A TWO-STREAM SIAMESE NEURAL NETWORK FOR VEHICLE RE-IDENTIFICATION BY USING NON-OVERLAPPING CAMERAS

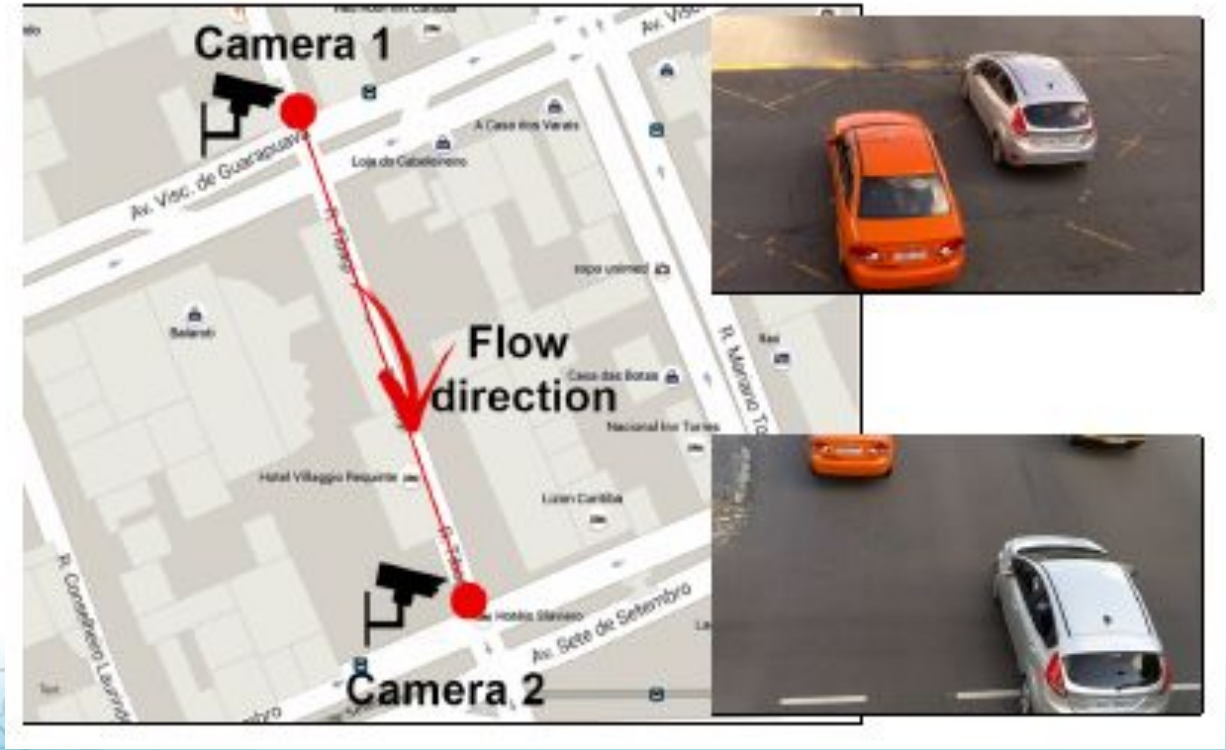
Icaro O. de Oliveira, Rodrigo Minetto and Keiko V. O. Fonseca
Federal University of Technology - Paraná (UTFPR), Curitiba, Brazil

Outline

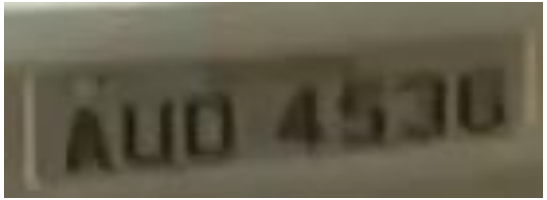
- Introduction
- Related Work
- Two-Stream Siamese Network
- Experiments
- Conclusions
- Sample Video

