

УДК 681.518:004.93.1

**ІНТЕЛЕКТУАЛЬНА СИСТЕМА ПІДТРИМКИ ПРИЙНЯТТЯ РІШЕНЬ
ДЛЯ ДІАГНОСТУВАННЯ ОНКОПАТОЛОГІЙ**

А. С. Довбиш, д-р техн. наук, професор;

В. О. Жаловага, магістр

Сумський державний університет, м. Суми

Досліджується у рамках інформаційно-екстремальної інтелектуальної технології метод розпізнавання морфологічних зображень тканин, одержаних за результатами біопсії для підвищення достовірності та оперативності установлення діагнозу за умов зсуву та повороту зображень.

Ключові слова: *гістологічні дослідження, морфологічні зображення тканин, система підтримки прийняття рішень, розпізнавання образів, навчання, екзамен, оптимізація, параметри функціонування, функціональна ефективність, інформаційний критерій.*

Исследуется в рамках информационно-экстремальной интеллектуальной технологии метод распознавания морфологических изображений тканей, полученных по результатам биопсии для повышения достоверности и оперативности установления диагноза при условии сдвига и поворота изображений.

Ключевые слова: *гистологические исследования, морфологические изображения, система поддержки принятия решений, распознавание образов, обучение, экзамен, оптимизация, параметры функционирования, функциональная эффективность, информационный критерий.*

ВСТУП

Через ускладненість раннього виявлення онкологічних патологій актуальним є створення інтелектуальної системи підтримки прийняття рішень (СППР) для розпізнавання морфологічних зображень тканин пацієнтів, одержаних за методом біопсії [1]. Це дозволить підвищити достовірність та оперативність діагностування онкопатологій за умов багатовимірності простору ознак. Невисока достовірність існуючих методів розпізнавання зображень у медицині [2,3] обумовлена відсутністю інваріантності навчальної вибірки через зсув, поворот і різний масштаб зображень, одержаних під час медичних дослідженнях, тобто прийняття рішень здійснюється за довільних початкових умов. Тому на практиці виникає необхідність розроблення адаптивних методів розпізнавання образів, які дозволяють здійснювати нормалізацію зображень у процесі навчання. Для вирішення цієї проблеми перспективним є використання ідей і методів інформаційно-екстремальної інтелектуальної технологій (ІЕІ-технології), яка ґрунтується на максимізації інформаційної

спроможності системи підтримки прийняття рішень (СППР), що навчається [4]. У працях [5,6] розглядається в рамках ІЕІ-технології питання побудови вирішальних правил для діагностування онкопатологій шляхом паралельної оптимізації в процесі навчання контрольних допусків на ознаки розпізнавання. При цьому одержані результати не дозволили побудувати безпомилковий за навчальною вибіркою класифікатор.

У статті розглядається питання інформаційно-екстремального синтезу СППР для діагностування онкопатологій шляхом послідовної оптимізації контрольних допусків на ознаки розпізнавання.

ПОСТАНОВКА ЗАВДАННЯ

Метою дослідження є розроблення інформаційно-екстремального методу аналізу та синтезу СППР, що навчається (самонавчається), який передбачає комплекс категорійних математичних моделей, оцінку функціональної ефективності та алгоритми оптимізації параметрів функціонування, засоби інформаційних технологій для діагностування онкопатологій.

Розглянемо постановку завдання загального синтезу системи розпізнавання [2]. Нехай ефективність навчання розпізнавання реалізацій класу X_m^o , $m = \overline{1, M}$, 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. Треба для структурованого вектора параметрів навчання системи розпізнавання $g_m = \langle g_{m,1}, \dots, g_{m,q}, \dots, g_{m,Q} \rangle$ шляхом організації послідовних ітераційних процедур знайти екстремальні значення координат вектора g_m , що забезпечують максимум КФЕ навчання системи розпізнавання:

$$E_{\max}^* = \max_G E_m, \quad (1)$$

де G – область допустимих значень параметрів навчання.

На етапі екзамену необхідно з максимальною достовірністю прийняти рішення про належність реалізації образу, що розпізнається, одному із класів із заданого алфавіту.

