

**ВИКОРИСТАННЯ АРХІТЕКТУР ГЛИБОКИХ  
ЗГОРТКОВИХ НЕЙРОННИХ МЕРЕЖ ТА TENSORFLOW  
У ЗАДАЧІ ВИЗНАЧЕННЯ БДЖІЛ НА ЗОБРАЖЕННІ**

*Анотація. Бджоли є важливим елементом сільськогосподарської промисловості, але через зменшення їх популяції росте необхідність відстежування стану бджолиних сімей, наприклад, за допомогою фото- та відеоматеріалів. У цій публікації наведено аналіз ефективності роботи архітектур глибоких згорткових нейронних мереж та фреймворку Tensorflow на задачі визначення бджіл на фотографіях. Для дослідження було використано п'ять різних моделей з відкритого репозиторія TensorFlow 2 Detection Model Zoo, що були треновані на анованому вручну датасеті з фотографій бджіл. З тренованих моделей найкращий результат 80% визначення бджіл та їх частин показала архітектура SSD MobileNet.*

*Ключові слова: нейронні мережі, бджоли, глибоке навчання, технології комп'ютерного зору, tensorflow, згорткові нейронні мережі, ідентифікація комах, ідентифікація бджіл, виявлення об'єктів.*

**Постановка проблеми.** Бджоли є важливими запилювачами різних видів рослин і відіграють важливу роль у підтримці біорізноманіття нашої планети, надаючи приблизно 14% запилювань у сільськогосподарському виробництві.[1] Але останніми роками популяція бджіл скорочується через різні фактори, зокрема через хвороби та паразитів.[2]

Через скорочення числа бджіл по всьому світу виникла потреба в моніторингу стану бджіл та бджолиних вуликів. Традиційні методи ідентифікації їх стану базуються на візуальному спостереженні бджолярем, що може зайняти багато часу, та такі потребують спеціалізованих знань від людини.

З розвитком глибокого навчання та технологій комп'ютерного зору, йде зростання інтересів до можливостей їх використання для автоматизації спостереження за складними біологічними системами, зокрема за бджолиними сім'ями.

Важливим плюсом використання нейронних мереж та технологій комп'ютерного зору є їх швидкість роботи та масштабованість. Системи, що використовують глибоке навчання, можуть бути точнішими у задачах підрахунку бджіл у вулику, а також їм не потрібна обов'язкова присутність людини для отримання результатів. На відміну від традиційних методів спостереження за бджолиними сім'ями, нейронні мережі можуть обробляти великі об'єми даних швидше та точніше, що робить їх потужним інструментом для дослідження життя бджолиних сімей у вуликах.

Ідентифікація бджіл на зображенні – це важливий перший крок, що дозволить використовувати результати для аналізу кількості бджіл та їх здоров'я на фото та відео.

**Аналіз досліджень і публікацій.** Як писалося вище, з ростом зацікавленості до застосування глибокого навчання та комп'ютерного зору, почали з'являтися дослідження на тему їх використання в області ідентифікації стану складних біологічних систем. Наприклад, були проведені дослідження ідентифікації метеликів[3], ос[4], павуків[5], та мух[6]. Також, були проведені дослідження використання цих технологій для контролю паразитів[7][8]. Всі ці дослідження успішно використали методи машинного навчання або комп'ютерного зору для ідентифікації об'єктів.

Схожі з поточною задачею ставились у статті з ідентифікації комах[9]. В ній автори теж використовували архітектури згорткових нейронних мереж, зокрема Faster R-CNN, SSD MobileNet та SSD Inception для ідентифікації та класифікації комах на фото.

Якщо розглядати використання нейронних мереж для визначення бджіл на зображенні, то існує дослідження[10], де згорткова нейронна мережа використовується у зв'язці з інфрачервоною камерою та роботом для визначення щільності бджіл на окремій зоні вулика.

Також якщо розглядати використання комп'ютерного зору для обчислення кількості бджіл, то автори цієї статті[11] за допомогою камери, мікрокомп'ютера Raspberry PI та бібліотеки OpenCV розробили систему для обчислення кількості бджіл на контрольованій ділянці зі статичним фоном поза вуликом.

З іншого боку, якщо розглядати безпосередньо дослідження що поєднують використання комп'ютерного зору та нейронних мереж для ідентифікації бджіл, то в цьому випадку [12] автори відстежували використання архітектури Faster R-CNN для визначення Мелопонінових бджіл на зображенні. В цьому до-

слідженні вони досягли результату точності визначення бджіл на зображенні у 74%.

