ML 24/25-10 Creating Text from images with OCR API

Team : TechTitans

Mithila Prabhu (1567111)

Taibaz Pathan (1564714)

Khushal Singh (1564813)

Table Of Contents

1. Objective

- The objective of the project **“Creating Text from Images with OCR API”** is to develop a console application using the **Tesseract SDK** for efficient text extraction from input images.
- The primary goal is to identify the most effective **image preprocessing technique** among various available methods to enhance OCR accuracy.
- Additionally, the project evaluates the impact of different preprocessing approaches on text extraction quality by comparing their effectiveness.

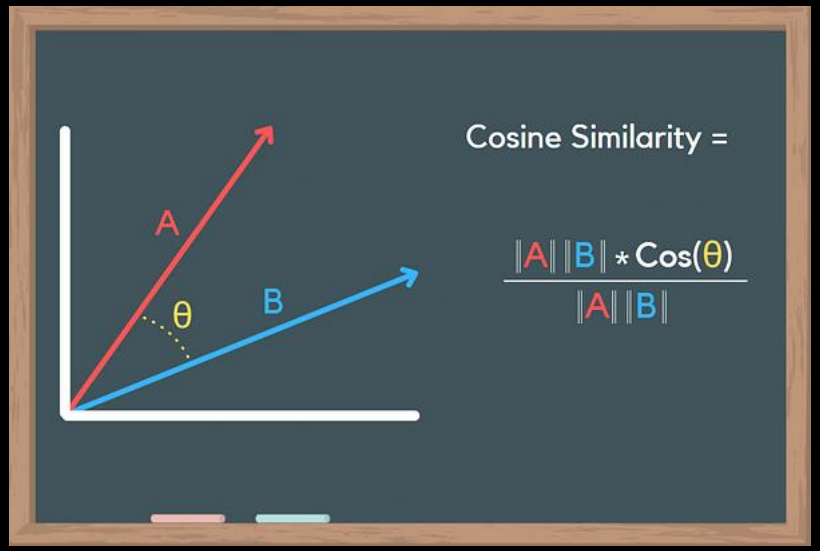
2. Introduction

- **Optical Character Recognition (OCR)** is a technology that converts printed or handwritten text from images into machine-readable text. It is widely used for document digitization, automated data entry, and text extraction. The accuracy of OCR depends on factors such as image quality and preprocessing techniques like resizing, noise removal, and contrast enhancement. By optimizing these steps, OCR performance improves, making it more effective for various real-world applications.

- Tesseract OCR is an open-source software that converts scanned documents, PDFs, and images into editable text. Developed by HP and maintained by Google, it supports over 100 languages and uses machine learning for accurate text recognition. It can handle complex layouts and noisy images, making it versatile for a variety of applications. Tesseract is widely used for document processing, text extraction, and digitizing information in both commercial and research fields.

2. Introduction

- Cosine similarity measures the similarity between two vectors by calculating the cosine of the angle between them. It's commonly used in text analysis and NLP to compare documents or word embeddings. The value ranges from -1 to 1, where 1 indicates identical vectors, 0 means no similarity, and -1 represents opposite vectors. It's especially useful for comparing documents of different lengths and is widely applied in tasks like document classification and recommendation systems.



2. Introduction

- Embeddings represent high-dimensional text data as lower-dimensional vectors, making it easier to analyze and compare extracted text. In this project, **text embeddings** are generated from OCR-extracted text using a **pretrained embedding model**. These embeddings help in evaluating the effectiveness of different **image preprocessing techniques** by capturing semantic relationships in the extracted text. **Cosine similarity** is used to compare embeddings, allowing for an objective assessment of text quality. This approach is crucial for analyzing OCR accuracy and improving text recognition in real-world applications.

- A similarity matrix represents pairwise similarity between a set of objects, with each element showing how similar two objects are. It's commonly used in machine learning, clustering, and recommendation systems, using metrics like cosine similarity or Pearson correlation. The matrix is symmetric, with diagonal elements typically showing maximal similarity .

3. Methodology

1. Create a Blank Text File

for storing extracted text.

2. Load Input Images

For Processing

3. Apply Preprocessing Techniques

Grayscale, Adaptive Thresholding and save the transformed image.

4. Extract Text using Tesseract OCR

store it in a text file and dictionary.

5. Track Time and Memory Usage

for preprocessing and OCR.

6. Generate Embeddings

using the *text-embedding-ada-002* model via ChatGPT API.

7. Compute Cosine Similarity

and store results in Excel.

8. Evaluate preprocessing techniques

based on similarity scores, time, and memory usage.

