

**Грабар О.І.**

Державний університет «Житомирська політехніка»

**Лисогор Ю.І.**

Державний університет «Житомирська політехніка»

**Скачков В.О.**

Державний університет «Житомирська політехніка»

## АНАЛІЗ АЛГОРИТМІВ ВІЗУАЛЬНОЇ СИСТЕМИ РОЗПІЗНАВАННЯ СОРТІВ РОСЛИН ДЛЯ СТВОРЕННЯ ВЕБСЕРВІСУ

Основна мета полягає у проєктуванні, розробленні й створенні вебсервісу для розпізнавання рослин за зображенням. Високоякісне розпізнавання зображень рослин – складне завдання для комп'ютера через нерівномірний вигляд і складну структуру рослин. Розглядається стан сучасного машинного навчання та задачі розпізнавання від ідентифікації конкретних органів рослини до загального розпізнавання всієї рослини «в дикій природі». З боку машинного навчання, розпізнавання рослин є дрібнозернистим класифікаційним завданням із високою мінливістю між класами й часто малими міжкласовими відмінностями, які пов'язані з таксономічною ієрархічною класифікацією. Для опису текстури незалежно від розміру візерунка й орієнтації на зображенні необхідно мати інваріантний опис для обертання та масштабу. Для практичного застосування також потрібне ефективне обчислення. Нижче наведено декілька методів для розпізнавання текстур. Показано локальне подвійне зображення та гістограма Фур'є, що описується на основі локальних двійкових шаблонів. Функції Фур'є та гістограми локальних двійкових шаблонів (LBP) можна узагальнити на будь-який набір рівномірних візерунків. У першому випадку використовується опис LBP-HF-SM, де для побудови дескриптора розраховуються функції Фур'є-гістограми як знаків, так і величини-LBP. Величина-LBP перевіряє, чи величина різниці сусіднього пікселя відносно центрального пікселя перевищує порогове значення  $t_p$ . У глибоких навчальних завданнях є поширеною практикою навчання декількох мереж на різних підмножинах навчальних даних. Об'єднання таких мереж зазвичай відбувається за допомогою простого механізму й прагне перевершити індивідуальні мережі. Визначення видів рослин із фотографій із використанням текстурного визнання за допомогою сучасних методів дає можливість досягти досить хороших результатів, водночас зберігаючи обчислювальні вимоги, що робить його придатним для обробки в реальному часі для створення вебсервісу.

**Ключові слова:** машинне навчання, розпізнавання, вебсервіс, полярні перетворення Фур'є, метадані, нейронна мережа.

**Постановка проблеми.** Інтерес до методів візуальної класифікації рослин стрімко виріс в останній час. Це пов'язано з глобальним розповсюдженням пристроїв, оснащених камерами, й розвитку вебсервісів у глобальній мережі. Саме розпізнавання рослин розглядалось, майже без винятків, як розпізнавання фотографій, що зображують виключно певний організм рослини, такий як квітка, кора, плід, лист або їхня комбінація.

Розпізнавання листів стало найпопулярнішим підходом до розпізнавання рослин, і в літературі зазначено широкий спектр таких досліджень. Розпізнавання листя зазвичай стосується лише визнання широких листів, голки розглядаються окремо. Для опису листів було запропоновано декілька прийомів, які часто ґрунтуються на об'єднаних

рисах різного характеру (особливості форми, особливості кольорів тощо).

Одним із хороших алгоритмів для розпізнавання листів є SIFT (Scale-invariant feature transform), методи геометричних особливостей, моментних інваріантів, моментів зерен і полярних перетворень Фур'є.

**Аналіз останніх досліджень і публікацій** у міжнародній літературі показав, що одним із найкращих із практичних рішень за останні роки являються дослідження Н. Кумара й співавторів [1], які описали алгоритми та розробили Leafsnap – комп'ютерну систему зору для автоматичної ідентифікації видів рослин, що була розроблена за допомогою більш ранньої системи ідентифікації рослин від Agarwaletal [3] та Belhumeural

[2]. Авторами [1] було представлено попередній фільтр на входних зображеннях, численні прискорення та додаткові постобробки в алгоритмі сегментації, використання простіших й ефективніших алгоритмів розпізнавання на основі кривини. По базі даних Leafsnap зі 184 порід дерев система розпізнавання знаходить правильні збіги серед перших 5 результатів для 96,8% запитів із набору даних.

