

**В. В. МОСКАЛЕНКО, М. О. ЗАРЕЦЬКИЙ, Я. Ю. КОВАЛЬСЬКИЙ,  
С. С. МАРТИНЕНКО**

*Сумський державний університет, Україна*

## **МОДЕЛЬ І МЕТОД НАВЧАННЯ КЛАСИФІКАТОРА КОНТЕКСТІВ СПОСТЕРЕЖЕННЯ НА ЗОБРАЖЕННЯХ ВІДЕОІНСПЕКЦІЇ СТІЧНИХ ТРУБ**

*Відеоінспекція часто використовується для діагностики дефектів стічних труб. Для коректного кодування знайдених дефектів згідно існуючих стандартів необхідно враховувати багато контекстної інформації щодо орієнтації та розташування камери в трубі. Запропоновано модель класифікації контекстів спостереження на кадрах відеоінспекції стічних труб та п'ятиетапний метод машинного навчання. Основна ідея запропонованого підходу полягає у поєднанні методів глибокого машинного навчання з принципами максимізації інформації та кодування самокоректуючими кодами Хеммінга. Запропонована модель складається з глибокої згорткової нейронної мережі із сигмоїдним та округленим вихідними шарами та інформаційно-екстремальних вирішувальних правил. Перші етапи методу полягають в аугментації даних та навчанні екстрактора ознак в складі сімської моделі з softmax триплетною функцією втрат. Наступні етапи пов'язані з обчислення двійкового коду для кожного класу розпізнавання, що використовується як мітка під час навчання з бінарною крос-ентропійною функцією втрат з метою підвищення компактності розподілу спостережень кожного класу в двійковому просторі Хеммінга. На останньому етапі методу навчання передбачається для кожного класу здійснювати оптимізацію параметрів радіально-базисних вирішувальних правил в просторі Хеммінга на основі інформаційного критерію. Інформаційний критерій, виражений як логарифмічна функція точнісних характеристик вирішувальних правил, забезпечує максимальну узагальнюючу здатність та достовірність моделі за найскладніших у статистичному розумінні умов. Перевірка ефективності даного підходу здійснювалась на даних, наданих компаніями Ace Pipe Cleaning (Канзас Сіті, США) та MPWiK (Вроцлав, Польща) шляхом порівняння результатів навчання за запропованою та традиційною моделями та схемами навчання. Отримана модель класифікатора кадрів зображення забезпечує прийнятну для практичного використання точність класифікації на тестовій вибірці, що становить 96,8 % і перевищує результат традиційної схеми навчання з softmax вихідним шаром на 6,8 %.*

**Ключові слова:** інспекція стічних труб; згорткові нейронні мережі; сімські мережі; інформаційно-екстремальне навчання; класифікація.

### **Вступ**

Каналізаційні труби є важливими об'єктами інфраструктури, які потребують частого моніторингу. Звичайний метод аналізу стану каналізаційних труб передбачає огляд цих труб за допомогою відеоспостереження на предмет дефектів усередині труби. Виявлені дефекти кодуються в рамках одного зі стандартів, серед яких найбільшого поширення набули британський MSCC5 та американський PACP6 або PACP7 стандарти [1]. При цьому для коректного кодування необхідно мати інформацію про локалізацію, орієнтацію, форму, тяжкість та близькість дефекту до вхідних-вихідних терміналів і відгалужень. Однак для екстракції цієї інформації необхідно мати дані про орієнтацію і розміщення камери спостереження в трубі, однак ці дані в явному вигляді, як правило, недоступні. Тому розпізнавання контексту

спостереження є актуальною задачею.

Формування звіту щодо стану стічних труб відповідно до стандартів вимагає ретельного вивчення та детального аналізу зібраних відеоматеріалів інспекції. Застосування методів комп'ютерного зору та машинного навчання для аналізу відеоспостережень дозволяє підвищити продуктивність та зменшити витрати [1].

Оцінювання орієнтації і позиції камери може бути реалізовано шляхом використання методів візуальної одометрії або одночасного картографування та локалізації (SLAM) [2]. Проте відео матеріал інспекції містить накладені надписи дистанції, вимірної механічним одометром. Також для кодування дефектів не існує потреби в дуже точних даних про кутову орієнтацію та зміщення камери відносно центру труби. Тому оцінювання орієнтації камери достатньо здійснювати на основі класифікатора зображень. При цьому в алфавітні класів можуть бути включені різноманітні

нестандартні ситуації, які необхідно правильно обробляти.

