

PROCESSAMENTO DE IMAGENS : RECONHECIMENTO E TRATAMENTO DE IMAGENS PARA TRADUÇÃO DE TEXTOS

Lucas Menduiña R. Evangelista

Abstract — O propósito central deste artigo consiste em explorar diversas abordagens e métodos destinados ao processamento de imagens. O estudo apresenta um sistema que visa reconhecer palavras em imagens utilizando Python, Tesseract e OpenCV e traduzi-las com o auxílio do Google Translate API. As técnicas aplicadas visam filtrar as imagens por meio de manipulações a fim de otimizar o reconhecimento de palavras.

I. INTRODUÇÃO

Processamento de imagem é o processo de transformar uma imagem para a forma digital e realizar determinadas operações a fim de extrair informações úteis. Com isso, algumas das operações contempladas pela área são : aprimoramento de imagem, restauração de imagem, compressão, processamento morfológico, segmentação e reconhecimento.

A implementação do processamento de imagens tem causado um impacto significativo no setor de tecnologia. Isso se reflete em diversas áreas beneficiadas pelo avanço dessas técnicas, tais como a medicina, sensores de tráfego, detecção de faces e reconstrução de imagens.

A digitalização de documentos é essencial para o desenvolvimento e organização de empresas, tanto pequenas quanto grandes. Nesse contexto, a necessidade de extrair textos a partir de imagens desempenha um papel fundamental na capacidade de armazenar e processar informações do mundo real. Esses avanços representam marcos significativos na globalização e na digitalização da informação. Dessa forma, por exemplo, livros antigos que estão em estado de deterioração podem ser digitalizados e eternizados no mundo digital.

O sistema abordado no artigo tem como objetivo a tradução de textos extraídos de imagens, utilizando o Tesseract, um mecanismo de reconhecimento de texto de código aberto OCR, que é o *Optical Character Recognition*. Ao longo do estudo, serão apresentadas as técnicas empregadas para aprimorar a extração de texto das

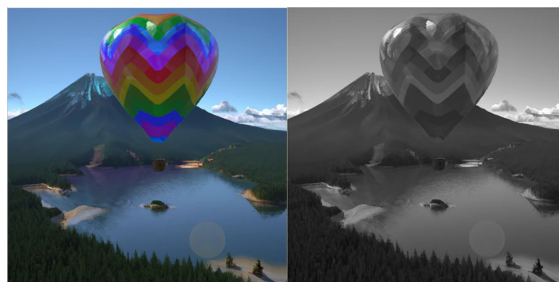
imagens, bem como a descrição do sistema proposto e a apresentação da conclusão final.

II. TÉCNICAS UTILIZADAS

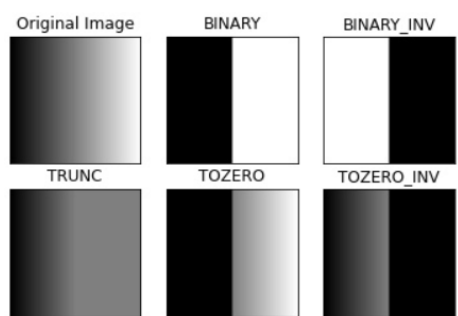
Nesta seção, apresentam-se as técnicas empregadas no processamento das imagens do sistema proposto. Serão elucidadas as explicações e a relevância de cada técnica para a obtenção dos resultados desejados. As técnicas utilizadas incluem: 1. `cvtColor()`, 2. `Threshold`, 3. `Closing`, 4. `Black Hat` e 5. `Median Filter`.

1. `cvtColor()`

Existem mais de 150 métodos de conversão de cores disponíveis no OpenCV. Neste artigo, será abordado mais atentamente a conversão da escala BGR (nativa do OpenCV) para a escala de cinza. A escala de cinza simplifica imagens representadas em *multi-channel (RGB)* para representação em *single-channel*. Com isso, imagens em escala cinza requerem menos memória e poder de processamento. Além disso, há melhoras significativas na detecção de recursos estruturais e textuais da imagem, como bordas e contornos.



