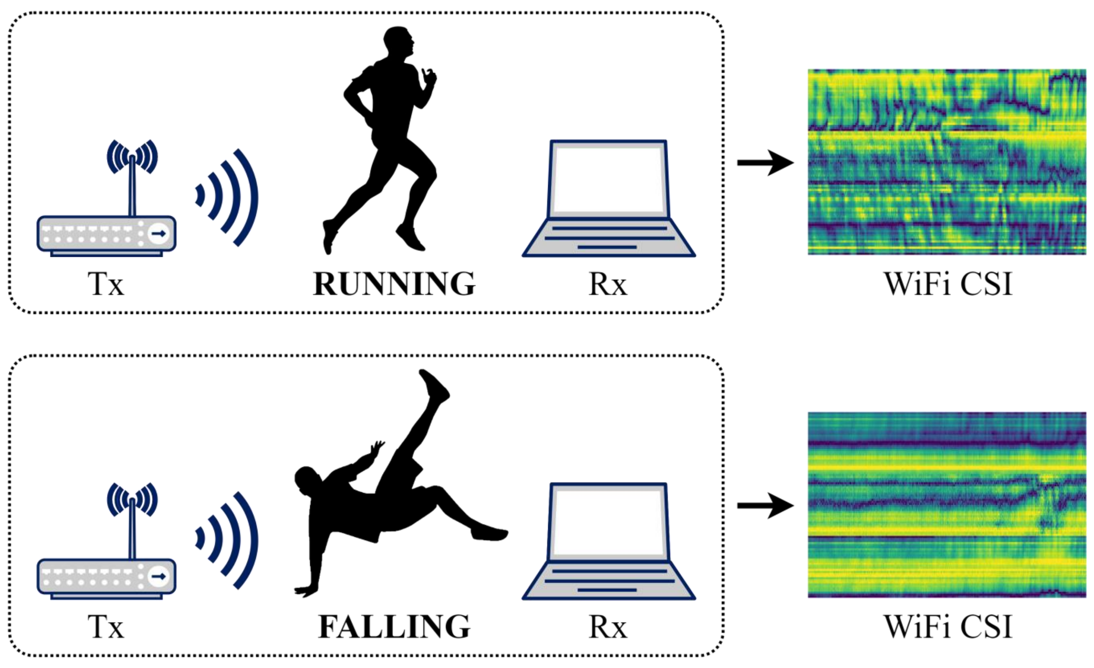
DiffAR: Adaptive Conditional Diffusion Model for Temporal-augmented Human Activity Recognition

Shuokang Huang[1], Po-Yu Chen[1,2], Julie A. McCann[1]

[1]Imperial College London, [2]JPMorgan Chase & Co.

BACKGROUND

WiFi channel state information (CSI) captures wireless signal variations caused by human interference to support device-free and non-intrusive human activity recognition (HAR).



