
СИСТЕМИ ТЕХНІЧНОГО ЗОРУ І ШТУЧНОГО ІНТЕЛЕКТУ З ОБРОБКОЮ ТА РОЗПІЗНАВАННЯМ ЗОБРАЖЕНЬ

УДК 004.032.26

А.В. КОЖЕМ'ЯКО, О.С. ВАСИЛЬКІВА, О.С. БЕЗКРЕВНИЙ, О.В. ПРОКОФ'ЄВ

ГЕОІНФОРМАЦІЙНА СИСТЕМА ДЛЯ КЛАСИФІКАЦІЇ ОБ'ЄКТІВ НА ГЕОГРАФІЧНИХ КАРТАХ

*Вінницький національний технічний університет
21021, Хмельницьке шосе, 95, м. Вінниця, Україна
E-mail: kvantron@hotmail.com*

Анотація: У статті розглянуті методи класифікації геометричних об'єктів на географічних картах для задач обробки, аналізу та розпізнавання геоінформаційних зображень. Застосовано ієрархічний підхід в реалізації алгоритму, що дозволяє здійснювати розпізнавання за короткий час, не вимагаючи значних обчислювальних ресурсів, шляхом активації обмеженого числа локальних ділянок зображення, наприклад географічних карт.

Ключові слова: геоінформаційні системи, комп'ютерний зір, ПЛІС, паралелізм, штучна нейронна мережа

Аннотация: В статье рассмотрены методы классификации геометрических объектов на географических картах для задач обработки, анализа и распознавания геоинформационных изображений. Применен иерархический подход в реализации алгоритма, позволяющего осуществлять распознавание за короткое время, не требуя значительных вычислительных ресурсов, путем активации ограниченного числа локальных участков изображения, например географических карт.

Ключевые слова: геоинформационные системы, компьютерное зрение, ПЛИС, параллелизм, искусственная нейронная сеть

Abstract: The article deals with the methods of classification of geometric objects on geographic maps for the tasks of processing, analysis and recognition of geographic information images. A hierarchical approach is implemented in implementing an algorithm that allows recognition in a short time, without requiring significant computational resources, by activating a limited number of local image areas, such as geographic maps.

Keywords: geoinformation systems, computer vision, FPGA, parallelism, artificial neural network

ВСТУП

Розпізнавання візуальних образів є одним із найважливіших компонентів систем управління та обробки інформації, автоматизованих систем і систем прийняття рішень. Завдання, пов'язані з класифікацією і ідентифікацією предметів, що характеризуються кінцевим набором деяких властивостей і ознак, виникають в таких галузях як: геосистеми, робототехніка, інформаційний пошук, моніторинг та аналіз візуальних даних, дослідження штучного інтелекту. Алгоритмічна обробка і класифікація зображень застосовуються в системах відеоспостереження, розпізнавання, системах віртуальної реальності та інформаційних пошукових системах. На даний момент в виробництві широко використовуються системи розпізнавання рукописного тексту, автомобільних номерів, географічних карт, що знаходять застосування в інтерфейсах програмних продуктів, ідентифікації географічних об'єктів, а також в інших прикладних цілях[1-3].

ОБҐРУНТУВАННЯ ВИБОРУ МЕТОДУ

Для розпізнавання та обробки зображень з'являється все більше методів і технологій для пошуку меж об'єкта і відділення його від основного фону. Але найвідомішими і часто використовуються донині є методи Робертса, Прюїтта і Собеля[4,5]. Кожен з методів базується на процесі просторової фільтрації. Просторова фільтрація заснована на пересуванні маски фільтра від першого пікселя до останнього пікселя зображення по порядку. Кожен піксель з координатами (x, y) розраховується як сума добутків

значення пікселів під маскою фільтра на відповідні їм коефіцієнти фільтра. Якщо на зображенні спостерігаються різкі перепади яскравості, то вдаються до похідних першого і другого порядку. Для того, щоб визначити першу похідну одновимірної функції $f(x)$ необхідно розрахувати різницю значень сусідніх елементів зображення:

$$\frac{df}{dx} = f(x+1) - f(x)$$

Приватна похідна застосовується для того, щоб зберегти ті ж асигнації для змінних $f(x, y)$, якщо необхідно працювати по двох осях координат з приватними похідними.

