How to Use DocScan and the ScanTent

Last update of this guide 24/09/2018

DocScan and the ScanTent are new tools to help you digitise historical documents and books with your smartphone. The **ScanTent** is a portable piece of equipment that holds your smartphone in place above a document. **DocScan** is a mobile app which helps you to take high quality images in an efficient manner. Users can upload images from DocScan directly to the Transkribus platform in order to work with Handwritten Text Recognition technology.

Download the DocScan app (Android only):

- https://play.google.com/store/apps/details?id=at.ac.tuwien.caa.docscan

Visit the website for DocScan and the ScanTent:

- https://scantent.cvl.tuwien.ac.at/en/

Download the Transkribus Expert Client, or make sure you are using the latest version:

- https://transkribus.eu/

Consult the Transkribus Wiki for further information and other How to Guides:

- https://transkribus.eu/wiki/

Transkribus and the technology behind it are made available via the following projects and sites:

- https://read.transkribus.eu/
- https://transcriptorium.eu/
- https://github.com/transkribus/

Contact:

- The Transkribus Team: email@transkribus.eu

Contents

Introduction

DocScan and the ScanTent allow you to digitize historical documents and books in excellent quality, without a special book scanner.

You can use your **smartphone to take high quality images** of books and of documents in a fast and reliable way. The DocScan app has a **'series mode'** which will take a photo every time a page is turned, leaving you with your hands free to hold and move documents.

You can edit your images and upload them directly to the Transkribus platform, where you can work with **Handwritten Text Recognition (HTR)** technology (for more information see How to Transcribe Documents with Transkribus - Introduction).

Working with DocScan and the ScanTent, it will take around 12-15 minutes to scan a book with 300 pages (=150 images). So as a rule of thumb, you will be able to take **more than 500 images within an hour**.

ScanTent

The ScanTent is a portable piece of equipment that holds your smartphone in place above a document.

The main features of the ScanTent are:

- **Professional photo environment** for taking high quality pictures without any extra light (nylon silk fabric)
- **LED lightning with USB power supply** for indirect illumination of the documents (if necessary)
- **Black felt fabric** at the base for a standardized background
- **Large base area** so that users can place their arms inside the ScanTent to hold bound documents open with both hands
- Scanning of documents of around **A3 size or even a bit larger**
- **Lightweight (500 gram)** and **collapsible** (fits into small suitcases)

The ScanTent will be available on the market in late 2018.

Figure 1 The ScanTent

If you would like to receive more information once the ScanTent is available for purchase or hire, please contact: scantent@caa.tuwien.ac.at

DocScan

The DocScan app was developed specifically for the task of digitizing books and archival documents with a smartphone. It is currently available for Android phones.

The main features of DocScan are:

- **Fast and reliable detection** of document pages
- **'Single mode'** for taking individual images manually
- **'Series mode'** for automated capturing of images (movement is detected automatically)
- Ability to **crop pages**
- **Direct upload** of documents to the **Transkribus** server

The DocScan app can be downloaded for free from the Google Play store:

- https://play.google.com/store/apps/details?id=at.ac.tuwien.caa.docscan

Figure 2 Logo for DocScan and the ScanTent

Use Cases

ScanTent and DocScan for archives and libraries

ScanTent / DocScan are excellent alternatives to the usual book scanning devices within archives and libraries.

The main advantages are:

- **High quality images** - modern smart phones deliver excellent image quality with high resolution
- **Cost efficient** – both for the end-user as well as for the library/archive
- **No support or licensing fees**
- **No user support** needed from the archive or library - users will get familiar with the DocScan app on their own
- **Copyright friendly** – users take and store images on their own device, rather than those owned by a library or archive

Further advantages for public domain material: user generated scans

DocScan offers an option for "crowd-scanning", where user-generated images can be added to the digital holdings of a library or archive.

The workflow is as follows:

- A library/archive creates automatically a **QR code with the main metadata of a record** (signature, record identifier, Transkribus identifier), which is displayed on the institution's website.
- **A user will scan the relevant QR code** with the DocScan app and then begin digitizing the document
- **A copy of each digitised image** and its appropriate metadata will then become **directly available to the library or archive**, from where it can be connected with their digital repository.

