УДК 681.007.05

КЛАСИФІКАЦІЯ ПОВЕРХНЕВИХ ТЕХНОЛОГІЧНИХ ДЕФЕКТІВ У МЕТАЛОПРОКАТІ ЗА ДОПОМОГОЮ ГЛИБОКОЇ НЕЙРОННОЇ МЕРЕЖІ

І. В. КОНОВАЛЕНКО, П. О. МАРУЩАК

Тернопільський національний технічний університет імені Івана Пулюя

Розроблено та досліджено автоматизований метод класифікації трьох типів технологічних дефектів металопрокату. Основою методу є класифікатор на базі нейронної мережі ResNet50, який дає змогу розрізняти на зображеннях плоских поверхонь пошкодження трьох класів з точністю 95,8% на тестових даних з незбалансованою кількістю зображень різних видів. Для навчання класифікатора використано ~88 тис. зображень. Показано, що застосування розробленої моделі на базі нейромережі ResNet50 забезпечує відмінні показники продуктивності, високу швидкодію і точність, що робить такий класифікатор дієвим інструментом для задач технічної діагностики і неруйнівних методів контролю для класифікації дефектів на поверхнях металопрокату.

Ключові слова: *прокат, поверхневі дефекти, розпізнавання, класифікатор, пошкодження.*

An automated method for detecting and classifying three types of rolled metal defects has been developed and investigated. The method is based on the ResNet50 neural network classifier, which allows three classes of damage to be recognized on images of flat surfaces with a total accuracy of 95.8%, calculated on the basis of analysis of test data with an unbalanced number of images of various types. 88 thousand images were used to train the classifier. It is shown that the use of the developed model based on the ResNet50 neural network provides high recognition quality, high speed and accuracy, which makes this classifier an effective tool for identifying and classifying defects on metal surfaces.

Keywords: *rolling strip, surface defects, recognition, classifier, damage.*

Вступ. Розвиток сучасної металургії обумовлює необхідність розвитку методів онлайн-контролю якості прокату. Одні з основних у цьому напрямку є автоматизовані методи оптико-цифрового контролю з використанням для обробки одержаних зображень нейронних мереж [1–3]. Перевагою такого підходу є можливість безперервного контролю дефектів поверхні прокату, виявлення їх основних груп, розроблення оперативних заходів щодо зменшення їх кількості. На сьогодні апробовані як лабораторні, так і промислові зразки [4–8]. При цьому основними напрямками їх поліпшення залишаються вдосконалення апаратної частини для одержання зображень кращої якості та підвищення ефективності програмних продуктів, зокрема точності розпізнавання та класифікації дефектів [9–11]. Для цього широко застосовують нейронні мережі, однією з переваг яких є одночасне функціонування елементів. Це суттєво поліпшує швидкодію системи і цим збільшує ефективність обробки зображень. Нейромережеві методи є стійкішими за інші статистичні підходи під час розпізнавання зображень за наявності шумів [12, 13].

На сьогодні існує достатньо велика кількість систем та параметрів автоматизованого розпізнавання та класифікації зображень [12, 14–16]. Водночас для оцінювання поверхні прокату важливо створити стійкі та швидкі системи, які здатні

Контактна особа: П. О. МАРУЩАК, e-mail: Maruschak.tu.edu@gmail.com

класифікувати якомога ширшу лінійку відомих видів дефектів. Відомо, що у ГОСТ 21014-88 описано та класифіковано 64 види дефектів металопрокату [14].

Слід зазначити, що на сьогодні відома значна кількість праць [1–6, 16, 17], де запропоновано низку технічних рішень щодо алгоритмів класифікування дефектів із використанням нейронних мереж та реалізації їх у вигляді програмного забезпечення. Основна проблема – точність класифікування подібних за ознаками дефектів різних груп, яка потребує розуміння основних принципів їх утворення. Тому важливо запропонувати нові алгоритми обробки зображень для майбутніх дефектометричних систем, які були б дієвими навіть за низької яскравості/контрастності зображень, а також шумів технологічного походження.

