

УДК 621.384

Опубліковано 21 жовтня 2020 року

ІНФОРМАЦІЙНА ТЕХНОЛОГІЯ РОЗРОБКИ БАЗИ ЗНАНЬ ПРО РОЗПІЗНАВАННЯ МІСЬКИХ БУДОВ НА ЦИФРОВИХ КОСМІЧНИХ ТА АЕРОФОТОЗНІМКАХ ПРИ АВТОМАТИЗОВАНОМУ МОНІТОРИНГУ МІСЬКОГО СЕРЕДОВИЩА

ПУСТОВАРОВ Володимир Володимирович 

начальник групи

Харківське представництво генерального замовника –
Державного космічного агентства України

УКРАЇНА

Анотація: У роботі розроблено та запропоновано архітектуру ансамблю нейронних мереж на основі модифікованої згорткової нейронної мережі для сегментації зображень (автокодувальника) U-net. Використання розробленої архітектури дозволить підвищити точність сегментації міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища. Також запропоновано технологію формального представлення побудови інформаційної технології розробки бази знань про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища. Розроблена технологія визначає взаємозалежну сукупність етапів, з можливістю подальшого розпаралелювання виконання підетапів, з урахуванням особливостей подання (формалізації) знань про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища.

ВСТУП.

Відомо, що за веденням постійного моніторингу за змінами міського середовища можливе ефективне управління сучасним містом [1, 2]. Основними вимогами, що пред'являються до такого моніторингу, є оперативність і достовірність до отриманої інформації.

Одним із ефективних підходів щодо підвищення оперативності і достовірності отримання інформації про зміни міського середовища при моніторингу – є розробка автоматизованої системи розпізнавання міських об'єктів (забудов і тощо) на цифрових космічних та

аерофотознімках з визначенням наявності їх змін як інтелектуальної автоматизованої системи.

При цьому, використовується база знань міського середовища, як елемент автоматизованої системи розпізнавання міських об'єктів. Така база – є сукупністю правил, фактів, механізмів виведення, реалізованих на основі використання глибоких нейромережових або гібридних (нечітких нейромережових) моделей. Однак, розробка бази знань – є одним з найбільш трудомістких етапів при створенні інтелектуальних систем. Тому, на даний час, питання комплексного підходу щодо розробки бази знань на промисловому рівні для подання і накопичення знань про процеси розпізнавання міських будов на цифрових космічних та аерофотознімках з використанням глибоких нейронних мереж і гібридних (нечітких нейромережових) моделей – є недостатньо дослідженими.

Таким чином, у предметній області існує протиріччя, яке, з одного боку, полягає у необхідності розробки бази знань для автоматизованих систем розпізнавання міських будов на цифрових космічних та аерофотознімках з використанням глибоких нейронних мереж і нечіткої логіки, а з іншого боку – у обмежених можливостях існуючих технологій розробки подібного класу бази знань при комплексному поданні і накопиченні знань як про процеси розпізнавання визначених об'єктів на цифрових космічних та аерофотознімках, так і про процеси виявлення змін міського середовища і аналіз причин їх виникнення при веденні відповідного моніторингу.

ОСНОВНА ЧАСТИНА.

Відомо багато публікацій, що присвячені питанням розробки баз знань за різними напрямками. В основі базової технології створення бази знань використовуються класичні технології розробки інтелектуальних систем.

При цьому, у якості базового підходу використовується метод прототипування, заснований на виконанні певної послідовності етапів розробки бази знань з проміжним поданням, уточненням і формуванням відповідного прототипу. А для етапів класичних технологій створення інтелектуальних систем характерно досить умовне, не формалізоване визначення меж виконання цих етапів і переходів між ними.

Дані обставини значно ускладнюють забезпечення технологічності

створення прототипу бази знань моніторингу міського середовища з промислової точки зору, не дозволяє мінімізувати необхідний обсяг робіт зі створення бази знань, ускладнюють управління веденням розробки, ускладнюють формулювання завдань і функціональних обов'язків розробників.

Формальне представлення інформаційної технології розробки бази знань про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища.

Формально інформаційну технологію розробки бази знань про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища з використанням глибоких нейромереж та нечіткої логіки пропонується представити у вигляді кортежу

$$T^{KBF} = \langle P^{KBF}, \{S_i^{KBF}\}, \{C_j^{KBF}\}, \{L_m^{KBF}\}, \{M_n^{KBF}\} \rangle \quad (1)$$

де P^{KBF} – постановка завдання щодо розробки бази знань про розпізнавання міських будов на цифрових космічних та аерофотознімках;

$\{S_i^{KBF}\}$ – множина етапів розробки відповідної бази знань;

$i = \overline{1, \dots, I}$, I – загальна кількість етапів розробки;

$\{C_j^{KBF}\}$ – множина зав'язків між етапами з множини $\{S_i^{KBF}\}$;

$j = \overline{1, \dots, J}$, J – загальна кількість зав'язків;

$\{L_k^{KBF}\}$ – множина мов представлення знань про розпізнавання міських будов на цифрових космічних і аерофотознімках;

$k = \overline{1, \dots, K}$, K – загальна кількість мов представлення знань;

$\{M_n^{KBF}\}$ – множина моделей представлення знань;

$n = \overline{1, \dots, N}$, N – загальна кількість моделей.

В якості базових етапів з множини $\{S_i^{KBF}\}$ розглядаються модифіковані етапи розробки класичних інтелектуальних систем:

– етап ідентифікації знань $S_1^{KBF} \in \{S_i^{KBF}\}$ щодо розпізнавання міських будов на цифрових космічних і аерофотознімках;

– етап концептуалізації знань $S_2^{KBF} \in \{S_i^{KBF}\}$;

– етап формалізації знань $S_3^{KBF} \in \{S_i^{KBF}\}$;

– етапи програмної реалізації (прототипування) $S_4^{KBF} \in \{S_i^{KBF}\}$;

– верифікації бази знань $S_5^{KBF} \in \{S_i^{KBF}\}$.

