

Gestión Documental

Integrantes:

Jean Paul Yepes Guarnizo - Ingeniería de Sistemas
jpyepes@unal.edu.co

Emmanuel Vargas Posada - Ingeniería Eléctrica.
evargaspo@unal.edu.co

Jean Carlos Jimenez Giraldo - Ingeniería Industrial
jecjimenezgi@unal.edu.co

Luis Eduardo Oliveros - Ingeniería Física
loeliveroso@unal.edu.co

Jhon Bryan Yopez Herrera - Ingeniería de Sistemas
jbyepesh@unal.edu.co

Visión Artificial

Docente John Willian Branch Bedoya
Universidad Nacional de Colombia
Sede Medellín
2019-I

1. DESCRIPCIÓN DEL PROYECTO

En todas las organizaciones es necesario realizar una gestión documental, en especial la labor de realizar un balance financiero, este es un proceso lento debido a un gran obstáculo que son las facturas, las facturas no están digitalizadas sino que solo tienen un soporte físico, esto es una dificultad para las empresas porque deben contratar personal para que ellos copien los datos en un computador para su posterior procesamiento, lo que se pretende con este proyecto es agilizar este proceso obteniendo los datos contenidos en la factura por medio de técnicas de visión artificial.

2. ESTADO DEL ARTE

Este problema cae en la categoría de reconocimiento óptico de caracteres o OCR por sus siglas en inglés (Optical Character Recognition), el OCR incluye la transformación de todos los elementos que el ser humano sea capaz de leer hacia una representación que pueda ser manipulada por un computador, el objetivo es traducir imágenes de caracteres impresos o escritos a mano hacia códigos de representación de caracteres como el ASCII, sin la intervención humana^[1].

El sistema para poder obtener los caracteres que puedan ser interpretados por un computador consta de dos partes: un dispositivo para la obtención de imágenes y un algoritmo para el reconocimiento de caracteres, los dispositivos ópticos para la obtención de imágenes en este problema son: una cámara fotográfica o un escáner, ya que tenemos como resultado imágenes del espectro visible que es el mismo en el que los seres humanos interpretan los caracteres.

Los algoritmos para el reconocimiento de caracteres deben seguir las siguientes tareas: el filtrado, la segmentación de caracteres, aclarar los caracteres y el reconocimiento de los caracteres usando inteligencia artificial^[1], de esto último existen muchas implementaciones algunos ejemplos son: amazon textract u office lens.

3. PROPUESTA DE SOLUCIÓN

Debido a esta problemática en las empresas, se tiene como propuesta es que a partir de imágenes tomadas con el celular o escaneadas mostrar las fases necesarias para el reconocimiento de caracteres mencionadas anteriormente, adicionalmente para usar estas técnicas se necesita de un preprocesamiento previo ya que algunas imagenes estan torcidas o en la imagen no se distinguen claramente los caracteres para una adecuada segmentación de las letras.

4. DESARROLLO DE SOLUCIÓN

4.1.1 OBTENCIÓN DE IMÁGENES

Para la obtención de imágenes se tiene considerado dos opciones: cámaras de teléfonos celulares y escáneres. Las cámaras de teléfono es una alternativa más rápida ya que el

proceso de la obtención dura muy pocos segundos, pero las muestras son diferentes en todos los casos, la sección de interés está rotado o se deforman los caracteres debido a que el proceso de obtención es diferente en todos los casos. La imagen 1 fue tomada con la cámara de un iphone 6, esta cámara cuenta con 8 megapíxeles de resolución, apertura de $f/2,2$ y estabilización óptica^[2].

