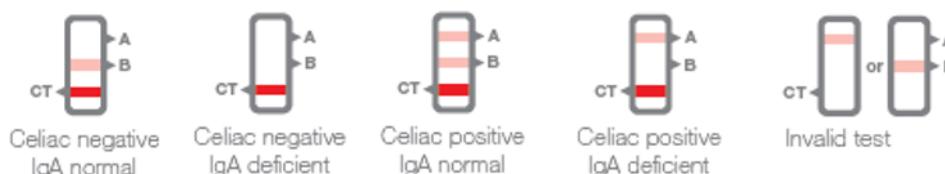


Ilustración 10 Posibles resultados de las tiras Simple CD2WB (Fuente: <http://augurix.com/simtoma/documentation/>)

En total se ha dispuesto de 21 tiras *SimtomaX Glutencheck*, 12 tiras *CD1WB* y 13 tiras *CD2WB* formando un total de 46 tiras reactivas. Sin embargo, de estas 46 tiras reactivas, algunas se encuentran en mal estado o deterioradas, ya que algunas llegan a tener 10 años de antigüedad. Las tiras que más se han deteriorado han sido las *SimtomaX Glutencheck* ya que sólo se han empleado 13 de las 21 tiras para el estudio. Por otro lado, los resultados de las tiras no están distribuidos de manera uniforme. En las tiras *CD2WB* no se tiene ninguna tira con el resultado de IgAs anti-gliadina. Además, en el hospital no se ha recogido ningún test inválido. Se pueden observar los resultados de las tiras empleadas en la tabla 2 y la tabla 3.

Tabla 2 Tiras *SimtomaX Glutencheck* disponibles (Fuente propia)

Tira reactiva	EC negativo, IgA negativo	EC negativo, IgA positivo	EC positivo, IgA positivo	EC positivo, IgA negativo	Total
<i>SimtomaX Glutencheck</i>	2	3	7	1	13

Tabla 3 Tiras del fabricante *Operon* disponibles (Fuente propia)

Tira reactiva	Positivo	Negativo	IgA anti-tTG	IgA anti-gliadina	IgA anti-tTG y anti-gliadina	Total
<i>CD1WB</i>	7	5	-	-	-	12
<i>CD2WB</i>	-	3	5	0	5	13

Se han tomado 5 fotos a cada una de las tiras reactivas. Las fotos han sido tomadas teniendo en cuenta diferentes parámetros. Cada una de las 5 fotos varía en el ángulo, la iluminación y el fondo para que el conjunto de datos sea lo más variado posible.

2.3. MODELO DE CLASIFICACIÓN DE IMAGEN

Otro de los pilares de este proyecto se centra en automatizar en la medida de lo posible la obtención, procesamiento y almacenamiento de los resultados obtenidos en la aplicación móvil. Uno de los puntos en los que más margen de automatización tiene es el de la

³ <https://operon.es/es/ceeliac-disease/>

interpretación del resultado del test serológico. Para realizar esta función se deberá realizar un modelo de clasificación de imagen mediante el cual se determinará el resultado del test. El software y procesamiento computacional disponible hoy en día facilitan la posibilidad de realizar este tipo de tareas.

Una de las herramientas más extendidas para la creación de modelos mediante machine learning es una librería de google llamada Tensorflow. Esta librería se puede obtener de manera gratuita y es compatible con el lenguaje de programación Python. Gracias a este tipo de herramientas los procesos de obtención de datos, entrenamiento de modelos y realizar predicciones son mucho más fáciles y rápidos.

Para poder utilizar Tensorflow se debe disponer de un entorno de desarrollo integrado que sea compatible con la librería. Para la realización del modelo de este proyecto, se ha optado por utilizar Python 3.6. La gran cantidad de guías, ejemplos y tutoriales disponibles en la red sobre cómo utilizar Tensorflow en Python 3.6 ha facilitado esta elección.

Para la creación de un modelo de manera supervisada, se debe tener un conjunto de datos. En este caso, el conjunto de datos son fotos de las tiras reactivas. Una vez obtenido el conjunto de datos, se deberá definir manualmente el resultado de cada una de las fotos. A partir de ahí, se podrá entrenar el modelo. Una vez el modelo se haya entrenado, se podrán realizar predicciones de imágenes nuevas.

2.3.1. ALGORITMO DE CLASIFICACIÓN

Otro parámetro que se debe definir para entrenar el modelo, es el algoritmo que se utilizará para realizar esta acción. Hoy en día, la manera más empleada para realizar esta tarea es la utilización de una red neuronal convolucional (CNN, por sus siglas en inglés). En lugar de preprocesar los datos para derivar características como texturas y formas, una CNN sólo toma los datos de píxeles de la imagen como entrada, “aprende” cómo extraer estas características y, en última instancia, infiere qué objeto constituyen. La entrada de un CNN es una matriz de datos tridimensional donde las dos primeras dimensiones hacen referencia a la altura y la anchura de la imagen y la tercera dimensión hace referencia al color de cada celda en formato RGB.

Una CNN está compuesta por una serie de módulos que cumplen funciones diferentes. El primer módulo por el que pasa la imagen es la convolución. Una convolución toma como entrada las diferentes casillas de la matriz de datos de la imagen y les aplica filtros para calcular nuevas características, produciendo un mapa de características de salida o una característica convolucionada. Las convoluciones se definen por el tamaño de las casillas de entrada y por la profundidad del mapa de salida (veces que se aplica el filtro). La CNN multiplica cada casilla de entrada por cada casilla del filtro para después sumar los elementos. Cada una de estas sumas forma una casilla de la nueva matriz convolucionada (Ilustración 11).

Durante el periodo de entrenamiento, la CNN obtiene los valores óptimos para el filtro convolucional mediante los cuales se obtienen mayor cantidad de características de la imagen de entrada. Cuanta mayor sea la cantidad de filtros que se aplican, mayor es la cantidad de características extraídas y mayor es el gasto computacional. Sin embargo, el gasto computacional va en aumento progresivamente mientras que cada vez se extraen menos características.

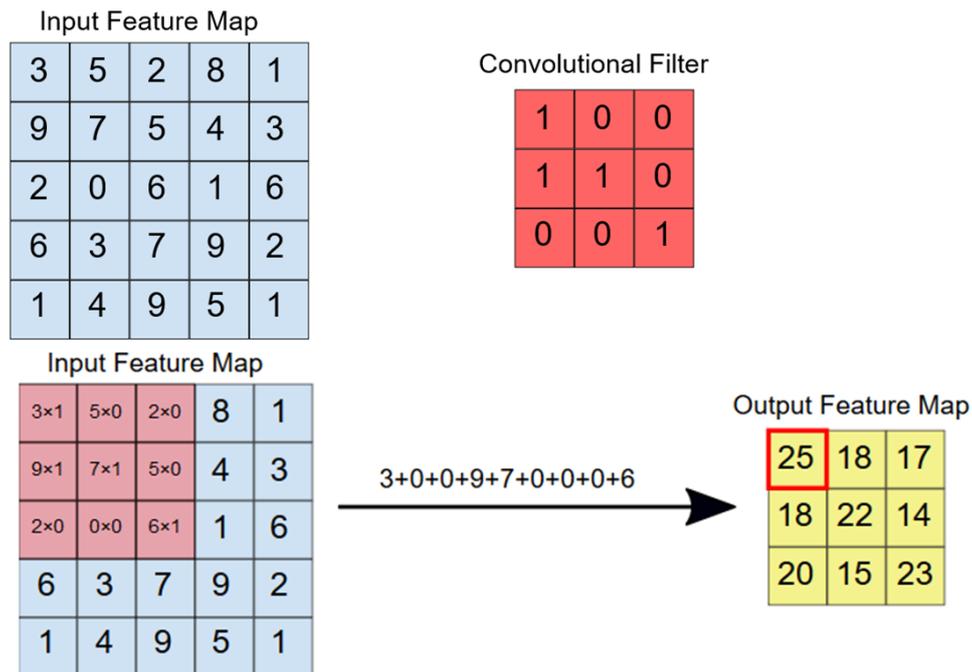


Ilustración 11 Ejemplo del proceso de convolución (Fuente: (Google, 2019^a))

El siguiente módulo se llama Unidad Linear Rectificada (ReLU, por sus siglas en inglés) y tiene como objetivo introducir no linealidad al modelo. El tercer y último módulo se llama Pooling. En este paso se reduce el número de dimensiones del mapa de características, al mismo tiempo que preserva la información más crítica de la característica. Un algoritmo común utilizado para este proceso se llama max pooling. El funcionamiento de este algoritmo es muy parecido a la convolución, solo que, en vez de realizar una multiplicación, se extrae el valor máximo de cada grupo de casillas a la salida.

Al final de una red neuronal convolucional hay una o más capas completamente interconectadas. Su trabajo consiste en realizar una clasificación basada en las características extraídas por las convoluciones. Normalmente, la capa final completamente conectada contiene una función de activación softmax, que produce un valor de probabilidad de 0 a 1 para cada una de las etiquetas de clasificación que el modelo está tratando de predecir (Google, 2019^a) (Ilustración 12).

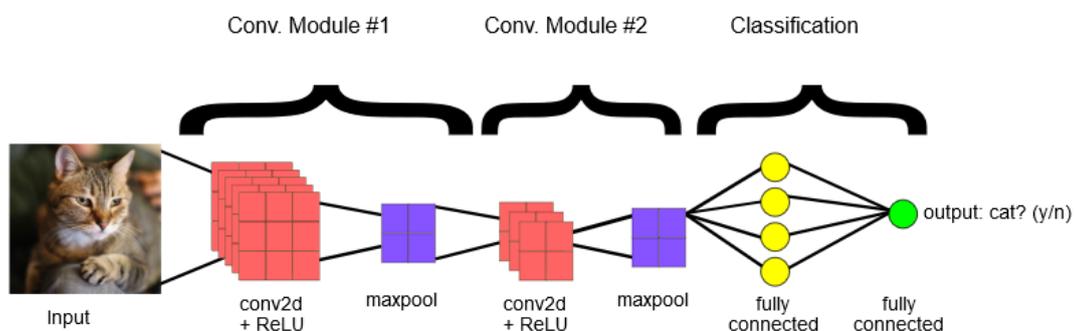


Ilustración 12 Esquema funcional general de una CNN (Fuente: (Google, 2019^a))

2.3.2. EVALUACIÓN DEL MODELO

Una de las técnicas para evaluar los modelos predictivos es la matriz de confusión. Una matriz de confusión es una herramienta que permite la visualización del desempeño de un algoritmo que se emplea en aprendizaje supervisado. Cada columna de la matriz representa el número de predicciones de cada clase, mientras que cada fila representa a las instancias en la clase real (Ilustración 13).

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Ilustración 13 Matriz de confusión (Fuente: <https://rpubs.com/chzelada/275494>)

En las celdas de la primera diagonal se sitúan los datos correctamente clasificados en sus correspondientes clases, y en la otra diagonal se encuentran los erróneos. Además, hay varios índices que se pueden obtener de estos datos:

- Exactitud: Valores verdaderos entre todos los valores.
- Tasa de error: Valores falsos entre todos los valores.
- Sensibilidad: Verdaderos positivos entre total de positivos.
- Especificidad: Verdaderos negativos entre total de negativos.
- Precisión: Verdaderos positivos entre total de clasificados como positivos.
- Valor de predicción negativo: Verdaderos negativos entre total de clasificados como negativos.

De los valores de sensibilidad y especificidad se puede obtener la curva ROC. La curva ROC muestra la relación entre falsos positivos ($1 - \text{especificidad}$) y verdaderos positivos (sensibilidad). Mide la capacidad de un modelo para distinguir entre clases. Cuanto más alta sea el área bajo la curva (AUC, por sus siglas en inglés), mejor será el modelo para predecir las clases correctamente (Ilustración 14). Por ejemplo, cuanto más alta sea el AUC, mejor será el modelo para distinguir entre pacientes con enfermedad y sin enfermedad.

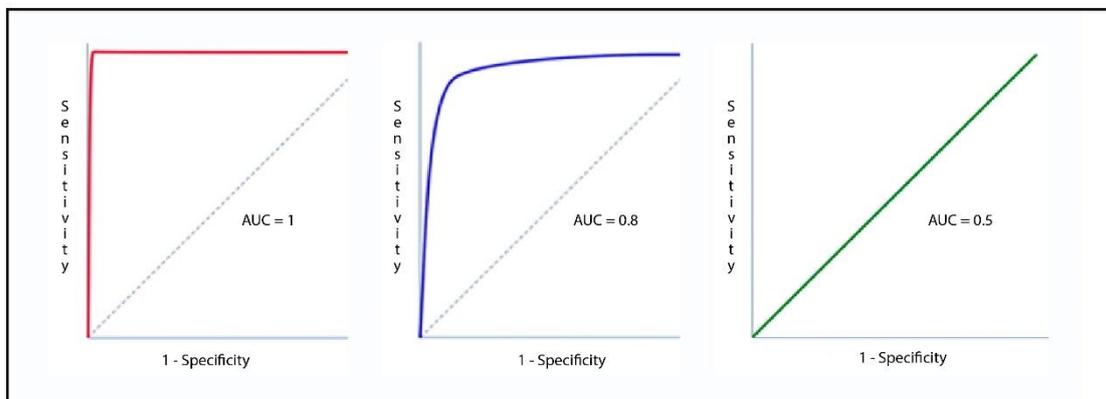


Ilustración 14 Curvas ROC con diferentes AUC (Fuente: <https://www.cienciasinseso.com/en/tag/roc-curve/>)

2.4. ESPECIFICACIONES INICIALES

2.4.1. BASE DE DATOS

Para realizar la base de datos de la aplicación móvil se deben definir el tipo de base de datos, las entidades que contendrá y la base de datos, los atributos de cada entidad y por último la relación entre las entidades.

Android Studio dispone de una librería nativa para trabajar con bases de datos relacionales SQLite llamada Room persistency. Esta librería facilita las operaciones con una base de datos interna en aplicaciones Android. Por ello, se ha decidido emplear este tipo de base de datos.

Las bases de datos relacionales tienen una serie de ventajas respecto a las bases de datos no relacionales. Son bases de datos sencillas, los datos se implementan y gestionan con facilidad y los datos se almacenan de manera organizada gracias a la estructura de tablas sobre la que se basan. Los datos almacenados son consistentes, es decir, gracias al tipo de almacenamiento de los datos es difícil encontrar contradicciones. Además, el sistema de base emplea un procesamiento orientado a conjuntos que subdivide cada entidad en valores mínimos. Esto permite conectar entidades diferentes por medio del contenido. Asimismo, el lenguaje de consultas para las bases de datos relacionales está estandarizado.