Threshold é uma das técnicas de segmentação que gera uma imagem binária a partir de uma imagem em tons de cinza, separando-a em duas regiões com base em um valor de limite. Portanto, os pixels com valores de intensidade superiores ao limite especificado serão tratados como brancos ou 1 na imagem de saída, enquanto os outros serão pretos ou 0. É uma técnica que separa uma imagem em regiões distintas com base nos valores de intensidade dos pixels. O objetivo é facilitar a compreensão da imagem, permitindo que um computador a analise e extraia informações úteis.

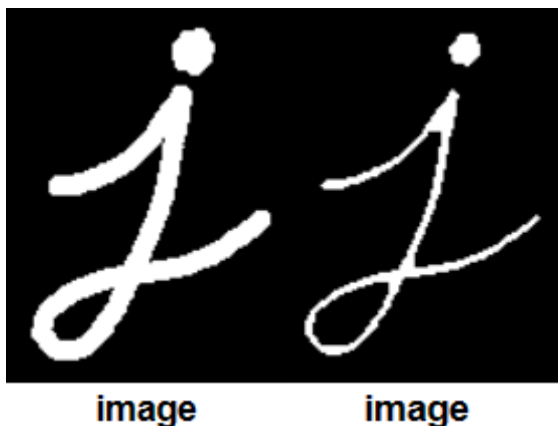


3. Closing

Fechamento é uma técnica que combina a dilatação seguida da erosão.

3.1. Erosion

Erosão desgasta as fronteiras do objeto em primeiro plano. Um pixel na imagem original será considerado 1 somente se todos os pixels sob o kernel forem 1, caso contrário ele é erodido (tornado zero).



3.2. Dilation

A dilatação é o oposto da erosão. Nesse caso, um elemento de pixel é '1' se pelo menos um pixel sob o kernel for '1'. Isso aumenta a região branca na imagem ou o tamanho do objeto em primeiro plano aumenta. Normalmente, em casos de remoção de ruído, a erosão é seguida pela dilatação. Isso ocorre porque a erosão remove ruídos brancos, mas também encolhe o objeto. Portanto, fazemos a dilatação. Uma vez que o ruído desapareceu, ele não voltará, mas a área do nosso objeto aumenta. A dilatação também é útil para unir partes quebradas de um objeto.

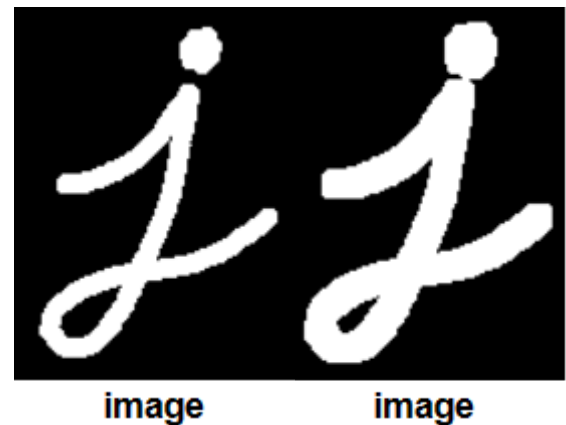


Fig 3.2 : Exemplo de dilatação de um caractere

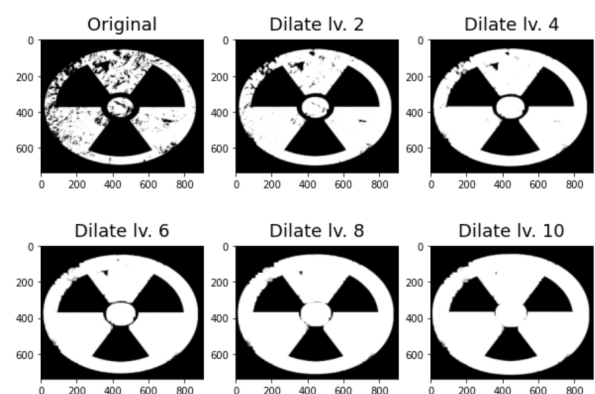


Fig 3.3 : Exemplo de dilatação de uma imagem

4. Black Hat

Black Hat é uma técnica que realiza a diferença entre o fechamento da imagem

(closing), com a imagem original. Essa técnica é eficaz para realçar letras ou detalhes na imagem, dependendo do objetivo do sistema em questão. Pode ser utilizada, por exemplo, para identificar placas de veículos.

