

УДК 528.72 : 004.932.4

М. С. ЛАВРЕНЮК

*Інститут космічних досліджень НАН та ДКА України,  
Національний технічний університет України «Київський політехнічний інститут  
імені Ігоря Сікорського»*

## МЕТОД ОБ'ЄКТНОЇ ПОСТОБРОБКИ КАРТ КЛАСИФІКАЦІЇ З УРАХУВАННЯМ СПЕЦИФІКИ КОЖНОГО КЛАСУ

*Запропоновано метод фільтрації карт класифікації земного покриву, отриманих на основі супутникових знімків високого розрізнення. Основна ідея методу - дослідження властивостей кожного об'єкта в цілому (групи пікселів з однаковим значенням класу) з врахуванням специфіки кожного класу і, як наслідок, підбором порогу фільтрації для кожного з них окремо. Враховуючи специфіку карт класифікації земного покриву, необхідно зберегти межі між сільськогосподарськими полями. Для цього запропоновано модифікований алгоритм Собеля. Точність та ефективність даного методу підтверджується тестуванням на незалежній вибірці, а також шляхом візуального порівняння з результатами загальновідомих методів фільтрації.*

**Ключові слова:** класифікація посівів, постобробка, фільтрація, маркування зв'язних компонент, алгоритм Собеля.

### Вступ

Більшість прикладних задач в сфері спостереження Землі спираються та використовують карти класифікації земного покриву [1 - 4]. Незалежно від того, які методи машинного навчання використовуються для її розв'язання: традиційні, чи найсучасніші методи глибинного навчання, на карті класифікації присутній шум — поодинокі пікселі чи групи пікселів, які віднесені до не правильного класу [5 – 8]. Детальний аналіз існуючих підходів до розв'язання даної проблеми розглянутий в [9]. Проте традиційні методи не враховують семантику зображення, а тому не є ефективними для фільтрації карт земного покриву. Найскладнішою проблемою при розв'язанні цієї задачі, яка має самостійне значення, є фільтрація орних (оброблюваних) сільськогосподарських земель, оскільки окремі поля розділяються межами і лісосмугами, які не можна розглядати як шум.

В даній статті запропоновано процедуру постобробки карти класифікації, що дозволяє отримати вищу точність, ніж за допомогою загальноприйнятих фільтрів, і в той же час зберегти форму об'єктів та невеликі за розміром об'єкти, такі як лісосмуги, вирубки та межі між полями. На відміну від загальновідомих методів фільтрації шуму на зображеннях, запропонований метод базується на об'єктному підході, враховує семантику зображення та полягає у виділенні

морфологічних ознак різних класів земного покриву, і в першу чергу, орних (оброблюваних) земель.

### 1. Алгоритм постобробки карти класифікації

Об'єктом даного дослідження є карта класифікації земного покриву  $I$ , що містить  $N$  класів. Така карта може бути отримана шляхом класифікації одного, групи, або часового ряду супутникових знімків з використанням будь-якого методу класифікації, наприклад методів машинного навчання [10 – 14]. Характерними особливостями такої карти є дискретна множина можливих значень (інтенсивностей) пікселів, а також наявність невірно класифікованих пікселів або цілих сегментів, які можна розглядати як шум на зображенні. Ці особливості відрізняють карти класифікації від інших типів графічних, відео- зображень, а також багатоспектральних супутникових знімків, тому для них не можна застосовувати традиційні методи фільтрації зображень. Більше того, оскільки класи на карті мають різний семантичний зміст, алгоритм постобробки повинен враховувати семантику зображення, і навіть параметри традиційних операцій фільтрації мають відрізнитись для різних класів.

Без втрати загальності будемо вважати, що базовими категоріями класів відповідно до

номенклатури класів European Land Use and Cover Area frame Survey (LUCAS) [15] є ліс (С), сільськогосподарські землі (В), водні об'єкти (W), штучні об'єкти (А), необроблювальні землі (Е) та відкритий ґрунт (F) [16]. Всі наявні класи можна поділити на категорії за рівнем достовірності класифікації (висока-низька) та за площею об'єктів (велика-мала). Представимо множину класів на карті у вигляді об'єднання [17]:

$$I = \{C, W, A, E\} \cup \{B, F\}. \quad (1)$$

З метою врахування семантичного змісту кожного із класів постобробку карти будемо здійснювати окремо для кожної із підмножин (1).

Для фільтрації невірно класифікованих пікселів на картах класифікації пропонуємо комплексний алгоритм, загальна схема якого показана на рис. 1. Розглянемо кожний етап постобробки окремо.

Операції, які виконуються в блоці 1 (детектування меж на основі модифікованого

оператора Собеля) описані в [9], тому детально зупинимось на етапах 2-7 (рис. 1).

### 1.1. Порогова фільтрація малих об'єктів з урахуванням специфіки кожного класу

Спочатку використовуємо стандартний фільтр, що ідентифікує всі об'єкти менші заданого порогу  $t_i$ , де  $i = \overline{1, N}$ , як шум. Після цього значення зашумлених пікселів замінюються на значення, що найчастіше зустрічається в диску circle радіусом  $r_i$  пікселів з центром в досліджуваному пікселі

$$I(x, y) = \operatorname{argmax}(\operatorname{hist}(\operatorname{circle}(x, y, r_i))) \quad (4)$$

Далі ця ж процедура повторюється  $\operatorname{num}$  разів з  $t_i^j$ , де  $i = \overline{1, N}$  та  $j = \overline{1, \operatorname{num}}$ .

Наступним кроком всі пікселі, що на початковій карті були віднесені не до класу С, а

