УДК 621.3

Н.О. Матвєєва, Ю.В. Лазоренко

ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ ОПТИМІЗАЦІЇНАВЧАННЯ ДЛЯ РОЗВ'ЯЗАННЯ ЗАДАЧІ КЛАСИФІКАЦІЇ СИГНАЛІВ

Анотація. Представлені результати дослідження алгоритмів зворотного розповсюдження помилки для навчання нейронних мереж, які виконують класифікацію сигналіву дефектоскопії. Моделювання багатошарового персептрона виконувалось в середовищі MATLAB.

Ключові слова:композитні матеріали, нейронні мережі, багатошаровий персептрон, алгоритми оптимізації навчання.

Вступ. Проблема класифікації є однією з найбільш часто виникаючих і розв'язуваних задач як при наукових дослідженнях, так і на практиці. Використання для цього нейронних мереж напряму пов'язане з вирішенням задачі їх навчання. Модель навчання нейронних мереж на основі корегування помилок є однією з основних [1,2].

При проведені неруйнівного контролю композитних матеріалів слід враховувати їх складний рельєф поверхні. Технологія виготовлення волокнистих композитів, звичайно, не передбачає механічну обробку, що ускладнює процес сканування поверхні та додає різні види шумів. Виникає задача – аналізуючи оброблювані сигнали, необхідно отримати інформацію щодо наявності та розмірів дефектів. Для розв'язання таких задач використовуються нейронні мережі [1, 2], які активно розвиваються останнім часом, володіють універсальними та адаптивними властивостями й забезпечують високу ефективність розпізнавання.

Метою роботиє знаходження оптимального алгоритму навчання багатошарового персептронапри розв'язанні задачі класифікації сигналів дефектоскопії.

Основна частина. Кожна штучна нейронна мережа являє собою множину простих елементів – нейронів, які сполучені певним чином. Конкретний вигляд виконуваного мережею перетворення даних обу-

 ISSN 1562-9945

мовлюється не тільки характеристиками нейронів, які входять до її структури, але і особливостями її архітектури, а саметопологією міжнейронних зв'язків, напрямом і способами передачі інформації між нейронами, а також засобами навчання мережі.

При проведені аналізу вхідних і вихідних даних значення ваги та зсуву нейронної мережі автоматично налагоджуються так, щоб мінімізувати різницю між бажаним сигналом та отримуваним на виході в результаті моделювання. Ця різниця називається помилкою навчання і для конкретної конфігурації нейронної мережі визначається шляхом пропускання через мережу всіх спостережень, які маються, та порівняння вихідних значень с бажанім, цільовим значенням. Тобто формується функція помилки (критерій якості навчання)..

Багатошарові нейронні мережі прямого розповсюдження являють собою нелінійні системи, які дозволяють краще кваліфікувати ніж звичайні статистичні методи. Багатошаровий персептрон (multilayerperceptron - MLP) складається з множини вхідних вузлів, які створюють вхідний шар, один або декілька скритих шарів й одного вихідного шару. При навчанні MLP використовується алгоритм зворотного розповсюдження помилки (back-propagationlearning) [3]. Кожний нейрон MLP, який навчається на основі зворотного розповсюдження має нелінійну гладку функцію активації, в якості якої часто використовують нелінійну сигмоїдальну функцію активації типа логістичної або гіперболічного тангенса

Алгоритми зворотного розповсюдження помилки аналогічні методам знаходження екстремуму функції декількох змінних і поділяються на тригрупи – нульового, першого та другого порядку [3, 4]. В алгоритмах нульового порядкудля знаходження екстремуму використовується тільки інформація о значеннях функції у заданих точках.

Алгоритми оптимізації навчання є стратегіями, заснованими на реалізації ідеї ітеративного спуску, які забезпечують мінімізацію функціонала навчання. У процесі роботи алгоритмів, як правило, виникає задача одновимірного пошуку мінімуму уздовж заданого напрямку. Це можуть бути антиградієнті або сполучені напрямки [4].

Градієнтні алгоритми навчання є специфічною реалізацією градієнтного спуску в просторі вагових коефіцієнтів і зміщень MLP і забезпечують рух по поверхні функціоналу помилки в напрямку, протилежному вектору градієнта:

$$x_{k+1} = x_k - \alpha_k g_k \,, \tag{1}$$

де x_k – вектор вагових коефіцієнтів; α_k – параметр швидкості навчання; g_k– вектор градієнта функціонала похибки.

