



**UNIVERSITÀ
DEGLI STUDI
DI PADOVA**



DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE

CORSO DI LAUREA IN INGEGNERIA INFORMATICA

“Deep Learning e Studio delle Ecocardiografie Fetali per la Diagnosi Prenatale di Cardiopatie Congenite”

Relatore: Prof. Loris Nanni

Laureanda: Serena De Antoni

ANNO ACCADEMICO 2021–2022

Data di laurea 14/03/2022

Indice

INTRODUZIONE.....	2
CAPITOLO 1 – L’INTELLIGENZA ARTIFICIALE.....	3
1.1 - Concetti principali.....	3
1.2 - Reti Neurali.....	4
1.3 - Reti Neurali profonde e Deep Learning.....	6
1.4 - Reti Neurali Convoluzionali (CNN).....	6
CAPITOLO 2 – STUDIO ED ELABORAZIONE DEI DATI.....	9
2.1 - Analisi dei dati.....	9
2.2 - Processamento dei dati.....	11
2.3 - Creazione dei dataset.....	13
2.4 - LIME (Local Interpretable Model-agnostic Explanations).....	15
2.5 - Descrizione della rete utilizzata e addestramento.....	16
CAPITOLO 3 – RISULTATI OTTENUTI.....	17
3.1 - Studio dei risultati.....	17
3.2 - Risultati ottenuti e confronto con i dataset precedenti	18
3.3 - Considerazioni sui risultati ottenuti.....	21
CONCLUSIONI.....	23
RIFERIMENTI.....	24

INTRODUZIONE

Le cardiopatie congenite rappresentano quasi un terzo delle principali anomalie congenite che possono verificarsi durante la gravidanza. Risultano essere tra le principali cause di mortalità infantile, infatti, ogni anno, causano la morte di circa 300.000 bambini in età prenatale o durante l'infanzia.

La diagnosi precoce di cardiopatie congenite porta a una riduzione della mortalità che ora ammonta al 25% nelle morti prenatali e al 50% nelle morti infantili [1].

Nel corso degli anni, la diagnosi precoce di queste malattie è aumentata di circa il 20%, arrivando circa al 40%, ma non raggiungendo ancora risultati ottimali [2].

Risulta quindi evidente l'importanza di uno studio e di una diagnosi precoce per consentire l'avvio tempestivo di terapie appropriate per limitare i rischi e favorire il normale sviluppo del cuore del feto o per consentire il rapido intervento chirurgico. Esistono dei fattori di rischio che aumentano la possibilità di cardiopatia congenita nel corso della gravidanza, qualora questi fattori fossero presenti vi è una indicazione ad una valutazione ecocardiografica precoce estesa del feto. La maggior parte delle cardiopatie però insorge in gravidanze prive di rischio e pertanto il sospetto diagnostico di cardiopatia viene principalmente posto nel corso dell'ecografia delle venti settimane, ovvero l'ecografia morfologica [3]. Molti casi di cardiopatie congenite, nonostante le ecografie morfologiche effettuate nelle settimane vicine alla ventesima, non vengono rilevate dai medici. Spesso i fattori decisivi nella non individuazione di una cardiopatia sono umani, come ad esempio la poca esperienza del medico, la mal interpretazione dei valori acquisiti oppure da ecografie poco nitide dovute ai movimenti del feto o dalle piccole dimensioni del cuore di quest'ultimo.

Negli ultimi anni, grazie al progredire della tecnologia, c'è stato un incremento nell'uso di tecniche di deep learning per assistere i medici durante la diagnosi di cardiopatie congenite prenatali che hanno portato a buoni risultati nel tasso di rilevazione di queste malattie.

Questa tesi si pone come obiettivo di illustrare un'applicazione di deep learning basata su Reti Neurali Convolutionali (CNN) che studia le ecocardiografie fetali per la diagnosi prenatale di cardiopatie congenite.

CAPITOLO 1

L'intelligenza Artificiale

1.1 Concetti principali

L'intelligenza artificiale si può definire come la riproduzione parziale dell'attività intellettuale propria dell'uomo (con particolare riguardo ai processi di apprendimento, di riconoscimento, di scelta) realizzata o attraverso l'elaborazione di modelli ideali, o concretamente con la messa a punto di macchine che utilizzano per lo più a tale fine elaborati elettronici [4].

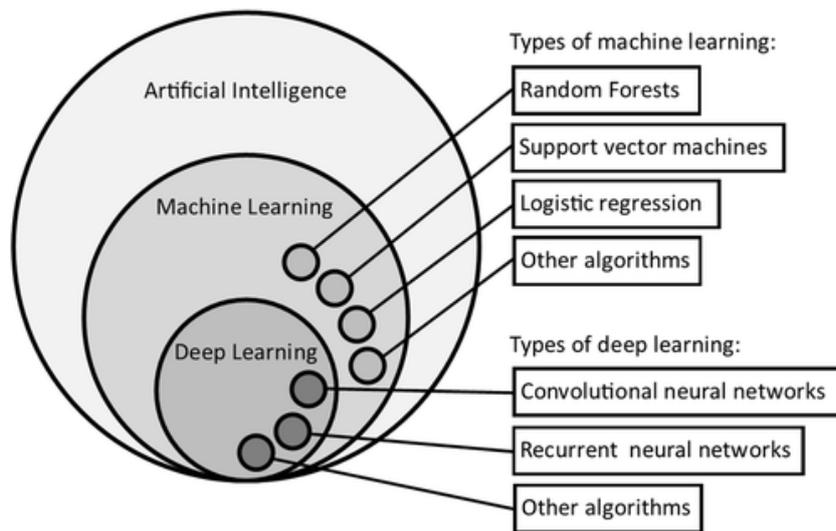


Figura 1. 1: rappresentazione della relazione tra intelligenza artificiale, machine learning e deep learning con esempi di machine learning e metodi di deep learning

Fonte: T. G. Day, B. Kainz, J. Hajnal, R. Razavi, J. M. Simpson, *Artificial intelligence, fetal echocardiography, and congenital heart disease*, 18 gennaio 2021

L'intelligenza artificiale è una disciplina molto vasta che copre diverse tematiche tra cui il Machine Learning, che è uno degli approcci ritenuti oggi più importanti dell'intelligenza artificiale [4]. Il Machine Learning è un sistema di apprendimento automatico che durante la fase di training (addestramento) apprende da esempi e successivamente è in grado di generalizzare e gestire nuovi dati nello stesso dominio applicativo. Quindi si può dire che la macchina impara dall'esperienza(E) per svolgere il compito(T) con risultati(P), i quali migliorano di volta in volta attraverso l'esperienza (E).