Друга ж похідна обчислюється тут як різниця між сусідніми значеннями першої похідної:

$$\frac{d^2 f}{dx^2} = f(x+1) + f(x-1) - 2f(x)$$

Обчислення першої похідної зображення засноване на дискретних наближеннях двовимірного градієнта. За визначенням, градієнт зображення $f(x, y)$ в пікселі (x, y) - це вектор:

$$\Delta f = \begin{pmatrix} G_x \\ G_y \end{pmatrix} = \begin{pmatrix} df/dx \\ df/dy \end{pmatrix}$$

загальновідомо, що траєкторія вектора градієнта збігається з траєкторією максимальної швидкості зміни функції f в пікселі (x, y) . Модуль цього вектора має головну роль при визначенні контурів на зображенні. Він позначається символом ∇f і прирівнюється:

$$\Delta f = |\Delta f| = \sqrt{G_x^2 + G_y^2}$$

У всіх методах визначення меж об'єкта використовується матриця пікселів розміром 3×3 :

$$\begin{pmatrix} z_1 & z_2 & z_3 \\ z_4 & z_5 & z_6 \\ z_7 & z_8 & z_9 \end{pmatrix}$$

Найлегший спосіб визначення перших приватних похідних в точці Z_5 полягає в використанні перехресного градієнтного оператора Робертса:

$$G_x = (z_9 - z_5); G_y = (z_8 - z_6);$$

Значення G_x і G_y можуть бути отримані за допомогою перетворення цілого зображення за допомогою однієї з ковазюють масок оператора Робертса:

$$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

можна помітити, що оператор Прюїтта, так само як і оператор Робертса, працює з плаваючою маскою розмірністю 3×3 пікселя. На відміну від оператора Робертса в операторі Прюїтта матриці задаються іншими формулами:

$$G_x = (z_7 + z_8 + z_9) - (z_1 + z_2 + z_3); G_y = (z_3 + z_6 + z_9) - (z_7 + z_4 + z_1);$$

Щоб застосувати ці формули для виділення країв на зображенні, використовується оператор Прюїтта[4,5]. Даний оператор задається наступними масками розмірністю 3×3 :

$$\begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}$$

Так само використовує матрицю розмірністю 3×3 і оператор Собеля. Головна його відмінність від оператора Прюїтта полягає в застосуванні вагового коефіцієнта 2 для центрових пікселів:

$$G_x = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3); G_y = (z_3 + 2z_6 + z_9) - (z_7 + 2z_4 + z_1);$$

Це штучне збільшення використовується для зменшення ефекту пом'якшення на результуючому зображенні. Маски, що застосовуються в методі Собеля розмірністю 3×3 :

$$\begin{pmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}$$

Як було описано вище, за допомогою матриць можна розрахувати складові величини градієнта G_x і G_y . Щоб визначити саме значення градієнта цими складовими потрібно оперувати одночасно:

$$\Delta f \approx |G_x| + |G_y|$$

Переглянувши результати роботи трьох алгоритмів на практиці, можна зробити висновок, що при використанні географічних карт місцевості найякіснішим методом виявився метод пошуку границь Собела. Оскільки саме цей метод дозволив яскраво виділити границю, знизити інтенсивність фону зображення і якісно позначити всі елементи зображення. Метод Робертса теж показав непогані результати, проте при його використанні можна помітити безповоротну втрату даних з вихідного зображення. Метод Прюїтта в цьому випадку виявився слабким і не якісним.

