

1 Riassunto

1.1 Introduzione e motivazioni

L'obiettivo del lavoro di ricerca presentato nel paper è quello di progettare un sistema automatico, chiamato Tortoise@, in grado di riconoscere l'attività di scavo dei nidi da parte delle tartarughe marine e di comunicarne quindi la posizione. Per lo scopo, si utilizzano un dispositivo dotato di accelerometro e segnalatore GPS, applicato al carapace del singolo esemplare, e una rete neurale. Tramite l'accelerometro si raccolgono dati riguardanti il movimento della tartaruga, mentre la rete neurale è istruita con il compito di riconoscere l'attività di deposizione delle uova: il segnalatore GPS è utilizzato per comunicare la posizione del nido. Le uova possono così essere raccolte e protette dai ricercatori fino alla nascita dei nuovi esemplari.

I dati riguardanti le popolazioni di rettili e specie anfibe sono, infatti, allarmanti: si registrano diminuzioni degli esemplari rispettivamente del 34% e 97%. Nel caso specifico delle tartarughe la diminuzione della popolazione dovuta in larga parte all'inquinamento, responsabile dell'infertilità delle uova, e all'attività predatoria.

Il monitoraggio, tuttavia, non può essere svolto su larga scala poiché dipendente dalla diretta osservazione degli esemplari nel loro habitat naturale. L'obiettivo della ricerca è dunque quello di automatizzare il processo di individuazione dei nidi e facilitare questa attività fondamentale nell'ottica di una prevenzione efficace.

1.2 Soggetto ed aspetti etici

Il progetto è stato autorizzato dal Museo di Storia Naturale dell'Università di Pisa e dall'Unione dei Comuni Montani Colline Metallifere. Il lavoro di ricerca ed osservazione degli esemplari è stato svolto all'interno del Centro di Protezione per le Tartarughe Mediterranee di Massa Marittima, Toscana, Italia. Nel processo di ricerca sono stati coinvolti 100 esemplari sani di tartarughe femmina provenienti da 3 specie diverse (*Turtles hermanni*, *Turtles graeca*, *Turtles marginata*). L'applicazione del dispositivo non è invasiva e non causa dolore o morte dell'animale. Per questi motivi non è stato necessario consultare la Commissione Etica dell'Università di Pisa in merito alla procedura (come descritto da protocollo n 86/609/CEE).

1.3 Hardware e scelta dei sensori

Il sistema di rilevazione è composto da due dispositivi MicaZ e un PC standard. MicaZ è un dispositivo programmabile a basso consumo con interfaccia di comunicazione a wireless a 2.4 GHz (con antenna MPR2400) compatibile con lo standard IEEE 802.15.4. Il dispositivo è stato equipaggiato con la scheda MTS310 comprensiva sensori di temperatura, sensori di luce e due accelerometri (modello ADXL202). Uno dei dispositivi MicaZ raccoglie i dati provenienti dai sensori: i dati sul movimento sono utilizzati per riconoscere le diverse attività dell'animale, mentre quelli riguardanti luce e temperatura dell'acqua sono raccolti con lo scopo di dare informazioni di contesto riguardanti la rilevazione. Il secondo dispositivo viene collegato al PC in modo che si possano trasmettere i dati rilevati.

La frequenza di campionamento dei sensori è di 4Hz (una rilevazione ogni 250 ms), i dati sono salvati nella memoria locale del dispositivo. Gli accelerometri hanno sensibilità $\pm 2g$ e descrivono i movimenti dell'esemplare sull'asse x e sull'asse y (rispettivamente lato corto e lato lungo del carapace).

Nell'ottica di un utilizzo reale del dispositivo sarà necessario ridurre le dimensioni così da aumentare il comfort per l'animale.

1.4 Protocollo di raccolta dei dati

Come già anticipato, le rilevazioni avvengono in un habitat semi naturale e sono supervisionate da un operatore umano che ha il compito di segnalare 3 attività specifiche dell'esemplare: mangiare, camminare e scavare. La procedura di raccolta si divide in 3 fasi:

- **Posizionamento:** l'operatore applica il dispositivo sul carapace dell'esemplare in una posizione non invasiva (la scelta della posizione richiede consistenza, così da avere omogeneità sui dati raccolti).
- **Monitoraggio:** l'operatore osserva l'esemplare annotandone le attività in ordine cronologico così da poterle associare ai dati campionati.
- **Memorizzazione:** il dispositivo viene rimosso dal carapace e si scaricano i dati delle rilevazioni.

Importante evidenziare che non è stato possibile conoscere a priori la durata delle singole attività e dunque non è stato definito nessuno slot temporale a priori per le rilevazioni.

1.5 Analisi dei dati

I dati provenienti dai sensori sono raccolti in serie temporali separate. Ogni serie corrisponde ad un'attività tra mangiare, scavare o camminare.