Мета статті – розробити метод для класифікації дефектів на основі використання глибокої згорткової нейронної мережі.

Дефекти та їх класифікація. Відомо, що дефекти прокату є стандартизованими. Автоматизований аналіз дефектів смуг прокату дозволив визначити морфологічні ознаки та встановити технологічні причини появи дефектів (табл. 1).

Таблиця 1. Дефектоскопічні ознаки дефектів [14]

Клас	Ознаки дефекту	Тип дефекту
I	Одиничні та множинні невеликі плями округлої форми	Вкатана окалина (Shells)
II	Множинні лінії та подряпини, переважно в напрямку вальцювання, проте інколи можлива й інша орієнтація	Механічні риски (подряпини) у вигляді проковзувань (Grooves / Scratches)
III	Мають вигляд “виривів” поверхні з нерівними краями	Прокатні плівки (Rolling skin)

Слід зазначити, що дефекти навіть в межах одного класу можуть відрізнятися за формою, зовнішнім виглядом та структурою, що ускладнює їх класифікацію, а інколи дефекти одного класу схожі на дефекти іншого. Окремим чинником є зміна освітленості під час отримання зображення в різні періоди доби, в результаті чого об’єкт може виглядати по-різному.

Важливою перевагою нейронних мереж є їх високі узагальнювальні властивості, які дають змогу під час навчання виділяти сукупність рис, характерних для різних класів дефектів. Це робить їх найкращим засобом для обробки зображень, який автоматично виокремлює особливості шуканих дефектів за високої варіативності вхідних зображень одного і того ж класу. В таких випадках необхідно передбачити процедуру попередньої обробки зображень [18, 19].

Навчальні дані. На основі фотографій плоских сталевих поверхонь підготовано навчальну вибірку, яка містить 87704 зображення пошкоджень трьох видів, а також неушкоджених поверхонь (рис. 1). Їх розмір 256×256 pixels. Частину зображень взято з Kaggle competition “Severstal: Steel Defect Detection” [18]. Загалом навчальна вибірка містить 1820 зображень з пошкодженнями класу I, 14576 зображень – II та 2327 зображень – III. У навчальній базі присутні 63 зображення з пошкодженнями одночасно I і II класів та 228 з пошкодженнями II і III класів. Навчальна вибірка також містить 69272 зображення неушкоджених поверхонь, що становить 79% від її обсягу. Таким чином, вона суттєво незбалансована з точки зору розподілу зображень різних класів. Поверхня досліджуваних сталевих зразків має різні текстуру та відтінок, а самі зображення отримували за різної освітленості. Тому навчальна вибірка характеризується різноманітністю не тільки самих дефектів, а й поверхонь та ступенів їх освітлення.

Для збільшення інформативності навчальної вибірки використовували техніку аугментації: нарощування навчальної вибірки шляхом модифікації її окремих зразків. Такий підхід часто застосовують, оскільки він поліпшує результати, а на невеликих вибірках – зменшує перенавчання [19, 20]. Для цього генератор даних, який постачав їх для навчання моделі, з певною імовірністю відображав зображення відносно горизонтальної та вертикальної осей та повертав на кут, кратний 90°.

