DEEP LEARNING AND ITS APPLICATIONS PROJECT PRESENTATION ON “MONOCULAR DEPTH ESTIMATION ON EGOCENTRIC DATASET”

GROUP-06

Anmol Passi (B16092)
Rajan Bajaj (B16072)
Sarthak Shekhawat (B16033)

PROBLEM STATEMENT

- We propose a novel training objective that enables our neural network to learn to perform depth estimation from egocentric dataset.

MOTIVATION AND APPLICATIONS

- Robotic Vision.
- Intelligent Visual Surveillance.
- Intelligent Driver assistance system.
- Fog Simulation.
- Object Trajectory Simulation.
- Absolute Distance Estimation.

PREVIOUS WORK

- There are some work on Depth Estimation by a variety of methods.

For example: Colour filtered dual camera, single camera video clip processing and dual camera mobile phone.

PREVIOUS RESEARCH

- Multirange real-time depth inference from a monocular stabilize footage using a fully convolutional neural network.
- End-to-end depth from motion with stabilized monocular videos.

(Clement Pinard, Laure Chevalley, Antoine Manzanera, David Filliat)

DATASET

- Egocentric Videos.
- UT Ego Dataset.
- Dataset can be generated if required.

PROPOSED METHODOLOGY

- Our network is inspired from DepthNet.
- Each convolution may be followed by a Spatial Batch Normalization and Relu or leaky Relu activation layer.
- Batch Normalization helps convergence and stability during training while normalizing a convolutional output over a batch.

PERFORMANCE PARAMETERS

- Like FlowNetS, we may use a multi-scale criterion, with a L1 reconstruction error for each scale.

$$Loss = \sum_{s \in scales} \gamma_s \frac{1}{H_s W_s} \sum_i \sum_j |\beta_s(i,j) - \zeta_s(i,j)|$$

$$Loss = \sum_{s \in scales} \gamma_s \frac{1}{H_s W_s} \sum_i \sum_j |\beta_s(i,j) - \zeta_s(i,j)|$$

- γ_s is the weight of the scale ,arbitrarily chosen.
- $(H_s, W_s) = (1/2^s H, 1/2^s W)$ are the height and weight of the output.
- ζ_s is the scaled depth groundtruth, using average pooling.
- β_s is the output of the network at scale s.

THANK YOU