Метод сполучених градієнтів дозволяє визначити необхідний мінімум набагато швидше. Всі алгоритми цього методуна першій ітерації починають пошук у напрямку антиградієнту:

$$p_0 = -g_0. \tag{2}$$

Для визначення розміру кроку уздовж сполученого напрямкувиконуються спеціальні одномірні процедури пошуку мінімуму. Коли вибрано напрямок спуску, потрібно визначити оптимальну відстань (крок пошуку), на величину якого слід змінити настроюванні параметри:

$$x_{k+1} = x_k + \alpha_k p_k. \tag{3}$$

Потім визначається наступний напрямок пошуку як лінійнакомбінація нового напрямку найшвидшого спуску і вектора руху:

$$p_k = -g_k + \beta_k p_{k-1}\,. \tag{4}$$

Різніалгоритми методу сполученого градієнта розрізняються способом обчислення константи $\beta_к$. Для всіх алгоритмів методу сполучених градієнтів напрямок пошуку періодично перевстановлюється заново на напрямок антиградієнту (рестарт). Це відбувається, коли виникають проблеми зі збіжністю. Цекоштовна в обчислювальному відношенні процедура.

Методи другого порядку вимагають знання других похідних функціоналу помилки. К цим методам відноситься метод Ньютона. Основний крок метода Ньютона знаходиться по формулі:

$$x_{k+1} = x_i - H_k^{-1} g_k\,, \tag{5}$$

де x_k – вектор значень параметрів на k-ї ітерації; H– матриця других часткових похідних цільової функції, або матриця Гессе; g_k– вектор градієнта на k-ї ітерації. Метод Ньютонав багатьох випадках сходиться скоріше, ніж методи сполученого градієнта, але потребують великих затрат через обчислення гессіана. Для того, щоб уникнути обчислення матриці Гессен, пропонуються різні засоби її заміщення приблизними виразами, це породжує так звані квазіньютонові алгоритми

 ISSN 1562-9945

(наприклад, алгоритм метода січних плоскостей або алгоритм Левенберга-Марквардта) [3, 4].

Навчання нейронної мережі припиняється при виконані однієї з умов: значення функції якості навчання стало менше граничного; градієнт критерію якості став менше; досягнуто граничне число циклів навчання; перевищено максимальний час, виділений на навчання.

Експериментальні дослідження. При проведені сканування композитних матеріалів за допомогою вихорострумового перетворювача отримуються триформи сигналів унімодальний, пологий унімодальний та бімодальній. Унімодальний сигнал з максимальною амплітудою характеризує дефекти, які перевищують зону контролю, а бімодальні з найбільшим провалом вершини належать точковим дефектам.

Для моделювання використовувалось обчислювальне середовище MATLAB R2010b. У якості навчальної множини для нейронної мережі запропоновані значення функцій, які відповідають модельним сигналам, одержаним присkануванні поверхні композитів [5], у точках $x = -2, -1.9, ..., 2$:

$$y(x) = \exp(-1{,}5x^2) - k \cdot \exp(-3x^2)\,, \tag{6}$$

де k змінюється від 0 до 1: при k = 0-0.35 одержуємо вузький унімодальний сигнал, котрий характерний для довгої тріщини, довжина якої перебільшує зону контролю. При зміні k = 0.35-0.55 отримуємо положистий унімодальний сигнал, характерний для тріщин меншої розмірності. Беручи k = 0.6-1 дістаємо бімодальний сигнал, який мають маленькі тріщини (при k = 1 – точковий дефект).

Для вирішенняпоставленого завданнявикористовувався багатошаровий пересептрон з 21нейроніву вхідномушарі(за кількістю компонентвхідноговектора), 10 нейронів у скритому шарі та 3 нейрониу вихідномушарі(за кількістю компонентвихідноговектора, тобто маємо три форми сигналу).

Для створення нейронної мережі прямого поширення застосовували функцію*feedforwardnet* (*hiddenSizes,trainFcn*), де *hiddenSizes* – вектор-рядок з одного або декількох розмірів прихованого шару; *trainFcn* – функція навчання.