МАТЕМАТИЧНА МОДЕЛЬ

Вхідний математичний опис СППР подамо у вигляді структури

$$\Delta_B = \langle G, T, \Omega, Z, Y; \Phi \rangle,$$

де G – простір вхідних сигналів (факторів), які діють на СППР; T – множина моментів часу зняття інформації; Ω – простір ознак розпізнавання; Z – простір можливих станів СППР; Y – вибіркова множина на вході СППР; $F : G \times T \times Z \times \Omega \rightarrow Y$ – оператор формування вхідної навчальної матриці.

Категорійну модель процесу навчання системи розпізнавання за інформаційно-екстремальним методом подамо у вигляді діаграми відображень множин (рис. 1), які застосовуються на етапі навчання при оптимізації геометричних параметрів контейнерів класів розпізнавання та контрольних допусків на ознаки розпізнавання:

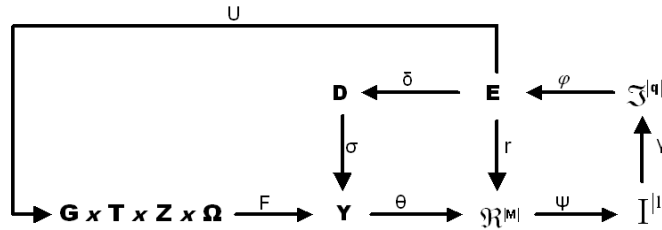


Рисунок 1 – Категорійна модель СППР у режимі навчання

На рис. 1 оператор $\theta : Y \rightarrow \mathfrak{R}^{|M|}$ буде у загальному випадку нечітке розбиття $\tilde{\mathfrak{R}}^{|M|}$, яке допускає перетин класів. Оператор класифікації $\psi : \mathfrak{R}^{|M|} \rightarrow I^{|l|}$, який перевіряє основну статистичну гіпотезу про належність реалізацій $\{x_m^{(j)} | j=1, n\}$ класу X_m^o формує множину гіпотез $I^{|l|}$, де l – кількість статистичних гіпотез. Оператор $\gamma : I^{|l|} \rightarrow \mathfrak{S}^{|q|}$ шляхом оцінки статистичних гіпотез формує множину точнісних характеристик $\mathfrak{S}^{|q|}$, де $q=l^2$ – кількість точнісних характеристик. Оператор $\varphi : \mathfrak{S}^{|q|} \rightarrow E$ обчислює множину значень інформаційного КФЕ, який є функціоналом точнісних характеристик. Контур оптимізації геометричних параметрів розбиття $\mathfrak{R}^{|M|}$, який реалізує базовий алгоритм навчання, шляхом пошуку максимуму КФЕ навчання розпізнавання реалізацій класу X_m^o замикається оператором $r : E \rightarrow \mathfrak{R}^{|M|}$. Оптимізація системи контрольних допусків здійснюється за ітераційною процедурою, в якій задіяні оператори $\theta, \psi, \gamma, \varphi, \delta, \sigma$. Оператор U регламентує процес навчання і дозволяє оптимізувати параметри плану навчання.

Як критерій оптимізації параметрів навчання СППР використовувалася модифікована інформаційна міра Кульбака

$$E = \log_2 \frac{P_t}{P_f} = \log_2 \frac{p(\mu_m)p(\gamma_{1,k}/\mu_m) + p(\mu_{m+1})p(\gamma_{2,k}/\mu_{m+1})}{p(\mu_m)p(\gamma_{2,k}/\mu_m) + p(\mu_{m+1})p(\gamma_{1,k}/\mu_{m+1})},$$

де $P_t^{(k)}$ – повна ймовірність правильного прийняття рішень про належність реалізацій класів X_m^o і X_{m+1}^o контейнера $K_m^o \in X_m^o$, побудованому на k -му кроці навчання СППР розпізнавати реалізації класу X_m^o ; $P_f^{(k)}$ – повна ймовірність помилкового прийняття рішень, гіпотези $\gamma_{1,k}$ та $\gamma_{2,k}$ – гіпотези про належність контейнера $K_{m,k}^o$ реалізацій відповідно класів X_m^o і X_{m+1}^o .

