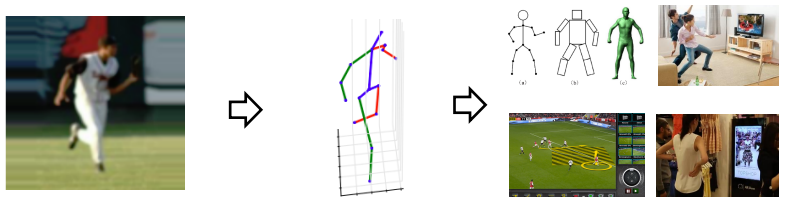
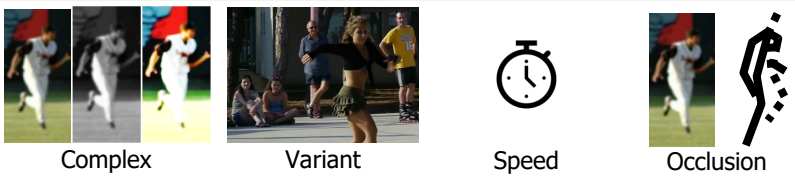
# Multi-Layer Multi-Task Detection Neural Network Based Real-Time Multi-Person Pose Estimation from Monocular Camera

修士課程卒業　羅　丁力

## Background

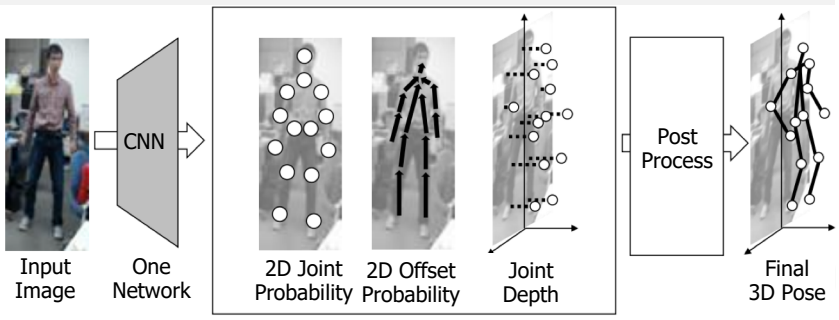



### Challenges

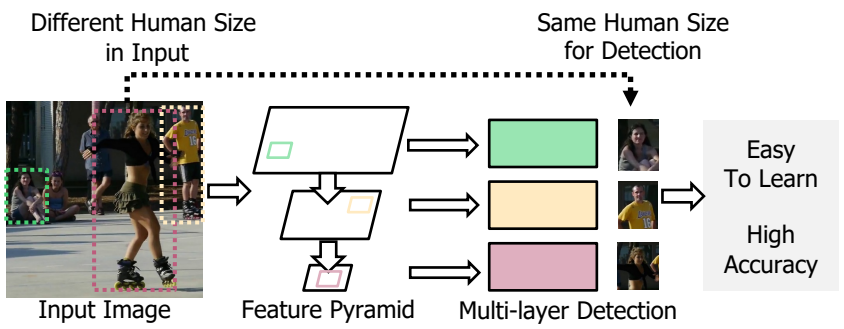



## Proposals

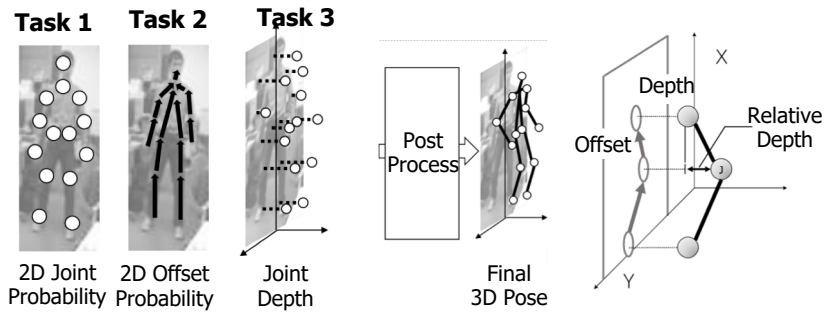
### Framework




### Proposal 1: Multi-layer detection based high-speed pose estimation network structure

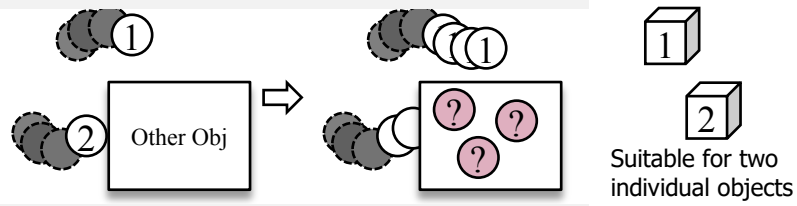



### Proposal 2: One-shot multi-task 3D pose estimation neural network

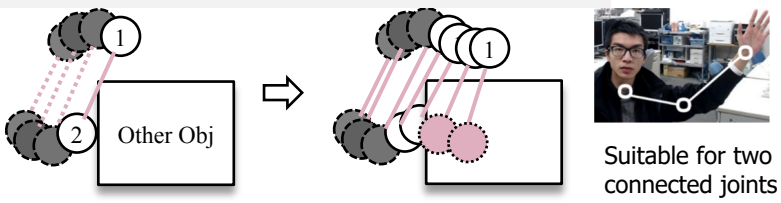



### Proposal 3: Physically-constrained temporal post processing

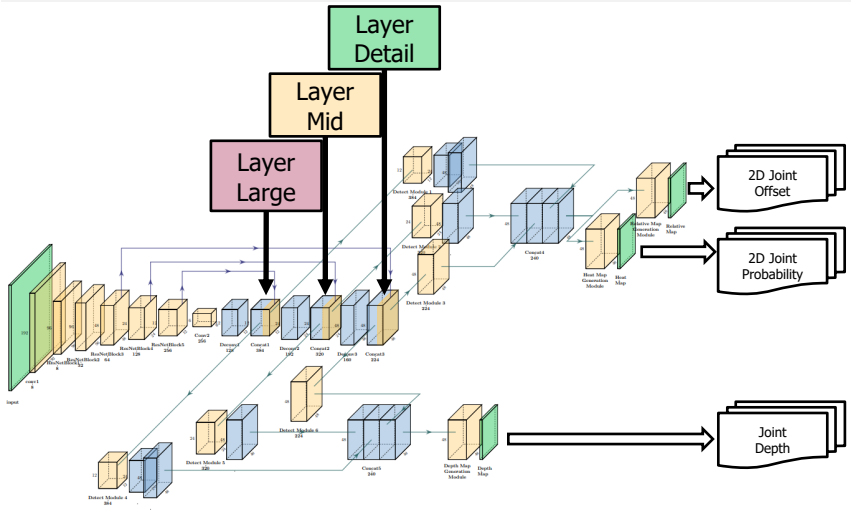
