Semantic Graph Convolutional Networks for 3D Human Pose Regression

Yeong Mo Kim

Seoul National University

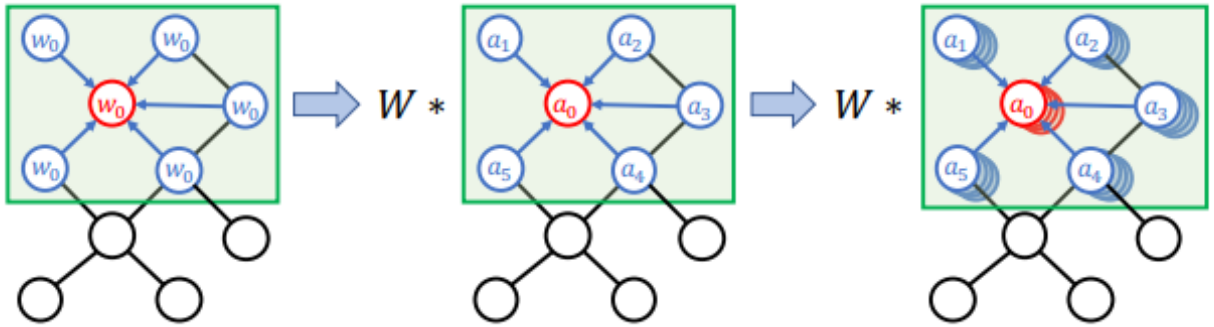
Semantic Graph Convolutions

- Convolution operation for “vanilla” GCN(ResGCN)

$$\mathbf{X}^{(l+1)} = \sigma\left(\mathbf{W}\mathbf{X}^{(l)}\tilde{\mathbf{A}}\right),$$

- Convolution operation for SemGCN

$$\mathbf{X}^{(l+1)} = \sigma\left(\mathbf{W}\mathbf{X}^{(l)}\rho_i\left(\mathbf{M}\odot\mathbf{A}\right)\right),$$

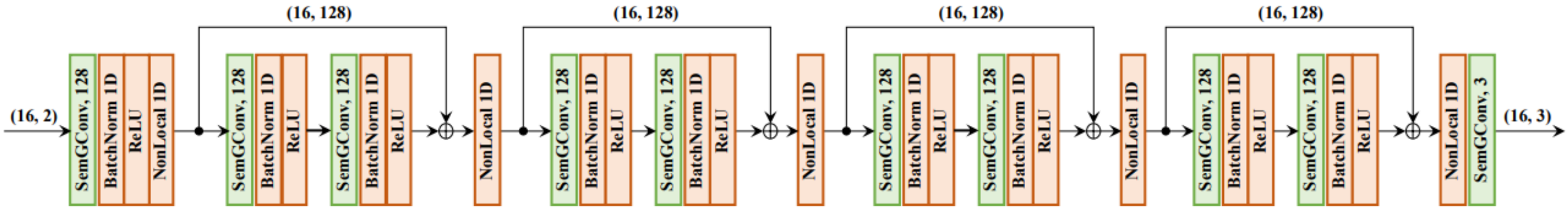


Semantic Graph Convolution

$$\mathbf{X}^{(l+1)} = \Big\Vert_{d=1}^{D_{l+1}} \sigma\left(\vec{\boldsymbol{w}}_d\mathbf{X}^{(l)}\rho_i\left(\mathbf{M}_d\odot\mathbf{A}\right)\right),$$

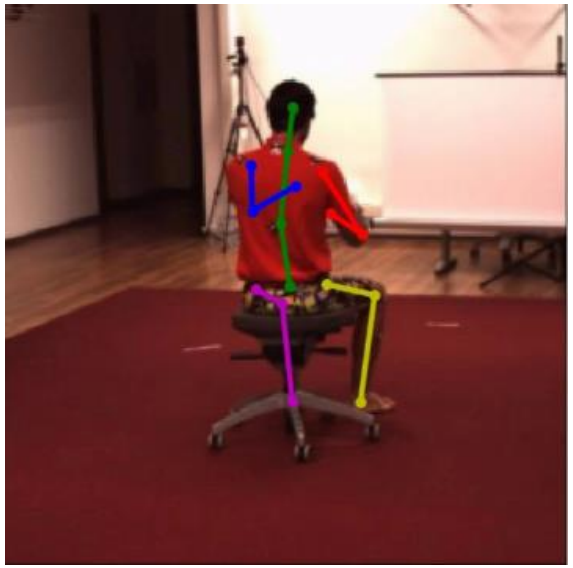
SemGCN Architecture

- One residual block built by two SemGConv layers
- One residual block is followed by one non-local layer