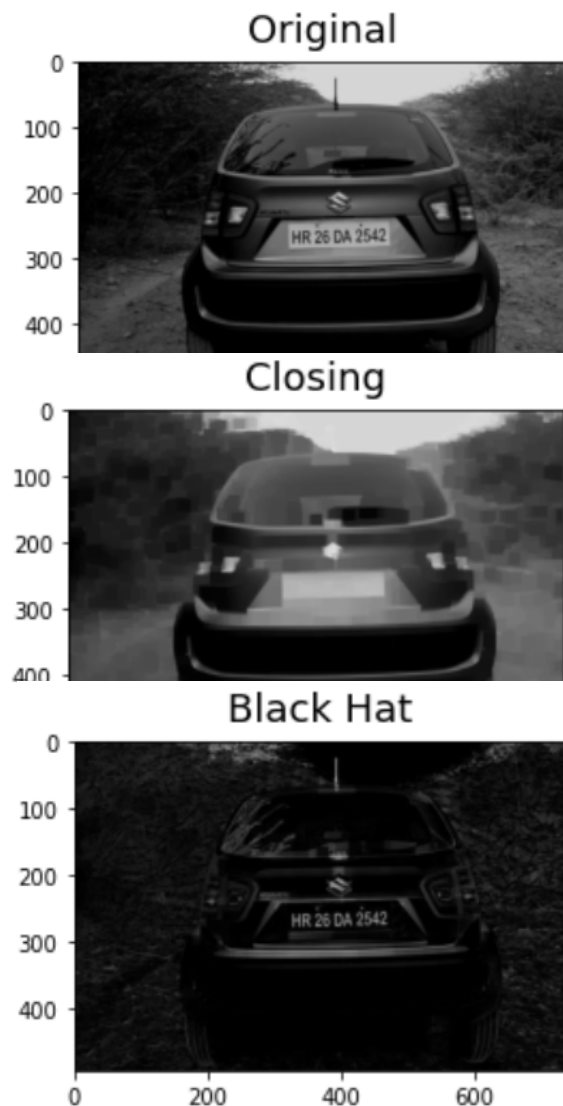


Fig 4: Exemplo de closing para realçar a placa de um veículo.

5. Median Filter

Ruído em imagens pode ser brevemente definido como variações aleatórias em alguns dos valores de pixel de uma imagem. Por exemplo, dada a matriz 3x3 :

[22, 24, 27]

[31, 98, 29]

[27, 22, 23]

Observe o pixel central: ele é claramente um valor atípico nesta matriz. Valores atípicos como esse podem criar o que é chamado de ruído sal e pimenta, que resulta em uma imagem que se assemelha :



A filtragem mediana é excelente para reduzir esse tipo de ruído. O algoritmo de filtragem examinará a imagem inteira, usando uma pequena matriz (como a 3x3 mostrada acima) e recalculará o valor do pixel central simplesmente tomando a mediana de todos os valores dentro da matriz. Com o exemplo acima, os valores ordenados são [22, 22, 23, 24, 27, 27, 29, 31, 98], e a mediana desse conjunto é 27. Aplicando o filtro na imagem acima, o resultado fica :



Fig 5: Exemplo de median filter para remoção de ruídos.

III. SISTEMA PROPOSTO

Nesta seção, será apresentado o sistema proposto, bem como seus objetivos, aplicação no mundo real e testes para diferentes tipos de imagens.

