

Manuale OCR Anastasis



Strumenti per l'apprendimento rapido e autonomo



INDICE

Il Programma.....	3
Installazione	3
La finestra DI AVVIO	4
Salvare il documento acquisito	6
MODALITA' AVANZATA	7
Menu File	8
MENU HOME.....	9
MENU MODIFICA	10
MENU SCANNER.....	12
Modalita' di acquisizione	12
MENU VISUALIZZA	13
La finestra Pagine	14
Finestra Immagine	15
Finestra testo	16



IL PROGRAMMA

L'introduzione dei libri digitali in formato PDF ha agevolato l'utilizzo dei software compensativi con sintesi vocale, ma non tutti i materiali e documenti esistono già in formato digitale. Si pone quindi il problema di come trasferire un testo cartaceo in formato digitale.

Per convertire i testi cartacei in digitali occorre dotarsi di uno scanner e di un software OCR.

Il processo di trasformazione da documento cartaceo a contenuto digitale avviene in due fasi:

1. Lo scanner acquisisce il testo cartaceo attraverso un sensore che "copia" la pagina e la trasforma in un'immagine digitale che viene inviata al computer.
2. Il software di OCR legge l'immagine digitale, riconosce i caratteri e trasforma l'immagine in testo digitale che è possibile salvare nei diversi formati possibili (p.es. TXT, RTF o PDF).
OCR è l'acronimo di "**Optical Character Recognition**", ovvero Riconoscimento Ottico dei Caratteri. L'OCR trasforma un testo su carta in un testo digitale.

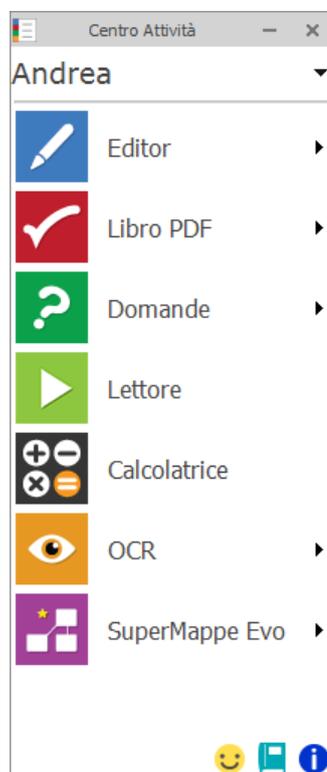
Di norma insieme allo scanner viene fornito gratuitamente un software di OCR, che però raramente è di qualità sufficiente a produrre testi idonei alla lettura.

L'utilizzo di un programma professionale permette di avere una percentuale di errori di riconoscimento estremamente bassa ed è molto più veloce.

INSTALLAZIONE



L'installazione del modulo OCR crea un'icona sul Desktop che serve per avviare il Centro Attività. Dopo l'installazione l'icona OCR sarà presente anche dentro il Centro Attività.

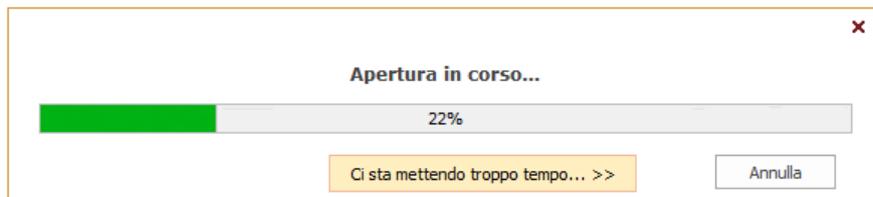




LA FINESTRA DI AVVIO



La scansione, l'apertura e l'elaborazione di file con molte pagine potrebbe richiedere del tempo quindi consigliamo di frammentare il documento in più file. In tutte le finestre di elaborazione è presente un bottone [Ci sta mettendo troppo tempo ... >>] che apre il pannello con le indicazioni da seguire per documenti molto grandi.

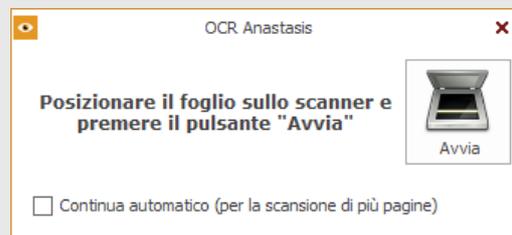


Nella finestra di avvio sono presenti i comandi:



Scansione

Avvia la scansione del documento presente nello scanner collegato al computer, con le impostazioni definite nella finestra **Avanzate**. Se il computer non è collegato ad uno scanner l'icona di questo comando è disabilitata.



Una volta posizionato il documento sullo scanner fare clic su pulsante **Avvia**.

NOTA BENE Se l'opzione "Continua automatico" è selezionata, ad ogni fine pagina riparte la scansione in modo automatico dopo un breve intervallo in secondi definito nelle impostazioni avanzate di scansione; quando è stata scansionata l'ultima pagina del documento si preme il tasto "Fine".



Elabora documento

Aprire la finestra in cui è possibile scegliere il file che si vuole importare. E' possibile aprire file .pdf, .csp e i file immagine (.jpg, .jpeg, .png, .tiff, .bmp, .gif). Terminata l'operazione di riconoscimento del file, il programma propone la finestra in cui scegliere in quale formato vogliamo salvare il documento. Le opzioni disponibili sono:





Una volta scelto il formato possiamo salvare facendo clic sul pulsante **Salva** o accedere alla modalità avanzata facendo clic sul pulsante **Vai ad Avanzate**.

