
This is the **published version** of the bachelor thesis:

Caravaca Hernández, Álvaro; Gil Resina, Debora, dir.; Cano Ribé, Pau, dir.
Detecció del bacteri *Helicobacter pylori*. 2023. (Enginyeria Informàtica)

This version is available at <https://ddd.uab.cat/record/280731>

under the terms of the  license

Histopatologia Digital.

Detecció del bacteri *Helicobacter pylori*.

Álvaro Caravaca Hernández

1 de juliol de 2023

Resum— En aquest treball s'ha desenvolupat un sistema per al diagnòstic de la presència d'H. Pylori mitjançant imatges d'histologia dels pacients. S'han explorat dos paradigmes diferents. En el primer, s'ha utilitzat la xarxa neuronal VGG-16 per extreure les característiques de les imatges i s'han entrenat diversos models de classificació per discernir la presència o absència d'H. Pylori en els pacients. En el segon paradigma, s'ha explorat l'ús d'un autoencoder per a la detecció d'anomalies en les mostres dels pacients, contribuint al diagnòstic basat en la presència d'anomalies detectades. Els resultats demostren que l'autoencoder és més efectiu en el diagnòstic global dels pacients.

Paraules clau— *Helicobacter Pylori*, Histologia, Xarxes Neuronals, VGG, Models de classificació, Ensemble, Autoencoder.

Abstract— In this work, a system has been developed for the diagnosis of the presence of H. Pylori through histology images of patients. Two different approaches have been explored. In the first one, the VGG-16 neural network has been used to extract the characteristics of the images and several classification models have been trained to discern the presence or absence of H. Pylori in the patients. In the second approach, the use of an autoencoder for the detection of anomalies in patient samples has been explored, contributing to the diagnosis based on the presence of anomalies detected. The results show that the autoencoder is more effective in the overall diagnosis of patients.

Keywords— *Helicobacter Pylori*, Histology, Neural Networks, VGG, Classification Models, Ensemble, Autoencoder.

1 INTRODUCCIÓ

EL bacteri *Helicobacter Pylori* és la principal causa de la gastritis, una inflamació de la mucosa gàstrica que pot conduir a altres malalties greus, com l'úlcera gàstrica i el càncer gàstric. Per tant, la detecció precoç d'aquest bacteri és fonamental per al diagnòstic i el tractament eficaç d'aquestes patologies. A més, estudis afirmen que més del 50% de la població mundial ha estat infectada pel bacteri, amb una prevalença que supera el 80% en adults de més de cinquanta anys [1]. Això posa de manifest la importància de desenvolupar mètodes de detecció eficients per a una millor detecció de la infecció. Actualment, el diagnòstic de la presència d'*Helicobacter Pylori* es realitza a través de diferents mètodes, entre els quals destaquen les proves serològiques, la prova ràpida de la ureasa i l'anàlisi histològica. En aquest treball s'han utilitzat imatges d'histologia ob-

tingudes a partir d'una gastroscòpia amb biòpsia gàstrica, on s'ha aplicat una tècnica de tinció immunohistoquímica per detectar la presència del bacteri. Aquestes imatges han estat proveïdes per l'Hospital Universitari General de Catalunya-Grupo Quirónsalud.

La gastroscòpia amb biòpsia gàstrica és una tècnica utilitzada per identificar la infecció per H. pylori i altres patologies relacionades amb la inflamació de la mucosa gàstrica. Una tècnica emprada en aquest procés és la tinció immunohistoquímica, la qual permet la visualització del bacteri a través de la tinció d'una proteïna específica present a la seva membrana. Les imatges obtingudes a través de la tinció immunohistoquímica són revisades per un patòleg amb l'objectiu de detectar la possible presència d'H. Pylori en la mostra. Aquesta tinció provoca que la proteïna específica de la membrana del bacteri es tingeixi d'una tonalitat vermellosa, mentre que la resta de teixit es manté en una tonalitat blava. L'ús d'aquesta tinció facilita la detecció del bacteri, permetent als patòlegs identificar la possible presència d'H. pylori a les mostres, tot i que augmenta el cost i el temps necessari per dur a terme la

-
- E-mail de contacte: alvaro.caravaca@autonoma.cat
 - Menció realitzada: Computació
 - Treball tutoritzat per: Débora Gil Resina i Pau Cano Ribé
 - Curs 2022/23

prova.

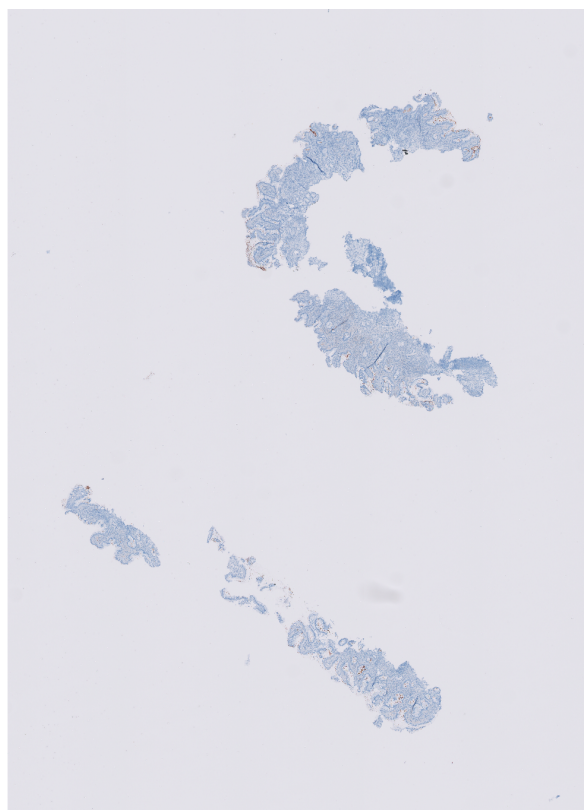


Fig. 1: Mostra d'histologia amb tinció immunohistoquímica

A la figura 1 es pot identificar la presència de diferents tonalitats de colors després d'aplicar la tinció immunohistoquímica en una imatge d'histologia obtinguda a partir d'una gastroscòpia amb biòpsia gàstrica.

Així i tot, la detecció d'H. Pylori de manera visual feta per un patòleg és costosa perquè la mostra del pacient és molt gran. És per això que recentment s'han començat a desenvolupar tècniques basades en intel·ligència artificial que pretenen reduir el cost i el temps necessari per a aquests diagnòstics. En aquest context, aquest treball té com a finalitat contribuir a la detecció de l'H. Pylori en imatges d'histologia usant intel·ligència artificial.

2 ESTAT DE L'ART

Existeix una àmplia diversitat de tipus d'imatges mèdiques utilitzades per a la detecció d'H. Pylori. Mentre alguns estudis opten per usar imatges d'histologia per detectar la presència del bacteri, d'altres prefereixen imatges obtingudes a través de tècniques com l'endoscòpia gàstrica. En els estudis que fan servir imatges d'histologia, aquestes s'aconsegueixen a partir de biòpsies gàstriques [2] [3]. És important assenyalar que només un d'aquests estudis [3] fa servir la tinció immunohistoquímica per a la detecció d'H. Pylori. D'altra banda, els estudis que no fan servir imatges d'histologia obtenen les seves imatges mitjançant l'endoscòpia gàstrica [4] [5] [6].

