

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ

НАУКОВА РОБОТА

для участі у Всеукраїнському конкурсі студентських та наукових робіт з
природничих, технічних і гуманітарних наук

Тема: CNN-технологія розпізнавання об'єктів за даними аерофотозйомки

Шифр роботи «PM2018-2»

2018-2019 н.р.

ВСТУП

Одними з найважливіших областей досліджень і розробок сучасної кібернетики є області машинного навчання, розпізнавання образів і комп'ютерного зору. Прискорюються темпи розвитку технологій інформаційного суспільства, розвиток робототехніки, розвиток концепцій «розумний будинок» і «розумне місто», розвиток інтернету речей і систем штучного інтелекту визначають цій області особливе місце в сучасному науковому знанні. У багатьох прикладних задачах в практиці сучасного програмування існують методи збору даних, кластеризації та класифікації, методи статистичного висновку. У повсякденному житті, як і в корпоративну, і в промислову середовища починають впроваджуватися технології, поступово руйнують границю між реальним і віртуальним простором, що вимагає нового якісного рівня повсюдно впроваджуваних технологій розпізнавання, чия сфера застосування в останні роки зростає колосально: вважалися складними раніше завдання розпізнавання сьогодні цілодобово вирішуються мобільними пристроями пересічних громадян. Комп'ютеризовані простори з вираженою топологією, такі як «розумний будинок» рядового користувача, розрахована на багато користувачів доповнена і розрахована на багато користувачів віртуальна реальності різного ступеня занурення, ускладнюючий штучний інтелект в комп'ютерних іграх різного призначення вимагають нових ідей і підходів, нового рівня точності та швидкості розпізнавання образів.

Ефективність розпізнавання образів в процесі роботизації економіки і виробництва, роботизації рятувальних заходів, роботизації збройних сил і приватних армій, що діють в інтересах держави та інших корпоратократичних об'єднань, стає питанням збереження людських життів.

Метою даної роботи є розробка проекту ефективної багатоступінчастої системи виявлення об'єктів військової техніки і військової інфраструктури, будинків, окопів, тощо. Визначення їх параметрів в умовах недостатньої

видимості, шумів різного походження, низької частоти кадрів і невисокою роздільної здатності реєструючого обладнання. Система повинна бути розроблена для аналізу кадрів відзнятих із БПЛА.

Одним з найбільш сучасних підходів до вирішення задачі розпізнавання образів є згортова нейронна мережа, що була представлена французьким дослідником Яном Лекуном як продовження неокогнітрону. Її ефективність була доведена у 2016 році, коли CNN Trimps-Soushen виграв змагання ILSVRC із похибкою розпізнавання в 2.991%, що є неможливим для людини.

В процесі виконання поставленого завдання необхідно:

1. Розглянути основні принципи та існуючі методи розв'язання задач машинного навчання, розпізнавання образів і комп'ютерного зору, провести аналіз доступних існуючих систем.
2. Розробити алгоритм і детально описати запропонований підхід до вирішення поставленого завдання з використанням моделей.
3. Навчити класифікатор з використанням моделей.
4. Провести випробування запропонованої підсистеми розпізнавання з використанням моделей.

РОЗДІЛ 1

ОСНОВНІ ТЕОРЕТИЧНІ ПОЛОЖЕННЯ

У даному розділі розглядаються основні види ШНМ, їх складові. Розглядаються основні області практичного застосування моделей штучних нейронних мереж, їх переваги та недоліки. Розглядаються алгоритми навчання моделей та оптимізації процесу навчання.

1.1. Що таке нейронні мережі

Дослідження по штучних нейронних мереж (далі — нейронні мережі) пов'язані з тим, що спосіб обробки інформації людським мозком в корені відрізняється від методів, застосовуваних звичайними цифровими комп'ютерами. Мозок є надзвичайно складний, нелінійний, паралельний комп'ютер (систему обробки інформації). Він має здатність організувати свої структурні компоненти, звані нейронами, так, щоб вони могли виконувати конкретні завдання (такі як розпізнавання образів, обробку сигналів органів почуттів, моторні функції) у багато разів швидше, ніж можуть дозволити найшвидкодіючі сучасні комп'ютери. Прикладом такого завдання обробки інформації може служити звичайне зір. У функції зорової системи входить створення кончини навколишнього світу в такому вигляді, який забезпечує можливість взаємодії з цим світом. Більш точно, мозок послідовно виконує ряд задач розпізнавання (наприклад, розпізнавання знайомого обличчя у незнайомому оточенні). На це він витрачає 100-200 мілісекунд, у той час, як виконання подібних задач на комп'ютері може зайняти декілька днів.

Іншим прикладом може служити локатор (сонар) кажана, який представляє собою систему активної ехолокації. Крім надання інформації про відстань до потрібного об'єкта цей локатор надає інформацію про відносну швидкість об'єкта, про його розміри, а також про азимут та висоту

руху об'єкта. Для виділення цієї інформації із отриманого сигналу крихітний мозок кажана проводить складні нейронні обрахунки. Ехолокація кажана за своїми характеристиками якості і швидкодії перевищує найскладніші прилади, створені інженерами [5].

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

Що ж дозволяє мозку людини або кажана досягати таких результатів? При народженні мозок має досконалу структуру, яка дозволяє будувати власні правила на основі того, що ми називаємо «досвідом». Досвід накоплюється з плином часу, і особливо найбільші зміни відбуваються у перші два роки життя людини.

Поняття розвитку нейронів пов'язано з поняттям пластичності мозку — здатністю настройки нервової системи відповідно до навколишніх умов. Саме пластичність грає найважливішу роль в роботі нейронів в якості одиниць обробки інформації в людському мозку. Аналогічно, в штучних нейронних мережах робота проводиться з штучними нейронами.

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

Нейронна мережа — це величезний розподілений паралельний процесор, що складається з елементарних одиниць обробки інформації, які накопичують експериментальні знання і представляють їх для подальшої обробки. Нейронна мережа схожа з мозком з двох точок зору.

- Знання надходять в нейронну мережу з навколишнього середовища і використовуються в процесі навчання;
- Для накопичення знань застосовуються зв'язки між нейронами, які називаються синаптичними вагами.

Процедура, яка використовується для процесу навчання, називається алгоритмом навчання. Ця процедура вибудовує в певному порядку синаптичні ваги нейронної мережі для забезпечення необхідної структури взаємозв'язків нейронів.

Зміна синаптичних ваг являє собою традиційний метод настройки нейронних мереж. Цей підхід дуже близький до теорії лінійних адаптивних фільтрів, яка вже давно заявила про себе і застосовується в різних областях

діяльності людини. Однак нейронні мережі можуть змінювати свою топологію. Це обумовлено тим фактом, що нейрони в людському мозку постійно відмирають, а нові синаптичні зв'язки постійно створюються.

Найважливіші властивості біологічних нейромереж:

1. Паралельність обробки інформації. Кожен нейрон формує свій вихід тільки на основі своїх входів і власного внутрішнього стану під впливом загальних механізмів регуляції нервової системи.

2. Здатність до повної обробки інформації. Всі відомі людині завдання вирішуються нейронними мережами. До цієї групи властивостей відносяться асоціативність (мережа може відновлювати повний образ по його частині), здатність до класифікації, узагальнення, абстрагування і безліч інших. Вони до кінця не систематизовані.

