

*Олексій Володимирович Коломійцев (доктор технічних наук, с.н.с.)<sup>1</sup>*  
*Володимир Володимирович Пустоваров<sup>2</sup>*

<sup>1</sup>*Національний технічний університет «Харківський політехнічний інститут», Харків, Україна*

<sup>2</sup>*Харківське представництво генерального замовника – Державного космічного агентства України, Харків, Україна*

## ПРОПОЗИЦІЇ ЩОДО ПІДВИЩЕННЯ ТОЧНОСТІ СЕГМЕНТАЦІЇ МІСЬКИХ БУДОВ НА ЦИФРОВИХ КОСМІЧНИХ І АЕРОФОТОЗНІМКАХ ПРИ АВТОМАТИЗОВАНОМУ МОНІТОРИНГУ МІСЬКОГО СЕРЕДОВИЩА

*У основі топографії, що використовується у військовій сфері, – є оволодіння способами вивчення місцевості. Тому, отримання кількісних та якісних характеристик місцевості по космічних і аерофотознімках – є однією з основних частин військової топографії. В роботі розроблені пропозиції щодо створення архітектури ансамблю нейронних мереж на основі модифікованої згорткової нейронної мережі (ЗНМ) для сегментації зображень міських будов (автоматичного кодувальника) U-Net у вигляді суперпозиції функцій. Розроблені пропозиції щодо модифікації (автоматичного кодувальника) U-Net, що складаються із: модифікації підмережі вилучення ознак з використанням у якості енкодера попередньо навченої глибокої ЗНМ VGG, а також реалізації класифікатора автокодувальника з використанням модифікованої нечіткої нейронної мережі Ванга-Менделя на основі ІНМТ2 для піпксельної класифікації визначених об'єктів та створення узагальненої топології нейромережевої моделі для сегментації міських будов. Модифікація (автоматичного кодувальника) U-Net дозволить підвищити точність сегментації міських будов на цифрових космічних та аерофотознімках при автоматизованому моніторингу міського середовища у реальному масштабі часу.*

**Ключові слова:** сегментація міських будов; класифікація; згорточна нейронна мережа; нечітка нейронна мережа; енкодер; декодер; автокодувальник; цифрові космічні та аерофотознімки.

### Вступ

**Постановка проблеми.** Ведення бойових дій (операцій) Збройними силами здійснюється на поверхні Землі та у просторі, що її оточує. Бій, операція, пересування та розміщення підрозділів, частин, з'єднань, їхні навчання нерозривно пов'язані з певною ділянкою, районом, смугою земної поверхні, якій притаманні свої властивості.

При цьому, у військовій справі, ділянка земної поверхні з розташованими на ній об'єктами, на якій війська виконують свої завдання, називається місцевістю. Властивості місцевості, способи їхнього визначення, орієнтування на місцевості та виконання вимірювань на ній вивчає військова топографія – це галузь воєнної науки, яка поєднує теоретичні основи наук про просторові характеристики земної поверхні та способи їхнього відображення [21].

Відомо, що топографічна карта – це документ, який створюється заздалегідь і відображає стан місцевості на час її видання. Тому, важливо

володіти і сучасної способами та засобами отримання інформації про місцевість.

Космічні і аерофотознімки, їх різновиди та способи використання – є другою важливою частиною військової топографії про місцевість як засобів управління при вирішенні тактичних завдань, способи орієнтування на місцевості та використання наземних навігаційних засобів.

Бурхливе зростання міських будов на території сучасної інфраструктури держави потребує вдосконалення систем управління містом. Ефективне управління сучасним містом (мегаполісом) залежить від своєчасного отримання даних про забудову міських територій, що забезпечується веденням відповідного моніторингу.

Підвищити значення оперативності і достовірності моніторингу міських забудов можливо за рахунок використання даних, що отримані від автоматизованого аналізу космічних і

аерофотознімків території міста. При цьому, найбільш важливим – є виявлення змін міського середовища і аналіз причин їх виникнення.

До змін міського середовища, що можуть контролюватися за даними автоматизованого аналізу космічних і аерофотознімків, можливо віднести: знесення і зведення будівель, зміни площі зелених насаджень, будівництво або розширення доріг, виявлення незаконного будівництва і тощо.

За умови аналізу космічних і аерофотознімків, при автоматизованому моніторингу міського середовища, можна розглядати перший етап – автоматичного розпізнавання різноманітних об'єктів на місцевості за двома класами (для здійснення розпізнавання): міська будова та не будова. При цьому, за базовий метод розпізнавання можливо обрати семантичну (смыслову) сегментацію зображень (міських будов), що полягає у виділенні на зображенні областей, кожна з яких відповідає певній ознаці.

На даний час для сегментації зображень широко використовуються глибокі нейронні мережі [1], які мають високу точність розпізнавання об'єктів сегментації.

Таким чином, необхідно розробити архітектуру нечіткої згорткової нейронної мережі (ЗНМ) на базі автоенкодера U-Net, де, у якості підмереж вилучення ознак (енкодера), необхідно використовувати попередньо навчену ЗНМ VGG16, а у якості класифікатора необхідно використовувати модифіковану нечітку нейронну мережу Ванга-Менделя [2, 3].

