

УДК 621.384

С.А. Олізаренко, В.А. Капранов, Р.В. Сафронов

Харківський національний університет Повітряних Сил імені Івана Кожедуба, Харків

РОЗРОБКА АРХІТЕКТУРИ НЕЧІТКОЇ ЗГОРТОЧНОЇ НЕЙРОННОЇ МЕРЕЖІ ДЛЯ РОЗПІЗНАВАННЯ КОМПАКТНИХ (ТОЧКОВИХ) ОБ'ЄКТІВ НА ЦИФРОВОМУ АЕРОФОТОЗНІМКУ

В роботі запропонована архітектура ансамблю нейронних мереж на базі нейромережевого детектора, згорточної нейронної мережі і модифікованої нечіткої нейронної мережі для розпізнавання компактних (точкових) об'єктів.

Ключові слова: згорточна нейронна мережа, нечітка нейронна мережа, детектор, підмережа вилучення ознак, класифікатор, карта ознак.

Вступ

Постановка проблеми (завдання). За останній час найкращі результати по розпізнаванню об'єктів на цифрових знімках дають автоматизовані системи, реалізовані на основі згорточних нейронних мереж (ЗНМ). Існуючі архітектури ЗНМ розрізняються як елементами структури (шарами, глибиною), так і функціональністю. Найефективнішою ЗНМ за критерієм оперативності і точності, що забезпечує одночасне виявлення і класифікацію об'єктів, є нейромережевий детектор SSD (Single Shot MultiBox Detector) [1]. У якості складової для вилучення ознак в SSD використовується ЗНМ VGG. У той же час більш глибокі ЗНМ з точки зору кількості шарів, наприклад ЗНМ на базі модулів Inception, забезпечують формування більш глибокої ієрархії карт ознак, що в свою чергу в подальшому безпосередньо впливає на ефективність розпізнавання. У якості класифікатора SSD виступають повнозв'язні шари, які перетворені в згорткові. Однак дослідження показують, що нечіткі нейронні мережі забезпечують більш високу точність розпізнавання порівняно з класичними повнозв'язними мережами [2].

У зв'язку з цим в роботі розроблена архітектура нечіткої ЗНМ Fuzzy Inception-SSD на базі детектора SSD, де в якості підмережі вилучення ознак запропоновано використання ЗНМ Inception-v2, а в якості класифікатора модифікованої нечіткої нейронної мережі Ванга-Менделя.

Аналіз публікацій. В даний час існує значна кількість публікацій, присвячених питанням розробки архітектури ЗНМ для розпізнавання об'єктів на зображення для різних цілей [1; 3–5]. При цьому питання комплексного застосування різних архітектур мереж для виявлення і класифікації компактних (точкових) об'єктів, в тому числі з використанням нечітких нейронних мереж, є мало розглянутими і вимагають подальших досліджень.

Метою статті є розробка архітектури нейромережевого ансамблю на базі згорткових і нечітких

нейронних мереж для розпізнавання компактних (точкових) об'єктів на цифровому аерофотознімку.

Основний матеріал

Особливістю використання ЗНМ Inception-v2 як підмережі вилучення ознак в Fuzzy Inception-SSD є, по-перше, відсутність додаткових класифікаторів, які в оригінальній версії ЗНМ Inception-v2 не впливають на якість розпізнавання і виступають в ролі регуляризатора, по-друге, у введенні додаткових шарів зниження розмірності карт ознак.

Основними принципами, що лежать в основі функціонування підмережі вилучення ознак Fuzzy Inception-SSD, є наступні:

в якості базових елементів підмережі використовуються як окремі типи шарів ЗНМ, і так модулі «inception» трьох видів (сукупність певних чином пов'язаних шарів ЗНМ різних типів і розмірності), що забезпечують зниження кількості параметрів підмережі і формування ознак різного рівня абстракції в рамках одного модуля;

за рахунок комбінування окремих шарів і модулів «inception» забезпечується відсутність різкого зниження розмірності представлення даних підмережі, а також забезпечення балансу між глибиною і шириною підмережі за рахунок рівномірного збільшення або зменшення глибини і ширини за допомогою використання відповідної кількості модулів;