У перших алгоритмах класифікаційного аналізу візуальних спостережень під час інспекції стічних труб для ознакового опису використовувалися методи виявлення країв та обробки контурів [3]. Цей метод вимагає ручного налаштування гіперпараметрів та ігнорує велику кількість контекстної інформації. Застосування фільтру Габор для екстракції візуальних ознак під час аналізу даних інспекції каналізаційних труб забезпечує більш гнучкий та теоретично обґрунтований підхід [4]. Однак неглибокі моделі цього типу мають низьку інформаційну ємність і не підходять для ефективного аналізу контекстів за умов складних конфігурацій дефектів. Крім того, вирішення задачі автоматичного аналізу даних відеоспостереження ускладнено наявністю великої кількості артефактів та шуму. Погана видимість під час огляду може виникнути з різних причин, таких як пар, густий газ, зміна рівня води чи положення джерела світла, розмитість внаслідок руху та низька роздільна здатність камери.

Прогресу в галузі аналізу візуальних даних досягнуто в рамках парадигми глибокого машинного навчання, що основана на використанні ієрархічного ознакового опису. Ієрархічні екстрактори ознак вирізняються більш високою інформаційною ємністю порівняно з однорівневими моделями [5, 6]. Проте деякі контексти спостереження чи їх частини з'являються рідко і можуть мати значну внутрішньокласову варіативність, що призводить до дисбалансу та малої кількості розмічених зразків, що відповідають складним і нерегулярним ситуаціям. Це накладає обмеження на застосування глибоких моделей, які є чутливими до обсягу та збалансованості розмічених навчальних даних.

Використання ідей та методів теорії інформації та синтезу вирішальних правил в рамках геометричного підходу знижує вимоги до кількості розмічених даних для навчання і робить його більш робастним. Сіамські нейронні мережі є прикладом такого підходу [7]. Сіамські нейронні мережі демонструють багатообіцяючі результати в умовах обмеженого обсягу навчальних даних і часто є основою для реалізації так званого навчання з першого погляду (few-shot learning). Коди, що виправляють помилки (error-correcting output codes), та інформаційно-екстремальні методи навчання поєднують принципи теорії інформації та геометричного підходу [8, 9]. Ці алгоритми найбільш корисні під час синтезу вирішальних правил з оптимальним словником ознак або для підвищення ефективності існуючих моделей. Таким

чином, найбільш перспективним шляхом подальшого вдосконалення моделей аналізу даних відеоінспекції труб є поєднання ідей та методів сіамських нейронних мереж та інформаційно-екстремального навчання.

**Мета статті** – розробити модель та метод машинного навчання на основі ідей і алгоритмів глибоких згорткових мереж та інформаційно-екстремальних вирішувальних правил задля підвищення ефективності візуальної класифікації контекстів спостереження під час інспекції стічних труб в умовах обмеженого обсягу розмічених даних.

## Модель та метод навчання

Для аналізу контекстів використовуються зображення (снєпшоти) приведені до розміру 160x160 пікселів. Як екстрактор ознак пропонується використати згорткову нейронну мережу MobileNet. При цьому коефіцієнт ємності такої мережі встановлюється на рівні 0,25, і використовується лише згорткова частина цієї моделі, без вихідного шару [10]. Структура моделі класифікатора показано на рис. 1.

MobileNet екстрактор ознак (backbone)
Шар Global Average Pooling 2D
Шар Dropout (rate=0.5)
Повноз'язний шар, Dense (128 вузлів)
Сігмоїдний шар, Sigmoid
Округлюючий шар
Радіально-базисний шар

Рис. 1. Структура моделі класифікатора

Шар Global Average Pooling 2D Layer використовується для зменшення розмірності, а Dropout-шар з 50% ймовірністю вимикання входів застосовується для регуляризації [5,9]. Повноз'язний шар з сигмоїдальною активацією формує вихідний шар екстрактора ознак.

Окрім того в моделі присутній шар з округленням (огрубленням) значень ознак для формування бінарного виходу. Після цього йде застосування радіально-базисної функції для визначення належності вхідного вектора до класу. Класи розмежовані гіперсферичними контейнерами в просторі Хеммінга. Кожен контейнер визначається базисним вектором (центром) а також радіусом, вираженим у відстані Хеммінга. В нашому випадку радіально-базисна функція  $\mu_z(b)$ , яка виражає

належність об'єкту до класу, для  $N$ -вимірного бінарного вектора  $b$  має вигляд

$$\mu_z(b) = 1 - \sum_{i=1}^N b_i \oplus b_{z,i}^* / d_z^* \quad (1)$$

де  $b_z^*$  – центр оптимального контейнера для класу  $X_z^0$ ;  $d_z^*$  – радіус оптимального контейнера для класу  $X_z^0$  в просторі Хеммінга. Запропонований метод складається з 5 основних етапів (фаз) (рис. 2). Перший етап включає в себе різноманітні операції з даними, такі як повороти, шкалювання, зміну яскравості і додавання шуму. Наступні 3 етапи – це безпосередньо навчання класифікатора. На останньому етапі вирішальні правила коригуються враховуючи дисперсію спостережень всередині класів в двійковому просторі Хеммінга.

