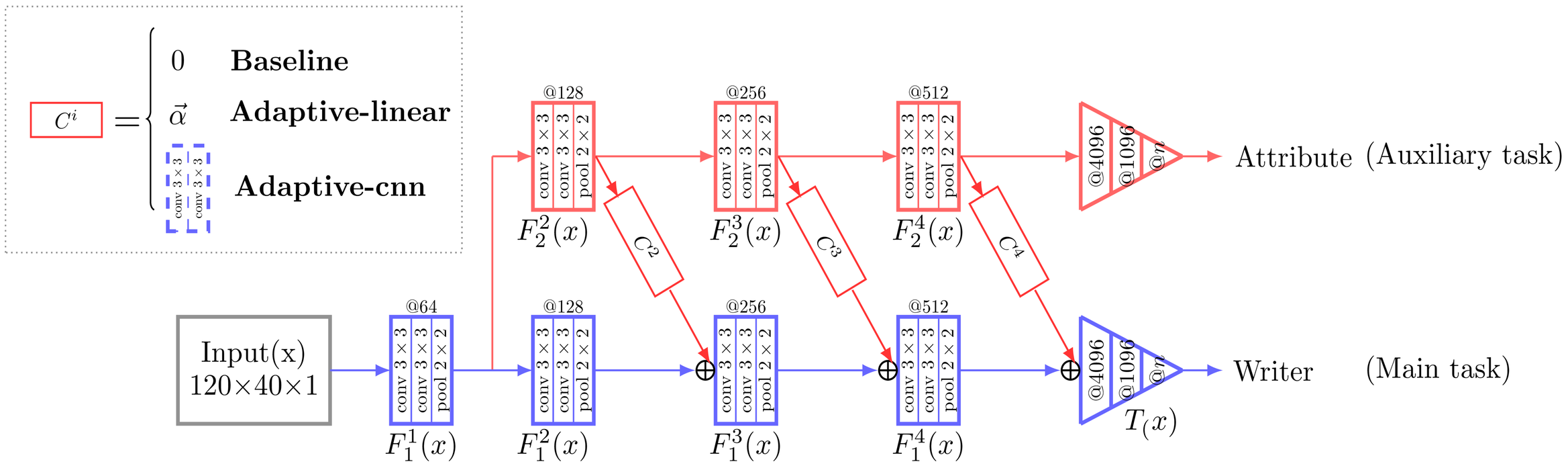
# Deep Adaptive Learning in Multiple Tasks

Sheng He and Lambert Schomaker
ALICE, University of Groningen, The Netherlands

## Network




## Results

| Model | Writer Identification | | | | Word Recognition | | | | Writer Identification | | | | Attribute Recog. | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CVL | | IAM | | CVL | | IAM | | CVL | | IAM | | CVL | IAM |
| | Top-1 | Top-5 | Top-1 | Top-5 | Top-1 | Top-5 | Top-1 | Top-5 | Top-1 | Top-5 | Top-1 | Top-5 | Precision | Precision |
| Baseline | 75.3 | 92.4 | 65.7 | 83.5 | 95.1 | 99.1 | 93.5 | 98.7 | 75.3 | 92.4 | 65.7 | 83.5 | 93.4 | 91.3 |
| Linear-adaptive | 77.0 | 93.1 | 68.0 | 84.7 | 94.1 | 98.9 | 91.3 | 98.1 | 75.3 | 92.4 | 65.5 | 83.4 | 82.8 | 77.9 |
| Deep-adaptive | **78.6** | **93.7** | **69.5** | **86.1** | 94.5 | 99.0 | 92.6 | 98.4 | **76.5** | **93.2** | **67.9** | **84.3** | 85.1 | 81.6 |

## Discussion

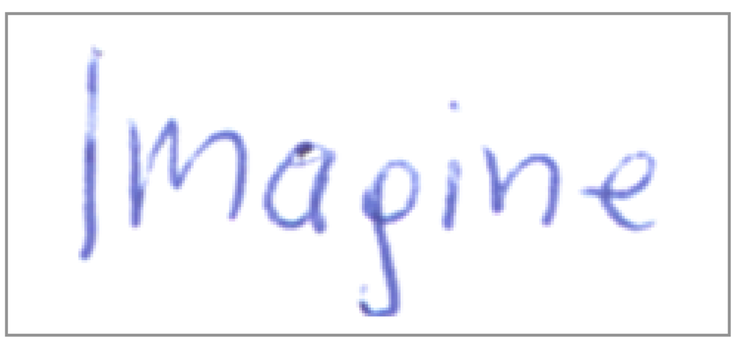
1, Generally, under the same conditions, recognizing the implicit information (writer identification) is more difficult than the explicit information (word recognition).

2, Adaptive learning can improve the performance of the main task.

3, Deep adaptive learning provides better performance than linear-adaptive learning.

4, The performance of the auxiliary task decreases because the information of the main task is back-propagated to the layers of the auxiliary task.

5, Transferring the features learned from other tasks can introduce an inductive bias which can reduce the risk of overfitting.




Implicit:
Writer identity: "bob"

Explicit:
Word content: 'Imagine'
Word length: 7 letters
Character attribute: a, e, g, i, m, n