trax

Trax's Data Quality
Enforcement Framework

Execution Summary

Dedicated to a single purpose – digitizing the physical world of retail, Trax was founded in 2010 with the objective of automating and optimizing in-store execution for the consumer goods industry.

Today, Trax is the leader in computer vision solutions for consumer goods and retail, with a focus in retail-specific innovations to the market. Being the leaders in this space, doesn't mean we stop innovating. Every day our technology teams are pushing the boundaries of our platform and anticipating our clients' needs for tomorrow.

In this document, we describe how Trax employs advanced data quality enforcement based on active learning, data science and unsupervised machine learning engines to ensure the accuracy and the completeness of the data.

Introduction

The need for data quality enforcement

There are over 30,000 SKUs in a typical grocery store and on average 30% of these SKUs changes throughout the year where new products and packaging designs are frequently being introduced to the market. To handle this huge amount of SKU in any given store and to overcome the different challenges of a retail environment such as imperfect shelves, narrow aisles and poor lighting conditions, the automated recognition platform must be able to meet these key criteria to ensure a high level of accuracy:

- Distinguish multiple products that are nearly identical in appearance
- Overcome obscure and reflective packaging, and poor photo angles
- Identify empty spaces, products in poor visual conditions, and partially obstructed products
- Detect changes in the product life cycle like new design versions
- Detect anomalies such as inconsistent data and incomplete shelf capture

Trax's Data Quality Enforcement Framework

At Trax, we have developed a breakthrough computer vision platform that combines several patent-pending approaches in numerous advanced engines spanning across recognition, stitching, geometry and quality enforcement - working in parallel to recognize products on the shelf at unprecedented SKU-level accuracy, speed and scale so your teams always have reliable shelf insights.

In the next few sections, we will provide an overview of the key components within the Trax's data quality enforcement framework and how it works to ensure the highest recognition standards to our clients.

1. Ensuring Good Quality Images

Trax applied several advanced image process techniques to ensure every image captured are in their best conditions before they are put through the recognition process. Through a blur prevention feature, our snapshot tool can prevent user from taking an image if the conditions such as blurry images or if a different orientation (landscape/portrait) from previous orientation is detected. For tasks which need more light, the camera flash will automatically be activated. And a built-in leveller tool would notify the users when their devices is skewed and help users orient themselves when they are between bays so they can take straight images from the same distance.

Photo-taking guidelines

To improve the efficiency and effectiveness of the digitization process, we have also established a set of photo-taking guidelines for clients to ensure that we will have the best possible images for recognition such as:

1. Photo Capture Order
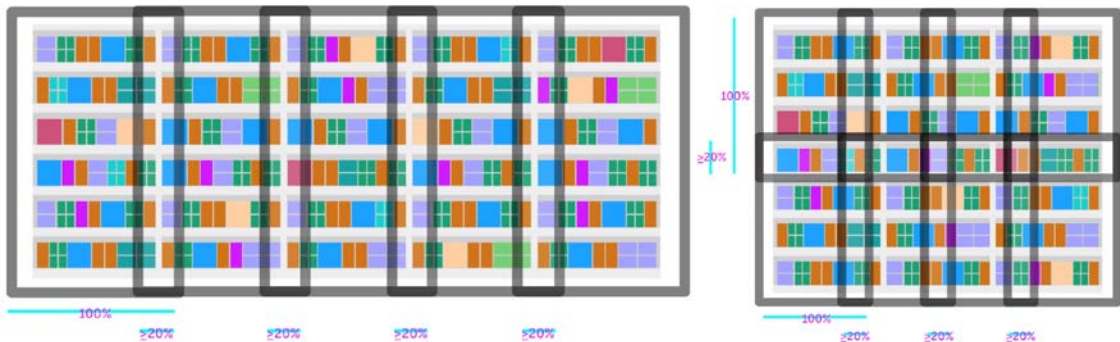
   a. The photographing order should always start from top-to-down and continue from left to the right of the shelves

   b. Allow at least one shelf to overlap vertically and 2 items overlap horizontally

   c. List prices must be clear and sharp to read from the image

2. Photo Capture Angle

   a. Ensure the camera is held straight, either horizontally or vertically

   b. Ensure the screen orientation within one display does not change.

   c. Ensure the distance from the shelf to the camera remains constant

These guidelines and best practices are available in an online learning tool - **Trax Mobile App Introduction and Training** - which is available to all users (auditor, sales reps) and is setup by Trax's implementation manager during the project.

