



Tiff to Text is designed to perform Optical Character Recognition (OCR) in a batch process. The program utilizes the OCR engine from Nuance (Owners of OMNI Page - formally ScanSoft) that is included with Microsoft Office Document Imaging (MODI).

With Tiff to Text the user has the option of processing all of the tif images in a file folder as well as tif images in all subfolders. The output will be a matching file folder structure with either the tif image along with a matching text file created from the OCR or just the text file.

Tiff to Text is easy to setup as the user only has to enter the input folder and output folder along with some simple choices.

## **Program uses:**

This program is designed to produce high quality OCR text from tif images for research purposes. It is ideal for law firms that receive CD's of images in the discovery process. An average CD contains up to 10,000 images and they can be processed at an average of 900 pages per hour. This average was calculated on typical single page discovery document in a G4 tiff format. The processor was a 2.4 mhz Pentium 4 with 512 of memory. Ideally the process will be run at night when the PC is not being used for another process as the OCR engine consumes most of the processing power of the PC.

Other uses for this product are all related to research projects where the quality of OCR text is important but where the formatting is not.

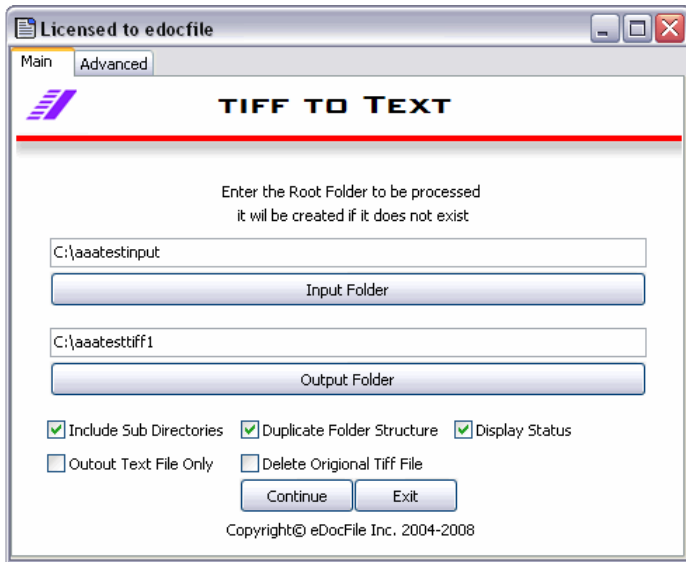
## **Scanning Images:**

This program does not include scanning software. It only processes images after they have been scanned. It assumes the images were created with a copier that scans to a file folder on a network such as those manufactured by Canon, Ricoh, Xerox, Sharp, HP, Kyocera, Konica, Toshiba, Brother, Lanier, Savin, Gestetner, Panasonic, HP, Oce, Kodak or other that has the ability to scan to a file folder on a network. It also processes tif images that were received in the discovery process or were batch scanned by other software.

## **Requirements:**

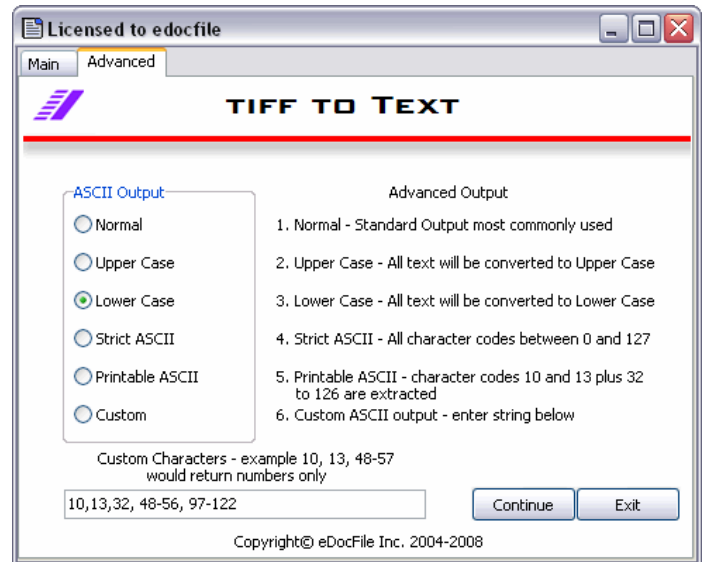
Microsoft Office Document Imaging

## Required Setup information:



- Include Sub Directories – this will OCR all the tif images in the Sub Directories of the root
- Duplicate Folder Structure – this makes a matching output file folder hierarchy of the OCR'd images, if not selected all output files will be placed in the output root.
- Display Status – allows the user to see what file is being OCR'd
- Output Text File Only – just creates a matching text file from the OCR contents
- Delete Original Tiff image – Deletes the file that was processed

## Advanced Options:



- Standard Output – Normal OCR text file
- Upper Case – All of the OCR text is converted to Upper Case
- Lower Case - All of the OCR text is converted to Lower Case
- Strict ASCII – Only outputs the OCR text that contains character codes between 0 – 127
- Printable ASCII – Only outputs the OCR text that contains character codes 10, (line feed), 13 (Carriage Return) and codes 32 to 126
- Custom - the user can input a string of character codes to be returned from the OCR such as 10, 13, 48-57 which would just output the numbers contained within the tif image.

Tiff to Text \$199.00  
Order through [www.edocfile.com](http://www.edocfile.com)

Price subject to change without notice.

For More Information visit [www.eDocFile.com](http://www.eDocFile.com)