#### **Аналіз останніх досліджень і публікацій.**

У [1] проведено аналіз глибоких нейронних мереж для семантичної сегментації у різних предметних галузях. Відмічено, що для вирішення завдань семантичної сегментації у рамках теорії глибокого навчання найбільш ефективними – є різні реалізації автокодувальника (autoencoder).

У [2] для вилучення ознак у U-Net у якості енкодера використовується ЗНМ, структура якої відповідна декодеру мережі. З точки зору технології передавання навчання, попередньо навчені глибокі ЗНМ, наприклад VGG, забезпечують формування більш якісних карт ознак, що безпосередньо впливає на ефективність сегментації зображень. У якості класифікатора U-Net виступають повнозв'язні шари, які перетворені у згорткові. При цьому, нечіткі нейронні мережі забезпечують більш високу точність розпізнавання порівняно з класичними повнозв'язними мережами [2, 3].

У [4-8] розглянути питання щодо розробки архітектури ЗНМ для сегментації об'єктів на

зображеннях за різною метою. Однак, питання щодо комплексного застосування різних архітектур мереж для сегментації міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища, у тому числі з використанням нечітких нейронних мереж, – не розглядаються.

У рамках глибокого навчання [1, 5, 9-18]. можливо реалізувати семантичну сегментацію за архітектурами 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.

Отже, у роботі, у якості базової архітектури глибокої нейронної мережі для сегментації міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища обрано мережу U-Net.

**Метою статті** є підвищення точності сегментації міських будов на цифрових космічних і аерофотознімках.

#### **Виклад основного матеріалу дослідження**

Архітектура U-Net, що запропонована у 2015 році, виявилася досить ефективною для вирішення різних завдань, таких як сегментація нейронних структур, рентгенографія і завдання відстеження клітин. Мережа характеризується кодером з послідовністю рівнів згортки і максимального пулу. Шар декодування містить дзеркальну послідовність транспонованих згорток. Він веде себе як традиційний автокодувальник.

Щоб розглянути різні рівні абстракції, U-Net реалізує пропущені з'єднання для копіювання незжатих активацій з кодуєть блоків в їх дзеркальні копії серед блоків декодування, як показано на рисунку 1.

Екстрактор функцій U-Net також може бути модернізований для поліпшення карт сегментації міських будов. Мережа під назвою «Тірамісу ...» застосувала концепцію U-Net з використанням екстрактора функцій на основі щільної мережі. Інші варіації включають використання капсульних мереж поряд з локальною маршрутизацією.

Тому, U-Net обраний у якості переможця конкурсу ISBI. У наборі даних PhC-U373 середнє значення IoU склало 0,9203, тоді як друге місце було на рівні 0,83. У наборі даних DIC-HeLa він набрав середнє значення IoU 0,7756, що було значно краще, ніж другий кращий підхід, який набрав лише 0,46.

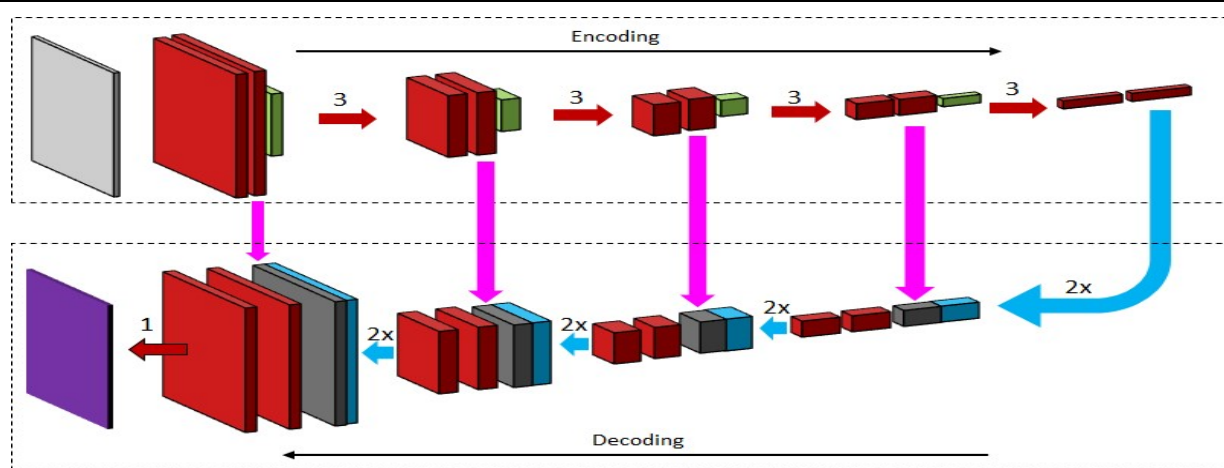


Рис. 1. Схематичне представлення U-Net

Для сегментації міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища у якості критеріїв вибору архітектури глибоких нейронних мереж можуть бути використані: прикладна спрямованість архітектури; забезпечення точності розпізнавання, що передбачається; мінімальна навчальна вибірка для забезпечення точності розпізнавання, що передбачається; мінімальна кількість епох для навчання; достатня швидкість розпізнавання у режимі реального часу.