МЕТОДИ РОЗПІЗНАВАННЯ ЗОБРАЖЕНЬ

Розроблений метод навчання системи розпізнавання являє собою модифікований варіант методу зворотного поширення помилки, що широко використовується. В силу специфічних характеристик системи, автором запропоновані додаткові кроки в розробці методу, що включають в себе використання критерію розрідженості, використання елементів вибірки для збільшення різноманітності навчального матеріалу і формування більш стійкої функції, що розпізнається системою для досягнення достатніх результатів реконструкції і розпізнавання.

Дана вибірка складається з локальних фрагментів зображень (x, y) і значень трансформацій s , які в загальному випадку можуть бути параметризовані довільним чином. При навчанні ставиться наступна ціль: при отриманні на вході фрагмента зображення x_i , і трансформації s_i представити на виході трансформований фрагмент зображення h_i , такий що $h_i = t_i$.

Метод навчання складається:

1. Для вихідного шару системи визначимо помилку мережі як середньоквадратичне відхилення від цільового результату $E = \frac{1}{2}(h - t)^2$, де h і t являють собою значення реконструкції мережі і цільове значення в векторній формі $h = (h_1, h_2, \dots, h_m)$, $t = (t_1, t_2, \dots, t_m)$, де m – число результатів вибірки.

2. Для кожного нейрона мережі, вихідне значення j -го нейрона або значення активації визначається як $a_j = \sigma(\sum_{i=1}^n \omega_{i,j} x_i)$, де σ - функція сигмоїди, $\omega_{i,j}$ - вага j -го нейрона, x_i - вхідні дані нейрона.

Нехай z_j - аргумент функції активації, тобто $z_j = \sum_{i=1}^n \omega_{i,j} x_i$, $a_j = \sigma(z_j)$. Визначимо приватну похідну для j -го нейрона по перемінним $\omega_{i,j}$:

$$\frac{\partial E}{\partial \omega_{i,j}} = \frac{\partial E}{\partial a_j} \frac{\partial a_j}{\partial z_j} \frac{\partial z_j}{\partial \omega_{i,j}}$$

Спростивши отриманий вираз, отримаємо значення помилки j -го нейрона δ_j шару l :

$$\delta_j = \frac{\partial E}{\partial a_j} \frac{\partial a_j}{\partial z_j} f(x) = \begin{cases} (a_j - t_j) \sigma(z_j) (1 - \sigma(z_j)), & l = 4 \\ \sum_{l \in L} \delta_j \omega_{jl}, & l = 3 \\ \sum_{l \in L} \delta_j \omega_{jl} \sigma(z_j) (1 - \sigma(z_j)), & l \in (0, 1) \end{cases}$$

відповідно до методу зворотного поширення помилки.

3. До кожного значення помилки нейрона δ_j додамо похідну від критерію розрідженості,

$$\delta_j = \delta_j + \beta \left(-\frac{s}{s_j} + \frac{1-s}{1-s_j} \right) \sigma(z_j) (1 - \sigma(z_j)),$$

де β і s є метапараметри розрідженості.

4. Для кожного j -го нейрона мережі (крім нейронів 2 шару) і для кожного ваги $\omega_{i,l}$: вводимо значення ваги на наступну величину:

$$\Delta \omega_{i,l} = -\alpha \delta_j,$$

де α константа швидкості навчання.

5. Після того, як ваги всіх нейронів оновлені, проводиться оцінка точності реконструкції системи. Мережа здійснює обчислення вихідного значення і новий розрахунок загальної помилки мережі \hat{E} . Метод може бути зупинений при дотриманні трьох критеріїв.

$$\Delta E = E - \hat{E} < \varepsilon_l,$$

тобто досягнута збіжність при оптимізації параметрів функції мережі.

$h - t < \varepsilon_r$ досягнута умова необхідної точності реконструкції.

Досягнуто максимальної кількості ітерацій алгоритму (як правило, ε ознакою неоптимізованої проблеми і неможливості досягнення збіжності).

6. Якщо метод не зупинено на попередньому кроці, наступна ітерація починається з кроку 1.