INTRODUCTION

System Setup



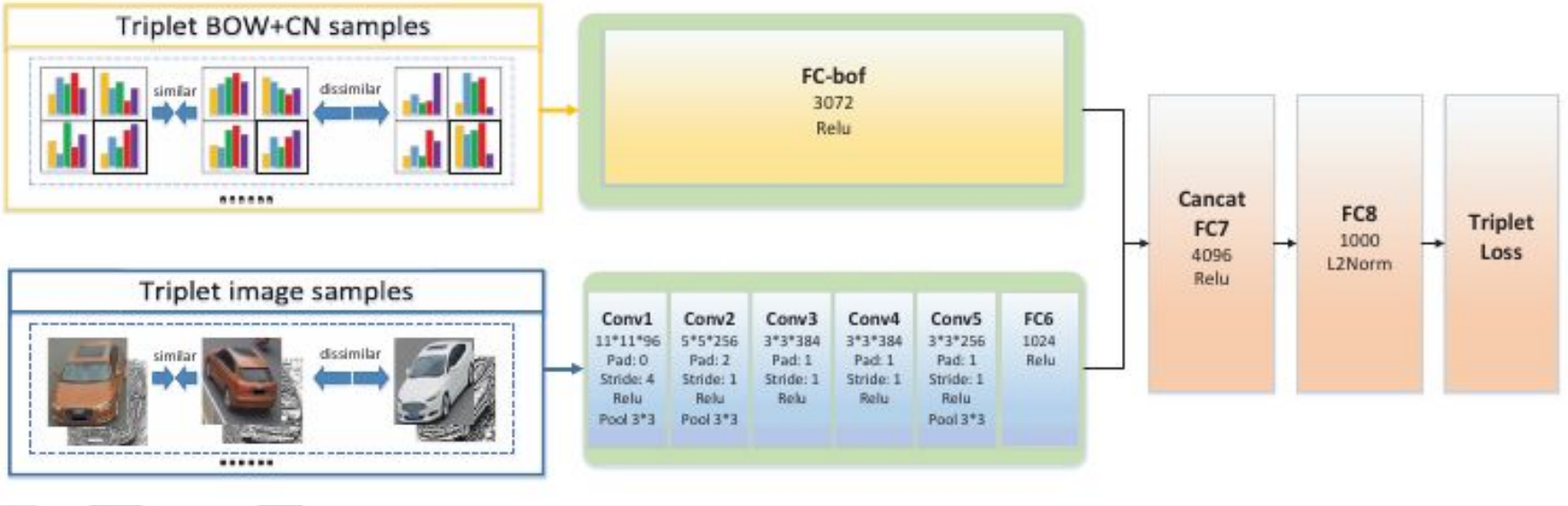
Problem



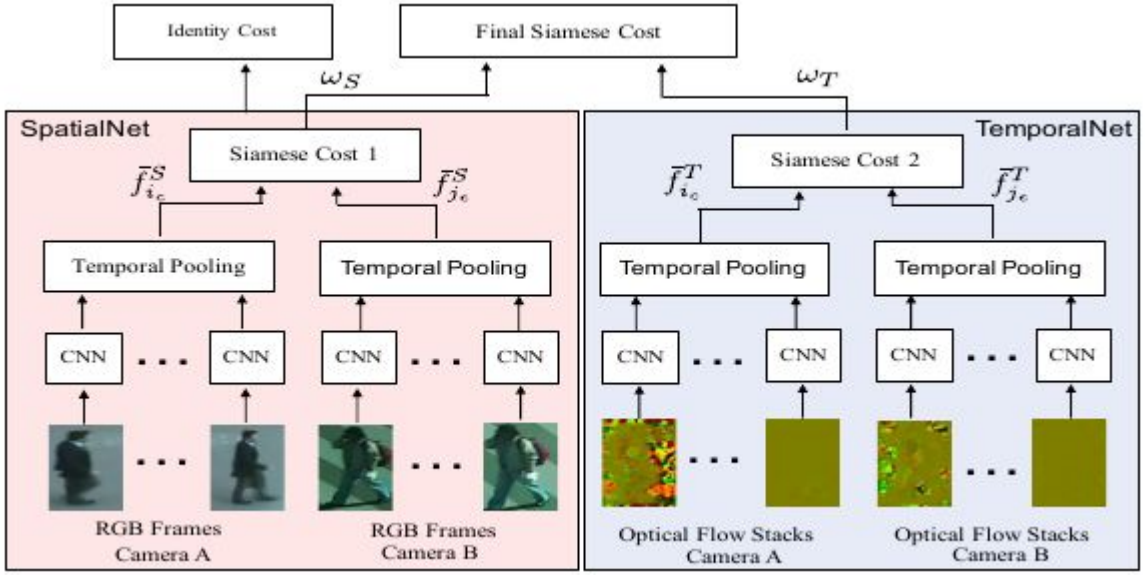
OCR: AUO-4530

RELATED WORK

Y. Tang, D. Wu, Z. Jin, W. Zou, and X. Li, “Multi-modal metric learning for vehicle re-identification in traffic surveillance environment,” ICIP 2017,pp. 2254–2258. (Related Work)

