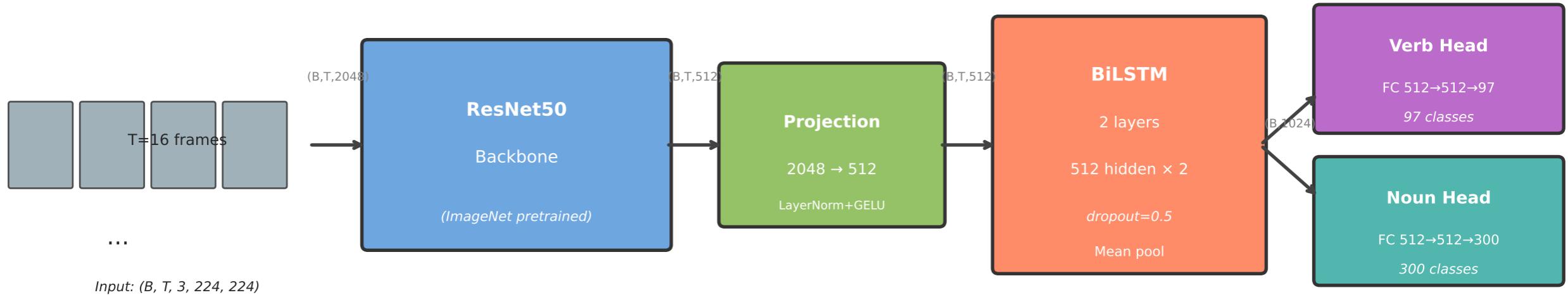


Model Architecture: ResNet50 + BiLSTM for Action Recognition



■ Spatial Encoder ■ Feature Projection ■ Temporal Model ■ Verb Classifier ■ Noun Classifier