**Мета і завдання досліджень.** Метою роботи є аналіз та визначення архітектури згорткової нейронної мережі, що найкраще підходить для задачі визначення бджіл на фотографіях.

**Матеріали, методи та методики досліджень.** У якості джерела даних для дослідження використовувалися фотографії, зроблені на камеру смартфона Samsung A20. Кожну фотографію розміром 4128x3096 пікселів без попередньої обробки було розділено на 9 рівних частин розміром 1376x1032 для полегшення процесу анотації та подальшої роботи з ними. Ці частини були використані для анотації вручну за допомогою програми Remo. Після процесу анотації було створено фінальний датасет шляхом вирізання анотованих зон з частин фотографій. Таким чином, кожне зображення містить окрему бджолу або її частину (якщо на фото була доступна тільки голова чи спина), у Таблиці 1 наведено приклади таких зображень.

Таблиця 1

Анотовані зображення бджіл



За підсумком, фінальний датасет містить 1246 анотованих зображень бджіл. З нього формуються дві вибірки: навчальна та тестова, що відповідно містять 1004 та 252 зображення.

Для навчання було обрано 5 моделей з відкритого репозиторія TensorFlow 2 Detection Model Zoo, детальна інформація про них міститься у Таблиці 2.

Для роботи з цими архітектурами, з датасету було згенеровано файл формату TFRecord, що далі буде використовуватися для їх навчання. Для порівняння результатів роботи вибраних архітектур згорткових нейронних мереж на задачі ідентифікації бджіл та подальшої побудови матриці невідповідностей використовуватимуться наступні метрики: істинно позитивні результати, хибно негативні, чутливість, та точність.

Інформаційний аналіз архітектур глибоких згорткових нейронних мереж

№	Архітектура	Стислий опис
1.	Faster R-CNN ResNet152 V1	<p>Faster R-CNN ResNet152 – це архітектура глибокої згорткової нейронної мережі, що була створена для роботи з задачами по виявленню об’єктів. Вперше така архітектура була представлена у 2016 році як розширення існуючого фреймворку Faster R-CNN, що являє собою конвеєр, який складається з сітки запропонованих зон (region proposal network, RPN) та базованої на зонах згорткової нейронної мережі (region-based convolutional neural network, RCNN)</p> <p>Архітектура ResNet була представлена у 2015 році як вирішення проблеми зникання градієнта у дуже глибоких нейронних мережах. ResNet містить залишкові з’єднання, які забезпечують потік інформації через мережу, навіть коли мережа стає дуже глибокою. Ця архітектура досягла найкращих результатів на різних задачах з розпізнавання зображень, включаючи ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2015, де вона перевершила інші моделі та здобула перше місце у завданні виявлення об’єктів [13].</p> <p>Успіх ResNet привів до її широкого використання у застосунках з комп’ютерного зору та стимулював подальше дослідження покращень архітектур глибоких нейронних мереж.</p> <p>У випадку Faster R-CNN ResNet152, ця архітектура містить 152 ResNet блоки.</p>
2.	Faster R-CNN Inception ResNet V2	<p>Faster R-CNN Inception ResNet V2 є глибокою нейронною мережею для виявлення об’єктів на зображеннях, яка поєднує в собі дві популярні архітектури – Faster R-CNN та Inception ResNet V2. Остання містить багато блоків, які називаються Inception ResNet блоками. Кожен блок містить різні типи згорток та пулінгів, що дозволяє збільшити кількість функцій, що можуть бути вивчені нейронною мережею. Крім того, блоки Inception ResNet містять також зворотні зв’язки, що допомагає запобігти проблемі зникаючого градієнта під час навчання глибоких нейронних мереж та покращити точність</p>

