

**В. В. МОСКАЛЕНКО, М. О. ЗАРЕЦЬКИЙ, А. С. МОСКАЛЕНКО,
А. Г. КОРОБОВ, Я. Ю. КОВАЛЬСЬКИЙ**

Сумський державний університет, Україна

БАГАТОЕТАПНИЙ МЕТОД ГЛИБИННОГО НАВЧАННЯ З ПОПЕРЕДНІМ САМОНАВЧАННЯМ ДЛЯ КЛАСИФІКАЦІЙНОГО АНАЛІЗУ ДЕФЕКТІВ СТИЧНИХ ТРУБ

Розроблено багатоетапний метод машинного навчання з попереднім самонавчанням для класифікаційного аналізу дефектів на стінках стічних труб за зображеннями відеоінспекції. Об'єктом дослідження є процес розпізнавання дефектів на стінках стічних труб. Предметом дослідження є метод машинного навчання для класифікаційного аналізу дефектів стічних труб на зображеннях відеоінспекції за умов обмеженого та незбалансованого набору розмічених навчальних даних. Запропоновано п'ятиетапний алгоритм навчання класифікатора. На першому етапі відбувається контрастне навчання з використанням екземпляр-прототипної контрастної функції втрат, де для вимірювання схожості закодованих зразків використовується нормалізована відстань Евкліда. На другому етапі розглядаються два варіанти регуляризованої функції втрат – триплетна функція NCA та контрастно-центрована функція. Регуляризуюча складова на другому етапі навчання використовується для штрафування за помилку округлення вихідного вектора ознак до дискретного виду і забезпечує реалізацію інформаційного пляшкового горла. На наступному етапі здійснюється обчислення двійкового коду кожного класу для реалізації кодів, що виправляють помилки, але з урахуванням структури класів і відношень між їх ознаками. Отриманий еталонний вектор кожного класу є цільовою розміткою зображення для навчання з використанням крос-ентропійної функції втрат. Останній етап навчання здійснює оптимізацією параметрів вирішувальних правил за інформаційним критерієм для врахування дисперсії розподілу класів в двійковому просторі Хеммінга. Для порівняння результатів навчання на різних етапах та в рамках різного підходу використовується мікро-усереднена метрика F1, що обчислюється на тестових даних. Результати, отримані на відкритому наборі даних Sewer-ML, підтверджують придатність запропонованого методу навчання до практичного використання, отримане значення F1-метрики дорівнює 0,977. Було показано, що запропонований метод забезпечує збільшення значення мікро-усередненої метрики F1 на 9 % порівняно з результатами, отриманими традиційним методом навчання.

Ключові слова: *стічні труби; інспекція; класифікаційний аналіз; згорткова нейронна мережа; навчання з самоучителем; функція втрат; регуляризація; інформаційно-екстремальне машинне навчання.*

Вступ

Система стічних труб, що прокладається під землею на глибині до 10 метрів, розміщується під розвиненою інфраструктурою міста. Ремонт стічної системи потребує значних матеріальних витрат, погіршує стан міського середовища та чинить незручності для жителів. Віддалена відеоінспекція стічних труб забезпечує вирішення більшості задач моніторингу стану системи транспортування стоків до очисних споруд. Проте аналіз даних відеоінспекції є трудомістким процесом, а звіти про інспекцію можуть мати помилки внаслідок суб'єктивного фактору оператора. Тому останнім часом у світі активно ведуться дослідження і розробки спрямовані на автоматизацію процесу аналізу відео даних інспекції з використанням технології штучного інтелекту [1, 2].

Однак роботизоване устаткування для інспекції стічних труб досі не є автономним, оскільки засмі-

чені чи пошкоджені труби можуть бути джерелом поломки і застрявання роботів. Тому оператор робота є невід'ємною частиною процесу інспекції. Оператор приймає рішення про зупинку, об'їзд, розчищення трубопроводу, а також повертає і фокусує камеру на підозрілих дефектах стічних труб. Задача технологій штучного інтелекту полягає в автоматичному розпізнаванні та кодуванні дефектів згідно діючих стандартів [3, 4].

У роботах [5-7] було розглянуто глибокі моделі детектування дефектів стічних труб. Даний підхід потребує складного процесу анування навчальних даних обмежувальними рамками дефектів. Проте враховуючи наявність інформації про орієнтацію камери при її повороті чи фокусуванні, дану задачу набагато легше вирішити з використанням звичайного класифікатора, дані для навчання якого можна взяти зі звітів про інспекцію без їх трудомісткої розмітки. Проте висока варіативність спостережень

дефектів і їх комбінацій, незбалансованість даних і наявність шуму та артефактів зображення призводять до недостатньо високої точності розпізнавання. Тому досі неможливо здійснювати процес повністю автоматично, що збільшує накладні витрати на додаткову перевірку згенерованих звітів людиною.

У працях [7, 8] запропоновано і досліджено нові архітектури (state-of-the-art) згорткових мереж для класифікаційного аналізу зображень. Проте продуктивність традиційних методів глибокого машинного навчання залежить від гіперпараметрів і початкової ініціалізації вагових коефіцієнтів. У праці [9] для зменшення експресії результатів залежно від гіперпараметрів та підвищення робастності моделі до шуму і артефактів пропонується кодувати мітки класів кодами, що виправляють помилки, наприклад за допомогою таблиці Адамара, однак такі коди не враховують особливостей даних і можуть нав'язувати їм невласливу структуру. У працях [10, 11] було запропоновано навчання дискретного ознакового подання шляхом максимізації інформаційної міри. Такий підхід забезпечує ефективну регуляризацию та високу узагальнюючу здатність завдяки реалізації принципу інформаційного пляшкового горла (information bottleneck). Для зменшення впливу початкової ініціалізації вагових коефіцієнтів і зниження впливу незбалансованості даних у працях [12] запропоновано багатофазний метод машинного навчання робастної моделі, де навчання з триплетною функцією втрат слугує для ініціалізації моделі перед подальшим навчанням з мітками, що коректують помилки. Однак в цих працях не було розглянуто ефективне використання нерозмічених навчальних зразків для початкової ініціалізації вагових коефіцієнтів нейромережі.