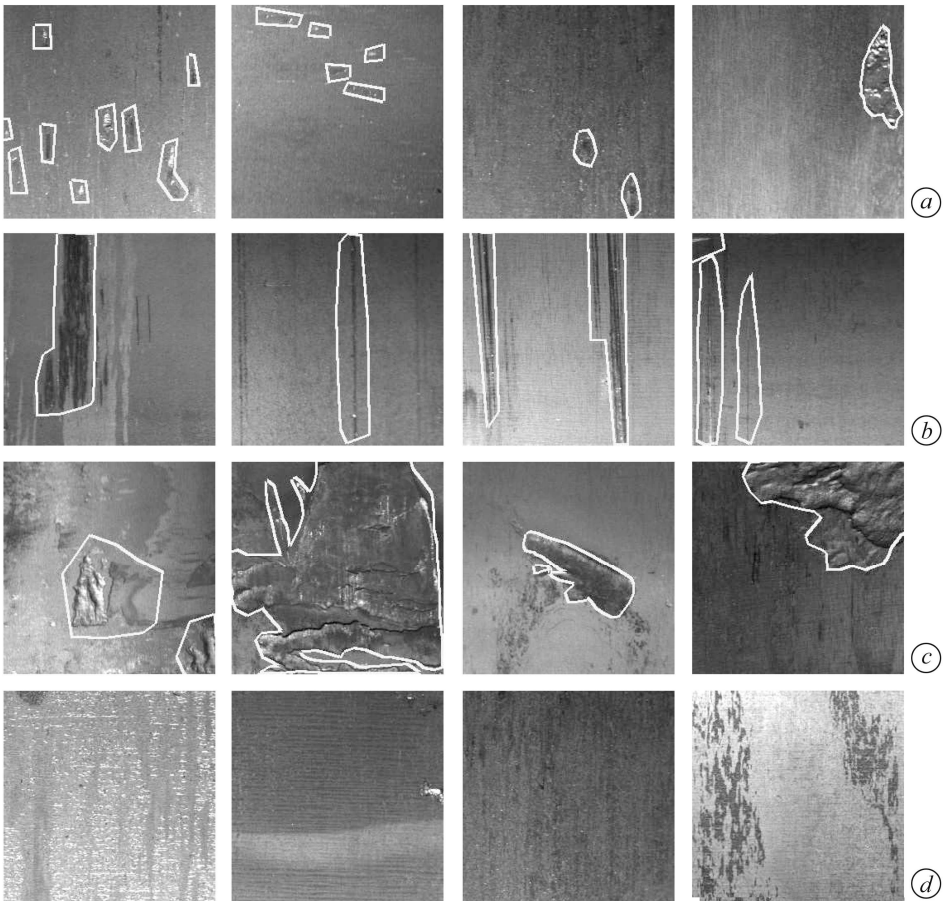

Рис. 1. Зображення з пошкодженнями класу I (*a*), II (*b*), III (*c*) та неушкоджених поверхонь (*d*).

Fig. 1. Images with class I (*a*), class II (*b*), class III (*c*) defects and intact surfaces (*d*).

Архітектура класифікатора. Як основу класифікатора використали залишкову нейронну мережу ResNet (рис. 2), запропоновану в 2015 році командою Microsoft Research [21]. Вона продемонструвала хороші узагальнювальні властивості з малими похибками для задач, пов’язаних з розпізнаванням зображень. Раніше дослідили залишкові нейронні мережі ResNet18, ResNet34, ResNet50, ResNet152, SeResNet50, SeResNet152. Мережі з глибиною до 50 шарів продемонстрували найгірші метрики продуктивності. Найкращого результату на тестовій вибірці досягнули за допомогою 50-шарової моделі ResNet.

Архітектура ResNet дає змогу навчати дуже глибокі нейромережі, які містять десятки і сотні шарів, і тому можуть формувати велику кількість ознак. Класифікатор на основі ResNet50 (рис. 2) складається з 50 послідовно з’єднаних шарів і загалом містить більше 23 мільйонів навчальних параметрів. Частина шарів об’єднана у блоки, які з’єднані “швидкими зв’язками” (shortcut connections). Така

будова здебільшого усуває проблему зникаючого градієнта під час навчання і дає змогу досягнути хороших результатів. “Швидкі зв’язки” передають вхідний сигнал відразу напряму до виходу залишкового блоку.

Нейромережа ResNet50 містить 4 каскади, кожен з яких складається із залишкових блоків (рис. 2) з трьома шарами, які виконують згортання фільтрами розміром 1×1 та 3×3. На вході моделі ResNet50 виконується згортання та макс-пулінг за допомогою ядер розміром 7×7 і 3×3 з кроком 2. Результат цієї операції подається на вхід першого каскаду, який складається з трьох залишкових блоків з трьома згортковими шарами. За переходу до кожного наступного каскаду ширина каналу подвоюється, а вхідний розмір зменшується у два рази. Криві лінії на рис. 2 відображають швидкі зв’язки.

