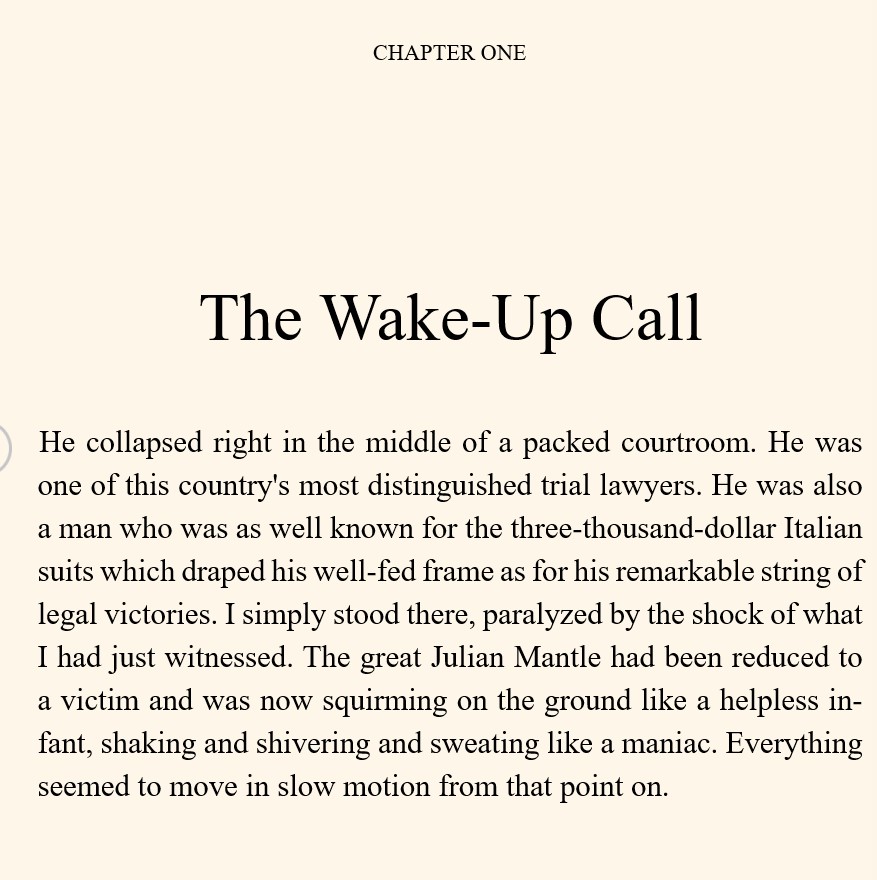
**PLAIN TEXT , IDEAL SITUATION ANALYSIS**

RAW IMAGE



**Ground Truth**

"CHAPTER ONE The Wake-Up Call He collapsed right in the middle of a packed courtroom. He was one of this country's most distinguished trial lawyers. He was also a man who was as well known for the three-thousand-dollar Italian suits which draped his well-fed frame as for his remarkable string of legal victories. I simply stood there, paralyzed by the shock of what I had just witnessed. The great Julian Mantle had been reduced to a victim and was now squirming on the ground like a helpless infant, shaking and shivering and sweating like a maniac. Everything seemed to move in slow motion from that point on."

**WORD ERROR RATE**

The WER calculation is based on a measurement called the “Levenshtein distance.”

The Levenshtein distance is a measurement of the differences between two “strings.”

To put it in a simple formula,

Word Error Rate = (Substitutions + Insertions + Deletions) / Number of Words Spoken

But how do you add up those factors? Let’s look at each one:

A substitution occurs when a word gets replaced (for example, “noose” is transcribed as “moose”)

An insertion is when a word is added that wasn’t said (for example, “SAT” becomes “essay tea”)

A deletion happens when a word is left out of the transcript completely (for example, “turn it around” becomes “turn around”)

**Character Error Rate**

This value indicates the percentage of characters that were incorrectly predicted.

The lower the value, the better the performance of the ASR system with a CharErrorRate of 0 being a perfect score.

The Character Error Rate (CER) compares, for a given page, the total number of characters (n),including spaces, to the minimum number of insertions (i),

substitutions (s) and deletions (d) of characters that are required to obtain the Ground Truth result.