Las bases de datos relacionales definen la relación entre las diferentes entidades de la base de datos. Para ello, se deben definir claves primarias y claves foráneas. Una clave primaria es una columna especial de la tabla de la base de datos relacional (o una combinación de columnas) designada para identificar de manera única todos los registros de la tabla. Una clave foránea es una columna o grupo de columnas en una entidad que proporciona un enlace entre los datos de dos tablas. Actúa como una referencia cruzada entre tablas porque hace referencia a la clave primaria de otra tabla, estableciendo así un vínculo entre ellas (Tutorialspoint, n.d.). Los tipos de datos que se pueden almacenar en una base de datos SQLite son:

- Null: El valor está vacío.
- Integer: El valor es un número entero con signo, almacenado en 1, 2, 3, 4, 6 u 8 bytes dependiendo de la magnitud del valor.
- Real: El valor es un valor de punto flotante, almacenado como un número de punto flotante IEEE de 8 bytes.
- Text: El valor es una cadena de texto, almacenada utilizando la codificación de la base de datos.
- Blob: El valor es grupo de datos, almacenados exactamente como han sido introducidos.

Las entidades hacen referencia a tablas que contienen diferentes atributos que las definen. Para la construcción de esta base de datos se han tenido que definir varias entidades y la relación entre ellas. La primera entidad definida es la del usuario que contiene los datos referentes a cada usuario. Esta entidad está compuesta por los atributos: ID usuario (autonumeric integer) (clave primaria), nombre de usuario (text), contraseña (text), nombre (text) y apellidos (text). La siguiente entidad definida es la del paciente que está compuesta por: ID paciente (integer) (clave primaria), ID usuario (integer) (clave foránea), edad (integer), sexo (text) y centro (text). La siguiente entidad definida es la de las respuestas del cuestionario. Esta entidad está compuesta por: ID cuestionario (integer) (clave primaria), ID paciente (integer)

(clave foránea), ID usuario (integer) (clave foránea), respuesta 1 (text), respuesta n (text), fecha prueba (text). La siguiente entidad es el tipo de test serológico y está relacionada con el resultado de la serología. Está compuesta por: ID test (integer) (clave primaria), nombre (text), tipo (text) y fabricante (text). La última entidad hace referencia al resultado de la serología y está compuesta por: ID serología (integer) (clave primaria), ID cuestionario (integer) (clave foránea), ID paciente (integer) (clave foránea), ID usuario (integer) (clave foránea), ID test (integer) (clave foránea), imagen (blob) y lote (text) (Ilustración 15).

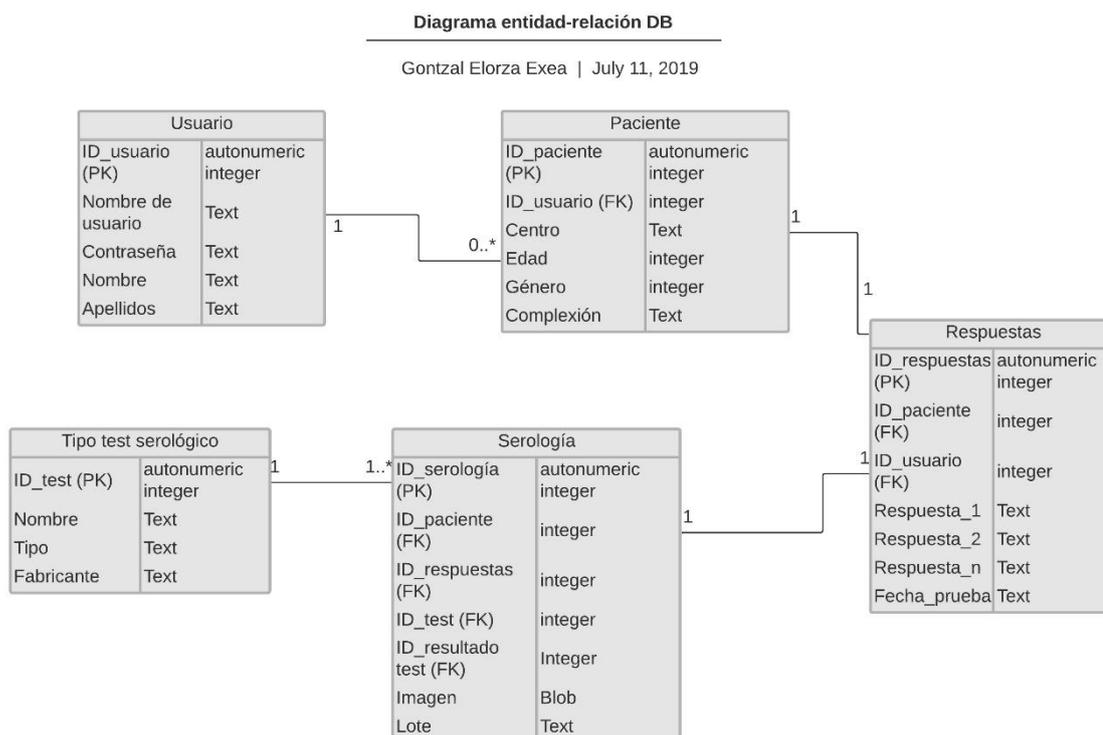


Ilustración 15 Diagrama de entidad – relación de la base de datos (Fuente propia)

Hay diferentes tipos de relaciones entre entidades:

- Uno a uno (1:1): Un registro se relaciona con otro registro. Por ejemplo, un cuestionario sólo puede estar relacionado con un paciente, y un paciente solo puede estar relacionado con un cuestionario.
- Uno a varios (1:N): Un registro se relaciona con cero o más registros de otra entidad. Por ejemplo, un usuario puede haber registrado cero o más pacientes, pero un paciente sólo puede haber sido registrado por un usuario.
- Varios a varios (N:M): Una entidad puede estar relacionada con cero o más entidades y viceversa.

2.4.2 CUESTIONARIO

El cuestionario que se deberá emplear para realizar el estudio y la aplicación ha sido realizado por profesionales del ámbito gastrointestinal y se nos ha proporcionado por la Dra. Ribes del Hospital Universitari i Politècnic La Fe de Valencia. Este cuestionario cuenta con 16 preguntas, 10 de ellas destinadas a los padres o tutores del niño que se someterá al test, 3 de ellas destinadas al profesional de la salud que realizará el test y las 3 últimas referentes al resultado de la tira reactiva. Estas son las preguntas que se deberán incluir en el cuestionario:

Destinadas a los padres o tutores:

1. ¿Sufre dolor abdominal durante periodos superiores a 3 semanas? (Si, No, NS/NC)
2. ¿Tiene el abdomen distendido? (Si, No, NS/NC)
3. ¿Esta estreñido a menudo y no responde a laxantes? (Si, No, NS/NC)
4. ¿Tiene diarrea durante periodos superiores a dos semanas? (Si, No, NS/NC)
5. ¿Sufre vómitos durante periodos superiores a 3 semanas? (Si, No, NS/NC)
6. ¿Considera que se cansa más rápido de lo normal? (Si, No, NS/NC)
7. ¿Tiene llagas en la boca con regularidad? (Si, No, NS/NC)
8. ¿Tiene la piel irritada con regularidad? (Si, No, NS/NC)
9. ¿Come gluten? (Si, No, NS/NC)
10. ¿Ha sido diagnosticado con enfermedad celiaca? (Si, No, NS/NC)

Destinadas al profesional de la salud:

11. ¿Está mermado el crecimiento (peso o altura)? (Si, No, NS/NC)
12. ¿Han dado los padres consentimiento para participar? (Si, No, NS/NC)
13. ¿Está el consentimiento escrito disponible?

Referentes al test serológico:

14. ¿Se ha realizado el test? (Si, No, NS/NC)
15. Razón por la que no se ha realizado el test: (Texto libre)
16. Resultado del test: (Positivo, Negativo, No válido)
17. ¿Derivar paciente al gastroenterólogo? (Si, No, NS/NC)

2.4.3. FLUJO DE TRABAJO DE LOS PROFESIONALES DE LA SALUD

Se ha definido el flujo de trabajo de los profesionales de la salud encargados en utilizar este sistema. En primer lugar, se define el flujo de trabajo con los diferentes componentes físicos (paciente, móvil, tira reactiva) y, en segundo lugar, se detalla el flujo inicial propuesto para la aplicación móvil a desarrollar. Los diagramas que se muestran a continuación son meramente informativos y tienen como finalidad explicar el diseño del sistema que se está proponiendo. No tienen finalidad de ser empleados para producción o testeo.

El primer paso que deberá realizar el usuario es preparar la tira reactiva. Para ello, deberá seguir las instrucciones de cada una de las tiras, ya que el tratamiento de la muestra de sangre varía entre ellas. Este debe de ser el primer paso ya que las tiras reactivas pueden llegar tardar 10 minutos hasta mostrar los resultados. Una vez tomada la muestra de sangre y preparada la tira reactiva el usuario deberá coger el móvil, abrir la aplicación y en caso de que ya esté registrado, iniciar sesión. Después, creará un nuevo paciente con los datos demográficos de este (edad, sexo y centro donde se realiza el cribado) y procederá a realizar el cuestionario

preestablecido para este proyecto. El cuestionario está compuesto por preguntas relacionadas con síntomas de la enfermedad celiaca e ingesta del gluten. Una vez rellenado el cuestionario y la tira reactiva muestre un resultado, esta deberá ser fotografiada. Con esto concluirá el estudio del primer paciente y podrá proceder a estudiar al siguiente (Ilustración 16).

DIAGRAMA DE SECUENCIA NIVEL ALTO

Gontzal Elorza Exea | July 11, 2019

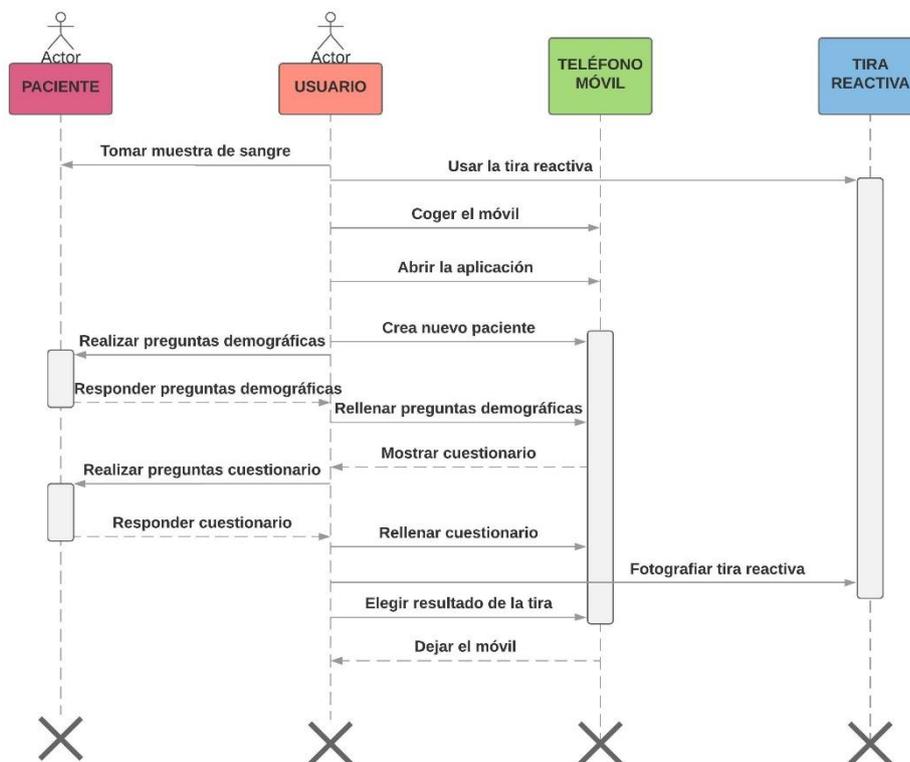


Ilustración 16 Diagrama de secuencia de componentes físicos (Fuente propia)

2.4.4. FLUJO DE LA APLICACIÓN MÓVIL

Se ha definido el flujo de pantallas dentro de la aplicación móvil con el fin de facilitar el proceso de desarrollo de esta. Para ello se han planteado una serie de casos de uso que se han englobado en un solo diagrama. Los diagramas que se muestran a continuación son meramente informativos y tienen como finalidad explicar el diseño del sistema que se está proponiendo. No tienen finalidad de ser empleados para producción o testeo.

La primera pantalla que será visible una vez se abre la aplicación es la pantalla de inicio de sesión. En esta pantalla se dispondrá de dos opciones. Por un lado, si el usuario no se ha registrado, tendrá que ir a la pantalla de registro. Por otro lado, si el usuario está registrado, deberá insertar sus credenciales y accederá a la pantalla principal.

La pantalla de registro sirve para registrar un nuevo paciente y una vez se registra se abre la pantalla de inicio de sesión.

La pantalla principal también dispondrá de dos opciones. Por un lado, se podrá iniciar el registro de un nuevo paciente. Por otro lado, se podrá consultar la base de datos para ver los datos referentes a los pacientes registrados.

La pantalla de registro de nuevo paciente dispondrá de ciertos campos a rellenar que serán los campos demográficos del paciente y al rellenar estos datos se podrá abrir la siguiente pantalla referente al cuestionario.

La pantalla del cuestionario estará compuesta por las preguntas descritas en el apartado 2.4.2. Una vez respondidas estas preguntas se podrá avanzar a la siguiente pantalla referente a la toma de la foto de la prueba serológica.

La pantalla referente a la foto de la prueba serológica da la opción de sacar una foto. Una vez tomada la foto esta será visible en la pantalla y se dará la opción de guardarla. Una vez se guarde la foto se finalizará el proceso de creación de paciente y se redireccionará al usuario a la pantalla principal. Durante las pantallas de creación de paciente, cuestionario y foto de la prueba serológica se podrá volver hacia atrás, pero directamente a la pantalla principal.

La pantalla de consulta de la base de datos mostrará los datos de los pacientes almacenados hasta el momento.

Toda esta información ha sido recogida y se ha creado un diagrama de secuencia para facilitar el proceso de desarrollo de la aplicación (Ilustración 17).

