Curriculum Learning for Multi-Task Classification of Visual Attributes

Nikolaos Sarafianos[1], Theodore Giannakopoulos[2], Christophoros Nikou[3] and Ioannis A. Kakadiaris[1]

[1]Computational Biomedicine Lab, University of Houston   [2]NCSR Demokritos   [3]University of Ioannina

ICCV17 International Conference on Computer Vision 2017

Introduction

Problem Statement
- Given an image of a human, the question is: how can someone effectively predict the corresponding visual attributes?
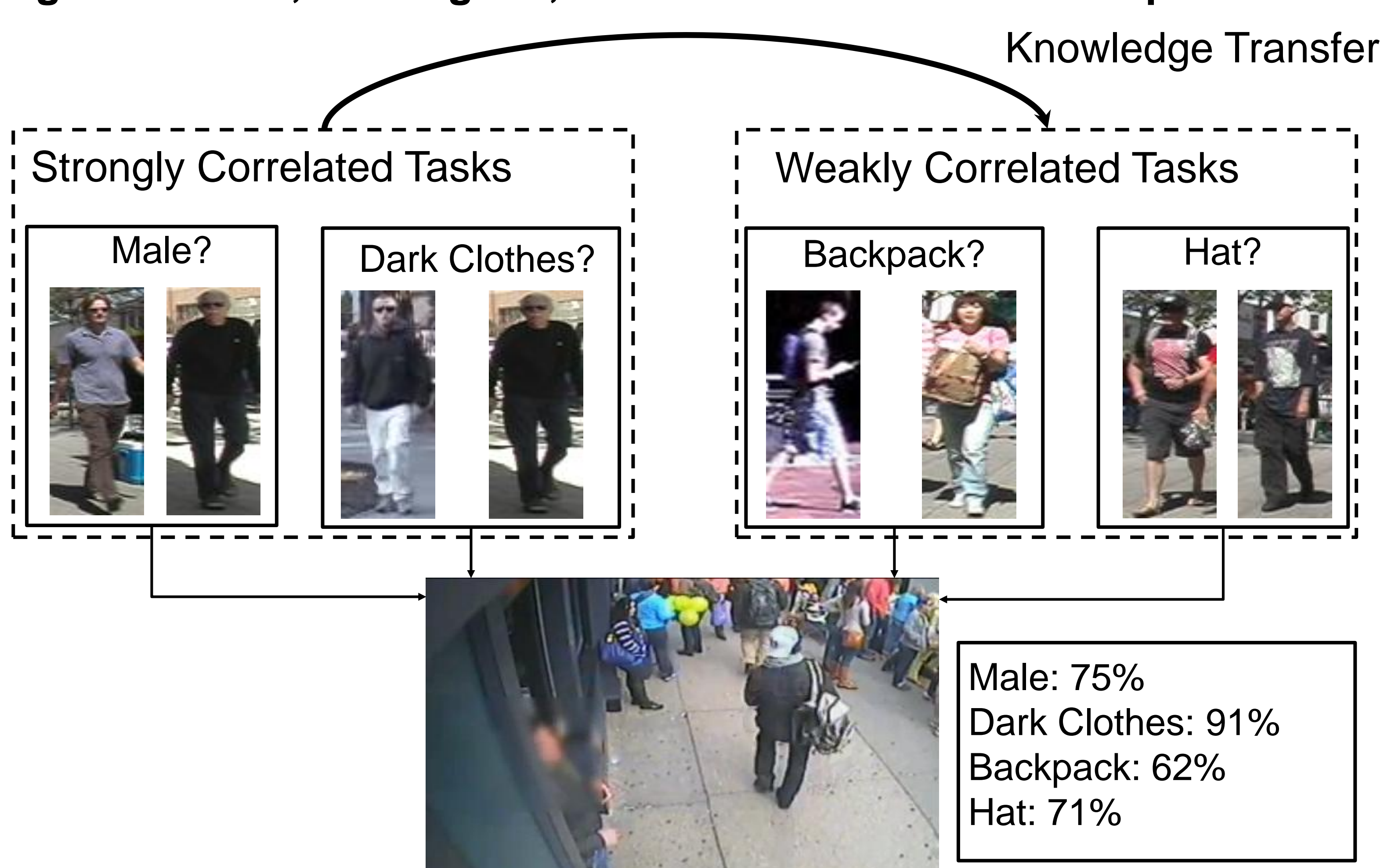
Motivation
- Recognize visual attributes using multi-task (MT) classification and curriculum learning

