

Применение методов распознавания в практикумах по программированию

Никита Бесшапошников, Михаил Дьяченко, Анатолий Кушниренко,
Михаил Кузьменко, Александр Леонов, Максим Матюшин, Кирилл
Прокин

Москва

Московский государственный университет,

ФГУ ФНЦ НИИСИ РАН,

Московский педагогический государственный университет,

АО “АйТи Сфера”

Построение базового курса программирования для детского сада и начальной школы – трудная задача

- Нужны усилия для облегчения и ускорения процесса усвоения материала на каждом этапе, по каждой теме
- Важнейший этап – первый десяток занятий
- На этом этапе бесценной находкой является работа с программой, как с материальным объектом, составленным из ПиктоКвадратиков или ПиктоКубиков

Ограничения при обучении в системе ПиктоМир

- Обучение рекомендуется начинать в возрасте 6 лет, что соответствует старшей группе дошкольных образовательных организаций
- Ограничения СанПиН на время работы с планшетными компьютерами
- Эргономические ограничения опорно-двигательного аппарата малышей

Как обойти ограничения

- Сократить часы непосредственного общения ребенка с планшетом
- Предлагается выкладывать программы управления виртуальным роботом из магнитных квадратиков (или деревянных кубиков), на которых напечатаны пиктограммы используемые для представления команда на экране ПиктоМира
- Дет с удовольствием собирают картинки из кубиков на столе даже в возрасте трех-четырёх лет
- Это дополнительно развивает еще и мелкую моторику
- Остается вопрос: как исполнить и проверить составленную программу на компьютере?
- Ответ: ее нужно «показать» компьютеру

Из реального в виртуальное

- Фотографируем мобильным устройством программу, составленную из кубиков на столе
- Распознаем пиктограммы и их расположение по фотографии
- Предоставляем возможность вставить полученный фрагмент в шаблон программы в ПиктоМире

Все сводится к вопросу эффективного распознавания картинок-пиктограмм на картинках одинакового размера, разложенных на плоскости при ручной съемке в 3-х мерном пространстве при различных сопровождающих искажениях.

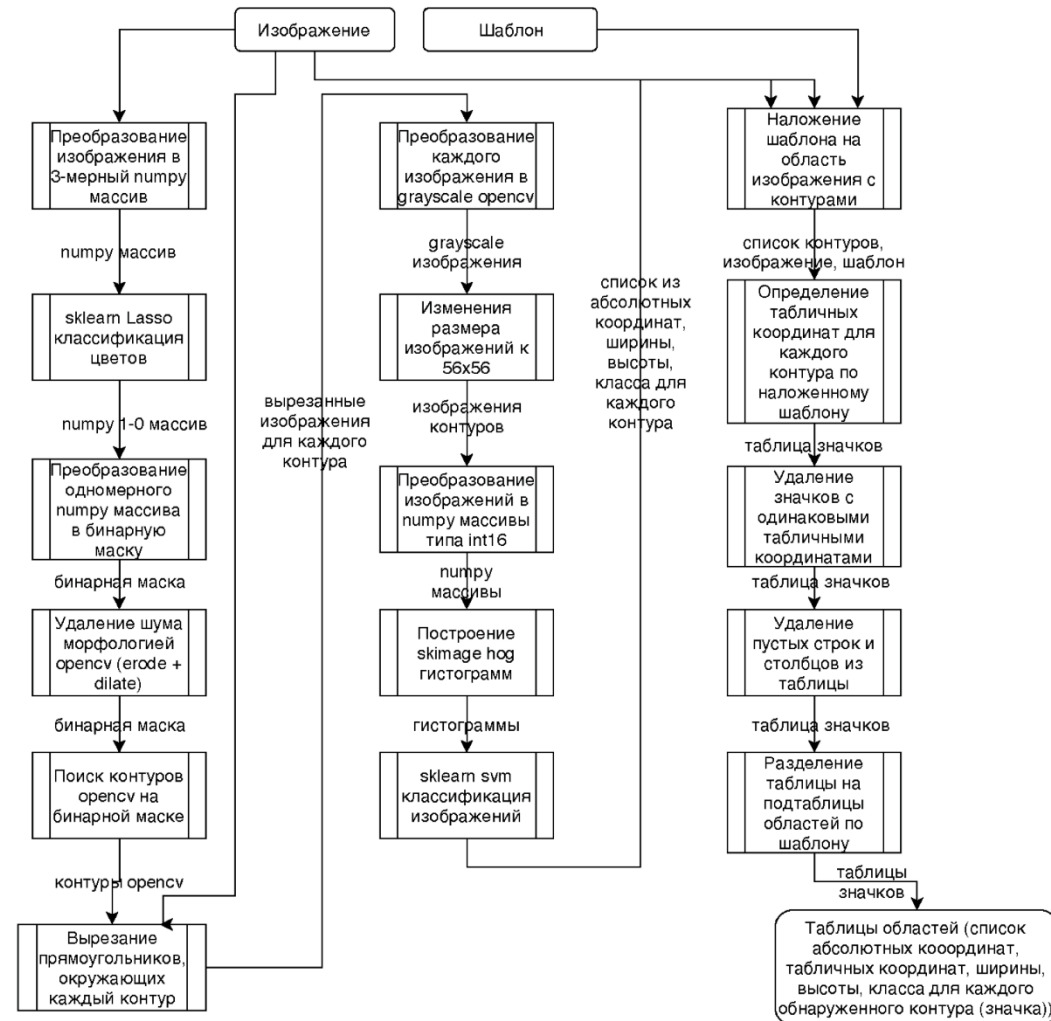
Задача распознавания программы

- Нахождение картинок-пиктограмм на фотографии
- Нахождение комбинаторной структуры расположения распознанных картинок-пиктограмм (в данной случае двумерной таблицы)
- Совмещение найденной структуры и шаблона в ПиктоМире
- Возможные варианты решения как с помощью эвристических методов, так и с помощью использования нейронных сетей
- Решение опираются на результаты исследования по гранту РФФИ 18-07-00901 “Исследование и разработка системы распознавания элементов рукотворного интерьера на базе нейронных сетей для построения дополненной реальности и выработка алгоритмов взаимодействия управляемых объектов с реально-виртуальным окружением”