Визначеним критеріям відповідає архітектура згорткової нейронної мережі U-Net, яка є повнозв'язаною згортковою мережею, що модифікована для сумісної роботи з меншою кількістю прикладів (навчальних образів) та робить точну сегментацію. Така мережа містить енкодер та декодер. 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. При цьому, розуміється, що кодер – це типова згорткова мережа, яка містить два послідовних шари згортки розміром  $3 \times 3$ , після яких йде шар ReLU і пулінг з функцією максимуму  $2 \times 2$  з кроком 2. У декодеру кожен крок містить шар, який зворотний пулінгу, що розширює карту ознак, після якого використовується згортка  $2 \times 2$ , що зменшує кількість каналів ознак. Виконується конкатенація з відповідним чином обрізаною картою ознак кодера і двох згорток  $3 \times 3$ , після кожної з якої йде шар ReLU. Однак, обрізка потрібна через те, що при конкатенації втрачаються граничні пікселі у кожній згортці. На останньому шарі використовується повнозв'язна згортка розміром  $1 \times 1$  для приведення кожного 64-компонентного вектора ознак до необхідної кількості класів [5].

До особливостей використання модифікованої ЗНМ U-Net як підмережі вилучення ознак у роботі є:

використання попередньо навченої ЗНМ VGG16 у якості енкодера; відсутність додаткових класифікаторів, які є у оригінальній версії ЗНМ VGG16 (табл. 1) та U-Net; у введенні додаткових шарів передискретизації розмірності карт ознак для реалізації конкатенації відповідних шарів енкодера (шари UpSampling). Класичний декодер збільшується у глибину за рахунок трьох послідовних шарів згортки розміром  $3 \times 3$  на кожному кроці (блоці декодера).

Отже, основними принципами, що лежать у навчанні модифікованої підмережі вилучення ознак U-Net, є:

- використання технології «bottleneck features», за якою, на виході попередньо навченої моделі знімаються ознаки (у роботі це енкодер у форматі VGG16) і подаються на вхід навченою частиною мережі (у роботі це декодер у форматі U-Net) і у подальшому навчається тільки друга частина (у роботі це тільки декодер). Особливістю моделі U-Net є наявність проміжних зв'язків між енкодером і декодером, тобто для декодера має місце кілька входів (у роботі це випадок 5), що значно ускладнює використання технології «bottleneck features» для семантичної сегментації (на відміну від нейронного класифікатора у якого тільки один вхід). Тому, розроблено підхід до вирішення даної проблеми за рахунок використання відповідних шарів VGG16 у рамках енкодера (табл. 2) та шарів передискретизації UpSampling у рамках структури декодера;

- використання технології тонкого налаштування моделі. У роботі послідовно разморозуються 1 та 2 шари енкодера, поєднуються ваги енкодера (попередньо навченої моделі VGG16 на бібліотеці зображень ImageNet) і отримані на першому етапі ваги декодера та виконується тонка настройка разморожених шарів енкодера і шарів 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

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:])

До основних елементів архітектури підмережі вилучення ознак модифікованої ЗНМ 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 та декодера;

– допоміжні шари у структурі декодера – BatchNormalization та UpSampling2D.

Згортковий шар можна представити у вигляді наступної функції:

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

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  $H^{out} \times W^{out} \times D^{out}$ .

З урахуванням формального представлення CONV-шару результат визначення елемента тензора  $Y_{H^{out}W^{out}D^{out}}$  можливо представляється як:

$$y_{i^{out}j^{out}k^{out}} = b_{k^{out}} + \sum_{i^{in}=0}^{H^f-1} \sum_{j^{in}=0}^{W^f-1} \sum_{k^{in}=1}^{D^{in}} f_{k^{out}}^{i^f j^f k^{in}} * x_{(i^{in}+i^{out})} \left( j_{+i^{out}}^{in} \right)_{k^{in}}, \quad (2)$$

де  $y_{i^{out}j^{out}k^{out}}$  – елемент тензора  $Y_{H^{out}W^{out}D^{out}}$ ,

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

$b_{k^{out}}$  – елемент множини  $B$ ,  $b_{k^{out}} \in B$ ;

$f_{k^{out}}^{i^f j^f k^{in}}$  – елемент комбінованого тензора

четвертого рангу  $F_{D^{out}}^{H^f W^f D^{in}}$ , що описує рецептивні

поля CONV-шару  $f_{k^{out}}^{i^f j^f k^{in}} \in F_{D^{out}}^{H^f W^f D^{in}}$ ;

$x_{(i^{in}+i^{out})} \left( j_{+i^{out}}^{in} \right)_{k^{in}}$  – елемент тензора

$X_{H^{in}W^{in}D^{in}}$ ,  $x_{(i^{in}+i^{out})} \left( j_{+i^{out}}^{in} \right)_{k^{in}} \in X_{H^{in}W^{in}D^{in}}$ ;

\* – операція згортки зв'язаних тензорів.