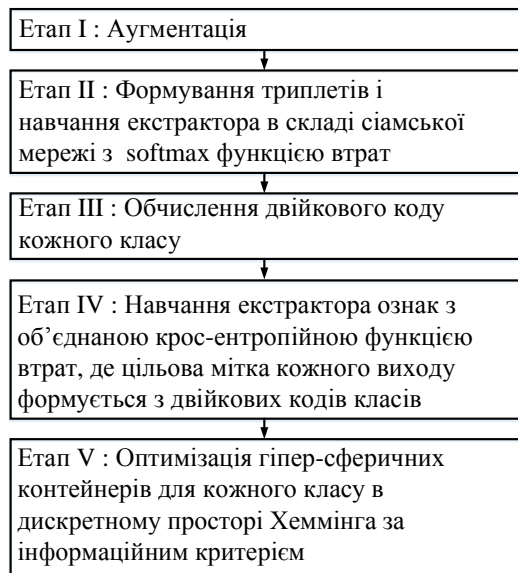


Рис. 2. Етапи запропонованого методу навчання

Другий етап навчання моделі аналізу зображення проводиться із застосуванням адаптивних алгоритмів зворотного поширення помилки, з яких найбільш популярним є Adam [5,6]. На вхід моделі подається міні-пакет, в якому  $M$  зображень кожного класу. Триплетна функція втрат розраховується за формулою

$$L = -\log \frac{\exp(\|f(x_a) - f(x_{ep})\|)}{\exp(\|f(x_a) - f(x_{ep})\|) + \exp(\|f(x_a) - f(x_{shn})\|)} \quad (2)$$

де  $f(x)$  – функція, що описує екстрактор та встановлює залежність між вхідним зображенням та вектором виходу сигмоїдного шару;

$x_a$  – зображення, що обрано випадковим чином з міні-пакету;  $x_{ep}$  – найближчий сусід з міні-батчу, що належить до того ж самого класу, тобто

$$x_{ep} = \arg \min_{x:C(x)=C(x_a)} \|f(x_a) - f(x)\|, \quad (3)$$

де  $C(x)$  – функція що повертає клас зображення;  $x_{shn}$  – зразок зображення з міні-пакету, що є найближчим серед зразків протилежних класів, проте знаходиться далі ніж зразок, тобто

$$x_{shn} = \arg \min_{\substack{C(x) \neq C(x_a) \\ \|f(x_a) - f(x)\| > \|f(x_a) - f(x_p)\|}} \|f(x_a) - f(x)\|. \quad (4)$$

Наступна фаза потрібна для перетворення в бінарну форму вихідного вектора відповідно до принципів error-correcting output codes, однак з урахуванням внутрішньої структури класів та відношень між зразками різних класів. Для цього навчальна вибірка  $\{x_{z,s} \mid z = \overline{1, Z}, s = \overline{1, n_z}\}$ , що містить  $n_z$  зразків  $z$ -го класу кодується дискретним поданням  $\{b_{z,s,i} \mid z = \overline{1, Z}, s = \overline{1, n_z}, i = \overline{1, N}\}$  з розмірністю  $N$  та обчислюється еталонний двійковий вектор кожного класу. Двійкове кодування здійснюється шляхом подавання зображень з множини  $\{x_{z,s}\}$  на вхід моделі і округлення виходу сигмоїдного шару до цілого числа

$$b_{z,s,i} = \begin{cases} 1, & \text{if } f(x_{z,s,i}) > 0,5; \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

Двійковий еталонний вектор  $b_z$  для  $z$ -го класу можна визначити шляхом порозрядного порівняння частоти двійкових одиниць в  $z$ -му класі з фоновією частотою одиниць в навчальній вибірці

$$b_{z,i} = \begin{cases} 1, & \text{if } \frac{1}{n_z} \sum_{s=1}^{n_z} b_{z,s,i} > \frac{1}{Z} \sum_{c=1}^Z \frac{1}{n_c} \sum_{s=1}^{n_c} b_{c,s,i}; \\ 0, & \text{otherwise.} \end{cases} \quad (6)$$

Обчислений еталонний вектор  $z$ -го класу  $b_z$  можна використовувати як розмітку (label) зразка для подальшого навчання з використанням Joint Binary Cross Entropy Loss, яка для кожного вхідного зразка  $x$  може бути обчислена за формулою

$$L = -\sum_{i=1}^N (b_i \log f_i(x) + (1-b_i) \log(1-f_i(x))), \quad (7)$$

де  $f_i(x)$  – значення  $i$ -го виходу сигмоїдного шару для вхідного зображення  $x$ ;  $b_i$  – значення  $i$ -го розряду еталонного вектора класу, до якого належить зображення  $x$ .

