PROPEL: Probabilistic Parametric Regression Loss for Convolutional Neural Networks

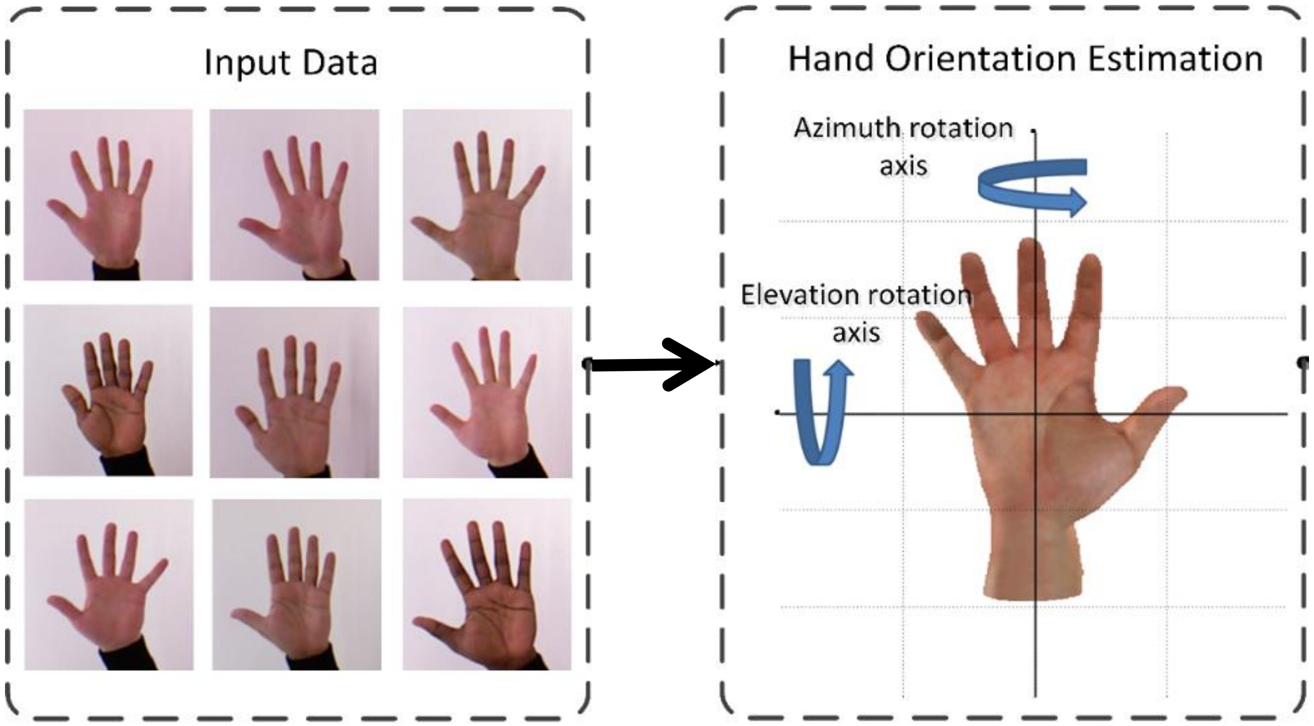
Muhammad Asad, Rilwan Basaru, S M Masudur Rahman Al Arif, and Greg Slabaugh

