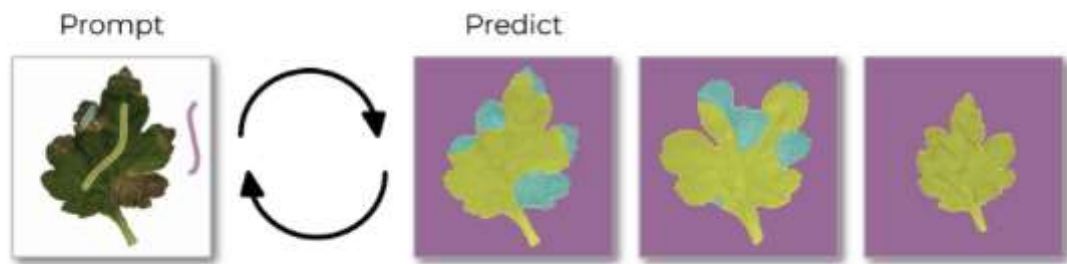


Introduction

Landing AI's [Visual Prompting](#) capability is an innovative approach that takes text prompting, used in applications such as ChatGPT, to [computer vision](#). The impressive part? With only a few clicks, you can transform an unlabeled dataset into a deployed model in mere minutes. This results in a significantly simplified, faster, and more user-friendly workflow for applying computer vision. View visual prompting examples below.

Visual Prompting Workflow



Visual Prompting builds on the successful use of text prompting in NLP

Traditionally, building a natural language processing (NLP) model was a time-consuming process that required a great deal of [data labeling](#) and training before any predictions could be made. However, things have changed radically. Thanks to large pre-trained transformer models like GPT-4, a single API call is all you need to begin using a model. This low-effort setup has removed all the hassle and allowed users to prompt an AI and start getting results in seconds.

Similarly to what has happened in NLP, large pre-trained vision transformers have made it possible for us to implement Visual Prompting. This approach accelerates the building process, as only a few simple visual prompts and examples are required. You can have a working computer vision system deployed and make inferences in seconds or minutes; this will benefit both individual projects and enterprise solutions.

How did we come up with Visual Prompting?

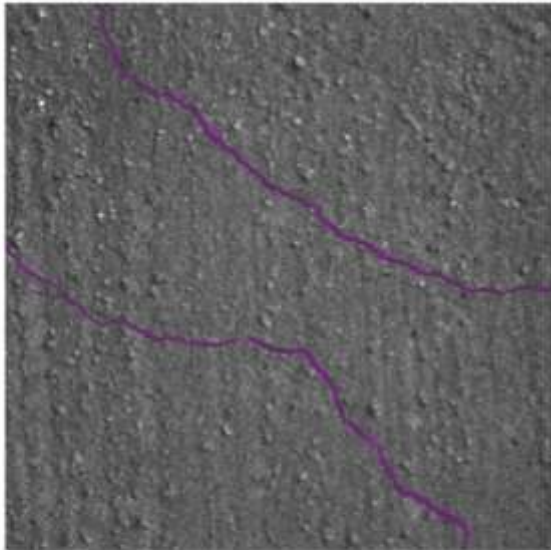
Visual Prompting is a method in which the user can label just a few areas of an image. This is a much faster and easier approach than conventional labeling, which typically requires completely labeling every image in the training set. Below are a few examples of Visual Prompting across verticals.

Similarly to what has happened in NLP, large pre-trained vision transformers have made it possible for us to implement Visual Prompting. This approach accelerates the building process, as only a few simple visual prompts are required. You can have a working computer vision system deployed

and make inferences in seconds or minutes; this will benefit both individual projects and enterprise solutions.

