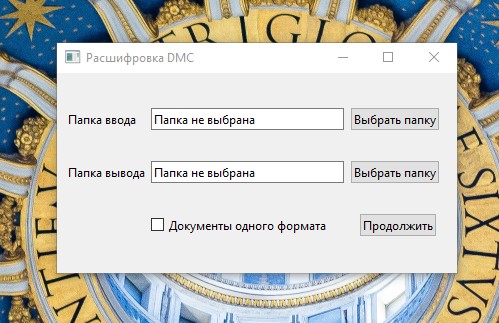
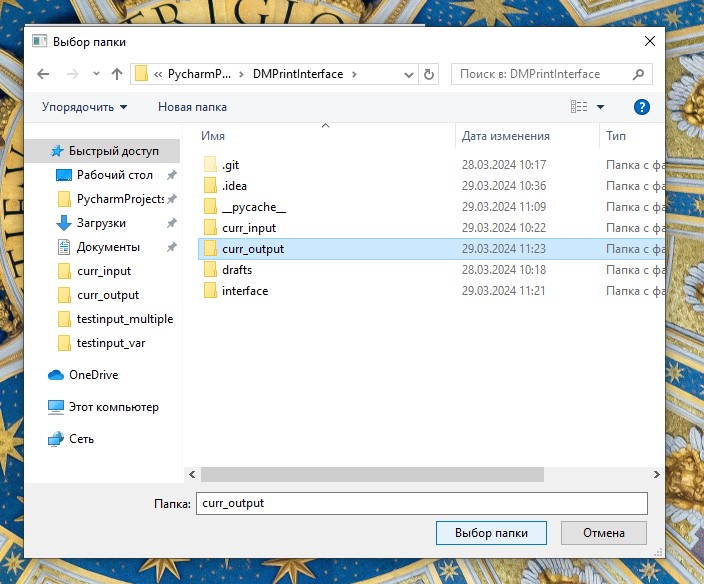
Перед вами первая версия декодера DMC из PDF с картинками в CSV с кодами. К сожалению, приходится представлять программу в таком виде: для декода все еще требуется участие пользователя, а внутри еще есть, что оптимизировать. Но со временем (и с вашей помощью) мы сделаем ее лучше.

Как пользоваться

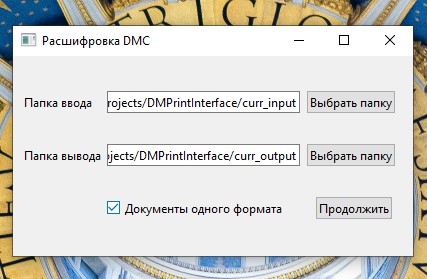
Когда вы запускаете программу, появляется главное меню. Здесь нужно выбрать, откуда вы хотите загрузить PDF-файлы для декода и куда хотите выгрузить результат.



Для этого достаточно выбрать папку и нажать на «Выбор папки». Ничего кроме папок пока выбирать нельзя.



Если во всех документах в папке расположение кодов одинаковое, рекомендую поставить галочку в боксе «Документы одного формата» -- так вам не придется вводить область распознавания для каждого документа отдельно. Когда папки и наличие/отсутствие галочки установлены, жмите «Продолжить».

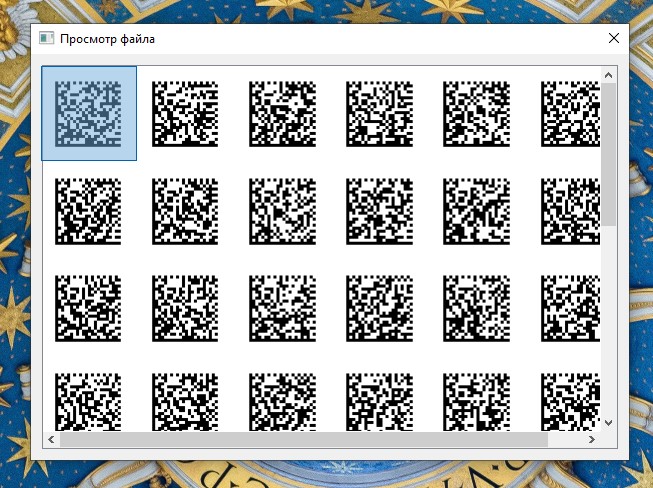


В ответ вы получите такое окно просмотра файла. Здесь ваша задача – выделить мышкой границы области с кодом. Тут внимательно: не самого кода, а всей области. Алгоритм возьмет эту область за шаг и будет смещать ее вправо и вниз в поисках кодов.