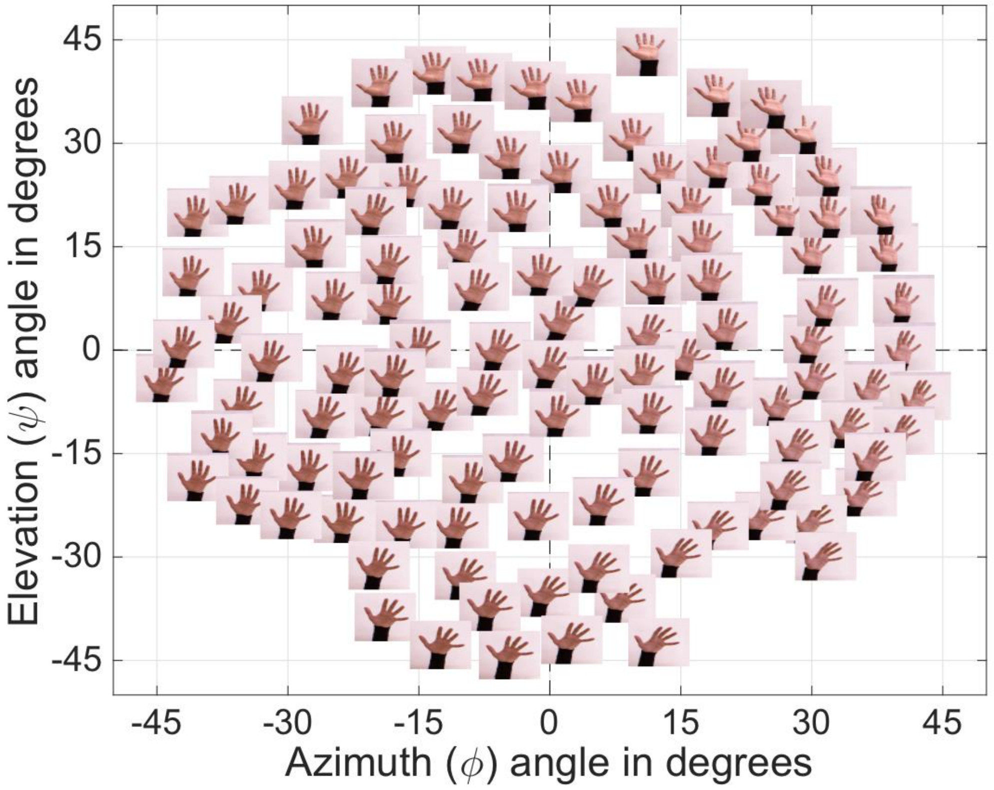
CITY
UNIVERSITY OF LONDON
EST 1894

ICPR 2020

Problem Definition

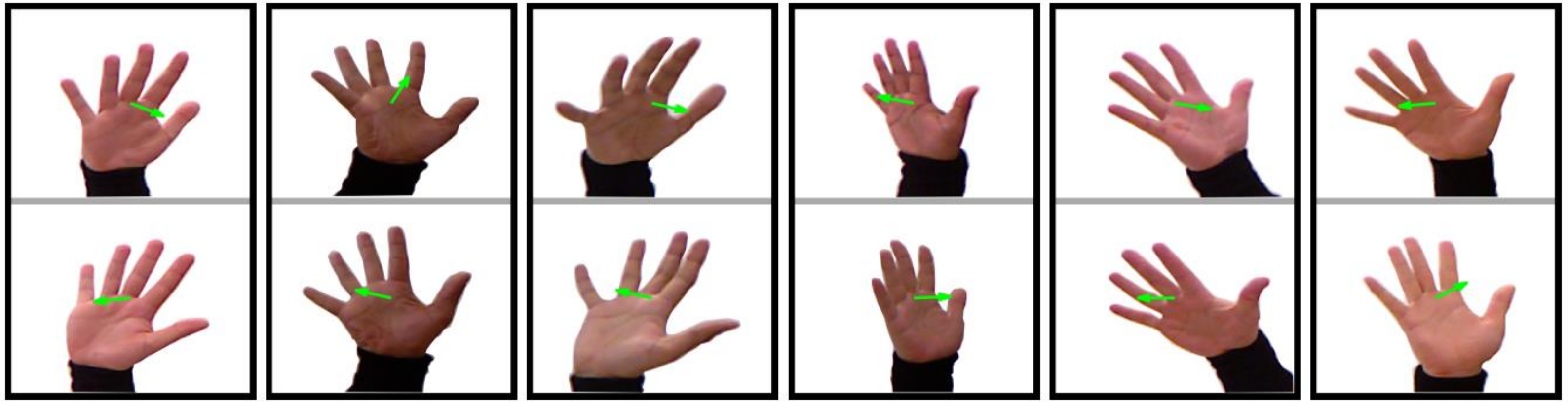
- Can we use a machine learning model to learn the mapping of 2D images onto 3D hand orientation?