9. Save final results

based on cosine similarity, time tracking, and memory usage, and save the final results in an Excel file.

4. Implementation

Image Preprocessing Techniques

1. Convert images to grayscale.
2. Apply Adaptive Thresholding
3. Apply Global thresholding.
4. Adjust Saturation.
5. Apply De-skew processing to images.

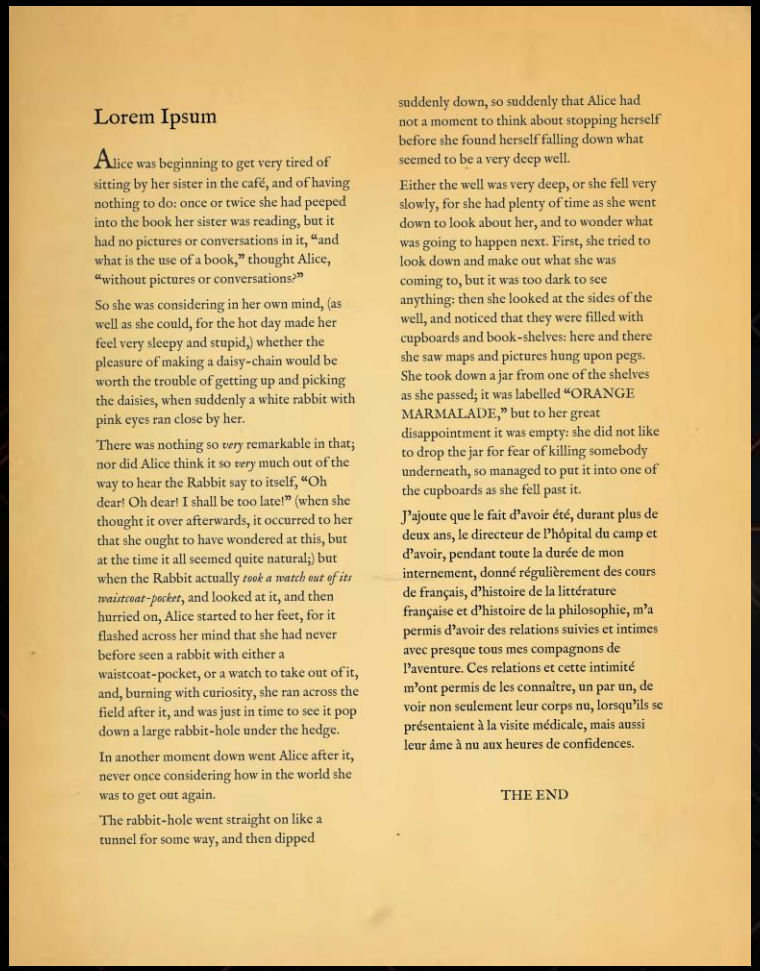
Performance Evaluation Metrics

- Calculate Average Cosine similarity across all models. Exclude the diagonal values (which are always 1 because they represent self-similarity).
- Cosine Similarity : normalizedCoSim = similarityScore / maxSimilarity
- Execution Time : normalizedTime = 1 - (timeTaken / maxTime)
- Memory Usage : normalizedMem = 1 - (memoryUsage / maxMemory)
- Final Score=(0.5×normalizedCoSim)+(0.3×normalizedTime)+(0.2×normalizedMem)
- 50% importance to similarity score. 30% to execution time. 20% to memory usage.

5. Results

- Comparison of different preprocessing techniques.
- Demonstration of extracted text from sample images.

Preprocessed Images



Lorem Ipsum

Alice was beginning to get very tired of sitting by her sister in the café, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?"

So she was considering in her own mind, (as well as she could, for the hot day made her feel very sleepy and stupid,) whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a white rabbit with pink eyes ran close by her.

There was nothing so *very* remarkable in that; nor did Alice think it so *very* much out of the way to hear the Rabbit say to itself, "Oh dear! Oh dear! I shall be too late!" (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural;) but when the Rabbit actually *took a watch out of its waistcoat-pocket*, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and, burning with curiosity, she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

In another moment down went Alice after it, never once considering how in the world she was to get out again.

The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down what seemed to be a very deep well.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything: then she looked at the sides of the well, and noticed that they were filled with cupboards and book-shelves: here and there she saw maps and pictures hung upon pegs. She took down a jar from one of the shelves as she passed; it was labelled "ORANGE MARMALADE," but to her great disappointment it was empty: she did not like to drop the jar for fear of killing somebody underneath, so managed to put it into one of the cupboards as she fell past it.