The formula to calculate CER is as follows: CER = [ (i + s + d) / n ]\*100

# TESSERACT OCR

EXTRACTION -

CHAPTER ONE The Wake-Up Call He collapsed right in the middle of a packed courtroom. He was one of this country's most distinguished trial lawyers. He was also a man who was as well known for the three-thousand-dollar Italian suits which draped his well-fed frame as for his remarkable string of legal victories. I simply stood there, paralyzed by the shock of what I had just witnessed. The great Julian Mantle had been reduced to a victim and was now squirming on the ground like a helpless in- fant, shaking and shivering and sweating like a maniac. Everything seemed to move in slow motion from that point on.

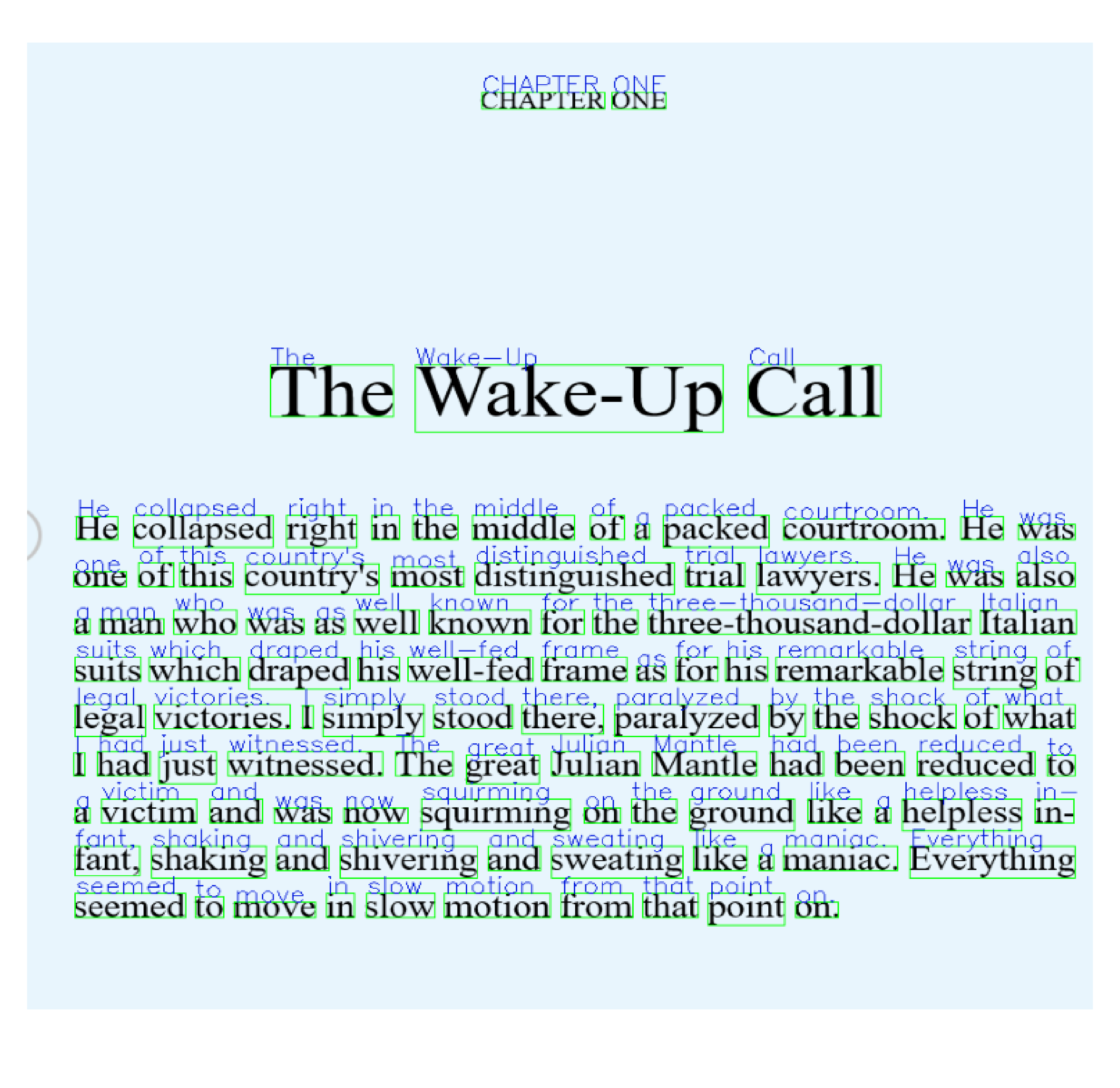
**#Detect Langauge other than English**

**# IT considers the special symbols,**

**# it is case sensitive as well**

**Word\_Error\_Rate: 0.018691588785046728**

**Char\_Error\_Rate: 0.006546644844517185**



￼