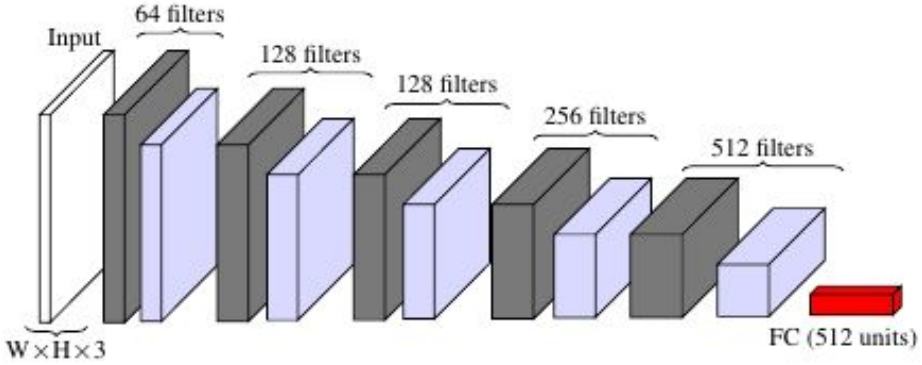


Dahjung Chung, Khalid Tahboub, and Edward J Delp, “A two stream siamese convolutional neural network for person re-identification,” ICCV, 2017. (Related Work)