J'ajoute que le fait d'avoir été, durant plus de deux ans, le directeur de l'hôpital du camp et d'avoir, pendant toute la durée de mon internement, donné régulièrement des cours de français, d'histoire de la littérature française et d'histoire de la philosophie, m'a permis d'avoir des relations suivies et intimes avec presque tous mes compagnons de l'aventure. Ces relations et cette intimité m'ont permis de les connaître, un par un, de voir non seulement leur corps nu, lorsqu'ils se présentaient à la visite médicale, mais aussi leur âme à nu aux heures de confidences.

THE END

Input Image

Lorem Ipsum

Alice was beginning to get very tired of sitting by her sister in the café, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?"

So she was considering in her own mind, (as well as she could, for the hot day made her feel very sleepy and stupid,) whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a white rabbit with pink eyes ran close by her.

There was nothing so *very* remarkable in that; nor did Alice think it so *very* much out of the way to hear the Rabbit say to itself, "Oh dear! Oh dear! I shall be too late!" (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural;) but when the Rabbit actually *took a watch out of its waistcoat-pocket*, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and, burning with curiosity, she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

In another moment down went Alice after it, never once considering how in the world she was to get out again.

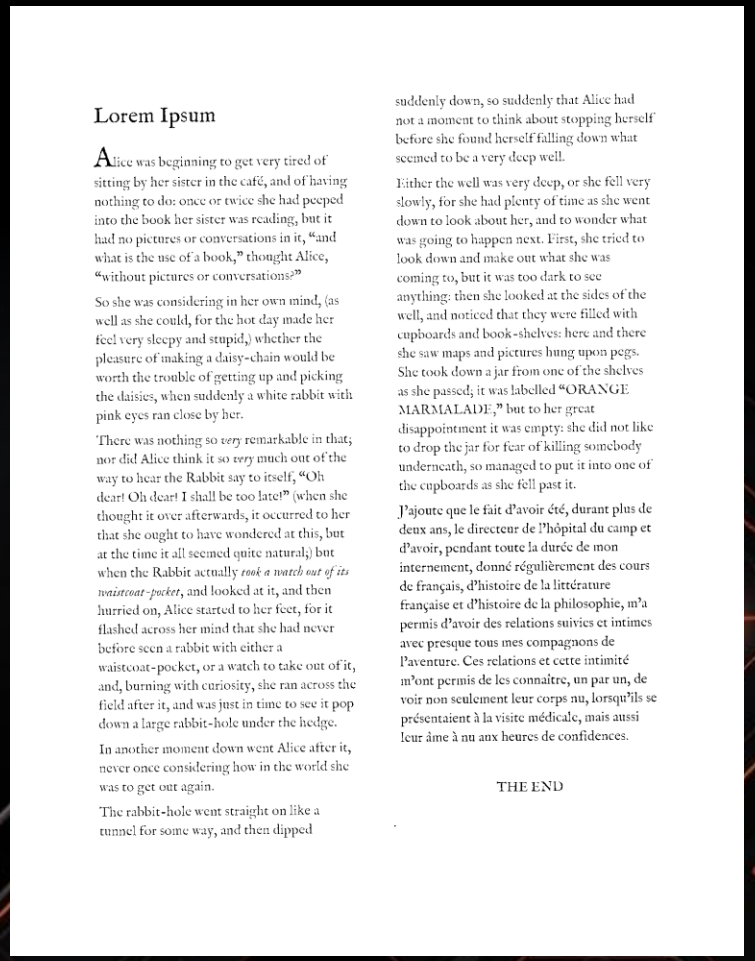
The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down what seemed to be a very deep well.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything: then she looked at the sides of the well, and noticed that they were filled with cupboards and book-shelves: here and there she saw maps and pictures hung upon pegs. She took down a jar from one of the shelves as she passed; it was labelled "ORANGE MARMALADE," but to her great disappointment it was empty: she did not like to drop the jar for fear of killing somebody underneath, so managed to put it into one of the cupboards as she fell past it.

J'ajoute que le fait d'avoir été, durant plus de deux ans, le directeur de l'hôpital du camp et d'avoir, pendant toute la durée de mon internement, donné régulièrement des cours de français, d'histoire de la littérature française et d'histoire de la philosophie, m'a permis d'avoir des relations suivies et intimes avec presque tous mes compagnons de l'aventure. Ces relations et cette intimité m'ont permis de les connaître, un par un, de voir non seulement leur corps nu, lorsqu'ils se présentaient à la visite médicale, mais aussi leur âme à nu aux heures de confidences.