В працях [13, 14] розглядаються різні варіанти навчання з самоучителем, найбільш популярним серед яких є методи контрастного навчання і їх модифікації. У загальному випадку функцією втрат контрастного навчання є NCE (тобто Noise Contrastive Estimation), що забезпечує оцінювання взаємної інформації між ознаковим поданням вибіркового зразка з ознаковим поданням решти зразків. Проте ефективність контрастного самонавчання для ініціалізації нейромережі не завжди передбачувана, може залежати від особливостей задач на кінці мережі (downstream task) і потребує дослідження в кожному конкретному випадку.

Метою статті є розроблення ефективного методу глибокого навчання з неповним залученням вчителя для класифікаційного аналізу дефектів стічних труб. Для досягнення поставленої мети пропонується вирішити наступні задачі: реалізувати глибоке навчання в рамках традиційного підходу та в рамках багатофазного методу з використанням

принципів інформаційно-екстремальної технології; здійснити навчання в рамках обох підходів з і без застосування попереднього контрастного самонавчання; оцінити результати і обрати найкращий варіант. Вирішення даних задач дасть змогу зробити висновок як про ефективність багатофазного методу навчання порівняно з традиційним, так і про вплив попереднього контрастного самонавчання на точнісні характеристики класифікаційного аналізу дефектів стічних труб.

Постановка задачі

Нехай дано вибірку $\{x_j^U \mid j = \overline{1, n}\}$ нерозмічених зображень в різних контекстах та вибірку $\{x_{m,j}^S \mid m = \overline{1, M}; j = \overline{1, n_m}\}$ розмічених зображень в контекстах камери направленої до стінок труби для навчання та тестування алгоритму класифікаційного аналізу дефектів, де n та n_m – обсяг вибірки нерозмічених зображень та обсяг вибірки розмічених зображень m -го класу відповідно. Нехай дано множину класів $\{X_m^O \mid m = \overline{1, M}\}$, що характеризує дефекти на стінках стічної труби. Крім того, відома структура вектора g параметрів моделі аналізу даних, що має вигляд

$$g = \langle e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1}, f_1, \dots, f_{\xi_2}, \dots, f_{\Xi_2} \rangle, \quad (1)$$

$$\Xi_1 + \Xi_2 = \Xi,$$

де e_{ξ_1} – ξ_1 -ий параметр, що впливає на екстракцію ознакового ознак з кадрів відеоінспекції, $\xi_1 = \overline{1, \Xi_1}$;

f_{ξ_2} – ξ_2 -ий параметр, що впливає на функціональну ефективність класифікатора, $\xi_2 = \overline{1, \Xi_2}$.

При цьому для параметрів моделі класифікаційного аналізу існують обмеження, що можуть бути задані у вигляді нерівностей $R_{\xi_1}(e_1, \dots, e_{\xi_1}, \dots, e_{\Xi_1}) \leq 0$, $R_{\xi_2}(f_1, \dots, f_{\xi_2}, \dots, f_{\Xi_2}) \leq 0$.

Необхідно знайти оптимальні значення параметрів вектора g (1), що забезпечують максимальне значення мікро-усередненого значення F1-міри для класифікатора дефектів на стінках труби

$$F1 = \frac{2 \sum_z TP_z}{2 \sum_z TP_z + \sum_z FP_z + \sum_z FN_z}, \quad (2)$$

$$g^* = \arg \max_G \{F1(g)\}. \quad (3)$$

де TP_z – значення лічильника правильно-позитивних рішень для z -го класу на тестовій вибірці;

FP_z – значення лічильника хибно-позитивних рішень для z -го класу на тестовій вибірці;

FN_z – значення лічильника хибно-негативних рішень для z -го класу на тестовій вибірці.

Модель та метод навчання для класифікаційного аналізу дефектів на стінках стічних труб

Для класифікаційного аналізу дефектів на кадрах пропонується застосовувати обчислювально ефективну згорткову мережу MobileNet з коефіцієнтом ємності, що дорівнює 0,5. При цьому використовується лише екстрактор ознак (backbone) без вихідних шарів [7]. Вхідна роздільна здатність мережі дорівнює 224x224 пікселів. На рис. 1, а показано класичний варіант згорткової мережі, а на рис. 1, б модифікований варіант для дослідження пропонованого методу навчання.

MobileNet екстрактор
Global Average Pooling 2D шар
Dropout шар (rate=0.5)
Dense шар (128 вузлів)
Сігмоїдний шар
Dense шар (Z вузлів)
Softmax шар

а

MobileNet екстрактор
Global Average Pooling 2D шар
Dropout шар (rate=0.5)
Dense шар (128 вузлів)
Сігмоїдний шар
Округлюючий шар
Радіально-базисний шар

б

Рис. 1. Архітектура згорткової мережі: а – традиційний варіант; б – модифікований варіант

Для зменшення розмірності на виході екстрактора ознак використовується шар Global Average Pooling. Для регуляризації додатково використовується шар Dropout [7, 8]. Ознаковий опис формується на виході сігмоїдного шару. Класифікаційні рішення моделі приймаються шляхом округлення ознакового опису до двійкового подання і застосування та радіа-

льно-базисних функцій належності до класів, що розділені гіперсферичними контейнерами. Кожен гіперсферичний контейнер класу описується двійковим еталонним вектором (прототипом класу), який визначає геометричний центр контейнера, та радіусом контейнера в кодових одиницях. В цьому випадку функція належності $\mu_z(b)$ до z -го класу для N -вимірного двійкового вектора b може бути обчислена за формулою

$$\mu_z(b) = 1 - \sum_{i=1}^N b_i \oplus b_{z,i}^* / d_z^*, \quad (4)$$

де b_z^* – оптимальний двійковий еталонний вектор класу X_z^0 ;

d_z^* – оптимальний кодовий радіус в просторі Хеммінга для контейнера для класу X_z^0 .

Традиційним підходом до навчання згорткової моделі класифікатора є додавання до екстрактора ознак повнозв'язного (Dense) шару і Softmax-нормалізації та використання оптимізатора, що здійснює пошук мінімуму крос-ентропійної функції втрат за методом зворотного поширення помилки (наприклад Adam).