Остання фаза машинного навчання пов'язана з оптимізацією радіусу контейнерів за інформаційним критерієм для врахування меж відхилення двійкового подання спостережень кожного класу від відповідних еталонних векторів

$$E_z^* = \max_{\{d\}} E_z(d), \quad (8)$$

де  $\{d\} = \{0, 1, \dots, \left(\sum_i b_{z,i} \oplus b_{c,i} - 1\right)\}$  – набір

концентричних радіусів з центром  $b_z$  розподілу даних в  $z$ -му класі, які обчислюються за правилом (6);  $E_z$  – інформаційний критерій для  $z$ -го класу, який є функцією від характеристик точності [11]

$$E_z = 1 + \frac{1}{2} \left( \frac{a_z}{a_z + D_{2,z}} \log_2 \frac{a_z}{a_z + D_{2,z}} + \frac{\beta_z}{D_{1,z} + \beta_z} \log_2 \frac{\beta_z}{D_{1,z} + \beta_z} + \frac{b_z}{D_{1,z} + b_z} \log_2 \frac{b_z}{D_{1,z} + b_z} + \frac{D_{1,z}}{D_{1,z} + b_z} \log_2 \frac{D_{1,z}}{D_{1,z} + b_z} + \frac{D_{2,z}}{a_z + D_{2,z}} \log_2 \frac{D_{2,z}}{a_z + D_{2,z}} \right), \quad (9)$$

помилка першого роду у вирішального правила для  $z$ -го класу;  $\beta_z$  – помилка другого роду;  $D_{1,z}$  – перша достовірність (чутливість);  $D_{2,z}$  – друга достовірність (специфічність).

Вихідний шар моделі має теоретико-інформаційний базис. Двійкові еталонні вектори класів використовуються по аналогії з самокоректуючими кодами Хемінга, а радіус контейнера характеризує максимальну кількість помилок, які можуть бути виправлені після отримання повідомлення по каналу зв'язку з завадами. Інформаційний критерій характеризує міру зняття невизначеності після отримання повідомлення з заданого алфавіту.

Таким чином, оптимізація радіусу гіперсферичних контейнерів дозволяє підвищити ефективність прийняття рішень за умов перетину класів з різними габаритами розподілу спостережень та виявляти

новизну в даних у випадку появи спостережень, що не потрапляють в контейнер жодного класу. При цьому критерієм оптимізації є логарифмічний інформаційний критерій, що підвищує узагальнюючу здатність моделі та забезпечує її робастність за найскладніших у статистичному розумінні випадках.

## Результати моделювання

Алфавіт класів [3] містить 9 основних контекстів спостереження (рис. 3). Клас  $X_1^0$  характеризує орієнтацію камери вперед вздовж труби. Класи  $X_2^0$ ,  $X_3^0$ ,  $X_4^0$  та  $X_5^0$  характеризують орієнтацію камери вліво, вправо, вгору та вниз відповідно, при якій спостерігач розуміє куди орієнтована камера, тобто поворот є неповним.

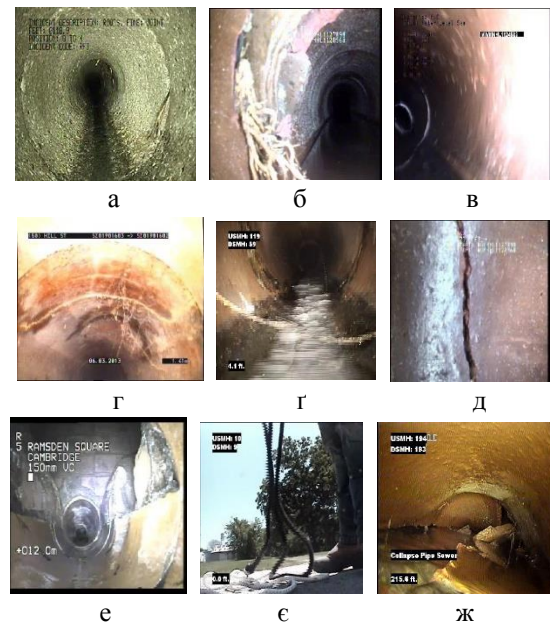


Рис. 3. Приклади зображень кожного з класів:

- а – клас  $X_1^0$ ; б – клас  $X_2^0$ ; в – клас  $X_3^0$ ;
- г – клас  $X_4^0$ ; г – клас  $X_5^0$ ; д – клас  $X_6^0$ ;
- е – клас  $X_7^0$ ; е – клас  $X_8^0$ ; ж – клас  $X_9^0$

Клас  $X_6^0$  характеризує орієнтацію камери впритул до стінки зліва, справа, знизу чи згори, коли важко зрозуміти яка саме це частина труби (верхня, нижня, ліва чи права). Клас  $X_7^0$  характеризує точку входу в трубу для інспекції, починаючи від люка до початку робочої області труби. Клас  $X_8^0$  характеризує ситуації, які потрібно ігнорувати і не обробляти. До цих ситуацій відносяться втрата видимості, кадри

відео до входу чи після виходу з труби. Клас  $X_9^0$  характеризує завал труби, що перешкоджає наступному руху вздовж неї.