THE END

After Converting to Grayscale



Lorem Ipsum

Alice was beginning to get very tired of sitting by her sister in the café, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?"

So she was considering in her own mind, (as well as she could, for the hot day made her feel very sleepy and stupid,) whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a white rabbit with pink eyes ran close by her.

There was nothing so *very* remarkable in that; nor did Alice think it so *very* much out of the way to hear the Rabbit say to itself, "Oh dear! Oh dear! I shall be too late!" (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural;) but when the Rabbit actually *took a watch out of its waistcoat-pocket*, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and, burning with curiosity, she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

In another moment down went Alice after it, never once considering how in the world she was to get out again.

The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down what seemed to be a very deep well.

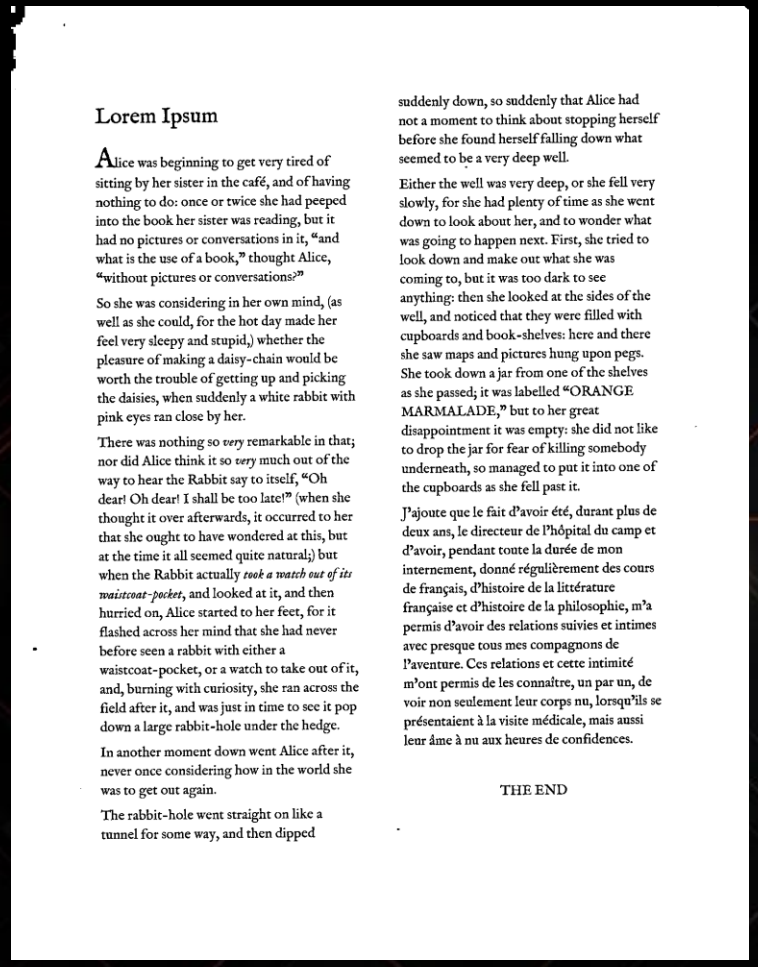
Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything: then she looked at the sides of the well, and noticed that they were filled with cupboards and book-shelves: here and there she saw maps and pictures hung upon pegs. She took down a jar from one of the shelves as she passed; it was labelled "ORANGE MARMALADE," but to her great disappointment it was empty: she did not like to drop the jar for fear of killing somebody underneath, so managed to put it into one of the cupboards as she fell past it.

J'ajoute que le fait d'avoir été, durant plus de deux ans, le directeur de l'hôpital du camp et d'avoir, pendant toute la durée de mon internement, donné régulièrement des cours de français, d'histoire de la littérature française et d'histoire de la philosophie, m'a permis d'avoir des relations suivies et intimes avec presque tous mes compagnons de l'aventure. Ces relations et cette intimité m'ont permis de les connaître, un par un, de voir non seulement leur corps nu, lorsqu'ils se présentaient à la visite médicale, mais aussi leur âme à nu anx heures de confidences.

THE END

After Global Thresholding at 80

Preprocessed Images



Lorem Ipsum

Alice was beginning to get very tired of sitting by her sister in the café, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?"

So she was considering in her own mind, (as well as she could, for the hot day made her feel very sleepy and stupid,) whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a white rabbit with pink eyes ran close by her.

There was nothing so *very* remarkable in that; nor did Alice think it so *very* much out of the way to hear the Rabbit say to itself, "Oh dear! Oh dear! I shall be too late!" (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural;) but when the Rabbit actually *took a watch out of its waistcoat-pocket*, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and, burning with curiosity, she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

In another moment down went Alice after it, never once considering how in the world she was to get out again.