1.5.1 Metodi di pre-processing

I dati raccolti vengono pre-processati riducendone il rumore e di normalizzando i valori in un range uniforme:

- **Filtraggio:** si applica un filtro mobile (appartenente alla tipologia Finite Impulse Response) con lo scopo di ridurre il rumore. Il filtro lavora facendo la media su un intervallo di 5 punti nella sequenza originale, per produrre ogni punto nella sequenza filtrata.
- **Normalizzazione:** i valori interi prodotti dall'accelerometro sono ridimensionati utilizzando la media di tutti i valori osservati. Questa non è una vera e propria normalizzazione. Lo scopo è quello di ottimizzare il consumo di memoria ottenendo uno stream di dati composto unicamente da interi.
- **Down-sampling:** vengono prodotte serie di valori con rate di campionamento differenti (da 4Hz a 1Hz). In questo modo si possono confrontare le prestazioni della rete neurale e cercare di raggiungere un trade-off tra occupazione della memoria e accuratezza del riconoscimento.

1.5.2 Input della rete neurale

Una sequenza dell'accelerometro è composta da una coppia di sequenze: una riguardante l'asse x l'altra l'asse y. Tuttavia grazie ai rilevamenti sull'asse x è stato possibile riconoscere l'attività di scavo. Infatti, l'alternanza delle zampe posteriori dell'animale durante lo scavo del nido produce una sequenza approssimante un'onda quadrata periodica. Per questo motivo il resto del paper si concentra sull'utilizzo di queste sequenze.

Il dataset è composto da 83 sequenze temporali divise come segue:

- 15 sequenze di camminata;
- 6 sequenze di nutrimento;

- 62 sequenze di scavo, 13 delle quali terminate con la deposizione di un uovo. Le restanti 49 non terminano con la deposizione (comportamento normale attribuibile agli esemplari in cattività).

La frequenza di sampling scelta di 1 Hz con un intervallo di un secondo (un time step). Ad ogni time step viene traslata una finestra (input window) di un intervallo pari ad un quarto della stessa, cos da ottenere nuovi valori in output. Ogni time step classificato dalla rete neurale analizzando i dati contenuti nella relativa input window. Questo procedimento noto come Input Delay Neural Network (IDNN).

Il periodo di tempo scelto per il pattern di 90 time step (secondi), corrispondente alla durata media di due onde quadrate durante l'attività di scavo. **La rete neurale ARS sar istruita con lo scopo di riconosce questo pattern caratteristico all'interno di sequenze pi lunghe ottenute durante la rilevazione.** Nell'ottica di avere un trade off tra occupazione della memoria e accuratezza degli output si deciso di utilizzare delle sotto-sequenze, estrapolate dai dati raccolti, lunghe 300 time step. Considerata la lunghezza del pattern (90 secondi) e la lunghezza della shift (un quarto di time step) la rete neurale produce in output 10 classificazioni per sotto-sequenza.

1.6 Tarta net activity recognition system

1.6.1 Modello di machine learning per TartaNet

Il modello di machine learning utilizzato Artificial Neural Network (ANN) nella sua ormai nota architettura ANNs/Multi-layer Perceptron (MLP). Un'architettura di questo tipo organizzata in un input layer, uno o pi hidden layer e un output layer. Il modello permette di applicare pesi diversi sugli archi che connettono le unit della rete, in modo da ottenere la migliore approssimazione degli output desiderati. Nel caso delle MLP si utilizza la tecnica della **backpropagation**: si basa sull'algoritmo di **discesa del gradiente** che viene ripetuto con lo scopo di minimizzare l'errore tra gli output della rete ed i valori attesi (modificando ad ogni iterazione i pesi).

1.6.2 Input Delay Neural Network

La rete neurale ha bisogno di rappresentare la relazione tra eventi nel tempo, senza per un particolare allineamento temporale. IL modello IDNN (Input

Delay Neural Network) risultato essere quello pi adeguato in relazione a questi requisiti. Si utilizza una shifting windows che, traslata sui segmenti della sequenza, nutre gli hidden layer. Una volta addestrato, il modello viene usato per leggere nuovamente i segmenti di sequenza nel tempo, cos da ottenere la propriet di traslation invariance (permette di identificare il pattern caratteristico a prescindere dalla sua posizione nella sequenza). Si utilizzano infine pesi diversi per ogni unit nascosta cos da nutrire gli output layer. BORDER BORDER

1.6.3 Local Receptive Field (LRF)

L'utilizzo di pesi diversi per ogni layer del modello IDNN pu portare ad un eccessivo consumo di memoria se relazionato alle capacit del dispositivo utilizzato per la ricerca. Per questo, il modello di base stato modificato facendo riferimento al modello IDNN MLP. Il primo accorgimento riguarda l'utilizzo di local receptive field applicato alla input windows del modello IDNN. Ogni neurone nascosto scansiona l'input utilizzando i propri local receptive field: in questo modo una unit dell'hidden layer riceve un inventario di feature con le quali rappresentare le caratteristiche di una sotto-finestra. Si identificano 4 sotto-sequenze all'interno del pattern caratteristico, due riguardanti la fase ascendente della forma d'onda quadrata, le altre la fase discendente. Si ottengono cos sotto-sequenze formate da 22 input.