Например, в этом документе область кода выглядит так:

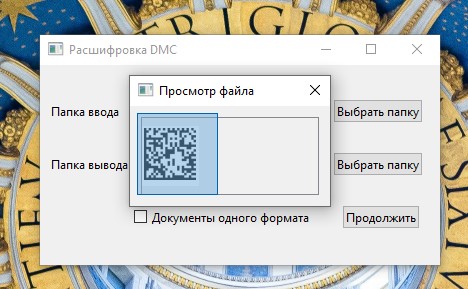


А в этом – так:

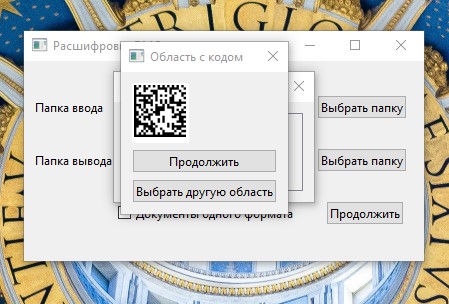


(Кстати, если передумали декодировать -- нажмите крестик на окне и закройте программу. Это еще не поздно сделать на этапе выбора области или на прогресс-баре декода. На обработке проблемных страниц уже не получится.)

Если файл содержит один код на страницу, выделите область немного поверх границ страницы. Тогда программа автоматически подставит границы страницы как область выделения.

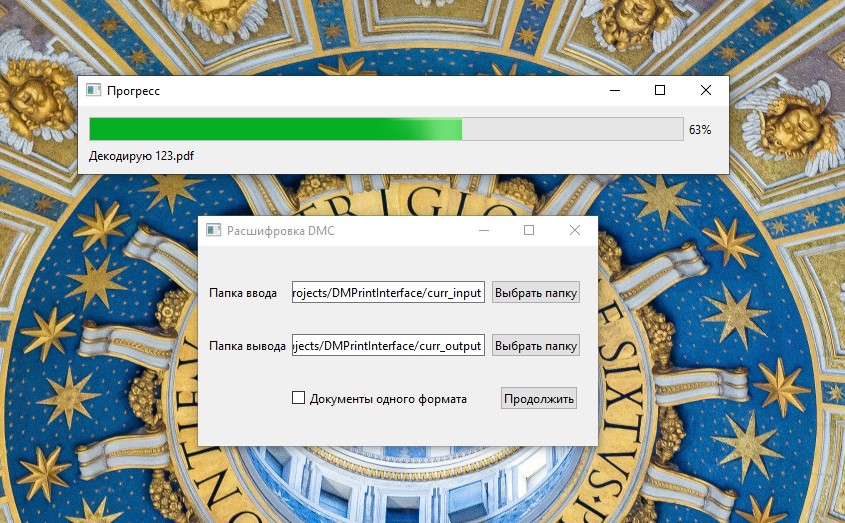


Если выделили что-то не то – не страшно. Следующим шагом программа покажет вам, что будет распознавать. Можно нажать «выбрать другую область» или крестик, тогда вы вернетесь к превью страницы и сможете выбрать область с кодом еще раз.



Если выберете «Продолжить», то возможны два варианта развития событий:

1. Появится предпросмотр следующего файла в папке – программа предложит ввести области для всех файлов перед декодом, и только потом начнет декодировать – чтобы вы могли поставить ее работать и пойти пить чай.
2. Начнется декод (если в папке только один файл, вы задали область последнему файлу в папке, либо выбрали один формат для всех файлов в папке). Вы увидите прогресс-бар, который будет заполняться для многостраничных файлов и не будет для одностраничных – в этом случае он просто покажет, что работает над ними.