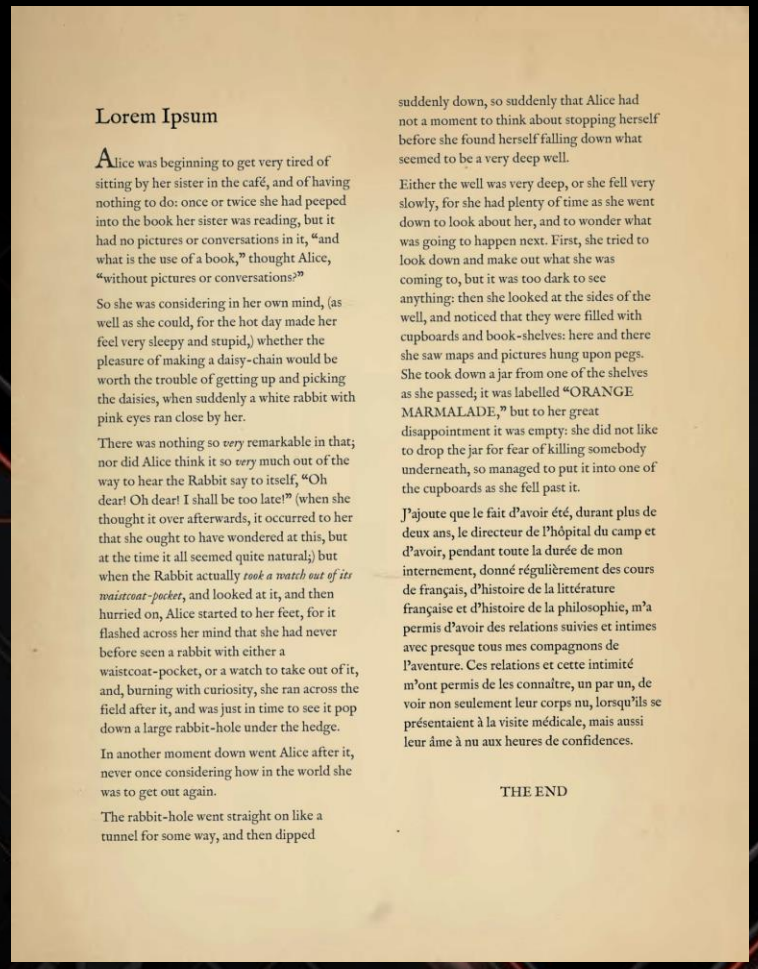
The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down what seemed to be a very deep well.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything: then she looked at the sides of the well, and noticed that they were filled with cupboards and book-shelves: here and there she saw maps and pictures hung upon pegs. She took down a jar from one of the shelves as she passed; it was labelled "ORANGE MARMALADE," but to her great disappointment it was empty: she did not like to drop the jar for fear of killing somebody underneath, so managed to put it into one of the cupboards as she fell past it.

J'ajoute que le fait d'avoir été, durant plus de deux ans, le directeur de l'hôpital du camp et d'avoir, pendant toute la durée de mon internement, donné régulièrement des cours de français, d'histoire de la littérature française et d'histoire de la philosophie, m'a permis d'avoir des relations suivies et intimes avec presque tous mes compagnons de l'aventure. Ces relations et cette intimité m'ont permis de les connaître, un par un, de voir non seulement leur corps nu, lorsqu'ils se présentaient à la visite médicale, mais aussi leur âme à nu aux heures de confidences.

THE END

Global Thresholding at 150



Lorem Ipsum

Alice was beginning to get very tired of sitting by her sister in the café, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?"

So she was considering in her own mind, (as well as she could, for the hot day made her feel very sleepy and stupid,) whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a white rabbit with pink eyes ran close by her.

There was nothing so *very* remarkable in that; nor did Alice think it so *very* much out of the way to hear the Rabbit say to itself, "Oh dear! Oh dear! I shall be too late!" (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural;) but when the Rabbit actually *took a watch out of its waistcoat-pocket*, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and, burning with curiosity, she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

In another moment down went Alice after it, never once considering how in the world she was to get out again.