3. Самоорганізація. В процесі роботи біологічні НМ самостійно, під впливом зовнішнього середовища, навчаються вирішення різноманітних завдань. Невідомо ніяких принципових обмежень на складність завдань, що вирішуються біологічними нейронними мережами. Нервова система сама формує алгоритми своєї діяльності, уточнюючи і ускладнюючи їх протягом життя. Ця властивість НМ народжує безліч питань. Адже кожна замкнута система в процесі розвитку спрощується, деградує.

4. Біологічні НМ є аналоговими системами. Інформація надходить в мережу по великій кількості каналів і кодується за просторовим принципом: вид інформації визначається номером нервового волокна, по якому вона передається. Амплітуда вхідного впливу кодується щільністю нервових імпульсів, що передаються по волокну.

5. Надійність. Біологічні НМ володіють фантастичною надійністю: вихід з ладу навіть 10% нейронів в нервовій системі не перериває її роботи.

1.2. Види багат шарових штучних нейронних мереж

Відмінності обчислювальних процесів в нейронних мережах часто обумовлені способом взаємозв'язків нейронів. За сукупністю критеріїв на сьогоднішній день багат шарові архітектури штучних нейронних мереж можна розділити на статичні і динамічні. Кожен з класів архітектур нейронних мереж може включати безліч підкласів, реалізуючи різні підходи, нижче будуть наведені основні з них.

До статичних архітектур відносять мережі прямого поширення, в яких реалізована однонаправлений зв'язок між шарами, відсутні динамічні елементи і зворотний зв'язок, а вихід навченої штучної нейронної мережі однозначно визначається входом і не залежить від попередніх станів мережі.

Статичні нейронні мережі прямого поширення:

- Персептрон
- Нейронні мережі Кохонена
- Когнітрон
- Неокогнітрон
- Сучасна згортова мережа

На противагу статичним архитектурам, існують динамічні архітектури штучної нейронної мережі, що реалізують рекурентну структуру з використанням зворотних зв'язків, завдяки чому стан мережі в кожному момент часу залежить від попереднього стану. Рекурентні штучні нейронні мережі як правило базуються на багат шаровому персептроні.

Динамічні нейронні мережі:

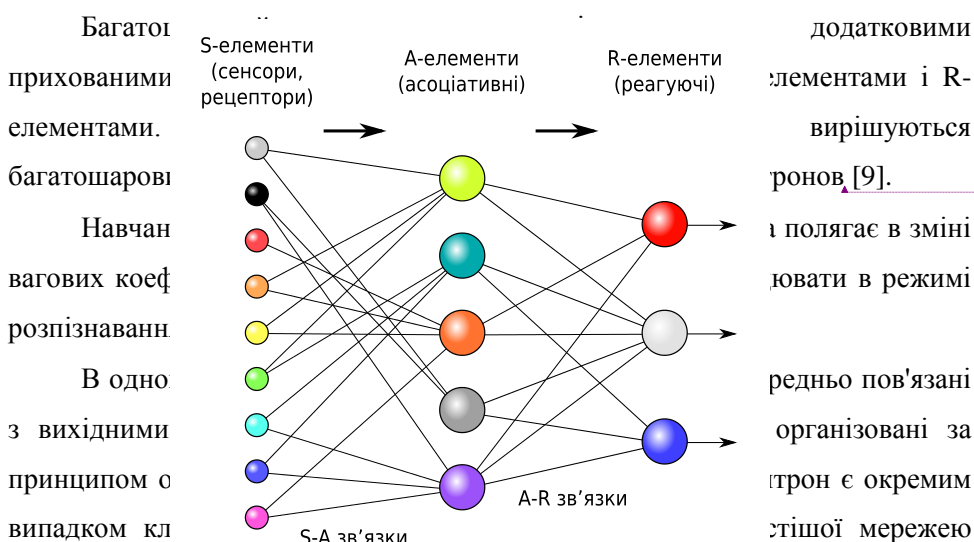
- Нейронна мережа Хопфілда
- Нейронна мережа Коско
- Нейронна мережа Джордана
- Нейронна мережа Елмана

Отформатовано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

- Нейронна мережа довгої короткочасної пам'яті

1.2.1. Персептрон

Елементарний персептрон організовується на основі сенсорних даних на вході - S-елементів, асоціативних елементів — A-елементів, і реагуючих елементів на виході — R-елементів. Набір S-елементів, пов'язаний з A-елементом утворює асоціацію, і A-елемент активізується після того як досягнуто певне число сигналів від S-елементів. A-елемент передає зважений сигнал на сумуючий R-елемент, і в залежності від того, чи перевищує зважена сума деякий поріг, R-елемент видає результат роботи персептрону (рис.1.1).



Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

прямого поширення — лінійним класифікатором, і має безліч принципів обмежень, таких як неможливість реалізації функції XOR.

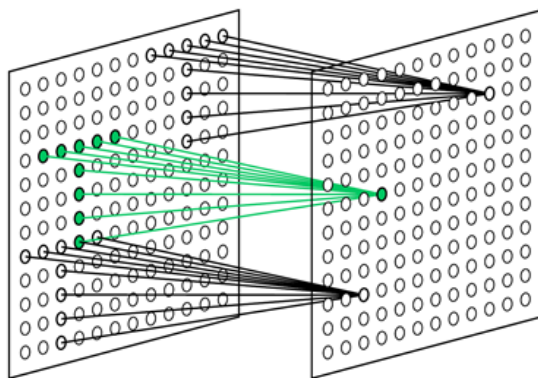
Отформатовано: Шрифт: полужирный, Цвет шрифта: Черный, Русский (Россия)

1.2.2. Когнітрон

Когніторон був розроблений на основі будови біологічної зорової кори, має ієрархічну принципово багат шарову архітектуру. Нейрони між шарами когнітрону пов'язані тільки локально, і кожен шар реалізує різні рівні узагальнення: вхідні шари сприймають прості образи, в разі зображень, такі

як лінії, великі однорідні ділянки, їх орієнтацію і локалізацію в просторі вхідних даних, в той час як глибокі шари сприймають складніші абстрактні структури, незалежні від локалізації та інших простих ознак способу.

Когнітро
гальмуючих ш
вході нейрона
когнітрон, що
конструювався
сигналів визн
постсинаптиче
когнітрону сп



їх збудливих і
рчих сигналів на
спрощені моделі
вчатку когнітрон
аптичне простір
вщині нейронів,
площині. Нейрон
сигналів, а тільки

його частину, чим і реалізується принцип локальної зв'язності. Область пресинаптичного простору, яка відповідає області постсинаптичного простору сигналу, що впливають на стан даного нейрона, називається його локальним рецептивним полем (рис.1.2). Рецептивні поля близьких один до одного постсинаптичних нейронів, звані зонами конкуренції, перекриваються, тому активність даного пресинаптичного нейрона позначається на все більшого поширення області постсинаптичних нейронів наступних шарів ієрархії. Розміри зон конкуренції обумовлюють кількість сприймаємих в просторової області ознак.