DIAGRAMA DE SECUENCIA FLUJO DE LA APLICACIÓN

Gonzal Elorza Exea | July 11, 2019

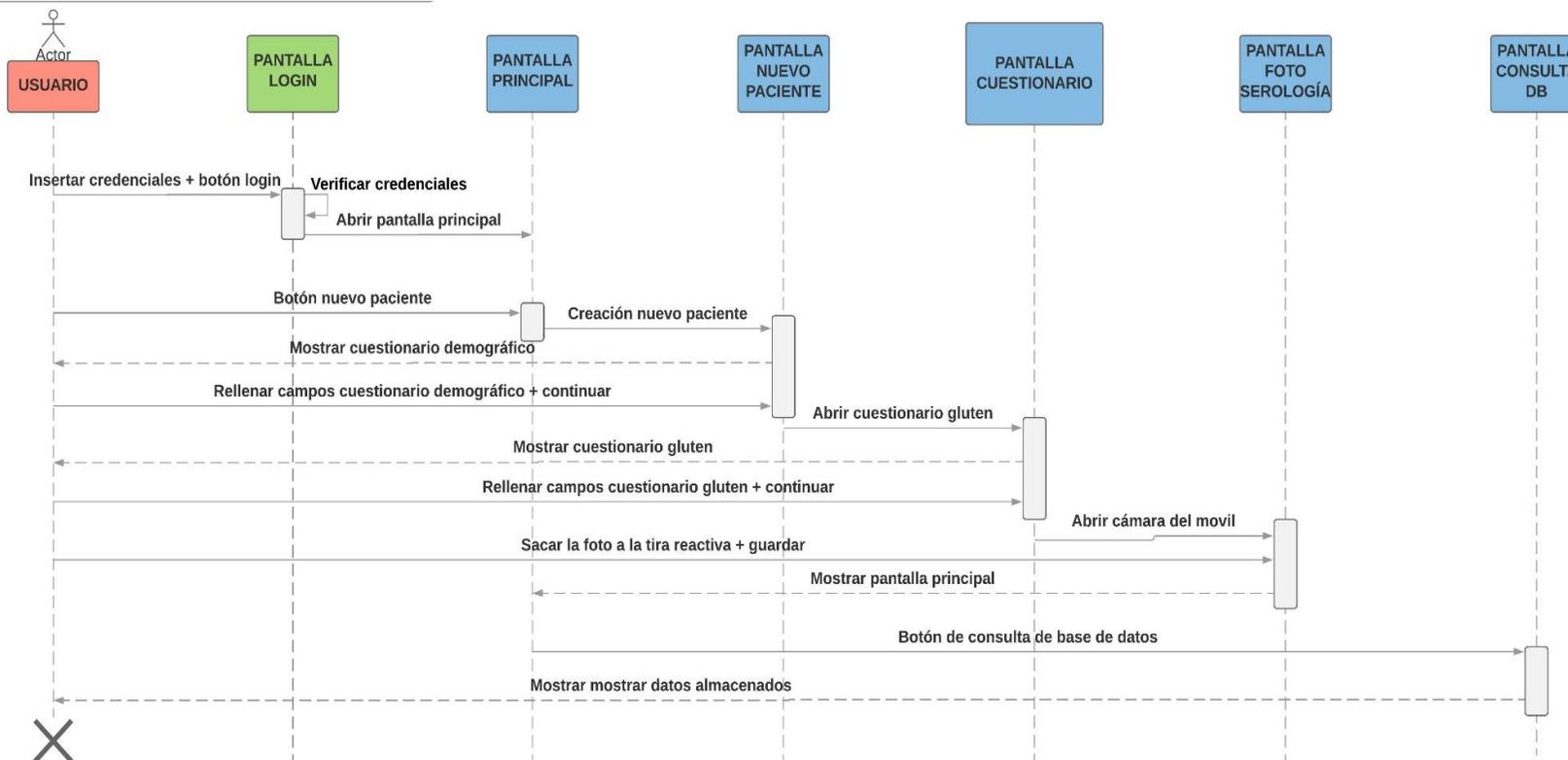


Ilustración 17 Diagrama de secuencia del flujo de la aplicación (Fuente propia)

2.4.5. REQUISITOS FUNCIONALES DE LA APLICACIÓN MÓVIL

Se han definido un total de 6 requisitos funcionales que la aplicación debe cumplir para poder verificar y validar su correcto funcionamiento. Para ello, se han creado 6 tablas (Tablas 4, 5, 6, 7, 8 y 9), una para cada requisito funcional. Cada una incluye:

- Identificador: Un código identificador para poder asociar el requisito durante la validación.
- Descripción: Breve explicación sobre el requisito.
- Entradas: Campos de entrada en el requisito funcional.
- Precondiciones: Requisitos para que la función se desempeñe de manera correcta.
- Postcondiciones: Requisitos para verificar que la función se haya desempeñado de manera correcta

Tabla 4 Requisito funcional 1: Registrar usuario (Fuente propia)

Registrar usuario	
Identificador	RF_1
Descripción	Para poder utilizar la aplicación un usuario debe registrarse introduciendo sus datos personales
Entradas	Nombre de usuario, contraseña, nombre y apellidos
Precondiciones	<ul style="list-style-type: none"> - El nombre de usuario empleado para el registro no puede coincidir con uno existente - Se deben rellenar todos los campos de entrada
Postcondiciones	<ul style="list-style-type: none"> - Se deben registrar los datos del usuario en la base de datos - La aplicación debe mostrar un mensaje de información sobre el proceso (Éxito/Fracaso) - Se debe abrir la pantalla de inicio de sesión

Tabla 5 Requisito funcional 2: Iniciar sesión (Fuente propia)

Iniciar sesión	
Identificador	RF_2
Descripción	Para acceder al menú principal de la aplicación el usuario debe iniciar sesión con sus credenciales
Entradas	Nombre de usuario y contraseña
Precondiciones	<ul style="list-style-type: none"> - El nombre de usuario y la contraseña introducidos deben coincidir con los de un usuario registrado - Se deben rellenar todos los campos de entrada
Postcondiciones	<ul style="list-style-type: none"> - Se debe abrir el menú principal - La aplicación muestra un mensaje de información sobre el proceso (Éxito/Fracaso)

Tabla 6 Requisito funcional 3: Registrar paciente (Fuente propia)

Registrar paciente	
Identificador	RF_3
Descripción	Una vez el usuario haya iniciado sesión, puede empezar a registrar pacientes para realizar el proceso de cribado
Entradas	Centro donde se realiza el cribado, edad y sexo
Precondiciones	<ul style="list-style-type: none"> - Se deben rellenar todos los campos de entrada - El campo de edad debe ser un valor numérico - Se debe haber iniciado sesión
Postcondiciones	<ul style="list-style-type: none"> - Se deben registrar los datos del paciente en la base de datos - La aplicación debe mostrar un mensaje de información sobre el proceso (Éxito/Fracaso) - Se debe abrir la pantalla del cuestionario

Tabla 7 Requisito funcional 4: Rellenar cuestionario (Fuente propia)

Rellenar cuestionario	
Identificador	RF_4
Descripción	Después de cada registro de paciente, el usuario debe rellenar el cuestionario referente a la EC
Entradas	Respuesta 1, respuesta 2, respuesta 3, respuesta 4, respuesta 5, respuesta 6, respuesta 7, respuesta 8, respuesta 9, respuesta 10, respuesta 11, respuesta 12, respuesta 13, respuesta 14, respuesta 15, respuesta 16 y respuesta 17
Precondiciones	<ul style="list-style-type: none"> - Se deben rellenar todos los campos de entrada menos el de la respuesta 15 - El paciente al que hace referencia el cuestionario debe haber sido registrado
Postcondiciones	<ul style="list-style-type: none"> - Se deben registrar los datos del cuestionario en la base de datos - La aplicación debe mostrar un mensaje de información sobre el proceso (Éxito/Fracaso) - Se debe abrir la pantalla de la cámara

Tabla 8 Requisito funcional 5: Tomar foto (Fuente propia)

Tomar foto	
Identificador	RF_5
Descripción	Una vez el cuestionario haya sido respondido, se debe tomar la foto de la tira reactiva
Entradas	Foto de la tira reactiva
Precondiciones	<ul style="list-style-type: none"> - Se debe haber rellenado el cuestionario - La foto tomada debe mostrar el resultado de la tira reactiva de manera clara
Postcondiciones	<ul style="list-style-type: none"> - Se deben registrar los datos la foto en la base de datos - La aplicación debe mostrar un mensaje de información sobre el proceso (Éxito/Fracaso) - Se debe abrir el menú principal

Tabla 9 Requisito funcional 6: Consulta de base de datos (Fuente propia)

Consulta de base de datos	
Identificador	RF_6
Descripción	Después de haber iniciado sesión el usuario puede realizar una consulta de los datos referentes a los pacientes guardados en la base de datos interna
Entradas	Ninguna
Precondiciones	<ul style="list-style-type: none"> - Se debe haber iniciado sesión - Debe haber algún paciente registrado
Postcondiciones	<ul style="list-style-type: none"> - Se deben mostrar los pacientes registrados y los datos de su cuestionario

CAPÍTULO 3: RESULTADOS

Los resultados de este proyecto están divididos en dos bloques principales. Por un lado, están los resultados referentes a la aplicación móvil desarrollada. Por otro lado, están los resultados referentes al modelo de clasificación de imagen construido.

3.1. APLICACIÓN MÓVIL

Se ha desarrollado una aplicación móvil en el sistema operativo Android con versión objetivo de Android 9.0 (nivel de API 28). La aplicación diseñada cuenta con 8 pantallas diferentes y diversas funcionalidades que se han dividido en los casos de uso definidos en el capítulo anterior. La aplicación desarrollada cuenta con una base de datos interna SQLite para el almacenamiento y gestión de los datos referentes a los usuarios y los pacientes registrados.

3.1.1. BASE DE DATOS

Para la creación de la base de datos SQLite se ha seguido el diagrama de entidad relación planteado en el capítulo anterior. Se ha utilizado la librería nativa de Android Room Persistency para facilitar el proceso y se han creado las diferentes tablas conforme a lo definido en el diagrama.

Se ha decidido eliminar una de las entidades definidas en el diagrama anterior (Ilustración 11) y mantener la misma estructura con unas ligeras modificaciones. La tabla eliminada ha sido la de “Tipo de test serológico”. Se ha construido un modelo que identifica directamente el tipo de test empleado por lo que la tabla no aportaba valor añadido. Se han eliminado varias claves foráneas (FK, por sus siglas en inglés) de la estructura. En la estructura final cada tabla sólo tiene una FK. Además, se ha eliminado el atributo “lote” de la tabla “Serología” y el atributo “complexión” de la tabla “Paciente”. De esta manera, la base de datos resultante contiene menos tablas y menos atributos por tabla (Ilustración 18).

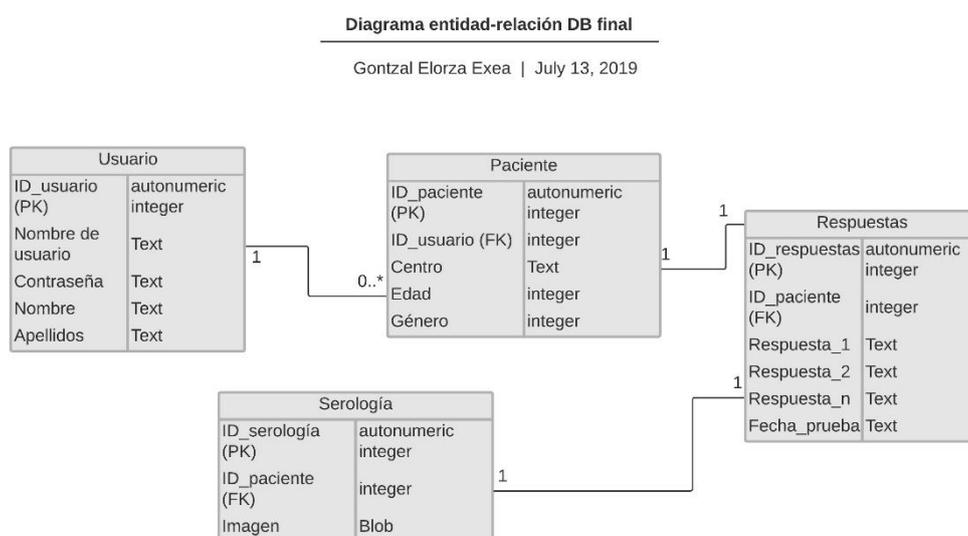


Ilustración 18 Diagrama entidad – relación final de la base de datos (Fuente propia)

3.1.2. DISEÑO DE LA APLICACIÓN

La aplicación móvil ha sido desarrollada utilizando el entorno de desarrollo integrado Android Studio. Está compuesta por 8 pantallas diferentes definidas en el capítulo anterior por las que se permite navegar al usuario para realizar el registro de nuevos usuarios, pacientes y pruebas. Se han empleado colores oscuros y pantalla completa vertical para el diseño de la aplicación. Se ha bautizado la aplicación con el nombre de “EuroGlutenScreen” haciendo referencia al proyecto europeo del que nace la idea de este trabajo. Para crear las diferentes pantallas diseñadas y gestionar sus elementos se han empleado fragmentos. Por ello cada vez que en este apartado se hace mención a una pantalla se hará referencia concretamente a un fragmento.

La primera pantalla que es visible al acceder a la aplicación es la de inicio de sesión. Esta pantalla cuenta con dos campos de texto editable que hacen referencia al nombre de usuario y a la contraseña para poder iniciar sesión. Además, hay un elemento de tipo card para iniciar sesión y un visor de texto para registrar el usuario. El elemento card no es un botón como tal, pero permite incluir funcionalidades a nivel de diseño (forma no rectangular) y a la vez que funcione como un botón. En la parte superior de la pantalla, se puede observar el logo de la aplicación con el nombre “EuroGlutenScreen” que es visible en todas las pantallas de la aplicación (Ilustración 19).

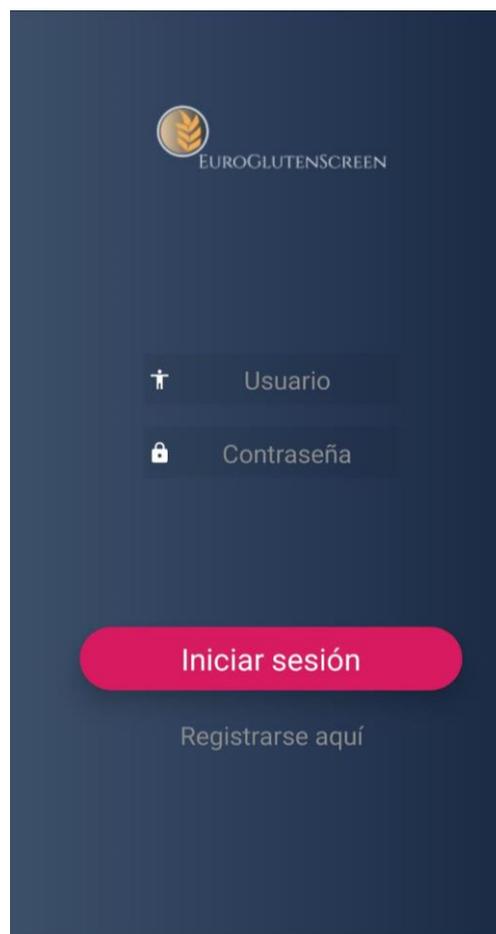


Ilustración 19 Pantalla de inicio de sesión (Fuente propia)

En caso de que el visor de texto “Registrarse aquí” sea pulsado, se avanzará a la pantalla de creación de usuario. Si el elemento tipo card con el texto “Iniciar sesión” es pulsado se recogen los datos de los dos campos de texto editable y se consulta la base de datos. En el caso de que el nombre de usuario y la contraseña coincidan con los de un usuario registrado, se avanzará a la pantalla del menú principal. Por el contrario, si el nombre de usuario o la contraseña no coinciden con el de algún usuario registrado, aparecerá un aviso en la parte inferior de la pantalla informando que alguno de los campos es erróneo (Ilustración 20).