Per afrontar la detecció del bacteri, s'han explorat dues tècniques de processament d'imatges diferents: les

tècniques clàssiques i les xarxes neuronals.

Les tècniques clàssiques de processament d'imatges es basen en algorismes i mètodes tradicionals que han estat desenvolupats i refinats al llarg del temps. Les xarxes neuronals són una branca de l'aprenentatge automàtic que ha revolucionat la forma en què les màquines processen i analitzen les dades. Els models de xarxes neuronals estan basats en la interconnexió de neurones artificials, que intenten imitar a les neurones del cervell humà. Aquestes xarxes poden aprendre patrons complexos i generar prediccions o classificacions a partir de les dades d'entrada.

Un dels estudis [4] fa ús de la tècnica de Support Vector Machine (SVM) per al diagnòstic d'H. Pylori, sense recórrer a l'ús de xarxes neuronals convolucionals (CNN). Aquest enfocament ha obtingut una sensibilitat del 83.9%. Cal remarcar que aquest estudi ha utilitzat un conjunt de dades altament equilibrat, amb un 45% de mostres negatives i un 55% de mostres positives en H. Pylori. Aquesta característica és notablement diferent dels altres estudis consultats, els quals presenten conjunts de dades molt desequilibrats, amb un nombre molt més gran de mostres negatives. En el cas més desequilibrat, s'arriba a un 75.7% de mostres negatives i un 24.3% de mostres positives en H. Pylori. Els altres estudis consultats [2] [3] [5] [6] s'han centrat en l'ús de xarxes neuronals convolucionals preentrenades per al diagnòstic d'H. Pylori. En tots aquests estudis, les mostres dels pacients són utilitzades com a entrada, mentre que la sortida és una probabilitat que indica la presència del bacteri. Mitjançant l'aplicació de xarxes neuronals convolucionals, s'han obtingut nivells de sensibilitat que oscil·len entre el 81.9% i el 90.0%.

Per tant, es pot identificar que s'ha abordat aquesta investigació des de diversos punts de vista, utilitzant diferents tipus d'imatges i diferents tècniques mèdiques, i que han aconseguit resultats prometedors fent servir tècniques d'aprenentatge profund, en particular les xarxes neuronals convolucionals.

3 OBJECTIUS

Aquest treball té un objectiu principal:

1. Determinar la presència d'H. Pylori en imatges d'histologia utilitzant diversos models d'aprenentatge profund.

La finalitat del treball és diagnosticar els pacients amb presència o absència d'H. Pylori. Per tal d'aconseguir aquest objectiu s'han desenvolupat i testejat dos paradigmes diferents. El primer consisteix en un classificador binari basat en models preentrenats. El segon consisteix en un autoencoder dissenyat i entrenat específicament per modelar teixit sa.

Amb l'ús d'aquests dos paradigmes diferents, queden definits tres subobjectius que seran necessaris per a assolir l'objectiu principal. Aquests subobjectius són:

1. Obtenir un sistema de classificació basat en les característiques d'un model preentrenat. Consistirà a entrenar un conjunt de classificadors fent servir les característiques extretes d'una xarxa neuronal VGG-16.

2. Implementar un ensemble d'aquests classificadors per incrementar la correctesa de la detecció del bacteri.
3. Obtenir un sistema basat en autoencoders. Consistirà a entrenar un autoencoder usant mostres sanes per a modelar teixit sa.

Els primers dos subobjectius pertanyen al primer paradigma per diagnosticar els pacients. El darrer subobjectiu pertany al segon paradigma de diagnòstic.

És important destacar que en tot moment es busca prioritzar la detecció d'H. Pylori, minimitzant al màxim el nombre de falsos positius. En aquest sentit, és preferible errar en el diagnòstic i identificar incorrectament un pacient amb H. Pylori, ja que aquest pacient continuarà experimentant símptomes i eventualment haurà de sotmetre's a altres proves per a una nova avaluació, en lloc de sotmetre a un pacient a un tractament innecessari en cas de diagnòstic erroni.

4 METODOLOGIA

Per la realització d'aquest treball, he optat per seguir la metodologia Kanban [7]. El motiu pel qual he optat per utilitzar aquesta metodologia és perquè ofereix una forma senzilla i efectiva de visualitzar el progrés del projecte. La utilització de taulers Kanban permet veure clarament en quin punt es troba cada tasca i quines són les següents accions a realitzar per avançar en el projecte.

A més, Kanban també permet prioritzar les tasques segons la seva importància i urgència en cada moment, fent que em pugui centrar en les tasques més rellevants per al projecte i evitar la sobrecàrrega de treball en tasques menys prioritàries. Això també m'ajuda a prendre decisions basades en dades reals i a optimitzar el meu temps i recursos per aconseguir els millors resultats possibles.

Per la implementació de la metodologia Kanban en el meu treball, he decidit utilitzar les columnes Backlog, To Do, In progress, Review, Done i Blocked.

La columna Backlog serà la primera, on inclouré totes les tasques a realitzar. A la columna To Do, prioritaré les tasques més importants per la fase del projecte actual i definiré les que es faran durant el sprint.

Les tasques In progress seran les que ja estiguin en marxa durant el sprint actual, i aniran passant a la columna Review a mesura que estiguin a punt per ser revisades. Una vegada revisades, passaran a la columna Done, indicant que s'han completat.

Finalment, la columna Blocked s'utilitzarà per a aquelles tasques que es trobin bloquejades per algun motiu, com ara la falta de recursos o la necessitat de més informació.

Per organitzar millor el temps i les tasques, he decidit dividir el projecte en sprints de tres setmanes cada un. Això em permetrà avaluar el progrés del projecte i realitzar ajustos en la planificació en funció dels objectius aconseguits fins a aquest punt. A continuació es mostra el tauler Kanban que seguiré durant la realització del treball, com es pot veure a la figura 2.

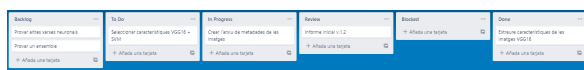


Fig. 2: Tauler Kanban

5 PLANIFICACIÓ

En aquest apartat s'exposa la planificació que es durà a terme durant l'execució del treball. A la figura 3 es pot veure el diagrama de Gantt, que serveix per veure gràficament les tasques a fer i la seva durada.

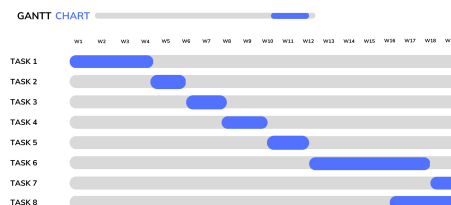


Fig. 3: Diagrama de Gantt

Task 1 (Setmana 1-4): Revisió bibliogràfica i documentació sobre l'ús de xarxes neuronals.

Task 2 (Setmana 4-6): Extracció de les característiques de les imatges dels pacients utilitzant la xarxa neuronal VGG-16.

Task 3 (Setmana 6-8): Entrenament i avaluació de diversos models de classificació a nivell de finestres d'imatges dels pacients.

Task 4 (Setmana 8-10): Detecció de presència o absència d'H. Pylori a nivell de pacient tenint en compte el conjunt de totes les seves finestres. Es podrà assolir el primer subobjectiu i part de l'objectiu principal.

Task 5 (Setmana 10-12): Implementació d'un ensemble de models per a millorar el rendiment i la robustesa de la detecció, i comparació del seu rendiment amb els rendiments dels models individuals. Es podrà assolir el segon subobjectiu.