Рис 1.2. Рецептивне поле. Кожний нейрон

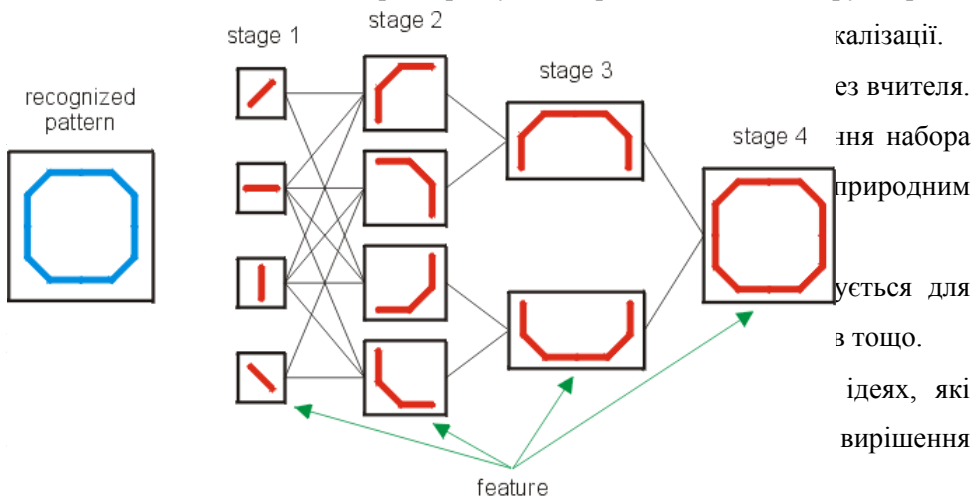
Когнітрон навчається на основі принципу самоорганізації, в класичному випадку реалізується алгоритм навчання без учителя.

1.2.3. Неокогнітрон та згорткові нейронні мережі

Неокогнітрон є прямим розвитком ідей, що лежать в основі когнітрон і точніше моделює структуру зорової кори головного мозку і є класифікатором, здатним до робастого розпізнаванню образів.

Кожен шар неокогнітрона складається з площини простих S-нейронів і площині складних C-нейронів, які також організують локальну зв'язність. Локальне рецептивне поле площині S-нейронів наступного шару формується

з пресинаптичних сигналів площини С-нейронів попереднього шару. Локальні ознаки способу сприймаються S-нейронами, а спотворення локальних ознак компенсуються С-нейронами. В результаті цього процесу кожен шар після вхідного має своїм входом все більш узагальнену картину, утворену С-нейронами попередніх шарів. З кожним рівнем глибини первинні прості ознаки детектуються у все більш складних сполученнях (рис.1.3). Площина S-нейронів можна розглядати як один нейрон, ваги якого визначають ядро згортки, що застосовується до попереднього шару у всіх можливих позиціях. Всі S-нейрони реагує на образ, відповідний ядру згортки,



1.3. Навчання нейронних мереж

Найважливішою особливістю нейронних мереж є їх здатність навчатися на основі даних оточуючого середовища і у результаті підвищувати свою продуктивність. Підвищення продуктивності мережі відбувається з часом у відповідно до певних правил. Процес навчання нейронної мережі відбувається за допомогою інтерактивного процесу коригування синаптичних ваг та порогів. В ідеальному випадку нейронна мережа отримує знання про навколишнє середовище на кожній ітерації процесу навчання. З поняттям навчання асоціюється досить багато видів

Отформатовано: Шрифт: 16 пт, полужирный, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 16 пт, полужирный, Цвет шрифта: Черный, Русский (Россия)

діяльності, тому складно дати точне визначення цього процесу. З точки зору нейронних мереж можна використовувати наступне визначення.

Навчання — це процес в якому вільні параметри нейронної мережі настраюються за допомогою моделювання середовища, в яку ця мережа вбудована. Тип навчання визначається способом підстроювання цих параметрів.

Це визначення процесу навчання припускає наступну послідовність подій.

1. У нейронну мережу надходять стимули із зовнішнього середовища.
2. В результаті цього змінюються вільні параметри нейронної мережі.
3. Після зміни внутрішньої структури нейронна мережа відповідає на збудження вже іншим чином.

Вищевказаний список правил вирішення проблеми процесу навчання називається алгоритмом навчання [5].

Отформатовано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

1.3.1. Парадигми навчання ШНМ

На сьогодні відомо три парадигми навчання нейронних мереж, в основу яких покладено особливості машинного навчання:

- Навчання з вчителем;
- Навчання без вчителя;
- Навчання з підкріпленням;

1.3.2. Правила навчання штучних нейронних мереж

Теорія навчання розглядає три фундаментальні властивості, пов'язані з навчанням за прикладами: ємність, складність зразків і обчислювальна складність. Під ємністю розуміється, скільки зразків може запам'ятати мережа, і які функції і межі ухвалення рішень можуть бути на ній сформовані. Складність зразків визначає число навчальних прикладів, необхідних для досягнення здатності мережі до узагальнення. Дуже мале

число прикладів може викликати «перенавчання» мережі, коли вона добре функціонує на прикладах навчальної вибірки, але погано — на тестових прикладах, які підлягають тому ж статистичному розподілу. Відомі 4 основних типи правил навчання: корекція за помилкою, навчання Больцмана, Дельта правило і навчання методом змагання. Ці правила є базовими і основними для навчання ШНМ, всі решта правила які існують є модифікаціями цих правил.

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

1.3.3. Метод зворотного поширення помилки

Метод зворотного поширення помилки — метод навчання багат шарового перцептрон. Це ітеративний градієнтний алгоритм, який використовується з метою мінімізації помилки роботи багат шарового перцептрон та отримання бажаного виходу. Основна ідея цього методу полягає в поширенні сигналів помилки від виходів мережі до її входів, в напрямку, зворотному прямому поширенню сигналів у звичайному режимі роботи. Барц і Охонін запропонували відразу загальний метод («принцип подвійності»), який можна застосувати до ширшого класу систем, включаючи системи з запізненням, розподілені системи, тощо. Для можливості застосування методу зворотного поширення помилки функція активації нейронів повинна бути диференційованою [5].

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Алгоритм зворотного поширення помилки застосовується для багат шарового перцептрон. У мережі є множина входів x_1, \dots, x_n , множина виходів (*Outputs*) і безліч внутрішніх вузлів. Перенумеруємо всі вузли (включаючи входи і виходи) числами від 1 до N (наскрізна нумерація, незалежно від топології шарів). Позначимо через w_{ij} вагу зв'язку, що з'єднує i -й і j -й вузли, а через o_i — вихід i -го вузла. Якщо нам відомий навчальний приклад (правильні відповіді мережі $t_k, k \in Outputs$), то функція помилки, отримана за методом найменших квадратів, виглядає так:

$$E(\{w_{ij}\}) = \frac{1}{2} \sum_{k \in Outputs} (t_k - o_k)^2. \quad (1.1)$$

Як модифікувати ваги? Ми будемо реалізовувати стохастичний градієнтний спуск, тобто будемо підправляти ваги після кожного навчального прикладу i , таким чином, «рухатися» в багатовимірному просторі ваг. Щоб «добратися» до мінімуму помилки, нам потрібно «рухатися» в сторону, протилежну градієнту, тобто, на підставі кожної групи правильних відповідей, додавати до кожної ваги w_{ij}

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}, \quad (1.2)$$

де $0 < \eta < 1$ — множник, що задає швидкість «руху».

Похідна розраховується таким чином. Нехай спочатку $j \in Outputs$, тобто вага, яка нас цікавить, входить в нейрон останнього рівня. Спочатку зазначимо, що w_{ij} впливає на вихід мережі лише як частина суми $S_j = \sum_i w_{ij} x_i$, де сума береться по входах i -го вузла. Тому

$$\begin{aligned} \frac{\partial E}{\partial S_j} &= \frac{\partial E}{\partial o_j} \frac{\partial o_j}{\partial S_j} = \left(\frac{\partial}{\partial o_j} \frac{1}{2} \sum_{k \in Outputs} (t_k - o_k)^2 \right) \left(\frac{\partial \sigma(S_j)}{\partial S_j} \right) \\ &= \left(\frac{1}{2} \frac{\partial}{\partial o_j} (t_j - o_j)^2 \right) (o_j (1 - o_j)) = -o_j (1 - o_j) (t_j - o_j), \end{aligned} \quad (1.3)$$

де σ — це функція активації, у даному випадку (стосовно обчислення похідної) являє собою Експоненційну сигмоїду (функцію Фермі) розглянуту вище.