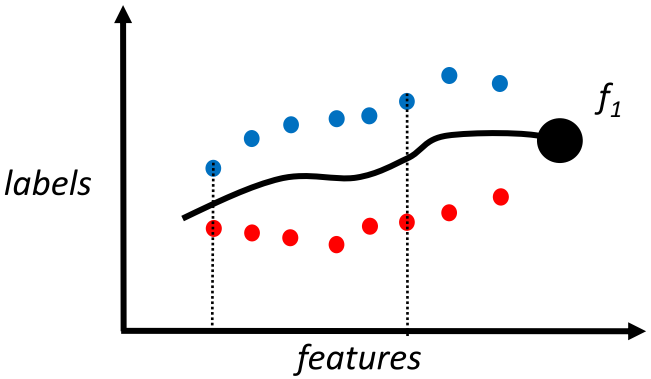


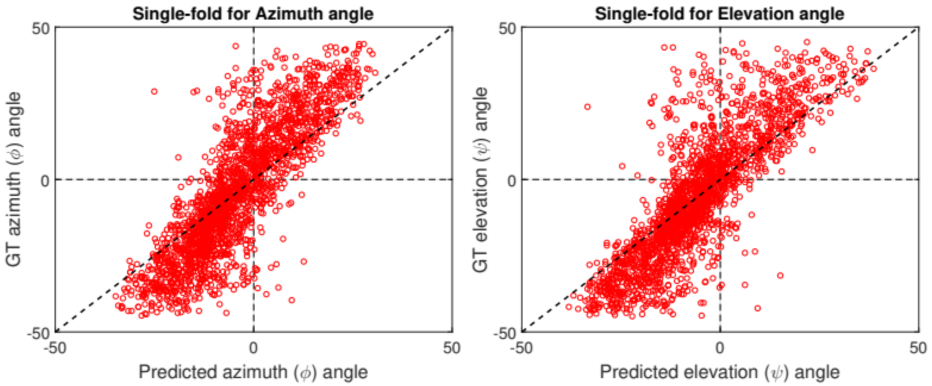
Why Probabilistic Regression?

- Symmetry problem: opposite orientation ←→ similar hand shapes



- Existing regression methods *try to fit* into the data [1]





- Motivates the need for probabilistic regression to handle ambiguity

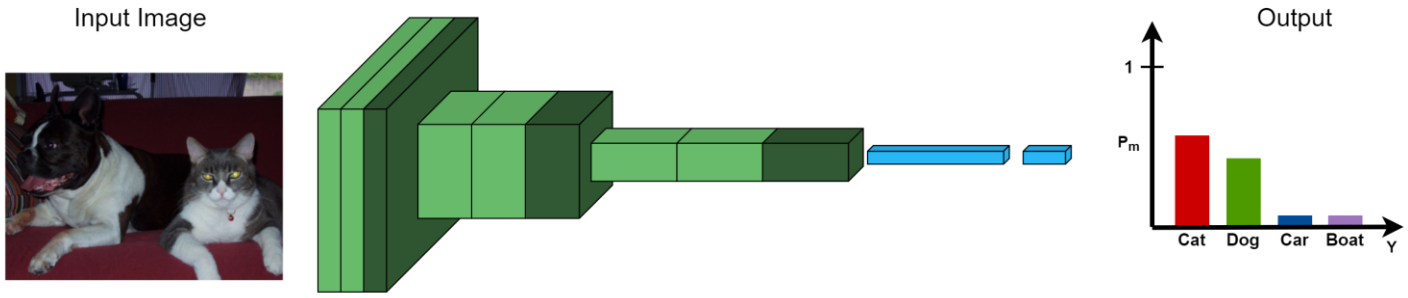
[1] M. Asad, G. Slabaugh. **"Learning marginalization through regression for hand orientation inference."** *CVPR Workshop*. 2016.