На етапі ідентифікації знань – S_1^{KBF} виконується загальна постановка задачі щодо розробки бази знань – P^{KBF} , визначаються необхідні ресурси і джерела знань.

На етапі концептуалізації – S_2^{KBF} формується концептуальна модель (поле знань) бази знань – $M_1^{KBF} \in \{M_n^{KBF}\}$ при виконанні наступних підетапів:

– за результатами аналізу предметної області – визначається склад основних знань у вигляді вербального опису (склад поля знань) про розпізнавання міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища – $S_2^{K(KBF)}$;

– за результатами аналізу знань – формується безпосередньо поле знань (структура предметної області) M_1^{KBF} з використанням відповідної мови описів поля знань (у даній технології використовуються конструкції подання концептуальних моделей на основі мови об'єктно-орієнтованого моделювання UML $L_1^{KBF} \in \{L_k^{KBF}\}$) – $S_2^{M_1(KBF)}$.

На етапі формалізації (подання) знань – S_3^{KBF} формується математична – $M_2^{KBF} \in \{M_n^{KBF}\}$ та логічна – $M_3^{KBF} \in \{M_n^{KBF}\}$ моделі бази знань при виконанні наступних підетапів:

– визначається формалізований склад знань про розпізнавання міських будов на цифрових космічних і аерофотознімках, а також – наявність змін у міському середовищі – $S_3^{KS(KBF)}$;

– визначаються способи представлення знань та обираються (розробляються) методи формалізації знань – $S_3^{KR(KBF)}$. Для представлення знань про розпізнавання міських будов на цифрових космічних і аерофотознімках, а також наявність змін у міському середовищі пропонується використання методів, що засновані на: інтервальних нечітких множинах другого типу – $L_2^{KBF} \in \{L_k^{KBF}\}$, інтервальних нечітких логічних системах другого типу – $L_3^{KBF} \in \{L_k^{KBF}\}$ і модифікованих глибоких нейронних мережах (згорткових нейронних мережах) на основі попередньо навчених моделей – $L_4^{KBF} \in \{L_k^{KBF}\}$ та моделей для семантичної сегментації – $L_5^{KBF} \in \{L_k^{KBF}\}$;

– виконується формалізований опис знань (формується математична модель – M_2^{KBF}) з використанням обраних способів подання і розроблених методів формалізації знань – $S_3^{KD(KBF)}$;

– розробляється узагальнений алгоритм рішення знань (у рамках

формування математичної моделі – M_2^{KBF}), враховуються формалізовані знання і забезпечується виявлення тенденцій перебігу процесу побудови бази знань і представлення результатів вирішення підзадач – $S_3^{AM(KBF)}$ для кожної формалізованої підзадачі (сегментація міських будов на цифрових космічних та аерофотознімках, виявлення наявності змін у міському середовищі);

– розглядаються як процедурні знання елементи математичної моделі – M_2^{KBF} , що перетворюються у логічну модель бази знань – M_3^{KBF} , яка описується з використанням конструкцій представлення логічних моделей мови об'єктно-орієнтованого моделювання UML – L_1^{KBF} – $S_3^{LM(KBF)}$.

На етапі програмної реалізації (попереднього прототипування) бази знань – S_4^{KBF} формується фізична модель бази знань – $M_4^{KBF} \in \{M_n^{KBF}\}$ і безпосередньо програмний прототип бази знань при виконанні наступних підетапів:

– розробляється архітектура бази знань, яка визначає структуру, функції і взаємозв'язок компонентів бази знань – $S_4^{BA(KBF)}$. При цьому, у якості основних елементів бази знань використовуються функції приналежності нечітких множин, навчальна вибірка (неявна база знань) і синаптичні карти нейромереж;

– виконується реалізація (прототипування) компонентів бази знань – $S_4^{BP(KBF)}$. При цьому, реалізація компонентів бази знань включає розробку проектних рішень (формування моделі M_4^{KBF}) щодо програмної реалізації бази знань з використанням фізичних моделей мови об'єктно-орієнтованого моделювання UML – L_1^{KBF} та безпосередню програмну реалізацію бази знань з використанням високорівневої мови програмування – $L_6^{KBF} \in \{L_k^{KBF}\}$ та (або) системи управління базою даних (знань) – $L_7^{KBF} \in \{L_k^{KBF}\}$;

– здійснюється наповнення бази знань на основі елементів математичної моделі – M_2^{KBF} , розглянутих як декларативні знання – $S_4^{BK(KBF)}$. Наповнення бази знань здійснюється як за результатами безпосереднього введення декларативних знань у базу, (наприклад, при введенні нової або розширення існуючої навчальної вибірки), так і за результатами навчання елементів бази знань (наприклад, при навчанні нейромереж формуються відповідні синаптичні карти). При цьому,

знання представлені у вигляді синаптичних карт пропонується визначати за результатами побудови і подальшого навчання нейромереж з використанням спеціалізованих засобів моделювання.

Тестування (верифікація) бази знань – $S_5^{KBF} \in \{S_i^{KBF}\}$ безпосередньо виконується або над проектними рішеннями, представленими у вигляді моделей UML, або над програмним кодом, які реалізують безпосередньо прототип або остаточну версію бази знань про розпізнавання. У разі виявлення похибок повернення може здійснюватися на будь-який попередній етап розробки бази знань залежно від типу виявленої похибки. При цьому, множина зв'язків $\{C_j^{KBF}\}$ між етапами з множини $\{S_i^{KBF}\}$ визначається порядком виконання цих етапів (прямі зв'язки $C_1^{KBF} - C_4^{KBF}$) та результатами тестування (зворотні зв'язки $C_5^{KBF} - C_8^{KBF}$).

