

ОГЛЯД НЕЙРОМЕРЕЖЕВИХ ПІДХОДІВ ДЛЯ АНАЛІЗУ АКУСТИЧНИХ СИГНАЛІВ ВОДНОГО СЕРЕДОВИЩА

А. О. Олексій^{1, а}, А. А. Верлань¹

¹ Національний технічний університет України
«Київський політехнічний інститут імені Ігоря Сікорського»
Теплоенергетичний факультет

Анотація

У даній роботі проаналізовано нейромережеві моделі, що можуть бути застосовані для класифікації гідроакустичних сигналів, а саме: кооперативний метод глибокого навчання, модифікована мережа LeNet, розріджений складений автоенкодер з багатодоменим набором характеристик, багат шаровий персептрон, навчений з використанням алгоритму рою сальпів.

Ключові слова: нейронні мережі, LeNet, розріджений складений автоенкодер, багат шаровий персептрон, кооперативний метод глибокого навчання

Вступ

На сьогодні, задача аналізу акустичних сигналів морського середовища має багато складних підзадач, серед яких є наступні проблеми: значний обсяг одержуваної інформації та висока швидкість зміни навколишнього середовища. При пасивних методах прийому сигналів, сигнали можуть характеризуватись невисоким співвідношенням сигнал/шум. Все це потребує ефективного та універсального рішення, яким є нейромережі. Будуть розглянуті можливі нейромережеві підходи та зроблені висновки щодо можливостей їх використання. В подальшому, обрані підходи можуть послужити основою для розробки прикладної програмної системи аналізу акустичних сигналів водного середовища.

1. Модифікована нейромережа LeNet

У статті [1] розглядається модифікована нейромережа LeNet. Мережа LeNet має 7 шарів, включаючи 2 шари згортки, 2 шари об'єднання та 3 повнозв'язні шари. Операція згортки є зваженим складенням пікселів на основі ядра згортки. Метою шару згортки є вилучення особливостей вхідного зображення. Перший згортковий шар може витягувати лише деякі, відносно низькорівневі характеристики, такі як краї, лінії тощо. Другий шар згортки отримує складніші характеристики на основі низькорівневих характеристик, вилучених першим шаром. Згортковий шар в архітектурі нейромережі LeNet, використовує ядра згортки 5×5 , і крок переміщення ядра згортки дорівнює 1. Роль шару підвибірки еквівалентна перетворенню зображення з великою роздільною здатністю на зображення з меншою роздільною здатністю, що дозволяє зменшити кількість параметрів. Шар підвибірки в нейромережі LeNet використовує вхідне

поле 2×2 з розміром кроку переміщення 2, тобто чотири пікселі попереднього шару вводяться в один піксель наступного шару, і поля введення не перекриваються. Повнозв'язний шар перетворює високорозмірні характеристики на низькорозмірні і зберігає корисну інформацію. Останній повнозв'язний шар виступає в ролі класифікатора. Набори чотирьох типів спектра подаються до нейромережі LeNet для навчання. Порівнюючи точність, було виявлено, що використання спектра LOFAR дало найкращі результати.

Модифікована нейромережа LeNet має перший згортковий шар, п'ять ядер згортки, кожне з яких має розмір 4×4 і ядро згортки з ковзним кроком 1, без відступу. Фільтр в шарі підвибірки має розмір 2×2 і розмір кроку 2. Мета полягає в тому, щоб зменшити розмірність карти ознак і розширити область сприйняття. Ефект зменшення розмірності полягає в тому, щоб зменшити вихідне зображення на чверть і залишити середній вихід. Існують три комбінації такого шару згортки та шару підвибірки. Щоб зробити мережу придатнішою для ідентифікації та класифікації спектра LOFAR для гідроакустичних сигналів, регулював розмір і кількість згорток ядра в кожному згортковому шарі.

Модифікована нейромережа LeNet має ряд переваг, серед яких простота у реалізації та задовільна точність класифікації при роботі з даними, що перетворені в аудіо спектр. В той самий час, у модифікованій нейромережі LeNet є багато недоліків. Нейромережа показує невисокі результати при роботі з кольоровими зображеннями. Також, LeNet не адаптована для роботи з великими наборами даних.

