



**Politecnico
di Torino**

POLITECNICO DI TORINO

Corso di Laurea in Ingegneria Aerospaziale
A.a. 2020/2021
Sessione di Laurea: Luglio 2021

Tecniche di Machine Learning per lo studio dei flussi turbolenti

Relatore

Prof.ssa Stefania Scarsoglio

Candidato

Francesco Verzobio

Alla mia famiglia

Indice

1	Introduzione	5
2	Contesto storico	7
3	Fondamenti di Machine Learning	10
3.1	Apprendimento Non Supervisionato	11
3.1.1	Riduzione della dimensionalità	12
3.1.2	Clustering spettrale	12
3.2	Apprendimento Semi-Supervisionato	13
3.2.1	Rete generativa avversaria (GAN)	13
3.3	Apprendimento Supervisionato	14
3.3.1	Rete neurale feed-forward	16
3.3.2	Rete neurale convoluzionale (CNN)	16
3.3.3	Rete neurale ricorrente	17
3.3.4	Long Short-Term Memory (LSTM)	18
3.3.5	Echo State Network (ESN)	19
4	Machine Learning per lo studio dei flussi turbolenti	20
4.1	Flussi turbolenti	20
4.1.1	Flussi analizzati	22
4.2	Principali applicazioni del Machine Learning ai flussi turbolenti	26
4.2.1	Apprendimenti Non Supervisionato e Semi-Supervisionato	27
4.2.2	Apprendimento Supervisionato	27
5	Apprendimenti Non Supervisionato e Semi-Supervisionato per i flussi turbolenti	29
5.1	Clustering Spettrale applicato alla convezione termica turbolenta tridimensionale	29
5.2	GAN per lo studio della turbolenza omogenea isotropa	30

6	Apprendimento Supervisionato: Reti neurali profonde per i flussi turbolenti	32
6.1	ESN per la previsione statistica di un flusso di taglio turbolento semplice	32
6.1.1	Confronto con i risultati del MLP e della LSTM	34
6.2	TBNN per la rappresentazione del tensore anisotropo degli sforzi di Reynolds	37
7	Apprendimento Supervisionato: Reti neurali convoluzionali per i flussi turbolenti	41
7.1	DCNN per lo studio della RBC turbolenta	41
7.2	CNN per la ricostruzione della turbolenza omogenea isotropa bidimensionale in decadimento	44
8	Conclusioni	49
	Bibliografia	51

Capitolo 1

Introduzione

I flussi turbolenti sono caratterizzati dalla presenza di fluttuazioni, sia spaziali che temporali, di vorticità, alti livelli di dissipazione e diffusività ed estrema non linearità del moto. Le fluttuazioni turbolente (spaziali e temporali) delle quantità del campo fluido rendono il moto turbolento tridimensionale e non stazionario. Questo anche se le condizioni al contorno sono stazionarie e bidimensionali. Le fluttuazioni sono irregolari, caotiche e non prevedibili [1].

In ambito ingegneristico è importante saper trattare i flussi turbolenti. Ad esempio, nella progettazione delle turbine eoliche la conoscenza della turbolenza nel flusso in ingresso e nello strato limite delle pale è importante per migliorare le prestazioni. Nei motori a combustione interna la forte turbolenza aumenta la miscelazione tra il carburante e l'aria, migliorando l'efficienza complessiva e riducendo le emissioni. Nella progettazione di aeroplani ritardare il verificarsi di turbolenze nello strato limite delle superfici alari porta ad una riduzione del consumo di carburante. Questi esempi, e un vasto numero di altre applicazioni, dimostrano quanto sia importante determinare l'effetto della turbolenza sulle prestazioni dei dispositivi ingegneristici e giustificano il continuo interesse nello sviluppo di tecniche per simulare e prevedere i flussi turbolenti [2].

Per la risoluzione delle equazioni che descrivono i flussi turbolenti è possibile utilizzare la simulazione numerica diretta (DNS). La DNS è il metodo che restituisce i risultati più accurati, ma ha un costo computazionale elevatissimo. Per questo motivo i ricercatori hanno iniziato ad usare modelli di turbolenza approssimati, come il RANS (Reynolds Averaged Navier-Stokes) e il LES (Large Eddy Simulation).

Non è sempre possibile ottenere risultati soddisfacenti con i metodi appena elencati. Per questo motivo i ricercatori hanno iniziato a sperimentare nuove tecniche. Negli ultimi anni, grazie alla disponibilità di una grande quantità di dati provenienti da simulazioni ed esperimenti e grazie alla crescita delle risorse computazionali, gli studiosi hanno cominciato ad usare modelli di turbolenza che si

basano sui dati, applicando quindi il Machine Learning (ML) ai flussi turbolenti. Il ML è una branca dell'Intelligenza Artificiale (AI) che si occupa dello sviluppo di algoritmi finalizzati all'apprendimento automatico mediante la statistica computazionale e l'ottimizzazione matematica [3]. Negli ultimi decenni l'intelligenza artificiale ha avuto una crescita esponenziale. Questo ha permesso il suo sviluppo anche in fluidodinamica. La meccanica dei fluidi si è tradizionalmente occupata di enormi quantità di dati ricavati da esperimenti, misurazioni e simulazioni numeriche. In fluidodinamica il ML offre una vasta gamma di tecniche di estrazione di informazioni dai dati, le quali vengono poi tradotte in conoscenze.

In questo lavoro inizialmente vengono presentati gli eventi storici che hanno portato all'incontro tra il Machine Learning e la fluidodinamica. Si passa poi alla spiegazione dei fondamenti dell'apprendimento automatico e in particolare vengono illustrate le tecniche di ML più utilizzate in fluidodinamica. Vengono quindi presentate le caratteristiche principali dei flussi turbolenti. Infine, sono mostrate diverse applicazioni del ML ai flussi turbolenti.

Capitolo 2

Contesto storico

All'inizio degli anni Quaranta, Kolmogorov, uno dei fondatori della teoria dell'apprendimento statistico, considerava la turbolenza come uno dei principali domini di applicazione del Machine Learning. I progressi del ML negli anni '50 e '60 sono stati caratterizzati da due sviluppi distinti. Da un lato è avvenuto lo sviluppo della cibernetica e dei sistemi esperti progettati per emulare il processo di pensiero del cervello umano. Dall'altro lato è avvenuto lo sviluppo di macchine, come il perceptron, mirate ad automatizzare processi come la classificazione e la regressione. Il perceptron (perceptrone) è stata la prima macchina di apprendimento inventata; era formato da una rete di unità decisionali binarie. L'uso dei perceptroni per la classificazione ha portato ad un'elevata eccitazione, che venne però smorzata dalla scoperta che le capacità di queste macchine erano limitate: i perceptroni a singolo strato erano in grado di apprendere solo funzioni separabili linearmente e non erano in grado di apprendere la funzione XOR¹. Era noto che i perceptroni multistrato potessero apprendere la funzione XOR, ma il loro sviluppo venne ostacolato dalle scarse risorse computazionali del tempo. Il ridotto interesse per i perceptroni è stato presto accompagnato da uno scarso interesse per l'intelligenza artificiale in generale.

Nei primi anni '60, un altro ramo del ML, strettamente correlato alla cibernetica, è stato sperimentato da Ingo Rechenberg e Hans-Paul Schwefel presso l'Università Tecnica di Berlino. Essi hanno eseguito esperimenti in galleria del vento su una struttura ondulata composta da cinque piastre collegate, con l'obiettivo di trovare gli angoli ottimali per ridurre la resistenza aerodinamica (vedi Figura 2.1). La svolta del loro lavoro avvenne con l'aggiunta di variazioni casuali agli angoli, dove la casualità è stata generata utilizzando una scheda Galton (un generatore di numeri analogici casuali). L'incremento o il decremento dell'angolo era auto-appreso

¹La disgiunzione esclusiva (XOR) è un operatore logico che restituisce VERO (V) se e solo se gli ingressi sono diversi tra di loro. Se gli ingressi sono uguali (V-V o F-F) restituisce FALSO (F).

dall'algoritmo attraverso il risultato negativo o positivo della precedente misura. Il lavoro prodotto da Ingo Rechenberg e Hans-Paul Schwefel ha ricevuto uno scarso riconoscimento all'interno della comunità della meccanica dei fluidi. Nonostante ciò, in un numero significativo di applicazioni in fluidodinamica vengono utilizzate idee che possono essere ricondotte al loro lavoro.

Un nuovo interesse per il potenziale dell'AI applicata alla fluidodinamica si è materializzato nei primi anni '80, quasi contemporaneamente ai primi sviluppi della fluidodinamica computazionale. L'attenzione venne focalizzata sui sistemi esperti per l'assistenza nella progettazione aerodinamica e nei processi di sviluppo.

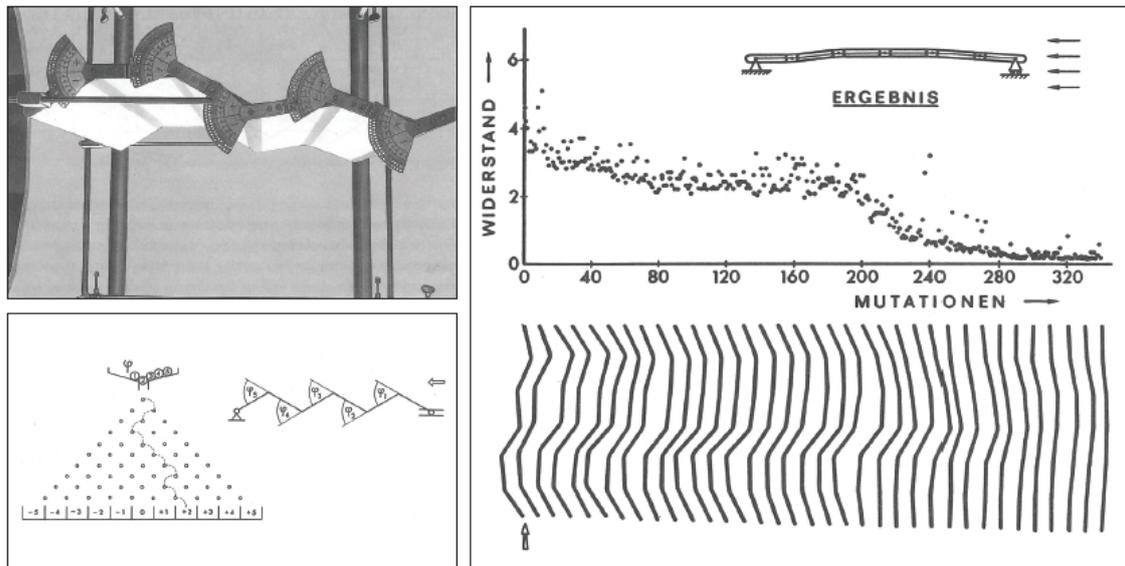


Figura 2.1: *Primo esempio di apprendimento automatico nella meccanica dei fluidi sperimentale: gli esperimenti di Rechenberg e Schwefel (1964) su lastre ondulate in modo ottimale al fine di ridurre la resistenza aerodinamica, utilizzando una scheda Galton come generatore di numeri analogici casuali [4].*

Un collegamento indiretto tra la meccanica dei fluidi e il ML fu il cosiddetto "Lighthill report" del 1974, il quale criticava i programmi di intelligenza artificiale nel Regno Unito, perché non soddisfacevano le loro grandi pretese. Questo report ha avuto un ruolo importante nella riduzione dei finanziamenti e dell'interesse per l'AI nel Regno Unito e successivamente negli Stati Uniti. L'argomentazione principale di Lighthill si basava sulla sua percezione che l'AI non sarebbe mai stata in grado di affrontare la sfida dell'esplosione combinatoria tra le possibili configurazioni nello spazio parametrico. Egli ha usato i limiti dei sistemi di elaborazione del

linguaggio di quel tempo come dimostrazione chiave dei fallimenti dell'intelligenza artificiale.

Un nuovo interesse per il ML, e in particolare per le reti neurali (NNs), avvenne alla fine degli anni '80 con lo sviluppo dell'algoritmo di backpropagation². Questo ha permesso il miglioramento delle NNs con più livelli, anche se inizialmente venivano usati al massimo due livelli. Altre fonti di stimolo furono i lavori di Hopfield (1982), Gardner (1988) e Hinton & Sejnowski (1986), che svilupparono collegamenti tra gli algoritmi del ML e la meccanica statistica. Tuttavia, questi sviluppi non hanno attratto molti ricercatori della meccanica dei fluidi.

All'inizio degli anni '90 sono state sviluppate numerose applicazioni delle reti neurali alla fluidodinamica. Alcuni esempi sono la Particle Tracking Velocimetry, una tecnica utilizzata per misurare le velocità e le traiettorie di particelle in movimento, e la Particle Image Velocimetry, un metodo ottico di misura del campo di moto di un fluido. Nello stesso periodo si iniziò ad usare le NNs per lo studio dei flussi multifase (flussi simultanei di materiali con due o più fasi termodinamiche). La Proper Orthogonal Decomposition (POD) consente di ridurre la complessità di simulazioni ad alta intensità di calcolo. Essa venne utilizzata, insieme alle NNs lineari, per ricostruire il flusso turbolento nell'intorno della parete di un canale, utilizzando solo le informazioni della parete. Questa applicazione è stata una delle prime ad utilizzare più strati di neuroni, segnando forse il primo utilizzo del Deep Learning, come è noto oggi, nel campo della meccanica dei fluidi.

Negli ultimi anni si è assistito ad un nuovo incremento delle applicazioni del ML alla meccanica dei fluidi. Questo grazie soprattutto al miglioramento delle prestazioni delle architetture del Deep Learning. L'applicazione delle tecniche di ML ai flussi può portare grossi benefici sia alla fluidodinamica stessa, che al Machine Learning [4].

²L'algoritmo di backpropagation confronta il valore di uscita del sistema in analisi con il valore desiderato (obiettivo). Sulla base della differenza così calcolata (errore), l'algoritmo modifica i pesi sinaptici della rete neurale, facendo convergere progressivamente il set dei valori di uscita verso quelli desiderati.

Capitolo 3

Fondamenti di Machine Learning

Il Machine Learning si occupa di algoritmi che apprendono da un insieme di dati e fanno delle predizioni su di essi, costruendo così in modo induttivo un modello basato su dei campioni.