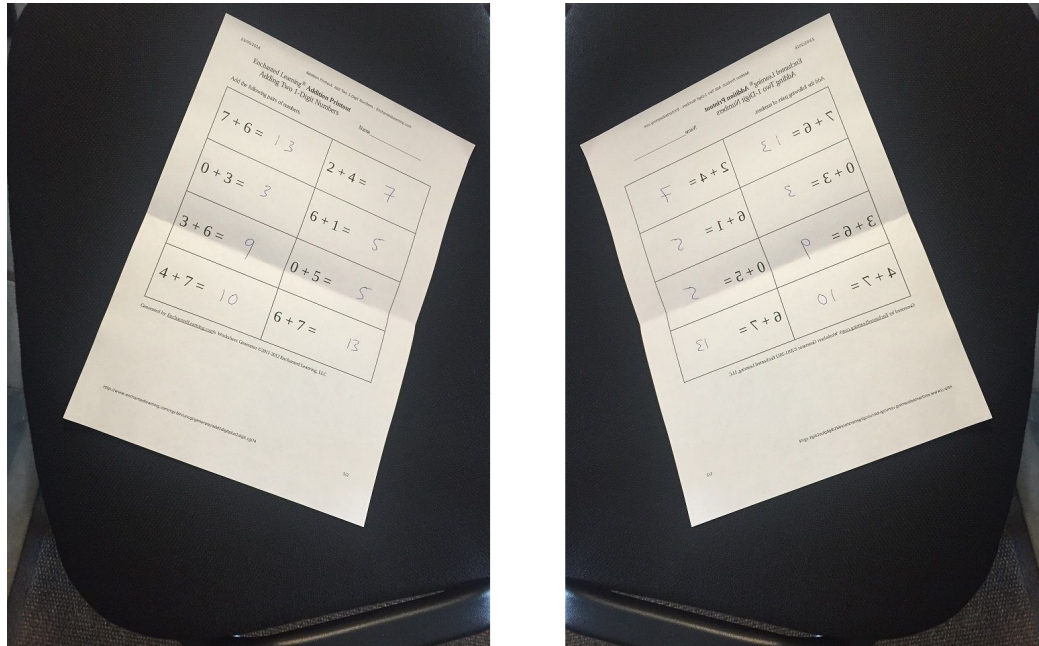


Imagen 1: obtención de imágenes por medio de cámaras de celular

En cambio, los escáneres son más lentos en comparación con los teléfonos ya que dependiendo la densidad de pixeles por pulgadas requieren de más tiempo para su digitalización pero en estos la toma de las muestras es no sufre de los problemas presentados en la cámara de celular hablados anteriormente. La imagen 2 fue tomada con el escáner de la impresora EPSON L4160, esta cuenta con 1200 dpi a una profundidad 48 bits de entrada y 24 bits de salida^[3].

Información importante: Todos los clientes que cuenten con la Tarjeta de Crédito Credencial migrarán al nuevo Programa de Lealtad de los Bancos Aval y podrán ser utilizados en opciones de recompensa que informaremos próximamente. Mayor Información en www.bancodeoccidente.com.co

| CUPO TOTAL | | CUPO AVANCES | | CUPO DISPONIBLE | | SOBRECUPU | | | | | | |
|----------------------------|-----------------|--------------|---------------------------|--------------------|-----------------|---------------------|-----------------|------|-------|--|--|--|
| 1873 | | 1873 | | 0 | | 6.12 | | | | | | |
| DESCRIPCION | No. DE COMPTE | CIUDAD/PAIS | FECHA | VALOR CANCELADO | CARGOS Y ADONOS | SALDO DEPENDIDO | PLAZO | TASA | TASA% | | | |
| CONSUMOS ANTERIORES | | | | | | | | | | | | |
| LINEO COLOMBIA | 817159 | 1100100 BO | 19/11 | 048,000.00 | 5,252.64 | 110,305.51 | 24 | 3 | 28.55 | | | |
| VIVA COLOMBIA | 719732 | 0561500 RI | 24/11 | 98,520.00 | 12,440.00 | 62,200.00 | 8 | 3 | 28.55 | | | |
| COLOMBIA TRAVEL SYSTEM | 711453 | 1100100 BO | 24/11 | 82,000.00 | 13,666.67 | 40,999.99 | 6 | 3 | 28.55 | | | |
| NUEVOS CONSUMOS FACTURADOS | | | | | | | | | | | | |
| IVA COSTO DE REEXPEDICION | 000969 | | 21/01 | 2,800.00 | 2,800.00 | 0.00 | 0 | 0 | 0.00 | | | |
| CARGO POR REEXPEDICION | 000969 | | 21/01 | 17,500.00 | 17,500.00 | 0.00 | 0 | 0 | 0.00 | | | |
| GRACIAS X SU PAGO - VM | 576229 | | 01/02 | 0.00 | -500,000.00 | 0.00 | 0 | 0 | 0.00 | | | |
| ABONO TRAS SLDO FAVOR | 000000 | | 04/02 | 0.00 | -12,333.31 | 0.00 | 0 | 0 | 0.00 | | | |
| GLA TECNOLOGIA | 222550 | 11001 | 11/02 | 1,555,555.56 | 44,444.44 | 1,555,555.56 | 36 | 1 | 35 | | | |
| SALDO A SU FAVOR | SALDO EN MORA | | MORA DESDE | PARA PAGO A PLAZOS | PARA PAGO TOTAL | PARA PAGO ANTES DEL | PARA PAGO | | | | | |
| | \$0.00 | | AÑO MES DIA | CANCELE | CANCELE | AÑO MES DIA | MINIMO ALTERNIO | | | | | |
| | | | | \$111,000.00 | \$1,879,122.90 | 2019 03 04 | \$86,000.00 | | | | | |
| PAQUETE DEL MES | | | | | | | | | | | | |
| SALDO ANTERIOR | \$ 757,120.12 | | | | | | | | | | | |
| (+) TOTAL CARGOS DEL MES | \$ 1,634,258.09 | | CUOTAS DEL MES | \$ 96,103.75 | | | | | | | | |
| (+) AVANCES | \$ 0.00 | | AVANCES | \$ 0.00 | | | | | | | | |
| (-) PAGOS Y CREDITOS | \$ -512,255.31 | | INTERESES CORRIENTES | \$ 13,958.09 | | | | | | | | |
| | | | INTERESES DE MORA | \$ 0.00 | | | | | | | | |
| NUEVO SALDO | \$ 1,879,122.90 | | PARA PAGO A PLAZOS | \$ 111,000.00 | | | | | | | | |
| | | | PARA PAGO MINIMO ALTERNIO | \$ 86,000.00 | | | | | | | | |

Si usted cancela un valor mayor o igual al Pago Mínimo Alternativo y menor al Pago a Plazo, amplía el plazo a 36 cuotas para el saldo a capital de las transacciones que aún tengan al día el día de corte de la obligación vencida de la transacción, conservando la tasa de interés placada. La ampliación de plazo aplica por única vez para cada transacción.