Existing Probabilistic Learning with CNNs

Let P_{gt} be ground truth target distribution, CNN learns P_m using loss functions:

Task | Network | Loss

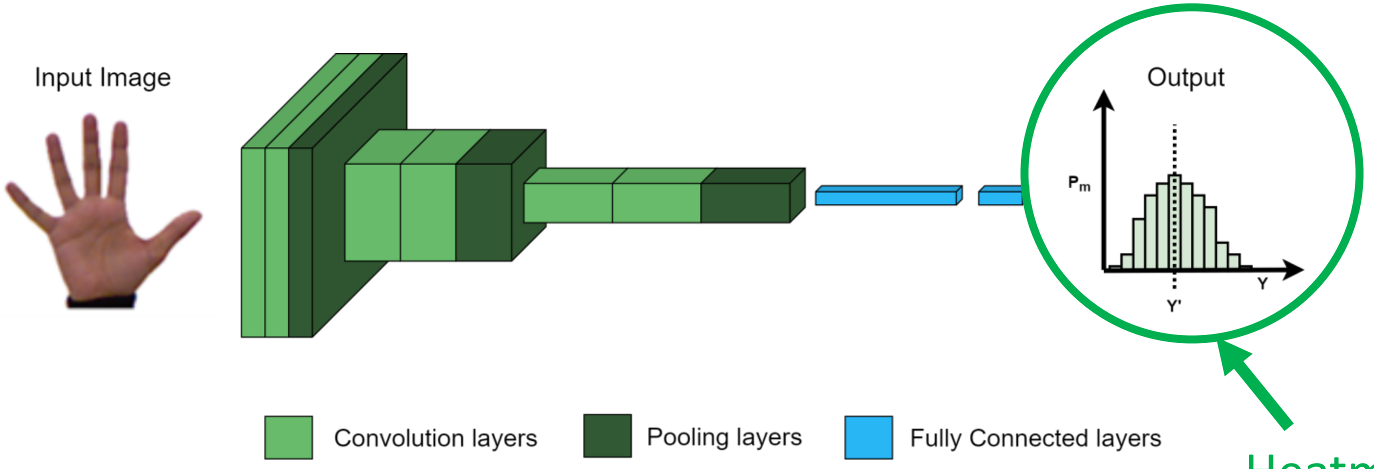
Classification



$$\mathcal{L}(P_{gt}, P_m) = -\sum_{i \in I} P_{gt}^i \log(P_m^i)$$

Cross Entropy

Regression



Heatmap

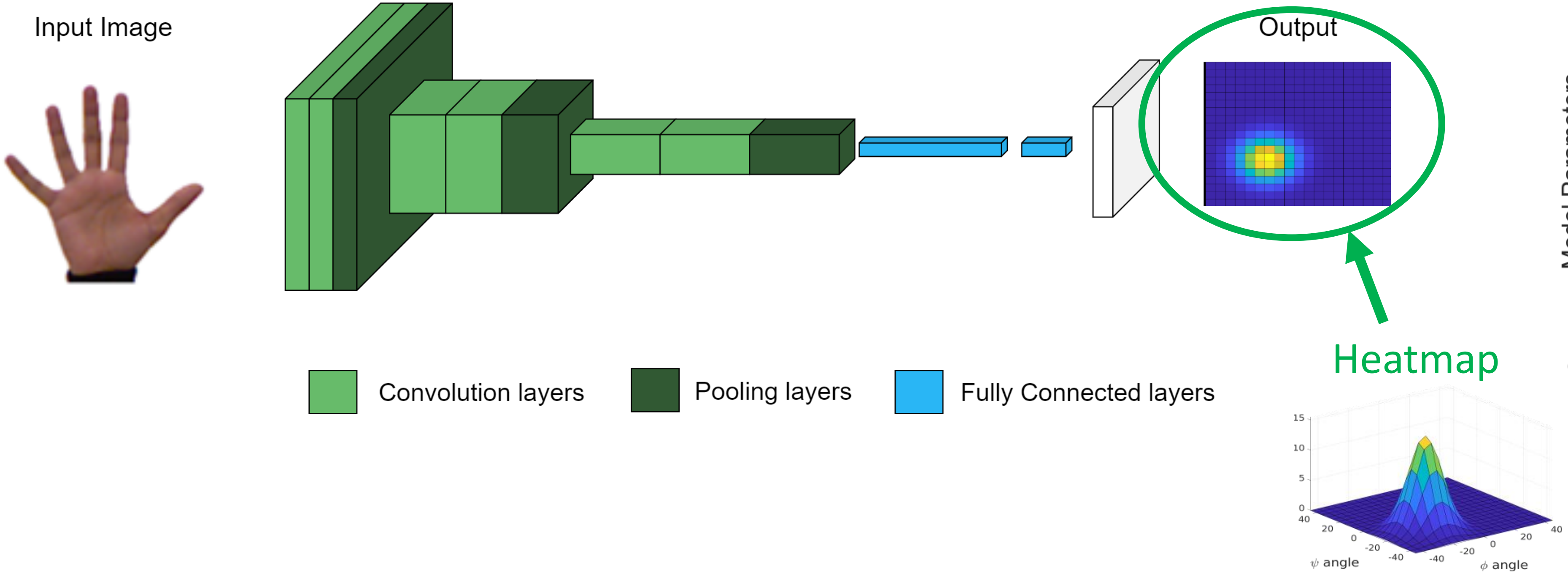
$$\mathcal{L}(P_{gt}, P_m) = \frac{1}{2}\sum_{i \in I} (P_{gt}^i - P_m^i)^2$$