I dati utilizzati per costruire il modello finale provengono da più dataset: training dataset, validation dataset, test dataset. Durante l'addestramento viene usato il training dataset, il quale serve al modello per imparare a creare delle relazioni tra i dati input e i dati output, se i dati di output sono noti. Se non sono noti, e quindi i dati input non sono etichettati, il modello riclassifica ed organizza i dati input. A questo punto si potrebbe incorrere nell'overfitting: il modello è in grado di utilizzare perfettamente i dati di addestramento, ma non è in grado di generalizzare su nuovi dati. Per evitare l'overfitting viene quindi utilizzato il validation dataset. Questo set di dati si chiama validation perché si occupa di validare i risultati ottenuti dal training dataset. Infine, una volta che il modello riesce ad avere delle buone performance sul training dataset, e soprattutto sul validation dataset, viene utilizzato il test dataset per testare il modello su nuovi dati.

Gli elementi fondamentali del ML sono un generatore di campioni (Sample generator), il sistema in questione (System) e una LM (Learning Machine), come si può vedere in Figura 3.1.

Le approssimazioni delle LMs sono fundamentalmente stocastiche e il loro processo di apprendimento può essere riassunto come la minimizzazione di un rischio funzionale:

$$R(w) = \int L[y, \phi(x, y, w)]p(x, y)dx dy,$$

dove i dati x (input) e y (output) sono i campioni di una distribuzione di probabilità p , $\phi(x, y, w)$ definisce la struttura e w i parametri della LM, la funzione di perdita L bilancia i vari obiettivi di apprendimento (come ad esempio: accuratezza, semplicità, ...). Il rischio funzionale è ponderato da una distribuzione di probabilità $p(x, y)$, che limita anche le capacità predittive della Learning Machine. Gli

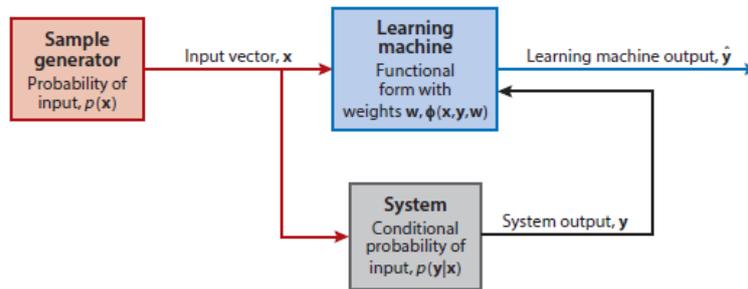


Figura 3.1: Il problema dell'apprendimento [4].

algoritmi di apprendimento possono essere raggruppati in tre categorie: apprendimenti supervisionato, non supervisionato e semi-supervisionato, come mostrato in Figura 3.2. Questa distinzione dipende dal fatto che i dati di addestramento siano o meno etichettati.

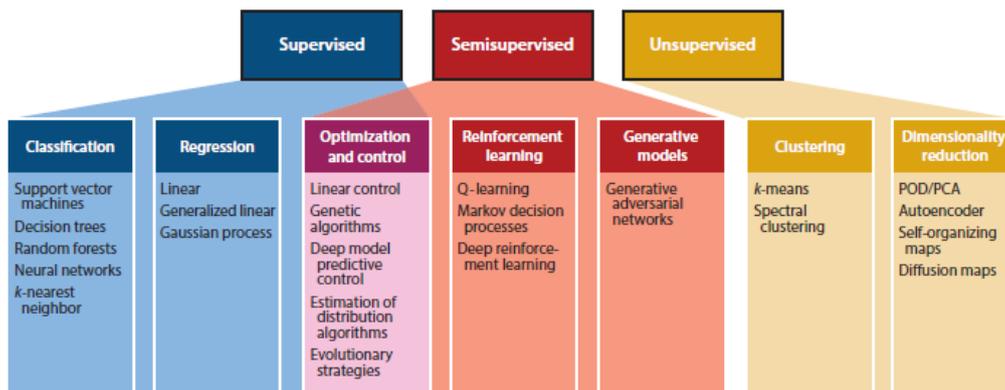


Figura 3.2: Gli algoritmi di apprendimento automatico possono essere classificati in supervisionati, non supervisionati e semi-supervisionati, a seconda dell'entità e del tipo di informazioni disponibili per il processo di apprendimento [4].

3.1 Apprendimento Non Supervisionato

Nell'apprendimento non supervisionato sono presenti dati senza etichetta (dati non strutturati). Con le tecniche di apprendimento non supervisionato è possibile osservare la struttura dei dati ed estrapolare informazioni cariche di significato.

Queste tecniche di ML non possono però contare su una variabile nota relativa al risultato o su una funzione di ricompensa. Le tecniche che fanno parte dell'apprendimento non supervisionato sono: riduzione della dimensionalità dei dati e clustering.

3.1.1 Riduzione della dimensionalità

L'identificazione di rappresentazioni a dimensione inferiore per dati ad alta dimensione può essere utilizzata come pre-elaborazione per tutte le attività degli algoritmi di apprendimento supervisionato. La riduzione della dimensionalità può essere vista anche come un collo di bottiglia per il filtraggio delle informazioni: i dati vengono elaborati attraverso una rappresentazione dimensionale inferiore, prima di essere mappati di nuovo alla loro dimensione originaria.

L'algoritmo POD (Proper Orthogonal Decomposition) appartiene a questa categoria di apprendimento. Il POD, chiamato anche Principal Component Analysis (PCA), è una rete neurale a due strati (un autoencoder) che applica una funzione di attivazione lineare ai suoi input linearmente pesati. L'autoencoder serve per comprimere e decomprimere dati di alta dimensione.

Gli algoritmi di apprendimento senza supervisione non sono mai stati largamente utilizzati in fluidodinamica. Negli ultimi anni però alcuni ricercatori hanno sviluppato numerosi autoencoder che, se adeguatamente associati all'ambito dei flussi, possono essere utili per la riduzione dell'ordine di dati sia stazionari, che dipendenti dal tempo.

Un metodo alternativo, rispetto all'autoencoder, per effettuare la riduzione della dimensionalità è quello di usare un encoder e un decoder. Le informazioni in input vengono filtrate attraverso un encoder. Successivamente, attraverso un decoder, vengono ricodificate solamente le informazioni utili. Questa riduzione della dimensionalità non è lineare e può essere svolta attraverso reti neurali profonde.

3.1.2 Clustering spettrale

Il clustering è una tecnica che consente di aggregare all'interno di gruppi (detti clusters) dei dati. Nei clusters sono quindi presenti dati che hanno alcune caratteristiche in comune. Il clustering è un'ottima tecnica che permette di scovare relazioni tra i dati.

Una tecnica di raggruppamento molto utilizzata è il clustering spettrale (vedi Figura 3.3). Le tecniche di raggruppamento spettrale fanno uso della matrice di similarità dei dati. Questa matrice è fornita come input. Essa rappresenta un grafo in cui ogni coppia di punti è collegata tramite un arco pesato. Punti distanti hanno un collegamento con un peso inferiore rispetto a quello di punti vicini. Il grafo viene suddiviso in più sottografi, in modo che la somma dei pesi dei bordi

che collegano i sottografi è la più bassa possibile e la somma dei pesi dei bordi presenti all'interno dei sottografi è la più alta possibile.

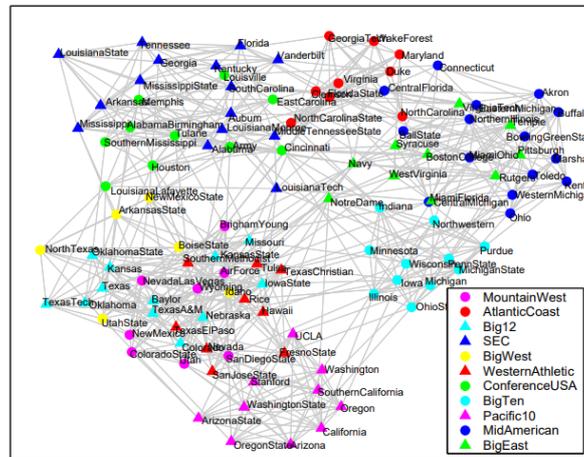


Figura 3.3: *Clustering spettrale applicato alle squadre di football americano del campionato NCAA Division I-A [5].*

3.2 Appendimento Semi-Supervisionato

L'apprendimento semi-supervisionato è una via di mezzo tra quello supervisionato e quello non supervisionato. Questo apprendimento combina, durante l'addestramento, una piccola quantità di dati etichettati con una grande quantità di dati senza etichetta. Due algoritmi di questa categoria sono: l'apprendimento per rinforzo (RL: Reinforcement Learning) e le Generative Adversarial Networks (GANs).

L'apprendimento per rinforzo viene utilizzato per la risoluzione di problemi in cui è presente un sistema (agente) che interagisce con il suo ambiente. In particolare, il sistema migliora le proprie performance grazie alle interazioni con l'ambiente. A differenza di ciò che accade nell'apprendimento supervisionato, l'agente non dispone di informazioni etichettate in riguardo alle azioni corrette, ma apprende dalle proprie esperienze tramite segnali di ricompensa (rinforzi).

Nel prossimo paragrafo sono illustrate le reti generative avversarie, frequentemente utilizzate in fluidodinamica.

3.2.1 Rete generativa avversaria (GAN)

Le reti generative avversarie (GANs) sono algoritmi di apprendimento in cui è presente un modello generativo, ossia un modello che produce dati secondo una

distribuzione di probabilità che imita quella dei dati utilizzati per il suo addestramento [4]. La Learning Machine è composta, oltre che dal modello generativo, da un discriminatore (vedi Figura 3.4). Lo scopo del generatore è quello di produrre nuovi dati, mentre il discriminatore apprende come distinguere i dati reali da quelli generati artificialmente. I pesi di queste reti sono ottenuti attraverso un processo, ispirato alla teoria dei giochi, chiamato apprendimento del contraddittorio. Il discriminatore mira a massimizzare la probabilità di discriminare tra dati reali e dati prodotti dal generatore, mentre il generatore mira a minimizzare la stessa probabilità. La funzione da minimizzare è la divergenza Kullback-Liebler¹ tra le due distribuzioni di probabilità. Le reti generative e discriminatorie si addestrano essenzialmente da sole: dopo l'inizializzazione con i dati di addestramento etichettati, avviene l'addestramento "autogestito". Questo processo di auto-formazione aumenta il fascino delle reti GANs. Bisogna però tenere in conto che non è sempre possibile raggiungere un equilibrio. Sicuramente una grande quantità di dati aiuta il processo, ma, al momento, non c'è alcuna garanzia di convergenza.

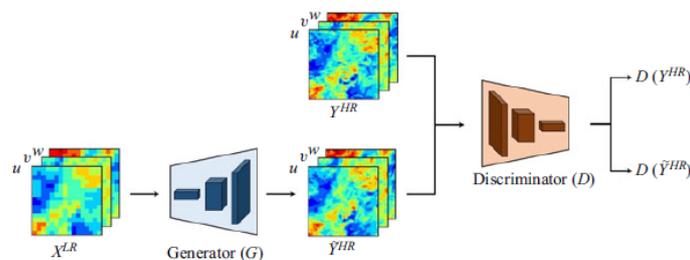


Figura 3.4: *Generative Adversarial Network* [6].

3.3 Apprendimento Supervisionato

Le tecniche di apprendimento supervisionato costruiscono un modello partendo da dati di addestramento etichettati; dopo l'addestramento il modello cerca di fare previsioni su dati non disponibili o futuri. Con supervisione si intende quindi che nell'insieme dei campioni di addestramento i segnali di output desiderati sono già noti [7].

La funzione di costo (o funzione di perdita) rappresenta la discrepanza tra le previsioni del modello che si sta addestrando e le istanze del problema reale. La

¹Nella teoria della probabilità la divergenza di Kullback-Leibler (anche detta divergenza di informazione, entropia relativa, o KLIC) è una misura non simmetrica della differenza tra due distribuzioni di probabilità.

minimizzazione della funzione di costo determina i parametri sconosciuti della LM. In questo contesto, l'apprendimento supervisionato risale ai metodi di regressione e interpolazione proposti secoli fa da Gauss. Una funzione di perdita comunemente impiegata è:

$$L[y, \phi(x, y, w)] = \|y - \phi(x, y, w)\|^2.$$

Funzioni di perdita alternative possono riflettere diversi vincoli sulla LM. Inoltre, la scelta tra metodi lineari e non lineari incide direttamente sul costo computazionale associato ai metodi di apprendimento [4].

Un primo metodo di apprendimento supervisionato è la classificazione. L'obiettivo di questa tecnica è quello di riuscire a prevedere l'etichettatura dei dati, sulla base dell'analisi di dati precedentemente etichettati. È probabilmente il metodo di apprendimento più antico, considerando che ne è un esempio il perceptron (Rosenblatt, 1958).

Un'altra tecnica di apprendimento supervisionato è la regressione. Essa consiste in un processo statistico che cerca di stabilire una relazione tra due o più variabili. Fornito ad un modello di regressione un valore x , esso restituirà il corrispondente valore y generato dall'elaborazione di x .

Le Reti Neurali (Neural Networks: NNs) sono i metodi più conosciuti ed utilizzati nell'apprendimento supervisionato. Sono approssimatori di funzioni fondamentali non lineari. Negli ultimi anni sono stati dedicati diversi sforzi per comprendere la loro efficacia. Il teorema di approssimazione universale (Hornik et al., 1989 [8]) afferma che qualsiasi funzione può essere approssimata da una rete sufficientemente ampia e profonda.

La potenza e la flessibilità delle NNs derivano dalla struttura modulare delle reti basata sul neurone come elemento centrale di costruzione. Nel cervello umano ogni neurone è collegato tramite l'assone ai dendriti di altri neuroni, con i quali è in grado di comunicare attraverso l'invio di un segnale elettrico che passa per le sinapsi. Un singolo neurone non ha grandi capacità e non è in grado di svolgere quasi alcun compito. Ma quando più neuroni vengono combinati assieme, si forma una rete neurale che riesce a svolgere compiti complessi e a raggiungere grandi risultati. Lo stesso vale nelle reti neurali artificiali. Ogni neurone riceve un input, lo elabora attraverso una funzione di attivazione e produce un output. La funzione di attivazione è applicata alla somma pesata dei valori in input. I pesi, che rappresentano il funzionamento delle sinapsi, svolgono un ruolo fondamentale: dicono al neurone quali segnali sono rilevanti e quali no. Nessun segnale in input deve sopraffare gli altri, pertanto si usa normalizzare o standardizzare il dataset.

Le reti neurali possono essere feed-forward o ricorrenti. Se all'interno della rete ci sono molti livelli, la rete viene definita profonda (si parla, in questo caso, di Deep Learning).

3.3.1 Rete neurale feed-forward

Nelle reti feed-forward ci sono neuroni organizzati in più strati. Ogni strato passa i segnali allo strato successivo. C'è uno strato di ingresso, che distribuisce i segnali di input ai neuroni del primo strato. Ci sono poi uno o più strati intermedi, detti strati nascosti, che effettuano una trasformazione generalmente non lineare dei dati. C'è infine uno strato di uscita, la cui funzione di attivazione caratterizza il compito realizzato dalla rete (vedi Figura 3.5)

I metodi di ottimizzazione non lineare, come la backpropagation, vengono utilizzati per identificare i pesi della rete, al fine di ridurre al minimo l'errore tra la previsione ottenuta e quella desiderata.

