КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (18), 2022
ISSN 2663 - 4023
DOI 10.28925/2663-4023.2022.18.2438
УДК 004.934

Терейковський Ігор Анатолійович
доктор технічних наук, професор,
професор кафедри системного програмування і спеціалізованих комп'ютерних систем
Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", Київ, Україна
ORCID ID: 0000-0003-4621-9668
tereikowski@ukr.net

Чернишев Денис Олегович
доктор технічних наук, професор, перший проректор
Київський національний університет будівництва і архітектури, Київ, Україна
ORCID ID: 0000-0002-1946-9242
taqm@ukr.net

Корченко Олександр Григорович
доктор технічних наук, професор, завідуючий кафедрою безпеки інформаційних технологій
Національний авіаційний університет, Київ, Україна
ORCID ID: 0000-0003-3376-0631
agkorchenko@gmail.com

Терейковська Людмила Олексіївна
кандидат технічних наук, доцент,
доцент кафедри інформаційних технологій проектування та прикладної математики
Київський національний університет будівництва і архітектури, Київ, Україна
ORCID ID: 0000-0002-8830-0790
tereikovskal@ukr.net

Терейковський Олег Ігоревич
аспірант кафедри безпеки інформаційних технологій
Національний авіаційний університет, Київ, Україна
ORCID ID: 0000-0001-5045-0163
tereikovskyio@gmail.com

ПРОЦЕДУРА ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ СЕГМЕНТАЦІЇ РАСТРОВИХ ЗОБРАЖЕНЬ

Анотація. В даний час засоби семантичної сегментації зображень, що базуються на використанні нейронних мереж, знаходять все ширше застосування в комп'ютерних системах різного призначення. Незважаючи на значні успіхи в даній галузі, однією з найважливіших невирішених задач є завдання визначення типу та параметрів згорткових нейронних мереж, що лежать в основі кодера та декодера. В результаті проведених досліджень розроблено відповідну процедуру, що дозволяє адаптувати нейромережевий кодер і декодер до таких умов задачі сегментації: розмір зображення, кількість кольорових каналів, допустима мінімальна точність сегментації, допустима максимальна обчислювальна складність сегментації, необхідність маркування сегментів, необхідність виділення кількох сегментів, необхідність виділення деформованих, зміщених та повернутих об'єктів, допустима максимальна обчислювальна складність навчання нейромережевої моделі; допустимий термін навчання нейромережевої моделі. Виконання процедури застосування нейронних мереж для сегментації зображень полягає у формуванні базисного математичного забезпечення, побудові основних блоків та загальної схеми процедури. Розроблена процедура верифікована експериментально на прикладах семантичної сегментації зображень, що містять об'єкти типу автомобіль. Отримані експериментальні результати свідчать, що застосування запропонованої процедури дозволяє уникнувши складних довготривалих експериментів, побудувати нейромережеву модель, яка при достатньо

© І.А. Терейковський, Д.О. Чернишев, О.Г. Корченко, Л.О. Терейковська, О.І. Терейковський 2022

короткому терміні навчання забезпечує досягнення точності сегментації зображень близько 0,8, що відповідає найкращим системам аналогічного призначення. Показано, що шляхи подальших досліджень у напрямку вдосконалення методологічного забезпечення нейромережевої сегментації растрових зображень доцільно співвіднести з обґрунтованим використанням у кодері та декодері сучасних модулів та механізмів, адаптованих до значимих умов поставленої задачі. Наприклад, використання модулю ResNet дозволяє за рахунок нівелювання ефекту падіння градієнта збільшити глибину нейронної мережі, а модуля Inception забезпечує зменшення кількості вагових коефіцієнтів та обробку об'єктів різного розміру.

Ключові слова: нейромережева модель, згорткова нейронна мережа; семантична сегментація; растрове зображення; розпізнавання зображень.

ВСТУП

В сучасних умовах засоби розпізнавання зображень знаходять широке розповсюдження в інформаційних системах різноманітного призначення. Так, подібні засоби використовуються в системах відслідковування руху автомобільного транспорту та в системах біометричної аутентифікації для розпізнавання особи та емоційного стану. Одним із основних етапів такого розпізнавання є реалізація семантичної сегментації зображення, результатом якого є виділення на вхідному растровому зображенні цільового об’єкта, наприклад, автомобіля або обличчя людини. Проблематика семантичної сегментації зображень ускладнюється через можливе часткове або повне перекриття цільових об’єктів, нечіткість їх кордонів, різноваріантність розмірів та розміщення. Крім того, варіативність умов розпізнавання, посилюється внаслідок можливої зміни джерела отримання зображення, в якості якого в більшості комп’ютерних систем загального призначення використовується відеокамера, що зумовлює можливу зміну основних параметрів піддослідного растрового зображення – розміру, роздільної здатності та кількості кольорових каналів. Також негативний вплив на результат розпізнавання може бути викликаний завадами та ракурсом відеореєстрації. Внаслідок описаних труднощів засоби семантичної сегментації зображень, що базуються на класичних методах типу «нарощування областей», «аналіз кордонів», «подрібнення областей» являються вузькоспеціалізованими і потребують значної модифікації навіть при незначних змінах умов застосування. Висновок щодо обмеженості класичних методів виділення підтверджується результатами аналізу робіт [8, 14], в яких наведено опис моделей відслідковування зображень об’єктів за допомогою інтегрованих фільтрів різного типу. Разом з тим, за останні декілька років зріс інтерес до нейромережевих засобів виділення, що пояснюється доведеною ефективністю нейронних мереж (НМ) при вирішенні подібних складноформалізуємих та багатофакторних завдань.

Постановка проблеми. Вдосконалення методологічного забезпечення нейромережевої сегментації растрових зображень.