В основу розробки пропонованого методу навчання покладено такі методологічні принципи :

- навчання в декілька послідовних етапів, де кожен попередній етап здійснює ініціалізацію вагових коефіцієнтів моделі для наступного етапу і може повторюватися для реалізації принципу відкладених рішень та узгодження різних методів навчання;
- на першому етапі здійснюється навчання з самоучителем для ефективного використання великого обсягу нерозмічених навчальних даних;
- поєднання дискретного ознакового подання (Deep Hashing) з кодуванням міток класів (Target Coding), тобто один з етапів навчання з учителем повинен сформувати еталонні вектори дискретних ознак, що стануть кодами (мітками) класів для наступного етапу;

– врахування під час навчання екстрактора ознак нелінійності вирішувальних правил, що будуються в радіальному базисі;

– оптимізація вирішувальних правил повинна здійснюватися на основі інформаційного критерію, оскільки це забезпечує більшу узагальнюючу здатність за умов незбалансованості та обмеженого обсягу розмічених зразків.

Запропонований метод включає 5 основних етапів (фаз) навчання (рис. 2). Перший етап полягає в навчанні на нерозмічених даних з використанням екземпляр-прототипної контрастної процедури тре-

нування (instance-prototype contrastive learning) за модифікованою функцією втрат, що має формулу

$$L_{\text{contrastive}} = -\frac{1}{n} \sum_{i=1}^n \log \frac{\exp\left(-\frac{1}{N} \|z_i - \bar{z}\|_2^2\right)}{\exp\left(-\frac{1}{N} \|z_i - \bar{z}\|_2^2\right) + \sum_{k=1}^K \exp\left(-\frac{1}{N} \|z_i - z_k\|_2^2\right)}, \quad (5)$$

де n – кількість зразків в міні-пакеті ($n = 32$);

z_i – ознаковий опис на виході екстрактора ознак $z_i = f(x_i)$ для вхідного x_i зразка з міні-пакету;

$\bar{z} = \frac{1}{n_a} \sum_{j=1}^{n_a} f(x_j)$ – усереднене ознакове по-

дання для аугментованих версій вхідного зразка x_j , що розглядається як позитивна пара для вхідного зразка x_j ,

де $f(x)$ – функція, що описує екстрактор ознак і встановлює залежність між вхідним зображенням і вектором на виході сигмоїдного шару;

K – кількість попередньо оброблених зразків, що розглядаються як негативні пари, ознакові подання яких зберігаються в неіндексованому буфері (довжиною $K = 1024$) і оновлюються після обробки кожного міні-пакету;

N – розмірність ознакового опису на виході екстрактора ознак;

$\|\cdot\|_2^2$ – квадрат відстані Евкліда.

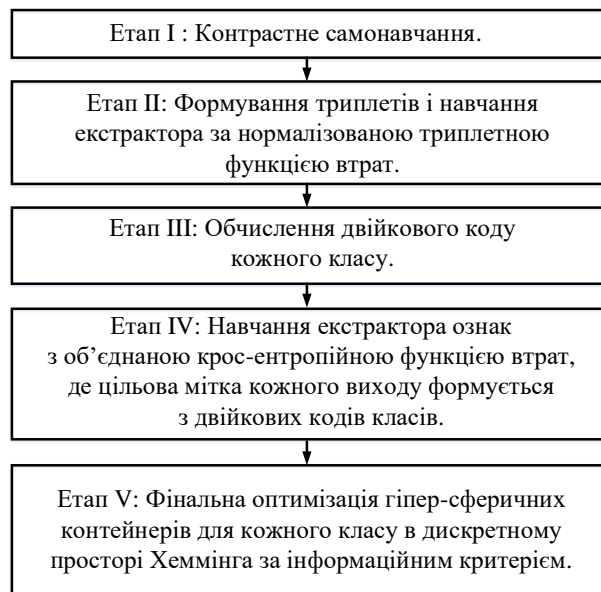


Рис. 2. Етапи запропонованого методу навчання

Друга фаза машинного навчання виконується на розміщеному датасеті з використанням модифікованої триплетної функції втрат. Розглянемо два варіанти функції втрат: модифікація NCA функції втрат (Neighbourhood Components Analysis) з майнінгом триплетів; контрастно-центрована функція втрат (contrastive-center loss).

Модифікована NCA функція втрат обчислюється на змішаних міні-пакетах, кожен з яких містить M зображень кожного класу, і має такий вигляд:

$$L_{\text{NCA}} = -\log \frac{\exp\left(-\frac{1}{N} \|z_a - z_{ep}\|_2^2\right)}{\exp\left(-\frac{1}{N} \|z_a - z_{ep}\|_2^2\right) + \exp\left(-\frac{1}{N} \|z_a - z_{shn}\|_2^2\right)}, \quad (6)$$

де $z_a = f(x_a)$, $z_{ep} = f(x_{ep})$ та $z_{shn} = f(x_{shn})$,

x_a – якірне зображення, обране з міні-пакету;

$$x_{ep} = \arg \min_{x:C(x)=C(x_a)} \|f(x_a) - f(x)\|, \quad (7)$$

$$x_{shn} = \arg \min_{\substack{C(x) \neq C(x_a) \\ \|f(x_a) - f(x)\| > \|f(x_a) - f(x_p)\|}} \|f(x_a) - f(x)\|, \quad (8)$$

де $C(x)$ – функція, що повертає мітку класу для конкретного зразка.

Контрастно-центрована функція втрат (contrastive-center loss), в більш явному вигляді забезпечує збільшення компактності розподілу класів і збільшення відстані між центрами розподілу класів. При цьому контрастно-центрована функція втрат в явному вигляді використовує еталонні вектори класів $\{\bar{z}_j, j=1, \overline{M}\}$, які оптимізуються в процесі машинного навчання. Контрастно-центрована функція втрат має такий вигляд

$$L_{\text{center}} = -\frac{1}{n} \sum_{i=1}^n \left(\frac{\|f(x_i) - \bar{z}_{y_i}\|_2^2}{\sum_{j=1, j \neq y_i}^k \|f(x_i) - \bar{z}_j\|_2^2 + \delta} \right), \quad (9)$$

де y_i – номер класу для x_i зразка даних; \bar{z}_{y_i} – еталонний вектор y_i класу, який оптимізується в процесі машинного навчання; \bar{z}_j – еталонний вектор j -го класу, відмінного від y_i -го класу; δ – константа для уникнення ділення на нуль (за замовчуванням $\delta=1$); n – кількість зразків в міні-пакеті.