A pilot project of this workflow is currently set up with a partner archive in Germany. Results will be made available during 2018.

Scanathons

ScanTents will also be available for rent for libraries and archives interested in organizing "Scanathons".

The event will work as follows:

- A library or archive contacts **potential volunteers** among their users and rents (or buys) e.g. 20 ScanTents
- A specific (historical) **collection is selected for scanning**, e.g. documents relating to a "local hero"
- **Volunteers come with their own smartphones** and receive instructions from the library or archive staff about using DocScan and the ScanTents
- A **scanning session** takes place for e.g. 2 hours
- The Scanathon could be repeated the next day or even become a weekly activity for volunteers

Within 2 hours, each volunteer will be able to scan up to 1000 images (or 2000 book pages) from. This means that **20 volunteers will produce up to 20.000 images** or 40.000 book pages within one session!

The speed may be slower with more delicate archival material but volunteers will still be able to generate tens-of-thousands of high quality images.

Our first Scanathons took place in mid-2018 at the National Archives of the UK, the National Archives of Finland and the State Archives of Zurich:

- Report on 2018 Scanathon

Assembling the ScanTent

The ScanTent includes the following components:

- 1 x Tent fabric
- 4 x Tent poles
- 1 x Pad for smartphone
- 1 x Felt background
- 1 x LED lights
- For a quick overview of ScanTent assembly, please watch our video: https://youtu.be/iL2WNNi5VEI

The ScanTent should be assembled as follows:

- Place the smartphone pad on a flat surface, with the circular part facing upwards

Figure 3 Smartphone pad

- Unfold the tent fabric and position it over the smartphone pad. The tent has a large opening at the front that should be pointing towards you. The circular part of the pad should poke through the circular hole at the top of tent fabric.

Figure 4 Tent fabric

Figure 5 Tent fabric on smartphone pad

- There are four holes at the sides of the circular part of the pad. Take one of the tent poles and insert it in top left hole. Then stretch the pole out and insert the other end at the top left corner of the base of the tent fabric. Repeat this process for the other tent poles.

Figure 6 Inserting tent poles

Figure 7 Inserting tent poles

Figure 8 Inserting tent poles

Figure 9 Fixing the tent poles

Figure 10 Fixing the tent poles

Figure 11 ScanTent with all tent poles fixed

- Close the belt at the front of the ScanTent.

Figure 12 Closing the belt

- Lay the felt background on your surface and place the assembled ScanTent on top

Figure 13 ScanTent on felt background

Once the ScanTent is set up, you can start with taking pictures. If the environmental light is not sufficient (which is very rarely the case) plug the lights into a laptop or USB socket to turn them on.
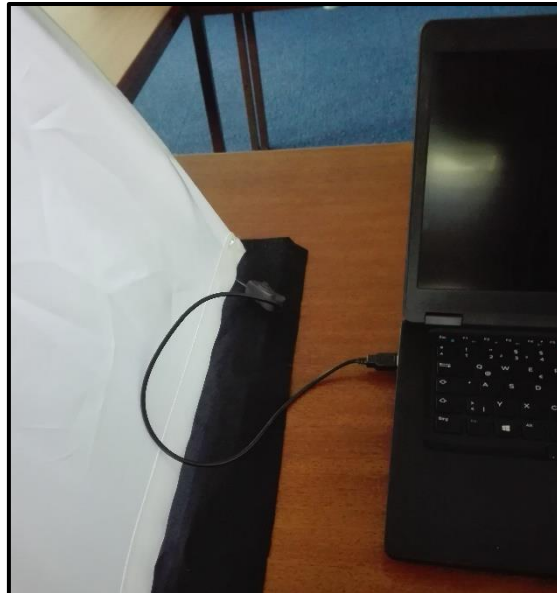
Figure 14 Light via USB

Place your smartphone on the pad at the top of the ScanTent, with the camera lens facing downwards. The lens should line up with the hole at the top of the ScanTent.

Figure 15 ScanTent ready for scanning

Working with DocScan

Login

If you are a Transkribus user, you should log into your Transkribus account in the app.