модулі «inception» розглядаються як «широкі» шари, при цьому «широкі» шари навчаються швидше, що особливо важливо на високих рівнях, і локально, що в свою чергу дозволяє після «широкого» шару здійснювати відповідно зниження розмірності нових карт ознак;

модулі «inception» одного типу мають послідовний взаємозв'язок, при цьому останній модуль в послідовності однакових модулів формує карти ознак, які використовуються як для виконання подальшого вилучення ознак в підмережі вилучення ознак Fuzzy Inception-SSD, так і для виявлення і кла-

сифікації об'єктів в рамках функціонування модифікованого детектора ЗНМ SSD.

В якості основних елементів архітектури підмережі вилучення ознак ЗНМ Fuzzy Inception-SSD виступають наступні шари:

згорткові шари (Convolution Layers, CONV-шари). CONV-шар обчислює значення виходів нейронів, які з'єднані з рецептивними полями (фільтрами, що навчаються) вхідного шару, з метою опису ознак (властивостей, особливостей) зображення у вигляді багатовимірного масиву карт ознак;

шари об'єднання (Pooling Layers, POOL-шари). POOL-шар забезпечує зниження просторового розміру зображення (карт ознак) з метою зменшення кількості параметрів і обчислювальної складності мережі;

шари активації: безпосередньо в роботі використовується шар випрямлення (Rectified Linear Unit, шари ReLU). Шар ReLU є лінійним блоком випрямлення і розглядається як функція активації для нейронів CONV-шару;

шари конкатенації або зв'язування (Concatenation, Concat-шар). Concat-шар забезпечує об'єднання різних вихідних багатовимірних масивів карт ознак, отриманих в результаті функціонування шарів модуля «inception», в один.

Формально згортковий шар можна представити у вигляді функції

$$Y_{H^{out}W^{out}D^{out}} = f_{CONV}(X_{H^{in}W^{in}D^{in}}), \quad (1)$$

де $X_{H^{in}W^{in}D^{in}}$ – коваріантний тензор третього рангу, що описує вхідні дані CONV-шару розмірністю $H^{in} \times W^{in} \times D^{in}$; $Y_{H^{out}W^{out}D^{out}}$ – коваріантний тензор третього рангу, що описує вихідні дані CONV-шару розмірністю $H^{out} \times W^{out} \times D^{out}$.

Шар об'єднання формально представляється у вигляді наступних функцій:

- для операції "max"

$$Y_{H^{out}W^{out}D^{out}} = f_{MaxPOOL}(X_{H^{in}W^{in}D^{in}}); \quad (2)$$

- для операції "avg"

$$Y_{H^{out}W^{out}D^{out}} = f_{AvgPOOL}(X_{H^{in}W^{in}D^{in}}). \quad (3)$$

Шар активації реалізує функцію активації ReLU

$$y_{i^{out}j^{out}k^{out}} = \mathbb{S}(x_{i^{in}j^{in}k^{in}}) = \max(0, x_{i^{in}j^{in}k^{in}}),$$

де $y_{i^{out}j^{out}k^{out}} \in Y_{H^{out}W^{out}D^{out}}$; $x_{i^{in}j^{in}k^{in}} \in X_{H^{in}W^{in}D^{in}}$;

$H^{out} = H^{in}$; $W^{out} = W^{in}$; $D^{out} = D^{in}$.

Математично модуль «inception» 1-го типу описується як суперпозиція функцій наступним чином

$$Y_{H^{med1}W^{med1}D^{med1}} = f_{CONV13}(f_{CONV12}(f_{CONV11}(X_{H^{in}W^{in}D^{in}}))); \quad (4)$$

$$Y_{H^{med2}W^{med2}D^{med2}} = f_{CONV22}(f_{CONV21}(X_{H^{in}W^{in}D^{in}})); \quad (5)$$

$$Y_{H^{med3}W^{med3}D^{med3}} = f_{CONV31}(f_{MaxPOOL}(X_{H^{in}W^{in}D^{in}})); \quad (6)$$

$$Y_{H^{med4}W^{med4}D^{med4}} = f_{CONV41}(X_{H^{in}W^{in}D^{in}}); \quad (7)$$

$$Y_{H^{out}W^{out}D^{out}} = Y_{H^{med1}W^{med1}D^{med1}} \circ Y_{H^{med2}W^{med2}D^{med2}} \circ Y_{H^{med3}W^{med3}D^{med3}} \circ Y_{H^{med4}W^{med4}D^{med4}};$$

