



НАУКА  
И ИННОВАЦИИ  
ВОЛОГОДСКОЙ ОБЛАСТИ

Областной конкурс научно-технических проектов  
Вологодской области «Потенциал будущего»  
Номинация «Ученик»

# Создание сверточной нейронной сети, способной распознавать математические символы

Работу выполнила ученица 11б класса МОУ  
«Гимназии №2»

Художилова Анна

2020 г.

# Актуальность

Искусственные нейронные сети прочно вошли в нашу жизнь и в настоящее время активно применяются там, где обычные алгоритмические решения оказываются неэффективными или вовсе невозможными.

Данный исследовательский проект направлен прежде всего на повышение качества образования. Целевой аудиторией проекта являются ученики и учащиеся.

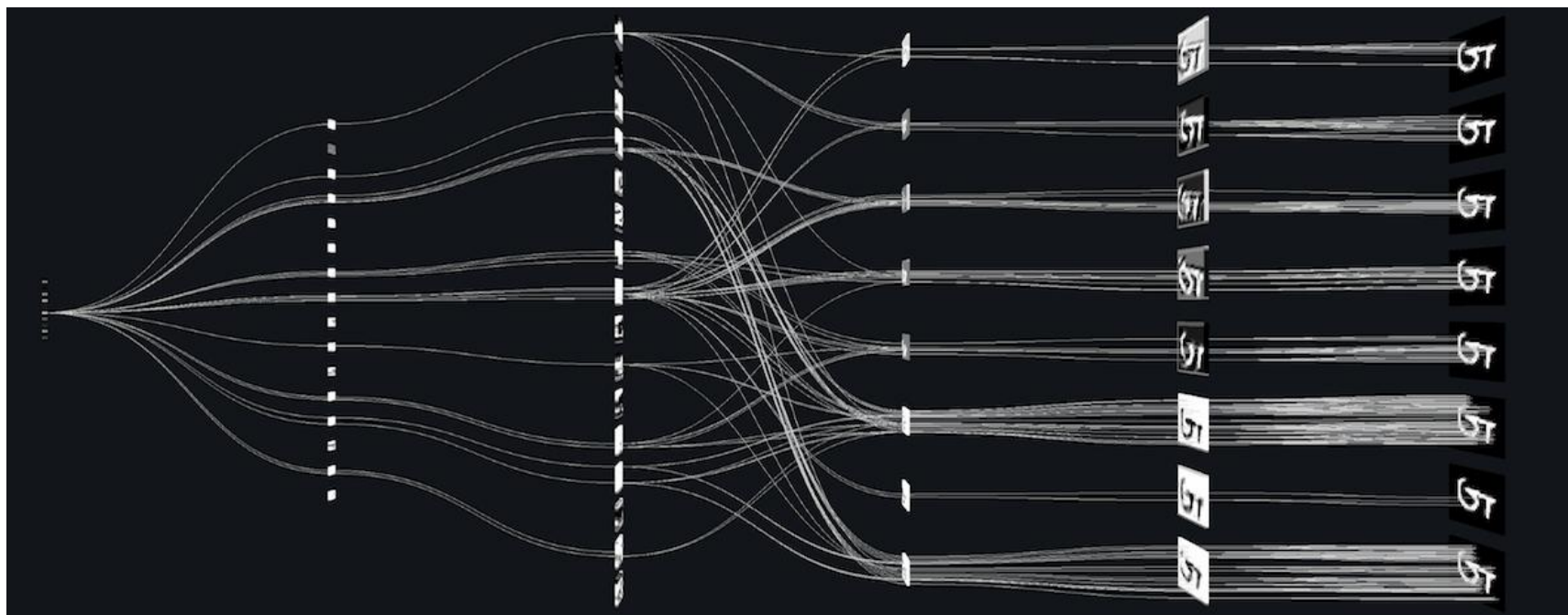
Актуальность заключается в том, что учителя часто задерживаются на работе до позднего вечера, выполняя проверку домашнего задания и контрольных работ. Каждый учитель сталкивается с неразборчивым почерком, но в гуманитарных предметах, непонятное слово можно понять из контекста, а в точных науках тот или иной символ может обозначать всё что угодно. От лица учеников могу сказать, что мы тоже сталкиваемся с непониманием написанного на доске. Поэтому в распознавании символов, точность распознавания – самая важная характеристика.

