

УДК 621.391

Д.Г. Медведєв, О.В. Бурлак, В.О. Головка
Національний авіаційний університет
Криворізький державний педагогічний університет
dimakrrog@rambler.ru

Дискримінантна функція параметризованих еозинофілів

В роботі представлено спеціалізовані способи та алгоритми параметризації цифрових зображень еозинофілів для задач медичної діагностики. Розроблено алгоритм отримання контуру еозинофіла як сплайн-функції та визначення його геометричних параметрів. Отримано дискримінантну функцію для параметризованих еозинофілів.

Розпізнавання образів, еозинофіл, сплайн

Постановка проблеми

За даними Міністерства охорони здоров'я України протягом останніх 10 років в Україні спостерігається поширення та ускладнення перебігу алергічних захворювань у дітей. Близько 18-20% їх мають різні за локалізацією та формами прояву алергічні захворювання. Підвищився рівень поширеності бронхіальної астми і в різних регіонах він коливається від 0,9 до 6,8 та atopічного дерматиту від 2,1 до 12,8 на 1000 дитячого населення. Достеменно доведений зв'язок зростання алергічних захворювань з погіршенням екологічної ситуації. [1]

Від запобігання та раннього виявлення цих захворювань залежить їх подальший прогноз та тактика запобігання та лікування.

На кафедрі педіатрії Дніпропетровської медичної академії розроблений метод прогнозування імунного статусу новонароджених. Цей метод включає збір та аналіз показників крові. Ці діагностичні показники додатково відбирають у великих гранульованих лімфоцитах периферійної крові, визначають серед них рівномірність забарвлення цитоплазми, характер зернистості, кількість гранул, як морфологічні ознаки [2]. Також розроблений метод прогнозування загострення алергічних процесів через виявлення таких елементів крові як еозинофіли. Цей метод прогнозування алергічного стану у дітей шляхом диференціювання показників молодих несементованих і зрілих еозинофілів з сегментованим ядром підвищує достовірність та оперативність раннього прогнозування [3].

Методи дозволяють економити час і кошти на аналіз та оптимізувати диспансеризацію дітей групи ризику.

Аналіз останніх досліджень і публікацій

Підвищити якість методу (знизити рівень помилок та зменшити час та трудомісткість налізу) можна за рахунок застосування

інформаційних технологій розпізнавання та параметризації. Розробками технологій розпізнавання, класифікації та параметризації біологічних об'єктів займаються в Донецькому технічному університеті [4], Biomedical Imaging Group [5]. Однак, застосувати ці результати безпосередньо для задач оцінки та прогнозування імунного статусу на основі морфології еозинофілів проблематично. Параметризація еозинофілів для задач визначення імунного статусу потребує математичного опису контурів еозинофіла та визначення на його основі прогностично важливих параметрів в автоматичному режимі.

Цілі статті. В процесі створення експертної системи розпізнавання та параметризації еозинофілів розв'язано наступні задачі:

1. Попередня обробка зображення з метою покращення його якості.
2. Розпізнавання еозинофіла по його цифровому зображенню.
3. Побудова сплайн-моделі контуру еозинофіла для його уточнення.
4. Визначення форми досліджуваного тіла.
5. Обчислення площі та довжини контуру клітини.
6. Визначення дискримінантної функції для отриманих коефіцієнтів параметризованих еозинофілів.

Основний матеріал дослідження

Метою попередньої обробки є покращення якості отриманого зображення (Рис. 1). Знімки, зазвичай, недостатньо чіткі та контрастні, а також несуть в собі багато зайвої інформації – шумів (пил, що осів на скельця мазків, інші чинники). Поліпшити чіткість та контрастність зображення можна апаратними методами, наприклад заміною відеокамери, на ту, що спеціалізується на зйомці зображень під мікроскопом. Але це дорогий спосіб, який не вирішує проблему виключення шумів. Тому доцільніше скористатися

алгоритмічними методами поліпшення якості зображення.

Зменшення ізольованих перешкод з мінімальним розмиванням контрасту забезпечується за рахунок одновимірної медіанної фільтрації по п'яти точках. Оконтурування еозинофіла виконується за допомогою нелінійних операторів Робертса і Собеля.

