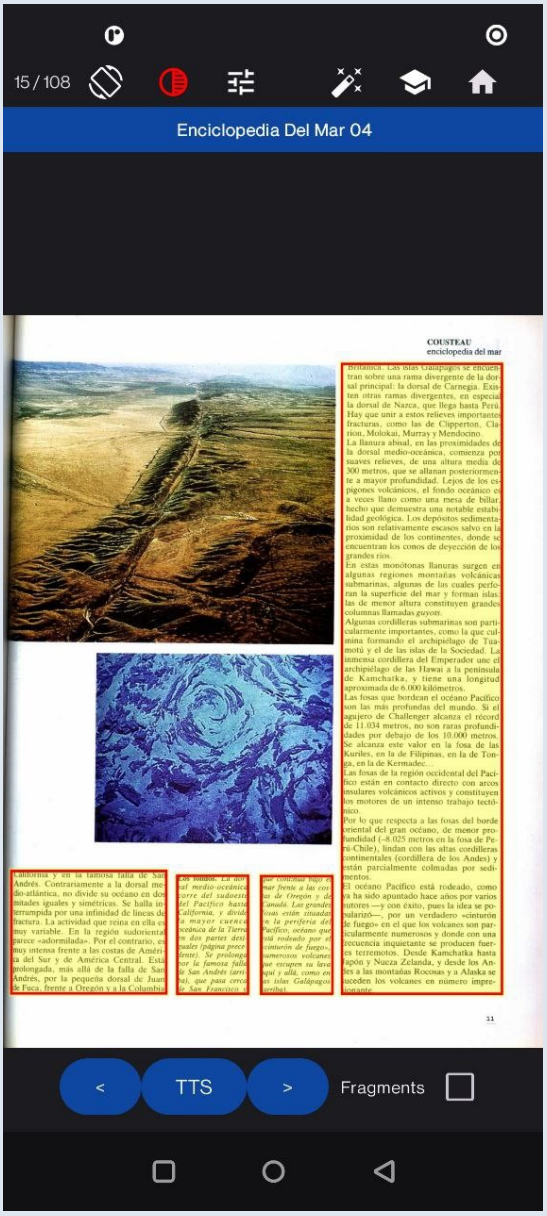
scan_ocr_tts

Introduction:
The program uses optical character recognition (OCR) technology to identify and extract text from documents. This recognition allows the program to detect text blocks on the page, facilitating reading. The text areas are then framed, allowing the user to visualize and select the parts they want to read or process.

We can then proceed with the audio playback of the text.

How it works:

When you open a PDF document and navigate to a page, the program attempts to detect text blocks and frame them with red rectangles on a yellow background.



It is possible to select/deselect the red rectangles by clicking on them.

Then, simply click the [TTS] button to start reading the blocks aloud.
At the end of the reading, the program will automatically move to the next page.

The reading of the blocks follows the order from top to bottom and left to right, which is the usual order for this type of layout, as seen in magazine PDFs.

Tip:

Sometimes, when the reading order is not correct, it may be desired to prioritize listening to certain

blocks first, and then the others later.

To do this, you can use the [sel. frames] checkbox.

Procedure :

- Select the first blocks to read
- Check the [sel. frames] checkbox
- The reading will start automatically
- At the end of the reading, all the blocks are deselected.
- The checkbox will be automatically unchecked
- Activate the new block to read
- Use the [TTS] button to finish reading

About text recognition:

The program uses optical character recognition (OCR) technology to identify and extract text from documents. This recognition allows the program to detect text blocks on the page, facilitating reading. The text areas are then framed, allowing the user to visualize and select the parts they want to read or process.

Given the great variation in PDF documents, especially in terms of contrast, the program offers several sliders to facilitate text recognition.