Для регуляризації розглядається штрафування за відхилення ознакового подання від двійкового виду. Ця регуляризація забезпечує інформаційне пляшкове горло (Information bottleneck), що сприяє підвищенню завадозахищеності та узагальнюючої здатності моделі. Регуляризуюча складова характеризується Липшицевою неперервністю (Lipschitz Continuous) і обчислюється за формулою:

$$R_{\text{bin}} = \frac{1}{n} \sum_{i=1}^n \lambda f(x_i)^T (e - f(x_i)), \quad (10)$$

де e – одиничний вектор; λ – коефіцієнт регуляризації ($\lambda = 0,01$).

Наступна фаза навчання використовується для обчислення двійкового коду (прототипу) класу для реалізації кодів, що виправляють помилки (error-correcting output codes), але з урахуванням структури класів і відношень між їх ознаками. Для цього навчальна вибірка для M класів $\{x_{m,s} \mid m = \overline{1, M}, s = \overline{1, n_m}\}$, що містить n_m зразків m -го класу, кодується дискретним поданням $\{b_{m,s,i} \mid m = \overline{1, M}, s = \overline{1, n_m}, i = \overline{1, N}\}$ з розмірністю N .

Кодування вхідного зображення $x_{m,s}$ двійковим поданням здійснюється операцією округлення i -го виходу сігмоїдного шару

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

Двійковий еталонний вектор b_m для m -го класу можна визначити шляхом порозрядного порівняння частоти двійкових одиниць в m -му класі з фоновою частотою одиниць в навчальній вибірці

$$b_{m,i} = \begin{cases} 1, & \text{if } \frac{1}{n_m} \sum_{s=1}^{n_m} b_{m,s,i} > \frac{1}{M} \sum_{c=1}^M \frac{1}{n_c} \sum_{s=1}^{n_c} b_{c,s,i}; \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

Еталонний вектор (прототип) z -го класу b_z використовується як цільова розмітка (label) кожного зразка даного класу на наступному етапі навчання з використанням крос-ентропійної функції втрат, яка для вхідного зображення x обчислюється за формулою

$$L = - \sum_{i=1}^N (b_i \log f_i(x) + (1 - b_i) \log(1 - f_i(x))), \quad (12)$$

де $f_i(x)$ – значення i -го виходу сігмоїдного шару для вхідного зображення x ;

b_i – значення i -го розряду еталонного вектора класу, до якого належить зображення x .

Останній етап машинного навчання призначений для оптимізації кодового радіусу контейнерів за інформаційним критерієм (14) для врахування дисперсії розподілу кожного з класів в двійковому просторі Хеммінга

$$E_m^* = \max_{\{d\}} E_m(d), \quad (13)$$

де $\{d\} = \{0, 1, \dots, \left(\sum_i b_{m,i} \oplus b_{c,i} - 1\right)\}$ – набір концентричних радіусів з центром в b_m ;

E_m – інформаційний критерій для m -го класу, який є функцією від характеристик точності [12].

Як інформаційний критерій оптимізації пропонується використати модифікацію ентропійної міри Шеннона, що виражена у вигляді функціоналу від точнісних характеристик рішень [12]. Оскільки оптимізація контейнерів відбувається окремо для кожного класу, то застосуємо формулу обчислення інформаційного критерію для дихотомічної задачі класифікації у такому вигляді [12]:

$$E_m^{(k)} = 1 + \frac{1}{2} \left(\frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{\alpha_m^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} + \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{\beta_m^{(k)}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} \log_2 \frac{D_{1,m}(d)}{D_{1,m}^{(k)}(d) + \beta_m^{(k)}(d)} + \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \log_2 \frac{D_{2,m}^{(k)}(d)}{\alpha_m^{(k)}(d) + D_{2,m}^{(k)}(d)} \right), \quad (14)$$

де $\alpha_m^{(k)}(d)$ – оцінка ймовірності помилок першого роду на k -му кроці навчання;

$\beta_m^{(k)}(d)$ – оцінка ймовірності помилок другого роду;

$D_{1,m}^{(k)}(d)$ – перша достовірність (чутливість);

$D_{2,m}^{(k)}(d)$ – друга достовірність (специфічність);

d – кодовий радіус гіперсферичних контейнерів в радіальному базисі простору Хеммінга.

Результати машинного навчання

Для підготовки вхідного математичного опису інтелектуальної інформаційної системи аналізу відео зображень було використано SEWER-ML набір

даних [8]. Набір даних містить 1,3 мільйони зображень з 75618 відеофайлів, зібраних з датських компаній, що обслуговують стічні трубопроводи. Проте датський стандарт, за яким здійснена розмітка вибірки не дозволяє поширити даний підхід на інші стандарти. Тому здійснена ручна перерозмітка даних. Було виділено 13 класів розпізнавання, 12 з яких показано на рис. 3. В більшості стандартів протокол інспекції передбачає, що оператор кравлера повертає камеру для детальнішого огляду дефекту і оцінювання його тяжкості, тому розглянемо задачу розпізнавання дефектів при повороті камери на стінку труби. Для задачі попереднього навчання без вчителя планується використати 20% набору даних. А для точної настройки під задачу розпізнавання дефектів планується обрати тільки ті зображення, що містять дефекти в контексті “вид збоку”. При цьому зображення приводяться до роздільної здатності 224x224 пікселів з метою зменшення витрат ресурсів на навчання та екзамени (inference). Алфавіт сформованих класів показано в табл. 1. Приклади зображень кожного класу показано на рис. 3.

Таблиця 1

Структура розміченої вибірки даних для розпізнавання дефектів на стінках стічних труб

Клас	Назва дефекту	Кількість розмічених зразків
X_1^0	Порушення цілісності прокладки на стику	250
X_2^0	Відкритий стик	200
X_3^0	Зміщений стик	331
X_4^0	Тріщина	270
X_5^0	Розлом	189
X_6^0	Розбита труба	306
X_7^0	Дірка в трубі	450
X_8^0	Пошкодження поверхні	210
X_9^0	Прикріплені відкладення	200
X_{10}^0	Осад	300
X_{11}^0	Тоненькі корені	500
X_{12}^0	Значне розростання коренів	310
X_{13}^0	Нормальний стан труби	500

Поділ вибірки на тестову і навчальну здійснюється у відношенні 70 % на навчання і 30 % на тестування. При цьому під час поділу використовується стратегія збереження пропорції кількості зразків кожного класу в тестовій і навчальній вибірці.