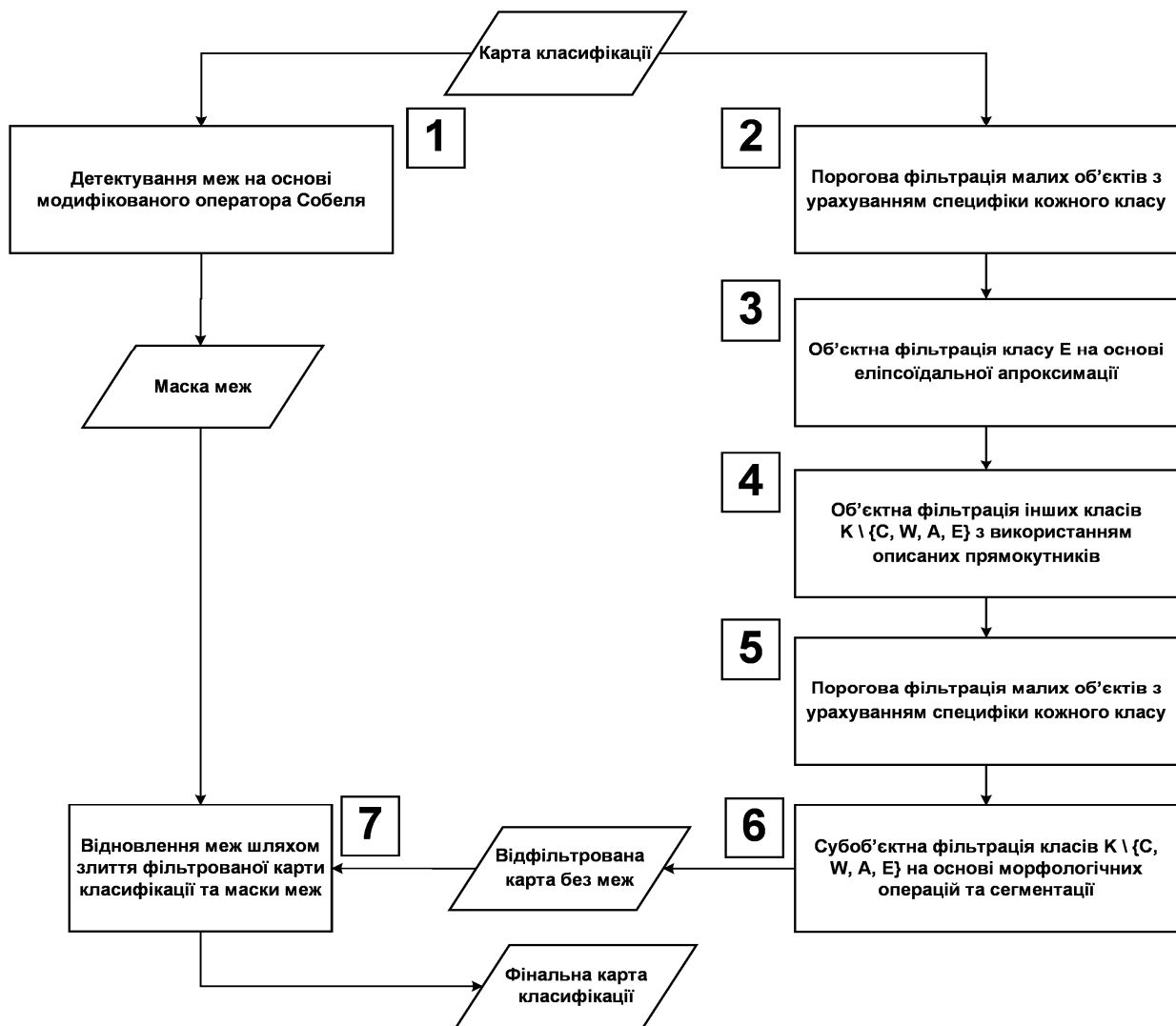


Рис. 1. Блок-схема запропонованого алгоритму фільтрації карти класифікації

після фільтрації отримали значення  $S$  переназначаються в клас  $E$ , тому що клас  $C$  завжди на картах класифікацій розпізнається з майже сто відсотковою точністю, незалежно від методу машинного навчання. Таким чином, вдається виявити вирубки в лісі, які при класифікації отримали значення не класу  $E$ , а будь-якого іншого класу та при фільтрації були ідентифіковані як шум.

## 1.2. Об'єктна фільтрація класу $E$ на основі еліпсоїдальної апроксимації

Беручи до уваги можливість присутності бур'яну на полях, що призводить до появи пікселів класу  $E$  посеред поля, на наступному етапі відбувається фільтрація лише класу  $E$ . Даний етап та всі наступні етапи фільтрації відбуваються не для кожного пікселя окремо, а для кожного об'єкта в цілому, щоб уникнути недоліків всіх загально-прийнятих фільтрів, які працюють саме для пікселів.

Означення.

Об'єктом будемо називати компоненту зв'язності графа – максимальний зв'язний підграф, вершини якого – це центри пікселів на зображенні  $I$ . Будемо вважати, що вершини  $u$  та  $v$  зв'язані між собою, якщо значення пікселів в них співпадають та відстань між ними не перевищує  $\sqrt{2}$ .

$$I(u) = I(v); \text{dist}(u, v) \leq \sqrt{2}. \quad (5)$$

Якщо об'єкт ідентифікований як шум, то для його видалення всі його пікселі замінюються на те значення, яке найчастіше зустрічається вздовж периметру цього об'єкта. Для розбиття карти класифікації на об'єкти використовується алгоритм маркування зв'язних компонент [18].

Наведемо псевдокод даного етапу фільтрації

*Будуємо бінарну маску класу  $E$*

*Визначаємо компоненти зв'язності  $\{CC^i\}$*

*Для  $i$  від 1 до  $|\{CC^i\}|$*

*Якщо  $S(CC^i) < T_E$*

*Визначаємо параметри еліпсоїдальної апроксимації об'єкта  $(\lambda_1, \lambda_2, E_x)$*

*Якщо  $E_x < T_{E_x}$*

*Об'єкт маркується як шум*

Важливо відмітити, що в даному алгоритмі видаляються всі об'єкти, площа яких менше заданого порогу, проте, щоб зберегти від видалення ті об'єкти зі значенням  $E$ , які є продовгуватими, але

вузькими, що відповідають дорогам чи межам між полями будуємо еліпс (рис. 2), що найкраще підходить для нашого об'єкта, тобто має такий же 2-й момент (коваріаційну матрицю), як і досліджуваний об'єкт [19, 20].

Для двовимірної неперервної функції  $f(x, y)$  момент порядку  $p+q$  визначається

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x, y) dx dy.$$

Для зображення функція  $f(x, y)$  – це інтенсивність пікселя  $(x, y)$ , тому дана функція є дискретною. В нашому випадку зображення є бінарним, тому функція  $f(x, y)$  приймає значення лише 1 в межах досліджуваного об'єкта, або 0 поза його межами. Для дискретної функції  $f(x, y)$  момент порядку  $p+q$  визначається так

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y).$$

Центральний момент  $\mu_{pq}$ , який обчислюється

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy,$$

для дискретної функції  $f(x, y)$  набуває вигляду

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y),$$

де  $\bar{x} = \frac{m_{10}}{m_{00}}$  та  $\bar{y} = \frac{m_{01}}{m_{00}}$  є координатами центроїда

даного об'єкта. Тоді коваріаційна матриця зображення  $I$  може бути представлена, як

$$\text{cov}(I(x, y)) = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{00} & \mu_{00} \\ \mu_{11} & \mu_{02} \\ \mu_{00} & \mu_{00} \end{bmatrix}.$$

Варто відзначити, що власні вектори цієї матриці відповідають великій та малій осям еліпса, а його орієнтація обчислюється як кут  $\theta$  між власним вектором, пов'язаним з найбільшим власним значенням та віссю, що є найближчою до цього власного вектора. Цей кут можна обчислити за

формулою [21]

$$\theta = \frac{1}{2} \arctan \left( \frac{2\mu_{11}}{\mu_{20} - \mu_{02}} \right).$$

Отже, враховуючи, що еліпс будувався таким чином, щоб його коваріаційна матриця була такою ж як і досліджуваного об'єкта, їхні орієнтації співпадають.