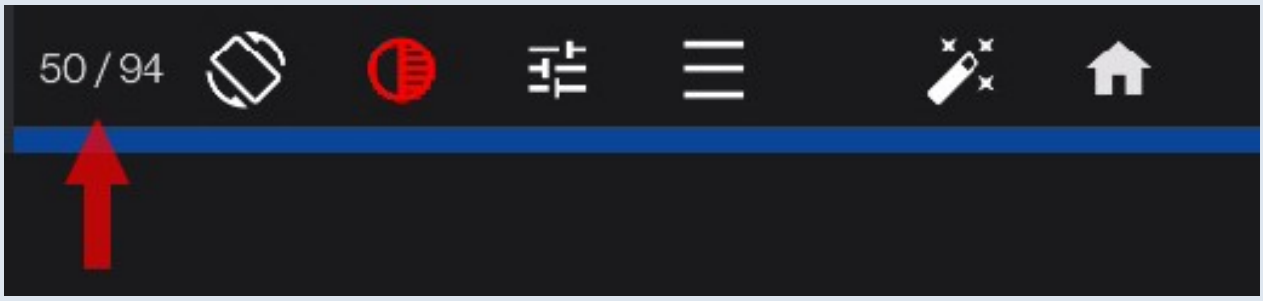
(See below for how these sliders work)

When closing the document, the slider values and the current page are saved, so the same situation is restored when reopening the document.

The toolbar.

Let's see what it consists of, from left to right.

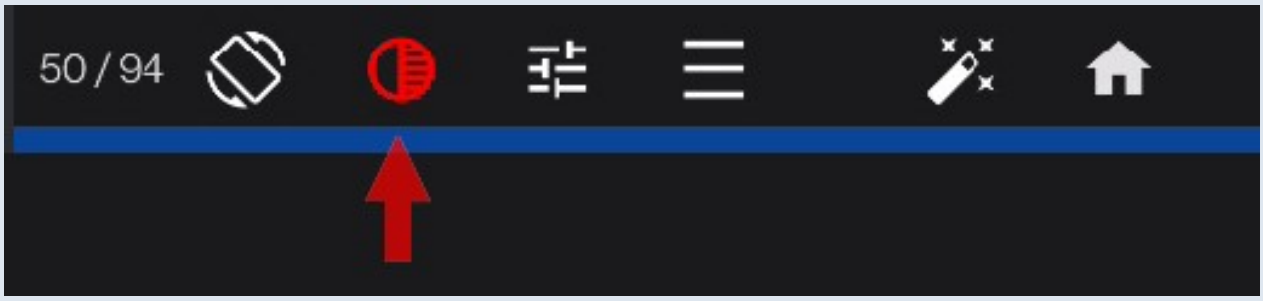
If you click on the current page indicator, it will be possible to jump to another page...



If you click on the flip icon, the screen will rotate 180°.

I use it to keep the power cable connected. It bothers me less when the connector is at the top rather than at the bottom. 😉

Increase the contrast.



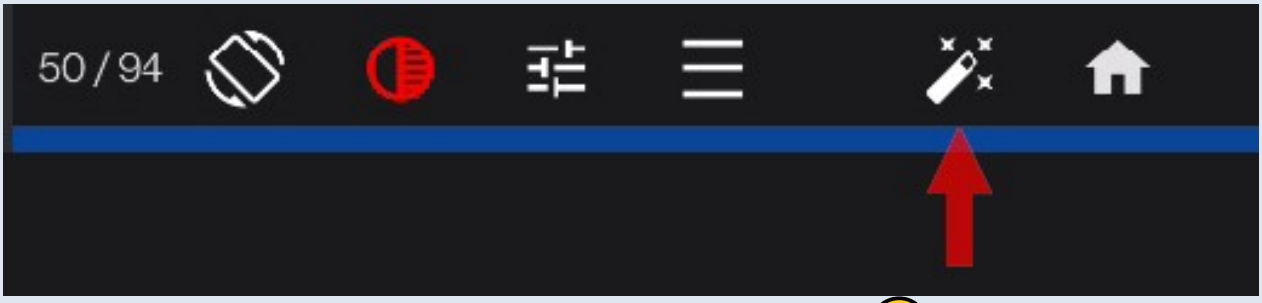
By default, the contrast is increased to improve text detection. It can be disabled if the text detection results are not satisfactory.