## KERAS OCR

**Word\_Error\_Rate: 0.6635514018691588**

**Char\_Error\_Rate: 0.46808510638297873**

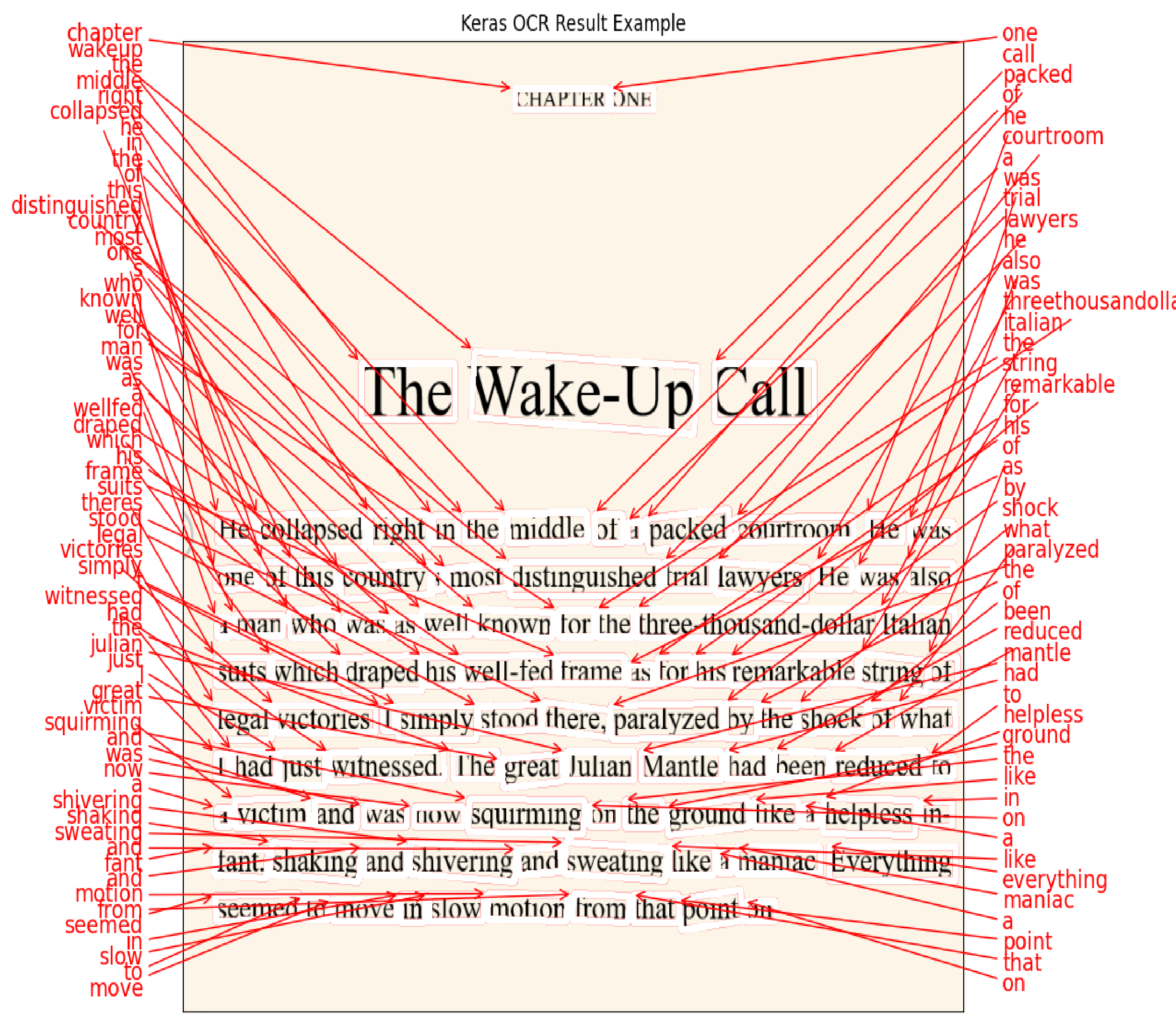
**EXTRACTION -**

chapter one the wakeup call middle collapsed right of packed he in the he courtroom a was of this distinguished trial lawyers he also country most s one was threethousandollar italian who well known for the man was a as which draped his wellfed frame for his remarkable of suits string as by the shock of what legal victories t simply stood theres paralyzed had witnessed the julian mantle had been reduced l just great to helpless victim and squirming the ground like in was on a a now shaking shivering like everything fant and and sweating maniac a seemed in slow motion from that point to move on

**# Case sensitivity is lost**

**# new lines is not detected**

**# Does not consider the special chars and punctuations**



# EASY OCR

**Word\_Error\_Rate: 0.04672897196261682**

**Char\_Error\_Rate: 0.011456628477905073**

**Extraction -**

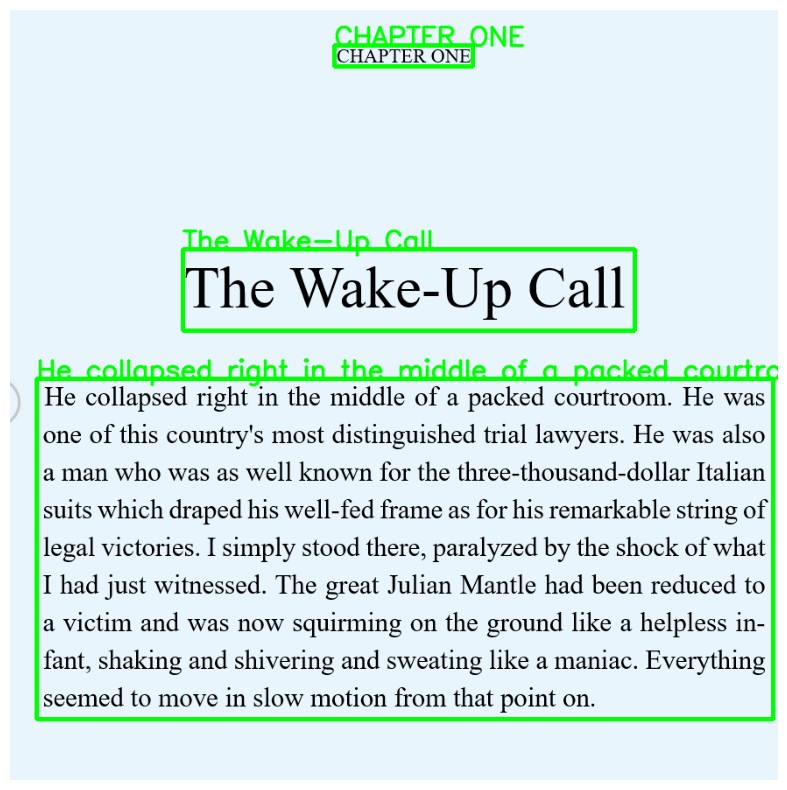
CHAPTER ONE The Wake-Up Call He collapsed right in the middle of a packed courtroom He was one of this country's most distinguished trial lawyers. He was also a man who was as well known for the three-thousand-dollar Italian suits which draped his well-fed frame as for his remarkable string of legal victories. I simply stood there, paralyzed by the shock of what I had just witnessed. The great Julian Mantle had been reduced to victim and was now squirming on the ground like a helpless in- fant; shaking and shivering and sweating like a maniac. Everything seemed to move in slow motion from that point on

**# either paragraph false or true, the prediction is found as line or paragraph**

**# special characters are also detected**

**# It is case sensitive**

**# provides confidence too.**



**PADDLEOCR (Parallel Distributed Deep Learning)**

**Word\_Error\_Rate: 0.07476635514018691**

**Char\_Error\_Rate: 0.011456628477905073**

**Extraction -**

CHAPTER ONE The Wake-Up Call He collapsed right in the middle of a packed courtroom. He was. one of this country's most distinguished trial lawyers.