Sottocategoria molto importante del Machine Learning è il Deep Learning (apprendimento profondo), indica quella parte dell'intelligenza artificiale che prende ispirazione dal funzionamento del cervello umano, utilizzando reti neurali artificiali. È un sistema che

riesce a classificare dati e strutturarli in modo gerarchico autonomamente, migliorando le prestazioni mediante un apprendimento continuo.

Negli ultimi anni, grazie al progredire della tecnologia, c'è stata un'esplosione nell'uso dell'intelligenza artificiale la quale ora si trova anche in oggetti di uso quotidiano. L'utilizzo in medicina però è stato più limitato sebbene i risultati ottenuti abbiano mostrato ottime prestazioni quando l'IA viene utilizzata per assistere un medico in un compito ben definito.

1.2 Reti Neurali

Una rete neurale è un modello di calcolo la cui struttura stratificata assomiglia alla struttura della rete di neuroni nel cervello, con strati di nodi connessi. Una rete neurale può apprendere dai dati, quindi può essere addestrata a riconoscere pattern, classificare i dati e prevedere eventi futuri [5]. Alla base del modello della rete neurale vi è il neurone artificiale, che si ispira alla struttura e al funzionamento del neurone biologico. Queste reti possono essere viste come funzioni matematiche non lineari che trasformano un insieme di variabili indipendenti $x = (x_1, \dots, x_n)$, chiamate ingressi della rete, in un insieme di variabili dipendenti $net_i = (net_1, \dots, net_n)$, chiamate uscite della rete che determinano il livello di eccitazione della rete. La precisa forma di queste funzioni dipende dalla struttura interna della rete e da un insieme di valori $w_I = (w_{I1}, \dots, w_{In})$ chiamati pesi applicati agli input. È presente anche un ulteriore peso detto "bias" $w_0 = (w_{01}, \dots, w_{0n})$ che viene assegnato ad un input fittizio avente valore 1 [6].

Questi elementi vengono combinati secondo la sommatoria:

$$net_i = \sum_{j=1}^n w_{Ij} * x_j + w_{0i} \quad (1.1)$$

Inoltre, il valore di output è dato dalla funzione di attivazione f applicato al risultato sopra ottenuto:

$$out_i = f(net_i) \quad (1.2)$$

Dove f non è lineare, continua e differenziabile.

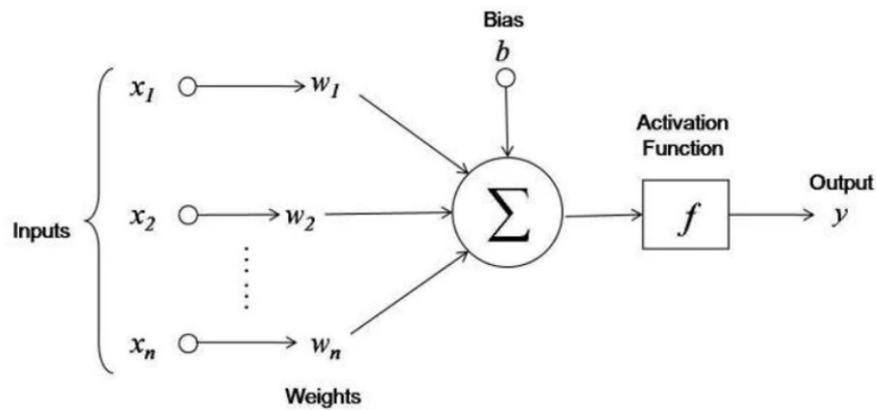


Figura 1. 2: operazioni eseguite da un neurone artificiale

Fonte: <https://www.intelligenzaartificialeitalia.net/post/spiegazione-della-pi%C3%B9-semplice-rete-neurale-per-principianti-con-codice-implementazione-python>

Il più semplice modello di rete neurale può essere considerato il Perceptron e solitamente le reti neurali sono composte da più perceptron divisi in vari livelli (Multilayer Perceptron), e ogni livello successivo riceve come dati iniziali di input i risultati forniti dal livello precedente. Ogni perceptrone è di tipo feed-forward in cui neuroni di un livello sono connessi soltanto a quelli del livello successivo.

Il Multilayer Perceptron (MLP) è una rete feed-forward avente almeno tre livelli e con una funzione di attivazione non lineare, può quindi distinguere dati che non sono linearmente separabili e può essere addestrato con l’algoritmo di propagazione dell’errore (“backpropagation”).

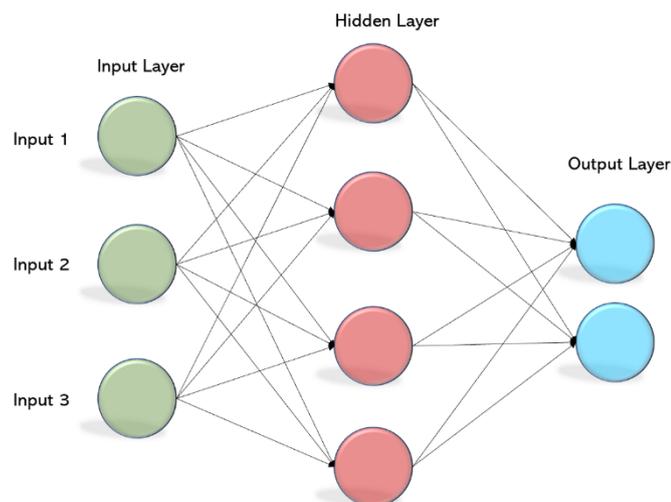


Figura 1. 3: esempio di Multilayer Perceptron

Fonte: <https://becominghuman.ai/multi-layer-perceptron-mlp-models-on-real-world-banking-data-f6dd3d7e998f>

Una rete neurale ben addestrata è in grado di svolgere compiti come la previsione di un valore di output, la classificazione di un oggetto, l'approssimazione di una funzione e il riconoscimento o il completamento di uno schema noto, e costituisce quindi un valido strumento matematico in molti campi soprattutto quello medico [7].

1.3 Reti Neurali Profonde e Deep Learning

Sottocategoria molto importante del Machine Learning è il Deep Learning (apprendimento profondo), alla base del quale sono presenti le reti neurali profonde (DNN), cioè reti neurali aventi un numero elevato di livelli di neuroni. Le reti neurali profonde che trovano maggior spazio al giorno d'oggi hanno un numero di livelli compreso tra 7 e 50, ma esistono reti che presentano oltre 150 livelli. Il numero di layer è direttamente proporzionale al miglioramento delle prestazioni e al livello di difficoltà del problema di riconoscimento. Quindi maggior numero di livelli implicano migliori prestazioni per problemi di riconoscimento complessi. Sottolineando i vantaggi apportati dal deep learning, e quindi dalle reti neurali, rispetto al machine learning: le reti neurali profonde scalano molto meglio con l'aumento dei dati, infatti a differenza che nel machine learning, nelle reti neurali profonde, non è ancora stata trovata la soglia per la quantità di dati oltre al quale le prestazioni saturano. Inoltre nel deep learning le feature vengono apprese dalla rete in modo automatico e non manualmente come nel machine learning. Le reti neurali profonde però necessitano di uno sforzo computazionale molto più elevato rispetto al machine learning.