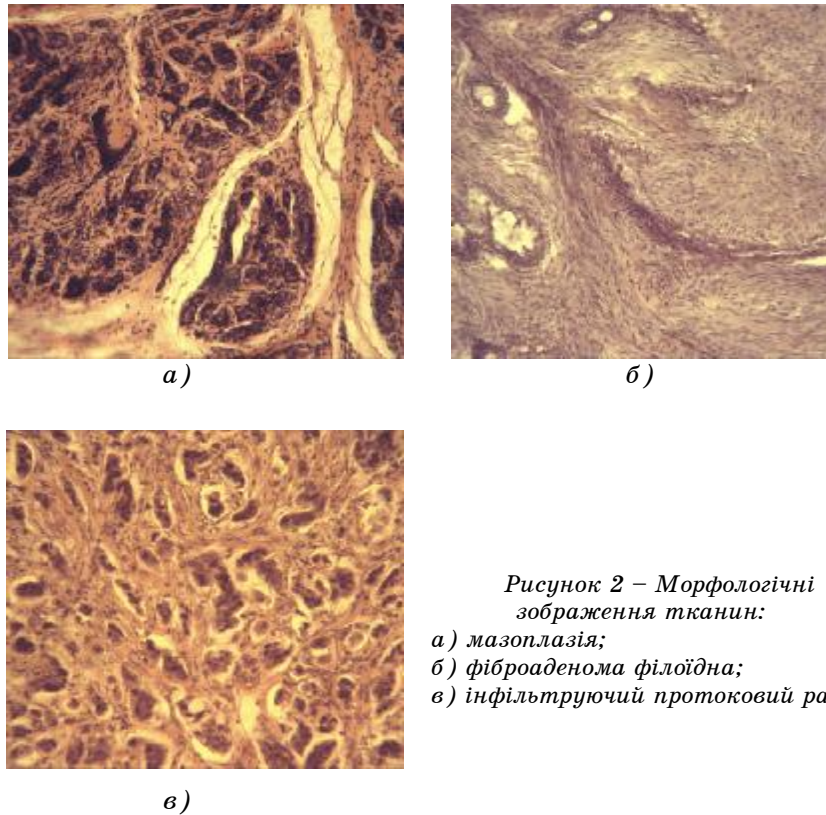
При допущенні, що $p(\mu_m) = p(\mu_{m+1}) = 0,5$, міра Кульбака остаточно набирає вигляду

$$\begin{aligned} E &= 0,5 \log_2 \left(\frac{D_1 + D_2}{\alpha^{(k)} + \beta^{(k)}} \right) \left[(D_1 + D_2) - (\alpha + \beta) \right] = \\ &= \log_2 \left(\frac{2 - (\alpha + \beta)}{\alpha + \beta} \right) \left[1 - (\alpha + \beta) \right], \end{aligned} \quad (2)$$

де D_1 , D_2 - перша та друга достовірності відповідно; α, β - помилка першого та другого родів відповідно.

ФОРМУВАННЯ ВХІДНОГО МАТЕМАТИЧНОГО ОПИСУ

Як вхідні дані використовувалися зображення тканин пацієнтів, одержані за методом біопсії при діагностуванні пацієнтів на онкологічні захворювання, які для прикладу показано на рис.2 [3].



Оброблення зображень здійснювалось у полярній системі координат, яка забезпечує інваріантність зображень до зсуву, повороту і масштабу за формулою

$$\Theta_j = \frac{\sum_{i=1}^N \theta_i}{2\pi R} , \quad (3)$$

де Θ_j – числове значення спектра для всіх пікселів кола j -го радіуса, $j = \overline{0, R}$; 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 відповідного

образу. На рис. 3 наведено спектрограми яскравості, одержані після оброблення зображення класу X_4 (фіброаденома філоїдна).