^аarturoleksii@gmail.com

2. Новий кооперативний метод глибокого навчання для розпізнавання підводних акустичних цілей

У статті [2] розглядається кооперативний метод глибокого навчання. Нейронна мережа автоенкодеру довгої короткочасної пам'яті (LSTM) будується з кількох LSTM шарів. У мережі автоенкодеру LSTM присутні кодувальник та декодувальник. Вони містять один або кілька LSTM шарів. Мережа автоенкодеру LSTM може усунути нерелевантну та надмірну інформацію та покращує здатність вилучення вхідних характеристик. Мета мережі автоенкодеру LSTM полягає в тому, щоб реконструйовані дані x були близькі до вхідних даних x . Мережа намагається вивчити функцію $f(x) = \hat{x}$, щоб $\hat{x}$ був близьким до x . В запропонованій архітектурі, кількість нейронів у прихованому шарі менше, ніж розмірність вхідного сигналу. На етапі кодування, кодувальник приймає вхідний сигнал x і відображає його в латентне уявлення h за допомогою детермінованої функції виду $x = f_{W,b} = (W \times x + b)$. Вагова матриця W та зсув b використовуються для лінійного перетворення вхідних даних, потім нелінійні функції активації $f(\bullet)$ використовуються для додавання нелінійних компонент. Нарешті, ми отримуємо латентне уявлення h з меншою розмірністю. На етапі декодування, ми використовуємо функцію $\hat{x} = f_{W',b'}(W' \times h + b')$ для реконструкції вхідних даних (W' і W зазвичай обмежуються формою $W' = W^T$).

На другому етапі, працює кооперативна глибока мережа, яка складається з попередньо навченого шару кодувальника і шару softmax. Деякі параметри кооперативної мережі LSTM ініціалізуються за допомогою автоенкодеру LSTM на першому етапі, а потім навчаються та оптимізуються протягом всього процесу навчання в контрольованому режимі. Кооперативна мережа вивчає властивий характер і мінливість за допомогою використання короткочасної пам'яті для оновлення довгої пам'яті, та подає неявну інформацію про категорію. Інформація про категорію складається з серії частотних компонент, що змінюються в часі. Характеристики, що вивчені автоенкодером LSTM, зберігають класифікаційну інформацію, яка може підвищити точність розпізнавання підводних акустичних цілей.

Використання блоків LSTM дає можливість роботи з послідовностями в якості вхідних даних. Також, нейромережа може приймати вхідні дані змінної довжини. Перевага перед звичайним автоенкодером полягає у можливості уникнути проблеми довгострокової залежності шляхом зберігання інформації на тривалі періоди часу. Недоліком є відсутність нелінійності, що може погіршити результат. Залежно від довжини послідовності, блоки LSTM схильні забувати деяку найменш значущу інформацію. Шари кодування намагаються зберегти кількість інформації, а не її якість, внаслідок чого виникають складнощі з визначенням релевантної інформації.

3. Класифікація підводних цілей на великих глибинах з використанням глибокої нейронної мережі із спільними багатодоменими характеристиками

У статті [3] запропонована глибока модель, що складається зі складеного розрідженого автоенкодеру та softmax-класифікатору. Спільні ознаки є низькорівневими та недостатньо надійними. Складений розріджений автоенкодер досить потужний для того, щоб вивчити структуру вхідних даних навчанням без вчителя. Група спільних ознак використовується як вхідний вектор ознак даних для генерації глибоких ознак на останньому прихованому шарі. Припустимо, що набір даних $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)}) = j\}$ з m навчальних вибірок, де $x^{(i)}$ і $y^{(i)}$ – вектор вхідних ознак і мітка категорії для i -ої вибірки, відповідно. Softmax-класифікатор здатний оцінити можливість класифікації кожного зразка, за допомогою якої ми можемо визначити категорію.

Розріджений автоенкодер здатний вивчити розріджене представлення з вихідних даних. Для вивчення більш дискримінаційного представлення даних пропонується побудувати модель складеного розрідженого автоенкодеру з багатшаровою структурою, яка будується шляхом об'єднання вхідного шару і прихованого шару складеного розрідженого автоенкодеру шар за шаром. Однак, вихідні шари розріджених автоенкодеру не потрібні для побудови складеного розрідженого автоенкодеру, тому вихідний шар кожного окремого розрідженого автоенкодеру видаляється. Після підсумовування, вихід прихованого шару першого розрідженого автоенкодеру $a^{(1)}$ розглядається як вхід другого розрідженого автоенкодеру. Повторюючи ці кроки, ми можемо отримати глибоку структуру моделі складеного розрідженого автоенкодеру з декількома розрідженими автоенкодерами. Якщо ми візьмемо L розріджених автоенкодерів для побудови моделі розрідженого складеного автоенкодеру, вихід останнього прихованого шару $a^{(L)}$ є глибокою ознакою, яку ми вивчаємо з необробленого вхідного сигналу за допомогою глибокої структури вхідних даних, яка також може бути використана для класифікації. Після навчання ознак моделі складеного розрідженого автоенкодеру, простір вибірок $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$ перетворюється на простір ознак $\{(a^{(L,1)}, y^{(1)}), (a^{(L,2)}, y^{(2)}), \dots, (a^{(L,m)}, y^{(m)}) = j\}$, де $a^{(L,1)}$ – вихід L -го прихованого шару для i -ї вибірки, $a^{(L,1)} \in R^{S_L}$, а s_L позначає кількість вузлів в L -му прихованому шарі складеного розрідженого автоенкодеру. Зверніть увагу, що в роботі використовується s_i для позначення кількості вузлів в першому прихованому шарі складеного розрідженого автоенкодеру. Для глибокої ознаки $a^{(L,1)}$, softmax-класифікатор може дати ймовірність того, що вхід належить до індексу категорії j . Також, softmax-класифікатор може бути навчений за допомогою методу L-BFGS. Вся модель складе-