Período por sus compras en 3 días mes: Acumulado Puntos al mes

En caso de mora inferior a dos (2) años, la permanencia del reporte negativo no podrá superar el doble de la mora, contada a partir de la fecha en que sean pagadas las cuotas vencidas o sea pagada la obligación vencida. En meses de mora igual o superior a dos (2) años, la permanencia del reporte negativo será de cuatro (4) años, contada a partir de la fecha en que sean pagadas las cuotas vencidas o sea pagada la obligación vencida.

Imagen 2: obtención de imágenes por medio de escáner

4.2 PREPROCESAMIENTO

4.2.1. Detección de bordes y contornos.

Para la identificación de bordes y contornos se utilizarán convoluciones, en específico los kernels de sobel, verticales y horizontales; y la suma de los dos para detectar bordes en ambos sentidos y también implementaciones que se encuentran en el opencv y son muy eficientes como la función Canny Edge Detection ^[4], que consiste de los siguientes pasos:

1. Aplicar un filtro gaussiano.
2. Encontrar la intensidad del gradiente en cada píxel de la imagen.
3. Aplicar una no máxima supresión que vuelve cero todos los píxeles que no son bordes.
4. Los píxeles que son lo suficientemente grandes para ser considerados bordes se llevan al negro mientras que los que no superen el umbral se llevan al blanco.

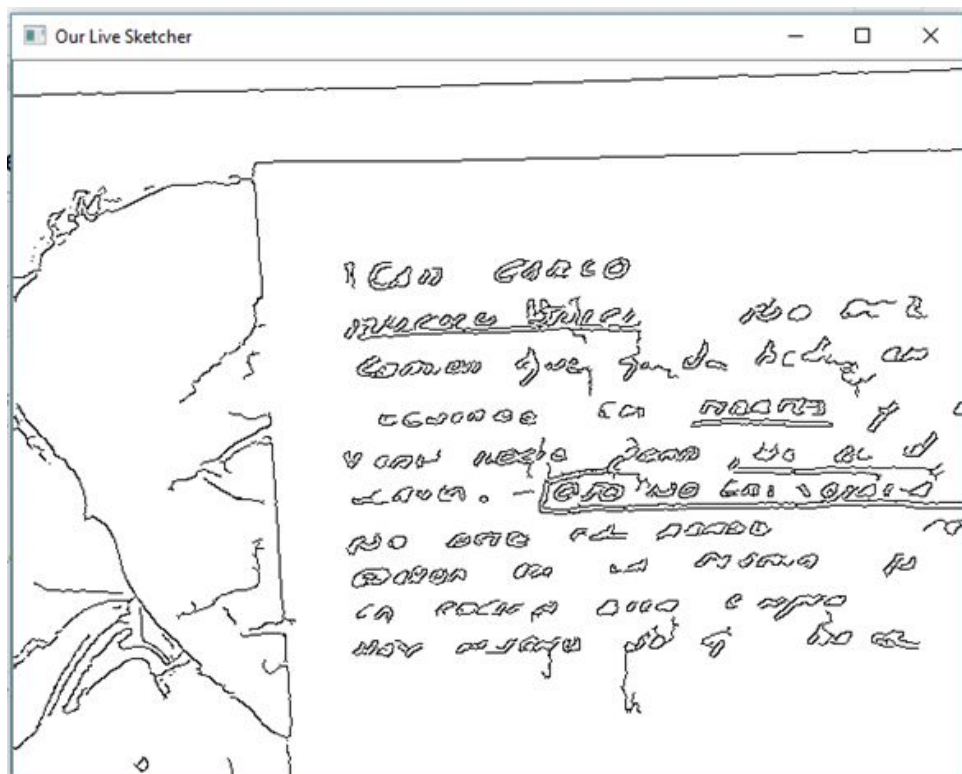


Imagen 3: resultado de la detección de bordes

Después de detectar los bordes es necesario obtener los contornos y clasificarlos, es decir identificar cuántas formas diferentes existen en la imagen, de que tipo son, determinar sus posiciones, sus tamaños, todo lo anterior para poder hacer un reconocimiento más profundo de qué significa cada cosa que aparece en la imagen.