		нейронної мережі [14].
3.	SSD MobileNet V2 FPNLite	<p>Ця модель поєднує у собі дві архітектури – SSD та MobileNet. Основна ідея за архітектурою SSD полягає в тому, що мережа використовує лише один прохід по зображенню для виявлення об'єктів та їх класифікації, тобто вона не потребує попереднього відокремлення регіонів зображення, як у інших методах об'єктного виявлення, наприклад, R-CNN. Ключовою особливістю цієї архітектури є використання багатомасштабних виходів згорткових рамок, приєднаних до набору масивів ознак у верхній частині нейронної мережі. Такий спосіб дозволяє ефективно змоделювати можливі місця рамок [15].</p> <p>MobileNet V2 – це нейронна мережа глибокого навчання, спроектована для роботи на мобільних пристроях та забезпечення високої точності класифікації зображень при мінімальних розмірі моделі та кількості операцій. Основна ідея MobileNet V2 полягає в розробці легкої та швидкої архітектури нейронної мережі за допомогою двох ключових технологій: residual connections та inverted residuals. Residual connections – це технологія, яка дозволяє передавати інформацію з попередніх шарів безпосередньо до наступних, що допомагає запобігти проблемі вивчення неглибоких функцій та забезпечити більш ефективне навчання нейронної мережі. Inverted residuals – це технологія, що дозволяє зменшити кількість параметрів мережі, зменшивши розмір шарів та застосовуючи їхнє збільшення з використанням 1x1 згортки та лінійних шарів [16].</p>
4.	SSD ResNet50 V1 FPN	<p>SSD ResNet50 – це глибока нейронна мережа, яка поєднує в собі дві архітектури: ResNet50 та Single Shot Detector (SSD). ResNet50 – це архітектура на базі ResNet з глибиною в 50 шарів, яка має високу точність класифікації зображень та виявлення об'єктів.</p>
5.	CenterNet Resnet50 V2	<p>CenterNet – це архітектура глибокої згорткової нейронної мережі, створена для задач з виявлення об'єктів. Це одноступінний детектор об'єктів, який використовує концепцію ключових точок та центроїдів для виявлення об'єктів. За-</p>

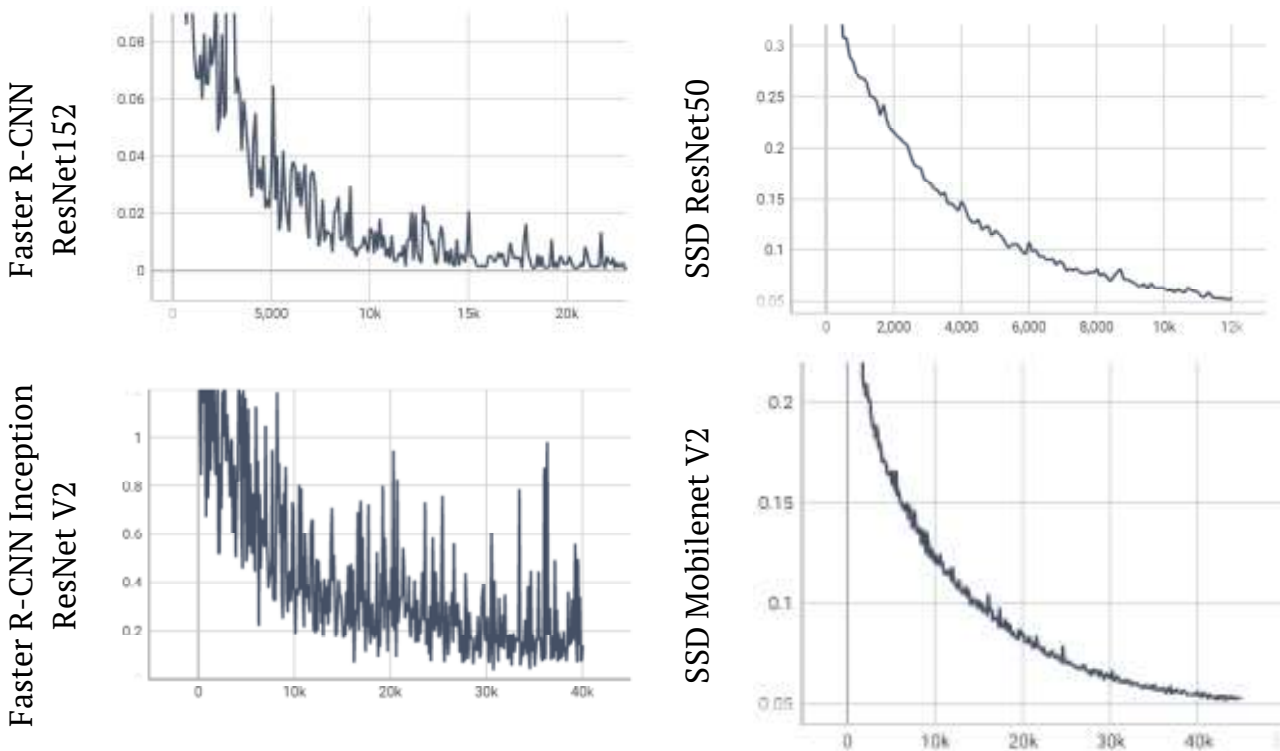
		<p>мість того, щоб використовувати детектор областей для локалізації об'єкта, CenterNet Resnet50 V2 використовує центроїди, які визначають центральну точку об'єкта. Це дозволяє мережі більш точно та ефективно визначати розміри об'єкта та його позицію на зображенні [17].</p>
--	--	--