Формально структуру шару активації можливо представити кортежем:

$$ACT = \left\langle X, Y, t^{ACT} \right\rangle \quad (3)$$

де  $X$  – багатовимірний масив вхідних даних, що представляє множину вхідних карт ознак

$$X \subset \mathbb{R}^{H^{in} \times W^{in} \times D^{in}};$$

$Y$  – багатовимірний масив вихідних даних, що представляє множину вихідних карт ознак

$$Y \subset \mathbb{R}^{H^{out} \times W^{out} \times D^{out}};$$

$t^{ACT}$  – множина типів функцій активації:

$$t^{ACT} = \left\{ \text{сігмоїд}, \text{гіперболічний тангенс}, \text{ReLU}, \text{SoftReLU}, \text{LeakyReLU} \right\}.$$

На рис. 2 подані основні активаційні функції, які використовуються на практиці.

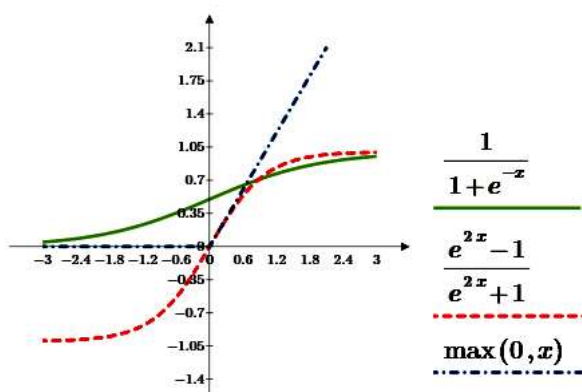


Рис. 2. Приклади активаційних функцій, що використовуються у шарі активації (сігмоїд, гіперболічний тангенс, RELU)

Шар активації, що реалізує функцію RELU, можливо представити у вигляді:

$$Y_{H^{out}W^{out}D^{out}} = f_{ReLU} \left( X_{H^{in}W^{in}D^{in}} \right) \quad (4)$$

Структуру POOL-шару можна представити у вигляді наступного кортежу:

$$POOL = X, Y, (H^{Pool}, W^{Pool}), (S_h, S_w), t^{Pool} \quad (5)$$

де  $X$  – багатовимірний масив вхідних даних, що представляє множину вхідних карт ознак

$$X \subset \mathbb{R}^{H^{in} \times W^{in} \times D^{in}};$$

$Y$  – багатовимірний масив вихідних даних, що представляє множину вихідних карт ознак

$$Y \subset \mathbb{R}^{H^{out} \times W^{out} \times D^{out}};$$

$(H^{Pool}, W^{Pool})$  – просторові розміри вікна об'єднання POOL-шару;

$(S_h, S_w)$  – крок зміщення рецептивного поля по висоті і ширині зображення (карти ознак) відповідно;

$t^{Pool}$  – множина типів операцій об'єднання

$t^{Pool} = \{ "max", "avg", "sum" \}$ . У роботі використовуються два типи операцій – "max" і "avg".

Характеристики вихідних даних POOL-шару можливо визначити наступним чином:

$$W^{out} = \frac{W^{in} - W^{Pool}}{S_w} + 1 \quad (6)$$

$$H^{out} = \frac{H^{in} - H^{Pool}}{S_h} + 1 \quad (7)$$

$$D^{out} = D^{in} \quad (8)$$

З урахуванням формального представлення POOL-шару, згідно з (5), результат визначення елемента тензора:  $Y_{H^{out}W^{out}D^{out}}$  можливо визначити

для операцій "max" і "avg", відповідно, як:

$$y_{i^{out}j^{out}k^{out}} = \max_{1 \leq i^{pool} \leq H^{pool}, 1 \leq j^{pool} \leq W^{pool}} x_{(i^{in}+i^{out}-1)(j^{in}+j^{out}-1)k^{in}} \quad (9)$$

$$y_{i^{out}j^{out}k^{out}} = \frac{1}{H^{pool}W^{pool}} \times \sum_{1 \leq i^{pool} \leq H^{pool}, 1 \leq j^{pool} \leq W^{pool}} x_{(i^{in}+i^{out}-1)(j^{in}+j^{out}-1)k^{in}} \quad (10)$$

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

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

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

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

Отже, структуру архітектури підмережі вилучення ознак у результаті модифікації архітектури ЗНМ U-Net можливо представити у вигляді таблиці 3.

Нечітку нейрону мережу Ванга-Менделя можливо використовувати у якості класифікатора модифікованого декодера U-Net, наступним чином:

- 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 декілька нейронів-нормалізаторів, кількість яких відповідає кількості пар третього шару.

