

Práctica 4: Reconocimiento de objetos

Sistemas de Percepción, 4º GIERM

En esta práctica, se trabajará el reconocimiento de objetos de cara a su clasificación entre diferentes clases. Concretamente, se pretende realizar el reconocimiento de dígitos que pudieran corresponder al contenido de una matrícula.

Se supondrá que la matrícula sólo puede contener los dígitos del 0 al 9 y que la tipografía es siempre la misma. Sin embargo, la "matrícula" puede aparecer con cualquier posición u orientación, así como con variedad de tamaños. Igualmente, se supone que, a la hora de validar el sistema, en una misma imagen pueden aparecer múltiples matrículas con diversos tamaños, orientaciones y posiciones.

1 Generación de imágenes de matrículas virtuales

Para centrar el estudio únicamente en las técnicas de reconocimiento, se evitará, en principio, el empleo de imágenes reales de vehículos en los que aparezca su matrícula. En lugar de ello, se recomienda usar el siguiente procedimiento para generar imágenes sintéticas de "matrículas" variadas, lo cual permitirá simular siempre vistas frontales y sin ningún fondo que dificulte la localización de los objetos de interés para la clasificación.

Haciendo uso de la página web: <https://www.customeuropeanplates.com>, se pueden generar imágenes de "matrículas" particularizadas, usando siempre la tipografía habitual en las matrículas alemanas. Una vez configurada una matrícula ejemplo, se puede realizar un zoom de la zona de la página web que muestra la matrícula, con idea de tener una versión con alta resolución de la misma, pudiendo almacenarla para formar parte de la batería de imágenes a emplear. Sirvan como ejemplo las imágenes mostradas en la figura siguiente.

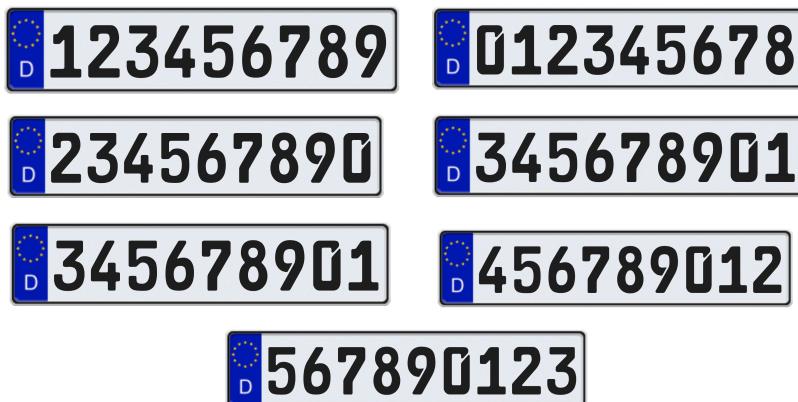


Figure 1: Diversos ejemplos de "matrículas" virtuales.

Siguiendo este procedimiento, se pueden preparar ejemplares de "matrículas" más adecuados para el proceso de entrenamiento del clasificador. Por ejemplo, se pueden preparar imágenes

con matrículas que sólo contengan el dígito correspondiente a cada una de las clases. Véase un ejemplo en la figura 2.