Таким чином, множину етапів інформаційної технології T^{KBF} можливо задавати наступним чином:

$$\{S_i^{KBF}\} = \{S_1^{KBF}, S_2^{KBF}, S_3^{KBF}, S_4^{KBF}, S_5^{KBF}\}, \quad (2)$$

$$S_2^{KBF} = S_2^{K(KBF)} \cup S_2^{M(KBF)}, \quad (3)$$

$$S_3^{KBF} = S_3^{KS(KBF)} \cup S_3^{KR(KBF)} \cup S_3^{KD(KBF)} \cup S_3^{AM(KBF)} \cup S_3^{LM(KBF)}, \quad (4)$$

$$S_4^{KBF} = S_4^{BA(KBF)} \cup S_4^{BP(KBF)} \cup S_4^{BK(KBF)}. \quad (5)$$

У подальшому буде розглянуто детально найбільш складні 2 і 3 етапи інформаційної технології розробки бази знань про розпізнавання міських будов на цифрових космічних та аерофотознімках, які пов'язані з формалізацією відповідних процесів.

Архітектура нечіткої згорточної нейронної мережі для сегментації міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища.

Першим етапом безпосередньо процесу аналізу космічних і аерофотознімків при автоматизованому моніторингу міського середовища можна розглядати етап автоматичного розпізнавання різноманітних об'єктів на місцевості.

У даному дослідженні визначені два класи для розпізнавання: міська будова та не будова. У якості базового методу у рамках інформаційної

технології розробки бази знань про розпізнавання міських будов на цифрових космічних і аерофотознімках використовується семантична сегментація.

Семантична (або смислова) сегментація зображень полягає у виділенні на зображенні областей, кожна з яких відповідає певній ознаці. У загальному вигляді завдання семантичної сегментації важко алгоритмізуються. Тому, для сегментації зображень на даний час широко використовуються глибокі нейронні мережі, які показують достатньо високу точність розпізнавання об'єктів сегментації.

Найбільш ефективними для вирішення завдань семантичної сегментації, у межах теорії глибокого навчання, – є різні реалізації автокодувальника (autoencoder). Для реалізації семантичної сегментації у межах глибокого навчання, був проведений аналіз основних архітектур глибоких нейронних мереж, кожна з яких у загальному випадку розглядається як мережа кодувальника, за якою слідує мережа декодера: FCN, SegNet, U-Net, FC-Densenet, E-Net & Link-Net, RefineNet, PSPNet, Mask-RCNN, G-FRNet, DecoupledNet, GAN Based Approaches.

Для сегментації міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища у якості критеріїв вибору архітектури глибоких нейронних мереж використовувалися:

- прикладна спрямованість архітектури;
- забезпечення передбачуваної точності розпізнавання;
- мінімальна навчальна вибірка для забезпечення передбачуваної точності розпізнавання;
- мінімальна кількість епох для навчання;
- швидкість розпізнавання у режимі реального часу.

За результатами проведеного аналізу основних архітектур глибоких нейронних мереж встановлено, що найбільш повно відповідає визначеним критеріям архітектура згорткової нейронної мережі U-Net [3].

Архітектура мережі U-Net – є повнозв'язана згорткова мережа, модифікована таким чином, щоб вона могла працювати з меншою кількістю прикладів (навчальних образів) і робила більш точну сегментацію. Мережа містить енкодер і декодер. При цьому, архітектура U-Net відрізняється від традиційного автоенкодера наявністю конкатенації у частині декодера з картами ознак із частини кодера.

На кожному кроці архітектури подвоюється кількість каналів ознак. Кодер – є типова згортчна мережа і містить два послідовних шари згортки розміром 3×3 , після яких йде шар ReLU і пулінг з функцією максимуму 2×2 з кроком 2. Кожен крок декодери містить шар, зворотний пулінгу, що розширює карту ознак, після якого використовується згортка 2×2 , що зменшує кількість каналів ознак. Конкатенація виконується з відповідним чином обрізаною картою ознак кодера і двох згорток 3×3 , після кожної з якої йде шар ReLU. Обрізка потрібна через те, що при конкатенації втрачаються граничні пікселі у кожній згортці. На останньому шарі використовується повнозв'язна згортка розміром 1×1 для приведення кожного 64-компонентного вектора ознак до необхідної кількості класів.

Особливістю використання модифікованої загорткової нейронної мережі U-Net як підмережі вилучення ознак у даному дослідженні є:

- використання попередньо навченої загорткової нейронної мережі VGG16 у якості енкодера,
- відсутність додаткових класифікаторів, які є в оригінальній версії загорткової нейронної мережі VGG16 (табл. 1) та U-Net;
- у введенні додаткових шарів передискретизації розмірності карт ознак для реалізації конкатенації відповідних шарів енкодера (шари UpSampling).

При цьому, класичний декодер збільшується у глибину за рахунок трьох послідовних шарів згортки розміром 3×3 на кожному кроці (блоці декодера).

Таблиця 1

Структура архітектури підмережі вилучення ознак VGG16 у якості енкодера (без класифікатора)

Layer (type)	Output Shape	Param
input_1 (InputLayer)	(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
...		
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0

Принципами навчання модифікованої підмережі вилучення ознак U-Net є наступні:

- використовується технологія «bottleneck features», згідно з якою, на виході попередньо навченої моделі знімаються ознаки (у даному випадку у енкодера у форматі VGG16) та подаються на вхід навченою частини мережі (у даному випадку – це декодер у форматі U-Net) та, у подальшому, навчається тільки друга частина (у даному випадку тільки декодер). Особливістю моделі U-Net, як вже зазначалось, – є наявність проміжних зв'язків між енкодером і декодером, тобто для декодера має місце кілька входів (у даному випадку 5), що значно ускладнює використання технології «bottleneck features» для семантичної сегментації (на відміну від нейронного класифікатора у якого тільки один вхід). Пропонується підхід щодо вирішення даної наукової задачі за рахунок використання відповідних шарів VGG16 у рамках енкодера (табл. 2) та шарів передискретизації UpSampling у рамках структури декодера;

- використовується технологія тонкого налаштування моделі. На даному етапі послідовно розморожуються 1 та 2 шари енкодера, поєднуються ваги енкодера (попередньо навченої моделі VGG16 на бібліотеці зображень ImageNet) і отримані на першому етапі ваги декодера і, виконується тонка настройка розморожених шарів енкодера і шарів декодера.

