

## НЕЙРОННІ МЕРЕЖІ ПРИ ОБРОБЦІ ЗОБРАЖЕНЬ В ІНТЕЛЕКТУАЛЬНИХ СИСТЕМАХ СПОСТЕРЕЖЕННЯ

**Фастовець В.І.**, доцент,

Харківський національний автомобільно-дорожній університет  
магістр факультету Інфокомунікацій  
Харківський національний університет радіоелектроніки, Україна

**Шуляков В.М.**, доцент

Харківський національний університет радіоелектроніки, Україна

***Анотація.** Проведений огляд нейронних мереж для вибору інтелектуальних систем відеоспостереження на їх основі.*

*Ключові слова: нейронні мережі, нейрон, синапс, Згорткові нейронні мережі, рекурентні нейронні мережі, навчання нейронної мережі.*

Мозок та нервова система людини складається з клітин, які називаються нейронами, і має надзвичайну складність. Унікальною здатністю нейрона є прийом, обробка і передача електрохімічних сигналів по нервовим шляхам. Від тіла нейрона до інших нейронів ідуть дендрити (деревоподібні відростки на входах нейрона). Вони приймають сигнали в точках з'єднання, які називаються синапсами.

Прийняті синапсом входні сигнали підводяться до тіла нейрона, де вони підсумовуються. Нейрон збуджується і посиляє по спеціальному відростку, так званому аксону, сигнал іншим нейронам.

Більшість нейронних мереж складаються з формальних нейронів (рис. 1). Цей примітивний обчислювальний пристрій (або його модель), що має кілька входів і один вихід, і є основним обчислювальним елементом нейромережі.

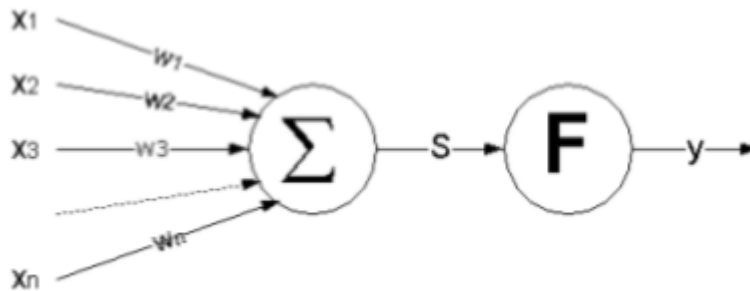


Рис. 1 Формальний нейрон.

Тут  $x_1..x_n$  - значення, що надходять на входи (синапси) нейрона,  
 $w_1..w_n$  - ваги синапсів (можуть бути як гальмуючі, так і підсилюючі),  
 $S$  - зважена сума входних сигналів.

$$S = \sum_{i=1}^n W_i X_i - T = (W, T) - T = |W| * |X| * \cos \alpha - T ,$$

де  $T$  - поріг нейрона,

$F$  - функція активації нейрона, що перетворює зважену суму в вихідний сигнал,  
 $\alpha$  - кут між векторами (вхідним вектором і вектором ваг).

Штучна нейронна мережа складається з набору пов'язаних формальних нейронів і здатна виконувати логічні операції і перетворення, реалізовані дискретними пристроями з кінцевої пам'яттю. Формальний нейрон моделює деякі властивості біологічного нейрона. Нейрони в природних нейронних мережах набагато складніші. Мозок може мати квантову структуру, а процес мислення може бути заснований на

квантових ефектах. Мозок людини містить 10 трильйонів нейронів, пов'язаних між собою більше ніж 1000 синапсами. Зв'язки між нейронами розвиваються і модифікуються протягом усього життя, тобто свій інтелектуальний досвід людина отримує в процесі навчання. Все це свідчить про перспективність розвитку штучних нейронних мереж.

Кожна нейронна мережа включає перший вхідний шар нейронів. Цей шар не виконує обчислень та перетворень, він приймає вхідні сигнали та розподіляє їх по іншим нейронам.

За структурою нейронні мережі розподіляються на такі типи:

1) Одношарова нейронна мережа. Сигнали з вхідного шару направляються на вихідний шар, який, перетворює сигнал і відразу видає відповідь. Вхідні нейрони є об'єднаними з основним шаром за допомогою синапсів з різними вагами для забезпечення якості зв'язків.