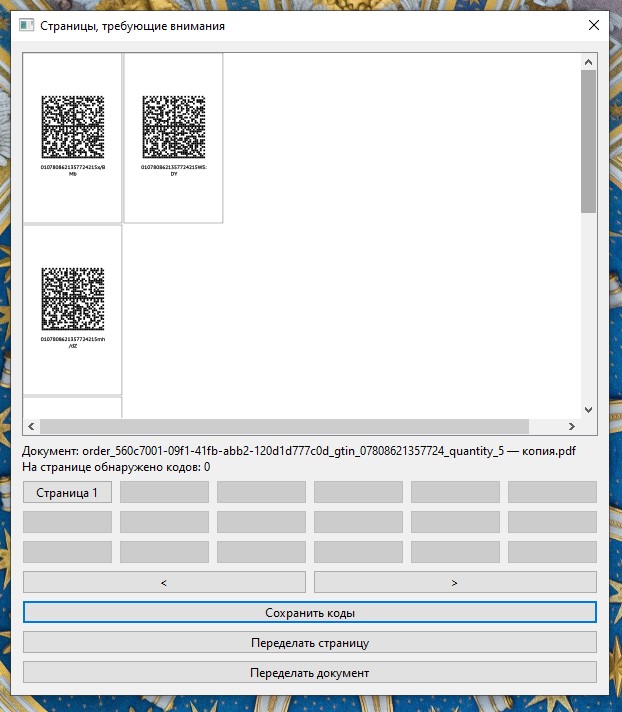
Алгоритм отслеживает ошибку в распознавании двумя способами:

1. Записывает количество кодов на первой странице, и, если на какой-то странице документа количество кодов не совпадает с этим значением, помечает страницу как возможно содержащую ошибку.
2. Если в области не обнаружен код, проверяет ее на наличие черной точки. Если находит ее, также помечает страницу как возможно содержащую ошибку.

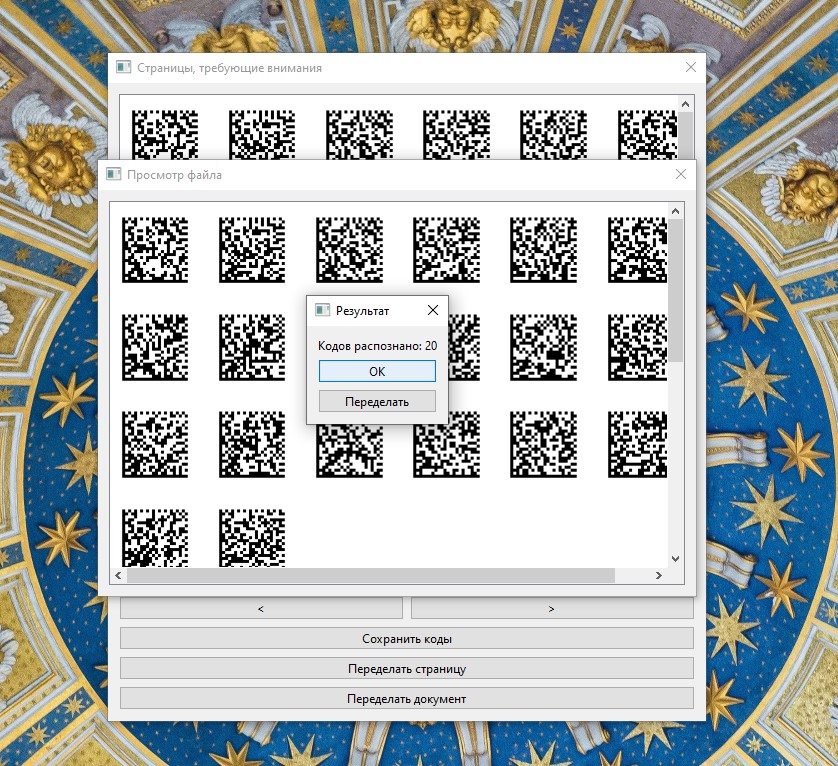
Если в документе нет потенциально ошибочных страниц, программа молча запишет коды из него в csv-файл.

Иначе – для каждого документа программа выведет окно со страницами, которые подозревает на наличие ошибки.

В примере ниже каждая страница файла содержит 48 кодов, кроме последней. Если вы видите, что ошибки нет и распознано верное количество кодов, вы можете просто сохранить коды в файл как есть (кнопка «Сохранить коды»). Либо, вы можете переопределить границы области с кодом и повторно запустить декод страницы или всего документа.