En la siguiente imagen se puede ver cómo se detectan los contornos de una hoja de papel. La Imagen y el código tomado; basado en videos de youtube de Rajeev Ratan ^[5]

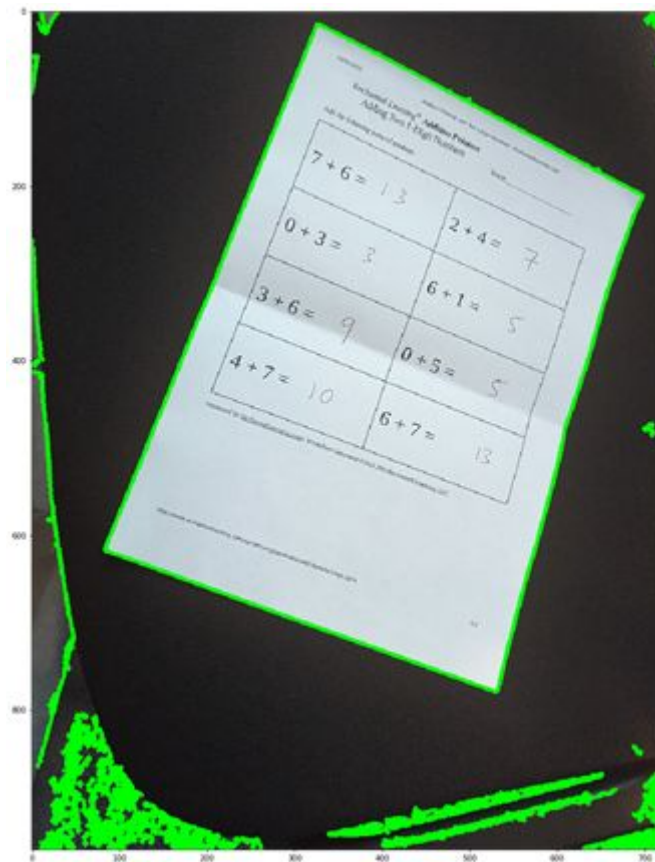


Imagen 4: Implementación de la función contour

Como se puede observar en la imagen anterior la función contour de opencv detecta el borde de la hoja y es posible que dicha función también detecte lo que hay en el interior de un contorno clasificandolo como padre e hijo, pero para este caso solo es de interés el contorno externo, esto se logra configurando ciertos parámetros de la función contour. Después de haber encontrado todos los contornos padres, nos interesa utilizar el de mayor área pues es el que corresponde a la hoja de papel, entonces para hacer esto se calculan las áreas de todos los contornos encontrados con la función contour donde la salida representa cada uno de los contornos dentro del arreglo contours, luego se ordenan de mayor a menor y se selecciona el primero y el resultado obtenido es el siguiente.

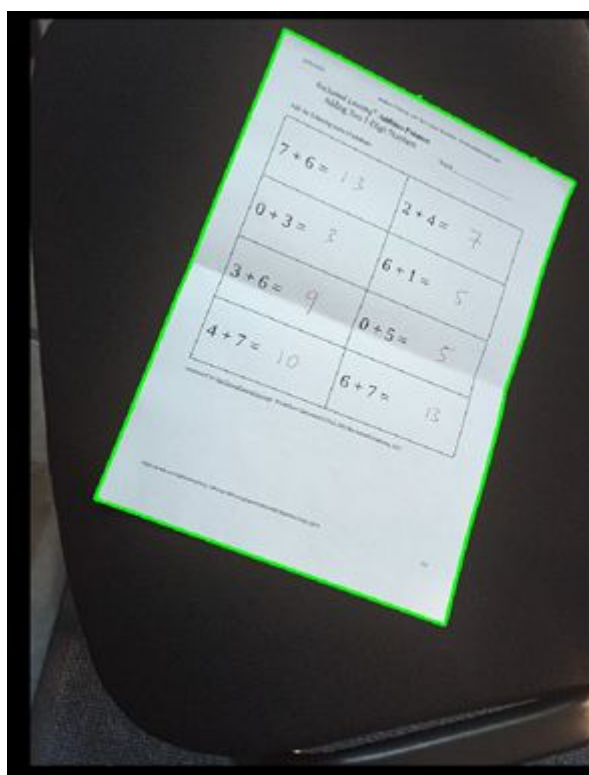


Imagen 5: Implementación de la función contour filtrando por contorno con area mas grande

Dado que los contornos están formados por muchos puntos es conveniente hacer un ajuste con un paralelogramo, ya que las transformaciones de perspectiva trabajan básicamente con solo 4 puntos.

Aunque en la mayoría de los casos la aproximación es suficiente en este no se usó la aproximación dado que el ajuste no resultó ser muy bueno, entonces hallando puntos extremos del contorno se hallaron las cuatro esquinas necesarias para la transformación de perspectiva. Luego el resultado final es la imagen bien orientada y lista para detectar caracteres en ella.