Інша група вчених – С.Г. Ву й співавтори [4] – запропонувала імовірнісну нейронну мережу для розпізнавання листів із використанням 12 цифрових морфологічних ознак, отриманих із 5 основних рис: діаметр, фізіологічна довжина, фізіологічна ширина, площа листа, периметр листа. Автори зібрали загальнодоступну базу рослинних листів «Flavia» [4].

Розпізнавання рослин, що дають декілька зображень різних типів змісту, таких як різні органи рослини чи рослина, в принципі, повинні бути більш надійними, ніж розпізнавання лише одним зображенням одного конкретного органу рослини, такого як лист або кора. Однак завдання складніше, якщо надається зображення не визначеного органу рослини.

Існує проблема відкритого розпізнавання, тобто проблема, в якій система розпізнавання повинна бути надійною для відомих і невідомих категорій. Кожне зображення в задачі належить до одного з 7 типів контенту: лист, сканування листів, квітка, фрукти, стебло або ціла рослина. Хоча тип змісту доступний у метаданих, найкращі результати розрахунку використовують однакові глибинні мережі для всіх типів змісту [5–7]. Z. Ge й співавтори [8] показали, що в цьому завданні загальні функції ConvolutionalNeuralNetwork (далі – CNN) виконують краще, ніж контент специфічні функції CNN, і що їхнє поєднання підвищує точність. S. Choi [7] показав, що мікшування кількох універсальних CNN-підписів покращує точність, а також виграє ChallengePlantCLEF 2015.

Pl@ntNet [9] – це інша система розпізнавання рослин на основі змісту. Це спільна інформаційна система, що забезпечує програму для обміну фотографіями й пошуку для ідентифікації установок. Вона була розроблена вченими чотирьох французьких дослідницьких організацій (Cirad, INRA, INRIA й IRD) і мережі TelaBotanica. База даних дерева ідентифікується шляхом поєднання інформації із зображеннями середовища існування, квітки, фруктів, листя та кори. Точні алгоритми, що використовуються в вебслужбі визначення Pl@ntNet [10], не задокументовані публічно. Існує

також мобільний додаток Pl@ntNet [11], програма для обміну й отримання зображень для ідентифікації рослин.

Текстурна інформація є важливою ознакою для розпізнавання багатьох органів рослин. Текстурний аналіз є усталеною проблемою з великою кількістю наявних методів, багато з яких були описані в оглядах [12–15]. Саму текстуру важко визначити. Є різні визначення візуальної текстури, але вони часто не мають формальності й повноти. Так, за визначенням Дж.К. Хокінса [16], поняття текстури залежить від трьох інгредієнтів:

- 1) місцевий порядок повторюється на великому в порівнянні з розміром рослини об'ємі;
- 2) порядок полягає в не випадковому розташуванні елементарних частин;
- 3) частини є приблизно однорідними об'єктами, що мають однакові розміри в текстурованому об'ємі.

Кілька недавніх підходів до розпізнавання текстури відрізняються відмінними результатами в стандартних наборах даних, багато з яких працюють лише з інтенсивністю зображення та ігнорують доступну інформацію про кольори. Ряд підходів базується на популярних локальних подвійних патернах (LBP) [17; 18], таких як недавні паралельні обертання інваріантних локальних шаблонів двійника (X. Qi та інші [19]) або особливості Фур'є-гістограми (Т. Ахонен і співавтори [20; 21]).

Глибинні згорткові нейронні мережі (CNNs) досягли ряду завдань комп'ютерного зосередження, зокрема тих, що пов'язані зі складним розпізнаванням і виявленням об'єктів із великими базами даних навчальних зображень. К. Сегеді й співавтори [22] вивчають шляхи ефективного розширення мереж шляхом факторизованих згортоків й агресивної регуляризації. Їхнє вивчення виконується на мережах початкового стилю (тобто мережах з архітектурами, аналогічними GoogleNet). Вони пропонують так звану архітектуру Inception v3. Крім того, К. Сегеді й співавтори показують, що тренування із залишковими з'єднаннями значно прискорюють підготовку мереж початкових операцій, а залишкові мережі початкових показників можуть перевершувати аналогічно дорогі мережі початківців без залишкових зв'язків за тонкого поля.

**Постановка завдання** дослідження полягає в описі текстури незалежно від розміру візерунка рослини й орієнтації на зображенні. Необхідний опис, інваріантний для обертання та масштабу для практичного застосування в режимі реального часу для розробки вебсервісу. Також потрібне ефективне обчислення.