Task 6 (Setmana 12-18): Utilització d'un autoencoder per poder detectar H. Pylori en els pacients. Es podrà assolir el tercer subobjectiu i acabar d'assolir l'objectiu principal.

Task 7 (Setmana 18-19): Anàlisi dels resultats obtinguts.

Task 8 (Setmana 16-19): Redacció de l'informe final.

6 DESENVOLUPAMENT

Les imatges proporcionades per l'Hospital Universitari General de Catalunya-Grupo Quirónsalud segueixen una estructura específica. En aquest conjunt de dades, cada pacient té dues mostres d'histologia, i cadascuna d'elles es divideix en diverses finestres. Les finestres representen zones d'interès dins de cada mostra, i això implica que el nombre de finestres pot variar entre els diferents pacients. En aquest treball, s'ha utilitzat exclusivament una de les dues mostres d'histologia de cada pacient per a l'anàlisi, fent ús d'aquesta mostra en particular i de les seves finestres associades. Per tant, un pacient és simplificat a una mostra d'histologia i diverses finestres que representen zones d'interès d'aquesta mostra.

Com es mostra a la figura 4, la mostra d'histologia d'un pacient es divideix en diverses finestres que representen zones d'interès.

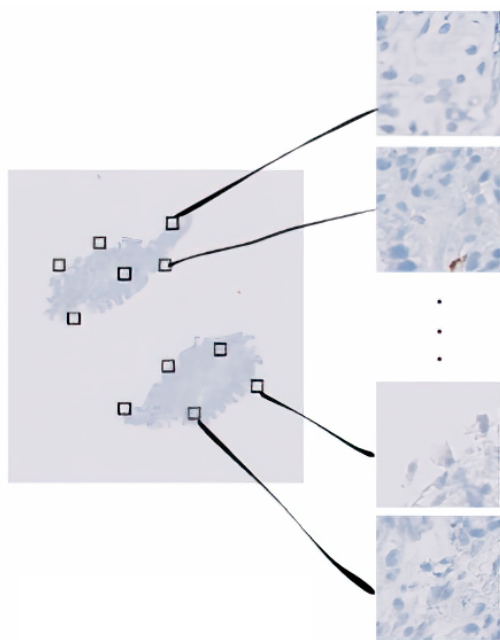
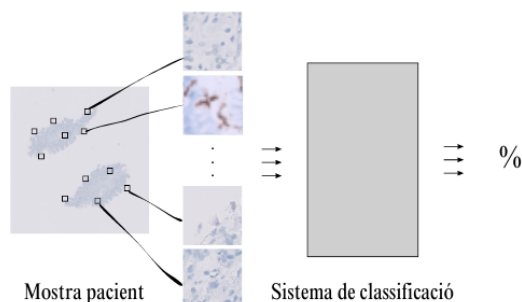


Fig. 4: Estructura de les imatges

Independentment del sistema utilitzat per detectar *H. Pylori*, l'entrada als sistemes sempre són les finestres associades a la mostra del pacient. La sortida d'aquests és la probabilitat que el pacient sigui positiu en *H. Pylori*. Aquesta probabilitat s'obté mitjançant l'agregació dels resultats de totes les finestres associades.

A la figura 5 es mostra la representació gràfica del procés d'obtenció d'aquesta probabilitat independentment del sistema de classificació utilitzat.

Fig. 5: Representació gràfica de l'obtenció de la probabilitat de ser positiu en *H. Pylori*

És important tenir en compte que cada sistema de detecció internament opera de manera diferent, encara que tots resulten en una única probabilitat. A més, l'ús de la probabilitat resultant també varia en la presa de decisions per al diagnòstic final de cada pacient depenent de cada sistema de detecció. Això implica que cada sistema té els seus propis criteris o llindars per determinar si un pacient és diagnosticat com a positiu o negatiu en *H. Pylori* basant-se en aquesta probabilitat.

Com s'ha especificat als objectius, s'han desenvolupat dos paradigmes de processament d'imatges diferents. Aquests paradigmes resulten en tres sistemes de classificació diferents, coincidint amb els tres subobjectius del treball. El primer paradigma consisteix en un classificador binari basat en models preentrenats. El segon paradigma

ma consisteix en un autoencoder dissenyat i entrenat específicament per modelar teixit sa.

Els sistemes de classificació resultants són:

1. Models de classificació basats en les característiques extretes d'una xarxa neuronal VGG-16.
2. Ensemble dels models de classificació anteriors.
3. Autoencoder basat en mostres sanes usat per modelar teixit sa.

A continuació s'exposa el desenvolupament intern de cadascun dels sistemes de classificació.

6.1 Models de classificació

Aquests models seran utilitzats per classificar les finestres individuals de la mostra d'histologia de cada pacient. Una vegada es tingui la classificació de totes les finestres, aquestes es combinaran per obtenir la probabilitat final relacionada amb la presència d'*H. Pylori*.

El primer pas d'aquest sistema consisteix a extreure les característiques de les finestres de la mostra d'histologia de cada pacient. Aquest procés es realitza utilitzant una xarxa neuronal VGG-16, que permet extreure una representació numèrica que reflecteix les propietats visuals de les finestres i que serà utilitzada per als següents passos del sistema.

La VGG-16 és una xarxa neuronal convolucional que està formada per 16 capes, incloent-hi capes convolucionals, capes de max-pooling i capes fully-connected. Les capes convolucionals s'encarreguen de detectar característiques visuals a diferents nivells d'abstracció a partir de petits filtres de 3x3. Les capes de max-pooling redueixen la dimensionalitat de la representació de les característiques. Finalment, les capes fully-connected realitzen la classificació final de la imatge. A la Figura 6 es pot veure l'arquitectura de la VGG-16.

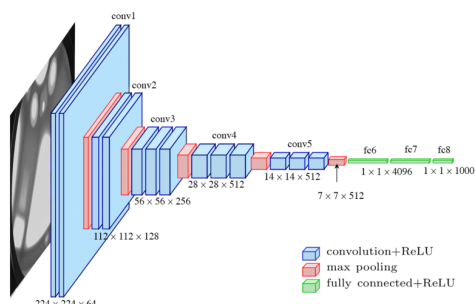


Fig. 6: Arquitectura de la VGG16

En lloc d'utilitzar el vector de 1000 característiques de la capa fully-connected final (fc8), en aquest treball s'ha optat per utilitzar el vector de 4096 característiques de la capa fully-connected anterior (fc7). Aquesta decisió s'ha pres tenint en compte que el vector de 4096 característiques pot proporcionar una representació més detallada de les imatges d'histologia, el que pot millorar la capacitat de discriminació i rendiment del sistema de detecció. No obstant això, no totes les 4096 característiques del

vector són essencials per determinar la presència d'H. Pylori en les finestres. Per aquest motiu, s'ha aplicat un algorisme de reducció de dimensionalitat anomenat mRmR (minimum redundancy maximum relevance) per seleccionar les característiques més rellevants per a la detecció del bacteri.

El mRmR (minimum redundancy maximum relevance) és un algorisme que s'utilitza per a la selecció de característiques en anàlisi de dades. Aquesta tècnica busca identificar aquelles característiques que presenten una elevada rellevància per a la classificació dels casos, alhora que es minimitza la redundància entre ells [8]. Així, s'obté un subconjunt reduït de les característiques més informatives per a la detecció de l'H. Pylori en les imatges d'histologia dels pacients.