Vi sono più tipologie di reti neurali profonde. Per questa tesi, in particolare, sono state utilizzate le reti neurali convoluzionali (CNN), che rientrano nei modelli feed-forward per problemi di classificazione di immagini.

1.5 - Reti Neurali Convoluzionali (CNN)

Una tipologia importante di rete neurale profonda è la Rete Neurale Convoluzionale (Convolutional Neural Network - CNN). Le CNN vengono utilizzate principalmente nella computer vision, cioè nel riconoscimento e identificazione delle immagini.

A differenza del Multilayer Perceptron (MLP) e delle reti neurali ordinarie, le CNN richiedono una pre-elaborazione dei dati molto minore, ogni neurone viene connesso solo a un numero ridotto di neuroni del livello precedente, detto campo recettivo, per cui avremo una forte riduzione del numero di connessioni, inoltre neuroni diversi dello stesso livello hanno lo stesso peso, individuano la stessa feature per locazioni diverse dell'immagine ed eseguono lo stesso tipo di elaborazione su porzioni diverse dell'input, che portano ad una

forte riduzione del numero di pesi. Quindi grazie alla diminuzione del numero di pesi e del numero di connessioni si avrà una rete molto più leggera e sarà quindi possibile aumentare la dimensionalità della rete.

L'architettura di una CNN è composta da una gerarchia di livelli: il layer di input è direttamente collegato ai pixel dell'immagine (non viene processata l'intera immagine ma una matrice di pixel), gli ultimi livelli sono generalmente fully-connected e operano come un classificatore MLP, mentre nei livelli intermedi si utilizzano connessioni locali e pesi condivisi per effettuare operazioni di convoluzione e riduzione del numero di parametri. I livelli di convoluzione estraggono delle caratteristiche (o feature) dalle immagini attraverso l'uso di filtri (o kernel) [4].

Un'immagine è una matrice di valori di pixel, avente tanti piani quanti sono i colori utilizzati per rappresentarla: un'immagine RGB presenterà tre piani, mentre un'immagine in bianco e nero avrà solo un piano. Un filtro digitale è fatto scorrere sulle diverse posizioni di input, per ogni sua posizione viene generato un output, eseguendo il prodotto scalare tra il filtro e la posizione dell'input coperta (entrambi trattati come vettori).

Il risultato della convoluzione è dato da:

$$net = \sum_{i=1}^n w_i * x_i \quad (1.3)$$

Dove w_i è il peso i-esimo del filtro e x_i è il valore i-esimo della porzione dell'input selezionata dal filtro.

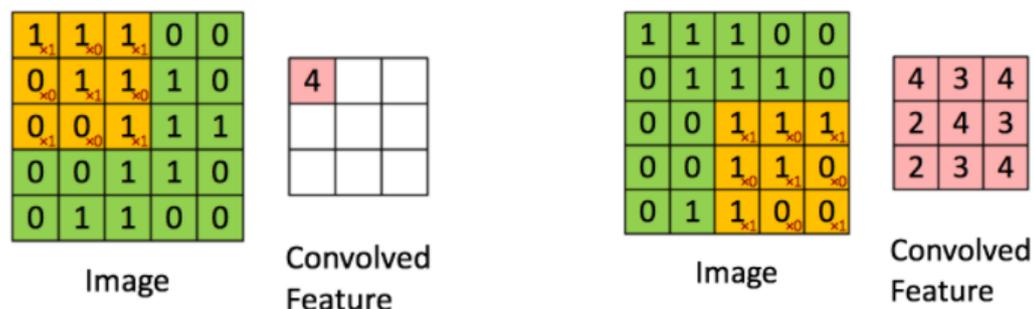


Figura 1.4: esempio di operazione di convoluzione

Fonte: <https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/>

Una CNN apprende da sola i valori dei filtri durante il training (addestramento), e maggiore è il numero di filtri a disposizione, più caratteristiche dell'immagine vengono estratte. Una CNN dunque studia e classifica il pattern basandosi su particolari caratteristiche.

I neuroni di ciascun livello di una CNN sono organizzati in volumi a tre dimensioni: nei piani width (larghezza) e height (altezza) vengono rappresentate le informazioni sulla struttura spaziale dell'immagine, invece nella terza dimensione depth (profondità) individua la Feature Map.

Il filtraggio crea come risultato una Feature Map la cui dimensione dipende da tre parametri:

- Profondità e dimensione del filtro: la profondità corrisponde al numero di filtri che vengono utilizzati per l'operazione di convoluzione.
- Stride: Stride è il numero di pixel con cui facciamo scorrere la nostra matrice di filtri sulla matrice di input. Utilizzare un passo maggiore riduce la dimensione delle Feature Map nel volume di output e riduce il numero di connessioni.
- Zero-Padding: un'ulteriore possibilità per regolare la dimensione della Feature Map è quella di aggiungere alla matrice di input degli zeri attorno al bordo. Con il parametro Padding si denota lo spessore (in pixel) del bordo

La funzione di attivazione utilizzata nelle CNN è la ReLU (Rectified Linear Unit), a differenza della sigmoide che viene utilizzata nelle reti MLP. Essa ci permette di risolvere il problema del vanish gradient e può essere scritta come:

$$f(net) = \max(0, net) \quad (1.4)$$

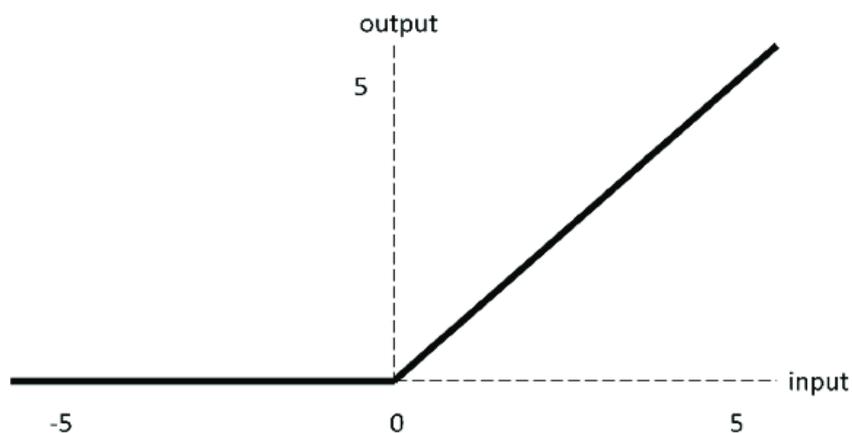


Figura 1.5: grafico funzione di attivazione ReLU

Il livello ReLU è inserito dopo i layer di convoluzione e rimuove tutti i valori negativi ponendoli a 0. Così facendo, la derivata vale 0 per i valori negativi o nulli di net e 1 per valori positivi [4].