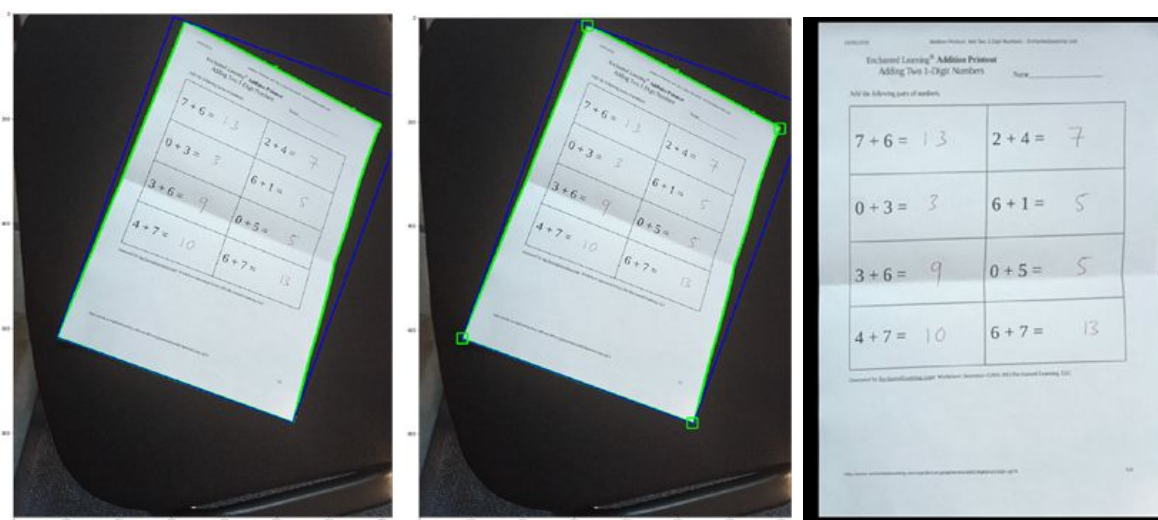


Imagen 6: Implementación de la función contour con puntos extremos y transformación de perspectiva

Lo anterior fue una implementación rudimentaria de la solución a un problema ya muy trabajado para el cual existen una gran cantidad de redes neuronales ya entrenadas. En las siguientes páginas se muestran una serie imágenes que dan muestra del procesamiento que realiza una red neuronal ya entrenada no sólo con el documento antes mostrado sino con otros tipos de documentos.



Imagen 7: Reconocimiento de contornos e inclinación con red neuronal

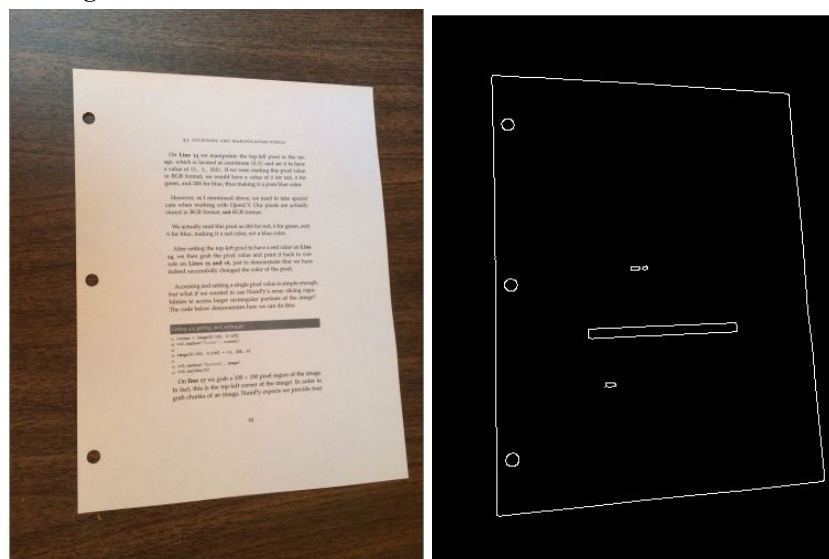


Imagen 8: reconocimiento de bordes

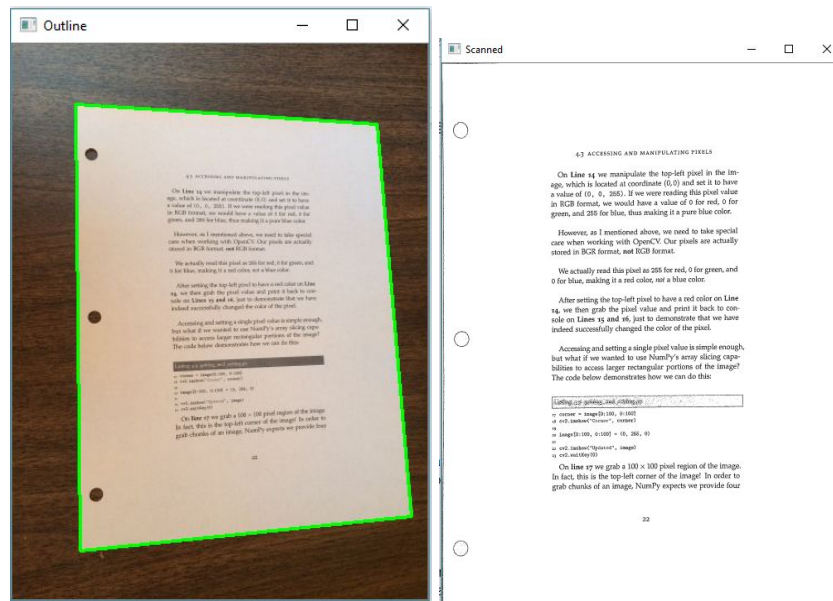


Imagen 9: Reconocimiento de de contorno y solución de perspectiva con red neuronal
En el siguiente grupo de imágenes queda demostrado la utilidad de las redes a lo hora de procesar los distintos tipos de documentos.