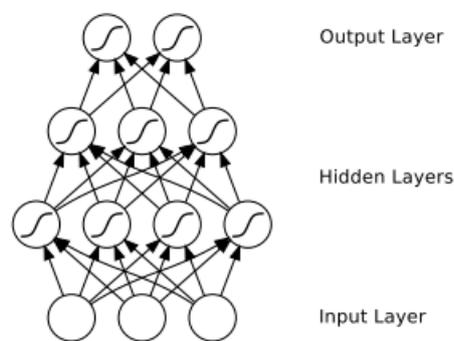


Figura 3.5: *Rete neurale feed-forward [9].*

Un esempio di rete neurale feed-forward è il multilayer perceptron (MLP), il quale è un'evoluzione del perceptron. Esso è formato da più strati di nodi che formano un grafo direzionato. Ogni strato è connesso al successivo. Trascurando i nodi in ingresso, ogni nodo è un neurone (elemento elaborante) con una funzione di attivazione non lineare. Il perceptrone multistrato usa la tecnica di backpropagation per l'addestramento della rete.

3.3.2 Rete neurale convoluzionale (CNN)

Un tipo particolare di rete neurale feed-forward è la rete neurale convoluzionale (CNN). In questa rete il modello di connettività tra i neuroni è ispirato dall'organizzazione della corteccia visiva animale, i cui neuroni individuali sono disposti in maniera tale da rispondere alle regioni di sovrapposizione che tassellano il campo visivo. Le applicazioni di queste reti sono nel riconoscimento di immagini e video, nei sistemi di raccomandazione, nell'elaborazione del linguaggio naturale e in bioinformatica.

Nel caso di analisi di immagini, la CNN applica più filtri alle immagini prima che esse vengano processate dalla rete neurale profonda.

Un esempio di filtro è il convolution, il quale passa su un'immagine, la elabora ed estrae le sue caratteristiche principali. L'applicazione del filtro consente di mettere in rilievo i dettagli delle immagini, al fine di identificare il loro significato.

Una seconda classe di filtri è definita pooling: i pixel dell'immagine sono raggruppati e filtrati in sottoinsiemi. Questo filtro permette di ridurre le dimensioni originali dell'immagine, mantenendo però le sue caratteristiche principali.

I filtri sono parametri i cui valori vengono appresi nel processo di apprendimento automatico. Quando l'immagine in analisi entra nel livello convoluzionale (convolutional layer) un numero definito a priori di filtri, inizializzati casualmente, viene applicato ad essa producendo un risultato, che passa poi al secondo livello della rete. Il processo si ripete finché la rete non individua i valori dei filtri tali per cui l'errore tra l'immagine in input e il risultato ottenuto è il minore possibile.

3.3.3 Rete neurale ricorrente

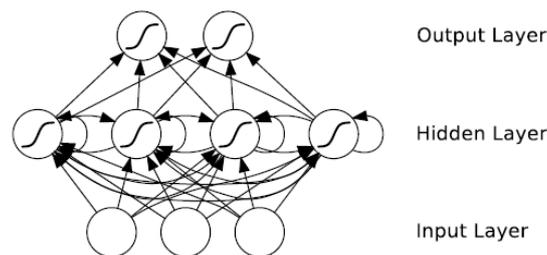


Figura 3.6: *Rete neurale ricorrente [9].*

Una classe di reti neurali molto utilizzata in fluidodinamica è rappresentata dalle reti neurali ricorrenti (RNNs; vedi Figura 3.6). In queste reti sono presenti neuroni collegati tra di loro in un loop (nelle reti feed-forward non sono invece presenti cicli). Tipicamente i valori di uscita di uno strato sono i valori di ingresso dello strato successivo. L'interconnessione tra gli strati permette l'utilizzo di uno dei livelli come memoria di stato e consente, avendo in ingresso una sequenza temporale di valori, di modellare un comportamento dinamico temporale dipendente dalle informazioni ricevute agli istanti di tempo precedenti. L'architettura della rete tiene conto dell'ordine intrinseco dei dati. Queste reti operano su sequenze di dati (come ad esempio immagini di un video, serie temporali, ecc.); i pesi della rete sono ottenuti tramite backpropagation nel tempo.

Le RNNs hanno avuto successo nell'elaborazione del linguaggio naturale e nel riconoscimento vocale. L'efficacia di queste reti è stata ostacolata dai problemi della scomparsa e dell'esplosione del gradiente², durante l'addestramento per backpropagation.

3.3.4 Long Short-Term Memory (LSTM)

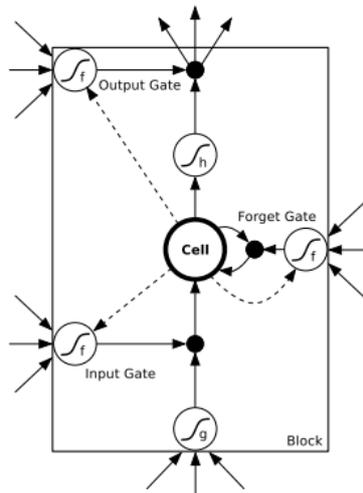


Figura 3.7: *Blocco di una LSTM con una cella [9].*

L'interesse per le RNNs è in gran parte attribuito allo sviluppo degli algoritmi di memoria a breve lungo termine (LSTM: Long Short-Term Memory).

La LSTM è costituita da un insieme di sotto-reti (blocchi di memoria) connesse in modo ricorrente. Questi blocchi hanno una struttura simile a quella dei chip di memoria di un computer. Ogni blocco contiene una o più celle di memoria auto-connesse e tre unità moltiplicative (le porte di input, output e forget), che attuano le operazioni di scrittura, lettura e azzeramento delle celle.

La Figura 3.7 fornisce un'illustrazione di un blocco di una LSTM con una singola cella. La rete LSTM è uguale alla RNN standard, tranne che per le unità

²Durante l'addestramento delle reti neurali profonde tramite retropropagazione dell'errore, ogni parametro del modello riceve ad ogni iterazione un aggiornamento proporzionale all'errore. Considerando che i gradienti delle funzioni di attivazione ai vari livelli vengono moltiplicati tramite la regola della catena, se si hanno funzioni di attivazione non lineari, il cui gradiente ha un valore compreso tra 0 e 1, il prodotto dei gradienti decresce esponenzialmente rispetto alla profondità della rete. Questo problema è chiamato scomparsa del gradiente. Quando invece il gradiente delle funzioni di attivazione assume valori elevati, il problema che si manifesta è quello dell'esplosione del gradiente.

di somma nello strato nascosto, che sono sostituite da blocchi di memoria. Le porte moltiplicative consentono alle celle di memoria di archiviare e accedere alle informazioni per lunghi periodi di tempo.

3.3.5 Echo State Network (ESN)

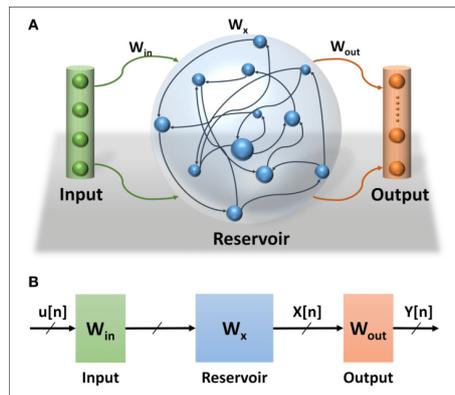


Figura 3.8: *Echo State Network* [10].

La Echo State Network è una rete neurale ricorrente formata da tre parti: input, reservoir (strati nascosti), output (vedi Figura 3.8). La connettività e i pesi dei neuroni nascosti sono fissi e assegnati in modo casuale. I pesi dei collegamenti tra i neuroni del reservoir e i neuroni di output sono appresi in modo che la rete possa produrre o riprodurre modelli temporali specifici.

Il funzionamento della rete ESN si divide in due momenti. Inizialmente il segnale di ingresso induce un segnale di risposta non lineare in ciascun neurone all'interno del reservoir. Successivamente viene generato il segnale di uscita tramite combinazione lineare dei segnali di risposta del reservoir.

L'interesse principale di questa rete è attribuito al fatto che sebbene il suo comportamento non sia lineare, gli unici pesi che vengono modificati durante l'addestramento sono quelli delle sinapsi che collegano i neuroni nascosti ai neuroni di output. Pertanto, la funzione di errore è quadratica rispetto al vettore del parametro e può essere facilmente differenziata in un sistema lineare.

Capitolo 4

Machine Learning per lo studio dei flussi turbolenti

4.1 Flussi turbolenti

Non esiste un'unica definizione di turbolenza. Nel corso dei secoli diversi scienziati hanno provato a darne una propria. Von Karman e Taylor, nel 1940, la definirono:

Turbulence is an irregular motion which in general makes its appearance in fluids, gaseous or liquid, when they flow past solid surfaces or even when neighboring streams of the same fluid flow past or over one another.

Il regime di un fluido può essere laminare o turbolento. Sia i flussi laminari, che i flussi turbolenti, sono soluzione delle equazioni di Navier-Stokes:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{V}) = 0, \quad (4.1)$$

$$\rho \left(\frac{\partial \vec{V}}{\partial t} + \vec{V} \cdot \nabla \vec{V} \right) = -\nabla p + \nabla \cdot [\mu(\nabla \vec{V} + \nabla \vec{V}^T) - \frac{2}{3}\mu(\nabla \cdot \vec{V})\bar{I}] + \rho \vec{f}, \quad (4.2)$$

$$\begin{aligned} \rho \left(\frac{\partial E}{\partial t} + \vec{V} \cdot \nabla E \right) = & -\nabla \cdot (p\vec{V}) + \nabla \cdot \left\{ [\mu(\nabla \vec{V} + \nabla \vec{V}^T) - \frac{2}{3}\mu(\nabla \cdot \vec{V})\bar{I}] \cdot \vec{V} \right\} + \\ & + \rho \vec{f} \cdot \vec{V} + \nabla \cdot (k\nabla T) + Q_v. \end{aligned} \quad (4.3)$$

Queste equazioni rappresentano le equazioni di bilancio della massa, della quantità di moto e dell'energia. Per alcuni semplici casi di flussi laminari è possibile risolvere analiticamente le equazioni di Navier-Stokes. Anche un flusso turbolento è soluzione di queste equazioni; il problema sta però nell'estrema complessità di questa ipotetica soluzione.

Il moto turbolento è caratterizzato dalla presenza di: fluttuazioni, sia spaziali che temporali, di vorticità; alti livelli di dissipazione e diffusività; estrema non linearità del moto. Le fluttuazioni turbolente (spaziali e temporali) delle quantità del campo fluido rendono il moto turbolento tridimensionale e non stazionario. Questo anche se le condizioni al contorno sono stazionarie e bidimensionali. Le fluttuazioni sono irregolari, caotiche e non prevedibili.

Il numero di Reynolds¹ misura l'importanza dei termini convettivi non lineari. Quando esso supera un certo valore critico avviene il processo di transizione: perturbazioni, inizialmente piccole, che evolvono linearmente nelle prime fasi della loro crescita, si sviluppano poi non linearmente. La turbolenza è la fase finale di questo processo non lineare di transizione.

La turbolenza è caratterizzata da fluttuazioni tridimensionali di vorticità. In una sezione del campo di moto turbolento si possono notare striature (streak), regioni con concentrazioni di sforzi (strain region), vortici (swirl) che si deformano, si uniscono, si dividono e ruotano per mutua induzione. Inoltre, si possono identificare regioni caratterizzate da particelle fluide dotate di vorticità, e quindi velocità angolare, denominate eddy.

Il moto turbolento è contraddistinto da un elevato livello di dissipazione. Inoltre è altamente diffusivo, in quanto la presenza di grandi fluttuazioni implica un tasso di mescolamento (diffusione) di quantità di moto, massa ed energia termica molto più alto rispetto al caso laminare, dove la diffusione è dovuta solo all'agitazione molecolare [1].

Un moto turbolento statisticamente stazionario può essere scomposto in un moto medio e nelle sue fluttuazioni attorno al valor medio, mediante la decomposizione di Reynolds. Attraverso questa decomposizione è possibile passare dalle equazioni di Navier-Stokes alle RANS (Reynolds Averaged Navier-Stokes equations). Nelle RANS le grandezze non sono più istantanee, ma mediate in un certo periodo di tempo, sufficientemente piccolo rispetto ai fenomeni che si vogliono studiare, sufficientemente grande rispetto ai disturbi della turbolenza. Nelle RANS compare il tensore degli sforzi di Reynolds, che rappresenta gli sforzi medi turbolenti dovuti agli effetti delle fluttuazioni turbolente sul campo di moto medio.

Per risolvere le equazioni di Navier-Stokes vengono usati modelli di turbolenza. Negli ultimi anni, grazie alla disponibilità di una grande quantità di dati, si è iniziato ad utilizzare modelli che si basano sui dati, e quindi tecniche di Machine Learning, per il calcolo dei flussi turbolenti.

¹Il numero di Reynolds è così definito: $Re = \frac{\rho \cdot V \cdot L}{\mu}$, dove ρ è la densità, V è la velocità del flusso, L è la lunghezza caratteristica e μ è la viscosità dinamica.

4.1.1 Flussi analizzati

In questo paragrafo vengono presentati i flussi che verranno trattati nei prossimi capitoli, ossia:

- flusso turbolento omogeneo e isotropo;
- convezione di Rayleigh-Bénard turbolenta;
- flusso di taglio turbolento tra due pareti.

Flusso turbolento omogeneo e isotropo

Un flusso turbolento omogeneo e isotropo è un caso ideale di flusso turbolento in cui la dinamica del moto del fluido non è influenzata dall'interazione del fluido con pareti solide o da eventuali flussi medi di velocità [11]. Questo flusso si può verificare solo nello spazio libero. In queste condizioni, infatti, il moto caotico del fluido può svilupparsi liberamente secondo la dinamica imposta dalle equazioni del moto. Il fluido è omogeneo e isotropo in quanto, data l'assenza di condizioni al contorno o di flussi medi imposti esternamente, la struttura del campo di velocità, in termini di grandezze statistiche, è invariante per traslazione (omogeneità) e rotazione (isotropia) del sistema di riferimento rispetto al quale si sta descrivendo il moto del fluido.