Аналіз останніх досліджень і публікацій. Використання нейромережевих моделей (НММ) для виділення об’єктів на зображеннях описане в науково-практичних роботах [3, 5, 14, 17]. Проведений аналіз дозволив побудувати показану на рис. 1 типову структурну схему функціонування нейромережевої системи семантичної сегментації растрових зображень. Також результати літературного аналізу дозволяють стверджувати про те, що у випадку використання нейромережевих технологій, використовується НММ, що складається із блоків – кодера та декодера. Завданням кодера є визначення багатовимірного масиву ознак початкового зображення, а завданням декодера являється

отримання обробленого зображення, у якого кожен із пікселів отримує маркер відношення до одного із виділених цільових об'єктів або фону. При цьому кодер та декодер представляють собою дещо модифіковані згорткові нейронні мережі (ЗНМ). Модифікація полягає у тому, що використовуються не повнорозмірні ЗНМ, а тільки блоки, які призначені для визначення значимих ознак. Тобто до складу таких ЗНМ входять тільки шари згортки та масштабування, а повнозв'язні шари вилучені.

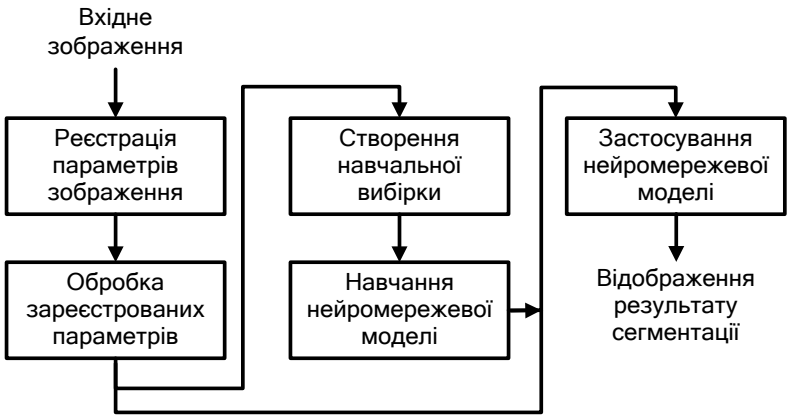

Рис. 1. Типова структурна схема функціонування нейромережевої системи семантичної сегментації растрових зображень

Як свідчать дані [11, 17], точність та обчислювальна ресурсоємність засобів виділення різнотипних об'єктів на зображеннях, що базуються на нейромережевих підходах, значно переважає інші технології виділення. Разом з тим, по причині новизни такого підходу та швидкого розвитку нейромережевих технологій, на сьогодні не вирішеними залишаються ряд окремих задач, розв'язання яких дозволить підвищити ефективність виділення. Відповідно до висновків [6, 16] та результатів аналізу [1, 11, 13], однією із таких задач являється адаптація типу та параметрів НММ до умов виділення об'єкта на зображенні. В свою чергу вирішенню цієї задачі має передувати визначення та формалізація означених умов виділення, розробка математичного апарату для оцінки точності та обчислювальної ресурсоємності моделі. Крім того, потребують уточнення критерії визначення ефективності НММ.

Мета статті. Розробка процедури застосування нейронних мереж для сегментації растрових зображень, що забезпечує достатню точність сегментації при застосуванні в системах біометричної аутентифікації.

РОЗРОБКА ТА ДОСЛІДЖЕННЯ ПРОЦЕДУРИ ЗАСТОСУВАННЯ НЕЙРОННИХ МЕРЕЖ ДЛЯ СЕГМЕНТАЦІЇ РАСТРОВИХ ЗОБРАЖЕНЬ

Базуючись на результатах [111] розробка процедури застосування нейронних мереж для сегментації зображень полягає у формуванні базисного математичного забезпечення, побудові основних блоків та загальної схеми процедури.

Формування базисного математичного забезпечення. Відповідно до [2, 9, 14], процедуру сегментації зображень на базі НМ можливо записати у вигляді:

$$Im_{in} \overset{Pr}{\rightarrow} Im_{pr} \overset{NC}{\rightarrow} Fm \overset{ND}{\rightarrow} Im_{out}, \quad (1)$$

ISSN 2663 - 4023

де Im_{in}, Im_{pr} – вхідне та оброблене зображення; $\overset{Pr}{\rightarrow}$ – оператор попередньої обробки зображення; $\overset{NC}{\rightarrow}$ – оператор нейромережевого кодування зображення; Fm – кортеж матриць ознак, отриманий в результаті операції нейромережевого кодування зображення; $\overset{ND}{\rightarrow}$ – оператор нейромережевого декодування зображення; Im_{out} – сегментоване зображення.

В свою чергу Im_{in}, Im_{pr} та Im_{out} представляють собою матрицю виду:

$$Im_x = \left\| \begin{matrix} pt_{1,1}(x) & \dots & pt_{n,1}(x) & \dots & pt_{N,1}(x) \\ \dots & \dots & \dots & \dots & \dots \\ pt_{1,m}(x) & \dots & pt_{n,m}(x) & \dots & pt_{N,m}(x) \\ \dots & \dots & \dots & \dots & \dots \\ pt_{1,M}(x) & \dots & pt_{n,M}(x) & \dots & pt_{N,M}(x) \end{matrix} \right\|, \quad (2)$$

де $pt_{n,m}(x)$ – опис кольору пікселя з координатами (n, m); N, M – розмір зображення по горизонталі та по вертикалі; x – етап обробки зображення.

Множина матриць ознак визначається так:

$$Fm = \{fm_1, \dots, fm_k, \dots, fm_K\}, \quad (3)$$

$$fm_k = \left\| \begin{matrix} \alpha_{1,1} & \dots & \alpha_{L,1} \\ \dots & \dots & \dots \\ \alpha_{J,1} & \dots & \alpha_{L,J} \end{matrix} \right\|_k, \quad (4)$$

де fm_k – матриця ознак, що співвідноситься з k-ою картою субдискретизації в останньому шарі ЗНМ, що використовується в якості кодеру; K – кількість карт субдискретизації в останньому шарі ЗНМ кодеру; $\alpha_{j,l}$ – значення ознаки в точці з координатами j, l для k-ої карти субдискретизації/згортки; L – розмір k-ої карти по горизонталі; J – розмір k-ої карти по вертикалі.