Aquesta selecció de característiques permet millorar l'eficiència dels models de classificació, ja que se centren en aquells aspectes de les imatges que són més rellevants per a la detecció del bacteri, fent que es redueixi el temps i el cost computacional necessari per a l'entrenament i l'execució dels models.

S'ha observat que utilitzar només les 10 característiques més importants, seleccionades mitjançant l'algorisme mRmR, és suficient per obtenir resultats satisfactoris en la classificació de la presència d'H. Pylori en les finestres. Així doncs, aquest conjunt de les 10 característiques més rellevants de cada finestra dels pacients serà usat com a entrada pels diferents models de classificació.

Pel que fa a l'entrenament dels models de classificació, s'ha utilitzat la tècnica de Calibrated Classifier per a la calibració de les probabilitats de les classes.

El Calibrated Classifier és una tècnica que ajuda a ajustar les probabilitats de les classes previstes pels models d'aprenentatge profund, amb l'objectiu d'obtenir una sortida més precisa i fiable. Aquesta tècnica funciona ajustant les probabilitats de sortida dels models a través d'un algorisme de regressió, el qual estima la probabilitat real d'una observació dins de cada classe [9].

Els models de classificació generen una decisió per a cada finestra, indicant si és positiva o negativa per a la presència d'H. Pylori. A partir d'aquestes decisions individuals, es calcula la probabilitat global de ser positiu per al pacient, mitjançant la determinació del percentatge de finestres positives en relació amb el total.

La figura 7 il·lustra el procés intern del sistema de classificació, des de l'entrada de les finestres de la mostra d'histologia del pacient fins a la generació de la probabilitat de ser positiu en H. Pylori.

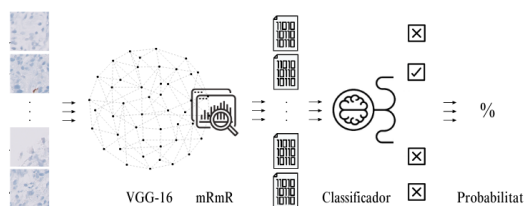


Fig. 7: Representació gràfica del procés intern del primer sistema de classificació

Per tal de determinar el llindar de separació òptim per

diagnosticar la presència o absència d'H. Pylori en els pacients, es duren a terme diversos experiments per trobar el millor criteri de classificació basat en les dades disponibles.

6.2 Ensemble de models de classificació

El segon sistema de classificació és un ensemble format pels models de classificació anteriors. Aquest sistema utilitza i combina els resultats de múltiples models per arribar a una decisió final. En aquest cas, aquest sistema comparteix gran part del funcionament intern amb el sistema dels models individuals, ja que es fan servir les mateixes característiques extretes de les finestres de la mostra d'histologia del pacient. L'ensemble té com a objectiu millorar la precisió i la capacitat de generalització del sistema de diagnòstic, aprofitant la diversitat d'opinions dels models individuals i realitzant una ponderació adequada dels seus resultats. D'aquesta manera, es busca obtenir un diagnòstic més robust i fiable per a la presència d'H. Pylori en els pacients.

La figura 8 il·lustra el procés intern d'aquest sistema de classificació, des de l'entrada de les finestres de la mostra d'histologia del pacient fins a la generació de la probabilitat de ser positiu en H. Pylori.

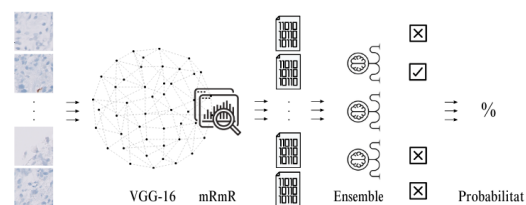


Fig. 8: Representació gràfica del procés intern del segon sistema de classificació

La principal diferència entre el segon sistema de classificació i el sistema anterior es troba en la manera en què es prenen les decisions per a cada finestra. Mentre que en el sistema anterior es basava en la classificació d'un únic classificador, en el segon sistema la decisió es pren d'acord amb el conjunt dels models que formen l'ensemble.

De la mateixa manera que en el primer sistema, per tal de determinar el llindar de separació òptim per diagnosticar la presència o absència d'H. Pylori en els pacients, es duren a terme diversos experiments per trobar el millor criteri de classificació basat en les dades disponibles.

6.3 Autoencoder

El tercer sistema de classificació es basa en l'ús d'un autoencoder basat en mostres sanes usat per modelar teixit sa.

Un autoencoder és un tipus de xarxa neuronal que té com a objectiu aprendre una representació latent de les dades d'entrada. Consisteix en dues parts principals: un codificador i un descodificador. El codificador transforma les dades d'entrada en una representació de dimensionalitat inferior anomenada codi latent. El descodificador intenta reconstruir les dades originals utilitzant aquest codi latent.

Durant el procés d'entrenament, l'autoencoder cerca minimitzar l'error de reconstrucció, és a dir, la diferència entre les dades originals i les dades reconstruïdes. Aquest procés de compressió i reconstrucció ajuda a l'autoencoder

a aprendre les característiques més rellevants de les dades d'entrada. Un cop l'autoencoder està entrenat, es pot utilitzar per reconstruir noves dades i detectar anomalies. Si les dades d'entrada no es poden reconstruir amb alta precisió, pot ser un indicatiu de presència d'anomalies o patrons no desitjats en les dades.

A la figura 9 es pot veure l'arquitectura d'un autoencoder.

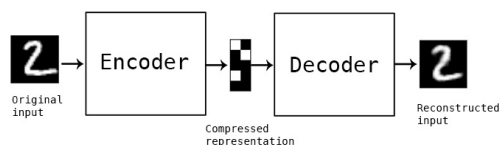


Fig. 9: Arquitectura d'un autoencoder

Aquest sistema també utilitza les finestres de la mostra d'histologia dels pacients com a entrada.

L'autoencoder s'entrena utilitzant finestres de teixit sa que no tenen la presència d'H. Pylori. Mitjançant aquest procés d'entrenament, l'autoencoder aprèn els patrons i les característiques del teixit sa i és capaç de generar una versió reconstruïda que és pràcticament igual a l'original. Quan l'autoencoder s'enfronta a una finestra que conté la presència d'H. Pylori, que és considerada una anomalia en comparació amb el teixit sa, es produeix un error de reconstrucció significatiu. Aquest error és detectat i utilitzat per identificar la finestra com a anomalia, indicant la presència del bacteri en el teixit.

Tal com s'ha definit anteriorment, després d'aplicar la tinció immunohistoquímica, es pot observar clarament una diferència de coloració entre la zona on es troba la proteïna de la membrana del bacteri H. Pylori i la resta del teixit. La zona amb la proteïna de la membrana del bacteri es tenyeix amb una tonalitat vermellosa, mentre que la resta del teixit es manté amb una tonalitat blava. Durant el procés d'entrenament de l'autoencoder, les finestres utilitzades provenen exclusivament de teixit sa, que només presenta la tonalitat blava. Això implica que, quan l'autoencoder ha de reconstruir una finestra sana, el resultat de la reconstrucció serà molt semblant a la finestra original, presentant també una tonalitat blava. No obstant això, quan l'autoencoder s'enfronta a la reconstrucció d'una finestra amb tonalitat vermellosa, la sortida de l'autoencoder mostrarà una conversió de coloració cap a la tonalitat blava. Aquesta diferència de coloració substancial entre la sortida reconstruïda i la finestra original actua com a indicador que aquesta finestra és una anomalia, és a dir, presenta la presència d'H. Pylori.