O sistema visa realizar o reconhecimento de texto em imagens usando a ferramenta *Tesseract* (OCR). Após a extração e tratamento do texto, o sistema traduzirá o texto para um idioma escolhido pelo usuário, utilizando a *API do Google Translate* por meio da biblioteca "googletrans".

Serão apresentados detalhes sobre quatro assuntos : 1. Imagens de entrada que foram utilizadas no algoritmo, 2. Pré-Processamento e Aplicação dos filtros, 3. Reconhecimento Textual, 4. Tradução dos textos extraídos.

IV. IMAGENS UTILIZADAS

Teste 1

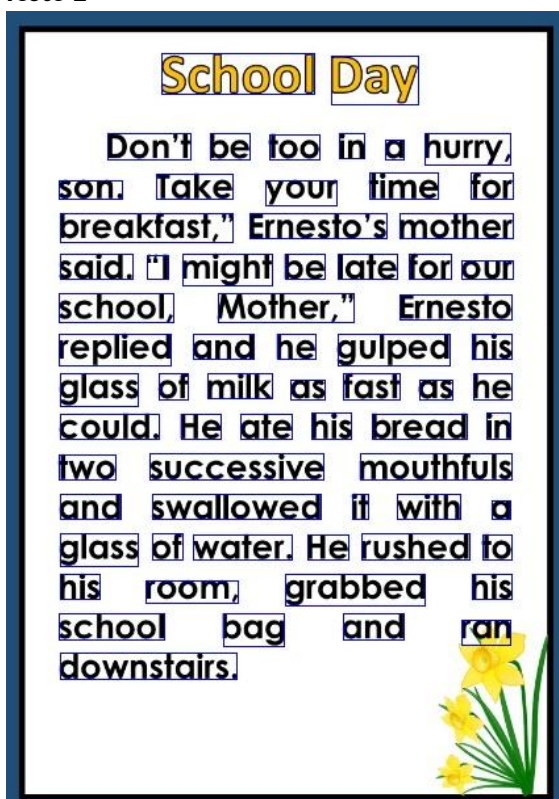


Fig 6.1 - Texto sem ruídos (antes)

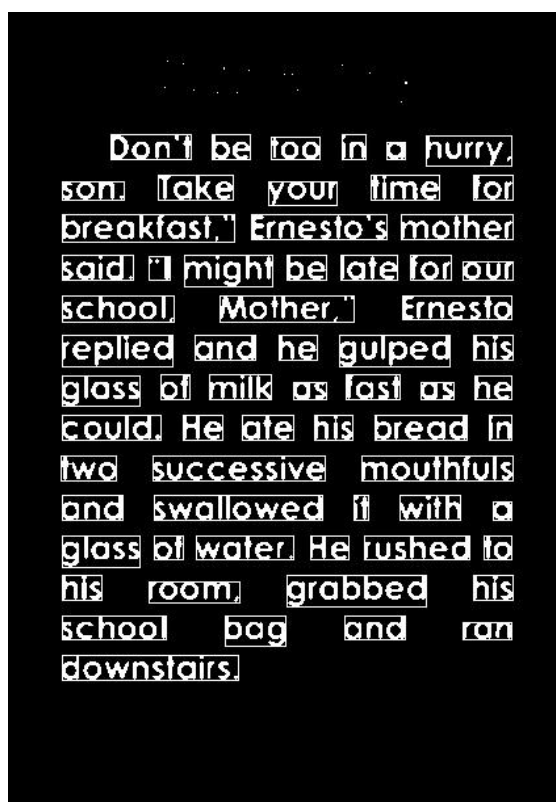


Fig 6.2 - Texto sem ruídos (depois)

V. Texto reconhecido (Inglês) :

" School Day Don't be too in a hurry, son. Take your time for breakfast," Ernesto's mother said. "I might be late for our school, Mother," Ernesto replied and he gulped his glass of milk as fast as he could. He ate his bread in two successive mouthfuls and swallowed it with a glass of water. He rushed to his room, grabbed his school bag and ran downstairs. a A "

