



ЗАДАЧА

Разработка
модуля
распознавания
сканированных
документов



Автор:
Мосгортранс





СИТУАЦИЯ

Обеспечение функционирования Московского Транспортного Комплекса требует обработки большого объема документации. От ускорения и автоматизации этого процесса напрямую зависит стоимость проезда для пассажиров. В связи с этим сформулирована задача.



Сейчас работа по обработке документов ведется специалистами в ручном режиме.

В ручном процессе большое количество ошибок из-за человеческого фактора, процесс выполняется медленно, а также требуется привлечение дополнительных человеческих ресурсов.

ПРОБЛЕМА



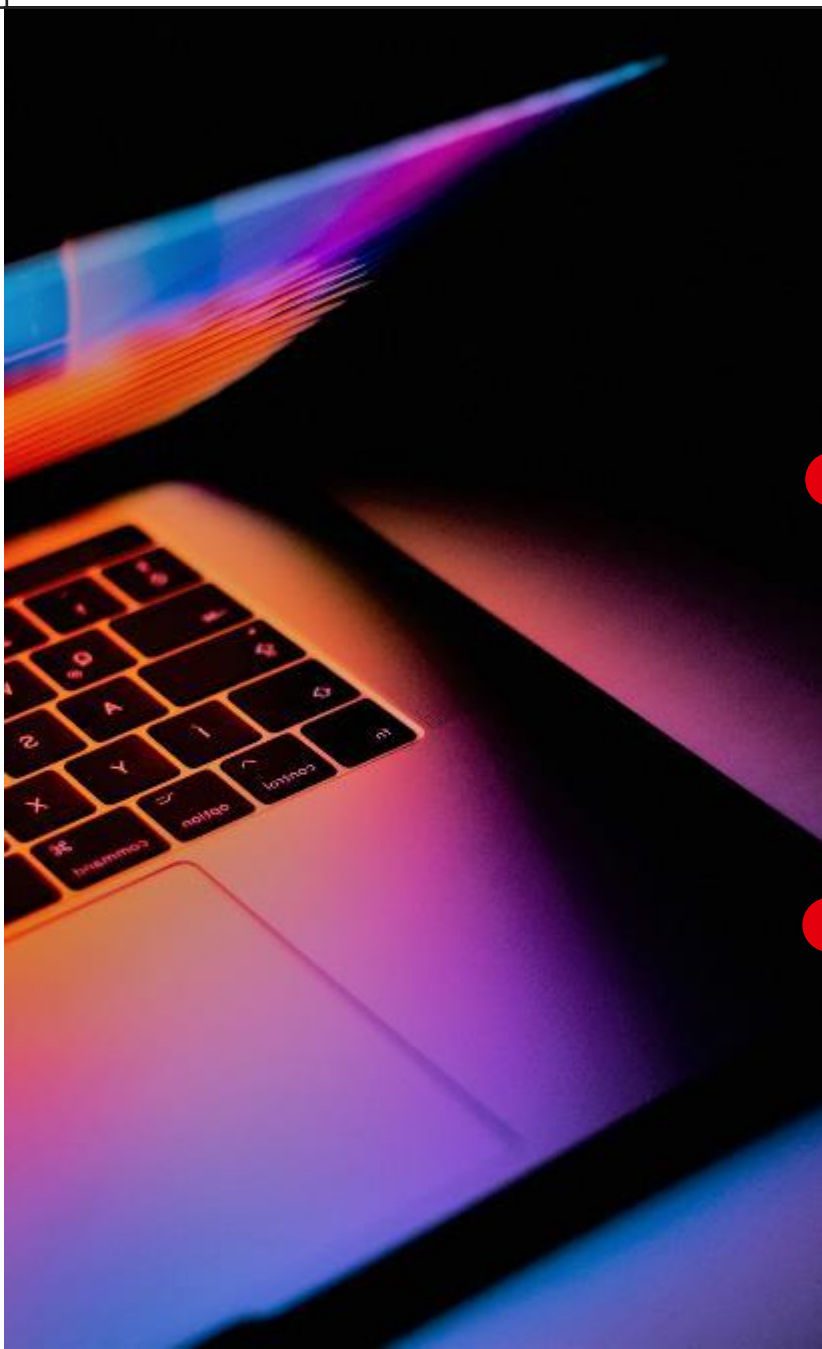
Направление поиска решения

Необходимо заменить человека, осуществляющего ввод информации из сканированных документов на программного робота, умеющего определять тип сканированного документа и передавать необходимую информацию в учетную систему предприятия.

Для этого нужно разработать алгоритм и приложение, которое сможет определить тип документа по представленному скану документа.



ДААННЫЕ ДЛЯ РЕШЕНИЯ



[Датасет с перечнем разных типов сканированных документов](#)

[Программа для разметки документа по блокам](#)