# Цель: Создание сверточной нейронной сети, способной распознавать математические символы.

## Задачи:

- 1. Теоретическая часть:
  - 1.1. Изучить глубокое обучение.
  - 1.2. Изучить архитектуру сверточных нейронных сетей.
  - 1.3. Изучить основы программирования, необходимые для создания сверточной нейронной сети.
- 2. Практическая часть:
  - 2.1. Подготовить тренировочные данные для проверки нейронной сети.
  - 2.2. Провести инициализацию, обучение и тренировку нейронной сети.
  - 2.3. Провести тестирование нейронной сети.
  - 2.4. Обучить нейронную сеть.
  - 2.5. Разработать программу для распознавания изображений с камеры.

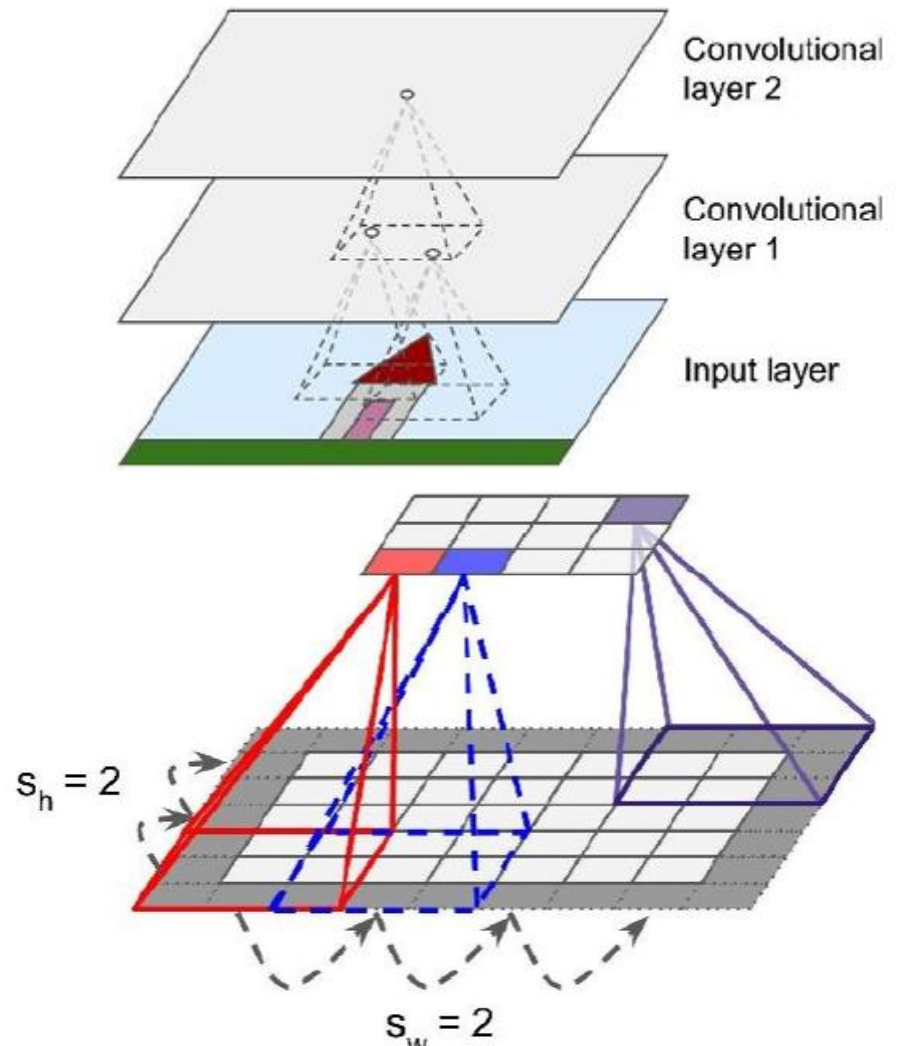
# Архитектура сверточных нейронных сетей