Власні числа матриці коваріації  $\lambda_{1,2}$  можна обчислити за формулою

$$\lambda_{1,2} = \frac{\mu_{20} + \mu_{02}}{2\mu_{00}} \pm \frac{\sqrt{4\mu_{11}^2 + (\mu_{20} - \mu_{02})^2}}{2\mu_{00}}.$$

Довжини малої та великої осей еліпса а та b знаходяться за формулами [22]

$$a = \sqrt{\lambda_1}; \quad b = \sqrt{\lambda_2}.$$

Тоді величина, що характеризує його витягнутість  $E_x$  – ексцентриситет еліпса визначається за допомогою власних чисел коваріаційної матриці

$$E_x = \sqrt{1 - \frac{b^2}{a^2}} = \sqrt{1 - \frac{\lambda_2}{\lambda_1}}; \quad E_x \in [0, 1).$$

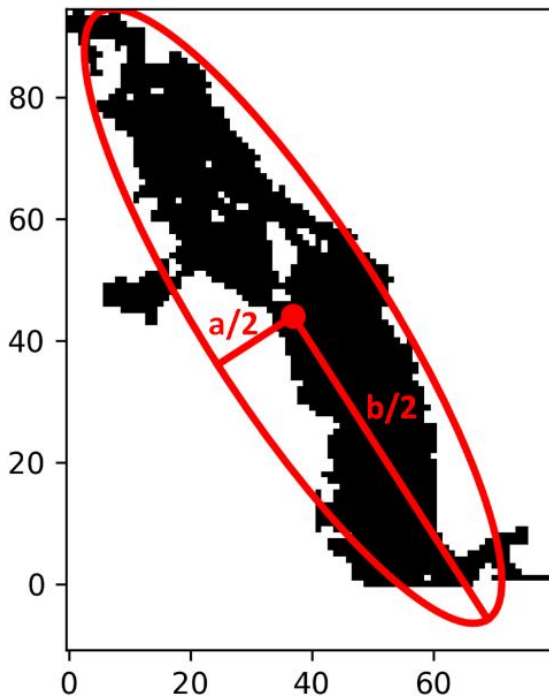


Рис. 2. Приклад об'єкта та його еліпсоїдальної апроксимації

### 1.3. Об'єктна фільтрація з використанням описаних прямокутників

Будемо вважати, що перша множина класів {C, W, A, E} вже відфільтрована належним чином і залишається лише фільтрувати решту класів, а саме  $K \setminus \{C, W, A, E\}$ . Основною ідеєю на даному етапі є ідентифікація об'єктів, що мають неправильну форму з посіченими та нерівними межами. Об'єкти такого типу на карті класифікації будемо вважати шумом та видаляти їх. Наведемо псевдокод запропонованого алгоритму для даного етапу фільтрації

Для кожного класу  $C_i$  з  $K \setminus \{C, W, A, E\}$

Визначаємо компоненти зв'язності  $\{CC_i^j\}$

Для j від 1 до  $|\{CC_i^j\}|$

Якщо  $S(CC_i^j) < T_{CC}$

Будуємо описаний прямокутник  $R_{CC}$

Будуємо контур об'єкта  $C_{CC_i^j}$

Спрощуємо контур  $\tilde{C}_{CC_i^j}$

Якщо  $S(CC_i^j) > T_{CC2}$

Якщо  $\frac{S(R_{CC})}{S(CC_i^j)} > T_{R_{CC}}$ , та  $|\tilde{C}_{CC_i^j}| > V_{CC}$

Якщо  $\frac{S(CC_i^j)}{S(CC_i^j \bullet s(r_{cc}))} < 1$

Об'єкт маркується як шум

Якщо  $\frac{S(CC_i^j)}{S(CC_i^j \circ s(r_{cc}))} > T_{s(r_{cc})}$

Об'єкт маркується як шум

Враховуючи той факт, що компактні об'єкти з непосіченими межами будуть представлятись у вигляді багатокутника з невеликою кількістю вершин, важливо використати таку процедуру спрощення контура, яка не згладить існуючий контур, а залишить з нього певну підмножину вершин. Саме тому для спрощення контуру  $C_{CC_i^j}$ , що складається з n вершин пропонується використовувати алгоритм Рамера-Дугласа-Пекера [23, 24]

$$C_{CC_i^j} = \{p_1, p_2, \dots, p_n\}; \quad p = (x_i, y_i).$$

Спрощений контур складається із меншого числа точок

$$\tilde{C}_{CC_i^j} \subseteq C_{CC_i^j}; n^* < nn^* < n;$$

$$\tilde{C}_{CC_i^j} = \{P_1^*, P_2^*, \dots, P_n^*\}; P^* = (x_i^*, y_i^*).$$

Новий контур  $\tilde{C}_{CC_i^j}$  розділяє початковий контур  $C_{CC_i^j}$  на  $S_k$  підмножин

$$S_k = \{P_i, P_{i+1}, \dots, P_j\}; P_i = P_{k-1}^*; P_j = P_k^*.$$

Для даних підмножин виконуються умови:

$$\bigcup_k S_k = C_{CC_i^j},$$

$$\left( \bigcup_k (S_k \cap S_{k+1}) \right) \cup (S_n \cap S_1) = \tilde{C}_{CC_i^j}.$$

Алгоритм Рамера-Дугласа-Пекера знаходить  $S_k$  таким чином, щоб

$$f(S_k) = \max \left( \text{dist} \left( P_i, \langle P_{k-1}^*, P_k^* \rangle \right) \right),$$

де  $e$  – заданий наперед поріг,  $\text{dist}$  – функція Евклідової відстані,  $P_i \in S_k$ ,  $\langle P_{k-1}^*, P_k^* \rangle$  – являє собою прямолінійний відрізок з кінцями в точках  $P_{k-1}^*$  та  $P_k^*$ .

Після закінчення роботи алгоритму даного етапу фільтрації застосовується блок, описаний в підрозділі 1.1.

