

МЕТОД РОЗПІЗНАВАННЯ ДОРОЖНІХ ЗНАКІВ З ВИКОРИСТАННЯМ ЗГОРТАЛЬНОЇ НЕЙРОННОЇ МЕРЕЖІ

У статті розглянутий підхід до реалізації системи детектування і розпізнавання дорожніх знаків, що базується на використанні згортальної штучної нейронної мережі. Використання такої архітектури штучних нейронних мереж дозволяє уникнути багатьох проблем з якими стикаються при розробці систем технічного зору. Згортальні нейронні мережі найкраще пристосовані для розпізнавання об'єктів на зображенні, тому вони цілком підходять для застосування в системах розпізнавання дорожніх знаків в режимі реального часу. В порівнянні з класичними, згортальні нейронні мережі менш чутливі до різного роду спотворення зображення, а саме поворотів, зсувів, масштабування та інших. Для навчання використовується класичний метод зворотного поширення помилки. Проте для ефективної роботи такої нейронної мережі необхідна велика вибірка зразків для навчання.

Ключові слова: розпізнавання дорожніх знаків, інтелектуальний автомобіль, згортальні нейронні мережі.

Вступ. На протязі останніх десятиліть в автомобільній промисловості було зроблено крок у бік інтелектуальних транспортних засобів, оснащених системами допомоги водієві. Прикладом комерційно успішної такої системи є GPS-навігація. Сучасні навігаційні пристрої допомагають не лише у виборі правильного маршруту, а також попереджають водія про

встановлені обмеження швидкісного режиму. Більшість навігаційних систем використовують статичну карту з інформацією про дорогу і встановлені на ній обмеження швидкості руху. У деяких ситуаціях, даних зі статичної карти не достатньо, наприклад, під час дорожніх робіт, коли тимчасово змінюють максимально допустиму швидкість руху. Рішення для таких ситуацій є використання системи технічного зору.

Постановка проблеми. Останнім часом в автомобільній промисловості впроваджують системи технічного зору у автомобілях топ-класу. Прикладами є BMW 7ER, Mercedes S-Class, Audi A8, Opel Insignia і VW Phaeton. Всі ці системи призначені для їх конкретного застосування в якості конвеєра ретельно налаштованих алгоритмів [1]. На першому етапі вхідні дані підлягають обробці з використання фіксованого набору алгоритмів, таких як корекція освітлення, гістограма вирівнювання, сегментація кольору, виділення країв, перетворення Хафа тощо. Результат цих кроків використовується для класифікації відповідним алгоритмом або методами машинного навчання, таким як штучні нейронні мережі (ШНМ) або метод опорних векторів. Проектування такого конвеєра алгоритмів займає досить багато часу і потребує перетворення для підтримки нових знаків.

Використання штучних нейронних мереж для детектування і розпізнавання дорожніх знаків є одним з досить ефективних методів. Проте використання класичних нейронних мереж для розпізнавання зображень ускладнюється, як правило, великим розміром вектора вхідних значень нейронної мережі, кількістю нейронів в проміжних шарах і, як наслідок, великими обчислювальними витратами на навчання і обчислення мережі [2,3]. Ще одним недоліком класичних нейронних мереж є відсутність можливості враховувати топологію вхідного зображення, тобто інформацію про взаєморозміщення пікселів зображення.

Виклад основного матеріалу дослідження. Доцільним є використання згортальних нейронних мереж. Згортальним нейронним мережам в меншій мірі притаманні описані вище недоліки. Ключовим моментом в розумінні згортальних нейронних мереж є поняття так званих “поділюваних” ваг, тобто частина нейронів деякого шару нейронної мережі може використовувати одні й ті ж вагові коефіцієнти. Нейрони, що використовують однакові ваги, об’єднуються в карти ознак (feature maps), а кожний нейрон карти ознак пов’язаний з частиною нейронів попереднього шару. При обчисленні мережі виходить, що кожний нейрон виконує згортку деякої області попереднього шару (що визначається множиною нейронів, які пов’язані з даним нейроном). Шари нейронної мережі, побудовані за описаним принципом, називаються згортальними шарами.

Загальну архітектуру згортальної нейронної мережі з двома згортальними шарами, що містять 5 і 10 нейронів відповідно, можна зобразити у вигляді схеми (рис. 1). Припустимо, що ми маємо вхідне зображення – 28x28 пікселів, розмір ядра згортки – 5x5, кількість класів – 10, фактор субдискретизації – 2. Шари субдискретизації, що йдуть після згортальних, зменшують розмір карти ознак і забезпечують інваріантність до масштабу. Після проходження кількох шарів карта ознак розкладається в вектор, який використовується в якості входу для класифікатора (повнозв’язної нейронної мережі).

