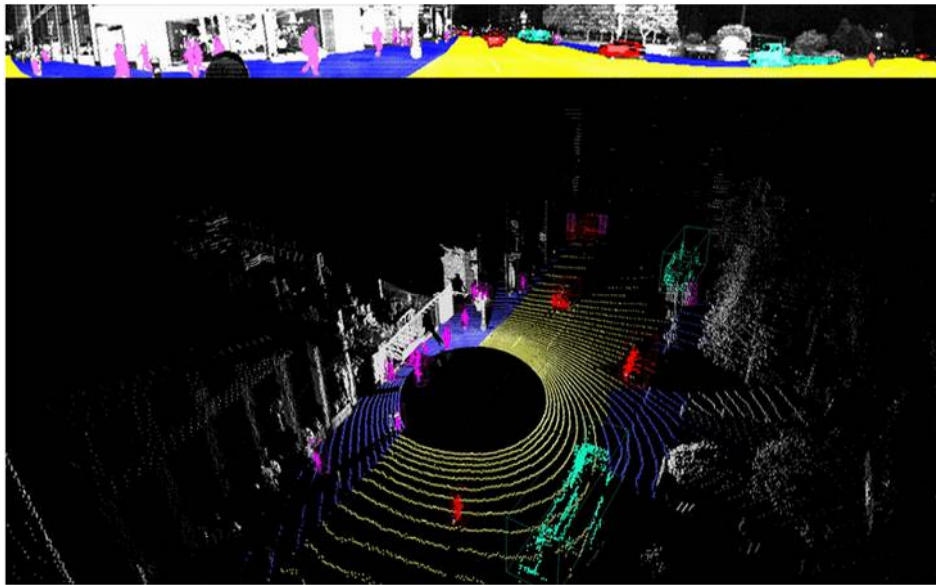


Ouster は、新ビジネスパートナーPLAYMENT 社と SCALE.AI 社と共に、ライダーデータのラベリングの効率化を追求する 2018 年 12 月 11 日 Angus Pacala



2D と 3D 投影の間で完全な相関関係のとれたピクセルごとに semantic masks を可視化する
上記の動画：<https://www.ouster.io/blog-posts/ouster-partners-playment-scaleai>

ロボティクスは難しい。データラベリングはそうあるべきではない？

我々は先週、PLAYMENT 社および SCALE.AI 社と提携して、Ouster の急速に増加する顧客（300 以上の顧客数で現在の増加中）の為に、ライダーのデータラベリングの効率化を促進するパートナープログラムを発表しました。我々は、これは Ouster 顧客のためだけでなく、機械学習業界に広く重要なステップとなると考え、本投稿で詳細に説明したいと思います。

過去数か月の間、PLAYMENT 社、SCALE.AI 社と一緒に、2D-3D データ構造を活用して、Ouster ライダーデータのラベリングのコストを 50%まで削減し、より高精度でより早いラベリングの作業を完了するため、データ取得からモデルの訓練開始までのプロセスを効率化する新しいラベリングツール開発をしてきました。

パートナープログラムは、また、全ての Ouster 顧客に対して、品質に関する高いレベルの標準化を支援し、ラベリング機能の共通セットが利用可能にします。Ouster 社のウェブサイトアクセスし、プログラムに参加するラベリングカンパニーの最新リストをご覧ください。リストは今後増える方向で、更新する予定です。我々は、顧客やラベラーにとっても有益となるライダーデータのオープンフォーマット化を開発し、ライダーデータの記録、保存、転送、読み込みの各作業を簡素化してきました。我々は、この業界の他の方々も、我々がスタートしたことに沿い標準化して頂けると希望して、本フォーマットのソースを公開しております（但し、ストラクチャライダーデータのみをサポートしています）。ライダーフォーマットは Ouster のオープンソースドライバーのアップデートに含まれ、顧客が直接、OS-1 センサーデータをこのフォーマットで書き込み、PLAYMENT 社や SCALE 社、あるいは他社のラベリング用ツールに転送できるようになる予定です。これに関しては 1 月以後、詳細に報告の予定です。