#### 1.4. Субоб'єктна фільтрація на основі морфологічних операцій та сегментації

На даному етапі нерозглянутою лишилась лише ситуація, коли до правильно класифікованого об'єкта класів  $K \setminus \{C, W, A, E\}$  дотикається об'єкт-шум (рис. 3). Таким чином об'єкт-шум та правильний об'єкт визначаються як один об'єкт і в такому випадку шум прибрати на попередніх етапах не вдалось. Для розділення об'єкта на шум, що дотикається, та правильну частину об'єкта застосовуємо ерозію з квадратом розміром  $r_{sq}$  на  $r_{sq}$ , і якщо кількість компонент зв'язності після цієї операції зростає, робимо висновок, що досліджуваний об'єкт містить в собі шум. Для розділення об'єкта на субоб'єкти будемо виконувати сегментацію методом водоподілу

(watershed) досліджуваного об'єкта [25]. В результаті всі субоб'єкти, площа яких не перевищує  $T_{seg}$ , маркуються як шум і видаляються. Наведемо псевдокод даного етапу фільтрації

Для кожного класу  $C_i$  з  $K \setminus \{C, W, A, E\}$

Визначаємо компоненти зв'язності  $\{CC_i^j\}$

Для  $j$  від 1 до  $|\{CC_i^j\}|$

Обчислюємо ерозію  $E(CC_i^j, sq(r_{sq}))$

Визначаємо компоненти зв'язності  $\{CCC_i^j\}$

Якщо  $|\{CCC_i^j\}| > 1$

Сегментація  $CC_i^j$  водорозділом  $\{Seg_i^j\}$

Для  $m$  від 1 до  $|\{Seg_i^j\}|$

Якщо  $S(Seg_i^j) < T_{seg}$

Об'єкт  $Seg_i^j$  маркується як шум

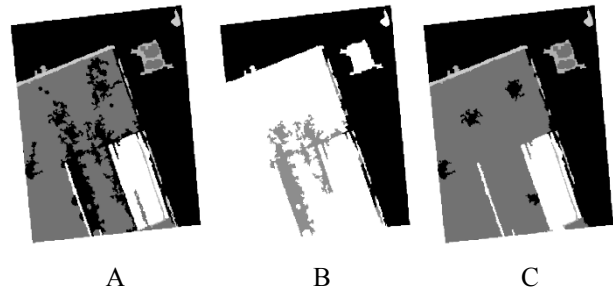


Рис. 3. Приклад об'єкта (чорним кольором) сполученого з шумом (А), сегментація запропонованого об'єкту (В), результат відділення та видалення шуму (С)

Субоб'єктна фільтрація є останнім етапом визначення шуму на карті класифікації. Після закінчення процедури постобробки бінарна маска пікселів з меж між полями, що отримана на першому етапі, повертається на фінальну відфільтровану карту класифікації.

## 2. Територія досліджень та використані дані

Запропонована методологія фільтрації застосована для карти класифікації посівів 2017 року на тестовому полігоні ЛЕСАМ, який включає Київську область (рис. 4).

Площа області складає 28,100 км<sup>2</sup>. Основними культурами є кукурудза, озима пшениця, соя,

соняшник, ярий ячмінь, озимий ріпак, цукровий буряк та горох. Для оцінки точності карти класифікації та ефективності різних процедур постобробки використовувались наземні дані, що були зібрані вздовж доріг, але не використовувались для побудови даної карти посівів. Тестова вибірка містить 455 полігонів, та включають основні класи земного покриву відповідно до стандартизованої номенклатури класів European Land Use and Cover Area frame Survey (LUCAS) (табл. 1). Загальна площа тестових даних становить 5858,16 га.

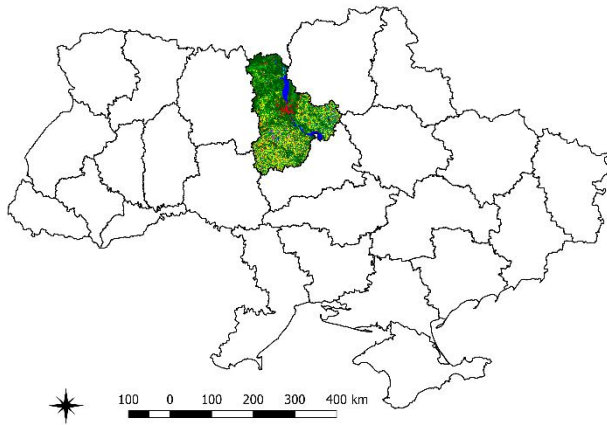


Рис. 4. Карта класифікації посівів в Київській області в 2017

Для даної карти класифікації, побудованої на основі радарних даних Sentinel-1 з розрізненням 10 м, підібрані були наступні значення параметрів

$$\begin{aligned} \text{num} &= 3; t_1^1 = 10; t_2^1 = 50; t_1^2 = 10; t_2^2 = 300; \\ r_1 &= 5; T_E = 300; T_{E_x} = 0,97; \\ T_{CC} &= 2000; T_{CC_2} = 300; T_{R_{CC}} = 1,2; \\ V_{CC} &= 9; r_{cc} = 3; T_{s(r_{cc})} = 1,2; r_{sq} = 3; T_{seg} = 1000. \end{aligned}$$

Для порівняння різних методів постобробки використовувались наступні метрики, які обчислювались на основі матриці невідповідностей: загальна точність класифікації (OA), індекс Карра, точність «виробника» (PA) та точність «користувача» (UA) [26]. Індекс Карра є статистичною мірою узгодженості між отриманими класами на карті класифікації та класами, що є реально (наземні дослідження). Точність «виробника» - це відношення правильно розпізнаних пікселів даного класу до загальної кількості пікселів даного класу в тестовій вибірці. Точність «користувача» - це відношення правильно розпізнаних пікселів даного класу до загальної кількості пікселів даного класу на карті класифікації. Для перевірки чи є різниця між отриманими точностями карт класифікацій статистично значимою використовувався статистичний тест McNemar [27, 28]. Тест McNemar – це непараметричний тест на основі стандартизованої статистики звичайних тестів

$$z = \frac{f_{12} - f_{21}}{\sqrt{f_{12} + f_{21}}}$$

Таблиця 1

Тестова вибірка відповідно до номенклатури класів LUCAS

№	Код	Клас	Полігон		Площа	
			No.	%	га	%
1	Axx	Штучні об'єкти	8	1,8	11,00	0,2
2	B11	Озима пшениця	75	16,5	1548,23	26,5
3	B32	Озимий ріпак	19	4,2	561,49	9,0
4	B16	Кукурудза	106	23,2	1386,53	23,6
5	B22	Буряк	12	2,6	458,38	7,8
6	B31	Соняшник	54	11,9	520,42	8,9
7	B33	Соя	57	12,5	848,87	14,5
8	C10, B60	Ліси	60	13,2	326,19	5,6
9	E01, E02	Луги	34	7,5	80,72	1,4
10	F00	Порожня земля	3	0,7	4,29	0
11	G01, G02	Водні об'єкти	16	3,5	89,15	1,5
12	H00	Болото	9	2,0	9,66	0,2
13	B43	Горох	2	0,4	13,23	0,2
		Разом	455		5858,16	

де  $f_{12}$  - це кількість зразків, які правильно класифікуються за допомогою першого класифікатора і неправильно класифікуються другим,  $f_{21}$  - це кількість зразків, які правильно класифікуються за допомогою другого класифікатора і неправильно класифікуються першим.