Le condizioni di turbolenza omogenea ed isotropa sono ideali e quindi raramente si verificano. Tuttavia, proprio grazie a tali condizioni è possibile utilizzare un maggior numero di strumenti sia per la simulazione che per l'analisi dei campi turbolenti di questo tipo. Un flusso omogeneo e isotropo è il candidato ideale per l'analisi del comportamento intrinseco della dinamica della turbolenza da un punto di vista fisico, in quanto tutti i risultati che si ottengono sono riconducibili esclusivamente alla dinamica non lineare del fluido.

La turbolenza omogenea ed isotropa può svilupparsi solo in condizioni di spazio libero. Si possono quindi risolvere le equazioni del moto in un dominio V , che rappresenta una porzione più piccola dello spazio libero dove la turbolenza si sviluppa. Se il dominio V è sufficientemente grande può accadere che il moto delle particelle di fluido che entrano nel dominio, a causa della natura caotica delle equazioni del moto, sia statisticamente scorrelato dal campo di moto che queste hanno quando abbandonano il volume V . Si può quindi assumere che le condizioni di ingresso siano uguali a quelle di uscita. Le condizioni al contorno naturali sono pertanto delle condizioni al contorno periodiche. L'unico vincolo che si deve rispettare è quello sulle dimensioni del dominio V . Il dominio V deve avere delle dimensioni caratteristiche maggiori della lunghezza di correlazione del campo di velocità, che, in prima approssimazione, si può considerare pari alla scala integrale.

Nella turbolenza omogenea ed isotropa si deve considerare anche un termine forzante, perché, a causa della cascata inerziale, l'energia viene trasmessa alle piccole scale e di conseguenza si osserva il fenomeno della dissipazione, che appare come una funzione decrescente nel tempo. La forzante agisce solo sulla scala più grande del sistema e la rifornisce dell'energia cinetica perduta a causa della dissipazione viscosa (vedi Figura 4.1). Senza una forzante non sarebbe possibile avere

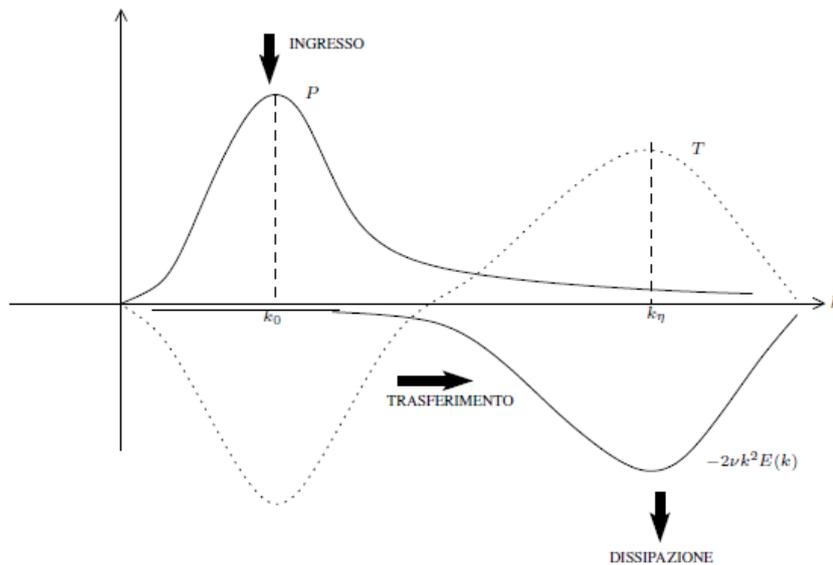


Figura 4.1: *Flusso di energia nella turbolenza omogenea isotropa. La curva P rappresenta il contributo dovuto al forzamento esterno; la curva $-2\nu k^2 E(k)$ rappresenta la dissipazione dell'energia cinetica dovuta alla viscosità. La funzione $T(k)$ è nota come trasferimento di energia. Ci sono tre intervalli di scale: un primo intervallo dove l'energia viene immessa nel sistema; un secondo intervallo dove l'energia viene dissipata; un terzo intervallo dove l'energia viene solo trasferita [11].*

un campo turbolento statisticamente stazionario perché l'energia andrebbe a zero per tempi lunghi. La produzione di energia è concentrata alle grandi scale mentre la dissipazione alle piccole. Questo tipo di comportamento fa sì che nell'intervallo inerziale² la dinamica della turbolenza non risenta del meccanismo di forzamento esterno e di dissipazione viscosa.

Nell'intervallo inerziale di un flusso stazionario omogeneo e isotropo il flusso di energia attraverso la scala è costante e pari al valore medio della dissipazione viscosa. Questo è dovuto al fatto che nelle scale intermedie non vi è né produzione

²Nell'intervallo inerziale sia la produzione che la dissipazione di energia sono assenti. In questo intervallo è presente solo il trasferimento di energia dovuto alle interazioni non lineari.

né dissipazione di energia; la prima si ha a livello della scala integrale mentre la seconda a livello della scala di Kolmogorov.

Convezione di Rayleigh-Bénard turbolenta

La convezione di Rayleigh-Bénard è un tipo di convezione naturale che si verifica in uno strato orizzontale piano di fluido riscaldato dal basso. Per numeri di Rayleigh³ poco superiori a 1700 si sviluppano le celle di Bénard: i moti ascendenti e discendenti si organizzano secondo uno schema ordinato e regolare costituito da corpi fluidi cilindrici controrotanti. Ciascuna cella convettiva è formata dallo strato ascendente, da quello discendente e da due correnti orizzontali di segno opposto che chiudono il circuito in alto e in basso. La struttura è bidimensionale. Accanto al fenomeno cinematico se ne rileva uno termico, consistente in un notevole aumento del flusso di calore trasmesso attraverso il fluido dalla parete inferiore a quella superiore [12].

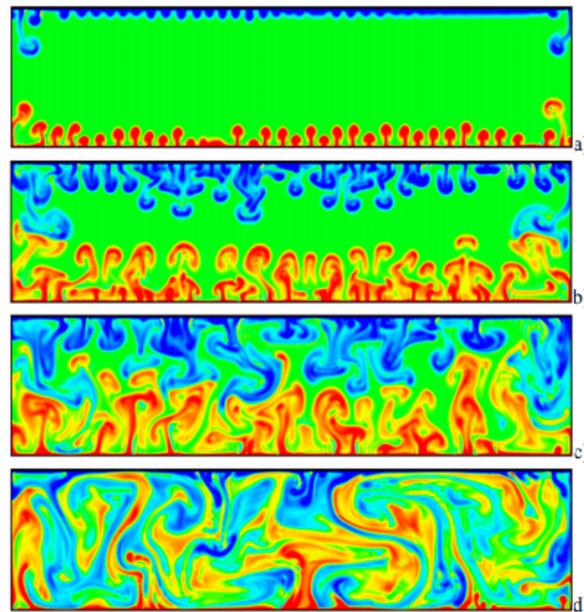


Figura 4.2: Istantanee della convezione di Rayleigh-Bénard turbolenta a diversi tempi (fluido newtoniano, $Pr = 15$, $Ra = 10^8$) [13].

³Il numero di Rayleigh è così definito: $Ra = \frac{g \cdot \alpha \cdot (T_i - T_s) \cdot d^3}{\nu \cdot k}$, dove g è l'accelerazione di gravità, α è il coefficiente di dilatazione termica, $T_i - T_s$ è la differenza di temperatura tra la parete inferiore e quella superiore, d è la distanza tra le pareti, ν e k sono i coefficienti di diffusione viscosa e termica.

Al crescere di Ra il campo di moto passa attraverso configurazioni successive. Dapprima ($Ra \approx 10^4 - 10^5$) appare la transizione ad una struttura completamente tridimensionale: le celle convettive perdono la forma cilindrica riorganizzandosi secondo uno schema regolare formato da punti di risalita verticale circondati da flusso in discesa.

Ulteriori aumenti di Ra producono il collasso anche di questa configurazione ed il graduale passaggio a moti sempre più intensi e disordinati (vedi Figura 4.2), nei quali si colloca infine la convezione turbolenta ($Ra \approx 10^6 - 10^8$). A causa del continuo rimescolamento di fluido il profilo medio di temperatura tra le pareti assume un andamento costante, completamente diverso da quello lineare che caratterizza il caso $Ra < 1700$, nel quale il flusso è fermo (vedi Figura 4.3).

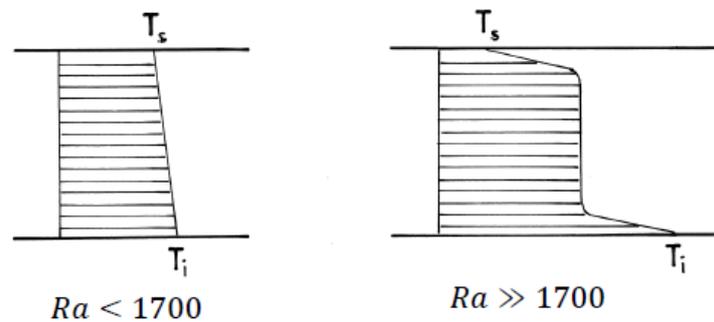


Figura 4.3: *Profili di temperatura per $Ra < 1700$ e per Ra molto elevato [12].*

Flusso di taglio turbolento tra due pareti

Tra due pareti a scorrimento libero si può instaurare un flusso di taglio, che può essere modellato tramite l'analisi di Fourier. Utilizzando il modello di Galerkin a 9 modi [14], il campo di velocità può essere scritto come:

$$u(x, t) = \sum_{k=1}^9 a_k(t) u_k(x),$$

dove i modi $u_k(x)$ indicano la dipendenza spaziale degli elementi costitutivi del flusso vicino alla parete e i 9 coefficienti reali di espansione $a_k(t)$ governano la loro dinamica [15]. Il modello, noto anche come unità minima di un flusso di taglio turbolento, rappresenta un ciclo di formazione della struttura che si suppone sia presente in ogni flusso di taglio: l'amplificazione non normale e la crescita transitoria delle perturbazioni nel profilo medio sotto forma di vortici streamwise,

la formazione di striature (streaks) streamwise e la loro instabilità e rottura, e la riformazione di vortici streamwise da interazioni non lineari. I campi di velocità istantanei possono essere ottenuti mediante la proiezione del modello di Galerkin in un dominio $L_x : L_y : L_z = 4\pi H : 2H : 2\pi H$ (vedi Figura 4.4), con condizioni al contorno periodiche nelle direzioni streamwise e spanwise (x e z); le condizioni al contorno a scorrimento libero sono applicate nella direzione normale alla parete (y).

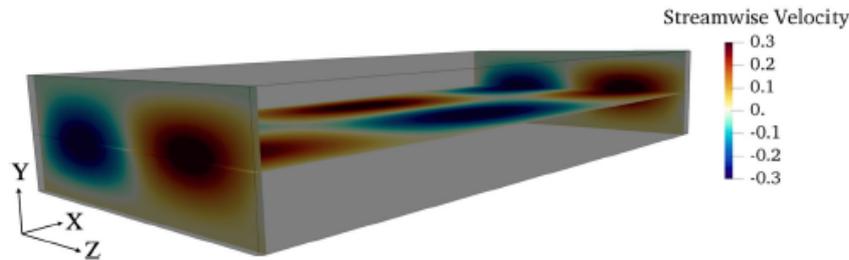


Figura 4.4: Istantanea con iso-contorni delle fluttuazioni della velocità del flusso, le quali mettono in evidenza le streaks a bassa e ad alta velocità. Il numero di Reynolds del flusso è $Re = 400$ e la distanza tra le pareti è $2H$ [15].

4.2 Principali applicazioni del Machine Learning ai flussi turbolenti

Lo studio della turbolenza consiste in gran parte nel rilevare schemi e correlazioni tra dati disponibili. I modelli di turbolenza possono essere geometrici (strutture coerenti su larga scala) o statistici.

In fluidodinamica non è solo importante avere una grande quantità di dati, ma è importante anche saperli organizzare, classificare e ridurre. Le conoscenze scientifiche non possono essere ottenute da enormi quantità di dati, che non sono stati prima analizzati.

I moderni sistemi informatici possiedono la velocità, la potenza e la flessibilità necessarie per accedere rapidamente a enormi quantità di dati, creando però problemi come l'archiviazione, il recupero e la post-elaborazione dei dati. Anche nello studio dei flussi turbolenti ci sono questi problemi, che possono però essere risolti grazie al Machine Learning [15].

4.2.1 Apprendimenti Non Supervisionato e Semi-Supervisionato

Gli algoritmi di apprendimento non supervisionato e semi-supervisionato possono facilitare nuove scoperte tramite l'analisi della struttura di set di dati esistenti.

Per quanto riguarda l'apprendimento non supervisionato, tramite il clustering spettrale è possibile analizzare le sovrastrutture turbolente di convezione, come è stato fatto in Schneide et al. [16].

La tecnica di apprendimento semi-supervisionato più utilizzata in fluidodinamica è la GAN. Nei flussi turbolenti si verifica una complessa sovrapposizione di strutture. Le tecniche di ML sviluppate per le immagini possono quindi essere utilizzate per apprendere la fisica dei flussi turbolenti. È stato dimostrato che le reti GANs possono essere addestrate per generare immagini sintetiche di alta qualità dei flussi turbolenti. In King et al. [17] è stata utilizzata una GAN per lo studio della turbolenza omogenea isotropa. In Kim et al. [18] è presentato un modello RNN-GAN in grado di generare una condizione al contorno di ingresso necessaria per lo sviluppo di una simulazione di un flusso turbolento in un canale.

L'aspetto negativo delle tecniche di apprendimento non supervisionato e semi-supervisionato è che sono metodi statici, che ignorano l'evoluzione dinamica dei flussi. Molto spesso in fluidodinamica si usa quindi abbinare a queste tecniche i metodi di ML supervisionati.

4.2.2 Apprendimento Supervisionato

Il ML supervisionato viene utilizzato per scoprire le relazioni tra un insieme di misurazioni e le variabili target. Questo tipo di apprendimento viene usato maggiormente, rispetto a quello non supervisionato e a quello semi-supervisionato, in fluidodinamica, perché può descrivere relazioni non lineari molto complesse tra gli input e gli output [15].

Il ML supervisionato si basa sul fatto che spesso è più facile addestrare un sistema con una serie di esempi etichettati di un comportamento input-output previsto, piuttosto che sviluppare un algoritmo per fornire la risposta corretta di tutti i possibili problemi.

I flussi turbolenti generalmente mostrano una fisica multi-scala (spaziale e temporale). Inoltre mostrano anche strutture intermittenti rotazionali e traslazionali. Queste particolari strutture possono essere analizzate tramite reti neurali profonde [19].

In letteratura si possono trovare diverse applicazioni del ML supervisionato ai flussi turbolenti. In Pandey et al. [15] è utilizzato il modello ESN per la previsione statistica di un flusso di taglio turbolento semplice. In Ling et al. [20] è proposto l'utilizzo di una rete neurale profonda, chiamata TBNN (Tensor Basis Neural Network), per la rappresentazione del tensore anisotropo degli sforzi di Reynolds.