Для кожного класу зібрано по 80 навчальних та тестових зразків. Кожне зображення використовується для розширення навчального набору в 4 рази шляхом застосування невеликих випадкових викривлень, до яких належать варіація яскравості, масштабу та поворотів до 10% від повного діапазону, а також накладання шумів типу сіль та перець.

Спочатку запропонований екстрактор ознак модифікується шляхом заміни сігмоїдного шару на softmax шар для реалізації класичного способу навчання з використанням categorical cross-entropy loss та оптимізатора Adam [5, 6]. Отримані результати будуть використовуватися для порівняння з запропонованим підходом.

На рис. 4 показано криві зміни точності (test\_acc та train\_acc) та функції втрат (test\_loss та train\_loss) на тестовій та навчальній вибірках залежно від кількості епох. При цьому розмір міні-батчу становить 32 зображення, а learning rate дорівнює 0,0001.

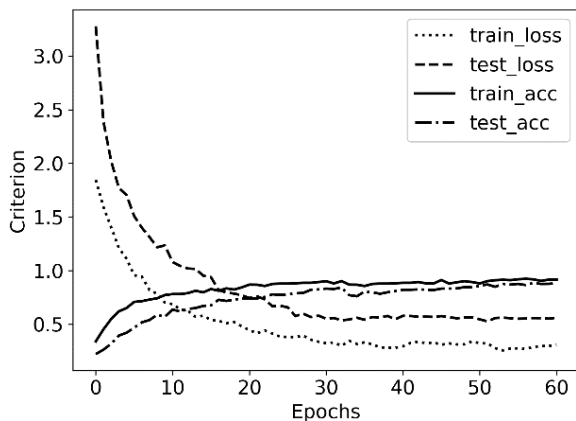


Рис. 4. Графік залежності точності та функції втрат від кількості епох навчання в рамках традиційного підходу

Аналіз рис. 4 показує, що зменшення помилки за тестовою (test\_loss) вибіркою припинилося на 31 епосі. При цьому точність отриманої моделі (test\_acc) становить 90 %.

На рис. 5 показано результат навчання екстрактора ознак з сігмоїдним вихідним шаром, де перші 30 епох навчання відбувається з використанням функції помилки (2) та наступні 30 епох з використанням функції помилки (7). При цьому кожні 5 епох будуються інформаційно-екстремальні вирішальні правила та обчислюється усереднене за алфавітом класів значення інформаційного критерію ефективності (9) за тестовою

(test\_info\_cri) та навчальною вибіркою (train\_info\_cri).

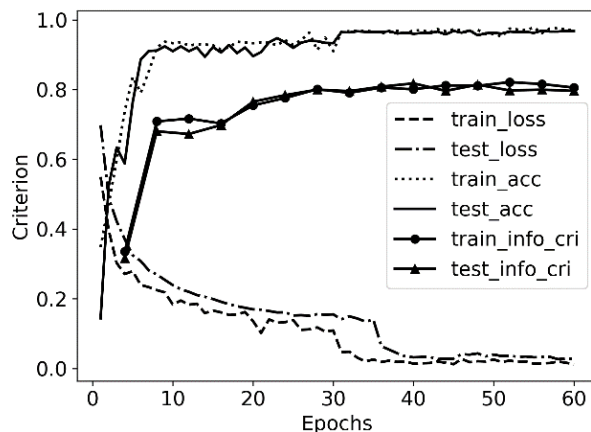


Рис. 5. Графік залежності точності, функції втрат та інформаційного критерію від кількості епох навчання в рамках запропонованого підходу

Аналіз рис. 5 показує, що починаючи з 10-ї епохи зменшення помилки за тестовою вибіркою дуже вповільнюється, а починаючи з 31-ї епохи відбувається різке підвищення точності вирішувальних правил більше ніж на 2%. Проте після 40-ї епохи навчання покращення результату практично припиняється. Для навченої моделі точність за тестовою вибіркою становить 96,8%, а значення інформаційного критерію за навчальною і тестовою вибірками мало відрізняються і рівне 0,81. При цьому середня відстань між центрами контейнерів, яка рівна 10, перевищує значення середнього радіуса контейнерів, який рівний 6,5. Це свідчить про високу завадозахищеність інформаційно-екстремальних вирішувальних правил.

Таким чином, використання багатофазної схеми навчання з інформаційно-екстремальними вирішувальними правилами дозволяє підвищити точність моделі для розпізнавання контекстів спостереження під час інспекції труб на 6,8 % порівняно з традиційною схемою навчання з softmax вихідним шаром.

