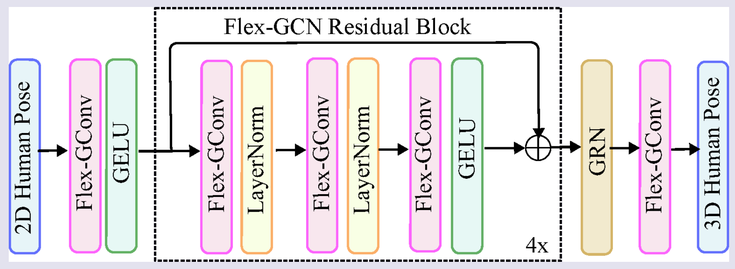
# Flexible Graph Convolutional Network for 3D Human Pose Estimation

**Authors** : Abu Taib Mohammed Shahjahan and A. Ben Hamza

## Contributions

- Present a flexible graph convolutional network (Flex-GCN), which captures highorder dependencies essential for reducing uncertainty due to occlusion or depth ambiguity in 3D human pose estimation.
- Designed network architecture that includes flexible graph convolutional layers and a global response normalization layer.

## Model Architecture




## Method

**Flexible Graph Convolutional Network.** Central to graph neural networks lies the fundamental concept of the feature propagation rule, which determines how information is transmitted among nodes in a graph. To this end, we propose a flexible graph convolutional network (Flex-GCN) with the following layer-wise update rule for node feature propagation:

$$\mathbf{H}^{(\ell+1)} = \sigma\Big(((1-s)\mathbf{I} + s\hat{\mathbf{A}})\hat{\mathbf{A}}\mathbf{H}^{(\ell)}\mathbf{W}^{(\ell)} + \mathbf{X}\widetilde{\mathbf{W}}^{(\ell)}\Big), \quad \ell = 0, \ldots, L-1$$

where $s \in (0,1)$ is a positive scaling parameter, $\mathbf{W}^{(\ell)}$ and $\widetilde{\mathbf{W}}^{(\ell)}$ are learnable weight matrices, $\sigma(\cdot)$ is an element-wise activation function, $\mathbf{H}^{(\ell)} \in \mathbb{R}^{N \times F_\ell}$ is the input feature matrix of the $\ell$-th layer with $F_\ell$ feature maps. The input of the first layer is $\mathbf{H}^{(0)} = \mathbf{X}$.

**Flex-GCN Network Design.**

- The input 2D pose undergoes a flexible graph convolutional layer, followed by a GELU activation function.
- Residual block consists of three flexible graph convolutional (Flex-GCNConv) layers.
- In each block, the first two convolutional layers are followed by layer normalization, while the third one is followed by GELU.
- This residual block is repeated four times. Then, a global response normalization (GRN) layer is applied after the residual blocks .
- The last flexible graph convolutional layer of the network generates the 3D pose.

## Results

**Comparison of our model and baseline methods in terms of MPJPE in millimeters, computed between the ground truth and estimated poses on Human3.6M under Protocol #1.**

| Method | Dire. | Disc. | Eat | Greet | Phone | Photo | Pose | Purch. | Sit | SitD. | Smoke | Wait | WalkD. | Walk | WalkT. | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Liu [2] | 46.3 | 52.2 | 47.3 | 50.7 | 55.5 | 67.1 | 49.2 | 46.0 | 60.4 | 71.1 | 51.5 | 50.1 | 54.5 | 40.3 | 43.7 | 52.4 |
| Zou [7] | 49.0 | 54.5 | 52.3 | 53.6 | 59.2 | 71.6 | 49.6 | 49.8 | 66.0 | 75.5 | 55.1 | 53.8 | 58.5 | 40.9 | 45.4 | 55.6 |
| Xu [4] | 47.1 | 52.8 | 54.2 | 54.9 | 63.8 | 72.5 | 51.7 | 54.3 | 70.9 | 85.0 | 58.7 | 54.9 | 59.7 | 43.8 | 47.1 | 58.1 |
| Zou [8] | 48.4 | 53.6 | 49.6 | 53.6 | 57.3 | 70.6 | 51.8 | 50.7 | 62.8 | 74.1 | 54.1 | 52.6 | 58.2 | 41.5 | 45.0 | 54.9 |
| Quan [3] | 47.0 | 53.7 | 50.9 | 52.4 | 57.8 | 71.3 | 50.2 | 49.1 | 63.5 | 76.3 | 54.1 | 51.6 | 56.5 | 41.7 | 45.3 | 54.8 |
| Zou [9] | 45.4 | 49.2 | 45.7 | 49.4 | 50.4 | 58.2 | 47.9 | 46.0 | 57.5 | 63.0 | 49.7 | 46.6 | 52.2 | 38.9 | 40.8 | 49.4 |
| Lee [1] | 46.8 | 51.4 | 46.7 | 51.4 | 52.5 | 59.7 | 50.4 | 48.1 | 58.0 | 67.7 | 51.5 | 48.6 | 54.9 | 40.5 | 42.2 | 51.7 |
| Zhang [5] | 45.0 | 50.9 | 49.0 | 49.8 | 52.2 | 60.9 | 49.1 | 46.8 | 61.2 | 70.2 | 51.8 | 48.6 | 54.6 | 39.6 | 41.2 | 51.6 |
| Ours | **40.2** | **45.8** | **45.0** | **46.8** | **48.6** | 54.0 | **42.4** | **42.1** | **53.2** | 66.7 | **45.6** | **45.4** | 48.8 | **38.4** | **40.1** | **46.9** |