La red neuronal para enderezar textos fue tomada del libro Deep Learning for Computer Vision with Python ^[6]

Como se puede apreciar las redes neuronales son supremamente eficientes segmentando una vez las imágenes son legibles y se ha superado el preprocesamiento. Otro ejemplo es una red neuronal que ha sido entrenada para reconocer 80 cosas distintas. Esta también fue tomada de [6]

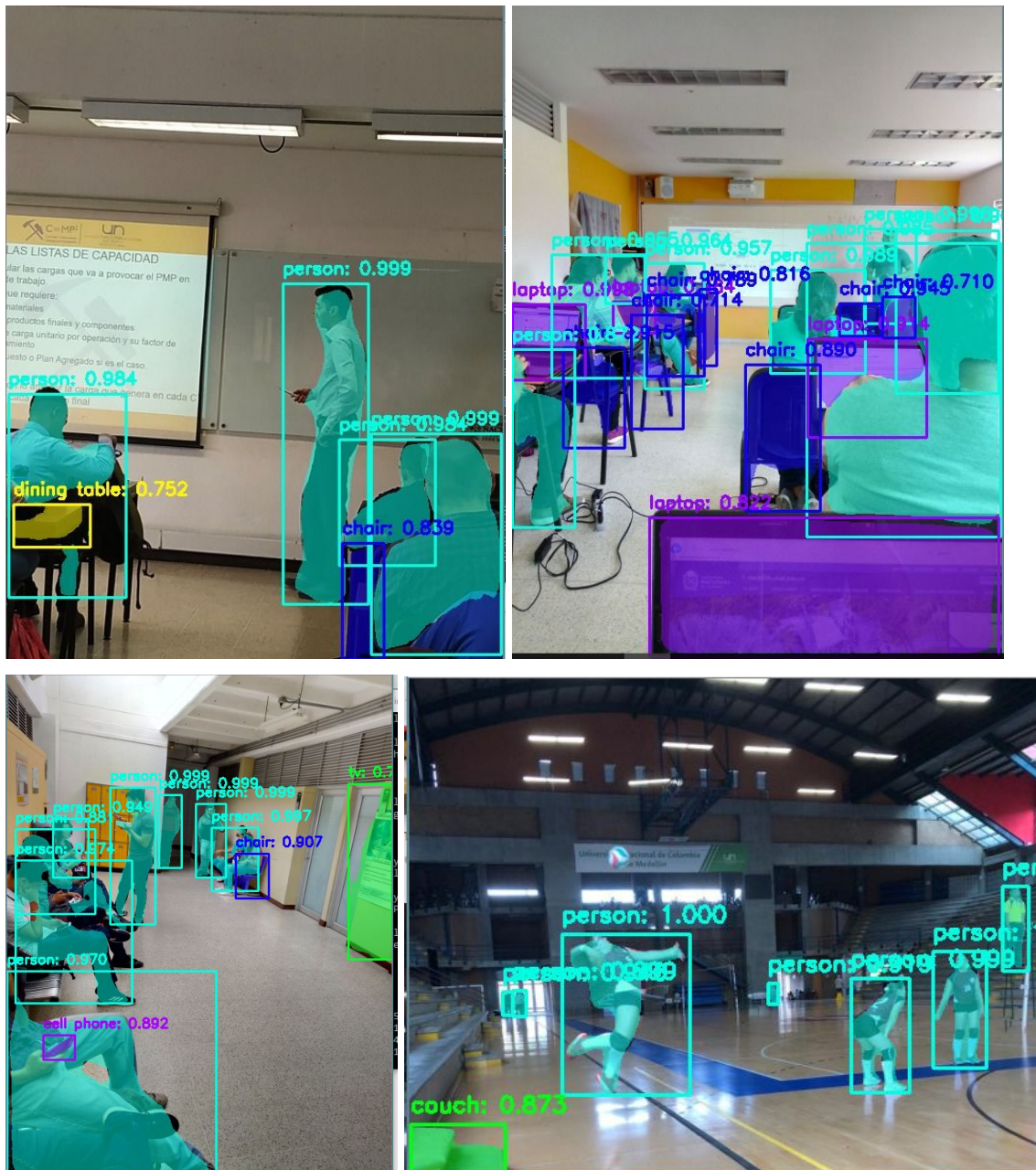


Imagen 11: Muestra de cómo funciona la identificación con la red R-CNN.

4.3. Detección de texto

La detección de texto consiste en ubicar las regiones de una imagen en las que hay texto, sin necesidad de descifrar los caracteres todavía. Dicha detección es indispensable si se quiere reconocer texto en vídeos, como prerequisite para el subsecuente procesamiento de dicho texto mediante reconocimiento de caracteres. En imágenes estáticas también es útil, para después ahondar los esfuerzos de reconocimiento de caracteres en áreas específicas de la imagen en lugar de buscar en la imagen completa. Estamos usando una implementación del detector de texto EAST de OpenCV.



Imagen 12: Identificación de palabras y caracteres.