### 3. Результати

Загальна точність початкової карти посівів для Київської області в 2017 році складала 91,9%. В якості еталонного загальноприйнятого фільтру було вибрано фільтр на основі голосування більшості з вікном форми диску. На рисунку 5 показано як змінюється загальна точність відфільтрованої карти класифікації посівів в залежності від розміру вікна даного фільтру. Бачимо, що спочатку загальна точність зростає, а при досягненні радіусом диска 10 пікселів відбувається насичення на рівні 93,6%.

Також протестовано метод фільтрації голосуванням з ваговими коефіцієнтами, що отримані були на виході ансамблю нейронних мереж при побудові карти класифікації для Київської області в 2017 році [29, 30]. Для диску радіусом 5 отримано загальну точність 93,4%, для радіусу 10 – 93,6%, результат такий же, як і для звичайного метода голосування. Причиною цьому є

незначна різниця ймовірностей правильно та не правильно класифікованих класів, що має значно менший вплив на вибір результуючого класу, ніж кількість пікселів кожного з класів.

Як видно з таблиці 2, найвища точність отримана для запропонованого методу, що на 2,3% вище, ніж точність початкової карти та на 0,6% вище в порівнянні з покращеним методом голосування. Варто відзначити, що зросла не лише загальна точність, але і точності кожного класу в порівнянні з початковою картою.

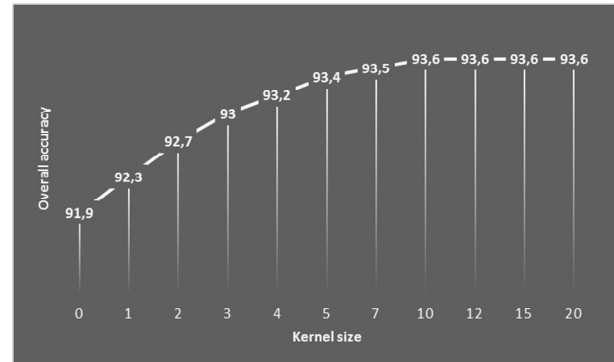


Рис. 5. Залежність загальної точності відфільтрованої карти класифікації від радіуса диска при фільтрації методом голосування

Таблиця 2

Порівняння точностей початкової карти класифікації (а), відфільтрованої за допомогою покращеного методу голосування (б) та запропонованого методу (с) (РА — точність «виробника»; UA — точність «користувача»)

№		а		б		с	
		РА, %	UA, %	РА, %	UA, %	РА, %	UA, %
1	Штучні об'єкти	87,7	71,5	100,0	83,3	95,9	74,0
2	Озима пшениця	95,9	98,5	96,7	98,8	96,5	98,8
3	Озимий ріпак	98,3	98,6	98,5	98,8	98,4	98,8
4	Кукурудза	93,0	84,9	97,3	85,3	98,0	86,2
5	Бурак	98,1	98,9	98,9	99,8	99,1	100,0
6	Соняшник	94,6	90,9	96,7	93,7	98,2	94,0
7	Соя	69,8	86,1	70,3	94,4	72,7	97,0
8	Ліси	99,5	99,0	100,0	99,3	99,7	99,2
9	Луги	94,9	60,1	98,0	66,6	96,8	66,0
10	Порожня земля	66,1	60,0	61,4	100,0	61,4	100,0
11	Водні об'єкти	99,9	99,9	100,0	100,0	100,0	100,0
12	Болото	78,5	83,5	94,4	100,0	83,7	100,0
13	Горох	98,0	97,8	99,3	99,8	100,0	99,8
	Загальна точність, %	91,9		93,6		94,2	
	Індекс Карра	0,9		0,92		0,93	



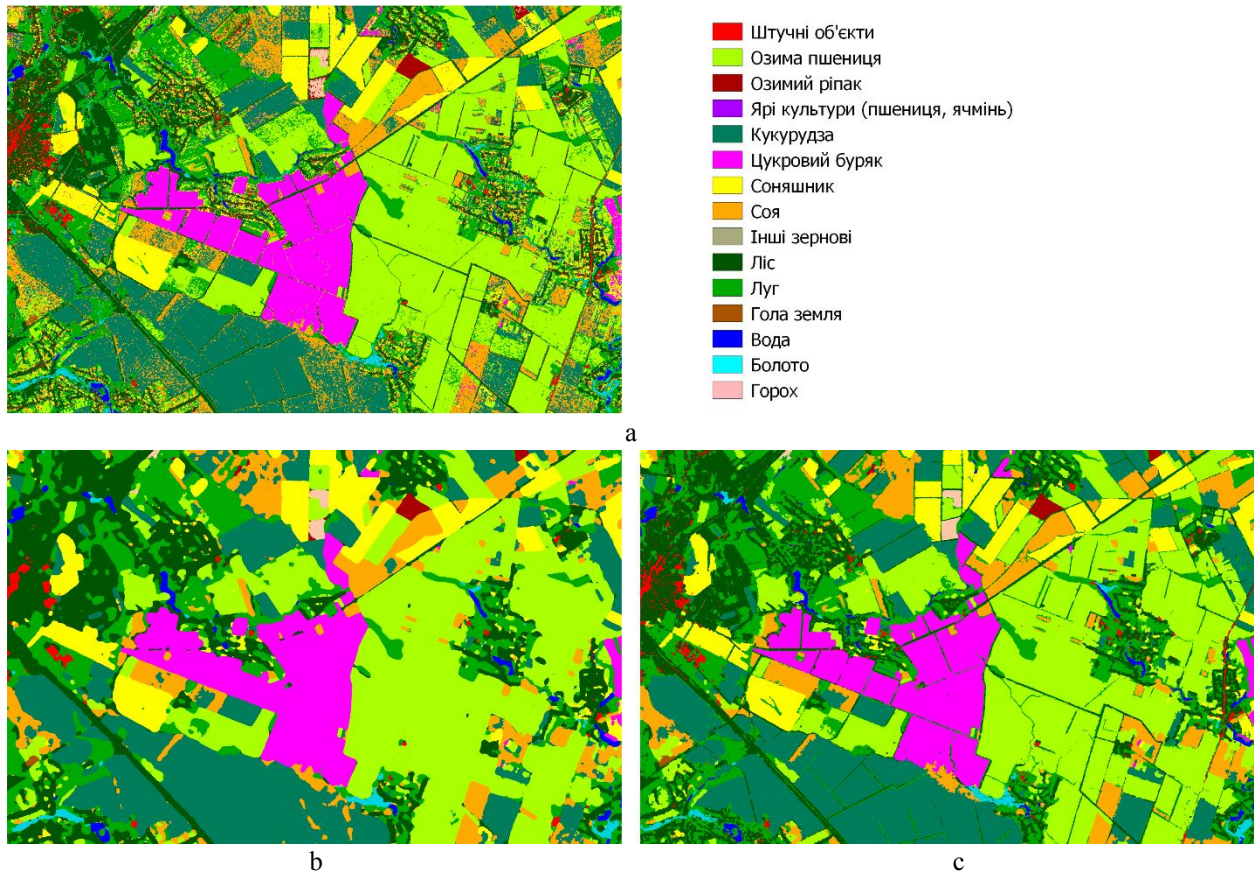


Рис. 6. Приклад початкової карти класифікації (а), застосування покращеного методу голосування (b) та запропонованого методу (с)

