

М.І. РИЩЕНКО, аспірант, НТУ «ХП»

МЕТОДИКА РОСПІЗНАВАННЯ ПЕТ-ЗНІМКІВ З МЕТОЮ ІДЕНТИФІКАЦІЇ РАКОВИХ ПУХЛИН

Предложена методика распознавания ПЕТ- снимков, позволяющая идентифицировать и наблюдать за развитием раковых опухолей с помощью комплекса алгоритмов кластеризации изображений.

Ключевые слова: алгоритм кластеризации k-средних, МРТ снимки, ПЕТ снимки, компьютерный анализ изображения.

Постановка завдання та аналіз літератури. У наш час активно використовуються томографічні методи діагностики захворювань, такі як позитронно-емісійна томографія (ПЕТ) або магнітно-резонансна томографія (МРТ). Ці методи дозволяють діагностувати поведінку організму хворого. Більш коректна зміна в поведінці організму може бути виконана за допомогою сегментації зображень. Вчені розробили методи, які дозволяють проаналізувати медичні знімки. Однак, багато з них не містять у собі повний обсяг даних, що надається знімками, тим самим піддають діагностику нечітким результатом. На практиці ефективним вирішенням цієї проблеми являється комплексне використання алгоритму кластеризації k-середніх (k-means).

Найбільш поширений серед неієрархічних методів алгоритм k-середніх [1], також званий швидким кластерним аналізом. Широко використовується в медицині, біології, інформатики. Алгоритм k-середніх будує k кластерів, розташованих на можливо великих відстанях один від одного. Основний тип задач, які вирішує алгоритм k-середніх, - наявність припущень (гіпотез) щодо числа кластерів, що задовольняє поставленим перед нами завданням.

Дія алгоритму така, що вона прагне мінімізувати середньоквадратичне відхилення на точках кожного кластера. Основна ідея полягає в тому, що на кожній ітерації переобчислювати центр мас для кожного кластера, отриманого на попередньому кроці, потім вектори розбиваються на кластери знову у відповідності з тим, який з нових центрів виявився ближчим за обраною метриці. Алгоритм завершується, коли на якийсь ітерації не відбувається зміни кластерів.

Основна частина. Знімки для аналізу були отримані в ході проекту між університетом «Paris-Est Creteil» та лікарнею «Paris-Est Creteil». Метою проведеного дослідження є кластеризація медичних знімків для виявлення злоякісних новоутворень. Медичні знімки були розділені на дві групи: знімки фантоми і знімки хворих. Перша група була створена штучно фахівцями лі-

карні і призначена для проміжного тестування обраної методики. Друга група представляла з себе томографічні знімки з різними стадіями захворювань пацієнтів, як це показано на рис. 1.

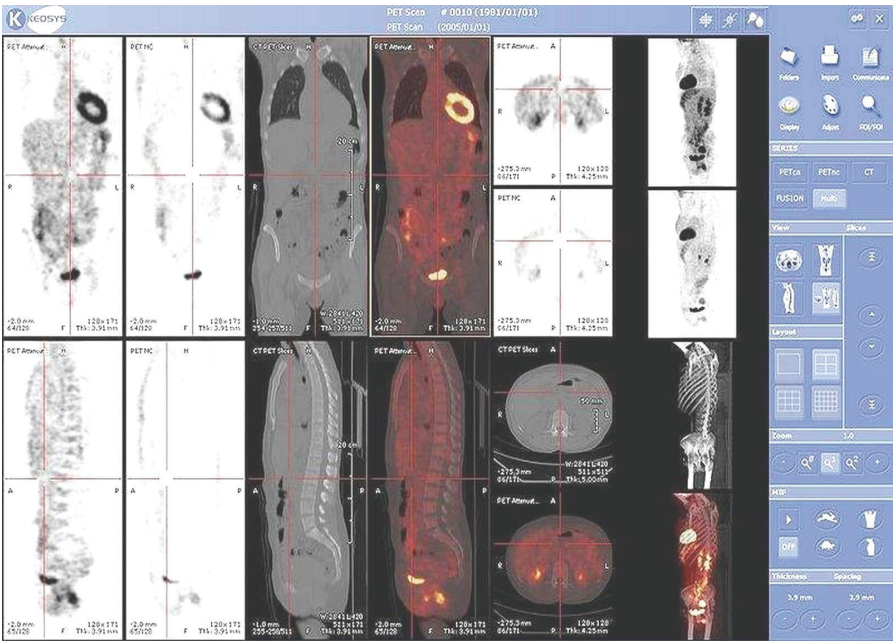


Рисунок 1 – Томографічний знімок пацієнта

Практична частина виконувалась за допомогою інструментарію ІТК[2]. Це сучасний крос-платформенний каркас, з відкритим вихідним кодом для розробки додатків, який широко використовуються при сегментації зображень та програми реєстрації зображення. Цей продукт був обраний в якості середовища для аналізу знімків, так як він реалізований в C++, є крос-платформенним [3]. Крім того, автоматизований процес упаковки створює інтерфейс між C++ та іншими мовами програмування, що дозволяє розробникам створювати програми з використанням різних мов програмування.

Було виділено два основних етапи за якими проводилась практична частина. Це кластеризація на основі вибраних розмірів пухлини і виділення кольорів, які характеризують злоякісні новоутворення. Обидва етапи можна реалізувати за допомогою алгоритму k-means. Різницею між двома підходами буде кількість кластерів та їх значення.

