

Министерство образования Республики Беларусь
Учреждение образования Белорусский государственный университет
информатики и радиоэлектроники

УДК 004.932.2

Попитич
Александр Яковлевич

Модели, методы и программное средство идентификации образов

АВТОРЕФЕРАТ

на соискание степени магистра технических наук
по специальности 1-40 80 05 Математическое и программное обеспечение
вычислительных машин, комплексов и компьютерных сетей

Научный руководитель
Серебряная Лия Валентиновна
к.т.н., доцент

Минск 2017

КРАТКОЕ ВВЕДЕНИЕ

Под идентификацией с точки зрения живого организма можно понимать способность организма обнаруживать в потоке информации определенные закономерности и явления. Идентификация может осуществляться на основе звуковой, тактильной или зрительной информации. В данной работе под идентификацией следует понимать классификацию и последующую локализацию на основе зрительной информации.

В общем случае для решения задачи идентификации применяют две не связанные между собой модели: одна – для классификации объектов на изображении, вторая занимается локализацией классифицируемых объектов на изображении. Это является серьезной проблемой на пути решения задачи идентификации, так как необходимо обучать две модели, что негативно сказывается как на времени обучения, так и на скорости работы модели.

В данной работе предложены решения, которые комбинируют решение двух задач (классификации и идентификации) на основе одной обученной модели. Данная модель представляет собой нейронную сеть.

Вторая проблема, которая присутствует во всех моделях, которые каким-либо способом используют нейронные сети – проблема подбора оптимальных параметров: количества слоев, количества нейронов в каждом из слоев и так далее. Для решения описанной проблемы в работе предложен метод автоматического построения оптимальной архитектуры нейронной сети на основе генетического алгоритма. Полученная модель способна подстраиваться под данные для обучения и под конкретно решаемую задачу, что позволяет в кратчайшие сроки добиться хороших результатов точности классификации и локализации объектов.

Таким образом работа посвящена оптимизации моделей идентификации образов, а также автоматизации процесса создания нейронных сетей.

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Цель и задачи исследования

Целью диссертационной работы является разработка моделей и алгоритмов для решения задач идентификации образов на изображении.

Для достижения поставленной цели необходимо решить следующие задачи:

- анализ существующих моделей и алгоритмов;
- разработка математической и функциональной моделей;
- анализ, выбор и реализация архитектуры нейронной сети;
- анализ и реализация алгоритма обучения выбранной нейронной сети;
- анализ и применение генетического алгоритма для решения задачи автоматического построения нейронной сети;

Объектом исследования является способы классификации и идентификации образов.

Предметом исследования являются модели и методы классификации и идентификации образов для решения задачи идентификации, а также способы оптимизации имеющихся моделей.

Связь работы с приоритетными направлениями научных исследований и запросами реального сектора экономики

Работа выполнялась в соответствии с научно-техническим заданием и планом работ кафедры «Программное обеспечение информационных технологий» по теме «Разработка моделей, методов, алгоритмов, повышающих показатели проектирования, внедрения и эксплуатации программных средств для перспективных платформ обработки информации, решения интеллектуальных задач, работы с большими массивами данных и внедрение в современные обучающие комплексы» (ГБ № 16-2004, № ГР 20163588, научный руководитель НИР – Н. В. Лапицкая).

Личный вклад соискателя

Результаты, приведенные в диссертации, получены соискателем лично.

Вклад научного руководителя Л. В. Серебряной, заключается в формулировке целей и задач исследования.

Апробация результатов диссертации

Основные положения диссертационной работы докладывались и обсуждались на 52-й научно-практической конференции БГУИР “Автоматическое построение архитектуры нейронной сети при помощи генетического алгоритма” (Минск, Беларусь, 2016); 51-й научно-практической конференции БГУИР “Обеспечение устойчивости сверточной нейронной сети при анализе рукописных текстов” (Минск, Беларусь, 2015).

Опубликованность результатов диссертации

По теме диссертации опубликовано 2 печатные работы – статьи для научно-практических конференций БГУИР.