**Виклад основного матеріалу дослідження.** Нижче приведемо декілька методів для розпізнавання текстур.

Використовується локальне подвійне зображення та гистограма Фур'є. Перше описується на основі локальних двійкових шаблонів (LBP). Загальний оператор LBP локально обчислює ознаки відмінностей між центральним пікселем і його  $P$  сусідів по колу радіуса  $R$ . З функцією зображення  $f(x, y)$  та точками координат точки  $(x_p, y_p)$ :

$$LBP_{p,r}(x, y) = \sum_{p=0}^{P-1} s(f(x, y) - f(x_p, y_p)) 2^p, s(z) = \begin{cases} 1 & \text{if } z \leq 0 \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

Для досягнення інваріантності обертання ми приймаємо так звані Фур'є-функції гистограми LBP (LBP-HF), представлені Ahonen et al. LBP-HF описують гистограму рівномірних візерунків, використовуючи коефіцієнти дискретного перетворення Фур'є (DFT). Уніфіковані LBP – це шаблони з максимум двома просторовими переходами (побітові 0–1 зміни). На відміну від простих інваріантів обертання з використанням LBP, яке об'єднує всі однорідні візерунки з таким самим числом 1s в один контейнер, функції LBP-HF зберігають інформацію про відносне обертання шаблонів.

Позначення рівномірної картини  $U_p^{n,r}$ , де  $n$  – це число «орбіти», що відповідає кількості бітів «1», а  $r$  – обертання шаблону, DFT для заданого  $n$  виражається як:

$$H(n, u) = \sum_{r=0}^{P-1} h_r(U_p^{n,r}) e^{-2\pi i u r}, \quad (2)$$

де значення гистограми  $h_r(U_p^{n,r})$  позначає кількість входжень рівномірного малюнка на зображенні. Функції LBP-HF дорівнюють абсолютному значенню величин DFT, і тому на них не впливає фазовий зсув, викликаний поворотом.

$$LBP - HF(n, u) = |H(n, u)| = \sqrt{H(n, u) \overline{H(n, u)}}, \quad (3)$$

Так,  $h_r$  реальні,  $H(p, y) = \overline{H(p, P - y)}$  для  $y = (1, \dots, P - 1)$ , і тому тільки  $\left[\frac{P}{2}\right] + 1$  від DFT магнітуди використовуються для кожного набору рівномірних візерунків з  $n$  «1» біт при  $0 < n < P$ . Три резус-представлення додається до трьох інших контейнерів, а саме: два для «1-рівномірних» візерунків (з усіма контейнерами одного значення) та одне для всіх неоднорідних візерунків.

Функції Фур'є та гистограми LBP можна узагальнити на будь-який набір рівномірних візерунків. У першому випадку використовується опис LBP-HF-SM, де для побудови дескриптора розраховуються функції Фур'є-гістограми як знаків, так і величини-LBP. Величина-LBP перевіряє, чи величина різниці сусіднього пікселя  $(x_p, y_p)$  відносно центрального пікселя  $(x, y)$  перевищує порогове значення  $t_p$ :

$$LBP - M_{p,r}(x, y) = \sum_{p=0}^{P-1} s(|f(x, y) - f(x_p, y_p)| - t_p) 2^p, \quad (4)$$

Прийнято загальну практику вибору порогового значення (для сусідів на  $p$  біт) як середнє значення всіх абсолютних різниць  $m$  у всьому зображенні:

$$t_p = \sum_{i=1}^m \frac{|f(x, y) - f(x_{ip}, y_{ip})|}{m}, \quad (5)$$

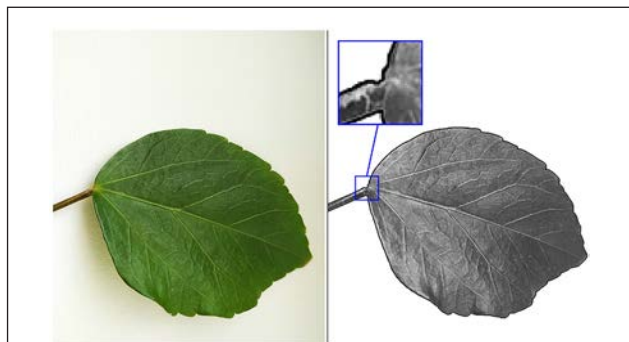
Гистограма LBP-HF-SM створюється шляхом об'єднання гистограм LBP-HF-S і LBP-HF-M (обчислено з рівномірного знаку-LBP і величини-LBP).