Anche le reti neurali convoluzionali vengono spesso utilizzate dai ricercatori per lo studio dei flussi turbolenti. In Fonda et al. [21] una rete CNN profonda è stata utilizzata al fine di identificare il ruolo delle sovrastrutture turbolente nel trasporto del calore nella convezione di Rayleigh-Bénard turbolenta. Fukami et al. [22] hanno utilizzato una rete neurale convoluzionale per la ricostruzione di dati di flussi turbolenti ad alta risoluzione, partendo da dati spaziali e temporali a bassa risoluzione.

In Kobayashi et al. [23] sono stati utilizzati i modelli di apprendimento MLP (Multilayers Perceptron) e LSTM (Long Short-Term Memory) per prevedere la riduzione della resistenza causata dai flussi pulsatili in tubi (pulsating pipe flows). Nakamura et al. [24] hanno utilizzato il modello di apprendimento automatico definito ML-ROM, per studiare il flusso turbolento in un canale.

Sono quindi molte le tecniche di ML applicabili ai flussi turbolenti. È importante considerare diversi fattori quando si sceglie il metodo da applicare, tra cui la qualità e la quantità dei dati, gli input e gli output desiderati e la funzione di costo da ottimizzare. Inoltre, è importante sviluppare e adattare algoritmi di ML che non siano solo informati sulla fisica, ma anche coerenti con la fisica dei flussi.

Capitolo 5

Apprendimenti Non Supervisionato e Semi-Supervisionato per i flussi turbolenti

5.1 Clustering Spettrale applicato alla convezione termica turbolenta tridimensionale

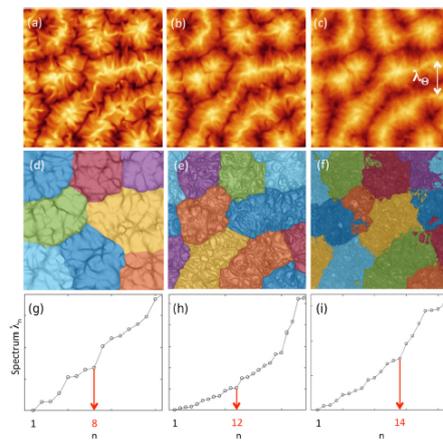


Figura 5.1: *Evoluzione temporale delle sovrastrutture lagrangiane turbolente di convezione. (a)-(c) Temperatura media nel tempo misurata nel piano medio. (d)-(f) Clusters delle traiettorie lagrangiane ottenuti tramite clustering spettrale. Le particelle che, nel tempo considerato, appartengono allo stesso cluster sono colorate con lo stesso colore. (g)-(i) Rappresentazione degli autovalori della matrice laplaciana [16].*

Schneide et al. [16] hanno utilizzato un metodo di apprendimento non supervisionato per lo studio della convezione di Rayleigh-Bénard turbolenta. In particolare, essi hanno applicato il clustering spettrale alla convezione termica turbolenta tridimensionale su un dominio di larga scala, al fine di analizzare l'evoluzione temporale di un insieme di particelle lagrangiane. Le singole traiettorie sono state tradotte in una rete ed è stata studiata la loro connessione con le sovrastrutture turbolente di convezione.

I ricercatori hanno rappresentato l'insieme delle particelle lagrangiane, in un determinato istante di tempo, attraverso un grafo non orientato. I vertici del grafo sono stati uniti tramite collegamenti. Per rappresentare l'evoluzione dinamica dell'insieme di particelle lagrangiane in un intervallo di tempo, sono stati assegnati dei pesi ai collegamenti. I pesi sono stati calcolati facendo l'inverso della distanza mediata nel tempo delle reciproche traiettorie e sono stati impostati a zero se la distanza superava una certa soglia. Dopo aver pesato la rete, è stata fatta la sparsificazione del grafo. Risolvendo il problema agli autovalori della matrice laplaciana del grafo, la rete è stata scomposta in k sottografi o clusters (vedi Figura 5.1). I clusters sono stati confrontati con i modelli di flusso convettivo su larga scala ottenuti nel sistema di riferimento euleriano. I ricercatori sono arrivati alla conclusione che le scale caratteristiche delle sovrastrutture nei sistemi di riferimento euleriano e lagrangiano concordano.

5.2 GAN per lo studio della turbolenza omogenea isotropa

I recenti progressi nel Deep Learning si sono dimostrati straordinariamente efficaci su problemi basati sulle immagini. Gran parte del successo delle tecniche di DL è dovuto alla loro capacità di identificare gerarchicamente e astrarre delle caratteristiche presenti nelle immagini.

Una complessa sovrapposizione di strutture si verifica anche nei flussi turbolenti. Per questo motivo, le tecniche sviluppate per le immagini vengono utilizzate anche per apprendere la fisica dei flussi turbolenti. È stato dimostrato che le reti GANs possono essere addestrate per generare immagini sintetiche di alta qualità dei flussi turbolenti.

King et al. [17] hanno implementato una GAN Convolutionale (C-GAN) sia per il generatore che per il discriminatore, al fine di studiare la turbolenza omogenea isotropa. I ricercatori hanno inserito nel loro modello sezioni bidimensionali di dati tridimensionali presi dal database di turbolenza Johns-Hopkins. I risultati, riassunti in Figura 5.2, suggeriscono che il generatore ha imparato con successo a campionare dalle funzioni di probabilità di distribuzione (PDFs: Probability

Distribution Functions) complete a più punti di campi di flusso turbolento. Il metodo che è stato utilizzato è statico e quindi ignora tutta l'evoluzione dinamica del flusso turbolento. È possibile migliorare l'analisi utilizzando algoritmi di ML supervisionati che incorporano informazioni dinamiche.

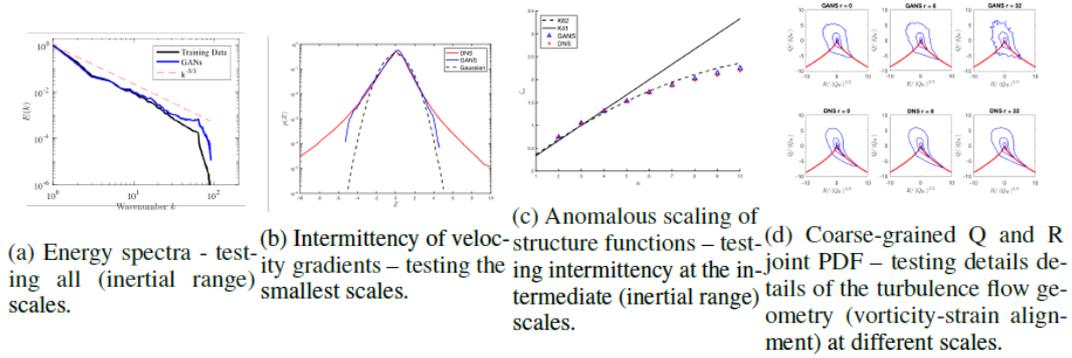


Figura 5.2: *Tecnica GAN applicata ad un flusso omogeneo isotropo turbolento. Nelle immagini (b) e (c) il rosso rappresenta i risultati della DNS, il blu rappresenta i risultati del metodo GAN [17].*

Capitolo 6

Apprendimento Supervisionato: Reti neurali profonde per i flussi turbolenti

6.1 ESN per la previsione statistica di un flusso di taglio turbolento semplice

Il flusso studiato è il flusso di taglio semplice descritto dal modello di Galerkin a 9 modi [14]. Il campo di velocità può essere scritto come:

$$u(x, t) = \sum_{k=1}^9 a_k(t) u_k(x),$$

dove i modi $u_k(x)$ indicano la dipendenza spaziale degli elementi costitutivi del flusso vicino alla parete e i 9 coefficienti reali di espansione $a_k(t)$ governano la loro dinamica. Il modello, noto anche come unità minima di un flusso di taglio turbolento, rappresenta un ciclo di formazione della struttura che si suppone sia presente in ogni flusso di taglio. Piuttosto che integrare il modello di Galerkin non lineare, che può essere scritto, per un dato numero di Reynolds, come

$$\frac{da_k}{dt} = F_k(a_m, t; Re),$$

con $Re = \frac{U_0 d}{2\nu}$ e $k, m = 1, \dots, 9$, è possibile risolvere il problema utilizzando una rete neurale [15].

Il dominio utilizzato nello studio è: $L_x : L_y : L_z = 4\pi H : 2H : 2\pi H$. Le direzioni streamwise e spanwise sono rappresentate dagli assi x e z rispettivamente; la direzione normale alla parete è l'asse y .

In questo paragrafo è sviluppato il modello ESN (Echo State Network); i suoi risultati sono poi confrontati con quelli ottenuti dalle tecniche MLP (Multilayer Perceptron) e LSTM (Long Short-Term Memory) [25].

Per l'addestramento delle tre diverse reti neurali sono state utilizzate 100 serie temporali, con 4000 fasi temporali ciascuna, e quindi un totale di 400.000 punti univoci. Successivamente, le reti addestrate sono state testate su una serie temporale completamente nuova, al fine di valutare la generalità e la robustezza dei modelli.

Parameter	MLP	LSTM	ESN
Input dimension	4500	90	9
Output dimension	9	9	9
Hidden layers	4	1	–
Neurons in hidden layers	90	90	–
Reservoir nodes	–	–	600
Activation function	Hyperbolic tangent	Hyperbolic tangent	Hyperbolic tangent
Optimization Algorithm	Adam [40]	Adam [40]	Ref. [39]
Counter overfitting strategy	Early stopping	Early stopping	Ridge regression

Figura 6.1: Parametri dei tre modelli (MLP, LSTM, ESN). Per quanto riguarda la tecnica ESN è stato scelto un raggio spettrale $\rho(A)=1.0$ e una densità dei nodi attivi nel reservoir di $D = 0.2$ [15].

Similmente alle reti ricorrenti, le reti ESN hanno uno strato di input e uno di output, che in questo specifico caso sono i vettori di dimensione 9 $a_{in}(t)$ e $a_{out}(t)$ rispettivamente. Lo strato di input è connesso ad un vettore di stato nel reservoir da $r(t) = W_{in}a_{in}(t)$, dove W_{in} è una matrice $N \times 9$, pesata ed inizializzata casualmente (con $N \gg 9$). Una tipica dinamica del reservoir durante la fase di addestramento è:

$$r(t + \Delta t) = (1 - \alpha)r(t) + \alpha \tanh(Ar(t) + W_{in}a_{in}(t) + \xi \mathbf{1}).$$

A rappresenta il reservoir ed è una matrice sparsa, $N \times N$ ed inizializzata casualmente; α rappresenta il tasso di perdita; $\xi \mathbf{1}$ rappresenta un vettore di polarizzazione contenente una costante scelta ξ in ciascun componente. Mentre le matrici W_{in} e A rimangono fisse nella procedura di addestramento, la matrice pesata degli output $9 \times N$, che è definita tramite $a_{out}(t) = W_{out}r(t) + c$, con bias c , viene aggiornata da una procedura di ottimizzazione standard che porta a W_{out}^* e c^* . Una volta completato l'addestramento dello strato di output, la dinamica è data da:

$$r(t + \Delta t) = (1 - \alpha)r(t) + \alpha \tanh(Ar(t) + W_{in}(W_{out}^*r(t) + c^*) + \xi \mathbf{1}).$$

La Figura 6.1 mostra i parametri utilizzati nel modello ESN, rispetto a quelli del MLP e dalla LSTM.

6.1.1 Confronto con i risultati del MLP e della LSTM

La Figura 6.2 riassume i modelli ottimizzati in termini di errore di addestramento (Training loss) ed errore di convalida (Validation loss). L'errore si riferisce all'errore quadratico medio tra i valori effettivi e quelli previsti. Tutti e tre i modelli raggiungono un livello accettabile di errore, il che suggerisce che sono ben addestrati e non soffrono di overfitting (adattamento eccessivo).

Network	Training loss	Validation loss	Training time (minutes)	$E(\bar{u}_x)$ (%)	$E(\overline{u'_x u'_x})$ (%)	$E(\overline{u'_x u'_y})$ (%)
MLP	2.0×10^{-5}	3.1×10^{-5}	$O(10^2)$	12.24	19.95	2.94
LSTM	2.8×10^{-8}	3.6×10^{-8}	$O(10^1)$	1.01	6.64	0.90
ESN	2.0×10^{-7}	2.4×10^{-7}	$O(10^0)$	3.96	22.33	0.98

Figura 6.2: *Riepilogo dei risultati di addestramento delle tre reti neurali (MLP, LSTM, ESN) [15].*

A livello computazionale, il MLP è il metodo più costoso a causa della grande dimensione dei dati di input e dell'elevato numero di strati nascosti. Il metodo LSTM ha i valori minori di errore; inoltre, il suo tempo di addestramento è inferiore rispetto al MLP. Tuttavia, la tecnica ESN supera sia il MLP, che la LSTM, in termini di tempo di addestramento. L'addestramento dell'ESN richiede una frazione di secondo su un normale computer, mentre l'addestramento della rete LSTM richiede un paio di minuti. Il problema affrontato è piccolo rispetto alla dimensione dei dati di addestramento di un tipico problema di turbolenza. Ci si aspetta quindi che l'ESN funzioni meglio se applicato ad un sistema con un numero di gradi di libertà molto maggiore rispetto ai 9 scelti per il caso in esame.

Dopo la fase di addestramento sono state somministrate nuove serie temporali (definite GT: Ground Truth). Queste nuove serie permettono di avviare la fase di previsione da parte dei modelli addestrati. Per quanto riguarda il MLP sono necessari dati per almeno 500 fasi temporali in ciascuna serie. Per il modello LSTM bastano solo 10 fasi temporali, a causa della sua memoria interna e del feedback. Il modello ESN ha bisogno solo di $a_k(0)$ per iniziare la sua previsione. La Figura 6.3 mostra un confronto delle previsioni di a_1 e a_9 . Dalla figura si può notare che i modelli addestrati non sono in grado di ricreare l'andamento esatto dei modi; l'errore quadratico medio nella fase di test è maggiore dell'errore di addestramento. Tuttavia, il risultato complessivo è molto vicino ai dati effettivi, ovvero tutte e tre le reti danno una dinamica non lineare che è vicina alla curva GT.