4.4 Reconocimiento de caracteres

Después de contornear la imagen, entonces se requiere de una red neuronal entrenada para reconocimiento de caracteres y texto, mejor llamadas OCR, las hay de muchos tipos, pero la más poderosa y que es libre es la tesseract, que fue desarrollada por HP en 1985 inicialmente solo para caracteres de la lengua inglesa, luego fue dejada y retomada más tarde por Google, entrenándola para reconocer casi todos los lenguajes del mundo, es libre y gratuita y se puede correr en python.

CONFECCIONES "SAN JORGE" S.A.C.
 Calle Real N° 1045 - Huancayo
 Huanayo, 15 de octubre del 2013
 Señor (es): Gloria Meza Artica RUC: 10198056570
 Dirección: Av. Giraldez N° 458 - Huancayo G. REMISIÓN

FACTURA
 N° 001 - 0000765

| CANTIDAD | DESCRIPCIÓN | PRECIO UNITARIO | VALOR DE VENTA |
|---|---------------------------|-----------------|----------------|
| 04 | Chompas de dralom | 70.00 | 280.00 |
| 02 | Chompas de Orlan | 50.00 | 100.00 |
| 05 | Chompas de lana de alpaca | 120.00 | 600.00 |
| Son: Un mil ciento cincuenta y seis con 40/100 nuevos soles | | Sub Total | 980.00 |
| | | IGV 18 % | 176.40 |
| | | Total S/ | 1,156.40 |

RUC 2054545476
 Serie 001 del 0001 al 1000
 F 117 08 2013
 NRO. AUT. 0123456789

EMISOR
SUNAT

Imagen 13: descripción

Windows PowerShell

```
PS C:\Users\Home\Documents\Universidad\2019-1\vision artificial\reconocimiento inteligente\opencv-text-recognition>
>> python text_recognition.py --east frozen_east_text_detection.pb --image images/factura.jpg
[INFO] loading EAST text detector...
OCR TEXT
=====
290248454757
```

Text Detection

CONFECCIONES "SAN JORGE" S.A.C.
 Calle Real N° 1045 - Huancayo
 Huanayo, 15 de octubre del 2013
 Señor (es): Gloria Meza Artica RUC: 10198056570
 Dirección: Av. Giraldez N° 458 - Huancayo G. REMISIÓN

FACTURA
 N° 001 - 0000765

| CANTIDAD | DESCRIPCIÓN | PRECIO UNITARIO | VALOR DE VENTA |
|---|---------------------------|-----------------|----------------|
| 04 | Chompas de dralom | 70.00 | 280.00 |
| 02 | Chompas de Orlan | 50.00 | 100.00 |
| 05 | Chompas de lana de alpaca | 120.00 | 600.00 |
| Son: Un mil ciento cincuenta y seis con 40/100 nuevos soles | | Sub Total | 980.00 |
| | | IGV 18 % | 176.40 |
| | | Total S/ | 1,156.40 |