A les figures 10 i 11 es pot veure la diferència en la reconstrucció d'una finestra sana i una finestra amb presència d'H. Pylori.

Com es pot veure, la reconstrucció feta per l'autoencoder en el cas de la finestra sana és pràcticament igual a l'original. Això és a causa de la reconstrucció utilitzant les característiques d'una finestra sana. En canvi, en la reconstrucció de la finestra infectada amb H. Pylori, l'autoencoder ha modificat la coloració del teixit.

Per determinar si una finestra és considerada una anoma-



Fig. 10: Reconstrucció d'una finestra sana



Fig. 11: Reconstrucció d'una finestra amb presència d'H. Pylori

lia, s'ha desenvolupat una funció que calcula la proporció de píxels vermellosos perduts entre la imatge original i la imatge reconstruïda. Aquesta proporció representa la quantitat de teixit infectat amb H. Pylori que ha estat reconstruïda com a teixit sa. Si la proporció de píxels vermellosos perduts és superior a 1, indica que s'ha produït una pèrdua de píxels vermellosos durant el procés de reconstrucció, i per tant s'ha identificat una finestra amb la presència d'H. Pylori. Aquesta finestra serà classificada com a positiva en H. Pylori.

Mitjançant aquesta proporció es genera una decisió per classificar cada finestra com positiva o negativa per a la presència d'H. Pylori. A partir d'aquestes decisions individuals, es calcula la probabilitat global de ser positiu per al pacient, mitjançant la determinació del percentatge de finestres positives en relació amb el total.

La figura 12 il·lustra el procés intern del sistema de classificació, des de l'entrada de les finestres de la mostra d'histologia del pacient fins a la generació de la probabilitat de ser positiu en H. Pylori.

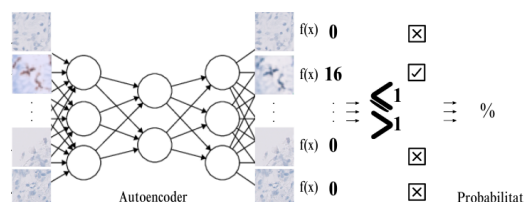


Fig. 12: Reconstrucció gràfica del procés intern del tercer sistema de classificació

De la mateixa manera que en els altres sistemes de classificació, per tal de determinar el llindar de separació òptim per diagnosticar la presència o absència d'H. Pylori en els pacients, es duren a terme diversos experiments per trobar el millor criteri de classificació basat en les dades disponibles.

7 EXPERIMENTS I RESULTATS

A continuació, es detallen les dades utilitzades per dur a terme els experiments.

L'Hospital Universitari General de Catalunya-Grupo Quirónsalud ha proporcionat un conjunt de 245 pacients amb etiquetes de densitat d'H. pylori, classificant-los com a NEGATIVA, BAIXA o ALTA. No obstant això, per simplificar el procés experimental, s'ha decidit convertir la variable de densitat en una classificació binària, distingint entre NEGATIVA i POSITIVA. Dels 245 pacients inclosos en l'estudi, 117 estan classificats com a NEGATIUS en densitat d'H. pylori, mentre que 128 estan classificats com a POSITIUS. Això representa un 47.8% de pacients negatius i indica que el conjunt de pacients presenta un equilibri notable entre les dues categories.

Com s'ha especificat anteriorment, el pacient és representat amb una mostra d'histologia, i aquesta mostra es divideix en diverses finestres que representen zones d'interès. En total, es disposen de 483.175 finestres entre tots els pacients. No obstant això, és important destacar que hi ha una gran variabilitat en el nombre de finestres per cada pacient. El pacient amb el menor nombre de finestres té només 172, mentre que el pacient amb el nombre més gran en té 21.794. Això posa de manifest una gran disparitat en la distribució de les finestres entre els diferents pacients, ja que cada mostra pot contenir diferents àrees d'interès per a l'estudi.

A més de proporcionar un conjunt de pacients etiquetats, l'Hospital Universitari General de Catalunya-Grupo Quirónsalud també ha subministrat un conjunt de finestres etiquetades. Aquestes finestres etiquetades constitueixen una petita mostra de tot el conjunt de finestres disponibles, on una patòloga ha realitzat una classificació de presència o absència d'H. Pylori, considerant-les com a positives o negatives. Aquesta informació de classificació és fonamental per entrenar els models de classificació, ja que requereixen dades etiquetades per aconseguir un aprenentatge adequat. El total de finestres etiquetades és de 1289. D'aquestes finestres, 1145 han estat classificades com a negatives i 144 com a positives. Això representa un 88.8% de finestres negatives, indicant un fort desequilibri entre les dues classes en aquest conjunt de dades. Aquest desequilibri pot dificultar l'aprenentatge dels models de classificació i limitar la seva capacitat per generalitzar correctament la classificació de la classe "presència d'H. Pylori" en les finestres. Això pot traduir-se en resultats poc satisfactoris per a la detecció del bacteri en casos positius.

A la taula 1 es mostra un resum de les dades utilitzades en els experiments realitzats.

TAULA 1: DADES UTILITZADES ALS EXPERIMENTS

	Pacients	Finestres	Finestres etiquetades
Negatius	117	-	1145
Positius	128	-	144
Total	245	483,175	1289

A continuació, es detalla el procés d'entrenament, validació i els resultats obtinguts per a cadascun dels sistemes de classificació utilitzats.

7.1 Models de classificació

Els models de classificació tenen com a entrada el vector de característiques més representatives de les finestres de la mostra del pacient després d'haver estat processada per la xarxa neuronal VGG-16, la qual té una entrada de dimensions 224x224. Per poder decidir quines són les 10 característiques més representatives, s'ha fet servir l'algorisme mRmR (minimum Redundancy Maximum Relevance) sobre els 1289 vectors de característiques de totes les finestres etiquetades. D'aquesta manera, s'aconsegueixen les posicions de les 10 característiques més rellevants que s'hauran d'extreure de cada nova finestra per poder-les emprar com a entrada als models de classificació.

Els models de classificació usats per aquest estudi inclouen el SVM, Random Forest, Logistic Regression, Decision Tree, kNN i Naive Bayes. Cadascun d'aquests models prendrà una decisió per a cada finestra, determinant si és positiva o negativa per a la presència d'H. Pylori.

Per a l'entrenament i la validació dels models de classificació, s'ha usat la tècnica de validació creuada k-fold estratificat amb un valor de k=6. Això significa que el conjunt de dades s'ha dividit en 6 subconjunts iguals, mantenint la proporció de les classes originals. S'ha realitzat l'entrenament i la validació dels models en 6 iteracions diferents, on a cada iteració, un subconjunt diferent ha estat usat com a conjunt de test, mentre que la resta han estat emprats per a l'entrenament.

Aquesta tècnica de validació creuada és útil per reduir l'overfitting dels models, ja que permet a cada subconjunt ser utilitzat com a conjunt de test i d'entrenament en diferents iteracions, de manera que es garanteix que els models no estan aprenent específicament per als casos concrets del conjunt de dades d'entrenament [10].