**Comparison of our model and baseline methods in terms of PA-MPJPE, computed between the ground truth and estimated poses on Human3.6M under Protocol #2.**

| Method | Dire. | Disc. | Eat | Greet | Phone | Photo | Pose | Purch. | Sit | SitD. | Smoke | Wait | WalkD. | Walk | WalkT. | Avg. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Liu [2] | 35.9 | 40.0 | 38.0 | 41.5 | 42.5 | 51.4 | 37.8 | 36.0 | 48.6 | 56.6 | 41.8 | 38.3 | 42.7 | 31.7 | 36.2 | 41.2 |
| Zou [7] | 38.6 | 42.8 | 41.8 | 43.4 | 44.6 | 52.9 | 37.5 | 38.6 | 53.3 | 60.0 | 44.4 | 40.9 | 46.9 | 32.2 | 37.9 | 43.7 |
| Xu [4] | 36.7 | 39.5 | 41.5 | 42.6 | 46.9 | 53.5 | 38.2 | 36.5 | 52.1 | 61.5 | 45.0 | 42.7 | 45.2 | 35.3 | 40.2 | 43.8 |
| Zou [8] | 38.4 | 41.1 | 40.6 | 42.8 | 43.5 | 51.6 | 39.5 | 37.6 | 49.7 | 58.1 | 43.2 | 39.2 | 45.2 | 32.8 | 38.1 | 42.8 |
| Quan [3] | 36.9 | 42.1 | 40.3 | 42.1 | 43.7 | 52.7 | 37.9 | 37.7 | 51.5 | 60.3 | 43.9 | 39.4 | 45.4 | 31.9 | 37.8 | 42.9 |
| Zou [9] | 35.7 | 38.6 | **36.3** | 40.5 | **39.2** | 44.5 | 37.0 | 35.4 | 46.4 | **51.2** | 40.5 | **35.6** | 41.7 | 30.7 | 33.9 | 39.1 |
| Lee [1] | 35.7 | 39.6 | 37.3 | 41.4 | 40.0 | 44.9 | 37.6 | 36.1 | 46.5 | 54.1 | 40.9 | 36.4 | 42.8 | 31.7 | 34.7 | 40.3 |
| Zhang [5] | 35.3 | 39.3 | 38.4 | 40.8 | 41.4 | 45.7 | 36.9 | 35.1 | 48.9 | 55.2 | 41.2 | 36.3 | 42.6 | 30.9 | 33.7 | 40.1 |
| Ours | **34.1** | **38.0** | 36.8 | **39.7** | **39.2** | **43.6** | **33.4** | **34.5** | **44.2** | 57.1 | **38.3** | 36.0 | 41.0 | **29.9** | **33.1** | **38.6** |

| Method | PCK (↑) | AUC (↑) |
|---|---|---|
| Xu *et al*. [18] | 80.1 | 45.8 |
| Zeng *et al*. [20] | 82.1 | 46.2 |
| Lee *et al*. [4] | 81.6 | 50.3 |
| Zhang *et al*. [21] | 81.1 | 49.9 |
| Ours | **85.2** | **51.8** |