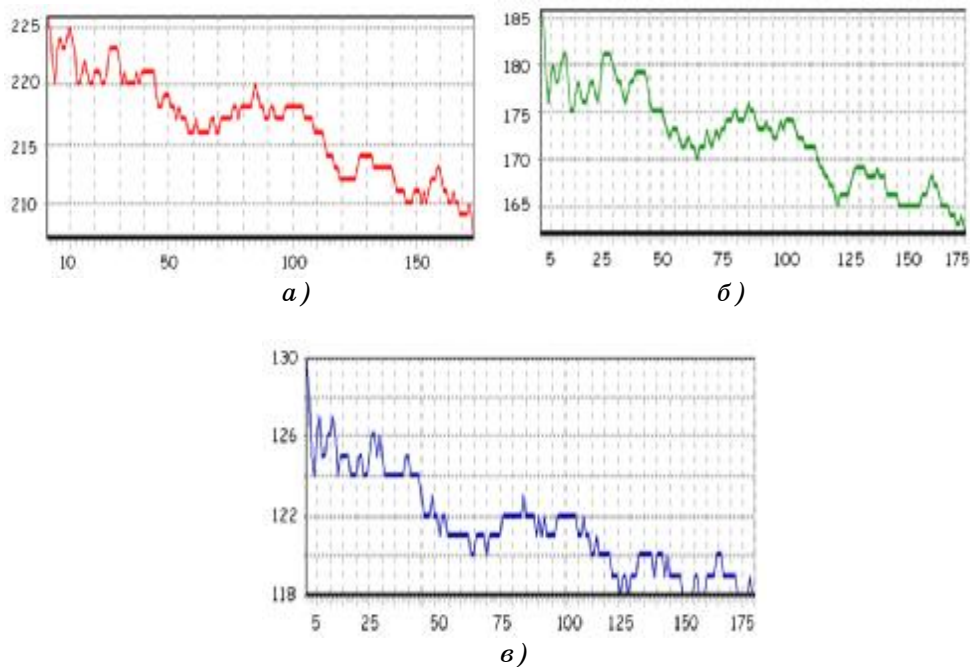


Рисунок 3 – Спектри складових RGB -зображення тканини для класу X_4 : а) червона складова; б) зелена складова; в) синя складова

РЕЗУЛЬТАТИ ФІЗИЧНОГО МОДЕЛЮВАННЯ

Базовий алгоритм навчання є ітераційною процедурою пошуку глобального максимуму інформаційного КФЕ в робочій області визначення його функції, де помилки першого та другого родів менші відповідно від першої та другої достовірностей:

$$d_m^* = \arg \max_{\{d\}} E_m, \quad (4)$$

де $\{d\} = \{0, 1, \dots, d < d(x_m \oplus x_c)\}$ - множина радіусів концентрованих гіперсфер, центр яких визначається вершиною $x_m \in X_m^0$; $d(x_m \oplus x_c)$ - кодова відстань між центрами класів X_m^0 і X_c^0 ; E_m - значення КФЕ навчання СППР розпізнавати реалізації класу X_m^0 в робочій області його визначення.

За результатами роботи базового алгоритму навчання [4] було побудовано графіки залежності усередненого значення критерію функціональної ефективності від радіуса контейнера (рис. 4).

Темною ділянкою графіка (рис. 4) позначено робочу область визначення функції критерію (2), в якій здійснюється пошук глобального максимуму критерію. Аналіз рис.4 показує, що максимальне значення КФЕ дорівнює $E_1^* = 3,85$. Для підвищення достовірності вирішальних

правил було проведено оптимізацію системи контрольних допусків (СКД) на ознаки розпізнавання [6]. На рис. 5 наведено графік залежності усередненого критерію (2) від параметра контрольного поля допусків delta під час реалізації алгоритму паралельної оптимізації СКД.

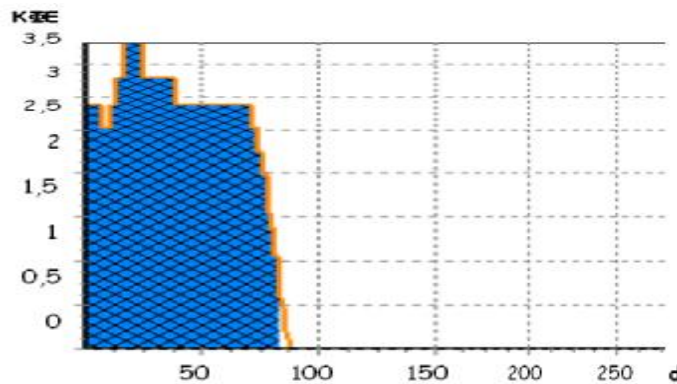


Рисунок 4 – Графік залежності КФЕ від радіуса контейнера для класу X_4

У результаті проведення паралельної оптимізації (рис.5) максимальне значення КФЕ дорівнювало 6,2; при значенні параметра delta 23.