Come risultato di questa divisione si ha un hidden layer composto da 4 unit, ognuna connessa ad una sotto-finestra mediante una diversa scelta dei pesi. Ogni unit nascosta processa 22 input e pertanto si ha bisogno di memorizzare 22 pesi per ognuna delle 4 unit nascoste (a differenza del modello IDNN classico che richiedeva la memorizzazione di 90 pesi per ogni unit nascosta). La seconda idea quella di condividere i pesi tra le unit nascoste forzando i vettori dei pesi ad essere uguali per le sotto-sequenze ascendenti e per le sotto-sequenze discendenti. In questo modo si riducono a 2 i nodi nascosti dimezzando lo spazio di memoria necessario per i relativi pesi.

1.6.4 Output Filter

A prescindere dal modello scelto, la rete neurale produce un output stream di valori compresi tra $[-1, 1]$ (valori positivi indicano segmenti simili al pattern caratteristico). Il filtro utilizza 3 diversi feature per classificare i segmenti.

- **Numero positivi (NP):** numero di valori positivi

- **Media positivi (MP):** media dei valori positivi su tutti i valori negativi
- **Mean positive threshold (MPT):** media dei valori positivi sopra la threshold τ
- La threshold τ calcolata come media dei valori positivi ottenuta con un set di validazione.
- Il filtro valuta ogni feature individualmente confrontandola con la soglia specifica con lo scopo di verificare se la risposta è positiva o negativa.

Il modello di rete neurale descritto produce un risultato per ogni segmento di input (ogni 300 secondi).

1.7 Risultati e discussione

In questa sezione sono confrontate le prestazioni dei modelli proposti (IDNN, IDNN LRF, IDNN LRF WS), con lo scopo di

1.7.1 Validazione del modello

Il dataset dei segmenti è stato diviso in 4 parti, ognuna di esse utilizzata per uno specifico scopo:

- **ANN training set:** segmenti utilizzati per generare i pattern positivi e negativi (134 pattern input bilanciati, 67 positivi e 67 negativi).
- **ANN validation set:** segmenti utilizzati per generare il validation set per i pattern, utilizzati per addestrare la rete (10 positivi e 10 negativi).
- **Output filter validation set:** Segmenti utilizzati per configurare le threshold del filtro (15 positivi e 15 negativi)
- **Test set:** segmenti di un dataset esterno, utilizzato per valutare le performance della rete (28 positivi e 28 negativi)

Tutti e tre i modelli di rete neurale proposti sono addestrati attraverso il training set. Le connessioni tra l'input layer e gli hidden layer sono inizializzate nel range $[-0.00001, 0.00001]$, mentre quelle tra hidden layer e output layer nell'intervallo $[-0.01, 0.01]$. Il validation set è utilizzato, invece, per

scegliere il modello pi performante sulla base dell'accuratezza delle classificazioni. Infine, l'output filter monitorato e quindi configurato valutandone il comportamento attraverso il relativo set di validazione. La fase di valutazione ci permette inoltre di scegliere la configurazione migliore rispetto al numero di hidden node, rate di apprendimento, il momentum e i parametri del weight decay.

1.7.2 Performance analysis

L'analisi delle performance valuta l'accuratezza della classificazione, misurata come percentuale sulla dimensione totale del dataset e della relativa matrice di confusione. I tre modelli proposti sono analizzati considerando l'errore medio, computato per 5 differenti inizializzazioni del vettore dei pesi. Il modello IDNN raggiunge il 96% di accuratezza nella fase di test, con un 4% dovuto a falsi positivi. Il modello IDNN LRF ha invece un'accuratezza del 95.5% ma con un considerevole risparmio in termini di memoria. Infine, il modello IDNN LRF WS raggiunge il 94% di accuratezza con errori dovuti anche in questo caso a falsi positivi (7%).

Importante specificare che la classificazione dei segmenti pu essere ripetuta pi volte (l'operazione di scavo del nido pu impiegare fino a 2 ore) e dunque ci ci permette di aumentare ulteriormente l'accuratezza dei modelli, evitando potenzialmente tutti i falsi positivi.

