



UNIVERSITÀ
DEGLI STUDI
DI PADOVA

UNIVERSITA' DEGLI STUDI DI PADOVA

Dipartimento di Ingegneria Industriale DII

Corso di Laurea in Ingegneria Meccanica

Sviluppo di una metodologia di "smart monitoring" di tipo non
intrusivo per carichi elettrici tramite l'utilizzo di intelligenza
artificiale

Prof.ssa Giovanna Cavazzini

Laureando: Alessandro Morello mat. 1205370

Anno Accademico 2021/2022

INDICE

Sommario	pag.3
Introduzione	pag.5
Capitolo 1: Raccolta ed elaborazione dati	pag.9
1.1 Raccolta dati da database	pag.9
1.2 Raccolta dati di consumo cumulati	pag.12
1.3 Elaborazione dati attraverso il software MATLAB	pag.13
Capitolo 2: Scelta della rete neurale e post-elaborazione dati	pag.17
2.1 Le reti neurali per il riconoscimento delle immagini	pag.17
2.2 Post-processing degli spettrogrammi e creazione dei database	pag.21
2.3 Creazione delle anchor box	pag.22
Capitolo 3: Training della rete neurale	pag.25
3.1 Opzioni di addestramento	pag.25
3.2 Plot dell'addestramento e prima verifica dei risultati	pag.27
3.3 Processo decisionale sui parametri di training	pag.31
Capitolo 4: Analisi segnali dei singoli elettrodomestici	pag.33
4.1 Segnali di lunga durata	pag.33
4.2 Segnali di durata medio-breve	pag.36
Capitolo 5: Sviluppo delle reti neurali	pag.45
5.1 Rete neurale per intervalli lunghi (4 ore)	pag.45
5.2 Rete neurale per intervalli brevi (1 ora)	pag.55
5.3 Analisi detection reti combinate	pag.59
Capitolo 6: Analisi risultati e confronto con rete Zorzi	pag.65
6.1 Analisi risultati	pag.65
6.2 Confronto con risultati degli algoritmi di disaggregazione	pag.75
6.3 Confronto con i risultati della rete Zorzi	pag.79
6.4 Confronto con i risultati ottenuti con un sistema commerciale (NED)	pag.80
6.5 Pro e contro dell'utilizzo della rete YOLO	pag.81
Conclusioni	pag.83
Bibliografia	pag.87

SOMMARIO

La tesi si propone come obiettivo lo sviluppo di una metodologia di “smart monitoring” di carichi elettrici (in ambito residenziale) mediante l'utilizzo di reti neurali. In particolare, si è puntato sull'utilizzo di reti neurali per il riconoscimento delle immagini in modo tale da individuare l'immagine dello spettro dei singoli elettrodomestici all'interno dello spettrogramma del segnale cumulato ottenuto da contatori domestici. Lo sviluppo di una metodologia efficace porterebbe ad ottenere queste informazioni in maniera relativamente semplice, senza la necessità di usare particolare sensoristica, risultando quindi fruibile e a basso costo. L'obiettivo secondario è quello di ottenere il maggior numero di informazioni utili possibili oltre al riconoscimento dei carichi: l'intervallo di funzionamento dell'elettrodomestico e conseguentemente il consumo totale dello stesso o comunque dare una linea per lo sviluppo di queste applicazioni.

Lo svolgimento dell'intero lavoro è stato possibile grazie alla fruizione di database dei carichi elettrici disponibili in internet (ad es. database ECO, tracebase, UK-Dale) e ottenuti dalle rilevazioni effettuate tramite il prototipo NILM UniPD-Lab.E.E. sviluppato in un lavoro precedente dal dott. Gianmarco Zorzi. Sono stati acquisiti segnali di consumo aggregato attraverso sistemi di monitoraggio installati in una utenza domestica. Gli spettrogrammi relativi ai segnali dei singoli elettrodomestici e ai segnali aggregati sono stati ottenuti tramite l'apposita funzione del software MATLAB, sul quale si è operato anche per lo sviluppo della rete neurale.

Una volta ottenuti i dati, si è proceduto con l'analisi delle varie reti neurali disponibili al fine di scegliere la migliore per lo scopo. Dopodiché si è andati a fare una prima prova sulla rete per verificare la bontà della scelta, identificare le problematiche da affrontare e porsi un target minimo realistico riguardo la precisione della rete stessa. Si è poi andati ad affinare il database degli spettrogrammi e i parametri di training della rete neurale (con un ragionevole rapporto tra precisione e tempo impiegato) per ottenere la miglior efficienza possibile nei limiti dello strumento e valutare limiti e possibili sviluppi della metodologia proposta. Si è posto come target il riconoscimento di almeno il 70% dei segnali, considerando i limiti della rete e quelli dei dati a nostra disposizione, che per alcuni elettrodomestici non sono molto numerosi.

Infine, raggiunti gli obiettivi relativi al riconoscimento, sono stati confrontati i risultati con quelli ottenuti dalla rete sviluppata da Zorzi. Il confronto ha permesso di analizzare le strade per la possibile implementazione tra i due metodi e, in generale, come migliorare ulteriormente i risultati ottenuti; questo consentirebbe di rendere fruibile ad un più ampio pubblico tale metodologia.

Introduzione

Quando si parla di tecnologie e dispositivi NILM (Nonintrusive Load Monitor) ci si riferisce a strumenti che dalla curva di consumo aggregata di un edificio riescono a identificare i singoli carichi che la compongono. Questa tipologia di tecniche viene definita non intrusiva per differenziarla dalle tecniche che prevedono l'applicazione di sensori sui singoli elettrodomestici^[1], in quanto questa non prevede l'utilizzo di alcuno strumento supplementare rispetto alla normale dotazione dell'impianto elettrico. L'invenzione del NILM risale all'inizio degli anni '80, grazie al lavoro svolto al MIT da George W. Hart, Ed Kern e Fred Schweppe^[2]. Lavoro che ha portato nel 1986 alla registrazione del primo brevetto, in cui si specifica che rientrano nella definizione di NILM tutte quelle misure di resistenze, reattanze, impedenze e in generale tutte le caratteristiche a due poli che si possono registrare nei circuiti^[3]. Così facendo si va, quindi, ad includere qualsiasi tipo di misurazione dell'energia, sia essa attiva, che noi prenderemo in considerazione e che è quella che noi effettivamente consumiamo (e paghiamo) all'interno delle nostre abitazioni, reattiva o apparente, che assumono, invece, rilevanza maggiore in ambito industriale o comunque in applicazioni che prevedano importante utilizzo di induttanze, come i motori elettrici.

Le motivazioni che spingono ad un continuo sviluppo di dispositivi NILM sempre più accurati sono da ricercare nella necessità nel mondo moderno di ricercare la maggiore efficienza energetica possibile, controllare i consumi ed evitare gli sprechi, anche migliorando le cattive abitudini delle persone, che sovente sono in contrasto con gli obiettivi appena citati. Il fatto di identificare i carichi elettrici dalla curva di consumo aggregata apre un mondo su tutto ciò che riguarda la corretta gestione energetica delle nostre case. In primis, permette di identificare quali elettrodomestici causano il maggior consumo e quindi di correggere le nostre abitudini nel loro utilizzo al fine di limitare sprechi e costi. Ad esempio, pensiamo ad uno degli elettrodomestici più diffusi ed energivori presenti nelle nostre abitazioni: la lavatrice. Il monitoraggio del suo consumo, e conseguentemente del costo dell'energia da essa consumata, può aiutare l'utilizzatore a modificare le sue abitudini, evitando, ad esempio, lavaggi non necessari o spostandoli in orari in cui il costo dell'energia è minore (solitamente le ore notturne). Un'altra applicazione correlata all'efficienza energetica è quella della gestione delle operazioni di manutenzione. Attraverso il riconoscimento dei segnali dei vari elettrodomestici, in un secondo momento, si possono analizzare quei segnali cercando di identificare le eventuali anomalie presenti all'interno degli stessi. Le anomalie comprendono consumi medi più alti o forme della curva diverse dal normale, cioè tutti gli indicatori di guasti o usura che possono avvisare l'utilizzatore della necessità di intervenire con la manutenzione, evitando sprechi o addirittura rotture per le quali non riusciremmo ad avere avvisaglie.

Questi algoritmi, oltre ai campi dell'efficienza e del risparmio energetico, trovano notevole utilizzo anche in altri ambiti. Queste sono in particolare legate al mondo delle cosiddette "ambient assisted residence"^[4], ossia tutte quelle applicazioni che servono a rendere più semplice l'utilizzo dei dispositivi presenti all'interno di una abitazione da parte di persone che non sono autonome, ad esempio anziani, disabili o persone affette da particolari disturbi. Le applicazioni in questo campo vanno dal monitoraggio delle attività svolte all'aiuto alla persona nell'utilizzo degli elettrodomestici, integrando un dispositivo NILM all'interno di un sistema di domotica.

Il processo alla base del NILM è quello della disaggregazione del segnale, ossia la ripartizione del segnale aggregato nei suoi componenti elementari, ossia i singoli segnali dei dispositivi elettrici che si trovano all'interno dell'utenza considerata. Il primo metodo, proposto da Hart, consisteva essenzialmente nella ricerca delle "firme" degli elettrodomestici presenti all'interno dell'aggregato, dove per firma intendiamo una caratteristica distintiva della curva di consumo. L'obiettivo finale è di riconoscere correttamente tutti i segnali che compongono l'aggregato senza la necessità di utilizzare alcun tipo di dispositivo, se non quello che registra i dati del contatore (quindi da considerare esterno all'utenza).

Rispetto agli anni '80, quando Hart propose i primi metodi di disaggregazione del segnale, la tecnologia ha fatto passi da gigante, consentendo misurazioni sempre più precise dei consumi effettuate con strumenti sempre più semplici e alla portata di tutti. La loro diffusione ha portato addirittura alla pubblicazione online di appositi toolkit scaricabili gratuitamente. Parallelamente sono stati creati o sviluppati anche gli algoritmi di disaggregazione e implementati nuovi metodi come quelli legati al deep learning e alle reti neurali, che, uniti a potenze di calcolo sempre maggiori, consentono di migliorare l'efficienza degli algoritmi e anche la loro velocità. Le tecniche utilizzate per

la disaggregazione dei carichi oggi sono numerose e molto diverse tra loro. Le principali categorie le possiamo vedere nella seguente immagine:

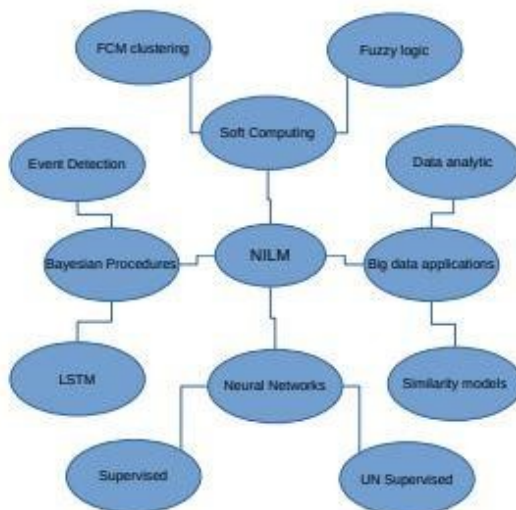


Figura i.1: schema che rappresenta le maggiori tecniche per la disaggregazione dei segnali in ambito NILM^[5]

La parte più interessante per il lavoro svolto è, sicuramente, quella relativa alle reti neurali. Questa è la tecnica che dà i migliori risultati sull'identificazione dei segnali sui quali non si hanno informazioni. Gli algoritmi più frequentemente utilizzati sono quelli delle recurrent neural network e LSTM (Long Short Term Memory). Nelle RNN i neuroni sono collegati tra loro in loop e i risultati ottenuti dai neuroni di output vengono salvati e riutilizzati nei neuroni di input per ottenere successivamente delle prediction su nuovi dati. Questo tipo di rete è comunemente utilizzata, ad esempio, per il riconoscimento vocale. Un importante difetto di queste reti è che nel lungo termine le prediction saranno basate su una quantità molto elevata di dati di output e i pesi della rete risulteranno sempre meno influenzati dalle nuove prediction. Di conseguenza, è difficile aggiornare la rete o correggere gli errori. Per sopperire, almeno in parte, a questo inconveniente, si introducono delle "celle di memoria" che decidono la quantità di informazioni da mantenere e quella da scartare. Queste celle sono chiamate Long Short Term Memory e sono quelle che caratterizzano le reti omonime. Questo tipo di algoritmo viene classificato anche come bayesiano, in quanto le probabilità del verificarsi di un evento non sono interpretate come frequenze, proporzioni o concetti simili, come avviene nella statistica classica, ma come livello di fiducia^[6]. Un altro tipo di rete utilizzata sono i cosiddetti autoencoder. Gli autoencoder sono stati introdotti all'inizio degli anni '80 per valutare la bontà del modello di addestramento non supervisionato. La rete, in sintesi, ha lo scopo di riprodurre i dati in input al fine di individuare le caratteristiche che meglio rappresentano il fenomeno di interesse^[7].

Le altre categorie, meno rilevanti ai fini di questo elaborato, sono il Soft Computing e l'applicazione di algoritmi basati sui cosiddetti Big-data. Le tecniche di soft computing uniscono le capacità di ottimizzazione degli algoritmi genetici, le capacità di apprendimento delle reti neurali e la capacità di modellare ed utilizzare insiemi incerti e complessi secondo la teoria degli insiemi fuzzy^[8]. Questa teoria, definita nel 1965 da Lotfi Zadeh, professore delle University of California, ridefinisce il concetto di appartenenza ad un insieme, rendendolo meno definito, legandolo ad un punteggio che vada 0 (non appartenente) a 1 (completamente appartenente) e dando quindi la possibilità agli elementi che non hanno tutte le caratteristiche proprie di quell'insieme di esserne parte in maniera parziale^[9]. Attraverso l'utilizzo di queste tecniche, come detto dal loro creatore Zadeh, ci si propone di riuscire ad operare meglio in tutti quegli ambiti in cui è necessario lasciare un margine di incertezza nelle definizioni e nelle operazioni svolte sui dati; ciò si traduce in una riproduzione o considerazione delle incertezze presenti nei dati reali in modo tale da irrobustire gli algoritmi e le reti neurali. Questi algoritmi possono consentire, quindi, un miglioramento nella classificazione di segnali simili appartenenti a elettrodomestici diversi o nella classificazione di segnali di elettrodomestici appartenenti a classi conosciute dalla rete, ma le cui curve non sono presenti nel database di addestramento.

Per quanto non esista una definizione universalmente riconosciuta di “big data”, possiamo dire che si tratta di database di grandi dimensioni con al loro interno molte categorie e strutture complesse. L’analisi dei big data è atta ad estrapolare informazioni utili allo scopo. Nel caso del NILM si applicano queste tecniche di analisi ai dati ottenuti da impianti elettrici con un layer che comunica tra l’impianto e il sistema di elaborazione dei dati. L’obiettivo è quello di ottenere ed elaborare una quantità molto elevata e variegata di dati nel minor tempo possibile (persino in tempo reale) e in maniera efficace, trovando correlazioni e strutture definite al loro interno per identificare i singoli carichi. Inoltre, questo apre all’utilizzo di dispositivi NILM anche per applicazioni più complesse come quelle industriali o relative a interi gruppi abitativi. Al momento però le tecniche di analisi dei big data non sono ancora state sfruttate in questo ambito in quanto ci sono ancora diversi problemi da risolvere dalle tecniche utilizzate alle sinergie^[10].

In generale, abbiamo visto come le tecniche più utilizzate prevedano l’utilizzo di algoritmi ricorrenti o di tipo LSTM per la disaggregazione dei dati e gli stessi algoritmi sono implementati in apposite reti neurali quando si passa ad una tecnica più avanzata. Nell’approccio che verrà affrontato all’interno dell’elaborato si cercherà di utilizzare un tipo di rete totalmente differente rispetto a quelle viste in questo capitolo. L’idea alla base del lavoro è quella di riuscire a identificare gli spettri dei singoli carichi all’interno dell’immagine dello spettro di potenza ottenuta dall’elaborazione del segnale aggregato. L’immagine che rappresenta lo spettro di potenza è detta spettrogramma e rappresenta l’andamento nel tempo della potenza assorbita alle varie frequenze di un segnale, nel nostro caso una curva di potenza attiva assorbita. Le novità rispetto alle reti più classiche sono essenzialmente due. La prima riguarda la struttura vera e propria della rete. Se solitamente si utilizzano reti che si basano su cicli in loop in cui i risultati vengono rimandati ai neuroni degli input, la YOLO (You Only Look Once) è una rete di tipo feed-forward, in cui i dati in ingresso viaggiano solamente in avanti e non vanno ad influenzare le elaborazioni successive. In questo modo si elimina il rischio che gli output della rete vengano influenzati da dati errati dovuti a errate classificazioni precedenti e consente di poter aggiornare i pesi della rete in maniera più agevole, senza il contributo degli output già generati. La seconda differenza consiste nel fatto che non si va ad applicare nessun tipo di algoritmo direttamente sulla curva aggregata, quindi non essendoci una disaggregazione vera e propria dei dati non c’è la necessità di ricostruire i segnali, con le conseguenti possibilità di errore legate ad una qualunque elaborazione, e si può risalire e operare direttamente sul segnale originale per ottenere altre informazioni (come i consumi o i costi dei singoli elettrodomestici) o effettuare altre analisi (come la ricerca di anomalie). Di contro, però, per effettuare qualsiasi di queste operazioni è necessario almeno un passaggio aggiuntivo per ricondurre lo spettrogramma dell’elettrodomestico alla rispettiva porzione di curva e ai relativi dati di potenza e tempo e solleva molti quesiti riguardo la capacità di identificare correttamente segnali contemporanei, che andremo ad analizzare all’interno dell’elaborato.

Allo stato attuale i dispositivi NILM veri e propri in commercio non sono moltissimi e hanno ancora grandi margini di miglioramento. Inoltre, sono presenti molti dispositivi che vengono identificati come NILM, ma che in realtà non sono fedeli a quella che è la definizione originaria del termine. Molti di questi, infatti, sono in realtà strumenti che vanno semplicemente a monitorare e registrare il consumo energetico di singoli elettrodomestici ai quali vengono collegati. Questi dispositivi sono molto accurati nelle misurazioni, sia di potenza attiva che reattiva, possono avere anche altre funzionalità molto utili e possono essere connessi con altri dispositivi, sia dello stesso tipo che esterni come Alexa, magari all’interno di applicazioni di domotica. Un esempio è lo Shelly EM[11], un congegno che viene collegato fino a due elettrodomestici contemporaneamente, ciascuno attraverso una pinza amperometrica, potendo misurare i valori di potenza attiva e reattiva, fattore di potenza, corrente, tensione e lo storico di questi ultimi due dati. Dà la possibilità di interrompere la corrente all’elettrodomestico su cui effettua le misurazioni sia in modo manuale che automaticamente quando supera una certa soglia preimpostata di potenza assorbita. Inoltre, lo si può connettere con altri dispositivi Shelly o di automazione domestica (come Alexa, Google Assistant e molti altri)^[11]. Questo dispositivo può sicuramente essere utile nel monitoraggio dei carichi, ma non rientra completamente nella definizione di NILM che abbiamo fornito all’inizio di questo capitolo.

Al contrario dell’esempio appena riportato, in commercio esistono dispositivi che corrispondono a pieno alla definizione di NILM. Questi, però, presentano ancora diversi inconvenienti. Il primo, non banale nell’ottica della commercializzazione verso un utente medio, è il prezzo, che spesso è superiore al centinaio di euro. Il secondo è legato alla fruibilità del prodotto. Molti presentano infatti un’interfaccia complessa, pienamente comprensibile da chi ha effettivamente delle competenze in questo ambito, ma non dall’utilizzatore generico, rendendo quindi difficile sfruttare in maniera

ottimale il prodotto. I dispositivi economici e semplici dal punto di vista dell'utilizzo presenti sul mercato non sono molti e come tutto ciò che riguarda queste tecnologie ha ancora notevoli margini di sviluppo. Uno di questi dispositivi è NED, sviluppato da una azienda torinese (Midori), di cui disponiamo di risultati a test effettuati su una abitazione grazie al lavoro precedentemente svolto da Zorzi. Questo dispositivo, pur disponendo di un dispositivo di campionamento ad alta frequenza, dà sì buoni risultati, ulteriormente migliorabili grazie alla possibilità di identificare manualmente i carichi da parte dell'utente, ma solo su un numero limitato di elettrodomestici, come vedremo al capitolo 6. Il progetto NED è stato poi ripreso e sviluppato da NEN, una "EnerTech" italiana, ossia una azienda che opera nello sviluppare tecnologie nel settore dell'energia, nata con lo scopo di semplificare servizio e gestione della fornitura domestica di luce e gas. Collaborando con Midori hanno dato vita ad un nuovo dispositivo, chiamato Robo, che è di fatto un aggiornamento di NED. Il principio di funzionamento non si discosta molto da quello dello Shelly, in quanto va a leggere e registrare i segnali di potenza, corrente, tensione, fase, ecc. attraverso l'utilizzo di una pinza amperometrica. Al contrario dello Shelly, però, esso è applicato direttamente al contatore e di conseguenza misura il consumo aggregato. Dopodiché, attraverso algoritmi di disaggregazione, esso riconosce i segnali presenti all'interno. L'elenco delle categorie di segnali riconosciuti fornito dall'azienda è: luci, stand-by, frigo e freezer (assieme), lavatrice, lavastoviglie, asciugatrice, forno, dispositivi di intrattenimento (come pc, console, tv). Il numero di segnali identificato è quindi abbastanza limitato e nemmeno molto diversificato. Non ci sono poi elettrodomestici con funzionamento breve. Tutti i dati, compresi quelli di consumo, costo dell'energia e gli altri resi disponibili da Robo sono facilmente accessibili tramite una applicazione. Queste tecnologie sono quindi ben lontane da un completo sviluppo e hanno margini di miglioramento grandissimi.

CAPITOLO 1

Raccolta ed elaborazione dati

1.1 Raccolta dati da database

Per massimizzare l'efficienza della rete neurale, a prescindere da quale venga scelta, oltre a trovare i corretti parametri per il training, è sempre necessario avere a disposizione una grande quantità di dati. Il modo più semplice e veloce per ottenere i dati necessari è quello di sfruttare i tanti database disponibili online che contengono i dati relativi alle curve di carico di numerose tipologie di elettrodomestici in commercio. I dati utilizzati sono gli stessi della tesi Zorzi e ciò renderà possibile un confronto sulle prestazioni dei metodi sviluppati.

1.1.1 Dataset ECO

Il dataset svizzero ECO^{[12][13]} contiene le curve di consumo in termini di potenza attiva rilevati con frequenza 1 Hz e viene fornita la potenza media all'interno del secondo. Il database completo comprende le misurazioni effettuate su 45 modelli di elettrodomestici. Sono state considerate solo quelle relative agli elettrodomestici che si trovano più frequentemente all'interno delle abitazioni.

Tipologia elettrodomestico	Numero modelli	Numero misurazioni
Asciugatrice	1	28
Bollitore	4	559
Forno a microonde	2	304
Freezer	4	508
Frigorifero	6	937
Lavastoviglie	1	53
Lavatrice	1	166
Macchina del caffè	3	664

Tab. 1.1: elenco misurazioni dataset ECO

1.1.2 Dataset GreenD

Il dataset GreenD^[14] è basato sui dati di consumo domestico di abitazioni in Italia e Austria. La frequenza di campionamento è di 1 Hz e sono disponibili le rilevazioni di potenza attiva relativi a 20 tipologie di elettrodomestici e 54 modelli, dei quali sono stati considerati solo quelli maggiormente presenti nelle abitazioni.

Tipologia elettrodomestico	Numero modelli	Numero misurazioni
Asciugacapelli	1	35
Asciugatrice	2	113
Aspirapolvere	1	17
Ferro da stiro	1	14
Forno	2	18
Lavastoviglie	5	464
Lavatrice	6	579

Tab. 1.2: elenco misurazioni dataset GreenD

1.1.3 Dataset REDD

Il dataset statunitense REDD^[15] è stato creato sulla base dei dati di consumi rilevati in zone residenziali delle città di Boston e San Francisco. Le misurazioni sono state fatte sia ad alta frequenza di campionamento (16 kHz) che a 1 Hz. Tra le due tipologie è stata considerata la seconda, in quanto una frequenza più alta è inutile ai fini del nostro lavoro. A questa frequenza sono stati rilevati i dati di consumo di 20 tipologie di elettrodomestici e 104 modelli, e di questi sono stati considerati solo quelli presenti nelle abitazioni con maggior frequenza.

Tipologia elettrodomestico	Numero modelli	Numero misurazioni
Asciugatrice	3	32
Forno	2	10
Forno a microonde	4	202
Frigorifero	5	65
Lavastoviglie	3	17

Tab. 1.3: elenco misurazioni dataset REDD

1.1.4 Dataset Tracebase

Tracebase^[16] è un dataset tedesco, sviluppato alla Technische Universität Darmstadt, contenente le misurazioni relative a 43 tipologie di elettrodomestici e 158 modelli. Come negli altri casi, sono stati considerati solo le curve relative agli elettrodomestici maggiormente presenti nelle abitazioni. Sono disponibili le frequenze di campionamento di 1 Hz e 0,12 Hz, tra le quali è stata considerata solamente la prima.

Tipologia elettrodomestico	Numero modelli	Numero misurazioni
Asciugatrice	2	15
Bollitore	8	353
Ferro da stiro	1	4
Forno a microonde	5	108
Freezer	1	9
Frigorifero	7	135
Lavastoviglie	3	88
Lavatrice	7	61
Macchina del caffè	5	102
Piano cottura	1	36
Tostapane	4	26

Tab. 1.4: elenco misurazioni dataset Tracebase

1.1.5 Dataset UK-Dale

Il dataset UK-Dale, creato da Kelly e Knottelbelt dell'Imperial College di Londra, rende disponibili le rilevazioni dei consumi di cinque abitazioni inglesi. Comprende sia le rilevazioni dei consumi complessivi delle abitazioni (rilevate a 16 kHz) che quelle dei consumi specifici di 103 modelli di elettrodomestici appartenenti a 40 tipologie differenti, campionate a 1 e 6 Hz^[17]. Tra le due frequenze di campionamento si considera la prima, mentre non sono stati considerati i dati aggregati.

Tipologia elettrodomestico	Numero modelli	Numero misurazioni
Bollitore	2	429
Ferro da stiro	1	4
Forno	1	31
Forno a microonde	1	140
Frigorifero	2	144
Lavastoviglie	2	34
Lavatrice	2	53
Macchina del caffè	1	5
Tostapane	2	248

Tab. 1.5: elenco misurazioni dataset UK-Dale

1.1.6 Dati raccolti con sistema UniPD-LAB.E.E.

L'ultimo database considerato deriva da misurazioni effettuate da un gruppo di ricerca universitario tramite un prototipo di dispositivo NILM (Non-Intrusive Load Monitoring) appositamente progettato, basato su una piattaforma Arduino. I dati non sono campionati a frequenza costante, anche se la frequenza media è comunque fissata a 1 Hz, e pertanto sono stati corretti tramite una apposita funzione dal dott. Zorzi all'interno del suo lavoro. Al contrario dei database già citati, i dati raccolti sono relativi alla potenza apparente assorbita dagli elettrodomestici.

Tipologia elettrodomestico	Numero modelli	Numero misurazioni
Asciugacapelli	1	6
Asciugatrice	1	3
Aspirapolvere	1	5
Bistecchiera	1	1
Bollitore	1	4
Cappuccinatore	1	1
Ferro da stiro	1	21
Forno	1	1
Forno a microonde	1	9
Frigorifero	1	2
Lavasciuga	1	3
Lavastoviglie	1	25
Lavatrice	1	10
Macchina del caffè	1	6
Piastra a induzione	1	1
Tagliaerba	1	2
Tostapane	1	1

Tab. 1.6: elenco misurazioni UniPD-LAB.E.E.

1.2 Raccolta dati di consumo cumulati

I dati relativi ai consumi cumulati sono stati ottenuti monitorando i consumi di una villetta e due condomini residenziali nell'ambito di uno studio sul monitoraggio di impianti asserviti a edifici reali "all-electric" nel settore residenziale effettuato da UniPD in collaborazione con ENEA e i consumi di un appartamento abitato da due adulti. I dati non sono a frequenza di campionamento costante, ma vengono rilevati in modalità spot, ossia attraverso rilievi di breve e medio periodo e pertanto dovrebbero essere elaborati e corretti per riempire i vuoti nella linea temporale ed ottenere una curva di consumo utilizzabile per la nostra rete neurale. Si può inoltre ipotizzare di aggiungere un segnale di rumore artificiale che simuli quello reale, che si presenterebbe come un segnale ottenuto a frequenza di campionamento fissata (ad esempio a 1 Hz, coerentemente con quella dei database dei singoli elettrodomestici). I dati provenienti dallo studio in collaborazione con ENEA sono salvati in file Excel contenenti gli "eventi" registrati (quindi accensioni e spegnimenti degli elettrodomestici), orario in cui avvengono e potenza attiva registrata dal contatore al momento dell'evento. Successivamente, i file sono convertiti in file MATLAB, poi corretti tramite apposite funzioni. Il metodo utilizzato per la raccolta dei dati sarà importante successivamente, quando si andrà ad elaborare una metodologia per l'elaborazione dei dati stessi.

Si è andati in seguito a sfruttare un apposito script per generare casualmente segnali aggregati della durata di 24 ore, composti dai segnali presenti all'interno dei dataset. Questi aggregati sono invece a frequenza di campionamento costante, coerente con quella dei database (1Hz). All'interno di ciascuno di essi, comparirà un solo segnale per tipologia di elettrodomestico tra quelle selezionate.

Il posizionamento temporale del singolo segnale è completamente casuale, consentendo quindi la presenza di elettrodomestici attivi contemporaneamente, con sovrapposizioni dei segnali sia totali che parziali, o segnali “troncati” alla scadenza dell’intervallo di 24 ore.

1.3 Elaborazione dati attraverso il software MATLAB

1.3.1 Elaborazione dati dei singoli elettrodomestici

Una volta raccolti i dati ottenuti dai database pubblici e quelli relativi alle misurazioni del prototipo NILM, si è proceduto con la correzione di questi ultimi, in modo tale da ottenere delle curve con frequenza di campionamento di 1 Hz. Da queste curve si vuole ottenere l’immagine dello spettro di potenza (spettrogramma). Per farlo si utilizza l’apposita funzione *pspectrum* di MATLAB, la quale va ad operare sugli spettri di segnali in tempo-frequenza.

pspectrum (signal, t, 'spectrogram')

pspectrum = funzione di Matlab che analizza i segnali in tempo-frequenza, elabora ed opera sullo spettro di potenza di un segnale;

Signal = vettore (o matrice) contenete il segnale in tempo-frequenza da analizzare;

t = tempo trascorso tra un campionamento e il successivo (costante);

'spectrogram' = opzione da attivare per ottenere lo spettrogramma anziché lo spettro di potenza del segnale

Una volta elaborato il segnale, il software restituisce un’immagine con in ascissa il tempo, in scala lineare, e in ordinata la frequenza, sempre in scala lineare. L’intensità viene rappresentata dal colore dell’immagine, come è possibile vedere da un’apposita barra presente nell’immagine.

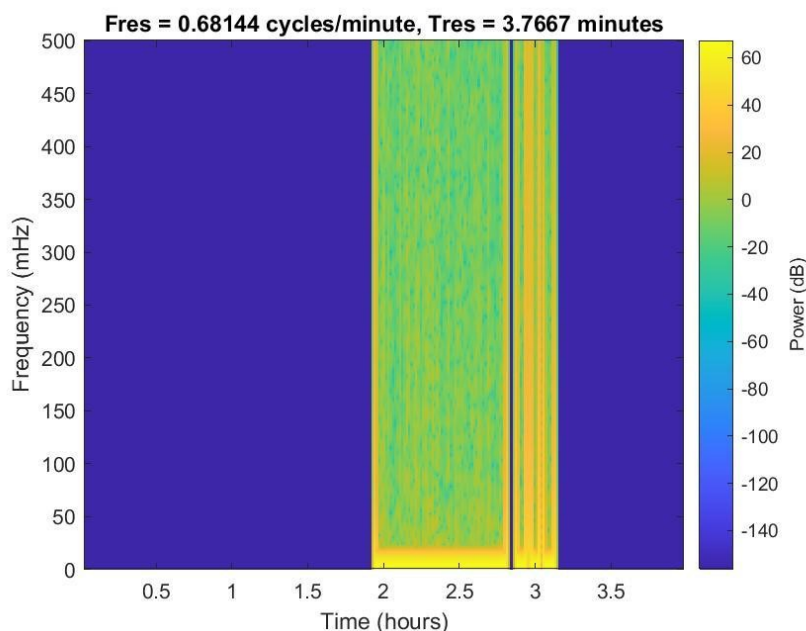


Figura 1.1: spettrogramma del segnale di un forno

Per ciascun segnale presente nei vari database sono stati creati uno o più spettrogrammi, posizionando il segnale in momenti differenti all’interno dell’intervallo temporale analizzato, per evitare il rischio che il carico venga riconosciuto solamente in specifiche posizioni dell’immagine (problema di overfitting).