де \circ – операція конкатенації тензорів карт ознак, для яких виконується наступна умова рівності розмірності

$$H^{med1} = H^{med2} = H^{med3} = H^{med4}; \quad (9)$$

$$W^{med1} = W^{med2} = W^{med3} = W^{med4}.$$

Математично три гілки модуля «inception» 2-го типу описується як суперпозиція функцій наступним чином:

$$Y_{H^{med1}W^{med1}D^{med1}} = f_{CONV15}(f_{CONV14}(f_{CONV13}(f_{CONV12}(f_{CONV11}(X_{H^{in}W^{in}D^{in}}))))); \quad (10)$$

$$Y_{H^{med2}W^{med2}D^{med2}} = f_{CONV23}(f_{CONV22}(f_{CONV21}(X_{H^{in}W^{in}D^{in}}))); \quad (11)$$

$$Y_{H^{med3}W^{med3}D^{med3}} = f_{CONV31}(f_{AvgPOOL}(X_{H^{in}W^{in}D^{in}})). \quad (12)$$

Четверта гілка модуля «inception» 2-го типу описується з використанням виразу (7), результат конкатенації з використанням (8) при виконанні умови відповідно до виразу (9).

Математично три гілки модуля «inception» 3-го типу описуються як суперпозиція функцій наступним чином:

$$Y_{H^{med1}W^{med1}D^{med1}} = f_{CONV131}(f_{CONV12}(f_{CONV11}(X_{H^{in}W^{in}D^{in}}))); \quad (13)$$

$$Y_{H^{med2}W^{med2}D^{med2}} = f_{CONV132}(f_{CONV12}(f_{CONV11}(X_{H^{in}W^{in}D^{in}}))); \quad (14)$$

$$Y_{H^{med3}W^{med3}D^{med3}} = f_{CONV221}(f_{CONV21}(X_{H^{in}W^{in}D^{in}})); \quad (15)$$

$$Y_{H^{med4}W^{med4}D^{med4}} = f_{CONV222}(f_{CONV21}(X_{H^{in}W^{in}D^{in}})); \quad (16)$$

$$Y_{H^{med5}W^{med5}D^{med5}} = f_{CONV31}(f_{AvgPOOL}(X_{H^{in}W^{in}D^{in}})); \quad (17)$$

$$Y_{H^{med6}W^{med6}D^{med6}} = f_{CONV41}(X_{H^{in}W^{in}D^{in}}); \quad (18)$$

$$\begin{aligned}
 & Y_{H^{out} W^{out} D^{out}} = \\
 & = Y_{H^{med1} W^{med1} D^{med1}} \circ Y_{H^{med2} W^{med2} D^{med2}} \\
 & \circ Y_{H^{med3} W^{med3} D^{med3}} \circ Y_{H^{med4} W^{med4} D^{med4}} \circ \\
 & Y_{H^{med5} W^{med5} D^{med5}} \circ Y_{H^{med6} W^{med6} D^{med6}} ,
 \end{aligned} \tag{19}$$

де \circ – операція конкатенації тензорів карт ознак, для яких виконується наступна умова рівності розмірності

$$\begin{aligned}
 & H^{med1} = H^{med2} = H^{med3} = H^{med4} = H^{med5} = H^{med6} ; \\
 & W^{med1} = W^{med2} = W^{med3} = W^{med4} = W^{med5} = W^{med6} .
 \end{aligned} \tag{20}$$

Остаточна структура архітектури підмережі вилучення ознак Fuzzy Inception-SSD в результаті модифікації архітектури ЗНМ Inception-v2 представлена в табл. 1.