Для додавання інваріантів обертання функції LBP-HF, які використовуються в запропонованому описі, зазвичай будуються з DFT величини різномірних рівномірних візерунків. Можна використовувати всі LBP, а не тільки підмножину уніфікованих візерунків. У такому випадку деякі орбіти мають меншу кількість моделей, оскільки деякі неоднорідні шаблони демонструють симетрії. Шаблони на одній орбіті відрізняються лише за обертком. Інші інваріанти обертання обчислюються з перших коефіцієнтів DFT для кожної орбіти:

$$LBP - HF^+ = \sqrt{H(n, 1) \overline{H(n+1, 1)}}, \quad (6)$$

$F^+$  позначає метод, що використовує повний набір моделей для функцій LBP-HF і додавання додаткових функцій LBP-HF<sup>+</sup>.

Для визначення сегментових текстурних об'єктів можна розширити  $F^+$  до сегментованих текстурних об'єктів, обробляючи кордон й інтер'єр об'єктного сегмента окремо. Якщо розглянути ділянку сегментованих об'єктів  $A$ , то можна описати лише ті точки, які мають всі сусіди в певній шкалі всередині  $A$ . Показано, що опис правильно сегментованих кордонів, тобто точки в  $A$  з одним або декількома сусідами поза  $A$  (див. рис. 1), додає додаткову дискримінаційну інформацію.



**Рис. 1.** Сегментація зображення листа (сірого кольору) й прикордонного регіону (чорного кольору) в різних масштабах за даними LBP радіусом  $R$ . Прикордонний регіон листа визначається як всі точки, які мають принаймні одного сусіда (у LBP,  $R$ ) поза сегментованою областю

Було проведено експерименти з 5 варіантами способу розпізнавання, що відрізняються при обробці прикордонного регіону:

1. Описує всі пікселі в  $A$  та максимізує оцінку задньої імовірності (тобто імовірнісний вихід SVM Platt) над усіма  $n$ -рівнями шкали.

2. Описує лише сегмент інтер'єру, тобто пікселі в  $A$  з усіма сусідами в  $A$ .

3. Описує лише межу сегмента, тобто пікселі в  $A$  з принаймні одним сусідом поза  $A$ .

4. Перший  $_{ib\Sigma}$  поєднує в собі опис  $F_{first_i}$  та  $F_{first_b}$  і максимізує суму їхніх оцінок задньої вірогідності над  $n_{conc}$  шкалами.

5. Перший  $_{ib\Pi}$  поєднує в собі опис  $F_{first_i}$  та  $F_{first_b}$  і максимізує продукт своїх оцінок задньої вірогідності над  $n_{conc}$  шкалами.

**Висновки.** У глибинних навчальних завданнях є поширеною практикою навчання декількох мереж на різних (але не обов'язково несумісних) підмножинах навчальних даних. Визначення видів рослин із фотографій із використанням текстурного визнання за допомогою сучасних методів дає можливість досягти досить хороших результатів, водночас зберігаючи обчислювальні вимоги невеликими. Це робить його придатним для обробки в реальному часі й дає змогу розробити вебсервіс для мобільного пристосування.

#### Список літератури:

1. Kumar N., Belhumeur P.N., Biswas A., Jacobs D.W., Kress W.J., Lopez I.C., et al. Leafsnap: a computer vision system for automatic plant species identification. *Computervision–ECCV*. Springer, 2012. P. 502–516.
2. Belhumeur P.N., Chen D., Feiner S., Jacobs D.W., Kress W.J., Ling H., et al. Searching the world's herbaria: a system for visual identification of plantspecies. *Computervision–ECCV*. Springer, 2008. P. 116–29.
3. Agarwal G., Belhumeur P., Feiner S., Jacobs D., Kress W.J., Ramamoorthi R., et al. First steps toward an electronic field guide for plants. *Taxon*. 2006. № 55 (3). P. 597–610.
4. Wu S.G., Bao F.S., Xu E.Y., Wang Y.X., Chang Y.F., Xiang Q.L. A leaf recognition algorithm for plant classification using probabilistic neural network. *IEEE international symposium on signal processing and information technology*, 2007. IEEE, 2007. P. 11–16.
5. Hang S.T., Tatsuma A., Aono M. Bluefield (KDE TUT) at Life CLEF 2016 plant identification task. *Working notes of CLEF 2016 : conference and labs of thee valuation forum*, 2016. 10 p.
6. Ghazi M.M., Yanikoglu B., Aptoula E. Open-set plant identification using an ensemble of deep convolutional neural networks. *Working notes of CLEF 2016 : conference and labs of thee valuation forum*, 2016. 7 p.
7. Choi S. Plant identification with deep convolutional neural network: SNU Medinfoat LifeCLEF plant identification task 2015. *Working notes of CLEF 2015 : conference and labs of thee valuation forum*, Toulouse, September 8–11, 2015. CEUR-WS, 2015.
8. Ge Z., McCool C., Sanderson C., Corke P. Content specific feature learning for fine-grained plant classification. *Working notes of CLEF 2015 : conference and labs of thee valuation forum*, Toulouse, September 8–11, 2015. CEUR-WS, 2015.
9. Barthélémy D., Boujemaa N., Mathieu D., Molino J.F., Bonnet P., Encficiaud R., et al. *ThePl@ntNet project: a computational plant identification and collaborative information system* : Tech. Rep. XIII World forestry congress, 2009.
10. Pl@ntNet-identify webservice. URL: <http://identify.plantnet-project.org/en/> (дата звернення: 14.12.2017).
11. Goëau H., Bonnet P., Joly A., Bakić V., Barbe J., Yahiaoui I., et al. Pl@nt net mobile app. *Proceedings of the 21st acminternational conference on multimedia*. ACM, 2013. P. 423–424.
12. Zhang J., Tan T. Brief review of invariant texture analysis methods. *PatternRecogn.* 2002. № 35 (3). P. 735–747. DOI: 10.1016/S0031-3203(01)00074-7.
13. Mirmehdi M., Xie X., Suri J. Handbook of texture analysis. London : ImperialCollegePress, 2009.
14. ChenCh, Pau L.F., WangPsp. Handbook of pattern recognition and computervision. Singapore : WorldScientific, 2010.
15. Pietikäinen M. Texturerecognition. *Computer vision: a reference guide* / editor K. Ikeuchi. Springer, 2014. P. 789–793.

