

# Advanced Action Recognition Technologies

---

## Taylor Videos for Action Recognition

通過泰勒級數分解提取動作特徵,過濾靜態和雜訊。適用於兒童運動分析與隱私保護。

- 來源: [Taylor Videos for Action Recognition](#)

## Fusing Higher-Order Features in Graph Neural Networks for Skeleton-Based Action Recognition

整合角度編碼(AGE)於骨架模型,增強相似動作辨識能力。

- 來源: [Higher-Order Features in Graph Neural Networks](#)

### 1. Angular Encoding (AGE):

- Third-order features measuring angles between three body joints
- Static angular encoding captures spatial relationships
- Velocity angular encoding captures temporal dynamics
- Four categories of angles: local, center-oriented, pair-based, and finger-based

### 2. Network Architecture:

- Backbone with three spatial-temporal blocks (STBs)
- Each STB contains:
  - Spatial Multiscale Graph Convolution (SMGC) unit
  - Temporal Multiscale Convolution (TMC) units
- Uses S2GC (Simple Spectral Graph Convolution) as default GNN

### 3. Feature Fusion:

- Combines joint coordinates (first-order)
- Bone vectors (second-order)
- Angular features (third/fourth-order)
- Two fusion strategies:
  - Direct concatenation
  - Model ensembling

### 4. Training Approach:

- Supervised learning with cross-entropy loss
- Unsupervised feature coding and dictionary learning
- MAML-inspired fusion of supervised and unsupervised learning

## Meet JEANIE: A Similarity Measure for 3D Skeleton Sequences via Temporal-Viewpoint Alignment

解決時序和視角不一致問題的新技術:

- 聯合時序與視角配準
- 優化少樣本動作識別
- 提升兒童運動影像分析準確度
- 來源: [JEANIE: 3D Skeleton Sequence Similarity](#)

#### 1. Core Components:

- Temporal-viewpoint alignment using soft-DTW to match query-support sequence pairs
- Camera viewpoint simulation using either Euler angles or stereo projections
- Graph Neural Network (primarily S2GC) with optional transformer for feature encoding
- Dictionary learning and feature coding for unsupervised learning

#### 2. Network Architecture:

- 3-layer MLP for temporal block encoding
- GNN backbone (S2GC, APPNP, SGC or GCN)
- Optional transformer block for capturing dependencies
- Final FC layer for feature extraction

#### 3. Learning Approaches:

- Supervised FSAR with similarity-based loss function
- Unsupervised FSAR using dictionary learning
- Fusion strategies:
  - MAML-inspired fusion of supervised/unsupervised losses
  - Feature alignment between supervised/unsupervised representations
  - Weighted score fusion
  - Finetuning unsupervised model with supervised loss

#### 4. Key Technical Innovations:

- Joint optimization of temporal and viewpoint alignments
- Smooth viewpoint warping with controlled maximum shift
- Local temporal block matching with neighboring views
- Differentiable soft minimum for optimal path selection