У кожному блоці використано дизайн у вигляді “шийки пляшки”. У межах кожного блоку три згорткові шари 1×1, 3×3, 1×1 з’єднані послідовно. У кінці модель містить шар усереднювального пулінгу, до якого приєднано кінцевий шар з трьома нейронами, кожен з яких представляє один клас аналізованих дефектів.

Класифікатор реалізовано за допомогою бібліотек Keras та TensorFlow.

Навчання. Для навчання класифікатора використали техніку “перенесення знань” (transfer learning). За основу взяли ваги, натреновані на 1,4 млн зображень з бази ImageNet, які місять об’єкти 1000 класів. Застосування такої моделі суттєво пришвидшує процес та дає змогу навчати нейромережу на меншій множині даних.

Всі зображення розділили на тестову (20% від загальної кількості) та навчальну (решта 80%) групу. Під час тренування моделі для валідації використовували 20% навчальної вибірки. Пошкодження всіх класів подані у кожній групі пропорційно до своєї частки у загальній кількості.

Процес досліджували за оптимізаторів Adam, RMSprop та SGD. Швидкість навчання на початку була 0,001 та зменшувалась на 25% після кожних 8 епох, якщо функція втрат не поліпшувала своє значення. Розмір пакета під час навчання стано-

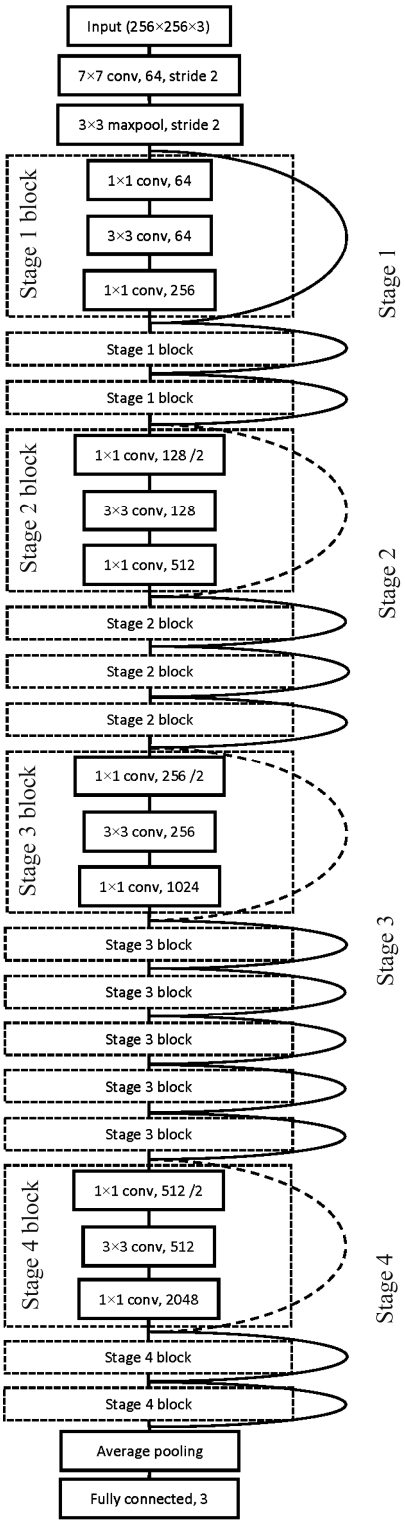

Рис. 2. Нейромережа ResNet50 з чотирма каскадами.

Fig. 2. ResNet50 neural network containing 4 cascades.

вив 16 та 20 зображень. Тренували доки функція втрат не покращувала своє значення упродовж 10 епох. У результаті навчання відбирали модель, для якої значення функції втрат на валідаційних даних найменше.

Навчали з використанням GPU NVIDIA GeForce GTX 1060 з 6 GB відеопам'яті. Встановили, що найкращий результат досягнуто за застосування фокальних функцій втрат, які добре зарекомендували себе саме для незбалансованих даних [22].