Структура и объем диссертации

Диссертация состоит из введения, общей характеристики работы, пяти глав, заключения, списка использованных источников, списка публикаций автора и приложения. В первой главе представлен анализ предметной области, выявлены основные существующие проблемы в рамках тематики исследования, показаны направления их решения, а также поставлены цели и задачи для работы. Вторая глава посвящена анализу и разработке модели классификации образов на основе сверточной нейронной сети. В третьей главе идет анализ и разработка модели локализации образов на изображении на основе модели для классификации. В четвертой главе предложена практическая реализация описанных двух моделей, приведены и описаны используемые алгоритмы. В пятой главе приведен разбор и анализ результатов разработанных моделей, а также предложены способы их улучшения.

Общий объем работы составляет 63 страницы, из которых основного текста – 57 страниц, 29 рисунков на 6 страницах, 4 таблиц на 1 страницах, список использованных источников из 31 наименования на 2 страницах и 1 приложение на 20 страницах.

ОСНОВНОЕ СОДЕРЖАНИЕ

Во **введении** определена область и указаны основные направления использования разрабатываемого решения. Также показана актуальность темы для решения современных задач в области машинного обучения и компьютерного зрения.

В **первой главе** проведен анализ известных моделей классификации и локализации объектов. Рассматриваются следующие модели классификации:

- логистическая регрессия;
- дерево принятия решений;
- случайный лес;
- нейронные сети.

Рассматриваются следующие модели идентификации образов:

- модели внешнего подобия признаков;
- модели, основанные на выделении отличительных особенностей.

В заключении в первой главе оформляются задачи и цели диссертационного проекта.

Во **второй главе** описывается разработка модели классификации образов на основе преимуществ уже известных для этого моделей. Исходя из рассмотренных аналогов моделей из предыдущего раздела, была разработана модель которая основана на сверточной нейронной сети. Была определена желаемая архитектура сети, которая состоит из следующих слоев:

- сверточный слой;
- слой уменьшения пространства признаков;
- нормализующий слой;
- полносвязный слой.

А также были определены основные свойства подобной сети, которые максимально полно впоследствии будут использоваться для решения двух задач – классификации и локализации образов:

- локальное восприятие признаков;
- наличие разделяемых весов;
- наличие субдискретизации.

Также в данном разделе предложены некоторые оптимизации, основанные на рассмотренных аналогах (сетей GoogLeNet, ResNet):

- использование сверток меньшего размера;
- применение слоев уменьшения размерности;
- использование сверток разного размера для устойчивости к изменению масштаба.

В качестве отдельного большого подраздела в данном разделе является разработка модели, которая способно автоматически строить архитектуру нейронной сети на основе входных данных и глобальных параметров. Для этого были рассмотрены различные генетические алгоритмы и способы применения подобных алгоритмов для решения задачи автоматического построения нейронных

сетей. Для примера на рисунке 1 изображен процесс скрещивания двух родительских нейронных сетей с целью получить две новые сети. Подробнее о применении и результатах использования подобных подходов будет отражено в главе с результатами [1, 2].

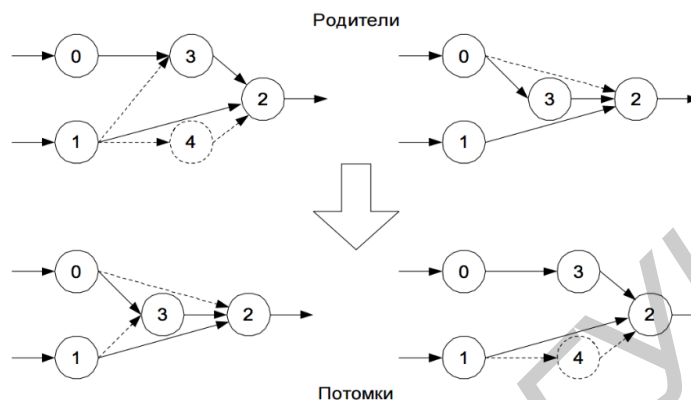


Рисунок 1 – Процесс скрещивания двух нейронных сетей

В **третьей главе** были детально проанализированы достоинства и недостатки описанных в первой главе способов локализации объектов на изображении, а также с учетом достоинств и недостатков разработана модель локализации, которая применяется в данной работе.