Press the menu icon (three lines) at the top left of the app to log in.

Opening documents

In order to start scanning with DocScan, you need to open a new or existing document.

Open a new document

To open a new document:

- Press the folder icon at the top right of the app.
- Then press "New document".
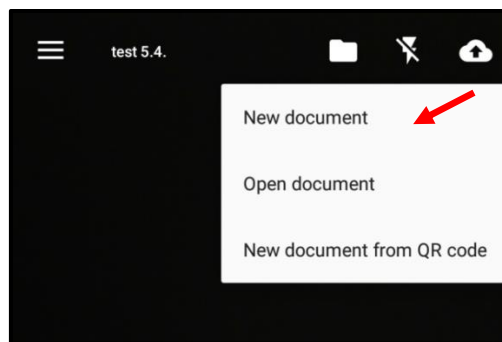
Figure 16 Open a document

- Assign a name to your new document.
- Press "Create Document"
- All the images you subsequently take will be stored under this name on your phone and when you upload them to Transkribus
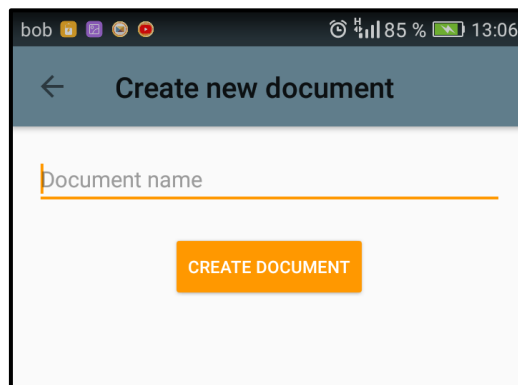
Figure 17 Assign a name to a new document

Open an existing document

To open an existing document:

- Press the folder icon at the top right of the app.
- Then press "Open document".
- Press the name of an existing document
- Press "Use selected document"
- All the images you subsequently take will now be assigned to this document

If you fail to open a new or existing document, your images will be stored under a document called "Untitled document".
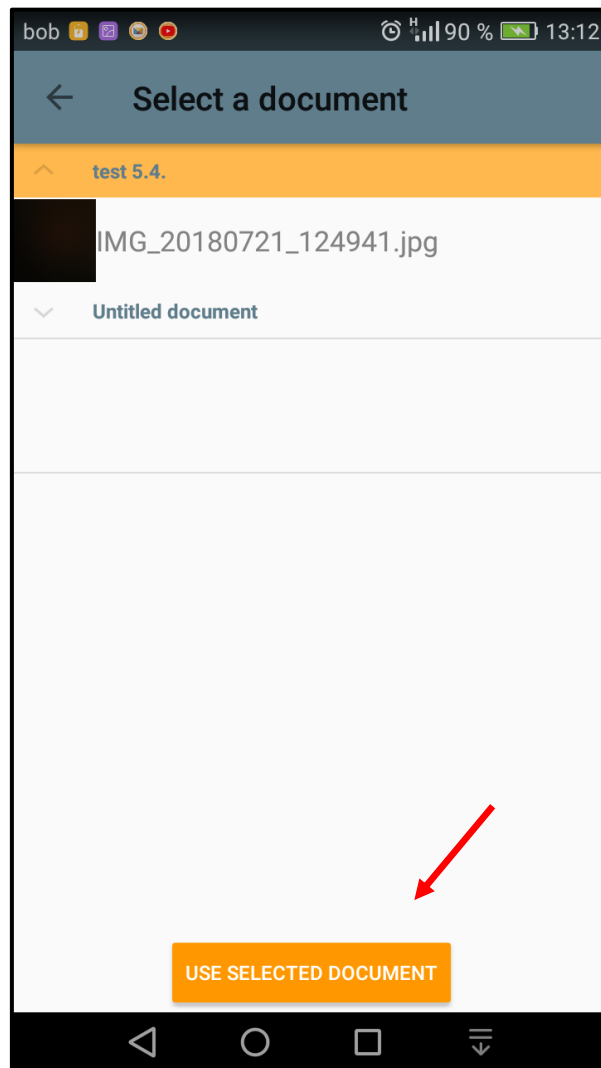
Figure 18 Add new images to existing document

Scanning

Once you have created a document, you can begin scanning!

