Long-Distance Gesture Recognition using Dynamic Neural Networks

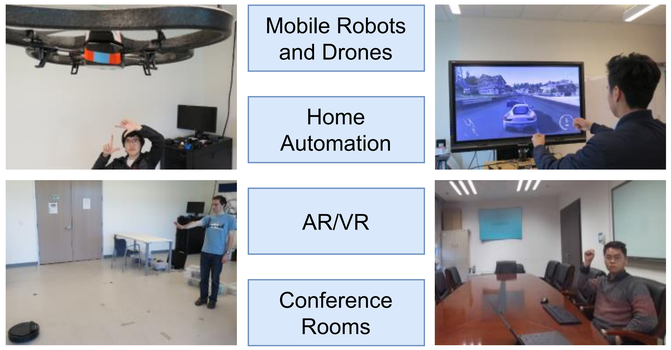
Shubhang Bhatnagar, Sharath Gopal, Narendra Ahuja, Liu Ren

BOSCH Invented for life

ILLINOIS

Scan To Know More

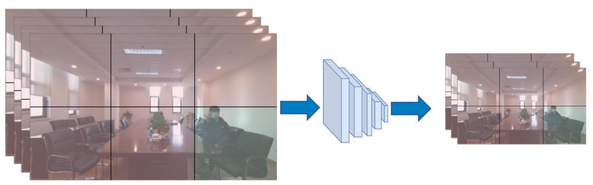
Motivation



- Gestures are a natural interface for communication between humans and machines
- Provide a convenient and contact-less way to communicate with robots

Challenges in Long Distance Recognition

Gesturing subject is small **&** 3D CNNs down sample input



Features might lose too much gesture information

Higher resolution video requires more compute, bandwidth!

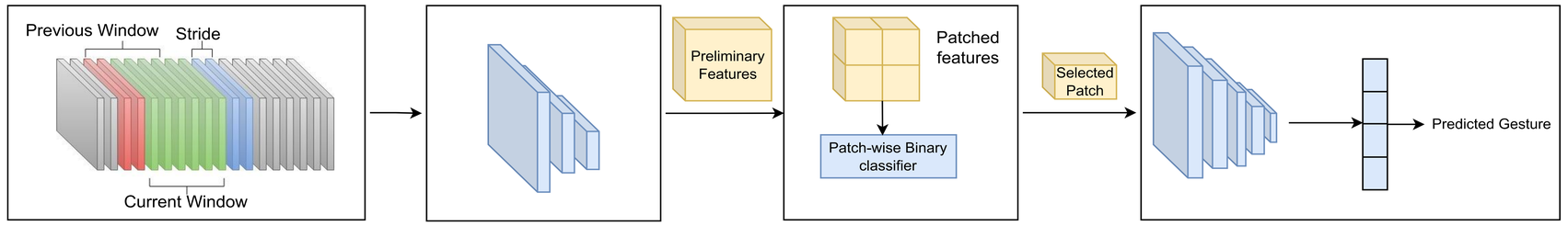
Dynamic Neural Network

Adapt computational graph to input at run-time

Discard background features

Preserve gesturing subject features

Proposed Spatially Dynamic Neural Network



Patch Selection Subnetwork

Binary Gesture Classifier

Core of Patch selection subnetwork

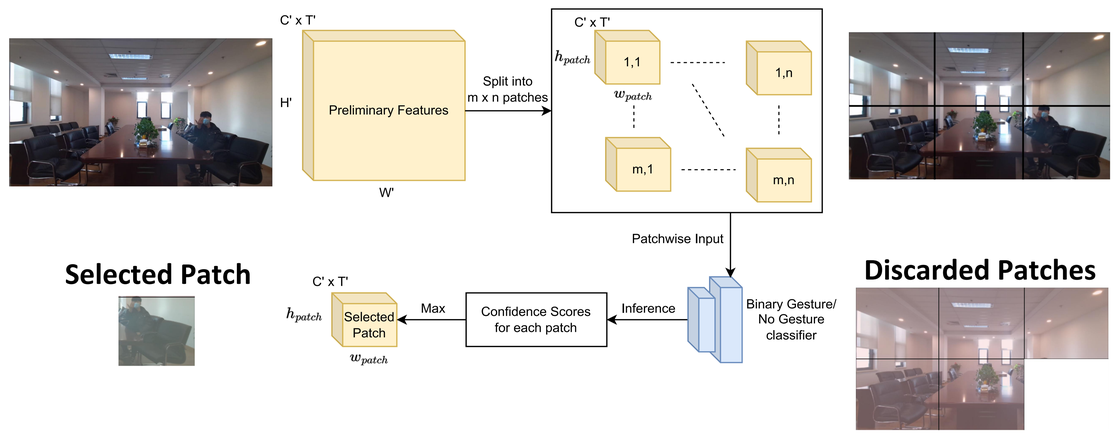
- Predicts if input patch features contain subject

Training Loss

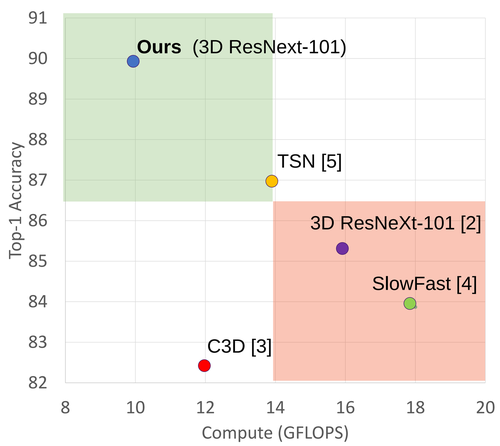
$L = L_{h_\phi} + \lambda L_{g_\psi}$

L_{h_ϕ} = Cross entropy for Gesture Recognition

L_{g_ψ} = Cross entropy for Patch Selection



Experiments and Results



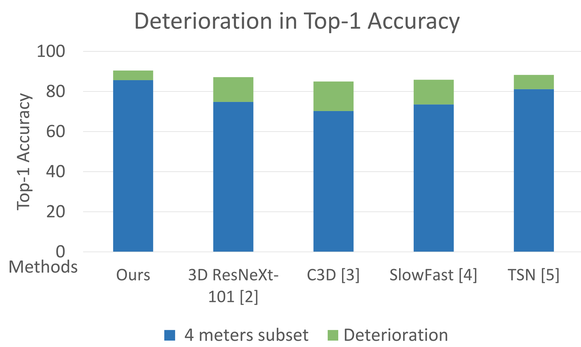
Lower compute & Better performance compared to state-of-the-art

Patches (m x n)	Compute (GFLOPS)	Accuracy %
1 x 2	26	86.48
2 x 2	18	88.67
2 x 3	**10**	**89.94**

Smaller patches improve accuracy and efficiency

Method	Compute (GFLOPS)	Top-1 Accuracy
Ours (3D MobileNet)	1.5	**76.68**
3D MobileNet [6]	1.5	65.33

Significant improvements in resource constrained robots



Lower performance deterioration with distance