Натренували два види класифікаторів: багатомітковий (multilabel) та багатокласовий (multiclass). Багатомітковий класифікатор передбачає, що на одному зображенні можуть бути присутні пошкодження кількох класів, тоді як багатокласовий орієнтований на те, що кожне зображення представляє пошкодження лише одного класу. У розглянутому випадку тільки 0,3% зображень містять пошкодження відразу кількох класів, тому їх загальний внесок у похибку незначний. Вихід багатоміткового класифікатора містив 3 нейрони, кожен з яких сигналізував про присутність чи відсутність пошкодження відповідного класу. На виході багатокласового класифікатора було 4 нейрони (по одному на кожен клас пошкодження та один для неушкоджених поверхонь).

Під час тренування багатоміткових класифікаторів використовували функції втрат бінарної крос-ентропії та фокальних втрат. Для багатокласових варіантів застосовували втрати категоріальної крос-ентропії та фокальні. В кінці кожної епохи зберігали значення таких метрик: кількість хибно позитивних, хибно негативних, істинно позитивних та істинно негативних випадків, а також точність (accuracy), влучність (precision), повноту (recall) та площу під ROC-кривою.

Кожен з видів класифікаторів – багатокласові та багатоміткові – мають свої переваги та недоліки. Дослідили кожен з цих видів за різних гіперпараметрів. Метрики продуктивності також відрізнялися. Розглянемо кожен з цих видів окремо.

Багатоміткова модель. Багатомітковий класифікатор навчали за зміни таких гіперпараметрів, як швидкість навчання, розмір пакета та кількість кроків за епоху. Його якість оцінювали за бінарною точністю. Дійсне значення подане на його вихід під час навчання і має вигляд $y^{gt} = [y_0^{gt}, y_1^{gt}, \dots, y_i^{gt}, \dots, y_{n_{cl}-1}^{gt}]$, де n_{cl} – кількість класів (тут $n_{cl} = 3$); y_i^{gt} – елемент вихідного вектора, який дорівнює 1, якщо клас з індексом i є на вхідному зображенні, та 0, якщо ні.

Вихідний вектор з передбаченням $y^{pr} = [y_0^{pr}, y_1^{pr}, \dots, y_i^{pr}, \dots, y_{n_{cl}-1}^{pr}]$, де $y_0^{pr} \in [0, 1]$. Якщо для елемента цього вектора виконується умова $y_i^{pr} \geq t_b$ (t_b – деяке граничне значення), то клас з індексом i є на зображенні, і навпаки. Означимо східчасту функцію, яка описує таку умову:

$$T_{bin}(y_i^{pr}) = \begin{cases} 0, & y_i^{pr} < t_b \\ 1, & y_i^{pr} \geq t_b \end{cases}.$$

Запишемо також функцію рівності двох аргументів, яка повертає 1, якщо вони одинакові, та 0 у протилежному випадку:

$$Eq(x_1, x_2) = \begin{cases} 0, & x_1 \neq x_2 \\ 1, & x_1 = x_2 \end{cases}.$$

Тоді бінарна точність для одного окремого передбачення становить:

$$a_{bin}^1 = \frac{1}{n_{cl}} \sum_i Eq(T_{bin}(y_i^{pr}), y_i^{gt}),$$

а бінарна точність передбачення пакета вхідних зображень розміром n_b:

$$a_{bin} = \frac{1}{n_b} \sum_j a^1_{bin\,j} \,.$$

Кращого результату досягнуто на моделі, натренованій за допомогою оптимізатора SGD з моментом 0,9 за кількості 3000 кроків за епоху.

