

[Forums](#) / [Semana 3](#)[Help Center](#)

Número de anotaciones

[Subscribe for email updates.](#)

UNRESOLVED

No tags yet. [+ Add Tag](#)Sort replies by: [Oldest first](#) [Newest first](#) [Most popular](#)[Francesc Sala](#) · 2 months ago

Hola,
en casos prácticos, ¿hasta cuantas muestras es necesario anotar para obtener un detector que dé buenos resultados?

Vale, ya me imagino que la respuesta bien puede ser "Depende". Ok, entiendo que la pregunta es muy genérica, y que dependerá mucho del problema en sí, del clasificador, del algoritmo de aprendizaje, y ves a saber de cuantas cosas más. Pero aún así ¿me podríais por favor dar una idea? No sé, poner el caso de un problema de detección fácil, otro medio y otro complejo, y indicar más o menos cuantas muestras se necesitan en cada uno.

Muchas gracias. Cordialmente, f

0 · flag

[Anonymous](#) · 2 months ago

Buenas,

el funcionamiento de un detector mejora logarítmicamente con el número de ejemplos. Es decir entrenando por ejemplo con 3 ejemplos ya empezaría a detectar algo de vez en cuando, usas 10 imágenes mejorará mucho, con 100 irá bastante decente, con 1000 irá más o menos bien y a partir de ahí ya tienes que añadir muchas más imágenes para mejorar un poco el clasificador.

Para tener un detector más o menos bueno necesitarías al menos unas 2500 imágenes.

Con imágenes me refiero a muestras positivas (peatones). En cuanto a los negativos necesitarías muuuuuchos más (15000) pero estos son muy fáciles de conseguir.

Para darte una idea con los peatones te comento de memoria:

- Base de datos de INRIA: 1280 peatones anotados para entrenar (Al hacerles el flip tienes el doble)
- Base de datos de Daimler: unos 15000
- Base de datos de Kitty: Unos 20000

Si usases técnicas de deep learning (Al final del curso se dará una pincelada sobre ello) ya no iríamos al orden de los millones de ejemplos...

Si quieres hacer tu propia base de datos te diría que empieces con 10 imágenes para ver que todo cuadra y cuando vaya todo más o menos bien amplíala a 100. Anotar más de eso es mucho trabajo...

↑ 0 ↓ · flag

Francesc Sala · 2 months ago

Muchas gracias por tu respuesta

↑ 0 ↓ · flag

[+ Comment](#)

Anonymous · 2 months ago

¿Que software utilizáis para anotar? ¿Conocéis alguno en Matlab?
¿Empleáis técnicas de bootstrapping o aprendizaje activo? Si es así que protocolo seguís? ¿Generáis muestras positivas, por ejemplo, las metéis al clasificador y con los resultados volvéis a anotar?

Gracias y saludos

↑ 0 ↓ · flag

[+ Comment](#)

David Vázquez Bermúdez STAFF · 2 months ago

Buenas,

nosotros para anotar tenemos nuestro propio software en C++. Pero tenemos uno en Matlab disponible que podéis usar:

<http://www.davidgeronimo.com/index.php?s=software>

El bootstrapping de negativos es imprescindible (en el código que os pasaremos de entrenamiento del detector aun no lo tiene implementado, pero os daremos los recortes negativos extraídos de un bootstrapping hecho con otro código. Seguramente para la semana que viene el código que os damos en Python ya tendrá el bootstrapping implementado).

En cuanto al active learning si que lo usamos. Te paso un paper en el que explicamos cómo lo hacemos:

<http://refbase.cvc.uab.es/files/VLP2011a.pdf>

En la siguiente versión del artículo igual hay más información:

<http://refbase.cvc.uab.es/files/vml2013.pdf>

Saludos

↑ 0 ↓ · flag

+ Comment



Adrian · 2 months ago

Muchisimas gracias David.

↑ 0 ↓ · flag

+ Comment

New post

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

B	<i>I</i>			Link	<code><code></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

