

УДК 681.518:004.93.1

А.С. ДОВБИШ, О.П. ЧЕКАЛОВ, С.С. МАРТИНЕНКО

Сумський державний університет, Україна

ІНТЕЛЕКТУАЛЬНА СИСТЕМА ДІАГНОСТУВАННЯ ОНКОПАТОЛОГІЙ

Розглядається у рамках інформаційно-екстремальної інтелектуальної технології синтез інтелектуальної системи підтримки прийняття рішень для діагностування онкопатологій на основі аналізу морфологічних зображень тканин, одержаних за методом біопсії. Діагностична система має два канали для розпізнавання образів, що формуються за результатами гістологічного аналізу і зображень тканин пацієнта відповідно. Побудована система забезпечує інваріантність алгоритму до деформацій як гістологічних образів, так і морфологічних зображень тканин, що обумовлено довільними початковими умовами зняття інформації.

Ключові слова: Інформаційно-екстремальна інтелектуальна технологія, критерій функціональної ефективності, розпізнавання онкопатологій, система підтримки прийняття рішень, алгоритм екзамену.

Вступ

Однією з причин ускладнення діагностування онкологічних патологій на ранніх стадіях є проблема багатоступовості, що виникає при аналізі векторів-реалізацій образів, які містять велику кількість ознак розпізнавання, що потребує розробки машинних алгоритмів класифікації для систем підтримки прийняття рішень (СППР).

Становлення сучасної теорії аналізу і синтезу інтелектуальних СППР відбувається, головним чином, шляхом подальшого розвитку ідей і методів як розпізнавання образів, так і статистичної теорії прийняття рішень [1 – 3]. Відомі детерміновані методи теорії розпізнавання образів [4,5] характеризуються високою оперативністю, але низькою достовірністю розпізнавання. Статистичні методи [6, 7] потребують наявності навчальних вибірок великих обсягів, що як правило ускладнено на практиці. Структурні (лінгвістичні) методи розпізнавання [8] характеризуються низькою оперативністю в режимі екзамену, тобто безпосереднього розпізнавання в робочому режимі, через необхідність формування образу, що розпізнається. Одними із перспективних методів розпізнавання є методи розроблені у рамках інформаційно-екстремальної інтелектуальної технології (ІЕІ-технологія), що ґрунтується на максимізації інформаційної спроможності системи розпізнавання шляхом введення додаткових інформаційних обмежень у процесі навчання [9 – 11].

У статті розглядається задача інформаційного синтезу в рамках ІЕІ-технології системи підтримки прийняття рішень, що навчається, для діагностування онкопатологій, за гістологічними образами та за

морфологічними зображеннями, з метою побудови безпомилкових, за навчальною вибіркою, вирішальних правил.

Структура та функції СППР при діагностуванні онкопатологій

На рис. 1 показано структуру інтелектуальної СППР для діагностування онкопатологій за результатами гістологічного аналізу. Основним призначенням блоку формування вхідного математичного опису є формування навчальної матриці. Для підвищення достовірності розпізнавання наведена на рис.1 СППР є двоканальною, що дозволяє обробляти як гістологічні дані, так і одержане від цифрового мікроскопа зображення тканини пацієнта, взятої за методом біопсії. Метою навчання СППР є побудова безпомилкових за навчальною вибіркою вирішальних правил шляхом оптимізації параметрів функціонування за відповідним методом ІЕІ-технології [9]. Спочатку оптимізуються вирішальні правила для гістологічних образів за попередньо побудованою ієрархічною структурою їх алфавіту [12]. Після цього будуються вирішальні правила для розпізнавання зображень, що входять в кожну страту відповідного рівня ієрархічної структури. При цьому страта верхнього рівня складається із двох найближчих класів різних типів (доброякісна та злоякісна патології). Сформована в процесі навчання база знань використовується в режимі екзамену, на якому спочатку розпізнається образ, сформований за результатами гістологічного аналізу, а після цього визначається клас зображення тканини пацієнта, що діагностується. Призначенням процедури узгодження є перевір-