Dopo una attenta analisi delle durate dei cicli di lavoro dei vari elettrodomestici analizzati, al fine di tenerne in considerazione il maggior numero possibile e di avere quindi un numero di dati ampio per poter procedere con un training efficace della rete (a prescindere dal tipo di rete che è stata successivamente scelta), si è deciso di elaborare i dati dei segnali all’interno di intervalli temporali di quattro ore. In questo modo si includono gli elettrodomestici con cicli più lunghi (di alcune ore, come

ad esempio le lavatrici) senza escludere quelli con cicli molto brevi (dell'ordine del minuto), che rischiano di diventare difficilmente distinguibili in intervalli eccessivamente ampi, in quanto l'immagine del segnale risulterebbe troppo "stretta".

Tutte le immagini create (oltre diecimila) vengono raccolte in un database MATLAB che servirà per il training della rete neurale. Il formato di salvataggio scelto è quello ".jpg" e le dimensioni delle immagini sono di 875x656 pixel.

Successivamente, date le difficoltà nel riconoscimento dei segnali brevi, si è proceduto allo stesso modo per generare spettrogrammi in intervalli di un'ora. Ovviamente sull'intervallo temporale ridotto non sono stati considerati i segnali degli elettrodomestici con cicli più lunghi, quindi lavatrice, lavastoviglie, frigorifero, asciugatrice e freezer e tutte le curve di durata superiore all'ora degli altri elettrodomestici.

1.3.2 Elaborazione dei dati di consumo aggregati

Il segnale di consumo aggregato è stato acquisito tramite un sistema di rilevazione di eventi (misurazioni spot). Il sistema di acquisizione va quindi a rilevare e registrare un evento nel momento in cui viene indotta una variazione nel segnale aggregato. In questo modo di fatto si registrano i momenti di accensione e spegnimento di un elettrodomestico, ma non la parte di funzionamento stazionaria (o quasi-stazionaria). La frequenza di campionamento, pertanto, non è costante. Essendo la trasformata di Fourier una funzione integrale, la frequenza di campionamento del segnale, che rappresenta nel nostro caso la funzione di cui calcolare lo spettro, impatta in maniera significativa sullo spettrogramma ad esso associato. Guardando l'immagine esemplificativa, risulta immediatamente evidente come il risultato sia completamente diverso, sia per la minor definizione dell'immagine sia per la scala delle frequenze, che è di un ordine di grandezza inferiore. È difficile pensare che la rete possa identificare immagini così diverse tra loro e perciò si è deciso di procedere parallelamente con altri metodi.

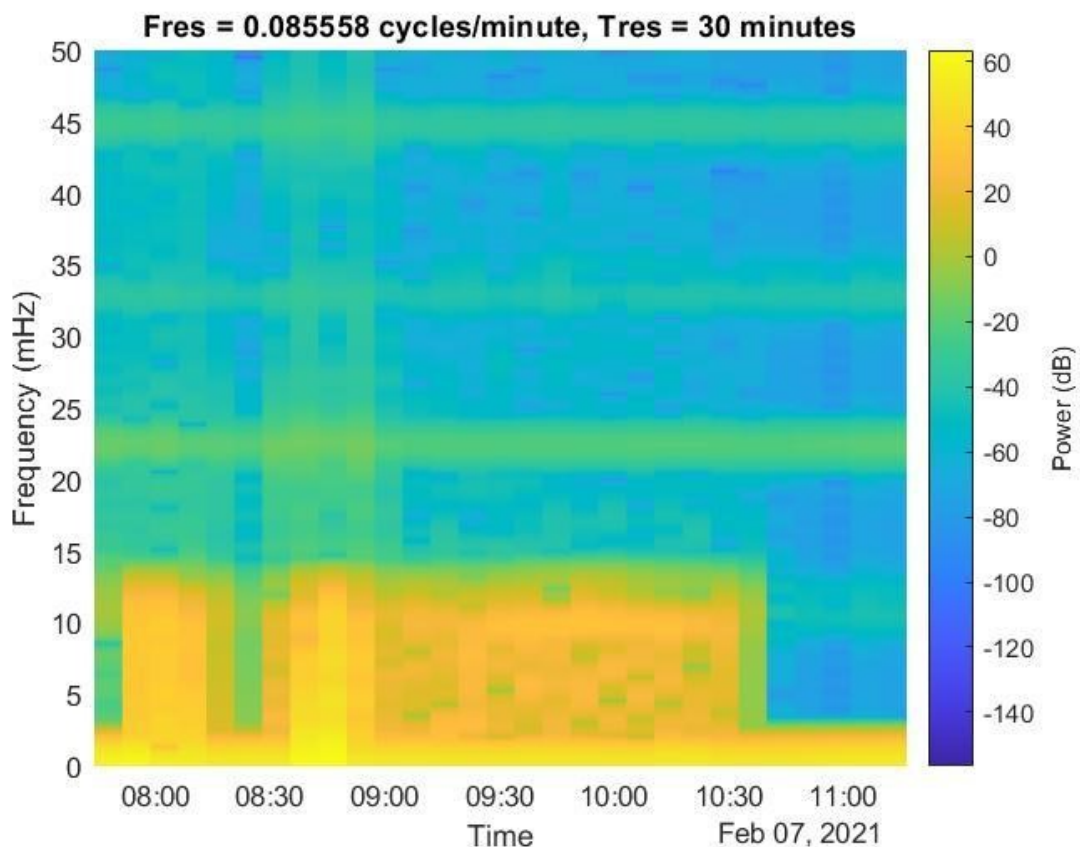


Figura 1.2: esempio di spettrogramma del segnale aggregato ottenuto tramite il sistema di registrazione eventi

L'alternativa è quella di utilizzare segnali aggregati simulati, ottenuti attraverso uno script usato per l'addestramento della rete Zorzi. Questi segnali, della durata di 24 ore, contengono al loro interno segnali pescati casualmente dai database per il training. L'operatore sceglie quali elettrodomestici includere nell'aggregato e quante attivazioni ha ciascuno di essi. Lo script va ad inserire i valori dei vettori dei singoli segnali all'interno del vettore aggregato di dimensione 86400×1 , in cui ogni elemento corrisponde ad un istante delle 24 ore totali. Oltre a ciò, va a porre pari a zero tutti gli istanti in cui la potenza registrata è inferiore a 20 W, simulando in questo modo l'eliminazione dell'eventuale rumore di fondo. I segnali presenti nell'aggregato finale saranno quindi in realtà diversi da quelli utilizzati nell'addestramento, aumentando la difficoltà di riconoscimento da parte della rete e la verosimiglianza della simulazione. La completa casualità dell'istante di attivazione del singolo elettrodomestico fa sì che si possano avere segnali sovrapposti. Non si è andato a inserire un tetto ai picchi massimi di potenza, non considerando quindi le limitazioni imposte da contratto alle utenze domestiche.

Il caso più gravoso per la rete neurale è quello che presenta segnali di durata molto diversa tra loro, un numero elevato di elettrodomestici e la presenza di elettrodomestici con caratteristiche simili in termini di tempo di funzionamento e di curva di consumo. La presenza di elettrodomestici con ciclo di funzionamento di durata molto differente rende difficile trovare la giusta durata degli intervalli dei singoli spettrogrammi su cui effettuare la detection. Un segnale molto breve, infatti, avrà un'immagine molto piccola nello spettrogramma e pertanto sarà molto più difficile da riconoscere. Un numero elevato di elettrodomestici fa sì che sia più facile avere segnali che si sovrappongono, sia totalmente che parzialmente, rendendo difficile se non impossibile la corretta identificazione di tutti i segnali interessati. Segnali simili per durata e curva, daranno invece luogo a spettrogrammi a loro volta simili tra loro, complicando di conseguenza il corretto riconoscimento da parte della rete neurale. Per questi motivi, negli aggregati simulati verranno inseriti tutti i segnali ritenuti interessanti perché relativi a elettrodomestici comuni o perché utili a capire il comportamento della rete su un numero ridotto di dati o su segnali difficilmente riconoscibili.

Tra i segnali lunghi sono stati inseriti lavatrice, lavastoviglie e asciugatrice, mentre tra i segnali medio-corti sono stati scelti asciugacapelli, aspirapolvere, bollitore, ferro da stiro, forno, macchina del caffè, microonde, piano cottura, piastra a induzione, robot da cucina, tagliaerba, tostapane. Tra questi vi sono elettrodomestici con numero di spettrogrammi nel database molto differenti, elettrodomestici dal segnale confondibile (esempio asciugacapelli, bollitore e piano cottura). Il robot da cucina, per numero di elementi nel dataset e tipologia di segnale sarà un elemento particolarmente critico. Tutto questo dovrà essere successivamente considerato al momento delle valutazioni sull'efficacia della rete e sulle sue potenzialità. Tra i segnali esclusi vi sono frigorifero e freezer. Nonostante siano elettrodomestici molto comuni, rimanendo funzionanti durante l'intero intervallo temporale, andrebbero a sovrapporsi con la maggior parte, se non la totalità, degli altri segnali presenti all'interno dell'aggregato, pregiudicando il riconoscimento degli stessi e rendendo molto complicato valutare nella maniera più corretta il comportamento della rete neurale. In ogni caso, sono stati creati degli aggregati di prova contenenti anche il frigorifero per dare una conferma, seppur parziale, ai risultati ottenuti sul set di addestramento.

CAPITOLO 2

SCELTA DELLA RETE NEURALE E POST-ELABORAZIONE DATI

La scelta della rete neurale è stata effettuata cercando di individuare quella che riuscisse a soddisfare nel miglior modo possibile tutti gli obiettivi, o quanto meno la maggior parte di essi, prioritizzando ovviamente l'identificazione corretta dei carichi rispetto agli obiettivi secondari. A seconda della scelta, poteva rendersi necessaria una post-elaborazione delle immagini, come effettivamente avvenuto.

2.1 Le reti neurali per il riconoscimento delle immagini

Il riconoscimento delle immagini è un problema da tempo affrontato con le reti neurali. Un facile esempio del loro impiego è quello dello strumento di ricerca per immagini fornito da Google. Essendo un problema già affrontato, sono disponibili vari tipi di reti neurali che adempiono al compito in modo diverso e con scopi differenti. Sono state prese in considerazione due tipi di rete: la rete googlenet e le reti YOLO. Una terza opzione, presa in considerazione e scartata per la complessità di implementazione, è la rete LSTM (Long Short-Term Memory network), che consente di individuare un oggetto in base alla sequenza in cui sono ordinati i dati in esame (ad esempio riconoscere l'autore di una frase basandosi sull'ordine dei caratteri e quindi delle parole che la compongono).

2.1.1 La rete GoogLeNet

La rete GoogLeNet è una rete convoluzionale, ossia una rete neurale di tipo "feed-forward" che riproduce il comportamento della corteccia visiva animale al fine di riconoscere un'immagine. Nella corteccia visiva animale, ciascun neurone riceve ed elabora gli stimoli provenienti da una parte del campo visivo. I segnali dei neuroni vengono poi combinati ed elaborati dal cervello che fornisce una risposta allo stimolo sensoriale. Analogamente, in una rete convoluzionale, ciascun neurone riceve un segnale da una parte dei neuroni dello strato precedente e li elabora mediante operazioni logiche. Il segnale in uscita viene inviato ai neuroni dello strato successivo, che operano allo stesso modo. Tra uno strato e l'altro può esserci un cosiddetto *pooling layer*, che va a ridurre la dimensione dei dati combinando gli output di gruppi di neuroni (detti *cluster*). Infine, i segnali arrivano a uno o più *fully connected layer*, in cui ciascun neurone è connesso a tutti i neuroni degli strati adiacenti. Questo layer è quello deputato alla classificazione dell'immagine di input.

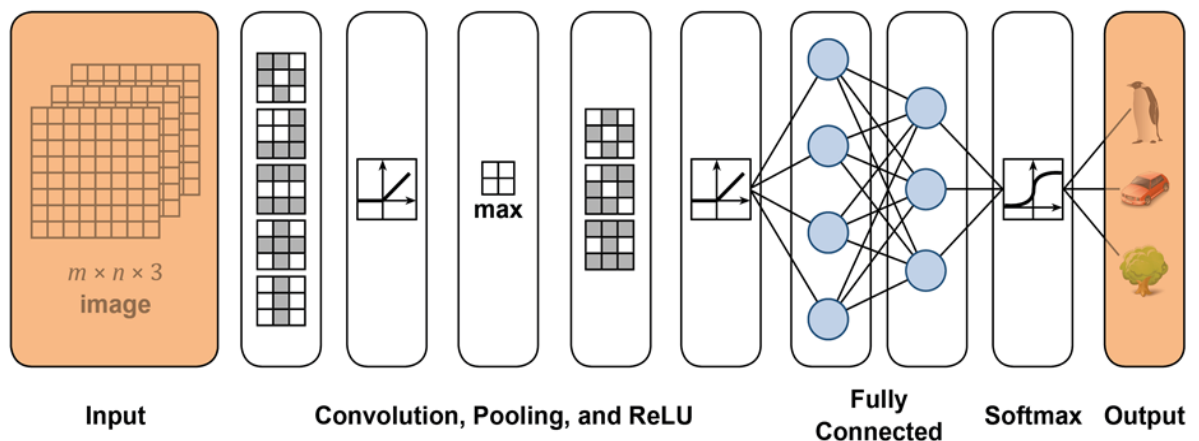


Figura 2.1: Rappresentazione della rete neurale GoogLeNet

Nella figura 2.1 si può vedere la rappresentazione semplificata degli strati che compongono la rete GoogLeNet^[18]. Lo strato di input riceve delle immagini di dimensioni $m \times n$ in formato rgb che vengono elaborate dai layer convoluzionali, di maxPooling (ossia che inviano allo strato successivo il segnale massimo tra quelli ricevuti dal cluster) e le cosiddette ReLU (*Rectified Linear Unit*) che settano i segnali di segno negativo in segnali nulli. Oltre questa parte, si trovano gli ultimi tre strati, composti da un *fully connected layer*, un Softmax layer e l'output layer. Il primo di questi riceve i segnali da

tutti i neuroni del layer precedente ed elabora un “punteggio” compreso tra 0 e 1 che identifica il grado di confidenza con la quale la rete identifica l’immagine di input come appartenente a ciascuna delle categorie conosciute dalla rete. Il Softmax layer identifica il segnale massimo tra quelli in ingresso e lo comunica all’output layer, il quale fornisce come risultato la categoria che ha ottenuto il massimo punteggio di confidence.

Questa rete va a identificare molto bene il soggetto dell’immagine, ma ha due grandi difetti per l’applicazione in esame. Il primo di essi è il fatto che fornisce un solo output e pertanto se all’interno di uno spettrogramma sono presenti più segnali la rete è in grado di identificarne solo uno. Il secondo è che non va ad identificare la zona in cui riconosce il segnale, che rappresenta l’intervallo di tempo da esso occupato. Di conseguenza, è difficile andare ad estrarre dal segnale altre informazioni quali ad esempio gli istanti di attivazione e spegnimento o il consumo complessivo. Si è quindi deciso di scartare questa opzione, in quanto inadatta al riconoscimento dei singoli carichi all’interno di un segnale di consumo aggregato.

2.1.2 Le reti di tipo YOLO

Le reti YOLO^[19] (You Only Look Once) sono anch’esse reti convoluzionali, in grado però di identificare ed etichettare le aree di una immagine all’interno delle quali riconosce un certo oggetto. La rete prende in esame l’immagine in input e la suddivide in regioni di interesse (o bounding box) che possono potenzialmente contenere al loro interno l’immagine di un oggetto che la rete è stata allenata a riconoscere. Questo processo avviene attraverso un solo passaggio di informazioni all’interno della rete, al contrario di quanto succedeva con le Region-based Convolutional Neural Network^[9], che svolgono lo stesso compito ma in più passaggi, dovendo estrarre le regioni di interesse per elaborarle, identificarle e poi andare a inserire le bounding box e affinarle nell’immagine originaria. Le reti YOLO, per fare questo, predicono simultaneamente le probabilità di appartenenza a ciascuna classe per multiple bounding box. Questa differenza, apparentemente banale, ha risvolti pesanti a livello di precisione ed efficienza della rete. Oltre ad essere molto più veloce, infatti, è anche molto più precisa, in quanto non estrae la regione di interesse dall’immagine intera e ciò permette di mantenerla nel suo contesto, evitando di conseguenza di confondere parti dello sfondo per oggetti appartenenti ad una delle classi.

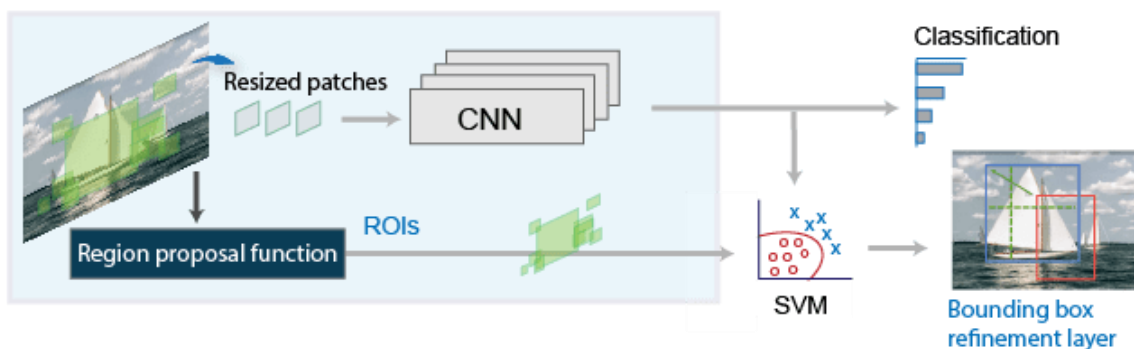


Figura 2.2: schema del funzionamento delle R-CNN ^[20]

In che modo vengono individuate le bounding box dalla rete YOLO? L’immagine in input viene suddivisa in una griglia formata da SxS riquadri. A quel punto predice simultaneamente le bounding box delle varie classi (dette anchor box) e le rispettive probabilità di trovarsi in quel riquadro, ripetendo l’operazione in ciascuno di essi. Una anchor box è di fatto una bounding box che è rappresentativa di una intera classe e viene definita per “riassumere” le dimensioni e le proporzioni degli oggetti appartenenti alla classe stessa. Se la rete individua il centro di un oggetto in uno dei riquadri questo diventerà il riferimento per individuare l’oggetto e quindi la rispettiva bounding box. Va poi a calcolare la probabilità della presenza dell’oggetto o la confidenza della identificazione dello stesso. La confidenza viene calcolata come prodotto tra la probabilità della presenza dell’oggetto e l’indice di Jaccard (o Intersection Over Union). L’IOU rappresenta la somiglianza tra la bounding box identificata e quella “teorica” derivante dal database di training, calcolata come il rapporto tra l’intersezione delle due aree e l’unione delle due.

$$Prob(classification) = Prob(object) * Prob(object) * IOU \quad (2.1)$$

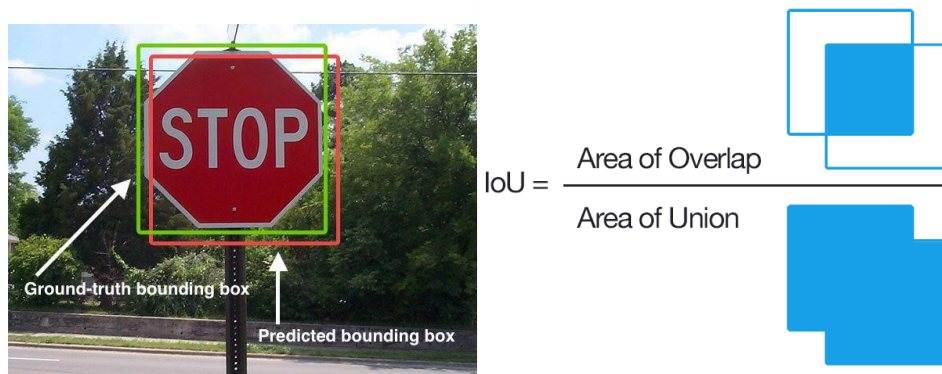


Figura 2.3 (a e b): rappresentazione dell'indice di Jaccard (IOU)^[22]

Calcolando in questo modo la confidenza della previsione della rete neurale, essa risulta sempre nulla se non esistono oggetti nella cella e, se diversa da zero, essa viene scalata in base alla corrispondenza tra la bounding box prevista e quella presente nel Ground Truth. Di conseguenza, perché venga individuato correttamente un oggetto, oltre a trovare delle feature che ne indichino la presenza (ad esempio l'intensità del colore), deve esserci anche una corrispondenza nella forma della bounding box e in particolare nelle proporzioni delle box. È importante notare che, lavorando sempre all'interno dell'immagine complessiva, non si perdono informazioni sul contesto, limitando quindi gli errori legati ad oggetti presenti nello sfondo, che nel caso degli spettrogrammi sono essenzialmente legati a segnali di disturbo o rumore.

$$IOU = \frac{|A \cap B|}{|A \cup B|} \quad (2.2)$$

I limiti della rete YOLO sono legati al fatto che ad ogni cella della griglia possono essere associate solo 2 bounding box e queste possono appartenere ad una sola classe. Ciò significa che se il centro di due oggetti diversi si trova nella stessa cella ne può essere identificato uno solo, quello la cui previsione ha l'indice di confidenza più elevato. Stessa cosa nel caso ci siano più oggetti di dimensioni ridotte raggruppati. Nel caso di studio questo fa sì che possano esserci difficoltà ad individuare segnali contemporanei o attivazioni di elettrodomestici molto ravvicinate (soprattutto se si tratta di segnali di breve durata temporale). Un ulteriore problema della rete è quello di non riuscire a riconoscere oggetti se hanno proporzioni diverse da quelle presenti nel database di training, ma per la tipologia di immagini in esame è un problema minore, in quanto interessa solo la larghezza delle box, quindi la durata temporale, che però difficilmente si discosta dai casi presenti nel database se questo è stato ben strutturato, e non l'altezza, che è legata alle frequenze e rimane invariata. La tipologia di errore più frequente è la scorretta localizzazione di oggetti di piccole dimensioni. Di conseguenza, ci si aspetta una difficoltà nell'individuare segnali di breve durata (ad esempio quelli di asciugacapelli o microonde) e/o segnali sovrapposti se questi sfortunatamente hanno il loro punto centrale ravvicinato.

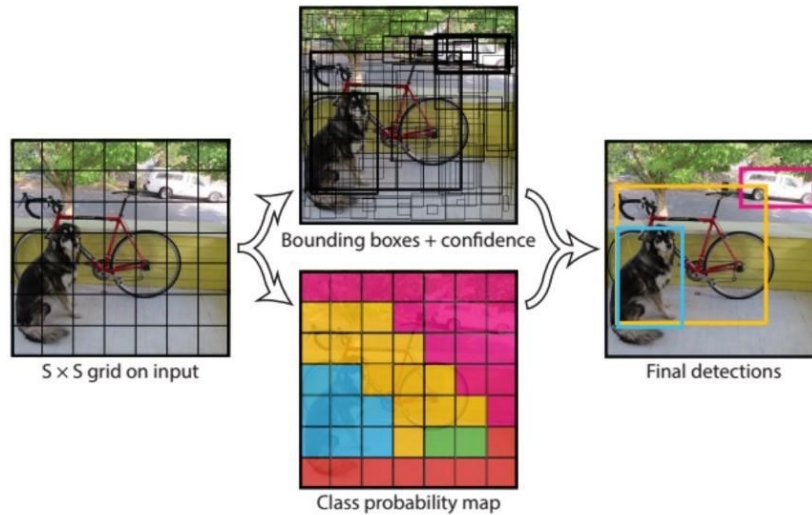


Figura 2.4: schema esplicativo del funzionamento della rete YOLO^[19]

Per la nostra applicazione, tra le varie tipologie di rete YOLO, si è scelto di utilizzare la versione YOLOv2, nonostante non sia l'ultima versione, in quanto è sufficientemente veloce e precisa e non introduce ulteriori complessità non necessarie, come ad esempio la classificazione multi-step della YOLOv3.

La rete è composta da due sottoreti: una convoluzionale (che può essere scelta a seconda delle necessità dall'autore) e una di detection. Come rete convoluzionale si è scelto di usare la rete RESNET-18, che è composta di 18 layer e riceve in input immagini di dimensione 224x224 pixel. La dimensione ridotta rende necessario il ridimensionamento del database (sia delle immagini che delle bounding box ad esse associate), ma velocizza il processo di training e detection senza andare a discapito della precisione.

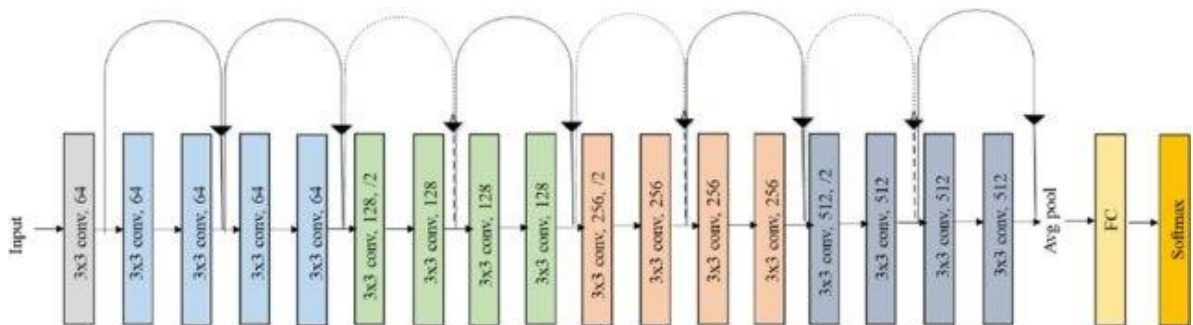


Figura 2.5: struttura della rete RESNET18^[21]

Concatenati alla rete RESNET-18, troviamo i layer cui è deputato il compito di riconoscere le immagini. In totale, la rete risulta essere composta da 78 layer, tra rete convoluzionale e rete di detection (compresi i layer di output). La struttura corrisponde a quella standard già implementata in MATLAB, in cui è stato solamente scelto di utilizzare la rete RESNET-18 come sottorete convoluzionale.

2.2 Post-processing degli spettrogrammi e creazione dei database

Come detto nel capitolo precedente, gli spettrogrammi creati sulla base di segnali della durata di quattro ore vengono creati con dimensioni di 875x656 pixel. Le operazioni necessarie per eseguire il training della rete sono quindi la creazione delle bounding box (e delle rispettive etichette), il ridimensionamento delle immagini con le box associate e creazione dei database che li contengono. Le operazioni sugli spettrogrammi di durata di un'ora sono le stesse.

Come prima cosa, si è proceduto alla creazione delle bounding box e delle rispettive etichette che servono ad andare a identificare direttamente a schermo a quale classe appartiene l'oggetto identificato. Le box non sono altro che le aree di interesse all'interno delle quali si trova l'oggetto da identificare, che nel nostro caso non è altro che una fascia colorata che corrisponde al segnale dell'elettrodomestico. In MATLAB possono essere create di forme diverse, ma nel nostro caso è sufficiente creare un riquadro attorno alla zona interessata dal segnale. Si va così a riconoscere esclusivamente il segnale dell'elettrodomestico, senza che la detection sia influenzata da ciò che avviene prima e dopo, ossia la presenza di altri segnali o di rumore. Le strade percorribili per la creazione delle box sono essenzialmente due. La prima, più semplice ed automatica, è quella di crearle attraverso una apposita app presente nel software chiamata Image Labeler App, che consente di applicarle direttamente sulle immagini scegliendo forma, colore della linea e il nome della classe di appartenenza da apporre sull'etichetta. Le immagini possono essere importate sia dalle cartelle del dispositivo in uso che da database caricato nel workspace di MATLAB. Una volta applicate su ciascuna immagine, le box possono essere esportate come Ground Truth Table, ossia una tabella che contiene i dati di ognuna delle box e delle etichette associate. Una volta creata, la tabella può essere facilmente convertita in un database di bounding box con un apposito comando:

$$Bbds = boxLabelDatastore (table)$$

I dati relativi alle box contenuti nella tabella sono il nome delle classi e i dati relativi alla localizzazione e alle dimensioni della zona di interesse. I primi sono contenuti nelle intestazioni delle colonne della tabella, mentre i secondi sono contenuti nelle celle che si trovano sulla riga corrispondente al nome del file dell'immagine su cui sono applicate e sulla colonna corrispondente al nome dell'etichetta. La cella contiene un vettore 4x1 del tipo $[x,y,w,h]$, in cui x,y sono le coordinate dell'angolo in alto a sinistra, w la larghezza e h l'altezza del riquadro.

	Asciugacapelli	Asciugatrice	Aspirapolvere
92	[435,48,16,535]	[]	[]
93	[690,48,17,535]	[]	[]
94	[179,48,17,535]	[]	[]
95	[428,48,21,535]	[]	[]
96	[683,48,19,535]	[]	[]
97	[172,48,19,535]	[]	[]
98	[]	[196,48,457,535]	[]
99	[]	[103,48,604,535]	[]
100	[]	[157,48,362,535]	[]
101	[]	[103,48,604,535]	[]
102	[]	[285,48,361,535]	[]
103	[]	[103,48,604,535]	[]
104	[]	[103,48,604,535]	[]
105	[]	[203,48,368,535]	[]
106	[]	[103,48,604,535]	[]

Figura 2.6: parte esemplificativa della tabella usata per creare il boxLabelDatastore

Nel nostro caso, data la semplicità delle box e la facile individuazione delle stesse, si può evitare il suo utilizzo andando a creare manualmente una tabella analoga alla Ground Truth Table, ricordando di inserire i dati in forma di *cell array of character vector*. Questo procedimento manuale è ovviamente sconsigliato, soprattutto quando si opera su database molto grandi.

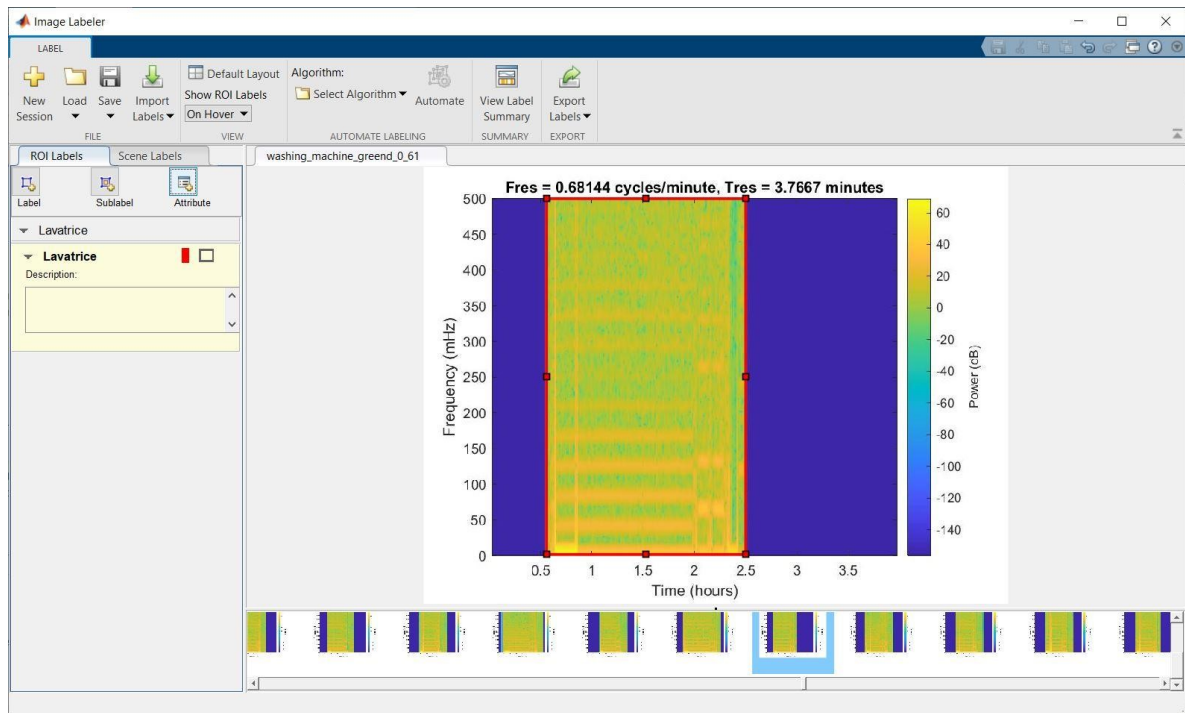


Figura 2.7: esempio di utilizzo della Image Labeler App con bounding box dello spettrogramma

Il database delle immagini può essere anch'esso creato con apposita funzione e a tal fine è importante salvare tutte le immagini all'interno di una cartella e suddividerle in sottocartelle nominate con il nome della classe. In questo modo si riesce a creare il database suddividendo i file nelle varie classi di appartenenza con il seguente comando:

```
Imds = imageDataStore(location, 'IncludeSubfolders', true, 'LabelSource', 'foldernames')
```

Una volta creati i due database li si va a combinare attraverso la funzione *combine*, andando quindi ad associare a ciascuna immagine la propria bounding box. Il nuovo database combinato servirà per il training della rete neurale. Prima di iniziare il training, il database completo va ridimensionato dalle dimensioni originali di 875x656 a quelle di input della rete di 224x224, per far sì che le immagini abbiano le dimensioni corrette per essere analizzate dalla rete e che le box vengano scalate allo stesso modo, così da non perdere la corrispondenza tra le due.