CAPITOLO 2

Studio ed Elaborazione dei Dati

2.1 - Analisi dei dati

Le immagini raccolte per questa tesi ammontano a 969 e rappresentano ecografie di feti tra la ventesima e la ventiduesima settimana di gravidanza, cioè nelle settimane in cui il feto presenta organi abbastanza sviluppati da poter essere studiati.



Figura 2. 1: esempio di ecocardiografia sana del piano “4 camere” (IMG_20210608_2_3)

Ogni immagine viene rinominata dai medici con un pattern ricorrente:

IMG_AAAAMMGG_<IdFeto>_<nFoto>

dove:

- AAAA sta per Anno
- MM sta per Mese
- GG sta per Giorno
- IdFeto è un numero intero di due cifre che rappresenta le immagini di uno stesso feto
- nFoto è un numero intero che rappresenta il numero della immagine effettuata in quel giorno per quel feto

La maggior parte delle ecografie si presenta come in Figura 2.1, con a sinistra il parametro “GA” che indica il numero della settimana e il giorno della gravidanza (GA20w5d ad

esempio indica il 5 giorno della 20esima settimana) e a destra dei parametri specifici dell'ecografo.

Alcune immagini però, soprattutto tra quelle patologiche, sono state catturate dai medici in maniera anomala: non presentano scritte e bordo nero, e presentano anche un nome non esplicativo come le altre.



Figura 2. 2: esempio di ecografia patologica del piano "efflusso destro" che si presenta diversamente (LVOT 2)

Le ecografie raccolte sono state suddivise dai medici in feti sani e feti affetti da cardiopatie, le immagini di cuori sani ammontano a 907, invece quelle di cuori malati ammontano a 62. Sia per i campioni sani che per quelli malati sono state raccolte immagini per quattro piani ecografici standard:

	Tre vasi	Quattro camere	Efflusso destro	Efflusso sinistro	totale
Campioni sani	175	364	143	225	907
Campioni malati	11	30	4	17	62

Tabella 2.1: numero di ecografie per ogni piano ecografico standard, sia per i cuori sani che per i cuori malati

Per un totale di

- Tre vasi: 186 ecografie
- Quattro camere: 394 ecografie
- Efflusso destro: 147 ecografie
- Efflusso sinistro: 242 ecografie

Non tutti i feti però presentano ecografie per ogni piano standard, e in contrapposizione alcuni feti presentano più immagini per un singolo piano, infatti non abbiamo una distribuzione equa di immagini. Questo perché alcune tipologie di scansioni, come la tre vasi, non è sempre possibile ottenerle vista la difficoltà di acquisizione.

Inoltre, 81 immagini sono ecografie di tipo “color”, cioè presentano delle parti rilevanti dell’ecografie evidenziate.

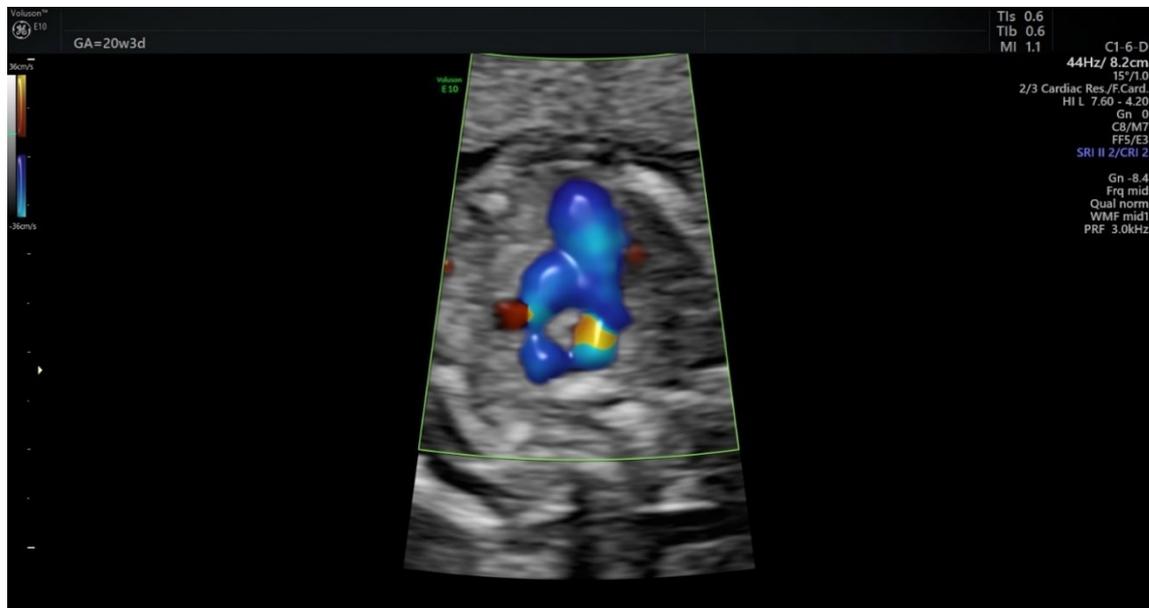


Figura 2. 3: esempio di ecografia patologica di tipo “color” del piano “3 vasi” (IMG_20210810_7_5)

2.2 - Processamento dei dati

Per poter processare le immagini è stato necessario ripulire le immagini da bordi e scritte. Sono stati utilizzati quindi due script MATLAB: la prima funzione ritaglia bordi e scritte e la seconda ritaglia in maniera più accurata solo il settore circolare dell’ecografia.

Quest’ultima funzione:

1. attraverso la funzione `imread()` individua i valori dei pixel dell’ecografia;
2. individua la prima riga e la prima colonna aventi pixel non neri;
3. individua l’ultima riga e l’ultima colonna aventi pixel non neri;
4. viene ritagliata l’immagine mantenendo la parte contenuta tra le righe e le colonne individuate nei punti 2) e 3).



Figure 2.4, 2.5, 2.6: esempio di ecografia sana del piano "4camere" (IMG_20210608_38_1) in alto , output della prima funzione in basso a sinistra e output della seconda funzione in basso a destra.

Nell'analizzare le immagini però, si è potuto notare che le funzioni di taglio utilizzate ritagliavano in maniera anomala alcune ecografie.