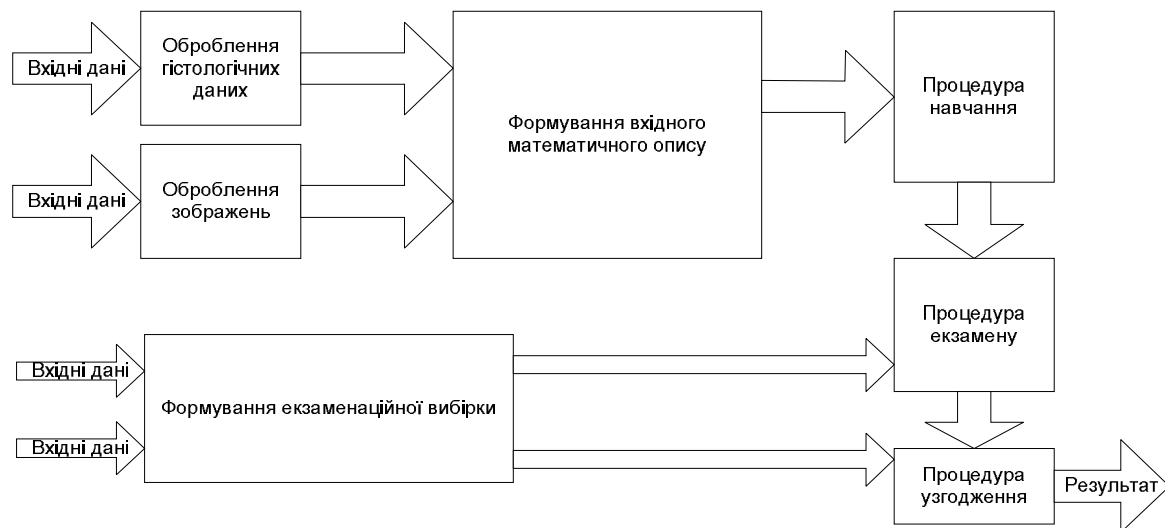


Рис. 1. Структура СППР при діагностуванні онкопатологій

ка відповідності результатів розпізнавання гістологічних реалізацій образу шляхом ідентифікації зображення, що розпізнається, з еталонним зображенням, сформованим на етапі навчання.

Постановка задачі синтезу СППР, що навчається

Розглянемо постановку задачі загального синтезу системи розпізнавання. Нехай функціональна ефективність навчання СППР розпізнаванню реалізацій класу X_m^0 , $m = \overline{1, M}$, характеризується значенням E_m критерію функціональної ефективності (КФЕ). Відома навчальна матриця $\|y_{m,i}^{(j)}\|$, $i = \overline{1, N}$, $j = \overline{1, n}$, де N, n – кількість ознак розпізнавання і випробувань відповідно. Строка матриці $\{y_{m,i}^{(j)} | i = \overline{1, N}\}$ утворює j -ту реалізацію образу, а стовпчик $\{y_{m,i}^{(j)} | j = \overline{1, n}\}$ – навчальну вибірку з генеральної сукупності значень i -ої ознаки розпізнавання.

Треба для структурованого вектора параметрів функціонування системи розпізнавання $g_m = \langle g_{m,1}, \dots, g_{m,q}, \dots, g_{m,Q} \rangle$, які будемо називати параметрами навчання і для яких відомі обмеження $R_q(g_1, \dots, g_Q) \leq 0$, шляхом організації багато циклічних ітераційних процедур знайти оптимальні значення координат вектора g_m , що забезпечують максимум КФЕ навчання СППР:

$$E_{\max}^* = \max_G E_m,$$

де G – область допустимих значень параметрів навчання.

У режимі екзамену СППР повинна приймати з достовірністю, наближеною до асимптотичної, рі-

шення про належність образу, що розпізнається, до відповідного класу із сформованого на етапі навчання алфавіту.

Опис алгоритму розпізнавання онкопатологій

Формування вхідного математичного опису та реалізація алгоритму навчання спочатку здійснюється для гістологічних образів, а потім – для зображень. Як вхідні дані розглядаються масиви результатів гістологічного аналізу та морфологічні зображення тканин, які було отримано за методом біопсії при діагностуванні пацієнтів на онкологічні захворювання. Наведемо узагальнені етапи реалізації інформаційно-екстремального алгоритму навчання СППР при діагностуванні онкопатологій.

Етап 1: Формується бінарна матриця гістологічних даних.

Етап 2: Формується ієрархічна структура класів розпізнавання за гістологічними ознаками [12].

Етап 3: Відновлюються у радіальному базисі простору ознак оптимальні контейнери класів розпізнавання, що входять в одну страту для кожного ієрархічного рівня.