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]
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]

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

$$R : IF \beta_1 \text{ is } \alpha_1 \text{ AND } \dots \text{ AND } \beta_n \text{ is } \alpha_n \\ THEN \beta_{n+1} = c_j^p, \quad (11)$$

де  $\beta_1$  – найменування вхідної лінійної згортки (ЛЗ), що задається кортежем  $\langle \beta_i, T_i, X_i, M_i \rangle$ ,  $i = 1, \dots, n$  і  $n$  – кількість вхідних ЛЗ;

$T_i = \{ \alpha_i \}$  – множина значень (термів) вхідної

ЛЗ правила  $R$ , кожне з яких є найменуванням нейронної згортки (НЗ) для опису значень параметрів сегментації;

$X_i$  – область значень НЗ, найменування яких входять в  $T_i$ ,  $M_i$  – 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 в рамках пропонованого підходу визначаються на етапі формування бази правил;

$\alpha_i$  – значення терма вхідної ЛЗ у вигляді найменування НЗ (лінгвістичного значення

параметра розпізнавання компактного (точкового) об'єкту), що задається кортежем  $\langle \alpha_i, X_i, \tilde{A} \rangle$ ,  $i = 1, \dots, n$ ,  $\tilde{A} \subseteq X_i$ ;

$\tilde{A} = \{ x, \mu_{\tilde{A}}(x) \}$  – IHMT2 на множині  $X_i$ , що описує можливі значення, які може приймати НЗ  $\alpha_i \in T_i$ ;

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

$Y$  – область значень термів, найменування яких входять в  $T_{n+1}$ , що представляють номер розпізнаваного класу;

$M_{n+1}$  – 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  $C_p = \{ c_j^p \}$ ,  $j = 1, \dots, m$ ,  $m$  – кількість класів сегментації.

Узагальнена топологія нейромережевої моделі для сегментації міських будов на цифрових космічних і аерофотознімках наведена на рис. 3.

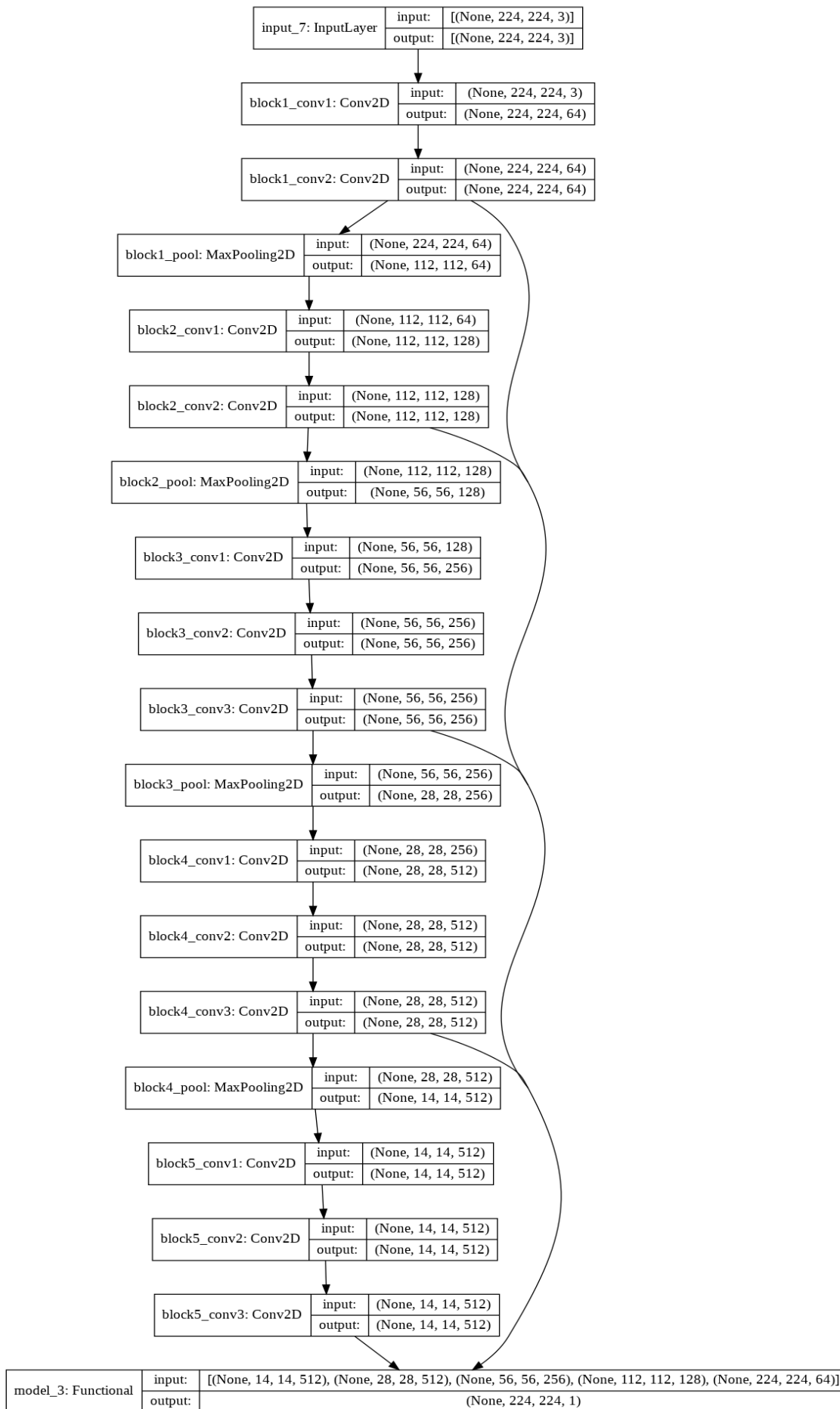


Рис. 3. Узагальнена топологія нейромережевої моделі для сегментації міських будов



## Висновки й перспективи подальших досліджень

Таким чином, запропоновано формальне подання нейромережевого ансамблю нейронних мереж на основі модифікованої згорткової нейронної мережі для сегментації міських будов. (автоматичного кодувальника) U-Net у вигляді суперпозиції функцій.