Nel dataset utilizzato, il numero di classi è pari al numero di tipologie di elettrodomestico presenti nei database delle curve di consumo, ossia 20.

2.3 Creazione delle Anchor Boxes

Una volta creato il database combinato e scalato, è necessario andare a calcolare le anchor box attraverso le quali la rete andrà a identificare gli oggetti come appartenenti ad una certa classe. Per farlo, si utilizza una specifica funzione di MATLAB, chiamata *estimateAnchorBoxes*, che restituisce le box ed eventualmente il valore medio dell'indice IOU.

```
[AnchorBoxes meanIOU] = estimateAnchorBoxes(combinedDataset, numAnchor)
```

In ingresso, oltre al dataset combinato e scalato, deve essere specificato il numero di anchor box che si vuole calcolare. Nel caso di un set con più di una classe, il numero minimo di box da creare è ovviamente il numero delle classi. Per valutare se il numero di anchor box è sufficiente, va valutato l'IOU medio, che è un numero compreso tra 0 e 1 che identifica quanto bene le anchor box si

sovrappongono alle bounding box del database. Se il valore supera 0,5 è accettabile, altrimenti va aumentato il numero di anchor box. Se il valore è troppo elevato (vicino all'unità) vuol dire che è insorto un problema di overfitting, quindi rischio di avere scarsi risultati nell'identificazione di oggetti non compresi nel database e va diminuito il numero di anchor box. Nel caso di studio è stato sufficiente creare una box per ciascuna delle 20 classi, ottenendo un valore di IOU medio di circa 0,84. Questo valore è molto alto, anche perché alcune classi, a causa dei pochi dati a disposizione, hanno un numero limitato di spettrogrammi. Ci si aspetta, quindi, che possa presentarsi un problema di overfitting, magari solo per alcune classi, anche se il modo in cui si sono creati gli spettrogrammi dovrebbe aver limitato il più possibile il problema.

CAPITOLO 3

TRAINING DELLA RETE NEURALE

All'interno di una rete neurale, le informazioni in ingresso vengono elaborate attraverso delle operazioni logiche (determinate dall'architettura della rete neurale) e parametri moltiplicativi, detti pesi, che vengono applicati alle informazioni (feature) mentre queste passano da un neurone al successivo. Il training, o addestramento, della rete neurale consiste nella regolazione delle feature e dei pesi al fine di garantire il corretto funzionamento della rete. Per fare ciò, si dà in pasto alla rete una serie di dati noti in ingresso e si fanno variare i pesi (solitamente in modo automatico attraverso un processo iterativo interamente a carico del software) finché la rete non riesce ad elaborare gli input in modo tale da ottenere un livello minimo di output corretti. Nel caso di una rete per il riconoscimento delle immagini, il database di training sarà composto da un adeguato numero di immagini già classificate. Oltre ad un set di addestramento, possono essere previsti anche un set di validazione e uno di testing. Il primo serve a valutare ed ottimizzare i parametri di training ed eventualmente l'architettura della rete, mentre il test set serve a valutare le performance della rete. Solitamente si procede creando un unico dataset e suddividendolo in tre parti in maniera casuale, secondo proporzioni definite dall'operatore (es. 80% training set, 10% validation set, 10% test set). La porzione maggiore va ovviamente destinata all'addestramento, che è la parte più importante e quella per cui serve il maggior numero di input. Una valutazione sulle prestazioni della rete può essere fatta anche testandola solamente sul training set, ma è meno precisa, in quanto mostrerebbe solamente la capacità della rete di elaborare correttamente i dati di addestramento e non il suo comportamento su dati sconosciuti. Ciò, oltre a sovrastimare la precisione delle previsioni della rete, nasconde anche altri eventuali problemi, in primis quello dell'overfitting, ossia il fatto che la rete riconosca perfettamente i casi del training set ma non riesca ad elaborare correttamente casi nuovi, anche solo leggermente differenti.

3.1 Opzioni di addestramento

L'addestramento viene fatto attraverso le apposite funzioni dal software, ma prima di procedere con lo stesso è necessario settare correttamente le opzioni con le quali verrà effettuato. La miglior combinazione delle opzioni di training non può essere determinata a priori. Per questo motivo, si procederà ripetendo l'addestramento con diverse impostazioni fino a trovare la migliore. Le impostazioni di primo tentativo vengono scelte in base all'esperienza e alla sensibilità dell'operatore, per poi andare a modificarle in base ai risultati ottenuti. L'obiettivo finale è quello di trovare le opzioni che ottimizzino l'efficienza della rete senza però allungare eccessivamente il tempo di training. Ora verranno analizzate quali sono le opzioni modificabili e in che modo si è operato su di esse.

3.1.1 Algoritmo di ottimizzazione dei parametri

Per ottimizzare il comportamento della rete viene introdotta una funzione di costo, il cui valore risulterà tanto maggiore quanto maggiore è l'errore commesso dalla rete nella prediction. Gli algoritmi di ottimizzazione vanno a modificare i pesi della rete neurale per minimizzare questa funzione di costo, che è rappresentata da una somma di addendi ed è funzione dei parametri. Gli algoritmi disponibili sono varianti del metodo della discesa del gradiente. Partendo da un certo set di parametri, l'algoritmo identifica la direzione di discesa massima del gradiente della funzione di costo (che porta quindi al minimo della funzione stessa) e va a modificare i pesi lungo quella direzione. La grandezza dello step di modifica dei parametri viene determinata dal *learning factor*. La scelta dell'algoritmo è ricaduta sull' "adam"^[23] (adaptive moment estimation), che considera i momenti primo e secondo del gradiente, considerando quindi le interazioni precedenti, e va ad adattare il learning factor per ciascun parametro.

3.1.2 Initial Learning Rate e Learn Rate Drop Factor

Il learning rate factor è un parametro che va a definire la grandezza degli step compiuti ad ogni interazione dall'algoritmo di minimizzazione della funzione di costo. Ovviamente, tanto maggiore è questo fattore di apprendimento tanto più velocemente ci si sposterà verso il minimo, ma se si aumenta eccessivamente si rischia di incorrere nel problema dell'overshooting. In questo caso il fattore fa sì che lo spostamento dal punto x_i a x_{i+1} (ricordando che, essendo funzione dei pesi, le coordinate dei punti sono i pesi utilizzati alla i -esima iterazione) lungo la direzione di massima

discesa del gradiente sia talmente ampio da superare il punto di minimo, che viene dunque “mancato” ad ogni iterazione, rendendo l’algoritmo non convergente. Il corretto learning rate sarà, dunque, quello che garantisce un compromesso tra la velocità di convergenza e l’evitare il problema dell’overshooting.

Un’altra impostazione molto importante per migliorare questo compromesso è il fattore di decadimento del fattore di apprendimento (learn rate drop factor). Questo fattore va a modificare il learning rate dopo un certo numero di interazioni o epoche. In questo modo, si può utilizzare un fattore di apprendimento grande nelle prime interazioni per avvicinarsi il più velocemente possibile alla zona adiacente al minimo, per poi diminuirlo e muoversi più lentamente verso il punto di minimo, o comunque ad una soluzione sufficientemente vicina, dato che spesso è molto difficile raggiungere il minimo in un numero accettabile di iterazioni e soprattutto, a livello di prestazione della rete, non ci sono differenze sensibili rispetto alla soluzione ottima. Così facendo, accelero il processo di apprendimento ed evito il problema di overshooting.

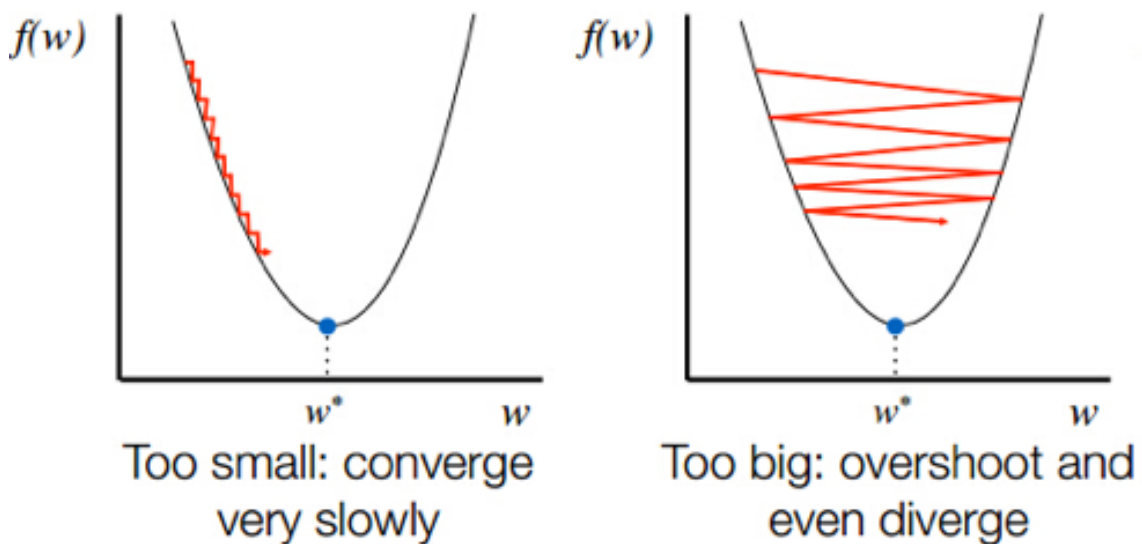


Figura 3.1: figura esplicativa del comportamento dell’ottimizzazione in funzione della scelta del fattore di apprendimento^[24]

3.1.3 Mini-batch, numero di epoche e shuffling

Se non specificato diversamente, l’aggiornamento dei pesi attraverso il metodo di discesa del gradiente viene effettuato applicando l’algoritmo sull’intero dataset. Ciò comporta che, se il database è molto esteso, ogni iterazione sia molto pesante a livello computazionale e quindi l’intero processo di addestramento richieda un tempo eccessivamente lungo per essere completato. Per rendere più veloce il training si potrebbe utilizzare l’algoritmo di discesa del gradiente stocastico, che aggiorna i parametri ad ogni elemento del database, anche se in questo modo la discesa è molto meno “diretta” verso il punto di minimo. Anche in questo caso esiste un compromesso, che garantisce una convergenza molto più veloce rispetto all’algoritmo classico e molto più diretta rispetto a quello stocastico, ed è il cosiddetto mini-batch gradient descent, in cui l’aggiornamento dei pesi viene fatto su una parte del database. Solitamente questi lotti di dati sono in numero compreso tra i 10 e i 100 elementi. Data la grandezza del database, si è scelto di utilizzare quest’ultimo algoritmo.

Un altro parametro su cui agire per portare a convergenza l’algoritmo e non allungare eccessivamente i tempi di addestramento, è il numero di epoche che questo deve durare. Per epoca si intende il numero di iterazioni necessario perché l’intero dataset venga interamente processato almeno una volta. Il numero “giusto” di epoche è quello minimo che porta vicino al punto di minimo la funzione di costo e anch’esso va valutato per tentativi.

Lo shuffling è un’opzione che consiste nel mescolamento dei dati all’interno dei “lotti”. Questo rende ancora più casuale il percorso verso il punto di minimo ed evita che il ripetersi di una combinazione di dati con un gradiente sfavorevole rallenti la discesa dello stesso e quindi renda più lungo il

processo di addestramento. Questa opzione è stata attivata, con il mescolamento che avviene alla fine di ogni epoca. L'altra opzione disponibile era di effettuare l'aggiornamento una sola volta, ma non era vantaggiosa.

3.1.4 Ambiente di esecuzione

Un'altra impostazione su cui lavorare per migliorare la prestazione dell'addestramento della rete è quella riguardante l'ambiente di esecuzione. Il settaggio standard è quello che fa eseguire l'algoritmo dalla CPU. Questo ambiente garantisce il funzionamento su qualsiasi dispositivo, ma ovviamente è molto meno prestazionale rispetto all'utilizzo della scheda grafica per la computazione. Al contrario, eseguire l'addestramento in ambiente GPU va ad aumentare nettamente la velocità di elaborazione, ma può imporre delle limitazioni legate alle prestazioni, in particolare sulla grandezza dei *mini-batch*. Lotti troppo grandi vanno infatti a occupare interamente la memoria e impediscono il completamento dell'addestramento. Inoltre, non tutte le schede video sono compatibili con tutte le versioni di MATLAB. Di conseguenza, va verificata la compatibilità del proprio hardware con la versione di software installata sul dispositivo.

3.1.5 Impostazioni di addestramento applicate

Per applicare le impostazioni personalizzate per l'addestramento della rete neurale si crea una apposita variabile tramite la funzione `trainingOption`. Le impostazioni finali per il training della rete neurale che opera sugli intervalli di quattro ore, scelte in base ai risultati ottenuti e alla velocità d'esecuzione tra varie combinazioni, sono state le seguenti:

```
trainingOptions("adam","InitialLearnRate",0.0001,"LearnRateDropFactor",0.1,"LearnRateDropPeriod",10,"Plots","training-progress","MiniBatchSize",32,"MaxEpochs",14,"Shuffle","every-epoch","ExecutionEnvironment","gpu")
```

- Algoritmo di ottimizzazione: "adam" (*adaptive moment estimation*)
- Numero di elementi dei "lotti" (*mini-batch size*): 32
- Fattore di apprendimento iniziale: 0,0001
- Numero massimo di epoche: 14
- Fattore di decadimento e periodo di decadimento: 0.1 (è un fattore moltiplicativo, pertanto deve essere <1); 10 epoche
- Shuffling alla fine di ogni epoca
- Ambiente d'esecuzione: gpu

Viene inoltre attivata la visualizzazione a schermo dei progressi dell'addestramento tramite grafici che vengono aggiornati in tempo reale, come vedremo successivamente.

3.2 Plot dell'addestramento e prima verifica dei risultati

Graficare i progressi dell'addestramento risulta importante per poter avere un primo riscontro sull'efficacia dell'addestramento stesso ed eventualmente per farsi un'idea preliminare sul tipo di problemi che affliggono la rete e intervenire in maniera mirata sulle impostazioni per correggerli. La finestra di plotting contiene informazioni riguardo l'avanzamento dell'addestramento (durata ed epoca, aggiornate in tempo reale), l'errore, l'hardware utilizzato e il fattore di apprendimento. I dati plottati sono l'errore e la radice dell'errore quadratico medio (*Loss, root mean square error*).

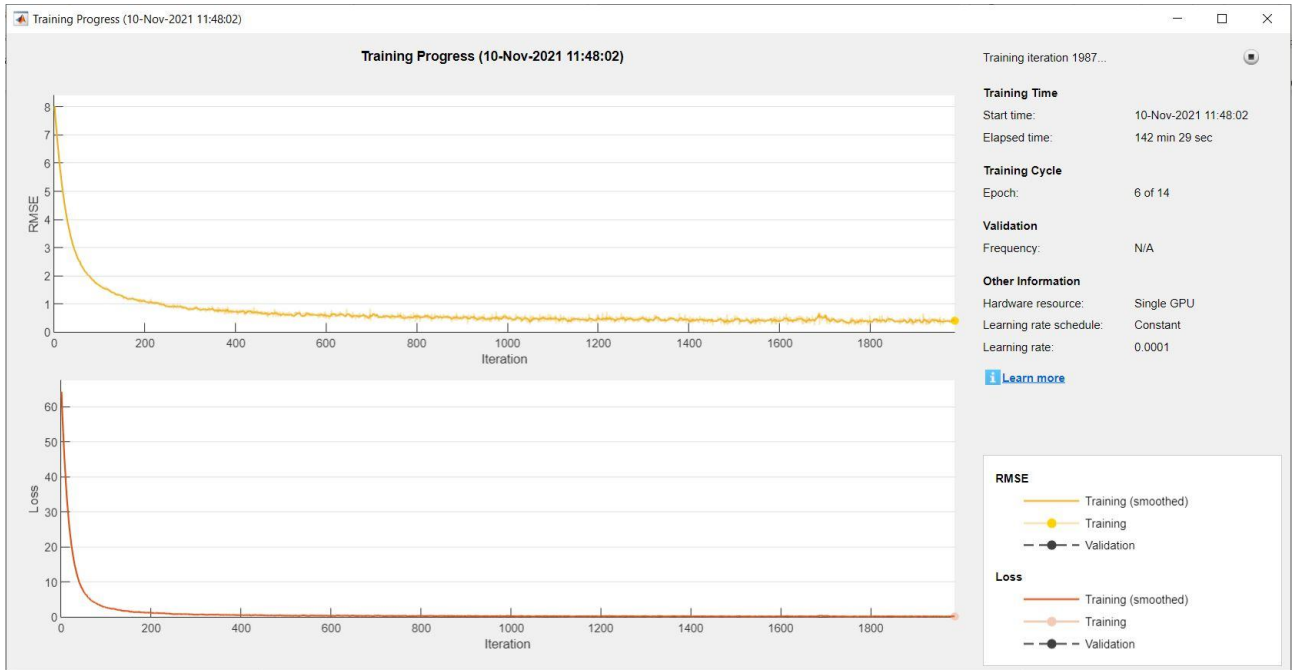


Figura 3.2: esempio della finestra in cui vengono plottati i progressi dell'addestramento

3.2.1 Errore ed errore quadratico medio

L'errore e l'errore quadratico medio, normalmente, sono intesi rispettivamente come la differenza tra la previsione e il dato reale e la media dei quadrati degli errori di ciascuna previsione. Nel caso in cui la previsione della rete non sia di tipo numerico, la definizione dell'errore, e di conseguenza della sua media quadratica, non è banale. Prima di analizzarne l'andamento e i possibili significati, è necessario quindi fare un approfondimento riguardo la definizione di errore per la rete di identificazione di immagini.

Per fare ciò, ci affidiamo alla guida del software MATLAB^[25], che va a definire questa funzione per calcolare l'errore (*training loss*):

$$\begin{aligned}
 & K_1 \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\
 & + K_1 \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} [(\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2] \\
 & + K_2 \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} (C_i - \hat{C}_i)^2 \\
 & + K_3 \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{noobj} (C_i - \hat{C}_i)^2 \\
 & + K_4 \sum_{i=0}^{S^2} 1_i^{obj} \sum_{c \in \text{classes}} (p_i(c) - \hat{p}_i(c))^2
 \end{aligned}$$

Figura 3.3: definizione della "training loss" per l'addestramento delle reti YOLO^[25]

In cui:

- S è il numero di celle della griglia.
- B è il numero di bounding box in ciascuna cella.
- $1obj_{ij}$: è 1 se la j -esima bounding box nella cella è responsabile dell'identificazione dell'oggetto. Altrimenti è pari a 0. Una cella i è responsabile della indentificazione dell'oggetto, se la sovrapposizione tra la ground truth e la bounding box in quella cella è maggiore o uguale a 0.6.
- $1noobj_{ij}$ è pari a 1 se la j -esima bounding box nella i -esima cella non contiene nessun oggetto, altrimenti è pari a 0.
- $1obj_i$ è pari a 1 se un oggetto viene identificato all'interno della i -esima cella, altrimenti è pari a 0.
- $K1, K2, K3,$ and $K4$ sono dei parametri che dipendono da una proprietà del layer di output della rete. $K1$ è il peso relativo all'errore di localizzazione. $K2$ e $K3$ sono pesi che valutano quando l'oggetto viene identificato ($K2$) o non viene identificato ($K3$) all'interno della sua bounding box. $K4$ è il peso relativo agli errori di classificazione (legato alla probabilità delle varie classi all'interno di una certa bounding box).
- (x_i, y_i) : coordinate relative al centro della j -esima box all'interno della i -esima cella. Se hanno l'apice $^{\wedge}$ sono relative alla ground truth box
- w_i, h_i : larghezza e altezza della j -esima box all'interno della i -esima cella. L'apice $^{\wedge}$ mantiene medesimo significato.
- C_i : la confidenza con la quale viene identificata la j -esima bounding box nella i -esima cella.
- $p_i(c)$: la stima della probabilità per la classe c all'interno della i -esima cella.

Questa definizione può essere divisa in tre parti, ciascuna con un suo specifico significato:

- Errore di localizzazione: è rappresentato dai primi due termini e misura l'errore tra la bounding box prevista dalla rete e quella della ground truth.
- Errore di confidenza: è rappresentato dal terzo e dal quarto termine. Il terzo misura l'errore nella confidenza quando un oggetto viene identificato all'interno della j -esima bounding box, mentre il quarto misura l'errore quando non viene identificato un oggetto all'interno della j -esima box. Questo errore può portare l'algoritmo di addestramento a divergenza se il numero di celle che non contengono oggetti sono in numero maggiore rispetto a quelle che ne contengono. In questo caso va incrementato il valore di $K2$ e diminuito quello di $K3$.
- Errore di classificazione: è rappresentato dal quinto termine. Se un oggetto viene identificato all'interno della i -esima cella, questo termine misura l'errore quadratico tra le probabilità assegnate a ciascuna classe nella i -esima cella.

La radice dell'errore quadratico medio serve ad aggregare gli errori delle varie previsioni e le loro grandezze ed è un indice della capacità predittiva della rete. Ovviamente, tanto minore è il valore assunto quanto migliore sarà la capacità predittiva.

$$RMSE = \sqrt{\frac{\sum_{k=1}^n (y_t - \hat{y}_t)^2}{n}} \quad (3.1)$$

Il calcolo di RMSE e Loss viene effettuato ad ogni iterazione. Oltre al valore assoluto assunto, è molto importante analizzare l'andamento di questi due parametri nel tempo per avere una prima idea di come sta lavorando la rete neurale. In primis, è fondamentale che i parametri siano funzioni decrescenti in funzione del numero di iterazioni (come nel grafico della fig. 3.2). Ciò significa che la rete diventa sempre più precisa a mano a mano che procede l'addestramento. In secondo luogo, non si devono presentare andamenti "irregolari", ma si deve cercare di ottenere una curva che si avvicini il più possibile a quella di una funzione monotona. Per fare questo, si deve osservare il comportamento generale, magari con una curva di tendenza, in quanto, soprattutto utilizzando il metodo mini-batch, delle piccole oscillazioni tra una iterazione e l'altra sono fisiologiche. Come si può osservare nel grafico, la funzione tende ad assumere carattere asintotico. Il tratto piatto della funzione non deve essere troppo corto (per essere sicuri che la funzione dell'errore non diverga e quindi di essere vicini ad un punto di minimo della funzione di costo), né troppo lungo, in quanto questo significherebbe "sprecare" tempo per miglioramenti non sensibili nel comportamento della rete. Se sto utilizzando l'intero dataset per l'addestramento si può banalmente decidere di fermare l'algoritmo dopo un certo numero di iterazioni consecutive in cui l'errore non ha variazioni

significative. Ben più difficile è valutare l'efficienza della rete addestrata con il metodo mini-batch. Infatti, soprattutto in casi come quello in esame in cui il dataset è molto grande, essendo la funzione di errore calcolata sul lotto usato in quella determinata iterazione, per avere una buona sicurezza sul buon esito dell'addestramento serve un numero molto maggiore di iterazioni (almeno un'epoca con andamento "stabile" è necessaria, ma non sufficiente a garantire la sicurezza riguardo la precisione sull'intero set).

3.2.2 Utilizzo dei set di validazione e test

Un altro valido strumento per valutare il comportamento della rete è quello di utilizzare un set di validazione e/o di test. Solitamente, questi devono essere differenti rispetto al set di addestramento per dare un'idea della reale efficacia della rete nell'identificare gli oggetti in contesti diversi da quelli già conosciuti dalla rete stessa. Nel nostro caso, essendoci classi con poche o addirittura una sola curva disponibile e conseguentemente un numero limitato di spettrogrammi, è pressoché impossibile riuscire a creare un set nuovo per la validazione, e per questo motivo si è preferito procedere con la validazione sull'intero test set effettuata alla fine di ogni epoca. Utilizzando il metodo mini-batch, questa validazione dà comunque un'indicazione migliore sul comportamento complessivo della rete, che invece è più difficilmente captabile dall'andamento sulle singole iterazioni, che considerano solo una parte del set, per giunta casuale a causa del mescolamento.

Anche per i set di validazione e/o di test vengono calcolati errore e radice dell'errore quadratico medio e ne va valutato attentamente il comportamento, in quanto non sempre è coerente con quello del set di addestramento. È naturale aspettarsi che, essendo la rete addestrata su dati totalmente o parzialmente diversi da quelli che si vanno ad analizzare nella validazione, l'errore sia maggiore. La cosa importante è che l'errore vada a decrescere a mano a mano che si procede con l'addestramento, proprio come per il test set, in quanto ciò comporta il sempre migliore comportamento della rete nell'elaborare dati fino a quel momento sconosciuti o diversi da quelli utilizzati nelle ultime iterazioni di addestramento. Se l'errore sui dati di validazione rimane molto maggiore significa che si ha varianza elevata (o overfitting) e quindi la rete non è in grado di riconoscere gli oggetti non appena subiscono piccole variazioni. Dal grafico si nota che il divario tra le curve del validation set e del training set rimane elevato o addirittura aumenta col progredire dell'addestramento. In questi casi, solitamente la soluzione più semplice è quella di ampliare il database o aumentare il fattore di apprendimento. Nel caso in cui sia l'errore sul training set che quello sul validation set rimangano elevati, siamo di fronte al problema di BIAS, per cui la rete fatica a riconoscere correttamente gli oggetti, siano essi noti o meno. In questo caso, invece, è consigliabile agire aumentando la complessità della rete (numero di layer, grado delle funzioni, numero di features) o diminuendo il fattore di apprendimento.

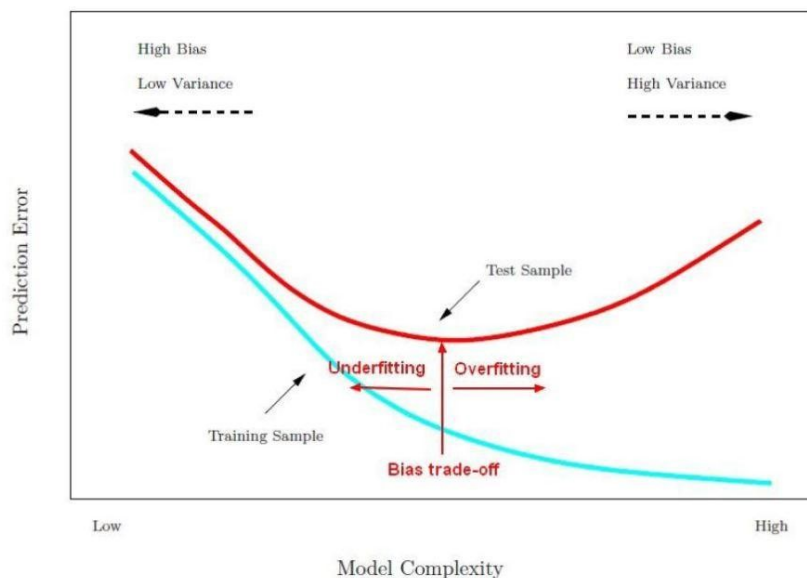


Figura 3.4: analisi del comportamento dell'algorithm in base all'andamento dell'errore in funzione della complessità del modello^[26]

3.3 Processo decisionale sui parametri di training

Per trovare il miglior setting possibile si è proceduto facendo delle analisi di primo tentativo su un database ridotto composto da spettrogrammi di durata di un'ora per analizzare il comportamento della rete su questo tipo particolare di immagini e per trovare le aree su cui lavorare e investire maggiori risorse in termini di tempo per cercare di ottenere il miglior risultato possibile.

3.3.1 Analisi preliminari

Come anticipato, una volta decisa l'architettura della rete si è iniziato a valutare il suo comportamento su un database limitato di spettrogrammi di durata di un'ora. Questo ha comportato l'esclusione di una parte delle classi di elettrodomestici e quindi ha notevolmente ridotto il numero totale di spettrogrammi utilizzati. Inoltre, i dati non sono stati elaborati in modo tale da distribuire gli spettri in maniera casuale su tutto l'intervallo temporale, ma ci si è limitati a considerare solamente quelli ottenuti dalla elaborazione diretta della curva estratta dai database. Si può prevedere l'insorgere del problema di overfitting, dato lo scarso numero di elementi presenti nel database, ma possiamo trarre delle conclusioni sulla capacità di identificazione della rete. Una volta addestrata la rete YOLO, si è notato come questa riuscisse a riconoscere i segnali del training set in modo molto efficace, anche quando questi erano molto simili in termini di durata ed intensità, avallando quindi la scelta della rete.

CATEGORIA	RISULTATO	PERCENTUALE	MISS	ERROR	NOTE
Air conditioner	2/2	100			
Asciugacapelli	14/16	87.5	2		2 doppie detection a confidence ca. 60%
Bollitore	5/6	83.33333333		1	2 doppie detection a confidence ca. 60%, errore con aspirapolvere
Coffee machine	10/10	100			
Forno	8/8	100			
Ferro da stiro	5/5	100			
Lavasciuga	2/2	100			
Microonde	17/18	94.44444444	1		1 confidence <60%
PC charger	2/2	100			
Piastra induzione	2/2	100			
Tagliaerba	4/4	100			
Tostapane	4/4	100			1 confidence <60%
TV	1/1	100			
Aspirapolvere	11/11	100			

Tabella 3.1: tabella riassuntiva dei risultati della rete di primo tentativo

Dalla tabella 3.1 si può notare come, su un dataset ridotto, la percentuale di immagini correttamente riconosciuta è elevatissima, superiore al 95%, con 3 mancate identificazione e una sbagliata. Inoltre, è risultata molto elevata anche la confidenza con la quale sono state effettuate, quasi sempre superiore al 75%, tranne in pochi casi in cui era comunque attorno al 60%. Si sono riscontrati quattro casi in cui sono state individuate delle bounding box "doppie", intese come box quasi sovrapposte che identificavano lo stesso spettro. Ottenuti i risultati sul training set, si è proceduto a verificare gli stessi su due tipi diversi di spettrogramma: uno ottenuto dai dati delle rilevazioni di consumo aggregato e uno costruito artificialmente inserendo casualmente segnali di elettrodomestici presenti nel set di addestramento. In entrambi i casi la rete ha fatto molta fatica a riconoscere i segnali. Nello specifico si è notato come venissero riconosciuti solo i segnali posti nelle immediate prossimità dell'istante in cui avvenivano anche nel test set, evidenziando una forte tendenza all'overfitting. Andando ad agire sul learning rate si è solo passati da un comportamento ad alta varianza ad uno con alto BIAS. Il passaggio successivo è stato, di conseguenza, quello di allargare il database. Per farlo e includere il maggior numero di dati possibili, è stato deciso anche di modificare l'intervallo di tempo su cui elaborare gli spettrogrammi, passando da 1 a 4 ore.

CAPITOLO 4

ANALISI DEI SEGNALI DEI SINGOLI ELETTRODOMESTICI

Per comprendere al meglio il comportamento della rete, è importante conoscere le caratteristiche degli elettrodomestici che si vuole riconoscere. Oltre alla durata, infatti, sono importanti l'intensità del segnale e come viene distribuita la potenza tra le varie frequenze. Di conseguenza influiranno, seppur in maniera indiretta, la potenza assorbita, l'andamento della curva del segnale (se è di tipo ON/OFF, presenta transitori o se è composto da più cicli concatenati), il modo in cui viene assorbita la potenza (se l'elettrodomestico è più resistivo o induttivo). Analizzando i vari segnali presenti nel database si può fare una previsione su quali saranno quelli più "singolari", quindi facilmente distinguibili dagli altri, e quali invece presentano similitudini e quindi possono risultare critici.

4.1 Segnali di lunga durata

4.1.1 Frigorifero e freezer

Come precedentemente anticipato, frigorifero e freezer sono due segnali simili. Il loro funzionamento avviene lungo l'intero arco della giornata. La curva presenta dei picchi, separati tra loro da intervalli a potenza nulla, di intensità e durata che variano da modello a modello. Tipicamente i picchi hanno durata inferiore agli intervalli che li separano ed entrambi sono dell'ordine di qualche minuto (o della decina di minuti)^[27]. In generale, i frigoriferi assorbono circa il 50% di potenza in più rispetto ai freezer ed entrambi presentano un fattore di potenza medio quasi unitario. Sono quindi elettrodomestici con comportamento resistivo, nonostante la presenza di un compressore che in avviamento dà un brevissimo picco di potenza apparente. Le curve sono tra loro molto simili, ma la differenza di potenza assorbita rende più facile il riconoscimento da parte della rete. Essendo attivi su tutto l'arco temporale, nelle prove finali sono stati esclusi dai segnali aggregati simulati in quanto andrebbero a coprire eccessivamente gli altri segnali presenti e quindi renderebbero più difficile una corretta valutazione del funzionamento della rete neurale.