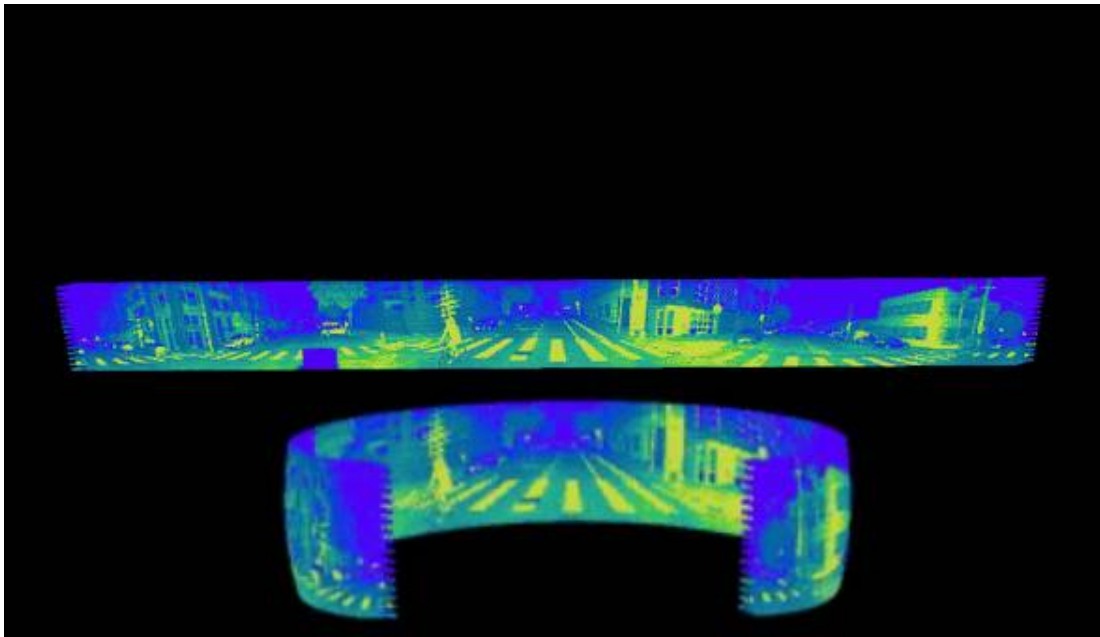
[Read more about Playment's integration with Ouster](#) ([和訳はこちらへ](#))

(Playment と Ouster は、提携しライダーデータのラベリングの効率化を図る。
データのラベリングに関して、Ouster ライダーは世界を一変させた)

[Read more about Scale.AI integration with Ouster](#) ([和訳はこちらへ](#))

(Scale.AI と Ouster データ統合に関する業務提携の発表)