ного розрідженого автоенкодеру, що складається з кількох прихованих шарів та softmax-класифікатора, може бути навчена за допомогою методу жадібного пошарового навчання. На етапі навчання без спостереження, приховані шари навчаються ієрархічно зі зразком вхідних даних, що означає, що навчання відбувається для кількох розріджених автоенкодерів окремо. Після навчання, параметри прихованого шару зберігаються як початкова конфігурація для тонкого налаштування. Метод L-BFGS використовується для оптимізації всієї моделі на етапі тонкого налаштування.

Розріджені автоенкодери мають штраф за розрідженість, що допомагає запобігти перенавчанню. Також, вони беруть найвищі значення активації у прихованому шарі та обнуляють інші приховані вузли. Це не дозволяє автоенкодерам використовувати всі приховані вузли одночасно і змушує використовувати зменшену кількість прихованих вузлів. У той же час, потреба мережі в пам'яті буде значною, знайти нижню межу тимчасової складності для розріджених автоенкодерів досить складно та залишається традиційна для автоенкодерів проблема розкодування даних.

4. Класифікація цілей пасивного гідролокатора за допомогою багатопарового перцептронну, навченого алгоритмом рою сальпів

4.1. Багатопаровий перцептрон

У статті [4] представлений багатопаровий перцептрон, навчений з використанням алгоритму рою сальпів. Нейромережа багатопарового перцептронну є найбільш поширеним типом нейронних мереж зі прямою передачею даних. Нейрони є основними елементами обробки багатопарового перцептронну. У багатопаровому перцептронні, нейрони односпрямовано пов'язані між собою за допомогою зв'язків, званих «вагами». Запропонована модель трипарового перцептронну, де n означає кількість вхідних вузлів, h – кількість прихованих вузлів і m означає кількість вихідних вузлів. Вихід мережі багатопарового перцептронну розраховується рівнянням:

$$S_j = \sum_{i=1}^n (W_{ij} \times X_i) - \theta_j, \quad j = 1, 2, \dots, h, \quad (1)$$

де W_{ij} – вага, що з'єднує i -ий вузол (у вхідному шарі) з j -им вузлом (у прихованому шарі), θ_j позначає j -й вузол зміщення (у прихованому і X_i є вхід i -го вузла (у вхідному шарі). Вихід кожного прихованого вузла розраховується за допомогою сигмоїдальної функції як у рівнянні:

$$S_j = \text{sigmoid}(s_j) = \frac{1}{1 + \exp(-s_j)}, \quad j = 1, 2, \dots, h. \quad (2)$$

$$o_k = \sum_{j=1}^h (W_{jk} \times S_j) - \theta'_k, \quad j = 1, 2, \dots, h. \quad (3)$$

Після обчислення значень прихованих вузлів, кінцевий вихід може бути визначений таким чином:

$$O_k = \frac{1}{1 + \exp(-o_k)}, \quad k = 1, 2, \dots, m, \quad (4)$$

де W_{jj} – вага, що з'єднує j -й вузол (у прихованому шарі) з k -м вузлом (вихідний шар), а θ_k представляє k -й вузол зміщення (вихідному шарі).

