



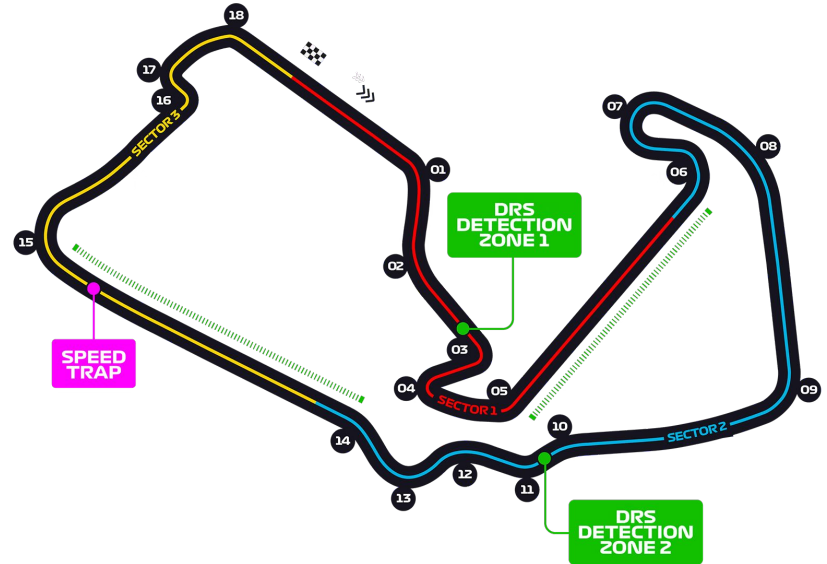
# How to Unlock More In Self Driving Datasets



**Daniel Gural**

Machine Learning & Developer  
Relations - Voxel51

# The Race for Autonomy



# How the Field Shapes Up

What is getting people ahead in the grid

## Powerful GPUs

- More powerful
- Easy to scale
- Libraries only getting better

## Boundless Amounts of Data

- Map whole cities to datasets
- Hours and hours of drives

## Top Tier Talent

- Incredibly talented engineers
- Behind closed doors ...



# How the Field Shapes Up

What is getting people ahead in the grid

## Powerful GPUs

- More powerful
- Easy to scale
- Libraries only getting better

## Boundless Amounts of Data

- Map whole cities to datasets
- Hours and hours of drives

## Top Tier Talent

- Incredibly talented engineers
- Behind closed doors ...

*Until Today!*



# Today's Outline

What to expect:



## Beginner Techniques

Curation, Visualization, and Dataset Management



## Advanced Techniques

Embeddings, Zero Shot Models, and Depth



## Expert Techniques

NerFs, Gaussian Splats, and Simulations

# Before We Get Started

## Who is Voxel51?

- > The lead maintainers of FiftyOne

## What exactly is a “voxel”?

- > A voxel is to a video what a pixel is to an image. It is literally, a “volume element” in the space-time volume of a video.

## So, what’s with the “51”?

- > Good question!