**Результати досліджень.** Виконано порівняння метрик результатів роботи 5 різних архітектур глибоких згорткових нейронних мереж на задачі ідентифікації бджіл. Однією з основних метрик під час навчання нейронних мереж є значення функції втрат (Loss function). Функція втрат дозволяє виміряти відстані або різницю між прогнозованим виходом моделі та основною істиною навченої моделі.

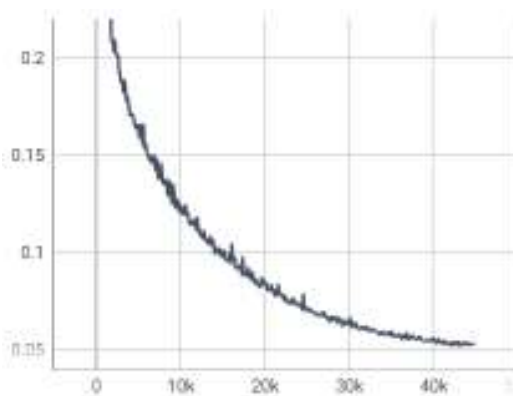
У таблиці 3 наведено прогрес навчання нейронних мереж та значення функції втрат на відповідних епохах.

Таблиця 3

Візуалізація значення функції втрат під час навчання моделей нейронних мереж



CenterNet Resnet50 V1



Модель R-CNN ResNet152 V1 спочатку навчалася 45 000 епох, проте після 25 000 епохи модель не показувала зниження значення функції втрат. Щоб уникнути присутності фактору перетренування в результатах, далі використовуватиметься чек-поінт з епохи 25 000. Основні гіперпараметри, що використовувались для тренування моделей, значення функції втрат та витрачений на навчання час, наведені в таблиці 4.

Для тестування результатів навчання моделей було використано 28 фотографій груп бджіл на сотах, що не входили до складу початкового датасету. За попереднім підрахунком, загальна кількість бджіл на цих фотографіях дорівнює 775. Результати аналізу у вигляді матриці невідповідності наведені у таблиці 5.

Таблиця 4

Гіперпараметри та підсумкове значення функції втрат моделей

Архітектура	Розмір групи	Кількість епох	Функція втрат	Час тренування
Faster R-CNN ResNet152	8	25 000	0.0024	2 години
Faster R-CNN Inception ResNet V2	32	40 000	0.1367	2 години
SSD ResNet50	8	40 000	0.0533	6 годин 40 хвилин
SSD Mobilenet V2	64	45 000	0.0528	4 години 20 хвилин
CenterNet Resnet50 V1	48	35 000	0.1685	5 годин 07 хвилин
















Матриця невідповідностей та середній час на аналіз тестових файлів

Архітектура	Істинно позитивних	Хибно негативних	Відсоток ідентифікованих	Точність	Середній час, мс
Faster R-CNN ResNet152	477	2	61.55	0.995	401
Faster R-CNN Inception ResNet V2	412	12	53.16	0.971	541
SSD ResNet50	474	4	61.16	0.991	695
SSD Mobilenet V2	620	27	80.00	0.958	675
CenterNet Resnet50 V1	532	5	68.64	0.990	662

Виходячи з наведених у матриці невідповідності результатів, SSD Mobilenet V2 показала найкращу продуктивність з результатом коректного визначення 80% бджіл, але також мало найгірший показник хибно негативних визначень, що у 2.7 разу вищий за середнє значення серед досліджених моделей. З наведених у таблиці 6 візуалізацій, ми можемо побачити, що такий поганий результат відбувся через те, що модель виділяла обмежувальною рамкою краї файлів, пусті місця, а також повторно відмічала ідентифікованих бджіл.

Також є сенс відмітити результати Faster R-CNN ResNet152 V1. Хоч ця модель і не знаходила дрібні частини бджіл і через це мала малий відсоток у підрахунку бджіл на фото, показала добрі результати у визначенні бджіл повністю.