PROBLEM

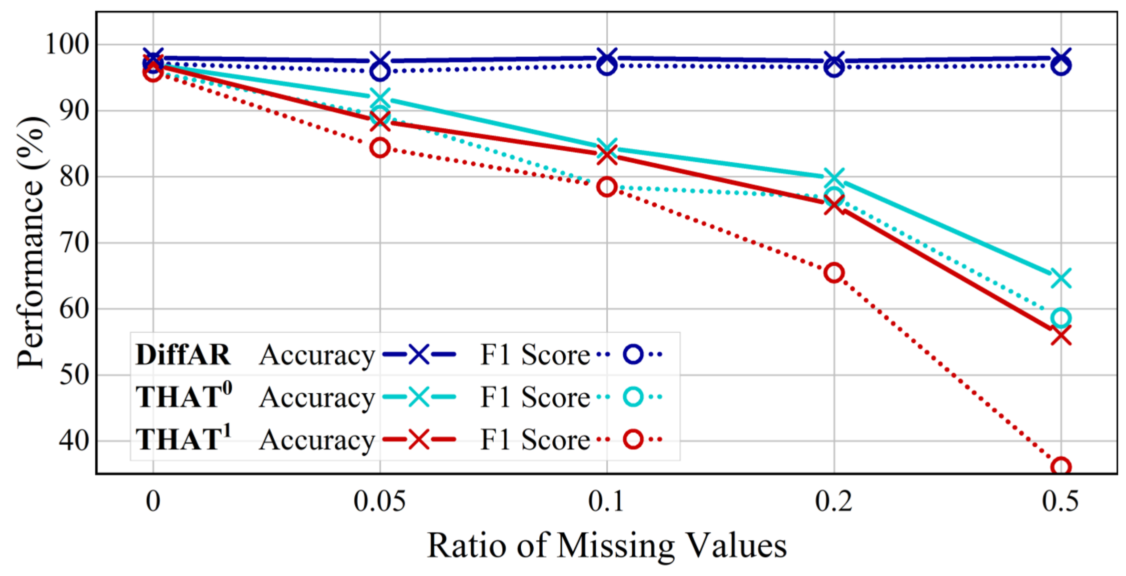
CSI-based HAR performance is hampered by incomplete CSI recordings due to fixed window sizes in CSI collection and human/machine errors that incur missing values in CSI.

CONTRIBUTIONS

- DiffAR strengthens CSI-based HAR using diffusion models.
- An adaptive conditioner guides the progressive steps with step-specific conditions for diffusion models to synthesize patterns of different granularity.