Якщо ж i -й вузол — не на останньому рівні, то у нього є виходи; позначимо їх через $Children(j)$. У цьому випадку

$$\frac{\partial E}{\partial S_j} = \sum_{k \in Children(j)} \frac{\partial E}{\partial S_k} \frac{\partial S_k}{\partial S_j}, \quad (1.4)$$

і

$$\frac{\partial S_k}{\partial S_j} = \frac{\partial S_k}{\partial o_j} \frac{\partial o_j}{\partial S_j} = w_{i,j} \frac{\partial o_j}{\partial S_j} = w_{i,j} o_j (1 - o_j). \quad (1.5)$$

А $\frac{\partial E}{\partial S_k}$ — це це аналогічна поправка, але обчислена для вузла наступного рівня (будемо позначати її через δ_k - від Δ_k вона відрізняється відсутністю

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

множника $(-\eta x_{ij})$ Оскільки ми навчилися обчислювати поправку для вузлів останнього рівня і виразити поправку для вузла нижчого рівня через поправки більш високого, можна вже створювати алгоритм навчання. Саме через цю особливість обчислення поправок цей алгоритм називається *алгоритмом зворотного поширення помилки*.

Коротке викладення вищесказаного:

- Для вузла останнього рівня

$$\delta_j = -o_j(1 - o_j)(t_j - o_j). \quad (1.6)$$

- Для внутрішнього вузла мережі

$$\delta_j = -o_j(1 - o_j) \sum_{k \in \text{Outputs}(j)} \delta_k w_{j,k}. \quad (1.7)$$

- Для всіх вузлів

$$\Delta w_{i,j} = -\eta \delta_j x_i. \quad (1.8)$$

На практиці дуже гарні результати показують мережі досить простої структури, що складаються з двох рівнів нейронів — прихованого рівня (hidden units) і нейронів-виходів (output units), кожен вхід мережі з'єднаний з усіма прихованими нейронами, а результат роботи кожного прихованого нейрона подається на вхід кожному з нейронів-виходів. У такому випадку досить подавати на вхід кількість нейронів прихованого рівня.

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

РОЗДІЛ 2

ОПИС АРХИТЕКТУРИ ЗГОРТКОВОЇ НЕЙРОННОЇ

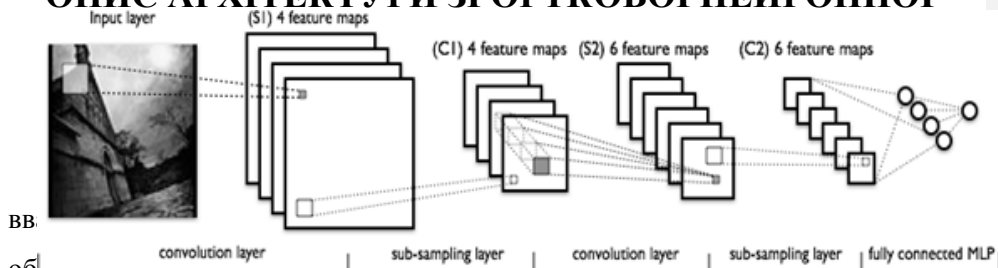


Рис. 2.2. Топологія згорткової нейронної мережі на відомому міжнародному конкурсі з розпізнавання образів ImageNet.

18 Саме тому в своїй роботі я використовував згорткову нейронну мережу, засновану на принципах неокогнітрона і доповнену навчанням за алгоритмом зворотного поширення помилки.

13.5

2.1. Згорткова нейронна мережа

9 ЗНМ складається з різних видів шарів: згорткові (convolutional) шари, агрегування (subsampling) шари і шари «звичайної» нейронної мережі.

4.5

Перехідні шари (subsampling), чергуючись між собою, формують піраміду розв'язування.

0

Згорткові мережі є далеким середнім між звичайними перцептронами і звичайним багатшаровим перцептроном. На сьогоднішній день

кращі результати в розпізнаванні зображень отримують з їх допомогою. В середньому точність розпізнавання таких мереж перевершує звичайні ШНМ на 10-15%. ЗНМ — це ключова технологія глибинного навчання.

Основною причиною успіху ЗНМ стало концепція загальних ваг. Незважаючи на великий розмір, ці мережі мають невелику кількість параметрів, що настроюються в порівнянні з їх предком — неокогнітроном. Є варіанти ЗНМ (Tiled Convolutional Neural Network), схожі на неокогнітрон, в таких мережах відбувається часткова відмова від пов'язаних ваг, але алгоритм навчання залишається тим же і ґрунтується на зворотному

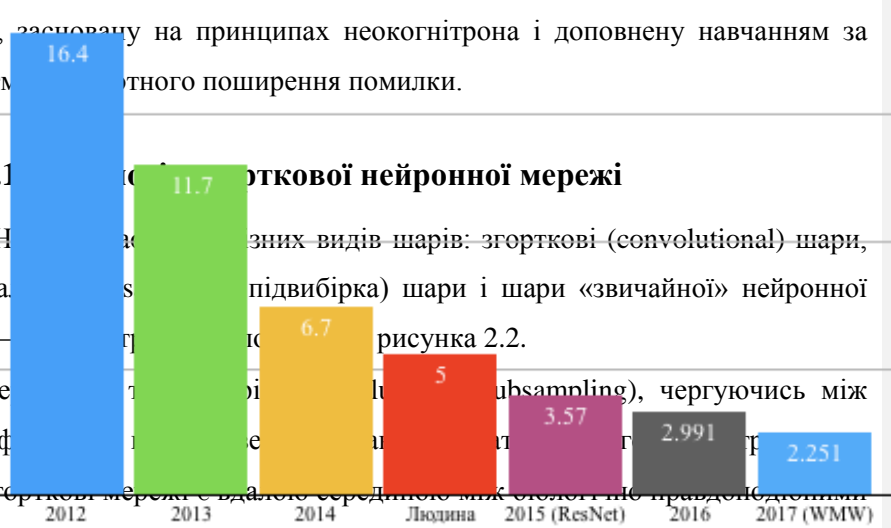


Рис. 2.1. Результати конкурсу ImageNet з 2012 по 2017 роки

Отформатировано: Шрифт: полужирный, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

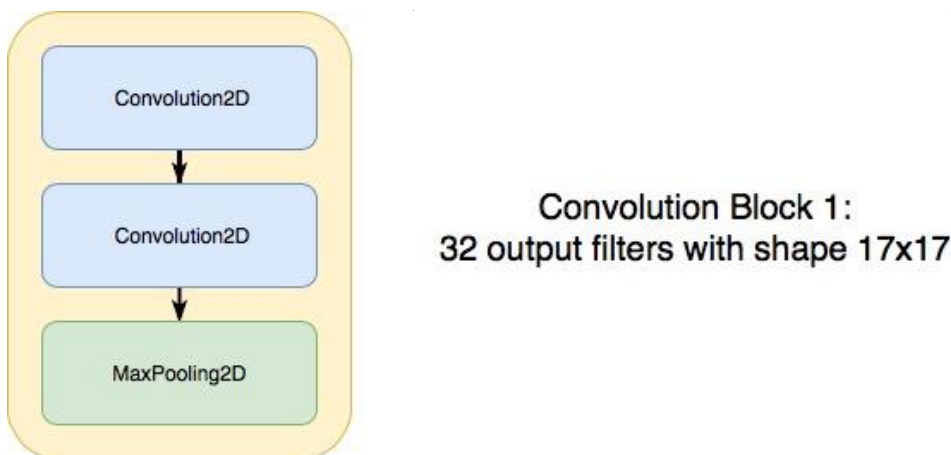
поширенні помилки. ЗНМ можуть швидко працювати на послідовній машині і швидко навчатися за рахунок чистого розпаралелювання процесу згортки по кожній карті, а також зворотної згортки при поширенні помилки по мережі.