Зазначимо, що у випадку, коли останнім шаром ЗНМ кодеру є шар згортки, то у виразах (3, 4) fm_k співвідноситься з k-ою картою згортки. При цьому в переважній більшості відомих ЗНМ в одному шарі субдискретизації/згортки використовуються однакові за розміром карти.

За аналогією з [9, 12, 15] процес вдосконалення нейромережевих моделей кодеру та декодера на етапі їх проєктування можливо визначити так:

$$E \rightarrow max, \quad (5)$$

$$E = \sum_{i=1}^{I} \alpha_i w_i, \;\; \alpha_i \in \{\alpha\}, w_i \in \{w\}, \quad (6)$$

де E – функція ефективності засобів сегментації; I – кількість параметрів ефективності k_i – значення і-го параметру ефективності; α_i – ваговий коефіцієнт і-го параметру ефективності; $\{\alpha\}$ – множина вагових коефіцієнтів параметрів ефективності; $\{w\}$ - множина параметрів ефективності.

Базуючись на результатах [2, 6, 10], отримано перелік із дев'яти параметрів оцінки ефективності, що наведено в табл. 1.

ISSN 2663 - 4023

Таблиця 1

Перелік параметрів оцінки ефективності нейромережевих засобів семантичної сегментації зображень

Позначення	Характеристика параметру
w_1	Точність сегментації
w_2	Обчислювальна складність сегментації
w_3	Можливість маркування сегментів зображення
w_4	Можливість виділення декількох сегментів, що відповідають різним об'єктам
w_5	Можливість виділення сегментів, що відповідають об'єктам, які перекриваються
w_6	Можливість виділення деформованих об'єктів
w_7	Можливість виділення зсунутих об'єктів
w_8	Можливість виділення повернутих об'єктів
w_9	Термін навчання нейромережевої моделі

Отриманий перелік параметрів ефективності носить попередній характер і може бути розширений. Значення кожного із параметрів ефективності та значення їх вагових коефіцієнтів можливо визначити за допомогою методів експертного оцінювання. В базовому випадку їх значення можливо оцінити по бінарній шкалі 0 або 1.

Зазначимо, що використання виразів (5, 6) доцільно при попередньому визначенні перспектив використання нейромережевих засобів на етапі проєктування, коли умови задачі семантичної сегментації не можуть бути чітко описані за допомогою числових величин. У випадку оцінки ефективності розроблених експериментальних зразків нейромережевих засобів сегментації зображень доцільно в якості показників ефективності використовувати точність (A) та обчислювальну ресурсоємність (Θ) нейромережевих засобів. Це дозволяє оцінити ефективність експериментальних зразків нейромережевих засобів за допомогою виразу виду:

$$\begin{cases} A \to max \\ \Theta \le \Delta \end{cases}. \tag{7}$$

Особливості задачі сегментації зображень зумовлюють необхідність застосування для оцінки точності показники, що відображають схожість геометричних об'єктів. Відповідно до [2, 14, 17] передбачено застосування коефіцієнту Жаккара:

$$J = \frac{|N \cap M|}{|N \cap M| + |N - M| + |B - \mathrm{N}|}, \tag{8}$$

де N, M – області, що підлягають порівнянню.

За умов використання one-hot кодування очікуваного вихідного сигналу та однакових розмірів вхідного та сегментованого зображення вираз (8) деталізується так:

$$J = \frac{\sum_{i=1}^{I} n_i m_i}{\sum_{i=1}^{I} n_i + \sum_{i=1}^{I} m_i - \sum_{i=1}^{I} (n_i - m_i)}, \tag{9}$$

де I – кількість точок, що описують очікуваний вихідний сигнал нейромережевої моделі; n_i – значення, характерне для і-го пікселя сегментованого зображення; m_i-значення, характерне для і-го пікселя очікуваного вихідного сигналу.

ISSN 2663 - 4023

Побудова нейромережевого кодеру. Дані літературних джерел [2-4, 9, 12] свідчать, що в нейромережевих засобах, призначених для виділення об'єктів на растрових зображеннях, доцільно використовувати кодер, розроблений на базі ЗНМ типу LeNet-5, VGG, ResNet та GoogLeNet, що пояснюється їх високою ефективністю, апробованістю та наявністю доступного інструментарію для комп'ютерної реалізації.

Функціонал кодеру на базі ЗНМ описується за допомогою виразу виду:

$$F_C\left(\|R\|_{H,L,K}\right) = \left\{\|\mathrm{c}_1\|_{X,Y}, \|\mathrm{c}_2\|_{X,Y}, \dots, \|\mathrm{c}_N\|_{X,Y}\right\}, \tag{10}$$

де F_C – функція кодування вхідного зображення; $\|R\|_{H,L,K}$ – трьохвимірний масив значень, елементи якого співвідносяться з вхідним зображенням; K – кількість кольорових каналів вхідного зображення; H, L – розміри вхідного зображення.

Враховуючи наведені в [4, 6] характеристики ЗНМ типу LeNet та особливості структури класичного нейромережевого кодера, вираз, що характеризує залежність показників точності та обчислювальної ресурсоємності кодеру від його конструктивних параметрів, можливо записати у вигляді:

$$\Theta, A = F(S, \|b\|, \|m\|, C, H, \|k\|), \tag{11}$$

де S – кількість стеків; $\|b\|$ – масив, що містить розміри ядер згортки; b_w – розмір ядра згортки для w-го стеку; $\|m\|$ – масив, що містить значення масштабних коефіцієнтів; m_w – значення масштабного коефіцієнту для w-го стеку; $\|k\|$ – масив, що містить кількість карт згортки; k_w – кількість карт згортки для w-го стеку; C – кількість кольорових каналів, що підлягають обробці; H – розмір вхідного зображення.