structured lidar data：ストラクチャ（構造化した）ライダーデータの特徴



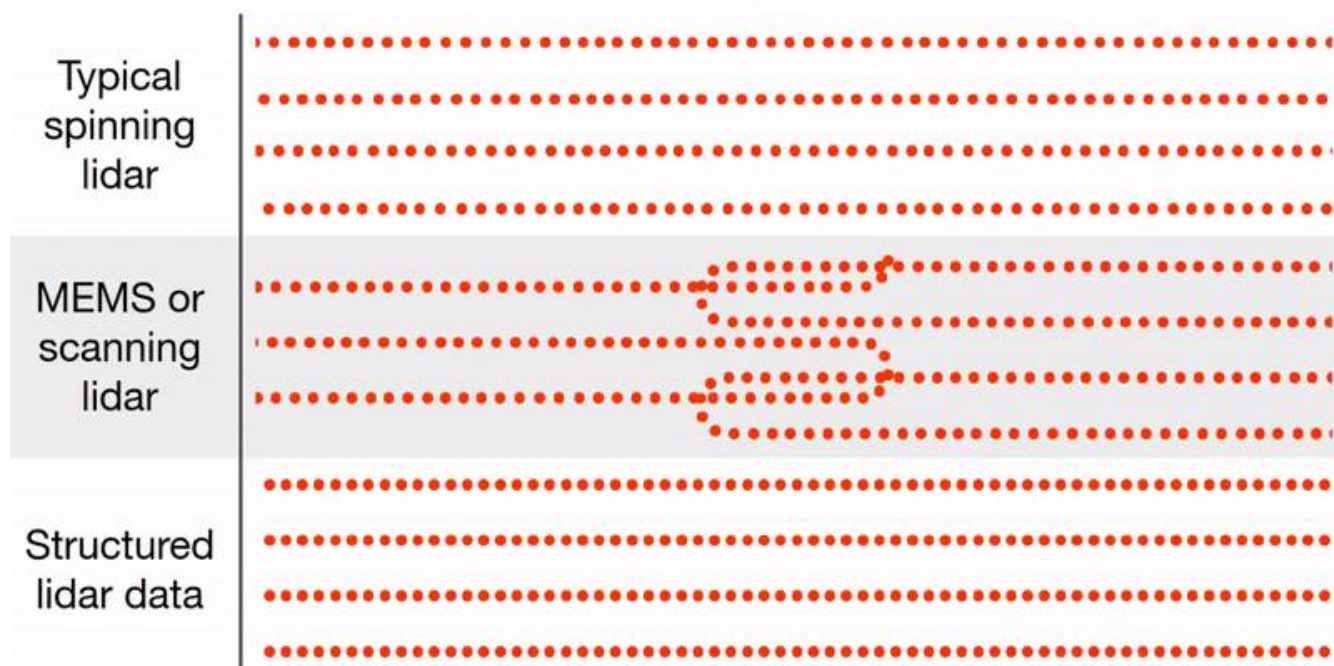
OS-1 センサーからのストラクチャライダーデータの可視化

上記の動画：<https://www.ouster.io/blog-posts/ouster-partners-playment-scaleai>

従来 ライダーは、スピンするユニットか、あるいはビームスキャンを行うユニットのどちらかでした。典型的なスピンライダーは、非一定速度でスピンし、点データ（ポイント）は全フレームに対して不均一に配置されています。これに対して、MEMS あるいはスキャンライダーは、非線形に S 字カーブにスキャンを行う複数（2-3 程度）のビームを有することが一般的です。これらのどちらもが、2D ディープラーニングに活用できる固定グリッドに、ライダーデータの保存をサポートしていません。このライダーデータは圧縮データフォーマットで、ラベリングが容易です。

我々の[マルチビームフラッシュライダー](#)のデザインのおかげで、Ouster OS-1 は、ちょうどカメラと同じように水平と垂直角距離を常に一定に保ちながら、**構造化したライダーデータ**を出力できます。これによりライダーは、2048×64 ピクセル固定サイズの深度マップを、全フレーム上に反射強度とアンビエントライト画像（周辺の画像）を出力することが出来、畳み込みニューラルネットワーク（convolutional neural networks）を利用可能となり、画像の保存やラベリングのワークフローを飛躍的に容易にすることができます。