Розроблено архітектуру ансамблю нейронних мереж та узагальнену топологію нейромережевої моделі для сегментації міських будов.

Використання розробленої модифікованої ЗНМ дозволить підвищити точність сегментації

міських будов на цифрових космічних і при автоматизованому моніторингу міського середовища.

Перспективами подальших досліджень може стати вдосконалення архітектури ансамблю нейронних мереж на основі подальшої модифікації ЗНМ U-Net та нечіткої нейронної мережі Ванга-Менделя, як класифікатора модифікованого декодера U-Net, для сегментації міських будов на цифрових космічних і аерофотознімках при автоматизованому моніторингу міського середовища.

## Література

1. A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation (2015), <https://arxiv.org/abs/1511.00561>.
2. Олізаренко, С.А., Капранов, В.А. & Сафронов, Р.В. (2016). "Розробка архітектури нечіткої згорткової нейронної мережі для розпізнавання компактних (точкових) об'єктів на цифровому аерофотознімку". Системи озброєння і військова техніка. № 4. С. 38-41.
3. 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. С. 141-146.
4. Семантическая сегментация: краткое руководство": [Электронный ресурс]. URL: <https://neurohive.io/ru/osnovnyy-data-science/semantic-segmentation/> (Дата звернення: 14.07.2019).
5. Ronneberger, O., Fischer, P. & Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention", Springer. p.p. 234-241.
6. Milletar, F., Navab, N. & Ahmadi, S. (2016). "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, 2016 Fourth International Conference on 3D Vision (3DV)". Stanford, CA., p.p. 565-571.
7. Иванов, Е.С., Тищенко, И.П. & Виноградов, А.Н. (2019). "Сегментация мультиспектральных снимков с применением сверточных нейронных сетей" Современные проблемы дистанционного зондирования Земли из космоса. № 1. С. 25-34.
8. Соловьев, Р.А., Тельпухов, Д.В. & Кустов, А.Г. (2017). "Автоматическая сегментация спутниковых снимков на базе модифицированной сверточной нейронной сети UNET". Инженерный вестник Дона. № 4 [ivdon.ru/ru/magazine/archive/n4y2017/4433](http://ivdon.ru/ru/magazine/archive/n4y2017/4433).
9. Fully Convolutional Networks for Semantic Segmentation", CVPR (2015), <https://arxiv.org/abs/1411.4038>.
10. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation" (2016), <https://arxiv.org/abs/1611.09326>.
11. ENet: "A Deep Neural Network Architecture for Real-Time Semantic Segmentation" (2016), LinkNet: "Feature Forwarding: Exploiting Encoder Representations for Efficient Semantic Segmentation" (2017), <https://arxiv.org/abs/1606.02147>.
12. Multi-Path Refinement Networks for High-Resolution Semantic Segmentation", CVPR (2017), <https://arxiv.org/abs/1611.06612>.
13. Pyramid Scene Parsing Network", CVPR (2017), <https://arxiv.org/abs/1612.01105>.
14. Mask-RCNN: extends Faster R-CNN" (2017), <https://arxiv.org/abs/1703.06870>.
15. Gated Feedback Refinement Network for Dense Image Labeling", CVPR (2017), <http://www.cs.umanitoba.ca/~ywang/papers/cvpr17.pdf>.
16. Decoupled Deep Neural Network for Semi-supervised Semantic Segmentation", NIPS (2015), <https://arxiv.org/abs/1506.04924>.
17. Semi and Weakly Supervised Semantic Segmentation Using Generative Adversarial Network" (2017), <https://arxiv.org/abs/1703.09695>.
18. Коломійцев, О., Третьяк, В., Закіров, З., Кукобко, С., Калачова, В., & Мартовицький, В. (2020). Оптимізація завантаження файлів сховища даних в olap-файли на основі рангового підходу. InterConf, (25), 108-117. вилучено із <https://ojs.ukrlogos.in.ua/index.php/interconf/article/view/4300>.
19. Коломійцев, О., Рябуха, Ю., Карлов, Д., & Третьяк, В. (2020). Особливості організації і класифікація сучасних технологій реплікації даних. InterConf, (14). вилучено із <https://ojs.ukrlogos.in.ua/index.php/interconf/article/view/2058>.
20. Коломійцев, О., Рябуха, Ю., Калачова, В., & Третьяк, В. (2020). Аналіз методів і процедур шкального оцінювання в задачах прийняття рішень при проектуванні і супроводженні розподілених автоматизованих інформаційних систем. InterConf, (15). вилучено із <https://ojs.ukrlogos.in.ua/index.php/interconf/article/view/2309>.
21. Ткачук П.П., Тревого І.С. Військова топографія Підручник / Т.М. Гребенюк, В.Д. Макаревич, І.С. Тревого, В.М. Корольов, В.М. Глотов, О.П. Полець, В.Ю. Жидков. – Львів: Видавництво Львівської політехніки, 2011. – 416 с.