2.2. Топологія розробленої нейронної мережі

У додатку А наведено граф створеної нейронної мережі. Дана нейронна мережа утворена з 2-х згорткових блоків та багат шарового перцептрону. Кожен згортковий блок складається з двох шарів згортки та одного шару пулінгу.

2.2.1. Перший згортковий блок

На рис. 2.3 зображено схему першого згорткового блоку. Блок складається з двох шарів згортки та одного шару максимального пулінгу.



входу і застосовує функцію активації ReLU для кожного нового значення, формуючи 2-вимірну карту збудження цього фільтру. Нижче наведено приклад вхідного зображення (рис. 2.4) та 3 карти ознак (рис. 2.5) отримані після згортки.

Надалі отримані карти до другого згорткового шару. На цьому згортковому шарі кількість фільтрів до 32 з тією ж розмірністю 3×3 та активація ReLU. На ньому відбуваються

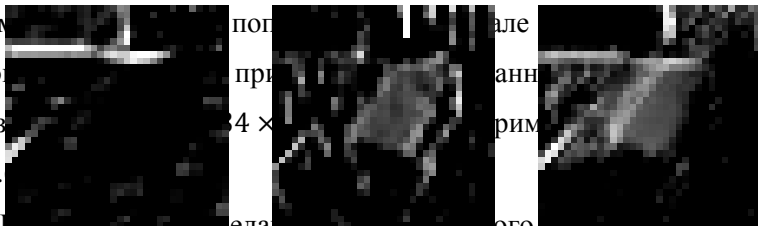


Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

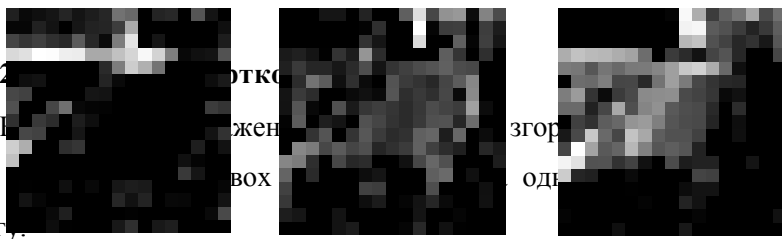
такі самі, але більшою частотою карт ознак (рис. 2.6).



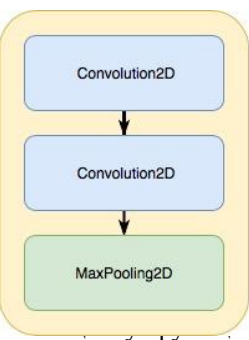
Ці карти ознак передаються до наступного шару пулінгу. Цей шар розділяє вхідне зображення на набір прямокутників розмірністю 2×2 без перекриттів, і для кожної такої підобласті виводить її максимум. Ідея полягає в тому, що точне положення ознаки не так важливе, як її грубе положення відносно інших ознак. Агрегувальний шар слугує поступовому скороченню просторового розміру представлення для зменшення кількості параметрів та об'єму обчислень у мережі, і відтак також для контролю перенавчання. Отримані зображення мають вже розмірність 17×17 пікселів.

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Розмір карти ознак згортки згорнутий до розміру 17×17 пікселів. Цей блок також виконує операцію максимального пулінгу.



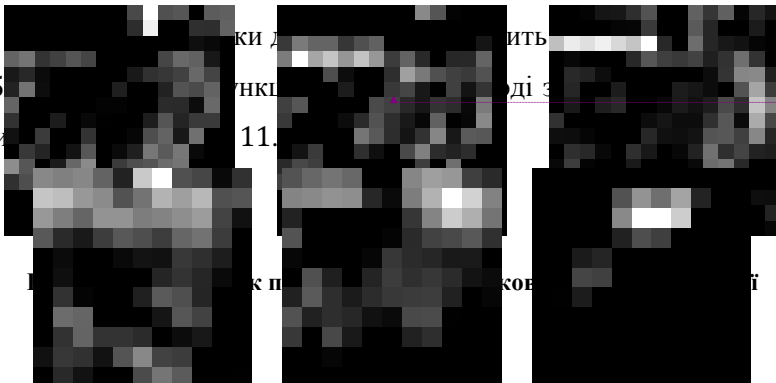
Розмір карти ознак згортки згорнутий до розміру 17×17 пікселів. Цей блок також виконує операцію максимального пулінгу. Розмір карти ознак згортки згорнутий до розміру 17×17 пікселів. Цей блок також виконує операцію максимального пулінгу.



Convolution Block 2: 64 output filters with shape 5x5

Рис. 2.8. Схема другого блоку згортки карти ознак розмірністю 15×15 .

Розмір карти ознак згортки згорнутий до розміру 17×17 пікселів. Цей блок також виконує операцію максимального пулінгу. Розмір карти ознак згортки згорнутий до розміру 17×17 пікселів. Цей блок також виконує операцію максимального пулінгу.



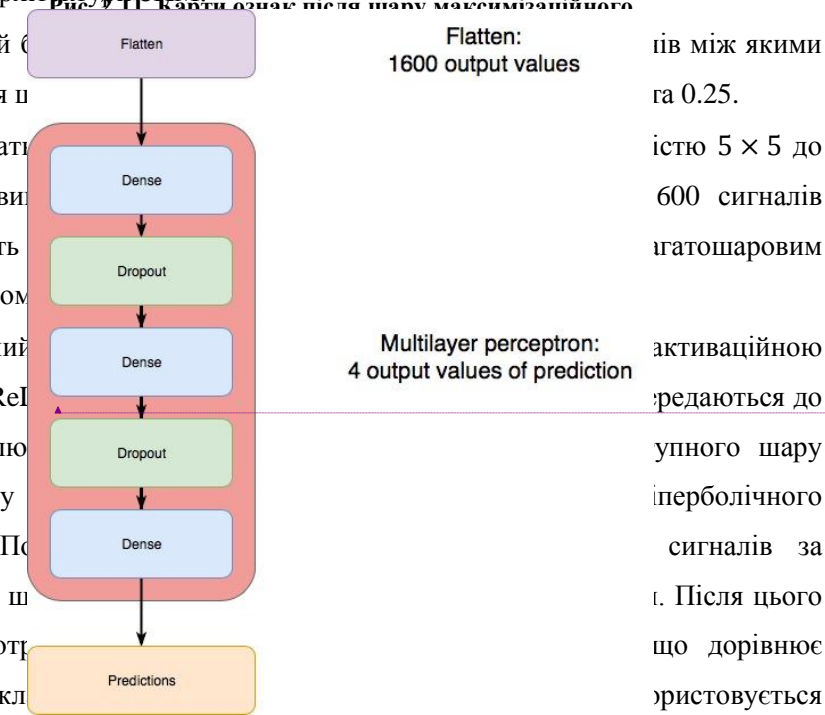
Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Далі знову застосовується шар максимізаційного пулінгу. Що повертає нам 64 карти

2.2.3.