### Daniel Gural

ML Engineer & Developer  
Evangelist at Voxel51



# Before We Get Started

Link to Code



LinkedIn



# Beginner Techniques





# Juggling Around Different Samples

How to Manage a Self Driving Dataset



Unstructured Data

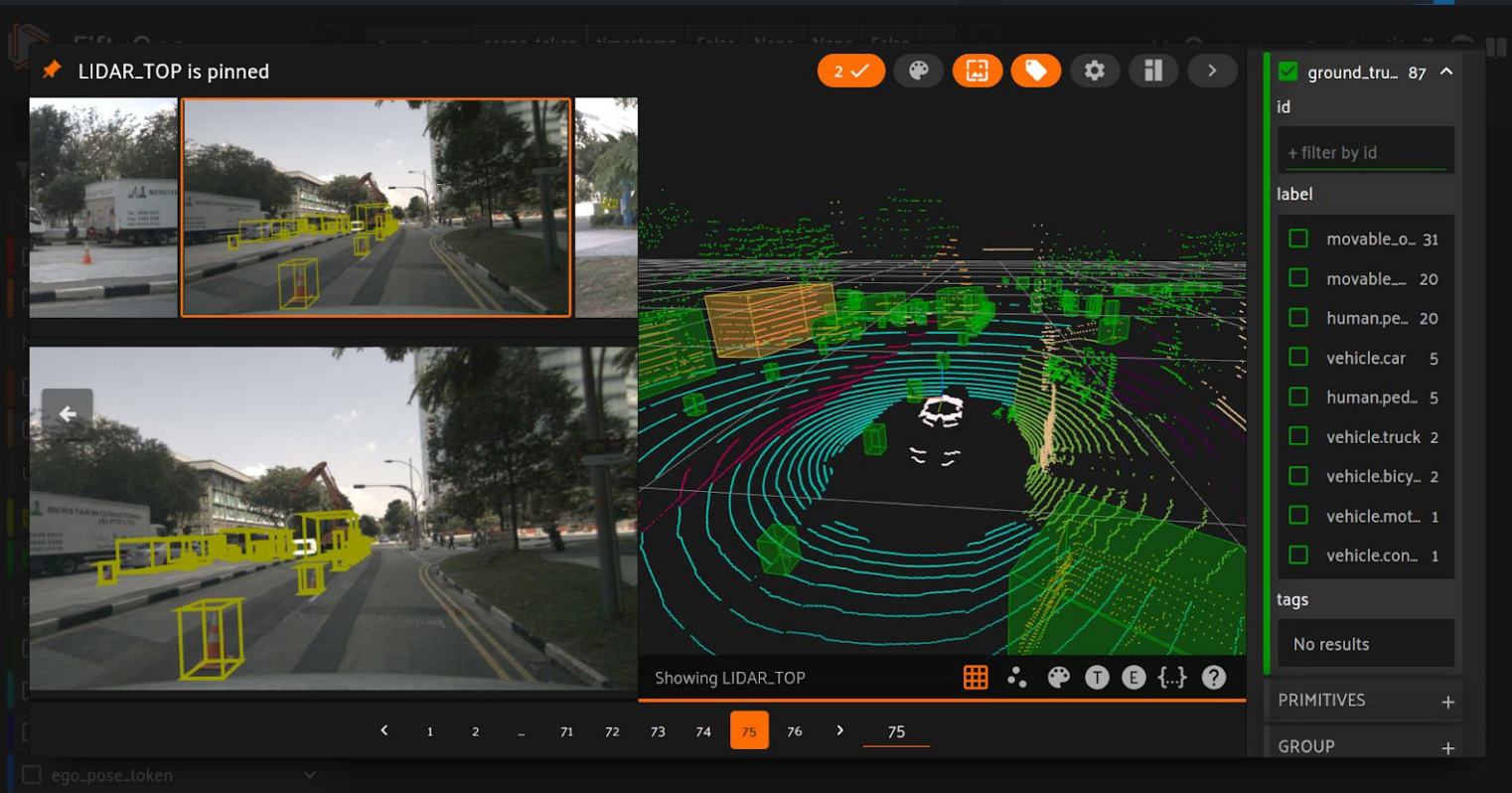


At a HUGE scale



# FiftyOne to the Rescue

✦ LIDAR\_TOP is pinned



The screenshot displays the FiftyOne web interface. On the left, a video player shows two frames of a street scene with yellow bounding boxes around objects. The top frame is a still image, and the bottom frame is a video frame. On the right, a 3D LIDAR point cloud visualization shows the same scene with colored lines representing the LIDAR data. The sidebar on the right lists object detection results for the selected frame (75). The results are categorized by label and tag.

Showing LIDAR\_TOP

id

+ filter by id

label

- ☐ movable\_o... 31
- ☐ movable\_... 20
- ☐ human.pe... 20
- ☐ vehicle.car 5
- ☐ human.ped... 5
- ☐ vehicle.truck 2
- ☐ vehicle.bicy... 2
- ☐ vehicle.mot... 1
- ☐ vehicle.con... 1

tags

No results

PRIMITIVES +

GROUP +

ego\_pose\_token

# What Is FiftyOne?

## In a Nutshell

- > Visualize, clean, and curate
- > Find hidden structure
- > Evaluate models

*Flexible. Customizable. Connected.*

*The open source tool for building high-quality datasets and computer vision models*



# Voxel51's Mission

is to bring transparency and clarity to the world's data



## Data Quality

Nothing hinders machine learning systems more than **poor quality data**

The secret to getting ML products into production lies in **improving datasets, not model architectures**




## Data-Centric Workflows

Companies are hiring ML talent as fast as they can, but they **need tools and workflows** to build high quality datasets and models

Our **open source approach** aims to bring data-centric ML to anyone who wants it



# Grouped Datasets

 FiftyOne quickstart-groups + add stage

× ? Have a Team? # 🔊 📄


Sample tags Label tags Labels Other fields Show ▾

FILTER

- SAMPLES TAGS
  - ☐ train 200
- LABEL TAGS
  - No label tags
- METADATA
  - ☐ metadatasize\_bytes 200
  - ☐ metadatamime\_type 200
- LABELS
  - ☒ ground\_truth 1,379
- PRIMITIVES
  - ☐ id 200
  - ☐ filepath 200
- GROUP
  - ☐ group.id 200
  - ☐ group.name 200