Етап 4: Обробляються вхідні морфологічні зображення в полярній системі координат.

Етап 5: За одержаними у процесі оброблення зображень спектрограмами яскравості для відповідних складових RGB формується вхідний математичний опис, який включає бінарну навчальну матрицю та еталонні вектори-реалізації образів.

Етап 6: За бінарною навчальною матрицею визначаються для кожного зображення найближчі за дистанційним критерієм (кодовою відстанню) класи розпізнавання.

Етап 7: Для кожної пари сусідніх класів реалізується двохциклічний ітераційний алгоритм навчання з

оптимізацією контрольних допусків на ознаки розпізнавання та геометричних параметрів контейнерів.

Етап 8: Відновлюються у радіальному базисі простору ознак оптимальні контейнери класів розпізнавання для заданого алфавіту.

Етап 9: Якщо значення інформаційного КФЕ не досягає максимального граничного значення, то здійснюється оптимізація інших параметрів функціонування СППР (рівні селекції, параметри словника ознак та інше) до тих пір, поки не буде побудовано безпомилкові за навчальною матрицею вирішальні правила.

Приклад розпізнавання зображень

Як вхідні дані розглядалися чотири морфологічні зображення тканин, одержаних за методом біопсії при діагностуванні пацієнтів на рак молочної залози, які показано на рис. 2.

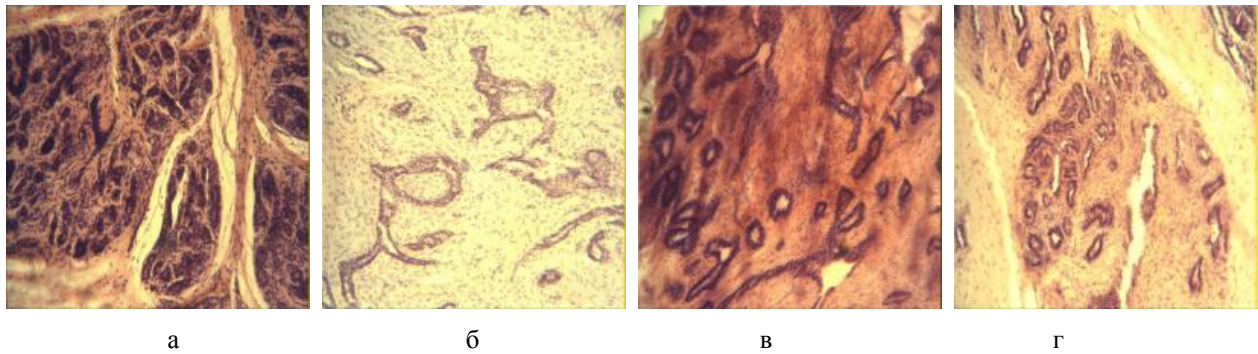


Рис. 2. Морфологічні зображення тканин молочної залози при: а - мазоплазії; б - інтраканалікулярній фіброаденомі; в - періканалікулярній фіброаденомі; г - протоковому фіброаденоматозі