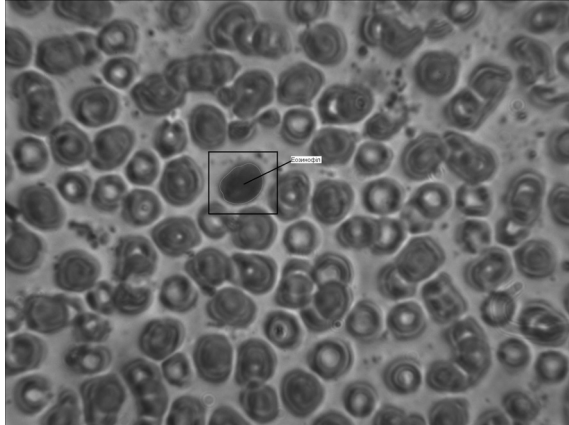


Рисунок 1 - Зображення еозинофіла.

Проте в загальному випадку це не дає змоги досягти оконтурювання саме еозинофіла (присутні й інші елементи крові). Окрім того, оператор потребує вибору порогового значення. Для роботи оператора зображення переводиться у формат YUV. Оператор застосовується до сигналу яскравості Y.

Еозинофіл відрізняється від загального фону більшою яскравістю, а від інших клітин – забарвленням (навіть візуально його можна виокремити від фону). Враховуючи лише яскравість не можна відрізнити еозинофіл та визначити його важливі характеристики. Тому необхідно приймати до уваги й компоненти кольоровості U (синьо-жовта) V (червоно-блакитна). Для цього до бінарного зображення яскравості (приймаємо за компоненту Y) додамо компоненти кольоровості у незмінному вигляді. Отримуємо кольорове зображення з іншим розподілом кольорів за рахунок бінарного перетворення яскравості. На ньому чітко виділені ті ж контури, але еозинофіл стане білого кольору, а інший фон має жовтуватий колір.

Завдяки різниці в кольорі виділяємо еозинофіл. Для цього розроблено алгоритм послідовної стратегії перебору. Знаходяться приблизні межі «білого тіла» – контур еозинофіла. Знаходяться білі точки, навколо яких є достатня кількість (не менше 5) таких же білих точок. Це виключає окремі білі точки серед кольорового фону. Далі рухаємось в чотирьох напрямках, шукаючи чорні точки.[6]

Оскільки контур є розмита лінія, що має розриви, то для його уточнення застосовуємо апроксимацію за допомогою сплайн-функцій.

Особливість контуру еозинофіла в тому, що він є замкнутою лінією. Тому не можливо використовувати декартову систему координат і виконується перехід до полярної системи координат (Рис.2).

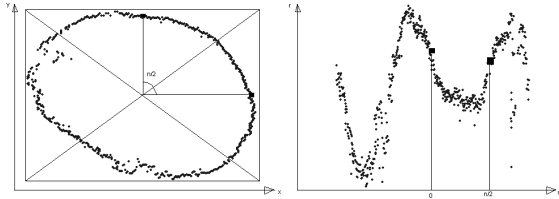


Рисунок 2 - Виділений контур еозинофіла в декартовій та полярній СК

Оскільки функцію контуру можна розглядати як періодичну модифікуємо розрахункові вирази [7] кубічного ермітового сплайну для періодичних крайових умов. Особливістю періодичного сплайну є те, що перед першим фрагментом знаходиться останній, а за останнім фрагментом іде перший. Для обчислень базисних функцій сплайну використовуються наступні формули:

$$\begin{aligned} X_0(x) &= -p_1 p_4 b, & X_1(x) &= p_1 c - p_2 d + b, \\ X_2(x) &= -p_4 c - p_3 d + a, & X_3(x) &= -p_3 p_2 d \\ a &= (x - x_j) / h_j, & b &= 1 - a, \\ c &= a b^2, & d &= a^2 b, \end{aligned}$$

$$p_1 = \frac{h_j}{h_{j-1}}, \quad p_2 = \frac{h_j}{h_{j+1} + h_j}, \quad p_3 = \frac{h_j}{h_{j+1}}, \quad p_4 = \frac{h_j}{h_{j-1} + h_j}$$

$$h_{j-1} = x_j - x_{j-1}, \quad h_j = x_{j+1} - x_j, \quad h_{j+1} = x_{j+2} - x_{j+1}.$$

Труднощі розрахунку на крайніх фрагментах розв'язуються наступним чином:

$$\text{для першого: } h_{j-1} = x_R - x_{R-1};$$

$$\text{для останнього: } h_j = x_1 - x_0, \quad h_{j+1} = x_2 - x_1;$$

$$\text{для передостаннього: } h_{j+1} = x_1 - x_0,$$

де R – кількість фрагментів сплайну,

j – номер поточного фрагмента.