- **Position your phone on top of the ScanTent**
- **Important**: The display of the smartphone needs to be aligned with the direction of the document. **If you stay in front of the ScanTent you must be able to read the messages of DocScan**
- DocScan may display feedback about focus, perspective and positioning
- When you are ready to take an image, press the "Play" button at the bottom of the app
- An image will be captured, as long as the ScanTent detects no other movement
- Turn to the next page of your document and continue scanning
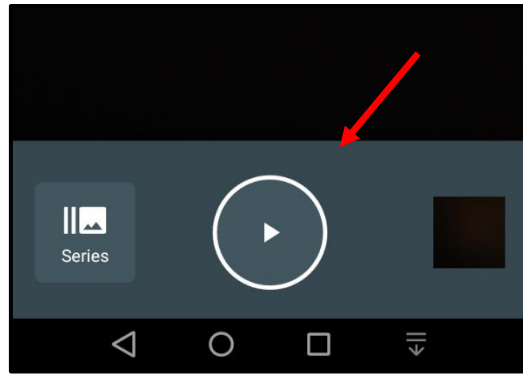
Figure 19 Start scanning

Shooting modes

You can choose to take single images manually, or set the app to automatically capture an image every time a page is turned. You can select from the "Single" or "Series" options at the bottom left of the app.
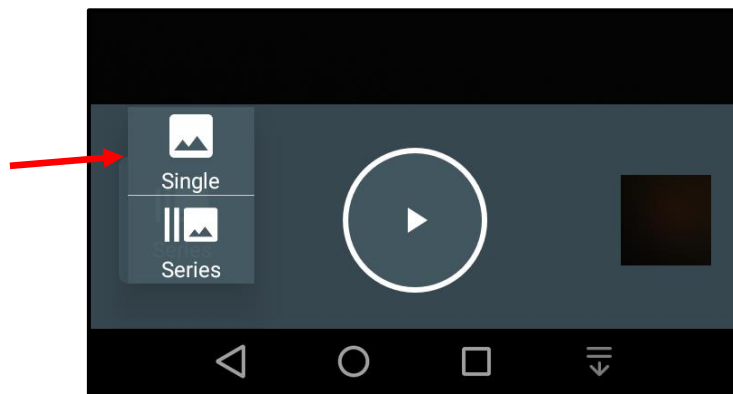
Figure 20 Single scan

Single mode

This mode allows you to take images manually by pressing the "Play" button. It is mainly used if no ScanTent is available and you just want to take some images of a document.

Series mode

This mode is used for efficient scanning. With this mode, DocScan will take an image every time a page is turned. The app will prompt you to turn the page if it has already been scanned. You now have your hands free to turn pages and move documents inside the ScanTent.

Important: In order to know when DocScan has taken an image automatically you have to **turn on the sound on your smartphone**. In this way you will "hear" when the image is taken and turn the page accordingly. **Note: It is expected that you sit comfortable in front of your ScanTent and that you need not to look at the display of your phone!**