Візуалізація тестових зображень з обмежувальними рамками

Архітектура	Зображення 1	Зображення 2	Зображення 3
Faster R-CNN ResNet152			
Faster R-CNN Inception ResNet V2			
SSD ResNet50			
SSD Mobilenet V2			
CenterNet Resnet50 V1			

**Висновки.** За результатами виконаних досліджень встановлено що мета-архітектура SSD MobileNet V2 показує найкращий результат у 80% у задачі ідентифікації бджіл на фотографії. Хоч Faster R-CNN ResNet152 показала найкращі результати у швидкості та точності ідентифікації бджіл на тестовій вибірці фо-

то, з її допомогою вдалось успішно визначити на 20% менше бджіл на фотографіях.

#### ЛІТЕРАТУРА

1. Garibalidi L. Wild pollinators enhance fruit set of crops regardless of honey bee abundance // Science. 2013., вип. 6127 Т. 339. С. 1608-1611
2. Worldwide occurrence records suggest a global decline in bee species richness // One Earth. 2021., вип. 1 Т. 4. С. 114-123.
3. The identification of butterfly families using content-based image retrieval // Biosystems Engineering. 2012., вип. 1 Т. 111. С. 24-32.
4. Weeks P.J.D., O'Neill M.A., Gaston K.J, Gauld I.D Species identification of wasps using principle component associative memories // Image and Vision Computing. 1999., вип. 12 Т. 17.
5. Tiscay-rivas J., Marcos del Pozo-Banos, Eberhard W, Alonso J Spider specie identification and verification based on pattern recognition of it cobweb // Expert Systems with Applications. 2013., вип. 10 Т. 40. С. 4213-4225.
6. Automatic identification of fruit flies (Diptera: Tephritidae) / Faria F. та ін. // Journal of Visual Communication and Image Representation. 2014., вип. 7 Т. 25. С. 1516-1527.
7. Wen C., Guyer D. Image-based orchard insect automated identification and classification method // Computers and Electronics in Agriculture. 2012.№ 89. С. 110-115.
8. Wen C., Guyer D., Li W. Local feature-based identification and classification for orchard insects // Biosystems Engineering. 2009., вип. 3 Т. 104. С. 299-307.
9. Patel D., Bhatt N. Insect Identification Among Deep Learning's Meta-architectures Using TensorFlow // International Journal of Engineering and Advanced Technology. 2019., вип. 1 Т. 9. С. 1910-1914.
10. Convolutional Neural Network for Honeybee Density Estimation / Luneckas T. та ін. // 2020 IEEE Symposium Series on Computational Intelligence (SSCI). 2020. С. 2558-2566.
11. Automatic behaviour analysis system for honeybees using computer vision / Jun tu G. та ін. // Computers and Electronics in Agriculture. 2016.№ 112. С. 10-18.
12. Aziah A. , Nizam A. , Mohd-isa W. Image Segmentation of Meliponine Bee using Faster R-CNN // 2019 Third World Conference on Smart Trends in Systems Security and Sustainability (WorldS4). 2019., С. 235-238.
13. Deep Residual Learning for Image Recognition / He K. та ін. // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016., С. 770-778.
14. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning / Szegedy S. та ін. // Proceedings of the AAAI Conference on Artificial Intelligence. 2016., вип. 1 Т. 31. С. 1-12.

15. SSD: Single Shot MultiBox Detector / Liu W. та ін. // Proceedings of the European Conference on Computer Vision (ECCV) (2016). 2016., С. 1-17.
16. MobileNetV2: Inverted Residuals and Linear Bottlenecks / Sandler M. та ін. // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018., С. 4510-4520.
17. Zhou X., Wang D., Krähenbühl P. Objects as points //arXiv preprint arXiv:1904.07850. 2019. S. 1-12