Structured vs unstructured lidar data



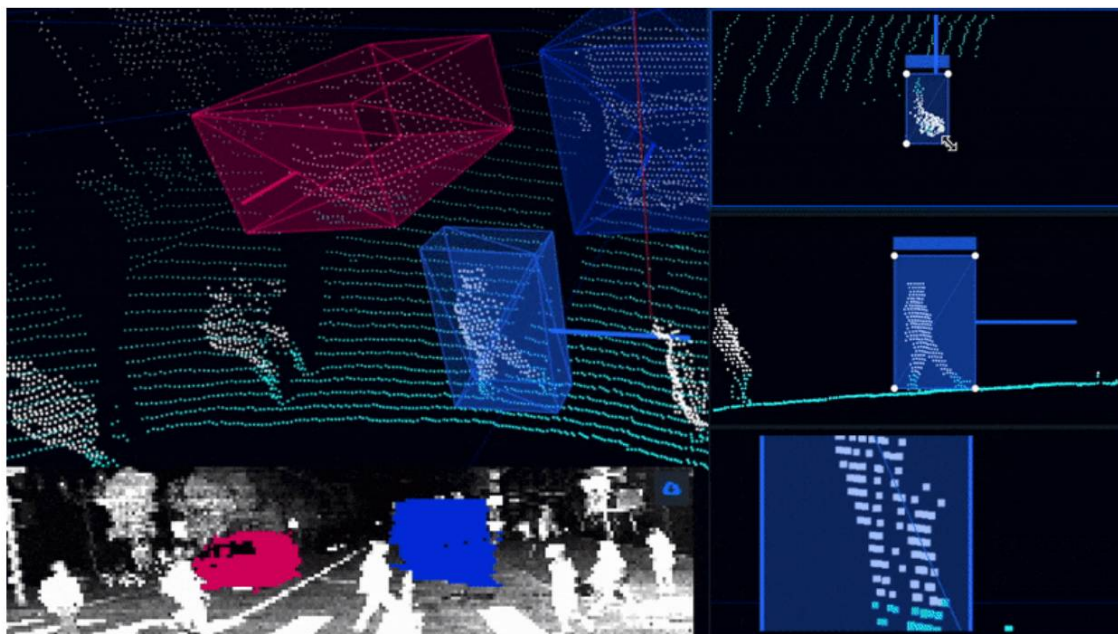
一方、RGB-D カメラや従来のフラッシュライダーも、構造化範囲データを出力できますが、どちらのセンサーも、OS-1と比較して、距離データ、分解能、視野、厳しい使用環境に耐えられる性能を有しておりません。しかし、これらのショートレンジの構造化 3D カメラは、我々が実施している分野でも利用価値はあります。そして我々は、これらの製造メーカに我々のアプローチを検討するよう提案を続けています。

ストラクチャライダーデータのラベリングワークフロー

我々は、我々のストラクチャデータを活用するために、ラベリングパートナー社のラベリングツールにおいて、ラベリングのコストを最小化し、機能を追加しています。また、アノテーションの正確性を大幅に改善する目的で、ラベリングパートナーと協力してきました。以下で、幾つかの例を紹介します。

- ・ カレントタスクの直感的で視覚的な確認作業として、アノテーターに同期化された 2D-3D ビューを提供する。
- ・ 2D ライダー画像にピクセルごとにマスクをラベル付けし、またマスク精度をチェックするために 3D 点群を使用する。マスクを 3D 点群内で、個別のポイントや選択されたポイント群のクラスタを追加したり除去したりし更新を行ない、マスクをアップデートする。
- ・ 3D 境界ボックスを使用して、100% 正確な 2D マスク、直方体、2D 画像での境界ボックスを自動的に生成する (semantic あるいは instance segmentation)