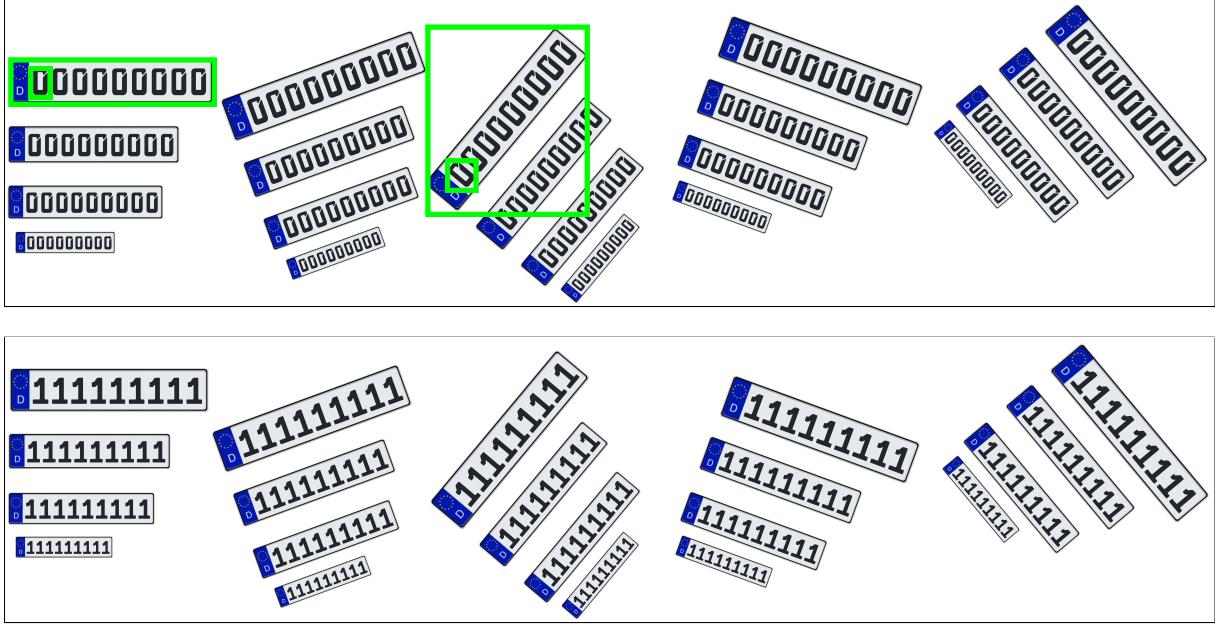


Figure 2: Ejemplos de imágenes conteniendo matrículas con patrones de cada clase para el entrenamiento.

Es interesante usar este tipo de imágenes, en lugar de imágenes de dígitos sueltos, para ayudarse del marco de la matrícula a la hora de reorientar ésta, como paso previo a la extracción de los dígitos. Esto será de gran ayuda para automatizar el proceso de entrenamiento. La idea es generar, para cada dígito, un elevado número de ejemplares, cubriendo diferentes posiciones, orientaciones y tamaños.

Estas imágenes se deben procesar automáticamente, de forma que, a partir de una o varias de ellas para cada clase, se pueda obtener toda la información necesaria para construir el clasificador.

De cara a la validación del clasificador, se empleará otro tipo de imágenes sintéticas, en las cuales se han incluido diversas matrículas completas, en gran variedad de orientaciones y tamaños. Ejemplo de una de estas imágenes se muestra en la figura 3.

2 Objetivo del trabajo

Se realizarán dos programas o funciones. El primero de ellos, se encargará de construir el clasificador, mientras que el segundo se encargará de la validación. Se pide la implementación de un *clasificador Bayesiano*. Como potenciales características para elegir, se propone el empleo de las relaciones de ocupación de subregiones en las imágenes de los dígitos normalizadas, como se detallará más adelante.