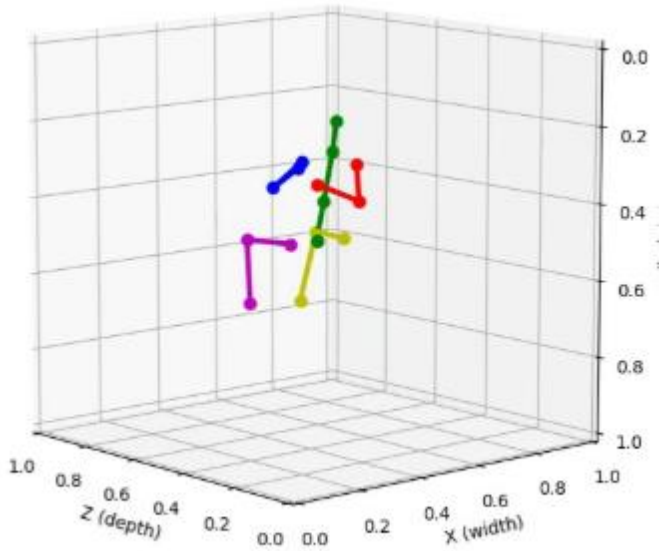


3D Human Pose Regression

- 3D pose estimation without depth imformation is really hard problem
- Recently, it has proven that 2D pose estimation is crucial for 3D pose estimation



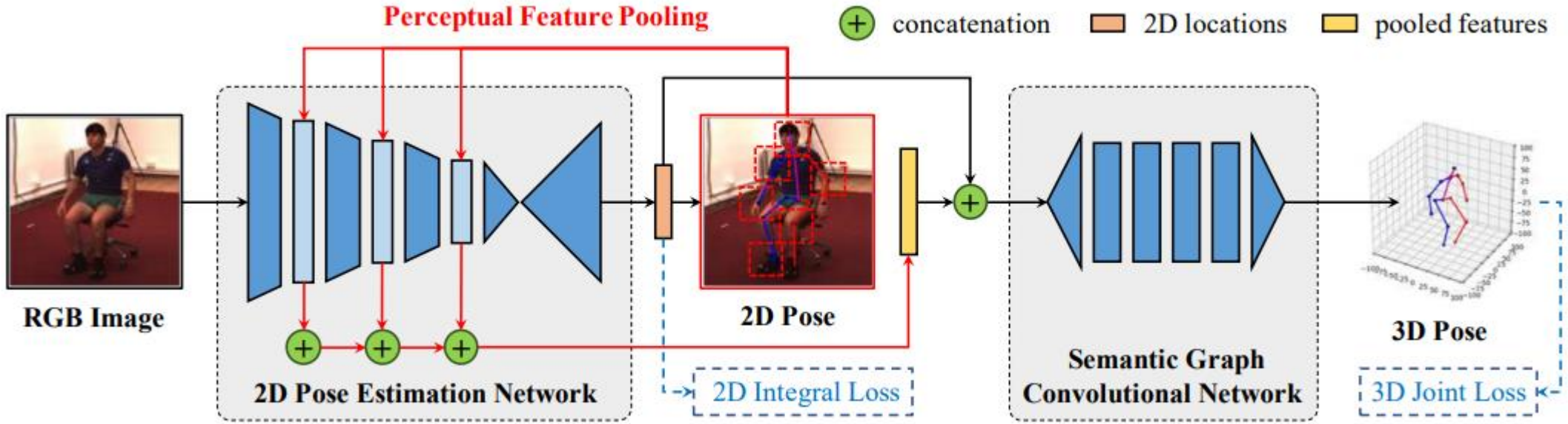
2D RGB image
+ 2D Pose estimation



3D Pose estimation

3D Human Pose Regression

- Whole framework consist of two neural network:
  - 2D Pose Estimation Network(ResNet, Hourglass)
  - 3D Pose Estimation Network(SemGCN)
- Pool features from multiple layers in ResNet
- Inputs of SemGCN: 2D locations(coordinates), pooled feature from ResNet



Results

Method	# of params	MPJPE (mm)
aGCN [68] / GAT [60]	0.16M	82.9
ST-GCN [67]	0.27M	57.4
FC [34]	4.29M	45.5 (62.9)
FC [34] w/ PG [13]	-	43.3 (60.4)
Ours	0.43M	43.8 (61.1)
Ours w/ PG [13]	-	42.5 (59.8)