## Висновки

Вперше розроблено модель візуальної класифікації контекстів спостереження для задач інспекції підземних труб на основі багатозарової згорткової мережі з інформаційно-екстремальними радіально-базисними вирішувальними правилами. Розроблено багатофазний метод навчання, що включає в себе аугментацію даних, попереднє навчання екстрактора ознак в складі згорткової сіамської мережі, обчислення двійкового коду для

кожного класу, точне налаштування екстрактора з використанням бінарної крос-ентропійної функції втрат та двійкового коду кожного класу як мітки, а також оптимізацію радіально-базисних вирішувальних правил за інформаційним критерієм.

Здійснена перевірка ефективності запропонованого підходу на даних, наданих компаніями Ace Pipe Cleaning (Канзас Сіті, США) та MPWiK (Вроцлав, Польща), шляхом порівняння результатів навчання за запропонованою та традиційною моделями та схемами навчання. Отримана модель класифікатора відео кадрів забезпечує прийнятну для практичного використання точність класифікації на тестовій вибірці, що становить 96,8% і перевищує результат традиційної схеми навчання з softmax вихідним шаром на 6,8%.

Розроблені модель та метод навчання можуть бути використані для вирішення широкого кола практичних задач класифікаційного аналізу зображень за умов обмеженого обсягу розмічених навчальних даних.

Подальші дослідження будуть спрямовані на дослідження регуляризованих складових функції втрат для підвищення ефективності формування дискретного ознакового подання спостережень відеоінспекції і інформаційно-екстремальних вирішувальних правил.

Робота виконана на базі лабораторії інтелектуальних систем кафедри комп'ютерних наук Сумського державного університету при фінансовій підтримці МОН України в рамках держбюджетної науково-дослідної роботи ДР № 0117U003934.

## Література

1. Moradi, S. Review On Computer Aided Sewer Pipeline Defect Detection And Condition Assessment [Text] / S. Moradi, T. Zayed, F. Golkho // *Infrastructures*. – 2019. – Vol. 4. – 10 p. DOI: 10.3390/infrastructures4010010.

2. Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age [Text] / C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, J. J. Leonard // *IEEE Transactions on Robotics*. – 2016 – Vol. 32, Iss. 6 – P. 1309–1332. DOI: 10.1109/TRO.2016.2624754.

3. Syahrian, N. M. Vision-Based Pipe Monitoring Robot For Crack Detection Using Canny Edge Detection Method As An Image Processing Technique [Text] / N. M. Syahrian, P. Risma, T. Dewi // *Kinetik*. – 2017. – Vol. 2. – P. 243–250. DOI: 10.22219/kinetik.v2i4.243.

4. Myrans, J. Automated Detection Of Fault Types In CCTV Sewer Surveys [Text] / J. Myrans,

R. M. Everson, Z. Kapelan // *Journal Of Hydroinformatics*. – 2018. – Vol. 21. – P. 153–163. DOI: 10.2166/hydro.2018.073.

5. Cheng, J. C. P. Automated Detection Of Sewer Pipe Defects In Closed-Circuit Television Images Using Deep Learning Techniques [Text] / J. C. P. Cheng, M. Wang // *Automation in Construction*. – 2018. – Vol. 95. – P. 155–171. DOI: 10.1016/j.autcon.2018.08.006.

6. Deep learning and image processing for automated crack detection and defect measurement in underground structures [Text] / F. Panella, J. Boehm, Y. Loo, A. Kaushik, D. Gonzalez // *ISPRS - International Archives Of The Photogrammetry, Remote Sensing And Spatial Information Sciences, XLII-2*. – Riva del Garda, 4-7 June, 2018. – P. 829–835. DOI: 10.5194/isprs-archives-xlii-2-829-2018.

7. Deepshoe: An Improved Multi-Task View-Invariant CNN For Street-To-Shop Shoe Retrieval [Text] / H. Zhan, B. Shi, L. Y. Duan, A. C. Kot // *Computer Vision And Image Understanding*. – 2019. – Vol. 180. – P. 23–33. DOI: 10.1016/j.cviu.2019.01.001.

8. Error-Correcting Output Codes In Classification Of Human Induced Pluripotent Stem Cell Colony Images [Text] / H. Joutsijoki, M. Haponen, J. Rasku, K. Aalto-Setälä, M. Juhola // *Biomed Research International*. – 2016. – Vol. 2016. – P. 1–13. DOI: 10.1155/2016/3025057.

9. The Model And Training Algorithm Of Compact Drone Autonomous Visual Navigation System [Text] / V. Moskalenko, A. Moskalenko, A. Korobov, V. Semashko // *Data*. – 2018. – Vol. 4. – P. 1–14. DOI: 10.3390/data4010004.

10. Michele, A. Mobilenet Convolutional Neural Networks And Support Vector Machines For Palmprint Recognition [Text] / A. Michele, V. Colin, D. D. Santika // *Procedia Computer Science*. – 2019. – Vol. 157. – P. 110–117. DOI: 10.1016/j.procs.2019.08.147.