Таблиця 1

Структура архітектури підмережі вилучення ознак Fuzzy Inception-SSD

№	Тип структурної одиниці мережі	Розмір / крок рецептивного поля	Вхідний розмір зображення
1	Свертковий шар (CONV- шар)	3×3/2	299×299×3
2	Шар випрямління (шар ReLU)	–	149×149×32
3	CONV- шар	3×3/1	149×149×32
4	Шар ReLU	–	149×149×32
5	CONV- шар	3×3/2	149×149×32
6	Шар ReLU	–	147×147×64
7	Шар об'єднання (MaxPOOL- шар)	3×3/2	147×147×64
8	CONV- шар	3×3/2	73×73×64
9	Шар ReLU	–	71×71×80
10	CONV- шар	3×3/2	71×71×80
11	Шар ReLU	–	35×35×192
12	CONV- шар	3×3/1	35×35×192
13	Шар ReLU	–	35×35×288
14	Модуль «inception» 1-го типу	Відповідно до структури модуля	35×35×288
15	Модуль «inception» 1-го типу	Відповідно до структури модуля	35×35×288
16	Модуль «inception» 1-го типу	Відповідно до структури модуля	35×35×288
17	Модуль «inception» 2-го типу	Відповідно до структури модуля	17×17×768
18	Модуль «inception» 2-го типу	Відповідно до структури модуля	17×17×768
19	Модуль «inception» 2-го типу	Відповідно до структури модуля	17×17×768
20	Модуль «inception» 2-го типу	Відповідно до структури модуля	17×17×768
21	Модуль «inception» 2-го типу	Відповідно до структури модуля	17×17×768
22	Модуль «inception» 3-го типу	Відповідно до структури модуля	8×8×1280
23	Модуль «inception» 3-го типу	Відповідно до структури модуля	8×8×1280
24	CONV-шар	1×1/1	8×8×1280
25	CONV- шар	1×1/1	5×5×1024
26	CONV- шар	1×1/1	3×3×612
27	Шар об'єднання (AvgPOOL-шар)	3×3	3×3×306

Особливостями використання нечіткої нейронної мережі Ванга-Менделя в якості класифікатора модифікованого детектора SSD в складі Fuzzy Inception-SSD є:

1) подання функцій приналежності в термінах інтервальних нечітких множин типу 2 (ІНМТ2) і реалізація операцій фазифікації, агрегування і активації з використанням операцій на ІНМТ2;

2) введення додаткової операції приведення типу в шарі дефазифікації вихідної змінної на основі класичного методу центру тяжіння (centroid);

3) введення декількох виходів мережі для розпізнавання відповідної кількості класів (підкласів) компактних (точкових) об'єктів. Для цього третій шар представляється як набір з декількох пар нейронів суматорів, а четвертий реалізує кілька нейронів-нормалізаторів, кількість яких відповідає кількості пар третього шару.

При цьому тензори карт ознак, що є результатом функціонування останніх модулів «inception» кожного типу і всіх додаткових шарів, виступають в якості вхідних тензорів карт ознак для модифікованого детектора ЗНМ SSD. Для вихідних тензорів карт ознак останніх модулів «inception» кожного типу виконується додаткова нормалізація з використанням шару локальної нормалізації (Local response normalization (LRN)). При цьому результат виконання нормалізації представляється у відповідності з наступним виразом:

$$y_{i^{out} j^{out} k^{out}} = x_{i^{in} j^{in} k^{in}} \left(k^{in} + \alpha \sum_{t \in G(k^{in})} x_{i^{in} j^{in} t^{in}}^2 \right)^{-\beta} , \tag{21}$$

де для кожного вихідного каналу k^{out} $G(k^{in}) \subset \{1, 2, \dots, D^{in}\}$ є відповідним підмножиною вхідних каналів, $H^{out} = H^{in}$; $W^{out} = W^{in}$; $D^{out} = D^{in}$.