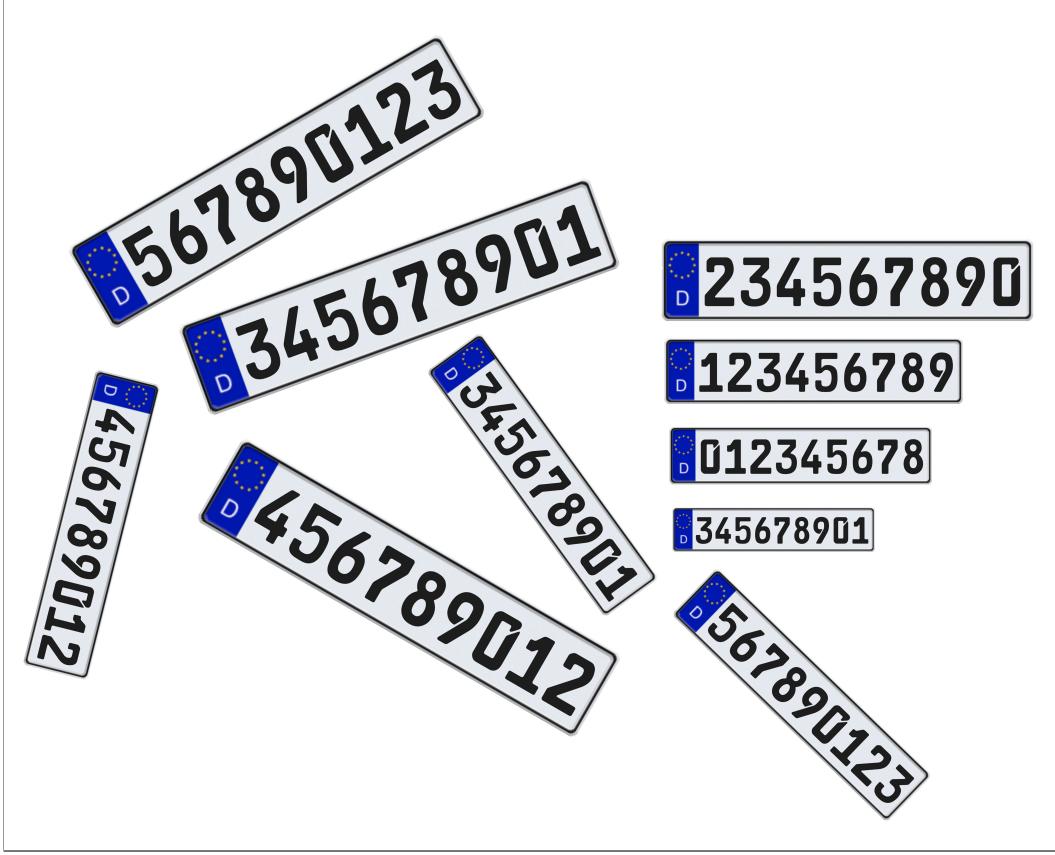


Figure 3: Ejemplos de imagen de validación del clasificador.

2.1 Necesidad de reorientación de la matrícula

A diferencia del uso de otras características para el reconocimiento de los dígitos, que pudieran ser independientes a la rotación, aquí se propondrá el empleo de características que sí dependen de la orientación del dígito. Como paso previo a la obtención de dichas características, resultará conveniente, por tanto, la normalización en orientación y tamaño del fragmento de imagen correspondiente a cada dígito extraído. Para cada matrícula localizada, se recomienda implementar la reorientación de la misma, de forma que quede lo más horizontal posible, con independencia de su orientación de partida. Como ayuda a este proceso, se puede hacer uso del marco de la matrícula, cuya orientación puede ser fácilmente deducible. En la figura 4, se muestra un ejemplo de reorientación de una matrícula, en base a la dirección principal de la misma.

En la versión básica del trabajo, se espera que la orientación con la que vengan todas las matrículas sea tal que no haya ambigüedad a la hora de obtener la versión horizontal de las mismas, no existiendo riesgo de que alguna de ellas pueda quedar del revés, como se muestra en la figura 5. En la parte izquierda de esta figura, se muestra un fragmento de imagen contenido una matrícula que aparece girada en exceso. Como consecuencia, cabe esperar que al reorientar la matrícula ésta quede invertida. Para evitar esto, basta que, a la hora de preparar las imágenes con las que se trabajará, ninguna matrícula esté girada respecto a la horizontal más allá del rango $\pm 90^\circ$.

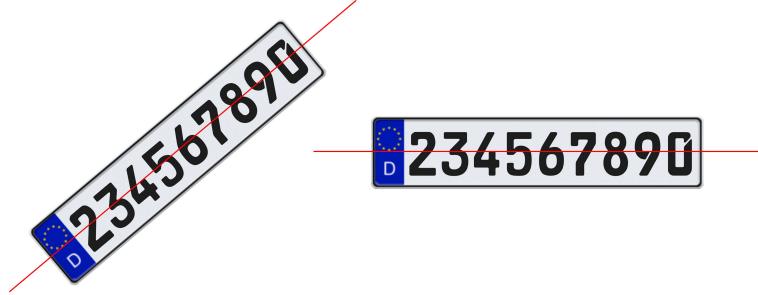


Figure 4: Ejemplo de detección de dirección principal de una matrícula y reorientación de la misma.



Figure 5: Muestra de reorientación incorrecta de matrícula.

2.2 Tipo de características a emplear

En este trabajo se propone el empleo de un subconjunto de características, consistentes todas ellas en el porcentaje de ocupación de una subregión de la imagen del dígito normalizada.

A partir de la imagen de una matrícula, llevada ya a una orientación horizontal, se propone la extracción de cada uno de los dígitos presentes en ella y la normalización del tamaño de cada uno de ellos a un tamaño específico que resulte conveniente. El tamaño a normalizar no tiene que mantener exactamente la misma relación de aspecto original. Por ejemplo, si se trabaja con imágenes en las que los dígitos, vistos en horizontal tienen unas dimensiones máximas y mínimas como las mostradas en la parte izquierda de la figura 6, puede verse que un tamaño promedio aproximado conveniente sería el que aparece en la parte central de la figura. Lo ideal es que las dimensiones horizontal y vertical elegidas para la normalización sean divisibles por 2 sucesivas veces (al menos divisible por 2 y 4), al tiempo que no distorsionen demasiado la relación de aspecto. En la parte derecha de la figura, se muestra el caso de una normalización más agresiva, en la que se ha forzado dimensiones coincidentes con potencias de 2 inferiores a las dimensiones promedio. Puede comprobarse que en este caso, la relación de aspecto se altera en mayor medida, pero aún así, sería una opción válida, con la ventaja de que puede subdividirse por dos de forma exacta indefinidamente, si ello fuera necesario.