У якості основних елементів архітектури підмережі вилучення ознак модифікованої загорткової нейронної мережі U-Net 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 у 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 у результаті шарів енкодера і декодера;
- допоміжні шари у структурі декодера – BatchNormalization та UpSamling2D.

Таблиця 2

Шари підмережі вилучення ознак VGG16, що використовуються для конкатенації з декодером

№	Layer output
1	vgg16_net_layers5_output = Input(vgg16_net.layers[17].output_shape[1:])
2	vgg16_net_layers4_output = Input(vgg16_net.layers[13].output_shape[1:])
3	vgg16_net_layers3_output = Input(vgg16_net.layers[9].output_shape[1:])
4	vgg16_net_layers2_output = Input(vgg16_net.layers[5].output_shape[1:])
5	vgg16_net_layers1_output = Input(vgg16_net.layers[2].output_shape[1:])

Можливо формально представити основні шари модифікованої мережі. Формально згортковий шар можна представити у вигляді наступної функції:

$$Y_{H^{out}W^{out}D^{out}} = f_{CONV}(X_{H^{in}W^{in}D^{in}}), \quad (6)$$

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 можливо формально представити у вигляді наступних функцій:

- для операції "max":

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

- для операції "avg":

$$Y_{H^{out}W^{out}D^{out}} = f_{AvgPOOL}(X_{H^{in}W^{in}D^{in}}). \quad (8)$$

Шар активації реалізує функцію активації ReLU:

$$y_{i^{out};j^{out};k^{out}} = \mathcal{S}(x_{i^{in};j^{in};k^{in}}) = \max_{ijk}(0, x_{i^{in};j^{in};k^{in}}), \quad (9)$$

де

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

$$x_{i^{in};j^{in};k^{in}} \in X_{H^{in}W^{in}D^{in}},$$

$$H^{out} = H^{in}, W^{out} = W^{in},$$

$$D^{out} = D^{in}.$$

Операцію конкатенації можливо математично описати як суперпозицію функцій, що описують CONV-шари енкодера та шари UpSampling2D декодера.

У результаті модифікації архітектури ЗНМ U-Net отримано структуру архітектури підмережі вилучення ознак, що представлена у таблиці 3.

До особливостей використання нечіткої нейронної мережі Ванга-Менделя у якості класифікатора модифікованого декодера U-Net відноситься наступне:

- подання функцій приналежності у термінах інтервальних нечітких множин типу 2 (ІНМТ2) і реалізація операцій фазифікації, агрегування і активації з використанням операцій на ІНМТ2;
- введення додаткової операції приведення типу у шарі дефазифікації вихідної змінної на основі класичного методу центру тяжіння (centroid);
- введення декількох виходів мережі для розпізнавання відповідної кількості класів (підкласів) предметної області. 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.

Таблиця 3

**Структура архітектури підмережі вилучення ознак
U-Net у якості декодера (без класифікатора)**

Layer (type)	Output Shape	Param	Connected to
input_2 (InputLayer)	(None, 14, 14, 512)	0	
input_3 (InputLayer)	(None, 28, 28, 512)	0	
up_sampling2d_4 (UpSampling2D)	(None, 28, 28, 512)	0	input_2[0][0]
concatenate_4 (Concatenate)	(None, 28, 28, 1024)	0	input_3[0][0] up_sampling2d_4[0][0]
conv2d_12 (Conv2D)	(None, 28, 28, 512)	4719104	concatenate_4[0][0]
batch_normalization_12	(BatchNo (None, 28, 28, 512)	2048	conv2d_12[0][0]
elu_12 (ELU)	(None, 28, 28, 512)	0	batch_normalization_12[0][0]
conv2d_13 (Conv2D)	(None, 28, 28, 512)	2359808	elu_12[0][0]

Таблиця 3 (продовження)

batch_normalization_13	(None, 28, 28, 512)	2048	conv2d_13[0][0]
elu_13 (ELU)	(None, 28, 28, 512)	0	batch_normalization_13[0][0]
conv2d_14 (Conv2D)	(None, 28, 28, 512)	2359808	elu_13[0][0]
batch_normalization_14	(None, 28, 28, 512)	2048	conv2d_14[0][0]
elu_14 (ELU)	(None, 28, 28, 512)	0	batch_normalization_14[0][0]
....			
elu_23 (ELU)	(None, 224, 224, 64)	0	batch_normalization_23[0][0]
spatial_dropout2d	(None, 224, 224, 64)	0	elu_23[0][0]

При цьому, тензори карт ознак, що є результатом функціонування шарів декодера, виступають у якості вхідних тензорів карт ознак для модифікованої нечіткої нейронної мережі Ванга-Менделя [4].

Таким чином, пропонується формальне подання нейромережевого ансамблю нейронних мереж на основі модифікованої згорткової нейронної мережі для сегментації зображень (автокодувальника) U-Net у вигляді суперпозиції функцій, де кожна функція представляє окрему згорткову нейронну мережу (модуль) у межах ансамблю мереж для сегментації міських будов на цифрових космічних та аерофотознімках:

$$\text{FuzzyVU} = F_3 \left(F_2 \left(F_1 (X_0) \right) \right), \quad (10)$$

де X_0 – вхідний багатовимірний масив даних, який представляє цифровий аерофотознімок;

F_1 – функція формального представлення результатів функціонування модифікованого екодера U-Net на базі ЗНМ VGG16;

F_2 – функція формального представлення результатів функціонування модифікованого декодера U-Net;

F_3 – функція формального представлення результатів функціонування модифікованої нечіткої нейронної мережі Ванга-Менделя, як класифікатора модифікованого декодера U-Net.

Далі буде більш детально розглянуто пропозиції щодо модифікації нечіткої нейронної мережі Ванга-Менделя.

