

How Dataloop Enables High-Quality Annotation at LinkedIn



Social Networks and the Challenge Today

As we stand today, LinkedIn has 740 million members in more than 200 countries worldwide. If you then add a couple of dozen different languages to the picture, you are left with a situation like other social networks where there is a huge responsibility of trying to keep nudity and profanity, as well as many other types of harmful content out of sight from their viewers. This is where machine learning

becomes a necessity to keep the site clean. With a site that relies on user-generated content, machine learning is used in order to automatically identify the quality of the content. This includes categorizing the content and preprocessing images for this purpose and of course, weeding out any inappropriate content.

LinkedIn has seen a tremendous growth of rich media content on the platform in the last few years. This has made robust image and video models for content filtering, recommendations, and search

indispensable. Training such deep learning models requires large amounts of labeled data in near real-time, meaning models must perform well at scale and in diverse environments. This growth was attainable with Dataloop's platform which offers a host of annotation tools and features made for fast, easy, and accurate labeling. We'll show how it powered LinkedIn AI teams to develop deep learning models using the annotated rich multimedia content.

How Dataloop Helps Solve Our Challenge

Dataloop's annotation tool is an AI-powered interactive multimedia labeling tool that enables annotation and management of rich-media content at scale with great efficiency and accuracy. The tool has contributed to the training and evaluation of models like scene-text recognition, multi-class classification models trained to extract rich-media metadata which are currently running in production. Furthermore, many others are in progress catering to the deeper semantic understanding of rich-media content.

Dataloop Annotation Platform Features

LinkedIn Multimedia AI relies heavily on Dataloop's annotation tool for multimedia labeling tasks. The annotation platform hosts a rich set of tools and features which makes the labeling very fast, easy and accurate. This section describes the various features that the platform supports.

User-Friendly Interface and Intuitive User Experience

The platform supports various ML use cases from classification, detection, segmentation to key-point extraction using a very user-friendly interface. The intuitive design and UX enables users to perform the tasks with ease and precision.



Classification

Classifying images is possible with the rapid ability to tag images, video frames, or audio clips, using single or multiple tags per item. Classification allows users to quickly identify the content inside each data item, categorize it into groups and clusters, and then translate the data distribution into immediate insights.

The Dataloop platform significantly expedites any classification task, by enabling users to select and tag bulks of items simultaneously, automatically switch between images upon completion, and identify similarities between pairs or sets of items in one click.

Furthermore, Dataloop uses ML models to automatically segment datasets into clusters in advance, thus ensuring that users spend their time validating, rather than manually tagging the data.



Fig-1: Dataloop's dataset management tools allow users to rapidly sort, filter, and query millions of items, as well as organize data into hierarchical directories and classification clusters.

Detection and Recognition

It is possible to use a variety of flexible and adjustable tools to detect, locate, and define objects in images and videos. Dataloop users can mix and match multiple detection applications - from 2D Bounding Boxes and Polygons, to Polylines and Ellipses, to 3D cuboids - all serving to mark and define objects based on the task at hand. Once detected, the object can be tracked across many frames and image sequences using unique identifiers. Association between objects can be defined using parenting and grouping, while advanced ontologies and unstructured descriptions can be applied to objects that require further categorization.

Furthermore, **Dataloop employs an unlimited variety of AI-Assisted tools and automation to speed up detection processes.** Users can work with readily available features such as one-to-many object recognition, or integrate their own ML models to trigger auto-annotation of the data. Over time, the platform utilizes active learning to progressively increase AI accuracy, eventually reaching a point where human intervention is required solely for edge case validation.



Fig-2: Using built-in object detection models to pre-label the data. Model results are visualized on top of the image and can be filtered based on model confidence scores, allowing human labelers to quickly validate and edit model outputs.

Semantic Segmentation

When computer vision applications require uncompromisable accuracy, Dataloop offers high-performance segmentation tools, ensuring pixel-level accuracy with maximum efficiency. With over x30 zoom, flexible opacity display, and adjustable tooling, users can work on granular details without losing sight of the larger image.

The platform utilizes a combination of AI-Assisted features and annotation-converters. For example, users can convert polygon annotations to semantic masks and vice versa, using points and vertices to mark larger areas, then switching to pixel-size brush to mark tiny details. These techniques enable users to reap the benefits of combining different tooling options, gaining better usability, comfort, and expedited efficiency.



Fig-3: Dataloop has a variety of AI-Assisted Semantic Segmentation tools, enabling users to achieve faster throughput while ensuring pixel-level efficiency.

Key-Points Extraction

Key points offer a quick and easy tool for marking locations, parts, or groups of objects with a single click. When combined together with object grouping, unique identifiers, and custom templates, key points become powerful tools for a wide variety of applications. For example, Dataloop users can model sets of key points and lines into structured templates. Those may serve anything from body pose estimation, to hand gesture detection, to face recognition, and much more.



Fig-4: Users can design custom templates for highlighting groups of annotations. For example, visually connecting key points in order to detect and analyze human poses.

Tackling LinkedIn's Rich Media Data Challenges

Various AI models were built leveraging capabilities of Dataloop's Annotation Platform for collecting good quality labeled data in an efficient manner.

Challenge #1

Scene Text Recognition (OCR) on Images and Videos

The task entails locating and recognizing the textual content embedded in the images and the frames of the videos. This is one of the most cumbersome tasks as professional images contain a lot of text like in the photos taken at conferences, lecture slides, quotes, etc which are abundantly prevalent on LinkedIn.

Dataloop provides the capability to draw bounding boxes around the text using a pre-trained AI model for each word present in the image or frame of the video.