Una vez normalizado el tamaño del dígito, se propone la subdivisión del mismo a distintos niveles. En la figura 7 se muestran tres posibles primeros niveles de subdivisión. El nivel 0 corresponde a la imagen del dígito completa, para este nivel se obtendrá una característica, x_1 , consistente en la proporción de píxeles ocupados (a blanco) frente al número total de píxeles

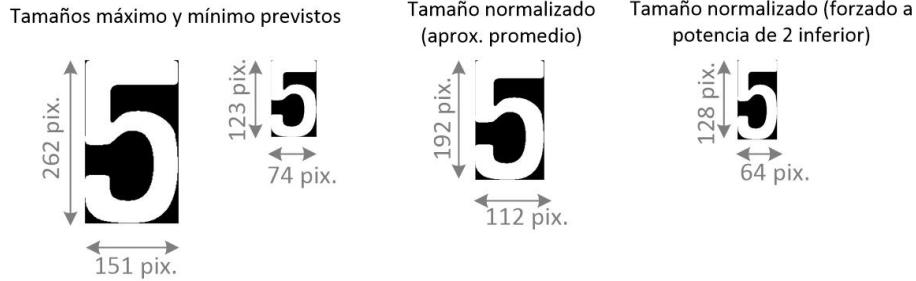


Figure 6: Ejemplos de normalización del tamaño del dígito.

del dígito normalizado. En el siguiente nivel, puede subdividirse el dígito en cuatro regiones iguales, para cada una de las cuales se obtendrá, nuevamente, la fracción de ocupación de la misma, dando lugar a cuatro nuevas características, $\{x_2, x_3, x_4, x_5\}$. En el siguiente nivel, nivel 2, podrían considerarse las ocho subdivisiones que se muestran a la derecha, dando lugar a 8 nuevas características, $\{x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}\}$. Este proceso de subdivisión y obtención de las características correspondientes podría repetirse sucesivamente, hasta donde fuera necesario. En la figura, también se muestra cómo sería la subdivisión de la imagen de algunos dígitos con esta técnica.

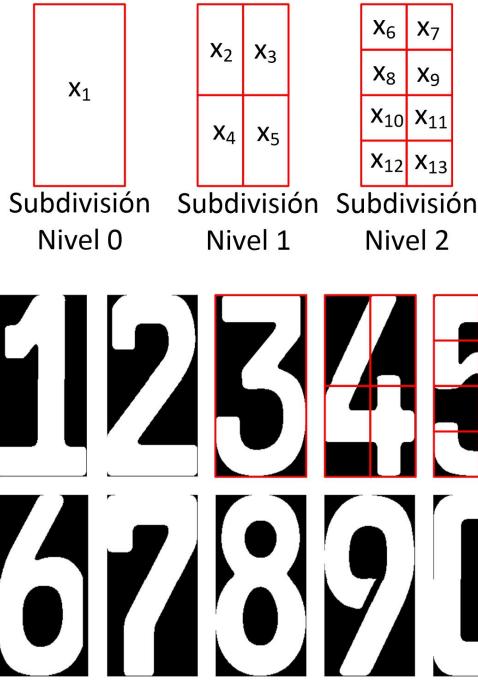


Figure 7: Primeros niveles de subdivisiones posibles de la imagen normalizada de un dígito.

La idea consiste en elegir un cierto número de características de uno o varios niveles de subdivisión, mediante las cuales pueda hacerse una correcta clasificación de los dígitos de la matrícula que aparezcan en la fase de validación.

No se pide el empleo de ninguna técnica específica de selección de características, más allá de la mera observación e intuición. Por otra parte, es deseable que el número de las características

finalmente elegidas no sea demasiado elevado, aunque **siempre intentando lograr 100% de éxito en el reconocimiento**.

Como muestra de la distribución de patrones y clases que se logra en un subespacio de características de dimensión tres (de cara a la visualización gráfica, se han elegido tres de las características seleccionadas), puede verse la figura 8.