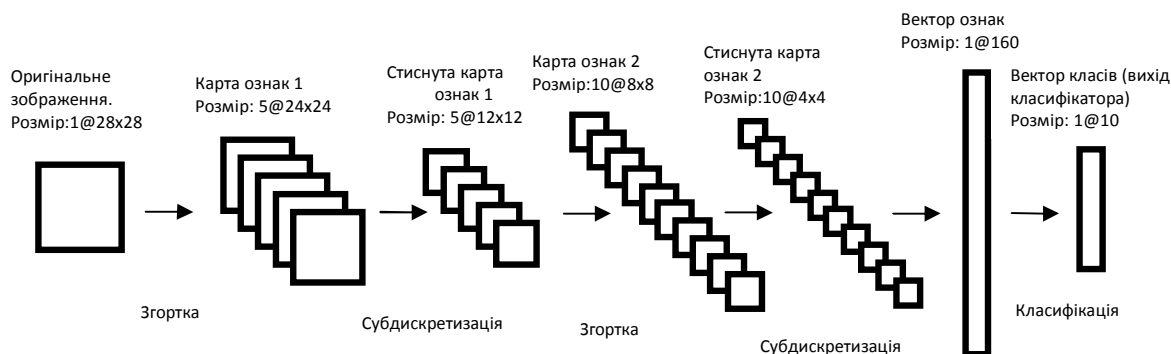


Рис. 1. Архітектура згортальної нейронної мережі

Карти ознак 1 і 2 містять нейрони, які приймають синаптичні входи від локальних рецептивних полів, тим самим визначаючи локальні ознаки. Ваги нейронів згортання в картах ознак використовуються спільні, таким чином, позиція локальних функцій стає менш важливою, що в свою чергу призводить до інваріантності зсуву. Для опису виходів нейронів згортання використовується наступний вираз:

$$y[m,n] = b + \sum_{k=0}^{K-1} \sum_{l=0}^{K-1} v[k,l]x[m+k,n+l]. \quad (1)$$

Вираз (1) описує роботу згортання на вхідному зображенні x з ядром згортання v . Єдина відмінність від стандартної згортки – граничне значення b , що додається до результату.

Після виконання операції згортання виконується операція субдискретизації – зменшення розмірів сформованих карт ознак. В даній архітектурі нейронної мережі вважається, що інформація про факт існування потрібної ознаки важливіша за точні її координати. Тому з декількох сусідніх нейронів карти ознак вибирається максимальний і приймається за один нейрон карти ознак меншого розміру. Інколи може застосовуватись операція знаходження середнього між сусідніми нейронами. За рахунок даної операції, окрім пришвидшення подальших обчислень, мережа стає більш інваріантною до масштабу вхідного зображення. Величина стиснення описується коефіцієнтом S . Для обчислення нейрона на даному шарі використовується наступний вираз:

$$y[m,n] = \phi(p) = \phi\left(b + u \sum_{k=0}^{S-1} \sum_{l=0}^{S-1} x[mS+k,nS+l]\right),$$

де
$$\phi(p) = \frac{1}{1 + \exp(-p)}.$$

Класифікація вхідного зображення відбувається на виході класифікатора. На даному шарі всі нейрони мають унікальний набір ваг, що дозволяє їм визначати складні ознаки і виконувати класифікацію. Вираз для обчислення цих класичний перцептронів визначається як:

$$y[n] = \phi(b[n] + \sum_{k=0}^{K-1} w[n,k]x[k]).$$

Важливою якістю архітектури згортальної нейронної мережі є те, що всім синаптичним вагам і значенням зміщення, мережу можна навчити використавши простий алгоритм зворотного поширення помилки, використавши навчальну вибірку.

Детектування і розпізнавання дорожніх знаків виконує повністю навчена згортальна нейронна мережа. Замість того, щоб фокусуватись на покращенні функцій і алгоритмів, простішим варіантом буде зосередити увагу на навчанні мережі використавши великий набір даних. В інших публікаціях неодноразово вказувалось, що використання великого набору даних для навчання класифікатора демонструє досить перспективні результати [4]. Тому колекція зразків навчальних наборів даних є важливою, так як безпосередньо впливає на продуктивність розпізнавання.