S'ha seleccionat un valor de k=6 per a la tècnica de validació creuada a causa del desequilibri existent entre les classes de les finestres etiquetades. Aquest desequilibri es manifesta en la petita quantitat de finestres positives, sent només 144 de les 1289 etiquetades. En cas de seleccionar un valor de k més gran, cada fold tindria una proporció encara menor de finestres positives. Això podria resultar en una manca d'aprenentatge adequat de les característiques de les finestres positives pels models de classificació.

Per tal de determinar l'efectivitat dels models de classificació, s'han obtingut les mètriques corresponents a cada iteració del k-fold, i s'ha calculat la mitjana per estimar el rendiment general del model. A més, s'ha realitzat una cerca dels millors hiperparàmetres per a cada model utilitzant un GridSearch.

A la taula 2 es mostren els millors hiperparàmetres per cada model, després d'haver realitzat un GridSearch.

Els resultats presentats a la taula 3 corresponen a la mitjana de les mètriques obtingudes per cada model després de ser entrenat amb els seus millors hiperparàmetres, i havent realitzat un total de 6 iteracions del k-fold estratificat.

La taula 4 mostra el recall de les dues classes tant en el conjunt d'entrenament com de test.

En general, tots els models han mostrat un rendiment

TAULA 2: MILLORS HIPERPARÀMETRES DELS MODELS

Model	Hiperparàmetres
SVM	C=10, gamma=1, kernel='rbf'
Random Forest	n_estimators=100, min_samples_split=2, min_samples_leaf=1, max_features='sqrt', max_depth=20, bootstrap=True
Logistic Regression	solver='liblinear', penalty='l2', C=1000
Decision Tree	splitter='best', min_samples_split=2, min_samples_leaf=4, max_features='sqrt', max_depth=None, criterion='entropy'
KNN	weights='distance', n_neighbors=10, metric='euclidean'
Naive Bayes	alpha=2

TAULA 3: AUC I F-1 EN TRAIN I TEST PER CADA MODEL

Model	Train AUC	Test AUC	Train f-1	Test f-1
SVM	0.916	0.879	0.851	0.806
Random Forest	0.982	0.868	0.981	0.803
Logistic Regression	0.885	0.874	0.810	0.800
Decision Tree	0.954	0.847	0.894	0.777
kNN	0.982	0.862	0.982	0.809
Naive Bayes	0.816	0.816	0.733	0.731

notable en les mètriques avaluades. No obstant això, s'ha detectat una tendència a obtenir pitjors resultats en la classe 'presència d'H. Pylori' en les dades d'entrenament i de test. Això s'explica pel desequilibri que hi ha en el conjunt de finestres etiquetades, ja que els models tenen dificultats per generalitzar en noves finestres en el cas de la classe minoritària.

A més, és important destacar que en la majoria dels casos les mètriques de test són lleugerament inferiors a les corresponents a l'entrenament. Aquesta diferència pot indicar la presència d'un cert grau d'overfitting, és a dir, els models podrien haver après específicament els detalls del conjunt d'entrenament i tenir dificultats per generalitzar de manera adequada en noves mostres.

Una vegada els models han estat entrenats, és possible classificar qualsevol nova finestra després d'extreure el vector de característiques requerit pels models. D'aquesta manera, es pot obtenir la classificació de cadascuna de les finestres de la mostra d'un pacient i obtenir la probabilitat de ser positiu en H. Pylori a partir del percentatge de finestres classificades com a positives respecte al total. En aquest punt, es pot aconseguir la probabilitat de ser positiu en H. Pylori per cadascun dels pacients.

L'eina utilitzada per determinar el diagnòstic dels pacients a partir de les seves probabilitats de ser positius ha estat

TAULA 4: RECALL EN TRAIN I TEST PER CADA MODEL

Model	Train No H.	Train H.	Test No H.	Test H.
SVM	0,904	0,780	0,859	0,756
Random Forest	0,982	0,981	0,837	0,771
Logistic Regression	0,893	0,729	0,886	0,720
Decision Tree	0,906	0,889	0,797	0,767
kNN	0,982	0,982	0,853	0,768
Naive Bayes	0,756	0,711	0,755	0,713

el llinard òptim obtingut a partir de la corba ROC (Receiver Operating Characteristic). Aquesta corba permet avaluar el rendiment del model i trobar el punt òptim per a la classificació, que és el punt més proper al punt (0,1). Un cop aconseguit el llinard òptim, es pot diagnosticar cada pacient.

A la taula 5 es presenten els resultats del diagnòstic final dels pacients, incloent-hi el llinard òptim usat i diverses mètriques per avaluar la precisió dels models.

TAULA 5: RESULTATS DEL DIAGNÒSTIC DELS PACIENTS

Model	Llinard	AUC	No H. Recall	H. Recall
SVM	18.969%	0.771	0.739	0.695
Random Forest	20.574%	0.725	0.678	0.695
Logistic Regression	21.469%	0.713	0.609	0.734
Decision Tree	25.132%	0.768	0.757	0.695
kNN	20.331%	0.753	0.739	0.695
Naive Bayes	10.825%	0.731	0.730	0.680

Es pot observar que els llinars òptims varien per a cada model, oscil·lant entre el 10.825% i el 25.132%. Aquestes diferències reflecteixen les característiques pròpies de cada model i la seva capacitat de discriminació. Per avaluar el rendiment global de cada model, s'utilitza l'àrea sota la corba ROC (AUC). Valors més alts d'AUC indiquen un millor rendiment en la classificació. En aquest cas, els models mostren un rang d'AUC entre 0.713 i 0.771. Això suggereix que els models tenen una capacitat limitada per distingir entre pacients amb presència i sense presència d'H. Pylori. A més, es mostra el recall de les dues classes, i es torna a posar de manifest que cap dels models aconsegueix un rendiment excepcional.

Veient aquests resultats es pot arribar a la conclusió que aquests models per separat no assoleixen la discriminació correcta dels pacients. És a dir, no són eficients a l'hora de diagnosticar un pacient amb presència o absència d'H. Pylori.

7.2 Ensemble de models de classificació

Per intentar millorar els resultats del diagnòstic dels pacients, s'ha utilitzat un ensemble de tots els models entrenats. En una primera versió, cadascun dels models té el mateix pes, per tant, la classificació de cada finestra es determina segons la classe que hagi estat més freqüent entre els models. En la segona versió, s'assigna un pes als models en funció de les seves mètriques, donant més pes als models amb millors resultats. Això permet que els models més rendibles tinguin més influència en la classificació final.

A la taula 6 es mostren els resultats per les dues versions.

Veient aquests resultats es pot comprovar que en fer l'ensemble assignant els pesos òptims a cada model millora les mètriques, respecte assignar el mateix pes a cadascun. Així

TAULA 6: RESULTATS DE L'ENSEMBLE DELS MODELS

Ensemble	Llindar	AUC	No H. Recall	H. Recall
Tots els models amb el mateix pes	21.793%	0.742	0.749	0.648
Pes òptim per cada model	18.324%	0.758	0.757	0.680

i tot, el rendiment de l'ensemble no millora els resultats obtinguts pels models de manera individual.

A la figura 13 es mostra la corba ROC resultant de l'ensemble amb els pesos òptims de cada model, on es pot observar el punt que representa el llindar òptim pel diagnòstic dels pacients.