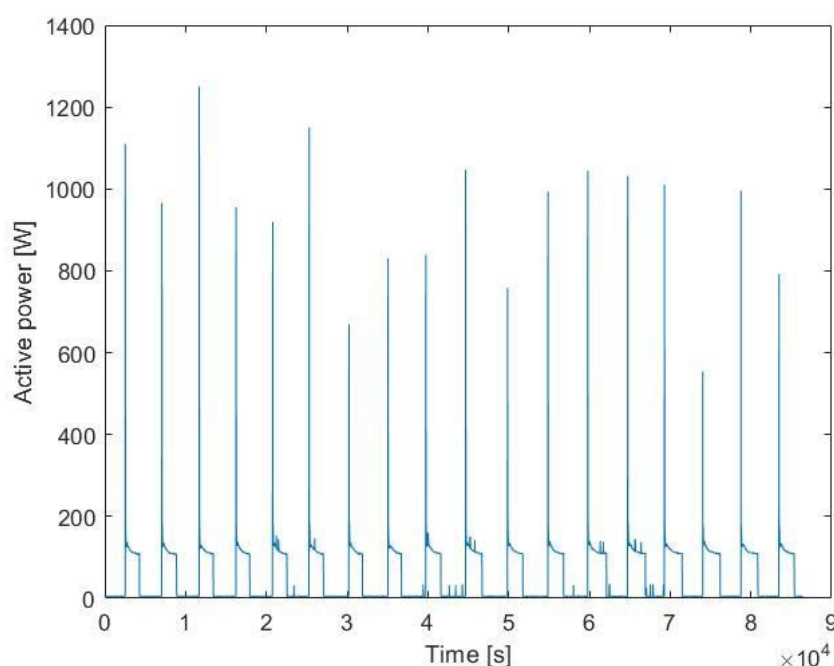


Figura 4.1: curva di potenza attiva di un frigorifero lungo l'arco di 24 ore

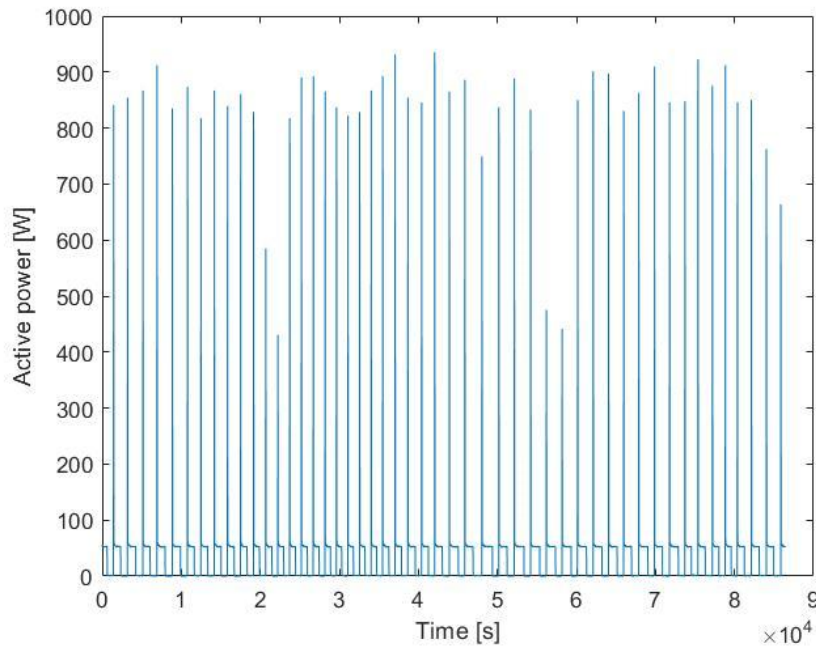


Figura 4.2: Curva di potenza attiva assorbita da un freezer durante l'arco di 24 ore

4.1.2 Asciugatrice

L'asciugatrice è un elettrodomestico il cui funzionamento consiste nel susseguirsi di diversi cicli ON/OFF che seguono un primo ciclo che si apre all'attivazione dell'elettrodomestico. Ciascuno dei cicli ha un suo fattore di potenza, che varia tra 1 e 0,4^[27] (relativo all'asciugatura senza apporto di calore), pertanto non ha senso parlare di un fattore medio lungo l'intero funzionamento dell'elettrodomestico. La potenza totale, la durata dei cicli e il tipo di cicli utilizzati variano a seconda del programma e del carico, essendoci un controllo della temperatura presente all'interno del cestello.

Nella curva in figura, si notano il susseguirsi di cicli ON/OFF durante il funzionamento e il tipico ciclo di asciugatura senza apporto di calore finale. Risulta evidente il basso fattore di potenza proprio di quest'ultimo.

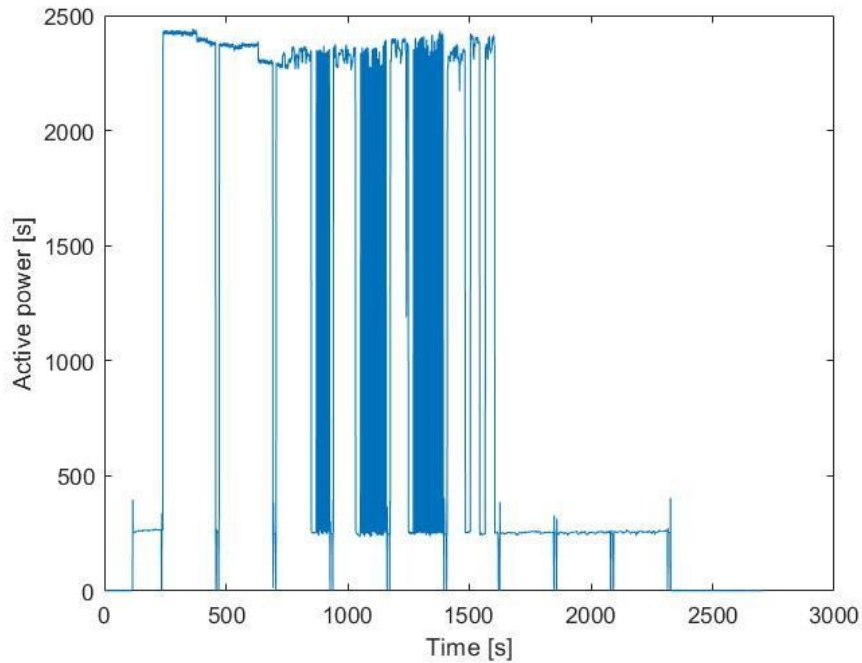


Figura 4.3: curva di potenza attiva di una asciugatrice

4.1.3 Lavastoviglie

Il funzionamento di una lavastoviglie consiste di tre cicli concatenati: lavaggio, risciacquo e asciugatura. All'interno dei cicli ci sono delle pause, a basso consumo energetico, in cui l'elettrodomestico va a scaricare i liquidi in quel momento al suo interno. Il fattore di potenza varia da 1, nelle fasi di riscaldamento dell'acqua e di asciugatura, fino a 0,6. Un programma di lavaggio dura di norma tra l'ora e le due ore, con picchi di energia assorbita fino ai 2,5/3 KW.

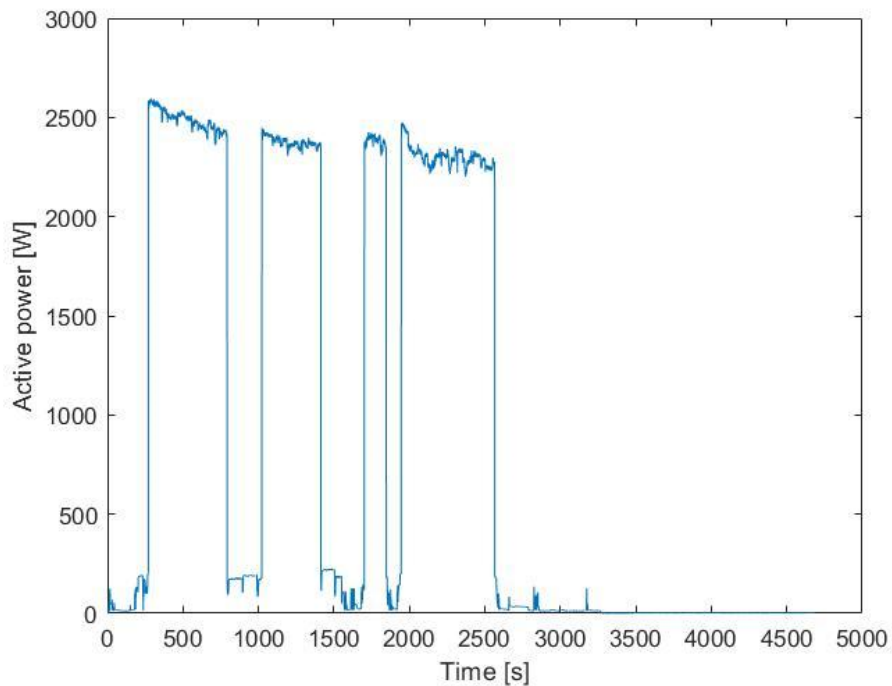


Figura 4.4: curva di energia attiva assorbita da una lavastoviglie. Si notano le "pause" per lo scarico dell'acqua e i picchi dovuti al riscaldamento dell'acqua e all'asciugatura

4.1.4 Lavatrice

La lavatrice, analogamente alla lavastoviglie, presenta una fase di lavaggio e una di risciacquo, che sono però seguite dalla centrifuga e dalla fase di scarico finale. Anche qui le fasi di riscaldamento hanno fattori di potenza unitari, essendo resistivi, mentre durante le fasi di movimentazione, a causa dell'utilizzo di motori elettrici induttivi, il fattore scende. In particolare, durante la fase di centrifuga scende sotto 0,6 e addirittura può arrivare a valori inferiori a 0,2. I valori di potenza assorbita, la durata del funzionamento e delle varie fasi sono soggette a variazioni a seconda del modello e dei programmi utilizzati, anche se in genere le fasi rimangono le stesse. La durata varia solitamente tra una e tre ore, mentre la potenza assorbita arriva fino a circa 3 KW.

Avendo delle similitudini con la lavastoviglie, si può prevedere che in casi particolari e con particolari sovrapposizioni di segnali i due elettrodomestici possano essere confusi tra loro.

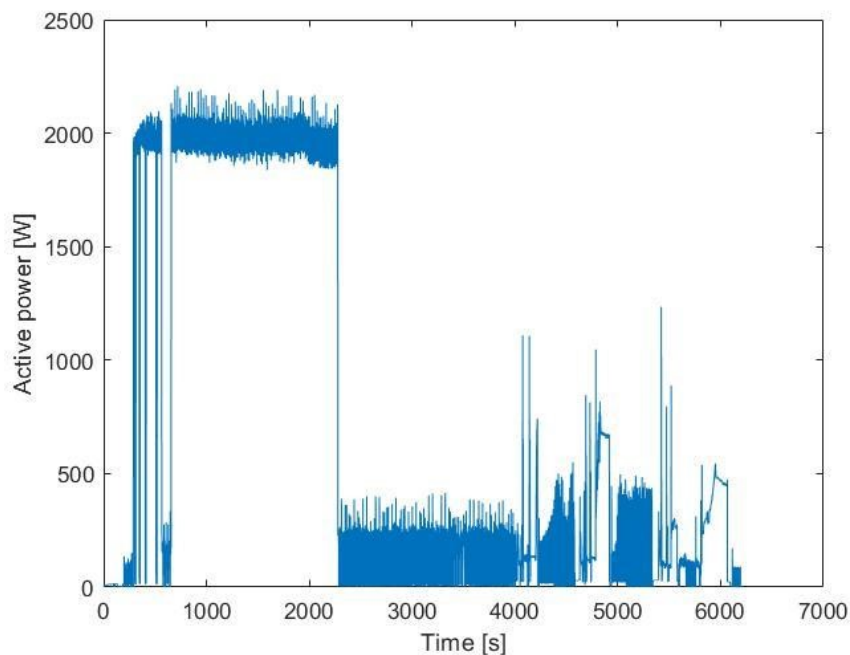


Figura 4.5: curva di consumo di una lavatrice

4.2 Segnali di durata medio-breve

Per la nostra applicazione, andiamo a definire segnali di medio-breve durata quelli inferiori all'ora o appena oltre a questa durata. Nel primo caso, i segnali brevi, si avranno sicuramente difficoltà nella loro identificazione con la rete neurale che identifica gli spettri su intervalli di quattro ore, mentre l'efficacia della rete sui secondi è da verificare. Se non si riuscisse ad ottenere una sufficiente precisione nelle prediction della rete neurale sui segnali corti, si renderebbe necessario andare ad implementare una seconda rete atta ad identificare i segnali su intervalli più corti, di un'ora, in modo da aumentare la dimensione degli spettrogrammi all'interno dell'immagine che rappresenta l'intero intervallo.

Gli elettrodomestici che rientrano in questa classificazione sono: asciugacapelli, aspirapolvere, bollitore, bisticchiera, cappuccinatore, ferro da stiro, forno, macchina del caffè, microonde, piano da cottura, piastra a induzione, robot da cucina, tagliaerba e tostapane. Nella composizione dei segnali aggregati sono stati esclusi cappuccinatore e bisticchiera, ritenuti meno interessanti per lo scarso numero di elementi presenti nel database e perché meno comuni rispetto, ad esempio, a microonde o macchina del caffè.

La maggior parte di questi possono essere semplicemente classificati come dispositivi ON/OFF, perlopiù resistivi. Questo, fa sì che le differenze tra un elettrodomestico e l'altro stiano nella potenza assorbita e nella durata del segnale. Quando tutti questi elementi sono simili, ovviamente, sarà più complicato il riconoscimento, in particolare quando la rete addestrata su intervalli di quattro ore prova a identificare segnali simili dell'ordine dei minuti. Di questa categoria non fanno parte il ferro da stiro, il microonde e il forno, che vengono analizzate separatamente.

4.2.1 Ferro da stiro

Il ferro da stiro è un elettrodomestico prevalentemente resistivo. La curva di consumo consiste in una serie di picchi di durata variabile ad intervalli variabili. La durata di utilizzo dipende dall'utente. In questo caso, è importante cercare di avere un'ampia varietà di tipologie di curve all'interno del database in modo tale da riuscire a coprire il più possibile la variabilità dei picchi e della durata totale di utilizzo. Questa variabilità fa sì che il segnale possa essere confuso con altri che presentano un funzionamento intermittente (asciugatrice, microonde, forno e, in particolari casi, macchina del caffè). La potenza massima può raggiungere circa 2,5 KW.

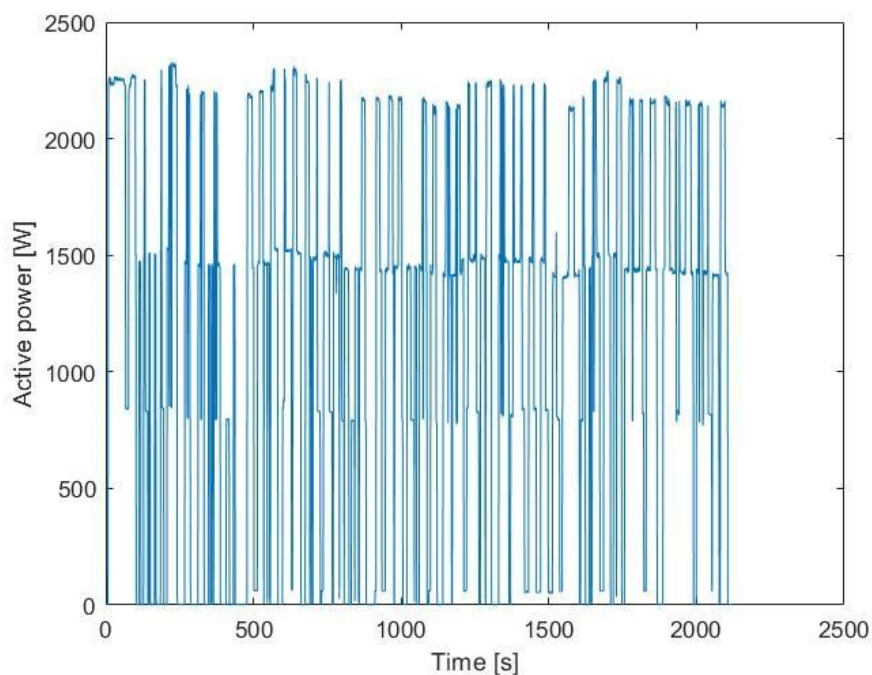


Figura 4.6: curva di consumo di un ferro da stiro. Durata e intervalli dei picchi possono variare, così come la durata totale, che dipende dall'utilizzatore

4.2.2 Forno

Il forno elettrico è un elettrodomestico quasi puramente resistivo, che presenta quindi un fattore di potenza pressoché unitario. La curva di consumo presenterà dei picchi, sempre alla stessa potenza, e degli intervalli. La durata degli stessi dipende di volta in volta dall'operatore e dall'utilizzo del forno. Anche in questo caso, esiste la possibilità che il segnale venga confuso con gli altri segnali intermittenti presenti nel database.

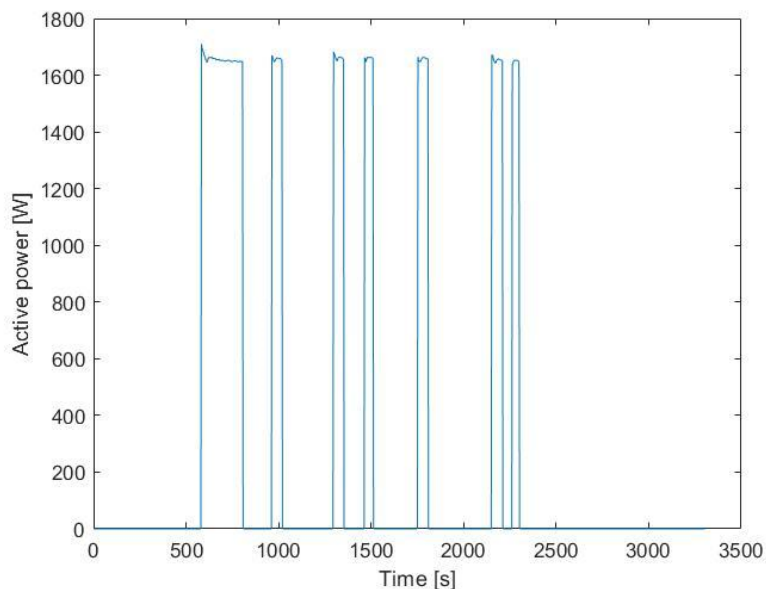


Figura 4.7: curva di consumo di un forno

4.2.3 Microonde

Il forno a microonde non ha una curva di consumo tipica ben definita, in quanto può sia somigliare ad un segnale a gradino, quindi avere un unico picco, sia essere un segnale intermittente, dipendentemente dal modello di forno. A differenza degli altri elettrodomestici, non è un elettrodomestico totalmente resistivo e presenta un fattore di potenza di circa 0,75^[28].

La variabilità della curva può rendere lo spettro più difficilmente riconoscibile. La rete, infatti, dovrà cercare di riconoscere sia i segnali a gradino che quelli intermittenti e, non potendosi “concentrare” unicamente sui secondi, questi ultimi potranno più facilmente essere confusi con altre curve di segnali intermittenti. Queste considerazioni hanno ovviamente peso maggiore su segnali modificati (come nel caso di quelli inseriti nell’aggregato simulato) e su segnali diversi da quelli del database, che potrebbero essere trovati in una abitazione reale.

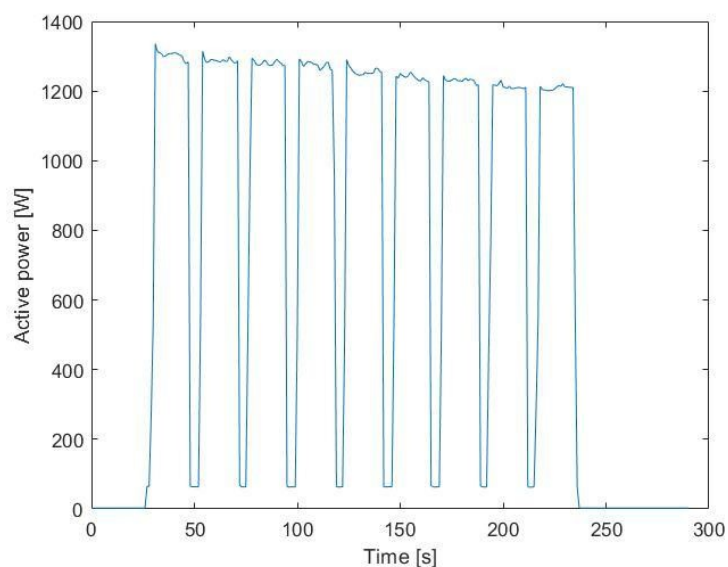


Figura 4.8: Curva di consumo di un forno a microonde

4.2.4 Curve elettrodomestici

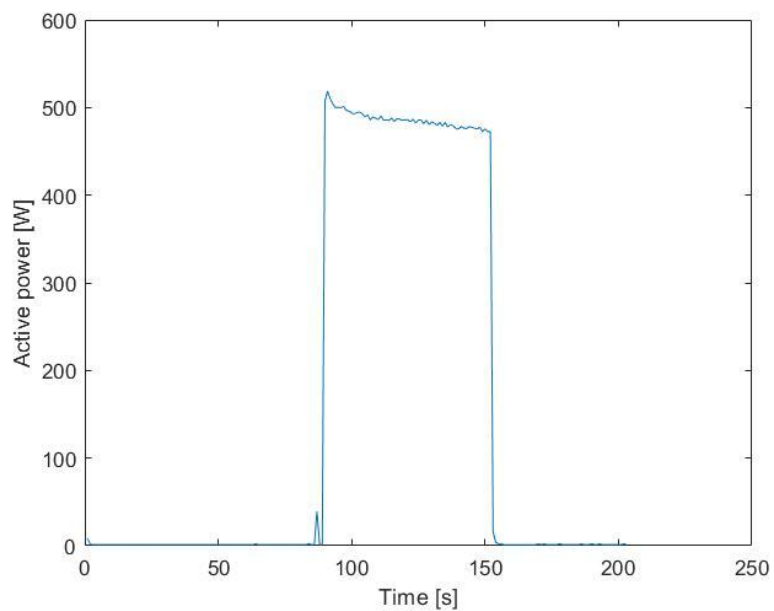


Figura 4.9: curva di consumo di un cappuccinatore

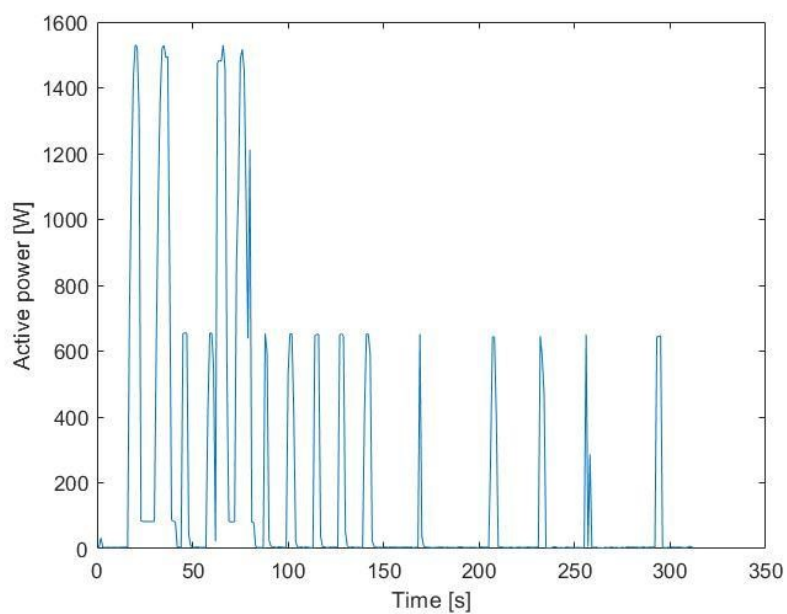


Figura 4.10: curva di consumo di una macchina del caffè. Ci sono anche curve di questo elettrodomestico che sono a gradino (quindi di tipo ON/OFF)