Contributions
- Introduced a curriculum learning paradigm to perform multi-task classification
- Converged faster than typical MT learning by employing prior knowledge
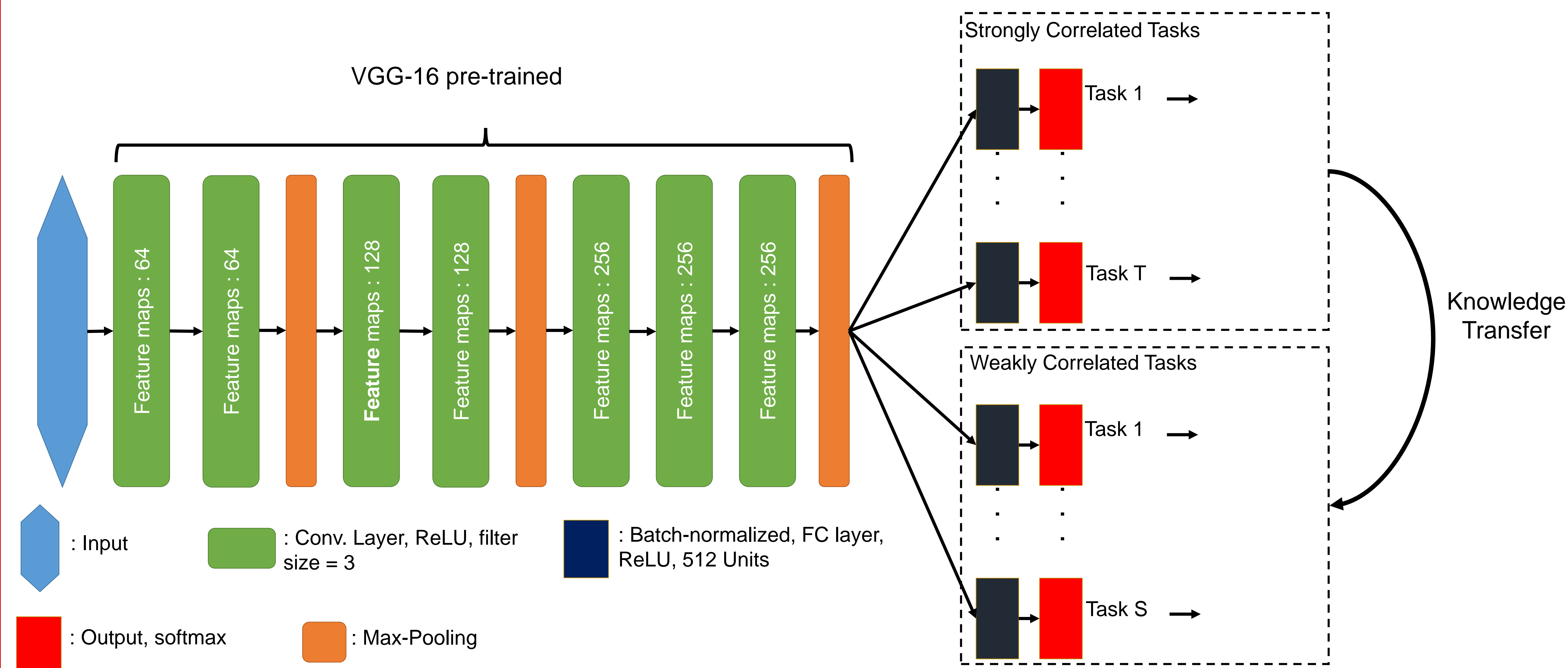
Overview

Images of males, wearing hat, with dark clothes and a backpack

Knowledge Transfer

Strongly Correlated Tasks: Male? | Dark Clothes?

Weakly Correlated Tasks: Backpack? | Hat?

Male: 75%
Dark Clothes: 91%
Backpack: 62%
Hat: 71%

Method

CILICIA Network Architecture

VGG-16 pre-trained

Feature maps : 64, Feature maps : 64, Feature maps : 128, Feature maps : 128, Feature maps : 256, Feature maps : 256, Feature maps : 256

Strongly Correlated Tasks: Task 1 ... Task T
Weakly Correlated Tasks: Task 1 ... Task S

Knowledge Transfer

- : Input
- : Conv. Layer, ReLU, filter size = 3
- : Batch-normalized, FC layer, ReLU, 512 Units
- : Output, softmax
- : Max-Pooling

Correlation-based Group Split

- Total dependency p_i of task t_i: $p_i = \sum_{j=1, j \neq i}^{T} \frac{cov(y_{t_i}, y_{t_j})}{\sigma(y_{t_i}) \cdot \sigma(y_{t_j})}, i = 1, \ldots, T$