**ПРЕДЛОЖЕНИЯ ПО ПОВЫШЕНИЮ ТОЧНОСТИ СЕГМЕНТАЦИИ ГОРОДСКИХ СТРОЕНИЙ НА ЦИФРОВЫХ КОСМИЧЕСКИХ И АЭРОФОТОСНИМКАХ ПРИ АВТОМАТИЗИРОВАННОМ МОНИТОРИНГЕ ГОРОДСКОЙ СРЕДЫ**

*Алексей Владимирович Коломийцев (доктор технических наук, с.н.с.)<sup>1</sup>*

*Владимир Владимирович Пустоваров<sup>2</sup>*

<sup>1</sup>*Национальный технический университет «Харьковский политехнический институт», Харьков, Украина*

<sup>2</sup>*Харьковское представительство генерального заказчика – Государственного космического агентства Украины, Харьков, Украина*

*В основе топографии, которая используется в военной сфере, - лежит овладение способами изучения местности. По этому, получение количественных и качественных характеристик местности по космическим и аэрофотоснимкам – является одной из основных частей военной топографии в настоящее время. В работе разработаны предложения по формальному представлению нейронного сетевого ансамбля нейронных сетей на основе модифицированной свертываемой нейронной сети для сегментации городских строений (автоматического кодировщика) U-Net в виде суперпозиции функций. Разработана архитектура ансамбля нейронных сетей на основе модифицированной сверточной нейронной сети (СНС) для сегментации городских строений (автоматический кодировщик) U-net. Разработаны предложения по модификации (автоматического кодировщика) U-Net которые включают в себя: модификацию подсети изъятия признаков с использованием в качестве энкодера предварительно обученной глубокой СНС VGG, а так же реализацию классификатора автоматического кодировщика с использованием модифицированной нечеткой нейронной сети Ванга-Менделя на основе ИНМТ2 для попиксельной классификации определенных городских строений и формирование обобщенной топологии нейросетевой модели для сегментации городских строений. Модификация (автоматического кодировщика) U-Net позволит повысить точность сегментации городских строений на цифровых космических и аэрофотоснимках при автоматизированном мониторинге городской среды.*

*Ключевые слова:* сегментация городских строений, классификация, сверточная нейронная сеть, нечеткая нейронная сеть, энкодер, декодер, автокодировщик, цифровые космические и аэрофотоснимки.

**SUGGESTIONS ON INCREASE OF EXACTNESS OF SEGMENTATION OF URBAN BUILDINGS ON DIGITAL SPACE AND AEROPHOTOS WITH AUTOMATED MONITORING OF THE URBAN ENVIRONMENT**

*Oleksii Kolomiitsev (Doctor of Technical Science, Senior Research Fellow)<sup>1</sup>*

*Volodymyr Pustovarov<sup>2</sup>*

<sup>1</sup>*National Technical University is the «Kharkiv Polytechnic Institute», Kharkiv, Ukraine*

<sup>2</sup>*Kharkov representative office of general customer – the State space agency of Ukraine, Kharkiv, Ukraine*

*In basis of topography that is used in a military sphere, - a capture lies the methods of study of locality. On it, receipt of quantitative and quality descriptions of locality on space and to the airphotos - is one of basic parts of military topography presently. Architecture of U-Net zdktncz effective enough for the decision of different tasks, such as segmentation of neuron structures, sciagraphy et cetera. A network is characterized by a coder with the sequence of levels of convolutional and maximal pool. Thus, the layer of decoding contains the mirror sequence of convolutional that transponirovan's. He behaves as traditional automatic encode. Extractor of functions of U-Net can be modernized for the improvement of maps of segmentation of municipal structures. Unclear to the neuron decorate a pattern Vanga-Mandelja it maybe to use the classifier of the modified decoder in quality U-Net. Formal presentation of neuron network ensemble of neuron networks is in-process offered on the basis of the modified rolled up neuron network for segmentation of municipal structures (automatic encode) of U-Net as super position of functions. Architecture of ensemble of neuron networks is worked out on the basis of the modified convolutional neuron network (CNN) for segmentation of municipal structures (automatic encode) of U-Net. Modification (automatic encode) of U-Net is conducted due to modification of subnet of withdrawal of signs with using as encoder preliminary trained deep CNN VGG, and similarly due to realization of classifier of automatic encode with the use of the modified unclear neuron network Vanga-Mandelja on the basis of INMT2 for on a pixel classification of certain municipal structures and the generalized topology of by a neuron network model is worked out for segmentation of municipal structures. Modification (automatic encode)*

of U-Net will allow to promote exactness of segmentation of municipal structures on digital space and airphotos at the automated monitoring of municipal environment.

**Key words:** segmentation municipal structures, classification, convolutional neural network, fuzzy neural network, encoder, decoder, autoencoder, digital space and airphotos.