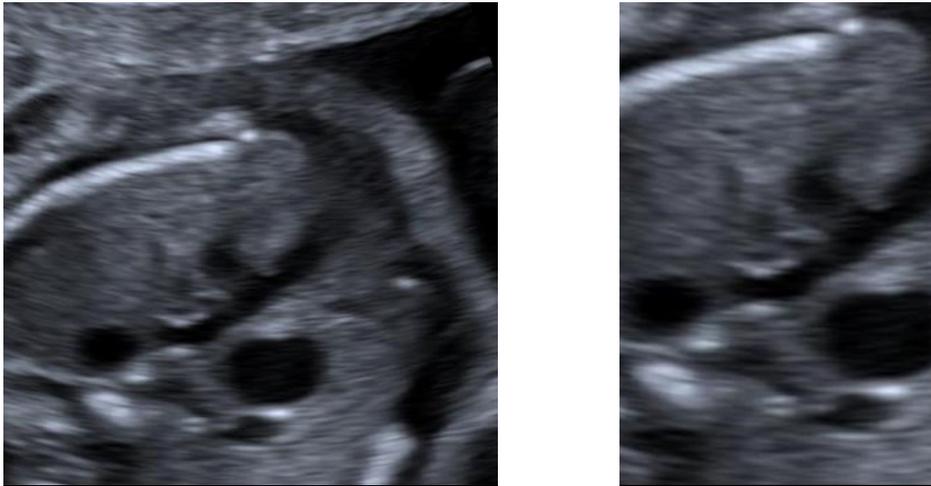


Figura 2.7: Esempio di immagine anomala (a sinistra), output dell'immagine dopo le due funzioni di taglio (a destra)

2.3 - Creazione dei dataset

Per l'addestramento della rete sono stati creati sette dataset con le seguenti caratteristiche:

1) Dataset 1 (SP0.33):

- Immagini utilizzate: tutte le 969 ecografie (62 anomale, 907 sane);
- Divisione dati: tre quarti delle immagini sono state usate nella fase di training e un quarto nella fase di test;
- Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.

2) Dataset 2 (SP0.33_NC):

- Immagini utilizzate: tutte le ecografie che non siano di tipo "color", per un totale di 888 ecografie (54 anomale, 834 sane);
- Divisione dati: tre quarti delle immagini sono state usate nella fase di training e un quarto nella fase di test;
- Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.

3) Dataset 3 (sSP0.33):

- Immagini utilizzate: un sottoinsieme delle immagini dei campioni sani (eccetto "cuori normali 1701") e tutte le immagini dei campioni malati, per un totale di 651 ecografie (62 anomale e 589 sane)

- Divisione dati: tre quarti delle immagini sono state usate nella fase di training e un quarto nella fase di test;
 - Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.
- 4) Dataset 4 (sSP0.33_NC):
- Immagini utilizzate: un sottoinsieme delle immagini dei campioni sani (eccetto “cuori normali 1701”) che non siano di tipologia “color” e tutte le immagini dei campioni malati che non siano di tipologia “color”, per un totale di 573 ecografie (54 anomale e 519 sane);
 - Divisione dati: tre quarti delle immagini sono state usate nella fase di training e un quarto nella fase di test;
 - Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.
- 5) Dataset 5 (sSP0.33_1°Funzione):
- Immagini utilizzate: un sottoinsieme delle immagini dei campioni sani (eccetto “cuori normali 1701”) e tutte le immagini dei campioni malati, per un totale di 651 ecografie (62 anomale e 589 sane);
 - Divisione dati: tre quarti delle immagini sono state usate nella fase di training e un quarto nella fase di test;
 - Funzioni utilizzate: è stata utilizzata soltanto la prima funzione per processare le immagini.
- 6) Dataset 6 (sSP0.5):
- Immagini utilizzate: un sottoinsieme delle immagini dei campioni sani (eccetto “cuori normali 1701”) e tutte le immagini dei campioni malati, per un totale di 651 ecografie (62 anomale e 589 sane);
 - Divisione dati: metà delle immagini sono state usate nella fase di training e metà nella fase di test;
 - Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.
- 7) Dataset 7 (sSP0.5_NC):
- Immagini utilizzate: un sottoinsieme delle immagini dei campioni sani (eccetto “cuori normali 1701”) che non siano di tipologia “color” e tutte le

immagini dei campioni malati che non siano di tipologia “color”, per un totale di 573 ecografie (54 anomale 519 sane);

- Divisione dati: metà delle immagini sono state usate nella fase di training e metà nella fase di test;
- funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.

Durante la divisione in dataset ecografie diverse dello stesso campione sono state inserite nello stesso set, o training o test, in modo da non alterare i risultati della rete.

Su consiglio dei medici e vista la scarsità delle immagini di tipo “color”, si è deciso di prendere in considerazione per l’addestramento solo i dataset 2,4 e 7.

A causa delle immagini anomale citate sopra (Figura 2.7), questi dataset sono stati poi ricreati con le stesse caratteristiche, ma le immagini ritagliate in maniera anomala sono state sostituite con le immagini originali. I nuovi dataset saranno individuati con gli acronimi SP0.33_NC_1, sSP0.33_NC_1, sSP0.5_NC_1.

Ogni dataset è rappresentato da due Cell Array realizzati in MATLAB, uno per il training set e uno per il test set. L’array ha un numero di celle pari al numero di ecografie che contiene e ogni cella è costituita da:

- L’ecografia processata dalla/e funzione/i;
- Un numero compreso tra uno e quattro che indica la tipologia di piano ecografico;
- Una label che indica se il campione considerato è sano o malato.

Sia il training set che il test set di ogni dataset comprende a sua volta un training set (o test set) per ogni piano ecografico, creando così quattro diverse CNN addestrabili. In questo modo, è stato possibile preparare un multiclassificatore composto da quattro CNN, ognuna delle quali ha la funzione di classificare le ecografie di un singolo piano standard.

2.4 - LIME (Local Interpretable Model-agnostic Explanations)

Per un’analisi più approfondita e per ulteriori studi da parte dei medici sono state passate le immagini di test del dataset sSP0.33 ad un codice MATLAB basato su LIME (Local Interpretable Model-agnostic Explanations) che mette in evidenza cosa la rete neurale considera importante durante la fase di test [8].