16. Hawkins J.K. Textural properties for pattern recognition. *Picture processing and psychopictorics* / editor B.S. Lipkin. Elsevier, 1970. P. 347–370.
17. Ojala T., Pietikainen M., Harwood D. Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. *Proceedings of IAPR*. Vol. 1. 1994. P. 582–585.
18. Ojala T., Pietikainen M., Harwood D. A comparative study of texture measures with classification based on featured distributions. *Pattern Recogn.* 1996. № 29 (1). P. 51–59. DOI: 10.1016/0031-3203(95)00067-4.
19. Qi X., Xiao R., Guo J., Zhang L. Pair wiserotation in variant co-occurrence local binary pattern. *Computervision—ECCV 2012*. Springer, 2012. P. 158–171.
20. Ahonen T., Matas J., He C., Pietikainen M. Rotation in variant image description with local binary pattern histogram Fourier features. *Proceedings of SCIA '09*. Springer-Verlag, 2009. P. 61–70.
21. Zhao G., Ahonen T., Matas J., Pietikainen M. Rotation-invariant image and video description with local binary pattern features. *IEEE TransImageProcess*. 2012. № 21 (4). P. 1465–1477.
22. Szegedy C., Vanhoucke V., Ioffe S., Shlens J., Wojna Z. Rethinking the inception architecture for computervision. 2015.

### **Grabar O.I., Lysogor Yu.I., Skachkov V.O. ANALYSIS OF ALGORITHMS OF VISUAL SYSTEM RECOGNITION SORTS OF PLANTS FOR DEVELOPMENT THE WEB-SERVICE**

*The main objective is to create a web service for plant recognition by image. High-quality artificial of plants perception is a difficult task for a computer through a various kind and difficult structure of plants. The state of modern machine studies and task of recognition is examined from authentication of concrete organs of plant to general recognition all the plants “in the wild nature”. From the point of view of machine studies, recognition of plants is a fine-grained classification task with high change ability between classes and often small interclass differences that is related to taxonomical hierarchical classification. For description of texture regardless of size a pattern and orientation dark-and-light, it is necessary to have invariant description for a rotation and scale. For practical application an effective calculation is also needed. A few methods are below given for recognition of textures. It is shown a local double image and histogram of Fourier that is described on the basis of local binary templates (LBP). The functions of Fourier and histograms of LBP can be generalized on any set of even patterns. Description of LBP is used in first case – HF – SM, where for the construction of descriptor the functions of Fourier settle accounts histograms of both signs and size-of LBP. Size of-LBP checks, or the size of difference of nearby pixel in relation to a central pixel exceeds the threshold value of  $t p$ : In deep educational tasks is widespread practice of studies of a few networks on the different subsets of educational data. Association of such networks, usually combine by means of simple mechanism and aims to surpass individual networks. Determination of types of plants from photos with the use of texture confession by means of modern methods gives to attain good results sufficiently, here keeping calculable requirements small, that does him suitable for treatment real-time.*

**Key words:** machine studies, recognition, web-service, polar transformations of Fourier, meta data, neural network.