З урахуванням (7, 11) отримано вираз, що являється підґрунтям визначення конструктивних параметрів нейромережевого кодера на базі LeNet:

$$\begin{cases} A(S, \|b\|, \|m\|, C, H, \|k\|) \rightarrow max \\ \Theta(S, \|b\|, \|m\|, C, H, \|k\|) \le \Delta_\Theta \end{cases}. \tag{12}$$

Визначити значення використаних в виразі (12) конструктивних параметрів можливо, базуючись на наведених в роботі [6] принципах:

1. Кількість стеків доцільно прирівняти до кількості рівнів розпізнавання експертом цільових об'єктів на зображенні, що підлягає аналізу;

2. Кількість карт згортки в певному згортковому шарі встановлюється рівним кількості значимих ознак, що розпізнаються на цьому ж рівні експертом.

Побудова нейромережевого декодеру. Функціонування декодеру, що полягає у відновленні сегментованого зображення можливо описати за допомогою виразу виду:

$$F_D\left(\left\{\|\mathrm{c}_1\|_{X,Y}, \|\mathrm{c}_2\|_{X,Y}, \dots, \|\mathrm{c}_N\|_{X,Y}\right\}\right) = \|r\|_{h,\mathrm{l},\mathrm{k}}, \tag{13}$$

де F_D – функція декодування сегментованого зображення; $\|\mathrm{c}_n\|_{X,Y}$ – масив значень, що співвідноситься з n-ою картою ознак останнього шару нейромережевого кодеру; N – кількість карт ознак в останньому шарі нейромережевого кодеру; X, Y – розмір карти ознак в останньому шарі нейромережевого кодеру; $\|r\|_{h,\mathrm{l},\mathrm{k}}$ – трьохвимірний масив значень, що співвідноситься з вихідним зображенням; k – кількість кольорових каналів вихідного зображення; h, l – розміри вихідного зображення.

ISSN 2663 - 4023

Відповідно до результатів [2, 3, 12] в основу розробки моделі нейромережевого декодеру, призначеного для використання в засобах семантичної сегментації зображень, може бути покладено:

1. Підхід на основі одноетапного ресамплінгу вихідного зображення.
2. Підхід на основі багатоетапного ресамплінгу вихідного зображення.
3. Підхід на основі багатоетапного симетричного ресамплінгу вихідного зображення з інтеграцією симетричних наборів вагових коефіцієнтів.

При побудові декодеру на базі першого підходу особливості його функціонування полягають в тому, що отриманий в результаті кодування вхідного зображення багатовимірний масив ознак за допомогою класичних процедур масштабування зображень, видозмінюється до масиву, призначеного для опису зображення з виділеним об'єктом. Однією із таких процедур є процедура білінійної інтерполяції [1, 6, 14].

Підхід на основі багатоетапного ресамплінгу вихідного зображення полягає у тому, що отриманий на виході багатовимірний масив ознак видозмінюється до масиву, призначеного для опису зображення з виділеним об'єктом, за ЗНМ застосування модулів виділення ознак ЗНМ. Слід зазначити, що в загальному випадку тип та параметри ЗНМ, що застосована для побудови декодера, можуть не співпадати з типом та параметрами ЗНМ, що використана для побудови кодеру. Таким чином, НММ, що використовується для виділення об'єкта на зображенні при використанні декодера на основі багатоетапного ресамплінгу може бути несиметричною. Підхід на основі багатоетапного симетричного ресамплінгу передбачає, що для розробки кодеру і декодеру використовується однакові ЗНМ з дзеркально відображеною структурою. Тобто структурні ЗНМ кодера в зворотному порядку відповідають структурі ЗНМ декодера, однак В декодері шари масштабування не використовуються. При цьому відповідні шари згортки кодера та декодера зв'язані між собою. Подібно до кодеру, ефективність декодеру залежить від типу ЗНМ, на якій він базується. Крім того, як точність, так і обчислювальна ресурсоємність декодера в значній мірі залежать від підходу, на базі якого він побудований. Таким чином, варіація побудови декодера на основі одного із трьох наведених підходів та можливості застосування в ньому різних типів ЗНМ потенційно дозволяє адаптувати декодер до умов поставленої задачі виділення об'єктів:

$$\begin{cases} A(S, \|b\|, \|m\|, C, H, \|k\|, T) \rightarrow max \\ \Theta(S, \|b\|, \|m\|, C, H, \|k\|, T) \leq \Delta_{\theta} \end{cases}, \quad (14)$$

де T – тип підходу до побудови декодеру.

При побудові декодера на основі багатоетапного ресамплінгу або багатоетапного симетричного ресамплінгу визначити приблизні значення використаних конструктивних параметрів можливо, базуючись на принципах, наведених в описі моделі кодеру.

Побудова загальної схеми процедури. Згідно з результатами аналізу технологій нейромережевого виділення об'єктів та розробленими моделями кодера та декодера, для опису особливостей процедури семантичної сегментації зображень використано вираз:

$$\langle \{u_{req}\}, \{u_{con}\}, \{NN_d\}, \{u_{CNN}\}, \{d\}, \{\alpha\}, \{w\}, \rangle \rightarrow \langle CNN_{type}^{enc}, CNN_{type}^{dec}, \{CNN^{enc}\}, \{CNN^{dec}\} \rangle, \quad (15)$$

де $\{u_{req}\}$ – множина параметрів реєстрації; $\{u_{con}\}$ – множина вимог до результатів розпізнавання; $\{u_{obj}\}$ – множина, що містить опис виділених об'єктів; $\{NN_d\}$ - множина

доступних типів ЗНМ; $\{u_{cnn}\}$ – множина параметрів доступних типів ЗНМ; $\{d\}$ – множина експертних даних, що можуть бути використані для побудови моделі кодеру та декодеру; $\{\alpha\}$ - коефіцієнти параметрів ефективності, що використовуються в виразі (6); $\{w\}$ – множина параметрів ефективності, що використовуються в виразі (6) та наведені в табл. 1; $CNN_{type}^{enc}, CNN_{type}^{dec}$ – тип ЗНМ, що використовуються як базис побудови кодеру та декодеру; $\{CNN^{enc}\}, \{CNN^{dec}\}$ – параметри ЗНМ типу $CNN_{type}^{enc}, CNN_{type}^{dec}$, відповідно.