Так как одной из задач было использование одной основы – нейронной сети – для классификации и идентификации, поэтому для этого было решено использовать нейронную сеть, разработанную в предыдущей главе. На основе архитектуры нейронной сети вычисляется так называемая активационная карта признаков, которая соответствует классифицируемому объекту. Данная карта признаков вычисляется на уже обученной нейронной сети на основе сверточных слоев, которые входят в состав разработанной модели нейронной сети.

Сама активационная карта вычисляется при помощи следующего выражения:

$$M_c(x, y) = \sum_k w_k^c f_k(x, y) \quad (1)$$

где x, y – координаты точки;

w_k^c – значение веса, которое соответствует классу C для модуля (свертки) нейронной сети k ;

$f_k(x, y)$ – значение активационной функции для свертки k .

В **четвертой главе** детально описаны все алгоритмы, которые используются в выбранных моделях:

- алгоритм автоматического построения архитектуры нейронной сети;
- алгоритм мутации и скрещивания нейронных сетей;
- алгоритм обучения нейронной сети;

– алгоритм построения активационной карты для определенного класса.

Более подробно с этими алгоритмами можно ознакомиться с в полной версии записки к диссертации.

В **пятой главе** дано детальное описание результатов работы разработанных моделей. Модель автоматического построения нейронных сетей была реализована и проверена на начальной популяции (наборе нейронных сетей) в размере 30 сетей на 100 обучающих примерах из базы изображений CIFAR10. Сам процесс эволюционного моделирования и построения оптимальной архитектуры занял около 7-и часов. После завершения работы алгоритма также были получены данные, которые прямым образом характеризуют всех особей в популяции. Полученные данные было решено представить в виде графика на рисунке 2:

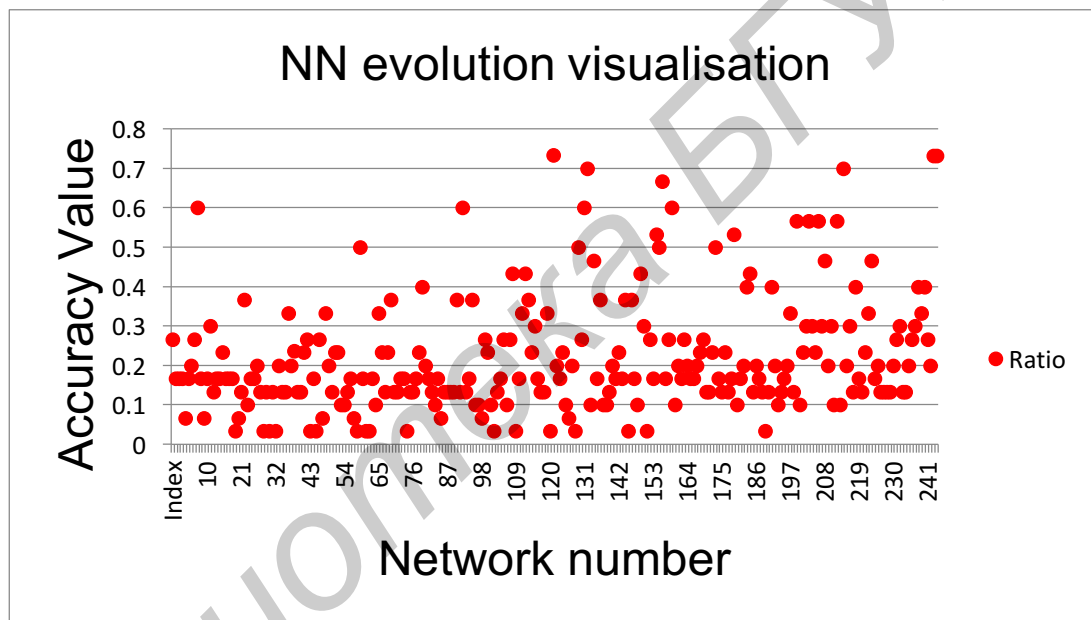


Рисунок 2 – Распределение точности нейронных сетей по номерам эволюционных циклов

По оси абсцисс расположены индексы сетей, на оси ординат – точность распознавания после обучения.

