MetaHTR

Towards Writer-Adaptive Handwritten Text Recognition


Ayan Kumar Bhunia [1]


Shuvozit Ghose*


Amandeep Kumar*


Pinaki Nath Chowdhury [1,2]

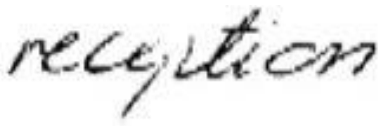

Aneeshan Sain [1,2]


Yi-Zhe Song [1,2]

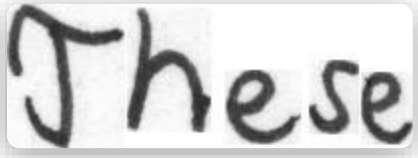
[1] SketchX Lab, CVSSP, University of Surrey, UK
[2] iFlyTek-Surrey Joint Research Centre on Artificial Intelligence
*Interned with SketchX Lab
http://sketchx.ai

CVPR VIRTUAL JUNE 19-25

Handwritten Text Recognition (HTR)

 → HTR Model → reputation

Problem – "People have diverse handwriting styles"






- **Idiosyncratic** style of writing characters
- **Cursive** or non-cursive

Traditional HTR Setup:

- Model is trained in a single flow
- No adaptation
- **Large dataset, synthetic** data
- Domain adaptation or generalisation

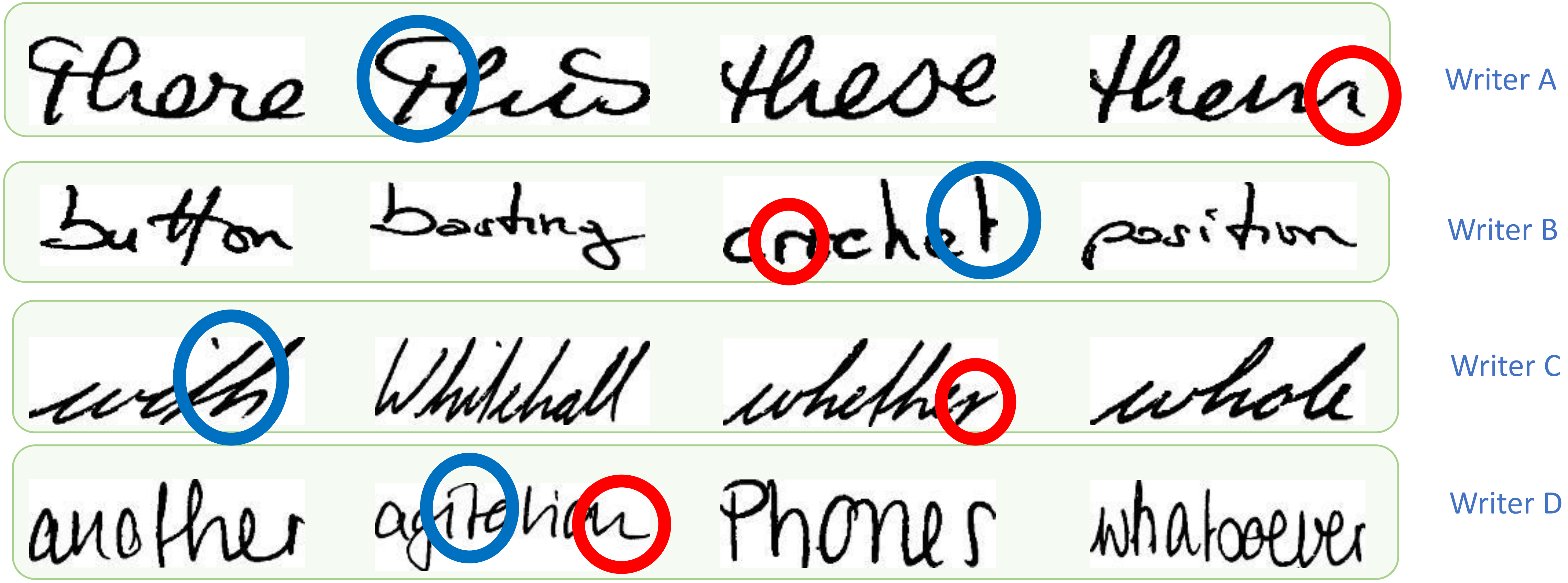
New style unobserved!