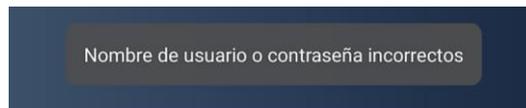


Ilustración 20 Aviso de identificación de usuario errónea (Fuente propia)

La pantalla de creación de nuevo usuario contiene cuatro campos de texto editable, cada uno de ellos referente a un atributo de la tabla usuario (nombre de usuario, contraseña, nombre y apellidos). Además, contiene un elemento de tipo card con el texto “Crear usuario” (Ilustración 21).

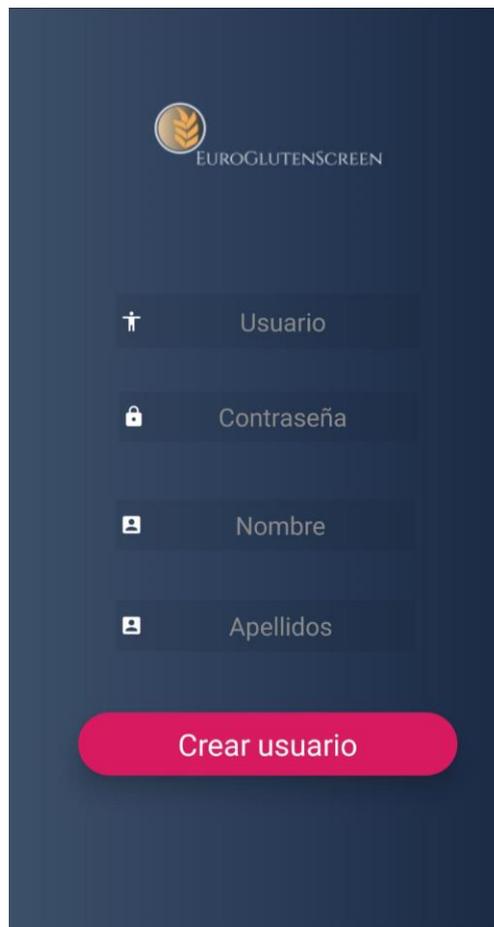


Ilustración 21 Pantalla de creación de nuevo usuario (Fuente propia)

Al pulsar el elemento de tipo card se realiza una consulta a la base de datos con el fin de detectar un nombre de usuario duplicado. Si el nombre de usuario escrito es existente, aparece un aviso indicándolo (Ilustración 22). Por el contrario, si el nombre de usuario escrito no existe en la base

de datos, se creará una nueva instancia de tipo usuario en la base de datos, se abrirá la pantalla de inicio de sesión y aparecerá un aviso de que el usuario ha sido creado satisfactoriamente (Ilustración 23).

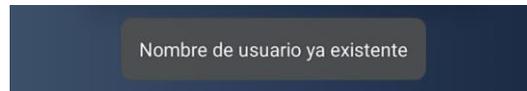


Ilustración 22 Aviso de nombre de usuario existente (Fuente propia)

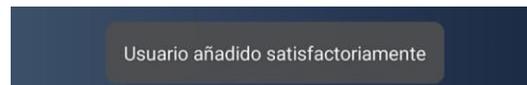


Ilustración 23 Aviso de creación de usuario satisfactoria (Fuente propia)

La pantalla del menú principal cuenta con un visor de texto que saluda al usuario por su nombre y apellidos, y dos elementos de tipo card, uno para crear un nuevo paciente, y otro para consultar la base de datos (Ilustración 24).

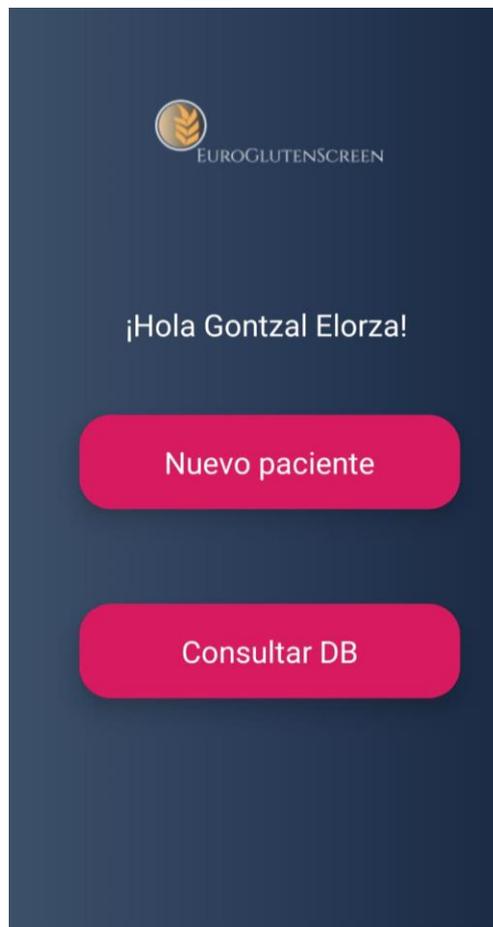


Ilustración 24 Pantalla del menú principal (Fuente propia)

En el caso de que el botón del nuevo paciente sea pulsado, se abre la pantalla de creación de nuevo usuario. En cambio, si se pulsa el botón de consultar DB se abrirá la pantalla de consulta de base

de datos. El elemento de visor de texto muestra el nombre y los apellidos introducidos en la pantalla de registro de usuario por lo que esta pantalla debe recibir el usuario que ha iniciado sesión.

La pantalla de creación de nuevo usuario está compuesta por dos elementos de texto editable, un elemento de tipo spinner en modo desplegable y un elemento de tipo card. Estos campos hacen referencia a los atributos de la tabla de paciente de la base de datos. El campo de “Centro” es de texto libre, el campo de “Edad” sólo deja insertar valores numéricos y el elemento spinner hace referencia al sexo y en el caso de que sea pulsado, mostrará las opciones masculino, femenino o no binario (Ilustración 25).

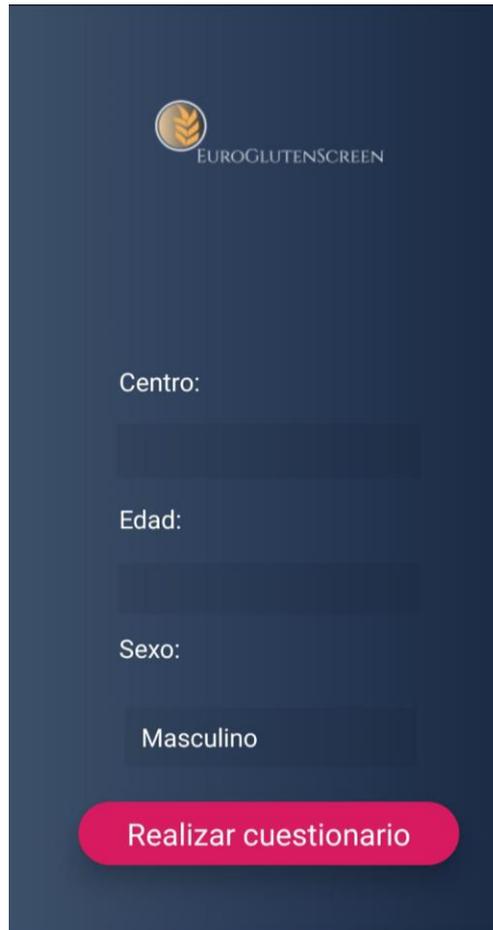


Ilustración 25 Pantalla de creación de nuevo paciente (Fuente propia)

En el caso de que el elemento tipo card con el texto “Realizar cuestionario” sea pulsado y alguno de los elementos de texto editable estén vacíos se mostrará un aviso indicando que faltan valores por rellenar en la parte inferior de la pantalla (Ilustración 26). Por el contrario, si todos los campos tienen algún valor, se creará una nueva instancia de tipo paciente en la base de datos aparecerá un aviso indicando que el paciente ha sido creado satisfactoriamente y se avanzará a la pantalla del cuestionario.

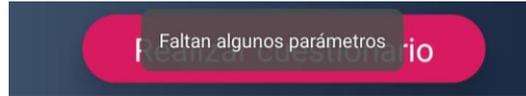


Ilustración 26 Aviso de campos nulos en creación del paciente (Fuente propia)

Para la creación de la nueva instancia de paciente, ha sido necesario habilitar la comunicación entre fragmentos, para así poder determinar qué usuario es el que está creando el nuevo paciente y poder rellenar el atributo de ID de usuario de la tabla de paciente (Ilustración 15).

La pantalla de cuestionario, ha sido diseñada utilizando un tipo de disposición diferente al resto. Esta pantalla es más extensa en cuanto a elementos que el resto por lo que se ha tenido que emplear una disposición de tipo ScrollView para poder ir navegando por la pantalla. Esta pantalla dispone de 17 visores de texto que muestran cada una de las preguntas del cuestionario especificado en el capítulo anterior. Además, para poder responder a las preguntas del cuestionario conforme a las respuestas descritas en el capítulo anterior se han empleado 16 elementos spinner en modo desplegable y un elemento de texto editable. El spinner que responde a la pregunta 15 despliega las opciones: positivo, negativo y no válido. Mientras, el resto de spinners despliegan las opciones: sí, no y NS/NC. Por último, dispone de un elemento tipo card (Ilustración 27).



Ilustración 27 Pantalla de cuestionario con disposición ScrollView (Fuente propia)

Al pulsar sobre el elemento card con el texto “Tomar foto de serología” se abrirá la pantalla de toma de foto, se guardarán los datos de todos los spinner y el editor de texto en una nueva instancia de tipo cuestionario en la base de datos. Para poder determinar el paciente al que corresponde el cuestionario, se ha posibilitado la comunicación entre fragmentos para recibir el paciente creado en la pantalla anterior, extraer su ID de paciente y almacenarla en la nueva instancia de cuestionario.

La pantalla de la cámara es la última en el flujo del proceso de la creación de un nuevo paciente y está formada por dos botones, y un visor de imágenes (Ilustración 28).

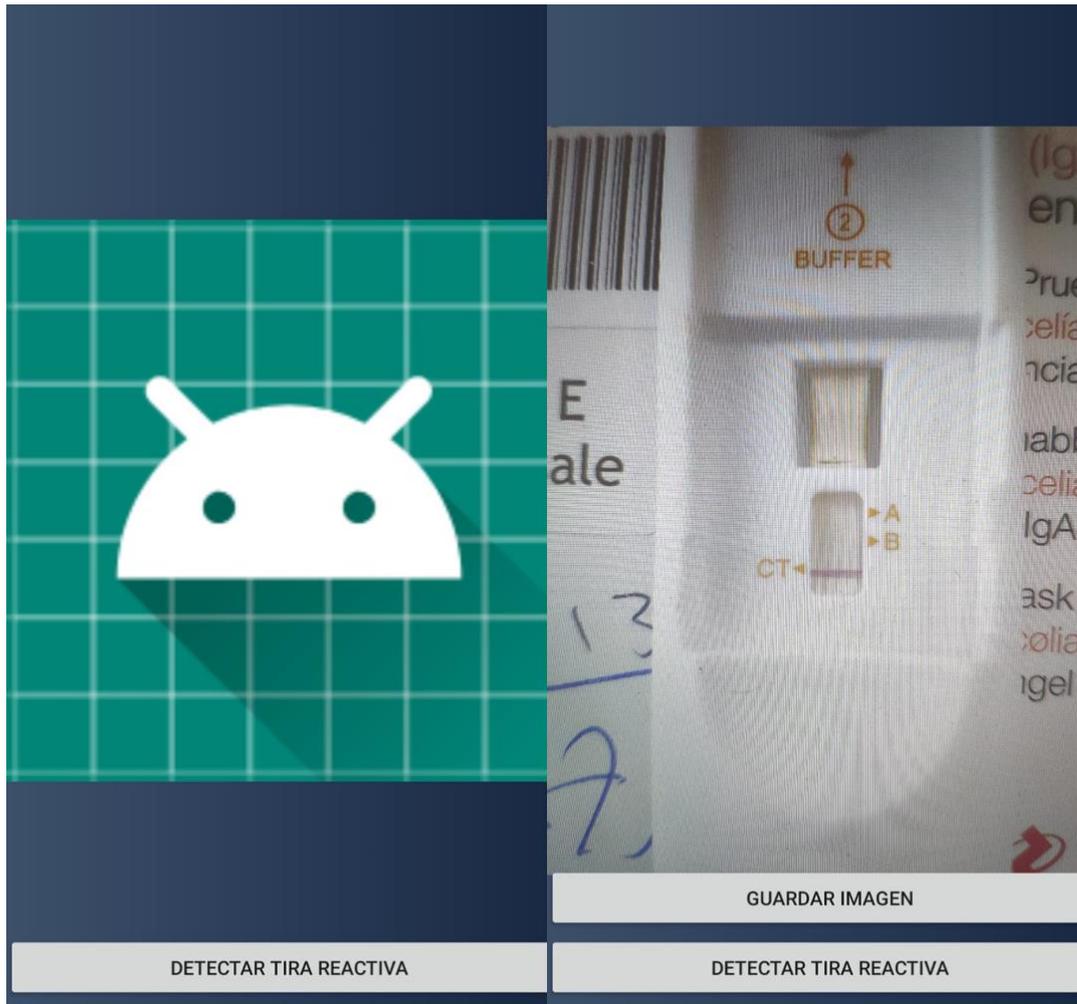


Ilustración 28 Pantalla de la cámara antes y después de tomar una foto (Fuente propia)

Al crearse esta pantalla sólo será visible el botón con el texto de “Detectar tira reactiva”. Una vez se pulse este botón, si es la primera vez que se está registrando un paciente desde que se ha instalado la aplicación, se solicitará permiso para acceder a la cámara y al almacenamiento interno del dispositivo (Ilustración 29). En el caso de que se acepten los permisos, se abrirá la cámara del dispositivo y permitirá tomar una foto. Una vez tomada la foto, se cerrará la cámara y se volverá a la pantalla de la cámara. El visor de imágenes mostrará la foto recién tomada y esta vez el botón con el texto de “Guardar imagen” será visible. En el caso de que se desee tomar una nueva imagen, se podrá pulsar de nuevo el botón “Detectar tira reactiva” y se repetirá al proceso previamente explicado. En el caso de que se pulse el botón “Guardar imagen” se almacenará la imagen visible en el visor de imagen en una nueva instancia de tipo foto de la base de datos, se abrirá de nuevo el menú principal, y aparecerá un aviso indicando la foto ha sido almacenada satisfactoriamente y que se ha terminado de realizar el cribado al paciente (Ilustración 30). Para poder enlazar la foto de la tira reactiva con el paciente creado, se ha posibilitado la comunicación entre fragmentos al igual que en la pantalla del cuestionario.