В загальному вигляді, для сплайна, який складається з R фрагментів, маємо вираз:

$$S(x) = \sum_{j=0}^R f(x_j) X_j(x), \quad (1),$$

де $X_j(x)$ локальна базисна сплайн-функція, що складається з чотирьох фрагментів:

$$X_j(x) = \begin{cases} X_{0,j-1}(x), & x \in [x_{j-1}, x_j] \\ X_{1,j}(x), & x \in [x_j, x_{j+1}] \\ X_{2,j+1}(x), & x \in [x_{j+1}, x_{j+2}] \\ X_{3,j+2}(x), & x \in [x_{j+2}, x_{j+3}] \\ 0, & x \notin [x_{j-1}, x_{j+3}] \end{cases}$$

Значення сплайну в довільній точці, що належить j-му фрагменту дорівнює $S(x) = f(x_{j-1})X_{0,j-1}(x) + f(x_j)X_{1,j}(x) + f(x_{j+1})X_{2,j+1}(x) + f(x_{j+2})X_{3,j+2}(x)$

Рухаючи вузли, крім фіксованих, знаходимо таке їх положення, при якому нев'язка мінімальна. Якщо задана точність досягнута, то оптимізація контуру завершується (рис 4). Якщо

точність не досягнута, то вибирається фрагмент з найбільшою нев'язкою та кількістю точок не менше десяти, і додається в нього новий вузол.

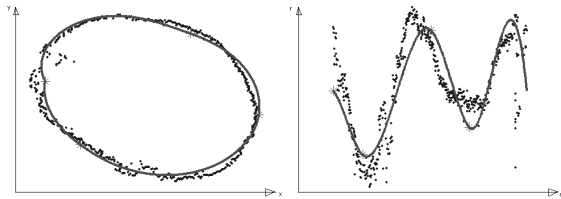


Рисунок 3 - Контур еозинофіла до оптимізації.

Завдяки використанню сплайнів для уточнення контуру еозинофіла не потрібно визначати похибку обчислень, тому що обчислення ведуться до заданої точності. На даному етапі використання сплайнів дало значну перевагу порівняно з іншими методами.

Подальша параметризація еозинофіла полягає у визначенні таких його морфологічних ознак, які визначають імунний статус дитини, а саме:

1. форма клітини (кругла, овальна, неправильна),
2. індекс видовженості клітини (відношення великого до малого діаметрів клітини),
3. площа еозинофіла,
4. довжина контуру.

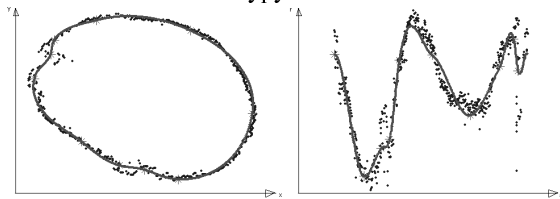


Рисунок 4 - Оптимізований контур еозинофіла.

Діаметром кола є відрізок, що проходить через центр кола та кінці якого належать колу. Коло має безліч діаметрів. Еліпс має два діаметра – велика та мала вісь. Діаметри еліпса перпендикулярні. Великий діаметр є найдовшим відрізком, кінці якого належать еліпсу. Малим діаметром є найдовший відрізок, перпендикулярний до великого діаметра. Еозинофіл здорової людини має форму круга. Якщо до організму потрапив вірус, то форма еозинофіла змінюється (в певних місцях розширюється, в інших – звужується) і може бути як овальною, так і неправильною (рис 5).

Тому великим діаметром еозинофіла вважатимемо найдовший серед відрізків, кінці якого належать контуру досліджуваного тіла. Пошук великого діаметра виконується перебором довжин всіх вказаних відрізків.

