

# Input

# Encoder Module

# Cross Modal Fusion Mudule

# Classification Module