Пропозиції щодо модифікації нечіткої нейронної продукційної мережі Ванга-Менделя для сегментації зображень.

Нечітка нейронна продукційна мережа Ванга-Менделя реалізує нечітку продукційну модель за правилами, умови і заключення яких формуються на основі нечіткої множини типу 1. При цьому, функції приналежності всіх нечітких множин представляються функцією Гауса. Тому, вихідний сигнал мережі Ванга-Менделя можливо розрахувати відповідно до наступних виразів:

$$y_{\text{class}}(x) = \frac{\sum_{i=1}^M c_i \prod_{j=1}^N \mu_{ij}(x_j)}{\sum_{i=1}^M \prod_{j=1}^N \mu_{ij}(x_j)}, \quad (11)$$

$$\mu_{ij}(x_j) = \exp \left[- \left(\frac{x_j - a_{ij}}{b_{ij}} \right)^2 \right], \quad (12)$$

де x_j – вхідний сигнал;

c_i – центр ширини функції Гауса, що представляє функцію приналежності нечіткої множини заключень правил;

μ_{ij} – функція Гауса з параметрами математичного очікування, яке визначає центр a_{ij} , і параметрами розкиду, які визначаються середнім квадратичним відхиленням, що представляє функцію приналежності нечіткої множини умов правил;

N – кількість вхідних змінних мережі;

M – кількість нейронів першого шару, що реалізує фаззіфікацію вхідних змінних.

Структура мережі Ванга-Менделя – є чотиришаровою нейронною мережею, в якій:

- у першому шарі виконується фаззіфікація вхідних змінних;
- у другому шарі здійснюється агрегація значень активації умови нечіткого продукційного правила;
- у третьому шарі виконується активізація заключень правил виведення;
- у четвертому шарі, що складається з одного нейрона, здійснюється дефаззіфікація вихідний змінної і формування вихідного сигналу.

Відповідно до структури мережі та виразів (11), (12) – перший і третій шари – є параметричними.

Отже, алгоритм нечіткого логічного виведення, який реалізований даною нечіткою нейронною продукційною мережею, базується на наступних основних положеннях:

- вхідні змінні є чіткими;
- функції приналежності всіх нечітких множин представляються функцією Гауса;
- акумуляція активізованих правил не здійснюється.

Таким чином, модифікація класичної мережі Ванга-Менделя для вирішення задачі для попиксельної класифікації визначених об'єктів здійснюється за наступною допомогою:

- представлення функцій приналежності у термінах ІНМТ2 та реалізації операцій фаззіфікації, агрегації і активації з використанням операцій над ІНМТ2 [5];
- введення додаткової операції приведення типу у шарі дефаззіфікації вихідної змінної на основі класичного методу центру тяжіння (centroid);
- введення декількох виходів мережі для розпізнавання відповідної кількості класів (підкласів) визначених об'єктів. При цьому, третій шар представляється набором з декількох пар нейронів-суматорів, а четвертий реалізує декілька нейронів-нормалізаторів, кількість яких відповідає кількості пар третього шару.

У загальному випадку, нечітке продукційне правило для модифікованої нечіткої нейронної продукційної мережі Ванга-Менделя можливо представити у наступному вигляді:

$$R: \text{IF } \beta_1 \text{ is } \alpha_1 \text{ AND } \dots \text{ AND } \beta_n \text{ is } \alpha_n \text{ THEN } \beta_{n+1} = c_j^p, \quad (13)$$

де β_1 – найменування вхідних даних (у вигляді чіткого (в окремому випадку) або лінгвістичного значення (ЛЗ)), що задається кортежем – $\langle \beta_i, T_i, X_i, M_i \rangle$, $i = 1, \dots, n$; n – кількість вхідних змінних;

$T_i = \{ \alpha_i \}$ – множина значень (термів) вхідної ЛЗ правила R , кожне з яких є найменуванням нечіткого значення (НЗ) для опису значень параметрів розпізнавання компактних (точкових) об'єктів повітряної розвідки;

X_i – область значень НЗ, найменування яких входять у T_i , M_i – семантична

процедура, що ставить у відповідність значенню ЛЗ нечітку множину. Синтаксична процедура генерування нових значень для ЛЗ G_i не використовується, оскільки всі значення ЛЗ у рамках пропонованого підходу визначаються на етапі формування бази правил;

α_i – значення терма вхідної ЛЗ у вигляді найменування НЗ (лінгвістичного значення параметра розпізнавання об'єкту), що задається кортежем $\langle \alpha_i, X_i, \tilde{A} \rangle$, $i = 1, \dots, n$ та $\tilde{A} \subseteq X_i$;

$\tilde{A} = \{x, \mu_{\tilde{A}}(x)\}$ – ІНМТ2 на множині X_i , що описує можливі значення, які може приймати НЗ $\alpha_i \in T_i$;

β_{n+1} – найменування вихідної ЛЗ, що задається кортежем у першому $\langle \beta_{n+1}, T_{n+1}, Y, M_{n+1} \rangle$;

Y – область значень термів, найменування яких входять у T_{n+1} , що представляють номер розпізнаваного класу (підкласу) об'єкту;

M_{n+1} – семантична процедура, що ставить у відповідність значенню ЛЗ одноточкову нечітку множину;

c_j^p – значення терма вихідний ЛЗ у вигляді найменування або номера класу з множини $C_p = \{c_j^p\}$, $j = 1, \dots, t$;

t – кількість класів (підкласів) компактних (точкових) об'єктів.