Protocol #1	Direct.	Discuss	Eating	Greet	Phone	Photo	Pose	Purch.	Sitting	SittingD.	Smoke	Wait	WalkD.	Walk	WalkT.	Avg.
Ionescu *et al.* [24] PAMI'16	132.7	183.6	132.3	164.4	162.1	205.9	150.6	171.3	151.6	243.0	162.1	170.7	177.1	96.6	127.9	162.1
Tekin *et al.* [57] CVPR'16	102.4	147.2	88.8	125.3	118.0	182.7	112.4	129.2	138.9	224.9	118.4	138.8	126.3	55.1	65.8	125.0
Zhou *et al.* [77] CVPR'16	87.4	109.3	87.1	103.2	116.2	143.3	106.9	99.8	124.5	199.2	107.4	118.1	114.2	79.4	97.7	113.0
Du *et al.* [11] ECCV'16	85.1	112.7	104.9	122.1	139.1	135.9	105.9	166.2	117.5	226.9	120.0	117.7	137.4	99.3	106.5	126.5
Chen & Ramanan [7] CVPR'17	89.9	97.6	89.9	107.9	107.3	139.2	93.6	136.0	133.1	240.1	106.6	106.2	87.0	114.0	90.5	114.1
Pavlakos *et al.* [41] CVPR'17	67.4	71.9	66.7	69.1	72.0	77.0	65.0	68.3	83.7	96.5	71.7	65.8	74.9	59.1	63.2	71.9
Mehta *et al.* [35] 3DV'17	52.6	64.1	55.2	62.2	71.6	79.5	52.8	68.6	91.8	118.4	65.7	63.5	49.4	76.4	53.5	68.6
Zhou *et al.* [75] ICCV'17	54.8	60.7	58.2	71.4	62.0	65.5	53.8	55.6	75.2	111.6	64.1	66.0	51.4	63.2	55.3	64.9
Martinez *et al.* [34] ICCV'17	51.8	56.2	58.1	59.0	69.5	78.4	55.2	58.1	74.0	94.6	62.3	59.1	65.1	49.5	52.4	62.9
Sun *et al.* [53] ICCV'17	52.8	54.8	54.2	**54.3**	61.8	53.1	53.6	71.7	86.7	61.5	67.2	**53.4**	47.1	61.6	53.4	59.1
Fang *et al.* [13] AAAI'18	50.1	54.3	57.0	57.1	66.6	73.3	53.4	55.7	72.8	88.6	60.3	57.7	62.7	47.5	50.6	60.4
Yang *et al.* [69] CVPR'18	51.5	58.9	**50.4**	57.0	62.1	65.4	49.8	52.7	69.2	85.2	**57.4**	58.4	43.6	60.1	47.7	58.6
Hossain & Little [21] ECCV'18	48.4	**50.7**	57.2	55.2	63.1	72.6	53.0	**51.7**	**66.1**	80.9	59.0	57.3	62.4	**46.6**	49.6	58.3
Ours (HG)	48.2	60.8	51.8	64.0	64.6	53.6	51.1	67.4	88.7	57.7	73.2	65.6	48.9	64.8	51.9	60.8
Ours (RN w/ FP)	**47.3**	60.7	51.4	60.5	**61.1**	**49.9**	**47.3**	68.1	86.2	**55.0**	67.8	61.0	**42.1**	60.6	**45.3**	**57.6**
Ours (GT)	37.8	49.4	37.6	40.9	45.1	41.4	40.1	48.3	50.1	42.2	53.5	44.3	40.5	47.3	39.0	43.8

Public Code

- Github public code link: https://github.com/garyzhao/SemGCN
- Dataset setup
  - Data 폴더에 h36m dataset zip 파일로 다운(license 획득에 1주일 정도 필요, 링크 상에 있는 아카이브 파일은 현재 다운 불가)
  - Dataset folder README 파일에 있는 Setup from original source를 보고 Human3.6m 데이터셋의 2D, 3D pose npy 파일을 생성
- Train&Evaluate
  - Evaluating pretrained model

```
python main_gcn.py --non_local --evaluate checkpoint/pretrained/ckpt_semgcn_nonlocal.pth.tar
python main_gcn.py --non_local --evaluate checkpoint/pretrained/ckpt_semgcn_nonlocal_sh.pth.tar --keypoints sh_ft_h36m
```

  - Train from scratch

```
python main_gcn.py --non_local --epochs 100 --keypoints sh_ft_h36m
python main_gcn.py --non_local --evaluate ${CHECKPOINT_PATH} --keypoints sh_ft_h36m
```

Code structure

- Main files: main_gcn.py(proposed method), main_linear.py(baseline)
- Network: model folder에 있음, sem_gcn.py(SemGCN 파일)에서 sem_graph_conv.py(SemGConv layer)를 불러와 SemGCN class를 구성 sem_graph_conv.py 대신 sem_ch_graph_conv.py를 사용할 경우 channelwise SemGConv layer를 사용가능
- Common: dataloader, train, loss, visualize 함수 등 train과 evaluation에 필요한 함수들 전반 포함
- Reproducing result example(train from scratch)