VI. Texto traduzido (Português-BR) :

" Dia Escolar Não tenha pressa, filho. Tome o café da manhã com calma", disse a mãe de Ernesto. "Posso chegar atrasado para a escola, mãe", respondeu Ernesto e engoliu o copo de leite o mais rápido que pôde. Ele comeu o pão em dois goles sucessivos e engoliu-o com um copo de água. Ele correu

para seu quarto, pegou sua mochila escolar e desceu correndo. um A “

Observações (Fig 6): Note que não há grandes desafios para o OCR extrair o texto contido na imagem IV.1. Os resultados extraídos são os mesmos para a imagem antes e depois do pré-processamento.

Teste 2



Fig 7.1 - HQ do Homem de Ferro, imagem com detalhes. (antes)

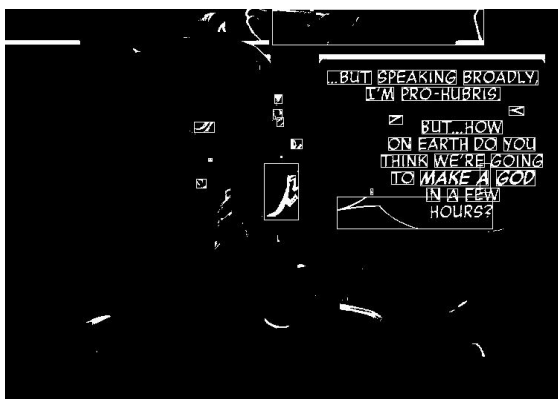


Fig 7.2 - HQ do Homem de Ferro, imagem com detalhes. (depois)

Teste 3

a. Texto reconhecido (Inglês):

“ a BUT SPEAKING BROADLY, ' T'M PRO-HUBRIS. ie <= a 2 7 BUT...HOW b ON EARTH 00 Y0u . : THINK WE'RE GOING a y; TO MAKE A GOD ' IN A FEW Hours? “

b. Texto traduzido (Português-BR):

“ a MAS FALANDO DE FORMA AMPLA, SOU PRO-HUBRIS. ou seja, <= a 2 7 MAS...COMO b NA TERRA 00 VOCÊ . : ACHO QUE VAMOS VAI; FAZER UM DEUS EM ALGUMAS HORAS? “

Observações (Fig 7): Note que o reconhecimento de texto em imagens com níveis mais elevados de detalhes representa um desafio para o algoritmo com as técnicas atualmente empregadas. Embora o reconhecimento permaneça eficaz, o Reconhecimento Óptico de Caracteres (OCR) pode, no entanto, confundir detalhes presentes na imagem com letras ou possíveis textos.