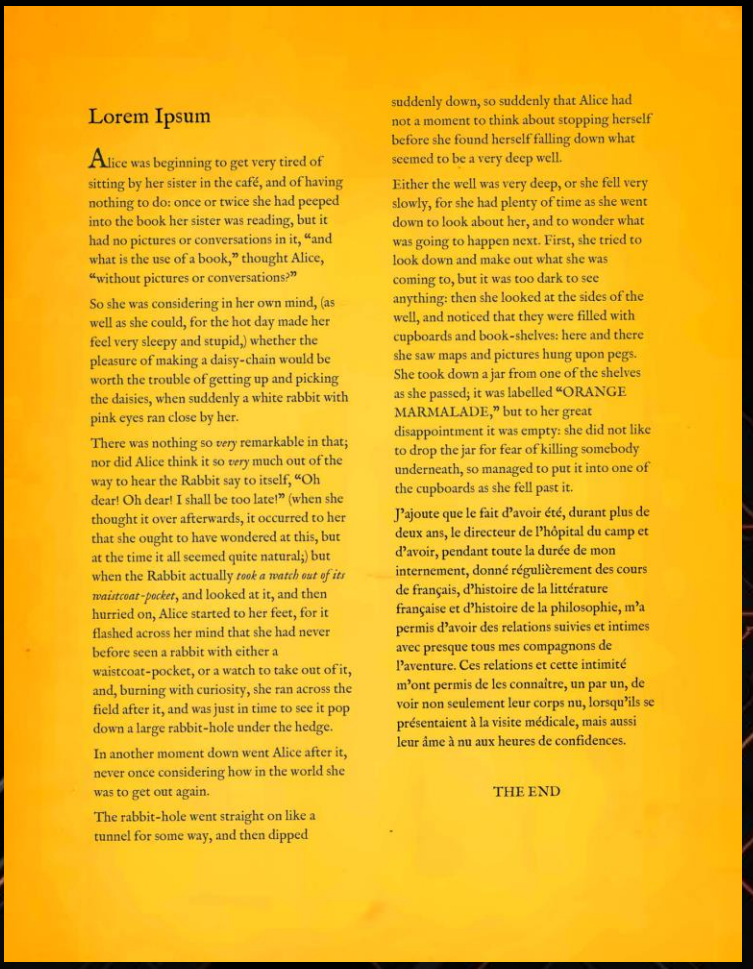
The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down what seemed to be a very deep well.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything: then she looked at the sides of the well, and noticed that they were filled with cupboards and book-shelves: here and there she saw maps and pictures hung upon pegs. She took down a jar from one of the shelves as she passed; it was labelled "ORANGE MARMALADE," but to her great disappointment it was empty: she did not like to drop the jar for fear of killing somebody underneath, so managed to put it into one of the cupboards as she fell past it.

J'ajoute que le fait d'avoir été, durant plus de deux ans, le directeur de l'hôpital du camp et d'avoir, pendant toute la durée de mon internement, donné régulièrement des cours de français, d'histoire de la littérature française et d'histoire de la philosophie, m'a permis d'avoir des relations suivies et intimes avec presque tous mes compagnons de l'aventure. Ces relations et cette intimité m'ont permis de les connaître, un par un, de voir non seulement leur corps nu, lorsqu'ils se présentaient à la visite médicale, mais aussi leur âme à nu aux heures de confidences.

THE END

Saturation adjustment at 0.6



Lorem Ipsum

Alice was beginning to get very tired of sitting by her sister in the café, and of having nothing to do: once or twice she had peeped into the book her sister was reading, but it had no pictures or conversations in it, "and what is the use of a book," thought Alice, "without pictures or conversations?"

So she was considering in her own mind, (as well as she could, for the hot day made her feel very sleepy and stupid,) whether the pleasure of making a daisy-chain would be worth the trouble of getting up and picking the daisies, when suddenly a white rabbit with pink eyes ran close by her.

There was nothing so *very* remarkable in that; nor did Alice think it so *very* much out of the way to hear the Rabbit say to itself, "Oh dear! Oh dear! I shall be too late!" (when she thought it over afterwards, it occurred to her that she ought to have wondered at this, but at the time it all seemed quite natural;) but when the Rabbit actually *took a watch out of its waistcoat-pocket*, and looked at it, and then hurried on, Alice started to her feet, for it flashed across her mind that she had never before seen a rabbit with either a waistcoat-pocket, or a watch to take out of it, and, burning with curiosity, she ran across the field after it, and was just in time to see it pop down a large rabbit-hole under the hedge.

In another moment down went Alice after it, never once considering how in the world she was to get out again.

The rabbit-hole went straight on like a tunnel for some way, and then dipped suddenly down, so suddenly that Alice had not a moment to think about stopping herself before she found herself falling down what seemed to be a very deep well.