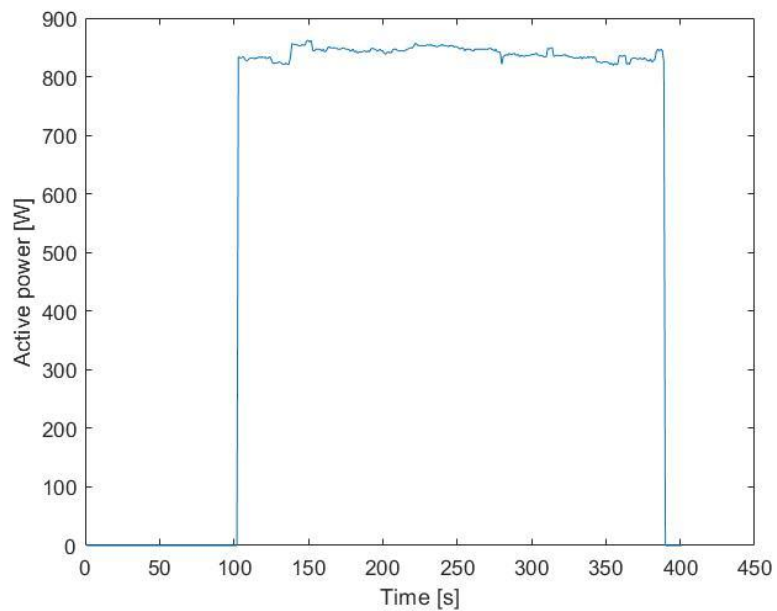


Figura 4.11: curva di consumo di un piano cottura

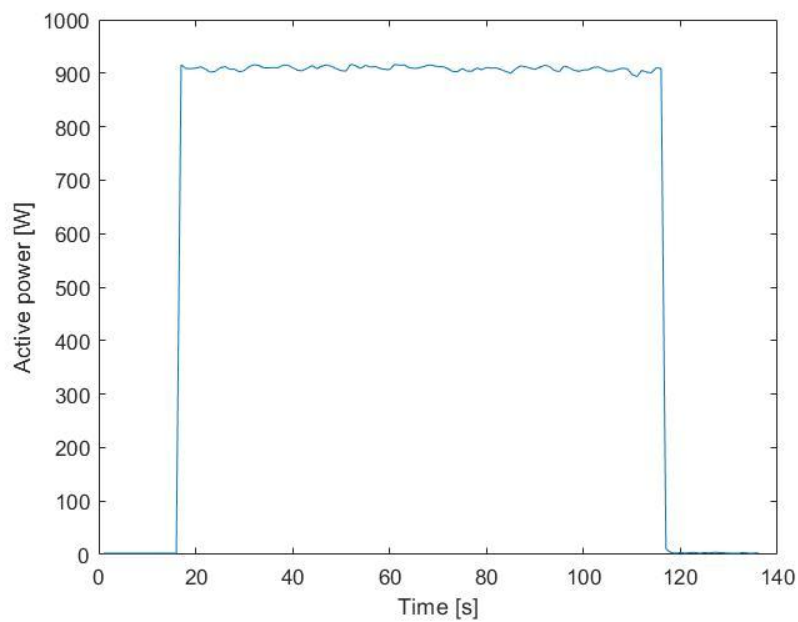


Figura 4.12: curva di consumo di un asciugacapelli

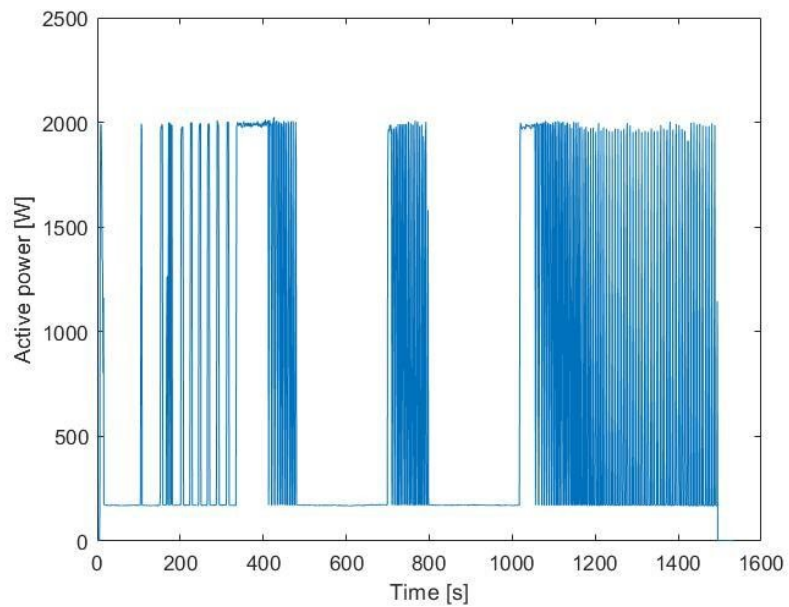


Figura 4.13: curva di consumo di una piastra a induzione

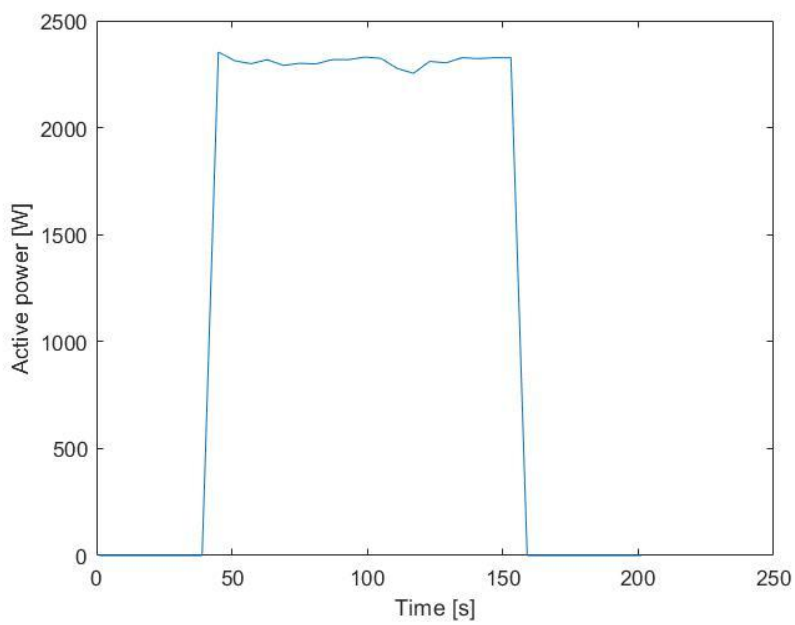


Figura 4.14: curva di consumo di un bollitore

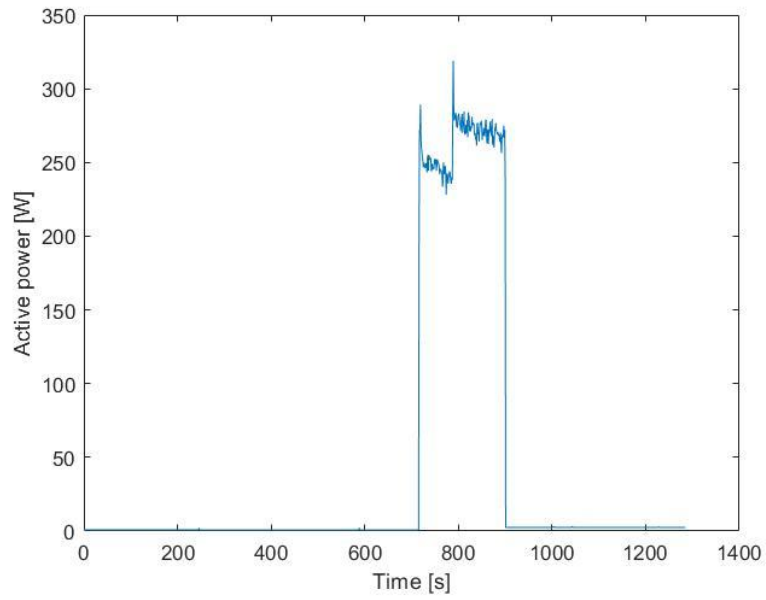


Figura 4.15: curva di consumo di un robot da cucina

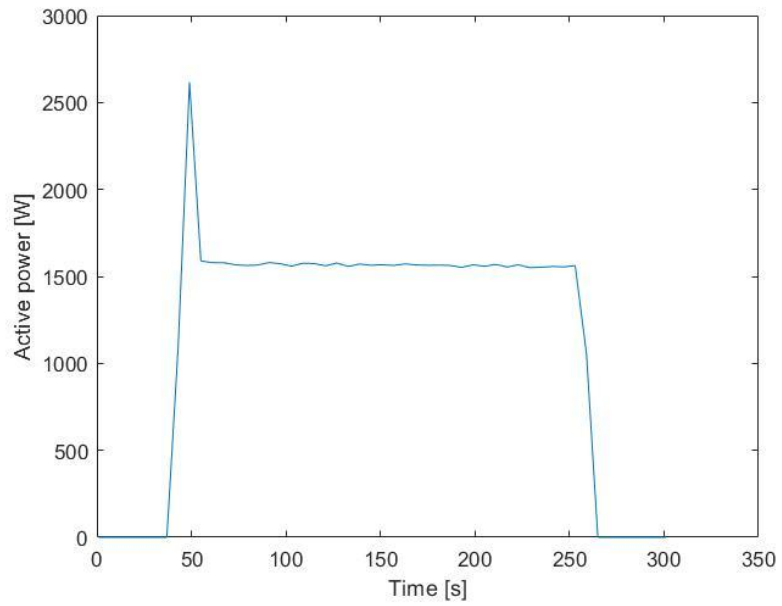


Figura 4.16: curva di consumo di un tostapane

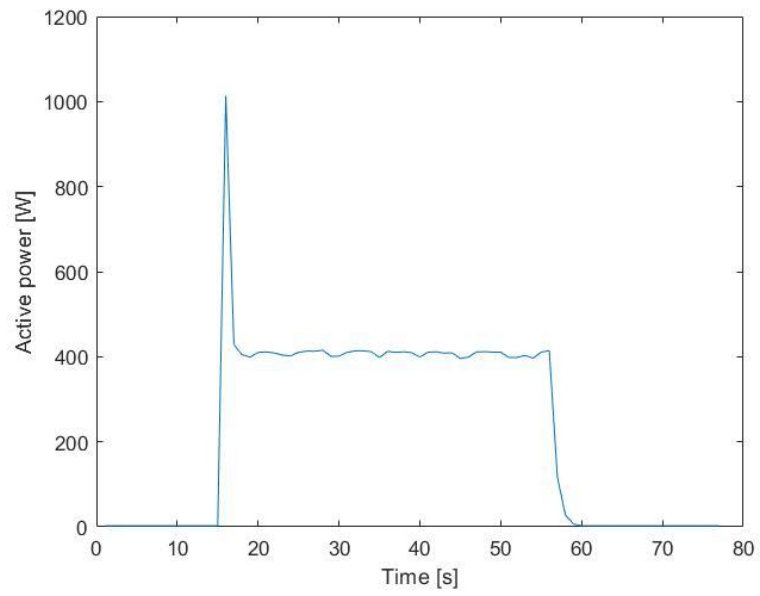


Figura 4.17: curva di consumo di un'aspirapolvere

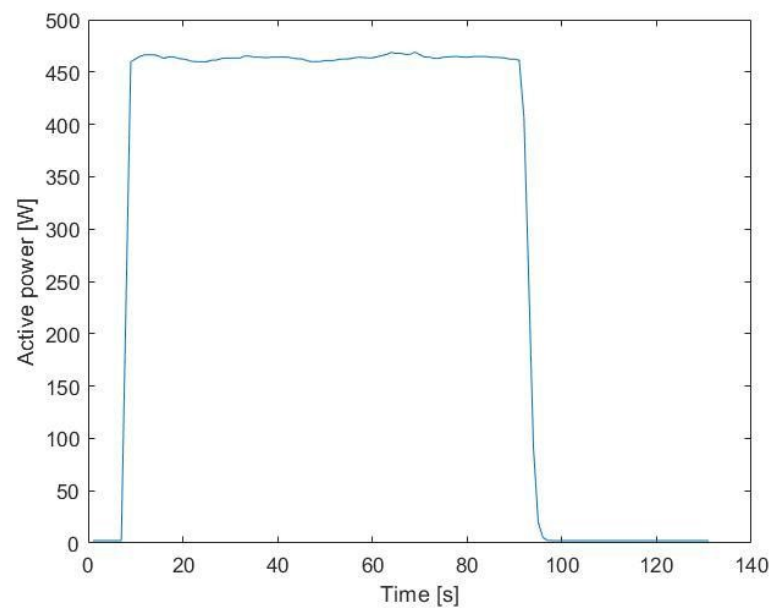


Figura 4.18: curva di consumo di un tagliaerba elettrico

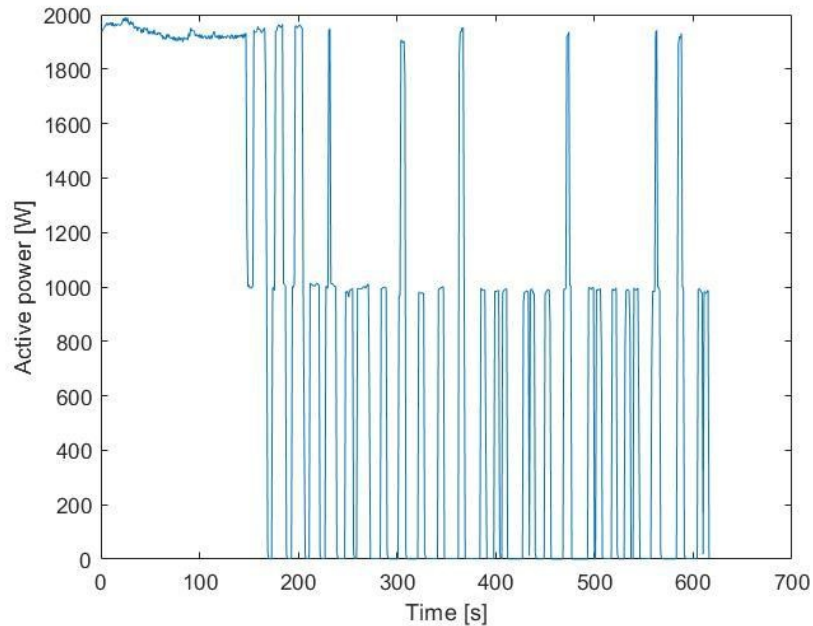


Figura 4.19: curva di consumo di una bistecchiera

Si può notare come quasi tutti i segnali siano dell'ordine dei pochi minuti e che alcuni di questi abbiano curve dall'aspetto molto simile tra loro. In particolare, bollitore, piano cottura e asciugacapelli hanno una curva a gradino e addirittura piano cottura e asciugacapelli hanno anche livelli di potenza simili. Non essendo elettrodomestici con un tempo di funzionamento definito, deve essere messo in preventivo che la rete possa confondere tra loro questi segnali. Analogamente, tra aspirapolvere e tostapane ci sono similitudini nella curva. Non essendoci grandissime differenze, se non nella parte iniziale, tra tutti questi, dipendentemente dai livelli di potenza, che variano anche da modello a modello, e dal tempo di utilizzo, che spesso dipende dall'utilizzatore, le possibilità di avere detection sbagliate aumenta ulteriormente. Gli unici segnali che si differenziano sono quello della piastra a induzione e quello della macchina del caffè, che però in particolari casi potrebbe confondersi con gli elettrodomestici a funzionamento intermittente.

Queste possibili difficoltà vengono ulteriormente accentuate negli intervalli di quattro ore, in quanto gli spettrogrammi di segnali così brevi si estendono per pochi pixel e le piccole differenze tra uno e l'altro, che vengono tramutate in ampiezza della fascia colorata (tempo) e differenza nel colore e nella sua distribuzione (potenza), diventano praticamente impercettibili anche per una rete neurale ben addestrata.

CAPITOLO 5

SVILUPPO DELLE RETI NEURALI

Una volta effettuate le analisi preliminari e analizzati i segnali degli elettrodomestici si è proceduto all'allargamento del database, come descritto nei capitoli precedenti. In particolare, si è cercato di posizionare con maggiore attenzione lungo tutto l'intervallo di quattro ore i segnali degli elettrodomestici, in modo tale da contrastare l'overfitting e slegare il più possibile l'identificazione dell'oggetto dalla sua posizione. Aumentando il numero di elementi, si è anche riusciti a diminuire il valore dell'IOU medio delle anchor box. Sono state aggiunte le classi degli elettrodomestici con cicli di durata maggiore dell'ora (ad esempio lavatrici e frigoriferi), portando dalle 14 classi presenti nel database precedenti alle 20 del database definitivo. Nonostante tutto, alcune classi sono rimaste "povere" di elementi a causa del piccolo numero di dati da elaborare e per questo come set di validazione si è andati ad utilizzare nuovamente il test set. Di queste classi sarà importante analizzarne il comportamento, attendendosi una maggiore possibilità di errori e mancate identificazioni anche sulle immagini del dataset. Avendo una base di partenza a livello di parametri di addestramento, si è potuto partire da quelli utilizzati nelle analisi preliminari e, una volta verificato che funzionassero, si è lavorato in un intorno di quei parametri per trovare la migliore combinazione.

5.1 Rete neurale per intervalli lunghi (4 ore)

5.1.1 Affinamento parametri training e analisi detection su training test

Il primo passaggio è stato quello di trovare il giusto compromesso tra dimensione del mini-batch e l'ambiente di elaborazione. Come visto in precedenza utilizzare la scheda grafica aumenta enormemente la velocità di computazione, ma di contro si può incorrere a problemi legati alla memoria necessaria per l'elaborazione o di compatibilità tra software e hardware presenti nel computer. Utilizzare la cpu aggira le limitazioni sulle dimensioni dei mini-batch a discapito della velocità. Si è notato come utilizzare "lotti" di dimensioni maggiori (anche di poco superiori ai 100 elementi, quindi al limite dei valori di solito utilizzati) su cpu non riesca a pareggiare le prestazioni del training su gpu nemmeno con un tempo doppio. La scelta è quindi ricaduta sull'addestramento in ambiente gpu con mini-batch composto da 32 elementi, che è un buon compromesso tra velocità e precisione della rete e rimane fruibile anche su dispositivi meno recenti o con prestazioni non molto elevate.

Si è provato anche a modificare il learning rate, ma aumentandolo la discesa del gradiente risulta troppo ripida, mentre diminuendolo converge troppo lentamente. Di conseguenza è rimasto invariato e pari a 0,0001, e si è solamente ad introdotto il decadimento con fattore 0,1 dopo l'ottava epoca per evitare il problema dell'overshooting. È risultato inutile anche provare a cambiare l'algoritmo di ottimizzazione, in quanto l'"adam" è risultato essere il più efficace, come preventivabile.

Come precedentemente scritto, data la mancanza di dati per alcune classi e l'impossibilità di creare in maniera automatica delle suddivisioni del database di addestramento, non si è proceduto con la creazione di un validation set completamente diverso, ma si è riutilizzato il set di addestramento. Questo anche perché le singole iterazioni considerano solo 32 elementi su 10982 e quindi, anche considerando il momento presente nell'algoritmo, i pesi non sono mai tarati sull'intero dataset. Nonostante queste considerazioni, le conclusioni tratte devono essere considerate parziali e solamente indicative. Si andrà successivamente a verificare l'efficacia del riconoscimento della rete attraverso gli spettrogrammi dei dati di consumo aggregati e soprattutto attraverso spettrogrammi simulati creati ad hoc con segnali noti dai database, anche modificati in intensità.

Detection	TOT	CORRECT (%)	MISSED (%)	WRONG (%)
Asciugacapelli	98	79.38	17.53	3.09
Asciugatrice	309	98.06	1.94	0
Aspirapolvere	132	86.36	12.12	1.52
Bistecchiera	26	96.15	3.85	0
Bollitore	1374	90.83	8.73	0.44
Cappuccinatore	30	76.67	23.33	0
Ferro da stiro	115	78.26	17.39	4.35
Forno	138	98.55	0.72	0.72
Freezer	990	100	0	0
Frigorifero	2679	99.96	0.037	0
Lavasciuga	8	0	0	100
Lavastoviglie	637	99.22	0.63	0.16
Lavatrice	965	99.69	0.21	0.10
Macchina del caffè	1408	96.31	3.13	0.57
Microonde	1280	84.77	13.44	1.95
Piano cottura	64	12.5	85.94	1.56
Piastra induzione	17	94.12	5.88	0
Robot da cucina	19	100	0	0
Tagliaerba	80	93.75	6.25	0
Tostapane	614	89.90	8.79	1.30

Tabella 5.1: tabella riassuntiva delle identificazioni effettuate dalla rete sul dataset di validazione

Possiamo dividere le classi in due categorie in base alla durata o al numero di spettrogrammi appartenenti alla classe stessa e andare ad analizzare i risultati secondo le caratteristiche più evidenti.

Si nota come i pochi casi di previsioni errate, ossia in cui viene prevista una classe sbagliata, siano quasi esclusivamente appartenenti agli elettrodomestici con segnali brevi, che vengono ovviamente confusi con altri elettrodomestici di quel tipo. In particolare, venivano erroneamente identificati altri elettrodomestici come microonde, macchina del caffè o bollitore, ossia le tre categorie con più esempi presenti all'interno del database. Si può ragionevolmente pensare che questi errori siano legati al fatto che le bounding box di questi elettrodomestici sono molto strette, data la durata breve del loro funzionamento, e l'intensità del segnale del carico sia simile, rendendo gli spettrogrammi pressoché identici. Essendoci pochi pixel valutati dalla rete neurale ci sono anche meno possibilità di avere una perfetta differenziazione tra le classi. Appartengono a queste anche la maggior parte dei casi di identificazioni mancate. Anche qui, la motivazione è probabilmente legata alle dimensioni delle bounding box. Ricordando quanto visto mentre si analizzavano pregi e difetti della rete, la YOLO fatica a riconoscere oggetti piccoli, soprattutto quando questi sono molto ravvicinati. Al contrario, gli elettrodomestici con funzionamento più lungo (forno, lavastoviglie, lavatrice, frigorifero e freezer) sono più facili da caratterizzare e infatti questi vengono identificati in maniera pressoché perfetta.

Si nota la presenza di due classi che non sono praticamente mai correttamente identificate: piano cottura e lavasciuga. Il primo appartiene alla categoria dei segnali brevi e i problemi legati alla sua corretta identificazione sono quelli già esplicitati in precedenza. Ad aggravare la situazione, c'è il basso numero di immagini presenti nel database. Diverso è invece il discorso legato alla lavasciuga. Oltre al numero limitatissimo di immagini, dovute ai pochi dati presenti nel database, va rimarcato come questo elettrodomestico vada a combinare le funzionalità di altri due presenti nel database, ossia lavatrice e asciugatrice, che hanno una presenza maggiore nel dataset. La conseguenza di ciò

è che i segnali della lavasciuga vengano identificati totalmente in maniera errata, solitamente ricondotti alla lavatrice. Date queste premesse, la lavasciuga non verrà considerata nella composizione dei segnali aggregati.

Se le identificazioni corrette rimanessero sopra il 60 – 70 % per la maggior parte delle categorie anche sugli aggregati, l'utilizzo di questa singola rete sarebbe sufficiente e si potrebbe andare solamente a migliorare ulteriormente il database e i parametri di training per alzarne il rendimento.

5.2.2 Analisi risultati detection su segnali aggregati reali

In queste prime analisi si sono analizzati i risultati sia sugli aggregati simulati che sugli aggregati derivati dalle rilevazioni da contatore. Come visto in precedenza, gli spettrogrammi ottenuti con l'ausilio del rilevatore di eventi risultano profondamente diversi rispetto a quelli ottenuti sui segnali simulati a frequenza di campionamento di 1 Hz. Nonostante vari tentativi, è risultato impossibile ottenere delle identificazioni su queste immagini. Per utilizzare questo tipo di rilevamento si sono valutate due opzioni. La prima, più semplice, quella di creare un nuovo database basato sui segnali degli elettrodomestici acquisiti con il rilevatore di eventi. Non avendo a disposizione dati di questo tipo né il tempo e il modo di ottenerne un numero sufficiente attraverso rilevazioni reali in laboratorio, è stata scartata. La seconda opzione, più complessa a livello operativo, è quella di trovare il modo di andare a “riempire” i momenti non rilevati attraverso un algoritmo. Essendo rilevate le variazioni sensibili di potenza, si può pensare di riempire gli istanti vuoti con un segnale costante o sinusoidale a bassa ampiezza. Per quanto questa possa essere una soluzione valida per tutti quei segnali e i cicli ON/OFF a potenza massima costante, non verrebbero considerati tutti quei transitori in cui il segnale non varia in maniera brusca. Essendoci il rischio di perdere troppe informazioni, si è preferito concentrarsi sui segnali aggregati simulati e sull'effettiva capacità della rete di identificare gli spettrogrammi dei singoli elettrodomestici, rimandando alle considerazioni finale tutte le valutazioni per l'implementazione del metodo per rilevazioni a frequenza di campionamento non costante.

5.2.3 Definizione del metodo e delle regole di identificazione

Prima di procedere con l'analisi vera e propria dei risultati ottenuti dalla rete è importante capire in che modo suddividere il segnale aggregato in intervalli di quattro ore, quando classificare lo spettrogramma come non identificabile (a causa di sovrapposizioni o perché troncato da uno degli estremi dell'intervallo temporale) e come operare quando un segnale appare in più intervalli ed è classificato identificabile in entrambi gli spettrogrammi.

Per prima cosa si deve decidere in che modo suddividere l'arco temporale di una giornata. La semplice suddivisione in sei intervalli di quattro ore renderebbe difficile andare a identificare tutti quei segnali che si trovano a cavallo di due intervalli differenti. Creando gli spettrogrammi del segnale aggregato ogni due ore, si riesce ad avere almeno una immagine in cui il segnale appare per intero o per la maggior parte della sua durata, in modo tale da renderlo classificabile in maniera corretta. Si può pensare di andare a diminuire l'intervallo temporale tra uno screenshot e l'altro, aumentando il numero di immagini detectabili su cui effettuare il riconoscimento. Ciò non viene fatto sin da subito e lo ritengo utile solo nel caso ci fossero molti segnali che vengono identificati in maniera diversa in due immagini consecutive.

Conseguenza di questa scelta, è il fatto che uno stesso segnale possa essere identificato all'interno di due o addirittura tre spettrogrammi. Nel momento in cui questo viene identificato una sola volta o viene identificato allo stesso modo in tutte le immagini, non sussiste alcun problema. Diversamente, invece, nel caso lo stesso segnale venisse identificato in maniera diversa su più immagini. In questo caso, la soluzione, banalmente, è quella di considerare solamente la detection con il punteggio di confidence maggiore. Si dovrà poi considerare quali soluzioni a livello operativo utilizzare per far effettuare questa scelta direttamente dal programma e poter poi ricostruire interamente la cronologia dei segnali durante la giornata.

Un altro passaggio fondamentale è quello di definire quando un segnale, all'interno di uno spettrogramma tratto dall'aggregato, è da ritenersi identificabile o meno, sia perché in presenza di una sovrapposizione (totale o parziale), sia per un posizionamento a cavallo di due intervalli. In quest'ultimo caso il segnale risulta “troncato” quando si va a generare lo spettrogramma. Dalle prime analisi sugli aggregati, si è notato come i segnali più corti non possono essere identificati se posti all'interno di un segnale di durata maggiore. Una completa sovrapposizione tra segnali di lunghezza

simile risulta più complessa da gestire. In generale, ne verrà identificato al massimo uno dei due, quello che risulta meno modificato o quello che mediamente viene rilevato con una confidence maggiore da parte della rete. In questo caso, si ritiene adeguata l'identificazione effettuata dalla rete. Se la sovrapposizione dei segnali è solo parziale (meno della metà del segnale più breve è "coperta" dall'altro elettrodomestico), entrambi i segnali possono essere identificati dalla rete. Questo perché, ricordando in che modo vengono assegnate le bounding box, se i centri delle due box si trovano in due celle diverse entrambi gli oggetti possono essere identificati dalla rete. Perché si verifichi questa condizione la sovrapposizione non può essere totale (da qui la "regola" di metà segnale non sovrapposto) e i segnali non devono essere troppo corti, altrimenti si troveranno con ogni probabilità totalmente all'interno di una singola cella e ovviamente i centri non potranno trovarsi in celle diverse. Nei rari casi in cui ci sono un segnale lungo e uno di media durata contemporaneamente attivi si possono identificare entrambi gli elettrodomestici, anche se quello di media durata è completamente inglobato da quello più lungo, purché si trovi all'inizio o alla fine della curva di durata maggiore.

Un problema collaterale alla presenza di segnali sovrapposti è quello relativo alla capacità della rete di identificare dei segnali nonostante l'intensità modificata. Un primo test, al di fuori degli aggregati, consiste nell'effettuare il riconoscimento su segnali noti presenti nel database modificati in intensità a varie percentuali. Questo test permette anche di avere un'idea del fatto che possano essere o meno identificati segnali simili a quelli del database ma di potenza diversa. Per fare questa verifica si è andati a utilizzare segnali dal database scelti a campione amplificati o attenuati fino al 50%. La verifica ha avuto esito positivo e pertanto non si considera un problema una differenza di potenza, purché non eccessiva, e quindi si ritiene che in buona parte dei casi di sovrapposizione almeno uno dei due segnali possa essere correttamente identificato.

5.2.4 Risultati detection sui segnali aggregati simulati

Gli aggregati simulati vengono costruiti inserendo casualmente i segnali di 15 elettrodomestici all'interno di un intervallo di una giornata, come spiegato nei capitoli precedenti. Si considerano identificabili i segnali che hanno le caratteristiche del paragrafo precedente. Per valutare la capacità della rete di effettuare le prediction si vanno a considerare sia il numero di spettrogrammi correttamente identificati che il numero di segnali. Ricordando che ciascun segnale appare almeno in due intervalli, ogni spettrogramma identificabile viene considerato come una immagine a sé stante. Si considera invece il segnale correttamente identificato se la prediction con il punteggio di confidence maggiore tra le immagini dello stesso segnale è corretta.

Sono stati simulati una decina segnali aggregati, utilizzati sia per confrontare vari detector addestrati con parametri differenti sia per ottenere delle indicazioni preliminari sul comportamento della rete. Il numero di dati resta ovviamente basso e verrà successivamente aumentato.

Come si può notare nella tabella alla pagina successiva, risulta evidente già con pochi casi come i segnali più lunghi siano riconosciuti molto meglio rispetto ai brevi. Analizzando il riconoscimento delle immagini, si vede come il forno e la lavatrice vengano praticamente sempre riconosciuti, mentre asciugatrice, lavastoviglie e piastra a induzione hanno circa il 50% di detection corrette sulle immagini. Tutti questi hanno un tempo di funzionamento medio superiore ai 1000 secondi. Anche per quanto riguarda il ferro da stiro, seppur con una percentuale di riconoscimenti inferiore, i segnali di durata più lunga vengono correttamente identificati. Tutti gli altri elettrodomestici con funzionamento breve hanno percentuali di immagini riconosciute ampiamente sotto il 30%. Per la maggior parte, questi segnali non vengono riconosciuti, mentre le identificazioni errate sono per la maggior parte segnali assegnati erroneamente al bollitore. Si è notato come i segnali di breve durata non vengano riconosciuti quando si trovano molto vicini ad altri segnali, siano essi lunghi o brevi. Si è anche verificato un caso di sovrapposizione pressoché perfetta tra i segnali di una lavatrice e di una lavastoviglie, con la rete che ha dato la priorità alla lavatrice, riconoscendola in entrambe le immagini in cui avveniva la sovrapposizione. Questo è giustificato dal fatto che la lavatrice viene riconosciuta più facilmente dalla rete, con una percentuale di immagini correttamente identificate che sfiora il 100%, e dal fatto che sommando i segnali la lavastoviglie non impone grandi modifiche alla forma della curva.

Gli unici due segnali che non vengono mai identificati sono l'asciugacapelli e il robot da cucina. Se quest'ultimo era già stato identificato come possibile punto critico, per l'asciugacapelli i problemi principali sono probabilmente la durata molto breve del segnale e la somiglianza con altri elettrodomestici, come visto precedentemente nell'analisi delle curve di consumo.

Il totale delle immagini correttamente identificate risulta basso (32%) e legato prevalentemente ai segnali lunghi, mentre più della metà delle immagini non vengono etichettate in alcun modo (54%). Una parte non indifferente (circa il 15%) delle immagini non riconosciute sono legate a segnali brevi ravvicinati e segnali leggermente sovrapposti.

Come detto nel capitolo iniziale, sono stati creati un piccolo numero di aggregati contenenti anche il segnale del frigorifero, per avere conferma della capacità della rete di identificarne il segnale essendo un elettrodomestico sempre presente nelle abitazioni. Avendo un funzionamento intermittente lungo tutta la giornata, il segnale sarà spesso contemporaneo a quello di altri elettrodomestici. Si è visto come i picchi vengano correttamente identificati nella maggior parte dei casi, se non quando si sovrappongono ad altri segnali, soprattutto quando questi sono quelli relativi ad elettrodomestici complessi e con funzionamenti lunghi come lavatrice o lavastoviglie.

IMMAGINI	%	Asciugacapelli	Asciugatrice	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Lavastoviglie	Lavatrice	Macchina del caffè	Micro onde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tosta pane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00				2												13	15	67.17	
Asciugatrice	50.00	7				1											6	14	3653.571	
Aspirapolvere	7.69			1													10	13	331.57	
Bollitore	27.27				3	1											7	11	119.29	
Ferro da stiro	29.41		2			5				1	1						8	17	1173.67	
Forno	84.62						11										2	13	2039.57	
Lavastoviglie	42.86		1					6	2								5	14	4097.25	
Lavatrice	93.33								14								1	15	5667.88	
Macchina del caffè	8.33									1							8	12	159.8	
Microonde	28.57										4						10	14	430.11	
Piano cottura	6.25											1					11	16	108.5	
Piastra induzione	50.00												8				8	16	1488	
Robot da cucina	0.00																5	7	184	
Tagliaerba	18.75										1						9	16	158.5	
Tostapane	20.00																10	15	176.5	

Tabella 5.2: matrice di confusione relativa alle immagini identificate dalla rete su spettrogrammi di 4 ore

SEGNALI	%	Asciuga capelli	Asciuga trice	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Lavastoviglie	Lavatrice	Macchina del caffè	Micro onde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tosta pane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciuga capelli	0.00				1													7	8	67,17
Asciugatrice	85.71		6			1													7	3653.571
Aspirapolvere	12.50			1														5	8	331.57
Bollitore	28.57				2	1												4	7	119.29
Ferro da stiro	44.44					4				1	1							2	9	1173.67
Forno	100.00						7												7	2039.57
Lavastoviglie	50.00							4	1									2	8	4097.25
Lavatrice	100.00								8										8	5667.88
Macchina del caffè	14.29									1								3	7	159.8
Microonde	33.33										3							6	9	430.11
Piano cottura	11.11											1						5	9	108.5
Piastra induzione	75.00												6					2	8	1488
Robot da cucina	0.00																	4	5	184
Tagliaerba	22.22										1							3	9	158.5
Tosta pane	37.50															2		4	8	176.5

Tabella 5.3: matrice di confusione relativa ai segnali identificati dalla rete su spettrogrammi di 4 ore

Dalla tabella 5.3 si nota, analizzando il numero di segnali identificati, che questi abbiano nel complesso percentuali di identificazione sensibilmente maggiori. In particolare, notiamo come anche la piastra a induzione e l'asciugatrice vengano identificate in almeno tre quarti dei casi. Questo aumento delle percentuali è dovuto al fatto che nelle due immagini in cui il segnale viene rappresentato spesso uno spettrogramma risulta essere più "pulito" dell'altro. I motivi di ciò sono la presenza di altri segnali vicini all'inizio o alla fine del segnale oppure un posizionamento vicino al bordo dell'immagine, che comporta una modifica dello spettrogramma dovuta al troncamento del segnale con la fine dell'intervallo o il suo inizio improvviso all'istante zero. A diminuire notevolmente, nella maggior parte dei casi, sono le detection mancate, mentre quelle errate rimangono pressoché invariate.

Il numero di segnali non identificati è del 40%, leggermente più alta invece quella delle detection corrette (42%). Andando ad analizzare l'impatto delle sovrapposizioni e dei segnali ravvicinati o troncati, si è notato che pesano per il 15% delle mancate o errate identificazioni. La prima possibilità per migliorare le prestazioni è quella di ampliare ulteriormente il database, in particolare per i segnali brevi. Questo comporterebbe però un notevole aumento del numero di immagini e quindi del tempo di addestramento della rete da parte del software. Prima di procedere su questa strada, si è preferito vagliare altre ipotesi meno dispendiose e, dati i limiti strutturali della rete su immagini di piccole dimensioni, presumibilmente più efficaci.

Prima di andare ad aumentare il numero di campioni su cui verificare l'efficacia della rete, date le difficoltà nell'identificazione dei segnali brevi, si è deciso di utilizzare una seconda rete, tarata su spettrogrammi della durata di un'ora, che vada a identificare solo i segnali medio-brevi in modo tale da migliorare la percentuale di detection corrette. L'utilizzo delle due reti apre diverse possibilità sul loro utilizzo combinato.

5.2.5 Seconda verifica su segnali aggregati e unione dei risultati

Dopo aver creato le reti per i segnali corti, si è continuata la verifica delle prestazioni di questa rete, anche per permettere di valutare le prestazioni delle reti combinate tra loro e le regole da seguire nella combinazione delle due. Per semplicità, andremo a valutare direttamente le prestazioni sui segnali, che sono di fatto lo scopo principale della rete.

Confrontando i dati con la serie precedente, si nota come siano quasi tutti in linea, soprattutto per quanto riguarda i segnali di lunga durata, mentre per i segnali brevi rimane il bollitore l'unico con un livello discreto di corrette identificazioni, anche per il fatto di essere quello con il numero maggiore di immagini presenti nel database. Migliora in modo sensibile l'identificazione del ferro da stiro, nonostante un numero non elevato di immagini nel database, il che è confortante, essendo un elettrodomestico al limite tra i lunghi e i corti, e il rischio che non venga identificato in modo sufficientemente preciso da nessuna delle due reti è concreto.

La lavastoviglie ha una percentuale più bassa di quanto preventivabile, anche se tre mancate detection sono dovute a sovrapposizioni con altri segnali, senza le quali la percentuale salirebbe quasi al 60 % (comunque non eccezionale).

Devices	%	Asciugacapelli	Asciugatrice	Aspirapolvere vere	Boillitore	Ferro da stiro	Forno	Lavastoviglie	Lavatrice	Macchina del caffè	Micro onde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tosta pane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00				8													14	22	70.16
Asciugatrice	84.62		22															4	26	3955.19
Aspirapolvere	15.38			4	5						4							13	26	371.45
Boillitore	43.48				10													13	23	119.65
Ferro da stiro	60.00		4			15												6	25	1325.23
Forno	100.00						25												25	2110.21
Lavastoviglie	50.00							12	3								2	7	24	4517.42
Lavatrice	100.00								26										26	4984.69
Macchina del caffè	0.00				4	4	3											15	22	578.07
Microonde	12.00				6	4	1				3							11	25	637.28
Piano cottura	3.85				5							1				1		19	26	92.03
Piastra induzione	42.31												11					15	26	1488
Robot da cucina	0.00													0				13	13	184
Tagliaerba	9.52				1						4							14	21	166.69
Tostapane	28.00				2						1					7		15	25	167.17

Tabella 5.4: matrice di confusione relativa ai segnali identificati dalla rete su spettrogrammi di 4 ore, seconda serie di dati

Devices	%	Asciuga capelli	Asciuga trice	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Lavasto viglie	Lavatrice	Macchina del caffè	Micro onde	Piano cottura	Piastra induzione	Robot da cucina	Taglia erba	Tosta pane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00				9													21	30	70.16
Asciugatrice	84.85		28			1												4	33	3955.19
Aspirapolvere	14.71			5	7						4							18	34	371.45
Bollitore	40.00				12	1												17	30	119.65
Ferro da stiro	55.88		5			19				1	1							8	34	1325.23
Forno	100.00						32												32	2110.21
Lavastoviglie	50.00		1					16	4								2	9	32	4517.42
Lavatrice	100.00								34										34	4984.69
Macchina del caffè	3.45		1		6	6	3			1								18	29	578.07
Microonde	17.65				6	4	1				6							17	34	637.28
Piano cottura	5.71				8							2				1		24	35	92.03
Piastra induzione	50.00												17					17	34	1488
Robot da cucina	0.00				1									0				17	18	184
Tagliaerba	13.33				4						5				4			17	30	166.69
Tostapane	30.30				3						1					10		19	33	167.17

Tabella 5.5: matrice di confusione relativa ai segnali identificati dalla rete su spettrogrammi di 4 ore, risultati combinati

5.2 Rete neurale per intervalli brevi (1 ora)

La rete viene creata secondo gli stessi principi che hanno guidato la creazione della rete per intervalli lunghi. Ovviamente dal database di addestramento vengono esclusi tutti gli elettrodomestici con funzionamento di durata superiore all'ora e si è cercato di aumentare il più possibile il numero di immagini relative ai segnali medio-brevi in modo da migliorare le prestazioni. Gli elettrodomestici presi in considerazione per l'addestramento della rete sono stati: asciugacapelli, aspirapolvere, bistecchiera, bollitore, cappuccinatore, ferro da stiro, forno, macchina del caffè, microonde, piano cottura, piastra a induzione, robot da cucina, tagliaerba, tostapane. Nei casi in cui erano presenti una o poche curve di consumo (ad esempio la piastra ad induzione) si è cercato di creare il maggior numero possibile di spettrogrammi andando a posizionare di volta in volta in momenti diversi dell'intervallo il segnale, in modo tale da evitare il più possibile la presenza di problemi di overfitting dovuti al posizionamento. Negli altri casi, ci si è limitati ad utilizzare al massimo un paio di volte lo stesso segnale, ma di distribuirli casualmente su tutto l'arco temporale. Il numero di spettrogrammi relativi a ferro da stiro e forno risulta essere basso per il semplice fatto che alcune curve superavano l'ora di durata mentre altre erano di poco inferiori, rendendo di fatto impossibile creare molte immagini tramite il riposizionamento del segnale. L'intero database si compone di 5627 immagini.