Багатокласова модель. Під час навчання багатокласової моделі змінювали такі ж гіперпараметри, як і у багатомітковій, а також ваги класів. Точність моделі оцінювали не лише за бінарною точністю, а й обчислювали категоріальну. Дійсне значення на виході багатокласового класифікатора має вигляд такий, як й для багатоміткового. Проте для кожного значення цього вектора лише один елемент y_i^{gt} може дорівнювати 1, що є індикатором присутності на зображенні об'єкта класу *i*. За таких умов категоріальна точність (categorical accuracy) становить:

$$a_{cat} = \frac{1}{n_b} \sum_j Eq(\arg\max(y^{pr}), \arg\max(y^{gt})) \,.$$

Найкращого результату досягнуто під час застосування оптимізатора SGD з моментом 0,9 за швидкості навчання 0,001, розміру пакета 16 та кількості кроків за епоху 2000. При цьому використовували ваги класів, розраховані на основі кількостей зображень певного класу у навчальній множині даних. Така модель навчалась суттєво довше за інші, але загалом продемонструвала найкращі показники точності (accuracy), повноти (recall) та влучності (precision). Зведені метрики якості найкращих розроблених моделей на основі нейромережі ResNet50 наведено у табл. 2.

Таблиця 2. Метрики якості класифікаторів

Модель	Бінарна точність	Категоріальна точність	Повнота	Влучність
Багатоміткова 1	0,9581	–	0,6765	0,8814
Багатоміткова 2	0,9552	–	0,6604	0,8517
Багатокласова 1	–	0,9482	0,5399	0,8737
Багатокласова 2	–	0,9571	0,6657	0,8934

Показано (рис. 3*a*, *b*) ROC-криві для двох кращих багатоміткових та багатокласових моделей. Інтегральний показник AUC-ROC для них становить 0,98…0,99. Найкраще ідентифікують пошкодження III класу (площа AUC-ROC дорівнює 0,98), які мають суттєві морфологічні відмінності порівняно з непошкодженими поверхнями. Добре розпізнаються також пошкодження II класу (площа AUC-ROC дорівнює 0,96). Найважче виявити дрібні пошкодження I класу, які часто зливаються з фоном (AUC-ROC найнижча – 0,89…0,90).

Загалом кращих показників вдалося досягти для багатоміткових моделей. Структурно така модель є комбінацією з чотирьох бінарних класифікаторів. На виході кожного із них використали власне граничне значення для прийняття рішення про наявність пошкодження певного класу на аналізованому зображенні. За влучністю та повнотою багатоміткова модель 1 краще виявляє пошкодження класу II, водночас як багатоміткова модель 2 – пошкодження класу III.

Класифікація поверхневих дефектів металопрокату за допомогою глибокої нейронної мережі забезпечила зв'язок морфологічних ознак дефектів з їх класами, щоб вивчити та систематизувати причини їх утворення [23]. Енкодер, використаний у класифікаторі для виявлення карт ознак, притаманних різним класам, у подальшому можна застосувати для побудови моделі сегментації, яка дасть

можливість обчислювати геометричні властивості знайдених дефектів (площа, координати тощо) [19, 20].