Нахождение программы на фотографии с помощью OpenCV

- Находятся границы пиктограмм по цвету (алгоритм сегментации по водоразделам, алгоритмы выделения цветов)
- В рамках границ проводится классификация (методы knn с подбором параметров кроссвалидацией, и линейная модель lasso, оба метода реализованы в библиотеке sklearn)
- Наложение шаблона программы на область изображения с контурами
- Определение табличных координат

Схема нахождения программы на фотографии с помощью OpenCV



Недостатки решения на OpenCV

- Сильная привязка алгоритма к цвету пиктограмм, что делает его чувствительным к смене освещения
- Неустойчивость работы при перспективных искажениях
- Неустойчивость работы при наличии шумов или размытия
- Было принято решение использовать нейронные сети

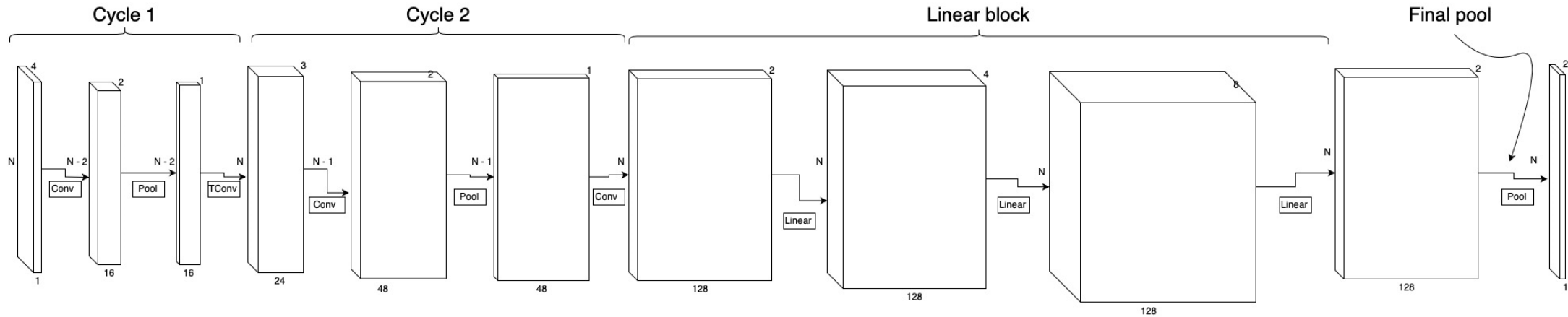
Распознавание пиктограмм на фотографии с помощью нейросетей

- Был отснят существенный набор данных для обучения нейронной сети: несколько видео в разных обстановках и уровнях освещенности, разбитое на кадры.
- Было обучено несколько стандартных моделей для нахождения объектов на изображении с помощью библиотеки Tensorflow
 - SSD Mobilenet
 - Faster RCNN Resnet
 - Faster RCNN Inception
 - Yolo with Darknet
- Лучший результат по соотношению скорость работы-точность показала модель Faster RCNN Inception (max proposals 50)

Нахождение табличной структуры

- Так как эвристические методы показали себя довольно неустойчивыми к геометрическим искажениям, была разработана модель нейронной сети для определения табличной структуры
- Данная модель по размерам таблицы (что легко можно найти эвристическим методом) и координатам пиктограмм на изображении находит относительные координаты в искомой таблице
- Данная модель названа нейротабулятором

Нейротабулятор



- Модель обучается на специально сгенерированных данных с различными искажениями
- На валидационной выборке набирает точность 0.9998

Варианты работы нейросетевых решений

- Возможны два варианта работы:
 - Серверное приложение
 - Работа на мобильном устройстве
- При работе на сервере, особых ограничений нет, но при отсутствии соединения с Интернетом, использование распознавания в системе ПиктоМир становится невозможным
- Для работы на мобильных устройствах были обучены более «легковесные» нейронные сети (SSD Lite для Android и CoreML модель для iOS), а модель нейротабулятора сконвертирована в код на C++
- Это позволяет довольно эффективно выполнять задачу распознавания на мобильном устройстве
- На обучение нейронных сетей было потрачено порядка 100 часов машинного времени мощного сервера с графическими ускорителями, установленного в НИИСИ РАН, и несколько десятков часов работы оператора по подготовке данных

Спасибо за внимание!

Никита Бесшапошников, Михаил Дьяченко, Анатолий Кушниренко, Михаил Кузьменко, Александр Леонов, Максим Матюшин, Кирилл Прокин

Москва

Московский государственный университет,

ФГУ ФНЦ НИИСИ РАН,

Московский педагогический государственный университет,

АО “АйТи Сфера”

Работы выполнены в рамках темы «Разработка, реализация и внедрение семейства интегрированных многоязыковых сред программирования с автоматизированной проверкой заданий для учащихся образовательных организаций, ДОО, младшей, основной и старшей школы и студентов педагогических университетов» (тема госзадания РАН 0065-2019-0010).