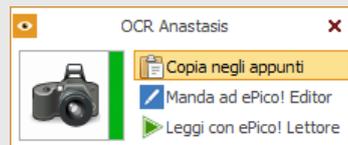
ATTENZIONE! L'apertura di file con molte pagine potrebbe richiedere del tempo quindi consigliamo di frammentare il documento in più file.



Cattura da schermo

Questo comando serve per “cattare” un testo presente sullo schermo e renderlo disponibile come testo editabile.

Nella finestra che appare, per prima cosa dobbiamo definire cosa fare con il testo che stiamo per fotografare.



Le opzioni sono:

- **Copia negli appunti:** mette il testo nella clipboard di Windows e lo rende disponibile per essere incollato in qualsiasi applicazione;
- **Manda ad ePico! Editor:** apre un nuovo documento nell'Editor di ePico!;
- **Leggi con ePico! Lettore:** il testo viene letto dalla sintesi vocale del modulo lettore.

Una volta definita la destinazione facciamo clic sul pulsante **Cattura da schermo**  presente nella finestra. Appare un selettore formato da una croce rossa, posizioniamo il selettore all'inizio della parte di schermo da “cattare”, teniamo premuto il tasto sinistro del mouse e trasciniamo il selettore fino a che non si è definita l'area desiderata.

A questo punto parte il riconoscimento del testo contenuto nell'area (il cui progresso è visibile nella barra verticale accanto al pulsante).

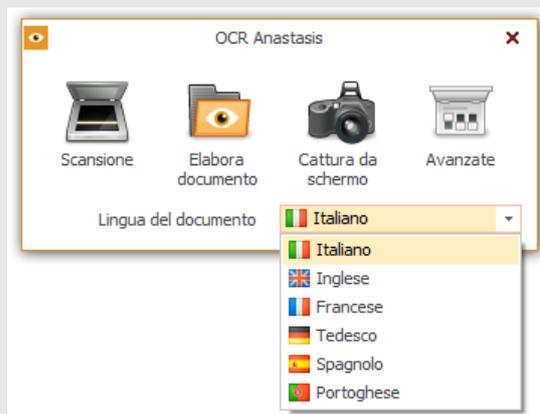


Avanzate

Aprire la finestra in cui è possibile definire le impostazioni di scansione e riconoscimento del documento ([vd. sezione MODALITA' AVANZATA](#)).

Lingua del documento

Aprire un menu a tendina in cui scegliere la lingua del documento che si vuole acquisire. Questa impostazione non agisce sulla lingua di lettura; per leggere in altre lingue è necessario impostare la lingua di lettura direttamente nei vari moduli di ePico!





SALVARE IL DOCUMENTO ACQUISITO

Dopo la fase di acquisizione, il programma procede con il riconoscimento del testo e il documento è pronto per essere salvato.

La finestra che appare propone i formati disponibili tra cui scegliere per salvare il documento.

Le opzioni disponibili sono:



- CSP (scelta di default): invia il testo acquisito al modulo PDF di ePico!
- CSD: invia il testo acquisito al modulo Editor di ePico!
- RTF: salva il documento in un file di testo .rtf
- PDF: salva il documento in un file .pdf
- DOCX: salva il documento in un file .docx di Microsoft Word
- EPUB: salva il documento in un file .epub per i libri digitali

Una volta definito il formato, confermare con il pulsante **Salva**.

Se non cambiamo impostazioni il documento viene salvato di default nella cartella utente **Documenti**.

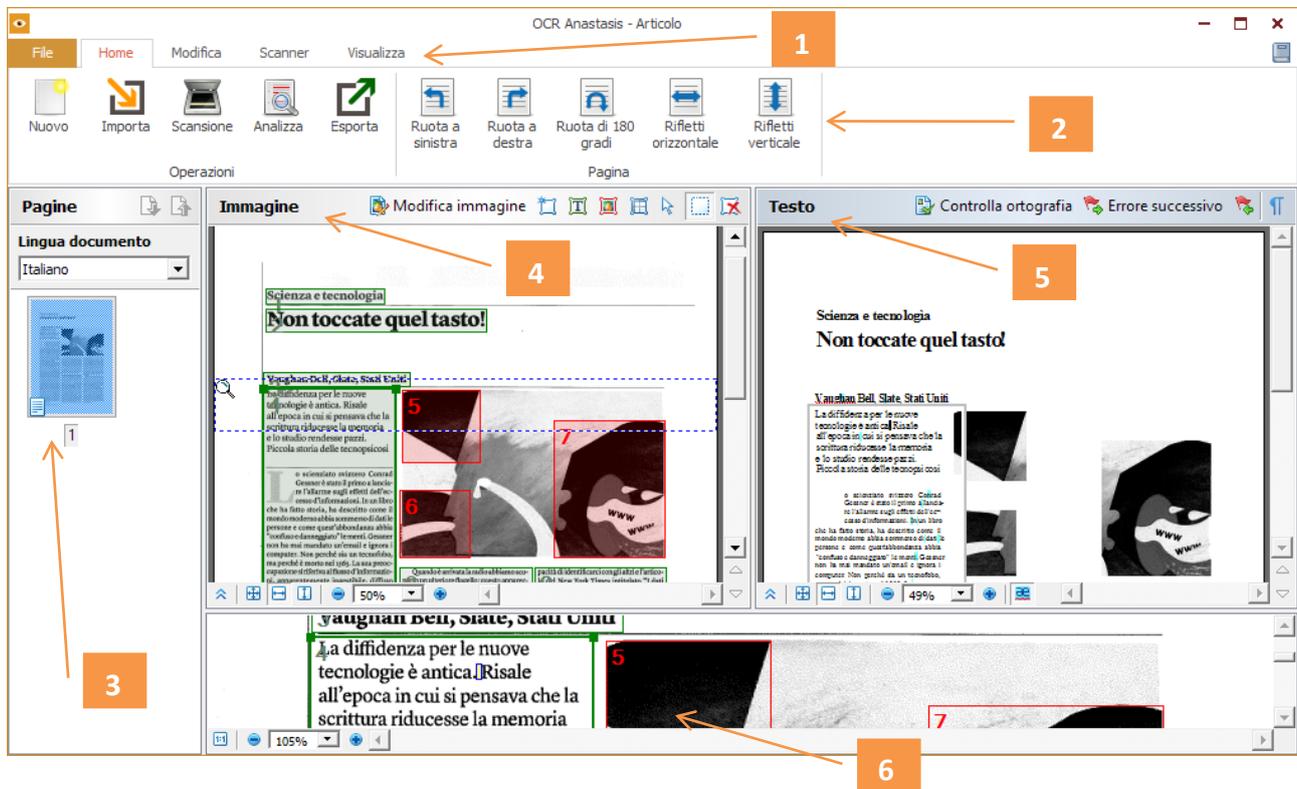
NOTA BENE Cliccando sul segno di spunta “Aprire il file al termine” il documento esportato verrà automaticamente aperto con il programma predefinito.