#### Point based temporal smooth




#### Ours: Physically-constrained temporal smooth

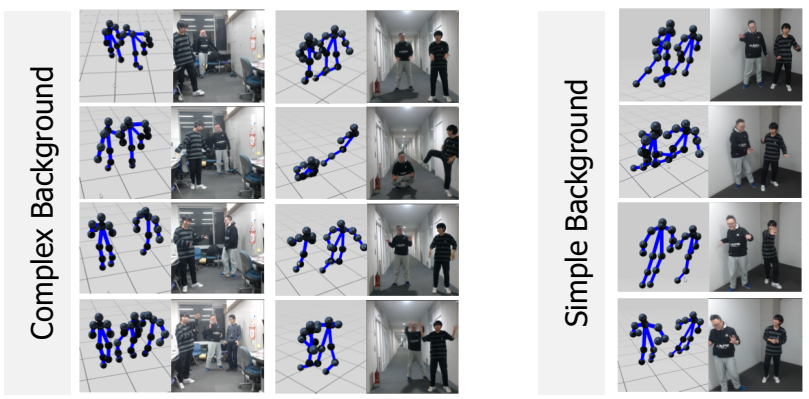



### Final Neural Network Structure




## Experiment Result

| Methods | Input | MPJPE(mm) | Speed(ms) |
|---|---|---|---|
| Martinez et al. | 2D Pose Input | 62.9 | - |
| Zhou et al. | Video | 64.9 | - |
| Open Pose (high accuracy mode) + Martinez et al. | RGB Image | 70.5 | 833.3 |
| Open Pose (high speed mode) + Martinez et al. | RGB Image | 82.7 | 113.6 |
| LCR-Net | RGB Image | 87.7 | - |
| Ours | RGB Image | 112.9 | 46.5 |




## Conclusion

Multi-person 3D pose estimation system from monocular camera in real-time speed

Graduate School of Information, Production and Systems
Waseda University