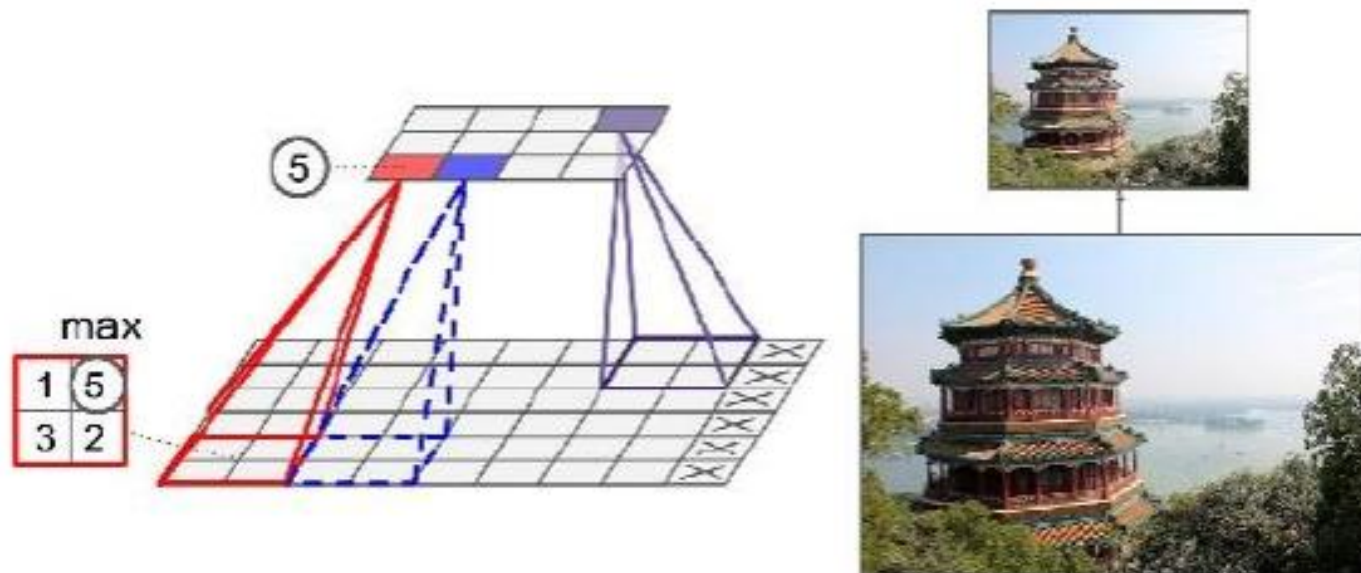
Изображение пропускается через серию свёрточных, нелинейных слоев, слоев объединения и полносвязных слоёв, и генерируется вывод. Выводом может быть класс или вероятность классов, которые лучше всего описывают изображение.

# Сверточный слой

Нейроны в сверточном слое связаны с выходами ограниченного числа нейронов из предыдущего слоя, которые расположены внутри небольшого прямоугольного рецепторного поля.