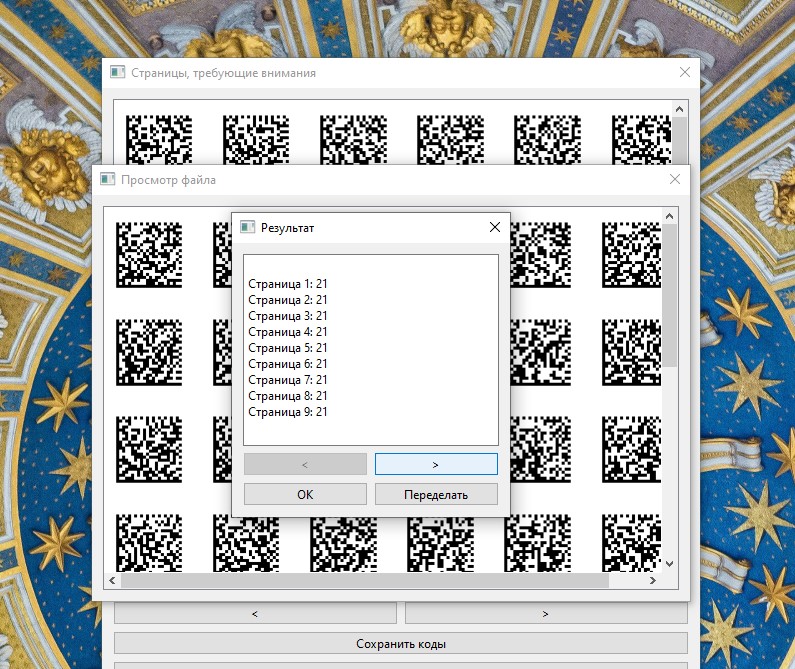
Соответственно, если вы выберете «Переделать страницу», то программа снова предложит вам задать область распознавания и прогонит страницу через декодер. Затем она покажет, сколько кодов удалось распознать.



Нажимаете «Переделать» – снова выбираете область и программа декодирует страницу еще раз. Нажимаете «ОК» -- программа сохраняет коды для страницы, но пока только внутри себя. Чтобы записать коды в файл, понадобится нажать еще и «Сохранить коды» в окне предпросмотра ошибок, которое появилось ранее.

Если выбираете «Переделать документ», программа предлагает заново выбрать область распознавания и повторно декодирует весь файл. В результате она выведет окно с количеством распознанных кодов на каждую страницу. А вы уже решайте – переделывать или нет.

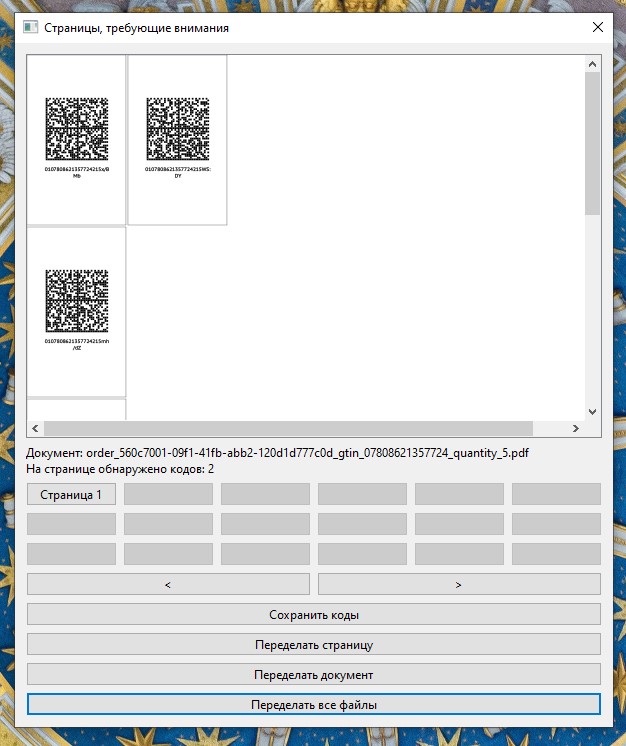
Главное, не забывайте: кнопка «ОК» в этом окне коды в файл не сохранит – обязательно нужно вернуться в окно предпросмотра страниц с потенциальными ошибками и нажать «Сохранить коды».



Будьте осторожны: в программе нет защиты от дублирования кодов в файл. Если каким-то образом вы дважды нажмете «Сохранить коды», то они запишутся дважды.

Когда нажмете «Сохранить коды», предпросмотр документа закроется и откроется предпросмотр следующего документа, в котором замечены потенциально проблемные страницы. Либо сообщение, что все файлы были декодированы.

Если вы отметили в боксе, что все документы в папке одного формата, то у вас будет опция переделать все документы в папке. Это, скорее всего, будет долго. А еще, придется проследить за задвоением кодов: вдруг для каких-то документов коды уже распознаны и записаны в файл. В общем, эта функция не отлажена, но она есть.



Это вся основная информация по работе с программой. Если у вас возникнут вопросы, жалобы и пожелания – пишите или говорите мне вслух. Ваша обратная связь очень важна.