Continue scanning

You can return to the scanning function at any time pressing the burger icon (three lines) at the top left of the app and then "Camera".
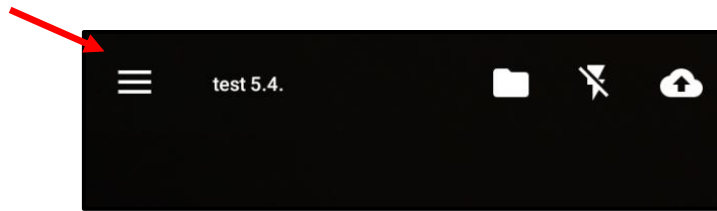
Figure 21 Menu
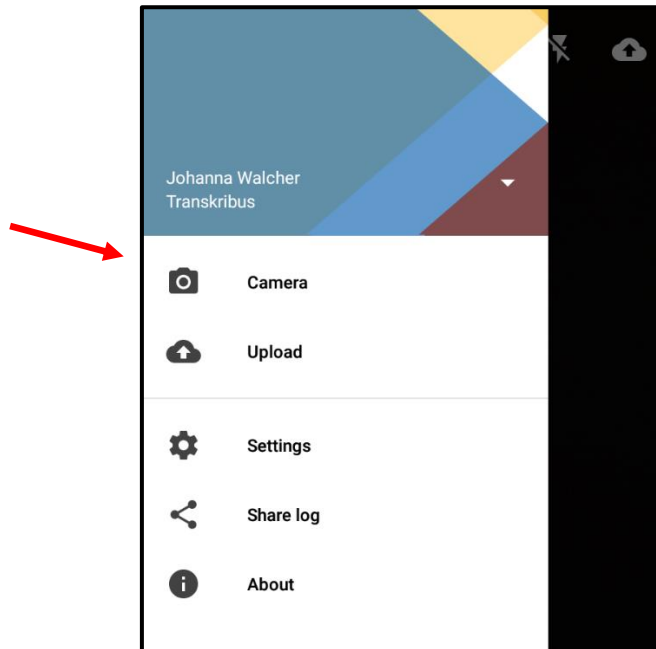
Figure 22 Camera

Uploading to Transkribus

Once you have finished scanning, you can upload your images directly to Transkribus.

- Press the cloud icon at the top right of the app
- Login to your Transkribus account if necessary
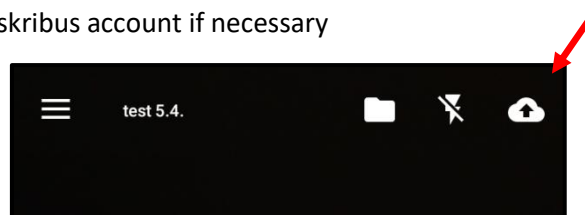
Figure 23 Uploading

- Select the document that you would like to upload to Transkribus
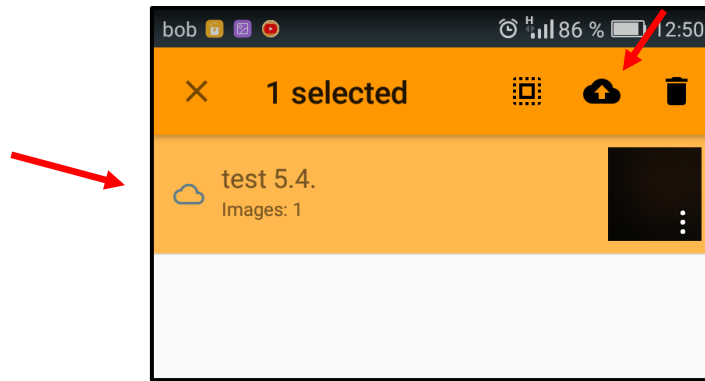- Press the cloud icon once more

Figure 24 Choose the document to be uploaded

- Open Transkribus on your computer
- You will find your uploaded documents in a collection called "DocScan – Uploads"
- Usually the upload will be ready in a few minutes. If you are uploading a large number of pages, it may take a little longer.

Other settings

Further settings can be found and adjusted by pressing the burger icon (three lines) at the top left and then "Settings".
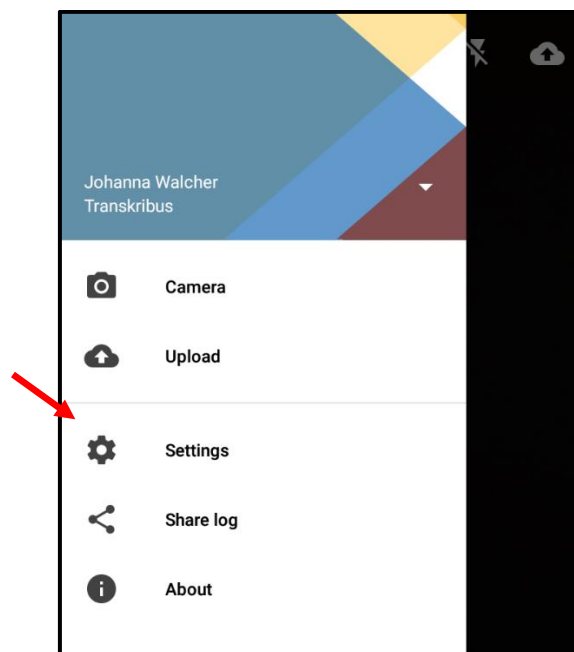
Figure 25 More settings

Light

The flash can be adjusted by pressing the flash icon at the top right of the app.