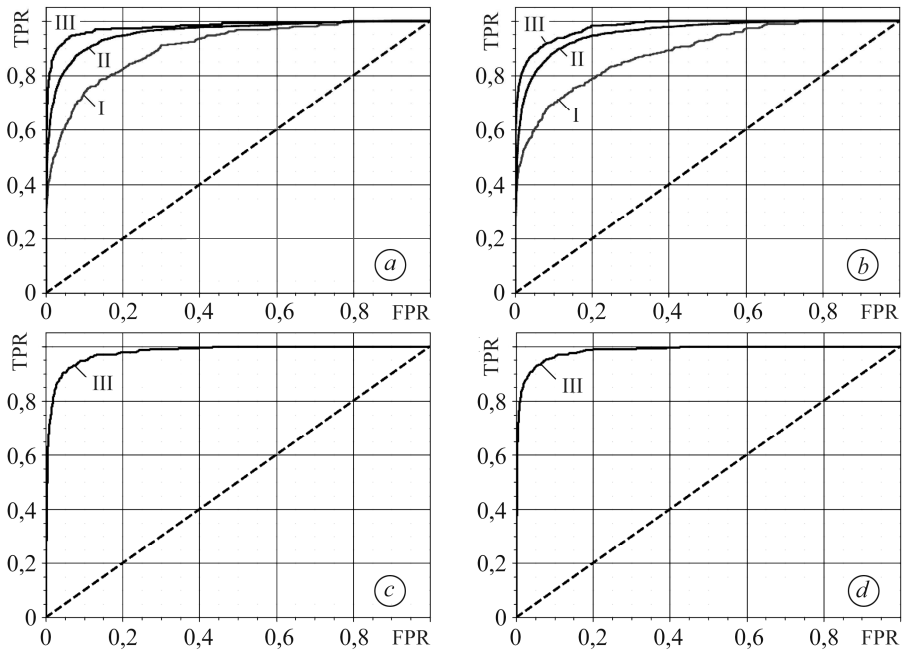

Рис. 3. ROC-криві багатоміткових (*a*, *b*) та багатокласових (*c*, *d*) моделей класифікаторів. Площа під ROC-кривою для класу I становить 0,89…0,90, II – 0,96, III – 0,98…0,99. TPR – істинно позитивний показник, FPR – хибно позитивний показник.

Fig. 3. ROC-curves of multilabel (*a*, *b*) and multiclass (*c*, *d*) classifier models. AUC-ROC for class I equals 0.89…0.90, II – 0.96, class III – 0.98…0.99. TPR – true positive rate, FPR – false positive rate.

ВИСНОВКИ

Виявлено основні закономірності класифікації поверхневих дефектів металопрокату за допомогою глибокої нейронної мережі зі застосування стандартизованих ознак. Розроблено та досліджено низку багатоміткових та багатокласових моделей на основі нейромережі ResNet50 для виявлення і класифікації поверхневих дефектів. Кращих результатів досягнуто для багатоміткових моделей. Оскільки область застосування класифікаторів характеризується суттєвою нерівномірністю розподілу поверхонь різних класів, під час навчання використали функцію фокальних втрат. Усереднена точність найкращої багатоміткової моделі за всіма видами поверхонь становить 0,9581. З багатокласових найкращий результат у моделі, для якої на під час тренування задавали ваги, які коригують розподіл зразків за класами. Виявлено, що найкраще розпізнаються пошкодження класів III та II з відносно великою площею. Для них також характерні значні морфологічні відмінності порівняно з неушкодженою поверхнею. Найважче розпізнаються дрібні пошкодження класу I, які найбільше схожі на поверхневі утворення, що часто зустрічаються на неушкоджених зразках. Розроблений автоматизований метод виявлення та класифікації трьох типів дефектів металопрокату забезпечує можливість дефектоскопії зі заданими параметрами ефективності та швидкодії. Одержані результати технічної діагностики і неруйнівних методів контролю є вихідною інформацією для коригування режимів вальцювання металевої смуги, що забезпечить зменшення кількості поверхневих дефектів та поліпшення якості металопрокату.