Зазначимо, що $\{u_{req}\}$ в значній мірі залежить від формату реєстрації (g_{rf}), роздільної здатності (g_{rs}) та розміру вхідного зображення (g_{sz}). Крім того, базуючись на результатах [11, 17], визначено необхідність застосувати в процедурі блоків, пов'язаних з формуванням навчальної вибірки та оцінкою точності виділення. Таким чином, перелік основних завдань процедури складається з:

- Визначення переліку параметрів зображень, що підлягають реєстрації.
- Формування параметрів навчальних прикладів для НММ.
- Визначення типу ЗНМ, що є найбільш ефективним в умовах поставленої задачі.
- Формування навчальної вибірки НММ, обсяг якої повинен бути достатнім.
- Обґрунтування параметрів, що мають бути використані для оцінювання точності.
- Розрахунок значень конструктивних параметрів ЗНМ визначеного типу.
- Оцінка точності виділення та за необхідності уточнення архітектури НММ.

Схема розробленої процедури показана на рис.2.

Рис. 2. Схема процедури застосування нейронних мереж для сегментації зображень

Комп'ютерні експерименти. З метою верифікації запропонованої процедури проведені комп'ютерні експерименти, що були реалізовані за допомогою спеціально розробленого програмного забезпечення, базою якого послужили запропоновані моделі кодера та декодера. Програмне забезпечення написане на мові програмування Python з

використанням бібліотеки TensorFlow. В процесі комп'ютерних експериментів досліджено ефективність застосування запропонованої процедури для семантичної сегментації кольорових зображень для виділення на них об'єкту типу автомобіль.

Як джерело формування навчальної вибірки нейронної мережі використано базу даних Carvana Image Masking Challenge, що являється вільно доступною за посиланням https://www.kaggle.com/c/carvana-image-masking-challenge. В цій базі даних містяться 5088 висококонтрастних зображень, записаних у форматі RGB. Розмір окремого зображення становить 256×256 пікселів. Враховуючи розмір і кольоровий формат оригінальних та сегментованих зображень, визначено доцільність подачі на вхід кодеру необробленого зображення. Відповідно множина зареєстрованих параметрів відповідає трьохканальному зображенню розміром 256x256 пікселів. При визначенні найбільш ефективного типу ЗНМ з урахуванням можливостей інструментального забезпечення визначено, що до складу $\{NN_d\}$ доцільно віднести ЗНМ типу LeNet, VGG, AlexNet, GoogLeNet та ResNet. Значення функції ефективності для цих типів ЗНМ при їх застосуванні в якості бази кодеру та декодеру наведено в табл. 2. Зазначимо, що розглянута можливість побудови декодеру на основі одноетапного, багатоетапного та багатоетапного симетричного ресамплінгу. В табл. 2 багатоетапному ресамплінгу відповідає колонка E_{dec1}, а багатоетапному симетричному ресамплінгу – колонка E_{dec2}. Для одноетапного ресамплінгу E_{dec}=0,75.

Таблиця 2

Значення функції ефективності для допустимих типів ЗНМ

Тип ЗНМ	E_{enc}	E_{dec1}	E_{dec2}
LeNet	0,55	0,65	0,7
VGG	0,75	0,7	0,75
AlexNet	0,5	0,4	0,5
GoogLeNet	0,6	0,5	0,55
ResNet	0,6	0,5	0,55

Значення функції ефективності для кожного із означених типів визначено шляхом експертного оцінювання з урахуванням виразу (6). З використанням виразу (5) визначено, що найбільш ефективним являється кодер на основі ЗНМ типу VGG, а найбільш ефективним типом декодеру – декодер на основі одноетапного ресамплінгу. Вхідні та вихідні параметри навчальних прикладів НММ відповідають оригінальному та сегментованому зображенню, показаному на рис. 6. Конструктивні параметри VGG були адаптовані до аналізу напівтонових зображень розміром 256x256 за рахунок збільшення вхідне поле ЗНМ. Навчальна вибірка сформована на основі описаної БД зображень та розділена на тренувальну, валідаційну та тестову вибірки. Навчальна вибірка NM доповнена шляхом аугментації навчальних прикладів. З них 8000 прикладів використано в якості тренувальних даних, а 1088 прикладів – для валідації та тестування. Відповідно до рекомендацій [5, 6, 13], обсяг тренувальної вибірки 8000, а загальний обсяг валідаційної та тестової вибірки 1680 прикладів. Для оцінки точності НММ використано коефіцієнт Жаккара. Графіки точності виділення автомобіля, що реалізоване за допомогою побудованої НММ, показано на рис. 3. Для порівняння проведені експерименти, пов'язані з виділенням маски автомобіля за допомогою НММ з іншими архітектурними параметрами. Так, на рис. 4 та на рис. 5 показані графіки точності виділення об'єктів за допомогою НММ, у якої блоки кодеру та декодеру побудовані на базі VGG.