2. Standardizing Recognition Rules

To be able to achieve a consistency level of recognition standards for clients and enables efficient implementation of the recognition engine, we have established standardized recognition rules for identifying SKU from shelf images. It also allows the usage of automated accuracy enhancement tools and enhances control over the reporting layer.

Below is the level of recognition we have established for SKU recognition:

SKU Recognition

If an SKU level recognition is needed for a brand, it is mandatory to recognize all SKUs within this brand. When product facing can't be recognized on "SKU Level", then it will be recognized on a higher level, such as "Brand Level" or "Category Level".

In this example, highlighted SKU will be recognized as "Bundaberg Pink Grapefruit 340ml"

Brand Recognition

When a product facing is not identified as a match to any existing SKU in the database but its brand is identified, it will be recognized on "Brand Level".

In this example, highlighted SKU will be recognized as "Bundaberg Other"

Category Recognition

When a product facing has no match SKU or Brand Other in the database and its category is in scope, then this product is identified on "Category Level.

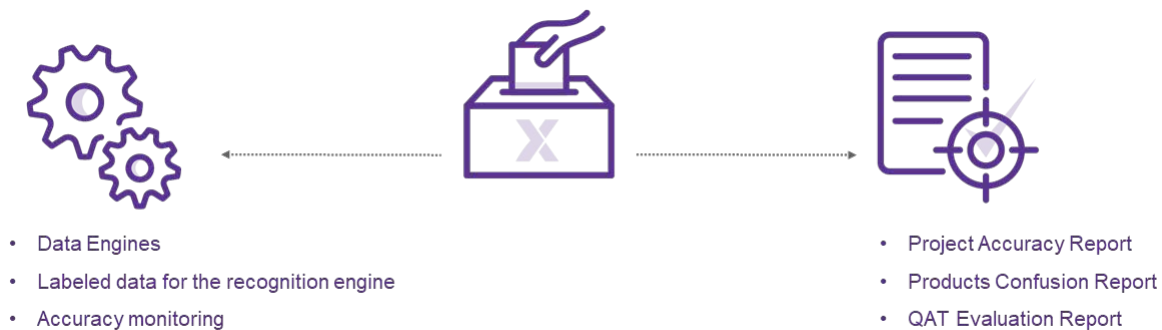*In this example, Highlighted SKU will be recognized as "Juice Other"*

Irrelevant

If a product's facing, brand or category information is not a match to information available in the database then the product will be identified as "Irrelevant".

In this example, Highlighted SKU will be recognized as "Irrelevant"

We have also established rules of recognition for empty shelves, placement (storage, stacking, reclining SKU), packaging (transparent/non-transparent packs, bundles) and irregular shelf conditions such as Clip Slips, Hoppers, Baskets, Tray and Displays.

3. Measuring Recognition Accuracy

At Trax, we leverage a robust quality assurance process known as voting to measure the accuracy of our image recognition engine. The voting results provide us with very high confidence on what was really in the image, and it's done using an active learning method that enables us to invest more resources in confusing products and less resources in in products that are easy to distinguish and recognize.

- Data Engines
- Labeled data for the recognition engine
- Accuracy monitoring

- Project Accuracy Report
- Products Confusion Report
- QAT Evaluation Report

Uses of Voting

To measure each client's project accuracy, we constantly sample the data we received and send them for manually validation. Every sample is sent to multiple data experts, where each project category has its own domain experts that are responsible to keep the accuracy high. The calculation of the sample size takes into consideration the number of SKUs and the project's real inflow to ensure that we will cover all the products that appear in the stores over a short period of time.

How it works - Voting

1. When data are received in Trax cloud, a sample of the received data goes into a Voting Pool.

2. The sample is initially sent to 3 domain experts and we compare their results

3. If majority is not reached, more experts are added. If majority is not reached, more experts are added. If majority is reached, the voting is concluded, and the domain experts who didn't get the same answer will get trained.

4. Once voting is completed for all products, the result is compared with the recognition engine's. We will then re-train the engines on the products it got wrong.