- Assign tasks with top 50% of p_i as strongly correlated, rest as weakly
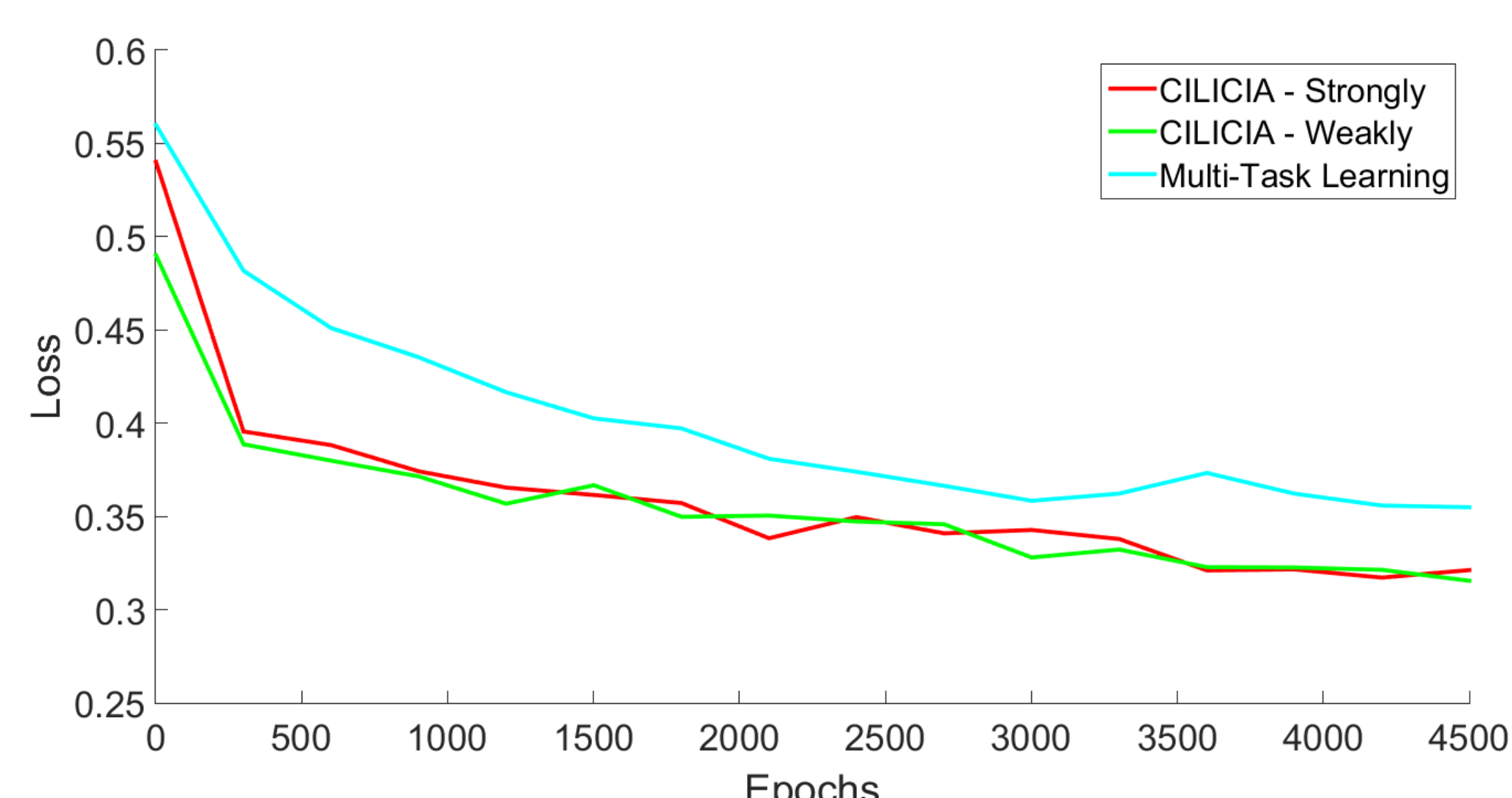
Multi-task Curriculum Learning

- Strongly-correlated loss: $L_s = \sum_{t=1}^{T} \lambda_t \cdot L_t$

- Weakly-correlated loss: $L_w = \lambda \cdot L_s + (1 - \lambda) L_w^f$

Ablation Studies

Group	Random Split	CILICIA	CILICIA (without knowledge transfer)
Strongly	65.36	**76.01**	**76.01**
Weakly	63.08	**71.80**	69.91
Total	64.22	**73.91**	72.95

Q_2: Is correlation-base split beneficial? Q_1: Is knowledge transfer beneficial?

CILICIA - Strongly
CILICIA - Weakly
Multi-Task Learning

Q_3: Does CILICIA converge faster?

Results

Group	SoBiR Dataset					VIPeR Dataset		
	SVM	Individual Learning	Multi-Task Learning	CILICIA		Multi-Task Learning	Zhu *et al.* [1]	CILICIA
Strongly Av.	58.3	69.3	71.3	**72.3**		73.4 ± 2.6	75.7 ± 3.2	**85.1 ± 1.0**
Weakly Av.	65.6	73.0	73.2	**73.7**		71.9 ± 1.8	72.5 ± 1.7	**74.8 ± 0.5**
Total Av.	61.9	71.2	72.3	**73.1**		73.2 ± 1.2	74.1 ± 1.0	**80.5 ± 2.6**

Key Takeaway

Combining curriculum learning with multi-task classification improves the classification accuracy up to 1% in the SoBiR dataset and by approximately 6% in the VIPeR dataset.

References

[1] J. Zhu, S. Liao, D. Yi, Z. Lei, and S. Z. Li. Multi-label CNN based pedestrian attribute learning for soft biometrics. In Proc. IEEE International Conference on Biometrics, Phuket, Thailand, May 19-22 2015.

UNIVERSITY of HOUSTON | CBL
Changing the way people look at computers computers people