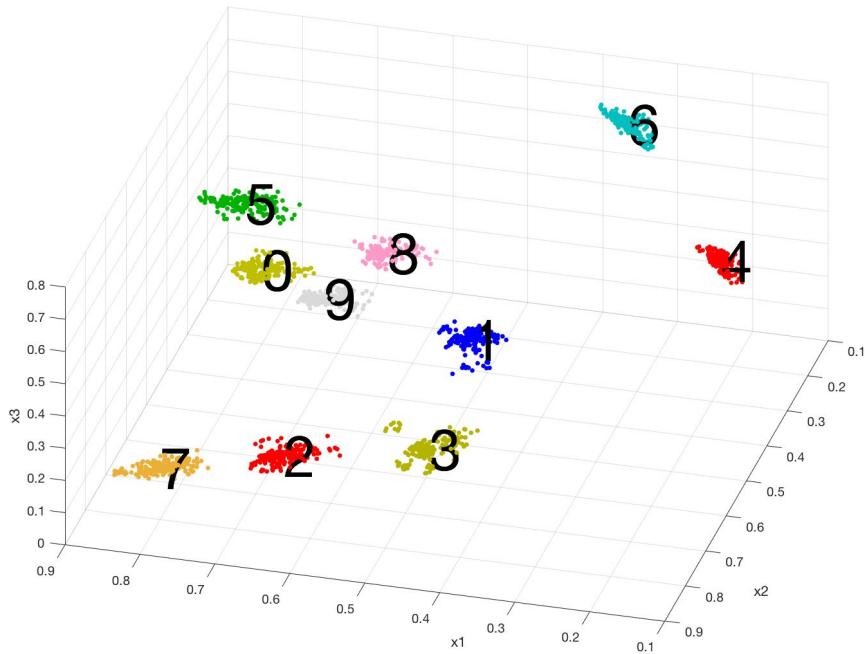


Figure 8: Ejemplo de distribución de patrones y clases en un subespacio de características de dimensión tres.

2.3 Procedimiento de validación

Ya se comentó anteriormente que la validación se realizará a través de imágenes del tipo mostrado en la figura 3.

Deberá encontrarse un conjunto de características que permitan lograr una clasificación 100% correcta. Para demostrar visualmente el éxito de una prueba de clasificación durante la fase de validación, pueden emplearse imágenes como la mostrada en la figura 9. Como puede verse, de la imagen dada, se extraen las matrículas, se reorientan y se sobreimprime, para cada una de ellas, en una posición aproximadamente coincidente con el centroide de cada dígito particular, el carácter que corresponde a dicho dígito. Como puede apreciarse, la primera de las matrículas, se ha reorientado de forma incorrecta, al aparecer en la imagen original con un giro excesivo, y el reconocimiento de los caracteres no es adecuado. Como se mencionó anteriormente, para la versión básica del trabajo, se supone que éste circunstancia no se dará.

3 Versión avanzada

Se pide resolver el problema de que una matrícula pueda resultar invertida en el proceso de reorientación, como ocurre en el caso de la primera matrícula extraída de la figura 9.

Para la versión avanzada se pide programar, además del clasificador Bayesiano, un clasificador de mínima distancia, con el mismo objetivo.

4 Puntuación

La versión básica de este ejercicio permitirá alcanzar hasta un 75% de la calificación máxima correspondiente, mientras que la realización de la versión avanzada, permitirá alcanzar el 100% de dicha calificación.