Il pulsante **Vai ad Avanzate** passa alla modalità avanzata dove sono disponibili tutti i comandi per modificare le impostazioni del documento.



MODALITA' AVANZATA

Selezionando il comando **Modalità avanzata** dal pannello di avvio o in fase di salvataggio del documento, si apre la finestra con i comandi per la gestione delle impostazioni del documento.



La videata principale è composta da:

1. la barra dei menu o barra multifunzione
2. la barra dei comandi relativi al menu attivo
3. la finestra **Pagine** dove sono presenti tutte le pagine del documento corrente. Sono disponibili due modalità di visualizzazione dell'elenco delle pagine: una è rappresentata dalle icone (come nell'immagine sopra – **Anteprime**), l'altra tramite un elenco delle pagine del documento e delle relative proprietà (**Dettagli**). È possibile passare da una modalità all'altra utilizzando il menu di scelta rapida della finestra (tasto destro del mouse) e scegliendo il comando "Visualizzazione finestra pagine" e poi Anteprime o Dettagli.
4. la finestra **Immagine** dove è visualizzata l'immagine della pagina corrente. In questa finestra sono presenti i comandi per la definizione delle aree dell'immagine.
5. la finestra **Testo** dove è possibile visualizzare il testo riconosciuto. Nella finestra Testo è possibile controllare l'ortografia e il formato, oltre che modificare il testo riconosciuto.
6. la finestra **Zoom** dove viene visualizzata l'immagine ingrandita della riga o della sezione dell'immagine elaborata che si sta modificando. Mentre nella finestra Immagine è disponibile una vista complessiva della pagina, nella finestra Zoom è possibile osservare l'immagine in modo più dettagliato, regolare il tipo di area e la sua posizione o confrontare i caratteri incerti con l'immagine ingrandita. Nella finestra Zoom è possibile modificare le proporzioni dell'immagine, tramite il riquadro disponibile nella parte inferiore della finestra.

Vediamo in dettaglio i menu presenti in alto sulla finestra: File, Modifica, Scanner e Visualizza.



MENU FILE

In questo menu sono presenti tutti i comandi per la gestione dei file del programma.

 **Nuovo** Crea un nuovo file vuoto.

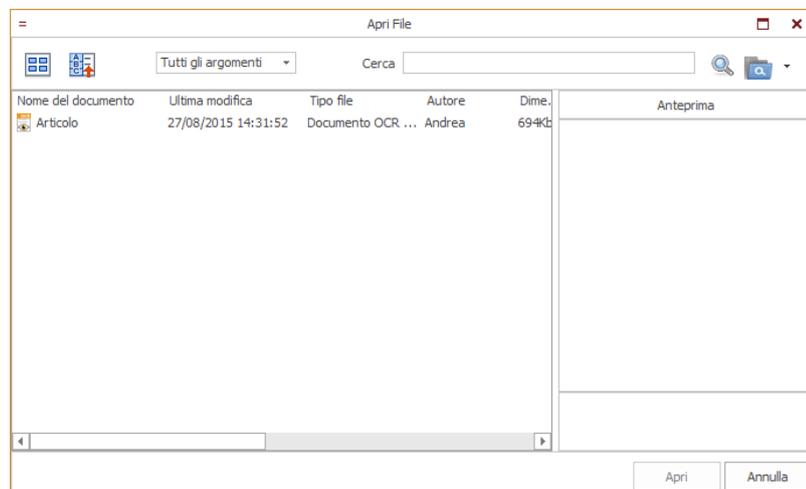
 **Apri** Importa un file già esistente. Il modulo OCR può importare file:

- **.cso**: formato proprio dell'OCR Anastasis
- **.pdf**
- **file immagine** (.jpeg, .jpg, .png, .bmp, .tif ...)

Per aprire un documento è sufficiente selezionarlo e fare clic sul pulsante **Apri**, oppure fare doppio clic sull'icona del file.

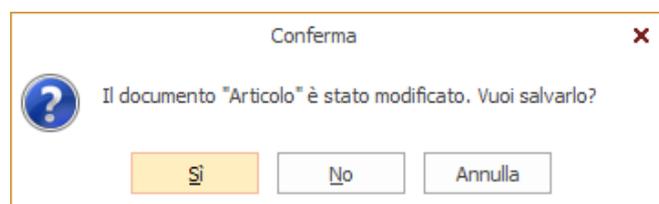
Se invece dobbiamo aprire un file che non è presente tra quelli nella finestra, dobbiamo fare clic sul pulsante **Sfogli**.