5.2.1 Analisi detection su training set

Dopo alcuni tentativi, sono risultate esserci due reti con risultati molto elevati su tutti gli elettrodomestici presenti nel database, che verranno rivalutate sui segnali aggregati, sia singolarmente che in combinazione con la rete neurale addestrata sugli intervalli di quattro ore. Il fatto che non possano essere identificati tutti gli elettrodomestici rende indispensabile l'utilizzo di entrambe le reti, che deve essere ovviamente regolamentato per andare a ottimizzarne l'apporto.

La prima rete, denominata "Detector 2", è stata addestrata utilizzando le seguenti impostazioni di training:

options = trainingOptions ("adam", "InitialLearnRate", 0.0001, "Plots", "training-progress", "Verbose", true, "MiniBatchSize", 32, "MaxEpochs", 8, "Shuffle", "every-epoch", "ExecutionEnvironment", "gpu", "GradientDecayFactor", 0.1, "LearnRateDropPeriod", 6)

Devices	TOT	% CORRECT	% MISSED	% WRONG
Asciugacapelli	431	95.36	3.25	1.39
Aspirapolvere	116	98.28	3.45	1.72
Bistecchiera	57	100.00	0.00	0.00
Bollitore	1411	99.79	0.00	0.21
Cappuccinatore	66	100.00	0.00	0.00
Ferro da stiro	66	100.00	0.00	0.00
Forno	60	96.67	3.33	0.00
Macchina del caffè	831	97.59	2.05	1.32
Microonde	1454	91.68	5.78	2.54
Piano cottura	125	72.80	17.60	9.60
Piastra induzione	45	100.00	0.00	0.00
Robot da cucina	102	100.00	0.00	0.00
Tagliaerba	120	99.17	0.83	0.00
Tostapane	743	98.38	1.62	0.00

Tabella 5.6: tabella riassuntiva dei risultati sul test set del "detector 2"

La seconda rete, denominata "Detector 4", invece, è stata addestrata con le seguenti opzioni:

options = trainingOptions ("adam", "InitialLearnRate", 0.0001, "Plots", "training-progress", "Verbose", true, "MiniBatchSize", 32, "MaxEpochs", 10, "Shuffle", "every-epoch", "ExecutionEnvironment", "gpu", "GradientDecayFactor", 0.1, "LearnRateDropPeriod", 8)

Devices	TOT	% CORRECT	% MISSED	% WRONG
Asciugacapelli	431	97.91	2.09	0.00
Aspirapolvere	116	100.00	0.00	0.00
Bistecchiera	57	100.00	0.00	0.00
Bollitore	1411	99.57	0.21	0.21
Cappuccinatore	66	100.00	0.00	0.00
Ferro da stiro	66	100.00	0.00	0.00
Forno	60	100.00	0.00	0.00
Macchina del caffè	831	97.23	1.56	1.20
Microonde	1454	97.94	1.93	0.34
Piano cottura	125	86.40	5.60	8.00
Piastra induzione	45	100.00	0.00	0.00
Robot da cucina	102	100.00	0.00	0.00
Tagliaerba	120	100.00	0.00	0.00
Tostapane	743	98.92	0.40	0.67

Tabella 5.7: tabella riassuntiva dei risultati sul test set del "detector 4"

Essenzialmente, l'unica differenza tra le due consiste nel numero di epoche di training. La seconda va a migliorare ulteriormente dei risultati già ottimi sul training set, ma questo non comporta necessariamente una migliore prestazione su immagini diverse da quelle dell'addestramento. Questi dati possono essere sintomo di una tendenza all'overfitting o semplicemente non portare ad un miglioramento nell'identificazione di immagini nuove, non giustificando le maggiori risorse di computazione spese per l'addestramento. Inoltre, come già detto in precedenza, nel nostro caso, oltre alla prestazione in sé della rete, va valutata la prestazione della stessa in combinazione con la rete addestrata sugli intervalli lunghi, data la necessità di utilizzarle entrambe per riuscire a ottenere prestazioni buone su tutti i tipi di segnali.

5.2.2 Analisi detection su segnali aggregati simulati

Per valutare le prestazioni delle reti, si è andati a simulare il segnale aggregato di 30 giorni, creandoli nello stesso identico modo in cui sono stati creati quelli per valutare la rete per i segnali lunghi. Anche i principi per ritenere identificabile o meno i vari segnali, così come la distinzione tra numero di immagini identificate e numero di segnali, non sono variati.

DET2	% correct	Asciugacapelli	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Macchina del caffè	Microonde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tostapane	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00	0		2			1	1	6				17	45	70.16
Aspirapolvere	45.00		18				1	1	5				5	40	371.45
Bollitore	73.08			38										52	119.65
Ferro da stiro	23.73			1	14	2	19	1						59	1325.23
Forno	58.62			1		34	2	6						58	2110.21
Macchina del caffè	42.37			6			25	4						59	578.07
Microonde	24.53				5		6	13					5	53	637.28
Piano cottura	33.33			1					19				8	57	92.03
Piastra induzione	56.86				1		9			29				51	1488
Robot da cucina	0.00								7	1	0		5	23	184
Tagliaerba	2.22								3			1	6	45	166.69
Tostapane	66.04				1								35	53	167.17

Tabella 5.8: matrice di confusione delle immagini identificate dal detector 2

DET4	% correct	Asciugacapelli	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Macchina del caffè	Microonde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tostapane	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00	0		2				4	16				15	45	70.16
Aspirapolvere	27.50		11				1	3	6					40	371.45
Bollitore	84.62			44										52	119.65
Ferro da stiro	27.12				16	2	15	1						59	1325.2
Forno	82.76					48	2	1						58	2110.2
Macchina del caffè	49.15			6		1	29	9						59	578.07
Microonde	26.42				4	3	8	14					1	53	637.28
Piano cottura	47.37			1				6	27					57	92.03
Piastra induzione	74.51							3		38				51	1488
Robot da cucina	0.00								10		0		1	23	184
Tagliaerba	2.22							4	3			1	4	45	166.69
Tostapane	56.60			4		1		1					30	53	167.17

Tabella 5.9: matrice di confusione delle immagini identificate dal detector 4

Il detector 2 presenta percentuali sensibilmente migliori per quanto riguarda aspirapolvere e tostapane, mentre il detector 4 risulta essere molto più efficace nell'identificazione degli spettrogrammi di bollitore, forno, piano cottura e piastra a induzione. La percentuale totale risulta ad appannaggio del secondo. Rispetto alle detection su intervalli più lunghi, aumenta il numero di identificazioni errate. I tipi di errore sono abbastanza in linea con quelli che ci si aspettava in sede di analisi delle curve di consumo, con segnali intermittenti confusi tra loro e segnali a gradino a loro volta confusi tra loro a seconda soprattutto della durata. Rimangono critici l'asciugacapelli e il robot da cucina, che oltre ai problemi legati alle similitudini con altri segnali hanno a loro sfavore lo scarso numero di immagini a loro attribuite nel database, a causa della scarsa quantità di dati a disposizione. In una situazione molto simile si trova il tagliaerba.

In sede di raccolta dati si è proceduto anche su una analisi del tipo di errore commesso dalle reti, cercando di trovare un filo rosso che unisse le immagini male identificate o del tutto ignorate dalla rete.

Si è evidenziato come le difficoltà su ferro da stiro e macchina del caffè siano principalmente legate ai segnali più brevi, che per la seconda significa un tempo inferiore ai 100 s. Per il primo ci sono molte immagini identificate come macchina del caffè, evidentemente legate ad un funzionamento intermittente comune, ma anche ai pochi dati presenti nel database. Le considerazioni da fare sulla macchina da caffè sono due. La prima riguarda i segnali molto brevi, sotto il minuto, che vengono ricondotti al bollitore. Come già detto, segnali molto brevi, a gradino, sono per la rete pressoché uguali e per questo è facile che vengano assegnati ad una sola categoria, che ha priorità per la rete. Per segnali un po' più lunghi, invece, entra in gioco il microonde. I due elettrodomestici possono avere entrambi una curva intermittente, che per segnali inferiori ai 300 s viene prevalentemente ricondotta al secondo. Viceversa, i segnali appena superiori verranno invece ricondotti alla macchina del caffè. Ciò risulta evidente nei risultati del detector 4 che vedono 9 immagini di macchina del caffè identificati come microonde e 8 immagini di microonde identificate come macchina del caffè. Curve di consumo del microonde di durata ancora maggiore vengono invece più facilmente confuse con altri segnali intermittenti, quali ferro da stiro e forno. Per quanto riguarda gli altri segnali corti a gradino, si nota come vengano identificati molto spesso come bollitore, piano cottura e tostapane, che sono tre elettrodomestici con un numero elevato di immagini nel database e con le percentuali di riconoscimenti più elevate.

immagini DET2	Asciuga capelli	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Macchina del caffè	Microonde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tostapane	TOT	DURATA MEDIA [s]
Asciugacapelli	0		2			1	1	6				17	45	70.16
Aspirapolvere		18				1	1	5				5	40	371.45
Bollitore			38										52	119.65
Ferro da stiro			1	14	2	19	1						59	1325.23
Forno					35	2	6						58	2110.21
Macchina del caffè			6			25	4						59	578.07
Microonde				5		6	13					5	53	637.28
Piano cottura			1					19				8	57	92.03
Piastra induzione				1		9			29				51	1488
Robot da cucina								7	1	0		5	23	184
Tagliaerba								3			1	6	45	166.69
Tostapane				1								35	53	167.17

Tabella 5.10: matrice di confusione relativa alle immagini identificate dal detector 2

immagini DET4	Asciuga capelli	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Macchina del caffè	Microonde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tostapane	TOT	DURATA MEDIA [s]
Asciugacapelli	0		2				4	16				15	45	70.16
Aspirapolvere		11				1	3	6					40	371.45
Bollitore			44										52	119.65
Ferro da stiro				16	2	15	1						59	1325.23
Forno					48	2	1						58	2110.21
Macchina del caffè			6		1	29	9						59	578.07
Microonde				4	3	8	14					1	53	637.28
Piano cottura			1				6	27					57	92.03
Piastra induzione							3		38				51	1488
Robot da cucina								10		0		1	23	184
Tagliaerba							4	3			1	4	45	166.69
Tostapane			4		1		1					30	53	167.17

Tabella 5.11: matrice di confusione relativa alle immagini identificate dal detector 4

Una valutazione ancora più precisa può essere fatta dopo aver analizzato l'identificazione dei segnali.

DET2	% correct	Asciugacapelli	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Macchina del caffè	Microonde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tostapane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00			1				1	4				12		7	25	70.16
Aspirapolvere	65.38		17				1	1	5						2	26	371.45
Bollitore	100.00			26												26	119.65
Ferro da stiro	37.04			1	10	2	7	1							6	27	1325.23
Forno	89.29			1		25	1	1								28	2110.21
Macchina del caffè	62.96			6			17	3							1	27	578.07
Microonde	22.22				5		2	6					5		9	27	637.28
Piano cottura	55.17			1					16				4		8	29	92.03
Piastra induzione	96.15						1			25						26	1488
Robot da cucina	0.00								7	1	0				4	12	184
Tagliaerba	4.17								3			1	6		14	24	166.69
Tostapane	78.57												22		6	28	167.17

Tabella 5.12: matrice di confusione relativa ai segnali identificati dal detector 2

DET4	%	Asciugacapelli	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Macchina del caffè	Microonde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tostapane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00	0		1				2	8				12		2	25	70.16
Aspirapolvere	30.77		8					6	6						6	26	371.45
Bollitore	92.59			25			1							1		27	119.65
Ferro da stiro	51.85				14	2	7	1						1	2	27	1325.23
Forno	96.43					27	1									28	2110.21
Macchina del caffè	55.17			6			16	5							2	29	578.07
Microonde	31.03				5	2	6	9							7	29	637.28
Piano cottura	65.52			1				6	19						3	29	92.03
Piastra induzione	96.15							1		25						26	1488
Robot da cucina	0.00								8		0		1		3	12	184
Tagliaerba	8.33							2	4			2	2	2	12	24	166.69
Tostapane	93.10					1							27		1	29	167.17

Tabella 5.13: matrice di confusione relativa ai segnali identificati dal detector 4

Si evidenzia subito come la percentuale di segnali correttamente identificati sia molto superiore rispetto a quella delle immagini, con diversi casi di elettrodomestici identificati correttamente con percentuali vicine o superiori al 90%. Il numero totale di segnali corretti tra le due reti è pressoché identico: il detector 4 presenta picchi più elevati nelle singole classi, mentre il 2 distribuisce maggiormente le detection corrette, portando anche aspirapolvere e macchina del caffè oltre il 60%. Esprimere una netta preferenza non è ancora possibile, data le piccole differenze tra i due, ed è preferibile rimandare la scelta dopo aver valutato come le due reti interagiscono con la rete per intervalli lunghi. Prima di fare ciò, è necessario andare a definire le regole di priorità sulle detection tra le reti.

5.3 Analisi detection reti combinate

Il modo più semplice ed efficace per unire i due tipi di rete è quello di assegnare la priorità nella detection di un certo elettrodomestico alla rete che meglio lo identifica all'interno dei segnali aggregati. La seconda rete entra in gioco solamente se la prima ignora completamente il segnale. Questo fa sì che aumenti il numero di segnali identificati, ma anche di errori, a discapito delle mancate identificazioni. La percentuale di segnali correttamente identificati relativa a ciascun elettrodomestico sarà sempre almeno uguale alla percentuale più alta fatta registrare tra le due reti.

Si deve mettere in preventivo che assegnando una priorità assoluta, senza considerare di volta in volta il punteggio di confidence, si possa incappare in qualche detection errata evitabile in favore di una più semplice operatività a livello di programmazione.

I segnali che ottengono la priorità sulla rete dei lunghi sono, oltre a lavastoviglie, lavatrice e asciugatrice, il ferro da stiro e il forno. Per tutti gli altri la priorità nella detection spetta alla rete addestrata sugli spettrogrammi di un'ora. Questo vale sia per il detector 2 che per il detector 4.

Analizzando i risultati contenuti nelle tabelle 5.12 e 5.13 si nota come il detector 2 abbia una migliore interazione con la rete per segnali lunghi e abbia dei significativi miglioramenti nelle percentuali, limando le differenze nelle classi in cui c'erano i deficit rispetto al detector 4 (in particolare ferro da stiro e piastra a induzione) e rimanendo più preciso dove già registrava prestazioni migliori (come per l'aspirapolvere).

Andando nel dettaglio, per quanto riguarda il detector 2, raddoppiano le detection del ferro da stiro, che arrivano al 66,67% e "guadagna" altre 6 identificazioni corrette distribuite tra microonde, tagliaerba e tostapane, che si porta così all' 80% di detection corrette. Il numero di identificazioni errate dovute all'intervento della rete secondaria è di poche unità. Non si sono registrate invece detection corrette della rete secondaria vanificate da errori della rete primaria.

Al contrario, il detector 4 conta cinque identificazioni dalla rete secondaria ma anche ben sette detection corrette bloccate da errori della rete primaria. Il numero più elevato di errati riconoscimenti rispetto ai segnali ignorati fa sì che ci sia una minore integrazione tra le due reti.

COMB 4+11	%	Asciugacapelli	Asciugatrice	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Lavasto viglie	Lavatrice	Macchina del caffè	Micro onde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tosta pane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00				1						2	8				12		2	25	70.16
Asciugatrice	84.62		22															4	26	3955.19
Aspirapolvere	40.00			12							6	6						6	30	371.45
Bollitore	92.59				25					1							1		27	119.65
Ferro da stiro	62.07		4			18	2			4	1								29	1325.23
Forno	100.00						29												29	2110.21
Lavastoviglie	50.00							12	3								2	7	24	4517.42
Lavatrice	100.00								26										26	4984.69
Macchina del caffè	57.14				6					16	4							2	28	578.07
Microonde	31.03				4	5	2			6	9							3	29	637.28
Piano cottura	65.52				1						6	19						3	29	92.03
Piastra induzione	96.15										1		25						26	1488
Robot da cucina	0.00											8				1		4	13	184
Tagliaerba	8.33											4			2	2		8	24	166.69
Tostapane	93.10						1											1	29	167.17

Tabella 5.13: matrice di confusione relativa ai segnali identificati dall'utilizzo combinato delle due reti (detector 4 + detector 11 per segnali lunghi)

COMB 2+11	%	Asciugacapelli	Asciugatrice	Aspirapolvere	Bollitore	Ferro da stiro	Forno	Lavasto viglie	Lavatrice	Macchina del caffè	Micro onde	Piano cottura	Piastra induzione	Robot da cucina	Tagliaerba	Tosta pane	ALTRO	MISSED	TOT	DURATA MEDIA [s]
Asciugacapelli	0.00				3						1	1				16		3	24	70.16
Asciugatrice	84.62		22															4	26	3955.19
Aspirapolvere	65.38			17						1	1	5						2	26	371.45
Bollitore	100.00				26														26	119.65
Ferro da stiro	66.67		4			20	2			1	1							2	30	1325.23
Forno	100.00						29												29	2110.21
Lavastoviglie	50.00							12	3								2	7	24	4517.42
Lavatrice	100.00								26										26	4984.69
Macchina del caffè	62.96				6					17	3							1	27	578.07
Microonde	27.59				4	5				3	8					4		5	29	637.28
Piano cottura	56.67				1							17			1	5		6	30	92.03
Piastra induzione	96.15									1			25						26	1488
Robot da cucina	0.00											6	1					5	12	184
Tagliaerba	13.04											3			3	5		12	23	166.69
Tostapane	79.31															23		6	29	167.17

Tabella 5.14: matrice di confusione relativa ai segnali identificati dall'utilizzo combinato delle due reti (detector 2 + detector 11 per segnali lunghi)

Riassumendo, abbiamo sei elettrodomestici che sfiorano o superano l'80% di identificazioni corrette, dei quali tre hanno il 100% (bollitore, forno, lavatrice) e uno il 96% (piastra a induzione). Questi quattro elettrodomestici hanno tutti almeno una di queste caratteristiche:

- Numero molto elevato di immagini presenti nel database
- Curva di consumo caratteristica per forma e/o durata

Vi sono poi tre elettrodomestici che la rete non riesce a identificare, che sono asciugacapelli, robot da cucina e tagliaerba. Tutti e tre sono caratterizzati da un numero limitato di segnali presenti nei database da noi considerati e di conseguenza da un numero basso di immagini presenti nel database degli spettrogrammi. Inoltre, per tipologia e durata del segnale, rientrano nel range monopolizzato da bollitore e tostapane, che infatti hanno le percentuali di riconoscimenti più elevate tra i segnali brevi.

Le restanti classi vengono identificate in percentuali comprese tra il 50 e il 70 %, che non sono disastrose ma vanno aumentate perché la rete possa essere considerata affidabile anche nel loro riconoscimento. Per molti di questi segnali sarebbe importante aumentare ulteriormente il numero di immagini disponibili nel database. Per gli elettrodomestici caratterizzati dalla possibilità di utilizzo con tempi e programmi differenti (ferro da stiro, aspirapolvere, microonde, macchina del caffè, lavastoviglie) è importante riuscire ad includere una varietà elevata di curve di consumo e di conseguenza di spettrogrammi, in modo tale da consentire una migliore capacità di riconoscimento su qualsiasi tipo di curva e limitare la polarizzazione verso quegli elettrodomestici con una sola curva come tostapane e bollitore, cui vanno attribuite buona parte delle detection sbagliate.

CAPITOLO 6

ANALISI RISULTATI E CONFRONTO CON RETE ZORZI

6.1 Analisi risultati

Grazie all'utilizzo delle due reti combinate, vengono identificati 245 segnali su 387 (il 63,3 %), conteggiando anche i tre casi di segnali di lavastoviglie sovrapposti alla lavatrice. Il risultato è discreto, considerando l'obiettivo iniziale del 70% di detection corrette e le problematiche legate ai segnali brevi, sia relative alle similitudini che al numero di curve di consumo a disposizione. Nell'insieme dei segnali non riconosciuti o identificati in modo errato sono compresi anche tutti quelli relativi ai tre segnali praticamente mai riconosciuti: asciugacapelli, robot da cucina e tagliaerba. Questi elettrodomestici, come visto nei capitoli precedenti, uniscono difficoltà nel riconoscimento dovute a somiglianze con altri segnali, una durata media di funzionamento molto breve e soprattutto un numero limitato di curve disponibili (anche se per l'asciugacapelli si è comunque riusciti ad ottenere un buon numero di spettrogrammi su cui addestrare la rete). Escludendo questi e le tre sovrapposizioni di cui sopra dalle identificazioni, la percentuale aumenta sensibilmente, fino al 74,5% (242 su 325).

Analizzando il comportamento della rete sui vari elettrodomestici, si nota che per segnali lunghi e caratteristici la rete neurale funziona in maniera pressoché perfetta. Per i segnali brevi, spesso simili tra loro, ci sono difficoltà legate sia alla dimensione nell'immagine dello spettrogramma sia alla necessità di avere un numero elevato di segnali per ciascuno degli elettrodomestici e, se questi hanno curve di consumo diverse a seconda dei modelli o del tipo di utilizzo, la stessa cosa deve valere anche per ciascun tipo di curva. In questo ultimo caso, la rete, anche avendo un buon numero di dati relativi alla classe, non riesce o fatica a trovare delle feature caratteristiche da attribuire alla stessa per identificarla in modo univoco. Di conseguenza, i componenti di queste classi si troveranno ad avere delle feature che li inseriscono in quelle che possiamo definire zone di conflitto o intersezione tra classi diverse, in cui confluiscono elettrodomestici differenti con spettrogrammi simili. Se in questa zona c'è una netta prevalenza di una certa classe, i nuovi segnali che appartenenti a questa zona verranno logicamente attribuiti alla classe predominante.

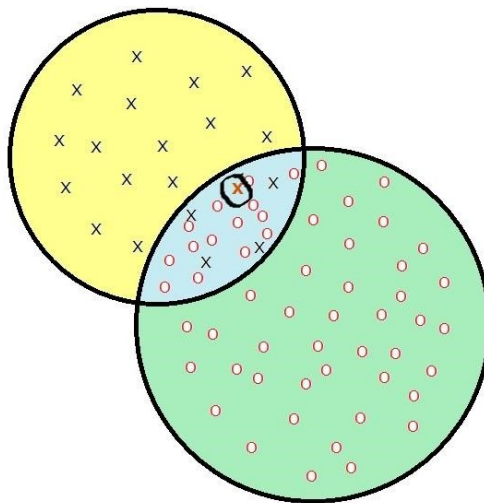


Figura 6.1: rappresentazione insiemistica di due classi. L'elemento cerchiato, non facente parte del database, appartiene all'intersezione tra i due insiemi. Essendo questa zona popolata perlopiù da elementi "O" piuttosto che elementi "X", facilmente il detector lo identificherà erroneamente come "O"

Un'altra misura indicativa sulle prestazioni della rete è il cosiddetto "F-Score" e i suoi parametri di precision e recall. Nell'ambito dell'Information Retrieval si definisce precision la percentuale di documenti attinenti recuperati da una ricerca sul totale dei documenti recuperati, ossia la percentuale dei segnali correttamente identificati come appartenenti ad una classe (veri positivi) sul totale di tutti i segnali identificati come appartenenti a quella stessa classe (veri positivi + falsi positivi). Per recall si intende invece la percentuale di documenti attinenti recuperati sul totale dei documenti attinenti, quindi la percentuale di segnali correttamente identificata come appartenente ad una classe sul totale di segnali appartenenti alla classe stessa (veri positivi + falsi negativi). Il parametro "F-Score", invece, è la media armonica di precision e recall.^[29] Tanto più vicino è all'unità e tanto migliore sarà la prestazione complessiva della rete.

$$precision = \frac{true\ positive}{true\ positive + false\ positive} \quad (6.1)$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative} \quad (6.2)$$

$$F1 = \frac{2}{precision^{-1} + recall^{-1}} \quad (6.3)$$

	True positive	False positive	False negative	Precision	Recall	F1 Score
Asciugacapelli	0	0	24	#DIV/0!	0	#DIV/0!
Asciugatrice	22	4	4	0.85	0.85	0.85
Aspirapolvere	17	0	9	1.00	0.65	0.79
Bollitore	26	14	0	0.65	1.00	0.79
Ferro da stiro	20	5	10	0.80	0.67	0.73
Forno	29	2	0	0.94	1.00	0.97
Lavastoviglie	12	0	12	1.00	0.50	0.67
Lavatrice	26	3	0	0.90	1.00	0.95
Macchina del caffè	17	6	10	0.74	0.63	0.68
Microonde	8	5	21	0.62	0.28	0.38
Piano cottura	17	15	13	0.53	0.57	0.55
Piastra induzione	25	1	1	0.96	0.96	0.96
Robot da cucina	0	0	12	#DIV/0!	0	#DIV/0!
Tagliaerba	3	1	20	0.75	0.13	0.22
Tostapane	23	30	6	0.43	0.79	0.56

Tabella 6.1: Parametri di precision, recall e F1 score relative alle detection delle due reti combinati sui segnali aggregati simulati

Come si nota dalla tabella, i segnali con i migliori risultati in termini di F1-Score sono asciugatrice, forno, lavatrice e piastra a induzione, che erano anche i migliori in termini di percentuale di identificazioni corrette. Le due categorie che risultano penalizzate in termini di precisione sono soprattutto bollitore e tostapane, che, come abbiamo visto in precedenza, sono i due segnali principali tra i corti e polarizzano le detection, ottenendo quindi un numero elevato di falsi positivi. Macchina del caffè e microonde hanno dati che si assomigliano, dovuti in buona parte al fatto che vengono confusi tra loro a seconda del tipo di segnale. In generale, i dati di F1-Score sono

abbastanza buoni, escluse le categorie critiche, e in linea con quanto si poteva immaginare dall'analisi delle matrici di confusione e da quella sugli errori fatta al capitolo precedente.

Oltre alla capacità di identificare correttamente la classe di appartenenza dei segnali, la rete deve essere valutata anche per la sua capacità di posizionare correttamente le relative bounding box. La corretta identificazione a livello geografico all'interno dell'immagine è l'unico modo per identificare i momenti di inizio e fine del segnale. La precisione assoluta non può essere garantita dalla rete e la risoluzione dell'asse delle ascisse impedisce in ogni caso il collocamento temporale preciso al secondo degli istanti di attivazione e spegnimento. La cosa importante è però che le bounding box vadano a riquadrare in maniera abbastanza precisa gli spettrogrammi, in modo tale da garantire comunque la possibilità di riportare in maniera abbastanza fedele gli istanti di inizio e fine del segnale, anche per eventuali operazioni post-identificazione. Questa valutazione non può essere fatta in maniera puramente numerica come per le identificazioni, ma devono essere valutati qualitativamente il posizionamento delle box e quanto la loro aderenza ai bordi dello spettrogramma. Nella valutazione viene dato molto più peso alla corretta delimitazione degli spettrogrammi in ampiezza, quindi sull'asse del tempo, che in altezza, quindi sull'asse relativo delle frequenze, che ha molta meno importanza per tutte le possibili operazioni successive, dato che è possibile ricondursi direttamente al segnale rilevato originale.

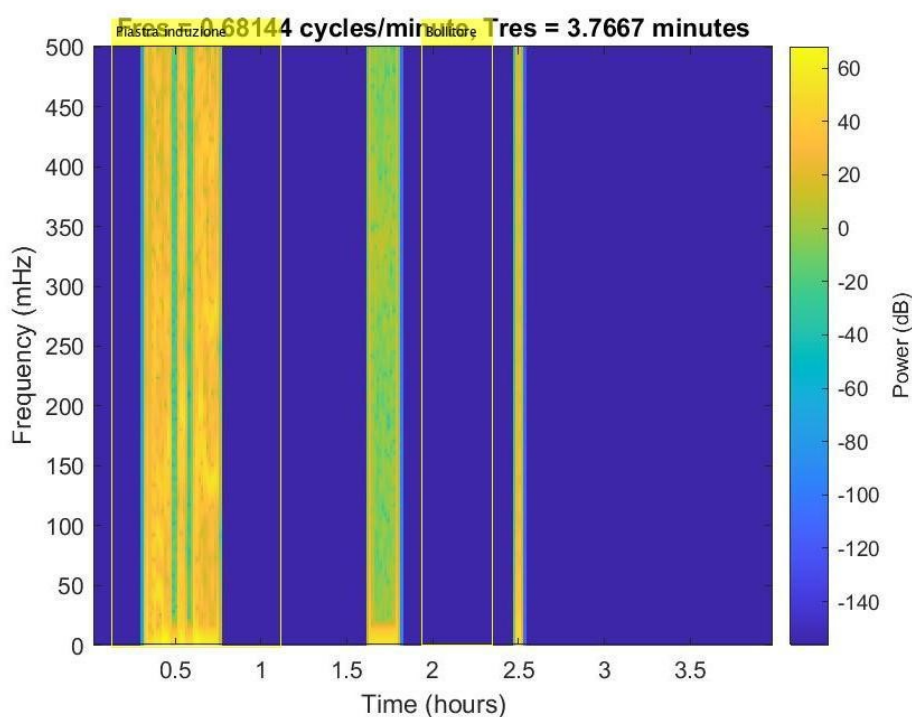


Figura 6.2: immagine con all'interno una bounding box ben posizionata e una con posizionamento completamente errato

Nell'esempio della figura 6.2 si possono vedere due bounding box. La detection della piastra a induzione è corretta e anche il posizionamento della stessa è abbastanza buono. La box è abbondante, ma include interamente il segnale e non sfiora nemmeno altri segnali. In post-elaborazione questo è buono, perché, se voglio calcolare i consumi andando ad integrare il segnale, gli istanti a potenza nulla non hanno alcuna influenza, mentre se voglio avere precisione nell'identificazione di inizio e fine segnale posso andare a considerare valore di potenza e della sua derivata temporale. Non essendoci altri spettrogrammi, non è complicato isolare il segnale. La seconda box, invece, è errata sia per la detection, in quanto si trova tra il segnale di un piano cottura e uno di un'aspirapolvere, sia per il posizionamento. La box non include nessuno dei due segnali, nemmeno parzialmente. Pertanto, anche ricondurla a uno dei due risulta impossibile. Una box di

questo tipo, anche nel caso identificasse una delle due classi in maniera esatta, non può essere considerata corretta.

