Identificación de placas vehiculares en Colombia a partir de reconocimiento óptico de caracteres ROC

Juan M. Aya Perlaza

Departamento de Ingeniería Electrónica Pontificia Universidad Javeriana

Laura Joaquí Muñoz

Departamento de Ingeniería Electrónica Pontificia Universidad Javeriana

Victor Múnera Rojas

Departamento de Ingeniería Electrónica Pontificia Universidad Javeriana

Inteligencia Artificial – Procesamiento de Imágenes y Visión 11 de septiembre de 2020

1. Introducción

La integración masiva de las tecnologías de la información, bajo diferentes aspectos del mundo moderno, ha llevado a tratar los vehículos como recursos conceptuales en los sistemas de información. Dentro de los equipos inteligentes se menciona el sistema de detección y reconocimiento de matrículas de vehículos.

Este sistema se utiliza para detectar las placas y luego hacer el reconocimiento de la matrícula que consiste en extraer el texto de una imagen y todo ello gracias a los módulos de cálculo que utilizan algoritmos de localización, segmentación de la matrícula y reconocimiento de caracteres. La detección y lectura de matrículas es un tipo de sistema inteligente y es considerable por las posibles aplicaciones en varios sectores, por ejemplo:

- *Cuerpo policial*: Este sistema se utiliza para la detección de vehículos robados y registrados. Las matrículas detectadas se comparan con las de los vehículos denunciados.
- Gestión de parqueaderos: La gestión de las entradas y salidas de los vehículos.
- Seguridad vial: Este sistema se utiliza para detectar matrículas que superan una cierta velocidad.

2. Objetivos

2.1. Objetivo General

• Implementar un sistema que permita reconocer los caracteres presentes en placas vehiculares colombianas.

2.2. Objetivos Específicos

- Identificar en la imagen la región donde se encuentra la placa del vehículo.
- Segmentar la placa en los diferentes caracteres que la componen.
- Utilizar un algoritmo de reconocimiento óptico de caracteres para obtener las letras y números de la placa.
- Diseñar un protocolo de pruebas que permita validar el correcto funcionamiento del sistema.

3. Marco teórico

3.1. Placas vehiculares colombianas

Los residentes en Colombia son requeridos por las autoridades correspondientes para registrar sus vehículos motorizados y exhibir la matrícula automovilística (o Placas, como se les conoce comúnmente). El actual diseño de placas fue introducido en los años 90, aunque algunas matrículas con el diseño anterior (usado entre 1972 y 1990) siguen en uso. [1]

Categoría	Imagen	Tipo	Colores	Formato de serie
Pasajeros	CAN-732	Particular	negro sobre amarillo	ABC·123 ¹
	XZF 607	Comercial	negro sobre blanco ^{1 2}	

Diseño de placas colombianas. Imagen de Wikipedia.

3.2. Reconocimiento automático de matrículas

El Reconocimiento Automático de Matrículas (ANPR) es un sistema de gran precisión capaz de leer las matrículas de los vehículos sin intervención humana mediante el uso de captura de imágenes a alta velocidad con iluminación de apoyo, detección de caracteres dentro de las imágenes proporcionadas, verificación de las secuencias de caracteres como si fueran de una matrícula de vehículo, reconocimiento de caracteres para convertir la imagen en texto; terminando así con un conjunto de metadatos que identifican una imagen que contiene una matrícula de vehículo y el texto decodificado asociado de esa matrícula. [2]

3.3. Segmentación

La segmentación de imágenes divide la imagen en sus partes constituyentes hasta un nivel de subdivisión en el que se aíslen las regiones u objetos de interés.