1. *Pernkopf F. and O'Leary P.* Image acquisition techniques for automatic visual inspection of metallic surfaces // NDT & E Int. – 2003. – 36(8). – P. 609–617.
2. *Song K. and Yan Y.* A noise robust method based on completed local binary patterns for hot-rolled. Part B // Appl. Surf. Sci. – 2013. – **285**. – P. 858–864.
3. *Semi-supervised* defect classification of steel surface based on multi-training and generative adversarial network / Y. He, K. Song, H. Dong, and Y. Yan // Opt. Lasers Eng. – 2019. – **122**. – P. 294–302.
4. *Surface* defect classification of steels with a new semi-supervised learning method / H. Di, X. Ke, Z. Peng, and Z. Dongdong // Opt. Lasers Eng. – 2019. – **117**. – P. 40–48.
5. *Song K., Hu S., and Yan Y.* Automatic recognition of surface defects on hot-rolled steel strip using scattering convolution network // J. Comput. Inf. Syst. – 2014. – **10**. – P. 3049–3055.
6. *Automatic* defect detectionon hot-rolled flat steel products / S. Ghorai, A. Mukherjee, M. Gangadaran, and P. K. Dutta // IEEE Transactions on Instrumentation and Measurement. – 2013. – **62**. – P. 612–621.
7. *Real-time* system for automatic cold strip surface defect detection / P. Kostenetskiy, R. Alkapov, N. Vetoshkin, R. Chulkevich, I. Napolskikh, and O. Poponin // FME Transact. – 2019. – 47(4). – P. 765–774.
8. *Steel* surface defects recognition based on multi-type statistical features and enhanced twin support vector machine / Maoxiang Chu, Rongfen Gong, Song Gao, and Jie Zhao // Chemometrics and Intelligent Laboratory Systems. – 2017. – **171**. – P. 140–150.
9. *Application* of RNAMlet to surface defect identification of steels / K. Xu, Y. Xu, P. Zhou, and L. Wang // Opt. Lasers Eng. – 2018. – **105**. – P. 110–117.
10. *Neogi N., Mohanta D. K., and Dutta P. K.* Review of vision-based steel surface inspection systems // EURASIP J. on Image and Video Proc. – 2014. – 50(1). – P. 1–19.
11. *Liu Y., Xu K., and Xu J.* An improved MB-LBP defect recognition approach for the surface of steel plates // Appl. Sci. – 2019. – **9**. – P. 4222. DOI: https://doi.org/10.3390/app9204222
12. *Li Yi, Guangyao Li, and Mingming Jiang* An end-to-end steel strip surface defects recognition system based on convolutional neural networks // Steel Research Int. – 2017. – 88(2). – P. 60–68.
13. *Automatic* metallic surface defect detection and recognition with convolutional neural networks / Xian Tao, Dapeng Zhang, Wenzhi Ma, Xilong Liu, and De Xu // Applied Sciences. – 2018. – 8(9). – P. 1575. DOI:10.3390/app8091575.
14. *ГОСТ 21014-88* Прокат черных металлов. Термины и определения дефектов поверхности. – М.: Изд-во стандартов, 1989. – 61 с.
15. *Коноваленко І. В., Марущак П. О.* Нові оптико-цифрові методи технічного діагностування матеріалів і конструкцій. – Тернопіль: Тернопільськ. нац. техн. ун-т ім. Івана Пулюя, 2018. – 201 с.
16. *Kapustiy B. E., Rusyn B. P., Tayanov V. A.* Peculiarities of application of statistical detection criteria for problems of pattern recognition // J. of Automation and Information Sci. – 2005. – 37(2). – P. 30–36.
17. *Image* segmentation based on the evaluation of the tendensy of image elements to form cluster with the help of point field characteristics / R. J. Kosarevych, B. P. Rusyn, V. V. Korniy, and T. I. Kerod // Cybernetics and Systems Analysis. – 2005. – 51(5). – P. 704–713.
18. *Kaggle* "Severstal: Steel Defect Detection. Can you detect and classify defects in steel?". – 2019. – https://www.kaggle.com/c/severstal-steel-defect-detection.
19. *Improved* relation classification by deep recurrent neural networks with data augmentation / Y. Xu, R. Jia, L. Mou, G. Li, Y. Chen, Y. Lu, Z. Jin. – 2016. – arXiv:1601.03651
20. *Deep* learning for atmospheric cloud image segmemtation / B. Rusyn, V. Korniy, O. Lutsyk, and R. Kosarevych // Proc. of the 11th Int. Scientific and Tech. Conf. on electronics and information technologies (September 16–18, Lviv). – Lviv, 2019 – P. 188–191. DOI:10.1109/ELIT.2019.8892285.
21. *Deep* residual learning for image recognition / K. He, X. Zhang, S. Ren and J. Sun // Proc. of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). – Las Vegas, NV, 2016. – P. 770-778, doi:10.1109/CVPR.2016.90.
22. *Focal* loss for dense object detection / T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar // Proc. of the IEEE Int. Conf. on Computer Vision (22–29 Oct. 2017). – 2017, doi:10.1109/ICCV.2017.324.
23. *Мазур В. Л., Добронравов А. И., Чернов П. П.* Предупреждение дефектов листового проката. – К.: Техника, 1986. – 141 с.

Одержано 21.02.2020