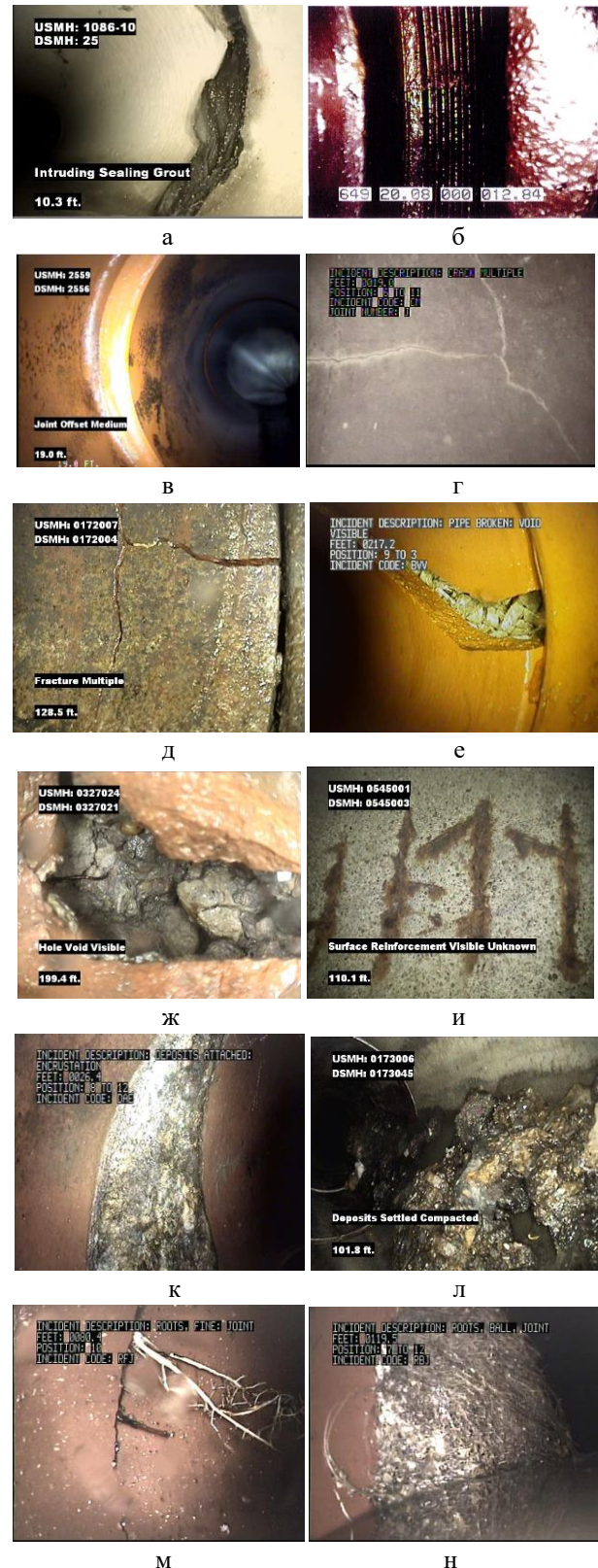


Рис. 3. Приклади зображень кожного з класів:

- а – клас X_1^0 ; б – клас X_2^0 ; в – клас X_3^0 ;
- г – клас X_4^0 ; д – клас X_5^0 ; е – клас X_6^0 ;
- ж – клас X_7^0 ; и – клас X_8^0 ; к – клас X_9^0 ;
- л – клас X_{10}^0 ; м – клас X_{11}^0 ; н – клас X_{12}^0

Балансування навчальної підвибірki здійснюється за рахунок аугментації класів з меншою кількістю зразків. Аугментація полягає у застосуванні операцій масштабування (збільшення з обрізанням) до 5 %, обертання з обрізанням на ± 18 градусів та збільшенням чи зменшенням яскравості до 5 %.

Розмір навчальної вибірки кожного класу після аугментації для класифікатора дефектів становить 500 зразків.

Процес навчання без вчителя пропонується продовжувати доки функція втрат не перестане помітно зменшуватися. Для вимірювання ефективності моделей під час навчання з учителем використовується критерій (2), що обчислюється на тестовій вибірці. При цьому після обробки кожного міні-паketу пропонується заново будувати інформаційно-екстремальні вирішувальні правила і тестувати їх на тестовій вибірці.

На рис. 4 показано графік зміни функції втрат (5) (Loss) від кількості ітерацій (Number of iterations).

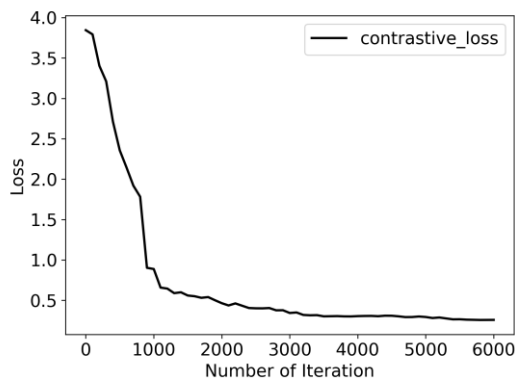


Рис. 4. Залежність контрастної функції втрат (5) на наборі даних від кількості ітерацій навчання на міні-пакетах нерозміченого набору даних Sewer-ML

Аналіз рис. 4 показує, що після 6 000 ітерацій функція втрат помітно не змінювалась.

На рис. 5 показано графік зміни значення метрики F1 (2) на тестовій вибірці (F1_val) під час навчання класифікатора дефектів за модифікованою триплетною функцією втрат NCA (6) (F1_val_nca) та контрастно-центрованою функцією втрат (7) (F1_val_cen).

Показані на рис. 5 криві F1_val_nca_pretrained та F1_val_cen_pretrained показують залежність метрики F1 (2) на тестовій вибірці від кількості епох навчання з функціями втрат (6) та (7) відповідно, але після контрастного самонавчання з контрастною функцією втрат (5).

Аналіз рис. 5 показує, що попереднє навчання без вчителя покращує точнісні характеристики мо-

делей, що навчаються як за функцією втрат (6), так і за функцією втрат (7). При цьому результат навчання з функцією втрат (6) після попереднього самонавчання забезпечує значення метрики F1=0,911, а результат навчання з функцією втрат (7) після попереднього самонавчання забезпечує значення метрики F1=0,927. Крім того помітно, що під час навчання з функцією втрат (6) метрика F1 зростає повільніше протягом 20 перших епох порівняно з метрикою F1 під час використання функції втрат (7).