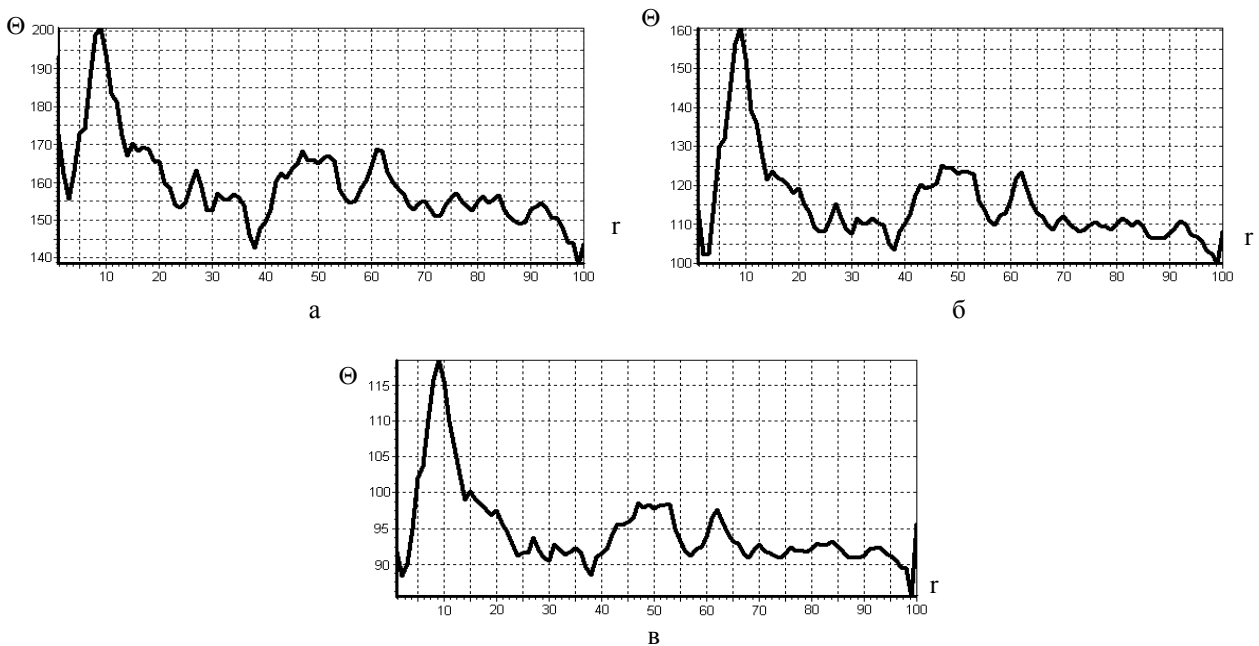


Рис. 3. Складові спектрограми яскравості зображення для класу X_1^0 (за кольорами): а - червона; б - зелена; в - синя

Оброблення зображення відбувалося у полярній системі координат, що робить їх інваріантними до зсуву, повороту та масштабу. Оброблення зображень відбувається шляхом усереднення значень яскравості відповідних складових RGB за формулою

$$\Theta_j = \frac{\sum_{i=1}^N \theta_i}{2\pi R},$$

де Θ_j – числове значення спектру у j -му радіусі, $j = \overline{0, R}$; θ_i – значення кольорової складової у i -му пікселі; N – загальна кількість пікселів в j -му колі; R – радіус кола зчитування.

У результаті оброблення зображень було сформовано спектрограми яскравості складових RGB, які розглядалися як реалізації відповідних образів.

На рис. 3 наведено спектрограми, одержані після оброблення зображення класу X_1^0 (мазоплазія).

Після формування навчальної матриці всі допустимі перетворення в процесі її нормалізації здійснювалися в дискретному просторі ознак за наведеним вище алгоритмом навчання (етапи 6-10). Як КФЕ навчання було використано модифікацію інформаційного критерію Кульбака [13]

$$E_m^{(k)} = \log_2 \left(\frac{2 - (\alpha^{(k)} + \beta^{(k)})}{\alpha^{(k)} + \beta^{(k)}} \right) \left[1 - (\alpha^{(k)} + \beta^{(k)}) \right],$$

де $\alpha^{(k)}$, $\beta^{(k)}$ – помилки першого та другого роду відповідно.

За результатами роботи базового алгоритму навчання було побудовано графіки залежності критерію функціональної ефективності від радіуса контейнера. Вони представлені на рис. 4.

Після аналізу одержаних результатів було при-

йнято рішення про необхідність проведення паралельної оптимізації контрольних допусків з метою підвищення значень КФЕ навчання СППР. На рис. 5 наведено графік залежності усередненого КФЕ від параметра поля допусків δ на ознаки розпізнавання. Аналіз рис. 5 показує, що оптимальне значення параметра поля допусків дорівнює $\delta^* = \pm 6$ градацій яскравості.

Графіки залежності КФЕ від радіуса контейнера, після проведення паралельної оптимізації наведено на рис. 6.

Після проведення навчання з оптимізацією СКД, було реалізовано алгоритм екзамону, результати якого наведено на рис. 7.

Аналіз результатів екзамону показав, що система правильно розпізнала морфологічне зображення і віднесла його до відповідного класу.