Для наповнення колекції даними для навчання нейронної мережі використовуються усі можливі варіанти: це можуть бути зображення дорожніх знаків знайдені за результатами пошуку в Інтернеті, самостійно зроблені фотознімки на дорогах, а також вже готові бази даних зображень дорожніх знаків. На рис. 2 показаний приклад зображень дорожніх знаків обмеження швидкості які використовуються для навчання нейронної мережі.



Рис. 2. Приклад зображень дорожніх знаків обмеження швидкості

Створення набору даних для навчання згортальної нейронної мережі лише із зображень дорожніх знаків недостатньо. Необхідний також, ще набір фонових зображень. Це дозволить запобігти хибні виявлення. На рис. 3 наведені приклади випадкових фонових зображень.



Рис. 3. Приклад фонових зображень

Навчальним алгоритмом, що використовується в роботі з згортальною нейронною мережею, буде алгоритм зворотного поширення помилки. Основною причиною вибору цього алгоритму є те, що він добре підходить при роботі в великих наборах навчальних даних. При реалізації зворотного поширення помилка вихідного шару може бути вибрана наступним чином:

$$E_n^p = \frac{1}{2} \cdot \sum_{k=1}^M (x_k - d_k)^2,$$

де M - кількість нейронів вихідного шару, k - номер вихідного нейрона, x_k - реальне значення вихідного сигналу нейрона, d_k - очікуване значення.

При навчанні мережі використовується метод найшвидшого спуску (градієнтний метод). Кореляція ваг здійснюється за формулою:

$$(\omega_n^{ij})_{new} = (\omega_n^{ij})_{old} - \eta \cdot \left(\frac{\partial E_n^p}{\partial \omega_n^{ij}} \right),$$

де $(\omega_n^{ij})_{new}$ - значення ваг після кореляції, $(\omega_n^{ij})_{old}$ - значення ваг до кореляції, η характеризує швидкість навчання (зазвичай спочатку вибирається рівним 0.0005, після чого поступово збільшується в процесі навчання). $\frac{\partial E_n^p}{\partial \omega_n^{ij}}$ обчислюється за наступною формулою:

$$\frac{\partial E_n^p}{\partial \omega_n^{ij}} = x_{n-1}^j \cdot \frac{\partial E_n^p}{\partial y_n^i},$$

де x_{n-1}^j - вихід j -го нейрона $(n-1)$ -го шару, y_n^i - скалярний добуток всіх виходів нейронів $(n-1)$ -го шару і відповідних вагових коефіцієнтів.

$$\frac{\partial E_n^p}{\partial y_n^i} = G(x_n^i) \cdot \frac{\partial E_n^p}{\partial x_n^i},$$

де $G(x_n^i)$ - похідна функції активації.

Помилку потрібно розповсюдити не попередні шари. Це можна зробити за наступною формулою:

$$\frac{\partial E_{n-1}^p}{\partial x_{n-1}^k} = \sum_i \omega_n^{ik} \cdot \frac{\partial E_n^p}{\partial y_n^i}$$

При реалізації процесу навчання для згортальних шарів мережі потрібно використовувати вищенаведені формули в матричному вигляді.

Значення початкових синапсичних ваг вибирається на основі рівномірного розподілу з нульовим математичним очікуванням і дисперсією, протилежним квадратному кореню із числа синапсичних зв'язків нейрона. Проте можна проводити початкову ініціалізацію ваг на основі дисперсії, рівній 0.05.

Для зручності і пришвидшення навчання мережі вхідне значення пікселів зображення нормується за формулою:

$$y_i = \frac{x_i}{128} - 1,$$

де x_i - значення пікселя початкового зображення, y_i - значення, що подається на вхід мережі.

Крім того, для пришвидшення навчання мережі можна пропустити те зразки, для яких помилка вже мала на даному етапі навчання. Це дозволить пришвидшити процес в кілька разів.

Описана вище методика навчання дозволяє досягти збіжності мережі, але є досить повільною. Тому актуальним є використання різноманітних методів другого порядку, які дозволяють прискорити навчання мережі в десятки разів.

Перш за все, слід відмовитись від ідеї пакетного навчання мережі, коли всі зразки із навчальної вибірки постійно подаються на вхід мережі. Значно швидше працюють стохастичні методи навчання, при яких на даному етапі із навчальної вибірки по визначеним критеріям вибирається лише частина зразків для навчання. Ефективний метод описаний в працях Яна ЛеКуна під назвою стохастичний діагональний метод Левенберга-Марквардта [5].

