

scan_ocr_tts

Introduction.

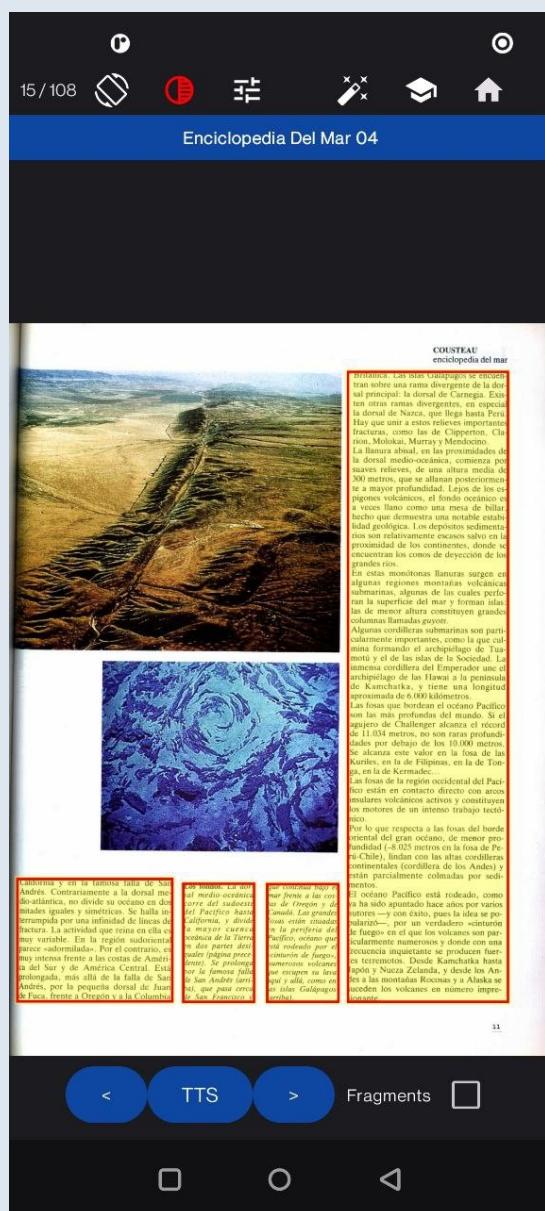
Le programme utilise une technologie de reconnaissance optique de caractères (OCR) pour identifier et extraire le texte des documents. Cette reconnaissance permet de détecter les blocs de texte sur la page, facilitant ainsi la lecture.

Les zones de texte sont ensuite encadrées, ce qui permet à l'utilisateur de visualiser et de sélectionner les parties à lire ou à traiter.

On peut ensuite procéder à l'écoute audio du texte.

Principe de fonctionnement :

Lorsque vous ouvrez un document PDF et que vous naviguez vers une page, le programme tente de détecter les blocs de texte et les encadre de rectangles rouges sur fond jaune.



Il est possible de sélectionner/désélectionner des rectangles rouges en cliquant dessus.

Ensuite, il suffit de cliquer sur le bouton [TTS] pour lancer la lecture audio des blocs.

À la fin de la lecture le programme passera automatiquement à la page suivante .

La lecture des blocs suit l'ordre de haut en bas et de gauche à droite, ce qui est l'ordre habituel pour ce type de mise en page par exemple dans les pdf de revues.

Astuce :

Parfois, lorsque l'ordre de lecture n'est pas correct, il peut être souhaité d'écouter une série de blocs en priorité, puis les autres ensuite.

Pour cela on utilise la casse à cocher [sel. frames]

Procédure :

- Sélectionner les premiers blocs à lire
- cochez la case [sel. frames]
- la lecture démarre automatiquement
- à la fin de la lecture tous les blocs sont désélectionnés
- la case est décoché automatiquement
- activez les nouveaux bloc à lire
- utilisez le bouton [TTS] pour terminer la lecture

À propos de la reconnaissance du texte :

Le programme utilise une technologie de reconnaissance optique de caractères (OCR) pour identifier et extraire le texte des documents. Cette reconnaissance permet de détecter les blocs de texte sur la page, facilitant ainsi la lecture. Les zones de texte sont ensuite encadrées, ce qui permet à l'utilisateur de visualiser et de sélectionner les parties à lire ou à traiter.

Étant donné la grande variation des documents PDF, notamment en termes de contraste, le programme propose plusieurs curseurs afin de faciliter la reconnaissance du texte.