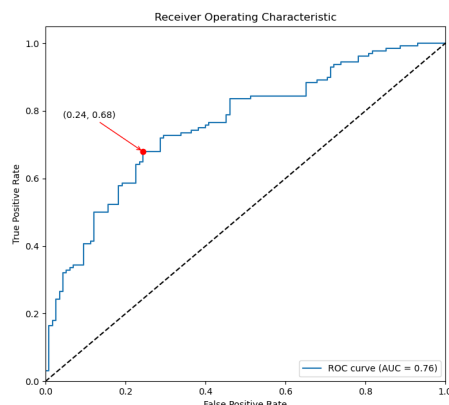


Fig. 13: Corba ROC de l'ensemble de models

Per tal de visualitzar gràficament la raó per la qual els models de classificació no són capaços de diagnosticar correctament els pacients, es presenta el boxplot de la figura 14. Aquesta representació mostra les probabilitats de ser positiu en H. pylori per a cada pacient, dividits segons el seu valor real.

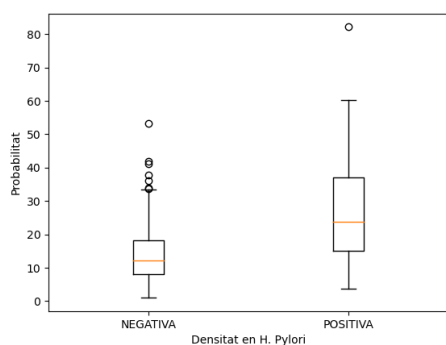


Fig. 14: Boxplot de les probabilitats d'H. Pylori amb l'ensemble de models

Com es pot veure, hi ha una certa tendència dels pacients sans a tenir un percentatge menor, però hi ha una superposició significativa entre els percentatges de pacients sans i infectats amb H. Pylori. Això implica que existeix una àrea ampla on els pacients poden ser diagnosticats tant com a sans com infectats. Com a resultat, és difícil determinar un llindar òptim per classificar aquests pacients, ja que sempre hi haurà un risc de falsos positius i falsos negatius.

TAULA 7: MATRIU DE CONFUSIÓ DEL DIAGNÒSTIC DELS PACIENTS AMB L'ENSEMBLE DE MODELS

Real / Predicció	No H. Pylori	H. Pylori
No H. Pylori	88 (TP)	28 (FP)
H. Pylori	41 (FN)	88 (TN)

A la taula 7 es mostra la matriu de confusió del diagnòstic dels pacients. Es pot observar que el nombre de falsos negatius i falsos positius és molt elevat, la qual cosa indica que, tot i que els models tenen resultats acceptables a nivell de finestres etiquetades, no són capaços de generar un bon diagnòstic a nivell de pacient.

7.3 Autoencoder

A diferència dels models de classificació, l'autoencoder utilitza directament les finestres com a dades d'entrada, en lloc de fer servir el vector de característiques extret de la xarxa neuronal VGG-16. A més, l'autoencoder és capaç de ser entrenat sense la necessitat de dades etiquetades. Això significa que pot aprofitar un conjunt més ampli de finestres pel seu entrenament, ja que no requereix l'assignació manual d'etiquetes.

En el cas de l'entrenament de l'autoencoder s'ha utilitzat una estratègia especial per a l'obtenció de dades. Com que l'objectiu és reconstruir teixit sa, s'han seleccionat finestres de pacients que han estat etiquetats com a sans, assumint que si un pacient és considerat sense H. Pylori, les seves finestres també seran considerades com a sanes. Aquest enfocament permet augmentar significativament el conjunt de dades disponibles, ja que no està limitat només a les finestres específicament etiquetades com a sanes, que són únicament 144. En aquest cas, s'han seleccionat 50 finestres aleatòries de cada pacient etiquetat com a sa, obtenint un total de 5850 finestres. Tot i que hauria estat desitjable fer servir un nombre encara més gran de finestres, s'han tingut en compte les limitacions de recursos disponibles. A més, a causa d'aquestes limitacions, les dimensions de les finestres han estat reduïdes de 224x224 a 28x28. Malgrat aquest canvi en les dimensions, el nombre de finestres utilitzades per a l'entrenament ha augmentat de manera significativa, passant de 1289 a 5850.

Un cop l'autoencoder ha estat entrenat, ja pot reconstruir qualsevol finestra nova. És en aquest punt on es pot calcular l'error de reconstrucció mitjançant una funció que calcula la proporció de píxels vermellencs que s'han perdut a la reconstrucció. Com s'ha especificat anteriorment, totes les finestres que superin el llindar d'1, seran considerades una anomalia, i per tant seran classificades com a positives en H. Pylori.

A partir de totes les finestres de la mostra d'histologia de cada pacient, es pot calcular la probabilitat de ser positiu en H. Pylori mitjançant el càlcul del percentatge de finestres classificades com a positives en relació amb el total de finestres.

Amb les probabilitats de cada pacient i la classificació real etiquetada es busca trobar el percentatge òptim

per diagnosticar finalment els pacients. Per fer-ho, a diferència dels models de classificació, s'ha utilitzat un k-fold estratificat, amb $k=10$, on en cada iteració es calcula l'àrea sota la corba ROC i es busca el punt en el qual es maximitza la diferència entre la taxa de positius veritables i falsos positius. D'aquesta manera s'intenta trobar el llindar en el qual el model discrimina correctament i redueix els falsos positius. El llindar per poder diagnosticar qualsevol altre nou pacient és la mitjana dels llindars obtinguts a cada iteració del k-fold estratificat.

A la taula 8 es mostren els resultats aconseguits en fer la classificació per pacient usant l'autoencoder.

TAULA 8: RESULTATS DEL DIAGNÒSTIC DELS PACIENTS AMB L'AUTOENCODER

	No H. Pylori	H. Pylori	Ponderat
Precision	0.862	0.957	0.911
Recall	0.957	0.859	0.906
f-1 score	0.907	0.905	0.906

El llindar òptim és 6.184% i l'AUC és 0.961.

Els resultats obtinguts amb l'autoencoder en el diagnòstic dels pacients són molt prometedors. Tant la precisió com el recall per a les dues classes són elevats, indicant que el model és capaç de distingir adequadament entre pacients sans i pacients amb la presència del bacteri. A més, la puntuació f-1 mostra un bon equilibri entre la precisió i el recall per a ambdues classes. L'AUC és també molt alta, suggerint un bon rendiment global en la classificació dels pacients.

A la figura 15 es mostra la corba ROC resultant de l'autoencoder, on es pot observar el punt que representa el llindar òptim pel diagnòstic dels pacients.

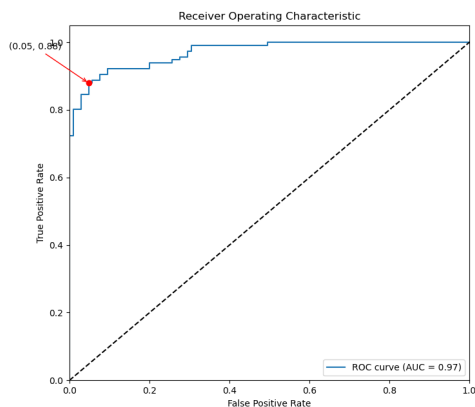


Fig. 15: Corba ROC de l'autoencoder

Per tal de visualitzar gràficament la raó per la qual l'autoencoder és capaç de diagnosticar amb bons resultats els pacients, es presenta el boxplot de la figura 16. Aquesta representació mostra les probabilitats de ser positiu en H. pylori per a cada pacient, dividits segons el seu valor real.