Для сільськогосподарських класів точність запропонованого методу не нижче точності покращеного методу голосування. Статистичний тест McNemar засвідчив, що різниця між запропонованим методом та покращеним методом голосування і між запропонованим методом та початковою картою класифікації є статистично значимою ( $p < 0,001$ ). Візуально результати порівняння цих методів свідчать про здатність запропонованого методу до збереження форм об'єктів та меж між ними, навіть коли ці межі на початковій карті не були правильно класифіковані (рис. 6).

### Висновки

В даній статті запропоновано метод постобробки карти класифікації земного покриву для зменшення наявного шуму на ній та збільшення загальної її точності. Ідея полягає в дослідженні кожної групи пікселів з однаковим значенням класу як цілісного об'єкта, на відміну від загально прийнятих методів, що базуються на принципі ковзаючого вікна. Даний підхід порівняно з загальноприйнятими методами фільтрації шуму і виявлено його перевагу як за допомогою оцінок

точностей та статистичного тесту McNemar, так і візуально для території Київської області за 2017 рік. Головною його перевагою є збереження форм об'єктів та меж між ними і запобігання зникненню об'єктів невеликих розмірів, які були достовірно класифіковані. Загальна точність карти класифікації зростає на 2,3% в порівнянні з початковою картою та на 0,6% в порівнянні з покращеним методом голосування. Метод є універсальним та може застосовуватись до карт класифікації, отриманих за допомогою будь-якого методу машинного навчання на основі як оптичних так і радарних даних.

*Публікація містить результати досліджень, проведених при грантовій підтримці Державного фонду фундаментальних досліджень за конкурсним проектом ДФФД Ф76/130-2017 від 17.10.2017 та в межах цільової комплексної програми НАН України з наукових космічних досліджень.*

### Література

1. *New postprocessing methods for remote sensing image classification: A systematic study [Text] / Xin Huang, et al. // IEEE Transactions on Geoscience and Remote Sensing. – 2014. – Vol. 52, Iss.11. – P. 7140-7159. DOI: 10.1109/TGRS.2014.2308192.*



2. Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine [Text] / A. Kolotii, et al. // *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. – 2015. – Vol. 40, Iss. 7. – P. 39-44.
3. Efficiency assessment of using satellite data for crop area estimation in Ukraine [Text] / F. Javier Gallego, et al. // *International Journal of Applied Earth Observation and Geoinformation*. – 2014. – Vol. 29. – P. 22-30.
4. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models [Text] / F. Kogan, et al. // *International Journal of Applied Earth Observation and Geoinformation*. – 2013. – Vol. 23. – P. 192-203.
5. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data [Text] / N. Kussul, et al. // *IEEE Geoscience and Remote Sensing Letters*. – 2017. – Vol. 14, Iss. 5. – P. 778-782.
6. Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping [Text] / A. Shelestov, et al. // *Frontiers in Earth Science*. – 2017. – Vol. 5. – P. 17.
7. Parcel-based crop classification in Ukraine using landsat-8 data and sentinel-1A data [Text] / N. Kussul, et al. // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2016. – Vol. 9, Iss. 6. – P. 2500-2508.
8. Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity [Text] / F. Waldner, et al. // *International Journal of Remote Sensing*. – 2016. – Vol. 37, Iss. 14. – P. 3196-3231.
9. Лавренюк, М. Метод об'єктної фільтрації карт класифікації земного покриття на основі морфологічних ознак [Text] / М. Лавренюк // *Індуктивне моделювання складних систем*. – 2017. – № 9 (в друці).
10. Löw, F. Decision fusion and non-parametric classifiers for land use mapping using multi-temporal RapidEye data [Text] / F. Löw, C. Christopher, M. Ulrich // *ISPRS Journal of Photogrammetry and Remote Sensing*. – 2015. – Vol. 108. – P. 191-204.
11. Geospatial Intelligence and Data Fusion Techniques for Sustainable Development Problems [Text] / N. Kussul, A. Shelestov, R. Basarab, S. Skakun, O. Kussul, M. Lavreniuk // *In ICTERI*. – 2015. – P. 196-203.
12. Land Cover Changes Analysis Based on Deep Machine Learning Technique [Text] / N. Kussul, N. Lavreniuk, A. Shelestov, B. Yailymov, I. Butko // *Jour. of Automation and Information Sciences*. – 2016. – vol. 48, iss. 5. – P. 42-54.
13. Deep learning approach for large scale land cover mapping based on remote sensing data fusion [Text] / N. Kussul, A. Shelestov, M. Lavreniuk, I. Butko, S. Skakun // *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. – 2016. – P. 198-201.
14. Along the season crop classification in Ukraine based on time series of optical and SAR images using ensemble of neural network classifiers [Text] / N. Kussul, M. Lavreniuk, A. Shelestov, B. Yailymov // *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. – 2016. – P. 7145-7148.
15. LUCAS 2009, Land Use / Cover Area Frame Survey [Electronic resource]. – Access mode: [http://ec.europa.eu/eurostat/documents/205002/208938/LUCAS2009\\_C1-Instructions\\_Revised20130925.pdf/](http://ec.europa.eu/eurostat/documents/205002/208938/LUCAS2009_C1-Instructions_Revised20130925.pdf/). – 12.06.2017.
16. Large-scale classification of land cover using retrospective satellite data [Text] / M. Lavreniuk, et al. // *Cybernetics and Systems Analysis*. – 2016. – Vol. 52, Iss. 1. – P. 127-138.
17. Лавренюк, М. С. Метод детектування меж на карті класифікації на основі модифікованого алгоритму Собеля [Текст] / М. С. Лавренюк // *Радіоелектронні і комп'ютерні системи*. – 2017. – № 4 (84). – С. 17-27.
18. Wu, Kesheng. Optimizing connected component labeling algorithms [Electronic resource] / Wu, Kesheng, Ekow, Otoo, Arie Shoshani // *Lawrence Berkeley National Laboratory*. – 2005. – Access mode: <https://escholarship.org/uc/item/7jg5d1zn>. – 12.08.2017.
19. Chrystal, George. On the problem to construct the minimum circle enclosing n given points in the plane [Text] / George Chrystal // *Proceedings of the Edinburgh Mathematical Society*. – 1885. – № 3. – P. 30-33.
20. Rocha, Lourena. Image moments-based structuring and tracking of objects [Electronic resource] / Lourena Rocha, Luiz Velho, Paulo Cezar Pinto Carvalho // *Computer Graphics and Image Processing, 2002. Proceedings. XV Brazilian Symposium on. IEEE, 2002*. – Access mode: <http://ieeexplore.ieee.org/document/1167130/> – 15.07.2017. DOI: 10.1109/SIBGRA.2002.1167130.
21. Stojmenovic, M. Direct ellipse fitting and measuring based on shape boundaries [Text] / Milos Stojmenovic, Nayak Amiya // *Advances in Image and Video Technology*. – 2007. – P. 221-235.
22. Vermeer, M. Statistical uncertainty and error propagation [Electronic resource] / Martin Vermeer. – Access mode: <https://users.aalto.fi/~mvermeer/uncertainty.pdf>. – 22.05.2017.
23. Ramer, Urs. An iterative procedure for the polygonal approximation of plane curves [Text] / Urs Ramer // *Computer graphics and image processing*. – 1972. – Vol. 1, iss. 3. – P. 244-256.
24. Douglas, David H. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature [Text] / David H. Douglas, Thomas K. Peucker // *Cartographica: The International Journal for Geographic Information and Geovisualization*. – 1973. – vol. 10, Iss. 2. – P. 112-122.
25. Beucher, S. Use of watersheds in contour detection [Electronic resource] / Serge Beucher, Christian Lantuéjoul // *Workshop on image processing*