Recognition Performance

We provide full transparency on the recognition performance process through a web dashboard, **Brand Statistics** in **Trax's Product and Brands application** where clients can visualize various brand-related statistics. The results are regularly review together with our clients and project teams to ensure that any correction actions are taken in the quickest and collaborative manner.

The performance indicators are provided at a brand and SKU level help us and our clients understand if there are any occurrences of issues over the period of the project and help to identify the root causes if there is a degeneration in recognition performance. Examples of the scene and product images is also provided to help users identify quickly store-visit issues.
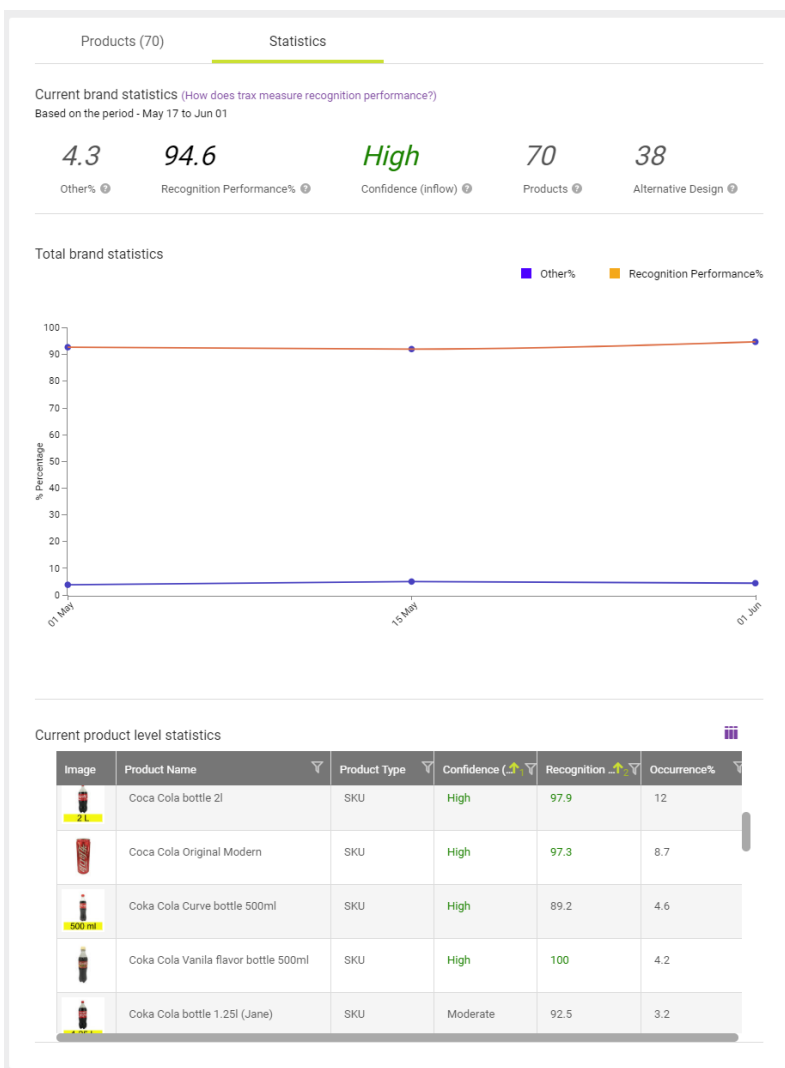
✓ **Recognition Performance %**

Percentage of products recognized accurately as compared to the voting results

✓ **Confidence**

Ranked according to the number of conclusive votes.

✓ **Others %**

Percentage of products of type "Others" recognized within the brand

✓ **Occurrence %**

Percentage of the brand distribution in the category

4. Ensuring Data Completeness

Incomplete data can lead to inaccurate results during the recognition and digitization process, and thus affecting the reporting of shelf metrics.

Augmented Reality (AR) assisted shelf capture

During a store audit, it is not unusual for the users to get interrupted by shoppers walking by. Through Trax's Mobile App, we leverage the immersive user experience of AR to guide users during their store audits to take images of a store shelf by marking out areas that they may have missed.

