

[Forums](#) / [Código ejemplo de detector](#)[Help Center](#)

## Cómo hacer un buen nuevo juego de imágenes de entrenamiento

[Subscribe for email updates.](#)

UNRESOLVED

No tags yet. [+ Add Tag](#)Sort replies by: [Oldest first](#) [Newest first](#) [Most popular](#)

Daniel Martínez · 2 months ago

Estoy intentando hacer funcionar el detector con un conjunto de imágenes nuevas.

Primero, el proceso de generación de imágenes de prueba, como ya avisaron, es verdaderamente tedioso. ¿Hay alguna utilidad más rápida que Gimp para generar los conjuntos de datos?

Segundo, el dataset Pedestrians-Dummy tiene como 50 imágenes. El Pedestrians tiene 5000. ¿Cuántas imágenes hacen falta para empezar a tener buenos resultados?

Estoy trabajando con este juego:

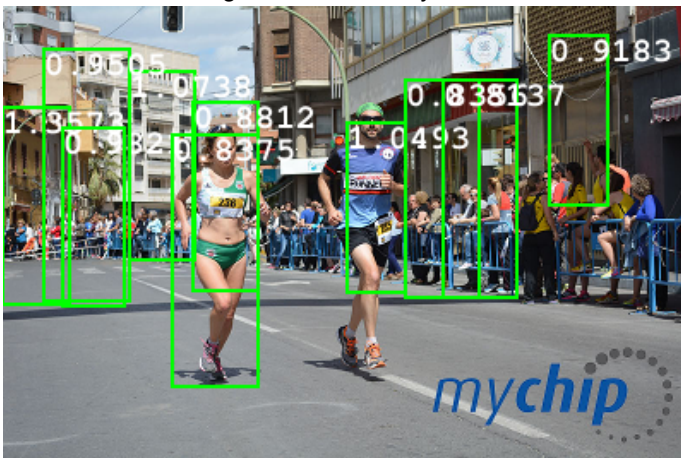
Positivos: 30 imágenes de 128x64, más las mismas invertidas horizontalmente.



Negativos: 20 imágenes de 128x64



De momento obtengo resultados muy malos.



Los resultados además varían mucho con el downScaleFactor.

¿Alguna otra recomendación?

↑ 0 ↓ · flag



Bikramjot Singh Hanzra · 2 months ago

Hi Daniel,

**Your dataset is too small.** The number of negative samples are too small to tune the SVM hyperparameters to get good results. There are 2 possible solutions that I can propose -

1. Increase the size of the dataset. Somewhere around 500 might make the results better.
2. Bootstrapping - Use hard negative Mining to reduce the rate of false positive. This is discussed in the class. Also, this post is [helpful](#). This is helpful when we generate the negative samples from the test images. We continue to retrain the classifier unless we reach good results.

Also, in case you are looking for ready-made datasets, I propose you should check the [UCI ML repository](#). I used the code provided in the class to train a car classifier.

Thank You

### ***Google Translation to Spanish***

Hola Daniel,

Su conjunto de datos es demasiado pequeño. El número de muestras negativas son demasiado pequeños para sintonizar la SVM hiperparámetros para obtener buenos resultados. Hay 2 soluciones posibles que puedo proponer -

1. Aumentar el tamaño del conjunto de datos. En alguna parte alrededor de 500 podrían hacer que los resultados mejor.
2. Bootstrapping - Utilice Minería fuerza negativa para reducir la tasa de falsos positivos. Esto se discute en la clase. Además, este [post](#) es útil. Esto es útil cuando se generan las muestras negativas de las imágenes de prueba. Seguimos a entrenar el clasificador a menos que lleguemos a buenos resultados.

Asimismo, en caso de que usted está buscando para conjuntos de datos ya hechos, propongo debe comprobar el [repositorio UCI ML](#). He utilizado el código proporcionado en la clase para entrenar un clasificador coche.

↑ 3 ↓ · flag

Jose Antonio Cisneros Huapaya · 2 months ago

Hi Bikramjot.

Did you get good results with your classifier? Which features and model did you worked with?

↑ 0 ↓ · flag



Bikramjot Singh Hanzra · 2 months ago

Hi Jose,  
The results were satisfactory. There were a few false positives, but overall I was satisfied. I had used HOG + Linear SVM.  
Thank You

0 · flag

+ Comment

New post

To ensure a positive and productive discussion, please read our [forum posting policies](#) before posting.

<b>B</b>	<i>I</i>			Link	<code>&lt;code&gt;</code>	Pic	Math		Edit: Rich ▼	Preview
<div></div>										

- ☐ Make this post anonymous to other students
- ☒ Subscribe to this thread at the same time

Add post