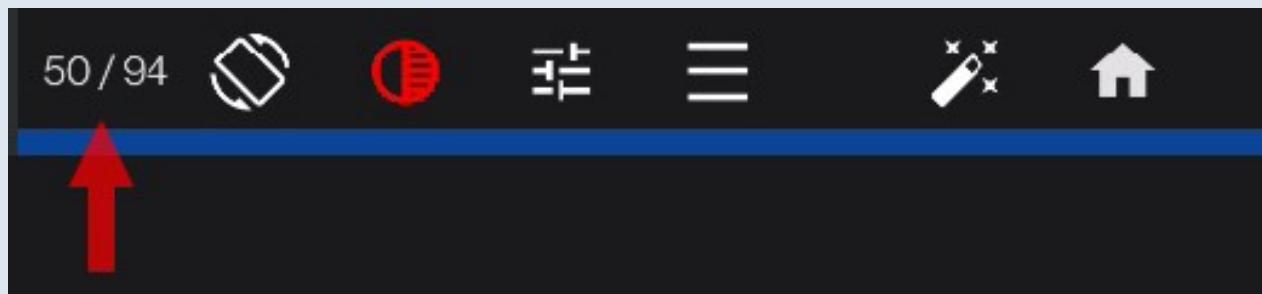
(Voir ci-dessous le fonctionnement de ces curseurs)

Lors de la fermeture du document, les valeurs des curseurs ainsi que la page en cours sont sauvegardées, de sorte à retrouver la même situation lors de la réouverture du document.

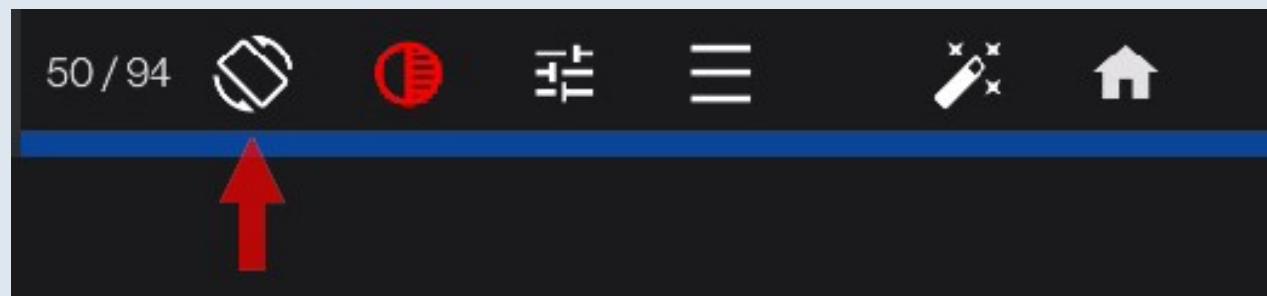
La barre d'outils.

Voyons en quoi cela consiste, de gauche à droite.

Si l'on clique sur l'indication de la page en cours, il sera possible de sauter vers une autre page...

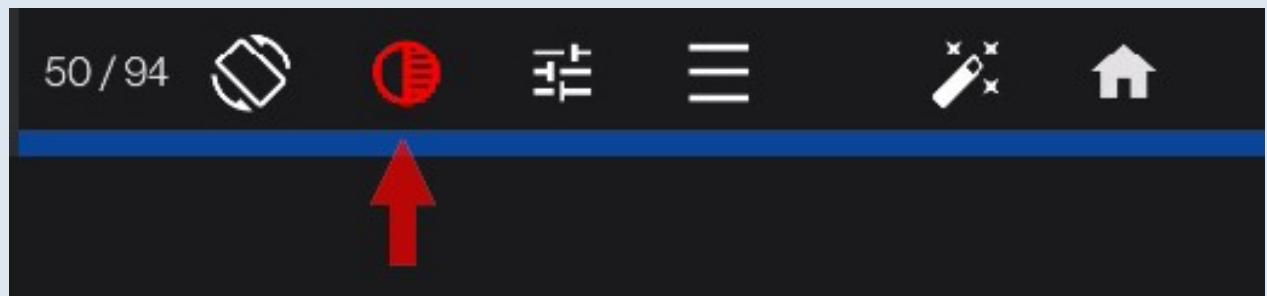


Si l'on clique sur l'icône de renversement, l'écran effectuera une rotation de 180°.