Під ІНМТ2 формально розуміється нечітка множина типу 2, або всі вторинні ступені $f_x(u) \equiv \mu_{\tilde{A}}(x, u)$, які дорівнюють 1 та, яке представляється у наступному вигляді:

– у разі безперервних (нескінченних) X та J_x , як:

$$\tilde{A} = \int_x \int_u \frac{1}{(x,u)}, \quad u \in J_x \subseteq U = [0,1], \quad x \in X, \quad (14)$$

або:

$$\tilde{A} = \int_x \left[\int_u \frac{1}{u} \right], \quad u \in J_x \subseteq U = [0,1], \quad x \in X; \quad (15)$$

– у разі дискретних (кінцевих) X та J_x , як:

$$\tilde{A} = \sum_{j=1}^N \frac{\left[\sum_{i=1}^{M_j} \frac{1}{u_i} \right]}{x_j}, \quad u \in J_x \subseteq U = [0,1], \quad x \in X \quad (16)$$

або, як множина:

$$\tilde{A} = \{((x, u), 1) | \forall x \in X, \forall u \in J_x \subseteq U = [0, 1]\} \quad (17)$$

де x – первинна змінна $x \in X$;

X – універсальна множина об'єктів наочної області;

u – вторинна змінна $u \in J_x$;

J_x – первинна приналежність $J_x \subseteq U$;

U – область визначення первинної приналежності $U = [0, 1]$;

$f_x(u) \equiv \mu_{\tilde{A}}(x, u)$ – вторинний ступінь, який дорівнює до 1.

Існує два основні способи формального представлення ІНМТ2 – вертикального та хвилястого зрізу. Для представлення функцій приналежності модифікованої нечіткої нейронної продукційної мережі Ванга-Менделя у термінах ІНМТ2 можливо розглянути базове поняття займаної площі невизначеності. Об'єднання усіх первинних приналежностей ІНМТ2 – є займаною площею невизначеності (footprint of uncertainty (FOU)) та може бути представлено у наступному вигляді:

$$FOU(\tilde{A}) = \cup_{\forall x \in X} J_x = \{(x, u) ; u \in J_x \subseteq [0, 1]\}. \quad (18)$$

Визначення FOU для ІНМТ2 і загальних НМТ2 співпадає. Однак, для ІНМТ2 займана площа невизначеності має особливе значення. Це пов'язано з тим, що, оскільки вторинні ступені ІНМТ2 не передають нової інформації про додаткові ступені невизначеності, то можна вважати, що FOU – є закінченим описом ІНМТ2. Для ІНМТ2 первинну приналежність можливо представити наступним чином:

$$J_x = \left[\bar{\mu}_{\tilde{A}}(x), \underline{\mu}_{\tilde{A}}(x) \right], \quad (19)$$

де $\bar{\mu}_{\tilde{A}}(x)$ – значення верхньої функції приналежності (upper membership function), що позначається $UMF(\tilde{A})$ або $\bar{\mu}_{\tilde{A}}$;

$\underline{\mu}_{\tilde{A}}(x)$ – значення нижньої функції приналежності (lower membership function), що позначається $LMF(\tilde{A})$ або $\underline{\mu}_{\tilde{A}}$.

Вкладеною інтервальною НМТ2 \tilde{A}_e для безперервних (нескінченних) множин X та J_x – є множина, у якій кожна первинна змінна $x \in X$ має тільки одну вторинну змінну $u \in J_x$ (тобто одне значення первинної

приналежності) з відповідним вторинним ступенем, який дорівнює до 1, тобто:

$$\tilde{A}_e = \int_{x \in X} \frac{[1]}{x}, \quad u \in J_x \subseteq U = [0,1]. \quad (20)$$

Вкладеною НМТ1 A_e з потужністю N для дискретних (кінцевих) множин X та J_x – є множина пар, де перші елементи пари представляють відповідні первинні змінні, а другі елементи точно одне значення з $J_{x_1}, J_{x_2}, \dots, J_{x_N}$, а саме u_1, u_2, \dots, u_N , тобто:

$$A_e = \sum_{i=1}^N \frac{u_i}{x_i}, \quad u_i \in J_{x_i} \subseteq U = [0,1], \quad x_i \in X. \quad (21)$$

Порівнюючи вирази (20) і (21), можливо представити вкладену ІНМТ2 через елементи вкладеної НМТ1 наступним чином:

$$\tilde{A}_e = \frac{1}{A_e}. \quad (22)$$

Відома представницька теорема (Representation Theorem), згідно з якою ІНМТ2 \tilde{A} може бути представлена як об'єднання вкладених у нього ІНМТ2, тобто:

$$\tilde{A} = \cup_{j=1}^n \tilde{A}_e^j, \quad (23)$$

де n – кількість вкладених ІНМТ2 (\tilde{A}_e^j представляється згідно з виразом (1.3.10))

$$n = \prod_{i=1}^N M_i;$$

N – кількість значень, на яку дискретизовано X ;

M_i – кількість значень, на яку дискретизовано J_x .

Представлення ІНМТ2 у вигляді виразу (23) – є представленням ІНМТ2 способом хвилястого зрізу.

За аналогією з виразом (23), займану площу невизначеності FOU можливо представити наступним чином:

$$FOU(\tilde{A}) = \cup_{j=1}^n A_e^j. \quad (24)$$

За виразом (24) та урахуванням виразу (23) можливо ІНМТ2 представити у наступному вигляді:

$$\tilde{A} = \cup_{j=1}^n \frac{1}{A_e^j} = \frac{1}{\cup_{j=1}^n A_e^j} \equiv \frac{1}{FOU(\tilde{A})} = \frac{1}{\cup_{\forall x \in X} [\mu_{\tilde{A}}(x), \mu_{\tilde{A}}(x)]}. \quad (25)$$

На рисунку 1 представлено підсумкову структуру модифікованої нечіткої нейронної мережі Ванга-Менделя на основі використання ІНМТ2 та з k виходами за кількістю розпізнаваних класів (підкласів) об'єктів.