New *MetaHTR* Setup:

- Meta-learn handwriting styles
- Writer specialised HTR model
- Style adaptation using a **few examples.**

Quick adaptation to new style!

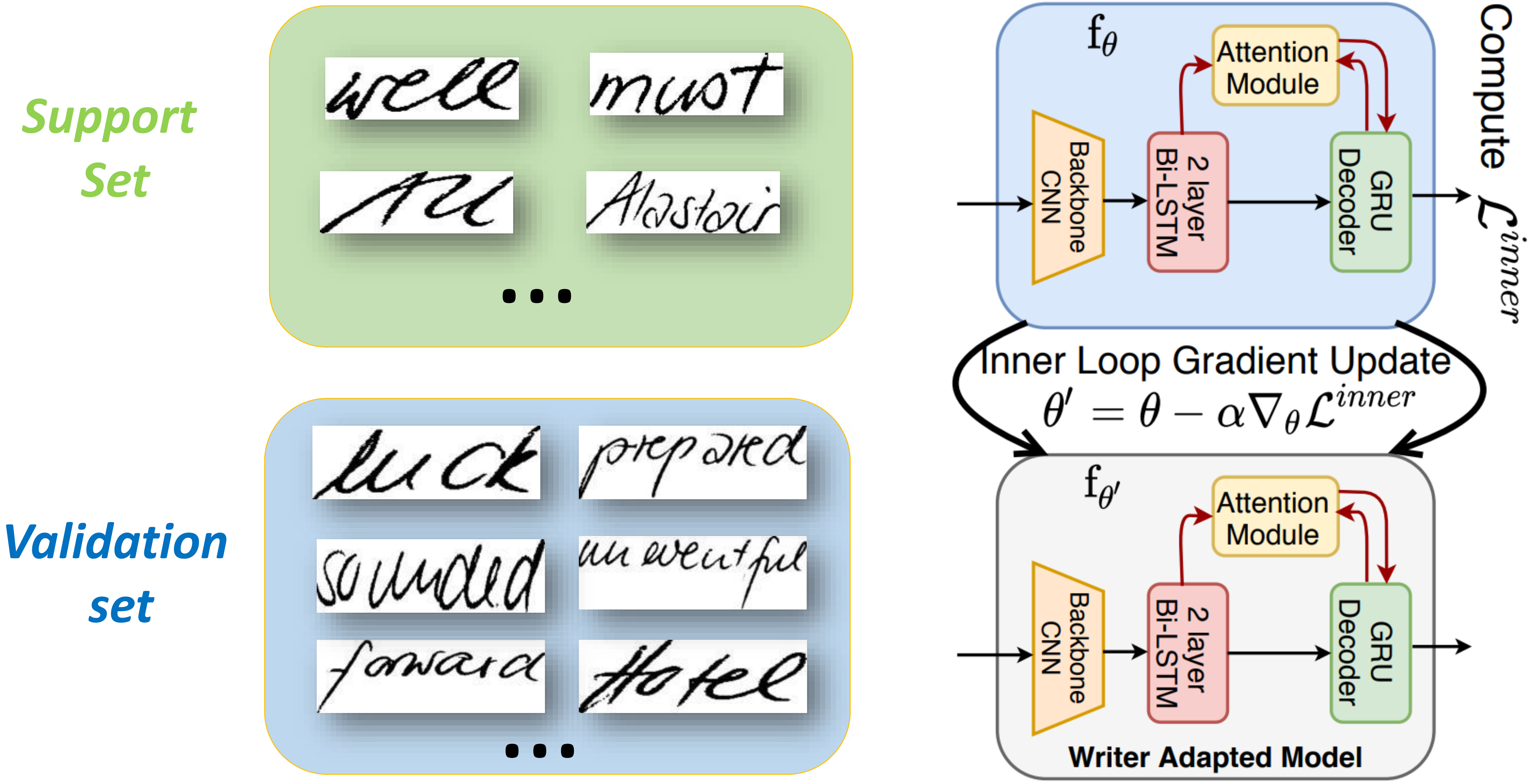
Challenges:



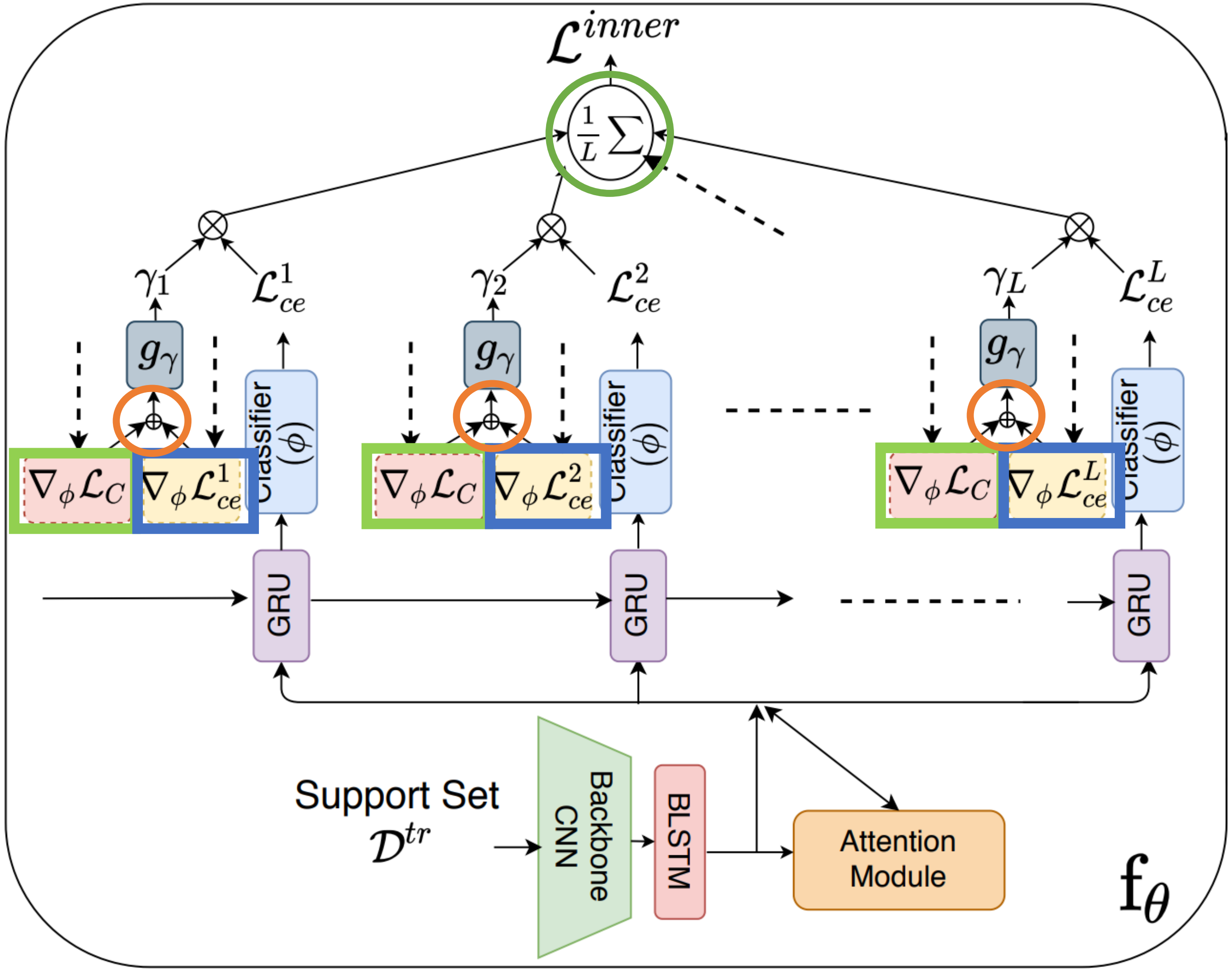
Contributions:

- Meta Handwritten Text Recognition (MetaHTR) for writer style adaptation.
- Meta-learn instance-wise weights, attending to character-specific adaptation.

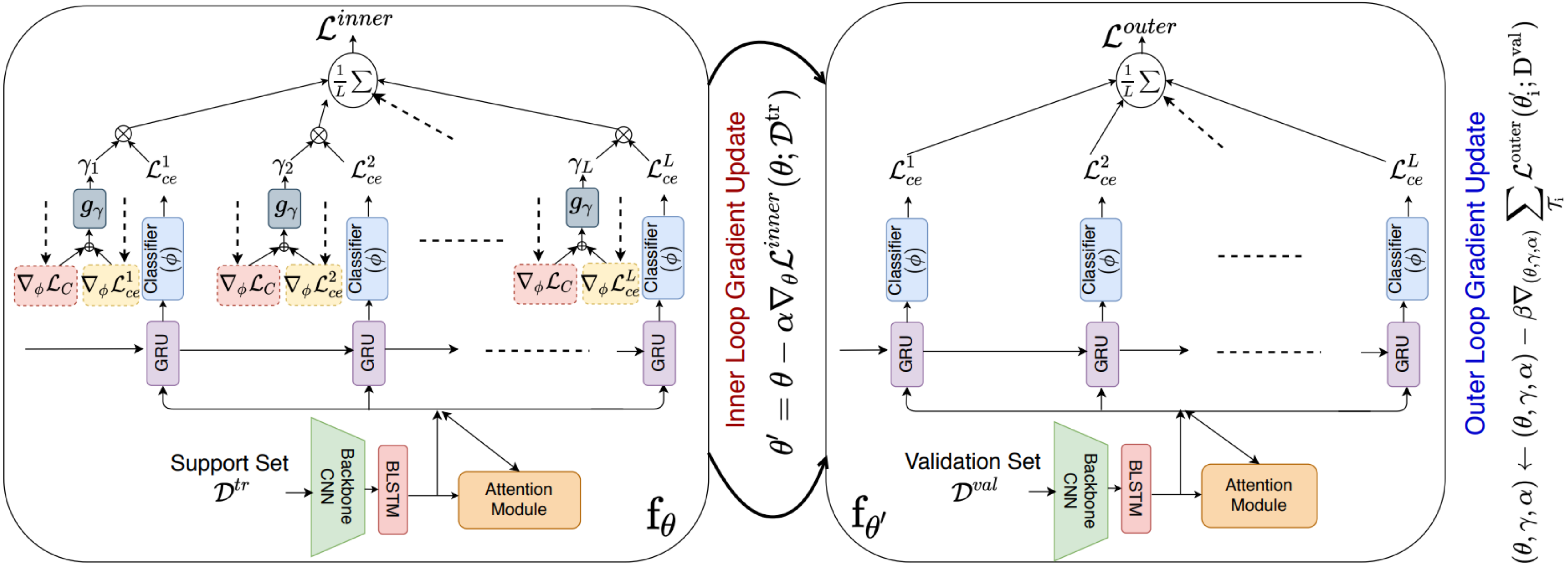
Meta Handwritten Text Recognition



Learning-to-Learn weights for Character-wise Loss



Bi-level Optimisation



Experiments

- **Datasets:** IAM[1] & RIMES[2]