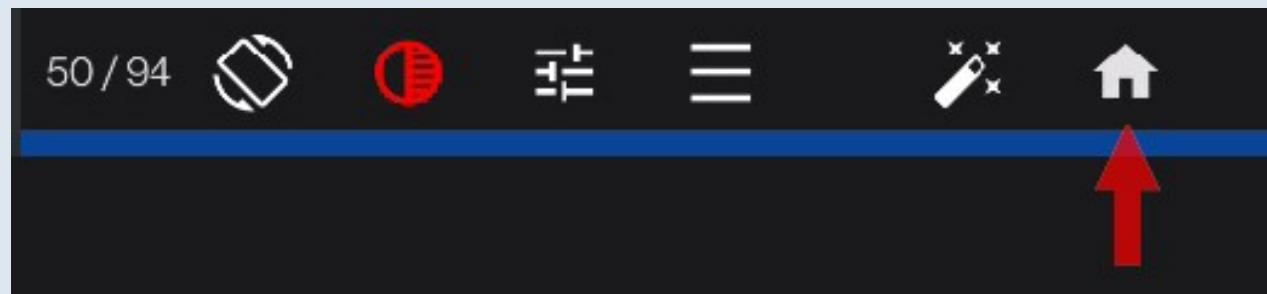
Je l'utilise pour garder le câble d'alimentation connecté. Il me gêne moins lorsque le connecteur est en haut plutôt qu'en bas. 😊

Augmente le contraste.



Par défaut, le contraste est augmenté pour améliorer la détection du texte. Il est possible de le désactiver si les résultats de la détection du texte ne sont pas satisfaisants.

Accueil.

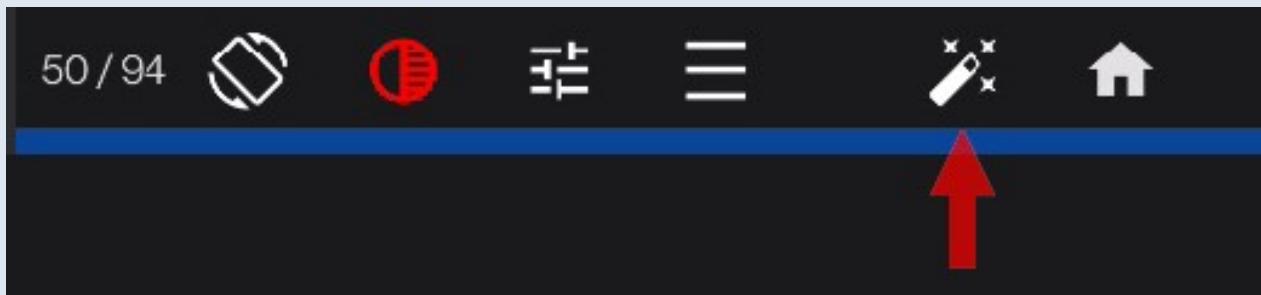


Retour à la page d'accueil pour changer de pdf ou quitter.

Remarque :

Les paramètres du PDF sont sauvegardés à ce moment précis. Il est donc préférable d'utiliser cette option plutôt que de réduire le programme.