Для підвищення значення інформаційного критерію (2) було здійснено послідовну оптимізацію СКД на ознаки розпізнавання за алгоритмом

$$\{\delta_{K,i}^*\} = \langle \arg\{\max_{G_{\delta_i}} \max_{G_E} \left[\bigotimes_{l=1}^L \max_{G_{d_1}} E_1^{(l)} \right] \} \rangle, \quad i = \overline{1, N}, \quad (5)$$

де $G_{\delta_i}, G_E, G_{d_1}$ – області допустимих значень поля контрольних допусків для i -ї ознаки, критерію оптимізації і кодової відстані d_1 відповідно; \otimes – символ операції повторення; L – кількість прогонів ітераційної процедури оптимізації СКД.

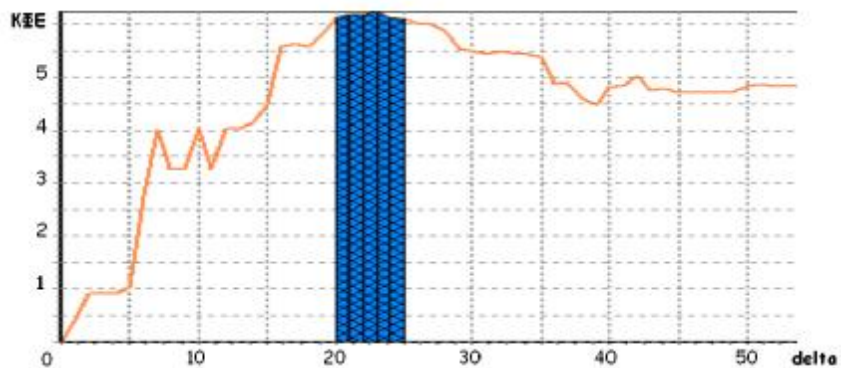


Рисунок 5 – Графік залежності КФЕ від параметра поля контрольних допусків

Після реалізації алгоритму (5) усереднене максимальне значення КФЕ дорівнює 7,02 замість 6,2 при паралельній оптимізації.

На рис. 6 наведено результати послідовної оптимізації СКД для кожної ознаки.

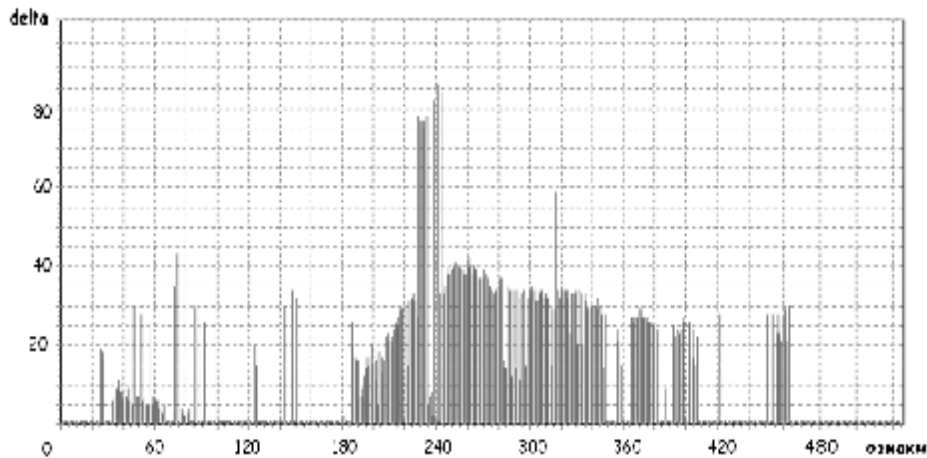


Рисунок 6 – Система оптимальних контрольних допусків на ознаки розпізнавання

Для проведення екзамену було вибрано реалізацію зображення класу “інфільтруючий рак”, яка належала іншому пацієнту і відрізнялася від реалізації, показаної на рис. 2 б. В режимі екзамену зображення, що розпізнавалося, було віднесене до правильного класу за геометричною (дистанційною) функцією належності реалізації образу. Наприклад, для класу X_m^o така функція належності для гіперсферичного класифікатора має вигляд

$$\mu_m = 1 - \frac{d(x_m^* \oplus x^{(j)})}{d_m^*}, \quad (6)$$

де x_m^* , $x^{(j)}$ – еталонний вектор-реалізація класу X_m^o і реалізація класу, що розпізнається, відповідно; d_m^* – оптимальний радіус контейнера класу X_m^o .