L'errore è dovuto alla natura altamente caotica del modello del flusso, che implica un'elevata sensibilità rispetto ai piccoli errori di arrotondamento. Con-

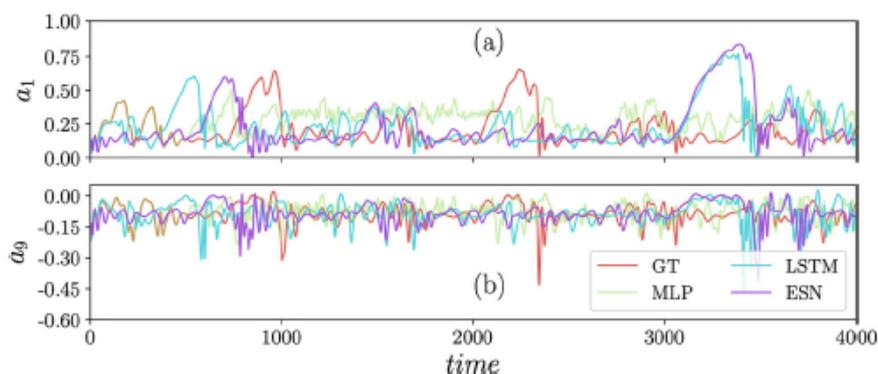


Figura 6.3: Confronto tra le serie temporali dei coefficienti a_1 e a_9 ottenute dalle tre reti neurali MLP, LSTM ed ESN e la serie prevista GT. I modelli sono stati addestrati con 400.000 punti e i risultati provengono da un test su una serie temporale con 4000 fasi [15].

siderando che lo scopo della ricerca era quello di ricreare le proprietà statistiche del sistema dinamico in esame, i ricercatori hanno analizzato il profilo medio e il profilo delle fluttuazioni delle componenti della velocità, al fine di quantificare l'errore.

Si è passati quindi all'analisi della velocità media e dello sforzo di taglio di Reynolds. In Figura 6.4 è presente un confronto tra i valori previsti dai modelli MLP, LSTM ed ESN e i valori GT, sia per quanto riguarda il profilo medio della velocità, che per quanto riguarda il profilo dello sforzo di taglio di Reynolds. Si può notare un'eccellente capacità di previsione dei modelli LSTM ed ESN.

Per quanto riguarda le fluttuazioni, nessuno dei modelli è riuscito a prevederle accuratamente nella direzione streamwise, mentre i risultati nella direzione normale alla parete e nella direzione spanwise sono ottimi (vedi Figura 6.5).

Il modello LSTM ha gli errori più bassi nella previsione della velocità media e dello sforzo di taglio di Reynolds. Il modello ESN si classifica al secondo posto. Il ridotto tempo di addestramento, la natura dei sistemi dinamici del reservoir, la maggiore accuratezza e il robusto framework favoriscono tuttavia il modello ESN. L'errore del modello ESN può essere ulteriormente ridotto utilizzando una rete profonda ESN.

In conclusione, si può affermare che l'approccio del reservoir computing, implementato qui come una Echo State Network, è in grado di prevedere statisticamente un flusso di taglio turbolento semplice senza costose operazioni di addestramento. Inoltre, la tecnica ESN ha grandi capacità anche nello studio di modelli di flusso che contengono un numero significativamente maggiore di gradi di libertà.

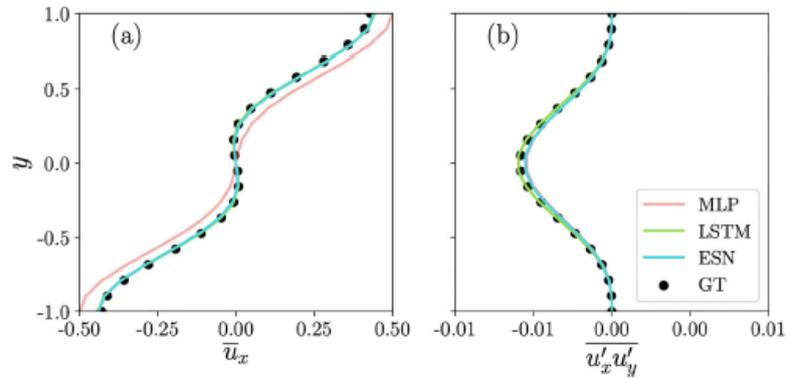


Figura 6.4: Confronto dei valori previsti dai modelli MLP, LSTM ed ESN con i valori attesi (GT) per il: (a) profilo medio della velocità; (b) profilo dello sforzo di taglio di Reynolds. I dati sono ottenuti da una media nelle direzioni streamwise e spanwise in combinazione con il tempo [15].

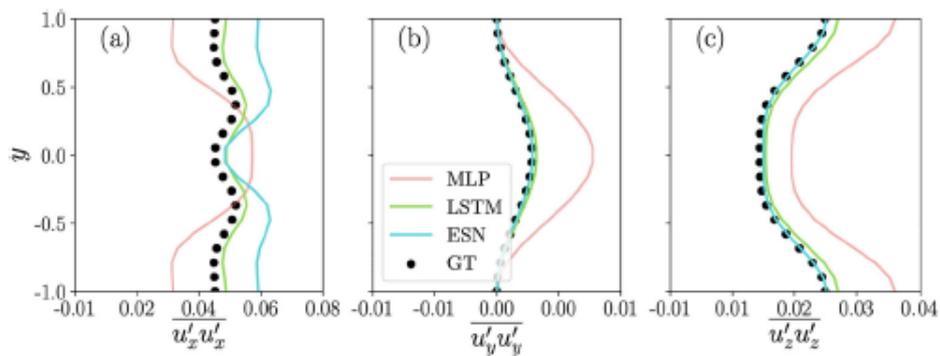


Figura 6.5: Confronto dei valori previsti dai modelli MLP, LSTM ed ESN con i valori attesi (GT) per le: (a) fluttuazioni della velocità nella direzione streamwise; (b) fluttuazioni della velocità nella direzione normale alla parete; (c) fluttuazioni della velocità nella direzione spanwise [15].

6.2 TBNN per la rappresentazione del tensore anisotropo degli sforzi di Reynolds

In Ling et al. [20] è proposto l'utilizzo di una DNN per la rappresentazione del tensore anisotropo degli sforzi di Reynolds.

È stata costruita una rete neurale profonda formata da 8 strati nascosti (hidden layers), la quale incorpora direttamente l'invarianza galileiana nelle sue previsioni. Questa rete è in grado di prevedere non solo gli autovalori del tensore anisotropo, ma anche l'intero tensore, preservando l'invarianza galileiana. Il preservarsi dell'invarianza è fondamentale per il rispetto delle proprietà fisiche. La DNN è stata addestrata, e successivamente testata, su un database di flussi per i quali sono disponibili sia i risultati delle RANS, che dati ad alta fedeltà.

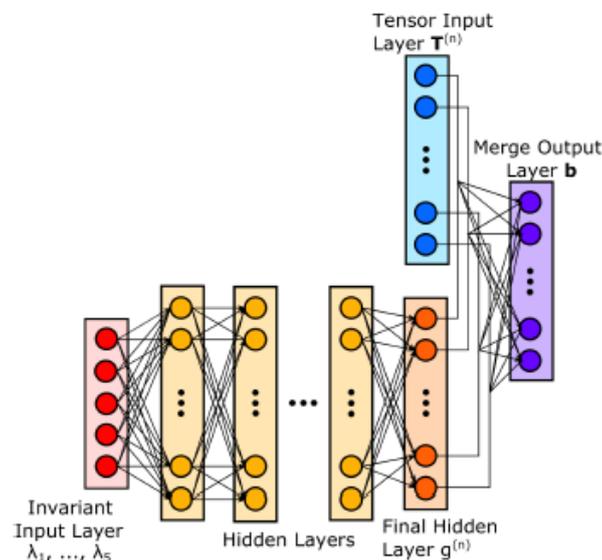


Figura 6.6: *Schema della Tensor Basis Neural Network [20].*

L'architettura progettata (vedi Figura 6.6) è stata definita Tensor Basis Neural Network (TBNN). Essa è in grado di incorporare l'invarianza rotazionale imponendo che il tensore anisotropo previsto si trovi su una base di tensori isotropi. L'invarianza rotazionale significa che la fisica del fluido non dipende dall'orientazione della struttura rispetto alle coordinate dell'osservatore. Questo è un principio fisico fondamentale ed è importante che qualsiasi metodo di chiusura della turbolenza lo rispetti. In caso contrario, un modello di ML valutato su flussi identici, ma con assi definiti in direzioni diverse, potrebbe produrre previsioni tra loro discordi.

L'applicazione dell'invarianza rotazionale alla DNN mostra un miglioramento sostanziale dei risultati, rispetto a quelli del MLP (Multilayer Perceptron), il quale non incorpora l'invarianza galileiana.

I modelli MLP e TBNN sono stati utilizzati per prevedere il tensore anisotropo degli sforzi di Reynolds normalizzato (b) per due tipi di flusso:

- flusso turbolento in un condotto;
- flusso su una parete ondulata.

I loro output sono stati confrontati anche con le soluzioni di due modelli RANS di base: Linear Eddy Viscosity Model (LEVM) e Quadratic Eddy Viscosity Model (QEVM).

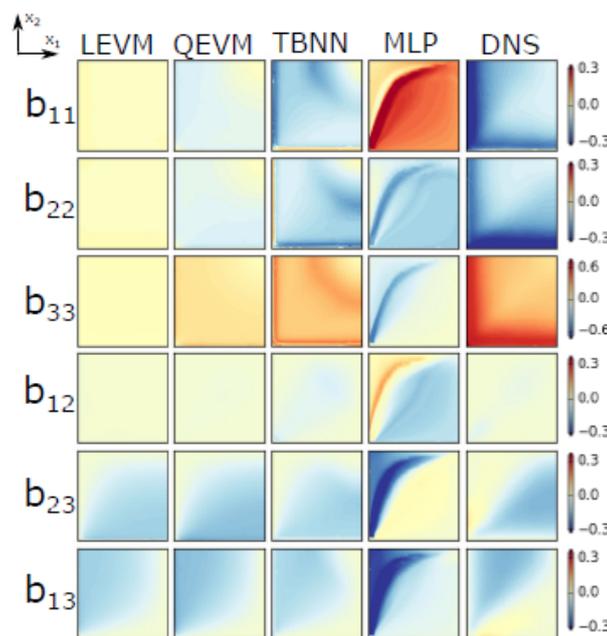


Figura 6.7: *Rapresentazione delle previsioni del tensore anisotropo degli sforzi di Reynolds (b) nel caso di flusso turbolento in un condotto. È mostrato solo il quadrante inferiore sinistro del condotto e la direzione streamwise del flusso è uscente dal foglio. Le colonne mostrano le previsioni dei modelli LEVM, QEVM, TBNN e MLP. I valori corretti del tensore (derivati dalla DNS) sono mostrati nella colonna più a destra [20].*

I risultati, per quanto riguarda il flusso turbolento in un condotto, sono mostrati in Figura 6.7. Il modello LEVM non è riuscito ad ottenere risultati soddisfacenti.

Il modello QEVM ha ottenuto risultati leggermente migliori, perché è riuscito a prevedere il segno corretto delle componenti del tensore. Il modello TBNN è stato il migliore, come si può vedere dalla figura. Il MLP ha fallito completamente nel prevedere il tensore anisotropo: le sue previsioni sono sia qualitativamente che quantitativamente imprecise.

La Figura 6.8 mostra le previsioni del tensore anisotropo nel caso di flusso su una parete ondulata. Le componenti b_{13} e b_{23} del tensore non vengono visualizzate perché, in questo caso, sono uguali a zero. Ancora una volta il LEVM non è stato in grado di prevedere nemmeno gli andamenti qualitativi delle componenti del tensore; il modello QEVM è riuscito a prevedere il segno corretto solo per le componenti sulla diagonale di b . Il modello TBNN ha previsto valori corretti per quanto riguarda le componenti sulla diagonale di b ; per quanto riguarda la componente b_{12} ha previsto dei valori troppo bassi. Anche in questo caso il MLP non è stato in grado di fare previsioni soddisfacenti.

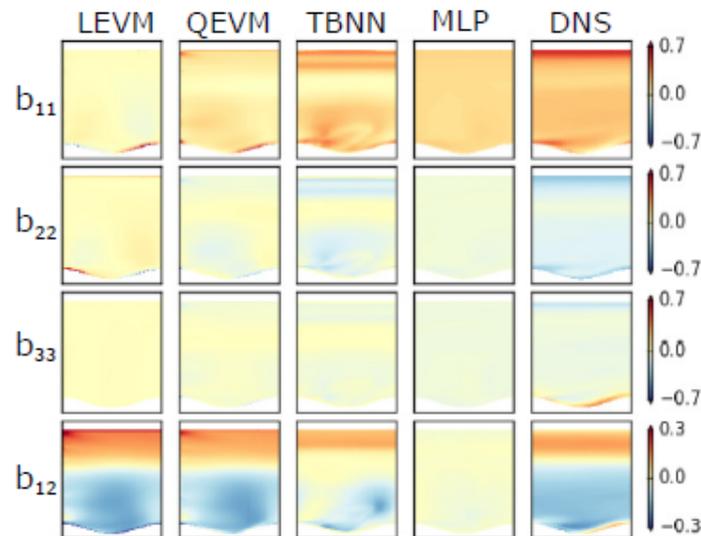


Figura 6.8: Rappresentazione delle previsioni del tensore anisotropo degli sforzi di Reynolds (b) nel caso di flusso su una parete ondulata. Le colonne mostrano le previsioni dei modelli LEVM, QEVM, TBNN e MLP. I valori corretti del tensore (calcolati tramite la DNS) sono mostrati nella colonna più a destra [20].

Un confronto più quantitativo delle prestazioni dei modelli è fornito dalla Figura 6.9, che mostra l'errore quadratico medio (RMSE) delle quattro tecniche rispetto ai dati ricavati dalla DNS. Dalla tabella si può notare che, nel caso di flusso su una parete ondulata, il MLP ha ottenuto un errore inferiore rispetto ai modelli

LEVM e QEVM. In entrambi i casi il modello TBNN ha ottenuto gli errori più bassi, sia rispetto ai due modelli RANS di base, che rispetto al MLP. Questo dimostra i vantaggi prestazionali dovuti all'incorporazione dell'invarianza galileiana nell'architettura di rete.