Home.

Return to the home page to change the PDF or exit.

Note:
The PDF settings are saved at this exact moment. Therefore, it is better to use this option rather than minimize the program.

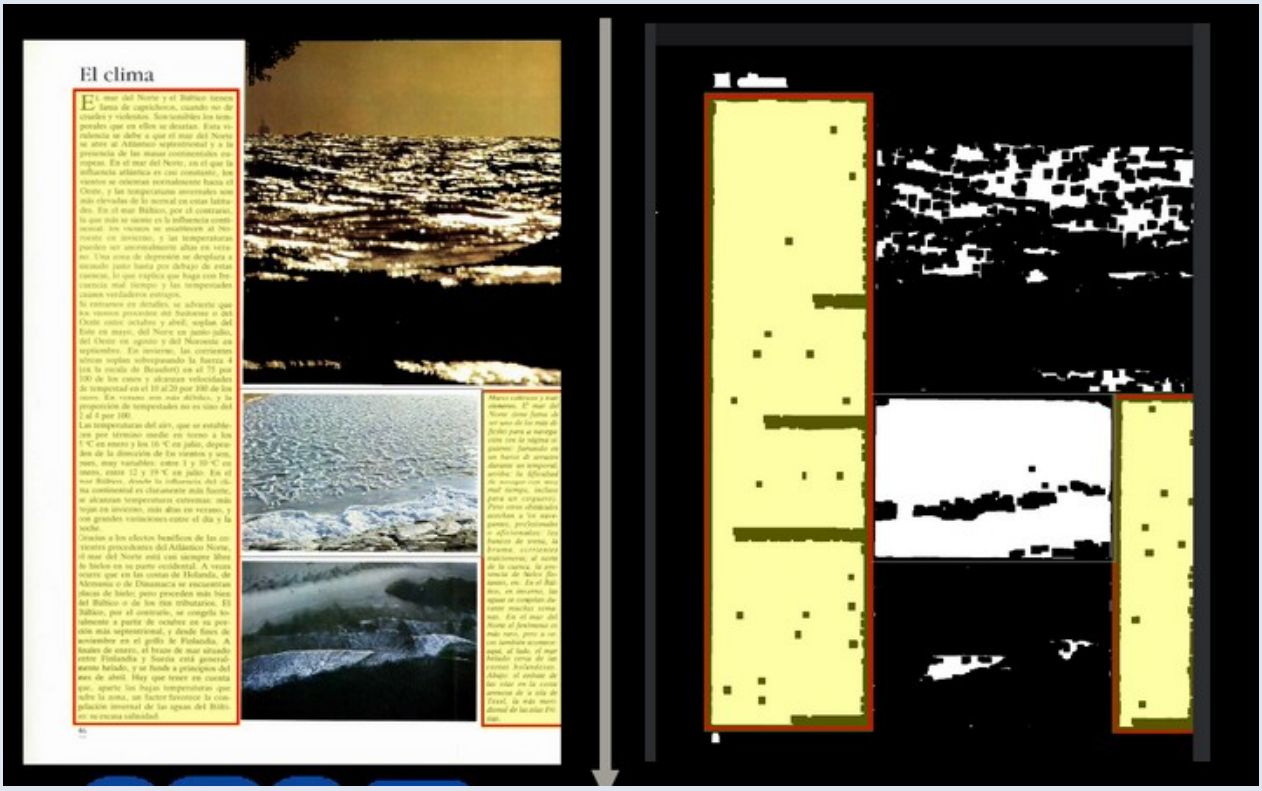
Working view.



This is where the "serious" stuff begins. 😉

To locate the text areas on the pages, the program converts the page to black and white, enhancing the details.