### References

1. A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation (2015), <https://arxiv.org/abs/1511.00561>.
2. Olizarenko, S.A., Kapranov, V.A. & Safronov, R.V. (2016). "Rozrobka arhitekturi nechitkoi zgotovnoyi neironoi meregi dlja rozpiznavannja kompaktnih (tochkovih) ob'ektiv na cifrovomy aerofotoznmiky". Sistemi ozbroennja i viickova tehnika. № 4. S. 38-41.
3. Mishenko, V.A., Astahova, I.F. & Krasnojarov, A.A. (2012). "Algoritm obuchnija nechitkoi nerjnoi seti Vanga-Mendelja dlja raspoznavanija rukopechnih snimkov v rabote pochtovoi slugbi" Vesnik Voronezskogo gosudarstvennogo universiteta. Serija: Sistemnij analiz I informacionne tehnologii. № 1. S. 141-146.
4. Semanticheskaja segmentacija: kratkoe rukovodstvo": [Elektronii resurs]. URL: <https://neurohive.io/ru/osnovy-data-science/semantic-segmentation/> (Data zvernennja: 14.07.2019).
5. Ronneberger, O., Fischer, P. & Brox, T. (2015). "U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention", Springer. P. 234-241.
6. Milletar, F., Navab, N. & Ahmadi, S. (2016). "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, 2016 Fourth International Conference on 3D Vision (3DV)". Stanford, CA., p.p. 565-571.
7. Ivanov, E.S., Tishenko, I. P. & Vinogradov, A. N. (2019). "Segmentacija myltispektralnih snimkov s primeneniem svertochnih neironih setei". Sovremennije problem distancionnogo zondirovanija Zemli iz kosmosa. № 1. S. 25-34.
8. SSolov'ev, R.A., Tel'puhov, D.V. & Kustov, A.G. (2017). "Avtomatičeskaja segmentacija sputnikovih snimkov na baze modifirovannoi svertočnoi neironnoi seti UNET". Inženierij vesnik Dona. № 4 [ivdon.ru/ru/magazine/archive/n4y2017/4433](http://ivdon.ru/ru/magazine/archive/n4y2017/4433).
9. Fully Convolutional Networks for Semantic Segmentation", CVPR (2015), <https://arxiv.org/abs/1411.4038>.
10. The One Hundred Layers Tiramisu: Fully Convolutional DenseNets for Semantic Segmentation" (2016), <https://arxiv.org/abs/1611.09326>.
11. ENet: "A Deep Neural Network Architecture for Real-Time Semantic Segmentation" (2016), LinkNet: "Feature Forwarding: Exploiting Encoder Representations for Efficient Semantic Segmentation" (2017), <https://arxiv.org/abs/1606.02147>.
12. Multi-Path Refinement Networks for High-Resolution Semantic Segmentation", CVPR (2017), <https://arxiv.org/abs/1611.06612>.
13. "Pyramid Scene Parsing Network", CVPR (2017), <https://arxiv.org/abs/1612.01105>.
14. "Mask-RCNN: extends Faster R-CNN" (2017), <https://arxiv.org/abs/1703.06870>.
15. "Gated Feedback Refinement Network for Dense Image Labeling", CVPR (2017), <http://www.cs.umanitoba.ca/~ywang/papers/cvpr17.pdf>.
16. "Decoupled Deep Neural Network for Semi-supervised Semantic Segmentation", NIPS (2015), <https://arxiv.org/abs/1506.04924>.
17. "Semi and Weakly Supervised Semantic Segmentation Using Generative Adversarial Network" (2017), <https://arxiv.org/abs/1703.09695>.
18. Kolomiicev, O., Tretjak, B., Zakirov, Z., Kukobko, S., Kalachova, V., & Martovickii, V. (2020). Optimizacija zavantagejnja failiv shovicha danih v olap-faili na osnovi rangovogo pidhodu. InterConf, (25), 108-117. Vilucheno iz: <https://ojs.ukrlogos.in.ua/index.php/interconf/article/view/4300>.
19. Kolomiicev, O., Rjabuha, Y., Karlov, D., & Tretjak, B. (2020). Osoblivosti organizacii i klasifikacija suchnih tehnologii replikacii danih. InterConf, (14). Vilucheno iz: <https://ojs.ukrlogos.in.ua/index.php/interconf/article/view/2058>.
20. Kolomiicev, O., Rjabuha, Y., Kalachova, V., & Tretjak, B. (2020). Analiz metodiv I procedur shkalnogo ocinyvannja v zadajah priinjattja rishen pri proektuvanni I suprovodženni rozpodilennih avtomatizovanih informacijnih sistem. InterConf, (15). Vilucheno iz: <https://ojs.ukrlogos.in.ua/index.php/interconf/article/view/2309>.
21. Tkachyk P, Trevogo I. Viiskova topografija. Pidrychnik. / T. Grebenyk, V. Makarevich, I. Trevogo, V. Korolov, V. Glotov, O. Polec, V. Gidkov. – Lviv: Vidavnistvo Lvivskoi politehniky. 2011. – 416 s.