Из графика видно, что с каждым новым поколением точность сетей после обучения возрастает. Появляется все больше сетей, которые обладают большей точностью распознавания. На графике представлено только 30 поколений. При дальнейшей работе алгоритма предсказуемо дальнейшее улучшение метрики точности распознавания и вырождение всей популяции в одну сеть.

В результате работы модели была получена архитектура, которая имеет точность распознавания 73%. Основным достоинством такой модели является

независимость от входных данных и решаемой задачи. Все параметры нейронной сети (количество слоев, веса, количество карт признаков) подбираются путем применения генетических операторов в ходе эволюционного процесса. Таким образом отпадает необходимость подбирать все параметры индивидуально для решаемой задачи и данных обучения, все это выполняется автоматически без участия экспертов предметной области.

Однако стоит отметить и недостатки полученной модели. Основным из них является необходимость большого количества вычислительных ресурсов, ведь каждая особь в популяции представляет собой нейронную сеть, которую необходимо обучить. А с учетом того, что в ходе работы генетического алгоритма особей необходимо обучать не один раз, а множество – время работы алгоритма и количество вычислительных ресурсов для этого возрастает в разы.

Но стоит отметить, что генетические алгоритмы очень легко преобразуются для выполнения распределенных вычислений, что может в разы уменьшить время работы и обучения всех сетей, так как в этом случае каждая нейронная сеть будет обучаться независимо друг от друга на отдельном узле.

Для оценки полученной автоматически сгенерированной нейронной сети было выполнено сравнение с уже известным аналогом, который так же решает задачу классификации – VGG16.

Для того, чтобы сравнение происходило в одинаковых условиях, было принято решение производить обучение двух таких сетей на одной базе изображений CIFAR-10. Для работы с изображениями из этой базы необходимы дополнительные изменения в архитектуре VGG16 – изменить размерность входного слоя на (3, 32, 32).

Сравнительные результаты приведена в таблице 1.

Таблица 1 – Сравнение точности классификации

Архитектура	Значение точности
VGG16	92,3 %
GNN	73,2%

Исходя из результатов, можно сделать вывод о том, что автоматически сгенерированная нейронная сеть уступает реализации VGG16 на 19% по точности распознавания. Поэтому подход автоматического построения сетей имеет место быть и может применяться, например, для реализации прототипов, где на первом месте стоит реализация работающей модели, нежели ее качественные

характеристики. В последующем автоматически сгенерированную модель можно подвергать дополнительным изменениям для улучшения ее определенных характеристик.

Полученная модель может применяться как основа для решения второй поставленной задачи – локализации объектов на изображении. Однако, в целях экономии времени и затрат на вычислительные ресурсы, эта модель будет рассматриваться отдельно от первой.

Вторая разработанная модель – модель локализации объектов, рассматривалась с точки зрения точности локализации классифицируемых объектов.

Для проверки работоспособности модели локализации объектов на изображении было решено использовать базу весов уже обученной нейронной сети VGG16. И уже на базе этой обученной нейронной сети проверить модель вычисления активационной карты на основе сверточных слоев.

Пример вычисления активационной карты для последнего сверточного слоя приведен на рисунке 3.



Рисунок 3 - Активационная карта для последнего сверточного слоя VGG16

Стоит отметить, что активационную карту можно вычислять для любого из сверточных слоев. Ниже приведены активационные карты для всех сверточных слоев.

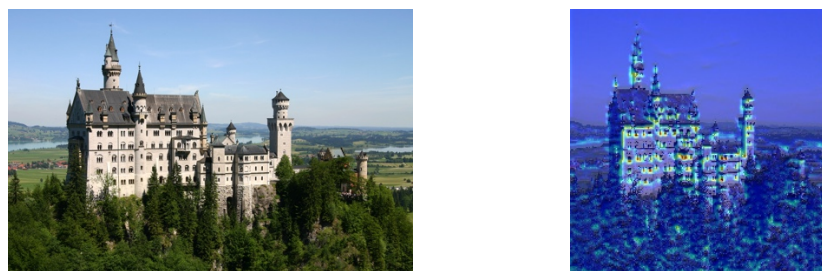


Рисунок 4 - Активационная карта для первого сверточного слоя VGG16

Анализируя активационную карту для первого слоя можно заметить, что при применении сверточных слоев первые слои определяют более специфические для анализируемого объекта признаки. В примере – это окна, крыши, деревья.