ISSN 2663 - 4023

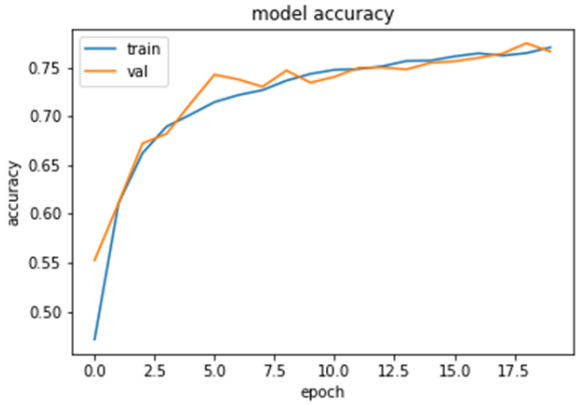

Рис. 3. Графіки точності сегментації допомогою HMM з кодером на основі VGG та декодером на основі одноетапного ресамплінгу

При цьому графіки точності, показані на рис. 4, відповідають декодеру з багатоетапним ресамплінгом, а графіки точності, показані на рис. 5 – декодеру з багатоетапним симетричним ресамплінгом.

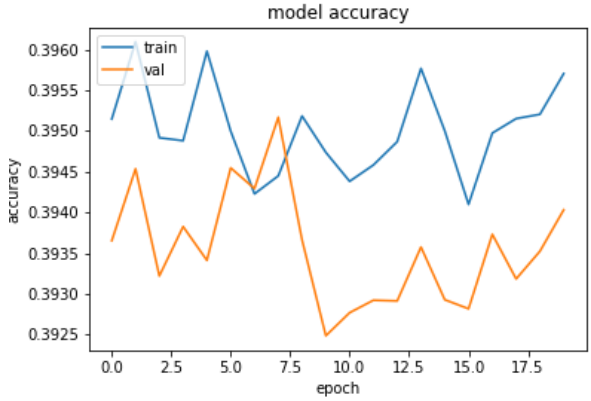

Рис. 4. Графіки точності сегментації за допомогою HMM з кодером на основі VGG та декодером з багатоетапним ресамплінгом

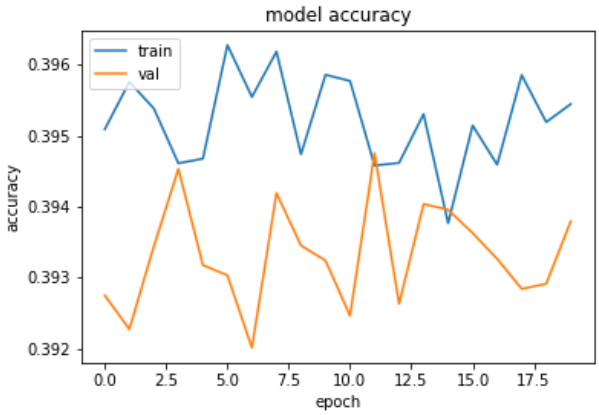

Рис. 5. Графіки точності сегментації за допомогою HMM з кодером на основі VGG та декодером з багатоетапним симетричним ресамплінгом

ISSN 2663 - 4023

Аналіз графіків, показаних на рис. 3-5, дозволяє стверджувати, що при використанні запропонованої НММ точність сегментації становить приблизно 0,8, що відповідно до результатів проведених експериментів, більше ніж в 2 рази перевищує точність виділення, яка може бути досягнута за допомогою інших НММ. Подальше підвищення точності, що може бути реалізоване за рахунок модифікації параметрів VGG, потребує проведення додаткових теоретичних досліджень. Крім того, підвищити точність виділення можливо за рахунок застосування при побудові кодеру та декодеру найбільш сучасних типів ЗНМ. Для ілюстрації використання розробленої процедури проведено виділення об'єкту (автомобілю) на зображеннях, що не були використані в процесі навчання НММ. Результат такого виділення показано на рис. 6.

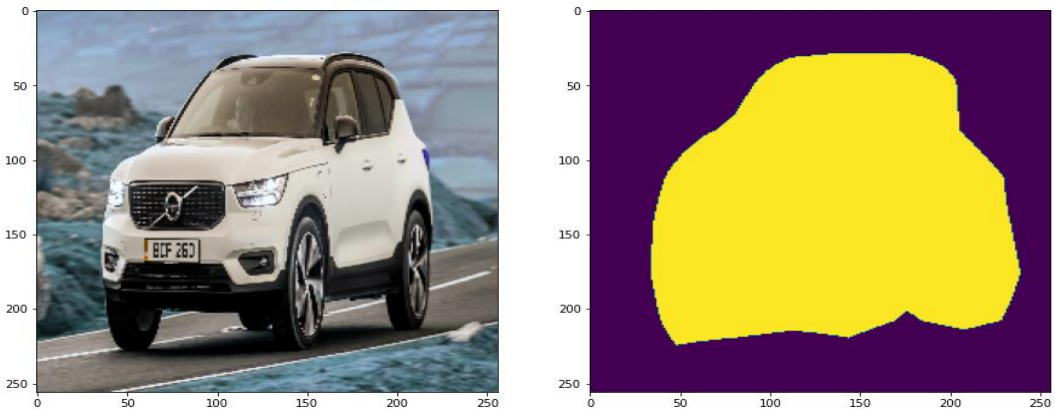

Рис. 6. Приклад семантичної сегментації

Отже результати експериментальних експериментів свідчать, що застосування запропонованої процедури дозволяє побудувати НММ, яка при достатньо короткому терміні навчання забезпечує досягнення точності сегментації зображень близько 0,8. При цьому вдалось уникнути проведення складних експериментів, спрямованих на визначення ефективної архітектури НММ, що надає можливість мінімізувати обсяг обчислювальних ресурсів, спрямованих на побудову НММ. Таким чином, запропонована процедура дозволяє при зменшенні обчислювальних ресурсів пов'язаних з побудовою НММ досягти точності сегментації, що співвідноситься з точністю кращих відомих систем аналогічного призначення [2, 9, 12].