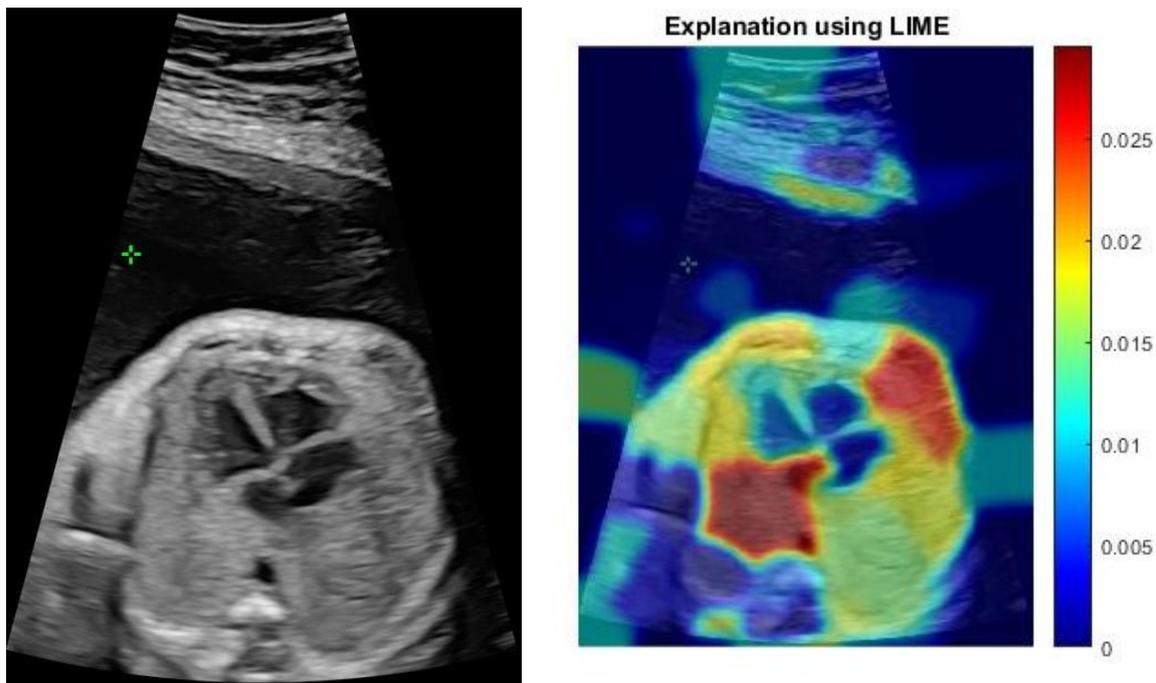


Figura 2.8: Esempio di immagine cuore sano a sinistra (IMG_20210818_23_4), output dell'immagine dopo la funzione a destra

2.6 - Descrizione della rete utilizzata e addestramento

La rete utilizzata per l'addestramento del sistema per il rilevamento delle cardiopatie è ResNet-50. Per questa tesi, è stato utilizzato un multiclassificatore composto da quattro reti ResNet-50 dove ogni rete serve a classificare le ecografie di un piano ecografico diverso.

La classificazione avviene compiendo i seguenti passi:

- Le ecografie fetali vengono date in input alle reti dei rispettivi piani ecografici;
- Calcolo della media dei valori output delle reti;
- Assegnazione della classe attraverso il valore più alto delle medie;

Per l'addestramento delle quattro CNN dei dataset è stato utilizzato il Transfer Learning prendendo come modello ResNet-50 (Residual Network a 50 livelli) e modificando i livelli di classificazione, che nel nostro caso saranno solo Sano e Malato. Con il Transfer Learning è stato possibile ridurre notevolmente i tempi di addestramento sfruttando i pesi preaddestrati e adattandoli al problema in esame.

Il valore di Batch Size è stato posto a 30 e il learning rate a 0.001.

Per l'addestramento della rete sono stati utilizzati solo i dataset SP0.33_NC, sSP0.33_NC, sSP0.5_NC nel primo caso e SP0.33_NC_1, sSP0.33_NC_1, sSP0.5_NC_1 nel secondo.

CAPITOLO 3

Risultati ottenuti

3.1 - Studio dei risultati

Per misurare le prestazioni delle reti è stata utilizzata la curva ROC (“Receiver Operating Characteristic”), cioè una curva che misura l’accuratezza con cui la rete è in grado di classificare i pattern a partire dai Falsi Positivi (FP) nell’asse delle ascisse e i Veri Positivi (TP) nell’asse delle ordinate.

Un Falso Positivo (FP) è un pattern negativo che è stato erroneamente assegnato ai positivi, quindi nel problema in esame è rappresentato un campione sano che viene classificato come malato.

Un Vero Positivo (TP) è un pattern positivo che è stato correttamente assegnato ai positivi, quindi nel problema in esame è rappresentato da un campione malato che è stato classificato come tale.

Il tasso di falsi positivi è dato da:

$$FalsePositiveRate(FPR) = \frac{NumberOfFalsePositive}{AllPositive} \quad (3.1)$$

Invece il tasso di veri positivi è dato invece da:

$$TruePositiveRate(TPR) = \frac{NumberOfTruePositive}{AllPositive} \quad (3.2)$$

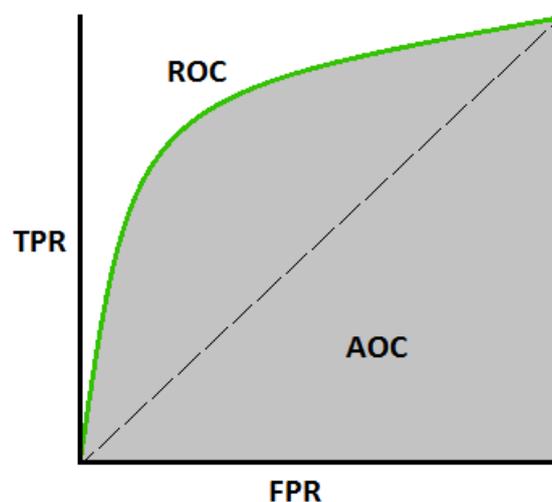


Figura 3.1 : esempio di curva ROC 1

Fonte: <https://www.domsoria.com/wp-conte-1>

L'area sotto la curva ROC detta AUC (Area Under The Curve) è una misura dell'accuratezza diagnostica. L'area sotto la curva può assumere valori tra 0 e 1 e tanto maggiore è l'AUC e tanto maggiore è il potere discriminante del test.

Per questa tesi è stato utilizzato il valore AUC per valutare le prestazioni delle reti.

3.2 - Risultati ottenuti e confronto con i dataset precedenti

I dataset analizzati precedentemente a questa tesi presentano immagini fino ad Agosto 2021. Evidenziando solo quelli utili ad un confronto, sono stati divisi in:

- SP0.33_NC_old:
 - Immagini utilizzate: tutte le ecografie che non siano di tipo “color”, per un totale di 563 ecografie;
 - Divisione dati: tre quarti delle immagini sono state usate nella fase di training e un quarto nella fase di test;
 - Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.
- sSP0.33_NC_old:
 - Immagini utilizzate: un sottoinsieme delle immagini dei campioni sani eccetto "immagini nuove normali" e senza le immagini “color” e tutte le immagini dei campioni malati che non siano di tipologia “color” per un totale di 371;
 - Divisione dati: tre quarti delle immagini sono state utilizzate nella fase di training e un quarto nella fase di test;
 - Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini.
- sSP0.5_NC_old:
 - Immagini utilizzate: un sottoinsieme delle immagini dei campioni sani eccetto "immagini nuove normali" e senza le immagini “color” e tutte le immagini dei campioni malati che non siano di tipologia “color” per un totale di 371;
 - Divisione dati: metà delle immagini sono state utilizzate nella fase di training e metà nella fase di test;
 - Funzioni utilizzate: sono state utilizzate entrambe le funzioni per il processamento delle immagini. [9][10].