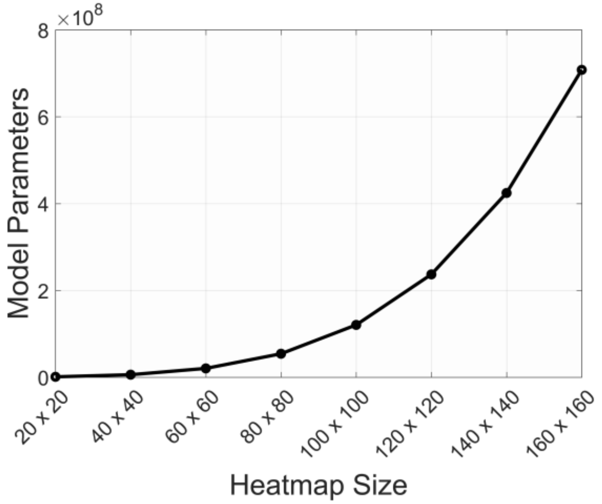
Mean Squared Error (MSE)

Existing Probabilistic Regression using CNNs

CNN learns probability heatmap distribution

- Require additional model parameters (as compared to directly learning target)
- Discretized target space → error in model output
- Higher dimensional target → exponential increase in parameters
- Increased model complexity → overfitting