На р... бр... ар... он, який є частиною архітектури ЗНМ.



Даний б... знаходяться п...

Спочатк... лінійного ви... потрапляють перцептроном

Перший... функцією ReL... шару відклю... перцептрону тангенсу. По... допомогою ш... сигнали потр... кількості кл...

Flatten

Flatten:
1600 output values

ів між якими га 0.25.

істю 5 × 5 до 600 сигналів гатошаровим

Multilayer perceptron:
4 output values of prediction

активаційною :редаються до упного шару іперболічного сигналів за і. Після цього що дорівнює ристовується

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

активаційна функція Softmax. Функція Softmax задається наступним чином:

Рис.2.12. Багатошаровий перцептрон

$$\sigma: \mathbb{R}^K \rightarrow [0,1]^K$$

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad (2.1)$$

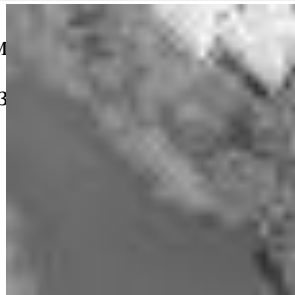
На виході з перцептрону ми отримуємо вектор значень ймовірностей з якими початкове зображення належить до кожного класу.

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

2.3. Вхідні дані та їх обробка

Як б... ньом... до нейронної мережі по... і роз... початковими



Отформатовано: Шрифт: 16 пт, полужирный, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: полужирный, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: полужирный, Цвет шрифта: Черный, Русский (Россия)

даними можуть бути зображення у форматі JPEG будь-якої розмірності. Для отримання даних для тренування та тестування мережі було написано ПЗ, яке зчитує зображення з файлової системи, конвертує зчитані дані до відтінків сірого, нормалізує значення, змінює розмірність до 38 × 38, присвоює класи зображенням в залежності від дерикторії в яких вони зберігались та зберігає усі дані у бінарний файл, що надає можливість швидкого зчитування даних у подальшому.

2.4. Результати тестування нейронної мережі

Було проведено 3 тести на загальну стійкість нейронної мережі. У кожному тесті було 5 разів випадковим чином перемішано дані для навчання та тестування. У кожному тесті змінювалась кількість даних для тренування та кількість епох навчання.

У першому тесті було розподілено вибірку наступним чином, 1660 тренувальних зображень та 184 тестових. Було проведено 60 епох навчання із батчами по 8 зображень. Результати наведені у таблиці 1.

Таблиця 1

Результати ШНМ після 60 епох навчання (1660 тренувальних даних, 184 тестових)

	1	2	3	4	5	Середнє
Точність	0.8152	0.7935	0.788	0.7935	0.7337	0.78478
Функція втрат	1.5942	1.3328	1.241	1.51	1.8594	1.50748
Час навчання, с	369	365	368	372	369	368.6

У другому тесті було розподілено вибірку наступним чином, 1476 тренувальних зображень та 368 тестових. Було проведено 80 епох навчання із батчами по 8 зображень. Результати наведені у таблиці 2.

Отформатовано: Шрифт: 16 пт, полужирный, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 16 пт, полужирный, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: (по умолчанию) Times New Roman, 12 пт, не полужирный, Цвет шрифта: Авто, Русский (Россия)

У третьому тесті було розподілено вибірку наступним чином, 1383 тренувальних зображення та 368 тестових. Було проведено 50 епох навчання із батчами по 8 зображень. Результати наведені у таблиці 3.

Таблиця 2

Результати ШНМ після 80 епох навчання (1476 тренувальних даних, 368 тестових)

	1	2	3	4	5	Середнє
Точність	0.7446	0.731	0.7663	0.7636	0.75	0.7511
Функція втрат	1.9422	1.8938	2.009	1.6949	2.0197	1.91192
Час навчання, с	561	560	539	576	545	556.2

Отформатировано: Шрифт: (по умолчанию) Times New Roman, 12 пт, не полужирный, Цвет шрифта: Авто, Русский (Россия)

Отже, за результатами тестів можна зробити висновок, що нейронна мережа є досить стійкою до зміни кількості тренувальних даних і досить непогано навчається при невеликій кількості епох навчання.

Таблиця 3

Результати ШНМ після 50 епох навчання (1383 тренувальних даних, 461 тестових)

	1	2	3	4	5	Середнє
Точність	0.7137	0.8004	0.7462	0.7527	0.7223	0.74706
Функція втрат	1.583	1.3692	1.5033	1.5289	1.8227	1.56142
Час навчання, с	321	308	316	349	358	330.4

Отформатировано: Шрифт: (по умолчанию) Times New Roman, 12 пт, не полужирный, Цвет шрифта: Авто, Русский (Россия)

Також було проведено тест стійкості розпізнавання класів. Під час тестування було випадковим чином перемішано вхідні дані для тренування та тестування. Було оцінено точність розпізнавання кожного класу після перемішування. Результати тесту наведено у таблиці 4.

За результатам тесту можна зробити висновок, що найбільш чутливим класом до перемішування вхідних даних є клас доріг.

А загальна точність в середньому дорівнює приблизно 75%.

Також було проведено 3 перенавчання із випадково перемішаними даними та підраховано ймовірності помилок першого та другого роду [1] по класам, загальні та середні значення помилок. Результати наведено у таблиці 5.

Результати ШНМ по класам після 5 перемішувань тренувальних даних

		Точність розпізнавання					Середнє
		1	2	3	4	5	
Класи	Вода	0.7241	0.7381	0.7292	0.8723	0.7692	0.76658
	Дорога	0.5556	0.5556	0.6207	0.7442	0.7857	0.65236
	Будівля	0.746	0.8814	0.7049	0.7959	0.7759	0.78082
	Траншея	0.9107	0.7632	0.7609	0.8	0.7778	0.80252
Середня точність		0.7341	0.734575	0.703925	0.8031	0.77715	0.75057

Таблиця 4

Таблиця 5

Ймовірність помилки першого та другого роду розпізнавання класу траншей

Клас	1		2		3		Середня	
	1 роду	2 роду	1 роду	2 роду	1 роду	2 роду	1 роду	2 роду
Вода	0.0377	0.0816	0.0638	0.0588	0.0889	0.0893	0.0635	0.0766
Дорога	0.0943	0.0612	0.0426	0.0784	0.0222	0.1607	0.053	0.1001
Будівля	0.0943	0.0204	0.0212	0.0588	0.0	0.0357	0.0385	0.0383
Загальна	0.2263	0.1632	0.1276	0.196	0.1111	0.2857	0.155	0.215

2.5. Опис інших архітектур та порівняння

У додатку Б та у додатку В наведено топології інших розроблених згорткових мереж. Позначимо архітектуру у додатку А як НМ0, архітектуру у додатку Б — НМ1 і архітектуру у додатку В — НМ2.

HM1 має 3 згорткових блоки, які у результаті утворюють 128 карт ознак розміром 1×1 .

HM2 є модифікацією HM1, але лише з двома згортковими блоками, що утворює 64 карти ознак розміром 8×8 .

Кожна модель навчалась на однаково перемішаних та розподілених даних, та на однаковій кількості епох. Результати наведено у таблиці 6.