Los algoritmos de segmentación se basan en una de estas dos propiedades básicas de los valores del nivel de gris: discontinuidad o similitud entre los niveles de gris de píxeles vecinos. [3]

- Discontinuidad: Se divide la imagen basándose en cambios bruscos de nivel de gris:
 - Detección de puntos aislados
 - Detección de líneas
 - Detección de bordes
- Similitud: Se divide la imagen basándose en la búsqueda de zonas que tengan valores similares, conforme a unos criterios prefijados:
 - o Crecimiento de región
 - Umbralización

3.4. Reconocimiento óptico de caracteres ROC

El reconocimiento óptico de caracteres (ROC), generalmente conocido como reconocimiento de caracteres y expresado con frecuencia con la sigla OCR (del inglés Optical Character Recognition), es un proceso dirigido a la digitalización de textos, los cuales identifican automáticamente a partir de una imagen símbolos o caracteres que pertenecen a un determinado alfabeto, para luego almacenarlos en forma de datos. [4]

4. Descripción de la solución

4.1. Propuesta

Se propone diseñar e implementar un sistema de software que permita, a partir de muestras fotográficas tomadas a vehículos en calles colombianas, identificar, segmentar y, para cada letra o número incluido en la placa de tránsito en cuestión, implementar un algoritmo de reconocimiento óptico de caracteres para poder conocer la matrícula del vehículo.



Imagen tomada de https://www.syscomblog.com/2018/10/software-para-reconocimiento-automatico.html

Se procederá a tomar la información de las bases de datos provistas por el docente, para luego realizar la identificación de los objetos en las imágenes y discriminar las placas. Luego de esto, se segmentará la imagen con el fin de procesar las respectivas partes que la componen, con el fin de conocer la placa del vehículo, empleando un algoritmo de reconocimiento óptico de caracteres.

4.2. Diagrama en Bloques

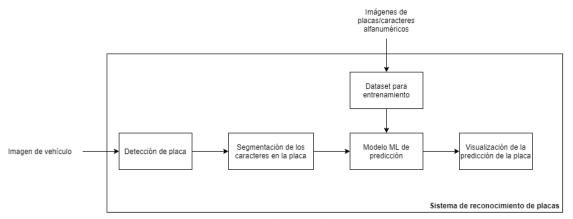


Diagrama en bloques del sistema

Como se puede observar en la imagen anterior, se diseñó un sistema que permita reconocer y predecir los caracteres de una placa vehicular colombiana a partir del procesamiento de una imagen inicial que contenga un vehículo y mediante el entrenamiento de algunos algoritmos de aprendizaje de máquina.

Para conseguir esto, se parte de la imagen inicial en la cual se identifica el lugar en el que se encuentra la placa a través del reconocimiento del contorno con un área y dimensiones relativas al tamaño de la imagen, luego, conociendo el lugar donde está la placa se toma este nuevo recorte de imagen y, mediante el mismo algoritmo, se detectan cada una de las letras/números presentes en ella. Esto permite generar 6 nuevos recortes de imagen, cada uno pasará por un modelo de Aprendizaje de Máquina previamente entrenado para poder predecir a qué caracter alfanumérico hace referencia.

4.3. Diseño de la solución

Para el diseño del sistema, se hizo uso de las librerías de OpenCV para toda la parte de Procesamiento de Imágenes, NumPY y Sklearn para el apartado de Inteligencia Artificial y, por último, glob para el manejo de directorios.

Los diagramas de flujo del sistema se presentan a continuación:

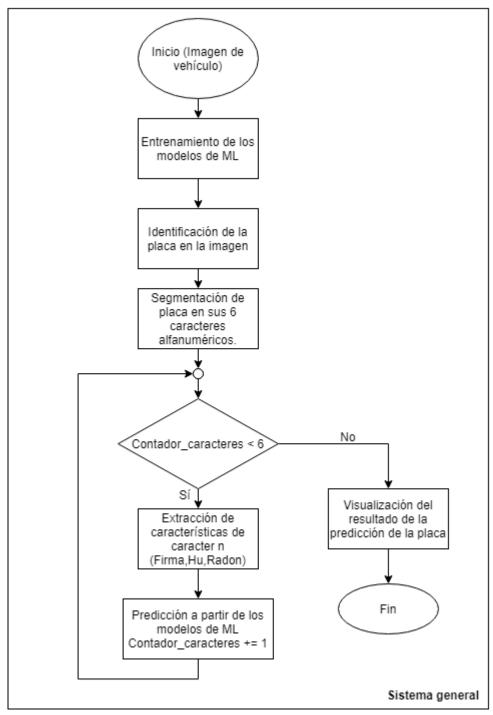


Diagrama de flujo del sistema general

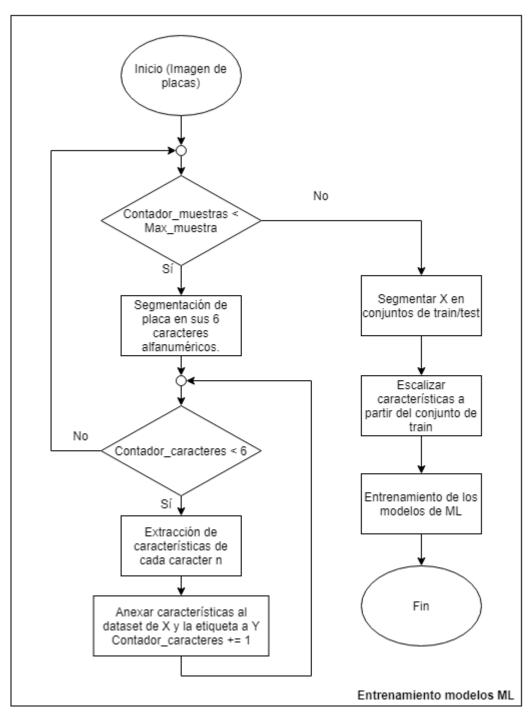


Diagrama de flujo del entrenamiento de los modelos de ML

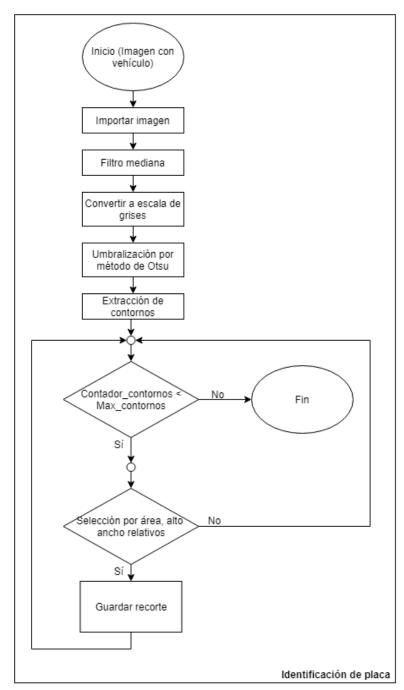


Diagrama de flujo identificación de placa en imagen

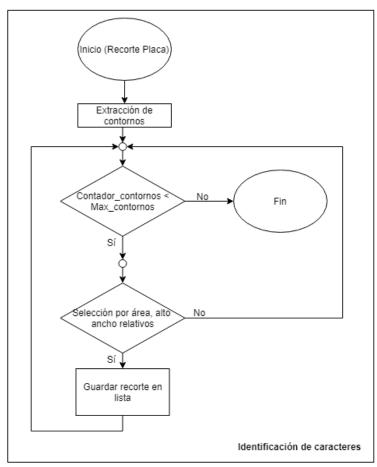
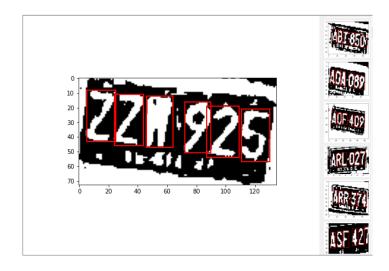


Diagrama de flujo identificación de caracteres en placa.

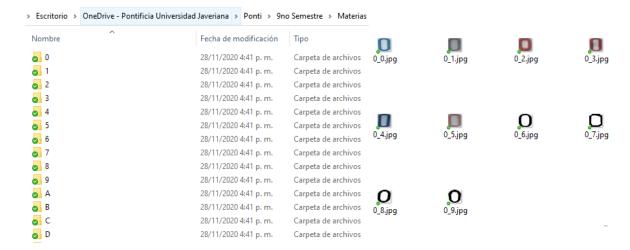
Entrenamiento de los modelos de ML

Para el entrenamiento, se utilizaron dos bases de datos:

Primero, se utilizaron 162 fotos de placas tomadas por cámaras en semáforos de la ciudad de Bogotá por la secretaría de movilidad. Esta base de datos resulta confidencial, por eso no se pueden subir las muestras, sin embargo, el método a utilizar fue segmentar cada imagen en los respectivos caracteres y, por último, a cada carácter sacarle características para conformar el dataset junto a sus etiquetas. Se obtuvieron un total de 972 letras.