The annotators then only have to review the annotations and correct them if required, thereby saving a lot of time and effort. Changes made are saved automatically.



Fig-6: In this example image there is text reading "8 REASONS WHY YOU'RE NOT GETTING HIRED FREE TUTORIAL WORK IT DAILY". Here, Dataloop's box & polygon annotations individually identified and labeled each word.

Challenge #2

Image Caption Quality Labeling

LinkedIn's image captioning feature automatically generates image captions for respective images. Image caption quality labeling tasks require evaluating specific image captions and classifying them into two or more categories based on their quality. In the following example, the image has two captions from different models, the labeler assigns one of the quality labels (for e.g good, very good, bad, etc) to each caption. This is a typical example of a multi-class annotation task.

Fast classification is used for quality labeling of images on Dataloop. Required labels are added to the recipe list. Annotators then select the desired label from the dropdown list, and in cases where feedback is required, the same can be added as an attribute to the label.



Fig-7: The image has a caption "graphical user interface, website" assigned from LinkedIn's captioning framework. The quality of the caption is marked "bad" when another annotation is added to the image.

Challenge #3

Video Closed Captioning

Closed captioning is a task where the audio portion of the video is captured and stored. This task requires the annotator to listen keenly and capture verbatim. Data generated from this task is used to evaluate the speech-to-text transcription models deployed in production, and hence, the accuracy is of prime importance.

Dataloop's platform has a very user-friendly interface for such speech transcription tasks. It also provides the annotator the flexibility to use the keyboard if required. They have the option to time-in and time-out the captions. These captions can also be edited and additional information can be added to the attribute section.

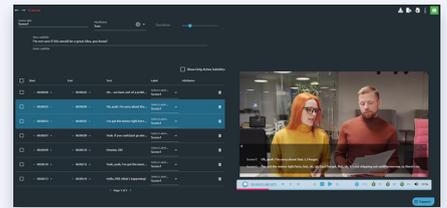


Fig-8: Video transcripts labeling in Dataloop, for a range of timestamps a text can be added along with a label.

Challenge #4

Image Logo Detection

The Logo Detection task is similar to scene-text detection which uses a bounding box tool for annotation. Annotators are required to identify the company brand and logos, locate and draw a bounding box around them with proper labels. The tool provides a wide range of drawing options for different shapes and sizes of logos such as polygons, circles. Dataloop's navigation from one image to the next is fast and auto-save is smooth.



Fig-9: Above is an example of logo detection with Dataloop's box annotation.

Challenge #5

Object Detection

Similar to other detection tasks, Dataloop comes very handy for object detection tasks. Annotators use this tool to draw boxes, which can be overlapping, around visible objects that can be identified in an image and the specific name of the object is tagged along with the box. The tool accepts hierarchical label taxonomy for classification, which is used to tag the objects during annotation. Each label is represented with a separate color code which makes it easy to distinguish among objects. These bounding boxes are color-coded for every identified object which makes it easy to distinguish each bounding box.

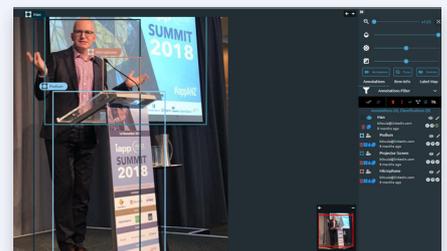


Fig-10: Labeling objects in an image along with the bounding box using Dataloop's annotation tools.

Challenge #6

Topic Tagging for Images and Videos

Topic tagging is similar to other multi-class hierarchical labeling tasks where a taxonomy of topic tags is defined and provided to Dataloop in JSON format. Annotators use fast classification to provide desired tags for an image from the dropdown list.

The major challenge is the large size of ontology. It is not only time-consuming but also infeasible to search among all the tags for an image. **Dataloop provides a very robust search tool to search among the large topic tags at different levels which makes the process very efficient and effective.**

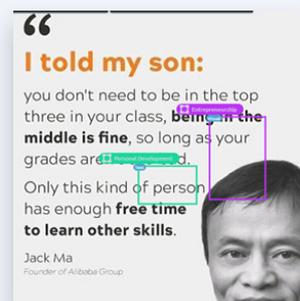


Fig-11: Label topics in a given image from a pre-defined set of topics that can be retrieved from the saved list of topic labels in Dataloop's annotation platform.

Great Customer Support

Dataloop's team provides great customer support on the issues, bug fixes, and new feature requests. LinkedIn's annotation team's experience working with Dataloop has been great. The Dataloop team is proactive in fixing bugs, suggesting new improvements, and helping with training sessions to improve the productivity and efficiency of labeling.

Conclusion

AI techniques have proven to be well suited to LinkedIn's rapidly growing data, by assisting and processing this data in order to identify patterns. The LinkedIn team building uses AI to automatically filter out any potentially harmful data that could save an organization from intentional or accidental harm. In addition, the need to moderate data in as close to real-time as possible keeps in line with users expecting their content to be available as soon as they post it, this way viewers can see content in real-time without being exposed to any harmful or inappropriate content.



"LinkedIn's network consists of over 740 million users across 200 countries in dozens of languages. When relying on user-generated content, it goes without saying that machine learning is crucial in helping us automatically identify the quality of content published on LinkedIn. That's where Dataloop comes in; their intuitive platform employs a wide variety of AI-assisted tools and automations, speeding up our detection processes and improving our model-generated content categorizations."

Shubham Agarwal, Senior Applied Research Engineer at LinkedIn