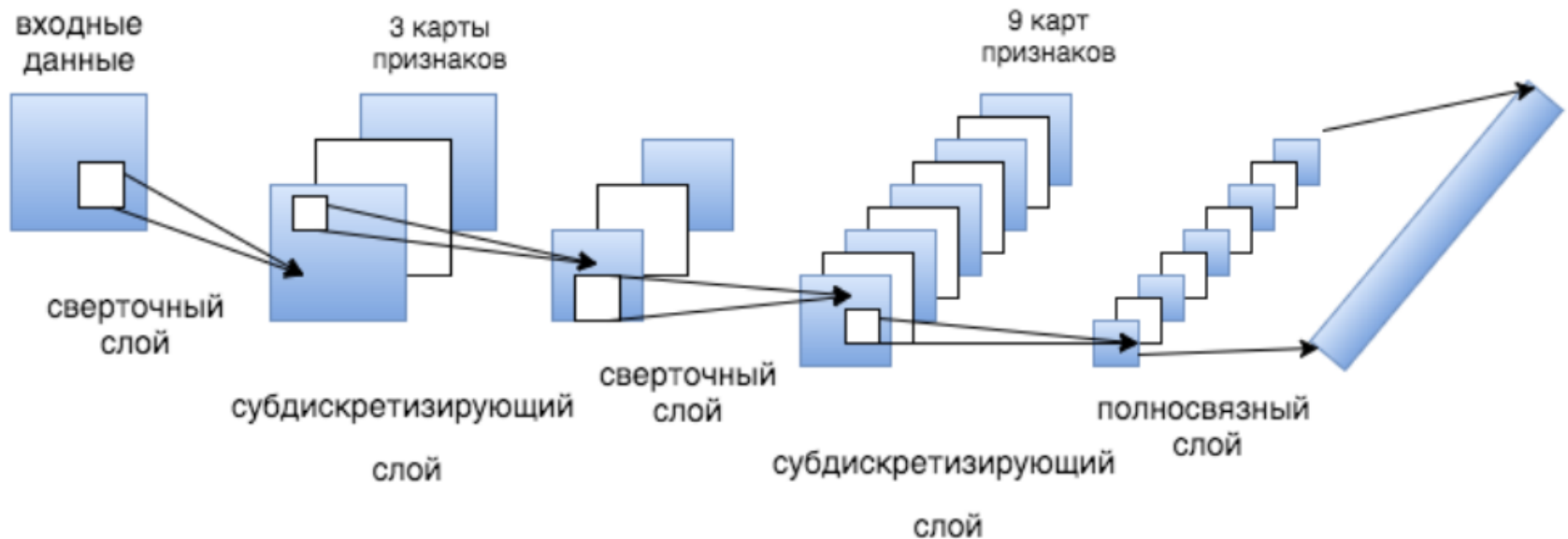


# Объединяющий слой



Цель данной группы слоев состоит в том, чтобы проредить входное изображение для сокращения нагрузки, расхода памяти и кол-ва параметров. Нейроны связаны также, как и в сверточном слое, но нейроны объединяющего слоя не имеют весов.

# Полносвязный слой



Этот слой берёт вводные данные и выводит  $N$ -пространственный вектор, где  $N$  — число классов, из которых программа выбирает нужный.

Способ, с помощью которого работает полносвязный слой — это обращение к выходу предыдущего слоя и определение свойств, которые больше связаны с определенным классом.

# Подготовка тренировочных данных для проверки нейронной сети

Для обучения сети использован набор NASTYv2, который содержит 168000 изображений 369 различных рукописных математических символов. Каждое изображение имеет размер 30x30 пикселей в формате RGB. Толщина линии составляет приблизительно 1/12 от высоты символа.

Для построения нейронной сети использовалась нейронная сеть со следующей архитектурой из 8 слоев:

Первый сверточный слой, который имеет размеры ядра 3x3 пикселя;

Первый объединяющий слой с размером окна 2x2;

Второй сверточный слой с размерами ядра 3x3;

Второй объединяющий слой с размерами окна 2x2;

Выравнивающий слой;

Первый полносвязный слой;

Исключающий слой («dropout»);

Второй полносвязный слой с 369 выходами;

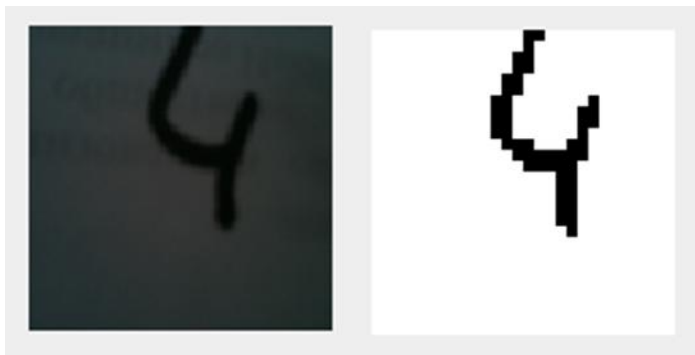


# Тестирование сети

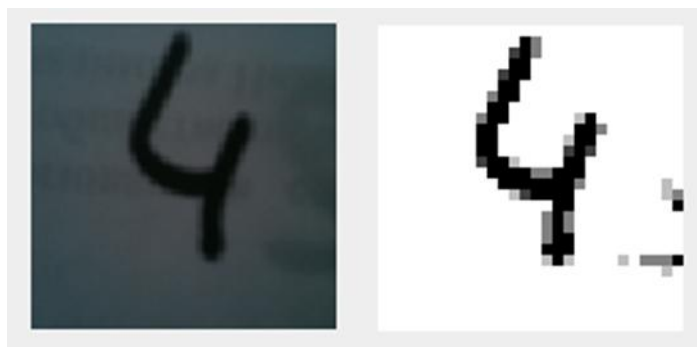
Тестирование проводилось с изображениями, полученными с камеры ноутбука. Так как изображения, полученные с камеры не соответствуют образцам из набора данных, на которых обучалась нейронная сеть, то прежде, чем подавать его на вход сети, необходимо провести обработку изображения – бинаризация.