Nella finestra che appare selezionare l'unità di memoria, la cartella e il file che vogliamo aprire. Infine, confermiamo facendo clic sul pulsante **Apri**.



Recenti Il riquadro Recenti, visualizza la lista degli ultimi documenti aperti. Per aprire un documento è sufficiente fare clic sull'icona del file.

 **Chiudi** Chiude l'ultimo documento che abbiamo aperto. Se il documento è stato modificato e non è stato salvato, appare un messaggio in cui viene richiesto se vogliamo o meno salvare le modifiche.

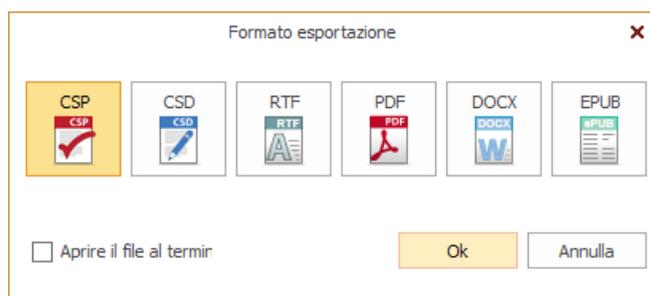


 **Salva** Salva le modifiche apportate al documento.

Se si tratta del primo salvataggio il programma richiede il nome da assegnare al file modificato. Il programma assegna al documento il proprio formato (.cso) Il documento viene salvato nella cartella **Documenti** dell'utente.



-  **Salva con nome** Salva il documento con un nome diverso da quello assegnato in precedenza.
-  **Esporta** Quando apportiamo delle modifiche al documento e poi lo salviamo, il programma assegna al file il proprio formato .cso. Questo comando permette di salvare il documento con un formato, un nome, e di decidere una destinazione, diversi da quelli assegnati in fase di salvataggio e propone di default la cartella **Scansioni** nella cartella Utente.
- Esporta è il comando che permette di fare una "copia" del documento in un altro formato e in un'altra destinazione.

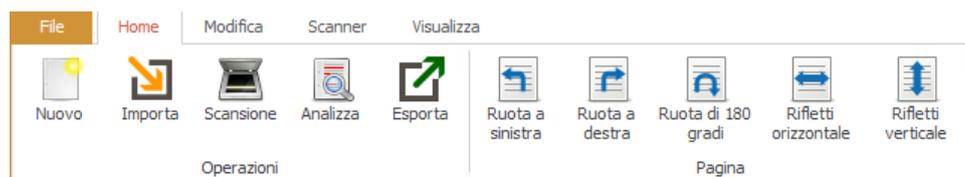


Informazioni Apre la finestra che contiene le informazioni del programma che si riferiscono a versione, copyright e licenza.

-  **Esci** Esce dal modulo OCR, chiede conferma per il salvataggio del documento aperto.

MENU HOME

Nel menu Home sono presenti i comandi per la gestione del documento.



Gruppo Operazioni

I comandi presenti in questa area sono:

-  **Nuovo** Crea un nuovo file vuoto.
-  **Importa** Aggiunge le pagine del file importato alla fine del documento corrente.
ATTENZIONE! L'importazione di file con molte pagine potrebbe richiedere del tempo quindi consigliamo di frammentare il documento in più file.
-  **Scansione** Avvia una nuova scansione. Uno scanner deve essere collegato al PC.
-  **Analizza** Esegue nuovamente l'analisi del documento corrente.
-  **Esporta** Apre la finestra di esportazione.



Gruppo Pagina

	Ruota sx	Ruota 90° a sinistra la pagina corrente.
	Ruota dx	Ruota 90° a destra la pagina corrente.
	Ruota 180	Ruota di 180° la pagina corrente
	Rifletti orizzontale	Riflette orizzontalmente la pagina corrente.
	Rifletti verticale	Riflette verticalmente la pagina corrente.

MENU MODIFICA



I comandi presenti in questo menu sono:

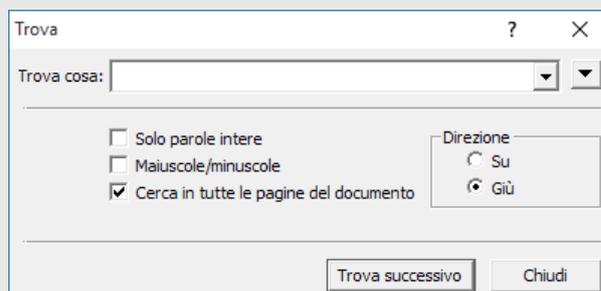
Gruppo Annulla

	Annulla	Annulla l'ultima operazione effettuata
	Ripristina	Ripristina l'ultima operazione annullata.

Gruppo Ricerca

I comandi presenti in questa scheda permettono di cercare parole/frasi all'interno del testo ed eventualmente di sostituirli in modo automatico.

Trova Apre la finestra in cui inserire un testo da ricercare all'interno della pagina attiva.



Per cercare una parola, scriverla nella casella **Trova cosa:** e fare clic su **Trova successivo**. Il programma evidenzia nel testo la prima corrispondenza trovata. Per procedere nella ricerca cliccare nuovamente su **Trova successivo**.



Le opzioni disponibili sono:

- **Solo parole intere:** quando l'opzione non è selezionata e noi inseriamo solo una parte della parola che cerchiamo, il programma propone tutte le parole che contengono le lettere da noi indicate. Se l'opzione è attiva il programma troverà solo il testo esattamente uguale a ciò che noi abbiamo scritto e solo se è una parola intera e non parte di essa.
- **Maiuscole/minuscole:** quando l'opzione è selezionata, il programma tiene conto delle lettere Maiuscole e minuscole del testo che abbiamo inserito nella casella Trova e proporrà solo le corrispondenze identiche.
- **Cerca in tutte le pagine del documento:** permette di circoscrivere la ricerca solo sulla pagina corrente o su tutte le pagine del documento.