#### REFERENCES

1. Garibalidi L. Wild pollinators enhance fruit set of crops regardless of honey bee abundance // Science. 2013., vip. 6127 T. 339. S. 1608-1611.
2. Zattara E. Worldwide occurrence records suggest a global decline in bee species richness // One Earth. 2021., vip. 1 T. 4. S. 114-123.
3. Wang J. The identification of butterfly families using content-based image retrieval // Biosystems Engineering. 2012., vip. 1 T. 111. S. 24-32.
4. Weeks P.J.D., O'Neill M.A., Gaston K.J, Gauld I.D Species identification of wasps using principle component associative memories // Image and Vision Computing. 1999., vip. 12 T. 17. S. 861-866.
5. Ticay-rivas J., Marcos del Pozo-Banos, Eberhard W, Alonso J Spider specie identification and verification based on pattern recognition of it cobweb // Expert Systems with Applications. 2013., vip. 10 T. 40. S. 4213-4225.
6. Automatic identification of fruit flies (Diptera: Tephritidae) / Faria F. та ін. // Journal of Visual Communication and Image Representation. 2014., vip. 7 T. 25. S. 1516-1527.
7. Wen C. , Guyer D. Image-based orchard insect automated identification and classification method // Computers and Electronics in Agriculture. 2012.# 89. S. 110-115.
8. Wen C. , Guyer D. , Li W. Local feature-based identification and classification for orchard insects // Biosystems Engineering. 2009., vip. 3 T. 104. S. 299-307.
9. Patel D. , Bhatt N. Insect Identification Among Deep Learning's Meta-architectures Using TensorFlow // International Journal of Engineering and Advanced Technology. 2019., vip. 1 T. 9. S. 1910-1914.
10. Convolutional Neural Network for Honeybee Density Estimation / Luneckas T. та ін. // 2020 IEEE Symposium Series on Computational Intelligence (SSCI). 2020. S. 2558-2566.
11. Automatic behaviour analysis system for honeybees using computer vision / Jun tu G. та ін. // Computers and Electronics in Agriculture. 2016.# 112. S. 10-18.
12. Aziah A. , Nizam A. , Mohd-isa W. Image Segmentation of Meliponine Bee using Faster R-CNN // 2019 Third World Conference on Smart Trends in Systems Security and Sustainability (WorldS4). 2019., S. 235-238.

13. Deep Residual Learning for Image Recognition / He K. ta in. // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016., S. 770-778.
14. Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning / Szegedy C. ta in. // Proceedings of the AAAI Conference on Artificial Intelligence. 2016., vip. 1 T. 31. S. 1-12.
15. SSD: Single Shot MultiBox Detector / Liu W. ta in. // Proceedings of the European Conference on Computer Vision (ECCV) (2016). 2016., S. 1-17.
16. MobileNetV2: Inverted Residuals and Linear Bottlenecks / Sandler M. ta in. // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2018., S. 4510-4520.
17. Zhou X., Wang D., Krähenbühl P. Objects as points //arXiv preprint arXiv:1904.07850. 2019. S. 1-12

Received 01.03.2023.

Accepted 03.03.2023.

### ***Using deep cnn architectures and tensorflow for bee identification on images***

*Problem statement. Bees are important pollinators for various plant species and are essential in maintaining our planet's biodiversity. With the decline in the number of bees worldwide, monitoring the bee hive state has become increasingly important. Purpose. Individual bee identification is an important task that, if performed well, will help to track the bee hive population and its health state in general. Related work. With the development of deep learning and computer vision technologies, several studies were done on their application in detecting insects like wasps, spiders, or fruit flies. Materials and methods. Photographs of bee hive frames with bees on them are being used as a data source for model training. These photographs were manually annotated using Remo annotation software. This study uses five widely used deep learning architectures (Faster R-CNN ResNet152, CenterNet ResNet50, SSD ResNet50, Faster R-CNN Inception ResNet V2, SSD MobileNet V2) for the purpose of bee detection on test images. These models were trained on the same dataset and evaluated on the same bee image set. Results and discussion. According to the confusion matrix, SSD MobileNet V2 architecture showed the best detection performance with 80% detected bees on test files, but it also had the falsest negative entries. On the other hand, Faster R-CNN ResNet152 model showed the best results in accuracy and evaluation speed, but successfully identified only 61% of the bees. Conclusions. The experiment showed that state-of-the-art CNN architecture SSD MobileNet V2 is a better performer with detecting 80% and 95% accuracy on bee images test set.*

**Жуков Олександр Олексійович** – аспірант кафедри програмної інженерії математичного факультету Запорізького національного університету.

**Горбенко Віталій Іванович** – к.ф.м.н, доцент, доцент кафедри Програмної інженерії математичного факультету Запорізького національного університету.

**Zhukov Oleksandr Oleksiiovych** – postgraduate student of the department of program engineering, faculty of Mathematics, Zaporizhzhia National University, Ukraine.

**Horbenko Vitalii Ivanovych** - candidate of physical and mathematical sciences, associate professor, associate professor department of program engineering, faculty of Mathematics, Zaporizhzhia National University, Ukraine.