Таким чином, відповідно до розроблених пропозицій пропонується формальне подання нейромережевого ансамблю нечіткої ЗНМ Fuzzy Inception-SSD у вигляді суперпозиції функцій, де кожна функція представляє окрему ЗНМ (модуль) в рамках ансамблю мереж для розпізнавання компактних (точкових) об'єктів:

$$\text{FuzzyCNN} = F_3(F_2(F_1(X_0))), \quad (22)$$

де X_0 – вхідний багатовимірний масив даних, який представляє цифровий аерофотознімок; F_1 – функція формального представлення результатів функціонування модифікованої ЗНМ Inception-v2 зі складу ЗНМ Fuzzy Inception-SSD; F_2 – функція формального представлення результатів функціонування модуля генератора прогнозу розташування з формуванням рамок виділення об'єкта зі складу модифікованого детектора SSD; F_3 – функція формального представлення результатів функціонування модифікованої нечіткої нейронної мережі Ванга-Менделя як класифікатора модифікованого детектора SSD.

Висновки

В роботі запропонована архітектура нечіткої згорткової нейронної мережі у вигляді ансамблю нейромереж на основі модифікованої архітектури нейромережевого детектора Single Shot MultiBox Detector. Модифікація детектора виконана за рахунок, по-перше, модифікації підмережі вилучення ознак з використанням більш глибокої архітектури на основі ЗНМ Inception-v2, по-друге, за рахунок реалізації класифікатора детектора з використанням модифікованої нечіткої нейронної мережі Ванга-Менделя на основі ІНМТ2. Використання розробленої архітектури дозволить підвищити точність розпізнавання компактних (точкових) об'єктів на цифрових аерофотознімках.

Список літератури

1. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg. SSD: Single Shot MultiBox Detector. *arXiv:1512.02325v2*, 2016.
2. Мищенко В.А. Алгоритм обучения нечеткой нейронной сети Ванга-Менделя для распознавания рукопечатных символов в работе почтовой службы / В.А. Мищенко, И.Ф. Астахова, А.А. Краснояров // Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии. – 2012. – № 1. – С. 141-146.
3. Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna. Rethinking the Inception Architecture for Computer Vision. *arXiv:1512.00567v3*, 2015.
4. Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *arXiv:1506.01497v3*, 2016.
5. Олизаренко С. А. Методика формирования платформы для моделирования глубокого обучения в интересах разработки систем автоматизированного дешифрирования аэрофотоснимков / С.А. Олизаренко, О.Ю. Лавров, В.А. Капранов // Системы обработки информации. – 2016. – Вып. 9(146). – С. 41-43.
6. Олизаренко С.А. Интервальные нечеткие множества типа 2. Терминология, представление, операции / С.А. Олизаренко, А.В. Перепелица, В.А. Капранов // Системы обработки информации. – X.: ХУПС, 2011. – Вып. 2(92). – С. 39-45.

Надійшла до редколегії 17.10.2016

Рецензент: д-р техн. наук, проф. О.І. Тимочко, Харківський національний університет Повітряних Сил ім. І. Кожедуба, Харків.

РАЗРАБОТКА АРХИТЕКТУРЫ НЕЧЕТКОЙ СВЕРТОЧНОЙ НЕЙРОНОЙ СЕТИ ДЛЯ РАСПОЗНАВАНИЯ КОМПАКТНЫХ (ТОЧЕЧНЫХ) ОБЪЕКТОВ НА ЦИФРОВОМ АЭРОФОТОСНИМКЕ

С.А. Олизаренко, В.А. Капранов, Р.В. Сафронов

В работе предложена архитектура ансамбля нейронных сетей на базе нейросетевого детектора, сверточной нейронной сети и модифицированной нечеткой нейронной сети для распознавания компактных (точечных) объектов.

Ключевые слова: сверточная нейронная сеть, нечеткая нейронная сеть, детектор, подсеть извлечения признаков, классификатор, карта признаков.

THE ARCHITECTURE DEVELOPMENT OF FUZZY CONVOLUTION NEURAL NETWORK FOR RECOGNITION OF COMPACT (POINT) OBJECTS ON DIGITAL AERIAL PHOTOGRAPHS

S.A. Olizarenko, V.A. Kapranov, R.V. Safronov

The architecture of the connectionist device based on connectionist detector, convolution neural network and modified fuzzy neural network for recognition of compact (point) objects.

Keywords: convolutional neural network, fuzzy neural network, detector, feature extractor, classifier, feature map.