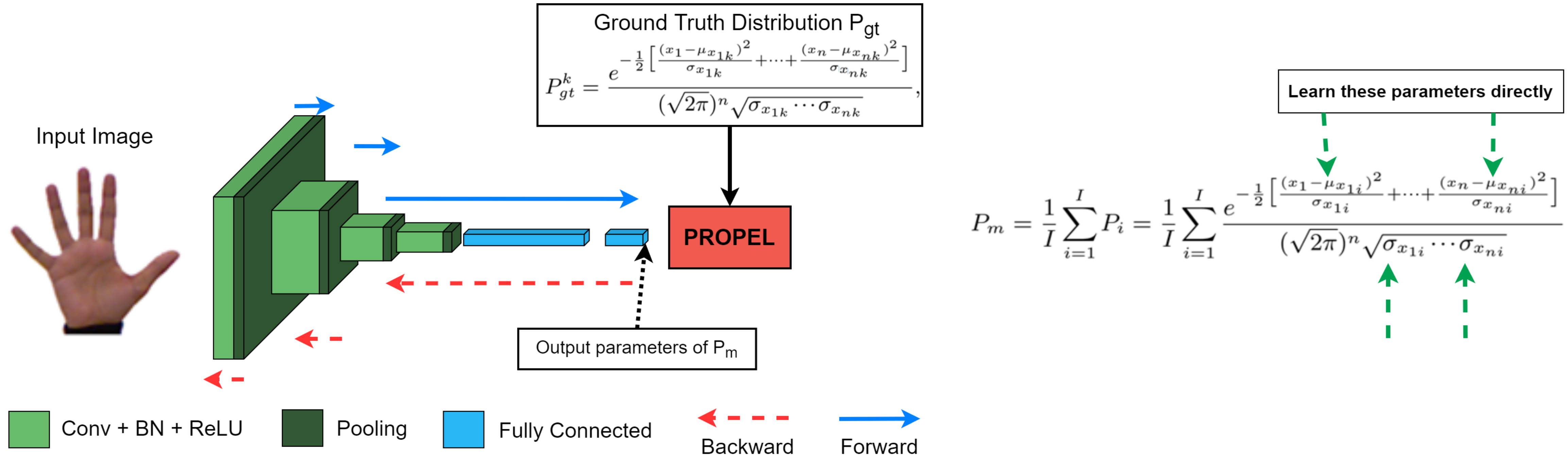




PRObabilistic Parametric rEgression Loss (PROPEL)

Contributions:

- Enables CNNs → learn parameters of a mixture of Gaussians probability distribution
- Fully-differentiable → analytic closed form solution → works with standard CNNs/optimizers
- Generalized to → higher dimensional targets → multi-modal distributions
- Better generalization with **10x less model parameters**



PROPEL Definition

- Let $\mathbf{x} = \{x_1, x_2, \cdots, x_n\}^\intercal \in \mathbb{R}^n$ define target prediction space
- PROPEL is defined as (using metric from [*]):

$$\boxed{L = -\log\left[\frac{2\int P_{gt} P_m \, d\underline{\mathbf{x}}}{\int ({P_{gt}}^2 + {P_m}^2) \, d\underline{\mathbf{x}}}\right]}$$

$$P_{gt}^k = \frac{e^{-\frac{1}{2}\left[\frac{(x_1 - \mu_{x_{1k}})^2}{\sigma_{x_{1k}}} + \cdots + \frac{(x_n - \mu_{x_{nk}})^2}{\sigma_{x_{nk}}}\right]}}{(\sqrt{2\pi})^n \sqrt{\sigma_{x_{1k}} \cdots \sigma_{x_{nk}}}},$$

P_{gt} : n-dimensional ground truth PDF

$$P_m = \frac{1}{I}\sum_{i=1}^{I} P_i = \frac{1}{I}\sum_{i=1}^{I} \frac{e^{-\frac{1}{2}\left[\frac{(x_1 - \mu_{x_{1i}})^2}{\sigma_{x_{1i}}} + \cdots + \frac{(x_n - \mu_{x_{ni}})^2}{\sigma_{x_{ni}}}\right]}}{(\sqrt{2\pi})^n \sqrt{\sigma_{x_{1i}} \cdots \sigma_{x_{ni}}}}$$