+ ADD GROUP

200 samples left





# Advanced Techniques

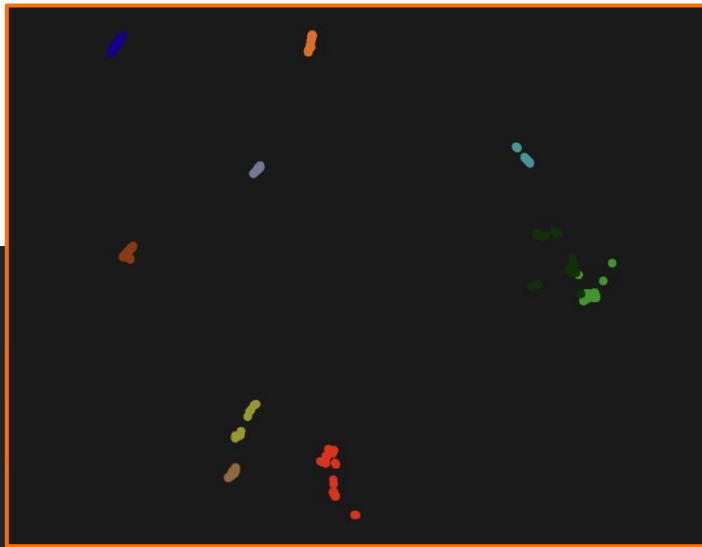


# Advanced Techniques

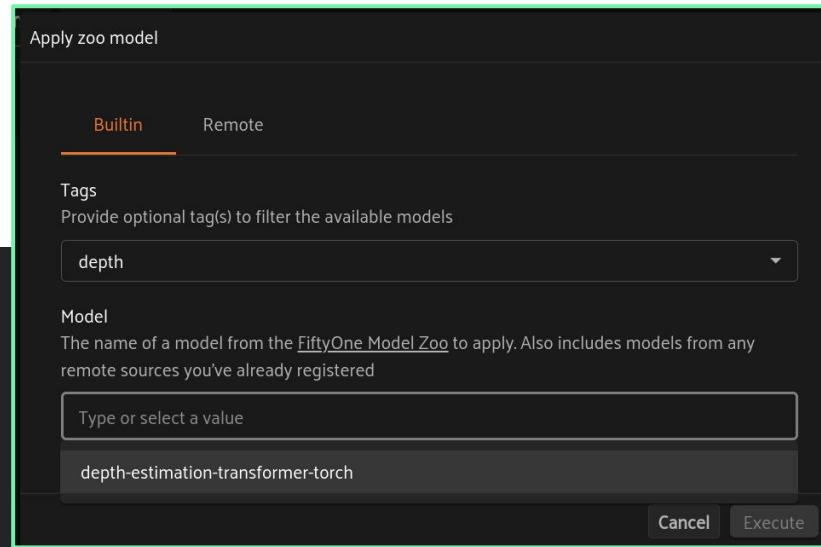


# Leveling Up Starts with More Information

How Can You Squeeze the Most Out of Your Dataset



Generating Embeddings

A screenshot of a web application interface for applying models from a model zoo. The interface has a dark theme. At the top, it says 'Apply zoo model'. Below this are two tabs: 'Builtin' (which is active) and 'Remote'. Under the 'Tags' section, there is a text prompt 'Provide optional tag(s) to filter the available models' followed by a dropdown menu containing the word 'depth'. The 'Model' section has a text prompt 'The name of a model from the [FiftyOne Model Zoo](#) to apply. Also includes models from any remote sources you've already registered' followed by a text input field containing 'depth-estimation-transformer-torch'. At the bottom right, there are two buttons: 'Cancel' and 'Execute'.

Applying Pretrained Models





# What Models Can Help?

A Short List of the Best Types



## Zero-Shot Predictors

Label just what you were looking for



## Depth

Know exactly how far away each car is



## Embedding Models

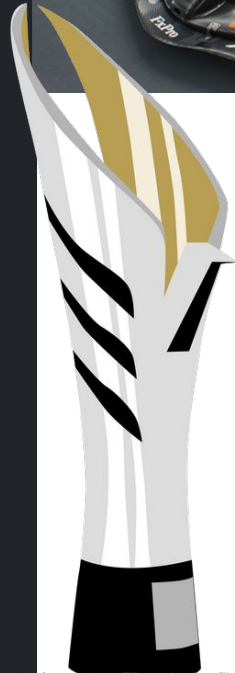
Find similar samples and explore unique samples



## 3D Embedding Models

Take it to the next level and generate embeddings on point clouds

# Expert Techniques



RACEFANS.NET

© XPB Images

# What's One of the Largest Blockers for Self Driving?



Time



# An Example



Remaining stopped at the zebra crossing



# How Do We Control Time?



# Just Control the World





# The Next Frontier



# Why Splats/NerFs aren't the full answer

What we learned from 2D  
Visual AI



Hard to Work With



Hard to Correct



Data Gaps Can Destroy  
Performance



# Just Create The World You Want





# Just Create The World You Want



# Before We Get Started

Link to Code



LinkedIn