1.7.3 Analisi di applicabilit

In questa sezione si valuta l'applicabilit dei modelli proposti su dispositivi a basso consumo, come quelli selezionati per le rilevazioni. Ognuno dei tre modelli rappresenta 90 interi per la rappresentazione della input window. Per quanto riguarda i pesi, il modello IDNN utilizza 90 pesi float per ogni hidden unit, mentre gli altri due modelli hanno ne utilizzano 22. Ovviamente il modello IDNN quello che ha bisogno di un utilizzo della memoria maggiore (1844 byte). Per quanto riguarda IDNN LRF WS e IDNN LRF si ha che il primo mostra il consumo di memoria minore (196 byte) mentre il secondo con un consumo leggermente maggiore (388 byte) rappresenta il miglior compromesso in termini di memoria/prestazioni.

1.8 Il sistema Tortoise@

Come già specificato TartaNet fa parte di un sistema autonomo applicabile su larga scala che ha lo scopo di individuare i nidi di Tartaruga marina. Il sistema è dunque volto all'attività di biologging basata su sensori di movimento. Le attività svolte sono:

- **Monitoraggio dell'ambiente (fase 1):** monitora l'ambiente tramite sensori di temperatura e luce. In questo modo è possibile attivare la fase 2 solo nel caso di condizioni favorevoli alla deposizione delle uova.
- **Monitoraggio dei movimenti (fase 2):** raccolta dei dati riguardanti il movimento di scavo.
- **Monitoraggio dei movimenti esteso (fase 3):** prolunga il monitoraggio di cui al punto precedente su base periodica, con lo scopo di aumentare la precisione complessiva del riconoscimento.
- **Comunicazione dei dati (fase 4):** invia le coordinate del nido in remoto.

1.9 Conclusioni

La ricerca mostra che un riconoscimento dell'attività di deposizione delle uova da parte delle tartarughe marine è possibile utilizzando un singolo sensore di movimento in coppia con uno dei modelli di rete neurale proposti. L'obiettivo della ricerca è stato raggiunto implementando un modello efficace ed efficiente specializzato nella rilevazione dei nidi. I risultati ottenuti sono supportati da una campagna di raccolta dati sperimentale sul campo. I risultati sono promettenti e danno l'impulso a possibili ulteriori migliorie, rivolte soprattutto alla riduzione di peso, dimensione e capacità della batteria. Inoltre, si potrebbe pensare di investigare l'utilizzo di una convolutional neural network aggiungendo ulteriori hidden layer. Un'altra possibilità potrebbe essere quella di comparare il modello ottenuto con un approccio alternativo basato sulle trasformate di Fourier.

2 Commento

La ricerca si inserisce in un ampio contesto di studio protezione di specie rettili e anfibi, lo dimostra la ricca documentazione citata. L'articolo si pro-

pone di fornire un'applicazione funzionante nel contesto del biologgin, capace di dare una possibile soluzione al problema della rilevazione dei nidi di *Tartaruga Marina* nel Mediterraneo. L'articolo ben presentato e le informazioni proposte sono consistenti ed esaustive. Nell'articolo allegata una puntuale documentazione di supporto che riguarda le autorizzazioni necessarie allo svolgimento delle attività di ricerca ed un dataset completo. I dati sono raccolti in modo congruo al loro utilizzo e sono presentati in modo coerente. La replicabilità degli esperimenti, tuttavia resa difficile dalla mancanza del codice relativo alla rete neurale. Le conclusioni presentate in istanza finale sono ben supportate dai dati e lasciano la strada aperta ad ulteriori migliorie. Da questo punto di vista poteva essere indagato maggiormente il problema delle dimensioni del dispositivo, applicato al carapace degli esemplari. Il requisito di ridurre al minimo l'ingombro dell'apparecchio risulta a mio parere centrale e non trascurabile rispetto al contesto di applicazione e allo scopo che l'articolo si prefigge. L'articolo facilmente comprensibile e favorisce una lettura interessata e partecipe, sia per le tematiche che tratta sia dal punto di vista delle soluzioni proposte. Un'appunto che mi sento di fare riguarda le informazioni di background rispetto all'argomento reti neurali: le conoscenze di base sono date per scontate ed un lettore non introdotto all'argomento potrebbe trovare difficoltà a comprendere il funzionamento complessivo dei modelli di rete neurale. Un'altra piccola mancanza riguarda il pad adesivo utilizzato per il carapace: il modello non specificato ed il link di riferimento non più disponibile.

In conclusione ritengo che il lavoro proposto sia valido e che possa dare un grosso contributo all'attività di protezione della biodiversità che popola i nostri mari. Il numero delle organizzazioni sensibili a queste problematiche considerevole, ma manca ancora un sistema su larga scala che possa rendere ancora più efficace gli sforzi finora messi in campo.