P_m : n-dimensional mixture of Gaussian learned model PDF

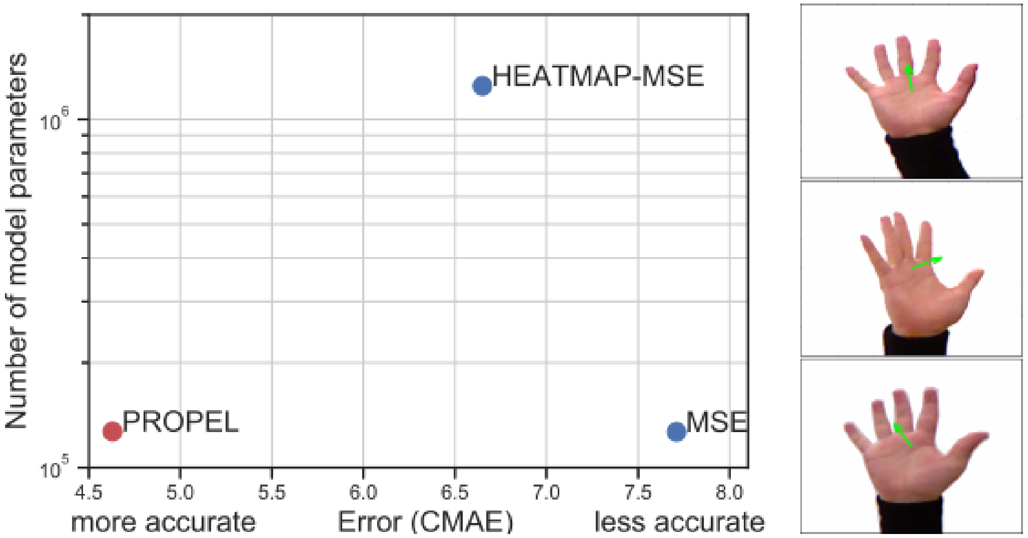
- Partial derivatives for optimizing each parameter in model PDF P_m:

$$\frac{\partial L}{\partial \mu_{x_{ni}}} = -\frac{1}{T1}\left[\frac{\partial G(P_{gt}, P_i)}{\partial \mu_{x_{ni}}}\right] + \frac{1}{T2}\left[\frac{2}{I^2}\sum_{i<j}^{I} \frac{\partial G(P_i, P_j)}{\partial \mu_{x_{ni}}}\right],$$

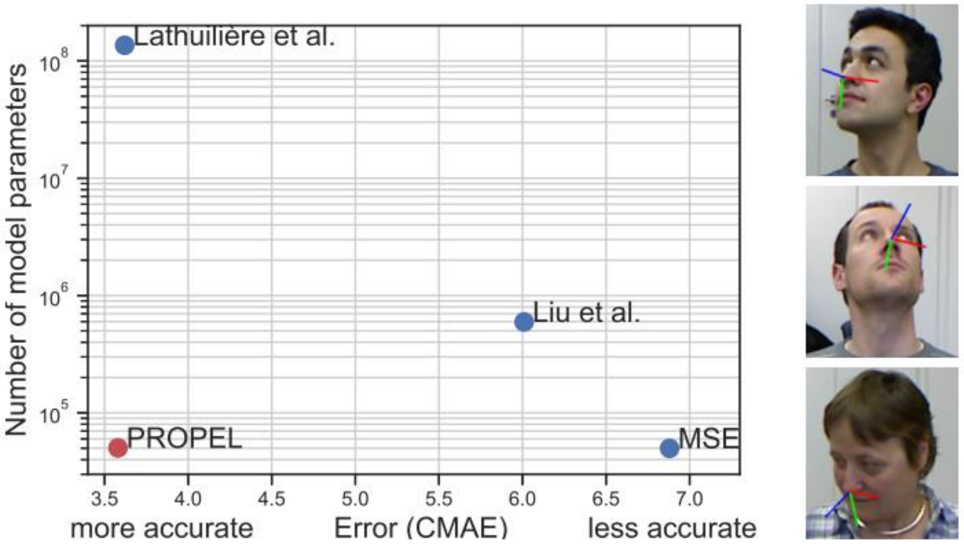
$$\frac{\partial L}{\partial \sigma_{x_{ni}}} = -\frac{1}{T1}\left[\frac{\partial G(P_{gt}, P_i)}{\partial \sigma_{x_{ni}}}\right] + \frac{1}{T2}\left[\frac{1}{I^2}\frac{\partial H(P_i)}{\partial \sigma_{x_{ni}}} + \frac{2}{I^2}\sum_{i<j}^{I} \frac{\partial G(P_i, P_j)}{\partial \sigma_{x_{ni}}}\right]$$

[*] S. Giorgos, et al. "**An analytic distance metric for Gaussian mixture models with application in image retrieval.**" *International Conference on Artificial Neural Networks (ICANN)*. 2005.