Малим діаметром, за аналогією з еліпсом, вважатимемо найдовший відрізок, серед перпендикулярних до великого діаметра. Його пошук також виконується перебором. Перпендикулярність перевіряється за допомогою скалярного добутку (рис 6).

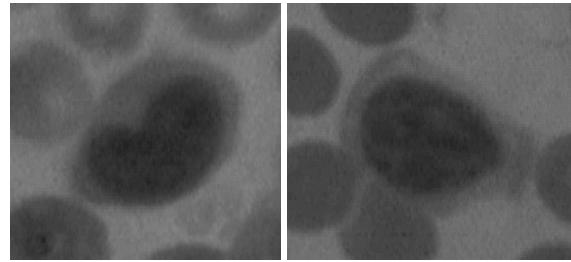


Рисунок 5 - Еозинофіли різної форми

Індекс видовженості клітини визначається відношенням великого діаметра до малого.

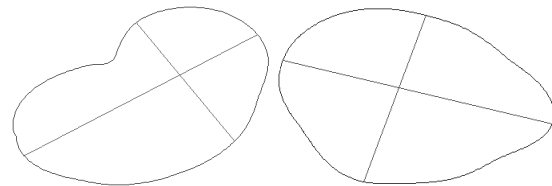


Рисунок 6 - Діаметри еозинофілів різних форм.

Якщо індекс видовженості еозинофіла належить проміжку (1, 1.1) і довжина контуру еозинофіла приблизно дорівнює πd_v , де d_v – великий діаметр, то форма є круглою. Якщо діаметри точкою перетину діляться навпіл, то форма є овальною. Форма є неправильною у всіх останніх випадках.

За формулою (1) сплайн представляється як сума базисних функцій. Але для визначення площі клітини та довжини його контуру необхідно перейти до представлення сплайну у вигляді кубічних поліномів окремо на кожному фрагменті.

Для спрощення цієї задачі було зафіксовано два вузли: самий лівий та самий правий. В полярній системі координат точки впорядковувалися по мірі зростання полярного кута. При переході до декартової системи координат впорядкованість не зазнавала змін. Тому, в більшості випадків, ті точки, що знаходяться від лівого зафіксовано до правого вузла, утворюють нижню частину контуру. Ті точки, що знаходяться від правого зафіксованого вузла до кінця сплайну утворюють верхню частину контуру.

Далі на кожному фрагменті вибираються чотири точки. Фрагмент містить від 5 до 150 точок. Якщо взяти підряд чотири точки, то різниця між їхніми значеннями може скласти лише одну десятитисячну. Це може призвести до погіршення точності обчислень. Якщо розбити фрагмент на чотири частини однакової довжини, щоб потім взяти по одній точці з кожної, то стає можливим випадок, коли в певну частину не потрапить жодної точки контуру досліджуваної клітини. Тому кожний фрагмент розділяється на чотири частини не по довжині, а по кількості точок. З кожної частини вибирається перша точка. По цих чотирьох точках виконується інтерполяція, в результаті якої отримуються

коефіцієнти кубічних поліномів сплайн-функції на кожному окремому фрагменті. На рис 7 зображено знайдений апроксимацією контур клітини та графіки отриманих кубічних поліномів.

Представлення сплайну у формі кубічних поліномів має наступні переваги. По-перше, тепер для побудови експертної системи та створення бази даних значно зменшується кількість інформації, яку потрібно зберігати. Оптимізований контур еозинофіла містить до двох тисяч точок. В середньому за описаним алгоритмом сплайн складається з 10 – 13 вузлів. Тобто тепер треба зберігати координати цієї кількості вузлів (від 10 до 13) та коефіцієнтів кубічних поліномів. Різниця в кількості даних, що зберігаються, очевидна.

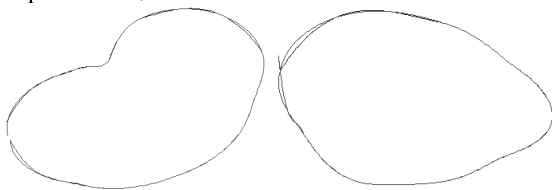


Рисунок 7 - Контур клітини у вигляді кубічних поліномів

По-друге, представляючи сплайн у вигляді окремої функції на кожному фрагменті, обчислення точного значення площі клітини та довжини її контуру зводиться до обчислення означеного інтеграла від кубічного поліному.