- **Models:** HTR models ASTER[3], SAR[4], and SCATTER [5]

- **Evaluation Metric:**
  - Word Recognition Accuracy with-Lexicon (L)
  - Word Recognition Accuracy with No-Lexicon(NL)

- **Baselines:**
  - Domain Adaptation Approach
  - Domain Generalisation Approach
  - Generative Approach
  - Learning Augmentation Approach
  - Meta-Learning based Adaptation

[1] U-V Marti and Horst Bunke. The iam-database: an english sentence database for offline handwriting recognition. IJDAR, 2002.
[2] Emmanuele Grosicki and Haikal El Abed. Icdar 2009 hand- ` writing recognition competition. In ICDAR, pages 1398– 1402, 2009.
[3] Baoguang Shi, Mingkun Yang, Xinggang Wang, Pengyuan Lyu, Cong Yao, and Xiang Bai. Aster: An attentional scene text recognizer with flexible rectification. T-PAMI, 2018.
[4] Hui Li, Peng Wang, Chunhua Shen, and Guyu Zhang. Show, attend and read: A simple and strong baseline for irregular text recognition (SAR). In AAAI, 2019.
[5] Ron Litman, Oron Anschel, Shahar Tsiper, Roee Litman, Shai Mazor, and R Manmatha. Scatter: selective context attentional scene text recognizer. In CVPR, 2020.

Quantitative Results

Comparison among Baselines, naive Fine-tuning, and MetaHTR for using L and NL.

Methods	IAM[4] (WRA)							RIMES[5] (WRA)						
	Baseline		Fine-tuning		MetaHTR			Baseline		Fine-tuning		MetaHTR		
	L	NL	L	NL	L	NL	GAP	L	NL	L	NL	L	NL	GAP
ASTER[1]	90.3	81.3	90.5	81.7	94.1	89.2	7.9 ↑	93.6	87.4	93.7	87.7	96.5	93.4	6.0 ↑
SAR[2]	91.6	84.4	91.7	84.7	94.8	91.5	7.1 ↑	93.8	88.7	93.8	88.8	96.5	93.7	5.0 ↑
SCATTER[3]	91.7	84.6	92.0	85.1	94.8	91.6	7.0 ↑	93.8	88.8	93.9	89.0	96.6	93.9	5.1 ↑

[1] Baoguang Shi, Mingkun Yang, Xinggang Wang, Pengyuan Lyu, Cong Yao, and Xiang Bai. Aster: An attentional scene text recognizer with flexible rectification. T-PAMI, 2018.
[2] Hui Li, Peng Wang, Chunhua Shen, and Guyu Zhang. Show, attend and read: A simple and strong baseline for irregular text recognition (SAR). In AAAI, 2019.
[3] Ron Litman, Oron Anschel, Shahar Tsiper, Roee Litman, Shai Mazor, and R Manmatha. Scatter: selective context attentional scene text recognizer. In CVPR, 2020.
[4] U-V Marti and Horst Bunke. The iam-database: an english sentence database for offline handwriting recognition. IJDAR, 2002.
[5] Emmanuele Grosicki and Haikal El Abed. Icdar 2009 hand- ` writing recognition competition. In ICDAR, pages 1398– 1402, 2009.