Результати порівняння після 60 епох навчання (1383 тренувальних даних, 461 тестових)

	HM0	HM1	HM2
Точність	0.7836	0.7328	0.7456
Функція втрат	1.5831	1.3692	2.009
Час навчання, с	327	345	303

Отформатовано: Шрифт: (по умовчаною) Times New Roman, 12 пт, не полужирный, Цвет шрифта: Авто, Русский (Россия)

Таблиця 6

За результатами порівняння різних архітектур можна зробити висновок, що найбільшу точність розпзнавання має HM0. Хоча HM1 має менше значення функції втрат.

2.6. Опис програмного забезпечення

Програмне забезпечення було реалізовано на мові програмування Python. Розробка і тестування проводилась у середі розробки Jupyter та текстовому редакторі Atom.

Було реалізовано засіб візуалізації роботи згорткової нейронної мережі (рис. 2.13) потрібно створить зчитує мс перегляду нейрон, відображає використання сервер, результатів, маузері. Для ні екрану та стини буде лінгу.

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатовано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Рис. 2.13. Вікно ПЗ для візуалізації роботи нейронної

Для візуалізації процесу навчання та дослідження збіжності функції втрат використовувалась додаткове програмне забезпечення TensorBoard.

Отформатовано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

ВИСНОВКИ

1. Надано історичну довідку про нейронні мережі, їх походження та біологічне обґрунтування.
2. Представлені базові моделі нейронних мереж, їх практичне застосування, недоліки та переваги.
3. Розглянуто основні теоретичні положення, парадигми та методи тренування штучних нейронних мереж.
4. Розглянуто основні принципи та існуючі методи розв'язання задач машинного навчання, розпізнавання образів і комп'ютерного зору.
5. Проведено аналіз доступних існуючих систем.
6. Розроблено модель згорткової нейронної мережі.
7. Детально описано запропонований підхід до вирішення поставленого завдання з використанням розробленої моделі згорткової нейронної мережі.
8. Навчено класифікатор з використанням розробленої моделі.
9. Проведено випробування розробленої системи розпізнавання об'єктів за даними аерофотозйомки з використанням CNN-технології.
10. Розроблено ПЗ для візуалізації роботи розробленої нейронної мережі на реальних даних.
11. За результатами тестування нейронної мережі доведено, що розроблена архітектура має середню точність розпізнавання приблизно 75%.
12. Пораховано ймовірності помилок першого і другого роду, які в середньому становлять 0.155 та 0.215 відповідно.

Отформатировано: Шрифт: 14 пт,
Цвет шрифта: Черный, Русский
(Россия)

ПЕРЕЛІК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Приставка П. О. Аналіз даних. Електронний посібник для студентів спеціальності «прикладна математика» / П. О. Приставка, О. М. Мацуга. — Київ, 2010. — 219 с.

2. Рутковская Д. Нейронные сети, генетические алгоритмы и нечеткие системы // Рутковская Д., Пилинский М., Рутковский Л. — М.: Горячая линия — Телеком, 2006. — 452 с.

3. Уоссермен Ф. Нейрокомпьютерная техника: Теория и практика / Ф. Уоссермен. — М.: Мир, 1992. — 240 с.

4. Форсайт Д. Компьютерное зрение. Современный подход / Д. Форсайт, Ж. Понс. — М.: Издательский дом «Вильямс», 2004. — 928 с.

5. Хайкин С. Нейронные сети: полный курс, 2-е издание / С. Хайкин. — М.: Издательский дом «Вильямс», 2008. — 1103 с.

6. Ясницкий Л. Н. Введение в искусственный интеллект, 2-е издание / Л. Н. Ясницкий. — М.: Издательский центр «Академия», 2008. — 176 с.

7. Яхьева Г.Э. Нечеткие множества и нейронные сети: Учебное пособие / Г.Э. Яхьева. — М.: БИНОМ, 2006. — 316с.

8. Subjectindependent facial expression recognition with robust face detection using a convolutional neural network / Matusugu, Masakazu, Katsuhiko Mori та ін.], 2003. — 559 с.

9. Логика мышления. Часть 3. Персептрон, сверточные сети [Электронный ресурс] — Режим доступа до ресурсу: <https://habr.com/post/214317/>.

10. Пару слов о распознавании образов [Электронный ресурс]. — 2014. — Режим доступа до ресурсу: <https://habr.com/post/208090/>.

11. Что такое свёрточная нейронная сеть [Электронный ресурс]. — 2016. — Режим доступа до ресурсу: <https://habr.com/post/309508/>.

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Отформатировано: Английский (США)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный, Русский (Россия)

Отформатировано

12. Backpropagation In Convolutional Neural Networks [Электронный ресурс]. — 2016. — Режим доступа до ресурсу: <http://www.jefkine.com/general/2016/09/05/backpropagation-in-convolutional-neural-networks/>.

13. LeCun Y. MNIST handwritten digit database [Электронный ресурс] / Y. LeCun, C. Cortes, C. Burges — Режим доступа до ресурсу: <http://yann.lecun.com/exdb/mnist/>.

14. Olah C. Understanding LSTM Networks [Электронный ресурс] / C. Olah. — 2015. — Режим доступа до ресурсу: <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>.

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, Цвет шрифта: Черный

Код поля изменен

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

Отформатировано: Английский (США)

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

Код поля изменен

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

Отформатировано: Английский (США)

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

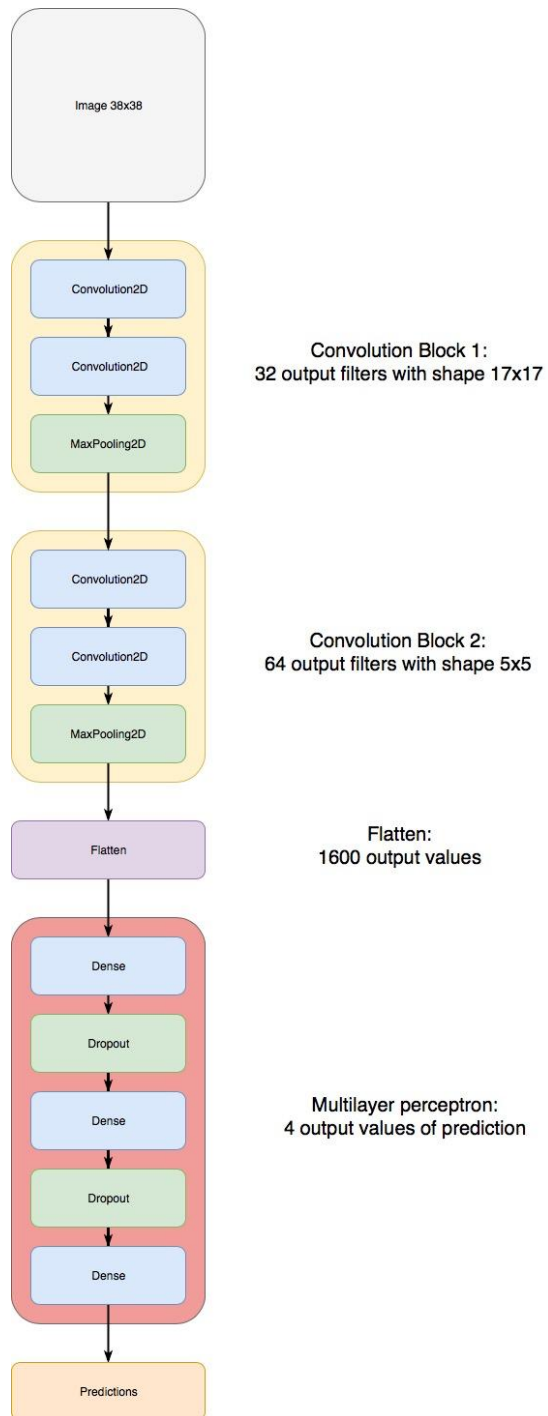
Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