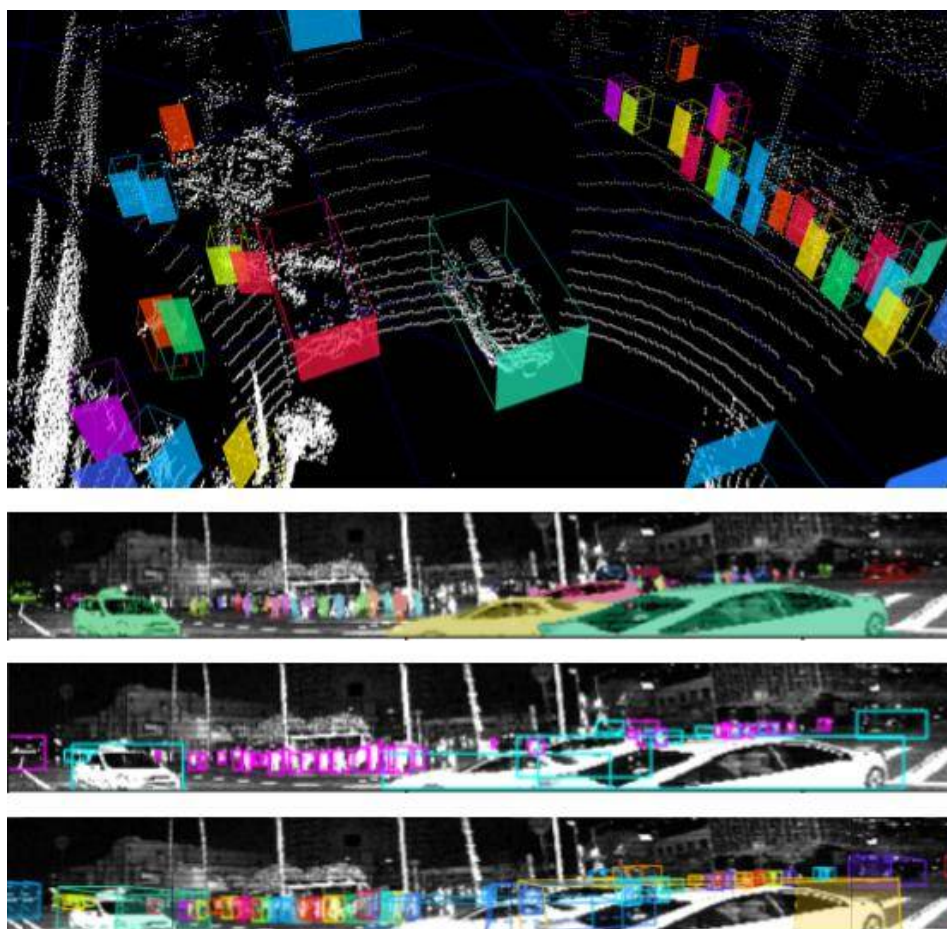
本ラベリングフローが実行されているのをご覧いただければ、いかに効率的で正確であるかお分かりいただけると思います。あえて代替案は、本ワークを 2D と 3D データに、それぞれ複製することかもしれません。



完全な 2D ピクセルごとのマスク（左下の画像）は 3D 境界ボックスから自動的に生成され、2D マスクは 3D 境界ボックス精度をリファインするのに使用されます。（動画は）

上記の動画：<https://www.ouster.io/blog-posts/ouster-partners-playment-scaleai>

本アプローチは、2D 境界ボックス、セマンティクス／インスタンスマスクから 3D 境界ボックス（2D あるいは 3D）、点群セマンティクス／インスタンスセグメンテーションまで、ラベルの最大の範囲を、可能な限り最小の努力で最大の精度で実現するものです。