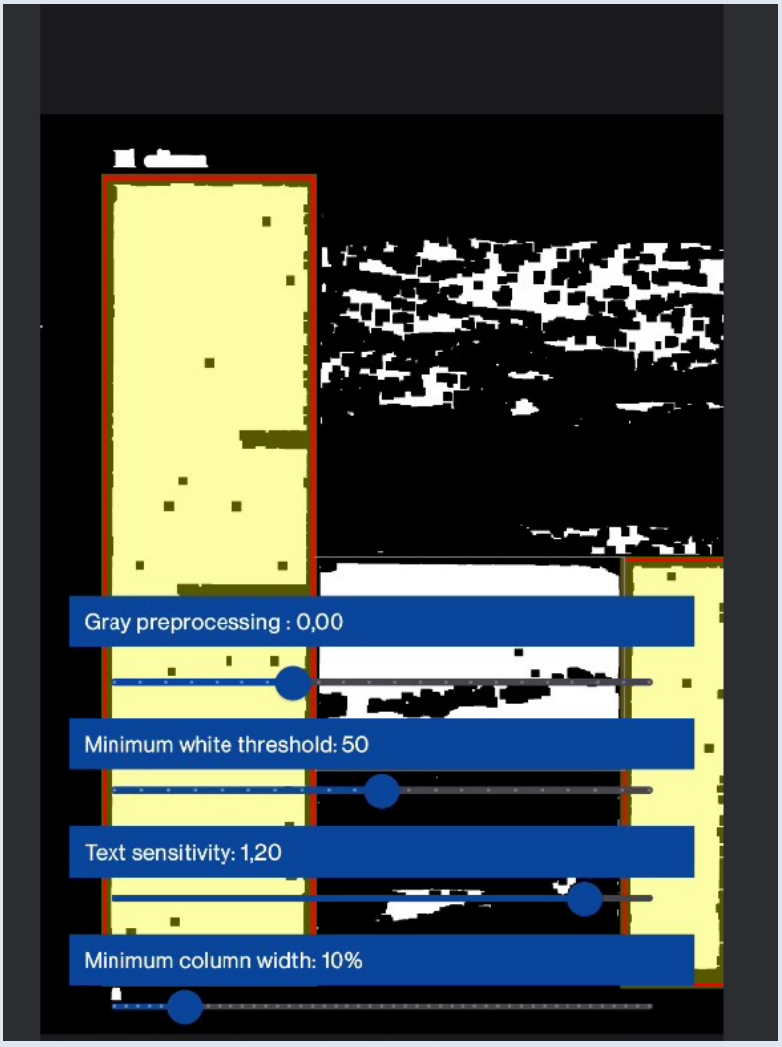
With this option, you can switch from one view to another.



It helps with the settings that will follow...

The sliders.

This button will overlay sliders on our page.



These are the sliders that we will need to adjust based on the quality of the PDF. The changes are applied in real time, and the result is immediately visible on the page (below the sliders).

Let's quickly go over them.

- **Gray preprocessing**
  Adjusts the gray level to optimize text detection.

- **Minimum white threshold**
  This value is expressed in %. Inside our red rectangles, it represents the minimum % of white required for them to be accepted.

- **Text sensitivity**
  This is the sensitivity for detecting text within the red squares.

- **Minimum column width**
  This is the minimum width the rectangles must have, below which they are ignored.

Each PDF document is a "special case," so I can't give you a precise method for adjusting the sliders. However, it's enough to experiment a bit and observe the real-time result to learn how to use them.

