**Introduction**

**Problems:**
- Stereo matching networks poorly reconstruct object boundaries due to smoothness bias, causing bleeding artifacts in 3D point clouds.
- Existing methods are limited to discrete predictions at pixel locations of a fixed resolution image grid.
- Absence of large-scale, realistic and high-resolution stereo datasets with pixel-accurate ground truth.

**Goals:**
- Predicting accurate and sharp depth boundaries.
- High-resolution outputs with constant memory.

**Key Contributions:**
- A bimodal mixture distribution as output representation such that sharp discontinuities can be regressed.
- A continuous function formulation aimed at estimating disparities at arbitrary spatial resolution with constant memory footprint.
- A new large-scale synthetic binocular stereo dataset with ground truth disparities at 8Mpx resolution.

**Our Method**
- SMD-Nets predict a bimodal (Laplacian) mixture distribution which allows to accurately capture uncertainty close to depth discontinuities. By doing so, sharp discontinuities can be regressed despite the underlying neural networks make smooth predictions.
- Our framework use a (2D or 3D) convolutional stereo backbone in combination with a shallow multi-layer perceptron head that regresses the five distribution parameters from interpolated features.

**Stereo Super-Resolution**
- Our continuous formulation allows us to exploit ground truth at higher resolution than the input.

**Sampling Strategy**
- We adopt a Depth Discontinuity Aware sampling approach during training that favors points located near object boundaries.

**UnrealStereo4K Dataset**
- We introduce a photo-realistic binocular stereo dataset at 3840 × 2160 resolution with pixel-accurate ground truth.

**Qualitative Results**
- Our continuous formulation allows us to exploit ground truth at higher resolution than the input.

**References**

**Links**
- Paper
- Supplement
- Code
- Video