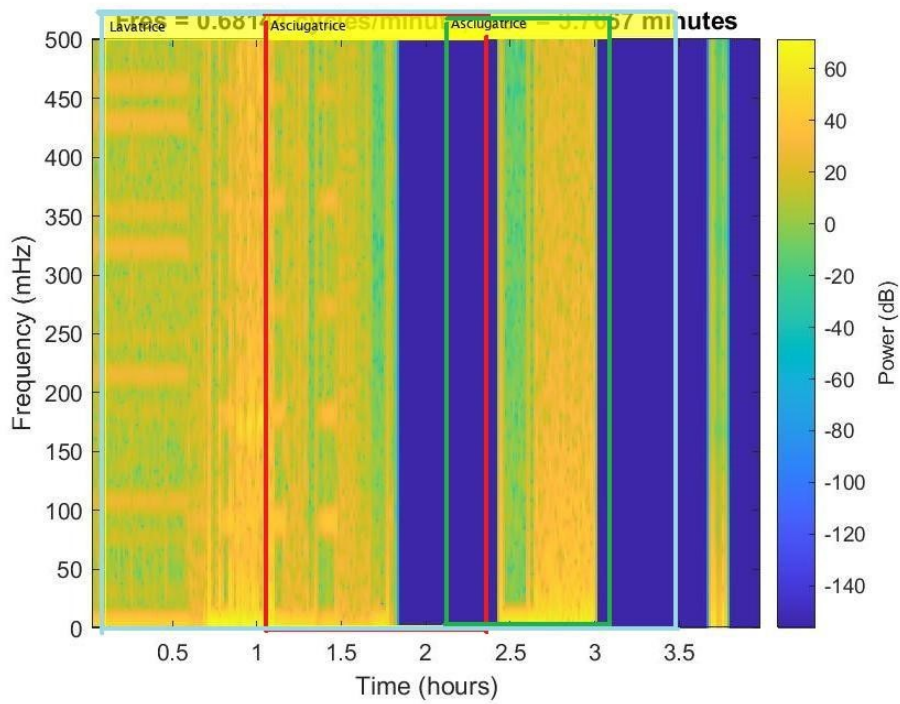


Figura 6.3: detection effettuate dalla rete su uno spettrogramma complesso con la presenza di 3 segnali distinti con una sovrapposizione

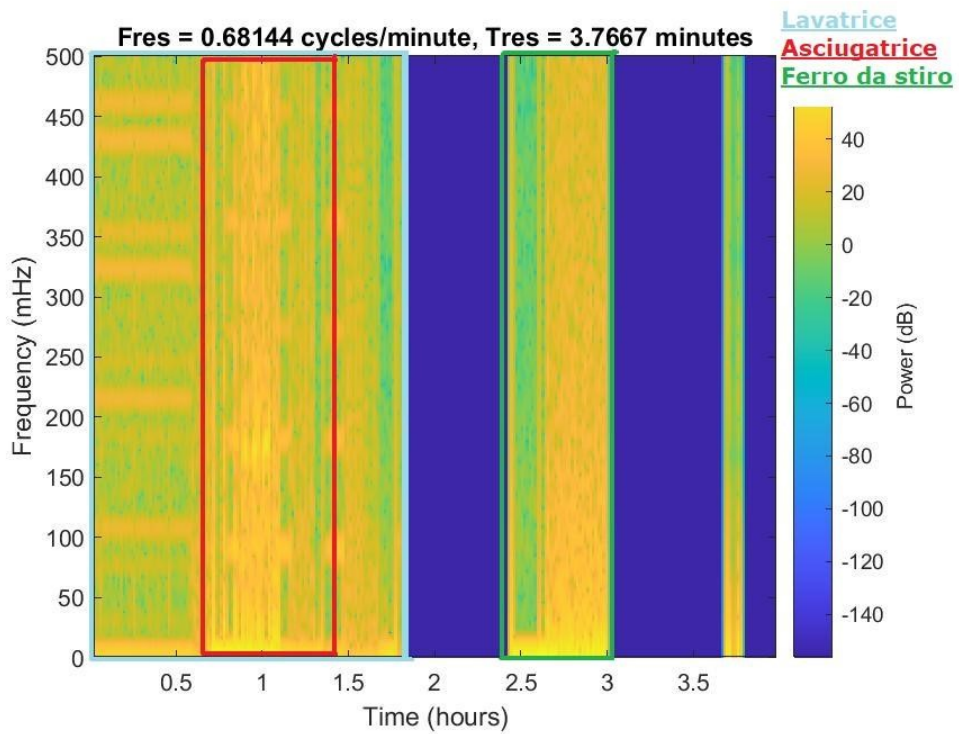


Figura 6.4: Bounding box corrette dei tre segnali presenti nell'immagine precedente

Le immagini precedenti si riferiscono a un caso tra i più complessi tra quelli che possono essere sottoposti ad una rete. Lo spettrogramma è stato analizzato dalla rete per segnali lunghi e non andremo ad analizzare il comportamento dell'altra rete sul segnale del ferro da stiro, in quanto su tutti i segnali è la rete dei lunghi ad avere la priorità nell'identificazione. Per prima cosa, possiamo notare come il segnale del ferro da stiro venga correttamente delimitato ma l'identificazione sia errata. Il ferro da stiro è di per sé un elettrodomestico soggetto ad errori di identificazione. Questo caso è anche un po' sfortunato in quanto fa parte di quel 33% di segnali non correttamente classificati e rappresenta la prima difficoltà a cui viene sottoposta la nostra rete. La seconda, ben più gravosa, è la presenza di due segnali sovrapposti tra loro. Come abbiamo già visto, se uno dei due fosse di durata molto breve, le possibilità di identificarlo sarebbero nulle. In questo caso, però, c'è il segnale di una asciugatrice, per quanto presente in uno dei suoi programmi più brevi (durata 40 minuti). Il segnale della lavatrice invece risulta troncato in testa. La parte mancante di segnale è di circa un terzo della durata totale. Lo spettrogramma viene interamente contenuto nell'intervallo precedente, rappresentato in un'altra immagine, in cui viene riconosciuto perfettamente il segnale della lavatrice. In questa situazione particolare, la rete riesce a identificare la presenza di due elettrodomestici differenti, in quanto i centri delle box si trovano in celle diverse. Il posizionamento delle box non è invece così preciso, in particolare quello della asciugatrice in realtà comprende solamente la seconda metà del segnale di quell'elettrodomestico, andando invece a inglobare l'ultima parte del segnale della lavatrice. Il posizionamento può comunque essere considerato accettabile, così come la detection. La box della lavatrice è invece meno precisa per la contemporanea presenza del troncamento del segnale, di una sovrapposizione e la presenza di un terzo segnale non troppo lontano. Tutti questi fattori fanno sì che di fatto la box non possa essere considerata corretta. Il fatto di aver generato gli spettrogrammi ogni due ore, fa sì che lo stesso segnale sia presente anche in una seconda immagine, in cui appare per intero e viene creata una bounding box ben posizionata e correttamente assegnata alla lavatrice. Avendo quest'ultima un punteggio di confidence maggiore rispetto a quella dell'immagine 6.3, avrà anche la priorità. Di conseguenza, la rete riesce a identificare correttamente due dei tre segnali lunghi, ma soprattutto riesce a identificare correttamente i due segnali sovrapposti, che rappresentano il più grande problema per qualsiasi sistema di riconoscimento dei carichi.

Questo genere di sovrapposizione tra segnali lunghi è avvenuto in altri sei casi, metà dei quali sono stati risolti correttamente. Nell'altra metà non è stato identificato nessuno dei segnali coinvolti. Questo perché la sovrapposizione può finire col creare un unico segnale che per lunghezza, distribuzione e valori di potenza assorbita o la combinazione di questi fattori, risulta essere completamente diverso da quelli generati da qualsiasi elettrodomestico presente nel database. Oltre che dalla posizione reciproca dei segnali, la capacità della rete di identificarli correttamente, in questi casi, dipende anche dal tipo di segnali che si sovrappongono. Se sono segnali simili c'è il rischio che questi vengano scambiati dalla rete come un unico segnale di una classe, in quanto la curva potrebbe risultare coerente con quelle già presenti nel database, seppur differente in termini assoluti di durata temporale e potenza assorbita. Nel caso fossero segnali molto diversi, come detto, c'è il rischio di ottenere un segnale talmente differente da quelli del database e quindi di non poter essere identificato correttamente.

Risulta essere oltremodo importante non limitarsi alla generazione degli spettrogrammi ogni una o quattro ore, ma generarne un numero maggiore, a costo di analizzare più volte gli stessi segnali, per avere più "fotografie" possibili del segnale aggregato e aiutare la rete a risolvere questi casi particolarmente critici. Questo "spostamento" dei segnali in punti diversi degli intervalli analizzati aiuta anche nei casi di sovrapposizioni in cui sono coinvolti segnali più corti. In questi casi, però, più che permettere una doppia detection, consente di isolare i due segnali attraverso i troncamenti a inizio o fine intervallo. Ovviamente rimangono valide le regole per cui le sovrapposizioni non possono essere totali e le bounding box devono andare a identificare parti diverse del segnale sovrapposto. Si può osservare un caso pratico nelle due immagini 6.5 e 6.6, di seguito analizzate.

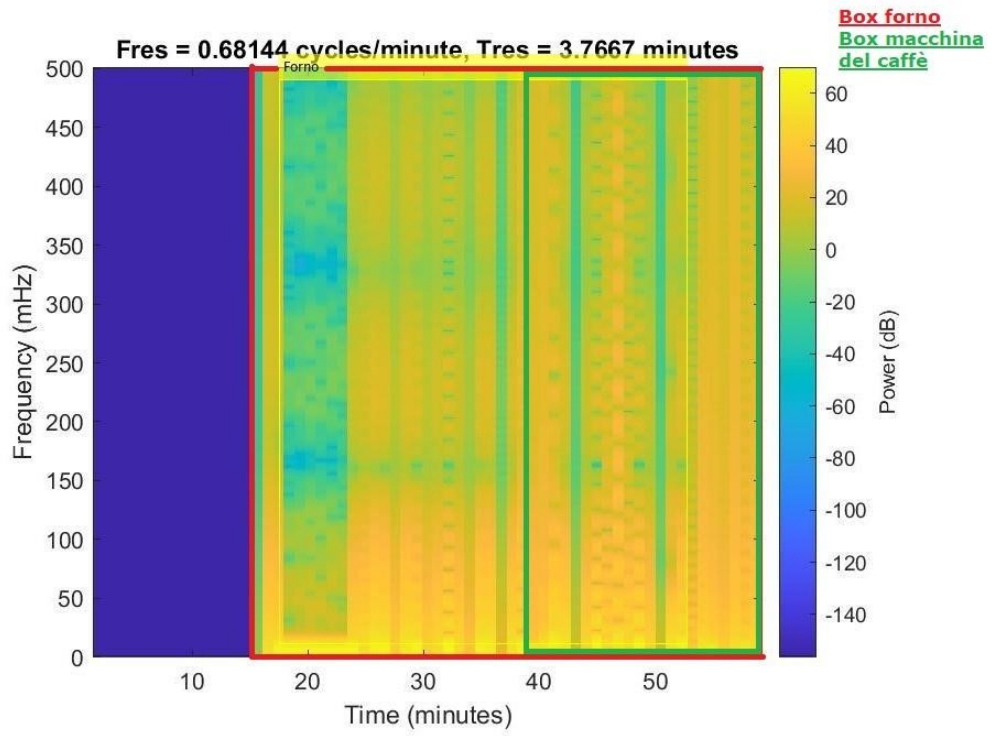


Figura 6.5: immagine contenente la detection su un segnale di un forno

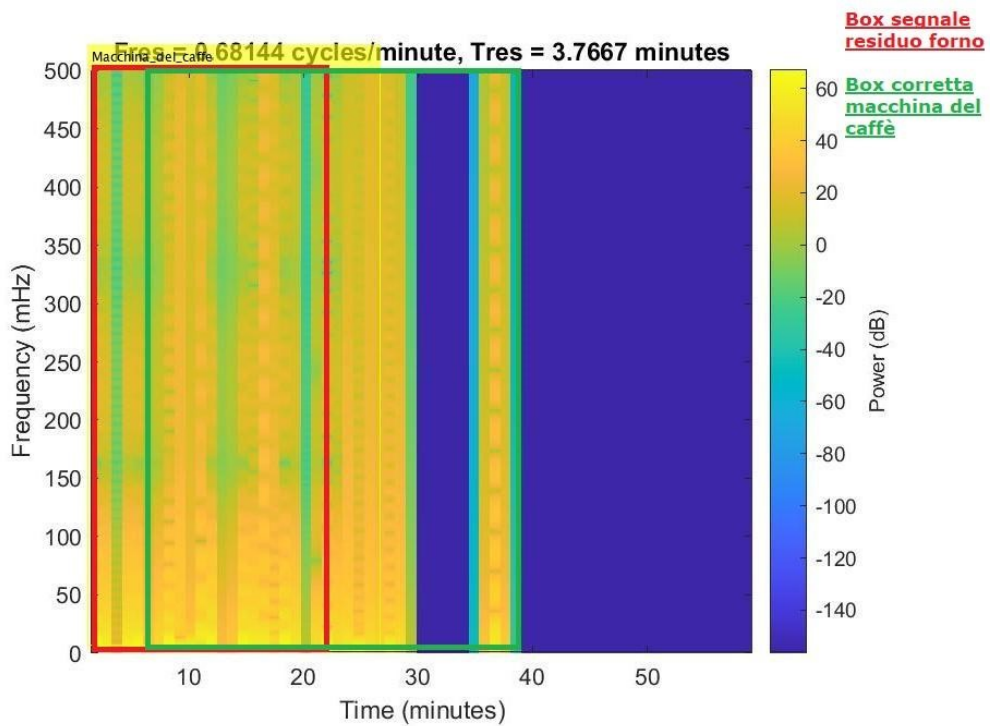


Figura 6.6: immagine contenente una parte del segnale del forno e il segnale della macchina del caffè

Nella prima immagine viene identificato il segnale di un forno, che risulta troncato alla fine della figura. La bounding box è abbastanza precisa, il segnale nella zona del troncamento non è mai ben rappresentato nello spettrogramma ed è anche inquinato dalla presenza del segnale della macchina del caffè, per cui la vera imperfezione è il posizionamento dell'inizio del segnale. Nella seconda immagine è presente una parte del segnale del forno e l'intero segnale della macchina del caffè. Si nota come all'interno di questo vi sia una pausa abbastanza lunga rispetto alla durata totale del funzionamento che rende difficile il riconoscimento dell'appendice, che la rete considera segnale a sé stante pur non riuscendo ad attribuirlo ad un elettrodomestico. Appendice a parte, anche qui il posizionamento tutto sommato è buono, considerando il disturbo dovuto alla presenza del segnale del forno che va a influenzare la detection, che è probabilmente la causa della traslazione verso sinistra della box rispetto al segnale reale. Come per l'esempio precedente, se non si fosse deciso di creare gli intervalli ogni 30 minuti, garantendo più di una detection ai vari segnali, la rete sarebbe risultata molto meno efficace. Come nel caso dei segnali lunghi, la capacità della rete di identificare i segnali in queste situazioni dipende anche dal posizionamento reciproco degli stessi e dal modo in cui si combinano.

Nell'analisi dei risultati ancora non è stato considerato, se non nella definizione delle regole di priorità, un parametro importante che è il punteggio di confidence delle identificazioni. Questo valore è in realtà indicativo, oltre che della certezza con cui la rete identifica il segnale, anche della qualità del posizionamento delle box. Un punteggio molto elevato, superiore a 0,85 sarà molto probabilmente legato ad una detection corretta e ad un posizionamento preciso della bounding box. Confidence superiori a 0,7 possono invece essere dovute a vari fattori. Il primo è quello di un segnale che la rete fatica un po' a riconoscere, come ad esempio i segnali della macchina del caffè (fig. 6.8). Il secondo caso è quello di una corretta detection della bounding box, ma un errata classificazione della stessa, soprattutto per quegli elettrodomestici che non vengono praticamente mai identificati (es. asciugacapelli, fig. 6.9). La terza possibilità è quella di una corretta identificazione con un posizionamento non perfetto della bounding box, magari dovuto alla presenza di altri segnali vicini, troncamenti o segnali contemporanei, come nei casi precedenti (v. fig.6.10). Una confidence compresa tra 0,7 e 0,5 è indice di un peggioramento di queste condizioni o una combinazione delle stesse. In particolare, la maggioranza delle detection sbagliate ha confidence inferiore a 0,6 (figura 6.11).

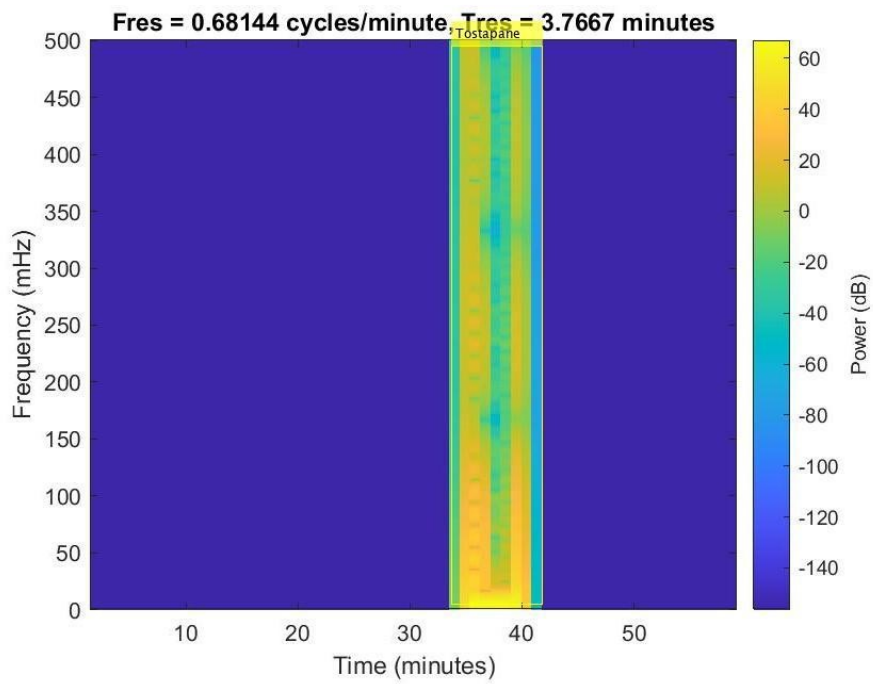


Figura 6.7: identificazione corretta di un tostapane con confidence pari a 0.8971. Il posizionamento della box risulta essere pressoché perfetto

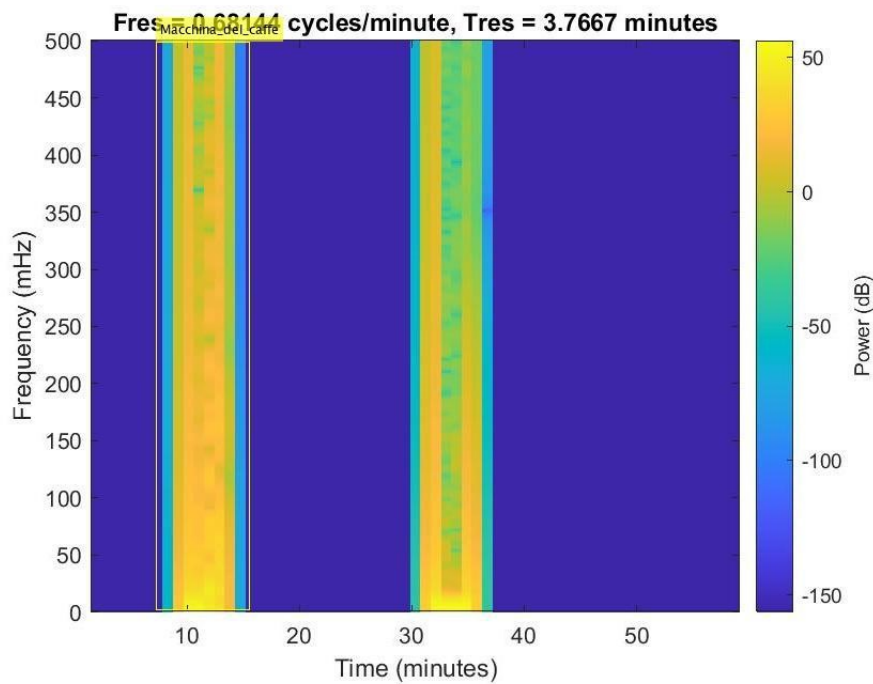


Figura 6.8: Identificazione corretta di un segnale di macchina del caffè con confidence pari a 0,7059. Il posizionamento della box è molto buono. La confidence relativamente bassa è dovuta con tutta probabilità alle difficoltà di riconoscimento della classe, che ricordiamo avere valore di F1-Score pari a 0,73 e percentuale di segnali correttamente identificati del 63%

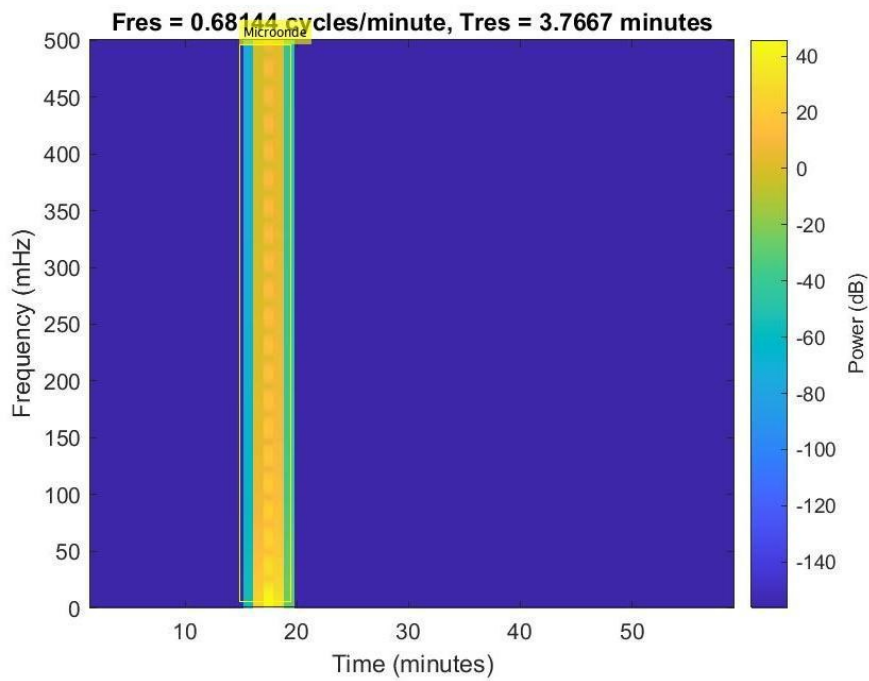


Figura 6.9: immagine relativa all'errata classificazione di un segnale di asciugacapelli. Il posizionamento della box risulta abbastanza buono. Il punteggio di confidence è di 0,7297

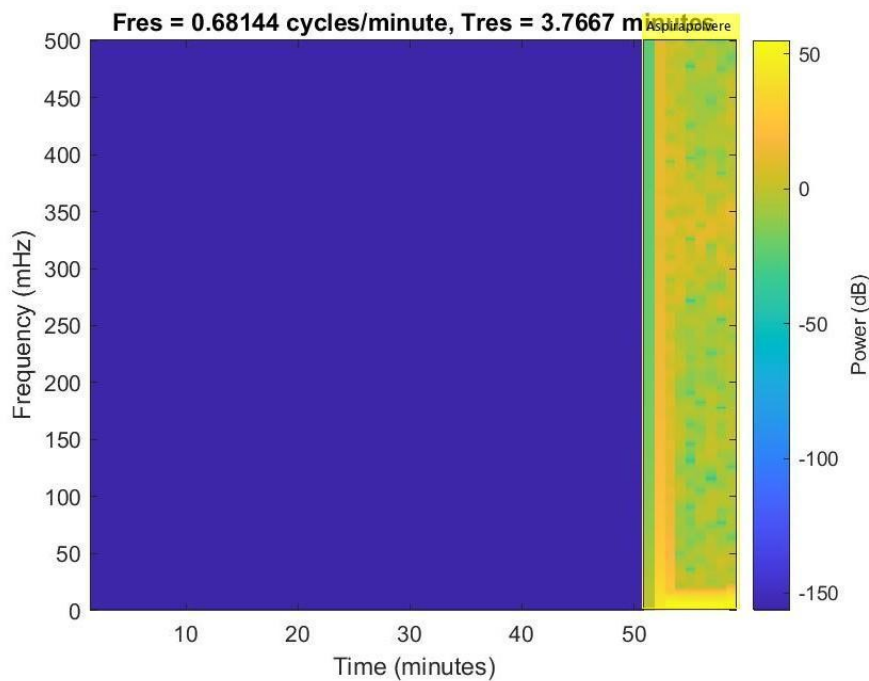


Figura 6.10: immagine relativa ad una identificazione pressoché perfetta del segnale di un'aspirapolvere, troncato a fine intervallo per una porzione inferiore al 20%. Il punteggio di confidence è pari a 0,7625

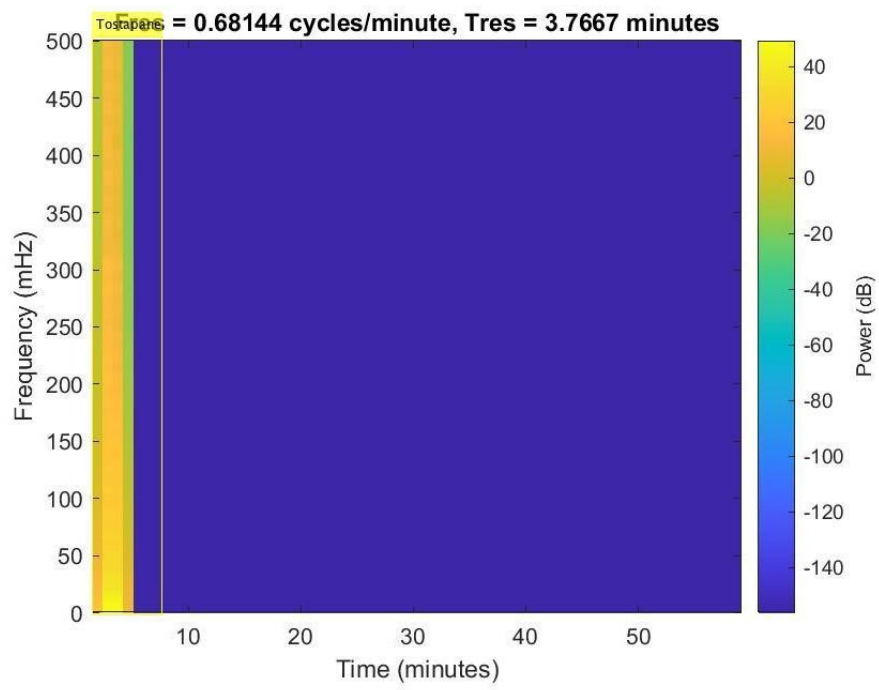


Figura 6.11: immagine che ritrae una detection sbagliata (tostapane anziché aspirapolvere) unita a segnale troncato e box imperfetta. Confidence = 0,5368

6.2 Confronto con risultati degli algoritmi di disaggregazione

Il lavoro presentato da Zorzi^[30] consiste nello sviluppo di due metodi per la disaggregazione e il riconoscimento dei segnali e il successivo calcolo del consumo dei vari elettrodomestici. Il primo metodo usato è il metodo Shift-NMF, il secondo è un metodo di detection avanzata su misurazioni spot degli aggregati (tramite registratore di eventi). Entrambi sono basati sull'algoritmo del *cocktail party*.

6.2.1 Risultati Shift-NMF

Il metodo, applicato sui segnali simulati, riconosce la totalità dei carichi, in quanto le feature della rete sono dati numerici delle curve di consumo e nell'analisi del carico aggregato può andare a ritrovare esattamente le stesse feature e di conseguenza non sbagliare. Una volta utilizzato su segnali aggregati reali le prestazioni vanno ovviamente a scendere. I risultati ottenuti dalla rete, in termini di precision e recall, sono stati i seguenti:

	EVENTI	EVENTI RILEVATI (TP)	FALSI POSITIVI	PRECISION	RECALL	F1
Lavatrice	23	22	4	0.846	0.957	0.898
Ferro da stiro	9	9	0	1.000	1.000	1.000
Microonde	28	28	16	0.636	1.000	0.778
Frigorifero	31	31	13	0.705	1.000	0.827

Tabella 6.2: Risultati ottenuti con metodo Shift-NMF sugli elettrodomestici più frequentemente utilizzati e su cui l'algoritmo è stato maggiormente addestrato

Questo algoritmo risulta molto efficace nell'identificare segnali che si ripetono in maniera molto simile nel tempo, quindi sia in termini di potenza che in termini di durata del funzionamento.

6.2.2 Risultati detection avanzata

Questo metodo è basato sulla pura identificazione degli eventi (rilevazioni di variazioni sensibili di potenza assorbita). Si va quindi a slegare il riconoscimento dalla durata degli eventi e si migliorano le prestazioni su tutti quegli elettrodomestici il cui funzionamento differisce in termini di tempo tra un utilizzo e l'altro. Di contro, però, ha difficoltà nel momento in cui ci sono eventi in contemporanea, che non riesce a riconoscere, e nell'identificare tutti quei segnali complessi o continui che non presentano variazioni improvvise o che presentano variazioni di potenza diverse tra accensione e spegnimento. Inoltre, fatica a riconoscere i carichi quando si sovrappongono, soprattutto quando queste sovrapposizioni riguardano segnali complessi (come, ad esempio, quello di una lavatrice). I risultati globali ottenuti con questo algoritmo sono stati i seguenti:

	<i>EVENTI</i>	<i>EVENTI RILEVATI (TP)</i>	<i>FALSI POSITIVI</i>	<i>DETECTION CORRETTE</i>	<i>PRECISION</i>	<i>RECALL</i>
Frigorifero	92	92	0	92	1.000	1.000
Lavatrice	72	69	0	0	1.000	0.958
Forno	3	3	0	3	1.000	1.000
Piastra capelli	3	0	0	0	-	0.000
TV	118	0	0	0	-	0.000
Aspirapolvere	43	30	0	21	1.000	0.698
Ferro da stiro	24	23	0	0	1.000	0.958
Microonde	92	89	0	70	1.000	0.967
Asciugacapelli	6	4	0	2	1.000	0.667

Tabella 6.3: risultati ottenuti con il metodo della detection avanzata su singoli elettrodomestici

Globalmente, sui 310 segnali analizzati, si ottiene un dato di recall pari a 0,68 e una percentuale totale di eventi correttamente identificati del 60%. Per definire un evento come correttamente identificato serve il corretto riconoscimento, ma anche la corretta ricostruzione del segnale. Come si evince dalla tabella, attraverso questo algoritmo è molto difficile se non impossibile riconoscere correttamente i segnali più complessi (lavatrice e ferro da stiro) e i segnali continui a basso impatto energetico (TV).

6.2.3 Risultati applicazione reti neurali ai metodi precedenti

Nel lavoro svolto da Zorzi si è andati ad applicare delle reti appositamente addestrate ai metodi di detection citati. L'applicazione ha riguardato nello specifico solamente i segnali che meglio venivano identificati da ciascun metodo. Oltre a questi dati, sono stati resi disponibili anche i risultati della rete sul set di addestramento. Si riporta di seguito la legenda della tabella 6.4

Numero Elettrodomestico:

- 1 Macchinetta del caffè
- 2 Piano cottura
- 3 Lavastoviglie
- 4 Asciugatrice
- 5 Freezer
- 6 Ferro da stiro
- 7 Bollitore
- 8 Microonde
- 9 Frigorifero
- 10 Tostapane
- 11 Lavatrice
- 12 Lavasciuga
- 13 Forno
- 14 Asciugacapelli
- 15 Cappuccinatore
- 16 Bistecchiera
- 17 Robot da cucina
- 18 Piastra ad induzione
- 19 Aspirapolvere
- 20 Tagliaerba

Output Class	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	Accuracy
1	748	0	0	0	0	0	0	19	0	2	0	0	6	0	0	0	0	1	0	0	96.4%
2	7	25	0	0	0	0	0	0	0	8	0	0	1	1	0	0	0	1	0	0	58.1%
3	0	0	678	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	41.9%
4	0	0	1	189	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	99.7%
5	0	0	0	0	516	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0.3%
6	0	0	0	1	0	35	0	2	0	0	0	0	0	0	1	0	0	0	0	0	99.8%
7	0	0	0	1	0	1	1343	8	0	0	0	0	1	0	0	0	0	0	0	0	0.2%
8	16	0	0	1	2	2	725	0	12	1	0	0	1	0	0	0	0	1	0	0	0.8%
9	0	0	0	0	2	0	0	0	1282	0	1	0	0	0	0	0	0	0	0	0	95.3%
10	6	10	0	0	3	0	7	0	251	0	0	0	7	0	0	0	0	2	2	0	4.7%
11	0	0	2	0	2	0	0	0	0	862	3	0	0	0	0	0	0	0	0	0	18.5%
12	0	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	3.6%
13	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0	0	0	100%
14	0	1	0	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	7	0	0.9%
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.3%
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0%
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0.0%
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0.0%
19	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	10	0	0	0.1%
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.0%
	96.3%	69.4%	99.6%	99.0%	99.4%	81.4%	99.9%	95.3%	99.9%	91.3%	99.2%	0.0%	100%	53.7%	0.0%	0.0%	100%	100%	45.5%	0.0%	97.6%
	3.7%	30.6%	0.4%	1.0%	0.6%	18.6%	0.1%	4.7%	0.1%	8.7%	0.8%	100%	0.0%	46.3%	100%	100%	0.0%	0.0%	54.5%	100%	2.4%
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	

Tabella 6.4: matrice di confusione dei risultati della rete neurale applicata all'intero set di addestramento^[30]

Dalla matrice di confusione si può notare come la precisione della rete sia in generale molto elevata, ma che elettrodomestici come asciugacapelli, aspirapolvere e tostapane presentino percentuali di riconoscimento relativamente basse e vengano confusi tra loro a causa delle similitudini tra le loro curve di consumo.

Applicando la rete neurale al metodo Shift-NMF ad elettrodomestici quali forno, frigorifero, lavatrice, microonde ed altri brevi (ad esempio aspirapolvere), si è notato come le identificazioni fossero buone nel caso in cui si fosse riuscito ad ottenere una buona disaggregazione dei segnali, ma che in molti casi finiva con l'attribuire gli eventi all'utilizzo della lavatrice, per colpa di alcune correlazioni tra le fasi finali del ciclo della lavatrice stessa con i picchi di potenza degli altri elettrodomestici.