Спочатку дослідження проводились з алгоритмами навчання першого та другого порядку [3, 4] на ідеальних сигналах без шуму (13 варіантів).На скритому та вихідному шарах застосовували в якості функцій активацій лінійну функцію (*purelin*), гіперболічну тангенціальну функцію (*tansig*) та логістичнусигмоїдальну функцію (*logsig*) в різних комбінаціях[6].

Кращі результати показали такі алгоритми навчання методу сполучених градієнтів*: traincgf, traincgp, traincgb, trainscg*; та квазіньютонові алгоритми з функціями:*trainlm*і*trainbr*.В якості функцій активацій застосовувались лінійна функція активації (*purelin*), гіперболічна тангенціальна функція (*tansig*) та логістична сигмоїдальна функція (*logsig*).

Для перевірки якості навчання проведено два експериментальних дослідження. У першому випадку навчання проводилось на сигналах без шуму, а тестування мережі – на зашумлених сигналах. Для другого дослідження навчання й тестування нейронних мереж відбувалось на зашумлених сигналах. Для моделювання зашумленнясигналів використовувався адитивний білий гауссів шум, який описується функцією*awgn(x, snr, 'measured'),* де *x* - вектор сигналу, скаляр *snr* задає відношення сигнал/шум в децибелах

$$snr(dB) = 20\log\left(\frac{A_{signal}}{A_{noise}}\right), \quad (7)$$

де A – середньоквадратичне значення амплітуди.

При проведені досліджень *snr*приймало значення 100, 50 і 30 децибел. Після проведення досліджень результати заносились у матрицю результатів. Оскільки нейронна мережа з тими самими функціями навчання та активації нейронів не даєоднаковий результат, якщо навчати її заново, то використовувалась середня похибка від результатів.

Дослідження для кожного алгоритму навчання та різних комбінацій функцій активації проводились 50 разів та підраховувалось середнє значення похибки при класифікації сигналів.

В таблицях 1-3 наведені середні помилки, отримані в процесі навчання нейронних мереж на ідеальних сигналах, а тестуванні – на сигналах з шумом.

 ISSN 1562-9945

Таблиця 1

Середні помилки при відношенні сигнал/шум *snr* =100Дб

Функ. тренування	trainlm	traincgb	trainscg	trainbr	traincgf	traincgp
Функ. активації						
'logsig','purelin'	0,021837	0,078571	0,075158	0,000227	0,046513	0,059822
'tansig','purelin'	0,02526	0,058195	0,072254	0,000543	0,073337	0,055787
'logsig','tansig'	0,020564	0,117692	0,142242	0,004949	0,176836	0,094574
'tansig','tansig'	0,072658	0,154414	0,107719	0,038328	0,148797	0,135683

Таблиця 2

Середні помилки при відношенні сигнал/шум *snr*= 50Дб

Функ. тренування Функ. активації	trainlm	traincgb	trainscg	trainbr	traincgf	traincgp
'logsig','purelin'	0,024717	0,041007	0,055923	0,001867	0,041415	0,074117
'tansig','purelin'	0,026983	0,131753	0,083315	0,002014	0,046484	0,038735
'logsig','tansig'	0,025729	0,151271	0,102915	0,004575	0,223883	0,136044
'tansig','tansig'	0,048479	0,131611	0,136155	0,037427	0,059263	0,110237

Таблиця 3

Середні помилки при відношенні сигнал/шум *snr* =30Дб

Функ. тренування	trainlm	traincgb	trainscg	trainbr	Traincgf	traincgp
Функ. активації						
'logsig','purelin'	0,027305	0,08119	0,04343	0,016679	0,045204	0,065496
'tansig','purelin'	0,035913	0,068405	0,148655	0,01662	0,077969	0,051315
'logsig','tansig'	0,043969	0,150888	0,184603	0,018301	0,163004	0,106515
'tansig','tansig'	0,084189	0,091729	0,130583	0,112113	0,148148	0,167698

Кращі результати показує нейронна мережа з функцією тренування trainbrта з функціями активації 'logsig', 'purelin'або 'tansig', 'purelin'.

Результати другого дослідження (навчання та тестування проводилось на зашумлених сигналах) наведені в таблицях 4-6.

Кращі результати показує знов нейронна мережа з функцією тренування trainbr та з функціями активації 'logsig', 'purelin' й 'tansig', 'purelin'.