Table 3:MPI-INF-3DHP

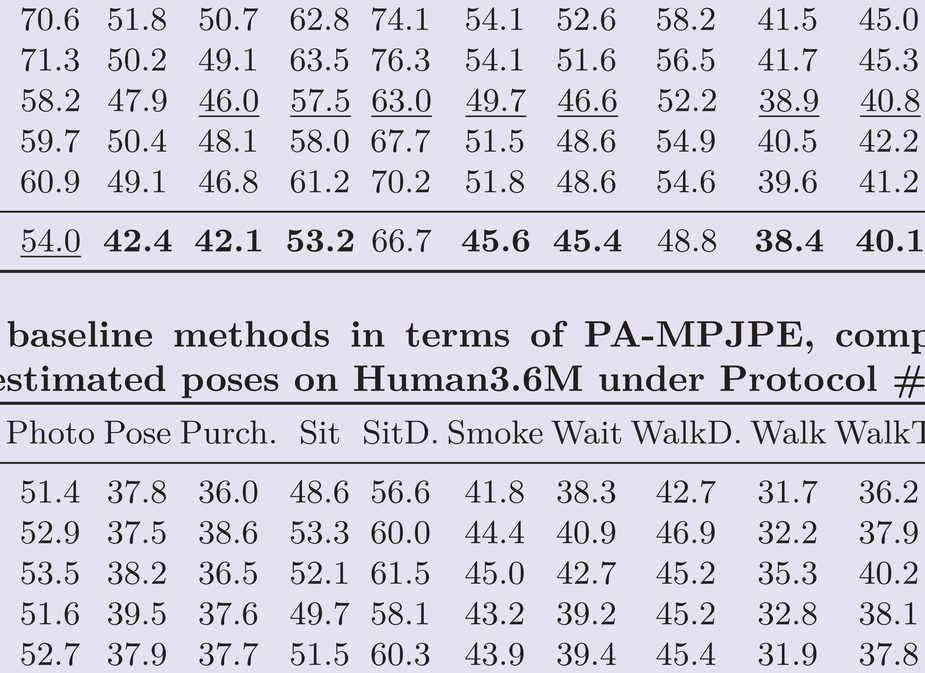



Figure : Visual comparison

## Results and Ablation study

**Performance comparison using the ground truth 2D pose as input.**

| Method | MPJPE (↓) | PA-MPJPE (↓) |
|---|---|---|
| SemGCN [6] | 42.14 | 33.53 |
| High-order GCN [7] | 39.52 | 31.07 |
| Modulated GCN [9] | 38.25 | 30.06 |
| Weight Unsharing [2] | 37.83 | 30.09 |
| Ours | **37.41** | **29.87** |

**Runtime analysis.**

| Method | Inference Time |
|---|---|
| High-Order GCN [7] | .013s |
| Weight Unsharing [2] | .032s |
| MM-GCN [1] | .009s |
| Modulated GCN [9] | .010s |
| Ours | 0.06s |

**Ablation study**

**Effect of initial residual connection (IRC).**

| Method | MPJPE (↓) | PA-MPJPE (↓) |
|---|---|---|
| Without IRC | 39.76 | 31.25 |
| With IRC | **37.41** | **29.87** |

**symmetry of modulation adjacency.**

| Method | MPJPE (↓) | PA-MPJPE (↓) |
|---|---|---|
| Without Symmetry | 37.99 | 30.11 |
| With Symmetry | **37.41** | **29.87** |

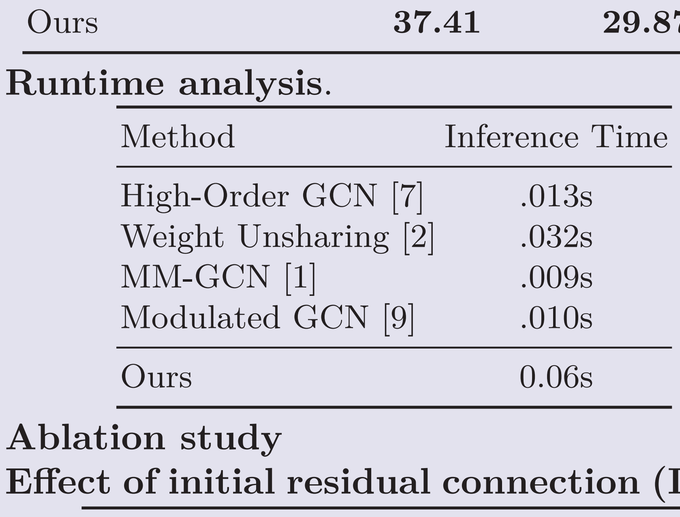



Figure :With and without pose refinement

Note on reproducibility: Code and pre-trained models are available:
https://github.com/shahjahan0275/Flex-GCN