Allo stesso modo è stata applicata la rete al metodo di detection avanzata, ottenendo ottimi risultati su frigorifero, microonde, ferro da stiro e forno, con oltre il 90% di segnali identificati, mentre asciugacapelli e aspirapolvere venivano confusi tra loro.

Attraverso i metodi di disaggregazione proposti si è andati anche a calcolare i consumi sia dei singoli elettrodomestici che dell'aggregato. I risultati sono influenzati dal fatto che viene utilizzato il consumo in termini di potenza apparente e dal fatto che le curve devono essere ricostruite e pertanto ci saranno curve non ottimamente ricostruite e altre non riconosciute. La misurazione della potenza

apparente porta ad una sovrastima importante del consumo totale e rende necessario delle correzioni successive per riportare il consumo ai valori di potenza attiva. Tutto questo, per quanto possa essere calcolata bene la fase dei vari elettrodomestici, porterà sicuramente ad avere un certo margine di errore, che Zorzi è riuscito a limitare a pochi punti percentuali.

6.3 Confronto con i risultati della rete Zorzi

Nonostante i due algoritmi siano molto diversi tra loro, presentano alcune similitudini abbastanza evidenti, che riguardano soprattutto i loro punti deboli. La prima riguarda la difficoltà nel distinguere segnali dalla curva molto simile tra loro. I due metodi di disaggregazione proposti nel lavoro di Zorzi vanno a identificare gli elettrodomestici attraverso feature legate numericamente alla curva, quali potenza massima, numero di picchi, potenza media e molte altre. Una volta disaggregato il segnale la rete cerca di riportare le varie parti ai segnali presenti in una matrice dizionario. Avere segnali simili, o anche solo con alcune correlazioni, fa sì che la rete possa confondere gli elettrodomestici. La rete di riconoscimento degli spettrogrammi, non essendo legata direttamente alle caratteristiche numeriche della curva, è meno influenzata dalle piccole correlazioni tra segnali, ma fatica molto nel distinguere elettrodomestici con curve simili e soprattutto durata di funzionamento simile.

La seconda similitudine è legata alla difficoltà nel riconoscere segnali contemporanei. Nel caso del metodo della detection avanzata si riesce in parte a risolverlo, in quanto si basa sulle variazioni discrete di potenza assorbita e riesce quindi a individuarne le attivazioni. Questo però non succede se uno dei due segnali è complesso, composto da diversi cicli ON/OFF, come ad esempio possono essere lavatrice, lavastoviglie e asciugatrice, o se gli istanti di ON o di OFF dei due segnali sono contemporanei. Su questo punto, la rete per il riconoscimento delle immagini risulta ancora più penalizzata, non potendo contare sulla rilevazione delle attivazioni. Come visto in precedenza, la rilevazione di segnali sovrapposti è molto difficile e può avvenire solo se le sovrapposizioni avvengono all'inizio o alla fine del segnale più lungo, ancora meglio se sono parziali. Inoltre, non si riesce a individuare la sovrapposizione di segnali brevi.

La terza similitudine, anche se per certi aspetti può sembrare banale, è quella delle classi di elettrodomestici che vengono ben identificate, che sono in gran parte le stesse per entrambe le reti. Per quanto riguarda la rete per la detection delle immagini, abbiamo visto ottenere ottimi risultati su asciugatrice, bollitore, forno, piastra a induzione e lavatrice (con più dell'85% di segnali correttamente individuati) e buoni risultati (> 66% di detection corrette) su aspirapolvere, ferro da stiro e tostapane. Per quanto riguarda i due algoritmi proposti da Zorzi, si sono ottenuti ottimi risultati su lavatrice, aspirapolvere, forno, microonde, frigorifero e ferro da stiro e buoni risultati anche nelle classificazioni dei segnali dell'asciugacapelli. Buona parte di questi elettrodomestici hanno ottenuto alte percentuali di detection da parte di entrambe le reti, dimostrando come curve di consumo caratteristiche e una grande presenza di dati su cui addestrare le reti aiutino molto nella corretta individuazione.

Tra le due reti vi sono però delle nette differenze, che possono portare a preferire una all'altra a seconda dell'ambito nelle quali le si vuole utilizzare, ma aprono anche la possibilità di trovare un modo per combinarle. La prima riguarda la capacità di individuare i segnali. La rete che opera sugli spettrogrammi manca il riconoscimento del 13,7% dei segnali identificabili e il 23,9% considerando anche quelli classificati come non riconoscibili. Se considero invece le immagini, il dato peggiora ulteriormente, andando a sfiorare il 40%, rendendo necessario analizzare più volte i segnali includendoli in più spettrogrammi. La capacità di riconoscere la presenza dei segnali della detection avanzata, invece, è decisamente migliore. Escludendo piastra per capelli e tv, che non sono state prese in considerazione nella nostra rete, vengono mancati poco meno del 7% dei riconoscimenti. Questo dato va ovviamente pesato, in quanto sono considerati segnali su cui il metodo è molto efficace e tendenzialmente andrà a peggiorare includendo le altre classi, ma rimane comunque un punto di forza a favore del metodo di disaggregazione.

Di contro, però, la rete per l'identificazione delle immagini è molto precisa nel posizionamento delle box, che significa delimitare correttamente nel tempo i segnali. Non andando a modificare la curva di consumo aggregato in fase di disaggregazione, si può, di conseguenza, facilmente risalire ad orario di inizio e fine del funzionamento e isolare l'intero segnale dell'elettrodomestico identificato. Il

grosso problema delle reti neurali applicate agli algoritmi Shift-NMF e di detection avanzata è che, nonostante riconoscano bene la presenza dei segnali, una buona percentuale di questi non viene rilevata correttamente, in quanto la curva di consumo non viene adeguatamente ricostruita. Dei segnali identificati solamente il 60% viene correttamente identificato. Questo avviene sia quando ci sono segnali contemporanei che per segnali complessi.

Un'altra differenza sta nel tipo di segnale che le reti vanno ad analizzare e il relativo campionamento utilizzato. La rete per il riconoscimento delle immagini è stata addestrata su spettrogrammi di segnali di potenza attiva ottenuti con campionamento a 1 Hz e non riesce a identificare segnali ottenuti con frequenze di campionamento molto più basse o, ancora peggio, con il sistema spot utilizzato nel rilevamento dell'aggregato reale. La rete sviluppata da Zorzi è stata addestrata sugli stessi segnali, ma opera su segnali campionati con il sistema di registrazione eventi, che sfrutta per identificare i momenti di ON e OFF dei segnali, e utilizza la potenza apparente, il che comporta la necessità di una rielaborazione per ottenere i dati di consumo.

6.4 Confronto con i risultati ottenuti con un sistema commerciale (NED)

Sugli stessi dati su cui è stata verificata la rete che applica gli algoritmi di disaggregazione basati sul cocktail party, è stato testato da Zorzi anche un sistema presente in commercio: NED. Il sistema, progettato dalla torinese Midori SRL, rileva il segnale direttamente dal contatore attraverso una pinza amperometrica (quindi senza modifiche nella linea elettrica)^[31]. Oltre al riconoscimento dei carichi, va anche a calcolare i consumi globali. All'interno del prodotto sono presenti anche un sistema per l'etichettatura manuale dei segnali e dei programmi di training. Non sono forniti dati relativi alle frequenze di campionamento utilizzate né sul sito né sul manuale del dispositivo, ma è stato ipotizzato che la rilevazione venga effettuata ad alta frequenza, nell'ordine del kHz^[30].

Il dispositivo NED è stato applicato allo stesso contatore del sistema UniPD-Lab.E.E. ed è quindi andato ad analizzare gli stessi segnali rielaborati dagli algoritmi di disaggregazione visti precedentemente. NED riconosce perfettamente frigorifero, lavatrice e forno, il quale però ha anche un numero rilevante di falsi positivi. Il ferro da stiro viene individuato quasi sempre, ma in meno della metà delle volte le detection si sono rilevate corrette. Aspirapolvere, microonde e asciugacapelli non vengono quasi mai riconosciuti. L'etichettatura manuale degli eventi non rilevati aiuta, in quanto ad ogni etichettatura manuale corrisponde un nuovo addestramento, ma non sempre questo è sufficiente per far apprendere alla rete il segnale. Le etichette poste dall'utilizzatore possono essere modificate in caso di errore in qualsiasi momento, mentre le etichette generate autonomamente da NED non possono essere modificate, per cui l'addestramento aiuta ad imparare segnali nuovi e ad adattarsi ai modelli presenti nell'abitazione, ma non aiuta a correggere le etichettature sbagliate in ottica delle detection future.

Anche il calcolo del consumo non è ottimizzato, in quanto va a sottostimare di circa il 30% i consumi globali. Questo è un difetto importante per un dispositivo che si propone di aiutare l'utente nell'attuare comportamenti atti al risparmio energetico.

In conclusione, possiamo dire che le difficoltà nell'identificazione dei segnali permangono anche in sistemi presenti in commercio.

6.5 Pro e contro dell'utilizzo della rete YOLO

Dopo aver analizzato caratteristiche e comportamento delle reti, possiamo andare ad analizzare quali sono i vantaggi e le note positive dell'utilizzo di un detector di immagini e quali sono i suoi punti deboli.

PRO:

- Semplicità della rete, sia per quanto riguarda la parte di progettazione e addestramento che per i principi e le nozioni necessarie per ottimizzarne il rendimento. A livello di progettazione si deve solamente effettuare una scelta sul tipo di rete da utilizzare, in quanto, essendoci già reti ben sviluppate per il riconoscimento delle immagini, non è necessaria una progettazione da zero. Per il resto, si deve decidere che tipo di intervalli temporali utilizzare per la creazione degli spettrogrammi e creare un database sufficientemente ampio e variegato. Si tratta quindi solamente di acquisire la giusta sensibilità nella creazione dei database e nel decidere i parametri di addestramento. Oltre a questo, serve solamente una conoscenza del comportamento dei vari elettrodomestici e delle loro curve. Per le reti di detection avanzata e Shift-NMF, invece, è necessario la conoscenza degli algoritmi di disaggregazione e i concetti relativi al cocktail party per progettare al meglio la rete.
- La rete riesce a riconoscere con ottimi risultati un buon numero di elettrodomestici anche all'interno di un aggregato con molti elettrodomestici attivi al suo interno. Forno, lavatrice, asciugatrice, tostapane, bollitore e piastra a induzione hanno percentuali di detection corrette superiori all'80%. A queste possiamo aggiungere ferro da stiro e aspirapolvere che vengono riconosciute in circa due terzi dei casi. Resta da testare su segnali non simulati.
- Il posizionamento delle bounding box, quindi l'identificazione degli istanti di attivazione e spegnimento dei segnali, ha una buona precisione. Unito al fatto che non vengono ricostruite le curve, fa sì che possa essere assegnato alla classe il suo segnale esatto. Di conseguenza, non ho problemi legati a detection non corrette a causa di errori nell'identificazione dei momenti di ON e OFF o per errori nella ricostruzione dei segnali, soprattutto nel caso di segnali complessi, che corrispondono solitamente anche agli elettrodomestici più energivori. Per associare i segnali ad un orario preciso del giorno anziché all'istante di un intervallo generico è sufficiente registrare i segnali all'interno di una timetable anziché una tabella o un vettore normale.
- Non utilizzando algoritmi di disaggregazione, il segnale aggregato rimane intatto e si può andare a calcolare facilmente il consumo di ciascun elettrodomestico semplicemente andando a ritroso dal posizionamento della bounding box all'orario di attivazione e spegnimento dei segnali.
- Non essendo legata alla rilevazione di eventi e quindi alla variazione di stato discreta degli elettrodomestici, la rete di detection delle immagini risulta essere più efficace su segnali complessi multi-stato o segnali a variazione continua.
- Rispetto ai dispositivi commerciali, non necessita di frequenze di campionamento particolarmente elevate, il che comporta un alleggerimento nella quantità di dati che devono essere inviati ed elaborati dalla rete.
- Operando su immagini, la rete fornisce risultati di immediato impatto visivo, che può agevolare l'utilizzo da parte di un possibile utente medio, che può anche comprenderne più facilmente l'operato.

CONTRO:

- I problemi relativi all'identificazione di segnali contemporanei non vengono risolti. Dall'analisi delle detection è emerso come la rete riesca a riconoscere segnali solo con sovrapposizioni parziali e che non riguardano segnali lunghi e nemmeno nella totalità dei casi. In qualsiasi altro caso di sovrapposizione solo il segnale più lungo o più "detectabile" potrà essere identificato. Inoltre, in nessuno di questi casi si riuscirà ad ottenere informazioni corrette in post-elaborazione, come ad esempio il consumo dei singoli elettrodomestici. Utilizzando il metodo della detection avanzata, invece, si può almeno in parte ovviare al problema, in quanto l'algoritmo va ad individuare i momenti di ON e OFF, salvo che in rari casi (eventi contemporanei o eventi legati a segnali complessi multi-stato).
- Non vengono risolti nemmeno i problemi relativi agli errori nell'identificazione dei carichi aventi curve simili per forma e durata temporale, tipicamente segnali brevi e del tipo ON/OFF, che sono già più difficili da identificare rispetto agli altri segnali per i limiti strutturali della rete YOLO.
- Il numero di segnali identificati dalla rete è minore rispetto alle reti che applicano gli algoritmi di disaggregazione ed è maggiore il numero di segnali erroneamente identificati.
- Essendo necessario dividere in più intervalli sovrapposti tra loro per permettere di posizionare almeno una volta ciascun segnale interamente all'interno di un intervallo (e quindi dello spettrogramma corrispondente) per avere una corretta identificazione, si deve prevedere un algoritmo ausiliario che riconduca il posizionamento delle varie bounding box ai corretti istanti del segnale aggregato e assegnare la priorità nel caso in cui venga identificato lo stesso segnale in più immagini.
- Per poter identificare segnali di lunghezze diverse, è necessario combinare l'utilizzo di due reti. Si rende necessario uno script ausiliario che vada ad assegnare le priorità della detection alla rete corretta, in maniera analoga a quanto scritto al punto precedente.
- Allargando il database per migliorare le prestazioni su tutte le classi o includere nuove classi, rischia di allungare notevolmente i tempi di addestramento e rendere molto pesante il database. Quest'ultima cosa renderebbe più difficile inserire all'interno di un dispositivo che utilizza questo tipo di rete un sistema che vada a ripetere l'addestramento in seguito all'inserimento manuale delle etichette dei segnali non identificati, impedendo o rendendo più difficoltoso il miglioramento delle prestazioni della rete "personalizzato" sugli elettrodomestici dell'utenza.
- Non operando sulla curva, necessita di operazioni ulteriori (per quanto semplici) per ricondurre la detection dell'immagine al segnale presente nella curva dell'aggregato e ottenere tutti quei parametri da esso derivati, quali ad esempio costi e consumi.

CONCLUSIONI

Come tutte le tipologie di rete neurale addestrate allo scopo del monitoraggio non intrusivo dei carichi elettrici, allo stato attuale anche questa rete presenta dei limiti. Presa coscienza di ciò, possiamo considerare i risultati ottenuti come buoni. La rete YOLO riesce a identificare con buone se non ottime percentuali un discreto numero di elettrodomestici. Ha maggiori difficoltà nell'identificare e classificare tutti gli eventi presenti all'interno dell'intervallo considerato rispetto agli algoritmi di disaggregazione, ma di contro ha un'ottima capacità di delimitarli correttamente a livello temporale. Non andando a modificare il segnale originale, l'identificazione degli istanti di inizio e fine del segnale è sufficiente per ottenere una corretta classificazione e identificazione dei segnali, cosa che avviene con molta meno frequenza nella rete che utilizza gli algoritmi di disaggregazione. Riuscire a identificare il 63% del totale dei segnali e il 75% dei segnali degli elettrodomestici identificabili risulta essere un risultato ottimo, potendo definire queste detection corrette sotto ogni punto di vista.

A rendere ancora più interessante questa metodologia, che non è tra le principali utilizzate attualmente nei vari dispositivi NILM^{[32][33]}, è la sua semplicità. Non è complessa da progettare, in quanto basata su reti già "preconfezionate", né da addestrare. Le maggiori complicazioni a livello di programmazione sono rappresentate dalle operazioni di post-processing da attuare per andare a riportare le detection sulla linea temporale e sul segnale aggregato completo. Queste operazioni si basano sulla correlazione tra il posizionamento all'interno dell'immagine dello spettrogramma e l'intervallo orario in esso rappresentato, che sono tra loro collegati da una proporzionalità che dipende dalla suddivisione del tempo da noi scelta e dalle dimensioni dell'immagine dello spettrogramma. Da qui, successivamente, si può procedere ancora in maniera piuttosto semplice al calcolo della potenza assorbita da ciascun elettrodomestico, dato che la curva di consumo rimane intatta (se non per l'eliminazione del rumore) e si hanno a disposizione gli istanti esatti di inizio e fine dei segnali. Utilizzando la curva di potenza attiva assorbita, non si rende necessario applicare correzioni al calcolo dei consumi, che sono ovviamente una ulteriore fonte di errore. Inoltre, ha un immediato impatto visivo nell'utente, cosa che può risultare utile per semplificarne la lettura e la comprensione e non complica l'eventuale introduzione di un processo di etichettatura manuale.

Rispetto ad altre reti ed altri algoritmi, non utilizzando misurazioni di tipo spot o comunque la rilevazione di eventi, risulta facilitata nel riconoscere correttamente segnali diversi da quelli di tipo ON/OFF, quali possono essere segnali a variazione continua o multi-stato, che sono tra i più difficili da identificare in maniera corretta. Di contro, ho maggiori difficoltà nell'identificazione dei segnali contemporanei, e più in generale si ottiene un maggior numero di segnali che non vengono classificati. In questo senso si può ipotizzare l'utilizzo combinato dei due diversi approcci, utilizzando un algoritmo di disaggregazione per identificare il maggior numero di eventi all'interno dell'aggregato e se possibile anche isolarli, per poi procedere con la classificazione dei rispettivi spettrogrammi con il detector di immagini. Questa soluzione potrebbe andare a mascherare di fatto i punti deboli dei due approcci, ma risulterebbe sicuramente molto più pesante da inserire all'interno di un singolo dispositivo, dovendo combinare l'utilizzo di almeno tre reti neurali. Inoltre, può rendere più facile "inquadrare" lo spettrogramma che si vuole analizzare all'interno dell'immagine dell'intero intervallo, creando spettrogrammi con i singoli segnali isolati dall'algoritmo in modo da effettuare le detection su segnali puliti e posizionati a piacimento, senza dover procedere con la creazione di un numero elevato di intervalli di osservazione, quindi di spettrogrammi, per poi inserire degli algoritmi che vadano ad individuare la detection con maggiore confidence tra quelle effettuate per assegnarla successivamente al segnale all'interno della curva di consumo completa. Ad alleviare il problema legato alla quantità elevata di dati elaborata dalle reti, c'è il fatto che non sia necessario andare a campionare i dati di consumo aggregati ad alte frequenze, diminuendo notevolmente la quantità di dati in ingresso.

Pensando ad una futura implementazione su dispositivi commerciali, questa non è l'unica complicazione. Come visto per NED, e in generale per lo stato dell'arte attuale delle tecnologie NILM^{[32][33]}, ci sono ancora difficoltà nel fornire un dispositivo con una rete addestrata in maniera tale da riconoscere con grande efficacia segnali ad essa sconosciuti (inteso come non presenti all'interno dei database di addestramento). È quindi importante dare la possibilità di una etichettatura manuale dei segnali non identificati dal dispositivo per permettere allo stesso di essere costantemente

aggiornato con nuovi addestramenti. Come abbiamo visto, per addestrare le reti sono servite circa 16500 immagini, per un peso totale di quasi 800 MB. Questa importante mole di dati va ovviamente ulteriormente ampliata per migliorare le prestazioni nel riconoscimento di quegli elettrodomestici non sufficientemente rappresentati e per includerne altri esclusi da noi all'inizio del lavoro, non essendo questo finalizzato alla creazione di una rete da commercializzare. Si può facilmente intuire che mettere a disposizione un database così pesante, che verrà ulteriormente ampliato dalle etichettature manuali, può rappresentare un primo ostacolo. Ancor più gravoso diventa il fatto di dover procedere con nuovi addestramenti ogni volta che un segnale identificato dall'utente viene aggiunto al database, anche solo banalmente per i tempi necessari all'elaborazione di una quantità così grande di immagini durante il processo di training. Ciò viene ulteriormente aggravato andando ad unire le due metodologie, in quanto ci si ritroverebbe a dover aggiornare i pesi tramite addestramenti di almeno tre reti distinte. La soluzione più logica a questo inconveniente sarebbe quella di garantire l'aggiornamento del dispositivo facendo fare gli addestramenti delle reti online anziché in locale. Per fare ciò, si può pensare anche alla creazione di un unico database comune presente in rete che viene aggiornato e aumentato dagli utenti, sulla falsa riga del funzionamento di Wikipedia. Basandosi sull'intervento di una terza parte, ossia l'utente, e facendo affidamento sulla sua responsabilità nell'inserire dati il più possibile corretti, dovrebbe però anche essere regolamentato e sorvegliato per evitare il deterioramento del database dovuto all'utilizzo scorretto dello strumento di etichettatura manuale.

Un altro problema, non secondario, ma comune a tutti i metodi utilizzati in ambito NILM, è quello relativo all'identificazione di segnali contemporanei e di segnali con curve simili, quindi confondibili. Riguardo i segnali contemporanei, la rete YOLO ha limiti strutturali che non possono essere risolti, se non con l'utilizzo di una seconda rete che vada a separare ed isolare i due segnali, cosa ancora molto difficile anche per gli algoritmi di disaggregazione. Questo perché a livello pratico, la presenza di due segnali contemporanei, che un algoritmo di disaggregazione può riuscire a leggere come una sovrapposizione dei segnali, per la rete YOLO, che opera sullo spettrogramma, la contemporaneità genera un "mescolamento" delle immagini delle due curve in una sola. Pertanto, nella maggior parte dei casi, uno dei due segnali verrà disperso nell'altro (il più lungo tra i due) e non verrà riconosciuto o verrà generata un'immagine di un segnale completamente diverso da entrambi i segnali originali, rendendo ancora una volta impossibile l'identificazione.

Per poter pensare concretamente all'implementazione su dispositivi commerciali, oltre alle classiche funzioni di identificazione e di calcolo dei consumi, vanno anche valutate le possibilità di utilizzo per quelle che sono le potenziali future applicazioni e direzioni di ricerca delle metodologie NILM. Ci si è limitati a considerare solamente applicazioni in ambito domestico relativo a singole utenze. I due principali campi di ricerca sono quello dell'*anomaly detection* e dell'*Ambient Assisted Living*^[33].

Per quanto riguarda il primo, consiste letteralmente nell'analisi dei segnali atta alla ricerca e all'individuazione di anomalie al loro interno, allo scopo di evidenziare cambiamenti nella curva di consumo che possono essere dovuti a malfunzionamenti o semplice usura. Questo è importante per evitare l'insorgere di guasti, permettendo di procedere tempestivamente con operazioni di manutenzione o sostituzione dell'elettrodomestico, e di evitare che il cattivo funzionamento dello stesso porti a consumi di corrente eccessivi, che causano sia un maggiore impatto ambientale che una maggiore spesa per l'energia da parte dell'utente, oltre ad un risparmio nella manutenzione dovuta alla tempestività dell'intervento. Ritengo che, in questo ambito, possa, almeno per il momento, essere una soluzione anche più efficace rispetto agli algoritmi di disaggregazione visti nei capitoli precedenti, grazie alla precisione con cui vengono identificati gli eventi e al fatto di non dover ricostruire la curva di potenza assorbita, quindi al mantenimento della curva reale del segnale dell'elettrodomestico. Unendo alla rete YOLO un algoritmo per il riconoscimento degli eventi che aiuti nel riconoscimento della presenza di segnali contemporanei si riuscirebbe anche ad evitare in gran parte gli errori che verrebbero commessi a causa della presenza di un secondo segnale che si somma a quello identificato, che risulterebbe a tutti gli effetti come un'alterazione della normale curva di consumo e sarebbe con ogni probabilità identificato come una anomalia dell'unico elettrodomestico identificato dalla rete YOLO.

L'Ambient Assisted Living è un programma dell'Unione Europea che si propone lo scopo di finanziare progetti di tecnologie innovative per l'assistenza degli anziani in ambito domestico^{[34][35]}. Si tratta, quindi, di tecnologie che fanno riferimento a sistemi intelligenti, come possono essere quelli domotici, che integrino meglio persone con difficoltà all'interno dei loro ambienti domestici per agevolare loro la quotidianità. L'applicazione dei dispositivi NILM in questo ambito serve a monitorare le attività delle persone e aiutare ad evitare comportamenti scorretti o anomali, quali ad esempio quelli attuati da persone con disturbi del sonno. La detection dei segnali può aiutare nell'individuare l'utilizzo di dispositivi in orari non consoni e coadiuvare l'utilizzo di altra sensoristica per avere un monitoraggio il più possibile ampio e completo^[33]. Anche in questo caso, la rete YOLO può tornare utile, potendo fornire identificazioni corrette e potendole correlare all'orario di utilizzo.

In generale, possiamo dire che la metodologia proposta può rappresentare una soluzione efficace da implementare in dispositivi NILM, anche se ha ancora limiti e difetti evidenti da limare, come per tutte le metodologie attualmente in voga in questo ambito. Essa presenta dei vantaggi evidenti nella corretta identificazione dei segnali, anche se è meno efficace nella loro rilevazione e classificazione, unita ad una semplicità di programmazione e all'impatto visivo sull'utente che possono renderla appetibile. Per migliorare in modo concreto le prestazioni ci sono diverse strade percorribili, ciascuna con motivazioni valide a seconda dello scopo per la quale la si vuole sviluppare. La più importante è quella di implementare un algoritmo di identificazione degli eventi che riesca ad aumentare la percentuale di segnali identificati e di isolarli con buona precisione consentendo alla rete di lavorare su spettrogrammi contenenti segnali "puliti", senza sovrapposizioni, rumori e problemi legati al posizionamento degli stessi all'interno degli intervalli di misura. Ha anche possibilità di sviluppo e applicazione nei maggiori ambiti di ricerca attuali per i dispositivi NILM, anche se ovviamente più specifico è il campo di utilizzo e maggiore dovrà essere l'adattamento della rete generica allo scopo, per cui per valutare al meglio quanto possano essere efficaci in specifiche applicazioni servono studi altrettanto specifici su reti ad esse adattate. Questo tipo di soluzione può quindi essere valutata concretamente, data la combinazione di buone prestazioni, efficacia e semplicità. Di contro, la pesantezza della tipologia di dati presenti nel database (ossia le immagini) e quindi le maggiori difficoltà che alla lunga si possono avere per l'aggiornamento dello stesso e quindi dell'apprendimento continuo tramite addestramenti automatici all'inserimento di nuovi dati nel database, può essere un grande freno per uno sviluppo di questo tipo di reti e fornire una maggiore potenzialità alle reti che utilizzano algoritmi di disaggregazione, nonostante anch'esse presentino ancora limiti e difetti.

Bibliografia

- [1] G. W. Hart, "Nonintrusive appliance load monitoring," in *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870-1891, Dec. 1992, doi: 10.1109/5.192069.
- [2] https://en.wikipedia.org/wiki/Nonintrusive_load_monitoring
- [3] <https://patents.google.com/patent/US4858141>
- [4] https://en.wikipedia.org/wiki/Assisted_living
- [5] arXiv:2102.12578; Akriti Verma, Adnan Anwar, M. A. Parvez Mahmud, Mohiuddin Ahmed and Abbas Kouzani, "A Comprehensive Review on the NILM Algorithms for Energy Disaggregation", 15 Apr. 2021.
- [6] https://it.wikipedia.org/wiki/Inferenza_bayesiana
- [7] Deep Learning per la disaggregazione di consumi elettrici di edifici residenziali; C. La Riccia, C. Liti, V. Piccialli, A. Pomente (DICII "Università degli Studi di Roma "Tor Vergata"), Settembre 2018, Report RdS/PAR2017/048; ENEA.
- [8] Choudhury, B., & Jha, R. (2016). Soft Computing Techniques. In *Soft Computing in Electromagnetics: Methods and Applications* (pp. 9-44). Cambridge: Cambridge University Press. doi:10.1017/CBO9781316402924.003
- [9] <https://www.britannica.com/science/fuzzy-logic#ref875418>
- [10] Zhang, Yang & Huang, Tao & Bompard, Ettore. (2018). Big data analytics in smart grids: a review. *Energy Informatics*. 1. 10.1186/s42162-018-0007-5.
- [11] <https://shelly.cloud/knowledge-base/devices/shelly-em/>
- [12] C. B. S. S. Wilhelm Kleiminger, «Household Occupancy Monitoring Using Electricity Meters», ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2015.
- [13] W. K. R. C. T. S. a. S. S. Christian Beckel, «The ECO Data Set and the Performance of Non-Intrusive Load Monitoring Algorithms», 1st AMC International Conference on Embedded System for Energy-Efficient Buildings, 2014.
- [14] A. Monacchi, D. Egarter, W. Elmenreich, S. D'Alessandro, A. M. Tonello. GREEND: An Energy Consumption Dataset of Households in Italy and Austria. In proceedings of the 5th IEEE International Conference on Smart Grid Communications (SmartGridComm 14 November 2014, Venice, Italy).
- [15] J. Zico Kolter and Matthew J. Johnson. REDD: A public data set for energy disaggregation research. In proceedings of the SustKDD workshop on Data Mining Applications in Sustainability, 2011.
- [16] P. B. D. B. M. H. H. C. M. W. R. S. Andreas Reinhardt, «On the Accuracy of Appliance Identification Based on Distributed Load Metering Data», 2nd IFIP Conference on Sustainable Internet and ICT for Sustainability, 2012.
- [17] Kelly, J., Knottenbelt, W. The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes. *Sci Data* 2, 150007 (2015).
- [18] <https://matlabacademy.mathworks.com/>, Deep Learning Onramp course.

- [19] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi University of Washington, Allen Institute for AI, Facebook AI Research. You Only Look Once: Unified, Real-Time Object Detection, CVPR 2016, 9 May 2016; <https://arxiv.org/pdf/1506.02640v5.pdf>
- [20] <https://it.mathworks.com/help/vision/ug/getting-started-with-r-cnn-fast-r-cnn-and-faster-r-cnn.html>
- [21] Ramzan, Farheen & Khan, Muhammad Usman & Rehmat, Asim & Iqbal, Sajid & Saba, Tanzila & Rehman, Amjad & Mehmood, Zahid. (2019). A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer's Disease Stages Using Resting-State fMRI and Residual Neural Networks. Journal of Medical Systems. 44. 10.1007/s10916-019-1475-2.
- [11] <https://www.pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/>
- [23] Kingma Diederik e Jimmy Ba, Adam: A method for stochastic optimization, arXiv:1412.6980v8 [cs.LG] 23 Jul 2015
- [24] fonte: <https://srdas.github.io/DLBook/GradientDescentTechniques.html>
- [25] <https://it.mathworks.com/help/vision/ref/trainyolov2objectdetector.html>
- [26] <https://ravedata.in/machine-learning/the-overfitting-problem/>
- [27] M. Pipattanasomporn, M. Kuzlu, S. Rahman e Y. Teklu, «Load Profiles of Selected Major Household Appliances and Their Demand Response Opportunities», IEEE, 2013.
- [28] Riga Technical University, Institute of Power Engineering, Department of Electric Power Supply, «Electrical Installation of residential buildings - Methodical Guidelines for Practical Works», RTU Press, 2019.
- [29] https://en.wikipedia.org/wiki/Precision_and_recall
- [30] Monitoraggio dei carichi non intrusivo (NILM) basato sulle reti neurali per l'analisi dei consumi elettrici nel settore residenziale, G. Zorzi, A.A. 2019/20
- [31] <https://ned.midorisrl.eu/>
- [32] "Non Intrusive Load Monitoring (NILM):A State of the Art", Jorge Revuelta Herrero, ´Alvaro Lozano Murciego,mAlberto Lopez Barriuso, Daniel Hernandez de la Iglesia,Gabriel Villarrubia Gonzalez, Juan Manuel Corchado Rodriguez,and Rita Carreira, Advances in Intelligent Systems and Computing, June 2018
- [33] R. Gopinath, Mukesh Kumar, C. Prakash Chandra Joshua, Kota Srinivas, Energy management using non-intrusive load monitoring techniques – State-of-the-art and future research directions, Sustainable Cities and Society, Volume 62, 2020, 102411, ISSN 2210-6707, <https://doi.org/10.1016/j.scs.2020.102411>.
- [34] <https://www.enea.it/it/internazionali/relazioni-unione-europea/old-30novembre2016/iniziative-strategiche/aal/ambient-assisted-living-aal/>
- [35] <http://www.aal-europe.eu/>