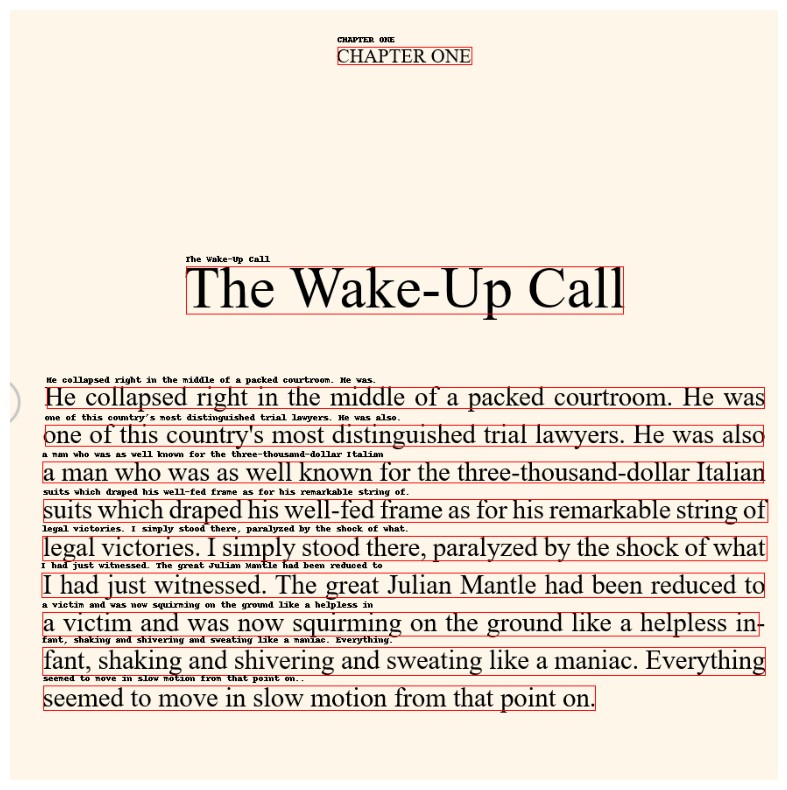
He was also. a man who was as well known for the three-thousand-dollar Italian suits which draped his well-fed frame as for his remarkable string of. legal victories. I simply stood there, paralyzed by the shock of what. I had just witnessed. The great Julian Mantle had been reduced to a victim and was now squirming on the ground like a helpless in fant, shaking and shivering and sweating like a maniac. Everything. seemed to move in slow motion from that point on..

**# it does acknowledge special characters**

**# it does acknoledge case sensitivity**

**# requires GPU**

**# it extracts sentences and not words**



## DONUTOCR

**Extraction** –

<s\_cord-v2><s\_menu><s\_nm> The Wake-Up Call</s\_nm><s\_num> He collapsed right in the middle of a packed courtroom. He was one of this country's most distinguished trial lawyers. He was also</s\_nm><s\_num> a man who was as well known for the three-thousand-dollar Italian</s\_nm><s\_num> suits which draped his well-fed frame as for his remarkable string of</s\_nm><s\_num> legal victories. Isimply stood there, paralyzed by the shock of what</s\_nm><s\_num> I had just witnessed. The great Julian Mantle had been reduced to a victim and was now squirming on the ground like a helpless in-</s\_nm><s\_num> a</s\_num><s\_price> a</s\_price><sep/><s\_nm> fant, shaking and shivering and sweating like a manac. Everything seemed to move in slow motion from that point on.</s\_menutype\_cnt></s\_total></s>

**After post - processing**

The Wake-Up Call He collapsed right in the middle of a packed courtroom. He was one of this country's most distinguished trial lawyers. He was also a man who was as well known for the three-thousand-dollar Italian suits which draped his well-fed frame as for his remarkable string of legal victories. Isimply stood there, paralyzed by the shock of what I had just witnessed. The great Julian Mantle had been reduced to a victim and was now squirming on the ground like a helpless in- a a fant, shaking and shivering and sweating like a manac. Everything seemed to move in slow motion from that point on.

**Word\_Error\_Rate: 0.08411214953271028**

**Char\_Error\_Rate: 0.03273322422258593**

**# it does acknowledge special characters**

**# it does acknoledge case sensitivity**

**# it extracts sentences and not words**

**NOTE – TESSERACT AND DONUT OCR CANNOT TO TOGETHER IN ONE ENV, THEY FACE SOME DEPENDENCY ISSUE**