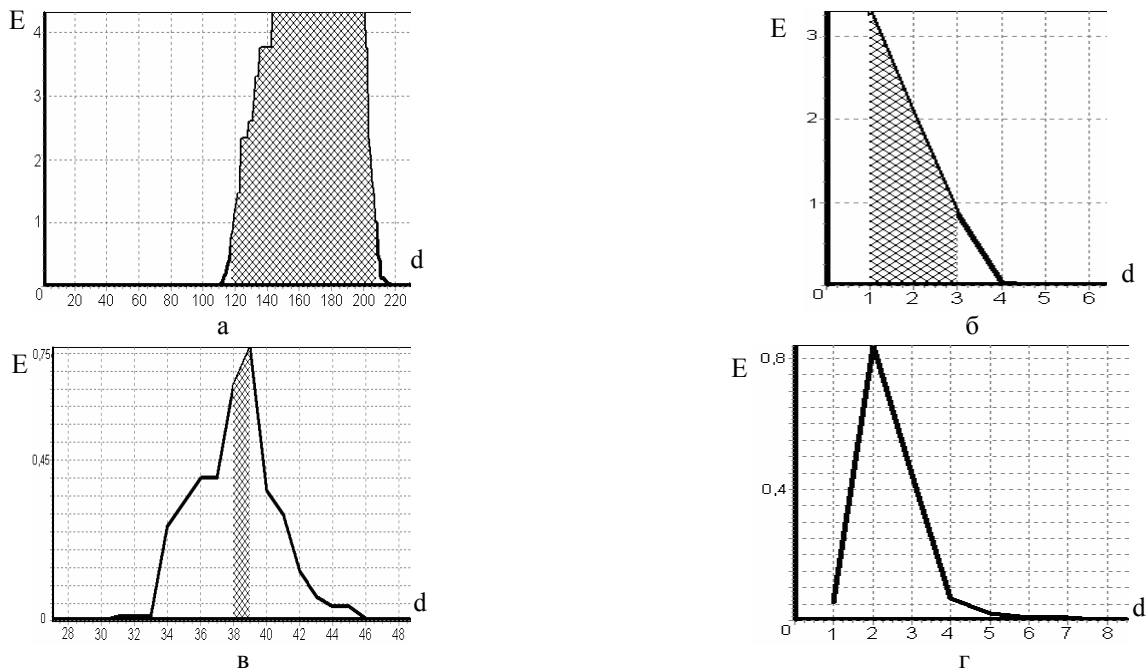


Рис. 4. Графіки залежності КФЕ від радіуса контейнера: а - для класу X_1^0 ; б - для класу X_2^0 ; в - для класу X_3^0 ; г - для класу X_4^0 .