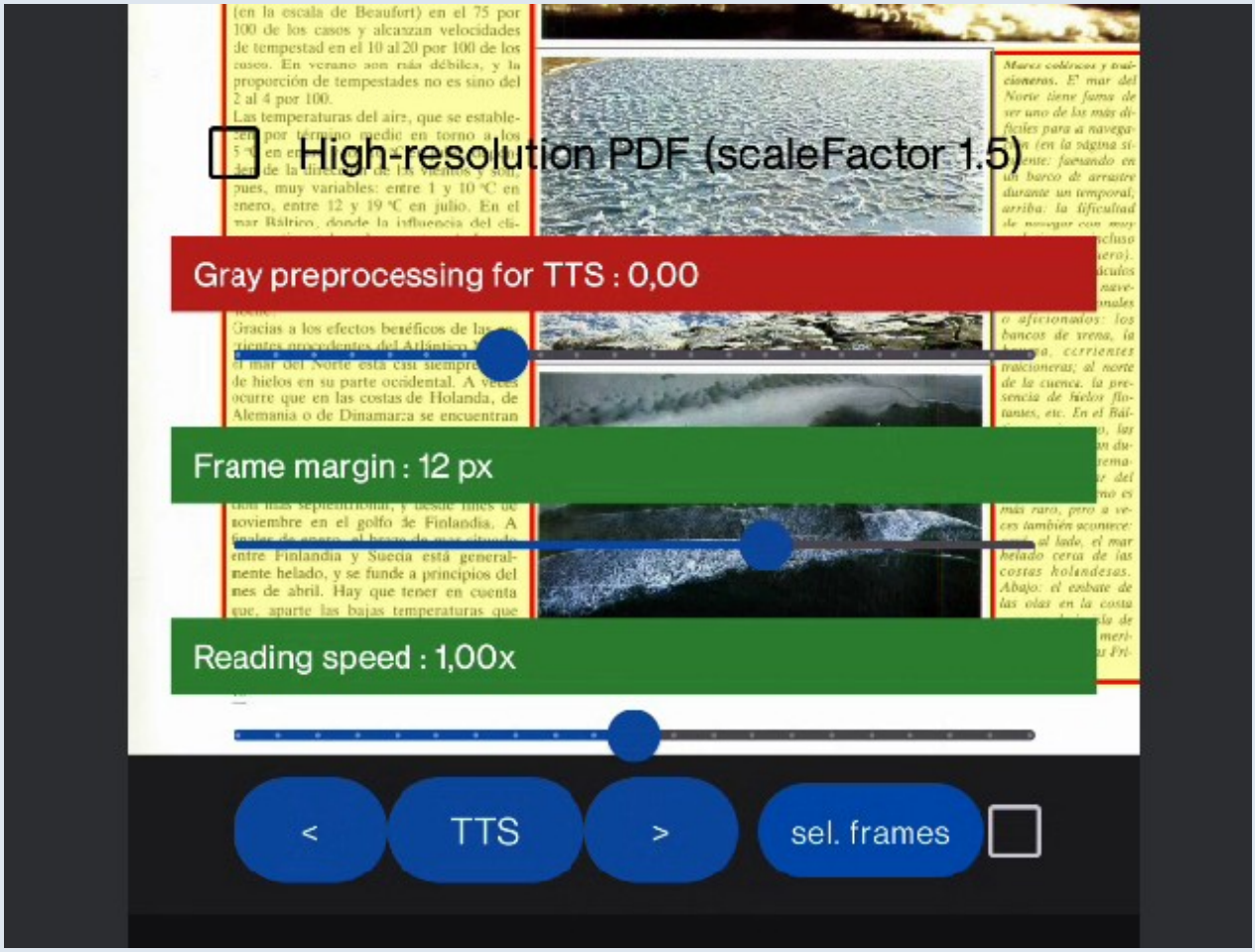
Contrary to what one might think, it's not that complicated!

In principle, once adjusted for one page of the PDF, the settings should be valid for the rest of the pages, with sometimes just a few minor adjustments.

Remember that when leaving the view with the home button, all settings are saved, and this applies individually for each PDF.

Options :

There are a few other additional options.



When we have a PDF of a scanned document with poorly resolved text, we can try these two options:

Check the box:

High-resolution PDF (scaleFactor 1.5)
This increases the resolution of the pages to try to read the text better. However, be careful, as it also slows down the program.

Same for the slider:
Gray preprocessing for TTS

Note: This setting only applies when performing OCR on the text within the rectangles.
It is unrelated to the previous sliders, which are used to detect where the text is located on the pages.

Use these two options as a last resort if you're having trouble obtaining accurate text.

Frame margin.
This is simply an option for fun; it draws the red frames a little smaller or larger around the text.

Reading speed.
Here, you can adjust the speed of the text-to-speech voice. This setting is also saved and depends on the language.