При цьому значення функції (6) становить 0,885; що свідчить про деформованість зображення щодо еталонного зображення.

Таким чином, розроблене інформаційне та програмне забезпечення дозволило реалізувати здатну навчатися в рамках інформаційно-екстремального методу систему підтримки прийняття рішень для діагностування онкопатологій. Результати контрольного прикладу в режимі екзамену СППР, тобто безпосереднього розпізнавання, підтвердили достовірність розпізнавання морфологічних зображень.

ВИСНОВКИ

1. Досліджено інформаційно – екстремальний метод розпізнавання морфологічних зображень тканин, одержаних за результатами біопсії, який складається з вхідного опису математичної моделі та алгоритму оптимізації параметрів функціонування СППР, що навчається.

2. Оброблення зображень здійснюється в полярних координатах, що забезпечує їх інваріантність до зсуву та зміни масштабу.

3. Розроблено і програмно реалізовано інформаційно-екстремальний алгоритм навчання системи підтримки прийняття рішень, що дозволило побудувати вирішальні правила для діагностування онкопатологій.

4. Реалізовано алгоритми паралельної та послідовної оптимізації системи контрольних допусків на ознаки розпізнавання та одержано задовільні результати, що дозволило підвищити достовірність розпізнавання морфологічних зображень.

5. Розроблено за інформаційно-екстремальним методом засоби інформаційної технології проектування системи підтримки прийняття рішень, що навчається.

6. Результати контрольного прикладу в режимі екзамену СППР, тобто безпосереднього розпізнавання, підтвердили достовірність розпізнавання гістологічних векторів-реалізацій класів розпізнавання і морфологічних зображень.

SUMMARY

INTELLECTUAL SYSTEM OF SUPPORT OF DECISION-MAKING FOR ONCOLOGIC DISEASES DIAGNOSTICATION

A.S. Dovbysh, V.O. Zhalovaga
Sumy State University, Sumy

The article describes a method of recognition of morphological images obtained through biopsy in context of information-extreme intellectual technology, to increase authenticity and timeliness of oncologic diseases diagnosing subject to the condition change and turn of images.

Key words: *histological researches morphological imafes, system of support decision-making, patterns recognition, learning, examination, optimization, parameters of functioning, functional efficiency informative criterion*

СПИСОК ЛІТЕРАТУРИ

1. Экспертная ситема для гистологической диагностики опухолей щитовидной железы / Ротин Д. Л., Петровичев Н.Н., Павловская А.И., Никитаев В.Г., Бердникович Е.Ю., Проничев А.Н., Попов Д.А. // Архив патологии, 2004. — № 6.
2. Шлезингер М.И. Математические средства обработки изображений / АН УССР, Ин-т кибернетики им. В.М. Глушкова. – Киев: Наук. думка, 1989. - 196 с.
3. Анисимов Б.В. Распознавание и цифровая обработка изображений / Б.В. Анисимов, В.Д. Курганов, В.К. Злобин.– М.: Высшая школа, 1983.–256 с.
4. Краснопоясовський А. С. Інформаційний синтез інтелектуальних систем керування: Підхід, що ґрунтується на методі функціонально-статистичних випробувань. - Суми: Видавництво СумДУ, 2004. - 261 с.
5. Довбиш А.С. Оптимізація параметрів навчання системи підтримки прийняття рішень / А.С. Довбиш, В.А. Тронь // Вісник СумДУ. Технічні науки. - 2007.– №2.– С. 154-162.
6. Краснопоясовський А.С. Оптимізація контрольних допусків на ознаки розпізнавання за МФСВ // Штучний інтелект. - 2003. – №1.– С. 53 – 62.

Надійшла до редакції 16 липня 2009 р.