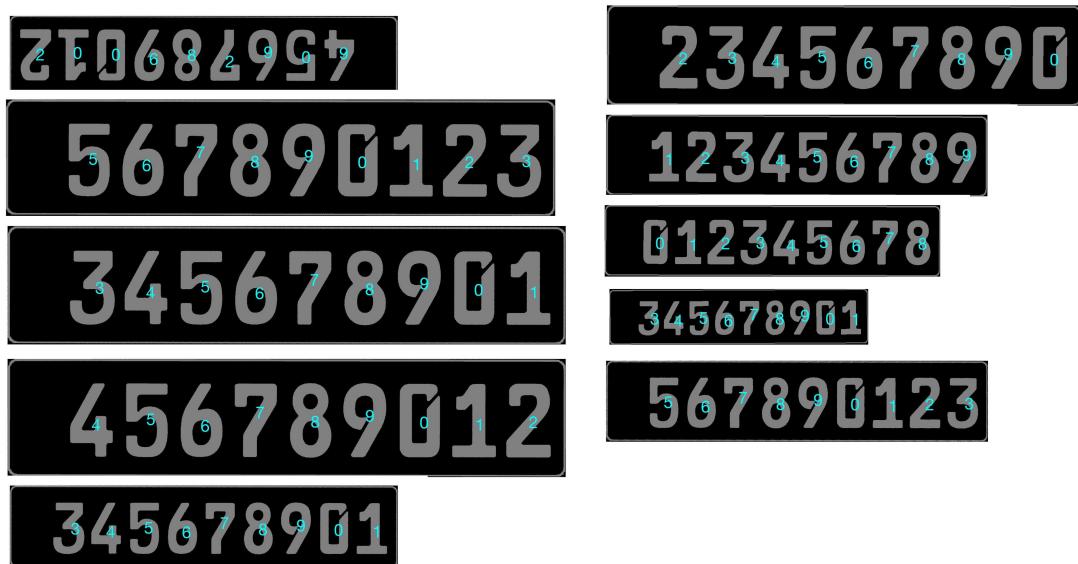
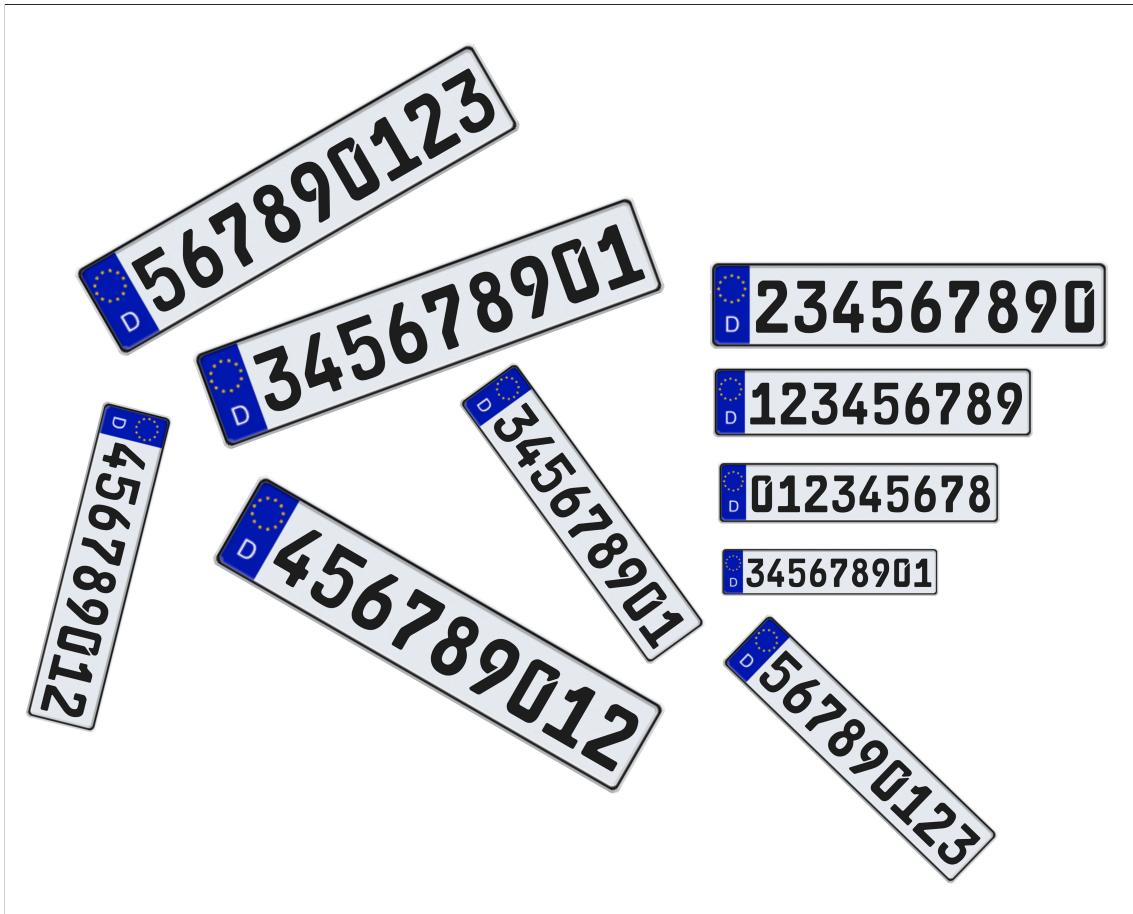


Figure 9: Muestra de resultado del proceso de validación del clasificador.