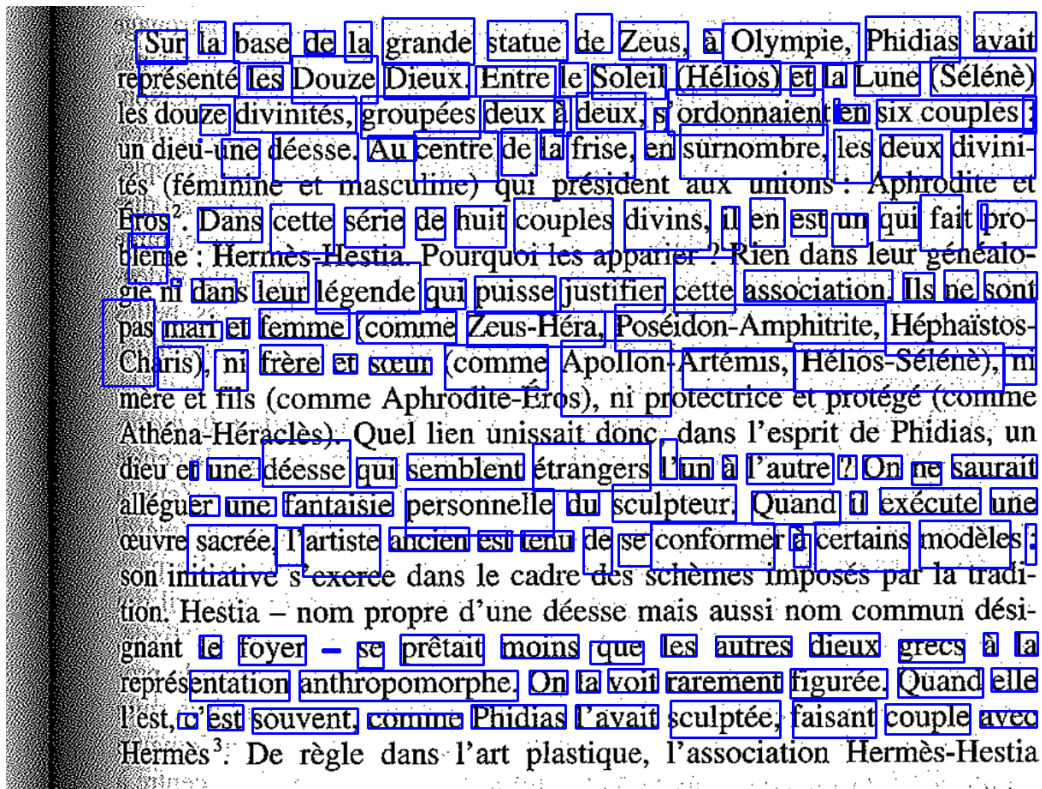


Fig 8.1 - Texto em francês (antes)

a. Texto reconhecido (Francês):

"sur jà base de la grande statue de Zeus, à Olympie Phidias ava.1t présenté les Douze Dieux. Entre le Soleil (Hélios) et la Luné (Selene) ze divinités, groupees deux à deux, s ordonna1ent 'eti six couplés : i ne deesse. Au centr e de la fnse en surnombre les deux d1vm1 – TOs Dans cette sene de hu1t couples d1vms, 11 en est un qui fa1t pro- èn ji dans leur legende qui puisse justifier cette association. Ils ne sont î>as Mari et femme (comme Zeus-Héia, Pos éxdon-Amphflnte HéphæstOs – aris) 'm fière et sœur (comme Apollon – Aflenns Héhos – Selcr'1e) m t'une" deesse qui semblent étrangers l un à l'autre ? On ne saurait er une' fantaisie personnelle du sculpteur. Quand il exécute uñe sacree T amste ancien est tenu de se conformer à certams modeles : le foyer – se prêtait moins 'que les autres dieux grecs à la entâtion anthropomorphe. On la voit rarement figurée. Quand elle 'c'est'souvent, comme Phidias l'avait sculptée;- faisant couple avec "

b. Texto traduzido (Português-BR):

"baseado na grande estátua de Zeus, no Olimpo Fídias apresentou os Doze Deuses. Entre o Sol (Helios) e a Lua (Selene) as divindades, agrupadas duas a duas, ordenam-se em seis casais: uma deusa. No centro da cena além dos dois d1v m1-TOs Nesta série de oito casais de d1vms, há um que deve ser pro-ènj em sua lenda que pode justificar esta associação. Eles não são como marido e mulher (como Zeus-Heia, Poséxdon-Amphflnte Héphaestos – aris) são orgulhosos e ir mäs (como Apolo – Aflenns Héhos – Selcr'1e), mas uma "deusa que parece estranha um ao outro". ? Não pode ser uma fantasia pessoal do escultor. Quando realiza um ato sagrado, o antigo Amste é obrigado

a conformar-se a certos modelos: a lareira - presta-se menos do que os outros deuses gregos à armadilha antropomórfica. Raramente o vemos retratado. Quando ela 'isso' é frequente, como Fídias a esculpiu; - fazendo um casal com (venv) “

Observações : Note que há muitos erros e palavras que não foram reconhecidas pelo Tesseract-OCR sem um pré-processamento da imagem.