Метод навчання передбачає роботу в онлайн-режимі, поповнюючи поступово зростаючу вибірку новими екземплярами локальних фрагментів зображення. Він також включає в себе стадію розширення вибірки шляхом проведення випадкових мутацій - так, маючи в наявності дві пари трансформованих фрагментів зображення (x_1, t_1) і (x_2, t_2) , характеризуються трансформаціями s_1 і s_2 , алгоритм комбінує з їх допомогою пари (x_1, t_1) і (x_2, t_2) , яка характеризується кумулятивною трансформацією $S = S_1 + S_2$.

Експериментальні результати показують, що розширення вибірки за допомогою таких комбінацій веде до збільшення різноманітності навчальних даних і сприяє формуванню стійкої функції розпізнавання системи, більш стійкої до перенавчання.

Для оцінки значення метапараметрів алгоритму β, s, α використовується методика певної валідації, що складається з наступного: накопичена в процесі онлайн-навчання досить масивна ($m > 100$ примірників) вибірка ділиться на три нерівні частини в соотношенні 7:2:1. Задаються допустимі інтервали значень метапараметрів, після чого проводиться серія повноцінних циклів навчання системи для кожного значення β, s, α .

Результатом навчання для кожного випадку є загальне середньоквадратичне відхилення мережі $E_{\beta, s, \alpha}$, оцінка якого проводиться на другій частині вибірки. За результатами певної валідації вибираються значення метапараметрів, які забезпечують мінімізацію помилки мережі, після чого продуктивність системи може бути протестована на третьому сегменті вибірки.

Використання трьох незалежних сегментів даних для навчання і тестування різних параметрів мережі являє собою основу методики певної валідації, що забезпечує незалежність вкладу метапараметрів в точність реконструкції системи.

Запропонований метод навчання задовольняє вимогам, пред'являються до функціонування системи. Її перевагою перед класичним методом зворотного поширення помилки є використання критерію розрідженості, що дозволяє комбінувати в одній вибірці зображення, візуально несхожі між собою і негативно впливають на роботу звичайних нейронних мереж, знижуючи чутливість до перенавчання. Використовуючи критерій розрідженості, метод дозволяє системі тимчасово пригнічувати активність більшої частини нейронів, використовуючи для репрезентації локальних ознак тільки невелику підмножину нейронів в кожному шарі, забезпечуючи тим самим незалежність нейронів мережі і здатність формувати більш складні уявлення. Використання методу мутації дозволяє збільшити вибірку, розширюючи обсяг даних для навчання. Нарешті, використання методики певної валідації дозволяє об'єктивно оцінювати використовувані мета параметри критерію розрідженості, визначаючи мінімізують загальну помилку значення мережі.

Розроблений метод розпізнавання зображень є ключовим елементом методу, представленого в даному дисертаційному дослідженні, здійснює безпосередню процедуру класифікації зображень по категоріям з використанням попередньо навченої моделі. Метод розпізнавання складається з послідовної активації детекторів моделі, починаючи з першого рівня, на заданому зображенні.

Нехай дана навчена модель M_i для деякого класу зображень c_i (наприклад, таким класом можуть виступати географічні карти), або кілька моделей для завдання мультикласового розпізнавання, і зображення I , яке необхідно розпізнати. Метод розпізнавання складається з наступних кроків:

1. Задамо загальну ієрархічну пам'ять НМ, представлену великою кількістю рівнів, розміром рівному максимальному числу рівнів навченої моделі.

2. Для L -того рівня моделі L_i , починаючи з першого, і для кожного детектора відповідного рівня

$$b_j^{(l)}(x, y, h, w)$$

застосуємо ідентифікуючу функцію детектора до всіх локальних ділянок зображення

$$I(x \dots x + w, y \dots y + h).$$

3. Якщо для деякого рівня L_i активації всіх детекторів цього рівня негативні (дорівнюють нулю), то зображення не належить до класу c_i .