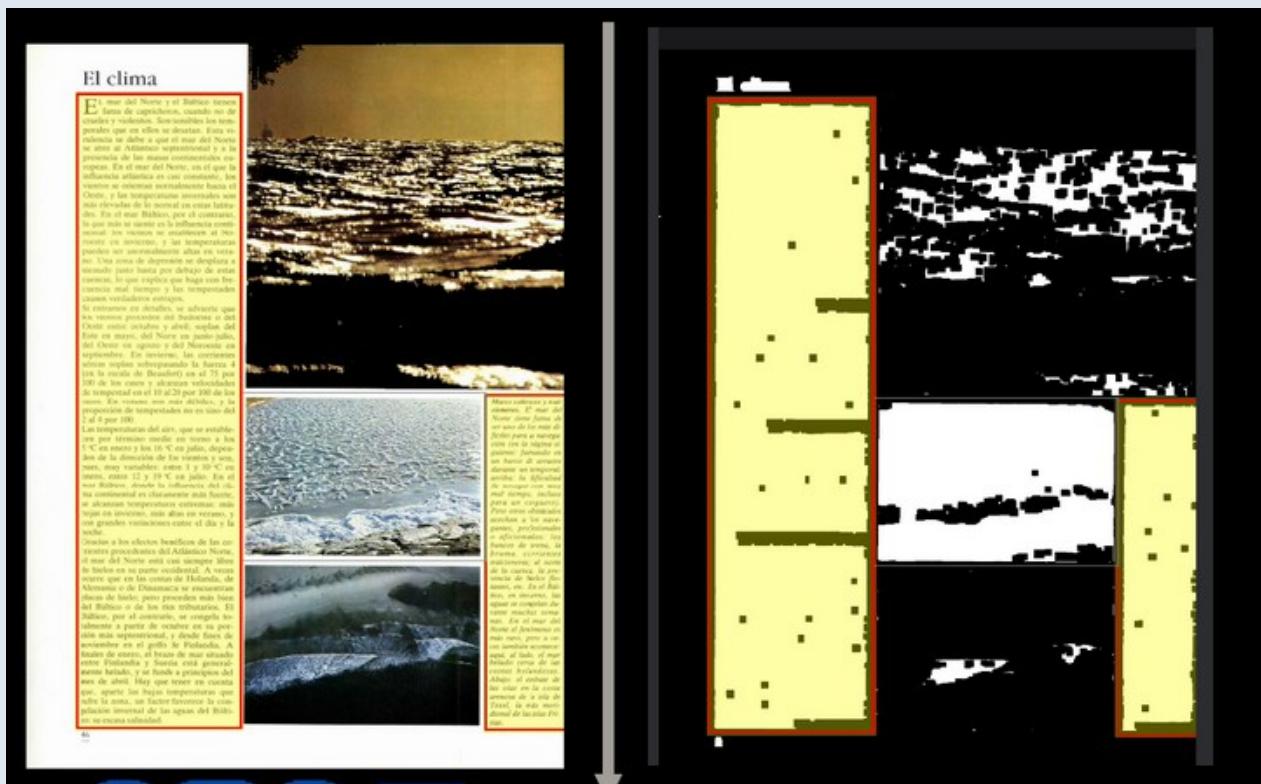
Vue de travail.



C'est à partir de maintenant que les choses "sérieuses" vont commencer.

Pour retrouver les zones de texte sur les pages, le programme transforme la page en noir et blanc en accentuant les détails.

Avec cette option vous pouvez basculer d'une vue à l'autre.