(1979). – Access mode: <http://cmm.ensmp.fr/~beucher/publi/watershed.pdf>. – 11.04.2017.

26. Congalton, Russell G. *A review of assessing the accuracy of classifications of remotely sensed data* [Text] / Russell G. Congalton // *Remote sensing of environment*. – 1991. – Vol. 37, Iss. 1. – P. 35-46.

27. Foody, Giles M. *Classification accuracy comparison: hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority* [Text] / Giles M. Foody // *Remote Sensing of Environment*. – 2009. – Vol. 113, Iss. 8. – P. 1658-1663.

28. Foody, Giles M. *Thematic map comparison* [Text] / Giles M. Foody // *Photogrammetric Engineering & Remote Sensing*. – 2004. – Vol. 70, Iss. 5. – P. 627-633.

29. *Efficiency assessment of multitemporal C-band Radarsat-2 intensity and Landsat-8 surface reflectance satellite imagery for crop classification in Ukraine* [Text] / S. Skakun, et al. // *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. – 2016. – Vol. 9, Iss. 8. – P. 3712-3719.

30. *Regional scale crop mapping using multi-temporal satellite imagery* [Text] / N. Kussul, et al. // *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. – Vol. 40, Iss.7. – 2015. – P. 45–52. DOI:10.5194/isprsarchives-XL-7-W3-45-2015.

## References

1. Huang, Xin, et al. New postprocessing methods for remote sensing image classification: A systematic study. *IEEE Transactions on Geoscience and Remote Sensing*, vol. 52, iss. 11, 2014, pp. 7140-7159.

2. Kolotii, A., et al. Comparison of biophysical and satellite predictors for wheat yield forecasting in Ukraine. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 40, iss.7, 2015, pp. 39-44.

3. Gallego, Francisco Javier, et al. Efficiency assessment of using satellite data for crop area estimation in Ukraine. *International Journal of Applied Earth Observation and Geoinformation*, vol. 29, 2014, pp. 22-30.

4. Kogan, Felix, et al. Winter wheat yield forecasting in Ukraine based on Earth observation, meteorological data and biophysical models. *International Journal of Applied Earth Observation and Geoinformation*, vol. 23, 2013, pp. 192-203.

5. Kussul, Nataliia, et al. Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data. *IEEE Geoscience and Remote Sensing Letters*, vol. 14, iss. 5, 2017, pp. 778-782.

6. Shelestov, Andrii, et al. Exploring Google Earth Engine Platform for Big Data Processing: Classification of Multi-Temporal Satellite Imagery for Crop Mapping. *Frontiers in Earth Science*, vol. 5, 2017, pp. 17.

7. Kussul, Nataliia, et al. Parcel-based crop classification in Ukraine using landsat-8 data and sentinel-1A data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, iss. 6, 2016, pp. 2500-2508.

8. Waldner, François, et al. Towards a set of agrosystem-specific cropland mapping methods to address the global cropland diversity. *International Journal of Remote Sensing*, vol. 37, iss. 14, 2016, pp. 3196-3231.

9. Lavrenyuk, Mykola. Metod ob'yektnoyi fil'tratsiyi kart klasyfikatsiyi zemnoho pokryvu na osnovi morfolohichnykh oznak [Method of object filtering maps of the classification of the earth's surface on the basis of morphological features]. *Induktyvne modelyuvannya skladnykh system*, 2017, no. 9 (unpublished).

10. Löw, Fabian, Christopher Conrad, and Ulrich Michel. Decision fusion and non-parametric classifiers for land use mapping using multi-temporal RapidEye data. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 108, 2015, pp. 191-204.

11. Kussul, N., Shelestov, A., Basarab, R., Skakun, S., Kussul, O., Lavreniuk, M. Geospatial Intelligence and Data Fusion Techniques for Sustainable Development Problems. *In ICTERI*, 2015, pp. 196–203.

12. Kussul, N., Lavreniuk, N., Shelestov, A., Yailymov, B., Butko, I. Land Cover Changes Analysis Based on Deep Machine Learning Technique. *Jour. of Automation and Information Sciences*, 2016, vol. 48, no. 5, pp. 42–54.

13. Kussul, N., Shelestov, A., Lavreniuk, M., Butko, I., Skakun, S. Deep learning approach for large scale land cover mapping based on remote sensing data fusion. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 198-201.

14. Kussul, N., Lavreniuk, M., Shelestov, A., Yailymov, B. Along the season crop classification in Ukraine based on time series of optical and SAR images using ensemble of neural network classifiers. *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2016, pp. 7145-7148.

15. *LUCAS 2009, Land Use / Cover Area Frame Survey*. Available at: [http://ec.europa.eu/eurostat/documents/205002/208938/LUCAS2009\\_C1-Instructions\\_Revised20130925.pdf/](http://ec.europa.eu/eurostat/documents/205002/208938/LUCAS2009_C1-Instructions_Revised20130925.pdf/) (accessed 12.06.2017).