Direzione Su/giù: definisce da dove comincia la ricerca rispetto alla posizione del cursore nel testo.



Trova successivo

Dopo aver inserito una parola o una stringa di testo nella finestra **Trova**, il comando **Trova successivo**, evidenzia nel testo della pagina attiva, la successiva corrispondenza al testo cercato.



Sostituisci

Apri la finestra **Sostituisci**.

Questo comando è particolarmente utile quando dobbiamo cercare una parola presente nel documento e la vogliamo sostituire con un altro testo.

Inseriamo nella casella **Trova cosa** la parola o la frase da sostituire e nella casella **Sostituisci con** inseriamo il nuovo testo.

Il programma propone la prima corrispondenza trovata e facendo clic sul pulsante **Sostituisci**, il testo viene sostituito dal testo presente nella casella Sostituisci con.

Il programma procede con le successive corrispondenze e quando non vogliamo sostituire facciamo clic sul pulsante **Trova successivo**.

Il comando **Sostituisci tutto** permette di sostituire tutte le corrispondenze trovate senza dover ogni volta dare la conferma.

Le opzioni presenti sono le stesse del comando Trova.



Vai alla pagina

Apri la finestra in cui possiamo indicare il numero di pagina che vogliamo visualizzare.



MENU SCANNER

In questo menu sono presenti i comandi per regolare più finemente i parametri di scansione. Il menu Scanner è attivo solo se è stato installato un driver di scansione.



Vediamo i comandi in dettaglio:

Scanner	E' possibile selezionare lo scanner che vogliamo utilizzare per la nostra scansione (utile se abbiamo più periferiche di scansione collegate al pc).
 Bianco/nero	Imposta il parametro di acquisizione dell'immagine in modalità bianco e nero.
 Scala di grigi	Imposta il parametro di acquisizione dell'immagine in modalità scala di grigi.
 Colore	Imposta il parametro di acquisizione dell'immagine in modalità Colore.
Risoluzione	Imposta il parametro che determina il numero di dpi da utilizzare durante la digitalizzazione. La risoluzione a 300 dpi è consigliata per testi con dimensioni carattere pari a 10 punti o superiori, mentre quella a 400 – 600 dpi è consigliata per testi con dimensioni carattere più ridotte (9 punti e inferiori).
Luminosità	Imposta il parametro di digitalizzazione che riflette il contrasto fra aree nere e aree bianche dell'immagine. Una corretta impostazione della luminosità migliora la qualità del riconoscimento.
Continua automatico	Quando l'opzione è attiva, lo scanner, dopo aver effettuato una scansione, non si ferma, ma attende il tempo definito nella casella Pausa (sec.) e riparte con un'altra scansione. Questa opzione è molto utile quando si devono digitalizzare più pagine.
Pausa (sec.)	Imposta l'intervallo di tempo che passa per la scansione automatica successiva.

MODALITA' DI ACQUISIZIONE

Ecco alcune indicazioni per la digitalizzazione di documenti cartacei.

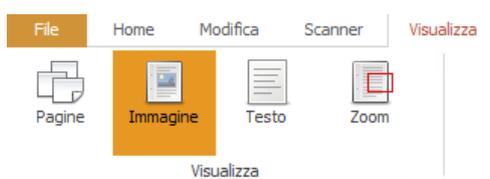
- Assicurarsi che lo scanner sia acceso e collegato.
- Per assicurarsi che lo scanner sia configurato correttamente, consultare la relativa documentazione. Accertarsi di installare il software fornito con lo scanner. Alcuni modelli di scanner devono essere accesi prima di accendere il computer.
- Inserire la pagina da digitalizzare nello scanner. Se lo scanner è munito di alimentatore automatico, è possibile inserire più pagine.
- Avviare la scansione del documento.

! Suggerimento: Per i documenti stampati tipici, la modalità di digitalizzazione ottimale è con la risoluzione di 300 dpi.

La qualità del documento di origine e le opzioni di digitalizzazione selezionate influiscono sulla qualità del riconoscimento. Una qualità scadente dell'immagine può influire negativamente sui risultati di riconoscimento. Per questo motivo è importante personalizzare le impostazioni di digitalizzazione in modo appropriato e verificare le caratteristiche del documento di origine.

MENU VISUALIZZA

In questo menu sono presenti i comandi per visualizzare o nascondere le finestre del programma.



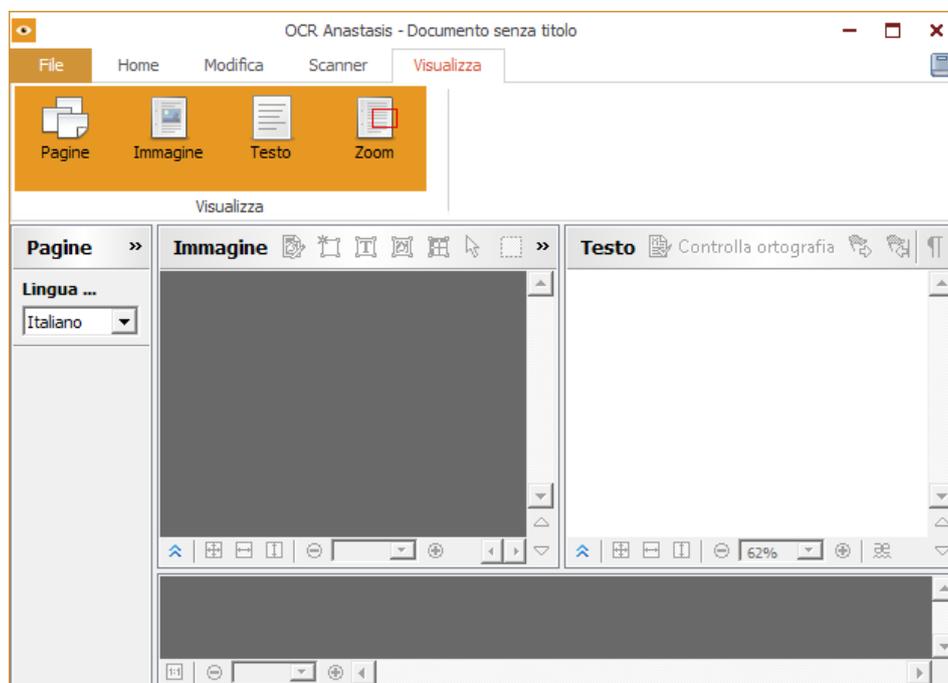
 **Pagine** Attiva/disattiva la finestra Pagine.