2) Багатошарова нейронна мережа. Тут, крім вихідного та вхідного шарів, є ще проміжні шари. Число цих шарів залежить від ступеня складності нейронної мережі. Вона вже більше нагадує структуру біологічної нейронної мережі. В процесі обробки даних кожен проміжний шар — це проміжний етап, на якому здійснюється обробка та розподіл інформації.

Нейронні мережі також можна класифікувати за напрямом розподілу інформації по синапсам між нейронами:

1) Нейромережі прямого поширення, або односпрямовані. У таких мережах сигнал переміщається від вхідного шару до вихідного. Сьогодні розробки односпрямованих нейромереж широко поширені і успішно вирішують завдання розпізнавання образів, кластеризації і прогнозування.

2) Рекурентні нейронні мережі (Recurrent Neural Network, RNN) - це мережі зі зворотніми зв'язками. Сигнал може рухатися і у прямому, і у зворотному напрямку. Результат виходу здатний повертатись на вхід. Вихід нейрона визначається ваговими характеристиками та вхідними сигналами, та може доповнюватись попередніми виходами, які знову повернулися на вхід. Цим нейромережам властива функція короткочасної пам'яті.

3) Радіально-базисні функції.

4) Картки, що самоорганізуються.

Але це далеко не всі варіанти класифікації та види нейронних мереж. Також їх розподіляють:

1) Залежно від типів нейронів: однорідні та гібридні.

2) За типом вхідної інформації: аналогові, двійкові та образні.

3) За характером налаштування синапсів: з фіксованими та динамічними зв'язками.

За характером навчання і формування зв'язків нейронні мережі можуть бути наступних видів:

– Навчання з учителем. Заздалегідь відомі еталонні значення результатів, зв'язки налаштовуються в процесі навчання.

– Самонавчання (навчання без вчителя). Еталонні значення результатів невідомі, мережа в процесі навчання повинна організувати вхідні образи на основі їх подібності.

– Фіксований зв'язок. Визначаються характером розв'язуваної задачі.

Також можна класифікувати нейронні мережі, як:

– Нейронні мережі прямого поширення (Feed Forward Neural Networks, FFNN).

– Згорткові нейронні мережі (Convolutional neural network, CNN).

– Рекурентні нейронні мережі (Recurrent Neural Network, RNN).

Для розпізнавання відео- і аудіо-потоків зазвичай використовують рекурентні мережі, а для розпізнавання образів згорткові мережі (наприклад, LeNet-5 у Франції та неокогнітрон у Японії).