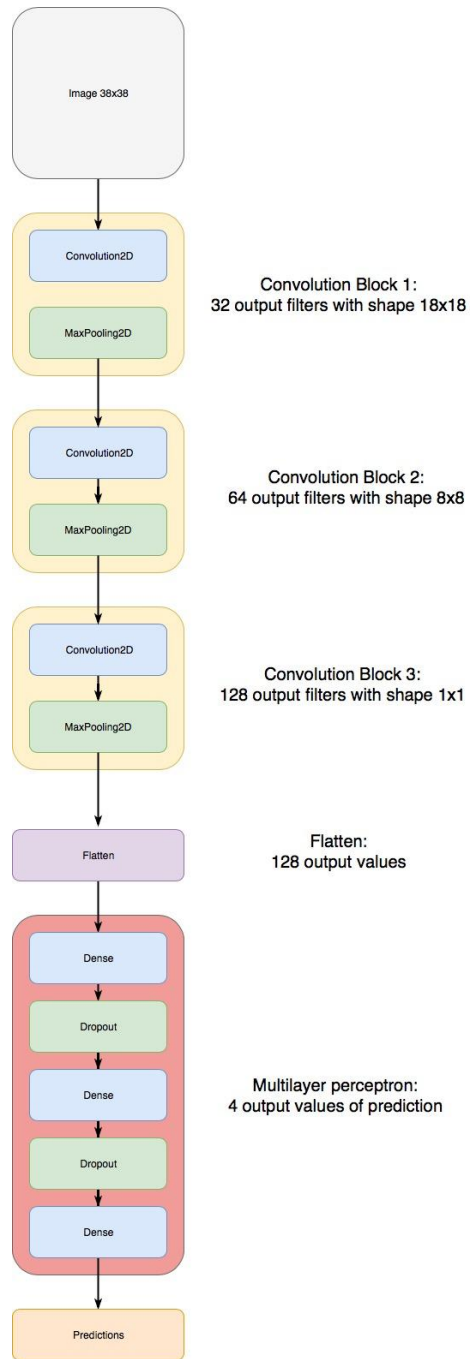
Код поля изменен

Отформатировано: Шрифт: 14 пт, без подчеркивания, Цвет шрифта: Черный

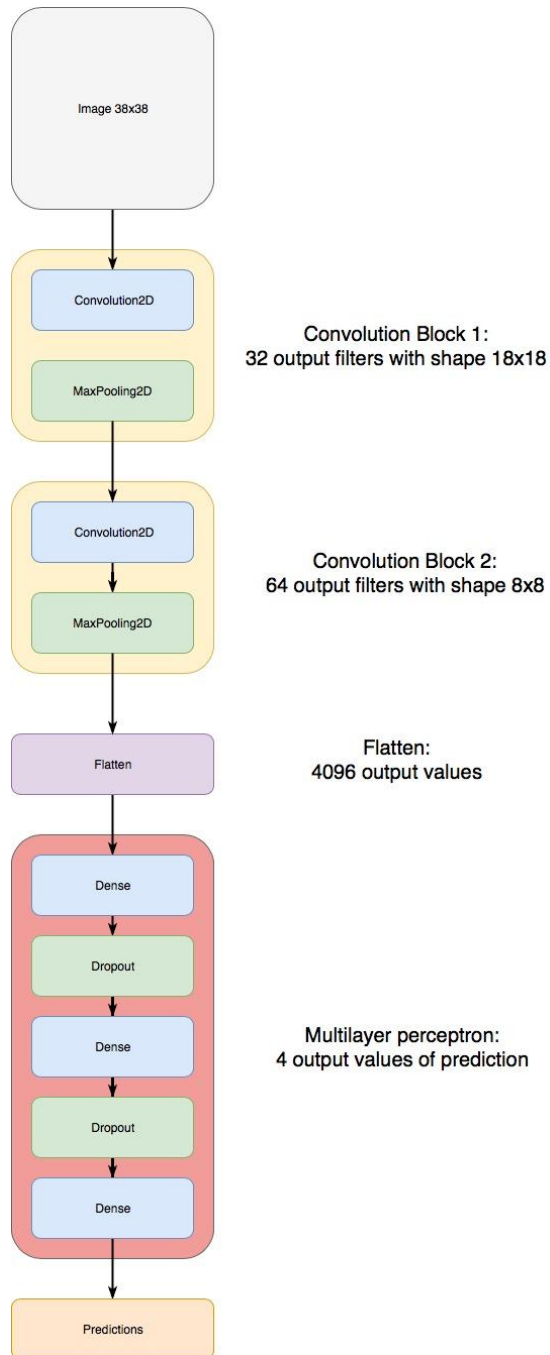
Отформатировано: Английский (США)



ДОДАТОК А



ДОДАТОК Б



ДОДАТОК В

АНОТАЦІЯ

на наукову роботу під шифром _____ РМ2018-2 _____

Актуальність роботи. Ефективність розпізнавання образів в процесі роботизації економіки і виробництва, роботизації рятувальних заходів, роботизації збройних сил і приватних армій, що діють в інтересах держави та інших корпоратократичних об'єднань, стає питанням збереження людських життів.

Метою даної роботи є розробка проекту ефективної багатовступінчастої системи виявлення об'єктів військової техніки і військової інфраструктури, будинків, окопів, тощо. Визначення їх параметрів в умовах недостатньої видимості, шумів різного походження, низької частоти кадрів і невисокою роздільної здатності реєструючого обладнання. Система була розроблена для аналізу кадрів відзнятих із БПЛА.

Мета роботи. Створити CNN-технологію для розпізнавання об'єктів за даними аерофотозйомки.

Методи дослідження. Щоб вирішити поставлену задачу було обрано штучні нейронні мережі, а саме згорткові нейронні мережі. Для побудови нейронної мережі та тестування було використано мову програмування Python, та додаткові модулі такі як бібліотека TensorFlow розроблена компанією Google для пришвидшення алгоритмів машинного навчання.

Структура і обсяг роботи. У першому розділі розглядаються основні види ШНМ, їх складові. Розглядаються основні області практичного застосування моделей штучних нейронних мереж, їх переваги та недоліки. Також розглядаються активаційні функції, алгоритми навчання моделей та оптимізації процесу навчання.

У розділі 2 представлено загальну топологію згорткової нейронної мережі, розповідається про розроблену архітектуру згорткової мережі, наведено приклади вхідних та вихідних даних для кожного шару згорткової нейронної мережі. Також у розділі 2 описано формат вхідних даних та описано створене ПЗ для конвертації зображень у потрібний формат, розповідається про тестування нейронної мережі та наведено таблиці з результатами тестів.

У висновках проаналізовано отримані результати роботи.

Ключові слова. ШТУЧНІ НЕЙРОННІ МЕРЕЖІ, ПЕРСЕПТРОН, ЗГОРТКОВІ НЕЙРОННІ МЕРЕЖІ, ФУНКЦІЯ ВИТРАТ, ГЛИБИННЕ НАВЧАННЯ, МАШИННЕ НАВЧАННЯ, РОЗПІЗНАВАННЯ ОБРАЗІВ, PYTHON, TENSORFLOW.