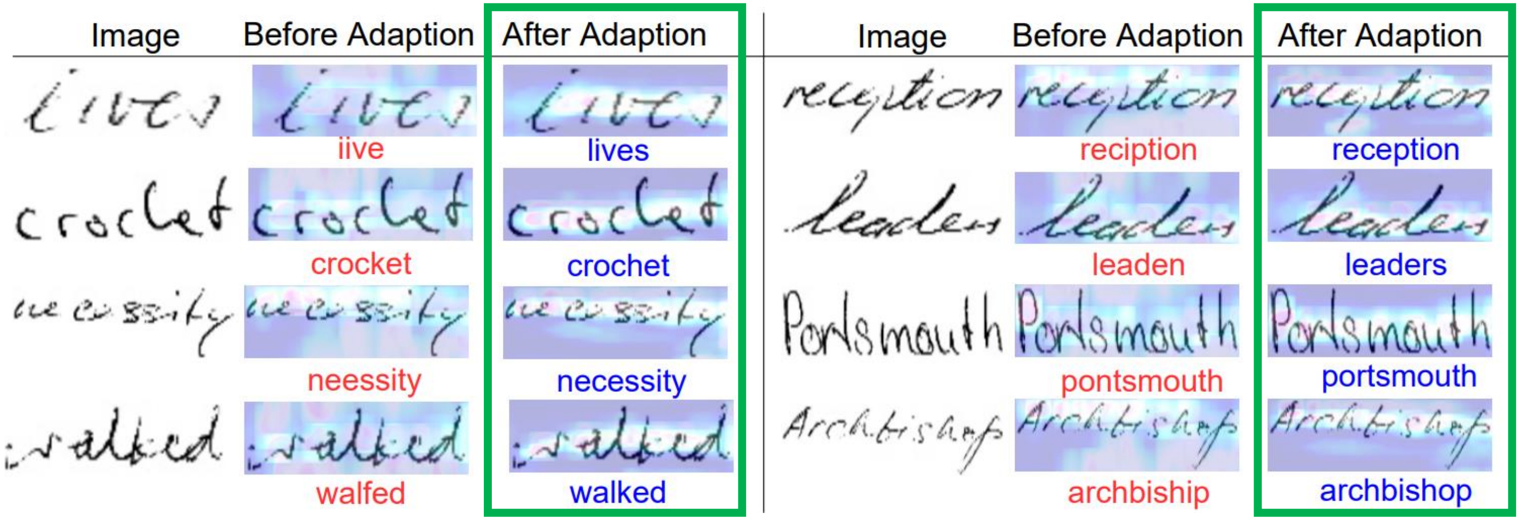
Quantitative Results:

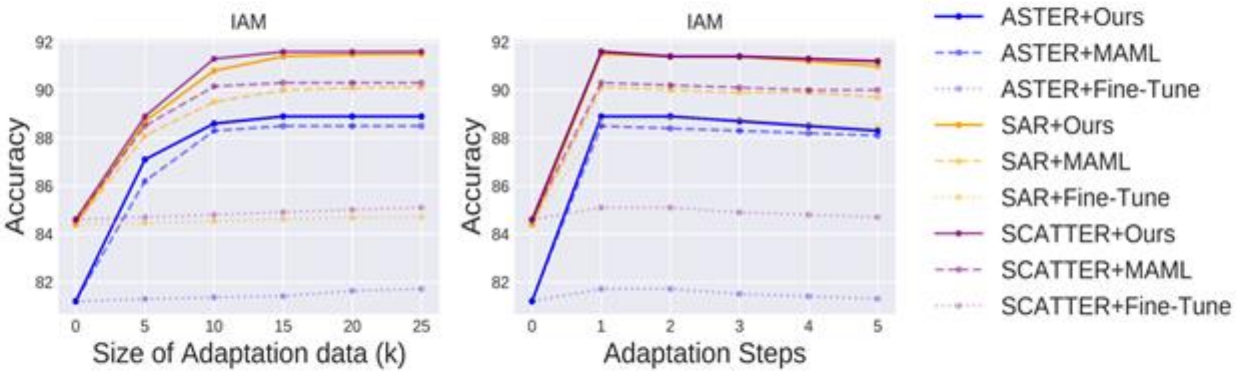
Performance analysis with different approaches