Either the well was very deep, or she fell very slowly, for she had plenty of time as she went down to look about her, and to wonder what was going to happen next. First, she tried to look down and make out what she was coming to, but it was too dark to see anything: then she looked at the sides of the well, and noticed that they were filled with cupboards and book-shelves: here and there she saw maps and pictures hung upon pegs. She took down a jar from one of the shelves as she passed; it was labelled "ORANGE MARMALADE," but to her great disappointment it was empty: she did not like to drop the jar for fear of killing somebody underneath, so managed to put it into one of the cupboards as she fell past it.

J'ajoute que le fait d'avoir été, durant plus de deux ans, le directeur de l'hôpital du camp et d'avoir, pendant toute la durée de mon internement, donné régulièrement des cours de français, d'histoire de la littérature française et d'histoire de la philosophie, m'a permis d'avoir des relations suivies et intimes avec presque tous mes compagnons de l'aventure. Ces relations et cette intimité m'ont permis de les connaître, un par un, de voir non seulement leur corps nu, lorsqu'ils se présentaient à la visite médicale, mais aussi leur âme à nu aux heures de confidences.

THE END

Saturation Adjustment 2.5

Similarity matrix

Model	Grayscale	GlobalThr	GlobalThr	GlobalThr	GlobalThr	GlobalThr	GlobalThr	Saturatior	Saturatior	Saturatior	Saturatior	Saturatior	Saturatior	Deskewec	AdaptiveThreshold
Grayscale	1	0.92562	0.93571	0.95091	0.93271	0.92114	0.9049	0.91801	0.90791	0.9092	0.93183	0.92329	0.92014	0.94148	0.88172
GlobalThreshold_80	0.92562	1	0.95949	0.9242	0.90188	0.89318	0.88504	0.8899	0.87723	0.88288	0.89892	0.89771	0.88404	0.92225	0.85245
GlobalThreshold_100	0.93571	0.95949	1	0.94309	0.93118	0.92468	0.91076	0.914	0.90421	0.91037	0.91848	0.91775	0.91075	0.93818	0.87532
GlobalThreshold_110	0.95091	0.9242	0.94309	1	0.92099	0.90012	0.88569	0.89367	0.8813	0.88076	0.90362	0.89956	0.90021	0.90922	0.86515
GlobalThreshold_120	0.93271	0.90188	0.93118	0.92099	1	0.95764	0.94456	0.95015	0.94604	0.93695	0.94881	0.94335	0.94839	0.94461	0.93175
GlobalThreshold_140	0.92114	0.89318	0.92468	0.90012	0.95764	1	0.98045	0.97651	0.98222	0.97407	0.97435	0.97265	0.97545	0.95963	0.94463
GlobalThreshold_150	0.9049	0.88504	0.91076	0.88569	0.94456	0.98045	1	0.97603	0.9806	0.96528	0.96848	0.97088	0.97243	0.94279	0.93605
SaturationAdjusted_1.2	0.91801	0.8899	0.914	0.89367	0.95015	0.97651	0.97603	1	0.98277	0.97168	0.97166	0.97122	0.96633	0.9525	0.92464
SaturationAdjusted_0.6	0.90791	0.87723	0.90421	0.8813	0.94604	0.98222	0.9806	0.98277	1	0.97771	0.97625	0.97284	0.97363	0.95142	0.93401
SaturationAdjusted_0.9	0.9092	0.88288	0.91037	0.88076	0.93695	0.97407	0.96528	0.97168	0.97771	1	0.96612	0.96111	0.95667	0.94967	0.91763
SaturationAdjusted_1.6	0.93183	0.89892	0.91848	0.90362	0.94881	0.97435	0.96848	0.97166	0.97625	0.96612	1	0.96866	0.97777	0.95556	0.92607
SaturationAdjusted_2	0.92329	0.89771	0.91775	0.89956	0.94335	0.97265	0.97088	0.97122	0.97284	0.96111	0.96866	1	0.96975	0.94765	0.92177
SaturationAdjusted_2.5	0.92014	0.88404	0.91075	0.90021	0.94839	0.97545	0.97243	0.96633	0.97363	0.95667	0.97777	0.96975	1	0.94394	0.92872
Deskewed	0.94148	0.92225	0.93818	0.90922	0.94461	0.95963	0.94279	0.9525	0.95142	0.94967	0.95556	0.94765	0.94394	1	0.91693
AdaptiveThreshold	0.88172	0.85245	0.87532	0.86515	0.93175	0.94463	0.93605	0.92464	0.93401	0.91763	0.92607	0.92177	0.92872	0.91693	1