4. Якщо завдання розпізнавання є мультикласовим, то повторюємо алгоритм, починаючи з кроку 1, для моделі наступної категорії M_{i+1} . В протилежному випадку метод вважається зупиненим.

5. Якщо детектори шару L_i активовані, то продовжимо послідовну активацію детекторів наступних шарів, повторюючи алгоритм з кроку 2.

6. Якщо активація останнього рівня позитивна, то зображення належить до класу c_i .

7. Для мультикласового розпізнавання перераховані вище кроки повторюються для безлічі навчених екземплярів моделі M_0, M_1, \dots, M_n , при цьому належність зображення до класу визначається по максимальному активному рівню.

Метод розпізнавання з використанням дворівневої моделі може ефективно застосовуватися для вирішення проблем розпізнавання таких об'єктів як географічні карти, в деяких випадках (певний масштаб) - силуети, автомобілі. У загальному випадку правило вибору кількості рівнів для моделі залежить від кількості ступенів свободи шуканого об'єкта, і не може бути однозначно встановлено по зображенню, бо залежить від впливу шуму і обмеженого простору спостереження.

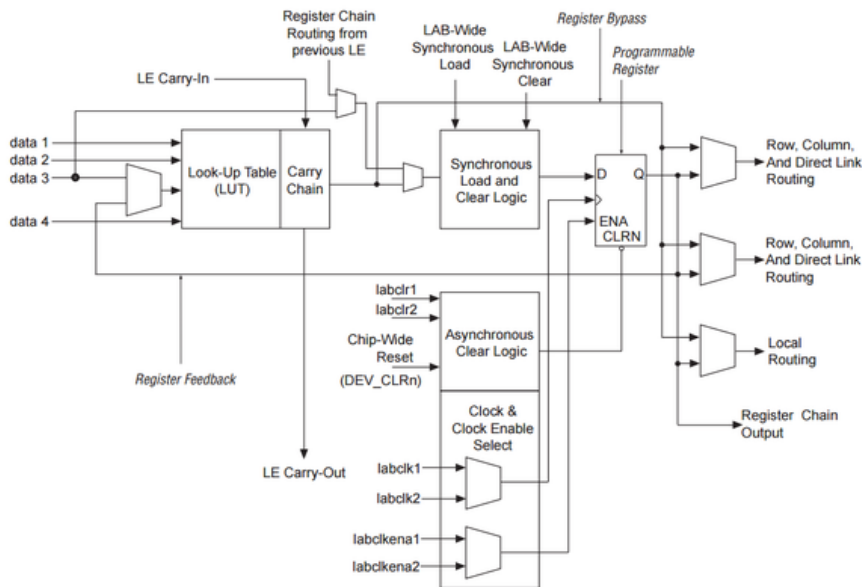


Рис. 1. Схема для обробки зображення

РЕЗУЛЬТАТИ ДОСЛІДЖЕННЯ

В ході моделювання було отримано результати розпізнавання знімків місцевості. Завдяки застосуванню оператора Собеля, що застосовується для виділення границь було отримано розпізнані карти місцевості. Метою було отримання карти, на якій будуть відмічені будівлі або інші об'єкти, які мережа навчена розпізнавати.



Рис. 2. Приклад розпізнаної карти



Рис.3. Приклад розпізнаної карти



Рис.4. Приклад розпізнаної карти

Як видно з прикладу було розпізнано більшість об'єктів на карті. Також помітно похибки розпізнавання – тіні від будівель, дерева, що тісно прилягають до дахів.

ВИСНОВКИ

Удосконалено метод навчання моделі за допомогою потоку даних, який дозволяє добувати і обробляти ключову для розпізнавання інформацію, локалізувати об'єкт на зображенні. Використовуючи запропонований метод, модель здатна навчатися на візуальному досвіді, оснащеного сенсором (відеокамерою), що робить її придатною для обробки інформації в різних додатках штучних зорових систем.

