**Задача 1**

В качестве тестового задания мы предлагаем Вам портировать участок кода (алгоритм обработки данных) с языка C# в язык C++. Данный алгоритм выполняет обработку результата работы нейронной сети по поиску фрагментов текста на изображении. Фрагменты — это обычно строки или части строк.

Особенностью работы нейронной сети является то, что она работает (предсказывает) неточно, часто избыточно, и находит гораздо больше фрагментов, чем нам необходимо. На это могут влиять поворот, шумы, искажения на изображении, либо попытка обработать изображение, сильно отличающееся от тех, на которых сеть училась.

Чтоб получить результат похожий на реальные строки (координаты описывающих прямоугольников - bounding box), разработан алгоритм пост-обработки результатов сети. Задача этого алгоритма отфильтровать результаты с помощью специально математической модели и набора пороговых значений и статистического анализа, а также объединять в некоторых случаях небольшие фрагменты текста в одну строку. Предоставленный заданием код решает эту задачу на языке C#.

1. Портировать математическую модель (алгоритм) с C# в C++.
2. Написать UnitTest для проверки алгоритма. Тест должен принимать входные данные из предоставленного с заданием текстового файла и формировать на выходе аналогичный файл с результатами в формате
3. Реализовать метод отрисовки результатов (bounding box строк) поверх изображения, предоставленного с заданием. Вы можете использовать этот метод для самопроверки - при корректной работе алгоритма будут корректно выделены строки текста или фрагменты строк.

Предоставленные данные:

1. Исходный код алгоритма на C#. Краткое описание точки входа и параметров.
2. Текстовый файл с набором входных данных для тестирования алгоритма.
3. Тестовое изображение для отрисовки результатов и самопроверки.

Требуемые результаты:

1. Необходимо собрать проект C++ для среды разработки Visual Studio с реализованным алгоритмом.
2. В проекте должен быть UnitTest для проверки работы алгоритма на наборе входных данных (из текстового файла)
3. Результат теста (выходные данные) в текстовом файле.
4. Изображение с прорисованными прямоугольниками.

Дополнительно можно:

Предложить варианты по оптимизации производительности и параллельному выполнению кода алгоритма.

**Задача 2**

Одной из задач OCR является преобразование найденных на изображении фрагментов текста в текстовую строку (собственно, само распознавание текста). Предлагаем вам

Разработать библиотеку распознавания изображений строк на китайском языке с использованием нейронной сети. Задание предусматривает использование готовой нейронной сети и фрагментов кода для преобразования входных данных (изображения) и обработки результата (декодирования в текст).

Необходимо:

1. Написать на С++ библиотеку (.dll \ .so), которая загружает изображение, инициализирует объекты ONNX runtime library, запускает методы распознавания загруженного изображения, получает и интерпретирует результат.
2. Написать консольное приложение, которое показывает пример использования этой библиотеки.

Предоставленные данные:

1. Нейронная сеть для распознавания символов китайского языка (приложение Neutron поможет ознакомиться с ней) в формате ONNX.
2. Алфавит для работы с данной нейронной сетью. Для декодирования результата в текст.
3. Набор констант, а также фрагменты кода, которые позволяют инициализировать некоторые объекты и корректно интерпретировать (декодировать) результат работы сети.
4. Изображение и примеры входных и выходных данных

Требуемые результаты:

1. Создать проект, подключить библиотеку ONNX Runtime, и используя данные фрагменты кода реализовать процесс распознавания изображения (загрузку изображения, использование нейронной сети, декодирование результата от сети).
2. Собрать проект C++ для среды разработки Visual Studio с реализованным алгоритмом в виде динамической библиотеки.
3. Собрать проект C++ для среды разработки Visual Studio в виде консольного приложения. Подключить собранную библиотеку (из п.2) и запустить процесс распознавания тестового изображения.
4. Получить результат в качестве строки распознанного текста   
   «今天来拜访我们，我们将»

**Дополнительное задание:**

Разработать алгоритм нарезки изображения на строки. Улучшить качество изображения, а именно - убрать шум и увеличить контрастность.   
Рекомендуется использовать библиотеку DLib.

Предоставленные данные:

Изображение

Требуемые результаты:

Приложение (код), которое выполняет поставленную задачу и сохраняет на диск массив изображений (строк) и исходное изображение улучшенного качества.