Shelf Capture Assistant

With Shelf Capture Assistant*, users can resume their photo capture tasks from the exact point at which they were interrupted using AR.

only available with selected Apple devices on iOS 11

Data Anomaly Detection

To ensure completeness of data during collection process, Trax has developed an unsupervised learning engine that actively monitors every session and detect for data anomalies in the collection layer such as:

Image and Scene Quality

Identify image and scenes quality issues based on reported flaws such as Blurry, Rotated, Reflection, Resolution, Capture angle, No Overlap

Image location
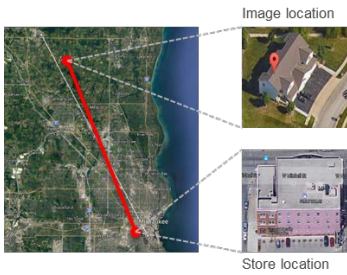
Store location

Inconsistent data

Identify visits suspected as fraud based on distance between known store location and photo's GPS.

April 12th: 6 facings

Incomplete data

Detect incomplete category and displays capture based on facings and leading brands.

January 11th: 31 Facings

Duplicate bays

Bays might be taken twice because of users' mistakes in same or nearby categories, and it might cause double facing count which lead to non-accurate reports.

Real-time warnings are sent also automatically when there is un-expected number of images detected during capturing of shelf images.

Additionally, Trax developed an accuracy enforcement engine to automatically identify for validation accuracy and stitching accuracy issues in the digitization layer, by leveraging a fusion of learning components to detect for any anomalies in the data.

5. Continuous tracking of new SKU/designs

An outdated SKU database can cause degeneration in the performance of the recognition engine, leading to inconsistent data and incomplete insight. And unlike many other domains, the data is 'live' in the retail industry – new products being launched, and existing products might have new packages and some products end their life.

To help our clients stay on top of all SKUs, Trax provides an on-going monitoring service known as **Brand Watch** that runs on an active learning engine to keep track of new packaging designs and/or new products as it also ensures the expected accuracy level of the recognition engine. Users can view new and changed SKUs within **Trax's Product and Brands application** and decide on actions such as adding or excluding detected objects from being tracked in the project.

How it works – Brand Watch
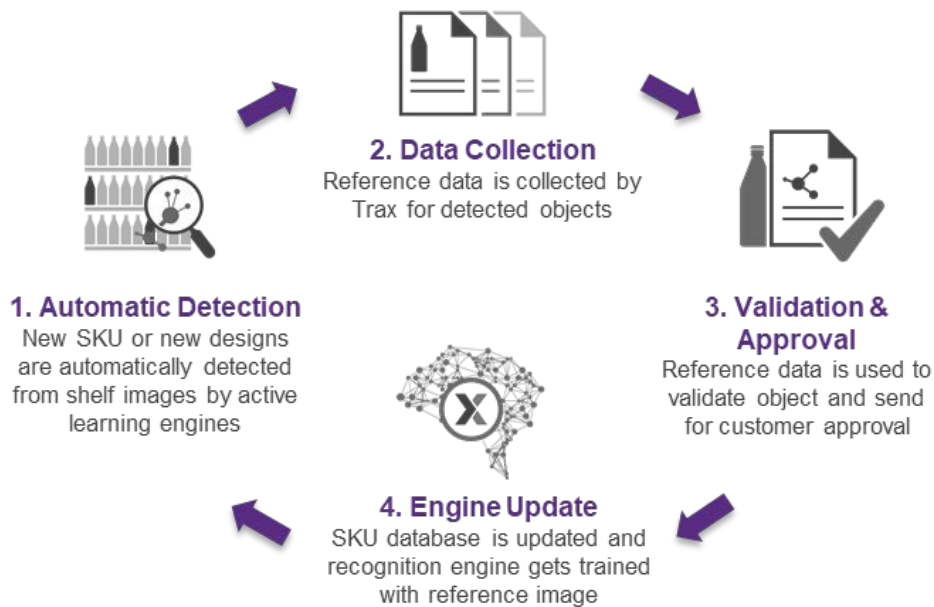
With this close-loop process, it helps our clients to alleviate the overhead from the customer of establishing and maintaining an up to date product repository for image recognition purposes. It also allows our client to quickly understand when new products and brands hit the shelves, and allows them to stay on top of new packaging design changes in the market, including competitors' brand and SKU.

Summary

Ensuring the trustworthiness and credibility of the data is extremely important for our clients to rely on the data generated by image recognition to make confident business decisions. Hence, Trax enforces rigorous data quality enforcement checks to ensure accuracy in image capture, recognition, calculation, and reporting of shelf metrics such as on-shelf availability, shelf share and more.