Com es pot veure, la tendència dels pacients etiquetats com a sans tenen percentatges de finestres anòmales més baixos que els pacients etiquetats amb H. Pylori. A diferència del boxplot de l'ensemble de models, les dues cai-

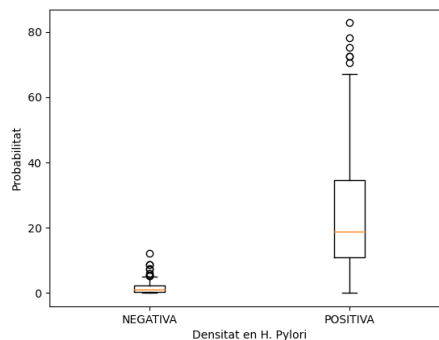


Fig. 16: Boxplot de les probabilitats d'H. Pylori amb l'autoencoder

xes estan molt més diferenciades, indicador que els resultats són millors amb aquest enfocament, ja que en escollir un llindar hi haurà menys risc de falsos positius i falsos negatius.

A la taula 9 es mostra la matriu de confusió del diagnòstic dels pacients.

TAULA 9: MATRIU DE CONFUSIÓ DEL DIAGNÒSTIC DEL PACIENT AMB L'AUTOENCODER

Real / Predicció	No H. Pylori	H. Pylori
No H. Pylori	112 (TP)	5 (FP)
H. Pylori	18 (FN)	110 (TN)

Del total de 245 pacients, s'ha observat una classificació incorrecta en 23 casos. És important destacar que, en la recerca del llindar òptim, s'ha donat prioritat a la reducció dels falsos positius. Aquests resultats són molt millors en comparació amb l'ensemble de models, però encara hi ha espai per a millores addicionals amb l'objectiu d'aconseguir un diagnòstic més precís i fiable.

8 CONCLUSIONS

Després de dur a terme aquest treball, es pot afirmar que els objectius establerts han estat satisfets dins dels límits de temps establerts. S'ha aconseguit l'assoliment exitós de tots els subobjectius que conduïen a l'objectiu principal. A través d'aquesta consecució, l'objectiu principal també ha estat satisfactòriament complert.

Primerament, s'ha emprat la xarxa neuronal VGG-16 per extreure les característiques de les imatges d'histologia dels pacients. Aquestes característiques extretes han estat utilitzades com a entrada als models de classificació. Seguidament, s'ha dut a terme l'entrenament de diversos models de classificació amb l'objectiu de discernir entre la presència i l'absència d'H. Pylori en cada finestra dels pacients. Cal destacar que aquests models han mostrat bons resultats a nivell de finestres, tal com s'ha comprovat amb el ground-truth de finestres etiquetades per una patòloga. No obstant això, no han mostrat resultats satisfactoris en el diagnòstic global d'un pacient, tenint en compte el conjunt de totes les seves finestres. Aquests resultats porten directament al segon subobjectiu, que consisteix a implementar un ensemble

dels models per intentar millorar els resultats. Malauradament, s'ha comprovat que, tot i assignar els pesos òptims a cada model, això no millora el seu rendiment de manera significativa. En conclusió, es pot afirmar que aquest paradigma del projecte, utilitzant la xarxa neuronal VGG-16 i diversos models de classificació, no ha aconseguit complir satisfactòriament l'objectiu principal, ja que no és capaç de diagnosticar correctament els pacients.

En contrast, el paradigma alternatiu basat en l'ús d'un autoencoder ha demostrat obtenir resultats satisfactoris. El tercer subobjectiu s'ha assolit mitjançant l'entrenament d'un autoencoder fent servir finestres sanes de pacients sans com a dades d'entrada. Aquest autoencoder ha demostrat ser eficaç en la detecció d'anomalies en noves finestres. Mitjançant l'anàlisi dels percentatges d'anomalies de cada pacient, s'ha assolit una discriminació precisa entre pacients amb presència o absència d'*H. Pylori*, fent que l'objectiu principal quedi satisfactòriament resolt. És important assenyalar que aquests són els resultats aconseguits amb unes limitacions de temps i recursos. Això implica que, amb l'accés a recursos millors i una dedicació temporal més extensa, és bastant probable que es puguin assolir resultats encara més favorables. Per exemple, una millora potencial per a l'entrenament de l'autoencoder és l'ús d'una quantitat significativament més gran de finestres sanes, així com la utilització de les finestres amb una millor resolució. Aquestes millores podrien reforçar la capacitat de l'autoencoder per detectar anomalies en noves finestres, proporcionant un diagnòstic més precís dels pacients.

En conclusió, aquest treball proporciona una base sòlida per a futurs desenvolupaments en l'àmbit del diagnòstic d'*H. Pylori* utilitzant imatges d'histologia. Amb la continuació d'aquesta recerca, s'espera poder contribuir a la millora de la detecció i tractament d'aquesta infecció bacteriana.

net/publication/51406923_Helicobacter_Pylori-Related_Gastric_Histology_Classification_Using_Support-Vector-Machine-Based_Feature_Selection.

- [5] S. Shichijo, "Application of Convolutional Neural Networks in the Diagnosis of Helicobacter pylori Infection Based on Endoscopic Images," 16 d'Octubre de 2017. [En línia]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2352396417304127>.
- [6] B.P. Mohan, "Convolutional neural networks in the computer-aided diagnosis of Helicobacter pylori infection and non-causal comparison to physician endoscopists: a systematic review with meta-analysis," 6 de Juliol de 2020. [En línia]. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7774656/>.
- [7] Atlassian, "Aplicación de la metodología kanban en el desarrollo de software." [En línia]. Disponible en: <https://www.atlassian.com/es/agile/kanban>.
- [8] Smazzanti, "MRMR: MRMR (minimum-redundancy-maximum-relevance) for automatic feature selection at scale." [En línia]. Disponible en: <https://github.com/smazzanti/mrmr>.
- [9] scikit, "1.16. probability calibration, scikit" [En línia]. Disponible en: <https://scikit-learn.org/stable/modules/calibration>.
- [10] scikit, "Sklearn.model_selection.Kfold" [En línia]. Disponible en: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html.

REFERÈNCIES

- [1] J.C. Yang, "Treatment of Helicobacter pylori infection: Current status and future concepts", 14 de Maig de 2014 [En línia]. Disponible en: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4017043/>.
- [2] W.G Gonçalves, "DeepHP: A New Gastric Mucosa Histopathology Dataset for Helicobacter pylori Infection Diagnosis," 23 de Novembre de 2022. [En línia]. Disponible en: <https://www.mdpi.com/1422-0067/23/23/14581>.
- [3] S. Zhou, "Deep learning assistance for the histopathologic diagnosis of Helicobacter pylori," 15 de Setembre de 2020. [En línia]. Disponible en: <https://www.sciencedirect.com/science/article/pii/S2666521220300041>.
- [4] C.R. Huang, "Helicobacter Pylori-Related Gastric Histology Classification Using Support-Vector-Machine-Based Feature Selection", Juliol de 2008 [En línia]. Disponible en: https://www.researchgate.net/publication/51406923_Helicobacter_Pylori-Related_Gastric_Histology_Classification_Using_Support-Vector-Machine-Based_Feature_Selection.