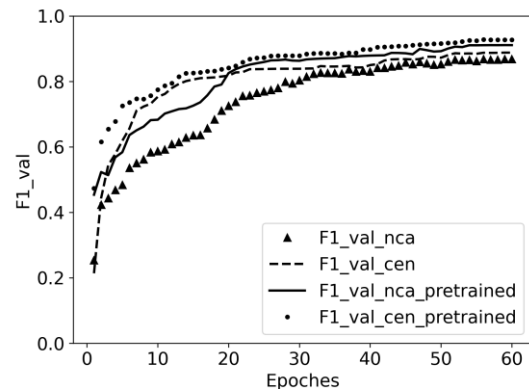


Рис. 5. Залежність значення метрики F1 (2) на тестовій вибірці від кількості епох навчання за функцією втрат (6) та (7) до і після попереднього навчання без вчителя з використанням функції втрат (5)

Для наступного етапу запропонованого методу використаємо моделі з переднього етапу навчання, заморожені після 30 епох, як стартові точки пошуку. Для порівняння здійснимо навчання моделей з традиційною структурою вихідного шару (див. рис. 1, б) і традиційним методом оптимізації. На рис. 6 показано графік зміни значення метрики F1 (2) на тестовій вибірці (F1_val) під час навчання класифікатора дефектів за традиційним методом без попереднього самонавчання (F1_val_baseline) та з попереднім контрастним самонавчанням (F1_val_baseline_pre). Крім цього на графіку показано зміни метрики F1 (1) на тестовій вибірці під час IV етапу навчання з крос-ентропійною функцією втрат (12) для моделей, що були попередньо навчені під час самонавчання та з учителем за модифікованою триплетною функцією втрат NCA (6) (F1_val_nca_pretrained) та контрастно-центрованою функцією втрат (7) (F1_val_cen_pretrained).

Аналіз рис. 6 показує, що запропонований метод навчання має переваги у порівнянні з традиційним методом навчання за метрикою F1 (2). При цьому найкращі результати забезпечує застосування на II етапі навчання контрастно-центрованої функції втрат (7). Запропонований метод навчання дозволяє підвищити метрику F1 на 9 % порівняно з традицій-

ним підходом. В результаті оптимальна модель класифікатора дефектів на стінках стічних труб має таке значення метрики $F1=0,977$, що свідчить про високу достовірність вирішувальних правил.

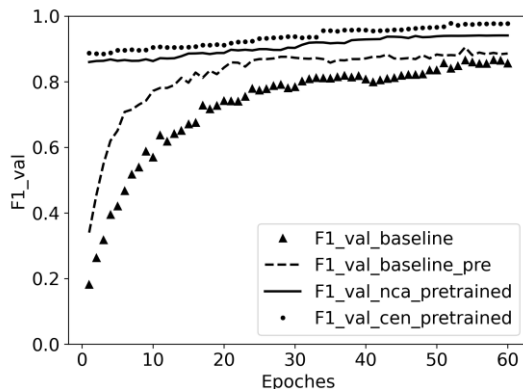


Рис. 6. Залежність значення F1-метрики (2) на тестовій вибірці від кількості епох навчання за традиційним методом та за запропонованим методом навчання з використанням функцій втрат (6) або (7)

Таким чином, експериментально доведено перевагу використання запропонованої моделі та багатофазного методу навчання для класифікаційного аналізу дефектів на стінках стічних труб. Показано, що використання попереднього контрастного самонавчання покращує результати як традиційного, так запропонованого методу навчання. Отримана точність прийнятна до практичного використання, оскільки подібні помилки допускають і самі оператори.

Висновки

1. Наукова новизна отриманих результатів :

- вперше розроблено багатоетапний метод машинного навчання, перший етап якого полягає в контрастному самонавчанні на нерозмічених даних, а наступні етапи пов'язані з визначенням двійкового коду кожного класу, що використовується як мітка класу під час точної настройки моделі;

- експериментально підтверджено перевагу запропонованого багатоетапного методу машинного навчання порівняно з традиційним підходом до глибокого навчання моделі класифікатора зображень;

- експериментально підтверджено, що використання попереднього контрастного самонавчання покращує наступні результати як традиційного, так запропонованого методу навчання.

2. Практичне значення роботи

полягає в підвищенні точнісних характеристик моделі класифікаційного аналізу дефектів на стінках стічних труб під час відеоінспекції за умов обмеженого обсягу

розмічених даних та неповної визначеності, пов'язаної з довільними умовами спостереження.

Подальші дослідження будуть спрямовані на поширення основних ідей запропонованого методу для навчання моделей детектування, локалізації та оцінювання розміру дефектів стічних труб.

Робота виконувалась у рамках науково-дослідного проекту "Виконання завдань перспективного плану розвитку наукового напрямку "Технічні науки" Сумського державного університету (ДР № 0121U112684).

Література

1. 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.
2. Moradi, S. Review on Computer Aided Sewer Pipeline Defect Detection and Condition Assessment [Text] / S. Moradi, T. Zayed, F. Golkhoo // *Infrastructures*. – 2019. – Vol 4, No. 1. – Article Id: 10. DOI: 10.3390/infrastructures4010010.
3. Haurum, J. B. A Survey on image-based automation of CCTV and SSET sewer inspections [Text] / J. B. Haurum, T. B. Moeslund // *Automation in Construction*. – 2020. – Vol. 111. – Article Id: 103061. DOI: 10.1016/j.autcon.2019.103061.
4. Visual-Based Defect Detection and Classification Approaches for Industrial Applications—A SURVEY [Text] / T. Czimmermann, G. Ciuti, M. Milazzo, M. Chiurazzi, S. Roccella, C. M. Oddo, P. Dario // *Sensors*. – 2020. – Vol. 20. – Article Id: 1459. DOI: 10.3390/s20051459.
5. Sewer pipe defect detection via deep learning with local and global feature fusion [Text] / Dawei Li, Qian Xie, Zhenghao Yu, Qiaoyun Wu, Jun Zhou, Jun Wang // *Automation in Construction*. – 2021. – Vol. 129. – Article Id: 103823. DOI: 10.1016/j.autcon.2021.103823.
6. Li, D. Sewer damage detection from imbalanced CCTV inspection data using deep convolutional neural networks with hierarchical classification [Text] / D. Li, A. Cong, S. Guo // *Automation in Construction*. – 2019. – Vol. 101. – P. 199-208. DOI: 10.1016/j.autcon.2019.01.017.
7. Deep Learning-Based Automated Detection of Sewer Defects in CCTV Videos [Text] / S. S. Kumar, W. Mingzhu, D. M. Abraham, M. R. Jahanshahi, I. Tom, J. C. Cheng // *Journal of Computing in Civil Engineering*. – 2020. – Vol. 34, Iss. 1. – Article Id: 4019047. DOI: 10.1061/(asce)cp.1943-5487.0000866.