There are four options:

- Flash off
- Flash auto
- Flash on
- Torch
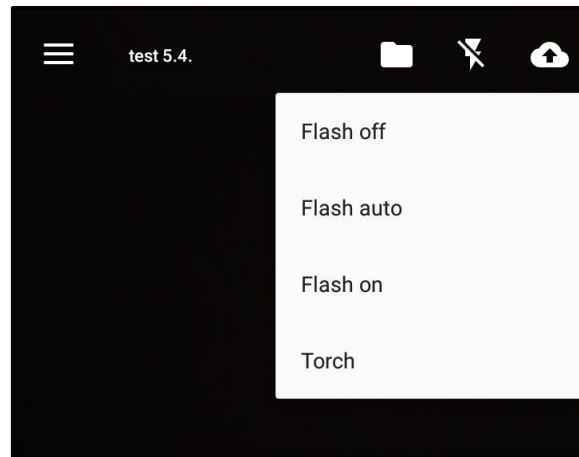
Figure 26 Flash settings

Editing your image

You can use DocScan to crop and rotate your images as required.

Once you have taken an image, press the thumbnail image at the bottom right of the app to open the edit settings.

Automated cropping, rotating and deleting

With the automated cropping function, you won't need to drag the frames into the right position, the app does it for you automatically.

- In order to use the automated cropping function first open the uploads, that is where you can find the scans you have taken.
- Choose the folder with the new scans to be cropped.
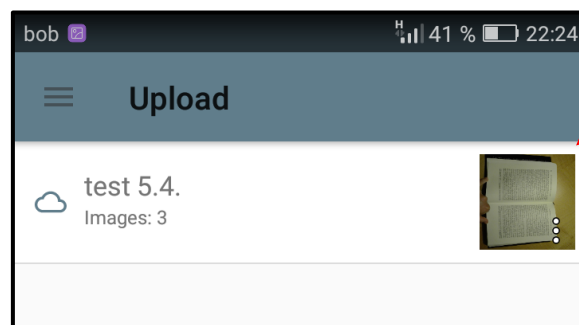
Figure 27 Open "Upload" to access the latest scans

- All pages will appear in yellow frames.
- Note: when cropping is activated several pixels will be added to the yellow frame, so the complete page is displayed in the image
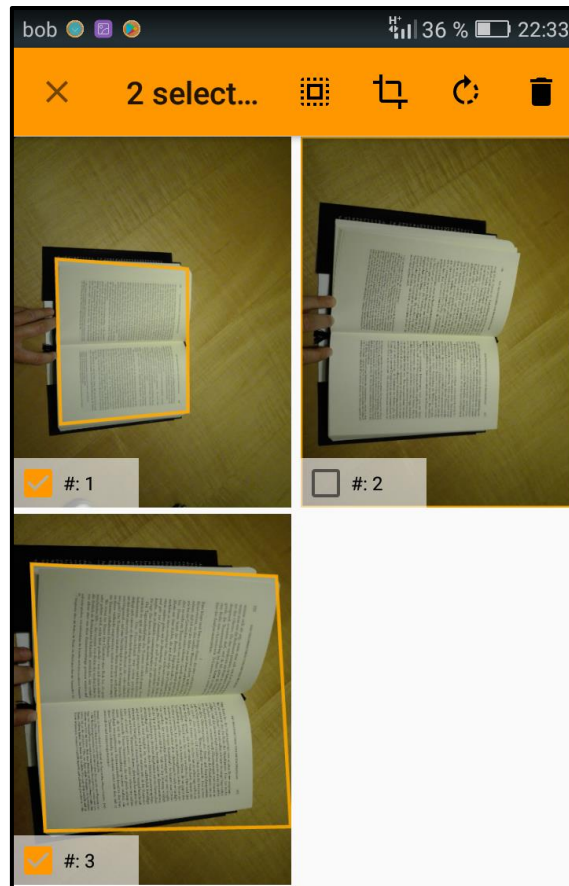- Mark the files you would like to be cropped.