Написання програми вироблялося на основі алгоритму[4], завданням якого є мінімізувати відстаней між об'єктами в кластерах. Останов відбувається, коли мінімізувати відстані більше вже неможливо. Мінімізуема функція у випадку k-means є така (1):

$$J = \sum_{k=1}^M \sum_{i=1}^N d^2(x_i, c_k), \quad (1)$$

де $x_i \in X$ – об'єкт кластеризації (точка), $c_j \in C$ – центр кластера (центроїд).

На момент старту алгоритму має бути відомо число C (кількість кластерів). Вибір числа C може базуватися на результатах попередніх досліджень, теоретичних міркуваннях або інтуїції.

Опис алгоритму. Початковий розподіл об'єктів по кластерах. Вибираються C точок. На першому кроці ці точки вважаються центрами кластерів. Вибір початкових центроїдів може здійснюватися шляхом підбору спостережень для максимізації початкової відстані, випадковим вибором спостережень або вибором перших спостережень.

1. Ітеративний перерозподіл об'єктів по кластерах. Об'єкти розподіляються по кластерам шляхом підрахунку відстані від об'єкта до центрів кластерів і вибору найменшого.

2. Коли всі об'єкти розподілені по кластерам, заново вважаються їх центри (2).

$$c_j = \frac{\sum_{i=1}^L x_i}{L}, \quad (2)$$

де $x_i \in C_j$, $|C_j| = L$ (можна вважати по кожній координаті окремо).

3. Якщо $c_j = c_{j-1}$, то це означає, що кластерні центри стабілізувалися і відповідно розподіл закінчено. Інакше переходимо до кроку 1.

Перевірка якості кластеризації [5]. Після отримань результатів кластерного аналізу методом k -середніх слід перевірити правильність кластеризації, тобто оцінити, наскільки кластери відрізняються один від одного. Для цього розраховуються середні значення для кожного кластера. Ознакою правильної кластеризації є середні для всіх вимірювань, які сильно відрізняються.

Після виконання практичних робіт слід виявити переваги алгоритму k -середніх – простота використання, швидкість використання, зрозумілість і прозорість алгоритму, може бути легко модифікований для вирішення різних завдань, таких як часткове навчання з вчителем або потокових даних. Результат роботи даного методу кластеризації ми можемо побачити на рис. 2.

Висновки. Запропонована методика розпізнавання ПЕТ-знімків дозволяє врахувати різні параметри, які впливають на якість аналізу, до них відносяться: розмір виявленої пухлини, колір який характеризує злоякісні новоутворення, правильну кількість кластерів в кожному з підходів розпізнавання.

Використання запропонованої методики розпізнавання медичних знімків дозволить значно покращити принципи аналізу та діагностування захворювань. Швидкість розбиття на кластери і масштабованість алгоритму роб-

лять дане дослідження актуальним на сьогоднішній момент.

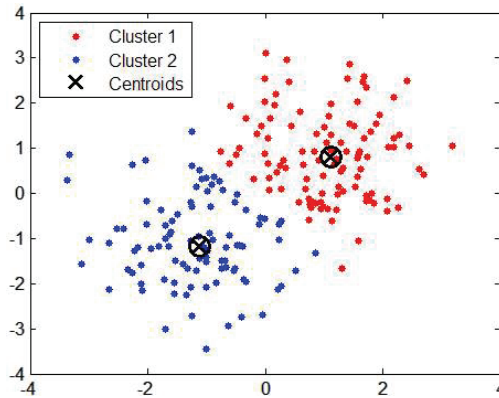


Рисунок 2 – Результат кластерного аналізу

Список літератури: 1. Pham DL: Spatial Models for FuzzyClustering // Computer Vision and Image Understanding. – 2001. – 84(2). – P. 285-297. 2. Luis Ibanez, Will Schroeder The ITK Software Guide/ Second Edition. Updated for ITK version 2.4 // November 21, 2005.3. <http://www.itk.org/ITK/project/about.html> – Insight Segmentation and Registration Toolkit (ITK). 4. http://zakup.vl.ru/132-metodi_klastern.html – методы кластерного анализа. 5. Sonka M, Hlavac V, Boyle R. Image Processing // Analysis and Machine Vision. – 2nd edn. – PWS Publishing, 1999.

Надійшла до редколегії 18.04.2013

УДК 621.391

Методика розпізнавання ПЕТ-знімків з метою ідентифікації ракових пухлин / М.І.Рищенко // Вісник НТУ «ХП». Серія: Техніка та електрофізика високих напруг. – Х.: НТУ «ХП», 2013. – № 27 (1000). – С. 145-148. – Бібліогр.: 5 назв.

Запропоновано методику розпізнавання ПЕТ-знімків, яка дозволяє ідентифікувати і спостерігати за розвитком ракових пухлин за допомогою використання комплексу алгоритмів кластеризації зображень.

Ключові слова: алгоритм кластеризації k-середніх, МРТ знімки, ПЕТ знімки, комп'ютерний аналіз зображення.

The method of recognizing the PET images, which allows us to identify and observe the development of cancerous tumors by using complex algorithms for clustering images.

Keywords: k-means Clustering Algorithm, MRI images, PET images, computer image analysis.