ВИСНОВКИ ТА ПЕРСПЕКТИВИ ПОДАЛЬШИХ ДОСЛІДЖЕНЬ

На підставі літературного аналізу сучасних робіт у галузі семантичної сегментації зображень показано, що найбільш перспективним напрямом їх удосконалення є використання нейромережевих технологій, ефективність яких багато в чому визначається рівнем їх адаптованості до умов поставленого завдання. В результаті проведених досліджень розроблено процедуру застосування нейронних мереж для сегментації растрових зображень, що передбачає адаптацію нейромережевого кодера та нейромережевого декодера до розміру зображення, кількості колірних каналів, допустимої мінімальної точність сегментації, допустимої максимальної обчислювальної складності реалізації процесу сегментації, необхідності маркування сегментів

зображення, необхідності виділення декількох сегментів, що відповідають різним об’єктам, необхідності виділення декількох сегментів, що відповідають об’єктам, які частково перекриваються між собою, необхідності виділення деформованих об’єктів, необхідності виділення зсунутих об’єктів, необхідності виділення повернутих об’єктів та допустимого терміну навчання нейромережевої моделі. Розроблена процедура верифікована експериментально на прикладах семантичної сегментації зображень, що містять об'єкти типу автомобіль. Отримані експериментальні результати свідчать, що застосування запропонованої процедури дозволяє уникнувши складних довготривалих експериментів, побудувати нейромережеву модель, яка при достатньо короткому терміні навчання забезпечує досягнення точності сегментації зображень близько 0,8, що відповідає найкращим системам аналогічного призначення. Шляхи подальших досліджень у напрямку вдосконалення методологічного забезпечення нейромережевої сегментації растрових зображень доцільно співвіднести з обґрунтованим використанням у кодері та декодері сучасних модулів та механізмів, адаптованих до значимих умов поставленої задачі. Наприклад, введення модулю ResNet дозволяє за рахунок нівелювання ефекту падіння градієнта збільшити глибину нейронної мережі, а введення модуля Inception забезпечує зменшення кількості вагових коефіцієнтів та обробку об'єктів різного розміру.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1 Abraham, J., Paul, V. (2019). An imperceptible spatial domain color image watermarking scheme. *Journal of King Saud University - Computer and Information Sciences*, *31*(1), 125–133. https://doi.org/10.1016/j.jksuci.2016.12.004.

2 Adithya, U., Nagaraju, C. (2021). Object Motion Direction Detection and Tracking for Automatic Video Surveillance. *International Journal of Education and Management Engineering (IJEME), 11*(2), 32-39. https://doi.org/10.5815/ijeme.2021.02.04.

3 Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. http://arxiv.org/abs/1511.0051.

4 Dmitry, A. (2018). Segmentation Object Strategy on Digital Image. *Journal of Siberian Federal University. Engineering & Technologies*, *11*(2), 213-220.

5 Cherrat, E. M., Alaoui, R., Bouzahir, H. (2020). Score fusion of finger vein and face for human recognition based on convolutional neural network model. *International Journal of Computing*, 11–19. https://doi.org/10.47839/ijc.19.1.1688.

6 Hu, Z., Tereykovskiy, I., Zorin, Y., Tereykovska, L., Zhibek, A. (2019). Optimization of convolutional neural network structure for biometric authentication by face geometry. *Advances in Intelligent Systems and Computing, 754*, 567-577.

7 Shen, J. (2014). Motion detection in color image sequence and shadow elimination. *Visual Communications and Image Processing, 5308*, 731-740.

8 Kong, T., et al. (2020). FoveaBox: Beyound Anchor-Based Object Detection. *IEEE Trans. Image Process, 29*, 7389–7398.

9 Liu, X.-P., Li, G., Liu, L., Wang, Z. (2019). Improved YOLOV3 target recognition algorithm based on adaptive eged optimization. *Microelectron. Comput, 36*, 59–64.

10 Prilianti, K. R et al. (2021). Non-destructive Photosynthetic Pigments Prediction using Multispectral Imagery and 2D-3HM. *International Journal of Computing, 20*(3), 391-399.

11 Reja, S. A., Rahman, M. M. (2021). Sports Recognition using Convolutional Neural Network with Optimization Techniques from Images and Live Streams. *International Journal of Computing, 20*(2), 276-285.

12 Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI), 9351*, 234-241.

13 Senocak, A., Oh, T.-H., Kim, J., Kweon, I. S. (2018). Part-Based Player Identification Using Deep Convolutional Representation and Multi-scale Pooling. У *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE. https://doi.org/10.1109/cvprw.2018.00225.

14 Shkurat, O. et al. (2020). Image Segmentation Method Based on Statistical Parameters of Homogeneous Data Set. *Advances in Intelligent Systems and Computing*, *902*, 271–281.

15 Simonyan, K., Zisserman, A. (2019). Very deep convolutional networks for large-scale image recognition. ArXiv1409.1556 Cs. http://arxiv.org/abs/1409.1556.

16 Tereikovskyi, O. (2022). The method of neural network selection of objects on raster images: master's thesis.: 123 Computer Engineering. Kyiv, 104 p.

17 Zhang, S. et al. (2018). Single-Shot Refinement Neural Network for Object Detection. ArXiv 1711.06897 Cs. http://arxiv.org/abs/ 1711.06897.