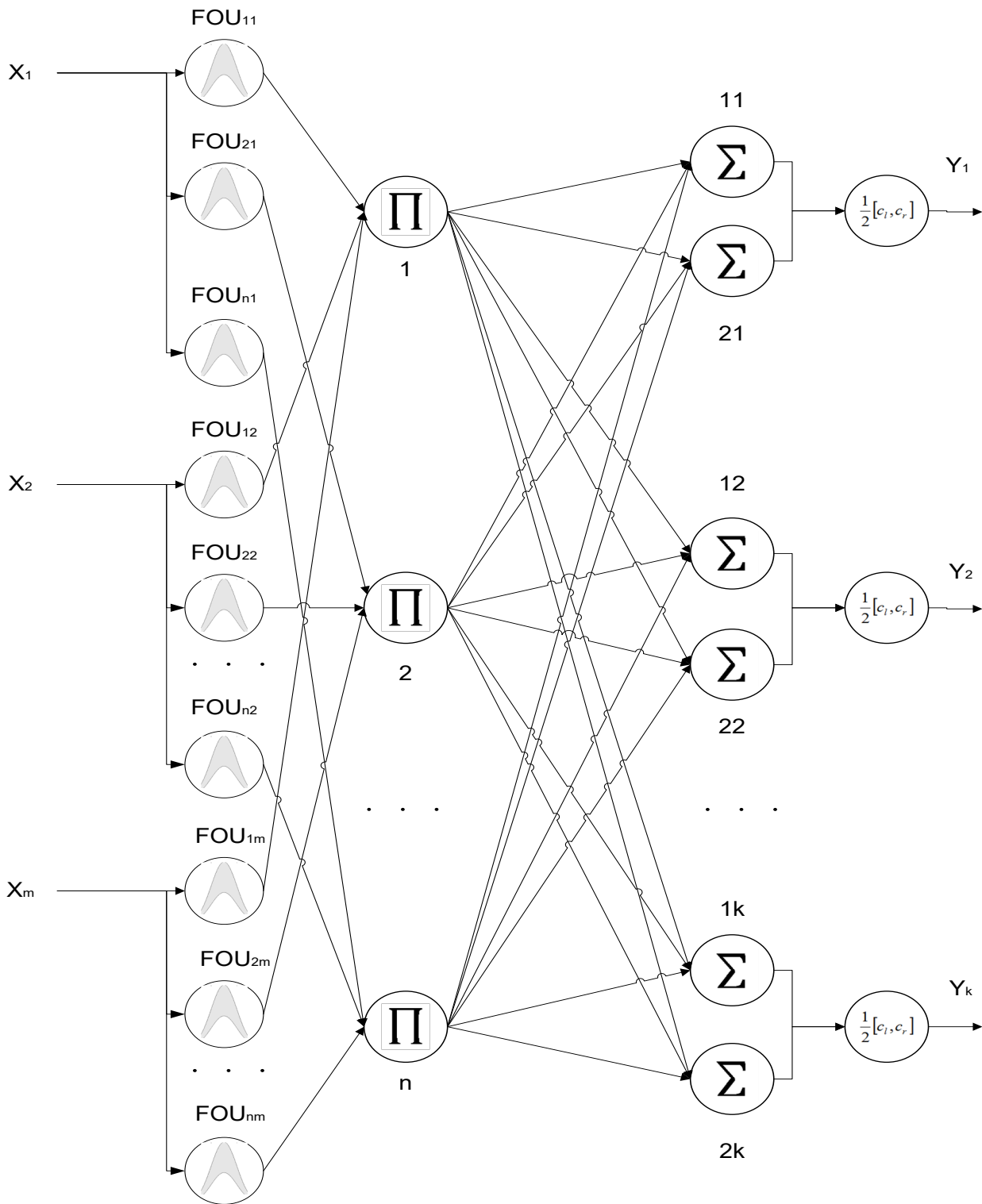


Рис. 1. Структура модифікованої нечіткої нейронної продукційної мережі Ванга-Менделя на основі ІНМТ2

При цьому, вихідний сигнал модифікованої мережі Ванга-Менделя можливо розрахувати за наступною низкою виразів:

$$\bar{y}_{\text{class}}(x) = \frac{\sum_{i=1}^M \bar{c}_i \prod_{j=1}^N \bar{\mu}_{ij}(x_j)}{\sum_{i=1}^M \prod_{j=1}^N \bar{\mu}_{ij}(x_j)}, \quad (26)$$

$$Y = J_x^{\text{class}} = \left| \bar{y}_{\text{class}}(x), y_{\text{class}}(x) \right|, \quad (27)$$

де x_j – вхідний сигнал;

\bar{c}_i, c_i - центр ширини верхньої та нижньої функції Гауса, що представляє функцію приналежності нечіткої множини заключень правил;

$\bar{\mu}_{ij}, \mu_{ij}$ - значення верхньої та нижньої функцій приналежності передумов правил;

J_x^{class} - первинна приналежність класу (підкласу), що розпізнається.

Таким чином, згідно представницької теореми (за виразом (25)), ІНМТ2 - є об'єднанням усіх вкладених НМТ1, які покривають його займану площу невизначеності. Важливість даного висновку полягає у наступному:

- підтверджується теза про те, що ІНМТ2 повністю описується (визначається) його FOU;

- з'являється можливість використовувати як операції над ІНМТ2 відповідні операції над НМТ1, що значно спрощує обчислювальну складність ІНМТ2 для їх реалізації у рамках, відповідних нечітких логічних систем. Тобто, використовуються функції приналежностей, які повністю описується відповідною FOU, у якої, у свою чергу, верхня і нижня межі можливо представити за функцією Гауса відповідно до виразу (12).

Пропозиції щодо виявлення змін на різночасних космічних і аерофотознімках за результатами автоматичного розпізнавання різноманітних об'єктів на місцевості.

Другим етапом аналізу космічних і аерофотознімків при автоматизованому моніторингу міського середовища можна розглядати етап виявлення змін на різночасних знімках за результатами автоматичного розпізнавання різноманітних об'єктів на місцевості.

Виявити зміни за різночасними знімками можливо завдяки використанню різних підходів. Один з них полягає у роздільній обробці вихідних або перетворених зображень для їх поелементного порівняння. Так, є метод диференціювання різночасових знімків (Image differencing),

який заснований на поелементному відніманні яскравості вихідних зображень. На ділянках без змін – різниця значень яскравості буде близька до нуля, а на тих, що змінилися – мати позитивні або негативні значення (у залежності від спрямованості змін).

