**Motivation**

- **Problem statement:** few-shot 3D human motion prediction

- **Input Sequence:** $X$

- **Prediction Sequence:** $\hat{Y}$

- **Time**

- **Key insight:** good generalization from few examples relies on
  - A generic initial model
  - An effective strategy for adapting this model to novel tasks

- **Our approach:** proactive and adaptive meta-learning (PAML)

  - **Model-agnostic meta-learning:** generic initialization through meta-learning [1]
  - **Model regression network:** learning-to-learn transformation from few-shot to many-shot models [2]
  - **Novel combination:** an integrated, end-to-end framework

**Meta-Learning for Human Motion Prediction**

- **Motion predictor $P_0$:** (e.g., learner; historical sequence $X \rightarrow$ future sequence $\hat{Y}$)

- **Setup:** learn from a meta-set—a collection of $k$-shot prediction tasks $T$

- **Meta-training:** train $P_0$ on small $\mathcal{D}_{train} = \{(X, Y^*)\}$ that achieves high performance on $\mathcal{D}_{val}$ for different prediction tasks on $\mathcal{D}_{test}$ of known action classes

- **Meta-testing:** adapt $P_0$ to solve prediction tasks on $\mathcal{T}_{test}$ of novel action classes

**Proactive Meta-Learner: General Model Initialization**

- **Model-agnostic meta-learning (MAML):** meta-learn a universal predictor under plain SGD updates [1]

- **Model adaptation to task $T_i$:** $\theta_i = \theta - \alpha \nabla_{\theta} L_{\mathcal{T}_i}(P_i)$

- **Model-objective:** maximal performance on $\mathcal{D}_{val}$ of task $T_i$

  $$ \min_{\theta} \sum_{i} \mathcal{L}_{\mathcal{T}_i}(P_i) = \min_{\theta} \sum_{i} \left( P_i - \nabla_{\theta} L_{\mathcal{T}_i}(P_i) \right) $$

- **Meta-objective:** a general-purpose initial $\theta$ across tasks: $\theta = \theta - \beta \nabla_{\theta} \sum_{i} \mathcal{L}_{\mathcal{T}_i}(P_i)$

**Adaptive Meta-Learner: Model Adaptation Strategy**

- **Model regression network (MRN):** guide model adaptation through meta-learning [2]

  - **Key insight:** a generic non-linear transformation $H$ from few-shot to many-shot model parameters

  - **Extension of tasks:** original image classification $\rightarrow$ our motion prediction

  - **Data:** explicitly leverage the original large training sets of known action classes

  - **Estimation of $H$ as a regression function during meta-training:** $\min_{\theta} \sum_i \mathcal{L}_{\mathcal{T}_i}(\theta_i - \theta_i')^2$

  $\theta_i$: learned on $\mathcal{D}_{val}$, by using SGD

  $\theta_i'$: learned on a large set of annotated sequences

**An Integrated, End-to-End Framework**

- **Integrated model adaptation during meta-training & meta-testing:** $\theta = H_{\theta}(\theta - \alpha \nabla_{\theta} L_{\mathcal{T}_i}(P_i))$

- **Integrated meta-objective during meta-training**

  $$ \min_{\theta} \sum_{i} \mathcal{L}_{\mathcal{T}_i}(P_i) = \min_{\theta} \sum_{i} \left( P_i - \nabla_{\theta} L_{\mathcal{T}_i}(P_i) \right) - \frac{1}{2} \lambda \left( \theta - \theta_i \right)^2 $$

**Ablation Analysis and Sanity Check**

- **Complementary components:** model initialization vs. adaptation

- **Improvements with more training sequences:** close to the oracle

- **Sanity check:** effectiveness for classification on mini-ImageNet

**Conclusions and Applications**

- **Extension of few-shot learning in a broader context:** image classification $\rightarrow$ motion prediction (imitation)

- **Meta-learning:** jointly learn generic model initialization & effective adaptation

- **Real-world scenario:** learn in an online, streaming manner with limited training data e.g., human-robot interaction and collaboration

---

**Recent Publications**