11. Information Synthesis Of Adaptive System For Visual Diagnostics Of Emotional And Mental State Of A Person [Text] / A. Dovbysh, I. Shelehov, D. Prylepa, I. Golub // *Eastern-European Journal Of Enterprise Technologies*. – 2016. – Vol. 4, N. 9(82). – P. 11–17. DOI: 10.15587/1729-4061.2016.756.

## References

1. Moradi, S., Zayed, T., Golkho, F. Review On Computer Aided Sewer Pipeline Defect Detection And Condition Assessment. *Infrastructures*, 2019, vol. 4, p. 10 p. DOI: 10.3390/infrastructures4010010.

2. Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J. J. Past, Present, and Future of Simultaneous Localization And Mapping: Towards the Robust-Perception Age. *IEEE Transactions on Robotics*, 2016, vol. 32, iss. 6, pp. 1309–1332. DOI: 10.1109/TRO.2016.2624754.

3. Syahrian, N. M., Risma, P., Dewi, T. Vision-Based Pipe Monitoring Robot For Crack Detection

Using Canny Edge Detection Method As An Image Processing Technique. *Kinetik*, 2017, vol. 2, pp. 243–250. DOI: 10.22219/kinetik.v2i4.243.

4. Myrans, J., Everson, R. M., Kapelan, Z. Automated Detection Of Fault Types In CCTV Sewer Surveys. *Journal Of Hydroinformatics*, 2018, vol. 21, pp. 153–163. DOI: 10.2166/hydro.2018.073.

5. Cheng, J. C. P., Wang, M. Automated Detection Of Sewer Pipe Defects In Closed-Circuit Television Images Using Deep Learning Techniques. *Automation in Construction*, 2018, vol. 95, pp. 155–171. DOI: 10.1016/j.autcon.2018.08.006.

6. Panella, F., Boehm, J., Loo, Y., Kaushik, A., Gonzalez, D. Deep learning and image processing for automated crack detection and defect measurement in underground structures. *ISPRS - International Archives Of The Photogrammetry, Remote Sensing And Spatial Information Sciences, XLII-2*. Riva del Garda, 4-7 June, 2018, pp. 829–835. DOI: 10.5194/isprs-archives-xlii-2-829-2018.

7. Zhan, H., Shi, B., Duan, L. Y., Kot, A. C., Deepshoe: An Improved Multi-Task View-Invariant CNN For Street-To-Shop Shoe Retrieval. *Computer*

*Vision And Image Understanding*, 2019, vol. 180, pp. 23–33. DOI: 10.1016/j.cviu.2019.01.001.

8. Joutsijoki, H., Haponen, M., Rasku, J., Aalto-Setälä, K., Juhola, M. Error-Correcting Output Codes In Classification Of Human Induced Pluripotent Stem Cell Colony Images. *Biomed Research International*, 2016, vol. 2016, pp. 1–13. DOI: 10.1155/2016/3025057.

9. Moskalenko, V. Moskalenko, A., Korobov, A., Semashko, V. The Model And Training Algorithm Of Compact Drone Autonomous Visual Navigation System. *Data*, 2018, vol. 4, pp. 1–14. DOI: 10.3390/data4010004.

10. Michele, A., Colin, V., Santika, D. D. Mobilenet Convolutional Neural Networks And Support Vector Machines For Palmprint Recognition. *Procedia Computer Science*, 2019, vol. 157, pp. 110–117. DOI: 10.1016/j.procs.2019.08.147.

11. Dovbysh, A., Shelehov, I., Prylepa, D., Golub, I. Information Synthesis Of Adaptive System For Visual Diagnostics Of Emotional And Mental State Of A Person. *Eastern-European Journal Of Enterprise Technologies*, 2016, vol. 4, no. 9(82), pp. 11–17. DOI: 10.15587/1729-4061.2016.756.