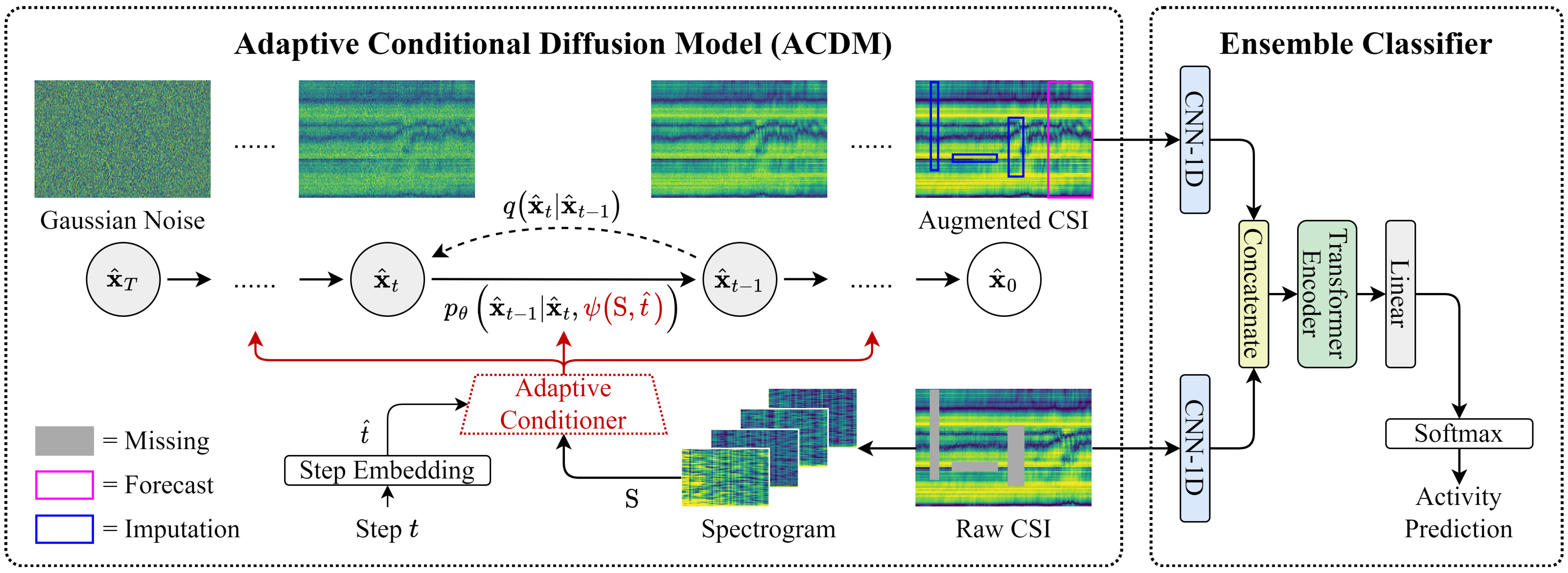
EXAMPLE

Comparison of **DiffAR** and **THAT** under different ratios of missing values in CSI. **THAT⁰** is tuned by samples with missing values, while **THAT¹** is not tuned by samples with missing values.



DiffAR

Overview of the proposed DiffAR. DiffAR augments incomplete WiFi CSI with diffusion models to improve the performance of CSI-based HAR.



FORMULATION

Adaptive Conditioner $\mathbf{c}_t = \psi(\mathbf{S}, \hat{\boldsymbol{t}}) = v(\mathbf{S}) \cdot \varphi(\boldsymbol{\omega}\hat{\boldsymbol{t}} + \boldsymbol{b})$

Objective $L^a(\theta) \coloneqq \mathrm{E}[\|\epsilon - \epsilon_\theta(\hat{\mathbf{x}}_t, t, \mathbf{c}_t)\|^2] = \mathrm{E}\left[\left\|\epsilon - \epsilon_\theta\left(\hat{\mathbf{x}}_t, t, \psi(\mathbf{S}, \hat{\boldsymbol{t}})\right)\right\|^2\right]$