Quindi riassumendo:

DATASET	Numero Immagini Cuori Sani		Numero Immagini Cuori Patologici		Totale immagini
SP0.33_NC_old	518		45		563
	3 vasi	74	3 vasi	4	
	4 Camere:	232	4 Camere:	26	
	Efflusso dx:	78	Efflusso dx:	4	
	Efflusso sx:	134	Efflusso sx:	11	
sSP0.33_NC_old	326		371		371
	3 vasi	38	3 vasi	4	
	4 Camere:	150	4 Camere:	26	
	Efflusso dx:	42	Efflusso dx:	4	
	Efflusso sx:	96	Efflusso sx:	11	
sSP0.5_NC_old	326		45		371
	3 vasi	38	3 vasi	4	
	4 Camere:	150	4 Camere:	26	
	Efflusso dx:	42	Efflusso dx:	4	
	Efflusso sx:	96	Efflusso sx:	11	

Tabella 3.1: Numero immagini per ogni piano di ogni dataset precedente

L'addestramento di questi dataset ha riportato come risultati:

DATASET	AUC (%)			
	Tre vasi	Quattro Camere	Efflusso destro	Efflusso sinistro
SP0.33_NC_old	90.00	80.15	86.11	98.53
sSP0.33_NC_old	60.00	90.71	90.00	80.77
sSP0.5_NC_old	92.31	86.50	64.29	75.00

Tabella 3.2: Valori di AUC in percentuale ottenuti da addestramenti delle quattro CNN di dataset precedenti

Invece i nuovi dataset utilizzati per l'addestramento della CNN hanno la seguente divisione delle ecografie:

DATASET	Numero Immagini Cuori Sani		Numero Immagini Cuori Patologici		totale immagini
sSP0.33	834		54		888
	3 vasi	113	3 vasi:	4	
	4 Camere:	361	4 Camere:	29	
	Efflusso dx:	142	Efflusso dx:	4	
	Efflusso sx:	218	Efflusso sx:	17	
sSP0.33_NC	519		54		573
	3 vasi	74	3 vasi	4	
	4 Camere:	232	4 Camere:	29	
	Efflusso dx:	78	Efflusso dx:	4	
	Efflusso sx:	135	Efflusso sx:	17	
sSP0.5_NC	519		54		573
	3 vasi	74	3 vasi	4	
	4 Camere:	232	4 Camere:	29	
	Efflusso dx:	78	Efflusso dx:	4	
	Efflusso sx:	135	Efflusso sx:	17	

Tabella 3.3: Numero immagini per ogni piano di ogni dataset

E l'addestramento delle quattro nuove CNN di ogni dataset ha riportato nella prima prova i seguenti risultati:

DATASET	AUC (%)			
	Tre vasi	Quattro Camere	Efflusso destro	Efflusso sinistro
SP0.33_NC	93.94	75.68	84.72	95.54
sSP0.33_NC	25.00	84.50	88.33	92.14
sSP0.5_NC	4.05	72.39	75.71	89.14

Tabella 3.4: Valori di AUC in percentuale ottenuti dall'addestramento delle quattro CNN di ogni dataset nella prima prova

E nella seconda prova i seguenti risultati:

DATASET	AUC (%)			
	Tre vasi	Quattro Camere	Efflusso destro	Efflusso sinistro
SP0.33_NC_1	51.89	89.79	93.97	93.94
sSP0.33_NC_1	90.00	82.32	88.89	90.00
sSP0.5_NC_1	39.19	85.71	87.14	88.73

Tabella 3.5: Valori di AUC in percentuale ottenuti dall'addestramento delle quattro CNN di ogni dataset nella seconda prova

3.3 - Considerazioni sui risultati ottenuti

Dalle tabelle 3.4 e 3.5 si può notare che l'addestramento di alcune reti ha portato a buoni valori di AUC. In particolare, si può notare che in analogia a quanto riportato nei dataset precedenti a questa tesi, il piano delle quattro camere ha ottenuto i valori più consistenti nei diversi dataset. Questi risultati possono avere spiegazione nel fatto che il piano ecografico delle quattro camere è quello che presenta il maggior numero di ecografie: le scansioni di questo piano, infatti, ammontano a 364 per i cuori sani e 30 per i cuori malati (vedi tabella 2.1), cioè circa il 40% della totalità delle ecografie dei campioni sani e circa la metà della totalità delle ecografie di campioni affetti da cardiopatie. Gli altri piani, tre vasi, efflusso destro e efflusso sinistro presentano oscillazioni maggiori tra dataset diversi, avendo un numero più ridotto di scansioni. Si può notare, infatti, che ad esempio i valori di AUC per il piano dei tre vasi hanno un valore minimo pari a 4.05 nel primo caso (e 39.19 nel secondo caso) e un valore massimo pari a 93.94 nel primo caso (e 90.00 nel secondo). Il piano dell'efflusso destro e dell'efflusso sinistro invece presentano meno oscillazione tra i risultati, ma presentano valori di AUC abbastanza alti per le poche immagini a disposizione [9][10].

Confrontando i risultati ottenuti dai dataset precedenti, si può notare come l'aumento del numero delle immagini non ha portato cambiamenti molto significativi nei valori di AUC ottenuti. Si può notare in media una lieve diminuzione dei valori di AUC, più accentuata nel primo caso.

Dalle immagini passate al codice MATLAB basato su LIME, si è potuto notare, grazie all'aiuto dei medici, che ciò che la rete neurale considera importante, nella maggior parte dei casi, non è la parte del cuore che l'immagine del piano ecografico rappresenta (vedi Figura

2.8). Inoltre ci sono alcuni casi estremi in cui ciò che viene messo in evidenza dalla macchina non è nemmeno una parte del cuore, ma il liquido amniotico o una parte del torace.