Time Tracking and Memory Usage Details

Model	Average Time (s)	Average Memory (MB)	
Grayscale	2024.865767	0.013807954	
GlobalThreshold_80	1676.491633	0.002740924	
GlobalThreshold_100	1701.283533	0.001915953	
GlobalThreshold_110	1733.68175	0.00275027	
GlobalThreshold_120	1655.744083	0.00275027	
GlobalThreshold_140	1670.999433	0.00275027	
GlobalThreshold_150	1693.374417	0.002975761	
SaturationAdjusted_1.2	2127.863867	0.001956601	
SaturationAdjusted_0.6	2134.148967	0.00275027	
SaturationAdjusted_0.9	2123.34895	0.002707142	
SaturationAdjusted_1.6	2145.41245	0.002723567	
SaturationAdjusted_2	2085.000333	0.002681414	
SaturationAdjusted_2.5	2074.20855	0.00275027	
Deskewed	1199.244817	2.110389561	
AdaptiveThreshold	4515.6431	0.002624088	

Model Evaluation

Rank	Model	Final Score
1	Grayscale	0.86416802
2	GlobalThreshold_110	0.86001553
3	GlobalThreshold_120	0.85609369
4	GlobalThreshold_100	0.85464952
5	GlobalThreshold_80	0.85116987
6	GlobalThreshold_140	0.8492935
7	GlobalThreshold_150	0.83966942
8	SaturationAdjusted_1.6	0.82312296
9	SaturationAdjusted_2	0.82287403
10	SaturationAdjusted_2.5	0.82200781
11	SaturationAdjusted_1.2	0.81745263
12	SaturationAdjusted_0.9	0.81327792
13	SaturationAdjusted_0.6	0.81191152
14	Deskewed	0.69106554
15	AdaptiveThreshold	0.64061227

Log File

```
[INFO] 21-03-2025 17:29:25: Processing 1 images...
[INFO] 21-03-2025 17:29:26: Saved processed image: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output
\ProcessedImage\ocr_preprocessing_input_Grayscale.png
[INFO] 21-03-2025 17:29:30: Saved processed image: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output
\ProcessedImage\ocr_preprocessing_input_GlobalThreshold.png
[INFO] 21-03-2025 17:29:31: Saved processed image: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output
\ProcessedImage\ocr_preprocessing_input_Shifted.png
[INFO] 21-03-2025 17:29:32: Saved processed image: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output
\ProcessedImage\ocr_preprocessing_input_SaturationAdjusted.png
[INFO] 21-03-2025 17:29:34: Saved processed image: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output
\ProcessedImage\ocr_preprocessing_input_Deskewed.png
[INFO] 21-03-2025 17:29:36: Successfully processed image: ocr_preprocessing_input.png
[INFO] 21-03-2025 17:29:36: Processing completed successfully.
[INFO] 21-03-2025 21:26:34: Starting Image Processing...
[INFO] 21-03-2025 21:26:34: Configuration loaded successfully.
[INFO] 21-03-2025 21:26:34: Deleted file: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output\ProcessedImage
\ocr_preprocessing_input_Deskewed.png
[INFO] 21-03-2025 21:26:34: Deleted file: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output\ProcessedImage
\ocr_preprocessing_input_GlobalThreshold.png
[INFO] 21-03-2025 21:26:34: Deleted file: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output\ProcessedImage
\ocr_preprocessing_input_Grayscale.png
[INFO] 21-03-2025 21:26:34: Deleted file: C:\Users\mithi\OneDrive\Desktop\SoftwareEngg\ocr-techtitans\src\OCRProject\Output\ProcessedImage
```

6. Conclusion

The project demonstrates the effectiveness of **OCR** in text extraction by analysing the impact of various **preprocessing techniques** on accuracy. **Time and memory tracking** provide insights into computational costs, while **evaluation using cosine similarity** ensures a comprehensive assessment. The results emphasize the importance of preprocessing in enhancing **OCR accuracy and efficiency**, making the system more reliable for real-world applications.

Future Work

1. **Enhancing Preprocessing Techniques** – Implementing advanced methods such as deep learning-based image enhancement and adaptive thresholding to further improve text recognition.
2. **Integration with Other OCR Models** – Exploring and comparing different OCR engines to analyse their effectiveness with varied preprocessing techniques.
3. **Automating Parameter Selection** – Developing an automated system to select the best preprocessing method dynamically based on image characteristics.
4. **Expanding Language Support** – Extending OCR capabilities to support multiple languages and handwritten text recognition.
5. **Real-time OCR Implementation** – Optimizing the system for real-time text extraction in applications like document scanning and automated data entry.

Thank You!