PIXANO, the CEA’s open-source automatic image annotation tool, delivers pixel-level precision

Automatic image annotation is a determining factor in the quality of machine learning, making it a key component of artificial intelligence for computer vision. The CEA has leveraged its broad, deep knowledge of AI to develop PIXANO\(^1\), a tool that effectively combines the speed of artificial systems with human validation. The tool will be available on an open-source basis so that members of the AI community from both academia and industry can develop annotation technologies and solutions faster and more efficiently.

Automatic image recognition tools are not new. The systems available 20 years ago, however, had to work hard to recognize something as simple as a hand-written number. Since then, the rise of artificial intelligence (AI) has created new opportunities to leverage the massive amounts of data contained in images in fields as diverse as medical imaging, in-line quality control in manufacturing plants, and autonomous driving systems. Today’s increasingly-powerful smart systems will be called upon to complete an ever-expanding number of tasks and take on growing responsibility. The industrial companies that use or sell these systems need relevant databases to train the systems, evaluate them, and test their capacities. In this context, the need for more efficient image annotation solutions is pressing.

List, the CEA Tech institute for smart digital systems, drew on its vast experience with AI to develop PIXANO (Pixel Annotation), an open-source solution for the efficient large-scale online annotation of images and videos.

Annotation consists of tagging or labelling the elements in images in a variety of ways and with varying levels of detail, from text descriptions and geometrical shapes drawn around objects to object properties and pixel labels.

\(^1\) https://github.com/pixano/pixano.github.io
PIXANO offers a wide range of tag types (bounding boxes, polygons, cubes, pixel masks, etc.) integrated into open-source, modular, reusable, and customizable web components. The modules are driven by artificial intelligence, which assists the human user with annotation tasks and reduces the number of clicks required to create a tag.

The relevance of the solution to industrial users has been confirmed. Partnerships and joint R&D projects with major industrial companies (including in the automotive industry) have provided List with opportunities to demonstrate the effectiveness of the solution\(^2\) and capitalize on prior development work. Because it is modular and can integrate new smart, customizable components, PIXANO can be used to create a broad range of solutions to meet the needs of AI designers.

The global data annotation market grew to more than €300 million in 2018 and is expected to continue to grow rapidly in the coming years. The internet has made the large-scale distribution of tasks (crowdsourcing) the leading annotation model.

About the CEA

Based in France, the CEA is a government research organization that runs programs in four areas: technology research for industry, low-carbon energy (nuclear and renewable), basic research (materials science and the life sciences), and defense and security. The CEA is also one of the world’s leading centers for high-tech R&D and innovation.

The CEA is France’s only government research organization to have earned a slot on the Derwent Top 100 Global Innovators list 2018-19. The organization’s mission is to create technological innovations to make France’s businesses more competitive through high-performance, differentiating products. The CEA patents, develops, and transfers key enabling technologies covering most conventional industrial activities as well as the most advanced high-tech industries. These technologies address the needs of companies of all sizes. The organization is active across France, supporting local stakeholders’ innovation strategies with the goal of creating value and long-lasting local jobs in all of France’s regional economies.

www.cea.fr

---

\(^2\) Including R&D conducted under the European CloudLSVA project (https://cloud-lsva.eu/).