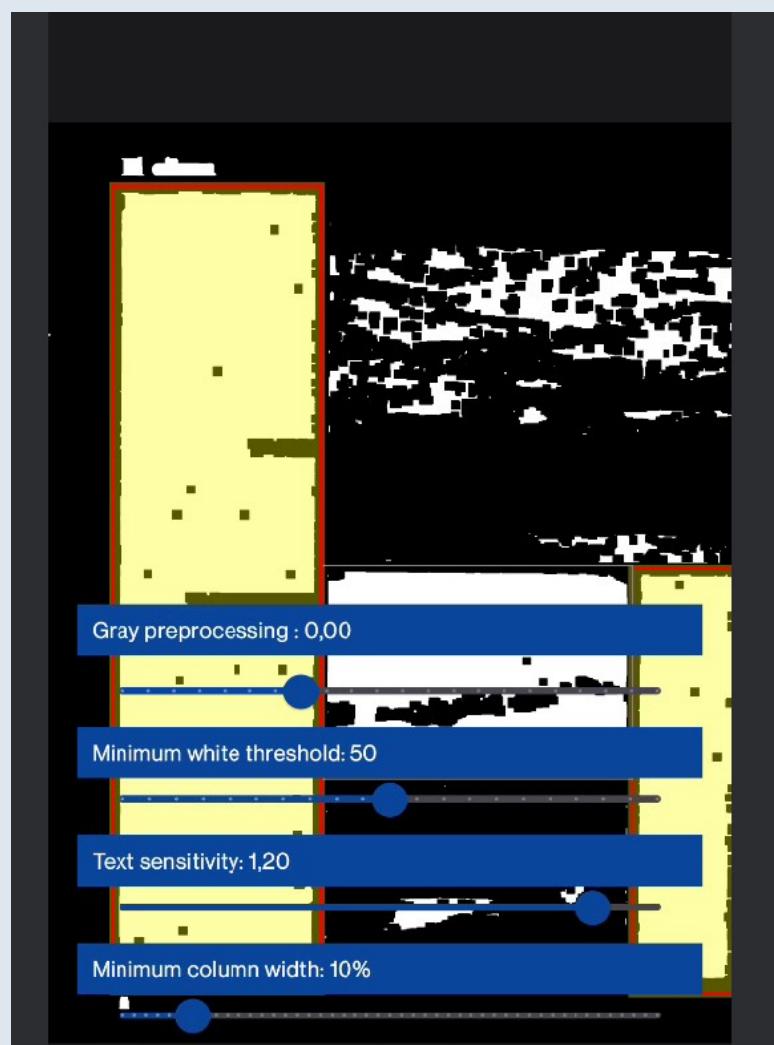


Çà aide pour les réglages qui vont suivre...

Les curseurs.



Ce bouton va superposer des curseurs sur notre page.



Ce sont ces curseurs que l'on devra ajuster en fonction de la qualité du PDF. Les modifications sont appliquées en temps réel et le résultat est immédiatement visible sur la page (sous les curseurs).

Passons-les en revue rapidement.

- **Gray preprocessing** (Prétraitement en niveaux de gris)

Ajuste le niveau de gris pour optimiser la détection du texte.

- **Minimum white threshold** (Seuil minimum de blanc)

Cette valeur est exprimée en %. À l'intérieur de nos rectangles rouges, il s'agit du % minimum de blanc pour que ceux-ci soient acceptés.

- **Text sensitivity** (Sensibilité du texte)

Il s'agit de la sensibilité pour détecter le texte à l'intérieur des carrés rouges.

- **Minimum column width** (Largeur minimale des rectangles)

Il s'agit de la largeur minimale que doivent avoir les rectangles, au-delà de laquelle ils sont ignorés.

Chaque document PDF est un "cas particulier", je ne peux donc pas vous donner de méthode précise pour régler les curseurs. Toutefois, il suffit d'expérimenter un peu et d'observer le résultat en temps réel pour apprendre à les utiliser.

Contrairement à ce qu'on pourrait penser, ce n'est pas si compliqué !

En principe, une fois réglés pour une page du PDF, les paramètres devraient être valides pour le reste des pages, avec parfois quelques ajustements mineurs.

Retenez qu'en quittant la vue avec le bouton d'accueil, tous les réglages sont mémorisés, et ce, pour chaque PDF de manière individuelle.

Options :



Il y a quelques autres options supplémentaires.

The screenshot shows a document viewer interface with several options displayed as colored bars:

- High-resolution PDF (scaleFactor 1.5)**: A red bar containing the text "High-resolution PDF (scaleFactor 1.5)".
- Gray preprocessing for TTS : 0,00**: A red bar containing the text "Gray preprocessing for TTS : 0,00".
- Frame margin : 12 px**: A green bar containing the text "Frame margin : 12 px".
- Reading speed : 1,00x**: A green bar containing the text "Reading speed : 1,00x".

Below these bars is a horizontal slider with a blue circular handle. At the bottom of the screen are four blue rounded rectangular buttons labeled "<", "TTS", ">", and "sel. frames".

The main content area displays a page from a document. The left side of the page contains text in Spanish. The right side features a photograph of a turbulent sea with large waves. Above the photograph, there is some descriptive text in Spanish.

Lorsque nous avons un PDF d'un document scanné dont le texte est mal résolu, on peut essayer ces 2 options :

Cocher la case :

High-resolution PDF (scaleFactor 1.5)

Cela augmente la résolution des pages pour tenter de mieux lire le texte. Cependant, attention, car cela ralentit également le programme.

Idem pour le curseur : **Gray preprocessing for TTS**
(Prétraitemet en niveaux de gris pour le TTS)

Remarque : Ce réglage ne s'applique que lors de l'OCR du texte dans les rectangles.

Il n'a donc rien à voir avec les curseurs précédents, qui servent à détecter où se trouvent les textes dans les pages.

Utilisez ces deux options en dernier recours si vous avez du mal à obtenir un texte correct.

Frame margin (Marge des cadres)

C'est simplement une option pour le fun ; elle dessine les cadres rouges un peu plus petits ou plus grands autour des textes.

Reading speed (Vitesse de lecture)

Ici, vous pouvez ajuster la vitesse de lecture de la voix de synthèse. Celle-ci est également sauvegardée et dépend de la langue.

Suite du document en cours de réalisation...