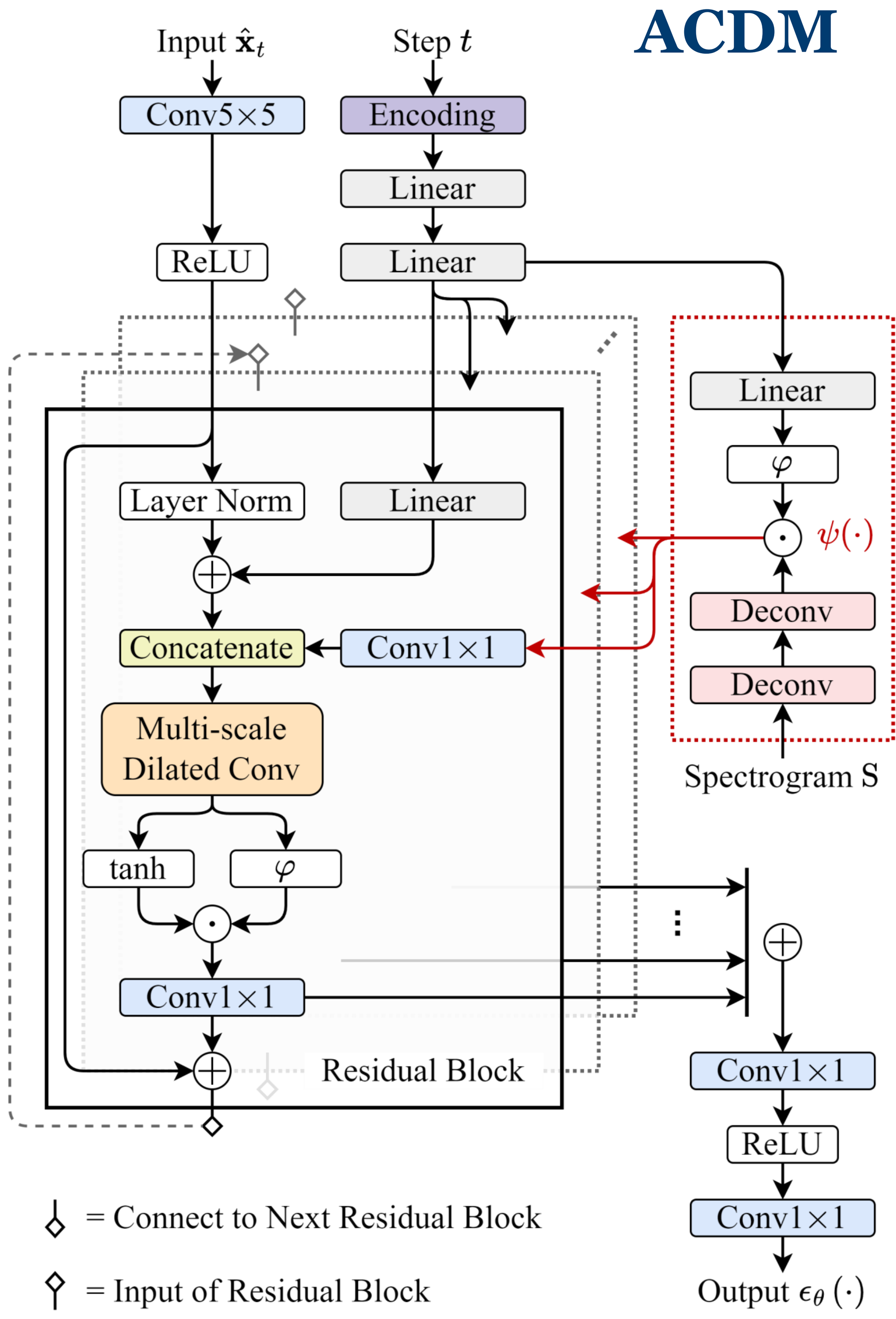
Algorithm 1 Training

repeat
1: $\hat{\mathbf{x}}_0 \sim q(\mathbf{x})$ # regard raw CSI as augmented CSI
2: $\mathbf{S}' = \mathrm{stft}(\mathbf{x}')$ where $\mathbf{x}' = \mathrm{mask}(\hat{\mathbf{x}}_0)$
3: $\hat{\boldsymbol{t}} = \mathrm{embed}(t)$ where $t \sim \mathrm{Uniform}(\{1, ..., T\})$
4: $\mathbf{c}'_t = \psi(\mathbf{S}', \hat{\boldsymbol{t}})$ # apply the adaptive conditioner
5: $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$
6: Take gradient step on
7: $\nabla_\theta \left\|\boldsymbol{\epsilon} - \boldsymbol{\epsilon}_\theta\left(\sqrt{\bar{\alpha}_t}\hat{\mathbf{x}}_0 + \sqrt{1-\bar{\alpha}_t}\boldsymbol{\epsilon}, t, \mathbf{c}'_t\right)\right\|^2$
until converged

Algorithm 2 Synthesis

Input: incomplete CSI $\mathbf{x} \in \mathbb{R}^{C \times N}$
1: $\mathbf{S} = \mathrm{stft}(\mathbf{x})$
2: $\hat{\mathbf{x}}_T \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ where $\hat{\mathbf{x}}_T \in \mathbb{R}^{C \times (1+\lambda_{\mathrm{fc}})N}$
3: **for** $t = T, ..., 1$ **do**
4: $\hat{\boldsymbol{t}} = \mathrm{embed}(t)$
5: $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ if $t > 1$ else $\mathbf{z} = 0$
6: $\mathbf{c}_t = \psi(\mathbf{S}, \hat{\boldsymbol{t}})$ # apply the adaptive conditioner
7: $\hat{\mathbf{x}}_{t-1} = \frac{1}{\sqrt{\alpha_t}}\left(\hat{\mathbf{x}}_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\boldsymbol{\epsilon}_\theta(\hat{\mathbf{x}}_t, t, \mathbf{c}_t)\right) + \sigma_t\mathbf{z}$
8: **end for**
return $\hat{\mathbf{x}}_0$

ACDM



CONCLUSION & FUTURE WORK

- DiffAR achieves the best quality of augmented CSI
- DiffAR outperforms state-of-the-art CSI-based HAR models
- Future work 1: Class imbalance in WiFi sensing
- Future work 2: Human sensing with dual-band WiFi singals