Вдосконалено алгоритм розпізнавання, що дозволяє як класифікувати зображення з точки зору його належності до категорії розпізнавання, так і визначити розташування зображеного об'єкта.

Розроблений метод використовує локально-ієрархічний підхід, що дозволяє підвищити як стійкість, так і швидкість роботи. Запропоновані алгоритми дозволяють вирішити поставлену задачу розпізнавання зображень з проведенням локалізації об'єкта.

Розроблено та реалізовано математичну модель розпізнавання геометричних образів на ПЛІС в середовищі розробки Quartus II.

СПИСОК ЛІТЕРАТУРИ

1. Старовойтов В. В. Локальные геометрические методы цифровой обработки и анализа изображений / В. В. Старовойтов. – Мн.: Ин-т техн. Кибернетики НАН Беларуси. - 1997. – 282 с.
2. Шлихт Г. Цифровая обработка цветных изображений / Г. Шлихт. – М: ЭКОМ. - 1997. – 336 с.
3. Гонсалес Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. - М: Техносфера. – 2004. — 1072 с.
4. Lee C. C. Fuzzy logic in control systems: fuzzy logic controller - part I, II // IEEE Trans, on Systems, Man and Cybernetics, Vol. 20, 1990. - №2 - P. 404-435.
5. Абламейко С. В. Обработка изображений: технология, методы, применение / С. В. Абламейко, Д. М. Лагуновский. – Мн.: Ин-т техн. кибернетики НАН Беларуси. - 1999. – 300 с.
6. Color in Electronic Displays / H.L. Widdel, D.L. Post, J.D. Grossman et al. – New York, London: Plenum press. – 1991. – P. 39-91. А.И. Власов

REFERENCES

1. Starovoytov V. V. Lokal'nye heometrycheskye metody tsyfrovoy obrabotky y analiza yzobrazhenyy / V. V. Starovoytov. – Мн.: Yn-t tekhn. Kybernetyky NAN Belarusy. - 1997. – 282 s.
2. Shlykht H. Tsyfrovaya obrabotka tsvetnykh yzobrazhenyy / H. Shlykht. – М: ЭКОМ. - 1997. – 336 s.
3. Honsales R. Tsyfrovaya obrabotka yzobrazhenyy / R. Honsales, R. Vuds. - М: Tekhnosfera. – 2004. — 1072 s.
4. Lee C. C. Fuzzy logic in control systems: fuzzy logic controller - part I, II // IEEE Trans, on Systems, Man and Cybernetics, Vol. 20, 1990. - №2 - P. 404-435.
5. Ablameyko S. V. Obrabotka yzobrazhenyy: tekhnolohyya, metody, prymenenye / S. V. Ablameyko, D. M. Lahunovskyu. – Мн.: Yn-t tekhn. kybernetyky NAN Belarusy. - 1999. – 300 s.
6. Color in Electronic Displays / H.L. Widdel, D.L. Post, J.D. Grossman et al. – New York, London: Plenum press. – 1991. – P. 39-91. А.У. Vlasov

Надійшла до редакції 8.06.2017 р.

КОЖЕМ'ЯКО АНДРІЙ ВІКТОРОВИЧ – к.т.н., доцент кафедри лазерної та оптико-електронної техніки, Вінницький національний технічний університет, Вінниця, Україна

ВАСИЛЬКІВА ОЛЕНА СЕРГІЇВНА – аспірант кафедри лазерної та оптико-електронної техніки, Вінницький національний технічний університет, Вінниця, Україна

БЕЗКРЕВНИЙ ОЛЕКСАНДР СЕРГІЙОВИЧ – аспірант кафедри лазерної та оптико-електронної техніки, Вінницький національний технічний університет, Вінниця, Україна

ПРОКОФ'ЄВ ОЛЕКСАНД ВОЛОДИМИРОВИЧ – магістрант кафедри лазерної та оптико-електронної техніки, Вінницький національний технічний університет, Вінниця, Україна