Model	Duct Flow	Flow over Wavy Wall
LEVM	0.23	0.18
QEVM	0.18	0.11
TBNN	0.13	0.08
MLP	0.33	0.09

Figura 6.9: *Errore quadratico medio [20].*

Capitolo 7

Apprendimento Supervisionato: Reti neurali convoluzionali per i flussi turbolenti

7.1 DCNN per lo studio della RBC turbolenta

In Fonda et al. [21] è presentato l'utilizzo di una rete neurale convoluzionale profonda (DCNN) per estrarre modelli, con gradi di libertà ridotti, dalla convezione di Rayleigh-Bénard (RBC) turbolenta ed estesa orizzontalmente. In particolare, l'obiettivo dello studio era quello di quantificare l'importanza delle sovrastrutture turbolente nel trasporto del calore. La convezione turbolenta su larga scala è organizzata in più sovrastrutture turbolente, le quali hanno un'estensione orizzontale maggiore dell'altezza dello strato di convezione (per questo motivo vengono chiamate "sovrastrutture"). Le sovrastrutture evolvono gradualmente nel tempo. È stato supposto che queste sovrastrutture abbiano un ruolo fondamentale nel trasporto turbolento del calore. Le conoscenze sui meccanismi di trasporto turbolento nello strato di convezione sono state utilizzate per pre-elaborare i dati, che sono stati poi inseriti nella rete neurale. È noto, ad esempio, che i massimi locali del flusso di calore convettivo coincidono con i minimi e i massimi locali delle fluttuazioni di temperatura.

I dati necessari per addestrare la rete sono stati ricavati dalla simulazione su larga scala delle equazioni del moto. I dati ottenuti sono stati analizzati in un dominio cartesiano con proporzioni $25H : 25H : H$, dove H è l'altezza dello strato di convezione. È stato utilizzato un numero di Prandtl¹ del fluido di $Pr = 7$ e sono stati utilizzati i numeri di Rayleigh $Ra = 10^5, 10^6, 10^7$. La DCNN è stata

¹Il numero di Prandtl è così definito: $Pr = \frac{\mu \cdot c_p}{k}$, dove μ è la viscosità dinamica, c_p è il calore specifico a pressione costante e k è la conducibilità termica.

addestrata al fine di ridurre il campo di temperatura tridimensionale in una rete planare temporale nel piano medio dello strato di convezione, come illustrato in Figura 7.1.

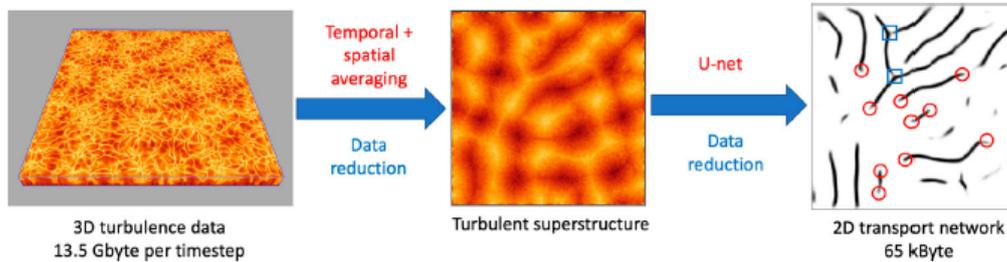


Figura 7.1: La sequenza di immagini rappresenta la riduzione dei dati di turbolenza 3D in una rete planare (2D). La massima compressione dei dati è stata di oltre un fattore di 100.000 (per il numero di Rayleigh più grande considerato). Si è passati da una sequenza di istantanee tridimensionali completamente risolte del campo di temperatura ad un modello di cresta che rappresenta le regioni con il flusso di calore convettivo in media più grande. Nell'immagine più a destra sono indicati i punti di difetto della rete: sono utilizzati i cerchi per rappresentare i punti finali delle creste e i quadrati per indicare i punti di fusione di due creste [15].

La rete planare ottenuta è stata utilizzata per quantificare il calore turbolento trasportato da una sovrastruttura turbolenta. I punti di difetto (indicati in Figura 7.1) scompaiono e formano nuove connessioni nel tempo; sono i "punti caldi", ovvero i punti di flusso di calore potenziato localmente. La conclusione alla quale sono arrivati i ricercatori è che la frazione di calore trasportata da una sovrastruttura diminuisce all'aumentare del numero di Rayleigh [21]. Una quantità crescente di calore è quindi trasportata dalla turbolenza di fondo su piccola scala, piuttosto che dalle sovrastrutture di convezione coerenti su larga scala, all'aumentare del numero di Rayleigh. La Figura 7.2 mostra i modelli di cresta ottenuti dalla rete neurale.

La rete utilizzata nello studio è la U-net, la quale combina il percorso di contrazione che avviene in una rete neurale convoluzionale standard, con un successivo percorso di espansione di concatenazioni e convoluzioni verso l'alto, al fine di ridurre la complessa sovrastruttura turbolenta 3D in una rete planare temporale nel piano mediano dello strato di convezione. L'utilizzo di una U-net è essenziale per lo studio di una RBC di grandi dimensioni, perché richiede piccoli set di dati per l'addestramento. La lenta evoluzione delle sovrastrutture in un grande dominio richiederebbe altrimenti simulazioni estremamente lunghe dell'intero flusso RBC per ottenere una quantità appropriata di dati di addestramento.

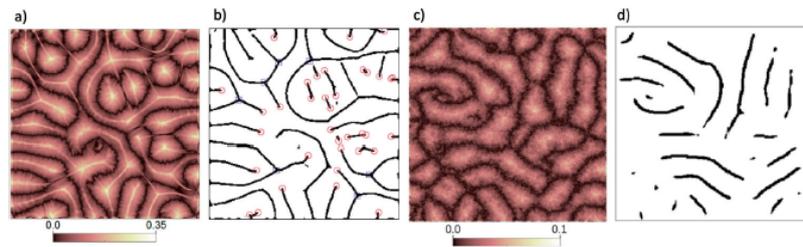


Figura 7.2: Modelli di cresta ottenuti dalla rete neurale profonda; l'output della rete è rappresentato nelle figure b e d. I numeri di Ra utilizzati sono stati $Ra = 10^5$ (a,b) e $Ra = 10^7$ (c,d); il numero di Pr è $Pr = 7$ [15].

La Figura 7.3 mostra la capacità della DCNN di estrarre i modelli di cresta a diversi numeri di Prandtl.

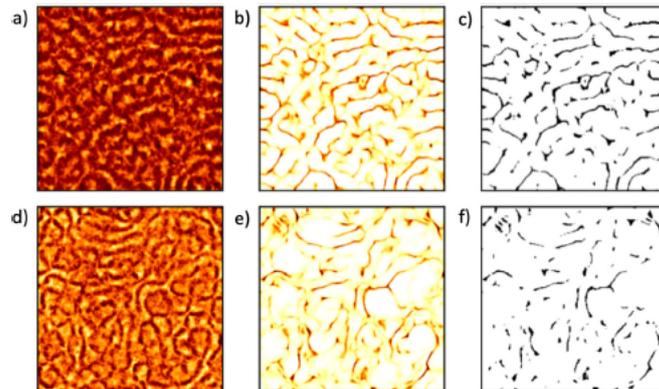


Figura 7.3: Estrazione delle creste per flussi con numeri di Prandtl diversi: (a) $Pr = 0,7$ e (d) $Pr = 0,021$. Le immagini (b) ed (e) mostrano l'output della U-net e le immagini (c) ed (f) la conversione in file binari. In entrambi gli esempi $Ra = 10^5$ e le proporzioni sono di $25H : 25H : H$ [15].

Sebbene sorprendentemente robusto, il metodo di riduzione dei dati utilizzato ha dei limiti, specialmente per i dati di convezione che sono caratterizzati da una forte inerzia del fluido.

7.2 CNN per la ricostruzione della turbolenza omogenea isotropa bidimensionale in decadimento

Fukami et al. [22] hanno utilizzato una rete neurale convoluzionale (CNN) per la ricostruzione spazio-temporale SR (Super Resolution) di flussi turbolenti. Il loro obiettivo era quello di ricostruire dati di flussi turbolenti ad alta risoluzione ($q(x_{HR}, t_{HR})$), partendo da dati spaziali e temporali a risoluzione estremamente bassa ($q(x_{LR}, t_{LR})$). Il processo di ricostruzione dei dati è stato diviso in due parti (vedi Figura 7.4): SR spaziale e SR temporale (inbetweening).

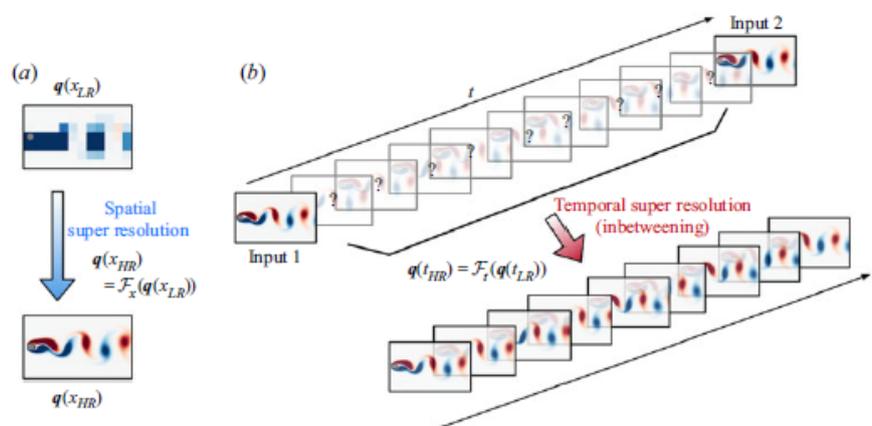


Figura 7.4: Ricostruzione dei dati tramite: (a) SR spaziale; (b) SR temporale (inbetweening) [22].

La CNN è stata addestrata tramite dati input x e dati output y . L'obiettivo dell'addestramento è quello di costruire una funzione di mappatura non lineare $y \approx F(x; w)$, dove il vettore w rappresenta i pesi all'interno della rete neurale. La rete deve quindi determinare i pesi w tali che $w = \operatorname{argmin}_w [E(y, F(x; w))]$, dove E è la funzione di perdita (o funzione di costo).

La CNN utilizzata è il modello ibrido DSC/MS (vedi Figura 7.5). Il modello DSC si oppone alla rotazione e alla traslazione degli oggetti all'interno delle immagini di input. Il modello MS è invece in grado di tenere conto delle proprietà principali del campo di flusso. Il modello DSC/MS viene quindi utilizzato per la ricostruzione di dati spazio-temporali.

In questo paragrafo è presentata la ricostruzione della turbolenza omogenea isotropa bidimensionale in decadimento. Il set di dati di addestramento è stato ottenuto risolvendo numericamente l'equazione di trasporto della vorticità bidi-

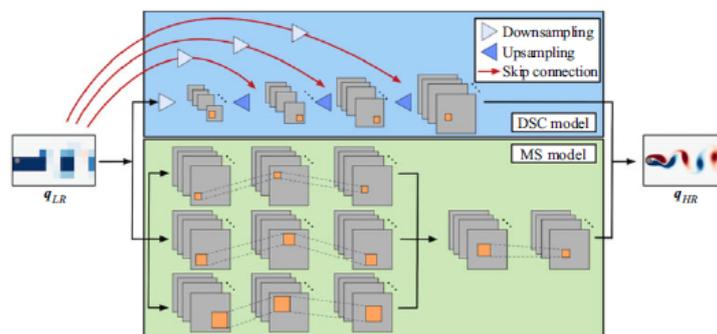


Figura 7.5: Modello ibrido DSC/MS [22].

mensionale:

$$\frac{\partial \omega}{\partial t} + \vec{u} \cdot \nabla \omega = \frac{1}{Re_0} \nabla^2 \omega,$$

dove $\vec{u} = (u, v)$ e ω sono rispettivamente la velocità e la vorticità.

La dimensione del dominio computazionale biperiodico e il numero di punti della griglia delle immagini utilizzati sono rispettivamente $L_x = L_y = 1$ e $N_x = N_y = 128$.

Il numero di Reynolds è definito come $Re_0 = u^* l_0^* / \nu$, dove u^* è la velocità caratteristica ottenuta dalla radice quadrata dell'energia cinetica iniziale mediata nello spazio, $l_0^* = [2\bar{u}^2(t_0) / \bar{\omega}^2(t_0)]^{1/2}$ è la lunghezza integrale iniziale e ν è la viscosità cinematica. I numeri di Reynolds iniziali utilizzati sono $Re_0 = 81,2$ per i dati di addestramento e convalida, e $Re_0 = 85,4$ per i dati di test.

Sono stati analizzati quattro casi che comprendono due livelli di risoluzione spaziale e due livelli di risoluzione temporale, come mostrato in Figura 7.6. Per l'analisi spaziale SR i dati in input sono stati divisi in due livelli di risoluzione: media (griglia 16×16) e bassa (griglia 8×8). Per l'impostazione della risoluzione temporale è stato definito un intervallo temporale medio ($\Delta T = 1, 0$) e un intervallo temporale ampio ($\Delta T = 4, 0$), dove ΔT rappresenta l'intervallo temporale tra il primo e l'ultimo fotogramma nell'analisi inbetweening.

Il decadimento del numero $Re_\lambda(t) = u^\#(t)\lambda(t)/\nu$, dove $u^\#(t)$ è la media quadratica spaziale della velocità in un campo istantaneo e $\lambda(t)$ è il valore della scala di lunghezza di Taylor² in un campo istantaneo, è mostrato al centro della Figura 7.6.

²La microscaletta di Taylor, chiamata anche scala di lunghezza della turbolenza, è una scala di lunghezza utilizzata per caratterizzare un flusso turbolento. Questa microscaletta prende il nome da Geoffrey Ingram Taylor.

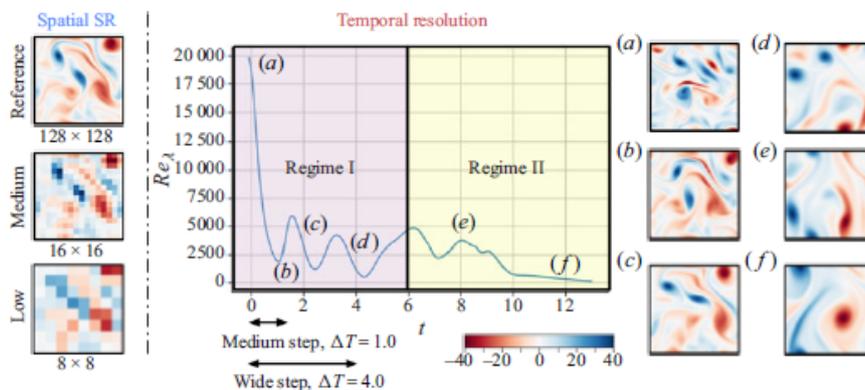


Figura 7.6: Ricostruzione spatio-temporale SR della turbolenza isotropa omogenea bidimensionale in decadimento. La curva situata al centro rappresenta il decadimento del numero Re_λ . I campi di vorticit  ω ai tempi (a)-(f) lungo la curva sono mostrati nelle immagini di destra [22].