Ilustración 29 Solicitud de permisos para acceder a los archivos del dispositivo y tomar fotos (Fuente propia)

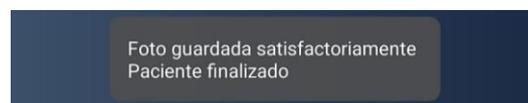


Ilustración 30 Aviso de guardado de foto satisfactoria y finalización de cribado (Fuente propia)

Si se pulsa el botón físico back en las diferentes pantallas de la aplicación hace volver al usuario a la pantalla anterior. Sin embargo, en las pantallas posteriores a la creación del paciente (cuestionario y cámara) el usuario ya está creado y almacenado en la base de datos, por lo que, si se pulsa este botón, se deberá borrar el paciente creado en la pantalla anterior. Con el fin de no tener más pacientes que cuestionarios, más cuestionarios que fotos y pacientes y cuestionarios sin una foto asignada, si se pulsa el botón físico back en la pantalla de cuestionario, aparecerá una ventana emergente indicando que si se vuelve atrás se perderán los datos referentes a este paciente (Ilustración 31). En el caso de que se acepte, se volverá al menú principal, se eliminará el paciente creado en la pantalla anterior y aparecerá un aviso indicándolo (Ilustración 32). En el caso de que no se acepte, no ocurrirá nada. Si se pulsa el botón de back en la pantalla de la cámara, ocurrirá lo mismo, sólo que esta vez si se acepta la ventana emergente, se eliminarán el paciente y el cuestionario creados recientemente.

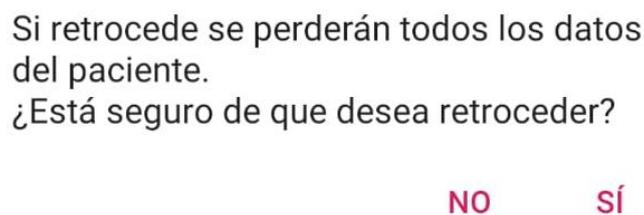


Ilustración 31 Ventana emergente de eliminación de paciente (Fuente propia)

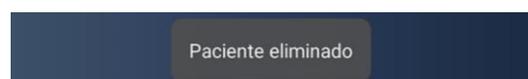


Ilustración 32 Aviso de eliminación de paciente (Fuente propia)

La pantalla de consulta de la base de datos muestra los datos almacenados referentes a los pacientes y sus respectivos cuestionarios. La dimensión de esta pantalla dependerá de la cantidad de

pacientes que haya almacenados. La pantalla está compuesta por un visor de texto, que se adapta a los datos disponibles en la base de datos. Primero, aparecen los datos demográficos e identificativos del paciente (ID, edad, centro, sexo e ID usuario). Después aparecerán las preguntas del cuestionario y sus respuestas (Ilustración 33).

Id : 1	No
Edad : 8	¿Está estreñado a menudo y no responde a laxantes?
Centro : Colegi Públic Federico García Lorca	No
Sexo : Femenino	¿Tiene diarrea durante periodos superiores a dos semanas?
Creado por usuario (id): 1	No
Fecha realización del cuestionario : domingo, 14 de julio de 2019	¿Sufre vómitos durante periodos superiores a 3 semanas?
¿Sufre dolor abdominal durante más de 3 semanas?	No
No	¿Considera que se cansa más rápido de lo normal?
¿Tiene el abdomen distendido?	Sí
No	¿Tiene llagas en la boca con regularidad?
¿Está estreñado a menudo y no responde a laxantes?	No
No	¿Tiene la piel irritada con regularidad?
¿Tiene diarrea durante periodos superiores a dos semanas?	Sí
No	¿Come gluten?
¿Sufre vómitos durante periodos superiores a 3 semanas?	Sí
No	¿Ha sido diagnosticado con enfermedad celíaca?
¿Considera que se cansa más rápido de lo normal?	No
Sí	¿Está mermado el crecimiento? (peso o altura)
¿Tiene llagas en la boca con regularidad?	No
No	¿Han dado los padres consentimiento para participar?
¿Tiene la piel irritada con regularidad?	Sí
Sí	¿Está el consentimiento escrito disponible?
¿Come gluten?	Sí
Sí	¿Se ha realizado el test POC?
¿Ha sido diagnosticado con enfermedad celíaca?	Sí
No	Razón para no realizar test:
¿Está mermado el crecimiento? (peso o altura)	Resultado test:
	Negativo
	¿Derivar paciente a gastroenterólogo?
	No

Ilustración 33 Pantalla de consulta de base de datos con datos referentes a un paciente (Fuente propia)

De esta manera, queda concluida la descripción del diseño de la aplicación y su funcionamiento. Se ha construido un archivo .apk para probar la aplicación en otros dispositivos con resultados satisfactorios. El peso total de la aplicación es de 1,8 MB. La aplicación no se ha subido a ningún servicio de descarga de aplicaciones como Google Play, ya que se trata de un prototipo.

3.1.3. IMPLEMENTACIÓN DE LA APLICACIÓN

Para poder crear la aplicación diseñada, ha sido necesario crear 13 clases java, 2 archivos tipo interfaz y 8 archivos XML de los cuales 7 son fragmentos y uno es la actividad principal. La aplicación está construida utilizando fragmentos en vez de actividades. De las 13 clases, una es la actividad principal, 7 definen la actividad de fragmentos, una define la base de datos y 4 son tablas de la base de datos.

La aplicación requiere de permisos para acceder al almacenamiento interno del dispositivo y para abrir la cámara.

Se han empleado varias librerías para el desarrollo de esta aplicación. Por un lado, se ha tenido que emplear la ya mencionada librería Room Persistence para poder crear la base de datos SQLite. Room Persistence ayuda a crear una caché de los datos de la aplicación en un dispositivo que esté ejecutándola. El funcionamiento de esta librería es simple. Para poder crear una base de datos, son necesarios tres elementos. En primer lugar, es necesario crear una clase de tipo abstracto que defina la base de datos y sus tablas. Después, se debe crear una interfaz llamada objeto de acceso a la base de datos, donde se definen las operaciones que se pueden hacer dentro de la base de datos (inserciones, eliminaciones, consultas...). Por último, es necesario crear una clase por tabla que se desee incluir en la base de datos. Dentro de esta clase, se deben definir los atributos por tabla y el tipo de dato que representan. Por otro lado, se han tenido que incluir las librerías:

- Cardview: Para insertar elementos tipo card (incluir botones de formas no cuadradas).
- ConstraintLayout: Para incluir disposiciones de tipo restricción.

La aplicación se ha construido sobre fragmentos en vez de sobre actividades por lo que la función de la actividad principal, es simplemente albergar estos fragmentos. Las clases de los fragmentos hacen referencia a cada una de las pantallas descritas en el apartado anterior. Dentro de ellas se detalla la actividad descrita en cada una de las acciones.

Al trabajar con fragmentos, se ocupa menos memoria en el dispositivo, sin embargo, algunas funcionalidades son difíciles de implementar. Para ello, se debe emplear la actividad principal para poder habilitar ciertas funcionalidades en los fragmentos. Un caso de este tipo es la comunicación de variables entre fragmentos. Un ejemplo de esta acción es cuando se recibe en el menú principal al usuario que ha iniciado sesión. Para poder llevar a cabo esta acción se debe escribir código en los 2 fragmentos que se comunican y en la actividad principal. En primer lugar, se debe crear un listener en el fragmento del que se deseen enviar los datos. Este listener llamará a una función de la actividad principal en la que se guardarán los datos recibidos. A su vez, en esta función se llamará a otra función en la que se reciban los datos en el fragmento de destino (Ilustración 34). En el caso del usuario que ha iniciado sesión, se manda un elemento de clase usuario con todos los datos que conlleva (id, nombre de usuario, contraseña, nombre y apellidos). Una vez se recibe el usuario, se extrae el nombre y los apellidos para poder mostrarlos en la pantalla del menú principal.

Fragmento Inicio de sesión

```
public interface Listener1{
    void onInput1sent (User input);
}
```

. . .

```
User input = user;
listener.onInput1sent(input);
```

Actividad principal

```
public void onInput1sent(User input) {

    mainmenu.updateUser(input);
    newpatient.updateUser(input);
    cuestionario.updateUser(input);
    camara.updateUser(input);

}
```

Fragmento Menú principal

```
public void updateUser ( User user1){
    name1=user1.getName();
    surname1=user1.getSurname();
}
```

Ilustración 34 Parte del código necesario para comunicar fragmentos (Fuente propia)

Este también es el caso de la funcionalidad descrita en las pantallas de cuestionario y cámara en la que, si se pulsa el botón back físico del móvil, sale una ventana emergente. Las actividades, tienen una función llamada OnBackPressed para esta acción, pero los fragmentos no. Por ello, se ha creado una interfaz que actúe de manera de enlace entre la actividad y el fragmento. De esta manera, cuando se pulse este botón desde alguna de estas ventanas se ejecutará un código creado por el usuario, en este caso se abrirá una ventana emergente (Ilustración 35).

Fragmento

```
public boolean onBackPressed() {
    String si = "Si";
    String no = "No";
    DialogInterface.OnClickListener dialogClickListener = (dialog, which) -> {
        switch (which){
            case DialogInterface.BUTTON_POSITIVE:
                int longitud;
                longitud = MainActivity.MyDB.myDao().getRowCount();
                List<Patient> patientlist = MainActivity.MyDB.myDao().getPatient();
                patient1 = patientlist.get(longitud-1);
                MainActivity.MyDB.myDao().deletePatient(patient1);
                dialog.cancel();
                dialog.dismiss();
                Toast.makeText(getActivity(), text: "Paciente eliminado", Toast.LENGTH_SHORT).show();
                hacerback();
                break;
            case DialogInterface.BUTTON_NEGATIVE:
                dialog.cancel();
                dialog.dismiss();
                break;
        }
    };
    AlertDialog.Builder altdial = new AlertDialog.Builder(getActivity());
    altdial.setMessage("Si retrocede se perderán todos los datos del paciente.\n" +
        "¿Está seguro de que desea retroceder?").setCancelable(false)
        .setPositiveButton(si, dialogClickListener)
        .setNegativeButton(no, dialogClickListener).show();
    return false;
}
```

Interfaz

```
public interface IOnBackPressed {
    boolean onBackPressed();
}
```

Actividad principal

```
public void onBackPressed(){
    Cuestionario cuestionario = (Cuestionario) getSupportFragmentManager().findFragmentByTag("cuestionario1");
    Camara camara = (Camara) getSupportFragmentManager().findFragmentByTag("camara1");
    if (cuestionario == null && camara == null)
    {
        super.onBackPressed();
    }
    else {
        if (cuestionario.isAdded() && cuestionario.isVisible() && cuestionario.getUserVisibleHint() && camara == null) {
            if (cuestionario.onBackPressed()) {
                super.onBackPressed();
            }
        }
        else {
            if (camara.isAdded() && camara.isVisible() && camara.getUserVisibleHint()){
                if (camara.onBackPressed()) {
                    super.onBackPressed();
                }
            }
            else{
                super.onBackPressed();
            }
        }
    }
}
```

Ilustración 35 Código necesario para realizar acciones con 45nter back en fragmentos (Fuente propia)

Las imágenes tomadas se almacenan en la base de datos interna de la aplicación. Para que esto sea posible, se deben almacenar primero en la memoria interna del dispositivo.

3.2. VALIDACIÓN DE LA APLICACIÓN MÓVIL

La validación de la aplicación desarrollada se ha llevado a cabo verificando el correcto funcionamiento de cada requisito funcional definido en el capítulo anterior (Tabla 10). Para validar el correcto funcionamiento de la aplicación en todos los dispositivos, se han incluido dispositivos de modelos diferentes.

Tabla 10 Cumplimiento de los requisitos en diferentes modelos de teléfono (Fuente propia)

	Modelo	Relación de aspecto	Versión	RF_1	RF_2	RF_3	RF_4	RF_5	RF_6
Móvil 1	Xiaomi mi 8	19:9	9.0	✓	✓	✓	✓	✓	✓
Móvil 2	Xiaomi mi A2	18:9	9.0	✓	✓	✓	×	×	✓
Móvil 3	Huwei P20 lite	19:9	9.0	✓	✓	✓	×	×	✓
Móvil 4	Pixel 3A	18:9	9.0	✓	✓	✓	✓	✓	✓
Móvil 5	Xiaomi mi A1	18:9	9.0	✓	✓	×	×	×	✓

El problema que presentan los móviles 2 y 3 está relacionado con la disposición de la propia aplicación. Todos los elementos están desplazados a la derecha y arriba por lo que algunos de ellos no son visibles en la aplicación. Por ello, no es posible pulsar el botón para guardar el cuestionario, ni se puede acceder a la cámara por la misma razón. En el móvil 5 ocurre lo mismo, solo que el botón para guardar el usuario casi no es visible.

3.3. MODELO DE CLASIFICACIÓN

La creación del modelo de clasificación para las imágenes de las tiras reactivas se ha llevado a cabo utilizando la herramienta Tensorflow y el lenguaje de programación Python. Para el entrenamiento del modelo se han empleado las imágenes descritas en el capítulo anterior. Tal y como se ha mencionado en el capítulo 2.3 hay diferentes opciones a la hora de definir cómo se desea que clasifique el modelo. Por esta razón, se ha decidido realizar dos modelos con el fin de determinar cuál da mejores resultados predictivos.

3.3.1. PROCESO DE CREACIÓN DEL MODELO

La creación de los modelos de clasificación ha sido llevada a cabo con la metodología propuesta por Gilbert Tanner⁴, divulgador y científico de datos.