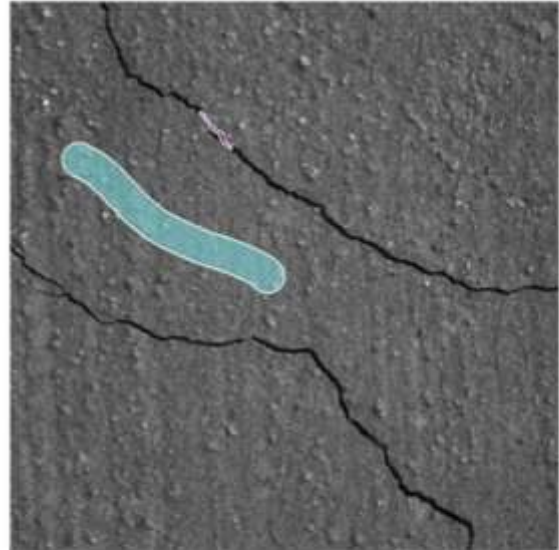
Crack Detection

Segmentation



2 masks in 15 minutes

Visual Prompting



2 strokes in 15 seconds

Vegetation Cover Estimation

Segmentation



13 masks in 12 minutes

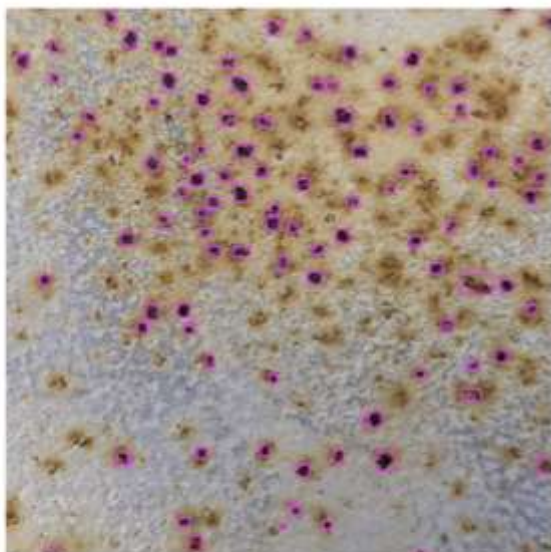
Visual Prompting



2 strokes in 10 seconds

Rust Detection

Object Detection



102 boxes in 9 minutes

Visual Prompting



2 strokes in 10 seconds

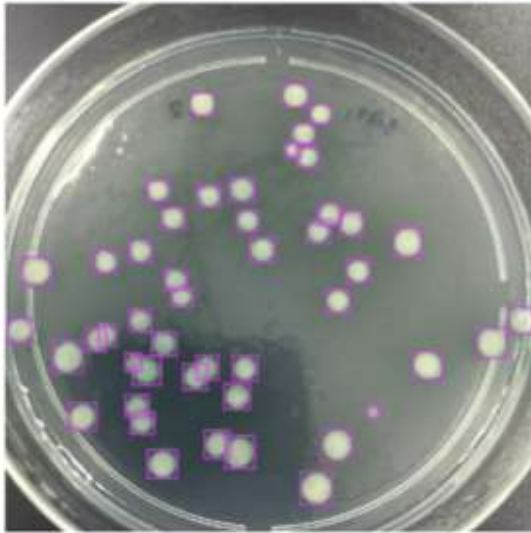
In 2020, the Landing AI team was working with a large [auto manufacturer](#) when a manufacturing staff member referred to “teaching” AI models rather than training them. Manufacturing inspectors are used to teaching other inspectors by pointing out a few things in an image – i.e., essentially by giving a visual prompt. This made us realize the importance of visual prompts in machine learning as well. The paradigm shift came when we understood that teaching machines was more about human teaching than machine training. Instead of laboriously labeling every single defect in an image of a car part, we could point out a few [defects](#) to an inspector, similar to teaching another human. If there were 100 defects within a single image of a sheet of metal, why did they have to label *every single one*? If they were teaching a fellow inspector how to find defects, they could just point to a few. We would like to make teaching an AI system more similar to teaching a human.

Enabling a customer to build a deployable cell detection model in just one sitting

[Visual Prompting](#) proved to be a game-changer for one of our life sciences customers, allowing them to develop a deployable cell detection model in just one sitting!