8. Haurum, J. B. Sewer-ML: A Multi-Label Sewer Defect Classification Dataset and Benchmark [Text] / J. B. Haurum, T. B. Moeslund // *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021. – P. 13456-13467.

9. Challenging the Adversarial Robustness of DNNs Based on Error-Correcting Output Codes [Text] / B. Zhang, B. Tondi, X. Lv, M. Barni // *Security and Communication Networks*. – 2020. – Vol. 2020. – Article Id: 8882494. DOI: 10.1155/2020/8882494.

10. Kolchinsky, A. Nonlinear Information Bottleneck [Text] / A. Kolchinsky, B. D. Tracey, D. H. Wolpert // *Entropy*. – 2019. – Vol. 21, Iss. 12. – Article Id: 1181. DOI: 10.3390/e21121181.

11. Contrastive Self-Supervised Hashing With Dual Pseudo Agreement [Text] / L. Yang, Y. Wang, Z. Miao, J. Wang, R. Zhang // *IEEE Access*. – 2020. – Vol. 8. – P. 165034–165043. DOI: 10.1109/ACCESS.2020.3022672.

12. Sewer Pipe Defects Classification Based on Deep Convolutional Network with Information-extreme Error-correction Decision Rules [Text] / A. S. Moskalenko, V. V. Moskalenko, M. O. Zaretskyi, V. Lysyuk // *Communications in Computer and Information Science*. – 2020. – Vol. 1158. – P. 253-263. DOI: 10.1007/978-3-030-61656-4_16.

13. Liu, X. Self-supervised Learning: Generative or Contrastive [Text] / X. Liu, F. Zhang, Z. Hou, L. Mian, Z. Wang, J. Zhang, J. Tang // *IEEE Transaction on Knowledge and Data Engineering*. – 2021. – DOI: 10.1109/TKDE.2021.3090866

14. Інформаційно-аналітична система оцінювання відповідності сучасним вимогам навчального контенту спеціальності кібербезпека [Текст] / А. С. Довбиш, І. В. Шелехов, Ю. О. Хібовська, О. В. Матяш // *Радіоелектронні і комп'ютерні системи*. – 2021. – № 1. – С. 70-80. DOI: 10.32620/reks.2021.1.06.

References

1. 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.

2. Moradi, S., Zayed, T., Golkhoo, F. Review on Computer Aided Sewer Pipeline Defect Detection and Condition Assessment. *Infrastructures*, 2019, vol 4, no. 1, article id: 10. DOI: 10.3390/infrastructures4010010.

3. Haurum, J. B., Moeslund, T. B. A Survey on image-based automation of CCTV and SSET sewer inspections. *Automation in Construction*, 2020,

vol. 111, article id: 103061. DOI: 10.1016/j.autcon.2019.103061.

4. Czimmermann, T., Ciuti, G., Milazzo, M., Chiurazzi, M., Roccella, S., Oddo, C. M., Dario, P. Visual-Based Defect Detection and Classification Approaches for Industrial Applications – A SURVEY. *Sensors*, 2020, vol. 20, article Id: 1459. DOI: 10.3390/s20051459.

5. Li, Dawei., Xie, Qian., Yu, Zhenghao., Wu, Qiaoyun., Zhou, Jun., Wang, Jun. Sewer pipe defect detection via deep learning with local and global feature fusion. *Automation in Construction*, 2021, vol. 129, article id: 103823. DOI: 10.1016/j.autcon.2021.103823.

6. Li, Duanshun., Cong, A., Guo, S. Sewer damage detection from imbalanced CCTV inspection data using deep convolutional neural networks with hierarchical classification. *Automation in Construction*, 2019, vol. 101, pp. 199-208. DOI: 10.1016/j.autcon.2019.01.017.

7. Kumar, S. S., Mingzhu, W., Abraham, D. M., Jahanshahi, M. R., Tom, I., Cheng, J. C. Deep Learning-Based Automated Detection of Sewer Defects in CCTV Videos. *Journal of Computing in Civil Engineering*, 2020, vol. 34, iss. 1, Article Id: 4019047. DOI: 10.1061/(asce)cp.1943-5487.0000866.

8. Haurum, J. B., Moeslund, T. B. Sewer-ML: A Multi-Label Sewer Defect Classification Dataset and Benchmark. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021, pp. 13456-13467.

9. Zhang, B., Tondi, B., Lv, X., Barni, M. Challenging the Adversarial Robustness of DNNs Based on Error-Correcting Output Codes. *Security and Communication Networks*, 2020, vol. 2020, article id: 8882494. DOI: 10.1155/2020/8882494.

10. Kolchinsky, A., Tracey, B. D., Wolpert, D. H. Nonlinear Information Bottleneck. *Entropy*, 2019, vol. 21, iss. 2, article id: 1181. DOI: 10.3390/e21121181.

11. Yang, L., Wang, Y., Miao, Z., Wang, J., Zhang, R. Contrastive Self-Supervised Hashing With Dual Pseudo Agreement. *IEEE Access*, 2020, vol. 8, pp. 165034 – 165043. DOI: 10.1109/ACCESS.2020.3022672.

12. Moskalenko, A., Moskalenko, V., Lysyuk, V., Zaretskyi, M. Sewer Pipe Defects Classification Based on Deep Convolutional Network with Information-extreme Error-correction Decision. *Communications in Computer and Information Science*, 2020, vol. 1158, pp. 253-263. DOI: 10.1007/978-3-030-61656-4_16.