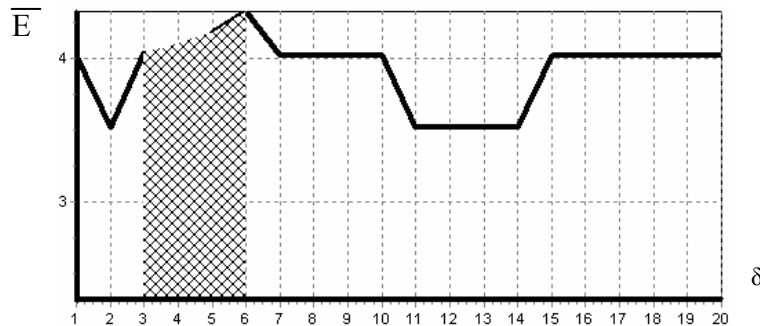


Рис. 5. Графік залежності середнього КФЕ від параметра δ

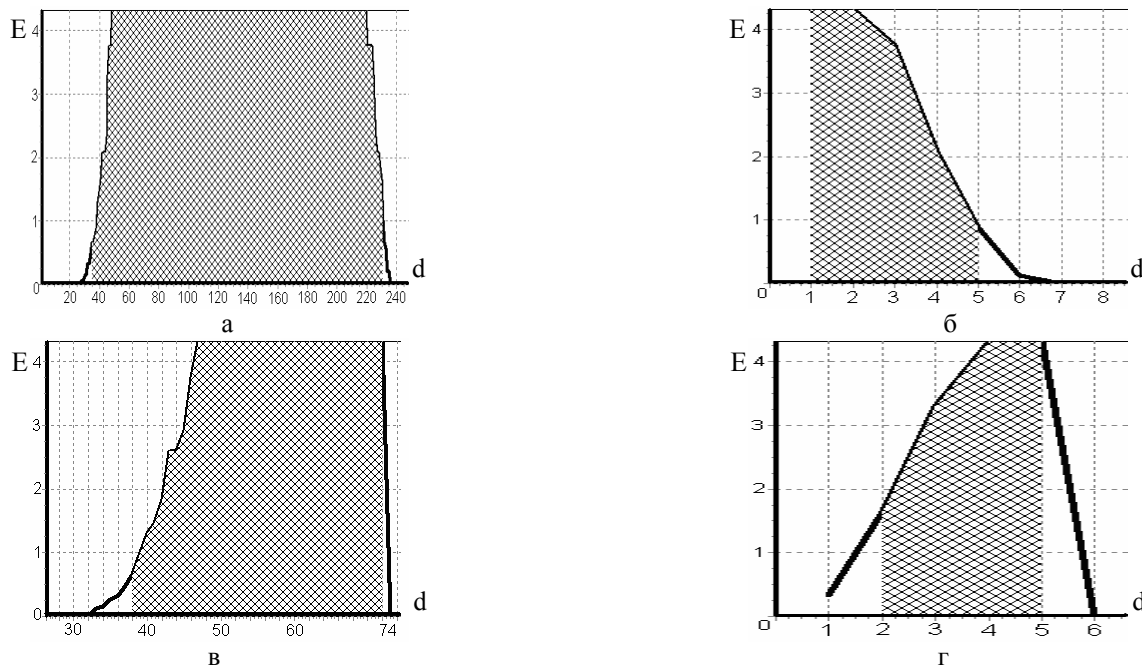
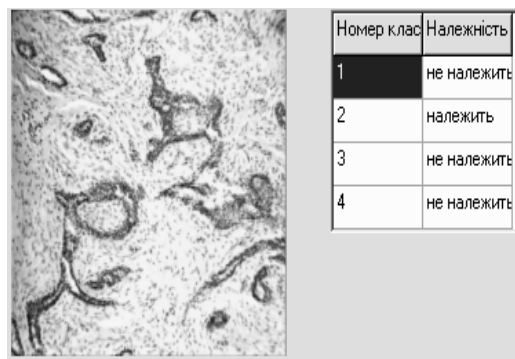


Рис. 6. Графіки залежності КФЕ від радіуса контейнера:

а – для класу X_1^0 ; б – для класу X_2^0 ; в – для класу X_3^0 ; г – для класу X_4^0 Рис. 7. Інтерфейс алгоритму розпізнавання реалізації класу X_2^0

Перспективи розвитку СППР

З метою застосування розробленого інформаційного та програмного забезпечення інтелектуальної системи підтримки прийняття рішень для діагностування онкопатологій необхідно вхідний математичний опис системи формувати за ієрархічною структурою. Це дозволить у значній мірі вирішити проблему багатовимірності алфавіту класів розпізнавання, оскільки відпадає необхідність оптимізації розвитку простору ознак для всього алфавіту. При цьому оптимізацію ієрархічної структури з метою підвищення достовірності та оперативності розпізнавання планується здійснювати так само у рамках ІЕІ-технології.

Висновки

1. Оброблення морфологічних зображень у полярних координатах на етапі формування вхідного

математичного опису СППР, що навчається, дозволяє забезпечити інваріантність алгоритму розпізнавання до деформації зображень типу зсуву, повороту та зміни масштабу.

2. Оптимізація системи контрольних допусків за інформаційно-екстремальним методом навчання інтелектуальної системи підтримки прийняття рішень дозволяє підвищити достовірність розпізнавання онкопатологій.

Література

1. Фукунага К. Введение в статистическую теорию распознавания образов / К. Фукунага – К.: Пер. с англ. – М.: Наука. Главная редакция физико-математической литературы, 1979. – 368 с.
2. Vapnik V. *Statistical Learning Theory*. / V. Vapnik. – John Wiley&Sons. New York, 1998. – 732 p.