		IAM		RIMES	
		L	NL	L	NL
DG	ASTER [40] + DG	91.7	84.6	93.9	89.1
	SAR [26] + DG	92.4	85.7	94.3	89.5
	SCATTER [28] + DG	92.5	85.9	94.6	89.7
DA	ASTER [40] + DA	90.3	81.1	93.8	87.4
	SAR [26] + DA	91.9	84.8	93.7	88.9
	SCATTER [28] + DA	92.0	84.6	93.6	89.2
GA	ASTER [40] + GA	91.2	83.6	93.7	88.0
	SAR [26] + GA	91.8	84.8	93.8	88.4
	SCATTER [28] + GA	92.0	85.2	93.8	88.7
Augmnt.	Luo *et al.* [29]	92.5	86.0	95.6	90.8
	AFDM [6]	91.2	83.6	95.2	88.2
	Luo *et al.* + AFDM [6]	92.7	86.7	96.1	91.3
Meta Learning based Adaptation	ASTER [40] Baseline	90.3	81.3	93.6	87.4
	ASTER [40] + MAML	93.0	87.1	96.3	91.9
	ASTER [40] + MAML-FO	92.9	86.9	96.2	91.6
	ASTER [40] + MetaSGD	91.1	83.4	93.7	88.0
	ASTER [40] + ANIL	93.0	87.0	96.2	91.7
	ASTER [40] + **Ours (MetaHTR)**	94.1	89.2	96.5	93.4
	SAR [26] Baseline	91.6	84.4	93.8	88.7
	SAR [26] + MAML	94.1	89.1	96.4	92.4
	SAR [26] + MAML-FO	94.0	88.8	96.3	92.2
	SAR [26] + MetaSGD	91.8	84.9	93.9	88.9
	SAR [26] + ANIL	94.0	88.9	96.3	92.3
	SAR [26] + **Ours (MetaHTR)**	94.8	91.5	96.5	93.7
	SCATTER [28] Baseline	91.7	84.6	93.6	88.8
	SCATTER [28] + MAML	94.1	89.3	96.4	92.5
	SCATTER [28] + MAML-FO	94.0	88.9	96.3	92.3
	SCATTER [28] + MetaSGD	92.0	85.2	93.9	89.1
	SCATTER [28] + ANIL	94.1	89.1	96.4	92.4
	SCATTER [28] + **Ours (MetaHTR)**	94.8	91.6	96.6	93.9

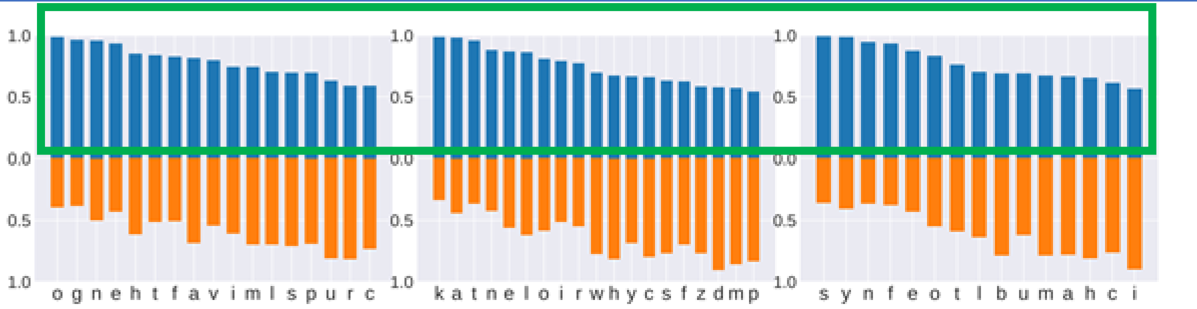
Ablative studies

Attention Map Visualization





Unconstrained W.R.A of MetaHTR with varying adaptation set-size (k), and adaptation steps.



Character-sets of three writers show varying W.R.A, signifying different levels of discrepancy against MetaHTR's initialization parameter. Characters having lower accuracy mostly get higher weights (orange) , and vice versa, in the inner-loop adaptation process.

SketchX

http://sketchx.ai