El proceso de creación de un modelo está dividido en varias partes. El primer paso es el de preparar el conjunto de imágenes con el que se desea entrenar al modelo. Para que todas las imágenes tengan la misma orientación y dimensiones se ha redefinido la dimensión de todas las imágenes a 800x600 píxeles. Una vez redimensionadas las imágenes, se debe realizar la partición de datos entre

⁴ <https://gilberttanner.com/2019/02/02/tensorflow-object-detection-tutorial-3-creating-your-own-object-detector/>

los datos de entrenamiento y los de testeo. Para este caso la partición de datos es de 80% para entrenamiento y 20% para test. A continuación, se debe definir el área de interés de cada imagen, especificando cuál es el resultado de esta área de interés (Ilustración 31). Para ello se ha empleado el software libre LabelImg disponible en GitHub y creada por el usuario Tzutalin. Cada imagen tendrá un archivo .xml con el mismo nombre donde se definirán las coordenadas de la zona de interés

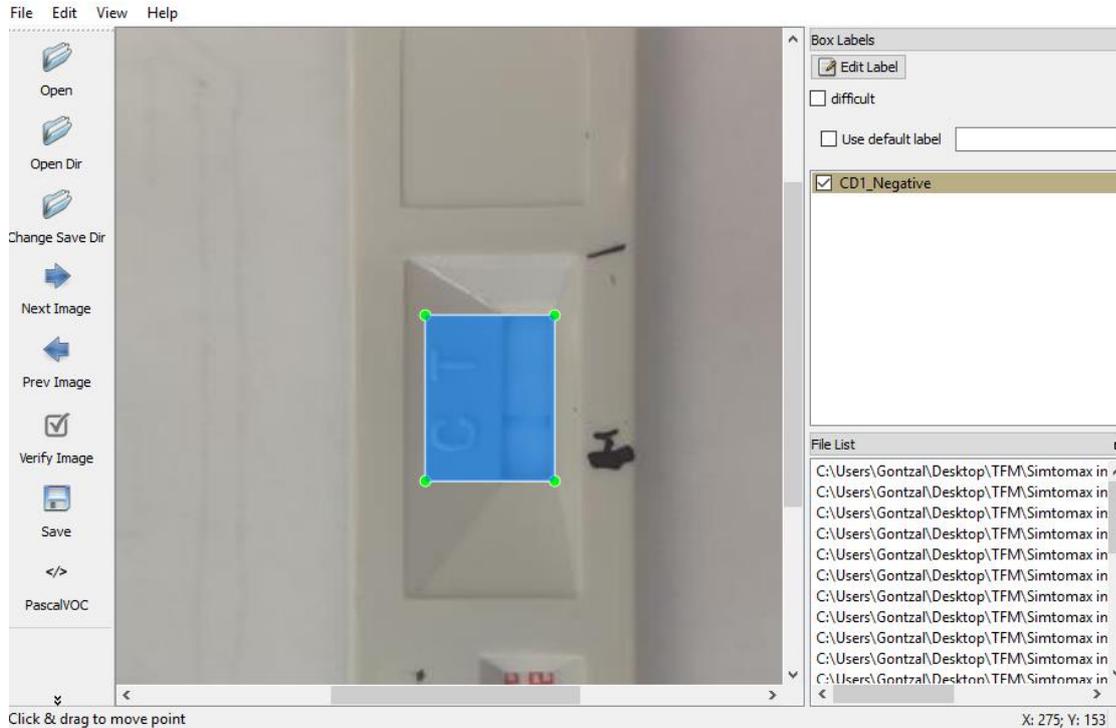


Ilustración 36 Herramienta LabelImg utilizada para delimitar el área de interés de una tira (Fuente propia)

El siguiente paso es crear dos archivos .csv que contengan la información de cada uno de los archivos, con las coordenadas de su zona de interés y su resultado. Uno estará formado por los archivos de testeo y el otro por los archivos de entrenamiento. Después se debe convertir este archivo en un registro que sea capaz de ser interpretado por la herramienta Tensorflow.

La siguiente parte de la creación es la configuración del entrenamiento. Para ello se debe crear un archivo que contenga las clases de salida del modelo que se desea construir. Además, se debe elegir el algoritmo que se desea emplear para la creación del modelo. Para este caso se empleará una CNN. Se deben modificar algunos parámetros del algoritmo de la red neuronal para que concuerden con las carpetas donde están almacenados los registros y el archivo de las clases de salida. Una vez preparada la configuración, se puede proceder a entrenar el modelo. Mientras el modelo se entrena, se puede observar la pérdida de la clasificación utilizando Tensorboard, una funcionalidad incluida en la herramienta Tensorflow (Ilustración 37). Una vez el valor de la pérdida no baje más se deberá parar manualmente el entrenamiento.

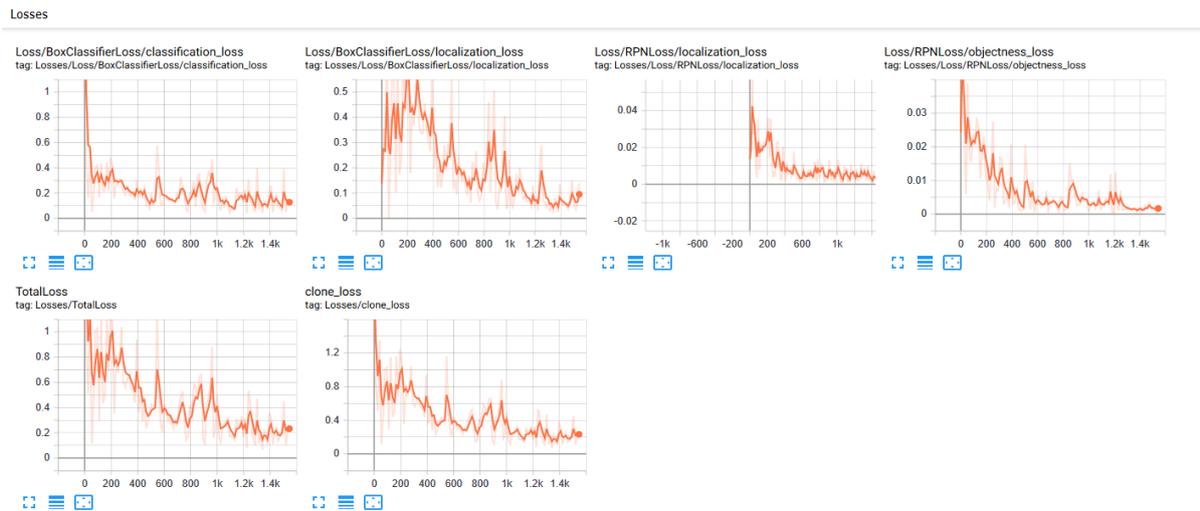


Ilustración 37 Pérdida durante el entrenamiento del modelo 2 en Tensorboard (Fuente propia)

Una vez se para el entrenamiento del modelo, se debe extraer el gráfico de inferencia congelado. El gráfico de inferencia congelado no puede volver a ser entrenado, sin embargo, es necesario que el gráfico de inferencia esté congelado para que pueda realizar predicciones con nuevos datos.

3.3.2. RESULTADOS DE LOS MODELOS

Se han creado dos modelos de clasificación siguiendo las pautas del punto anterior. La única diferencia entre estos dos modelos son la cantidad de clases de salida que tienen. El primer modelo es el más simple, tiene sólo dos salidas, mientras que el segundo modelo tiene seis.

- Clases del modelo 1: Positivo, negativo.
- Clases del modelo 2: CD1 positivo, CD1 negativo, CD2 positivo, CD2 negativo, SimtomaX positivo, SimtomaX negativo.

Con el fin de determinar el valor predictivo de los modelos, cada uno ha sido probado con 18 imágenes aleatorias (generadas mediante la aleatorización de contexto de las tiras reactivas disponibles). Se ha intentado que la muestra aleatoria de imágenes sea homogénea en cuanto a las imágenes por tipo de tira. Cada imagen es evaluada por el modelo que identifica la zona de interés y determina el resultado de esta indicando el porcentaje de certeza de la predicción (Ilustraciones 36 y 37).

Se han creado dos tablas donde se muestran los datos relacionados con cada una de las imágenes empleadas para validar el modelo (Tablas 11 y 12). La tabla muestra el resultado real del test, el resultado predicho, el porcentaje de certeza de la predicción y si la predicción ha acertado. A su vez, en la tabla referente al modelo 1, se especifica el tipo de tira de la imagen. En la tabla referente al modelo 2 esto está especificado en el resultado.

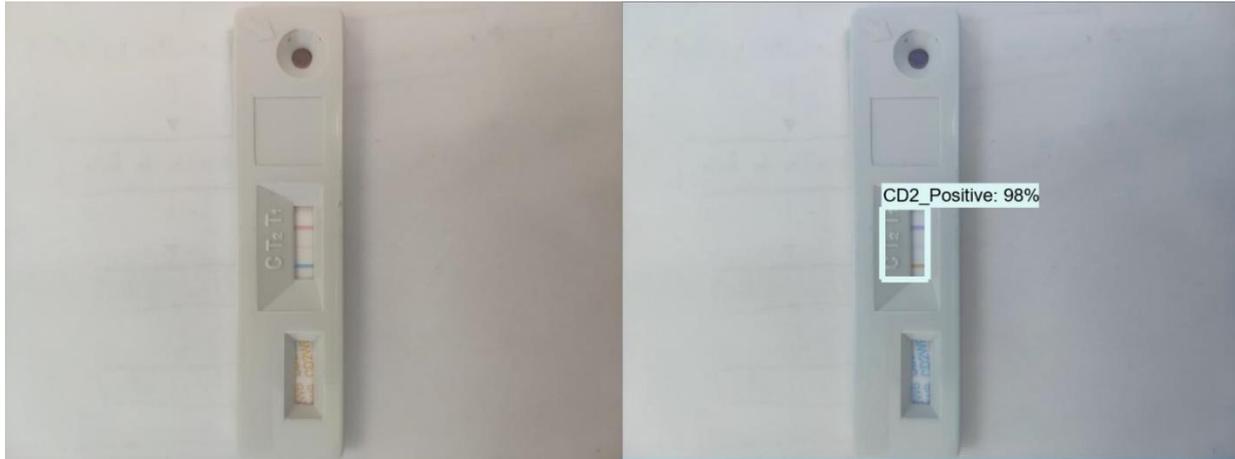


Ilustración 38 Foto de tira reactiva CD2 (izquierda) y la predicción realizada por el modelo 2 (derecha) (Fuente propia)



Ilustración 39 Foto de tira reactiva SimtomaX (izquierda) y la predicción realizada por el modelo 1 (derecha) (Fuente propia)

El modelo 1 ha predicho correctamente el resultado de 14 de las 18 imágenes que se han empleado para la validación. Se han empleado 12 tiras *SimtomaX*, 4 tiras *CD2* y 2 tiras *CD1* para la validación. El modelo ha predicho correctamente el 75% de las tiras *SimtomaX*, el 75% de las tiras *CD2* y el 100% de las tiras *CD1*. El valor predictivo negativo del modelo es del 50% mientras que el valor predictivo positivo el del 91,6%. Los valores de predicción del modelo oscilan entre el 54% y el 99%. El valor predictivo medio en las tiras en las que se ha predicho el resultado de manera correcta es del 93,7% mientras que en las tiras en las que no se ha predicho el resultado de manera correcta es del 81%. Se puede observar que el valor predictivo es mayor en tiras en las que se ha predicho el resultado de manera correcta.

Tabla 11 Resultados de la validación del modelo 1 (Fuente propia)

Modelo 1					
	Resultado	Predicción	Valor (%)	Acierto	Tira
Imagen 1	Positivo	Positivo	95	✓	SimtomaX
Imagen 2	Negativo	Negativo	54	✓	SimtomaX
Imagen 3	Negativo	Negativo	95	✓	SimtomaX
Imagen 4	Positivo	Negativo	90	×	SimtomaX
Imagen 5	Negativo	Positivo	79	×	SimtomaX
Imagen 6	Negativo	Positivo	92	×	SimtomaX
Imagen 7	Positivo	Positivo	97	✓	SimtomaX
Imagen 8	Positivo	Positivo	99	✓	SimtomaX
Imagen 9	Positivo	Positivo	99	✓	SimtomaX
Imagen 10	Positivo	Positivo	95	✓	SimtomaX
Imagen 11	Positivo	Positivo	98	✓	SimtomaX
Imagen 12	Positivo	Positivo	93	✓	SimtomaX
Imagen 13	Positivo	Positivo	99	✓	CD2
Imagen 14	Positivo	Positivo	98	✓	CD2
Imagen 15	Negativo	Negativo	92	✓	CD2
Imagen 16	Negativo	Positivo	63	×	CD2
Imagen 17	Positivo	Positivo	99	✓	CD1
Imagen 18	Positivo	Positivo	99	✓	CD1

El modelo 2 ha predicho correctamente el resultado de 14 de las 18 imágenes que se han empleado para la validación. Se han empleado 6 tiras *SimtomaX*, 6 tiras *CD2* y 6 tiras *CD1* para la validación. El modelo ha predicho correctamente el 83,3% de las tiras *SimtomaX*, el 66,7% de las tiras *CD2* y el 83,3% de las tiras *CD1*. Más concretamente se han predicho de manera correcta el 100% de las tiras *SimtomaX* con resultado positivo, el 66,7% de las tiras *SimtomaX* con resultado negativo, el 100% de las tiras *CD2* con resultado positivo, el 33,3% de las tiras *CD2* con resultado negativo, el 66,7% de las tiras *CD1* con resultado positivo y el 100% de las tiras *CD1* con valor negativo. El valor predictivo negativo del modelo es del 66,7% mientras que el valor predictivo positivo el del 88,9%. Los valores de predicción para las predicciones correctas oscilan entre el 59% y el 99% mientras que para las predicciones no correctas son entre el 60% y el 79%. El valor predictivo medio en las tiras en las que se ha predicho el resultado de manera correcta es del 76,6% mientras que en las tiras en las que no se ha predicho el resultado de manera correcta es del 68,5%. Se puede observar que el valor predictivo es mayor en tiras en las que se ha predicho el resultado de manera correcta.