I. Depois (Texto em francês)

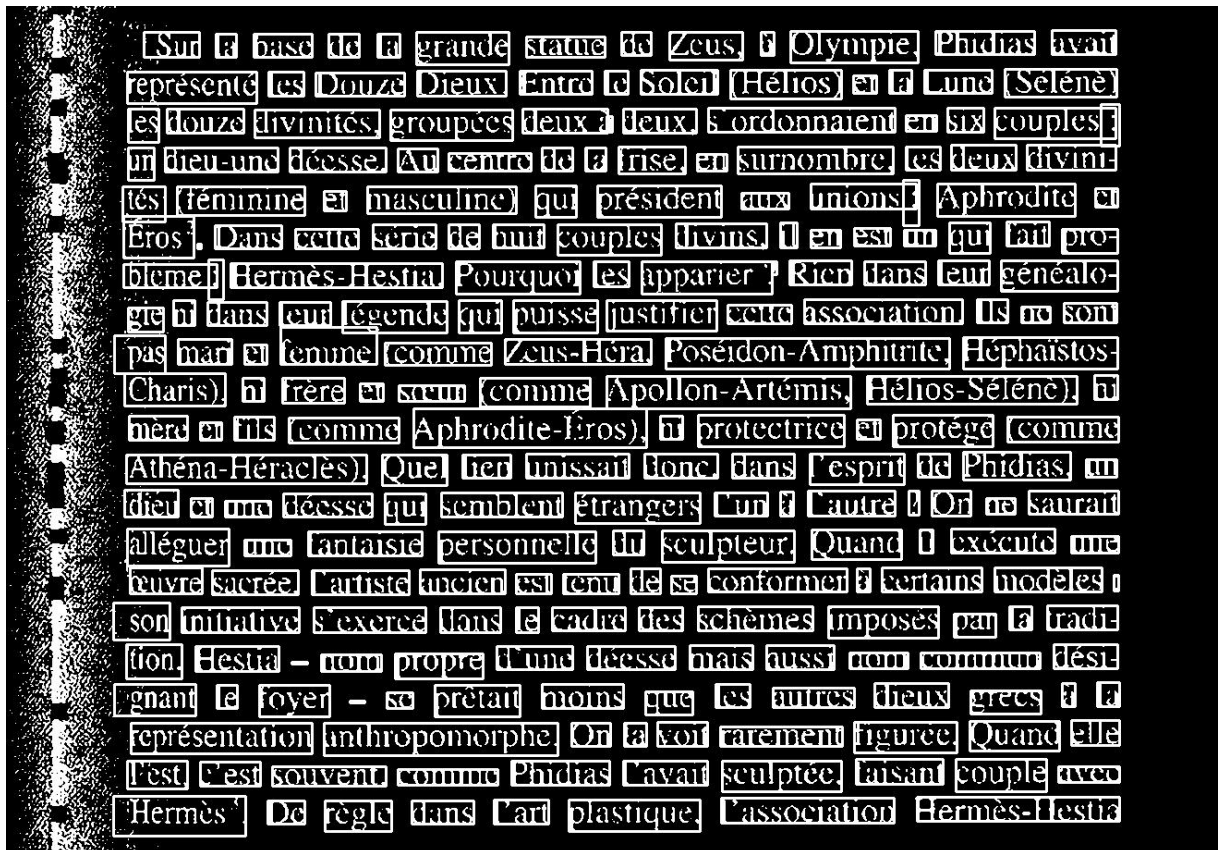


Fig 8.2 - Texto em francês (depois)

a. Texto reconhecido (Francês):