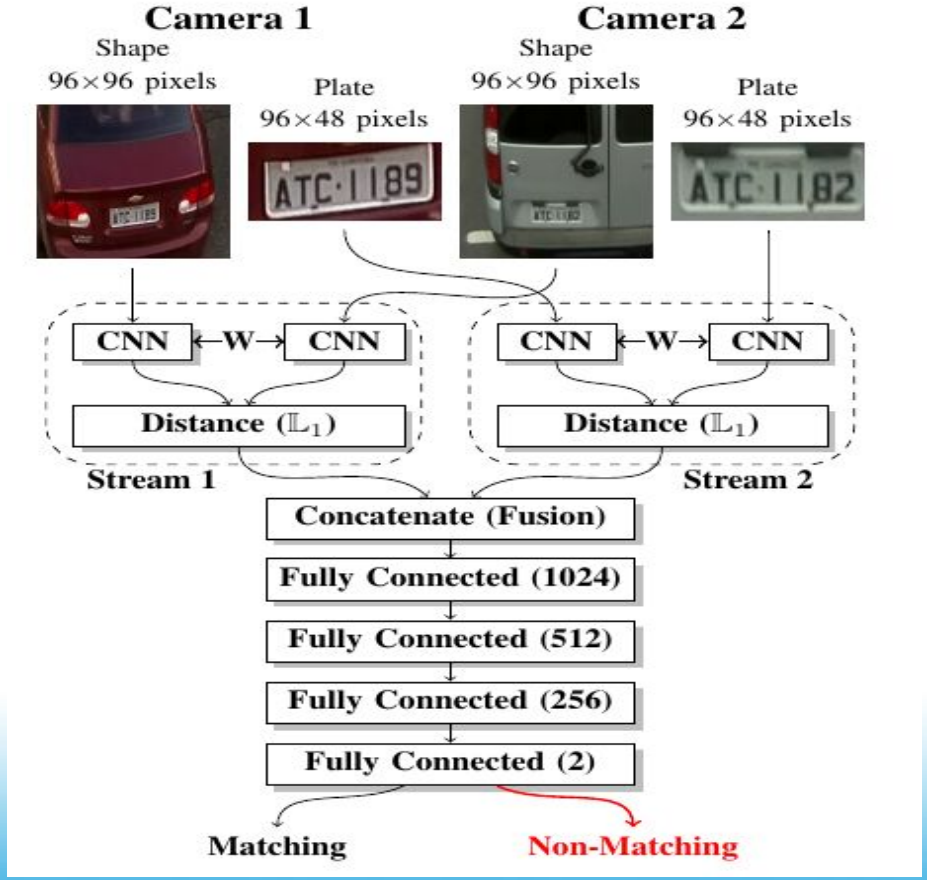


TWO STREAM SIAMESE NETWORK

Our Approach



Our Approach



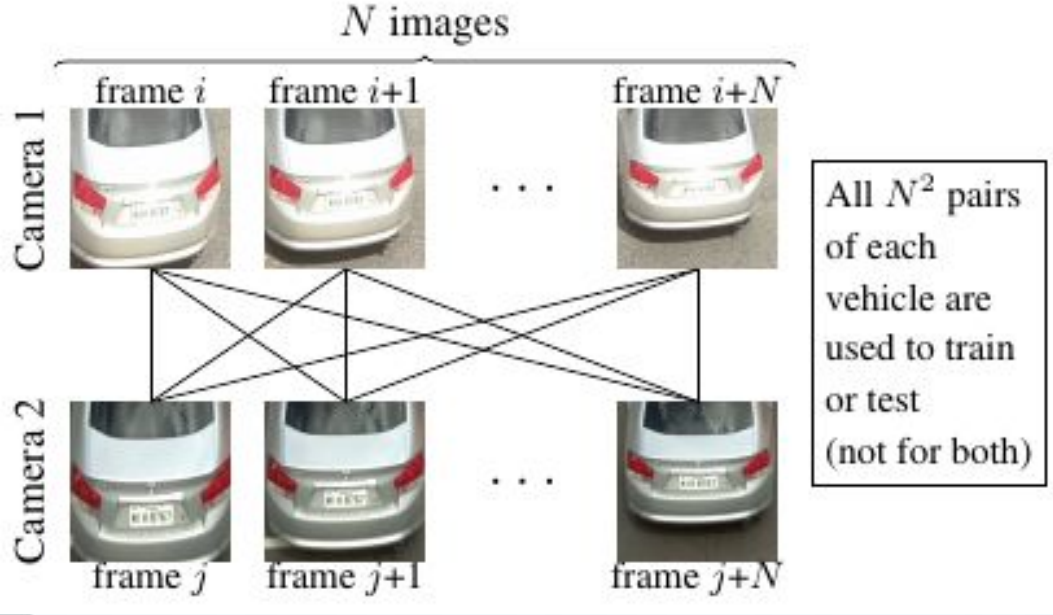
EXPERIMENTS

Experimental Setup: Dataset used

Table 1. Dataset information: number of vehicles and number of vehicles with a visible license plate in Camera 1 and 2; number of vehicles matchings between Camera 1 and 2.

Set	Camera 1		Camera 2		No.Match.
	#Vehicles	#Plates	#Vehicles	#Plates	
01	389	343	280	245	199
02	350	310	244	227	174
03	340	301	274	248	197
04	280	251	233	196	140
05	345	295	247	194	159
Total	1704	1500	1278	1110	869

Experimental Setup: Dataset used



Experimental Setup: Dataset used

Table 2. Parameter settings used in our experiments.

Settings	Training		Testing	
	#positives	#negatives	#positives	#negatives
$N = 3, \lambda = 5$	3867	3867	3903	19515
$N = 10, \lambda = 10$	42130	42130	42707	427070

Experimental Setup: Different Settings

	$N = 3, \lambda = 5$			
Algorithm	P	R	F	A
Siamese-Car (Stream 1)	85.8%	93.1%	89.3%	96.3%
Siamese-Plate (Stream 2)	75.9%	81.8%	78.8%	92.6%
Siamese (Two-Stream)	92.7%	93.0%	92.9%	97.6%
	$N = 10, \lambda = 10$			
Algorithm	P	R	F	A
Siamese-Car (Stream 1)	92.4%	83.5%	87.8%	97.9%
Siamese-Plate (Stream 2)	86.8%	59.5%	70.6%	95.5%
Siamese (Two-Stream)	94.7%	90.6%	92.6%	98.7%

Experimental Setup: Different Settings

	$N = 10, \lambda = 10$			
Siamese (Two-Stream)	P	R	F	A
CNN = Lenet5	89.6%	85.2%	87.3%	97.8%
CNN = Matchnet [22]	94.5%	87.1%	90.7%	98.4%
CNN = MC-CNN [23]	89.0%	90.1%	89.6%	98.1%
CNN = GoogleNet	88.8%	81.8%	85.1%	97.4%
CNN = AlexNet	91.3%	86.5%	88.8%	98.0%
CNN = Small-VGG	94.7%	90.6%	92.6%	98.7%

Experimental Setup: Complex Architectures

Vehicle (96×96 pixels)

Plate (96×48 pixels)

Siamese Two-Stream
(**Small-VGG**)
F = 92.6% and A 98.7%

VS

Vehicle (patches 224×224 pixels)

Siamese-Car (**Small-VGG**):
F = 88.1% and A = 97.9%
Siamese-Car (**Resnet50**):
F = 81.2% and A = 97.1%

Experiment Scenarios

Siamese-Car (Stream 1): **non-matching** ✗
Siamese-Plate (Stream 2): **non-matching** ✗
Siamese (Two-Stream): **non-matching** ✗

Experiment Scenarios

Siamese-Car (Stream 1): **matching** ✗
Siamese-Plate (Stream 2): **non-matching** ✓
Siamese (Two-Stream): **non-matching** ✓

Experiment Scenarios



Siamese-Car (Stream 1): **non-matching** ✓
Siamese-Plate (Stream 2): **matching** ✗
Siamese (Two-Stream): **non-matching** ✓

Experiment Scenarios




Siamese-Car (Stream 1): **matching** ✓
Siamese-Plate (Stream 2): **matching** ✓
Siamese (Two-Stream): **matching** ✓

CONCLUSIONS

Conclusions

- Fast Two-Stream Siamese;
- Tests: network more robust than other One-Stream Siamese architectures.
- Evaluation of simple and complex CNNs, used by the Siamese Network, to find a trade-off between efficiency and performance.

Future Works

- Combine OCR with the two distinctive features (shape and plate);
- More evaluations on the other public datasets;
- Combine spatial-temporal informations;
- Create a N Temporal Two Streams Siameses;

SAMPLE VIDEO

Thank you for your attention !
Questions ?