3. Красноясовський А.С. Статистична оцінка вибіркової послідовності в задачах автоматичної класифікації. / А.С. Красноясовський, В.В. Марченко // *Современные технологии машиностроения: Тематический сборник научных статей. Отв. ред. Н.В. Захаров.* – К.: ИСМО, Сумы, СумГУ, 1997. – Вып. 1. – С. 141-145.
4. Фу К. Структурные методы в распознавании образов / К. Фу – М.: Мир, 1977. – 320 с.
5. Браверман Э. М. Структурные методы обработки эмпирических данных / Э. М. Браверман, И. Б. Мучник – М.: Наука, 1983. – 464 с.
6. Промышленные роботы: Конструирование, управление, эксплуатация / В.И. Костюк, А.П. Гаврилик, Л.С. Ямпольский, А.Г. Карлов – К.: Вища шк. Головное изд-во, 1985. – 359 с.
7. Васильев В.И. Распознающие системы. Справочник. 2-е изд. / В. И. Васильев – К.: Наук. думка, 1983. – 422 с.
8. Шлезингер М.И. Математические средства обработки изображений / М.И. Шлезингер. – К.: АН УССР, Ин-т кибернетики им. В.М. Глушкова, Наук. думка, 1989. – 196 с.
9. Красноясовський А.С. Інформаційний синтез інтелектуальних систем керування: Підхід, що ґрунтується на методі функціонально-статистичних випробувань / А.С. Красноясовський – Суми: Видавництво СумДУ, 2004. – 261 с.
10. Красноясовський А.С. Класифікаційний аналіз даних: Навчальний посібник / А. С. Красноясовський – Суми: СумДУ, 2002. – 159 с.
11. Красноясовський А.С. Інформаційний синтез прогнозуючої системи підтримки прийняття рішень, що навчається / А.С. Красноясовський // *Проблеми біоники.* – 2002. – Вып. 57. – С. 17-21.
12. Довбиш А.С. Оптимізація параметрів плану навчання системи підтримки прийняття рішень / А.С. Довбиш, В.А. Тронь // *Вісник СумДУ. Серія Технічні науки.* – 2007. – №2 – С. 154-162
13. Красноясовський А. С. Оцінка інформативності ознак розпізнавання у рамках методу функціонально-статистичних випробувань / А. С. Красноясовський, О.Ю. Сластувеский // *Современные технологии машиностроения: Тематический сборник научных статей. Отв. ред. Н. В. Захаров.* – К.: ИСМО, 1997. – Вып.1. – С. 146 – 151.

Надійшла до редакції 16.07.2009

Рецензент: д-р техн. наук, проф., зав. каф. інформатики О.Ю. Соколов, Національний аерокосмічний університет ім. М.Є Жуковського «Харківський авіаційний інститут», Харків, Україна.

ИНТЕЛЛЕКТУАЛЬНАЯ СИСТЕМА ДИАГНОСТИРОВАНИЯ ОНКОПАТОЛОГИЙ

А.С. Довбиш, А.П. Чекалов, С.С. Мартыненко

Рассматривается в рамках информационно-экстремальной интеллектуальной технологии синтез интеллектуальной системы поддержки принятия решений для диагностирования онкопатологий на основе анализа морфологических изображений тканей, полученных методом биопсии. Диагностическая система имеет два канала для распознавания образов, которые формируются по результатам гистологического анализа и изображений тканей пациента соответственно. Для иерархической структуры построены решающие правила, которые обеспечивают инвариантность алгоритма от деформаций как гистологических образов, так и изображений.

Ключевые слова: информационно-экстремальная интеллектуальная технология, критерий функциональной эффективности, распознавание онкопатологий, система поддержки принятия решений, алгоритм экзамена.

ONCOPATHOLOGY DIAGNOSIS INTELLIGENCE SYSTEM

A.S. Dovbysh, O.P. Chekalov, S.S. Martynenko

Synthesis of intelligence decision support system for oncopathology diagnosis, in bounds of information-extreme intelligence technology, on basis of analysis of morphological tissue images, gained by biopsy method. Diagnostic system has two channels for pattern recognition, which formed by histological analysis results and patient tissue images accordingly. For hierarchical structure decision rules was built, it secures algorithm invariance under deformations both histological patterns and images.

Key words: information-extreme intelligence technology, functional efficiency criteria, oncopathology diagnosis, decision support system, exam algorithm.

Довбиш Анатолій Степанович – д-р. техн. наук, професор, завідувач кафедри інформатики Сумського державного університету, Суми, Україна, e-mail: kras@id.sumdu.edu.ua.

Чекалов Олександр Петрович – канд. техн. наук, доцент, доцент кафедри інформатики Сумського державного університету, Суми, Україна, e-mail: AP.Chek@sumdu.edu.ua.

Мартиненко Сергій Сергійович – аспірант кафедри інформатики Сумського державного університету, Суми, Україна, e-mail: smart@unesco.sumdu.edu.ua.