


Конкурс алгоритмов классификации изображений

Дано:

- Коллекция картинок размеров 28x28 пикселей.
- На каждой картинке изображена одна рукописная цифра. Например: 
- Часть картинок распределены по папкам с именами, соответствующими изображенной на картинке цифре. Например в папке с именем «0» хранятся изображения нуля. Будем называть все эти изображения *классифицированными* или *тренировочными*. Они хранятся в папке `training`.
- Другая часть изображений не классифицирована. Будем называть эти изображения *тестовыми*. Они хранятся в папке `testing`.

Что со всем этим делать?

- Создать алгоритм классификации изображений.
- Оттестировать и настроить алгоритм на тренировочных изображениях.
- Применить алгоритм к тестовым картинкам.
- Сохранить результат работы алгоритма с тестовыми картинками в файл (см. *Формат файла с ответами* ниже).
- Отправить алгоритм (исходный код и, при необходимости, исполняемый файл) и файл с результатами организаторам.
- Содержимое файла должно совпадать с результатом работы отправленного алгоритма.

Как будет оцениваться работа алгоритма?

- Организаторам известны все изображённые на не классифицированных картинках цифры.
- Будет вычислена доля правильно классифицированных изображений. Победит тот алгоритм, который даст больше правильных ответов.
- Если два алгоритма дадут одинаковый результат, то они дополнительно будут проверены на секретном наборе картинок.
- Если же и в этом случае результат будет одинаковым, то будет дана экспертная оценка самому алгоритмам.

Требования к алгоритму

- Получение результата за разумное время на современных ПК (минуты, не более 15 минут).
- Желательна индикация прогресса работы программы.
- Расход оперативной памяти не более 4 Гб.
-
- Язык программирования: Python 3.7, Pascal (PascalABC), Java, C++, C#.
- Использование сторонних библиотек допускается только для вспомогательных действий: загрузка изображений, сохранение данных в файл, простые операции с векторами или матрицами (умножение, вычитание, вычисление длины и т. п.).
- Используемые библиотеки должны входить в стандартную библиотеку языка или устанавливаться через менеджер пакетов Pip (для Python)

- Кодирование алгоритма связывающего исходя из результатов ручной классификации.
- Комментарии к коду желательны.

Формат файла с ответами. Пример

#имя файла, цифра

128-29527-31746.png, 0

121-29903-14590.png, 5

14124-2731-150.png, 9

Файлы иного формата не принимаются.

FAQ

- Да, дипломы и сертификаты будут. Призы возможно тоже.
- Нет, мы не сами рисовали все эти цифры.
- Да, все картинки одинакового размера.
- Да, все цифры помещаются на картинки целиком.
- Нет, в именах файлов не закодировано их содержимое.
- Да, больше ~10% правильно угаданных цифр на картинках не такой уж и плохой результат :)
- Да, мы можем *немного* исказить секретный набор картинок.

UPD:

Дедлайн перенесён

Алгоритмы и результаты принимаются до
31 октября 2019 года, 6:00 (по Читинскому времени)