Площа досліджуваної клітини дорівнює різниці інтегралів від верхньої та нижньої частини його контуру.

Довжина лінії будь-якої функції визначається за формулою:

$$l = \int_m^n \sqrt{1 + (f'(x))^2} dx.$$

В даному випадку $f(x) = ax^3 + bx^2 + cx + d$ і формула для обчислення довжини контуру прийме наступний

вид:
$$l = \int_m^n \sqrt{9a^2x^4 + 12abx^3 + (4b^2 + 6ac)x^2 + 4bcx + c^2 + 1} dx$$

(2)

Але не можливо взяти інтеграл (2) аналітично. Тому визначення довжини контуру еозинофіла виконується за допомогою чисельних методів. Інтеграл (2) обчислюється за методом трапеції.

Досліджуючи морфологічні ознаки великих гранульованих лімфоцитів (еозинофілів) можна визначити імунний стан людини, тобто стадію боротьби організму з вірусами. Морфологічні ознаки еозинофілів здорової та хворої людини суттєво відрізняються.

Для побудови експертної системи розпізнавання та параметризації еозинофілів треба визначити дискримінантну функцію. Її завдання полягає в тому, щоб за морфологічними ознаками еозинофіла $(x_1, \dots, x_p, p = 5)$ віднести

деяку дитину до одного з $k=2$ класів: хворих або здорових.[8]

Шукана дискримінантна функція має наступний вигляд:
$$Y(x) = \sum_{i=1}^p \beta_i x_i$$

Дана задача розв'язується в припущенні багатовимірних нормальних розподілів з невідомими статистичними параметрами розподілів.

На даній стадії дослідження було визначено такі морфологічні ознаки:

1. Форма (кругла, овальна, неправильна);
2. Великий діаметр;
3. Малий діаметр;
4. Площа клітини;
5. Довжина контуру.

Маємо об'єкт, якому відповідає вектор станів $x = (x_1, \dots, x_p)$, і який треба віднести або до класу W_1 з розподілом $N(m_1^{p \times 1}, \Sigma^{p \times p})$, або $W_2: N(m_2^{p \times 1}, \Sigma^{p \times p})$. Так як x_{11}, \dots, x_{1n_1} та x_{21}, \dots, x_{2n_2} – незалежні випадкові вибірки із W_1 та W_2 відповідно, то можна оцінити m_i

вибірковим вектором середніх $\bar{x}_i = (x_{i1}, \dots, x_{ip})'$, $i = 1, \dots, k$, а Σ – об'єднаною вибірковою коваріаційною матрицею $S = (s_{jh})$, $j = 1, \dots, p$, $h = 1, \dots, p$. Елементи коваріаційної матриці обчислюються за формулою:

$$S_{jh} = \sum_{a=1}^k \sum_b^{n_a} (x_{jab} - \bar{x}_{ja})(x_{hab} - \bar{x}_{ha}) \quad [9]$$

Для знаходження коефіцієнтів β_i необхідно розв'язати наступну систему рівнянь:

$$\begin{cases} \beta_1 s_{11} + \dots + \beta_p s_{1p} = m_{11} - m_{21}, \\ \beta_1 s_{21} + \dots + \beta_p s_{2p} = m_{12} - m_{22}, \\ \dots, \\ \beta_1 s_{p1} + \dots + \beta_p s_{pp} = m_{1p} - m_{2p} \end{cases}$$

Далі знаходимо середнє значення розподілу величини Y :

$$\bar{Y}_i = \frac{1}{n_i} \sum_{t=1}^{n_i} Y_{it},$$

та її дисперсію:

$$S_Y^2 = \sum_{j=1}^p \sum_{h=1}^p \beta_j s_{jh} \beta_h$$

Таким чином, узагальнена байєсівська процедура оцінювання полягає у віднесенні $x = (x_1, \dots, x_p)$ до W_1 , якщо

$$Y = \sum_{i=1}^p \beta_i x_i \geq \frac{\bar{Y}_1 + \bar{Y}_2}{2} + \ln \frac{q_2 C(1|2)}{q_1 C(2|1)}$$

та до W_2 в протилежному випадку, де $q_i = \frac{n_i}{n}$, C – ймовірність помилкової класифікації. В даному випадку $C(2|1) = C(1|2) = 1$, оскільки немає об'єктивних причин для надання їм інших значень.

