УДК 004.032.26

Козлов В.В.1, Льовкін В.М.2, Олійник А.О.3

1 студ. гр. КНТ-137 НУ «Запорізька Політехніка»

2 канд. техн. наук, доц. НУ «Запорізька Політехніка»

3 канд. техн. наук, доц. НУ «Запорізька Політехніка»

**ВИКОРИСТАННЯ НЕЙРОМЕРЕЖЕВИХ МОДЕЛЕЙ VGGNET  
В ЗАДАЧІ КЛАСИФІКАЦІЇ ЗОБРАЖЕНЬ**

На даний момент застосування та використання нейронних мереж для вирішення задачі класифікації зображень є одним з передових напрямків розвитку технологій машинного навчання. Для вирішення такої задачі прийнято використовувати підвиди згорткових нейронних мереж.

Згорткова нейронна мережа – це алгоритм глибокого навчання, який може приймати вхідне зображення, призначати важливість різним аспектам або об'єктам на зображенні та здатний диференціювати один об’єкт від іншого за рахунок механізмів, подібних зоровій корі [1].

Головною особливістю мережі є наявність операції згортки [2]. В такій архітектурі кожен мережевий рівень виступає в якості фільтра виявлення на наявність певних ознак або шаблонів, присутніх у вхідних даних.

Згорткова нейронна мережа є багатошаровою мережею без зворотних зв'язків. В загальному випадку її структура складається з таких шарів (рис. 1): вхідний (input), згортковий (convolutional), агрегувальний (pooling), згладжувальний (flatten), повнозв'язний (fully connected), вихідний (ouput).

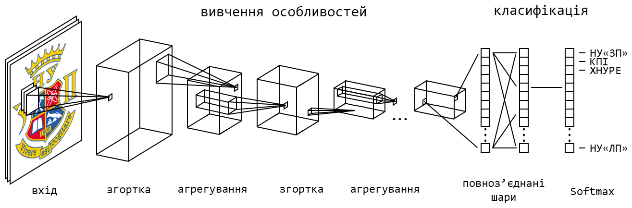


Рисунок 1 – Загальна архітектура згорткової нейронної мережі

В якості функцій активації нейронів застосовується ReLU та Softmax (для вихідного шару). Для навчання найчастіше використовується метод зворотного поширення помилки.

Підвидом згорткової нейронної мережі є сімейство моделей VGGNet, запропонованих вченими Оксфордського університету Кареном Сімоняном та Ендрю Зіссерманом в 2014 році [3].

Усі конфігурації VGG моделей мають загальну архітектуру і розрізняються лише кількістю шарів з вагами (рис. 2).

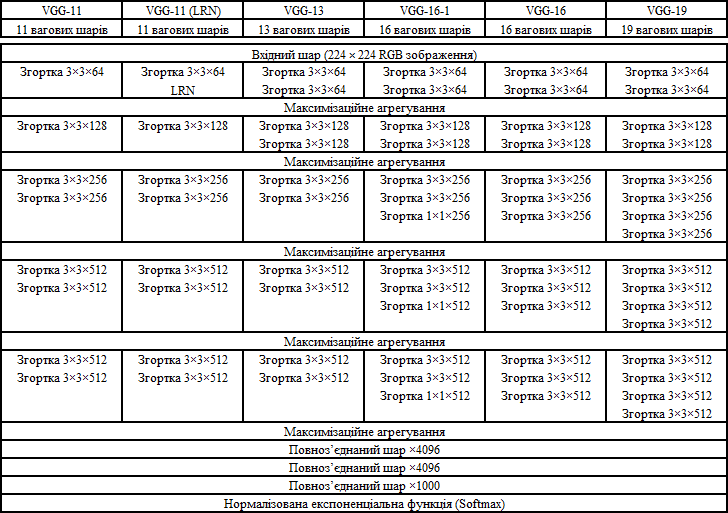


Рисунок 2 – Архітектура моделей VGGNet

В якості вхідних даних модель приймає RGB зображення розміру 224х224 пікселів. Зображення проходить через послідовність згортальних шарів, в яких використовуються фільтри з дуже маленьким рецептивним полем розміру 3х3. Просторова агрегація здійснюється за допомогою п'яти max pooling шарів на вікні розміру 2х2 з кроком 2, які розташовано за одним із згортальних шарів. Після згортальних шарів розміщено три повнозв'язних шари та softmax шар. В якості функцій активації нейронів застосовується ReLU.

Ефективність та перспективність використання VGG архітектури для вирішення задачі класифікації зображень доведено в ILSVRC [4], де моделі сімейства VGGNet досягли одного з найкращих результатів точності розпізнавання, випереджаючи навіть мережу GoogleNet.

**СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ**

1. Saha S. A Comprehensive Guide to Convolutional Neural Networks – the ELI5 way [Electronic resource] / S. Saha. – Access mode: https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

2. Субботін С.О. Нейронні мережі: теорія та практика: навч. посіб. / С.О. Субботін. – Житомир: Вид. О.О. Євенюк, 2020. – 184 с.

3. Very Deep Convolutional Networks [Electronic resource]. – Access mode: https://arxiv.org/abs/1409.1556/

4. ImageNet [Electronic resource]. – Access mode: http://www.image-net.org/challenges/LSVRC/