 **Immagine** Attiva/disattiva la finestra Immagine.

 **Testo** Attiva/disattiva la finestra Testo.

 **Zoom** Attiva/disattiva la finestra Zoom.

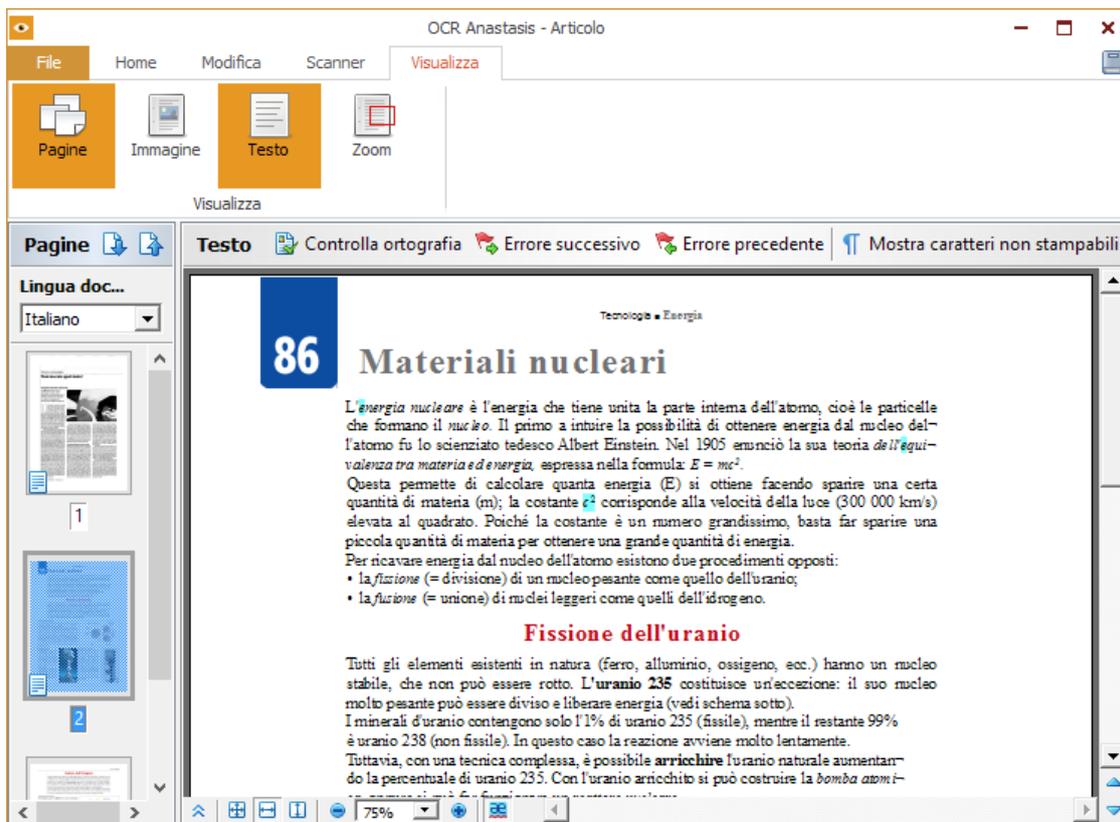
Se tutti i pulsanti sono attivati la finestra del programma è la seguente:





LA FINESTRA PAGINE

Nella finestra Pagine, sono presenti tutte le pagine che compongono il documento attivo.



Spostarsi nelle pagine

Per posizionarsi su una pagina basta fare clic sulla sua icona oppure fare clic sui pulsanti  **Pagina precedente** e  **Pagina successiva**.

Modificare l'ordine delle pagine

E' possibile modificare l'ordine delle pagine facendo clic sulla pagina da spostare tenere premuto il tasto sinistro del mouse e trascinare la pagina nella nuova posizione. E' anche possibile modificare l'ordine scrivendo nella casella del numero di pagina, il numero corrispondente alla posizione in cui vogliamo spostarla.

Lingua documento

Apri il menu a tendina in cui indicare in che lingua è scritto il documento attivo. La corretta selezione della lingua permette un riconoscimento corretto del testo.

Eliminare pagine

Per eliminare una pagina basta selezionarla e premere il tasto CANC sulla tastiera, oppure aprire il menu contestuale con il tasto destro del mouse e selezionare il comando **Elimina pagina dal documento**.

Quando eliminiamo una pagina non è possibile annullare il comando.