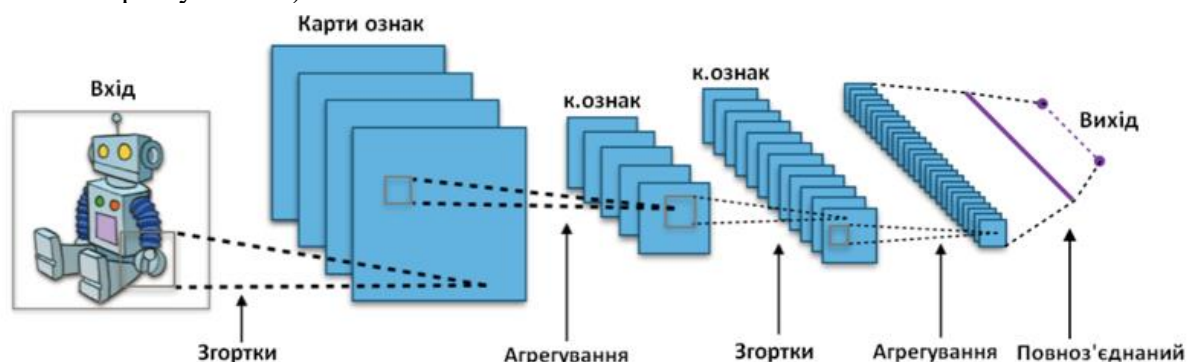


Рис 2. Типова архітектура загорткової нейронної мережі

У звичайному перцептроні, який представляє собою повнозв'язну нейронну мережу, кожен нейрон пов'язаний з усіма нейронами попереднього шару, причому кожний зв'язок має свій персональний ваговий коефіцієнт. У згортковій нейронній мережі в операції згортки використовується лише обмежена матриця ваг невеликого розміру, яку також називають ядром згортки. Її інтерпретують як графічне кодування якої-небудь ознаки, наприклад, наявності похилої лінії під певним кутом. Наступний шар, що вийшов в результаті операції згортки такою матрицею ваг, показує наявність даної ознаки в оброблюваному шарі і її координати, формуючи так звану карту ознак (англ. Feature map). Звісно, в згортковій нейронній мережі набір ваг не один, а ціла гама, що кодує елементи зображення (наприклад, лінії і дуги під різними кутами). При цьому ядра згортки не закладаються заздалегідь, а шляхом навчання мережі класичним методом зворотного поширення помилки формуються самостійно. Прохід кожним набором ваг формує свій власний примірник карти ознак, при цьому на одному шарі формується багато незалежних карт ознак, що робить нейронну мережу багатоканальною.

Згорткова нейронна мережа складається з шарів входу та виходу, а також із спеціалізованих прихованих шарів, які складаються зі згорткових шарів, агрегувальних шарів, повноз'єднаних шарів та шарів нормалізації.

Згортковий шар (convolutional layer) застосовує до входу операцію згортки, передаючи результат до наступного шару:

$$Z^l = h^{l-1} * W^l,$$

де  $Z^l$  – постсинаптичний рівень збудження нейронів  $l$ -го шару нейромережі,

$h^{l-1}$  – активаційний рівень нейронів  $(l-1)$ -го шару нейромережі,

$W^l$  – налагоджуваний параметр  $l$ -го шару нейромережі,

\* – дискретний оператор згортки.

Параметри шару складаються з набору ядер, або фільтрів для навчання, які мають досить вузьке рецептивне поле, але простягаються на всю глибину вхідної ємності. При прямому проходженні кожен фільтр обчислює скалярний добуток даних входу та фільтру, здійснює згортку за шириною та висотою вхідної ємності, і формує двовимірну карту збудження цього фільтру.

При переборі шаром матриці ваг її пересувають зазвичай не на розмір цієї матриці, а на невелику відстань. Так, наприклад, при розмірності матриці ваг  $6 \times 6$  її зрушують на один або два нейрона (пікселя) замість шести, щоб не пропустити шукану ознаку.

Далі виконується зменшення розмірності сформованих карт ознак за допомогою операції підвибірки, або сабдискретизації (англ. Subsampling, Pooling), виконується зменшення розмірності карт ознак, що були сформовані на попередньому етапі. Далі приймається положення, що інформація про факт наявності шуканої ознаки важливіше точного знання її координат, тому з кількох сусідніх нейронів карти ознак вибирається максимальний і приймається за один нейрон ущільненої карти ознак меншої розмірності. Після цього мережа стає більш інваріантною до масштабу вхідного зображення.

Початковий шар нейромережі приймає вхідне зображення, потім сигнал проходить серію згорткових шарів, в яких чергується власне згортка і субдискретизація (пулінг). В результаті обробки на кожному наступному шарі карта ознак зменшується в розмірі, натомість збільшується кількість каналів, що означає здатність розпізнавання складних ієрархій ознак. Після проходження декількох шарів карта ознак може виродитися у вектор або скаляр, але таких карт ознак стає дуже багато. На виході згорткових шарів мережі додатково встановлюють кілька шарів повнонейронної мережі, або перцептрон, на вхід якого подаються кінцеві карти ознак.

Максимізаційне агрегування (max pooling) використовує максимальне значення з кожного з кластерів нейронів попереднього шару. Воно розділяє вхідне зображення на набір прямокутників без перекриттів, і для кожної такої підобласті виводить її максимум:

$$h_{x,y}^l = \max_{i,j} \{h_{(x+i)(y+j)}^{l-1}\}, i = 0, \dots, s, j = 0, \dots, s,$$

де  $h_{x,y}^l$  – рівень активації нейрона з індексами (x, y) l-го шару нейромережі.

Існує спосіб пулінга, званий агрегуванням областей інтересу (Region of Interest pooling, RoI pooling) – це різновид максимізаційного агрегування, в якому розмір виходу фіксовано, а прямокутник входу є параметром.