“ Sur la base de la grande statue de Zeus, à Olympie, Phidias avait repr<cnlg les Douze Dieux. Entre le Soleil (Hélios) et la Lune (Sélénè) les douze divinités, groupées deux à deux, s'ordonnaient en six couples : un dieu-une déesse. Au centre de la frise, en surnombre, les deux divinités- (féminine et masculine) qui président aux unions : Aphrodite et Éros . Dans cette série de huit couples divins, il en est un qui fait problème : Hermès-Hestia. Pourquoi les apparier ? Rien dans leur généalogie ni dans leur légende qui puisse justifier cette association. Ils ne sont pas mari et femme (comme Zeus-Héra, Poséidon-Amphitrite, Héphestos-Charis), ni frère et sœur (comme Apollon-Artémis, Hélios-Sélénè), ni mère et fils (comme Aphrodite-Éros), ni protectrice et protégé (comme Athéna-Héraclès). Quel lien unissait donc, dans l'esprit de Phidias, un 'dieu et une déesse qui semblent étrangers l'un à l'autre ? On ne saurait alléguer une fantaisie personnelle du sculpteur. Quand il exécute une œuvre sacrée, l'artiste ancien est tenu de se conformer à certains modèles : -Son initiative s'exerce dans le cadre des schémas imposés par la tradition. Hestia – nom propre d'une déesse mais aussi nom commun désignant le foyer – se prêtait moins que les autres dieux grecs à la représentation anthropomorphe. On la voit rarement figurée. Quand elle l'est c'est souvent comme Phidias l'avait sculptée, faisant couple avec Hermès” De règle dans l'art plastique, l'association Hermès-Hestia

cadre des schèmes imposés par la tradi- 'ion. Hestia — norn propre d'une d'csse mais aussi nom commun dési- "énant le foyer — se prêtait moins que les autres dieux grecs à la représentation anthropomorphe. On la voit rarement figurée. Quand elle Test, c'est souvent, comme Phidias l'avait seulptée, faisant couple avec Hermès". De règle dans l'art plastique. l'association Hermès-Hestia "

b. Texto traduzido (Português-BR):

" Na base da grande estátua de Zeus, em Olímpia, Fídias reproduziu os Doze Deuses. Entre o Sol (Helios) e a Lua (Selene) as doze divindades, agrupadas duas a duas, foram dispostas em seis taças: um deus-um falecido. No centro do friso, em excesso, as duas divindades (feminina e masculina) que presidem as uniões: Afrodite e Eros. Nesta série de oito casais divinos, há um pró-Dlème: Hermes-Héstia. Pomqu01 emparelhá-los? Nada na sua genealogia ou na sua lenda que possa justificar esta associação. Eles não são marido e mulher (como Zeus-Hera, Poseidon-Anfitrite, Hefesto-'Charis), nem irmão e irmã (como Apolo-Ártemis, Hélios-Selene). nem mãe e filho (como Afrodite-Eros), nem protetor e protegido (como 'Atena-Héracles). Que vínculo uniu então na mente de Fídias, um deus e uma deusa que parecem estranhos um ao outro? Não podemos alegar uma fantasia pessoal do escultor. Quando executa uma obra doce, o artista antigo é obrigado a conformar-se com certos modelos: -A sua iniciativa exerce-se no quadro dos esquemas impostos pela tradição. Héstia – nome próprio de um falecido, mas também um nome comum que designa o lar – prestava-se menos do que os outros deuses gregos à representação antropomórfica. Raramente o vemos retratado. Quando o faz, muitas vezes, como Fídias apenas imaginara, forma um casal com Hermes. Regra na arte plástica. a associação Hermès-Héstia "

Observações : Note que a precisão e a redução de erros no reconhecimento do texto contido na imagem foram significativamente melhoradas mediante a aplicação de técnicas simples de pré-processamento.

V. CONCLUSÃO

Não existe um padrão ou uma "silver bullet" (bala de prata) que funcione para qualquer tipo de imagem. Depende muito de como a imagem original é composta. Com isso, não é verdade que as mesmas técnicas que funcionaram para uma determinada imagem funcione para alguma outra.

VI. REFERÊNCIAS

<https://www.youtube.com/@murtazasworkshop>
[OpenCV Lessons - OpenCV Tutorial C++ \(opencv-srf.com\)](#)
[Visão Computacional: reconhecimento de texto com OCR e OpenCV | Alura Cursos Online](#)