Por otra parte, se utilizó un dataset encontrado en github de un proyecto similar, el cual contenía 10 imágenes de 20x20 pixeles para cada carácter alfanumérico.



Validación de modelos.

Se implementaron 3 diferentes algoritmos de aprendizaje de máquina, a saber: KNN, SVM y Red neuronal.

Para cada una, se variaron sus respectivos hiperparámetros para encontrar cuál era el mejor modelo para el problema en cuestión. Finalmente, se obtuvieron los siguientes resultados:

```
El mejor resultado fue: [1, 0.9647058823529412]
[[4 0 0 ... 0 0 0]
[0 4 0 ... 0 0 0]
[0 0 2 ... 0 0 0]
...
[0 0 0 ... 3 0 0]
[0 0 0 ... 0 3 0]
[0 0 0 ... 0 0 4]]
```

Modelo KNN. Mejor modelo con k=1 y accuracy del 96% en conjunto de validación.

```
El mejor Kernel es linear con un accuracy de: 0.9882352941176471
Accuracy of SVM classifier on test set: 0.99
```

Modelo SVM. Mejor modelo con kernel lineal y accuracy de 99% en conjunto de validación.

```
Accuracy por Red Neuronal con sklearn en conjunto de entrenamiento: 1.00
Accuracy por Red Neuronal con sklearn en conjunto de validación: 0.72
[[3 0 0 ... 0 0 0]
[0 2 0 ... 0 0 0]
[0 0 2 ... 0 0 0]
...
[0 0 0 ... 1 0 0]
[0 0 0 ... 0 3 0]
[0 0 0 ... 0 0 4]]
```

Modelo NN. Modelo con 5 capas de (40,40,30,15,8) neuronas y accuracy de 72% en conjunto de validación.

Detección de placas

A continuación, se pueden observar los resultados de la identificación de las placas en las fotos a partir de las dimensiones en pixeles de los contornos encontrados.



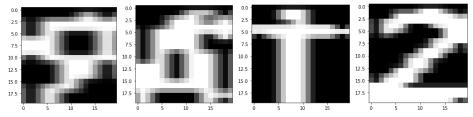
Segmentación de placas en caracteres

Con un procesamiento similar, en cada recorte de placa identificada, se procede a detectar cada uno de los 6 caracteres presentes en la matrícula. Los resultados se observan a continuación:



Predicción

Para la predicción, cada una de las letras detectadas se reescalizaban a un tamaño de 20x20, se les sacaba las características (Momentos de Hu, Transformada de Radón y Firma), se escalizaban las características a partir del conjunto de entrenamiento previo y se ingresaban a la función predict de cada modelo.



Resultados

Después de todo el proceso previamente mencionado, se obtuvieron las siguientes predicciones para las respectivas placas:



5. Conclusiones

- Para poder encontrar los contornos es preciso convertir la imagen a un formato binario, razón por la cuál fue necesario encontrar un método de umbralización que permitiera realizar este proceso. Después de distintas pruebas se encontró con que el método de Otsu era el que arrojaba mejores resultados, ya que los otros perdían ciertos objetos importantes de la foto o se mezclaba la placa con el carro.
- De los distintos métodos para identificar los posibles contornos de placas en una imagen, se decidió utilizar el área, alto y ancho de contornos para poder distinguirlos de los demás. Además, como las imágenes eran de diferentes tamaños, era preciso definir unos umbrales relativos al tamaño de esta, no absolutos.
- Por último, una vez se tenían los recortes por carácter, se tenía que definir cuáles iban a ser las características que se tomarían de cada uno estos para luego poder realizar la inteligencia artificial poder predecir qué letra o número estaba presente en dicha imagen. Fue aquí cuando inicialmente se decidió tomar los 7 momentos de Hu, sin embargo, estas resultaban muy pocas características, razón por la cual se decidió contar también con la transformada de Radón y la firma de cada letra, que era simplemente tomar toda la matriz 20x20 y hacerla de una sola dimensión.
- Las características tomadas a cada recorte de imagen (caracter) hacían referencia a 3 métodos distintos, los momentos de Hu, Transformada de Radón y la firma de la imagen. Primero, se realizó un entrenamiento con cada uno de los métodos independientemente y, por último, se juntaron todos los métodos en el dataset de características, arrojando los siguientes resultados:

Método	Accuracy en conjunto de test	
Momentos de Hu	38%	
Transformada de Radón	66%	
Firma	66%	
Concatenación de todos	66%	

A raíz de estos resultados, al ver que juntando todos los métodos no aumentaba la exactitud del modelo, entonces se decidió utilizar únicamente la firma que consiste en aplanar la matriz de imagen a una sola dimensión y así no aumentar el tiempo de procesamiento del algoritmo con los otros métodos que sí implican el llamado a otras funciones.

- Por otra parte, inicialmente se realizó la creación del dataset de entrenamiento a partir de la base de datos dada por el docente, con imágenes confidenciales de placas de la ciudad. Sin embargo, de todo el conjunto de imágenes, solo 162 funcionaron y estas no fueron suficientes para lograr un buen entrenamiento de los modelos, razón por la cual el valor máximo de exactitud al hacer la validación era de 66%. Para dar solución a este inconveniente, se procedió a buscar otro dataset en internet con imágenes de caracteres y se encontró una carpeta en GitHub con 10 muestras para cada letra y número. Al hacer el entrenamiento y validación con este nuevo dataset se obtuvieron exactitudes de 98 y 99%.
- Consecuentemente, la predicción de las letras cuando se hizo la prueba del sistema con diferentes placas fue mucho más acertada, logrando predecir entre 3 y 4 caracteres de los presentes en cada placa.

Referencias

- [1] Ministerio de Transporte Colombiano, Resolución No. 1690 del Ministerio de Transporte de Colombia, "Por la cual se adopta la ficha técnica Placa Única Nacional para vehículos de servicio Diplomático, Consular y de Misiones Especiales acreditados ante el Gobierno de Colombia." Bogotá, DC.
- [2] M. SYSTEMS, "What is Automatic Number Plate Recognition (ANPR)?" https://www.anprcameras.com/about-us/understanding-anpr/ (accessed Sep. 10, 2020).
- [3] "Tema 4:Segmentación de imágenes."