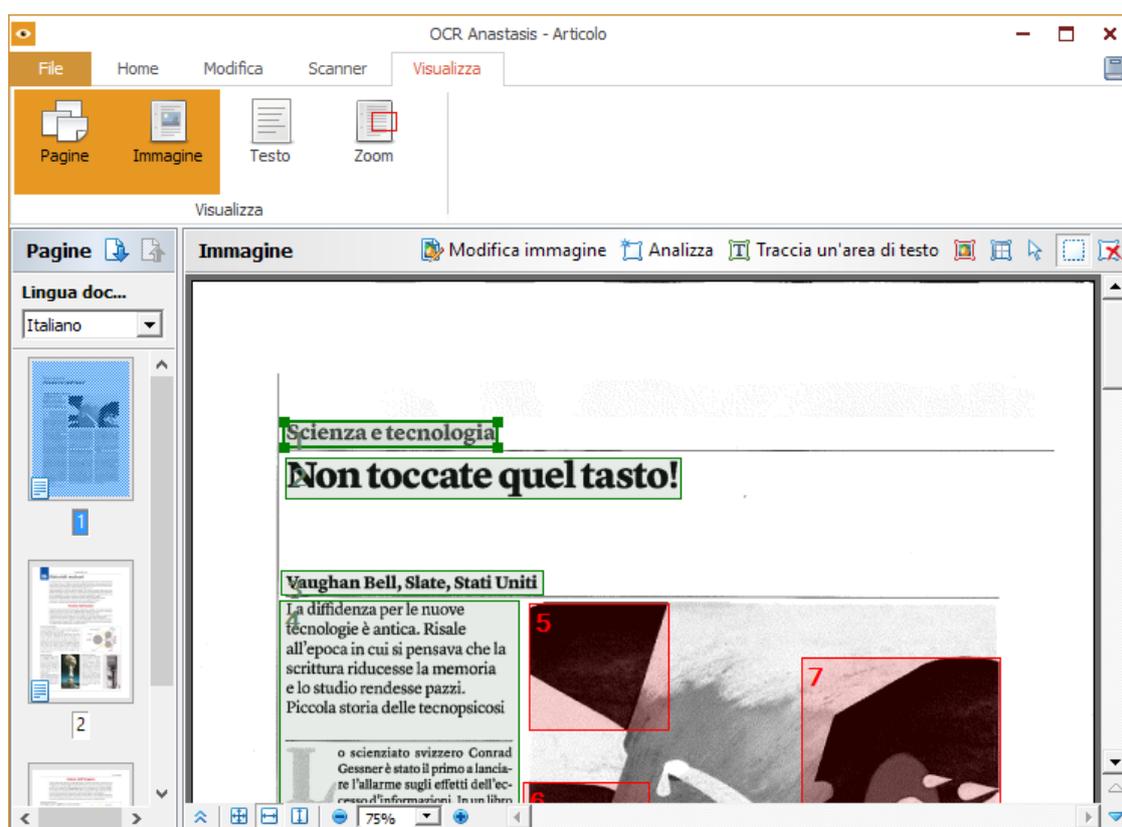
FINESTRA IMMAGINE

In questa finestra è presente l'immagine della pagina selezionata come risultato della scansione. Prima di eseguire la procedura di riconoscimento del documento, il programma analizza la struttura logica di quest'ultimo e rileva le aree contenenti testo, immagini, tabelle e codici a barre.

Queste informazioni sono utili per riprodurre la formattazione del documento originale.

Per impostazione predefinita, è selezionata automaticamente l'opzione di analisi del documento. Tuttavia una formattazione molto complessa potrebbe causare il rilevamento non corretto di alcune aree. La correzione delle sole aree che presentano problemi costituisce un approccio molto più veloce rispetto alla tracciatura manuale di tutte le aree.

I bordi di aree di tipo diverso presentano colori diversi.



È possibile correggere il layout contrassegnando e modificando manualmente le aree utilizzando gli strumenti e i comandi indicati di seguito:

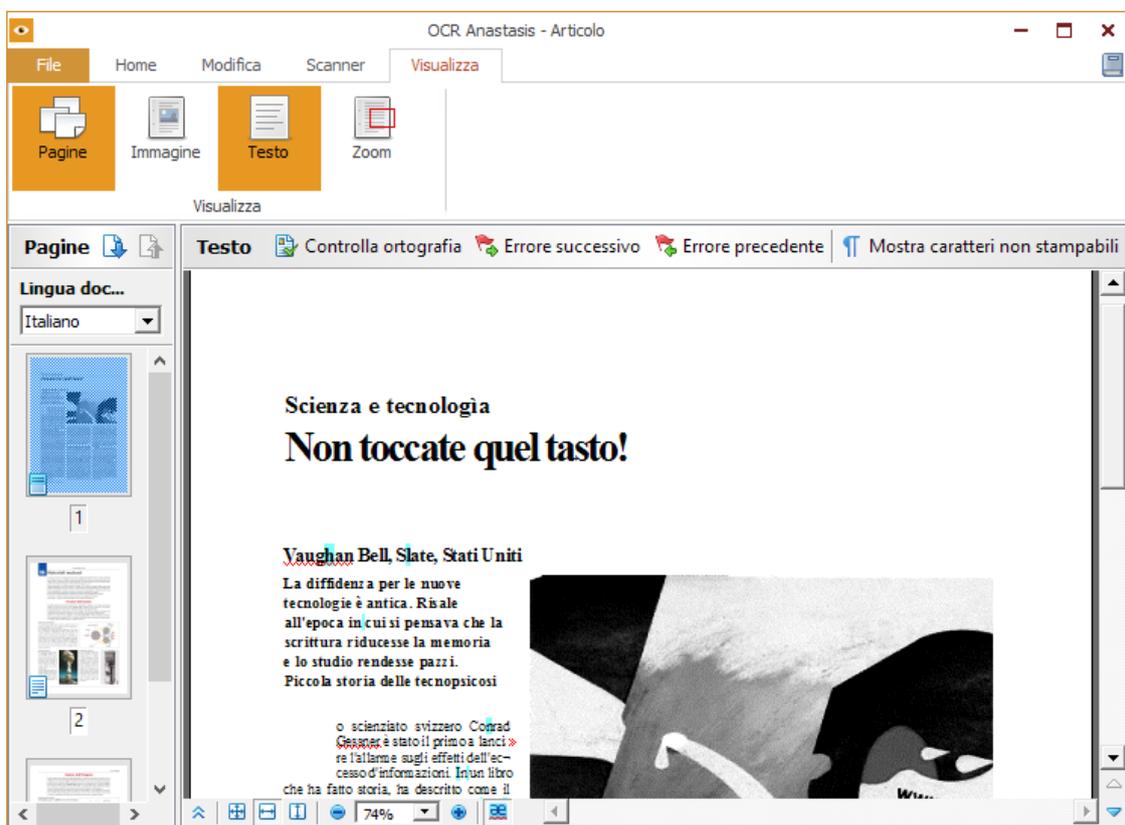
	Modifica immagine	Apre la finestra in cui è possibile modificare le impostazioni dell'immagine.
	Analizza	Esegue il riconoscimento del contenuto della pagina e la suddivide automaticamente in aree di testo, immagine, ecc.
	Traccia un'area di testo	Traccia un'area che definisce la presenza del testo nell'immagine.
	Traccia area di immagine	Traccia un'area che definisce la presenza di un'illustrazione nell'immagine della pagina selezionata. Serve per segnalare che nell'area selezionata è presente un'illustrazione e non deve essere fatto il riconoscimento del testo.