Варто зазначити що, так як знімки представлені у вигляді окремих спектральних каналів, – є можливість вибору найбільш інформативного каналу.

Отже, можливо виявити зміни, що відбулися або у одному, що представляє інтерес каналі, або у кожному каналі, а потім провести злиття даних одноканальних зображень.

Існує метод зіставлення знімків після класифікації (Post-classification comparison), який полягає у поелементному порівнянні різночасових класифікованих зображень. В ньому використовуються алгоритми класифікації як з навчанням, так і без нього. При цьому, з використанням традиційних підходів виникають серйозні труднощі у зіставленні результатів класифікації, що пов'язані з розбіжністю як кількості класів, так і їх співвіднесення з об'єктами. Ці складності можна подолати шляхом перекодування отриманих растрів, тим не менш, дана процедура досить трудомістка, а результат залежить від безлічі факторів. Так само і для поліпшення візуального сприйняття класифікації можливо скористатися функцією генералізації зображень, яка дозволяє максимально усунути острівні полігони і дрібні за площею зміни.

Методи, які засновані на використанні індексних зображень, також полягають у поелементному порівнянні, відповідно, після математичного перетворення (розрахунку індексу).

Виявлення змін за різночасними зображеннями, які перетворені за методом головних компонент, базується на поелементному відніманні значень яскравості відповідних пікселів головних компонент. Однак, існує й інший підхід, який полягає у спільній обробці сукупності знімків за різні дати, що геометрично поєднуються і обробляються разом, подібно багатозональному знімку. До нього також відноситься і метод створення композитних та синтезованих зображень. При цьому, потрібно взяти до уваги те, що одночасний аналіз трьох і більше знімків – є задача досить складна і трудомістка. Тому, одночасно обробляються два знімки, хоча це і не обов'язкова умова.

Таким чином, з урахуванням результатів сегментації, що отримані на попередніх етапах використання запропонованої інформаційної

технології, для порівняння сігментованих різночасових зображень доцільно використовувати метод зіставлення знімків після класифікації (сігментації).

При цьому, для відображення змін дані використовуються такі методи:

– віднімання:

$$R = A - B_A, \quad (28)$$

де R – результат виконання;

A – вихідне зображення (щодо якого проводиться пошук змін);

B_A – зображення зі змінами;

– ділення:

$$R = A / B_A. \quad (29)$$

ВИСНОВКИ З ТЕМИ ДОСЛІДЖЕННЯ.

Таким чином, за результатами розглянутих пропозицій щодо побудови інформаційної технології розробки бази знань про розпізнавання міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища можна зробити наступні висновки:

1. Запропонована інформаційна технологія дозволяє реалізувати промислову розробку бази знань про розпізнавання міських будов на цифрових космічних і аерофотознімках з використанням нейромереж та нечіткої логіки, з чітко визначеною і взаємозалежною сукупністю етапів, з можливістю подальшого розпаралелювання виконання підетапів, з урахуванням особливостей подання (формалізації) знань про розпізнавання міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища. Перспективами подальших досліджень у цьому напрямку може бути розробка пропозицій щодо використання засобів автоматизованого проектування баз знань у автоматизованих системах моніторингу міського середовища.

2. Запропонована архітектура ансамблю нейронних мереж на основі модифікованої згорткової нейронної мережі для сегментації зображень (автокодувальника) U-Net. Модифікація U-Net виконана за рахунок, по-перше, модифікації підмережі вилучення ознак з використанням у якості енкодера попередньо навченої глибокої ЗНМ VGG, по-друге, за рахунок

реалізації класифікатора автокодувальника з використанням модифікованої нечіткої нейронної мережі Ванга-Менделя на основі ІНМТ2 для попіксельної класифікації визначених об'єктів. Використання розробленої архітектури дозволить підвищити точність сегментації міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища.

3. Замість НМТ1 модифікованої нечіткої нейронної продукційної мережі Ванга-Менделя запропоновано використання ІНМТ2. Розглянуті ІНМТ2, з одного боку, забезпечують формалізацію більшої кількості додаткових ступенів невизначеності у порівнянні з НМТ1, з іншого боку, є такими, що "реалізуються" при розробці нечітких систем (моделей) та мають меншу обчислювальну складність, у порівнянні з НМТ2.

4. Для порівняння сігментованих різночасових зображень за результатами розпізнавання міських будов на цифрових космічних і аерофотознімках доцільно використовувати метод зіставлення знімків після класифікації (сігментації).

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ:

- [1] Бровко, Е.А., Кандоба, И.Н., Корнилов, Ф.А. & Перевалов, Д.С. (2015). Оперативный мониторинг структурных изменений изображений объектов на космических снимках земной поверхности. *Программные продукты и системы*, 1(109), 79-86.
- [2] Гук А.П., Евстратова Л.Г., Хлебникова Е.П., Алтынцев М.А., Арбузов С.А., Гордиенко А.С. & Гук А.А. (2013). Разработка методик автоматизированного дешифрирования аэрокосмических снимков. Выявление изменения состояния территорий по многозональным космическим снимкам, полученным на разные даты. *Геодезия и картография*, (8), 33-41.
- [3] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. *In International Conference on Medical image computing and computer-assisted intervention* (pp. 234-241). Springer, Cham.
- [4] Мищенко, В.А., Астахова, И.Ф. & Краснояров, А.А. (2012). Алгоритм обучения нечеткой нейронной сети Ванга-Менделя для распознавания рукопечатных символов в работе почтовой службы. *Вестник Воронежского государственного университета. Серия: Системный анализ и информационные технологии*, (1), 141-146.
- [5] Mendel J.M., Nagras H., John R.I. Standard Background Material About Interval Type-2 Fuzzy Logic Systems. *IEEE CIS Standards Committee*. Retrieved from <http://iee-cis.org/technical/standards>.