# **VideoCapsuleNet: A Simplified Network for Action Detection (Summary)**

## Abstract

The main subject of the paper is **Action Detection**. The paper mainly used **Capsules** in their network to address this problem. Their network results **pixel-wise action segmentation** and **action classification**. The network is a generalized from 2D to 3D and that issued a problem of increasing the network parameters then the network has been a computationally expensive. They addressed the previous issue with **Capsule-Pooling.**

## Introduction

**Capsules** is one of the main parts used in VideoCapsuleNet, and we can define Capsules as a group of neurons which can model different or parts of entities. The network relies on **routing-by-agreement** as a routing algorithm between capsules. One of the advantages of the capsules is that **viewpoint invariant.** The network also used 3D convolutions. In general, the network aims to learn the **semantic information** necessary for action detection.