Для решения данной задачи изначально был использован алгоритм Бредли, который работает на основе интегрального изображения (сумма яркости ячейки основывается на сумме яркостей предыдущих ячеек). Если яркость превышает пороговое значение, то в ней устанавливается максимальное значение яркости. Сложность состояла в том, чтобы точно подобрать пороговое значение яркости, откуда и были выявлены следующие недостатки работы алгоритма: часть изображения стиралась и становилась фоном или наоборот, фон не был вычищен до конца. Поэтому было принято решение использовать для задачи очистки изображения упрощенный алгоритм, работа которого основывается на работе алгоритма Вульфа, где пороговое значение адаптируется к изображению.

# Алгоритм Бредли



Черно-белое изображение. Нет градаций серого



Не вычищен фон.



Фон удалён вместе с фрагментами цифр.

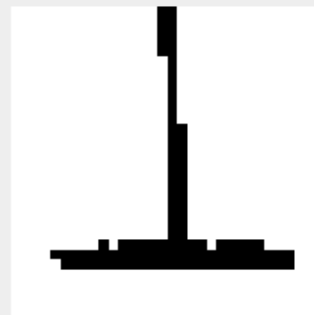
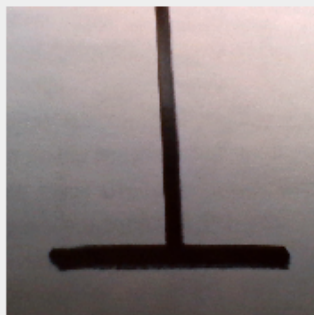
# Алгоритм Вульфа



Распознавание математических символов



Селфи!



Проверить

\perp

# Разработка программ для распознавания изображений с камеры



$\pm$	<code>\pm</code>	$\cap$	<code>\cap</code>
$\mp$	<code>\mp</code>	$\cup$	<code>\cup</code>
$\times$	<code>\times</code>	$\oplus$	<code>\oplus</code>
$\div$	<code>\div</code>	$\sqcap$	<code>\sqcap</code>
$*$	<code>\ast</code>	$\sqcup$	<code>\sqcup</code>
$\star$	<code>\star</code>	$\vee$	<code>\vee</code>
$\circ$	<code>\circ</code>	$\wedge$	<code>\wedge</code>
$\bullet$	<code>\bullet</code>	$\setminus$	<code>\setminus</code>
$\cdot$	<code>\cdot</code>	$\wr$	<code>\wr</code>
$+$	<code>+</code>	$-$	<code>-</code>

Обученная нейронная сеть сохранялась в файл и использовалась в специально написанной демонстрационной программе. Программа получает изображение с камеры ноутбука, бинаризует его, приводит к размеру изображения обучающей выборки и передает его на вход нейронной сети. Нейронная сеть выдает результат в виде наибольшей вероятности соответствия одному из 369 классов. Программа выводит символ на форматированном языке LaTeX.

# Практическая значимость

В результате проделанной работы была создана сверточная нейронная сеть, способная распознавать 369 различных рукописных математических символов. Главной задачей проекта являлась бинаризация изображения, которая и стала частью исследования. В ходе которого было установлено, что алгоритм Вульфа справляется с очисткой изображения лучше, чем алгоритм Бредли, благодаря чему удалось значительно повысить качество распознаваемых изображений.

Таким образом, благодаря созданной нейронной сети, распознающей математические символы, являющиеся частями формул, уравнений в текстовых документах, учащиеся и преподаватели смогут упростить перевод рукописных записей в цифровую форму.

# Контактная информация

Телефон: 8-921-830-10-94, Художилова Анна Вадимовна

Эл. почта: [anna250303@outlook.com](mailto:anna250303@outlook.com)