4.2. Алгоритм рою сальпів

У цій моделі населення сальпів ділиться на дві основні категорії: лідер і послідовник. Лідер – це перший сальп у ланцюжку, а інші сальпи вважаються послідовниками. У цьому алгоритмі сальпи розташовуються у n -мірному просторі пошуку, де n – число проблемних змінних. В результаті, позиції сальпів зберігаються у двовимірній матриці під назвою x , а матриця F розглядається як вихідна харчова матриця або ціль у задачах оптимізації. З кожною ітерацією, відбувається оновлення позиції лідера:

$$[(x_j)^l] = \begin{cases} F_j + c_l((ub_j - lb_j)c_2 + lb_j) & c_3 \geq 0 \\ F_j - c_l((ub_j - lb_j)c_2 + lb_j) & c_3 < 0 \end{cases}, \quad (5)$$

$$c_l = 2e^{\left(\frac{-(4l)}{L}\right)^2}, \quad (6)$$

де l позначає поточну ітерацію, а L – максимальну кількість ітерацій. Рівняння (7) використовується для оновлення послідовників:

$$x_j^i = \frac{1}{2}a \times t^2 + v_0t, \quad (7)$$

$$x_j^i = \frac{1}{2}(x_j^i + x_j^{i-1}). \quad (8)$$

Як було сказано раніше, ub_j і lb_j позначають верхню межу та нижню межі j -го виміру, відповідно. Якщо якийсь із пошукових агентів виходить за межі простору пошуку, він буде повернутий назад. Крім того, F визначається як джерело їжі. Якщо джерело їжі розглядається як кінцева мета або кінцевий оптимум, ця проблема розглядатиметься як проблема оптимізації. Проте, кінцевий оптимум оптимізаційних завдань невідомий. У цьому випадку, найкраще рішення розглядається як кінцевий оптимум та матриця F .

В цьому випадку, використання алгоритму рою сальпів дозволило вирішити проблему низької точності класифікації та застрягання в локальному мінімумі. Серед недоліків можна відмітити те, що вказана модель багатопарового перцептронна не є глибокою, тоді як при роботі зі значними вибірками краще працювати з глибокими нейромережами.

Табл. 1. Порівняння характеристик нейронних мереж

	Кооперативний метод глибокого навчання	Модифікована мережа LeNet	Розріджений складений автоенкодер з багатодоменим набором характеристик	Багат шаровий перцептрон, навчений з використанням алгоритму рою сальпів
Адаптована до роботи з даними, що записані у складних умовах навколишнього середовища	+	–	+	+
Робота з даними, що мають різні значення сигнал/шум	–	–	+	+
Висока точність класифікації	–	+	+	+
Значне навантаження процесора	–	+	+	–
Робота зі великими обсягами даних	+	–	+	–

Висновки

Модифікована нейромережа LeNet показала задовільну точність класифікації при роботі з даними за звичайних умов, але не досліджена робота нейромережі в умовах різних значень сигнал/шум та в складних умовах навколишнього середовища, що не дозволяє обрати її в якості можливого рішення. Кооперативний метод глибокого навчання показує хороші результати в складному та змінюваному середовищі, але нижча точність класифікації в порівнянні з іншими розглянутими підходами та не досліджена робота нейромережі з даними, що мають різні значення сигнал/шум є значними мінусами. Багат шаровий перцептрон, навчений з використанням алгоритму рою сальпів відповідає більшості вимог, але при роботі зі значними об'ємами даних краще використовувати більш глибокі архітектури. Тестування роботи нейромережі на даних, що були згенеровані в штучних умовах ставить під питання ефективність роботи нейромережі з даними, що записані в реальних умовах. Розріджений складений автоенкодер показує задовільні результати при різних значеннях сигнал/шум та на різних глибинах. Також, він є достатньо глибоким для роботи зі значним об'ємом даних. Недоліки, що полягають у ви-

можливості до ресурсів комп'ютера та часі навчання нейромережі не є ключовими, що дозволяє зазначити цей підхід як можливе рішення поставлених задач. Порівняння методів виконане в таблиці 1.

Перелік використаних джерел

1. Wu H, Song Q, Jin G. Underwater acoustic signal analysis: preprocessing and classification by deep learning // Neural Network World. — 2020. — Vol. 30, no. 2. — P. 85–96.
2. A new cooperative deep learning method for underwater acoustic target recognition / Yang Honghui, Xu Guanghui, Yi Shuzhen, and Li Yiqing // OCEANS 2019-Marseille / IEEE. — 2019. — P. 1–4.
3. Underwater target classification at greater depths using deep neural network with joint multiple-domain feature / Cao Xu, Zhang Xiaomin, Togneri Roberto, and Yu Yang // IET Radar, Sonar & Navigation. — 2019. — Vol. 13, no. 3. — P. 484–491.
4. Khishe Mohammad, Mohammadi Hassan. Passive sonar target classification using multi-layer perceptron trained by salp swarm algorithm // Ocean Engineering. — 2019. — Vol. 181. — P. 98–108.