上段：マニュアル操作でラベル付けされた 3D 境界ボックスの例、中下段：2D フレームで、自動に生成されたピクセルごとのマスク、境界ボックス、直方体を示す。

これはいったい何を意味するのか？ 10-50%低いコストと速いラベリング

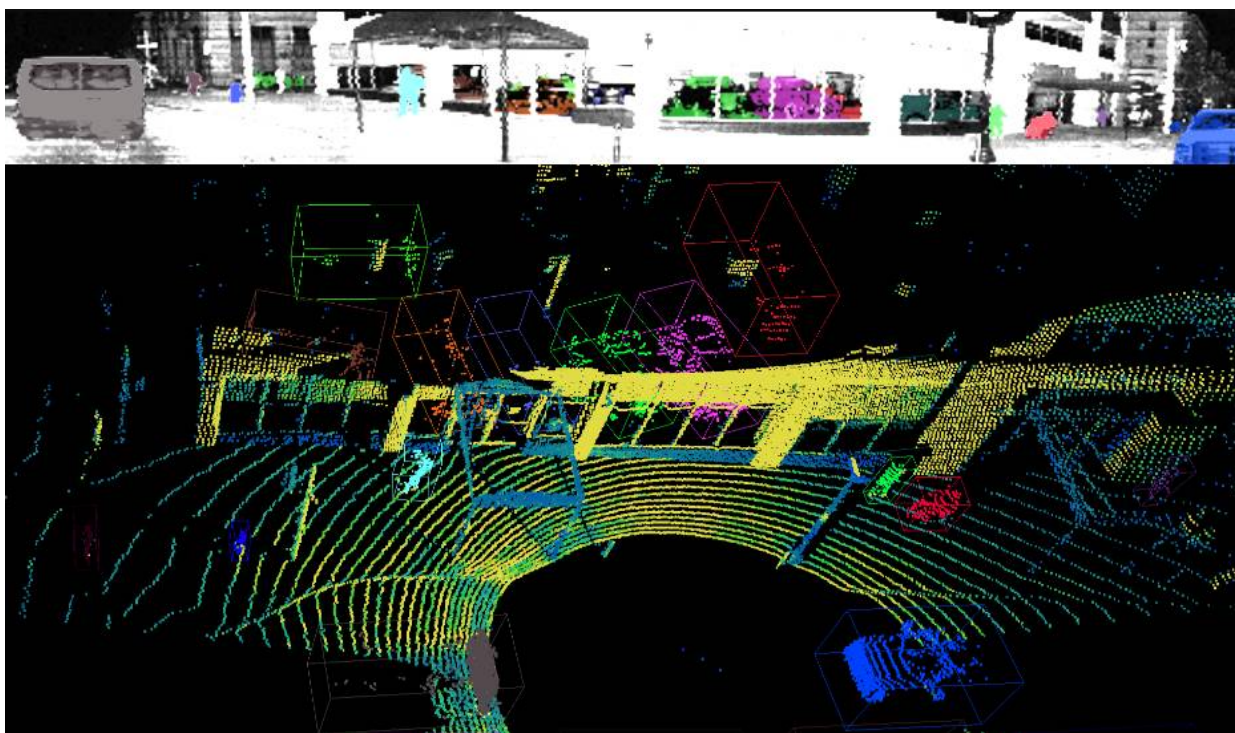
2Dと 3Dのデータセットの両方を個々にラベル付けする必要性が無くなり、我々はアノテーターが必要になるラベル付けの量を2倍まで減らすことが出来、顧客のコストを大幅に低減できます。