13. Liu, X., Zhang, F., Hou, Z., Mian, L., Wang, Z., Zhang, J., Tang, J. Self-supervised Learning: Generative or Contrastive. *IEEE Transaction on Knowledge and Data Engineering*, 2021. DOI: 10.1109/TKDE.2021.3090866.

14. Dovbysh, A., Shelechov, I., Khibovska, Ju., Matiash, O. Informatsiyno-analitychna systema otsinyuvannya vidpovidnosti suchasnym vymoham navchal'noho kontentu spetsial'nosti kiberbezpeka [Information and analytical system for assessing the com-

pliance of educational content specialties ciber security with modern requirements]. *Radioelectronic and computer systems*, 2021, vol. 1, pp. 70-80. DOI: 10.32620/reks.2021.1.06

Надійшла до редакції 11.10.2021 розглянута на редколегії 26.11.2021

МНОГОЭТАПНЫЙ МЕТОД ГЛУБИННОГО ОБУЧЕНИЯ С ПРЕДВАРИТЕЛЬНЫМ САМООБУЧЕНИЕМ ДЛЯ КЛАССИФИКАЦИОННОГО АНАЛИЗА ДЕФЕКТОВ СТОЧНЫХ ТРУБ

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

Разработан многоэтапный метод машинного обучения с предварительным самообучением для классификационного анализа дефектов на стенках сточных труб по изображениях видеоинспекции. Объектом исследования является процесс распознавания дефектов на стенках сточных труб. Предметом исследования является метод машинного обучения для классификационного анализа дефектов сточных труб на изображениях видеоинспекции в условиях ограниченного и несбалансированного набора размеченных обучающих данных. Предложено пятиэтапный алгоритм обучения классификатора. На первом этапе происходит контрастное обучение с использованием экземпляра-прототипной контрастной функцией потерь, где для измерения сходства закодированных образцов используется нормализованное расстояние Эвклида. На втором этапе рассматривается два варианта регуляризованных функций потерь – триплетная функция NCA и контрастно-центрированная функция. Регуляризирующая составляющая на втором этапе обучения используется для штрафования за ошибку округления выходного вектора признаков к дискретному виду и обеспечивает реализацию принципа информационного бутылочного горла. На следующем этапе производится вычисление двоичного кода каждого класса для реализации кодов, исправляющих ошибки, но с учетом структуры классов и отношений между их признаками. Полученный эталонный вектор каждого класса является целевой разметкой изображения для обучения с использованием кросс-энтропийной функции потерь. Последний этап обучения осуществляет оптимизацией параметров решающих правил по информационному критерию для учета дисперсии распределения классов в двоичном пространстве Хемминга. Для сравнения результатов обучения на разных этапах и в рамках разного подхода используется микро-усредненная метрика F1, которая вычисляется на тестовых данных. Результаты, полученные на открытом наборе данных Sewer-ML, подтверждают пригодность метода обучения к практическому использованию, полученное значение F1-метрики равно 0,977. Было показано, что предложенный метод обеспечивает увеличение значения микро-усредненной метрики F1 на 9 % в сравнении с результатами, полученными традиционным методом.

Ключевые слова: сточные трубы; инспекция; классификационный анализ; свёрточная нейронная сеть; обучение с самоучителем; функция потерь; регуляризация; информационно-экстремальное машинное обучение.

MULTI-STAGE DEEP LEARNING METHOD WITH SELF-SUPERVISED PRETRAINING FOR SEWER PIPE DEFECTS CLASSIFICATION

*V. Moskalenko, M. Zaretskyi, A. Moskalenko,
A. Korobov, Y. Kovalskyi*

A machine learning semi-supervised method was developed for the classification analysis of defects on the surface of the sewer pipe based on CCTV video inspection images. The aim of the research is the process of defect detection on the surface of sewage pipes. The subject of the research is a machine learning method for the classification analysis of sewage pipe defects on video inspection images under conditions of a limited and unbalanced set of labeled training data. A five-stage algorithm for classifier training is proposed. In the first stage, contrast training occurs using the instance-prototype contrast loss function, where the normalized Euclidean distance is used to measure the similarity of the encoded samples. The second step considers two variants of regularized loss functions – a triplet NCA function and a contrast-center loss function. The regularizing component in the second stage of training is used to penalize the rounding error of the output feature vector to a discrete form and ensures that the principle of information bottlenecking is implemented. The next step is to calculate the binary code of each class to implement

error-correcting codes, but considering the structure of the classes and the relationships between their features. The resulting prototype vector of each class is used as a label of image for training using the cross-entropy loss function. The last stage of training conducts an optimization of the parameters of the decision rules using the information criterion to consider the variance of the class distribution in Hamming binary space. A micro-averaged metric F1, which is calculated on test data, is used to compare learning outcomes at different stages and within different approaches. The results obtained on the Sewer-ML open dataset confirm the suitability of the training method for practical use, with an F1 metric value of 0.977. The proposed method provides a 9 % increase in the value of the micro-averaged F1 metric compared to the results obtained using the traditional method.

Keywords: sewer pipes; inspection; classification analysis; convolutional neural network; self-learning; loss function; regularisation; information-extreme machine learning.

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

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

Москаленко Альона Сергіївна – канд. техн. наук., ст. викладач каф. комп'ютерних наук, Сумський державний університет, Суми, Україна.

Коробов Артем Геннадійович – канд. техн. наук, асистент каф. комп'ютерних наук, Сумський державний університет, Суми, Україна.

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

Viacheslav Moskalenko – PhD, associate professor of Computer Science Department of Sumy State University, Sumy, Ukraine,

e-mail: v.moskalenko@cs.sumdu.edu.ua, ORCID: 0000-0001-6275-9803, Scopus Author ID: 57189099775.

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

Alona Moskalenko – PhD, senior lecturer of Computer Sciences Department of Sumy State University, Sumy, Ukraine,

e-mail: a.moskalenko@cs.sumdu.edu.ua, ORCID: 0000-0003-3443-3990, Scopus Author ID: 57148522500.

Artem Korobov – PhD, assistant of Computer Science Department of Sumy State University, Sumy, Ukraine, e-mail: a.korobov@cs.sumdu.edu.ua, ORCID: 0000-0003-3239-1977.

Yaroslav Kovalskyi – PhD student of Computer Science Department of Sumy State University, Sumy, Ukraine,

e-mail: socratus43@gmail.com, ORCID: 0000-0002-5345-5186.