RUC 2054545476
 Serie 001 del 0001 al 1000
 F 117 08 2013
 NRO. AUT. 0123456789

EMISOR
SUNAT

Imagen 14: descripción

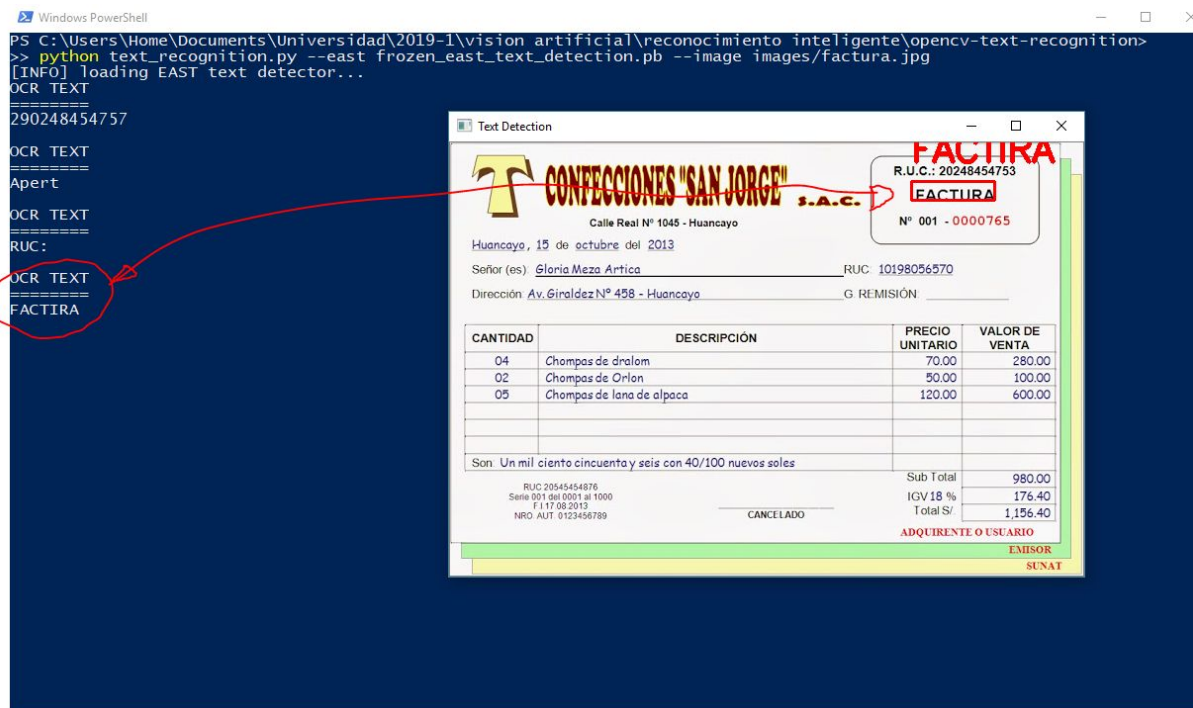


Imagen 15: descripción

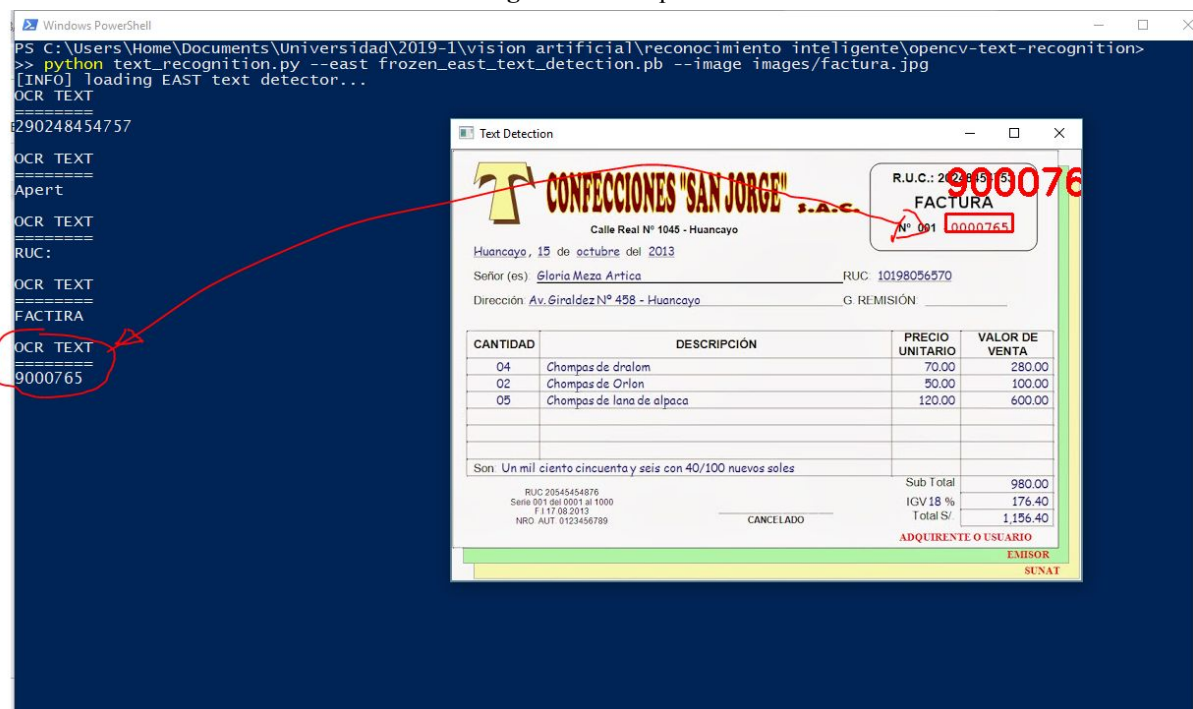


Imagen 16: descripción

5. REFERENCIAS

- [1] Shah, P., Karamchandani, S., Nadkar, T., Gulechha, N., Koli, K., & Lad, K. (2009). OCR-based chassis-number recognition using artificial neural networks. In ICVES 2009 - 2009 IEEE International Conference on Vehicular Electronics and Safety (pp. 31–34). <https://doi.org/10.1109/ICVES.2009.5400240>
- [2] Soporte apple. (2014). iPhone 6 - Especificaciones técnicas. [online] Disponible en: https://support.apple.com/kb/SP705?locale=en_US&viewlocale=es_ES [Accedido 28 Jun. 2019].
- [3] Soporte Epson. (2018). [online] Disponible en: https://mediaserver.goepson.com/ImConvServlet/imconv/bd1eade187740bc38d3560c3f56113e78988f6a6/original?assetDescr=ESPA%C3%91OL-CATSHEET-L4160_final [Accedido 29 Jun. 2019].
- [4] John F. Canny, “OpenCV: Canny Edge Detection,” 1986. [Online]. Available: https://docs.opencv.org/3.1.0/da/d22/tutorial_py_canny.html. [Accessed: 29-Jun-2019].
- [5] R. Ratan, *Ratan, Rajeev*. 2019.
- [6] D. A. Rosebrock, *Deep Learning for Computer Vision with Python*, 1st Editio. PUBLISHED BY PYIMAGESEARCH, 2017.