より小さなファイルサイズ

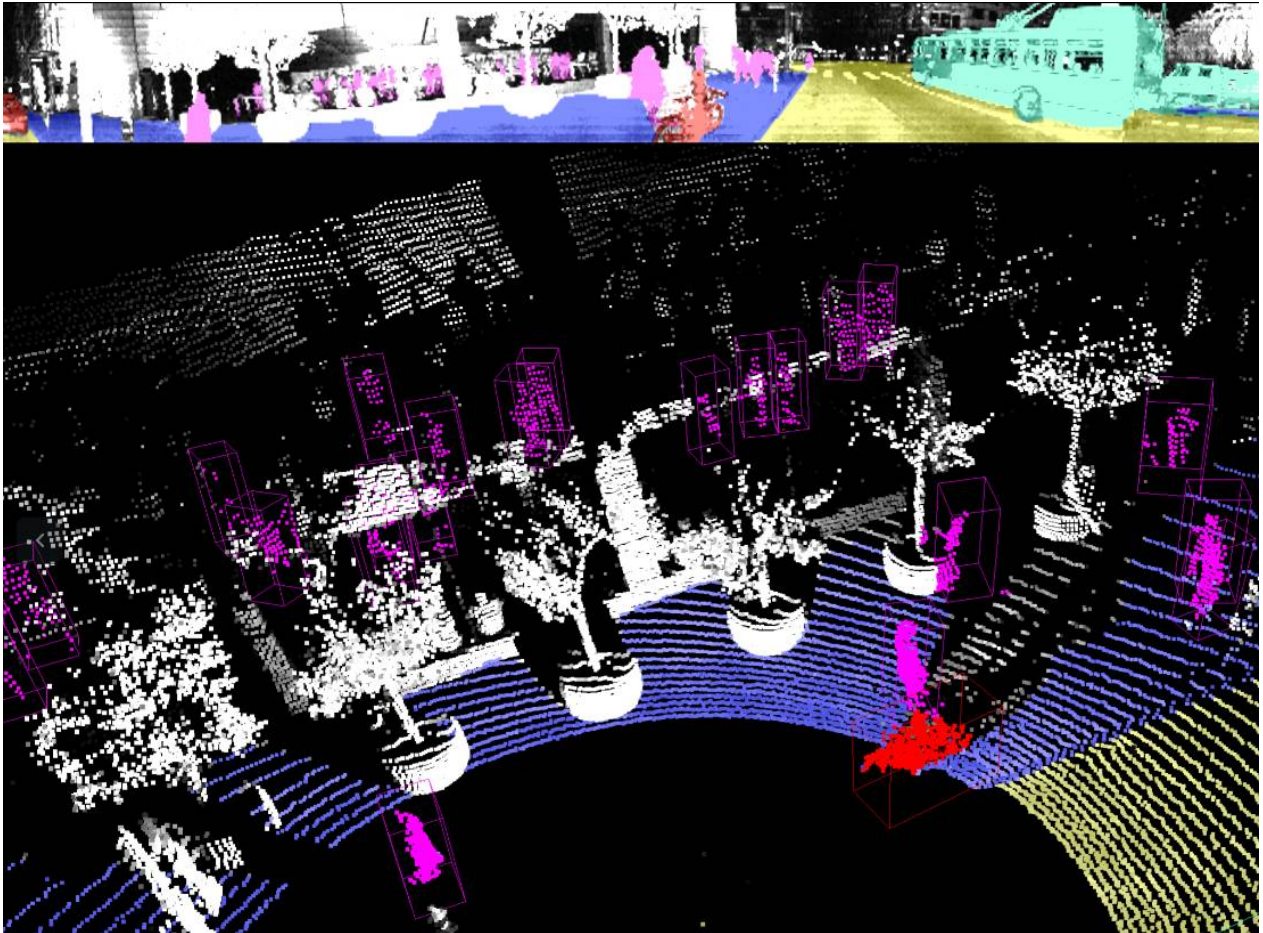
我々のオープンソースのストラクチャデータフォーマットはファイルサイズを減らし、転送の時間を最大97%まで削減します。これまでデータボリュームは既に大きく、ラベリングパートナーにハードドライブを物理的に送付されるお客様もおられますが、このようなことをする必要はもうありません。

ラベル精度の改善

2Dと3Dの両方のラベルの精度は、アノテーターに対してより視覚的な補助ツールと100%正確な2Dセマンティクスマスクがあれば、2Dと3D間のラベルの可動性(transferability)によって改善されます。セマンティクスセグメンテーションにはこれ以上、境界の問題はありません。歩行者やカーディーラやギャラリー内の車でさえも、容易かつ正確にラベル付けできます。より正確なラベルとは、アルゴリズムをテストする際に、より正確で測定可能な意味のあるメトリクスを作成できることを意味します。



この自動車販売店の窓枠が、車両のマスクから自動的に排除されていることをご覧いただけます。



ここでは、小売店の客は、人々の前後周りに散在する木やウィンドウ、家具にも影響されず完全にセグメントされています。

Rotoscoping and green screen : ロトスコーピングと緑のスクリーン

以上の全てが、少しでも親しみを感じられるのならば、本当にそうなのでしょう。大きな 2D 画像セット中の正確なマスクのピクセルを生成するテクノロジーとツールに数十年を費やしている巨大な産業がもう 1 つ存在します：フィルム産業です。

ロトスコーピングと緑のスクリーンは、長い歴史があり、しばしば高度な手作業を伴い（緑のスクリーンに関しては、全然そうでもないが）、合成のため、特定の俳優やオブジェクトをカメラシーンから区別する方法です。そして、ロボティクス産業のように、彼らは最近になって、3D カメラとディープラーニングを活用した自動ロトスコーピングテクノロジーを開発しつつあります。