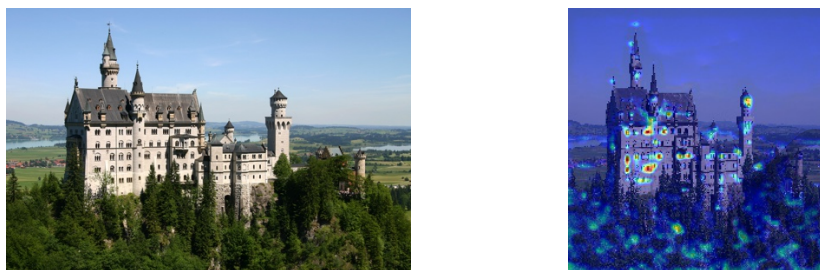


Рисунок 5 - Активационная карта для второго сверточного слоя VGG16



Рисунок 6 - Активационная карта для третьего и четвертого сверточного слоя VGG16

Исходя из приведенных выше примеров активационных карт различных сверточных слоев, можно сделать вывод о том, что наиболее глубокие сверточные слои выявляют наиболее общие признаки для анализируемого объекта. И именно такие признаки и необходимы для успешной локализации объектов.

Также была выполнена сравнительная характеристика разработанной модели и уже имеющихся дескрипторов локализации объектов. Для этого реализованная модель была обучена на CUB200 базе изображений для обучения, которая имеет также и разметку объектов.

Для сравнения необходимо было определить, насколько точно модель локализует объекты на изображении. Для этого каждое изображение разбивалось на 16 одинаковых частей. Далее происходила оценка того, находится ли объект внутри текущего сегмента. Результатом такой оценки было два значения – координаты занятых сегментов и их количество. После этого каждый сегмент из набора исходных сегментов нумеровался, и после этого вычислялось значение как

сумма номеров этих сегментов. В результате представлялось возможным сравнить полученную область с разметкой, которую предоставлял сама база.

Исходя из способа определения точности локализации, который описан выше, результаты приведены в таблице 2.

Таблица 2 – Сравнение точности локализации

Модель	Точность локализации
GNN-GCAM	53,7%
PDP	51,0%
PANDA-PCNN	76,4%

Анализируя результаты, можно сделать вывод о том, что разработанная модель локализации не только экономична с точки зрения использованных ресурсов, но и обладает приемлемой локализирующей способностью.

ЗАКЛЮЧЕНИЕ

1. Предложена модель автоматического построения нейронных сетей для решаемой задачи. Достоинства модели – автоматизация построения оптимальной архитектуры, недостаток – более низкая точность автоматически построенных сетей по сравнению с аналогами, разработанными вручную.

2. Предложена модель классификации объектов на изображении на основе сверточной нейронной сети. Были сделаны некоторые замечки о свойствах сверточных слоев и их обобщающей способности.

3. Предложена модель локализации образов. Основное достоинство – минимальное расходование ресурсов, так как модель основана на той же нейронной сети, что и модель классификации. Недостатки – менее точные результаты по сравнению с некоторыми аналогами.

Рекомендации по практическому использованию результатов

Полученные результаты формируют теоретическую и практическую базу для решения задач в области компьютерного зрения и машинного обучения, которые прямым образом связаны с задачами классификации и локализации объектов.

Разработанные методы могут применяться, к примеру, в системах анализа дорожной обстановки, самоуправляемых машинах, для распознавания бумажной документации и в подобных задачах.

СПИСОК ОПУБЛИКОВАННЫХ РАБОТ

[1] Попитич, А.Я. Автоматическое построение архитектуры нейронной сети при помощи генетического алгоритма / А.Я. Попитич // 52-я научно конференция аспирантов, магистрантов и студентов БГУИР. – 2016, С. – 77-79.

[2] Попитич А.Я. Обеспечение устойчивости сверточной нейронной сети при анализе рукописных текстов / А.Я. Попитич // 51-я научно конференция аспирантов, магистрантов и студентов БГУИР. – 2015. С. – 56-58.

Библиотека БГУИР