Tabla 12 Resultados de la validación del modelo 2 (Fuente propia)

Modelo 2				
	Resultado	Predicción	Valor (%)	Acierto
Imagen 1	SimtomaX Positivo	SimtomaX Positivo	78	✓
Imagen 2	SimtomaX Negativo	SimtomaX Positivo	74	×
Imagen 3	SimtomaX Negativo	SimtomaX Negativo	73	✓
Imagen 4	SimtomaX Positivo	SimtomaX Positivo	64	✓
Imagen 5	CD2 Positivo	CD2 Positivo	99	✓
Imagen 6	CD2 Positivo	CD2 Positivo	95	✓
Imagen 7	CD2 Negativo	CD2 Negativo	66	✓
Imagen 8	CD2 Negativo	CD2 Positivo	79	×
Imagen 9	CD1 Positivo	CD1 Positivo	71	✓
Imagen 10	CD1 Positivo	CD1 Positivo	63	✓
Imagen 11	CD1 Negativo	CD1 Negativo	74	✓
Imagen 12	CD1 Negativo	CD1 Negativo	59	✓
Imagen 13	SimtomaX Positivo	SimtomaX Positivo	83	✓
Imagen 14	SimtomaX Negativo	SimtomaX Negativo	84	✓
Imagen 15	CD2 Positivo	CD2 Positivo	96	✓
Imagen 16	CD2 Negativo	CD2 Positivo	61	×
Imagen 17	CD1 Positivo	CD1 Negativo	60	×
Imagen 18	CD1 Negativo	CD1 Negativo	67	✓

Los dos modelos han predicho de manera correcta el mismo número de imágenes. Se puede observar, que los valores de predicción del modelo 1 son mayores que los del modelo 2.

Se ha creado una tabla comparativa de los indicadores de cada uno de los modelos con el fin de facilitar la visualización y comparación de cada uno de estos valores (Tabla 13).

Tabla 13 Comparación entre los indicadores de los modelos (Fuente propia)

	Exactitud	Tasa de error	Sensibilidad	Especificidad	Precisión	Valor predictivo negativo
Modelo 1	0,78	0,22	0,92	0,50	0,79	0,75
Modelo 2	0,78	0,22	0,89	0,67	0,73	0,86

Con todos estos datos, se ha realizado una matriz de confusión con su curva ROC para cada uno de los modelos. De esta manera, se puede obtener una representación visual del rendimiento de cada uno de los modelos.

Para la matriz de confusión del modelo 1, se han designado los resultados positivos con el valor 1 y los valores negativos con el valor 0. Se puede observar que la especificidad de este modelo con las tiras de validación es tan solo de 0,5 (Ilustración 40).

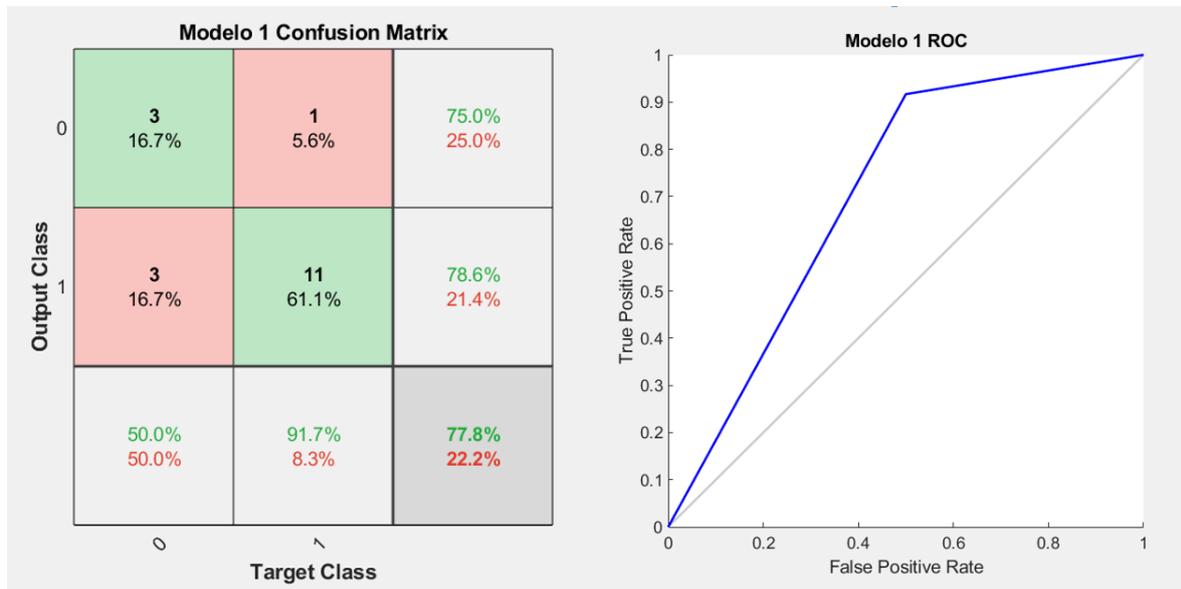


Ilustración 40 Matriz de confusión y curva ROC del modelo 1 (Fuente propia)

Para realizar la matriz de confusión del modelo 2, se han simplificado las 6 clases de salida a resultados positivos, con valor 1, y resultados negativos, con valor 0 (Ilustración 41). El hecho de que el modelo haya predicho el tipo de tira en las 18 imágenes empleadas para la validación permite realizar esta simplificación.

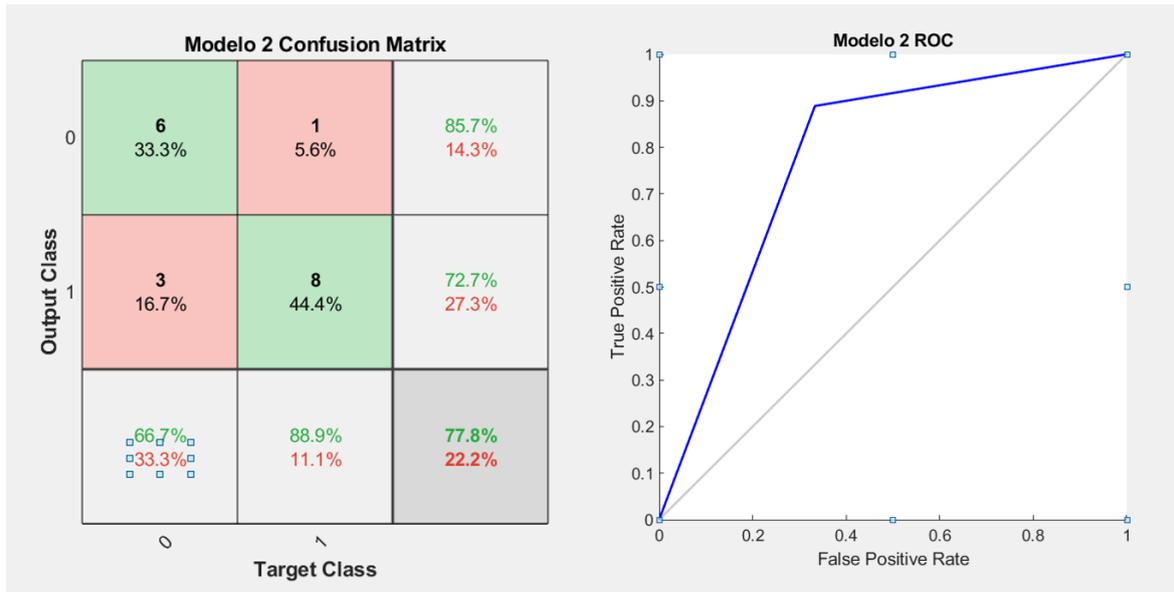


Ilustración 41 Matriz de confusión y curva ROC del modelo 2 (Fuente propia)

CAPÍTULO 4: DISCUSIÓN

En el presente TFM se han obtenido una serie de resultados que contribuyen a la creación de una herramienta para el *screening* masivo de la EC. Por un lado, se ha diseñado una aplicación móvil funcional capaz de registrar pacientes que se someten a un proceso de cribado de la EC. Por otro lado, se han creado dos modelos diferentes capaces de clasificar el resultado de las tiras reactivas mediante técnicas avanzadas de análisis de imagen. Desde un punto de vista holístico, el proyecto ha experimentado algunos cambios desde las especificaciones iniciales propuestas. Algunos de estos cambios vienen precedidos por dificultades técnicas y el proceso de aprendizaje con tecnologías desconocidas hasta el momento por parte del autor, algunos otros por mejoras que se han ido descubriendo y añadiendo durante este proceso. El caso de la aplicación web para almacenar y mostrar los datos en la nube es un claro ejemplo ya que, al sobrepasar la carga de trabajo de un TFM, se ha relegado para futuros trabajos. El caso de la base de datos es otro ejemplo de cómo el propio desarrollo del proyecto ha ayudado a comprender mejor su funcionamiento, posibilitando así la optimización de la base de datos y desarrollar competencias de análisis y propuesta de soluciones.

El proyecto sobre el que se basa este trabajo propone diseñar una herramienta que ayude en el cribado masivo de la EC. Para esto, se propone la utilización de tiras reactivas. No cabe duda de que la propuesta de realizar un *screening* masivo a la población global en edad pediátrica es un avance claro de cara a poder diagnosticar muchos de los casos no diagnosticados o que quedan por debajo del nivel del agua en el “iceberg” (Ilustración 2). En este contexto, las especificaciones del proyecto indicaban que en análisis debía hacerse mediante toma de fotografías a las tiras reactivas, sin embargo, el desarrollo de un sensor reutilizable mediante el cual se puedan realizar las mediciones de anticuerpos realizadas por las tiras reactivas podría suponer un gran avance de cara al cribado poblacional. Hoy en día, son cada vez más los biosensores que se pueden encontrar en el mercado. Los más comunes son los medidores de glucosa en sangre, que permiten monitorizar de manera continua la concentración de la glucosa. El desarrollo de un biosensor capaz de medir de manera efectiva los anticuerpos específicos de la EC presentes en una muestra de sangre sería de gran ayuda para el cribado. Por un lado, agilizaría el proceso de cribado ya que no se tendrían que preparar y almacenar los test serológicos. Por otro lado, la digitalización del proceso de cribado podría ser total, al posibilitar una comunicación de los resultados entre el teléfono móvil donde se registre al paciente y el biosensor. En cualquier caso, bien sea por toma de imágenes o sensor instrumentalizado el cribado poblacional masivo en sí puede ser de gran ayuda para los muchos pacientes no diagnosticados de la EC, ya que facilitaría el diagnóstico precoz de la enfermedad evitando así las complicaciones crónicas que puede suponer el consumo de gluten para estas personas.

Un gran dilema del diseño de la herramienta ha sido el almacenamiento de los datos referentes a los pacientes y dónde realizar la predicción de la imagen. Si los datos se almacenan y se muestran en un servidor, lo lógico es crear un modelo para PC o servidor, y si los datos se almacenan y se muestran en la propia aplicación móvil, lo lógico sería integrar el modelo en la aplicación. Sin embargo, analizando el proyecto desde un punto de vista holístico, se debe tener en cuenta cuál será la funcionalidad final de este (de cara a la aplicación real). Para poder

integrar los datos obtenidos en diferentes puntos por profesionales de la salud, el proyecto final debe incluir una base de datos centralizada. Por ello, el modelo que se ha decidido construir funcionará en PC o servidor. Aun así, el desarrollo de una aplicación web para consultar los datos obtenidos excede el tiempo disponible para la realización del trabajo por lo que se ha decidido almacenar los datos de manera local. El desarrollo de la aplicación web para la consulta de datos se realizará cuando se implante el proyecto real sobre el que se basa este trabajo, ya que se dispondrá de recursos humanos y técnicos para la implementación de un sistema distribuido.

En cuanto al desarrollo de la aplicación, se han empleado fragmentos en vez de actividades. La tendencia en la programación Android indica que los fragmentos son el futuro, sin embargo, ha habido varios problemas a la hora de implementar algunas de las funcionalidades que se tenían previstas en este tipo de clases. Emplear fragmentos facilita que la aplicación pueda ser utilizada en tabletas además de los teléfonos móviles. Además, se almacena el registro de pantallas por las que se ha pasado, por lo que es más fácil volver a la pantalla anterior. También posibilita reutilizar pantallas si es necesario. Asimismo, muchas de las funcionalidades que son simples en actividades, se han vuelto complicadas en fragmentos, como puede ser la comunicación de las variables entre pantallas. Por ello, creo que podría haber sido más conveniente utilizar actividades para el caso de esta aplicación.

Gracias al proceso de validación que se ha llevado a cabo con la aplicación, se han podido detectar varios fallos que no eran visibles en el móvil en el que se ha realizado el testeo de la aplicación durante su desarrollo. Dependiendo de la relación de las dimensiones de la pantalla del dispositivo en el que se ha instalado la aplicación, la disposición puede estar movida o no. Por ello, en un futuro se deberá posibilitar el autoajuste de la pantalla a las dimensiones de cada dispositivo. Además, no es posible instalar la aplicación en dispositivos con versiones antiguas de Android. Esto se podría arreglar definiendo la versión mínima necesaria en una más antigua, para poder dar soporte a dispositivos más antiguos. Asimismo, se ha detectado un error a la hora de girar la pantalla de los dispositivos ya que se desplazan todos los elementos. Para arreglar este problema, se debería realizar la misma acción que en el caso anterior y posibilitar el autoajuste de la pantalla.

En líneas generales, la aplicación es intuitiva, fácil de utilizar y perfectamente funcional. Puede ser una aplicación útil a la hora de agilizar el proceso de cribado, y sobre todo puede ayudar a eliminar papeleo al digitalizar el registro de los pacientes. Aun así, la aplicación tiene margen de mejora. El principal punto a mejorar de la aplicación es el acceso a un servidor en la nube. Una base de datos local no permite la interconexión entre los datos recogidos en los diferentes centros de salud y escolares. Por ello sería conveniente almacenar los datos recogidos en una base de datos en la nube porque por un lado se podrán analizar los datos recogidos de manera mucho más efectiva y, por otro lado, se podrán realizar consultas de los datos desde cualquiera de los dispositivos con la app instalada. Una vez este problema estuviese resuelto, la aplicación podría ser perfectamente funcional en el ámbito clínico. Otro punto de mejora de la aplicación es que sólo esté diseñada para dispositivos Android. Aunque en Europa la hegemonía de los dispositivos Android sea evidente, en Estados Unidos el sistema operativo más utilizado en dispositivos móviles es iOS. Para facilitar la incorporación del sistema a este tipo de países sería necesario diseñar la aplicación también para iOS.

Con respecto a los modelos de análisis de la imagen, los dos modelos propuestos presentan la misma precisión a la hora de predecir el resultado de las tiras. Sin embargo, si se analizan el resto de indicadores (Tabla 13) el modelo 2 presenta mejores valores. Aunque la especificidad de ambos modelos sea baja, la del modelo 1 es de 0,5. Cabe destacar que el modelo 2 ha predicho el tipo de tira en todas las imágenes empleadas para la validación. Esto se puede deber a que al definir el área de interés de cada una de las imágenes se ha incluido la parte de la carcasa de la tira donde están las letras de referencia (Ilustración 36) Al no tratarse de una enfermedad tan crítica como puede ser un cáncer, el error que presentan los modelos es asumible. Aun así, se han propuesto varias posibles mejoras de cara a los futuros modelos que se vayan a realizar una vez se implemente el proyecto.