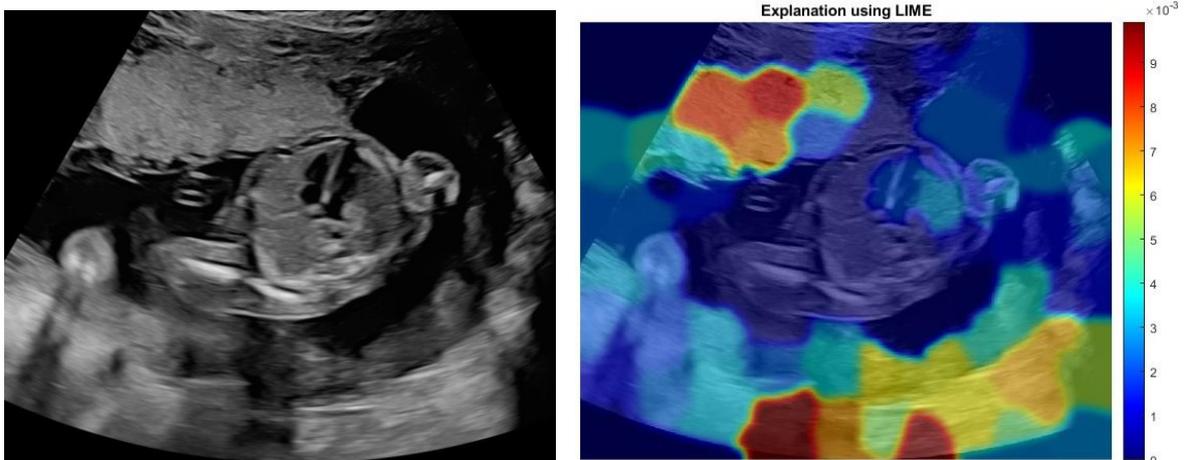


Figura 3.2 : Esempio di immagine '4 camere sano' (IMG_20210608_6_4) dove si può notare che ciò viene messo in evidenza dalla macchina non è una parte del cuore ma il liquido amniotico e il torace.

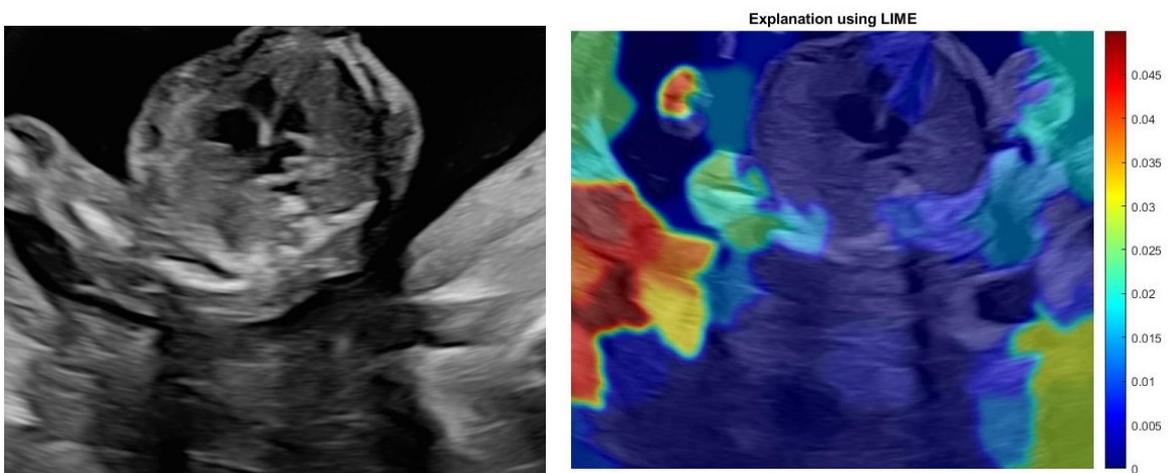


Figura 3.3 : Ulteriore esempio di immagine '4 camere sano' (IMG_20210610_23_2) dove si può notare che ciò viene messo in evidenza dalla macchina non è una parte del cuore ma il liquido amniotico e il torace.

CONCLUSIONI

Questa tesi ha lo scopo di fornire un'applicazione di deep learning, basata su CNN, mirata ad assistere i medici nella diagnosi prenatale di cardiopatie.

I risultati di AUC ottenuti mostrano una buona prestazione della rete, ma non possiamo garantirne l'affidabilità a causa:

- del numero molto basso di ecografie di feti affetti da cardiopatie;
- della disparità tra il numero di immagini per ogni piano ecografico;
- della modalità casuale di individuazione delle aree importanti da parte della rete neurale.

Sarà pertanto necessario, al fine di poter decretare l'attendibilità dei risultati, effettuare numerosi altri test su un numero maggiore di immagini di campioni malati e raccogliere i dati in modo che ci sia almeno un'immagine per piano ecografico per ogni feto. Inoltre sarà essenziale trovare una soluzione al problema di rilevazione del cuore e delle parti importanti da parte dell'intelligenza artificiale, effettuando numerosi altri test e addestramenti.

RIFERIMENTI

- [1] Rik Vullings, *Fetal Electrocardiography and Deep Learning for Prenatal Detection of Congenital Heart Disease*, Eindhoven University of Technology, Eindhoven, The Netherlands.
- [2] van Velzen CL, Clur SA, Rijlaarsdam ME, Bax CJ, Pajkrt E, Heymans MW, Bekker MN, Hrudá J, de Groot CJ, Blom NA, Haak MC. *Prenatal detection of congenital heart disease--results of a national screening programme*. BJOG. 2016 Feb;123(3):400-7. doi: 10.1111/1471-0528.13274. Epub 2015 Jan 27. PMID: 25625301.
- [3] ARCA - Associazione Ricerche Cardiopatie Aritmiche, “*Giornata Mondiale delle Cardiopatie Congenite 2021 - Paola Veronese*”. YouTube, 14 febbraio 2021, video, https://www.youtube.com/watch?v=A0cW_d_yYOM .
- [4] Davide Maltoni, *Machine Learning*, Università di Bologna: (http://bias.csr.unibo.it/maltoni/ml/DispensePDF/1_ML_Introduzione.pdf).
- [5] MathWorks, *Cos'è una rete neurale?* (<https://it.mathworks.com/discovery/neural-network.html>).
- [6] Bizopoulos P, Koutsouris D. *Apprendimento profondo in cardiologia. IEEE Rev Biomed Ing* . 2019; 12 (c): 168-193. 10.1109/RBME.2018.2885714.
- [7] Francesco Rigobello, *NEURAL NETWORKS FOR MEDICAL DECISION, tesi di laurea triennale*, Università degli Studi di Padova, Facoltà di ingegneria, Corso di Laurea in Ingegneria dell'Informazione, 27 Settembre 2012, G.M. Toffolo.
- [8] Kenta (2022). *Explainable AI: interpreting the classification using LIME* (<https://github.com/KentaItakura/Explainable-AI-interpreting-the-classification-performed-by-deep-learning-with-LIME-using-MATLAB/releases/tag/v2.1>), GitHub. Retrieved February 25, 2022.

- [9] Edoardo Bastianello, *Deep Learning ed Ecocardiografia Fetale per Riconoscimento di Cardiopatie Prenatale*, tesi di laurea triennale, Università degli Studi di Padova, Dipartimento di ingegneria dell'informazione, Corso di Laurea in Ingegneria Informatica, 21/09/2021, Loris Nanni.
- [10] Zebin Duan, *Identificazione di Cardiopatie Prenatali usando IA*, tesi di laurea triennale, Università degli Studi di Padova, Dipartimento di ingegneria dell'informazione, Corso di laurea in Ingegneria Informatica, 22/11/2021, Loris Nanni.