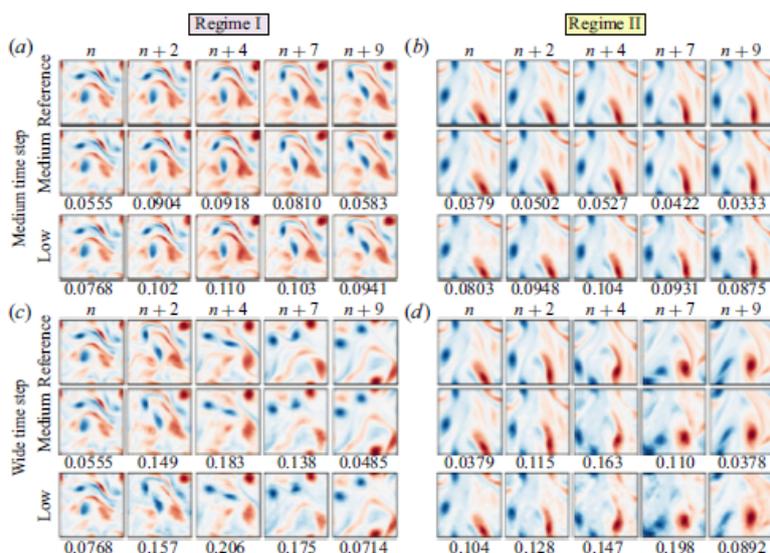


Figura 7.7: Ricostruzione spatio-temporale SR per la turbolenza omogenea isotropa bidimensionale in decadimento. Il campo di vorticit  ω   mostrato con lo stesso contour level della Figura 7.6. Sono rappresentati i risultati ottenuti da input spaziali a media e bassa risoluzione in un: (a) intervallo temporale medio per il regime I; (b) intervallo temporale medio per il regime II; (c) intervallo temporale ampio per regime I; (d) intervallo temporale ampio per il regime II. I valori sotto i campi di flusso riportano la norma L_2 dell'errore [22].

Per il processo di addestramento è stato considerato un numero fisso di istantanee: $(n_{snapshot,x}, n_{snapshot,t})=(10000, 10000)$. I quattro modelli sono stati costruiti separatamente.

L'analisi spaziale SR permette di ricostruire il primo e l'ultimo fotogramma. L'analisi inbetweening ha il compito di ricostruire otto istantanee tra il primo e l'ultimo fotogramma.

I campi di flusso ricostruiti dall'analisi spazio-temporale SR sono riassunti in Figura 7.7. Sul lato sinistro dell'immagine sono mostrati i campi ricostruiti nel regime I. Come si può notare dai risultati, l'evoluzione temporale della complessa dinamica della vorticità può essere accuratamente ricostruita dalla CNN. Per quasi tutti i casi, le norme L_2 dell'errore $\epsilon = \|q_{DNS} - q_{ML}\|_2 / \|q_{DNS}\|_2$, elencate sotto le immagini, sono maggiori nel caso di input spaziale a bassa risoluzione, rispetto al caso di input a media risoluzione.

Sul lato destro della Figura 7.7 sono mostrati i risultati nel regime II. Analogamente ai risultati ottenuti nel regime I, i campi di flusso ricostruiti sono in accordo con i dati DNS di riferimento. È importante notare che il picco della norma L_2 dell'errore è di 0,198 e si trova all'istante di tempo $t = (n + 7)\Delta t$, nel caso di dati input a bassa risoluzione e ΔT ampio. Questo fatto è in contrasto con i risultati ottenuti negli altri casi, che hanno il picco dell'errore al tempo $t = (n + 4)\Delta t$. Il motivo di questa disuguaglianza è che il modello di ML, che non incorpora informazioni sulle condizioni al contorno, deve gestire l'evoluzione temporale di una struttura relativamente grande su un dominio biperiodico. Inoltre, la ricostruzione temporale risente dell'errore derivante dalla ricostruzione spaziale. Per queste ragioni il picco dell'errore è spostato nel tempo rispetto agli altri casi.

In Figura 7.8 sono mostrate le norme L_2 dell'errore, al variare del regime. Per tutti i casi, gli errori nel regime I sono maggiori di quelli nel regime II. Uno dei motivi è che il cambiamento relativo delle strutture dei vortici nel regime II è inferiore rispetto a quello nel regime I.

Le ricostruzioni sono influenzate anche dalla diversa risoluzione dei dati input spaziali. Per notare questo fatto basta confrontare le immagini (a) e (b) della Figura 7.8.

La Figura 7.9 mostra l'andamento dell'energia cinetica totale, E_{tot} , per i quattro casi in esame.

In conclusione, è possibile affermare che la CNN è stata in grado di ricostruire dati ad alta risoluzione del decadimento della turbolenza omogenea isotropa bidimensionale, partendo da dati spazio-temporali a bassa risoluzione.

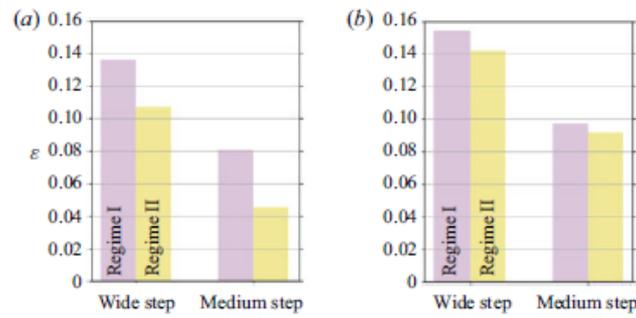


Figura 7.8: Norme L_2 dell'errore per i casi: (a) input spaziali a risoluzione media; (b) input spaziali a risoluzione bassa [22].

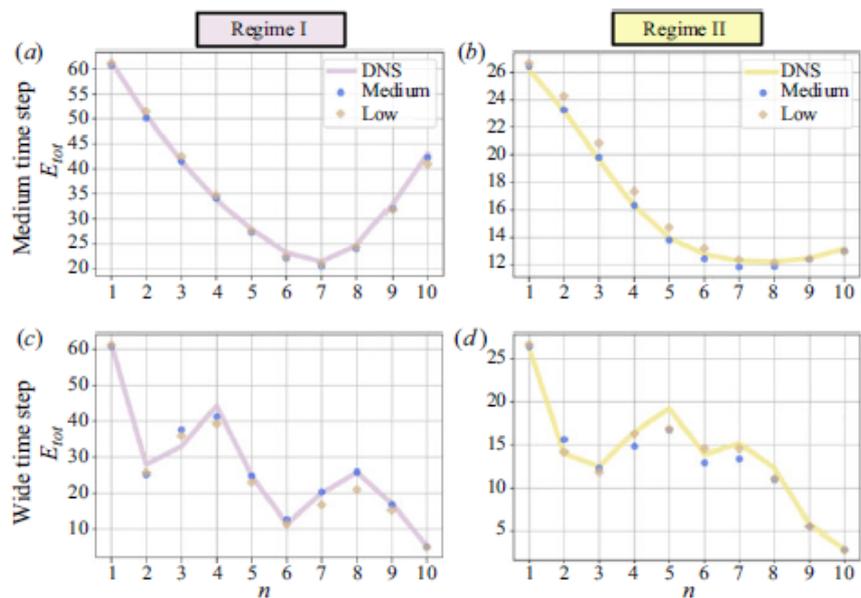


Figura 7.9: Decadimento dell'energia cinetica totale, E_{tot} , nel dominio del tempo [22].

Capitolo 8

Conclusioni

In questo lavoro è stato mostrato l'utilizzo delle tecniche di Machine Learning per lo studio dei flussi turbolenti. L'insieme dei metodi di ML descritti è molto ampio: sono stati applicati sia algoritmi di clustering (apprendimento non supervisionato), sia metodi semi-supervisionati, come le GANs, sia le reti neurali (apprendimento supervisionato).

In particolare, i metodi più utilizzati in fluidodinamica sono le reti neurali profonde. Sia le reti neurali convoluzionali, che quelle ricorrenti, possono prevedere la dinamica dei flussi turbolenti senza conoscere le equazioni matematiche che li descrivono [15].

L'applicazione delle DNNs ai flussi turbolenti ha rappresentato un cambio di paradigma per i ricercatori. In passato molte innovazioni sono state ispirate da intuizioni e da modelli interpretabili fisicamente. Oggi le DNNs stanno sfidando le tecniche tradizionali, costruendo modelli molto più performanti [19].

Nonostante il successo che il ML ha raggiunto in fluidodinamica, rimangono aperte ancora tante domande, come ad esempio:

- Partendo da un determinato set di dati, quanti nodi e quanti strati deve avere la rete da addestrare?
- Per addestrare adeguatamente una rete neurale, quanto deve essere grande il set di dati?
- Ci sono garanzie che un determinato modello di ML possa portare ad un buon risultato?
- Qual è l'incertezza degli output delle tecniche di ML?
- Si può garantire il non overfitting dei dati?

Il Machine Learning ha sicuramente ampi margini di miglioramento. In futuro dovrà essere ridotto il costo computazionale associato ai modelli di apprendimento. Inoltre, dovranno essere fornite garanzie per quanto riguarda la robustezza, le prestazioni e la convergenza delle tecniche [4].

Considerando lo sviluppo che il ML ha avuto negli ultimi anni, in futuro ci si aspetta una crescita pervasiva di modelli basati sui dati per lo studio dei flussi turbolenti.

Bibliografia

- [1] Renzo Arina. *Fondamenti di aerodinamica*. Levrotto & Bella, 2015.
- [2] Karthik Duraisamy, Gianluca Iaccarino, Heng Xiao. *Turbulence Modeling in the Age of Data*. Annual Review of Fluid Mechanics, 2019.
- [3] Dizionario Treccani.
- [4] Steven L. Brunton, Bernd R. Noack, Petros Koumoutsakos. *Machine Learning for Fluid Mechanics*. Annual Review of Fluid Mechanics, 2020.
- [5] Scott White and Padhraic Smyth. *A Spectral Clustering Approach To Finding Communities in Graphs*. SIAM, 2005.
- [6] Hyojin Kim, Junhyuk Kim, Sungjin Won and Changhoon Lee. *Unsupervised deep learning for super-resolution reconstruction of turbulence*. J. Fluid Mech., 2021.
- [7] Jacopo Kahl. *I tre principali tipi di Machine Learning*. Medium, 2020.
- [8] Hornik K., Stinchcombe M., White H. *Multilayer feedforward networks are universal approximators*. Neural Networks 2:359-366, 1989.
- [9] Alex Graves. *Supervised Sequence Labelling with Recurrent Neural Networks*. Springer, 2012.
- [10] Dhireesha Kudithipudi, Qutaiba Saleh, Cory Merkel and James Thesing. *Design and Analysis of a Neuromemristive Reservoir Computing Architecture for Biosignal Processing*. Frontiers in Neuroscience 9, 2016.
- [11] *Turbolenza omogenea isotropa*. Sapienza, Università di Roma, 2015.
- [12] Gianfranco Chicocchia, Massimo Germano. *Termofluidodinamica*. Levrotto & Bella, 2007.
- [13] Marcello Lappa. *On the Nature of Fluid-dynamics*. Nova Science Publishers Inc., 2019.

-
- [14] Jeff Moehlis, Holger Faisst and Bruno Eckhardt. *A low-dimensional model for turbulent shear flows*. New Journal of Physics, 2004.
- [15] Sandeep Pandey, Jörg Schumacher & Katepalli R. Sreenivasan. *A perspective on machine learning in turbulent flows*. Journal of Turbulence, 2020.
- [16] Christiane Schneide, Ambrish Pandey, Kathrin Padberg-Gehle and Jörg Schumacher. *Probing turbulent superstructures in Rayleigh-Bénard convection by Lagrangian trajectory clusters*. Physical Review Fluids, (2018).
- [17] Ryan King, Oliver Hennigh, A. Mohan & M. Chertkov. *From Deep to Physics-Informed Learning of Turbulence: Diagnostics*. physics.flu-dyn, 2018.
- [18] Junhyuk Kim, Changhoon Lee. *Deep unsupervised learning of turbulence for inflow generation at various Reynolds numbers*. Journal of Computational Physics, 2020.
- [19] J. Nathan Kutz. *Deep learning in fluid dynamics*. Cambridge University Press, 2017.
- [20] Julia Ling, Jeremy Templeton and Andrew Kurzawski. *Reynolds Averaged Turbulence Modeling using Deep Neural Networks with Embedded Invariance*. J. Fluid Mech, 2016.
- [21] Enrico Fonda, Ambrish Pandey, Jörg Schumacher and Katepalli R. Sreenivasan. *Deep learning in turbulent convection networks*. PNAS, 2019.
- [22] Kai Fukami, Koji Fukagata and Kunihiro Taira. *Machine-learning-based spatio-temporal super resolution reconstruction of turbulent flows*. J. Fluid Mech., 2020.
- [23] Wataru Kobayashi, Takaaki Shimura, Akihiko Mitsuishi, Kaoru Iwamoto, Akira Murata. *Prediction of the drag reduction effect of pulsating pipe flow based on machine learning*. International Journal of Heat and Fluid Flow, 2021.
- [24] Taichi Nakamura, Kai Fukami, Kazuto Hasegawa, Yusuke Nabae and Koji Fukagata. *Convolutional neural network and long short-term memory based reduced order surrogate for minimal turbulent channel flow*. Physics of Fluids, 2021.
- [25] P. A. Srinivasan, L. Guastoni, H. Azizpour, P. Schlatter and R. Vinuesa. *Predictions of turbulent shear flows using deep neural networks*. PHYSICAL REVIEW FLUIDS 4, 2019.

Ringraziamenti

Vorrei dedicare questo spazio alle persone che mi sono state vicine durante il percorso universitario e hanno contribuito alla realizzazione di questo elaborato.

Grazie alla relatrice, Professoressa Stefania Scarsoglio, che con pazienza e disponibilità mi ha seguito durante la stesura del lavoro. La ringrazio per avermi dato la possibilità di svolgere la Prova Finale in un ambito dell'ingegneria che mi ha sempre affascinato, ossia la fluidodinamica.

Grazie ai miei genitori, Clara ed Alberto, che da sempre mi sostengono nella realizzazione dei miei progetti. Grazie per avermi trasmesso la passione per la scienza, fondamentale in tutto il mio percorso di studi.

Grazie ai miei fratelli, Andrea e Matteo, e a mia sorella, Anna, per essere stati sempre presenti e pronti ad aiutarmi quando ne ho avuto bisogno. Grazie inoltre per tutti i momenti di spensieratezza.

Grazie a mio gemello, Luca, punto di riferimento insostituibile in ogni cosa che faccio.

Grazie alla mia ragazza e migliore amica, Irene. Grazie per il tempo che mi dedichi ogni giorno. Grazie per avermi sostenuto, tranquillizzato e incoraggiato in questi anni. Grazie per tutto l'affetto e la forza che mi dai ogni giorno.