Los modelos que se han construido presentan unos resultados relativamente buenos, pero no lo suficientemente buenos como para aplicarlos a una aplicación real. Una precisión del 78% para ambos modelos no es un resultado aplicable a un caso real. Sin embargo, para evaluar este resultado, se deben tener en cuenta las limitaciones que se han tenido a la hora de construir el modelo y, sobre todo, los datos de las imágenes. Por un lado, la cantidad de tiras de cada tipo con las que se han construido los modelos ha sido bastante pequeña. Se han tomado 5 imágenes distintas de cada tira en condiciones distintas con el fin de agrandar la cantidad de datos disponibles para crear el modelo, sin embargo, sería mucho más conveniente tener un set de tiras más amplio para mejorar la robustez del modelo. Por otro lado, muchas de las tiras elaboradas para el modelo están deterioradas debido al paso del tiempo. Algunas de estas muestras han sido tomadas hace más de diez años, por lo que el color de la tira transversal que debe aparecer indicando el resultado se ha atenuado de manera significativa en muchas de ellas. Disponer de tiras nuevas también mejoraría los datos disponibles para realizar el modelo. A todo esto, se le debe sumar la falta de algún tipo de resultado de tiras que ha habido. No se ha dispuesto de ninguna tira catalogada como no válida para ninguno de los 3 tipos de tira. Además, en las tiras *CD2WB* no se ha dispuesto de ningún resultado de IgA anti-gliadina (Tabla 2). Esto también podría ser mejorado disponiendo de una cantidad mayor de tiras para el estudio. Otra posible mejora del modelo es la utilización de técnicas de validación cruzada, como puede ser el K-fold. De esta manera, se mejoraría la robustez del modelo evitando el sobreajuste del modelo al set de datos de entrenamiento.

Este TFM me ha servido para desarrollar mis capacidades personales de autoaprendizaje y de organización. He adquirido conocimientos en el campo de la programación de aplicaciones, en el de machine learning y en de diseño de software y he podido aplicar uno de mis campos favoritos, como es la programación, al entorno sanitario. Previamente, no había creado ningún modelo de clasificación de imágenes ni ninguna aplicación, por lo que considero que los conocimientos que he adquirido durante estos 3 meses me han servido para hacer un trabajo del que estoy satisfecho.

4.1. LIMITACIONES

Este trabajo no está exento de limitaciones. La cantidad y la calidad de las tiras con las que se han realizado los modelos ha sido una de las principales limitaciones para entrenar y testear el modelo. En esta línea, sólo se ha trabajado con los tres tipos de tiras descritos. Se desconoce si más tipos de tiras son empleadas en Europa para realizar el cribado de pacientes con la EC. Además, solo se ha probado un modelo basado en CNN, por lo que se desconocen los

resultados que podrían ofrecer otras técnicas de Deep Learning. La aplicación móvil por su parte sólo ha sido validada en dispositivos con versión Android 9.0 y no ha sido validada por los usuarios finales de la aplicación, es decir, por los profesionales de la salud. Sólo se ha probado el diseño final de la aplicación sin tener en cuenta otros tipos de disposición de pantalla posibles. No se han probado distintos sistemas de gestión de bases de datos que podrían ser útiles de cara a la futura migración de los datos a un servidor en la nube. Otra de las limitaciones más importantes de este proyecto es que el sistema no se encuentra integrado. La aplicación y el modelo están separadas y no crean un sistema unificado por lo que el producto final de este proyecto está dividido en dos partes.

4.2. LÍNEAS FUTURAS

Se han definido una serie de trabajos futuros que se pueden llevar a cabo para continuar con la línea de este trabajo. Para poder solventar las limitaciones que han supuesto la falta de tiras reactivas, se pueden considerar dos opciones. En primer lugar, involucrar a asociaciones de pacientes con EC y centros de salud para poder obtener el mayor número de muestras posibles podría ser una opción válida para llevar a cabo un estudio. En segundo lugar, diseñar un protocolo mediante el cual se centralice el almacenamiento de los datos referentes a las tiras y a los pacientes celíacos el futuro acceso a los datos necesarios podría ser mucho más fácil. En cuanto al desarrollo de un sistema totalmente funcional en el ámbito clínico, se debe crear una base de datos centralizada para el almacenamiento de los datos referentes al cribado. Para ello, por un lado, se debe habilitar un servidor en la nube para el almacenamiento de estos datos y, por otro lado, se debe habilitar la comunicación entre los dispositivos con la aplicación instalada y el propio servidor. Además, una vez se disponga de las tiras necesarias para realizar un modelo robusto para la clasificación de las imágenes de las tiras, se podrán analizar diferentes maneras de abordar mejorar los resultados de este, como puede ser el empleo de herramientas para la validación cruzada. Una vez se disponga de un modelo robusto y una aplicación conectada a un servidor en la nube se podrá considerar que el sistema para el cribado masivo y sistematizado de la EC está completo y preparado para cumplir todas las funciones que necesite cumplir, pudiendo ser empleado en estudios clínicos y epidemiológicos de gran escala.

CAPÍTULO 5: CONCLUSIONES

En este capítulo, se revisan los objetivos marcados al comienzo de este proyecto. Primero, se analizará el cumplimiento de los objetivos secundarios propuestos. Finalmente, se analizará el cumplimiento del objetivo principal.

Se ha estudiado la EC centrándose el estudio en los preocupantes datos sobre su epidemiología y diagnóstico. Se ha podido corroborar que el surgimiento de un proyecto para llevar a cabo el *screening* masivo y sistematizado está totalmente justificado.

Se ha diseñado y desarrollado una aplicación totalmente funcional, capaz de recolectar información sobre los pacientes, sobre los cuestionarios que responden y sobre las tiras reactivas.

Se ha implementado y validado la aplicación en diferentes teléfonos móviles para demostrar su funcionamiento correcto y concordancia con el diseño propuesto. Se ha obtenido buen resultado con la mayoría de los teléfonos móviles, pero se debe tener cuidado con la versión del sistema operativo.

Se han implementado y validado dos modelos predictivos para automatizar el reconocimiento del resultado de las diferentes pruebas serológicas empleadas en el cribado. Se han detectado carencias en la manera en la que se han construido los modelos, pero se han propuesto diversas mejoras para poder crear un modelo robusto aplicable a la práctica clínica.

Teniendo en cuenta esto, se considera que se ha diseñado y desarrollado un sistema capaz de facilitar el proceso de cribado para la enfermedad celíaca en centros de salud o escolares. Aunque el sistema no cubra todas las necesidades de un sistema aplicable a la práctica clínica, se han definido las pautas necesarias para poder llevarlo a cabo. Este sistema puede considerarse un primer acercamiento válido para abordar un el problema que supone la baja tasa de diagnóstico de la EC.

CAPÍTULO 6: BIBLIOGRAFÍA

- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Retrieved from <http://cds.cern.ch/record/998831>
- Briani, C., Samaroo, D., & Alaedini, A. (2008). Celiac disease: From gluten to autoimmunity. *Autoimmunity Reviews*, 7(8), 644–650. <https://doi.org/10.1016/j.autrev.2008.05.006>
- Butterworth, J., & Los, L. (2019). Coeliac disease. *Medicine (United Kingdom)*, 47(5), 314–319. <https://doi.org/10.1016/j.mpmed.2019.02.003>
- Catassi, C., Fabiani, E., Ratsch, I.-M., Coppa, G., Giorgi, P. L., Pierdomenico, R., ... Ventura, A. (1996). The coeliac iceberg in Italy. A multicentre antigliadin antibodies screening for coeliac disease in school-age subjects. *Acta Paediatrica*, 85, 29–35. <https://doi.org/10.1111/j.1651-2227.1996.tb14244.x>
- Comisión europea. (2014). *Green Paper on mobile health*. <https://doi.org/10.5437/08956308X5704235>
- de Sousa Moraes, L. F., Grzeskowiak, L. M., de Sales Teixeira, T. F., & Gouveia Peluzio, M. do C. (2014). Intestinal Microbiota and Probiotics in Celiac Disease. *Clinical Microbiology Reviews*, 27(3), 482 LP – 489. <https://doi.org/10.1128/CMR.00106-13>
- Fasano, A., & Catassi, C. (2001). Current approaches to diagnosis and treatment of celiac disease: An evolving spectrum. *Gastroenterology*, 120(3), 636–651. <https://doi.org/10.1053/gast.2001.22123>
- Google. (2019a). Introducing Convolutional Neural Networks. Retrieved July 11, 2019, from ML Practicum: Image Classification website: <https://developers.google.com/machine-learning/practica/image-classification/convolutional-neural-networks>
- Google. (2019b). Platform Architecture.
- Green, P. H. R., & Cellier, C. (2007). Celiac disease. *The New England Journal of Medicine*, 357(17), 1731–1743. <https://doi.org/10.1056/NEJMra071600>
- Green, P. H. R., Jabri, B., & Kasarda, D. D. (2005). Innate and adaptive immunity : the Yin and Yang of celiac disease. *Immunological Reviews*, 206, 219–232.
- Guandalini, S. (2008). Historical perspective of celiac disease. *Pediatric and Adolescent Medicine*, 12, 1–11. <https://doi.org/10.1159/000128267>
- John M. Eisenberg Center for Clinical Decisions and Communications Science, J. M. E. C. for C. D. and C. (2007). Diagnosis of Celiac Disease: Current State of the Evidence. *Comparative Effectiveness Review Summary Guides for Clinicians*, 108(5). Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/27583323>
- Kelly, C. P., Bai, J. C., Liu, E., & Leffler, D. A. (2015). Advances in diagnosis and management of celiac disease. *Gastroenterology*, 148(6), 1175–1186. <https://doi.org/10.1053/j.gastro.2015.01.044>

- Martinez-Millana, A., Jarones, E., Fernandez-Llatas, C., Hartvigsen, G., & Traver, V. (2018). App Features for Type 1 Diabetes Support and Patient Empowerment: Systematic Literature Review and Benchmark Comparison. *JMIR MHealth and UHealth*, 6(11), e12237. <https://doi.org/10.2196/12237>
- Mustalahti, K., Catassi, C., Reunanen, A., Fabiani, E., Heier, M., McMillan, S., ... Maki, M. (2010). The prevalence of celiac disease in Europe: results of a centralized, international mass screening project. *Annals of Medicine*, 42(8), 587–595. <https://doi.org/10.3109/07853890.2010.505931>
- Riznik, P., De Leo, L., Dolinsek, J., Gyimesi, J., Klemenak, M., Koletzko, B., ... Dolinsek, J. (2019). Diagnostic Delays in Children with Coeliac Disease in the Central European Region. *Journal of Pediatric Gastroenterology and Nutrition*. <https://doi.org/10.1097/MPG.0000000000002424>
- Singh, K., & Landman, A. B. (2017). Chapter 13 - Mobile Health. In A. Sheikh, K. M. Cresswell, A. Wright, & D. W. Bates (Eds.), *Key Advances in Clinical Informatics* (pp. 183–196). <https://doi.org/https://doi.org/10.1016/B978-0-12-809523-2.00013-3>
- Tutorialspoint. (n.d.). SQLite tutorial. Retrieved July 11, 2019, from <https://www.tutorialspoint.com/sqlite/index.htm>

2. PRESUPUESTO

1. INTRODUCCIÓN

La realización de un presupuesto de un trabajo es fundamental a la hora de analizar la viabilidad económica de cualquier proyecto. Este Trabajo Final de Master, es un trabajo centrado en la investigación y el estudio del presupuesto se hace con el fin de comprender el coste que ha supuesto llevarlo a cabo. El análisis presupuestario de este proyecto se divide en dos puntos: los presupuestos parciales y el presupuesto final.

2. PRESUPUESTOS PARCIALES

Se han realizado dos presupuestos parciales. Por un lado, se ha realizado el presupuesto referente al personal necesario para realizar el presente trabajo. Por otro lado, se ha realizado el presupuesto referente al material físico y software necesario para realizar el presente trabajo.

2.1. COSTE DE PERSONAL

Se ha calculado el coste que supone las horas empleadas por cada uno de las personas implicadas en el desarrollo de este trabajo. Por un lado, están el tutor del trabajo, Antonio Martínez Millana, y el cotutor, Álvaro Fides Valero, ambos investigadores senior en el departamento ITACA-SABIEN. Por otro lado, está la Dra. Carmen Ribes, gastroenteróloga en el Hospital Universitari i Politècnic La Fe. Por último, está el alumno encargado de realizar el TFM, Gontzal Elorza. Se ha considerado que el autor del TFM es un investigador junior.

Tabla de presupuesto 1 Coste del personal

Personal	Uds.	Cantidad	Coste unitario (€)	Coste total (€)
Tutor del TFM (Investigador senior)	h	30	25,00	750,00
Cotutor del TFM (Investigador senior)	h	10	25,00	250,00
Dra. Ribes	h	10	45,00	450,00
Autor del TFM (Estudiante)	h	400	15,00	6.000,00
			TOTAL	7.450,00

2.2. COSTE DE MATERIAL

Se ha calculado el coste del material necesario para la realización de este trabajo. Dentro de este material se ha incluido también el software empleado.

Tabla de presupuesto 2 Coste de material

Material	Uds	Cantidad	Coste unitario (€)	Coste total (€)
Ordenador Asus Essentid	€	1	420,00	420,00
Teléfono móvil Xiaomi mi 8	€	1	329,00	329,00
Teléfono móvil Xiaomi mi A2	€	1	189,00	189,00
Cable USB tipo C	€	2	8,99	17,98
Monitor PC Philips 193V5LSB2	€	1	69,99	69,99
Office Profesional 2019	€	1	579,00	579,00
			TOTAL	1.604,97

3. PRESUPUESTO GENERAL

Para llevar a cabo el cálculo del presupuesto total del presente TFM, se han sumado los dos presupuestos parciales. A esta cifra debe sumársele un 20% de gastos indirectos relacionados con el puesto de trabajo (electricidad, puesto de trabajo en el CPI, ...). A todo ello se le debe aplicar el 21% de IVA para poder obtener el resultado del coste que supone la realización de este Trabajo de Final de Master.

Tabla de presupuesto 3 Presupuesto general

CATEGORÍA	COSTE (€)
Coste de personal	7.450,00
Coste de materiales	1.604,97
TOTAL DIRECTO	9.054,97
20 % de gastos indirectos	1.810,99
TOTAL (SIN IVA)	10.865,96
21 % de IVA	2.281,85
PRESUPUESTO TOTAL	13.147,81