[4] "Reconocimiento óptico de caracteres - Wikipedia, la enciclopedia libre." https://es.wikipedia.org/wiki/Reconocimiento_óptico_de_caracteres (accessed Sep. 10, 2020).

```
# -*- coding: utf-8 -*-
Created on Thu Nov 26 11:54:16 2020
@author: JuanManuel
from skimage.io import imread
from skimage.filters import threshold otsu
from skimage.transform import resize
from skimage import measure
from skimage.measure import regionprops
import matplotlib.pyplot as plt
import numpy as np
import cv2
import matplotlib.patches as patches
import os
import glob
from sklearn import svm
from skimage.transform import radon
# Modelos a comparar
from sklearn.neural network import MLPClassifier
from sklearn.neighbors import KNeighborsClassifier
from sklearn.preprocessing import MinMaxScaler
from sklearn.model selection import train test split
# Métricas a utilizar
from sklearn.metrics import classification report
from sklearn.metrics import confusion matrix
from sklearn.metrics import matthews corrcoef
letters = ['0', '1', '2', '3', '4', '5', '6', '7', '8', '9', 'A', 'B', 'C', 'D', 'E', 'F', '
def KNN(X train, X test, y train, y test):
    #Obteniendo el mejor valor de K a partir del conjunto de validación con mejor accuracy
    scores = []
    maxscore = [0,0] # [k,score]
    for k in range(1, 40):
        knn = KNeighborsClassifier(n neighbors = k,weights='distance',metric='euclidean', п
        knn.fit(X_train, y_train)
        score = knn.score(X test, y test)
        scores.append(score)
        if score > maxscore[1]:
            maxscore = [k, score]
    plt.figure()
    plt.xlabel('Parametro k')
    plt.ylabel('Accuracy')
    plt.scatter(range(1,40), scores)
    print('El mejor resultado fue:', maxscore)
    knn = KNeighborsClassifier(n neighbors = maxscore[0], weights='distance', metric='euclide'
    knn.fit(X train, y train)
    pred = knn.predict(X test)
    # Matriz de confusion
    print(confusion_matrix(y_test, pred))
    # Reporte de clasificación
```

```
print(classification report(y test, pred))
    # Matthews
    print('Matthews: ',matthews_corrcoef(y_test, pred))
    return knn
def SVM(X_train, X_test, y_train, y_test):
    kernels=['linear', 'poly', 'rbf', 'sigmoid']
    max score = 0
    max kernel = 500
    max degree = 500
    for Kernel in range(4):
        if Kernel == 1:
            for Degree in range(1,11):
                msv = svm.SVC(kernel=kernels[Kernel], degree=Degree)
                msv.fit(X_train, y_train)
                score = \overline{msv.score}(\overline{X}_{test}, y_{test})
                if score > max score:
                    max score = score
                    max kernel = Kernel
                    max degree = Degree
        else:
            msv = svm.SVC(kernel=kernels[Kernel])
            msv.fit(X train, y train)
            score = msv.score(X test, y test)
            if score > max score:
                    max score = score
                    max_kernel = Kernel
    print('El mejor Kernel es', kernels[max_kernel],'de grado',max_degree,'con un accuracy
    msv = svm.SVC(kernel=kernels[max kernel])
    msv.fit(X_train, y_train)
    print('Accuracy of SVM classifier on test set: {:.2f}'
         .format(msv.score(X_test, y_test)))
    return msv
def NN(X_train, X_test, y_train, y_test):
    clf = MLPClassifier(hidden layer sizes=(40,40,30,15,8),activation='relu',random state=1
    clf.fit(X train, y train)
    print('Accuracy por Red Neuronal con sklearn en conjunto de entrenamiento: {:.2f}'.form
    print('Accuracy por Red Neuronal con sklearn en conjunto de validación: {:.2f}'.format(
    # Se generan las métricas
    pred = clf.predict(X test)
    # Matriz de confusion
    print(confusion matrix(y test, pred))
    # Reporte de clasificación
    print(classification_report(y_test, pred))
    # Matthews
    print('Matthews: ',matthews corrcoef(y test, pred))
    return clf
def read training data(training directory):
    image data = []
    target data = []
    for each letter in letters:
        for each in range(10):
            image_path = os.path.join(training_directory, each_letter, each_letter + '_' +
            # read each image of each character
```

```
img details = imread(image path, as gray=True)
            binary image = img details < threshold otsu(img details)</pre>
             flat bin image = binary image.reshape(-1)
            image data.append(flat bin image)
             target data.append(each letter)
    abc = ['0', '1', '2', '3', '4', '5', '6', '7', '8', '9', 'A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K', 'L', 'M', 'N', '0', 'P', 'Q', 'R', 'S' 'U', 'V', 'W', 'X', 'Y', 'Z']
    nums = [0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29]
    for letra in range(len(target data)):