16. Lavreniuk, Mykola, et al. Large-scale classification of land cover using retrospective satellite data. *Cybernetics and Systems Analysis*, vol. 52, iss. 1, 2016, pp. 127-138.

17. Lavreniuk, M. S. Metod detektuvannya mezh na karti klasyfikatsiyi na osnovi modyfikovanoho alhorytmu Sobelya [Borders detection method based on a modified sobel algorithm for crop classification maps]. *Radioelektronni i komp'uterni sistemi - Radioelectronic and computer systems*, 2017, no. 4 (84), pp. 17–27.

18. Kesheng, Wu., Ekow, Otoo., Shoshani, Arie. Optimizing connected component labeling algorithms.

Lawrence Berkeley National Laboratory. Available at: <https://escholarship.org/uc/item/7jg5d1zn> (accessed 12.08.2017).

19. Chrystal, George. On the problem to construct the minimum circle enclosing n given points in the plane. *Proceedings of the Edinburgh Mathematical Society*, 1885, vol. 3, pp. 30-33.

20. Rocha, Lourena, Luiz Velho, and Paulo Cezar Pinto Carvalho. "Image moments-based structuring and tracking of objects. *Computer Graphics and Image Processing, 2002. Proceedings. XV Brazilian Symposium on. IEEE*. Available at: <http://ieeexplore.ieee.org/document/1167130/> (accessed 15.07.2017). DOI: 10.1109/SIBGRA.2002.1167130

21. Stojmenovic, Milos., Nayak, Amiya. Direct ellipse fitting and measuring based on shape boundaries. *Advances in Image and Video Technology*, 2007, pp. 221-235.

22. Vermeer, Martin. "Statistical uncertainty and error propagation. Available at: <https://users.aalto.fi/~mvermeer/uncertainty.pdf> (accessed 22.05.2017).

23. Ramer, Urs. An iterative procedure for the polygonal approximation of plane curves. *Computer graphics and image processing*, 1972, vol. 1, iss. 3, pp. 244-256.

24. Douglas, David H., Peucker, Thomas K. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. *Cartographica: The International Journal for*

*Geographic Information and Geovisualization*, 1973, vol. 10, iss. 2, pp. 112-122.

25. Beucher, Serge, and Christian Lantuéjoul. "Use of watersheds in contour detection. *Workshop on image processing*. Available at: <http://cmm.enscm.fr/~beucher/publi/watershed.pdf> (accessed 11.04.2017).

26. Congalton, Russell G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing of environment*, 1991, vol. 37, iss. 1, pp. 35-46.

27. Foody, Giles M. Classification accuracy comparison: hypothesis tests and the use of confidence intervals in evaluations of difference, equivalence and non-inferiority. *Remote Sensing of Environment*, 2009, vol. 113, iss. 8, pp. 1658-1663.

28. Foody, Giles M. Thematic map comparison. *Photogrammetric Engineering & Remote Sensing*, 2004, vol. 70, iss. 5, pp. 627-633.

29. Skakun, Sergii, et al. Efficiency assessment of multitemporal C-band Radarsat-2 intensity and Landsat-8 surface reflectance satellite imagery for crop classification in Ukraine. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, vol. 9, iss. 8, pp. 3712-3719.

30. Kussul, Natalia, et al. Regional scale crop mapping using multi-temporal satellite imagery. *The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2015, vol. 40, iss. 7, pp. 45-52. DOI:10.5194/isprsarchives-XL-7-W3-45-2015.

Поступила в редакцию 12.12.2017, рассмотрена на редколлегии 14.02.2018

## МЕТОД ОБЪЕКТНОЙ ПОСТОБРАБОТКИ КАРТ КЛАССИФИКАЦИИ С УЧЕТОМ СПЕЦИФИКИ КАЖДОГО КЛАССА

*Н. С. Лавренюк*

Предложен метод фильтрации карт классификации земного покрова, полученных на основе спутниковых снимков высокого разрешения. Основная идея метода - исследование свойств каждого объекта в целом (группы точек с одинаковым значением класса) с учётом специфики каждого класса и, как следствие, подбором порога фильтрации для каждого из них в отдельности. Учитывая специфику карт классификации земного покрова, необходимо сохранить границы между сельскохозяйственными полями. Для этого предложен модифицированный алгоритм Собеля. Точность и эффективность данного метода подтверждается тестированием на независимой выборке, а также путём визуального сравнения с результатами общеизвестных методов фильтрации.

**Ключевые слова:** классификация посевов, постобработка, фильтрация, детектирование границ, алгоритм Собеля.

## OBJECT BASED POSTPROCESSING METHOD FOR CROP CLASSIFICATION MAPS ACCORDING TO EACH CLASS SPECIFICITY

*M. Lavreniuk*

Obtaining reliable and accurate crop classification and land cover map based on satellite data, in particular high resolution data, is one of the most important tasks in remote sensing. Such maps provide basic information for many other applied problems and are vital in remote sensing studies. Despite of which machine learning methods were utilized for maps obtaining: traditional (Random Forest, Support Vector Machine, Multi-layer perceptron, logistic regression) or state-of-the-art approaches (autoencoder, convolutional neural network, recurrent neural network) there is some noise (single pixels or groups and clusters of pixels that wrong classify) on such maps. There are traditional methods for noise reduction, however these methods do not take into account image semantics.

Therefore, they are not effective for filtration land cover and crop classification maps based on satellite images. The most complicated task in the filtering such maps is to preserve boundaries between different agricultural fields and to remove quite big clusters of incorrect classified pixels (objects) and at the same time save small farmer fields that are right classified. Thus, in this paper we proposed new method for postprocessing crop classification map based on algorithm that takes into account each class specificity and as a result utilizes different thresholds for different classes. We proposed investigate each object in classification map independently and decision should be: "is this whole object a noise or not?". We consider each class independent and use connected component labeling technique for discriminating objects from classification map. Further different types of conditions based on sharpness and compactness where proposed for the investigated object.

Accuracy and efficiency of this method with the proposed filtration method have been tested on the independent set and using the visual comparison with the results of utilizing common filters. Also, McNemar statistical test has been conducted to prove the statistically significant gain of utilizing proposed filtration methodology compare to common voting filter.

**Keywords:** crop classification, postprocessing, filtration, connected component labeling, Sobel algorithm

**Лавренюк Микола Сергіович** – аспірант, Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського»; мл. науч. співр., Інститут космічних досліджень НАН України та ДКА України, Київ Україна, e-mail: nick\_93@ukr.net.

**Mykola Lavreniuk** – PhD student at the National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Junior Scientist at Space Research Institute NAS Ukraine and SSA Ukraine, Kyiv, Ukraine, e-mail: nick\_93@ukr.net.