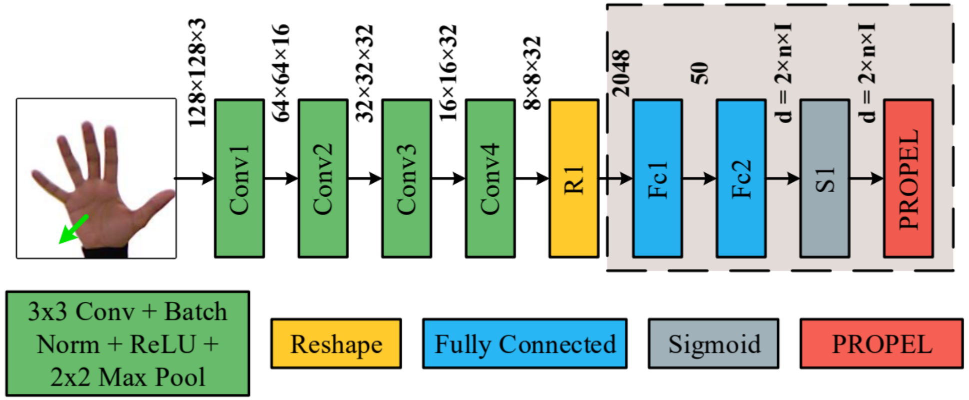
Experimental Validation: Accuracy + Efficiency

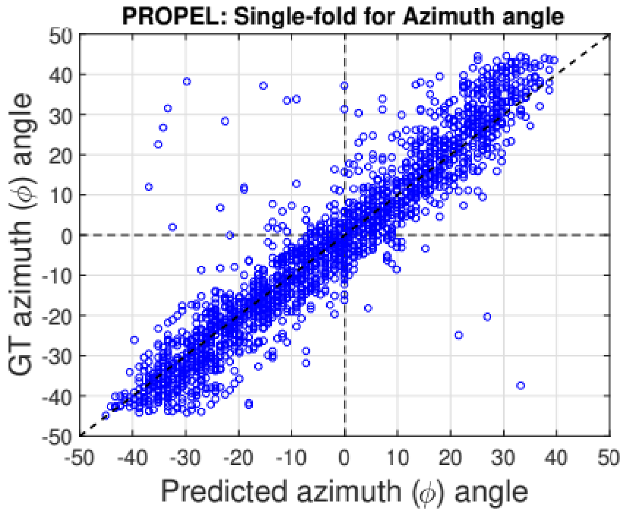


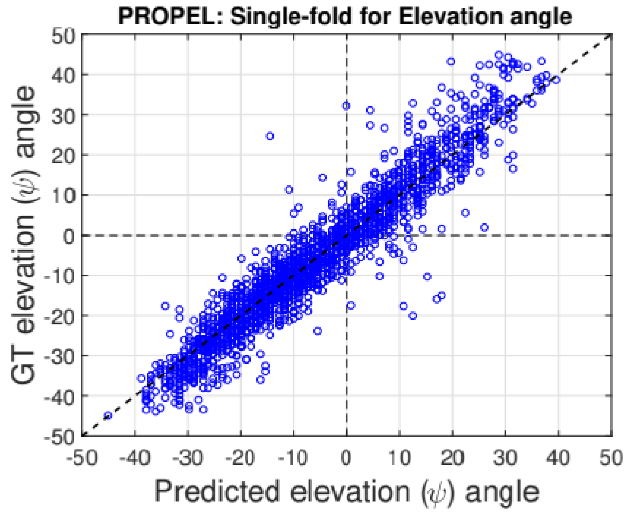
(a) Hand orientation estimation



(b) Head orientation estimation







Conclusion

- Importance of Probabilistic Regression
- Limitations with existing heatmap based CNN Regression
- PROPEL: enables learning parameters of probability distribution, achieves state-of-the-art accuracy with 10x less model parameters

Future Work:

- Look at higher dimensional targets learning, e.g. human body/hand pose estimation
- Selecting the number of Gaussians in model distribution
- Introduce covariance to learn covarying targets

See you at poster session T1.1 ☺