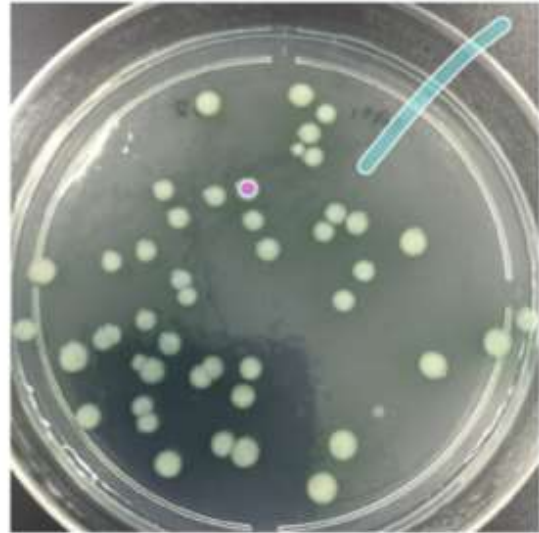
Cell Detection

Object Detection



47 boxes in 5 minutes

Visual Prompting



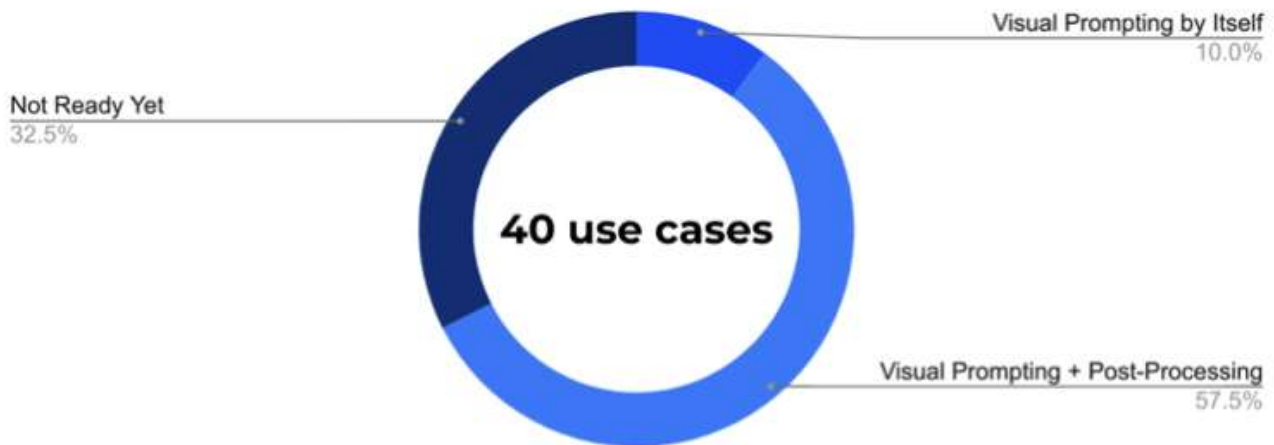
2 strokes in 10 seconds

This is how it went down: Initially, the customer attempted object detection for several weeks but encountered several challenges. There were too many cells to label, leading to the omission of some. At the same time, borderline and ambiguous cases caused confusion during labeling. So frustrating, right?

But then came Visual Prompting. The customer was amazed at how quickly they got the hang of it and the useful results they got, all within 10 delightful minutes! And they weren't alone – another customer couldn't help but rave about how similar the experience was to ChatGPT.

What's next? Make Visual Prompting practical through post-processing

In a quest to make Visual Prompting more practical for customers, we studied 40 projects across the [manufacturing](#), [agriculture](#), [medical](#), [pharmaceutical](#), [life sciences](#), and satellite imagery verticals. Our analysis revealed that Visual Prompting alone could solve just 10% of the cases, but the addition of simple post-processing logic increases this to 68%.



The reason behind this was clear. Customers wanted help mapping Visual Prompting's predictions to specific actions within their workflows. And most of them asked for help with post-processing to classification outputs. We got the message loud and clear from over 30 customer interviews conducted during our product discovery phase. This led us to devise a no-code solution for post-processing.

Beyond this, we see computer vision reaching an exciting inflection point; we are excited to collaborate with the computer vision community while continuing to adapt the latest technical innovations to make computer vision easy for everyone.

Conclusion

NLP has undergone transformative changes with the advent of text prompting. Now, an emerging avenue for computer vision is Visual Prompting. Why not give our new beta [Visual Prompting a try](#)? Revolutionize your computer vision processes with Visual Prompting!

To learn more, watch [Andrew Ng's Visual Prompting Livestream](#).