Figure 28 Choose the files you would like to be cropped

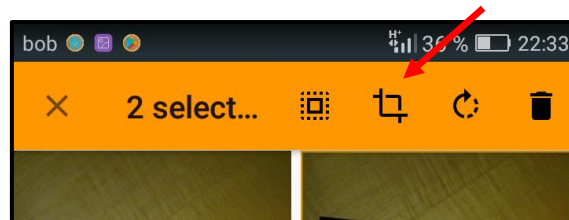- Finally click the cropping symbol to get the process started.

Figure 29 start the process by pressing the cropping button

Manual Cropping

It can happen, that the automatic cropping has got troubles with special kinds of pictures, in this case you can always use the manual cropping.

- Press the cropping icon at the bottom of the screen
- Drag the corners of the image to the desired position
- Press the cropping icon at the top right of the screen to save your cropped image
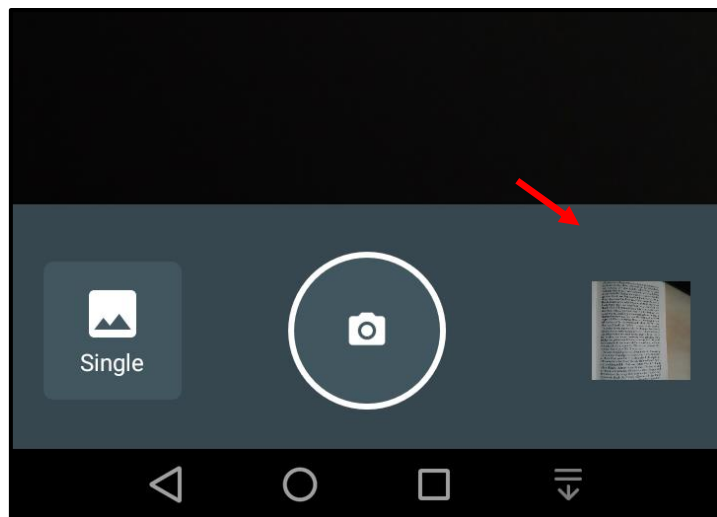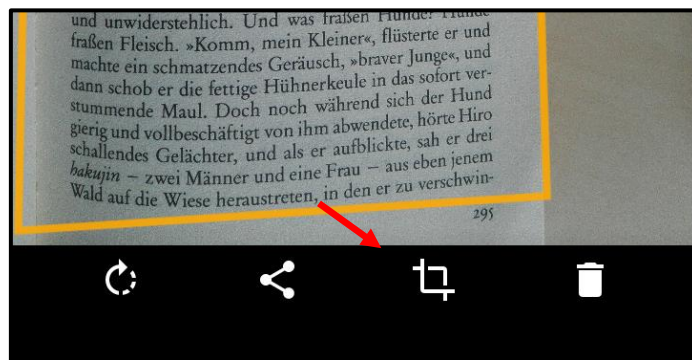- In the next screen, press the save icon
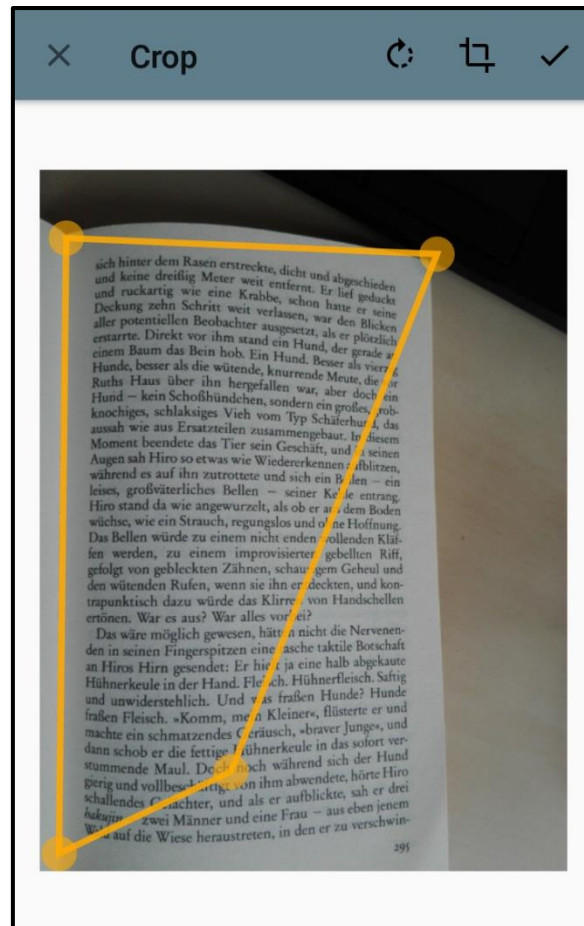
Figure 30 Cropping function