Таблиця 4

Середні помилки при відношенні сигнал/шум *snr* =100Дб

Функ.тренування	trainlm	traincgb	trainscg	trainbr	Traincgf	traincgp
Функ. активації						
'logsig','purelin'	0,051722	0,082656	0,049814	0,000241	0,052541	0,037667
'tansig','purelin'	0,025354	0,121009	0,076401	0,000325	0,06339	0,042976
'logsig','tansig'	0,023918	0,094399	0,100799	0,003348	0,180841	0,167962
'tansig','tansig'	0,04066	0,198937	0,13036	0,00388	0,140132	0,051592

Таблиця 5

Середні помилки при відношенні сигнал/шум *snr* =50 Дб

Ф-я тренування	trainlm	traincgb	trainscg	trainbr	Traincgf	traincgp
Ф-ії активації						
'logsig','purelin'	0,02032	0,03698	0,034756	0,001268	0,06526	0,057423
'tansig','purelin'	0,04931	0,188069	0,078399	0,001346	0,066582	0,094244
'logsig','tansig'	0,036197	0,151694	0,14985	0,003933	0,082759	0,106405
'tansig','tansig'	0,072793	0,047446	0,175436	0,072777	0,128994	0,118038

Таблиця 6

Середні помилки при відношенні сигнал/шум *snr* =30 Дб

Ф-я тренування	trainlm	traincgb	trainscg	trainbr	Traincgf	traincgp
Ф-ії активації						
'logsig','purelin'	0,023128	0,068147	0,043215	0,013492	0,04199	0,034587
'tansig','purelin'	0,036833	0,078169	0,100336	0,014493	0,182083	0,072564
'logsig','tansig'	0,017595	0,24615	0,242536	0,043339	0,166891	0,11076
'tansig','tansig'	0,072348	0,074369	0,151991	0,043805	0,097286	0,209213

Функція *trainbr* виконує процедуру навчання при умові, що функції зважування, нагромадження та активації мають похідні. Ця функція реалізована на базі алгоритмуЛевенберга-Марквардта, котрий відноситься до самих ефективних та швидкодіючих. Додаткове використання методу регуляризації за Байєсом дозволяє знаходити оптимальні значення для параметрів, які налагоджуються, і як наслідок, знаходження оптимального розміру мережі. Це дозволяє успішно справитися з проблемою перенавчання. Для цього здійснюється мінімізація комбінованого функціоналу якості навчання, яке враховує не тільки суму квадратів помилок навчання, але і квадрати ваги.

 ISSN 1562-9945

Висновки.Порівняльний аналіз алгоритмів оптимізації навчання нейронних мереж показав, що кращі результати отримуємо при навчанні мереж на сигналах з шумом. Для цього рекомендовано використовувати функцію тренування *trainbr*сумісно з функціями активації нейронів *tansig*', '*purelin*'або '*logsig*', '*purelin*'.

ЛІТЕРАТУРА

1. Хайкин Саймон. Нейронные сети: полный курс, 2-е издание.: Пер. с англ./СаймонХайкин– М.: «Вильямс». 2006.
2. Аксенов С.В. Организация и использованиенейронныхсетей (методы и технологии) / С.В. Аксенов, В.Б. Новосельцев – Томск: НТЛ. 2006. – 128 с.
3. Медведев В.С. Нейронные сети. MATLAB 6 / В.С. Медведев, В.Г. Потемкин – М.: ДИАЛОГ-МИФИ, 2002. – 496 с.
4. Дьяконов В., Круглов В. Математические пакеты расширения MATLAB. Специальный справочник. – СПб.: Питер, 2001. – 480 с.
5. Хандецкий В.С. Спектральная идентификация сигналов в дефектоскопии композитов с использованием теории статистических испытаний / Хандецкий В.С., Герасимов В.В. //Вісник ДНУ: Фізика. Радіоелектроніка. – Дніпропетровськ: – 2003. № 10. – С. 128 – 132.
6. Матвеева Н.А. Моделирование нейросети для решения задачи классификации в дефектоскопии // Системні технології. Регіон. міжвуз. зб. наук. праць. - Дніпропетровськ: ДНВП «Системні технології», 2011. -Вип. 1(72). - С. 37-44.