КІБЕРБЕЗПЕКА: освіта, наука, техніка
CYBERSECURITY: EDUCATION, SCIENCE, TECHNIQUE
№ 2 (18), 2022
ISSN 2663 - 4023

Ihor A. Tereikovskyi
Doctor of Technical Sciences, Professor, Professor of the Department of System Programming and Specialized Computer Systems, National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine
ORCID ID: 0000-0003-4621-9668
tereikowski@ukr.net

Denys O. Chernyshev
Doctor of Technical Sciences, Professor, First vice-rector, Kyiv National University of Construction and Architecture, Kyiv, Ukraine
ORCID ID: 0000-0002-1946-9242
taqm@ukr.net

Oleksandr G. Korchenko
Doctor of Technical Sciences, Professor, Head of the Department of Information Technology Security, National Aviation University, Kyiv, Ukraine
ORCID ID: 0000-0003-3376-0631
agkorchenko@gmail.com

Liudmyla A. Tereikovska
Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Information Technologies of Design and Applied Mathematics, Kyiv National University of Construction and Architecture, Kyiv, Ukraine
ORCID ID: 0000-0002-8830-0790
tereikovskal@ukr.net

Oleh I. Tereikovskyi
Postgraduate student of the Department of Information Technology Security, National Aviation University, Kyiv, Ukraine
ORCID ID: 0000-0001-5045-0163
tereikovskyio@gmail.com

PROCEDURE FOR USING NEURAL NETWORKS FOR SEGMENTATION OF RASTER IMAGES

Abstract. Currently, means of semantic segmentation of images, based on the use of neural networks, are increasingly used in computer systems for various purposes. Despite significant successes in this field, one of the most important unsolved problems is the task of determining the type and parameters of convolutional neural networks, which are the basis of the encoder and decoder. As a result of the research, an appropriate procedure was developed that allows the neural network encoder and decoder to be adapted to the following conditions of the segmentation problem: image size, number of color channels, permissible minimum accuracy of segmentation, permissible maximum computational complexity of segmentation, the need to label segments, the need to select several segments, the need to select deformed, displaced and rotated objects, the maximum computational complexity of learning a neural network model is permissible; admissible training period of the neural network model. The implementation of the procedure of applying neural networks for image segmentation consists in the formation of the basic mathematical support, the construction of the main blocks and the general scheme of the procedure. The developed procedure was verified experimentally on examples of semantic segmentation of images containing objects such as a car. The obtained experimental results show that the application of the proposed procedure allows, avoiding complex long-term experiments, to build a neural network model that, with a sufficiently short training period, ensures the achievement of image segmentation accuracy of about 0.8, which corresponds to the best systems of a similar purpose. It is shown that the ways of further research in the direction of improving the methodological support of neural network segmentation of raster images should be correlated with the justified use of modern modules and mechanisms in the encoder and decoder, adapted to the significant conditions of the given task. For example, the use of the ResNet module allows you to increase the depth of the neural network due to the leveling

of the gradient drop effect, and the Inception module provides a reduction in the number of weighting factors and the processing of objects of different sizes.

Keywords: neural network model, convolutional neural network; semantic segmentation; raster image; image recognition.

REFERENCES (TRANSLATED AND TRANSLITERATED)

1 Abraham, J., Paul, V. (2019). An imperceptible spatial domain color image watermarking scheme. *Journal of King Saud University - Computer and Information Sciences*, *31*(1), 125–133. https://doi.org/10.1016/j.jksuci.2016.12.004.

2 Adithya, U., Nagaraju, C. (2021). Object Motion Direction Detection and Tracking for Automatic Video Surveillance. *International Journal of Education and Management Engineering (IJEME), 11*(2), 32-39. https://doi.org/10.5815/ijeme.2021.02.04.

3 Badrinarayanan, V., Kendall, A., Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. http://arxiv.org/abs/1511.0051.

4 Dmitry, A. (2018). Segmentation Object Strategy on Digital Image. *Journal of Siberian Federal University. Engineering & Technologies, 11*(2), 213-220.

5 Cherrat, E. M., Alaoui, R., Bouzahir, H. (2020). Score fusion of finger vein and face for human recognition based on convolutional neural network model. *International Journal of Computing*, 11–19. https://doi.org/10.47839/ijc.19.1.1688.

6 Hu, Z., Tereykovskiy, I., Zorin, Y., Tereykovska, L., Zhibek, A. (2019). Optimization of convolutional neural network structure for biometric authentication by face geometry. *Advances in Intelligent Systems and Computing, 754*, 567-577.

7 Shen, J. (2014). Motion detection in color image sequence and shadow elimination. *Visual Communications and Image Processing, 5308*, 731-740.

8 Kong, T., et al. (2020). FoveaBox: Beyound Anchor-Based Object Detection. *IEEE Trans. Image Process, 29*, 7389–7398.

9 Liu, X.-P., Li, G., Liu, L., Wang, Z. (2019). Improved YOLOV3 target recognition algorithm based on adaptive eged optimization. *Microelectron. Comput, 36*, 59–64.

10 Prilianti, K. R et al. (2021). Non-destructive Photosynthetic Pigments Prediction using Multispectral Imagery and 2D-3HM. *International Journal of Computing, 20*(3), 391-399.

11 Reja, S. A., Rahman, M. M. (2021). Sports Recognition using Convolutional Neural Network with Optimization Techniques from Images and Live Streams. *International Journal of Computing, 20*(2), 276-285.

12 Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI), 9351*, 234-241.

13 Senocak, A., Oh, T.-H., Kim, J., Kweon, I. S. (2018). Part-Based Player Identification Using Deep Convolutional Representation and Multi-scale Pooling. У *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. IEEE. https://doi.org/10.1109/cvprw.2018.00225.

14 Shkurat, O. et al. (2020). Image Segmentation Method Based on Statistical Parameters of Homogeneous Data Set. *Advances in Intelligent Systems and Computing, 902*, 271–281.

15 Simonyan, K., Zisserman, A. (2019). Very deep convolutional networks for large-scale image recognition. ArXiv1409.1556 Cs. http://arxiv.org/abs/1409.1556.

16 Tereikovskyi, O. (2022). The method of neural network selection of objects on raster images: master's thesis.: 123 Computer Engineering. Kyiv, 104 p.

17 Zhang, S. et al. (2018). Single-Shot Refinement Neural Network for Object Detection. ArXiv 1711.06897 Cs. http://arxiv.org/abs/ 1711.06897.

(cc) BY-NC-SA

This work is licensed under Creative Commons Attribution-noncommercial-sharealike 4.0 International License.