Figure 31Cropping

Figure 32 Adjusting frames

Rotating

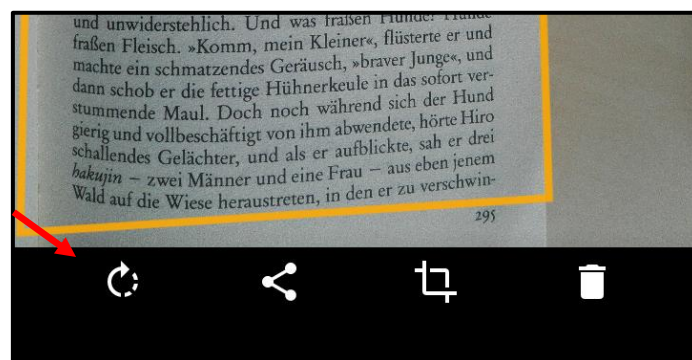- Press the circular arrow icon to rotate your image 90 degrees

Figure 33 Rotate the image

Sharing

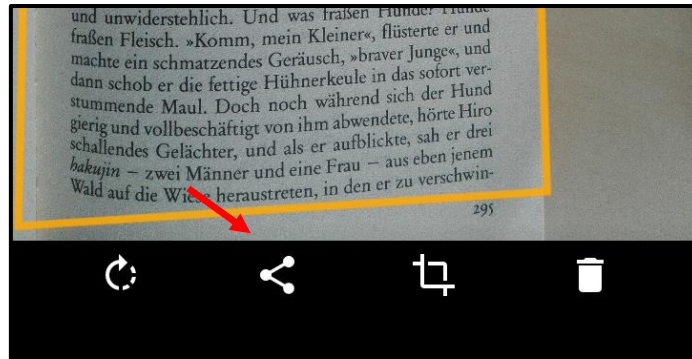- Press the share icon to share your images to other apps.

Figure 34 Share image

Deleting
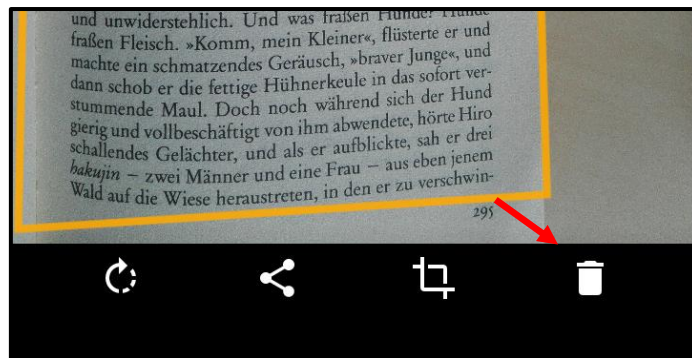- Press the bin icon to delete the image.

Figure 35 Delete image

Further information

If you would like to receive more information once the ScanTent is available for purchase or hire, please contact: scantent@caa.tuwien.ac.at

Credits

We would like to thank the many users who have contributed their feedback to help improve the Transkribus software.

Transkribus is made available to the public as part of H2020 e-Infrastructure Project READ (Recognition and Enrichment of Archival Documents) which received funding from the European Commission under grant agreement No 674943.