        for x in range(len(abc)):
            if target data[letra] == abc[x]:
                 target data[letra] = nums[x]
    return (np.array(image data), np.array(target data))
def Imp Placa(knn, msv, clf):
    print('\nPredicciones:')
    print('----')
    print('KNN\t |SVM\t |NN')
    for x in range(len(knn)):
        print(knn[x],end='')
    print('\t',end='|')
    for x in range(len(knn)):
        print(msv[x],end='')
    print('\t',end='|')
    for x in range(len(knn)):
        print(clf[x],end='')
    print('\n-----')
# MAIN
print('Reading data')
#Xtotal, ytotal = Base Datos()
training dataset dir = './training images'
Xtotal, ytotal = read training data(training dataset dir)
print('Reading data completed')
# Se generan los conjuntos de entrenamiento y validación
X_train, X_test, y_train, y_test = train_test_split(Xtotal, ytotal,random_state = 1)
scaler= MinMaxScaler()
scaler.fit(X_train)
X train = scaler.transform(X train)
X test = scaler.transform(X test)
knn = KNN(X_{train}, X_{test}, y_{train}, y_{test})
msv = SVM(X_train, X_test, y_train, y_test)
clf = NN(X train, X test, y train, y test)
for name in glob.glob('placas/*'):
    car image = imread(name)
    gray = cv2.cvtColor(car image, cv2.COLOR BGR2GRAY)
    gray = cv2.GaussianBlur(gray, (1,1),0)
    ret, binaria = cv2.threshold(gray, 0, 255, cv2.THRESH BINARY INV+cv2.THRESH OTSU)
    threshold value = threshold otsu(gray)
    binary_car_image = gray > threshold_value
    # fig, (ax1, ax2) = plt.subplots(1, 2)
```

```
# ax1.imshow(gray, cmap="gray")
# ax2.imshow(binary car image, cmap="gray")
# this gets all the connected regions and groups them together
label image = measure.label(binary car image)
plate\_dimensions = (0.06*label\_image.shape[0], 0.1*label image.shape[0], 0.06*label image.shape[0]
min_height, max_height, min width, max width = plate dimensions
plate objects cordinates = []
plate like objects = []
fig, (ax1) = plt.subplots(1)
ax1.imshow(car image, cmap="gray")
for region in regionprops(label_image):
    if region.area <50:</pre>
        continue
    min row, min col, max row, max col = region.bbox
    region_height = max_row - min_row
    region width = max col - min col
    if region height >= min height and region height <= max height and region width >=
        plate like objects.append(binary car image[min row:max row,
                                   min col:max col])
        plate objects cordinates.append((min row, min col,
                                          max row, max col))
        rectBorder = patches.Rectangle((min col, min row), max col - min col, max row -
        ax1.add patch(rectBorder)
plt.show()
for i in range(len(plate like objects)):
    license plate = np.invert(plate like objects[i])
    labelled plate = measure.label(license plate)
    character dimensions = (0.3*license plate.shape[0], 0.55*license plate.shape[0], 0.
    min height, max height, min width, max width = character dimensions
    fig, ax1 = plt.subplots(1)
    ax1.imshow(license plate, cmap="gray")
    characters = []
    counter=0
    column_list = []
    for regions in regionprops(labelled plate):
        y0, x0, y1, x1 = regions.bbox
        y0 -= 1
        region_height = y1 - y0
        region width = x1 - x0
        if region height > min height and region height < max height and region width >
            roi = license plate[y0:y1, x0:x1]
            rect\_border = patches.Rectangle((x0, y0), x1 - x0, y1 - y0, edgecolor="red"
                                         linewidth=2, fill=False)
            ax1.add_patch(rect_border)
```

```
# resize the characters to 20X20 and then append each character into the ch
         resized_char = resize(roi, (20, 20))
         image data = resized char.flatten()
         # Tradon = radon(resized char, theta=np.arange(0,210,30), circle=True).flat
         # Firma = resized char.flatten()
         # Hu = cv2.HuMoments(cv2.moments(resized char)).flatten()
         # image data = np.concatenate((Tradon, Firma, Hu), axis=None)
         characters.append(image data)
plt.show()
X = np.array(characters)
X = scaler.transform(X)
knn pred = list(knn.predict(X))
msv pred = list(msv.predict(X))
clf pred = list(clf.predict(X))
letters = ['0', '1', '2', '3', '4', '5', '6', '7', '8', '9', 'A', 'B', 'C', 'D', 'E', 'F', 'G', 'H', 'I', 'J', 'K', 'L', 'M', 'N', '0', 'P', 'Q', 'R', 'U', 'V', 'W', 'X', 'Y', 'Z']
nums = [0,1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,2]
for letra in range(len(knn pred)):
    for x in range(len(nums)):
         if knn pred[letra] == nums[x]:
             knn pred[letra] = letters[x]
         if msv pred[letra] == nums[x]:
             msv pred[letra] = letters[x]
         if clf_pred[letra] == nums[x]:
             clf pred[letra] = letters[x]
```