Агрегування, яке використовує усереднене значення з кожного з кластерів нейронів попереднього шару, називається усереднювальним агрегуванням (average pooling):

$$h_{x,y}^l = \frac{1}{s^2} \sum_{i=0}^s \sum_{j=0}^s h_{(x+i)(y+j)}^{l-1}$$

L -нормове агрегування ( $L^2$  -norm pooling) базується на Евклідовій відстані:

$$h_{x,y}^l = \sqrt{\sum_{i=0}^s \sum_{j=0}^s (h_{(x+i)(y+j)}^{l-1})^2}$$

Шар зрізаних лінійних вузлів (Rectified Linear Units layer, ReLU) застосовує ненасичувальну передавальну функцію ReLU і посилює нелінійні властивості функції ухвалення рішення і мережі в цілому. При цьому цей шар не зачіпає рецептивних полів згорткового шару. Інші функції, такі як насичувальний гіперболічний тангенс та сигмоїдна функція, також можуть використовуватись для посилення нелінійності.

Згладжувальний шар (flatten layer) згортає просторові розмірності входу в розмірність каналу.

Повноз'єднаний шар (fully connected layer) забезпечує високорівневий зв'язок кожного нейрона одного шару з кожним нейроном наступного шару:

$$Z^l = W^l h^{l-1}.$$

У багатьох нейромережах завершальним є шар втрат (loss layer), який визначає, як навчання штрафує відхилення між передбаченими та справжніми мітками класів.

Основними рисами згорткових нейромереж є:

1) Просторова організація, тобто згорткові шари мають нейрони, впорядковані у трьох вимірах: ширина, висота та глибина; нейрони всередині шару з'єднані лише з невеликою областю попереднього шару, яка називається рецептивним полем.

2) Для формування архітектури ЗНМ складають різні типи шарів, як локально-, так і повноз'єднані.

3) Для кожного рецептивного поля шару використовується один і той же банк ваг, або фільтр. Кожен фільтр повторюється на всьому зоровому полі, а повторні вузли використовують спільну параметризацію (вектор ваг та упередженості) й формують карту ознак. Це означає, що всі нейрони в заданому згортковому шарі реагують на одну й ту ж саму ознаку в межах свого рецептивного поля. Повторювання вузлів таким чином дозволяє ознакам бути виявленими незалежно від їхнього положення в зоровому полі, забезпечуючи таким чином властивість інваріантності відносно зсуву.

4) У шарах об'єднання карти ознак поділяються на прямокутні підрегіони, а ознаки кожного прямокутника зменшуються до одного значення, (середнього або максимального). Спільне використання ваг різко зменшує кількість вільних параметрів, на яких вчиться мережа. Це зменшує обсяг необхідної пам'яті та поліпшує продуктивність згорткової нейромережі.

5) Навчання згорткових нейромереж найчастіше здійснюють шляхом використання техніки зворотного поширення помилки. Для навчання глибинної мережі необхідні великий набір даних та великі обчислювальні потужності.

Досить швидко та ефективно можуть ідентифікувати та обробляти дані зображення згорткові нейронні мережі. Експерти Facebook AI Research (FAIR) створили глибокого навчання під назвою Mask R-CNN. Для кожного об'єкта, який присутній на зображенні, вона може створити піксельну маску. Ця архітектура є розширеною версією архітектури виявлення об'єктів Faster R-CNN.

FAST R-CNN використовує координати рамки та клас об'єкта в якості двох частин даних для кожного об'єкта на зображенні. Mask R-CNN надає додатковий поділ та виводить маску об'єкта після виконання сегментації. Спочатку ConvNet, до якої передається вхідне зображення, генерує карту об'єктів для зображення. Потім до карт об'єктів система застосовує мережу регіональних ознак (RPN) і генерує їх з оцінками об'єктивності. До ознак застосовується шар об'єднання для приведення їх до одного розміру. На останньому етапі система передає ознаки повнозв'язному шару для класифікації та генерує вихід з обмежувачими контурами для кожного об'єкта[1].

## Література

1. Субботін С. О. Нейронні мережі : теорія та практика: навч. посіб. / С. О. Субботін. – Житомир : Вид. О. О. Євенок, 2020. – 184 с.