Суть даного алгоритму полягає в наступному: перед кожним навчальним циклом із бази вибирається випадковим чином 500 зображень. Для кожного з них обчислюється Гессіан, потім визначається їх сума і визначається середнє значення для всіх 500 зразків. Далі виконується визначення коефіцієнта η , що характеризує швидкість навчання:

$$\eta_{ki} = \frac{\varepsilon}{\frac{\partial^2 E}{\partial \omega_{ki}^2} + \mu},$$

де ε характеризує швидкість навчання, μ - коефіцієнт, що перешкоджає надмірному збільшенню швидкості навчання, $\frac{\partial^2 E}{\partial \omega_{ki}^2}$ - гессіани, що визначаються окремо для кожного вагового коефіцієнта.

Ці обчислення проводяться для випадково вибраних 500 зразків перед кожним навчальним циклом.

Висновки. Використання згортальних нейронних мереж для систем технічного зору в автомобілях, а конкретніше для розпізнавання дорожніх знаків, є ефективним і достатньо гнучким варіантом. Основним недоліком можна вважати довгий і клопіткий процес навчання мережі, а також необхідність великого набору даних для навчальної вибірки. Проте результати розпізнавання таких систем є досить хорошими. Як підтверджують неодноразові дослідження, згортальні нейронні мережі здатні розпізнати дорожні знаки в середньому з ймовірністю 98.8 %.

ЛІТЕРАТУРА:

1. M.L. Eichner, T.P. Breckon, "Integrated Speed Limit Detection and Recognition from Real-Time Video", Proc. IEEE Intelligent Vehicle Symposium, 2008.
2. Хайкин, С. Нейронные сети: полный курс / С. Хайкин. – М.: Вильямс, 2006. – 1104 с.
3. Круглов В.В. Искусственные нейронные сети. Теория и практика. / В.В. Круглов, В.В. Борисов М.: Телеком, 2001.
4. D. Nistér, H. Stewénius, "Scalable recognition with a vocabulary tree", CVPR, issue 2, pp. 2161-2168, 2006.
5. Y. LeCun, L. Bottou, G. Orr, and K. Muller, "Efficient BackProp," in Neural Networks: Tricks of the trade, (G. Orr and Muller K., eds.), 1998.

Рецензент: д.т.н., проф. Катеринчук І.С., начальник кафедри зв'язку та інформатизації, Національна академія Державної прикордонної служби України імені Б. Хмельницького

к.т.н., доц. Джулий В.Н., Жигальский П.В., Солодеева Л.В.

МЕТОД РАСПОЗНАВАНИЯ ДОРОЖНЫХ ЗНАКОВ С ИСПОЛЬЗОВАНИЕМ СВЕРТЫВАЮЩЕЙ НЕЙРОННОЙ СЕТИ

В статье рассмотрен подход к реализации системы детектирования и распознавания дорожных знаков, основанный на использовании свертывающей искусственной нейронной сети. Использование такой архитектуры искусственных нейронных сетей позволяет избежать многих проблем с которыми сталкиваются при разработке систем технического зрения. Свертывающиеся нейронные сети лучше приспособлены для распознавания объектов на изображении, поэтому они вполне подходят для применения в системах распознавания дорожных знаков в режиме реального времени. По сравнению с классическими, свертывающей нейронные сети менее чувствительна к разного рода искажения изображения, а именно поворотов, оползней, масштабирование и др. Для обучения используется классический метод обратного распространения ошибки. Однако для эффективной работы такой нейронной сети необходима большая выборка образцов для обучения.

Ключевые слова: распознавание дорожных знаков, интеллектуальный автомобиль, свертывающей нейронные сети.

НАЦІОНАЛЬНИЙ
УНІВЕРСИТЕТ

Ph.D Julie V.M., Zhyhalsky P.V., Solodeeva LV

THE METHOD OF RECOGNIZING ROAD SIGNS COAGULATION USING NEURAL NETWORK

The article describes the embodiment of the detection and recognition of road signs based on the use of convolution artificial neural network convolution. Using this architecture of artificial neural networks to avoid many of the problems faced in the development of vision systems. Convolutional neural networks are better suited for recognizing objects in an image, so they are well suited for use in recognizing traffic signs in real time. Compared to classical, convolutional neural networks less susceptible to all sorts of image distortion, namely rotations, offsets, scaling, etc.. For training using the classic method of back propagation. However, for efficient operation of such a neural network requires a large sampling for training, then it has a direct impact on the recognition results.

Keywords: road sign recognition, intelligent car, roll neural network.