	Traccia un'area tabella	Traccia un'area di tabella nell'immagine della pagina selezionata.
	Seleziona oggetti	Attiva il selezionatore di oggetti.
	Traccia un'area di riconoscimento.	Traccia un'area che racchiude una sezione di immagine da analizzare automaticamente.
	Elimina area	Elimina l'area selezionata.

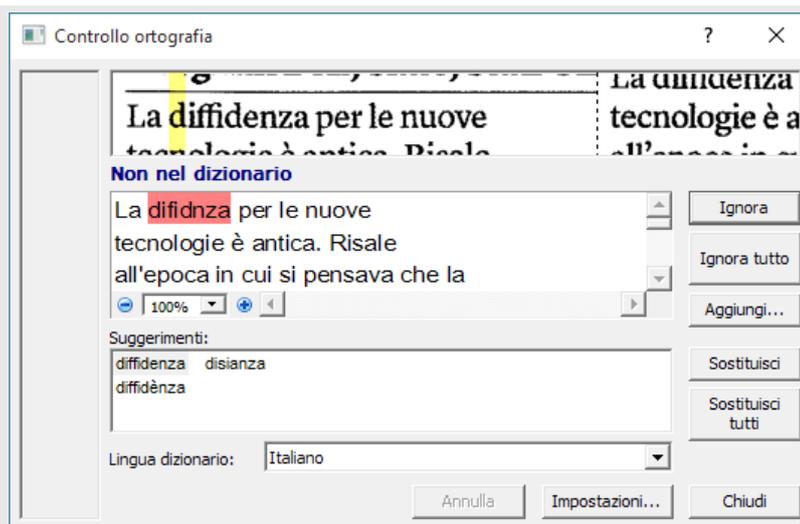
FINESTRA TESTO

In questa finestra è presente il testo riconosciuto dall'analisi dell'immagine presente nella finestra Immagine.



Il testo è a tutti gli effetti un testo digitale e può essere modificato, cancellato o editato esattamente come in un editor di testi. Inoltre è possibile eseguire il controllo ortografico per controllare che in fase di riconoscimento non si siano commessi errori.

	Controllo ortografia	Avvia il controllo ortografico del testo.
---	-----------------------------	---



La parola che non è stata riconosciuta risulta selezionata. Apportare la correzione se individuate un errore e fare clic sul pulsante **Conferma**. E' anche possibile selezionare una delle parole proposte nei Suggerimenti e premere il pulsante **Sostituisci**.



Errore successivo

Individua l'errore successivo.



Errore precedente

Individua l'errore precedente.



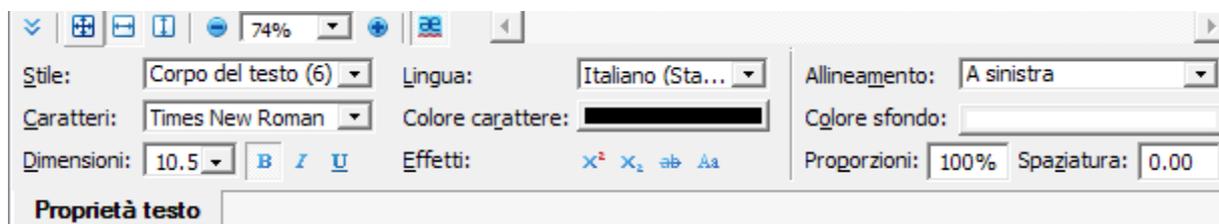
Mostra caratteri nascosti

Mostra/nasconde i caratteri non stampabili del testo.

In basso nella finestra Testo sono presenti i comandi per la formattazione del testo che si attivano facendo clic sul pulsante **Visualizza proprietà testo**.



Appare la barra con i comandi per la formattazione il testo.



In alto a destra è presente il pulsante che apre il manuale del programma.

**AVVERTENZA:**

I marchi registrati sono proprietà dei rispettivi detentori.

Bologna - Settembre 2022

TUTTI I DIRITTI SONO RISERVATI