Висновки

Розроблений алгоритм перевірено на реальних зображеннях. Алгоритм продемонстрував стійкість у роботі та досягненні заданої точності.

Вибірка досліджуваних об'єктів складалася з 300 зображень. Всі зображення були отримані авторами самостійно за допомогою мікроскопа та

цифрової фотокамери, призначеної для зйомок зображень з мікроскопа. Отримані зображення мали збільшення в 1000 разів. Мазки крові та їхній опис, розподіл зображень по класам, був отриманий на кафедрі педіатрії Дніпропетровської медичної академії.

Подальші дослідження полягають у виділенні візуальних параметрів еозинофіла, а саме:

1. Цитоплазматичні гранули.
2. Інтенсивність окраски.
3. Характер зернистості.
4. Розташування гранул

Список літератури

1. Про розвиток та удосконалення лікувально-профілактичної допомоги дітям з алергічними захворюваннями: наказ від 20.02.1995 № 33. – Режим доступу: www.moz.gov.ua/ua/main/?docID=9606
2. Спосіб прогнозування імунного статусу новонароджених: пат. на винахід № 62672 Україна / Мокія С.О., Шелевицький І.В., Василенко Н.В. – Опубл. 15.12.2005, Бюл. №12.
3. Спосіб прогнозування алергічних захворювань у дітей: пат. на винахід № 51144 Україна / Мокія С.О., Литвинова Т.В., Василенко Н.В., Сітало С.Г., Пономарьова Л.І. – Опубл. 15.11.2002, Бюл. №11.
4. Коков А.А. Автоматизированная подсистема распознавания и оконтуривания клеток / А.А. Коков. – Режим доступу: <http://www.masters.donntu.edu.ua/2003/kita/kokov/library/pub1.htm>
5. Biomedical Imaging Group: Gradient (index) – Режим доступу: <http://bigwww.epfl.ch/algorithms.html>
6. Медведєв Д.Г. Алгоритм оконтурювання еозинофілів за допомогою сплайнів / Д.Г. Медведєв // Комп'ютерні науки та інженерія: матеріали IV Міжнародної конференції молодих вчених CSE-2010. – Львів: Вид-во Львівської політехніки, 2010. – С.274-275
7. Сплайни в цифровій обробці даних і сигналів / Шелевицький І.В., Шутко М.О., Шутко В.М., Калганова О.О. – Кривий Ріг: Видавничий дім, 2008. – 232с., іл.
8. Афифи А. Статистический анализ: подход с использованием ЭВМ / А. Афифи, С. Эйзен ; пер. с англ. – М.: Мир, 1982. – С. 322-324.
9. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч.У. Мьюллер, У.Р. Клекка и др.; под ред. И.С. Енюкова ; пер. с англ. – М.: Финансы и статистика, 1989. – С. 90-93.

Надійшла до редколегії 10.06.2011

Е. В. БУРЛАК, В. А. ГОЛОВКО, Д.Г. BURLAK O.V., GOLOVKO V.O., MEDVEDEV D.G.

Национальный авиационный университет,
Криворожский государственный педагогический университет

National aviation university,
Kriviy Rig state pedagogical university

Дискриминантная функция параметризованных эозинофилов

Discriminant function for parameterized eosinophils

В работе представлены специализированные способы и алгоритмы параметризации цифровых изображений эозинофилов для задач медицинской диагностики. Разработан алгоритм получения контура эозинофила как сплайн-функции и определения его геометрических параметров. Получена дискриминантная функция для параметризованных эозинофилов

The specialized methods and algorithms of parameterization of digital representations of eosinophils for the tasks of medical diagnostics are presented. The algorithm of obtaining the contour of eosinophils as a spline function and determination of its geometrical parameters is developed. Discriminant function for parameterized eosinophils is derived.

Распознавание образов, эозинофил, сплайн

Recognition of patterns, eosinofil, spline