Надійшла до редколегії 16.07.2020 розглянута на редколегії 15.09.2020

## МОДЕЛЬ И МЕТОД ОБУЧЕНИЯ КЛАССИФИКАТОРА КОНТЕКСТА НАБЛЮДЕНИЙ НА ИЗОБРАЖЕНИЯХ ВИДЕОИНСПЕКЦИИ СТОЧНЫХ ТРУБ

**В. В. Москаленко, Н. А. Зарецкий, Я. Ю. Ковальский, С. С. Мартыненко**

Видеоинспекция часто используется для диагностики дефектов сточных труб. Для корректного кодирования найденных дефектов согласно существующих стандартов необходимо учитывать большое количество контекстной информации об ориентации и расположении камеры в трубе. Предложено модель классификации контекстов наблюдения на кадрах видеоинспекции сточных труб и пятиэтапный метод машинного обучения. Основная идея предлагаемого подхода заключается в сочетании методов глубокого машинного обучения с принципами максимизации информации и кодирования самокорректирующимися кодами Хемминга. Предложенная модель состоит из глубокой свёрточной нейронной сети с сигмоидным и округляющим выходными слоями и информационно-экстремальных решающих правил. Первые этапы метода заключаются в аугментации данных и обучении экстрактора признаков в составе сиамской модели с softmax триплетной функцией потерь. Следующие этапы связаны с вычислением двоичного кода для каждого класса распознавания, который используется как метка во время обучения с бинарной кросс-энтропийной функцией потерь с целью повышения компактности распределения наблюдений каждого класса в двоичном пространстве Хемминга. На последнем этапе метода обучения предусматривается для каждого класса осуществлять оптимизацию параметров радиально-базисных решающих правил в пространстве Хемминга на основе информационного критерия. Информационный критерий, выраженный как логарифмическая функция точностных характеристик решающих правил, обеспечивает максимальную обобщающую способность и достоверность модели в самых сложных в статистическом смысле условиях. Проверка эффективности данного подхода осуществлялась на данных, предоставленных компаниями Ace Pipe Cleaning (Канзас Сити, США) и MPWiK (Вроцлав, Польша) путем сравнения результатов обучения по предложенной и традиционной моделями и схемами обучения. Полученная модель классификатора кадров изображения обеспечивает приемлемую для практического использования точность классификации на тестовой выборке, которая составляет 96,8 % и превышает результат традиционной схемы обучения с softmax выходным слоем на 6,8 %.

**Ключевые слова:** инспекция сточных труб; свёрточные нейронные сети; сиамские сети; информационно-экстремальное обучение; классификация.

**MODEL AND METHOD OF TRAINING THE CLASSIFIER OF OBSERVATION CONTEXT  
ON VIDEO INSPECTION IMAGES OF SEWER PIPES****V. Moskalenko, M. Zaretskyi, Y. Kovalskyi, S. Martinenko**

Video inspection is often used to diagnose sewer pipe defects. To correctly encode founded defects according to existing standards, it is necessary to consider a lot of contextual information about the orientation and location of the camera from sewer pipe video inspection. A model for the classification of context on frames during observations in the video inspection of sewer pipes and a five-stage method of machine learning is proposed. The main idea of the proposed approach is to combine the methods of deep machine learning with the principles of information maximization and coding with self-correcting Hamming codes. The proposed model consists of a deep convolutional neural network with a sigmoid layer followed by the rounding output layer and information-extreme decision rules. The first stages of the method are data augmentation and training of the feature extractor in the Siamese model with softmax triplet loss function. The next steps involve calculating a binary code for each class of recognition that is used as a label in learning with a binary cross-entropy loss function to increase the compactness of the distribution of each class's observations in the Hamming binary space. At the last stage of the training method, it is supposed to optimize the parameters of radial-basis decision rules in the Hamming space for each class according to the existing information-extreme criterion. The information criterion, expressed as a logarithmic function of the accuracy characteristics of the decision rules, provides the maximum generalization and reliability of the model under the most difficult conditions in the statistical sense. The effectiveness of this approach was tested on data provided by Ace Pipe Cleaning (Kansas City, USA) and MPWiK (Wroclaw, Poland) by comparing learning results according to the proposed and traditional models and training schemes. The obtained model of the image frame classifier provides acceptable for practical use classification accuracy on the test sample, which is 96.8 % and exceeds the result of the traditional scheme of training with the softmax output layer by 6.8 %.

**Keywords:** sewer pipe inspection; convolutional neural networks; Siamese networks; information and extreme training; classification.

**Москаленко В'ячеслав Васильович** – канд. техн. наук, доцент каф. комп'ютерних наук, Сумський державний університет, Україна.

**Зарецький Микола Олександрович** – аспірант каф. комп'ютерних наук, Сумський державний університет, Україна.

**Ковальський Ярослав Юрійович** – аспірант каф. комп'ютерних наук, Сумський державний університет, Україна.

**Мартиненко Сергій Сергійович** – канд. техн. наук, асистент кафедри комп'ютерних наук, Сумський державний університет, Україна.

**Viacheslav Moskalenko** – PhD, associate professor of Computer Sciences Department of Sumy State University, Sumy, Ukraine,  
e-mail: v.moskalenko@cs.sumdu.edu.ua, ORCID Author ID: 0000-0001-6275-9803,  
Scopus Author ID:57189099775.

**Nikolay Zaretskyi** – PhD student of Computer Sciences Department of Sumy State University, Sumy, Ukraine,  
e-mail: n.zaretskij@gmail.com, ORCID Author ID: 0000-0001-9117-5604, Scopus Author ID: 57213687285.

**Yaroslav Kovalskyi** – PhD student of Computer Sciences Department of Sumy State University, Sumy, Ukraine,  
e-mail: socratus43@gmail.com, ORCID Author ID: 0000-0002-5345-5186.

**Serhii Martynenko** – PhD, assistant professor of dept. of computer science Sumy State University, Sumy, Ukraine,  
e-mail: serg.martynenko@gmail.com, ORCID Author ID: 0000-0003-0735-9246, Scopus Author ID: 36053218200